ASEFEH ASEMI

A NOVEL COMBINED INVESTMENT RECOMMENDER SYSTEM USING ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

Department of Information Systems
Doctoral School of Economics, Business, & Informatics

A Novel Combined Investment Recommender System Using Adaptive Neuro-Fuzzy Inference System

Ph.D. Dissertation in Business Informatics

Supervisor Prof. Andrea Ko

Budapest 2023

@ Asefeh Asemi

Corvinus University of Budapest
Doctoral School of Economics, Business, and Informatics

Supervisor Prof. Andrea Ko

©Asefeh Asemi

CONTENTS

List of Figures

LIST OF TABLES

LIST OF ABBREVIATIONS

3D	THREE-DIMENSIONAL
AICC	CORRECTED AKAIKE INFORMATION CRITERION
ANFIS	ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM
ANFIS	ADAPTIVE NETWORK-BASED FUZZY INFERENCE SYSTEM
ANN	ARTIFICIAL NEURAL NETWORK
ASR	AUTOMATIC SPEECH RECOGNITION
B2C	BUSINESS-TO-CONSUMER
BIC	BAYESIAN INFORMATION CRITERION
CCC	CUBIC CLUSTER CRITERION
CEE	CENTRAL AND EASTERN EUROPEAN
DITS	DIFFERENTIAL INVESTMENT TYPE SERVICE
DSS	DECISION SUPPORT SYSTEM
FIS	FUZZY INFERENCE SYSTEM
GAUSSMF	GAUSSIAN MEMBERSHIP FUNCTION
IITS	INTEGRATED INVESTMENT TYPE SERVICE
IRS	INVESTMENT RECOMMENDER SYSTEM
LR	LOW RISK
LRLC	LOW RISK-LOW COST
MF	MEMBERSHIP FUNCTION
MFT	MULTILATERAL TRADING FACILITY
NA	NO ANSWER (MISSING VALUE)
NFS	NEURAL-FUZZY SYSTEMS
OLR	OPPORTUNITY FOR HIGH RETURNS-LOW RISK
OLRLC	OPPORTUNITY FOR HIGH RETURNS-LOW RISK-LOW COST
OTF	ORGANIZED TRADING FACILITY
RMSE	SQUARE ROOT MEAN ERROR
RS	RECOMMENDER SYSTEM
SLR	STATE AID-LOW RISK
SLRLC	STATE AID-LOW RISK-LOW COST
SO	STATE AID-OPPORTUNITY FOR HIGH RETURNS
SOLC	STATE AID-OPPORTUNITY FOR HIGH RETURNS-LOW COST
SOLRLC	STATE AID-OPPORTUNITY FOR HIGH RETURNS-LOW RISK-LOW COST
SOM	SELF-ORGANIZING MAP
SSO	SOCIAL SECURITY ORGANIZATION
SSE	SUM OF SQUARED ERRORS
SVM	SUPPORT VECTOR MACHINES
TRAPMF	TRAPEZOIDAL MEMBERSHIP FUNCTION
TRIMF	TRIANGULAR MEMBERSHIP FUNCTION

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Andrea Ko, and my advisor, Dr. Adeleh Asemi, for their unwavering guidance, encouragement, and support throughout my doctoral journey. Their invaluable insights and expertise have been instrumental in shaping my research and have helped me to become a better scientist.

I would also like to acknowledge Prof. Peter Racskó for providing data for this research and Prof. Peter Feher for allowing me to publish data.

Furthermore, I extend my thanks to the members of my dissertation committee for their time, feedback, and constructive criticism. Their invaluable input has been instrumental in the success of my research.

I am also grateful to dr. Zoltán Szabó and Mrs. Arany Zsuzsanna for their help and support throughout my research. I would like to express my gratitude to my colleagues and friends for their support and for providing a stimulating and collaborative environment.

I am grateful to Corvinus University of Budapest for providing me with the resources and support necessary to complete my research. In particular, I would like to thank the Doctoral School of Economics, Business, & Informatics, and the Institute of Data Analytics and Information Systems for their assistance and support.

Last but not least, I express my appreciation to my family for their love, support, and understanding. They have been my source of strength and inspiration throughout this journey.

Thank you all for your contributions to my research and for helping me to reach this point in my academic career.

Asefeh Asemi
30 Jan 2023

CHAPTER I INTRODUCTION

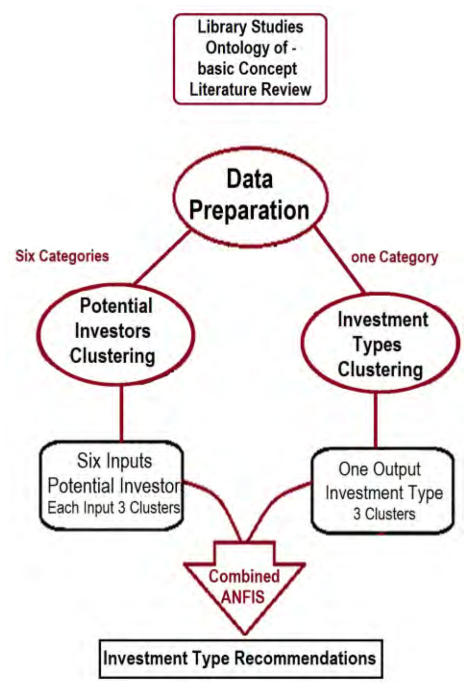
Information overload is a prevalent issue in today's digital age, and the need for intelligent recommender systems that can filter and deliver relevant information to users is becoming increasingly important. According to Hanani et al. (2001), "An information filtering system removes redundant or unwanted information from an information stream using (semi) automated or computerized methods before presentation to a human user." In situations where there is an overwhelming number of choices, recommender systems can help prioritize and efficiently deliver relevant information to alleviate the problem of information overload. Recommender systems can also be used in customer relationship management to create a 364-degree view of the customer. Research has shown that building relationships is a crucial factor in the success or failure of organizations, customers, and transactions (Woodcock et al., 2011). Customer interaction is also important, with trading places taking second place in terms of importance (Van Looy, 2016). With the increasing use of the Internet for purchasing information or services, e-commerce models such as Business-to-Consumer (B2C) have become more prevalent. B2C models involve using the Internet to market and sell products and services to consumers (Masoud, 2013). However, the study of customer behavior in management is still a relatively new field, and there are limited resources available to study and learn from the experiences of others. The importance and dimensions of paying attention to customer behavior have not yet received enough attention. One specific area where recommender systems can be useful is in the field of hedge fund investments. As Isinkaye et al. (2015) stated, "One of the hardest duties for hedge funds investors is choosing a perfect fund with simply the proper degree of risk." A combined Investment Recommender System (IRS) framework using Adaptive Neuro-Fuzzy Inference System (ANFIS) can be used to make investment recommendations, using a combination of Artificial Neural Networks (ANN) and fuzzy logic to adapt and optimize the system's parameters for more accurate and efficient recommendations. The problem of quantifying hedge fund risk is not only about the numbers, but also about how investors perceive what is "just right" (Tejeda-Lorente et al., 2019). With the increasing amount of data being generated every day from various sources, such as reviews, ratings, feedback, trading details, and investing data, the challenge is to extract meaningful insights from this vast amount of structured, semi-structured, and unstructured data, also known as Big Data. Recommendation systems are tools that can assist customers in finding products, items, or services that match their needs and preferences (Kanaujia et al., 2017). This dissertation aims to propose a novel framework for an IRS that utilizes adaptive neuro-fuzzy inference solutions and customized investment types. The proposed system uses machine learning and ANFIS to analyze the investor's past behavior and provide relevant and accurate investment recommendations. This new method of investment service customization can provide suitable and novel investment-type services and support the recommendation process in investment companies.

1.1. Research Problem

Investments can be made on various online platforms, and many of these platforms offer free consulting services to attract and retain investors. Online recommendations play an important role in online investment decisions, as they can be more effective than traditional offline expert advice (Senecal & Nantel, 2004). Recommender systems are particularly important in investment because they provide tailored advice to investors based on their individual needs and characteristics (Scheinbaum, 2017). In financial and investment companies, the customer relationship management system is a repository of investor information and experience and includes all investors' profiles. This information can be used to customize products or services for each unique investor, based on their personal needs. In investment services, the analysis of investor characteristics before receiving services or making investment decisions is crucial. Potential investors face a challenging task when deciding on investments, as there are many possibilities with different background information and options. Suitable recommendation systems can help them make better decisions based on their circumstances, reducing the risk of investment. Researchers are working to develop IRSs that consider demographic characteristics of investors such as age, gender, occupation, education, specialization, and other factors. This allows for a professional and effective approach to investment recommendations. A combined IRS framework using ANFIS is a system that utilizes a combination of ANN and fuzzy logic to make investment recommendations. ANFIS is used to adapt and optimize the system's parameters, making it more accurate and efficient in recommending investments. This type of system could potentially be used to help individuals or organizations make informed investment decisions based on market trends, historical data, and other important factors such as the financial situation, investor experience in previous investments, and their personality traits. According to cited sources, the rapid growth and increasing use of modern technologies have made it possible for managers of financial companies and real and legal investors to communicate more efficiently through the Internet and make the process of information exchange more efficient. Therefore, investing in the Internet using the appropriate computer application and devices is one of the major developments in this field in the world. Also, it needs to design recommendation systems based on investor experience and their feedback for proper investor service and they are important to companies in achieving their goals. It is stated that available systems that suggest to investors how to invest their funds are rare. Investors need a place to make investments and counselors have an especially key role for them. There is a need to know how to manipulate expenditures in an uncomplicated way for investors without high complexity. The benefit of the proposed recommender device is that it affords better pointers to an individual for saving, expenditure, and investment of their income which in turn maximizes their wealth. These systems act usefully if they are implemented based on potential investor experiences. The investment recommendation system is facing a significant challenge due to the unreliable nature of user data. The data collected from users may not be accurate, making it difficult to provide reliable investment recommendations. To tackle this issue, it is essential to implement advanced data analysis techniques to better understand user experiences. The proposed ANFIS-based recommender system is designed to address the challenges posed by unreliable data. By leveraging the power of ANFISs, the system can analyze data and provide accurate investment recommendations to individuals. The recommendations will help maximize their wealth by guiding them towards saving, expenditure, and investment options that are most suited to their needs and goals. Not only does this system benefit investors, but it also provides value to investment service providers. By reducing the time and cost involved in making investment decisions, it allows investment service providers to focus their resources on more important aspects of their business. The proposed ANFIS-based recommender system can play a crucial role in streamlining the investment process and providing reliable recommendations that benefit both investors and investment service providers.

One of the main problems in categorizing and clustering potential investors is the availability of data. The data used for categorization and clustering must be relevant, accurate, and up to date. Incomplete or inaccurate data can lead to incorrect categorization and clustering, which will negatively impact investment recommendations. Another problem is the diversity of potential investors. Different investors may have different investment goals, risk tolerance, and financial capacities, making it difficult to accurately categorize and cluster them. One of the main challenges in offering customized investment-type services is developing systems that are both accurate and adaptable. In addition, these systems must be able to effectively handle large amounts of data, process it quickly, and provide reliable results. Another challenge is ensuring that the customized investment-type services are relevant and useful to potential investors. The solution must be able to consider the unique characteristics of each potential investor and provide personalized investment recommendations. The main challenge in developing a combined recommender system is integrating different methods into a single solution. This requires a deep understanding of the different methods and how they interact with each other. The combined recommender system must also be able to effectively handle the large amounts of data generated by the system. Additionally, the recommender system must be able to provide relevant, accurate, and up-to-date investment recommendations to all kinds of investors. This requires constant monitoring and updating of the system to ensure that it remains effective and relevant over time. Using validated categorization and clustering, present research helps investment advisors and firms better understand their clients' investment preferences and behavior, thus solving related problems. This can be accomplished by using available data such as investment history, financial information, demographic information, and risk tolerance. By grouping similar investors together, investment advisors can make more accurate investment recommendations based on the characteristics of the group. also, by offering customized investment-type services using adaptive neural-fuzzy inference solutions provide personalized investment recommendations for different categories of investors. The use of these technologies allows for real-time analysis and adaptation of investment recommendations based on changes in market conditions and the investor's financial situation. This can result in more efficient investment strategies and improved investment outcomes for the individual investor. Furthermore, proposing a combined recommender system provide a comprehensive and cohesive investment recommendation solution for all categorized and clustered potential investors. The recommender system would bring together data from various sources, including financial information, investment history, and market conditions, to provide tailored investment recommendations based on the individual investor's specific needs and goals. The combination of multiple recommendation methods will provide a more robust investment recommendation solution, increasing the chances of successful investments and improved financial outcomes for the investor. To design and propose a combined recommender system framework for providing appropriate investment type recommendations, this research utilizes a series of conceptual stages, as outlined in Table 1-1. These stages were developed based on the initial findings of the research and aim to provide a solution to the problems faced in investment recommendation systems.

Table 1-1. Conceptual Stages of Proposing Combined IRS

Literature Review	Review the concepts of the research Recommender system ontology (Protégé Software) Review the past research: (Library studies)	
Data Preparation	Online Questionnaire: (Data Collection) ETL: (Data Cleaning) Coding: (Convert to Numeral Data)	
Input JMP	On category "Investment Type", for all proposed systems Six categories of "Potential Investor Type" for proposed combined system	
Machine Learning JMP Output	Three clusters "Investment Type" Clustering Six categories of "Potential Investor Type"	
Input ANFISs	One column three clusters Investment Type Six columns three clusters Potential Investors Type	
Proposing ANFISs	Data Preparation, Designing ANFIS, Proposing Structure Model (MATLAB Software)	
ANFISs Output	Investment Type Recommendations	

The proposed system provides an innovative approach to investment recommendation by combining the knowledge of experts and human intervention with mechanical analysis of data, to provide accurate and efficient recommendations. Not only does this system benefit investors, but it also provides value to investment service providers. By reducing the time and cost involved in making investment decisions, it allows investment service providers to focus their resources on more important aspects of their business. The proposed ANFIS-based recommender system can play a crucial role in streamlining the investment process and providing reliable recommendations that benefit both investors and investment service providers.

1.2. Research Objective & Research Questions

The main research question for this dissertation is "How can an ANFIS be utilized to propose an effective and efficient investment recommendation system?" The main objective of the dissertation is to propose a combined IRS using ANFIS to provide accurate and efficient investment recommendations for potential investors. To achieve this objective, the research will address the following specific sub-goals:

1. Categorization and clustering of potential investors based on available data to make accurate investment recommendations.
2. Offering customized investment-type services using adaptive neural-fuzzy inference solutions for different categories of potential investors.
3. Proposing a combined recommender system to provide appropriate investment type recommendations for all categorized and clustered potential investors.

The dissertation covers preparation, designing, and proposing the combined IRS to achieve the main objective and answer the main research question.

1.3. Chapterization of Dissertation

Chapter I: Introduction, this chapter provides an overview of the main research question and objectives of the dissertation. It includes a brief explanation of the motivation behind the research and the significance of the topic.

Additionally, it provides a clear outline of the structure of the rest of the dissertation.

Chapter II: Theoretical Framework and Literature Review, this chapter provides basic theoretical concepts and a comprehensive review of the existing literature on the use of ANFIS in investment recommendation systems. It covers previous research on the topic, including relevant studies and approaches used to develop IRSs. It highlights the strengths and limitations of existing methods and identifies areas for improvement in future research.

Chapter III: Research Methodology, this chapter explains the methodology used to achieve the main objectives of the research. It outlines the steps taken to develop the combined IRS using ANFIS. It includes a description of the data collection process, the categorization and clustering of potential investors, the development of the combined ANFIS model, and the evaluation of the proposed system.

Chapter IV: Experimental Results and Analysis, this chapter presents the results and analysis of the combined IRS. It includes a detailed explanation of the effectiveness of the model in providing appropriate investment-type recommendations for categorized and clustered potential investors. It also includes a comparison of the results with other existing methods and a discussion of the limitations and challenges faced during the development of the system.

Chapter V: Discussion and Conclusion, this chapter provides a comprehensive discussion of the research findings and their implications. It includes a discussion of the contribution of the research to the existing literature on the use of ANFIS in investment recommendation systems. It also highlights the limitations and challenges faced during the development of the system and provides suggestions for future research. Finally, it provides a conclusion to the main research question and objectives of the dissertation.

References: This chapter provides a list of the references cited throughout the dissertation. It includes a comprehensive list of the sources used in the literature review and the methodology sections.

Attachments: This chapter includes any relevant attachments or appendices that support the findings and results of the research. it may include investment questionnaire, generating rules process for ANFISs, and researcher's publications related to the dissertation that help to present the results of the study in a clear and concise manner. Additionally, it may include any additional information or data that may be useful to the reader but cannot be included in the main text of the dissertation.

In the following chapter, the theoretical aspects of the research will be explored and analyzed. Additionally, a review of previous studies that are relevant to the research topic will also be presented.

CHAPTER II THEORETICAL FRAMEWORK

This chapter presents the theoretical foundations of the research, including key concepts and a literature review of previous related research. The key concepts include user (customer or investor), company (investment services and products), and tool & technology (IRS). The literature review will provide an overview of previous research on IRSs, specifically those that use ANFIS and its applications in the field of investment. This will serve as a foundation for the development of the proposed combined IRS framework using ANFIS. Additionally, the literature review will provide a comprehensive understanding of the state of the art of research in this area and help identify the gaps in the existing research that this study aims to fill.

2.1. Investor (Customer)

According to the Encyclopedia of Health Care Management (2004), a customer is defined as "an individual or entity that is the recipient of a good or service made available by a supplier or provider, usually in exchange for something of value that is generally but not always monetary in nature." In the context of this research, the customer is the investor. The investor, as a customer, purchases services or investment products from investment companies or investment consulting firms. The Cambridge Dictionary (2023) defines an investor as "a person who puts money into something to make a profit or gain an advantage." Customer service refers to the provision of services to customers before, during, and after a purchase. The success of these interactions depends on the staff's ability to adjust to the personality of the guest (Buchanan, 2011). Customer service also encompasses organizational decisions regarding aspects such as product innovation and pricing. Organizations that prioritize providing good customer service typically invest more in staff training and actively solicit feedback from customers. From an overall sales process engineering perspective, customer service plays a key role in an organization's ability to generate revenue (Selden, 1998). A study by Watermark Consulting found that from 2007 to 2013, companies with better customer service outperformed their peers in terms of total shareholder return, posting a 26-point higher return than the S&P 500 (Tarnowska et al., 2020). Thus, customer service should be considered as part of an overall approach to systematic improvement. A positive customer service experience can also significantly impact a customer's overall attitude towards the organization (Teresa Swartz, 2002). Many companies have implemented customer feedback mechanisms that allow them to capture comments at the point of experience. This approach has been found to be beneficial as it allows organizations to improve their customer-provider relationship before the customer disengages, making it more likely that the customer will return in the future. Advances in technology have made it increasingly easy for companies to obtain feedback from their customers.

In commerce, the customer experience is created based on an interplay between a seller and a buyer during their interaction. This relationship is comprised of three parts: the customer journey, the brand touchpoints the customer interacts with, and the environments the customer experiences during their experience. As Verhoef, et al. (2009) argue, an ideal experience is achieved when all the customers' expectations are met during their experience. Customer experience also implies customer involvement at various levels, such as rational, emotional, sensorial, physical, and spiritual (Janakiraman et al., 2006). Customers can respond to a company in a direct or indirect manner. A direct response typically occurs when a customer initiates a purchase or receives a service from the company. Indirect responses, on the other hand, can include interactions such as advertisements, news reports, unscheduled encounters with sales representatives, word of mouth, or criticism (Meyer, 2007). The creation of direct interactions when customers buy, use, and receive services can be seen through customer interactions with retail employees. Indirect relationships, on the other hand, can take the form of unexpected interactions through a company's product representatives, certain services or brands, and positive advice - or even through "criticism, advertising, news, reports" and more (Andajani, 2015). The customer experience is created not only through the customer's values but also through the actions of the company providing the experience (Gentile et al., 2007). All interactions that customers experience before and after the purchase are part of the customer experience. In the retail industry, both companies and customers play a significant role in creating a customer experience (Andajani, 2015). The customer experience encompasses every aspect of a company. Customer feedback is data supplied by clients about their experience with a product or service. Its purpose is to reveal the degree of satisfaction and assist product, customer success, and marketing teams to identify areas for improvement. Companies can gather customer feedback proactively through polling and surveying customers, conducting interviews, or asking for reviews. Feedback can also be passively gathered by providing customers with an area in the product where they can provide comments, complaints, or compliments ("Customer Feedback Definition," 2023). Gartner (2019) believes that "the company's customer experience greatly influences their long-term exchange behavior and reflects the true drivers of loyalty". The question is, what are these outstanding experiences? Of course, these experiences can be in the form of customer feedback. Customer feedback is a crucial aspect of the customer experience, as it provides companies with valuable insights into the perceived value of their products and services. According to research by Forrester, companies that prioritize customer feedback see a significant increase in customer retention and loyalty. Additionally, customer feedback helps to measure satisfaction with a company's products and services, making it a vital tool for driving continuous improvement. There are various ways to collect customer feedback, such as through surveys, focus groups, and online reviews. The appropriate method will depend on the customer group, type of service or product, and the company's goals. Furthermore, it is important to ensure that the feedback collection process is tailored to the customer's needs and circumstances, and that it is accessible and easy to use. One issue that needs to be addressed is the proper and timely use of customer feedback. According to a study by McKinsey, companies that effectively leverage customer feedback see a 4-8% increase in revenue. Using intelligent systems for customer feedback analysis can help companies effectively utilize customer feedback to improve their products and services.

2.2. Investor Behavior

Understanding the investor's behavior as a customer is complex in the decision-making process when buying a product or service. Investor behavior is the result of various cognitive processes, social interactions, and social institutions, and the ability of investment firms to predict investor behavior is important. A deep understanding of investors' behavior creates more opportunities to predict and guide their behavior. The use of intelligent recommender systems is also an effective tool in predicting investor behavior. Several factors influence the analysis of investor behavior. One of these factors refers to the experience that the customer or investor gains in using the services or products of an investment company. This experience has had a significant impact on both his loyalty and attraction to new investors. One of the important dimensions of customer behavior is its social nature. Although the data collection data from investors about their behavior, the influence of other investors, social institutions, and social regulations governing society are also important in these behaviors. Therefore, the investors can only be understood and examined based on their relationships with other investors and in the framework of a larger social environment. "Customer engagement behavior can serve as a useful framework for classifying and segmenting customers, based on their propensity to engage and the types of engagement behaviors they display" (van Doorn et al., 2010). Of course, investors can be either individuals or organizations. Due to the differences between these two types of investors, there is a lot in common between them. In this research, potential investors are different people who answer "Investment Questionnaire" questions. These people are not necessarily current investors. In general, individuals in a community or organization can be considered potential investors. Here, the potential investor is a person, not an organization, and the present study is based on the demographic characteristics of people who answered the "Investment Questionnaire". Most investors do not act individually in decision-making and consider the opinions of different people in the investment process. Depending on the cognitive aspects of individuals and their characteristics, how they consult with different people in decision-making is different. People who are involved in decision-making may even come from a variety of backgrounds. Depending on the types of potential investors, they have different investment needs. Accordingly, the types of investments they choose are different. For example, income, savings, and jobs can be the most key factors in choosing the type of investment.

2.3. Investor Behavior in Investment Decision Making

Most investors do not make investment decisions in isolation and often consider the opinions of others in the process. In families, different individuals may be involved in various stages of the decision-making process. The level of involvement in the decision-making process often varies depending on the size of the investment. Additionally, individuals have varying cognitive aspects and characteristics that affect how they consult with others during decision-making. The individuals involved in the decision-making process may come from diverse backgrounds. In a family setting, the number and type of individuals involved in the decision-making process is usually consistent. Investor behavior also varies in different investment scenarios and the decision-making process. This behavior includes decisions on the type of investment, how and where to invest, review of different portfolios, and evaluation of services and products offered by investment companies. The decision-making method differs depending on whether the investor is making a new investment or extending an existing one. In simpler investment scenarios, the investor only needs to take a series of straightforward steps, but in more complex scenarios, more information and time is required to ensure the investment decision. It is worth noting that some investments, such as those in cryptocurrency, require a significant amount of information and technical and fundamental analysis. As Slovic (1972) notes, "the basic tenet of those in charge of helping the investor to make market decisions seems to be 'the more information, the better'." Different key factors play a role in investor behavior when making investment decisions. These factors include the opinions of specialists and experts in the field of investment, as well as the opinions of those who have invested in a particular field for the first time. Direct or indirect marketing through media and social networks can also have a significant impact on investor behavior. The performance of investment executives and agents in various fields is also a factor that affects the decision-making process. Ultimately, the opinions of investors who directly use the products and services of an investment company are the most important and effective factor. Investment companies must have a clear understanding of the needs of investors and the investment decision-making process to be successful. This includes recognizing the need, gathering information about the investment field of interest to the investor, evaluating different options, making investment decisions, and understanding significant issues in investor behavior post-investment. Adequate knowledge and understanding of investment companies helps them to design effective and successful portfolios for investment. As Christensen and Bower (1996) noted, "technological advances can exceed the required performance in a market, technologies that can initially only be used later in emerging markets can attack major markets and move incoming companies to victory over established companies." The design of investment recommending systems is one of these technical and effective advances in the investment market.

2.4. Investors (Costumers) experience & feedback

An investor's experience as a customer is the result of the investor's interaction with the company that assists an individual or organization in investing and uses the company's products and services in the investment process. This investment can be made directly by the investor or by an intermediary, and the experience gained can be during and after the investment process. According to Verhoef et al. (2009), "this interaction is made up of three parts: the customer journey, the brand touchpoints the customer interacts with, and the environments the customer experiences (including digital environment) during their experience. Good customer experience means that the individual's experience during all points of contact matches the individual's expectations." Gartner (2019) highlights the importance of managing the customer's experience, as it greatly influences their long-term exchange behavior and reflects the true drivers of loyalty. Customer experience implies customer involvement at various levels, such as rational, emotional, sensorial, physical, and spiritual (Janakiraman et al., 2006). The experience of investors may be gained directly or indirectly. In direct experience, the process of interaction starts with the investor, while in indirect experience, the investor gains the experience from news media in different contexts or through verbal interaction with other investors. According to Gentile et al. (2007), "customer experience is created by the contribution of not only the customers' values but also by the contribution of the company providing the experience." All the events experienced by customers before and after a purchase are part of the customer experience, and what constitutes customer experience is personal and may involve sensory, emotional, rational, and physical aspects to create a memorable experience. In the retail industry, both companies and customers play a significant role in creating customer experience (Andajani, 2015). The investor's experiences can be in the form of "investor feedback." Customer feedback exposes their degree of satisfaction and assists product, customer success, and advertising groups to recognize where there is room for improvement. Companies can gather customer feedback proactively via polling and surveying customers, interviewing them, or asking for reviews (Pendo.io Glossary). The investor feedback helps to measure the satisfaction of the investment company's products and services. Without investor feedback, no company can be assured of the value of the product or service it offers. The more importance given to investor feedback, the easier it is to retain the investor and the higher the investor's loyalty. It is possible to receive feedback in diverse ways, and the method of receiving investor feedback should be commensurate with their needs and conditions and should be at any time and in the simplest viable way with proper access. Another critical issue is the proper and timely use of investor feedback in the use of products and services. The use of intelligent systems is highly effective in the skillful and timely analysis of investor feedback.

2.5. Investment

The history of investment dates to around 1700 B.C. with the Code of Hammurabi (2008), the first known document that details the rights and relations of individuals with one another, highlighting important economic factors of the time. Investing refers to the allocation of funds with the expectation of achieving a benefit in the future. In finance, this benefit is known as a return. Generally, investors anticipate higher returns from riskier investments, while low-risk investments often yield lower returns (“Investment,” 2020). The term "investment product" encompasses a wide range of financial instruments, such as stocks, bonds, options, derivatives, and others, that individuals and institutions invest in with the goal of earning profits. The types of investment products available to individual and institutional investors may vary, but the basic profit motive is consistent across all of them (Cai et al., 2019). Investment can take many forms, including short, medium, and long-term changes in assets. This process may involve generating income from sales or dividends, rental properties, or a combination of various methods. Returns may also include positive impacts from foreign currency or losses resulting from fluctuations in foreign exchange rates. Investors typically seek higher returns from riskier investments. Low-risk investments, on the other hand, tend to have lower returns. As a result, investors, particularly beginners, are often advised to adopt a specific investment strategy and diversify their portfolios. Diversification has a statistical impact on reducing overall risk. "Investment Services" refers to the provision of investment advisory or investment management services, or any other related services, such as the management of an investment account or fund, providing advice on the investment or reinvestment of assets or funds, or otherwise acting as an "investment advisor" as defined by the Advisers Act, and performing activities related or incidental to these services. According to the Financial Conduct Authority (FCA) Handbook (2023), an "investment service" includes a wide range of activities related to financial instruments. These activities include, but are not limited to: (a) the reception and transmission of orders for one or more financial instruments; (b) the execution of orders on behalf of clients; (c) dealing on own account; (d) portfolio management; (e) the making of personal recommendations; (f) underwriting of financial instruments and/or placing of financial instruments on a firm commitment basis; (g) placing of financial instruments without a firm commitment basis; (h) operation of a multilateral trading facility; and (i) operation of an organized trading facility (FCA Handbook, 2023). In addition to these specific activities, investment services also encompass any services related to the management of an investment account or asset fund. This management may include providing investment advice or investment advice for cost or compensation, either directly or indirectly. Also, it can state that investment services include any investment services and related services provided by a contractor in relation to a program provided under an agreement.

2.6. Investment Recommender System (IRS)

The investment information system can be enhanced with the implementation of recommender systems, which are software tools and techniques that provide personalized suggestions for items that are likely to be of interest to a particular user (Burke, 2007; Resnick et al., 1994; Resnick & Varian, 1997). According to Liang (2008), recommender systems are a type of decision support system (DSS) that analyzes user behavior and proposes recommendations based on its results. They serve as a digital solution that supports financial investments by providing customized offers to investors according to their needs. Recommender systems are complex solutions with many potential applications. They are currently successful in facilitating access for online users to information that fits their preferences and needs in overloaded search spaces (Yera and Martin, 2017). These systems assist investors in determining the best items and services and make it easier to find favored objects. To implement the core function of a recommender system and identify beneficial items for the customer, it must be able to predict the utility of some items and services or at least examine the utility of some items and services, and then determine which items to recommend primarily based on this comparison. Special suggestion techniques are used to predict items and services based on the customer's needs or preferences. Knowledge-based systems provide items that are based on the knowledge gained from user behavior in the system. In this type of system, aspects of matching existing or potential items with customer needs and preferences are considered. A similarity characteristic estimates how much the customer's wishes match the items. In other words, it first finds the problem and presents the solutions to the problem. Multi-criteria rating recommenders, which provide recommendations by modeling a customer's utility for an item as a vector of scores along with several criteria, are a popular area of research. This approach uses multi-criteria rankings for calculating the rating predictions and producing recommendations. Cross-domain recommender systems handle the consumer desire aggregation and mediation techniques for the cross-system personalization problem in customer modeling, as an attainable solution to mitigate the cold-start and sparsity problems in recommender systems, and as a practical application of knowledge transfer in machine learning. As previously mentioned, an information system is any system that stores a large amount of information. These systems can be equipped with recommender systems, such as IRSs. IRSs are a type of investment DSS that analyze investor behavior and propose suitable investments for the customer or new investor based on the results. In these systems, the program uses techniques and methods of the recommender system to meet the information needs of the customer in investing. A recommender system is designed for the user, and customer or investor behavior plays a key role in evaluating the system. The behavior of the user (investor) is considered as the main feature of the investor when using the recommender system. A recommender system, or a recommendation system, is a subclass of information filtering systems that seeks to predict the "rating" or "preference" a user would give to an item. These systems are used in a variety of areas, such as Netflix, YouTube, Tinder, and Amazon. One popular method of designing a recommendation system is using a popularity-based approach. This solution can be identified by using customer experience, which can be the internal and subjective response investors have to any direct or indirect contact with a company. Investment services include making, organizing, and managing investments, such as stocks, bonds, and cash equivalents like cryptocurrencies. To design an effective IRS, it is important to consider customer feedback. One suitable solution is to use fuzzy neural inference systems. Knowledge-based recommendation systems can be designed using these systems, where investor feedback can lead to valuable information. The system designer can use experts' knowledge about this feedback to make the necessary changes to the recommender system and generate new rules. This way, the system can be updated, and dynamic based on investor feedback. Therefore, it is crucial to design recommendation systems based on customer experience and feedback for proper customer service. These systems are particularly important for companies in achieving their goals.

2.7. Recommender System and its Ontology

In the field of investment, investment advisory systems can play a key role in providing valuable information and in the development of the system and the company. These systems can provide a large amount of investor profile data, data from investor behavior, data from investor experiences and feedback, and information related to the capital market. They can also evaluate and filter all data stored and available in the system. General requirements for a system ontology include being coherent, comprehensive, consistent, concise, and essential. Lee et al. (2006) presented an ontology-based product-recommender system that can be implemented on a practical ontology system powerful enough to assist in filtering strategies that exploit the semantics formalized in an ontology to link items and their features to time functions. They stated that building an accurate profile for the user plays an essential role and that an ontology works very well to characterize the users' profile involved in the process of generating recommendations. Orciuoli and Parente (2017) proposed a Context-Aware Recommender System to assist indoor shopping by localizing shoppers and providing them with suggestions on finding suitable offerings related to products that meet their Wishlist. The ontology engineers often try to determine which are the highest categories and how they form a classification system that provides a comprehensive classification of all entities. These categories usually include the main entities, their properties, subsets, and the relationships between them. The concept of ontological dependence determines whether the entities of a group exist at the most basic level. Differences in the ontology are often about one entity belonging to a particular group and how they relate to other entities (Hofweber, 2021). Today, web platforms provide a powerful gateway for investment, especially in the field of cryptocurrencies, and providing fast and efficient services to investors is of particular importance. Due to the role of IRSs in this field, their improvement based on semantics is highly effective. This development can be based on a filtering approach aware of the time, knowledge, and needs of the investor to attract potential investors. Figure 2-1 illustrates the relationship between the basic sections of the ontology with the recommender system. An information system is a system that stores a large amount of information, and it is equipped with recommender systems, which are a subset of information systems that can interact with users and provide product recommendations. Recommender systems use techniques and methods to meet the information needs of users. These systems interact with other information systems and users to receive information and send output as recommendations. Ontology development is a crucial aspect of recommender system design. There are several basic rules related to the design of ontologies, including the determination of the ontology development methodology, ontology language, and ontology development environment (tool). The ontology development process is usually repetitive and iterative as it requires consensus among users. In this study, the researcher adopted the Menthology approach (Fernández, et al., 1997) for ontology development. This methodology includes the stages of specification, knowledge acquisition, conceptualization, integration, implementation, and evaluation, with emphasis on the evaluation stage. The researcher in this study focused on the conceptualization of the recommender systems and their properties. The ontology was implemented in Protégé 5.5, and figures and visualization were prepared in OntoGraf and OWLViz. The main elements of the recommender systems' properties ontology, their relations, and descriptions are presented. The general recommender systems' ontology includes the following objects: Axiom count 287, Logical axioms count 98, Declaration axioms count 103, Class counts 87, Object property count 9, Data property count 3, Individual count 4, Annotation Property count 2, and Sub Object Property Of 2.

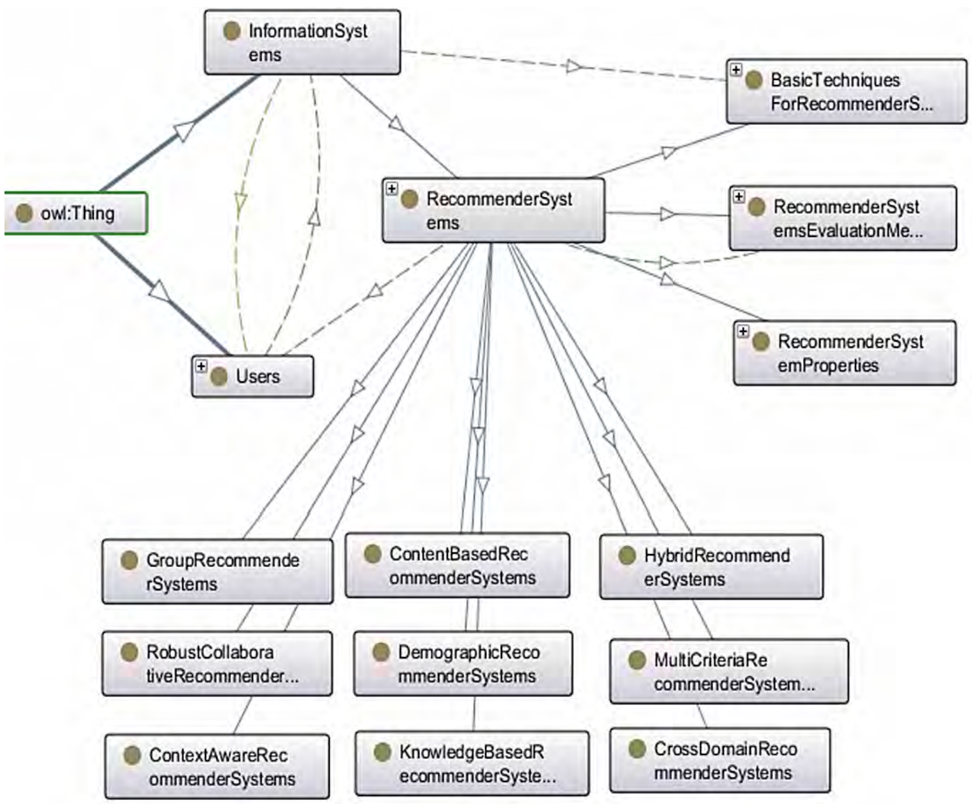


Figure 1-1. General ontology of recommender systems (OntoGraf)

Investors often seek out information to guide their investment decisions. A combined IRS framework using ANFIS can assist in this process by utilizing a combination of ANNs and fuzzy logic to make recommendations. This type of system can be used by individuals, organizations or even investment robots to make informed decisions based on market trends and historical data. Evaluation is a crucial aspect of this system, as it is done to assess the system's performance in providing relevant and accurate recommendations. IRSs assist investors in making informed decisions by identifying suitable investment options that align with their preferences. These systems analyze market trends, historical data, and the investor's needs and preferences to predict the usefulness of potential investments. Various recommendation techniques, such as knowledge-based and case-based methods, are employed to recommend investments that meet the investor's specific needs and objectives. Multi-criteria ranking recommenders, which provide recommendations by modeling user preferences as a ranking vector with multiple criteria, are a popular focus in many studies of IRSs. Such systems help investors make smarter investment decisions by providing them with a tailored selection of options that meet their unique needs and goals. This proposed method employs a multi-criterion ranking approach to calculate predictions and generate investment recommendations. The cross-domain IRS addresses the challenge of aggregating and mediating investor preferences for personalized cross-system modeling, providing a potential solution to the cold-start and sparsity issues commonly encountered in IRSs. Furthermore, it represents a practical application of knowledge transfer in the realm of machine learning. One specific type of recommender system is the context-aware recommender system, which is of particular importance in the field of investing. There has been extensive research conducted in context-aware recommender systems, including the fundamental understanding of context modeling and the development of various recommender algorithms. For example, these systems can analyze textual information provided by the customer or investor and use it as a valuable method for creating personalized investment advice. Additionally, research has also been conducted in the field of evaluation systems, specifically in the context of context-aware systems. Here, the researcher have conducted a thorough evaluation of various proposed approaches and techniques, taking into consideration their benefits and limitations. Specifically, in the realm of investment, this recommender system can be applied to portfolio design, goal-oriented frameworks, and approaches to assist in investment selection and development, as well as providing context-aware advisory capabilities. The system learns to offer investors items that are like those they have previously expressed a preference for. The similarity of the recommendations is calculated based on the characteristics of the compared items. For example, if a customer has given a positive rating to a certain type of investment, then the system can recommend other investments of that same type. Additionally, the recommender system employs classic content-based techniques to match the user's profile specifications with the properties of the items. In most cases, item attributes are extracted from item descriptions as keywords, but semantic indexing techniques can also be used, which involve indexing concepts instead of keywords. When deciding on a recommendation approach for investment, certain factors are taken into account. Some systems propose a group model of desired features for a jointly planned application and aid a group of investors in reaching a consensus. These recommender systems utilize a combination of techniques, such as a hybrid system which combines various methods to address the limitations of one technique with the strengths of another. These systems also complement each other. IRSs recommend items based on the demographic characteristics of the investor or customer, with the assumption that different recommendations should be made for various categories of people such as age, gender, occupation, level of education, amount of capital, and other factors. It can be said that a recommender system is tailored to the user, with an IRS being investor-based. The behavior of the investor, whether an individual or legal entity (organization), plays a crucial role in evaluating the system. Figure 2-2 illustrates the basic techniques for recommender systems. To effectively implement the core function of an IRS, identifying useful items for the client or investor, the system must be able to predict the profitability of certain items for investment and compare their usefulness against other options. Based on this comparison, the system then makes recommendations for investment based on the potential investors' group.

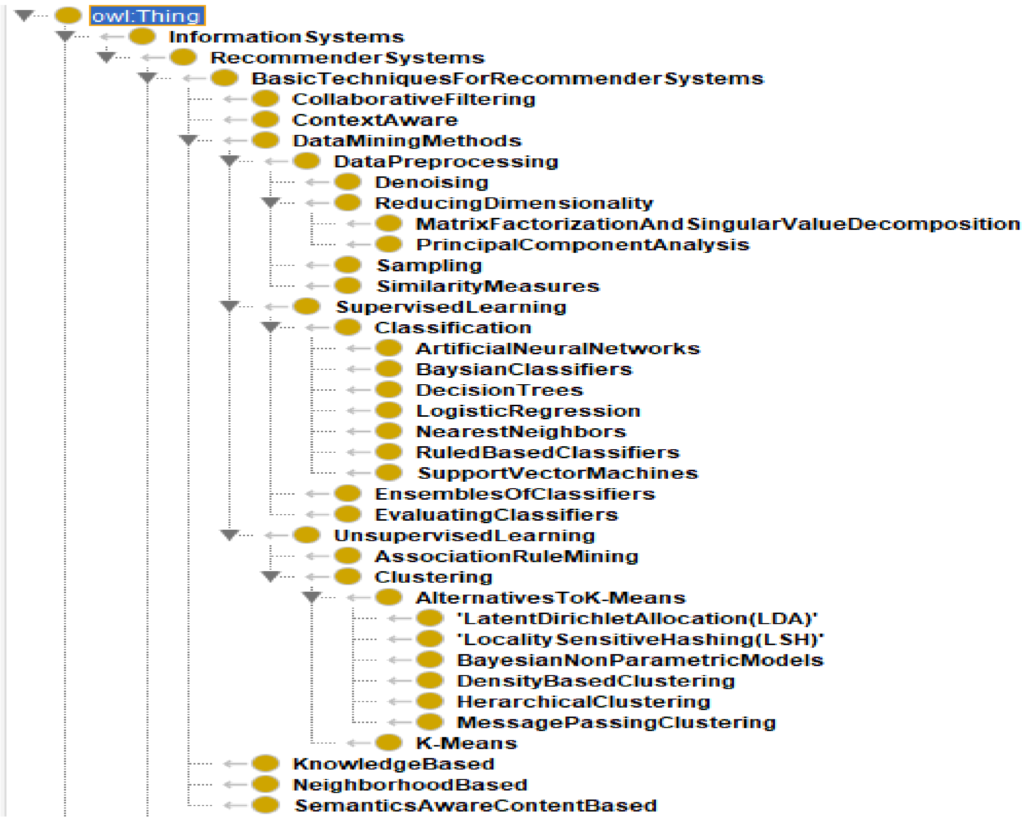


Figure 2-2. Basic techniques for recommender systems

Different methods of recommendation are employed to predict the most suitable items for an investor based on their specific needs and preferences. Figure 2-3 illustrates the general ontology of recommender system's types. The selection of a recommender system type is based on its properties. It is essential to evaluate the system at various stages and intervals throughout its lifecycle, for various purposes. The figure presents an OWL visualization of the properties of recommender systems. User preferences play a crucial role in determining the effectiveness of the recommendations provided by the system. The evaluation process should assess how well the suggestions align with the preferences of the system's users, and how much they aid in decision-making. Additionally, the system's ability to facilitate the discovery of preferred items should be evaluated.

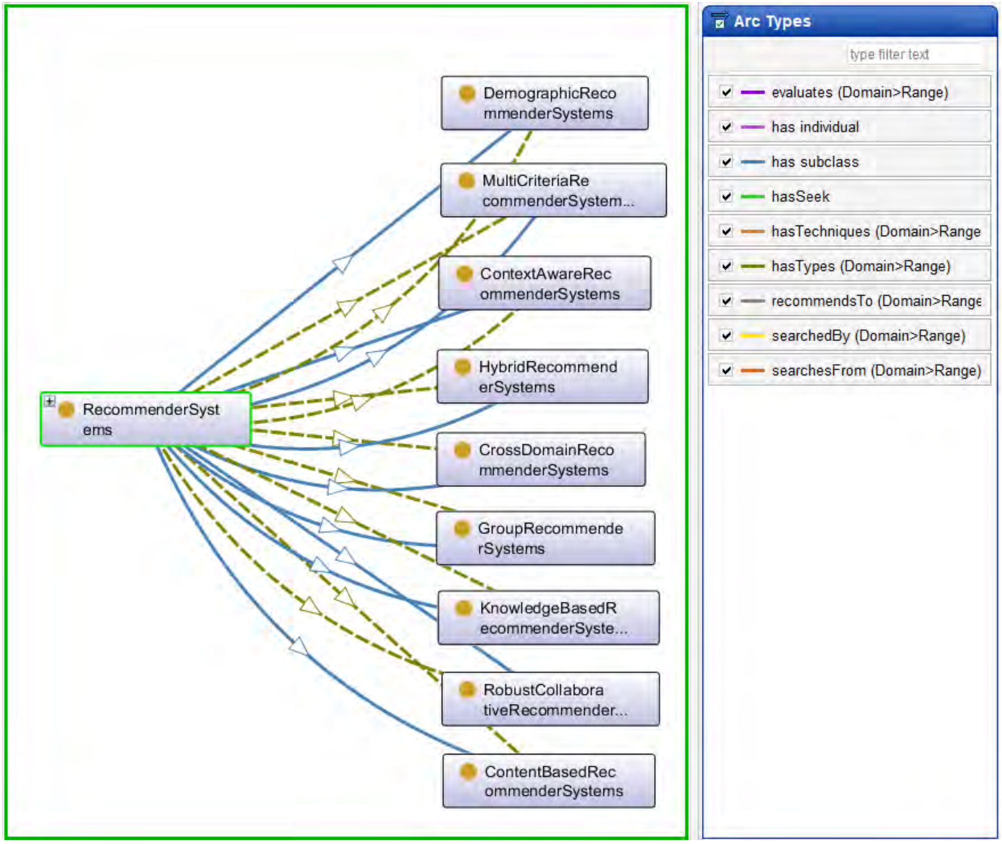


Figure 2-3. General ontology of recommender system's types

In the implementation of recommender systems, it is essential to identify useful items for users based on their current and potential needs. The value of recommending an item to a user should also be considered by the system's ability to predict the usefulness of the proposed items. The minimum usage of the items is compared and then items that align with the user's preferences are recommended.

In knowledge-based recommender systems, the features of the recommended items are considered for both potential and actual users. These features are based on the specific domain knowledge of the items and are chosen to match the user's preferences as much as possible. The recommendations are then evaluated based on their usefulness to the users. The goal of the recommender system is to provide utility to the users. The behavior of the user plays a crucial role in the evaluation of a recommender system. It is important to consider the individual needs and preferences of users as separate classes, as these can greatly impact the effectiveness of the system. The goal of a recommender system is to meet the needs of the user in a timely and efficient manner. When this is achieved, the user will have a greater sense of confidence in the system. Confidence is therefore an essential property that must be considered when designing a recommender system. Knowledge-based recommender systems operate by using past cases as a reference. The degree of compliance with the recommendations is evaluated based on how well they meet the needs or preferences of the user. One way to measure the adaptability of a recommendation is through serendipity, which examines the amount of information required before making a recommendation. Another way to evaluate adaptability is to compare the compatibility of the recommendations with the user's personalized preferences in their profile. The assurance of a recommendation, or the level of trust in the system's predictions, is also an important aspect to consider. Generally, the more adaptable a system is, the greater the level of trust in its recommendations. Coverage is another important property to consider. It refers to the range of issues and items that the system recommends to the user. It can vary depending on the goals of the system and can include the proportion of items recommended to the user, as well as the number of users that the system is able to serve. One sub-property of coverage is the "cold start" problem, which refers to the difficulty that a system may have in recommending items to new users or when the preferences and needs of existing users change over time. The coverage property of a recommender system is often hindered by a cold start, where the system must restart from scratch and upload updated data to provide appropriate recommendations. This can also be measured by the ratio of users or their interactions with the items recommended by the system. In addition to coverage, variety is another important property to consider. Both items, users and the recommendations provided to users should be diverse to cater to the diverse needs of the system's users.

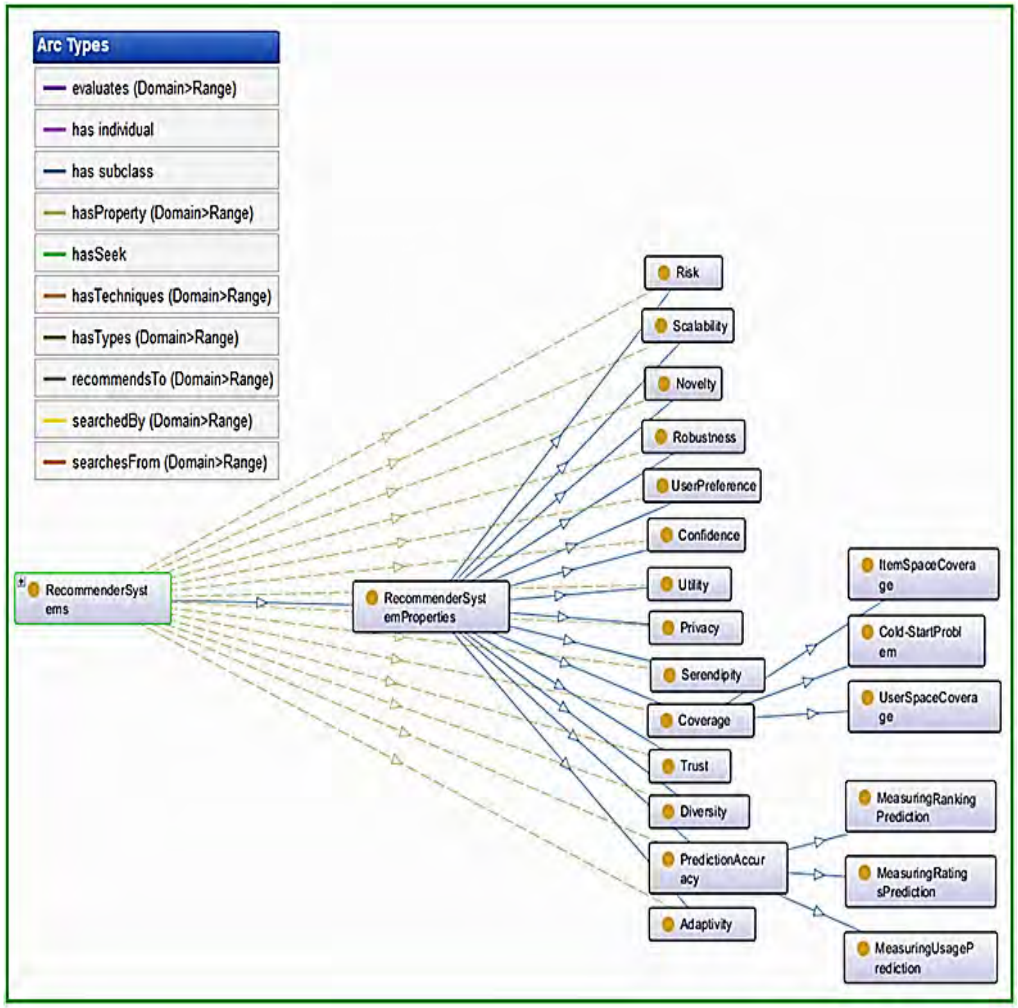


Figure 2-4. General ontology of recommender systems' properties (OntoGraf)

Furthermore, predictive accuracy is a key property of recommender systems, as seen in Figure 2-4. The accuracy of the predictions must consider the potential needs of users and can be improved using feedback from users on the recommendations provided. In this case, the prediction will be based on user preferences, with an emphasis on ranking, rating, and usage of the recommendations provided by the recommender system. The more accurate these metrics are, the more accurate the overall prediction will be. As technology and data continue to evolve, it is essential for recommender systems to be innovative to provide user-friendly services. This means not only being unique, but also utilizing algorithms that are robust enough to minimize errors in making recommendations.

In assessing the robustness of the system, the performance is evaluated under different conditions to ensure that it performs well in all scenarios. Additionally, it is important to avoid over-recommendation, as excessive recommendations can become tedious for users and negatively impact their experience. These recommendations may not always align with the user's current needs and can become a source of annoyance for the user. To address this issue, recommendation systems must incorporate serendipitous items, which are items that are unexpected but still relevant to the user's interests. However, as Kotkov et al. (2016) point out, incorporating serendipity into a recommendation system poses certain challenges. It requires careful selection of appropriate objectives and algorithms. Despite these challenges, serendipity can be a measure of the success of a recommendation system. Another important aspect of recommendation systems is privacy. It is crucial to protect users' information and preferences from third-party access. Additionally, the risk properties of the recommendations made by the system can be a decisive factor in determining whether a user continues to use the system. In some cases, even a recommendation from the system may be associated with potential risk. With the abundance of fake information, it is important to consider the strength and stability of recommendations in the presence of this type of data. The scalability of the system is also an important consideration, especially as the amount of data increases. Increasing the number of recommendations can also be an effective way to scale the system.

2.8. Decision Making based on the Fuzzy Inference

Fuzzy logic is a powerful tool that can be applied in a wide range of decision-making processes, from data processing to data analysis (Srivastava et al., 2013). Decision-making is a fundamental human activity that can take many forms, from individual choices to social decisions. Researchers have long recognized the importance of decision-making and have conducted numerous studies to improve our understanding of the process. At its core, decision-making involves choosing one option from among multiple possibilities. This function is critical in many different contexts. Fuzzy logic is a set of rules that are expressed in natural language and translated into mathematical equivalents by fuzzy systems. This conversion makes it easier to design systems and helps computers to process information more effectively. Fuzzy logic can accurately represent the behavior of a system in the real world. One of the most significant advantages of fuzzy logic is its ability to handle incomplete data, which is a common problem in decision-making. Making decisions based on incomplete information can lead to poor outcomes, so fuzzy logic can be a valuable tool for overcoming this challenge. ANFIS is a powerful method for adapting a set of input-output data to a fuzzy system. This approach allows for the use of fuzzy decision theories, which are based on FISs. Decisions can be modeled using FISs, making ANFIS a promising technique for improving decision-making in a variety of contexts.

2.9. Literature Review

According to Edmondson and McManus (2007), there are two main approaches to conducting a literature review. The first approach is to review existing literature and then develop a conceptual model that integrates and expands upon previous research. The second approach is to review the current literature on an emerging subject and use it to develop a new conceptual model (Webster & Watson, 2002). In this research, a combination of both approaches is used. A review of the scientific literature suggests that various types of recommender systems have been studied in the past. Both quantitative and qualitative research studies have highlighted the importance of implementing combined recommender systems. In this study, the literature review is divided into several main sections, each related to the research topic. These sections include Investment & Recommender Systems, ANFIS, Customer Service & ANFIS, and Investment & ANFIS (as seen in Table 2-1).

Table 2-1. Literature review sections

Literature review Sections (Qualitative Review)	Investment & Recommender Systems
	ANFIS
	Customer Service & ANFIS
	Investment & ANFIS
Quantitative Review Analysis	Brief systematic review on Investment Recommender Systems
Comparison & Innovation	Proposed Recommender System vs Existing Recommender Systems

2.9.1. Investment & Recommender Systems

According to Paranjape-Voditel and Umesh (2013), a stock market portfolio recommender system based on association rule mining was proposed. This system analyses stock data and suggests a ranked basket of stocks to support stock market traders, individual buyers, and fund managers in their investment decisions. The objective of this recommender system is to suggest an investment in a team of equity stocks when strong evidence of potential profit from these transactions is available. The system finds the correlation between stocks and recommends a portfolio. In 2017, a collaborative filtering-based recommender system for financial analysis using Apache Hadoop and Apache Mahout was proposed. The large amount of data involved in this study required the use of the Apache Hadoop framework for distributed processing. The researchers used collaborative filtering and Apache Mahout to analyze the data and implement the recommender system (Kanaujia et al., 2017). In a study by Hernández et al. (2019), the current state of financial technology was evaluated to design a new recommender system. The proposed system is a social computing platform that utilizes Virtual Organizations to improve the user experience in funding recommendations. The system utilizes data on the user's characteristics, asset classes, profitability, historical market data, and economic information found in the media. In a separate study, Tejeda-Lorente et al. (2019) presented a new recommender system that considers the risks associated with individual hedge funds. The system considers various factors such as current yields, historical performance, and diversification by industry, and uses fuzzy linguistic modeling to capture the preferences of traders. To demonstrate the effectiveness of their approach, the study first profiled over 4,000 top hedge funds based on their composition and performance, and then created simulated investment profiles to test the system's recommendations. In the field of financial technology, Hernández et al. (2019) proposed a multi-factor IRS that utilizes a social computing platform based on virtual organizations. This platform aims to improve the user experience in the investment recommendation process by utilizing a recommender agent that is responsible for the case-based reasoning system. The data used in the system includes information on user characteristics, asset classes, profitability, interest rates, historical stock market data, and financial news from various media sources. Faridniya and Faridniya (2019) presented a model that employs data envelopment analysis for resource allocation and investment type selection. Their study was a case study of the Social Security Organization (SSO) in Iran, and the results showed that the current investment strategies employed by the SSO were leading to bankruptcy. They concluded that a change in investment strategy was necessary to avoid this outcome. Sulistiyo and Mahpudin (2020) examined the demographic factors that influence the choice of investment type. They conducted a study on amateur golfers in Karawang City and found that investment type is divided into two categories: real estate and financial assets, and demographic factors play a role in the choice of investment. Their research showed that demographic factors affect the choice of investment type. Among the factors studied, five factors had the most impact. They included gender, occupation, education, number of family members, and income. Their findings showed that age did not affect the choice of investment type. The research by Tarnowska et al. (2020) investigated the impact of demographic factors on the choice of investment type. They found that five factors had the most significant impact, which were gender, occupation, education, number of family members, and income. However, their findings also indicated that age did not play a significant role in affecting the choice of investment type. The system addressed several key issues, including providing a favored framework for managers to make decisions on which actions are most likely to have the greatest impact on the internet promoter rating, using data mining techniques that allow investors to "learn" from the experiences of others without sharing proprietary information, thereby increasing the system's power, Supplementing traditional text mining options by allowing users to view specific, anonymous feedback related to individual customers. This can provide valuable insights into steps that can be taken to improve customer satisfaction and, offering a sensitivity assessment feature that allows managers to weigh different actions and determine which ones are likely to have the greatest effect. Kovács et al. (2021) examined the use of a two-stage clustering method for identifying investment patterns of potential retail banking customers, which can help improve marketing policies and strategic planning in the industry. Thompson et al. (2021) proposed a system that utilized clustering and association rule mining techniques to identify patterns in the preferences of individual clients and make personalized recommendations. Li et al. (2021) also used a similar approach by combining clustering and collaborative filtering to provide personalized recommendations to users. Both studies formed clusters of similar clients or users based on their preferences and then used the respective techniques to make personalized recommendations. For instance, Thompson et al. (2021) applied clustering and association rule mining to identify patterns and form clusters of similar clients. They then used the clusters to make recommendations based on the preferences of each cluster. Similarly, Li et al. (2021) used clustering to form clusters of users based on their preferences and then used collaborative filtering to make personalized recommendations. Pemisindo (2020) and Koosha et al. (2022) also employed a combination of decision trees and clustering to make recommendations to users. These studies used decision trees to determine the relevant attributes of the users and then used clustering to form clusters of similar users. Overall, clustering-based systems have been shown to be effective in making personalized recommendations based on user preferences. The combination of clustering with other techniques such as association rule mining or collaborative filtering can further improve the accuracy and relevance of the recommendations.

2.9.2. Adaptive Neural Fuzzy Inference Solution

In a study by Siddiquee et al. (2015), a film recommendation system was developed using the Fuzzy Inference System (FIS) and ANIS. The study utilized two similarity criteria, one based on the selection of similar users and the other on matching similar genres of user-rated movies. Four different techniques were used to calculate similarity, with FIS and ANFIS being utilized in the decision-making process. The results of the study showed that ANFIS performed better than FIS in most cases when the Pearson correlation criterion was used to calculate similarity. Yera and Martin (2017) conducted a literature review to identify common research topics and research gaps in the use of fuzzy tools in proposed systems. They focused on articles available on the Thomson Reuters Web of Science, analyzing them based on key features, evaluation strategies, datasets used, and application areas. In another study, Asemi et al. (2019) designed an ANFIS algorithm to evaluate Automatic Speech Recognition (ASR) systems as a case study in MVML-based ASR. The proposed algorithm was used to measure the performance of two dysarthric ASR systems based on MVML and MVSL active learning theories. The results of the study showed the effectiveness of the developed method. In their study, Szafranko et al. (2022) utilized ANFIS to aid in the generation of expert opinions and assessment of variations in building design. This approach proved effective in handling large volumes of data, highlighting the potential of ANFIS in the field of building design evaluation.

2.9.3. Customer Service & Adaptive Neural Fuzzy Inference Solution

In their study, Isakki et al. (2011) employed data warehousing and data mining technologies to analyze customer behavior, to create customer profiles. They found that this approach allowed them to provide the best service model according to customer orientation and develop effective marketing strategies. Zahin et al. (2013) conducted a comparison of different forecasting techniques for electricity generation. Using a dataset of five years of annual electricity demand in Bangladesh, they used Year, irrigation season, temperature, and rainfall as input parameters in ANFIS and load demand as the output. They also used another artificial intelligence method, ANN, to validate the results. The researchers found that ANFIS had superior predictive power for generation, as determined by various error measurements, when compared to both ANN and seasonal forecasting.

2.9.4. Investment & Adaptive Neural Fuzzy Inference Solution

Erdogan et al. (2016) proposed an alternative model for predicting the failure of enterprises by conducting a study on a sample of 356 business enterprises listed on the Istanbul Stock Exchange. The firms were classified into three levels using 18 parameters each, and the study employed differential analysis and ANFIS methods. The findings of this study support the creation of a balanced financial environment and aid in determining suitable enterprises for credit loans. Rajab and Sharma (2017) conducted a program-based research study in the field of Neural-Fuzzy Systems (NFS) by reviewing research articles published in prestigious international journals and conferences between 2005 and 2014. The study identified finance, marketing, distribution, business planning, information systems, manufacturing, and operations as the core business applications of NFS during this period. With the abundance of customer data received from various sources, it is crucial to classify potential investors based on their characteristics and experiences. Sedighi et al. (2019) proposed a new integrated approach for accurate stock price forecasting using ABC, ANFIS, and Support Vector Machines (SVM). The model outperformed other methods in accuracy and quality, and can be used to identify stock price trends, making it an innovation in algorithmic trading. The study used the 50 largest companies in the U.S. Stock Exchange from 2008 to 2018 for the evaluation. Hussain et al. (2022) proposed a Clustered Induced Ordered Weighted Averaging ANFIS for fuzzy time series prediction of cloud Quality of Service dataset. The method employs an intelligent sorting mechanism, fuzzy clustering, and ANFIS structure to provide understandable rules for cloud stakeholders while dealing with uncertain occurrences of data. The proposed CI-ANFIS model outperformed all current techniques, demonstrating its potential applicability in various complex prediction problems. Birim et al. (2022) utilized artificial learning applications to predict cryptocurrency return rates considering the complex and unstable financial system. The ANFIS approach was used to train the network using PSO for Ethereum, Bitcoin, and Tether. The ANFIS-PSO approach yielded strong results in cryptocurrency rate of return estimation with RMSE and MAPE used as performance indicators. A review of the research background shows that various machine learning methods and algorithms have been used for classification and clustering of diverse factors. Additionally, different studies have been conducted using the ANFIS technique in various subjects. However, there is a lack of research on using an ANFIS-based combined recommender system to recommend investment services based on customer experiences. Despite the potential of modern technologies, studies indicate that independent research in this area has not been undertaken.

2.9.5. Brief systematic review on Investment Recommender Systems

On March 4th, 2023, a search was conducted in Scopus, utilizing the following formula, to retrieve relevant documents on the research topic of interest:

(TITLE **(** *"recommender"* **OR** *"recommendation"* **OR** *"decision"* **)** **AND** TITLE **(** *investment* **AND** *system* **))** **AND** **(** LIMIT-TO **(** PUBYEAR **,** *2023* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2022* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2021* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2020* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2019* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2018* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2017* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2016* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2015* **)** **OR** LIMIT-TO **(** PUBYEAR **,** *2014* **))**

After the search, 154 documents were found, out of which 44 full records were imported into Zotero. Figure 2-5 displays the distribution of the documents by subject area, indicating that most of them belong to the field of computer science.

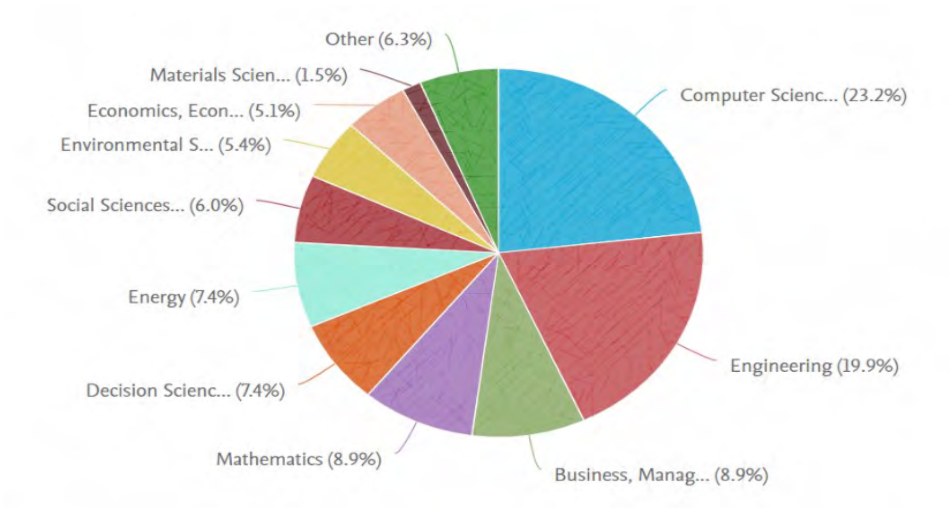


Figure 2-5. Documents by subject area in IRS by Scopus Analysis Tool

According to Figure 2-6, the documents in IRS are categorized by their funding sponsors. Most of the documents are supported by the National Natural Science Funds of China, followed by the Fundamental Research Funds for the Central Universities. The research projects supported by the Fundamental Research Funds for the Central Universities mainly focus on books and symposium papers.

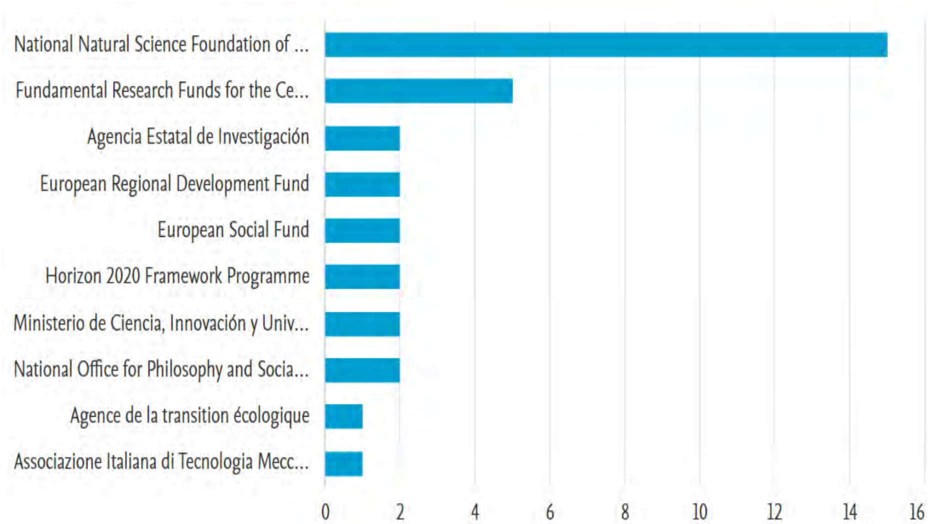


Figure 2-6. Documents by Funding Sponsor in IRS by Scopus Analysis Tool

After conducting an analysis, it was determined that there are 3947 data repository files in Mendeley Data related to "recommender" or "recommendation" or "decision" and investment and system, which are available in various formats such as Dataset (1267), Tabular Data (639), Document (366), Collection (285), Text (212), Software/Code (109), Image (57), File Set (41), Slides (26), Video (26), Audio (2), Geospatial Data (1), and others (2342). The author and index keywords from Scopus documents were analyzed using Voyant, revealing a total of 2640 words and 1509 unique keyword forms. The Vocabulary Density was calculated as 0.572, and the Readability Index was 68.616. Among the most frequently occurring words in the corpus were investments (94), decision making (60), investment decisions (39), decision support systems (36), and multi (26). To provide a comprehensive overview of all keywords, a Cirrus was created and included in Figure 2-7, which displays a keyword cloud view of the most commonly occurring Author & Index keywords. Additionally, Figure 2-8 shows the co-occurrence of the retrieved keywords in Scopus (154 documents) using VosViewer.

Figure 2-7. Keyword cloud view of the most frequently occurring Author& Index keywords by Voyant

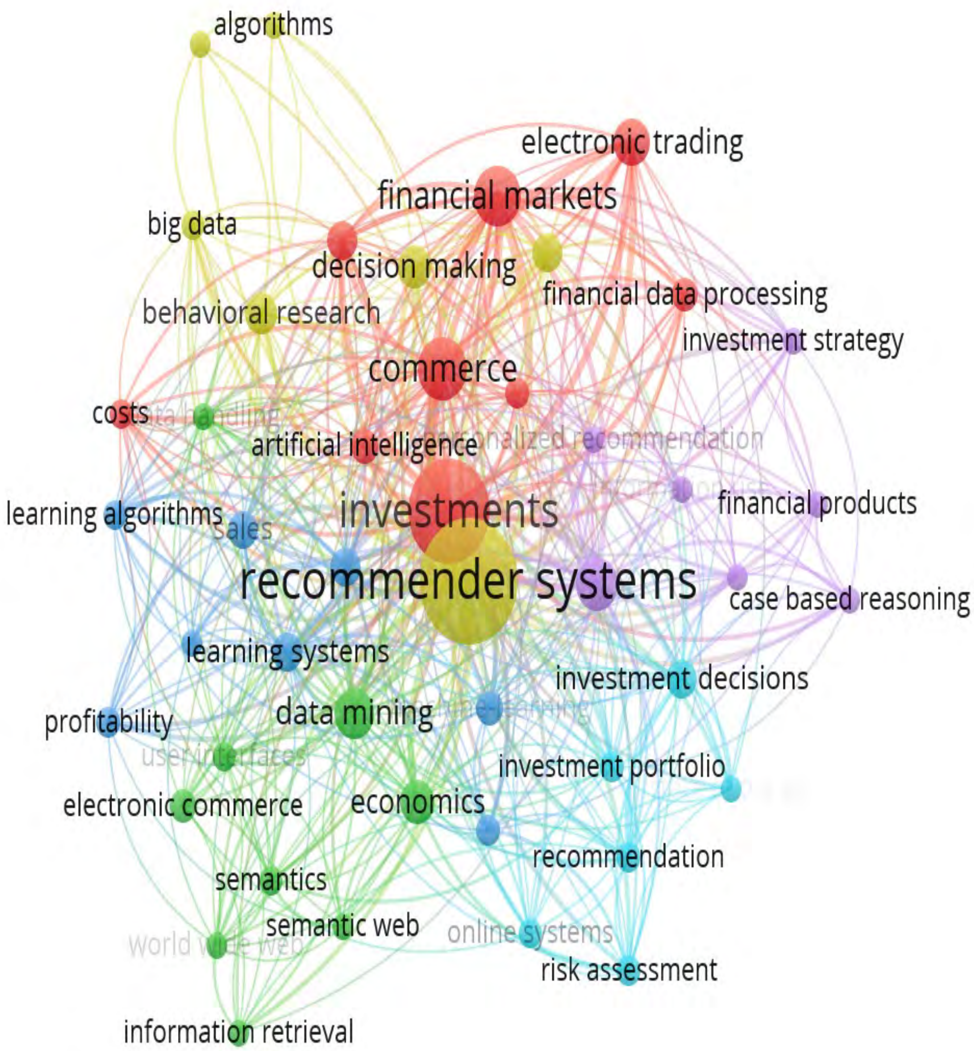


Figure 2-8. Co-occurrence of the keywords in IRS by VosViewer

VosViewer was used to investigate co-authorship patterns in a set of 154 documents retrieved from the IRS database. Figure 2-9 displays the co-authorship relationships among the documents, with most of them being attributed to Wang Y. The authors of these articles represent a diverse range of countries, including the United States, China, Turkey, and Poland, among others.

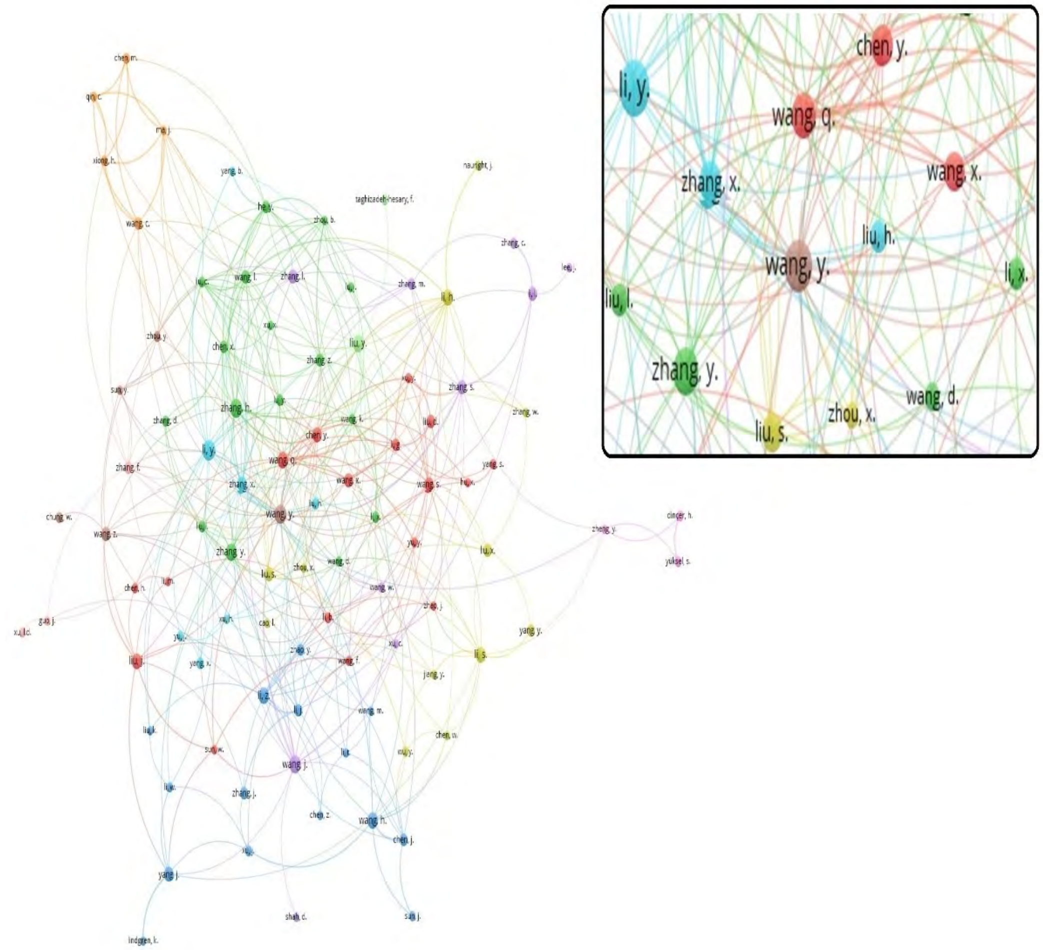


Figure 2-9. Co-authorship in IRS by VosViewer

Upon reviewing previous studies, it is evident that several topics gained popularity during the period from 2019 to 2023. These topics included Artificial Intelligence (AI) and Machine Learning, Blockchain Technology, Renewable Energy Systems and Investments, Decision Support Systems and Models, Digital Transformation and Industry 4.0, Sustainability and Green Investments, Big Data Analytics and Predictive Modeling, Cybersecurity and Risk Management, Peer-to-Peer Lending, and Alternative Financing Models, as well as Real Estate Investments and Portfolio Optimization. Table 1 displays the most frequent subjects based on the years 2019 to 2023 as per the IRS:

Table 2-2. Most frequent subjects in IRS (2019-2023)

Subject	**Frequency**
Decision Support System	28
Renewable Energy	9
Power Systems	8
Artificial Intelligence	6
Real Estate	4
Multi-Criteria Decision Analysis	3
Peer-to-Peer Lending	3
Machine Learning	2
Carbon Neutrality	2
System Dynamics	2

The most discussed topics include real estate investment, decision support systems, investment decision-making, and renewable energy systems. In 2020, decision support systems and real estate investment were frequently discussed, while in 2021, the focus shifted to renewable energy systems and investment decision-making. In 2022, renewable energy systems and decision support systems were still popular subjects. Looking ahead to 2023, the fintech ecosystem, decision support systems, and investment decision-making are expected to be the most talked about topics.

Table 2 displays the documents published in the last decade that have received at least five citations from the IRS. The document with the highest number of citations is Gottschlich & Hinz's 2014 paper. The data in the table has been retrieved from Scopus.

Table 2-3. Most cited documents in IRS (2014-2022)

Authors	Year	Title	Citation
Gottschlich & Hinz	2014	A decision support system for stock investment recommendations using collective wisdom	68
Salge et al	2015	Investing in information systems: On the behavioral and institutional search mechanisms underpinning hospitals' is investment decisions	59
Zhou et al	2019	Effects of a generalized dual-credit system on green technology investments and pricing decisions in a supply chain	40
Starita & Scaparra	2016	Optimizing dynamic investment decisions for railway systems protection	37
Ullah & Sepasgozar	2020	Key factors influencing purchase or rent decisions in smart real estate investments: A system dynamics approach using online forum thread data	35
Kovačić et al	2017	Optimal decisions on investments in Urban Energy Cogeneration plants – Extended MRP and fuzzy approach to the stochastic systems	33
Del Giudice et al	2019	Real estate investment choices and decision support systems	32
Geressu & Harou	2015	Screening reservoir systems by considering the efficient trade-offs - Informing infrastructure investment decisions on the Blue Nile	31
Yan et al	2017	Pre-disaster investment decisions for strengthening the Chinese railway system under earthquakes	28
Naranjo & Santos	2019	A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition	27
Fang et al	2021	Assessment of safety management system on energy investment risk using house of quality based on hybrid stochastic interval-valued intuitionistic fuzzy decision-making approach	24
Babaei & Bamdad	2020	A multi-objective instance-based decision support system for investment recommendation in peer-to-peer lending	24
Lakhno et al	2017	Development of the decision making support system to control a procedure of financial investment	24
Teotónio et al	2020	Decision support system for green roofs investments in residential buildings	23
Mo et al	2015	Delaying the introduction of emissions trading systems-Implications for power plant investment and operation from a multi-stage decision model	23
Kamari et al	2018	A hybrid decision support system for generation of holistic renovation scenarios-Cases of energy consumption, investment cost, and thermal indoor comfort	22
von Appen & Braun	2018	Interdependencies between self-sufficiency preferences, techno-economic drivers for investment decisions and grid integration of residential PV storage systems	17
Renna	2017	A Decision Investment Model to Design Manufacturing Systems based on a genetic algorithm and Monte-Carlo simulation	17

Flora & Vargiolu	2020	Price dynamics in the European Union Emissions Trading System and evaluation of its ability to boost emission-related investment decisions	15
Ali et al	2019	Does sustainability reporting via accounting information system influence investment decisions in Iraq?	14
Kafuku et al	2015	Investment decision issues from remanufacturing system perspective: Literature review and further research	14
Keding & Meissner	2021	Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions	13
Jankova et al	2021	Investment decision support based on interval type-2 fuzzy expert system	12
Ribas et al	2015	A decision support system for prioritizing investments in an energy efficiency program in favelas in the city of Rio de Janeiro	12
Akhmetov et al	2019	Mobile platform for decision support system during mutual continuous investment in technology for smart city	11
Quitoras et al	2021	Towards robust investment decisions and policies in integrated energy systems planning: Evaluating trade-offs and risk hedging strategies for remote communities	10
Akhmetov et al	2019	Decision support system about investments in smart city in conditions of incomplete information	10
Bruaset et al	2018	Performance-based modelling of long-term deterioration to support rehabilitation and investment decisions in drinking water distribution systems	10
Li et al	2016	Risk decision-making based on Mahalanobis-Taguchi system and grey cumulative prospect theory for enterprise information investment	10
Akhmetov et al	2018	Model for a computer decision support system on mutual investment in the cybersecurity of educational institutions	9
Cano et al	2017	A strategic decision support system framework for energy-efficient technology investments	9
Al-Augby et al	2016	Proposed investment decision support system for stock exchange using text mining method	9
Cabrera-Paniagua et al	2021	A novel artificial autonomous system for supporting investment decisions using a Big Five model approach	8
Tao et al	2021	Review and analysis of investment decision making algorithms in long-term agent-based electric power system simulation models	8
Khalatur et al	2020	Multiple system of innovation-investment decisions adoption with synergetic approach usage	8
Papapostolou et al	2018	Optimisation of water supply systems in the water – energy nexus: Model development and implementation to support decision making in investment planning	8
Siejka	2017	THE ROLE OF SPATIAL INFORMATION SYSTEMS IN DECISION-MAKING	8

		PROCESSES REGARDING INVESTMENT SITE SELECTION	
Hu & Zhou	2014	A decision support system for joint emission reduction investment and pricing decisions with carbon emission trade	8
Rühr et al	2019	A classification of decision automation and delegation in digital investment management systems	7
Ortner et al	2017	Incentive systems for risky investment decisions under unknown preferences	7
Li et al	2022	Shared energy storage system for prosumers in a community: Investment decision, economic operation, and benefits allocation under a cost-effective way	6
Sun et al	2020	Decision-making of port enterprise safety investment based on system dynamics	6
Xue et al	2019	Multi-scenarios-based operation mode and investment decision of source-storage-load system in business park	6
Thomas et al	2019	A decision-support tool for investment analysis of automated oestrus detection technologies in a seasonal dairy production system	6
Mutanov et al	2018	Investments Decision Making on the Basis of System Dynamics	6
Luo	2020	Application of improved clustering algorithm in investment recommendation in embedded system	5
Wei et al	2019	Joint optimal decision of the shared distribution system through revenue-sharing and cooperative investment contracts	5
Ren & Malik	2019	Investment recommendation system for low-liquidity online peer to peer lending (P2PL) marketplaces	5
Kozlova et al	2018	New investment decision-making tool that combines a fuzzy inference system with real option analysis	5
Niu et al	2017	Improved TOPSIS method for power distribution network investment decision-making based on benefit evaluation indicator system	5
Scaparra et al	2015	Optimizing investment decisions for railway systems protection	5

2.9.6. Proposed Recommender System vs Existing Recommender Systems

This research aims to assist companies in making informed investment decisions by utilizing fuzzy logic techniques based on customer characteristics and experiences.

Table 2-4. Differences in the proposed recommender system compared with existing recommender systems

Phase	Existing Recommender Systems	Proposed Recommender System
Data Gathering/Information Collection	- Typically uses data from actual customers - May use demographic or purchase history data	- Incorporates potential investor data, including fuzzy or uncertain data - Utilizes expert knowledge in the form of fuzzy data
Data Analysis	- May use techniques such as fuzzy linguistic modeling or case-based reasoning - Often employs collaborative filtering, association rule mining, or other traditional methods	- Utilizes ANFIS for investment type recommendations - Combines and utilizes multiple tools and methods for data analysis
Decision Prediction/Recommendation	- May be used for a variety of decision predictions such as behavior prediction, cost prediction, or product recommendation	- Specifically tailored for investment recommendations and tailored to specific investment service types and investor groups.

The proposed recommender system, as outlined in Table 2-4, offers a range of novel capabilities compared to existing systems, including:

- The ability to utilize a dataset to provide input data
- The ability to process potential investor data, including fuzzy or uncertain data
- Tailored recommendations for specific investment types
- Investment recommendations based on investment service types and potential investor groups
- The ability to combine and utilize various tools and methods for data analysis

- The ability to incorporate expert knowledge in creating new rules based on data from the preliminary dataset
- The ability to incorporate expert knowledge (fuzzy data) in creating new rules after analyzing data from investor feedback, allowing for a focus on specific types of investments
- Utilization of ANFIS solutions
- The ability to employ multiple types of recommendation systems including collaborative filtering, knowledge-based, and content filtering.

The next chapter deals with research methodology. In this chapter, the used methods, techniques, and tools to achieve the main goal and specific objectives of the research are discussed.

CHAPTER III RESEARCH METHODOLOGY

This research utilizes an applied-experimental methodology, incorporating a combination of both qualitative and quantitative methods. The research methodology is tailored to each of the research objectives, utilizing a variety of data collection and analysis tools. Quantitative data collection methods, such as database tools and questionnaires, were utilized to gather numerical values. On the other hand, qualitative methods, such as interviews, focus groups, and observations, were employed to consider factors beyond numerical values. The combined approach, utilizing both qualitative and quantitative methods, was employed in this study to provide a comprehensive understanding of the research topic.

In this case, some of the data used is not precise, but rather qualitative in nature. Some of this qualitative data was obtained using a scale, making it fuzzy and not crisp. Additionally, human knowledge and resulting data may not always be completely accurate, also resulting in fuzzy information. To account for this, the recommender system must be designed to handle fuzzy data. The FIS is utilized in this design to effectively handle the uncertainty of the data. When there is no dataset available and decisions are based solely on human knowledge, FIS is used alone. However, when a dataset is present and human knowledge is also utilized, ANFIS is employed to effectively incorporate both sources of information. In this study, the researcher proposes an IRS that utilizes ANFIS to analyze customer data and provide personalized investment recommendations. The goal of this system is to assist investment companies, individual investors, and fund managers in making informed decisions by recommending investment products and services tailored to the specific needs, experiences, and characteristics of potential investors. This system is unique in that it considers the correlation between potential investors' demographic and personality traits, investor groups, and investment services to recommend a portfolio that best fits the individual's needs. An intelligent fuzzy framework is used to generate association rules. The proposed system utilizes both machine learning and fuzzy logic to make investment recommendations. It is tested using a portfolio dataset and a web-based investment questionnaire. The research approach demonstrates the application of soft computing techniques, such as machine learning and fuzzy logic, in the design of recommender systems. The research process can be broken down into four main functions, as illustrated in Figure 3-1. The pre-processing steps consist of two parallel processes, where investment types and potential investor types are clustered using unsupervised machine learning techniques based on multiple variables. In the design step, ANFIS is used to create the combined IRS. The system is then used to make proposals during the proposal stage.

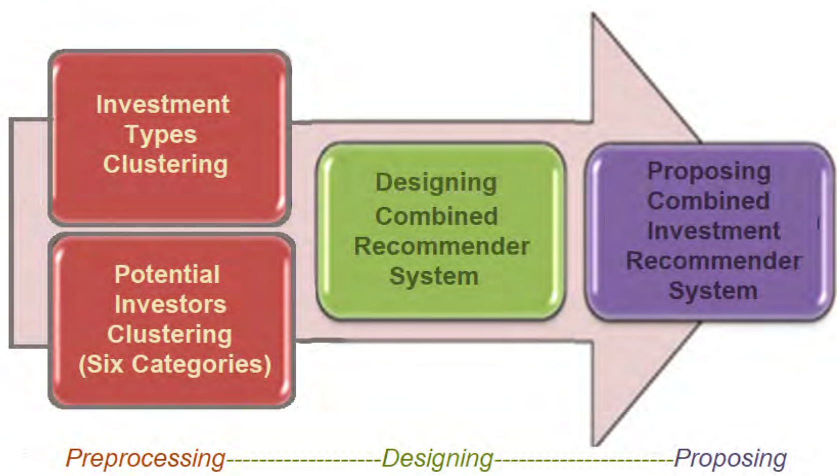


Figure 3-1. Research process to propose a combined recommender system

In this step, the proposed IRS provides tailored recommendations for the investor. The inputs to the ANFIS are the summary of factors related to the potential investor from the previous step and the output is a cluster of recommended investment types. The ANFIS utilizes two types of rules for prediction: those derived from training data and those based on the investor's feedback and expert opinions. The Membership functions (MFs) of ANFIS are designed based on the nature of the input factors and their measurement scales. MFs are used to model fuzzy sets, which are sets that do not have a sharp boundary. The MFs are designed based on the degree of membership of an element to a fuzzy set, which can vary from 0 (not a member) to 1 (a full member). The different levels of agreement that were designed for the MFs include:

a) Strong agreement (1.0): This level of the agreement indicates that an element is a full member of the fuzzy set, and there is no ambiguity or uncertainty.

b) Moderate agreement (0.5 to 0.9): This level of agreement indicates that an element is a partial member of the fuzzy set. There is some degree of ambiguity or uncertainty in the degree of membership.

c) Weak agreement (0.1 to 0.4): This level of agreement indicates that an element is a weak member of the fuzzy set. There is a high degree of ambiguity or uncertainty in the degree of membership.

d) No agreement (0.0): This level of the agreement indicates that an element is not a member of the fuzzy set.

The development of the investment recommender system heavily relies on the appropriate design of MFs and the selection of agreement levels. It is crucial to ensure that the MFs are designed to effectively capture the investor's traits and investment experiences, as this will guarantee that the recommendations provided are both relevant and accurate. The term "MFs" in this study refers to the different choices that potential investors may consider. By converting numerical inputs into fuzzy values using MFs, the fuzzy logic system can then facilitate decision-making. Each input is linked with a set of membership functions representing the various options that potential investors can select. The parameters of the membership functions will vary depending on the input and the specific needs of the system. Some common parameters that might be used in defining membership functions include:

- The shape of the function: This can be triangular, trapezoidal, Gaussian, or other shapes, depending on the nature of the input and the desired degree of fuzziness.
- The center of the function: This represents the point at which the input value has a membership value of 1.0. Depending on the shape of the function, this could be a single point or a range of values.
- The width of the function: This defines the extent to which the input value has a non-zero membership value. For example, a wide membership function will have non-zero membership values over a broader range of input values, while a narrow membership function will have non-zero membership values over a smaller range of input values.

By defining the parameters for each input and membership function, the system can accurately represent the range of potential investor preferences. This, in turn, enables effective decision-making based on those preferences. Here are 3-5 possible performance metrics that could be used to evaluate the overall research:

a) Accuracy: This metric measures how well the model can predict the output values based on the input data. A high accuracy score indicates that the model is making accurate predictions, while a low score suggests that the model is not performing well.

b) Precision and Recall: Precision and recall are commonly used metrics in classification tasks to evaluate the performance of the model. Precision measures the proportion of correctly predicted positive samples among all predicted positive samples, while recall measures the proportion of correctly predicted positive samples among all actual positive samples. These metrics provide a good balance between the rate of correctly identified positive samples and the rate of false positives.

c) F1 Score: The F1 score is the harmonic mean of precision and recall and is often used to evaluate the overall performance of the model in a classification task. It provides a good balance between precision and recall.

d) Mean Squared Error (MSE): MSE is a commonly used metric for evaluating the performance of regression models. It measures the average of the squared differences between the predicted values and the actual values. A lower MSE indicates better model performance.

e) R-Squared (R2) Score: R2 score measures the proportion of variance in the target variable that is explained by the model. It provides a measure of how well the model fits the data. A higher R2 score indicates a better fit.

Table 3-1 provides a brief overview of the research type, data analysis tools, and methods used in this study, in relation to the research questions.

Table 3-1. Research Methodology

No	Objective	Type	Data	Technique/ Method	Tool
1	Proposing primary Framework	Qualitative	Literature Review Library Studies Documents	Review Ontology	Mind-Mapper Protege
2	Preparation Data	Qualitative/ Quantitative	Portfolio dataset	Translation Cleaning Data Coding Data Convert Data Clustering	Translator Excel ETL Tools JMP
3	Designing System	Qualitative	Prepared data from the previous step	Preparation data Designing Sugeno FIS	MATLAB Fuzzy toolbox
4	Proposing System	Qualitative	Trained data Fuzzy rules	Training data Generating FIS Testing FIS Export model	MATLAB Fuzzy toolbox
5	System Evaluation	Quantitative Qualitative	Trained data Library studies	Root Mean Square Error (RMSE) F1 Score Comparison	MATLAB Fuzzy toolbox

3.1. Data

The dataset used in this study is a ready dataset in the Hungarian language that was collected through an online investment questionnaire published by a leading Hungarian financial portal called "Portfolio." The questionnaire, which is accessible in a web-based format at https://www.portfolio.hu/befektetesi-kerdoiv/?page=1, is designed to gather information about portfolio investments. Portfolio investments are defined as investments in the form of a group (portfolio) of assets, including transactions in equity, securities, such as common stock, and debt securities, such as banknotes, bonds, and debentures (World Bank, 2018). This type of investment covers a range of securities, such as stocks and bonds, as well as other types of investment vehicles. A diversified portfolio helps spread the risk of loss because of the below-expectations performance of one or a few of them ("Portfolio Investment," 2019). The original dataset was translated into English, corrected, and cleaned multiple times. After cleaning and translation, data were categorized based on the subject into 7 categories. The data was prepared for clustering and provided inputs and output for ANFIS (Attachment 1 translated to English).

Table 3-2. Investment questionnaire sections

Page No.	Subject	No. of Questions	Question Type	
P1	Customer's digital financial solutions	6	Closed-ended	Y/N
P2	Customer's financial awareness & risk appetite	11	Closed-ended	Multiple choice Y/N Numeric / fill-in-the-blank
P3	Customer's current savings and financial situation	6	Closed-ended	Multiple choice
P4	Customer's characteristics	19		5-point Likert scale One choice
P5	Customer's financial plans	2	Open-ended	Open box
P6	Investment services & tools And customer satisfaction	24	Closed-ended	Numeric / fill-in-the-blank One choice 5-point Likert scale Multiple choice
P7	Customer's demographic data (Gender, age, living location, education, job)	5	Closed-ended	Numeric / fill-in-the-blank Multiple choice
Total		73 Questions		

The portfolio is an online financial newspaper in Hungary with a user count of one million per month as of 2018. Portfolio Group was once ranked among the ten most-read news websites and the 15 most visited websites in Hungary (TOP15, 2018). The portfolio has a different emphasis on business, financial, and economic news. Besides its online media platforms, the enterprise offers several buying and selling offerings and presents a personal analysis of financial markets. The company also has a tournament commercial enterprise line, which organizes industry forums yearly in the fields of agriculture, insurance, lending, asset management, company finance, capital markets, the car sector, monetary IT (Information Technology), and real estate (Portfolio.hu, 2018). The data used in this research was collected from a questionnaire published by Portfolio.hu in 2019. The questionnaire, which was designed in partnership with the Corvinus University of Budapest and Dorsum, a leading provider of innovative investment software, was used to gather information on the financial consciousness of 1542 respondents. The data was initially collected in Hungarian and then translated, cleaned, and converted to numerical format by Asefeh Asemi. The questionnaire was accessible on the portfolio.hu website and collected data in various structured and unstructured formats (as described in Table 3-2 of the study). The purpose of the research was to gain insight into the financial consciousness of the respondents and the data was coded based on the questions asked. To prepare the data for analysis, the responses were converted to the numerical format. The research utilized two categories of data: 1) data obtained from the investment questionnaire and 2) opinions of experts in the field for the design of the system. The data here don't need to be analyzed frequently. The data is only analyzed (training fetching rules from data) once. Then the created rules will be applied to any new set of inputs. Therefore, computational time is considered one time.

3.2. Community of research & sampling

The research community for this study was divided into two sections. In the first section, data preparation, the research community consisted of all users of the

Portfolio who responded to the online investment questionnaire. These individuals were considered potential investors in this study. In the second section, system design, the research community included experts in investment and FISs, as well as relevant literature on the subject. The sample of potential investors involved in the dataset consisted of 1542 respondents, and the sampling method used was a random census. The gender breakdown of the respondents was 87% male and 13% female, as demonstrated in Table 3-3 and Figure 3-2.

Table 3-3. Gender of the respondents

Gender	Amount
Male	1307
Female	191

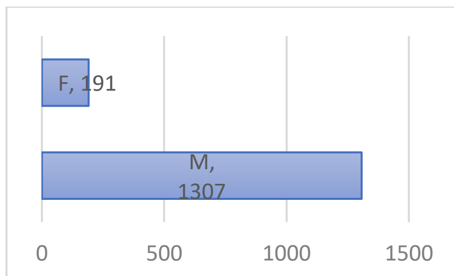


Figure 3-2. Gender of the respondents

According to the data analyzed, 21% of the respondents were in the age range of 15-19 years old, as reported in Table 3-4 and illustrated in Figure 3-3. Conversely, a minimal percentage (near 0%, or specifically 3 respondents) belonged to the age range of 75-79 years old.

Table 3-4. Age of the respondents

No.	Age/ Year	Percentage
1	15-19	21%
2	24-24	4%
3	25-29	9%
4	34-34	14%
5	35-39	13%
6	44-44	12%
7	45-49	11%
8	54-54	5%
9	55-59	3%
10	64-64	4%
11	65-69	3%
12	74-74	1%
13	75-79	0%

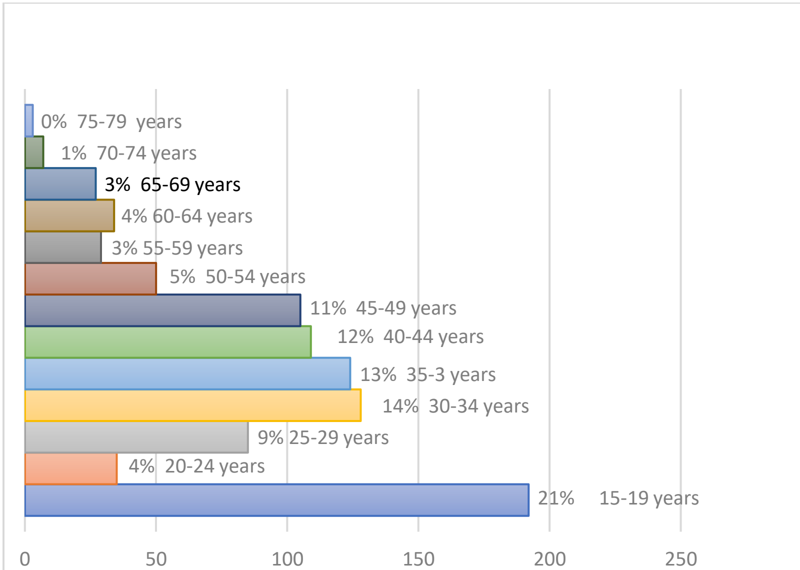


Figure 3-3. Age of the respondents

According to the data presented in Table 3-5 and Figure 3-4, the living location of the respondents was analyzed. The results indicate that 784 individuals were residing in Budapest, while 122 individuals were residing in rural villages.

Table 3-5. Living location of the respondents

Location	Amount
Budapest	784
City	268
County Town	314
Village	122

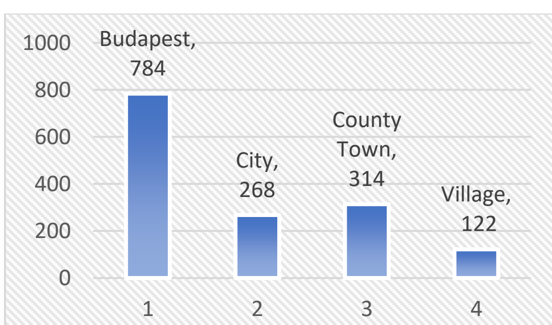


Figure 3-4. Living location of the respondents

The level of education of the respondents in the study, as reported in Table 3-6 and illustrated in Figure 3-5, consisted of 596 individuals who held a college or university level degree in a field outside of economics, and 2 individuals who held only a primary school level of education.

Table 3-6. Level of the respondents' education

Level of Education	Amount
College or university/ economics	564
College or university/ non-economics	596
Postgraduate training	73
High school secondary education	237
Currently studying	39
Primary school	2

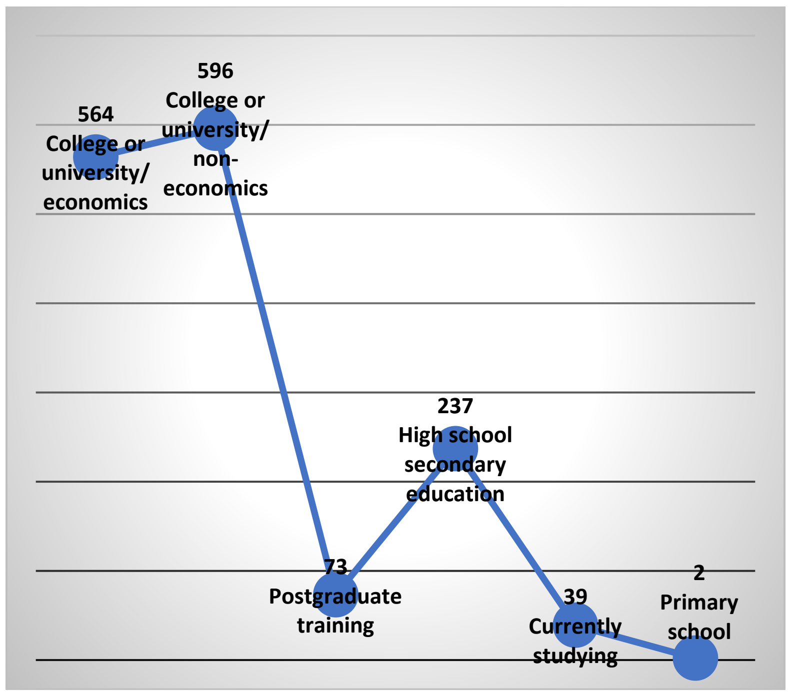


Figure 3-5. Level of the respondents' education

According to data presented in Table 3-7 and Figure 3-6, the study surveyed a total of 559 respondents who identified as "subordinate intellectual workers," and an additional 3 respondents who identified as "Large Contractors."

Table 3-7. Respondents’ Job

Code	Respondents’ Job	Amount
1	Employee middle management	228
2	Small-medium business	115
3	Graduate freelance	69
4	Employed lower manager	138
5	Subordinate intellectual worker	559
6	Skilled worker	51
7	Employed senior management	67
8	Micro or self-employed	88
9	Other intellectual services, traders (employees)	60
10	Trained or auxiliary	9
11	Large Contractor	3
12	Farmworker / agricultural/ season worker	8

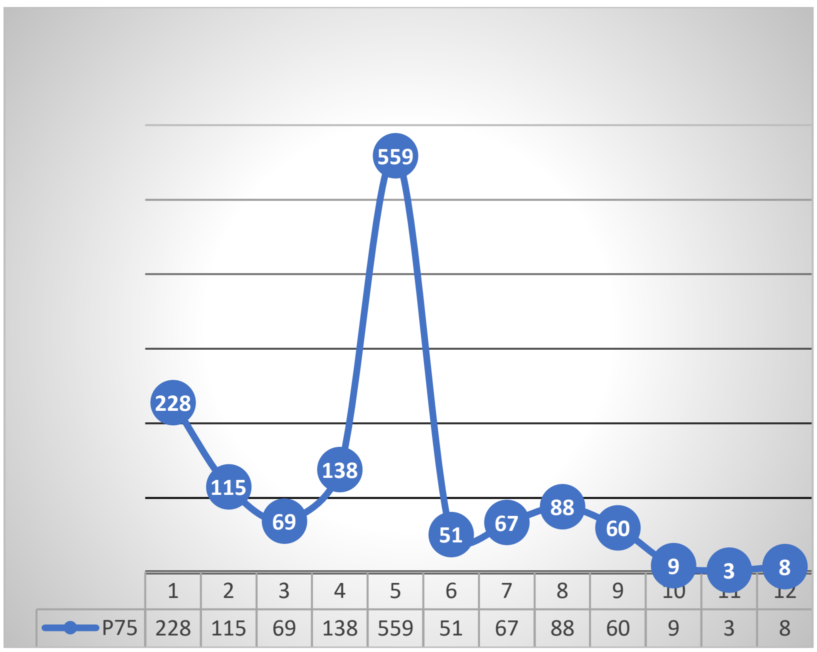


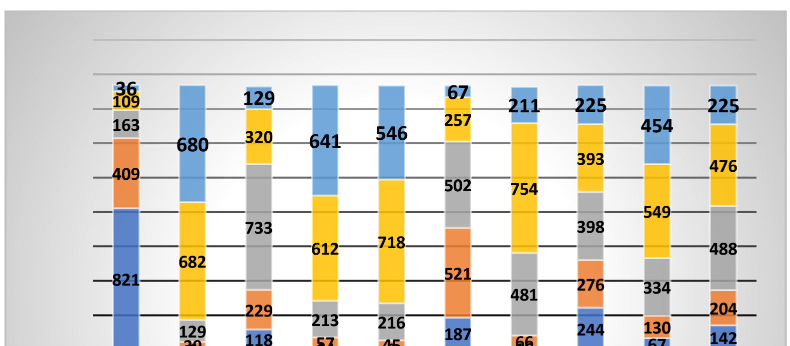
Figure 3-6. Respondents’ Job

A series of questions were administered to evaluate the personal characteristics of potential investors. The responses obtained from these questions provide insight into the correlation between individual behaviors and attitudes toward finance, and how these factors impact the selection of appropriate savings and investment products.

Table 3-8. Personal characteristics of potential investors (1)

Question	Q. Code	Sub-Question
How true are these statements to you?	P411	I do not plan my future, I prefer drifting with events, I plan flexibly
	P412	When I set a goal for myself, I usually plan the steps to get there
	P413	If I feel like my job is getting too risky, I do not waste my time on it
	P414	My destiny is in my own hands
	P415	It is up to me how I reach my goals
	P416	If my plan does not go as I expected, I will let it go
	P417	The factors that ensure my success are in my hands
	P418	I keep a detailed list of my plans
	P419	I like working in teams and getting help and assistance from the right professionals
	P4110	When I reach my goal, I reward myself

In this study, a five-point Likert scale was employed to gather responses from participants on 10 questions. The scale ranged from "not at all true" to "very true" and aimed to gauge the degree to which certain statements were perceived as accurate by the participants (as displayed in Table 3-8). The results revealed that out of the 821 participants, most of them strongly agreed that they do not plan and prefer to go with the flow. Conversely, only a small minority of 6 participants strongly agreed that they set goals for themselves and typically plan the steps to achieve them (as depicted in Figure 3-7).



	P411	P412	P413	P414	P415	P416	P417	P418	P419	P4110
5	36	680	129	641	546	67	211	225	454	225
4	109	682	320	612	718	257	754	393	549	476
3	163	129	733	213	216	502	481	398	334	488
2	409	39	229	57	45	521	66	276	130	204
1	821	6	118	14	9	187	16	244	67	142

Figure 3-7. Personal characteristics of potential investors (1)

In the second phase of the study, participants were presented with two options as answers to seven questions, as detailed in Table 3-9. The objective of these questions was to gather insight into the personal characteristics of the participants. Upon analyzing the responses, it was found that a total of 1380 participants reported that a plan influenced their options, while 159 respondents stated that their imagination was the only limitation on their possibilities, as depicted in Figure 3-8.

Table 3-9. Personal characteristics of the potential investors (2)

Question	Q. Code	Sub-Question	Options
Choose from the following options	P421	When I decide	(1) I tend to be nervous afterward if I have made the right decision/ (2) Instead, he worries if I am going to make the right decision
	P422	When I plan my day	(1) Rather, I focus on the tasks I see given in the day and organize my other activities around them/ (2) I would rather imagine what my day should be like and shape my business
	P423	When I do a task	(1) I work hastily than comfortably/ (2) I work more comfortably than a little rush
	P424	What are your characteristics?	(1) I try to influence the course of things / (2) I let things happen around me, I adjust to them
	P426	When I plan	(1) I can see what options I can choose from / (2) Only my imagination limits my possibilities
	P427	During my work	(1) Rather, I see dynamic work, clear assignment of tasks / (2) I find myself disorganized, deconstructive, passive
	P428	It bothers me more	(1) If the work I am doing seems pointless/ (2) If the work I am doing does not satisfy me mentally

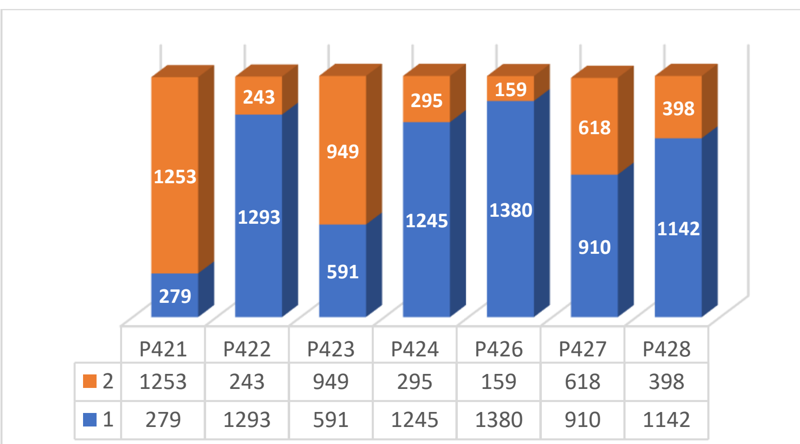


Figure 3-8. Personal characteristics of the potential investors (2)

3.3. Research Methods

A primary framework for a combined IRS is proposed and developed with practical results. The data analysis employs various methods, including machine learning, cluster analysis, and ANFIS. Seven ANFIS models are developed and proposed, each utilizing one independent variable, the type of investment, and several dependent variables. The seventh ANFIS model serves as the final system, which is arrived at by exporting seven group factors and incorporating one independent variable and six dependent variables.

3.3.1. Machine Learning

Machine learning is a field of artificial intelligence that utilizes data mining techniques to improve the capability and performance of a system based on its past performance. According to Garbade (2021), "machine learning can be loosely interpreted to mean empowering computer systems with the ability to “learn”. The main goal of machine learning is to enable machines to learn by themselves using the provided data and make accurate predictions." Alpaydin (2020) notes that machine learning involves computers discovering how to perform tasks without being explicitly programmed to do so, by learning from the data provided. The main fields related to machine learning in data sciences include artificial intelligence, natural language processing, data mining, mathematics, statistics, computer science, deep learning, and science, as depicted in Figure 3-9 of Sarkar et al. (2018).

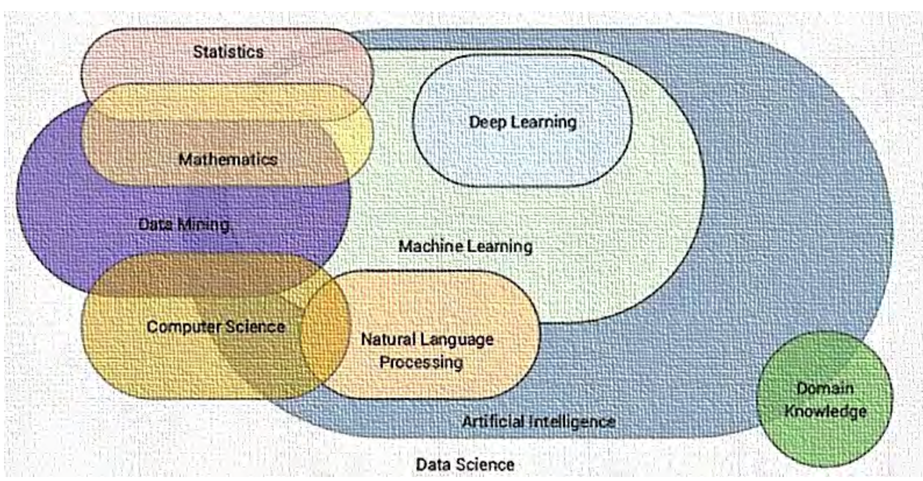


Figure 3-9. The main fields related to machine learning in data science

The use of machine learning methods, such as supervised, semi-supervised, and unsupervised learning, is an appropriate approach when there is limited prior knowledge of data patterns (Baştanlar & Ozuysal, 2014). One of the advantages of machine learning is that it is typically fast and cost-effective. In this research, unsupervised learning is utilized as it allows the system to learn from unlabeled data and categorize similar concepts together. It is important to note that the concepts within each category differ from those in other categories.

3.3.2. Cluster Analysis

Clustering, an unsupervised learning method in machine learning, is a process of grouping similar data points together without prior knowledge of the categories. One popular clustering algorithm is the k-means method, which is widely used for grouping unlabeled data (Han et al, 2012). The goal of unsupervised learning, such as clustering, is to discover the underlying structure of the data by grouping similar data points together. Clustering is a common descriptive method used to identify homogeneous groups of objects based on their properties, and the groups can be either segregated or overlapping (Agrawal et al., 2005). In the context of investment, K-Means and Self-Organizing Map (SOM) clustering are used to group potential investors and investment types/products. The K-Means algorithm, which is one of the simplest machine learning algorithms without supervision, determines the number of k centroids and assigns each data point to the nearest cluster, while minimizing the distance between the data points and the centroids. Elbow curves and silhouette plots are commonly used to identify the optimal number of clusters. Elbow curves help to determine the point at which increasing the number of clusters no longer results in a significant decrease in within-cluster sum of squares. Silhouette plots calculate the average silhouette width for each cluster and indicate the degree of separation between clusters. Using these techniques, potential investors can be segmented into different categories. The resulting groups are used as metrics for the ANFIS model, which in turn provides investment recommendations based on the clustering results (Kumar, 2020). Finally, the potential investors are categorized into several groups, and the results from these groups determine the metrics that enter the ANFIS for investment recommendations.

3.3.3. Adaptive Network-Based Fuzzy Inference System

ANFIS is an ANN developed based on the Takagi-Sugeno fuzzy inference system (Jang, 1991, 1992, 1993). It simplifies the deployment of fuzzy logic judgment in comparison to traditional neural network simulations (Asemi & Asemi, 2014). A recommender system that utilizes ANFIS examines the user's past behavior and recommends relevant and accurate information from a wide range of sources, including the user's interests and needs, as well as the products and services provided by the system. By identifying comparable items that align with the user's search and comparing them with similar users, ANFIS-based recommender systems can filter and provide relevant and accurate recommendations. The characteristics of users, their previous experiences, their history of activity in the system, and their profile information are highly effective in designing the recommender system (Jain & Gupta, 2018). The more accurate information leads to more appropriate recommendations for the user. Recommender systems involve a combination of techniques, such as hybrid systems, which combine techniques A and B to try to use the advantages of A to overcome the disadvantages of B. Knowledge-based systems recommenders recommend items based on domain-specific knowledge of how to provide some of the features needed by users and preferences, and how useful this is for the user (Abraham, 2005). In these systems, a function matches the similarity of the user needs (problem description) with the recommendations (problem solutions) using fuzzy logic, which is widely used in system design to control uncertainty, inaccuracy, and ambiguity in case characteristics and user behavior. There is a single framework for neural networks and fuzzy logic principles, known as ANIS, which conforms to a set of ambiguous IF-THEN rules that could approximate learning nonlinear functions. According to Tahmasebi (2010), ANFIS architecture consists of several layers. The first layer, also known as the fuzzification layer, takes the input values and determines the MFs that belong to them. In this layer, the degree of membership of each function is calculated using a set of preliminary parameters. The second layer, known as the "rule layer," is responsible for generating rules. The role of the third layer is to normalize the calculated firepower by dividing it into any value for the entire firepower. The fourth layer normalizes the values and takes the result parameter set as input. The values returned by this layer are fuzzy, and these values are transferred to the last layer to return the final output (Kamal et al, 2018). According to Karaboga and Kaya (2018), neural networks work with a data preprocessing step in which properties are converted to normal values between 0 and 1. The ANFIS neural network does not require the sigmoid function and processes by converting numerical values to fuzzy values. Asemi et al. (2019) also states that the robust generalizations and accurate forecasts provided by ANFIS make it possible to properly control the uncertainty in any conventional system. Engineers have made extensive use of ANFIS features in systems reconditioning design (Petkovi´c and ´Cojbaši´ c 2012; Petkovi´c et al.2012).

3.4. Proposed the Investment Recommender System framework

The proposed recommender system framework is built on the methods and techniques outlined in prior research. The data was analyzed in a step-by-step manner across different layers. The primary framework is divided into two main sections: data analysis and decision making. Figure 3-10 illustrates the primary framework and how it addresses the research questions of the study. The research design is divided into three phases:

- Data gathering: The data gathering phase includes the data acquisition and storage layers.
- Data analysis: The second phase includes two functions: (a) machine learning techniques such as clustering and (b) the use of ANIS. The system includes multiple parallel ANFISs, each with different inputs and a varying number of MFs.
- Decision making: The final phase includes the recommendation layer, which presents information to the customer and receives their feedback. This feedback is used to evaluate the probability errors, which are then referred to the data analysis phase for correction.

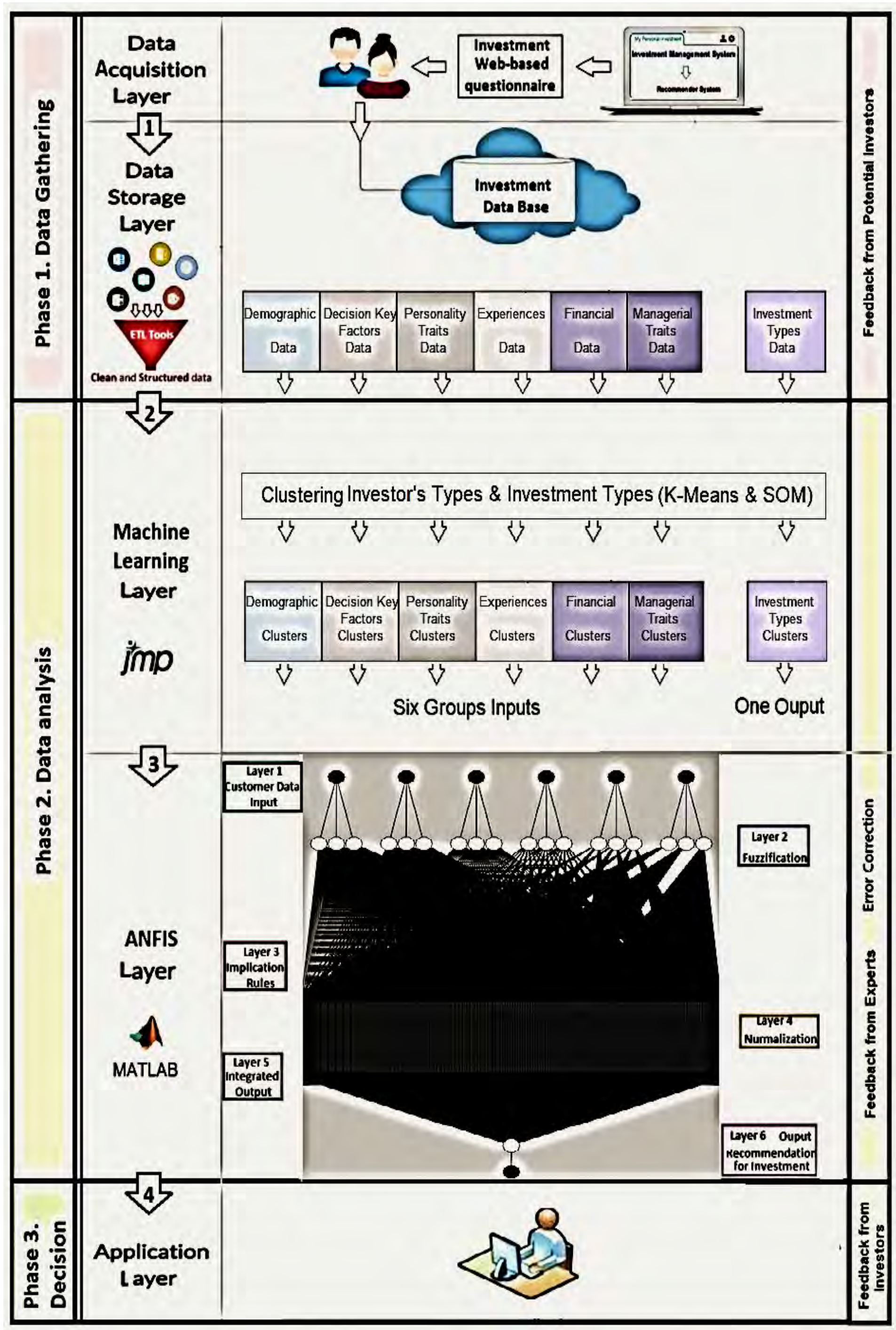


Figure 3-10. The primary IRS framework using ANFIS

The proposed IRS framework is composed of several layers, each with a specific function. These layers include 1. Data Acquisition: This layer is responsible for collecting and preprocessing the data required for the system, 2. Data Storage: This layer stores the data in a format that is easily accessible and retrievable for the system, 3. Machine Learning: This layer utilizes various techniques and algorithms to analyze and extract meaningful insights from the data, 4. ANFIS Layers: This layer comprises multiple sub-layers such as Fuzzification, Implication Rules, Normalization, Defuzzification, Integration, and Aggregated Output MFs, which are used to adapt and optimize the system's parameters, 5. Investment Recommendation and Feedback: This layer, also known as the Application layer, generates investment recommendations based on the insights gained from the previous layers, and also receives feedback from the user to improve the system's performance.

3.4.1. Data Acquisition Layer

The purpose of this layer is to gather information from users regarding their financial awareness and habits. This is accomplished through a web-based investment questionnaire, which includes questions about savings, spending practices, utilization of digital financial solutions, communication preferences, satisfaction with financial institutions and organizations, perceptions of future economic conditions, and demographic data. This information is then transferred to the next layer, which is responsible for storing the data in the cloud. The system uses this information to provide personalized recommendations based on user demographics. The basic concept behind these recommender systems is to provide different recommendations for different groups of users. Many websites today use simple and effective methods for providing recommendations based on user demographics and personal information. For example, users are directed to specific websites based on their language or country and offers may also be customized according to the user's age. While these approaches have been commonly used in marketing, there is limited research on demographic-based recommender systems.

3.4.2. Data Storage Layer

The original dataset was organized based on seven pages of a questionnaire. This study utilized a selection of this data, and during the preparation phase, it was divided into seven distinct categories (as outlined in Table 3-10). The proposed system focuses on one category of investment type as the output and utilizes six categories of potential investor types as inputs. The data storage layer of the proposed system stores all potential investor data on a private server, utilizing various data processing formats for efficient storage. In the data reformatting phase, the data is transformed and structured in a format that is suitable for the machine learning layer. This process involves cleaning, normalizing, and organizing the data to ensure that it can be effectively utilized by the machine learning models. Once the data is reformatted, it is transferred to the machine learning layer for further processing and analysis.

Table 3-10. Data categorization of potential investors

Data Type	Data Category	No. of Clusters	No. of Columns	No. of Rows	**No. of Questions**
Output	Investment Types' experiences of the potential investors	3	1	1542	4
Input 1	Demographics of the potential investors	3	1	1542	6
Input 2	Key Decision Factors of potential investors	3	1	1542	4
Input 3	Personality Traits of potential investors	3	1	1542	6
Input 4	Investment Experiences by the potential investors	3	1	1542	7
Input 5	Financial Situation of the potential investors	3	1	1542	4
Input 6	Managerial Traits of the potential investors	3	1	1542	9
Total	One output & six inputs	7×3	7	1542	40

3.4.3. Machine Learning Layer

The clustering component is a crucial aspect of the system's data analysis layer, which utilizes data mining and machine learning algorithms. This layer receives data from the data storage layer and processes it through clustering techniques. The data is then prepared for analysis in the machine learning layer, where attributes are grouped and clustered to provide inputs and outputs for the proposed ANFIS system. JMP software is utilized for clustering due to its wide range of features, capability to compare outputs in different ways, user-friendly interface, compatibility with Windows, and comprehensive documentation. The K-Means, SOM, Silhouette, and Elbow methods are employed to cluster the data, which is then transferred to the ANFIS layer for system design and implementation.

3.4.4. ANFIS Layers

In this layer, the ANFIS will utilize the Sugeno fuzzy model. The system is composed of five layers, as illustrated in Figure 3-11.

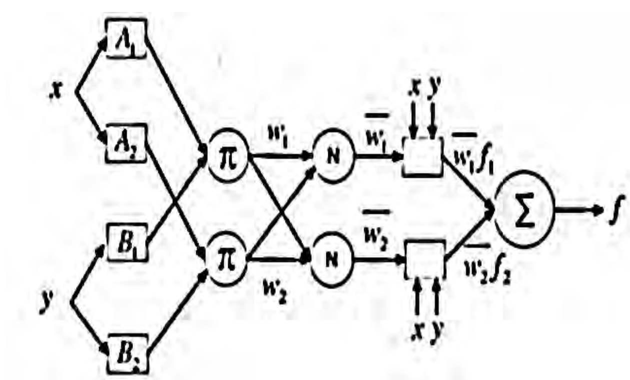


Figure 3-11. The architecture of a simple two-rule Sugeno-type ANFIS (Asemi, et al. 2019)

In this research, a combined IRS framework using ANFIS is proposed. The system includes parallel ANFISs, each with different inputs and MFs. The ANFIS model consists of five layers:

1. The first layer, known as the fuzzy layer, receives external input signals and directs them to the next layer. Neurons in this layer only serve as a conduit for the input signals.

2. The second layer, called the fuzzy rule layer, is the second hidden layer. Each neuron in this layer is associated with a single fuzzy Sugeno law.
3. The third layer, known as the normalization layer, is the third hidden layer. Each neuron in this layer receives and calculates signals from all neurons in the previous layer, resulting in a normalized firing strength value. This value determines the validity of the relevant rule for the given inputs.
4. The fourth layer, called the diffusion layer, is the fourth hidden layer. Each neuron in this layer is related to the corresponding normalized neuron in the previous layer, and it also receives the first input signals (x1, x2, etc.). The defuzzified neuron in this layer calculates the weight of the result of a rule.
5. The fifth and final layer is the output layer. In this layer, the neurons from the previous stage are added together and converted to numeric outputs through defuzzification using the Centre of Gravity method. There is only one neuron in this layer.

This proposed ANFIS model aims to adapt and optimize the system's parameters, making it more accurate and efficient in recommending investments. It could potentially be used to help individuals or organizations make informed investment decisions based on market trends and historical data. MFs play a crucial role in fuzzy sets. These functions form the foundation of fuzzy set theory, making it essential to accurately determine their shape. The shape of MFs can significantly impact the performance of a FIS in solving a specific research problem. There are several shapes of MFs, including triangular (trimf), trapezoidal (trapmf), and gaussian (gaussmf). The triangular shape is commonly used in fuzzy system design due to its ability to fuzz input into three specified parameters. The trapezoidal shape is useful in situations where data is ambiguous and is a result of converting linguistic variables into numerical variables. Gaussian shape is a popular method for determining the shape of MFs in fuzzy sets, and it is non-zero at all points. The shape of MFs is determined based on the degree of certainty of an input and its defined meaning of linguistic values. In this research, the number of MFs was determined based on the questions and answer options of the questionnaire, and their shape was determined through consultation with experts and designers of fuzzy systems. This model provides customized investment recommendations based on the needs of potential investors by using their input preferences to generate personalized output recommendations. The ANFIS model uses a combination of fuzzy logic and neural networks to model the relationship between the input preferences of potential investors and investment recommendations. It does this by using the fuzzy logic system to interpret the input preferences of the potential investor, and the neural network system to make investment recommendations based on those preferences. Specifically, the ANFIS model takes in the input preferences from the investor, which are fuzzified and passed through the fuzzy logic system to generate a set of fuzzy rules. These rules are then passed through the neural network system, which calculates the output investment recommendations based on the combination of the fuzzy rules and the input preferences. The output recommendations are also fuzzy and are then defuzzified to provide a final set of investment recommendations tailored to the needs of the potential investor. In this way, the ANFIS model can provide customized investment recommendations to potential investors based on their individual needs and preferences.

3.4.5. Application Layer

In the application layer, the system is tailored to meet the specific needs of potential investors and presents recommendations for investment products and services. This layer is connected to the data analysis phase, enabling end-users to access the source of investment recommendations within their investment platform via investment applications. The necessary applications for utilizing the recommendations within the investment platform are in this section of the model. Additionally, the system can receive feedback from customers in this layer and any probability errors are referred to the data analysis section for detection.

In the following chapter, the main research question and related sub-questions are addressed using experimental results. Six categories are utilized to develop six separate ANFIS models, and a combined ANFIS is also presented.

CHAPTER IV EXPERIMENTS RESULTS AND ANALYSIS

In Chapter IV, Experiments and Results, the main research question and objectives of this dissertation are explored through comprehensive experimentation and analysis. The main research question aims to investigate the utilization of an ANFIS in creating an effective and efficient investment recommendation system. The main objective is to propose a combined IRS that uses ANFIS to provide accurate and efficient investment recommendations to potential investors. To achieve this objective, the research focuses on three specific sub-goals. The first sub-goal involves categorizing and clustering potential investors based on available data to make more accurate investment recommendations. The second sub-goal involves offering customized investment-type services using adaptive neural-fuzzy inference solutions for different categories of potential investors. Finally, the third sub-goal proposes a combined recommender system that provides appropriate investment type recommendations for all categorized and clustered potential investors. The results and findings of the experiments conducted to address these sub-goals will be presented and analyzed in detail in this chapter.

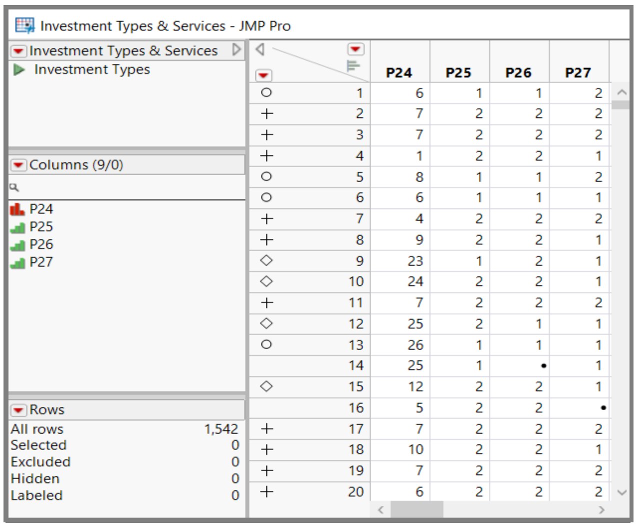
4.1. Clustering Output Investment Type/Product

To develop the system, output was initially defined for all ANFIS models. The output represents the investment type or product, as per the research objectives. The questionnaire consisted of four questions (pages 2-4 to 2-7) related to investment type/product and collected data from 1542 respondents. The first question asked about the investment products used by the potential investors, with options including listed stock, mutual fund, voluntary pension fund, government securities, and other financial products. As multiple answers were allowed, 31 categories were created based on the multiple choices. The second question asked whether the respondents had invested in the stock market in the last 3 years, with options of "Yes" or "No". The third question inquired about the regular monitoring of stock performance, with options of "Yes" or "No". The fourth question asked about the investment in government bonds, with options of "Yes" or "No". These questions serve as a sample to propose the system, and the company may modify the questions based on its objectives. Table 4-1 shows the missing values and the variance explained by the four principal components of the investment type/product data columns.

Table 4-1. Missing values and variances of investment type data

Questions	P24 (Nominal)	P25 (Ordinal)	P26 (Ordinal)	P27 (Ordinal)
Missing values Count in columns	198	4	2	2
Percentage of missing values	12.8%	0.3%	0.1%	0.1%
Individual variance	0.95	0.04	0.01	0.00
Cumulative variance	0.95	0.99	1.00	1.00

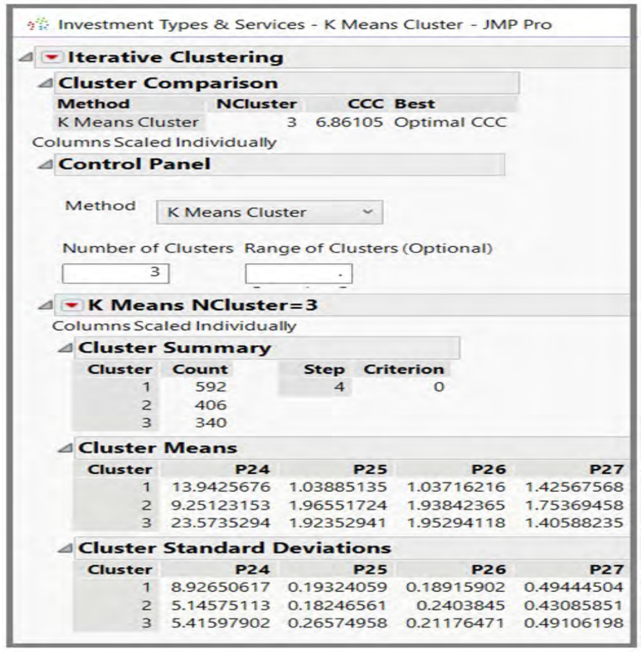
Clustering, an unsupervised machine learning technique, involves dividing data into similar groups. “The K-Means method, a simple distance-based clustering technique, has been widely applied in clustering investment type/product data. To evaluate the performance of K-Means, various software applications such as Python, R Studio, RapidMiner, Tableau, and JMP were tested. JMP was eventually selected due to its better results” (Clustering Methods for Unsupervised Machine Learning, 2019). By default, JMP does not utilize the Silhouette Score or Elbow method. However, researchers can write scripts to implement these methods. JMP offers access to three clustering algorithms, in addition to latent class analysis and cluster variables. The three clustering algorithms are Hierarchical clustering, K-means, and normal mixtures. The criterion for determining the optimal number of clusters in Hierarchical clustering and K-means is the Cubic Cluster Criterion (CCC), while normal mixtures utilize the Corrected Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). In JMP, the number of clusters is initially selected randomly. The data is then assigned to one of these clusters based on the degree of proximity or similarity. “JMP uses the CCC to select the optimal number of clusters that fit the data best . The CCC is used to estimate the number of clusters using Ward's minimum variance method, K-Means, or other methods that minimize the within-cluster sum of squares. The performance of the CCC is evaluated using Monte Carlo methods” (SAS Help Center: Cubic Clustering Criterion, 2015). The JMP software is then used to cluster the data and generate output for the ANFIS system.



		P24	P25	P26	P27
○	1	6	1	1	2
+	2	7	2	2	2
+	3	7	2	2	2
+	4	1	2	2	1
○	5	8	1	1	2
○	6	6	1	1	1
+	7	4	2	2	2
+	8	9	2	2	1
◇	9	23	1	2	1
◇	10	24	2	2	1
+	11	7	2	2	2
◇	12	25	2	1	1
○	13	26	1	1	1
	14	25	1	•	1
◇	15	12	2	2	1
	16	5	2	2	•
+	17	7	2	2	2
+	18	10	2	2	1
+	19	7	2	2	2
+	20	6	2	2	2

Figure 4-1. A part of imported investment type data in JMP

In the first step of the data clustering process, investment type data is imported into JMP as four columns with 1542 rows. As depicted in Figure 4-1, the imported data in JMP is used to cluster the investment type information. Afterward, the data is pre-processed using the K-Means technique within JMP.



Investment Types & Services - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	6.86105	Optimal CCC

Columns Scaled Individually

Control Panel

Method: K Means Cluster

Number of Clusters: 3 Range of Clusters (Optional): .

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count	Step	Criterion
1	592	4	0
2	406		
3	340		

Cluster Means

Cluster	P24	P25	P26	P27
1	13.9425676	1.03885135	1.03716216	1.42567568
2	9.25123153	1.96551724	1.93842365	1.75369458
3	23.5735294	1.92352941	1.95294118	1.40588235

Cluster Standard Deviations

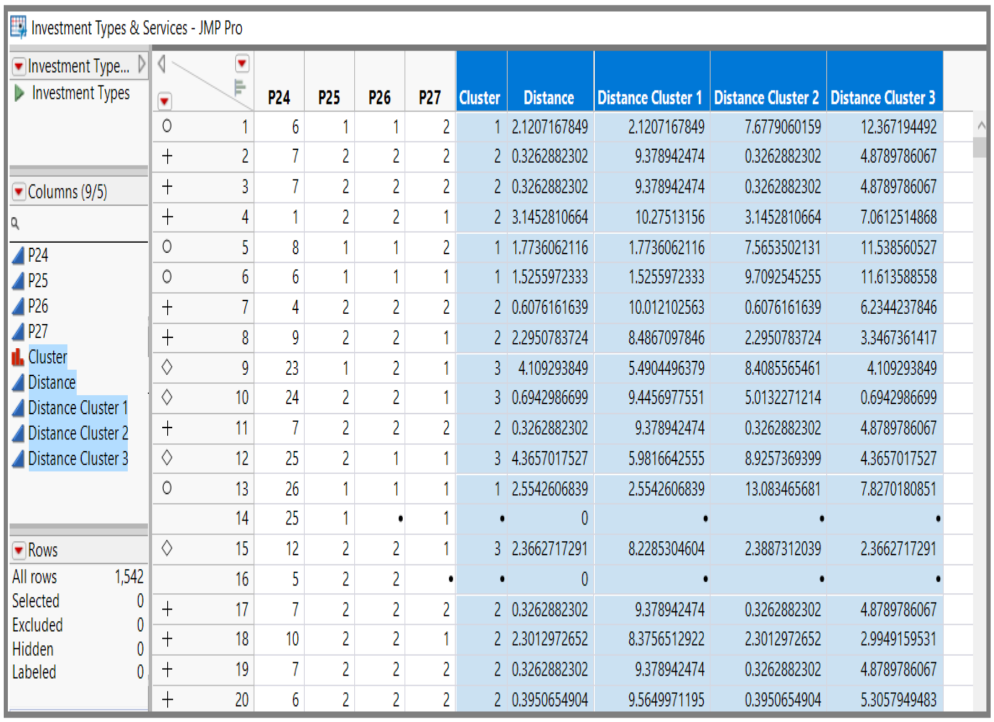
Cluster	P24	P25	P26	P27
1	8.92650617	0.19324059	0.18915902	0.49444504
2	5.14575113	0.18246561	0.2403845	0.43085851
3	5.41597902	0.26574958	0.21176471	0.49106198

Figure 4-2. Iterative Clustering of investment type data by K Means technique in JMP

Figure 4-2 depicts the results of the iterative clustering of investment type data using the K-Means technique in JMP. The number of clusters is determined based on the CCC, with three clusters being indicated for the investment type data. The figure displays the cluster summary, which includes the count of each cluster, with 592 observations in the first cluster, 406 in the second, and 340 in the third. The mean and standard deviation values are also displayed for each cluster. The script for the K-Means clustering process, including the parameters for single-step iteration, the number of clusters, and various visualizations of the data, is provided for reference:

```
K Means Cluster(
 Y( :P24, :P25, :P26, :P27 ),
 {Single Step( 0 ), Number of Clusters( 3 ), K Means Cluster,
 Go(
        Show Biplot Rays( [0, 0, 1] ),
        Biplot( 1 ),
        Scatterplot Matrix( 1 ),
        Biplot 3D( 1 )
 )},
 SendToReport(
        Dispatch( {}, "Cluster Comparison", OutlineBox, {Close( 1 )} ),
        Dispatch(
                {"K Means NCluster=3"},
```

```
                "Cluster Standard Deviations",
                OutlineBox,
                {Close( 0 )}
        ),
        Dispatch(
                {"K Means NCluster=3", "Biplot"},
                "2",
                ScaleBox,
                {Format( "Fixed Dec", 12, 1 ), Min( -1.5 ), Max( 1.5 ), Inc( 0.5 ),
                Minor Ticks( 0 )}
        )
    )
)
```


Investment Types & Services - JMP Pro

Investment Type...
Investment Types

Columns (9/5)
P24
P25
P26
P27
Cluster
Distance
Distance Cluster 1
Distance Cluster 2
Distance Cluster 3

Rows
All rows 1,542
Selected 0
Excluded 0
Hidden 0
Labeled 0

		P24	P25	P26	P27	Cluster	Distance	Distance Cluster 1	Distance Cluster 2	Distance Cluster 3
○	1	6	1	1	2	1	2.1207167849	2.1207167849	7.6779060159	12.367194492
+	2	7	2	2	2	2	0.3262882302	9.378942474	0.3262882302	4.8789786067
+	3	7	2	2	2	2	0.3262882302	9.378942474	0.3262882302	4.8789786067
+	4	1	2	2	1	2	3.1452810664	10.27513156	3.1452810664	7.0612514868
○	5	8	1	1	2	1	1.7736062116	1.7736062116	7.5653502131	11.538560527
○	6	6	1	1	1	1	1.5255972333	1.5255972333	9.7092545255	11.613588558
+	7	4	2	2	2	2	0.6076161639	10.012102563	0.6076161639	6.2344237846
+	8	9	2	2	1	2	2.2950783724	8.4867097846	2.2950783724	3.3467361417
◇	9	23	1	2	1	3	4.109293849	5.4904496379	8.4085565461	4.109293849
◇	10	24	2	2	1	3	0.6942986699	9.4456977551	5.0132271214	0.6942986699
+	11	7	2	2	2	2	0.3262882302	9.378942474	0.3262882302	4.8789786067
◇	12	25	2	1	1	3	4.3657017527	5.9816642555	8.9257369399	4.3657017527
○	13	26	1	1	1	1	2.5542606839	2.5542606839	13.083465681	7.8270180851
	14	25	1	•	1	•	0	•	•	•
◇	15	12	2	2	1	3	2.3662717291	8.2285304604	2.3887312039	2.3662717291
	16	5	2	2	•	•	0	•	•	•
+	17	7	2	2	2	2	0.3262882302	9.378942474	0.3262882302	4.8789786067
+	18	10	2	2	1	2	2.3012972652	8.3756512922	2.3012972652	2.9949159531
+	19	7	2	2	2	2	0.3262882302	9.378942474	0.3262882302	4.8789786067
+	20	6	2	2	2	2	0.3950654904	9.5649971195	0.3950654904	5.3057949483

Figure 4-3. A part of distances for each cluster of investment type data

Figure 4-3 displays the addition of a column in the data table which represents the cluster number assigned to each data point based on its distance to the respective cluster center. The distances to each cluster center are also saved as separate columns in the same data table. This clustering of rows is performed on numerical variables with a specified number of clusters for investment type data.

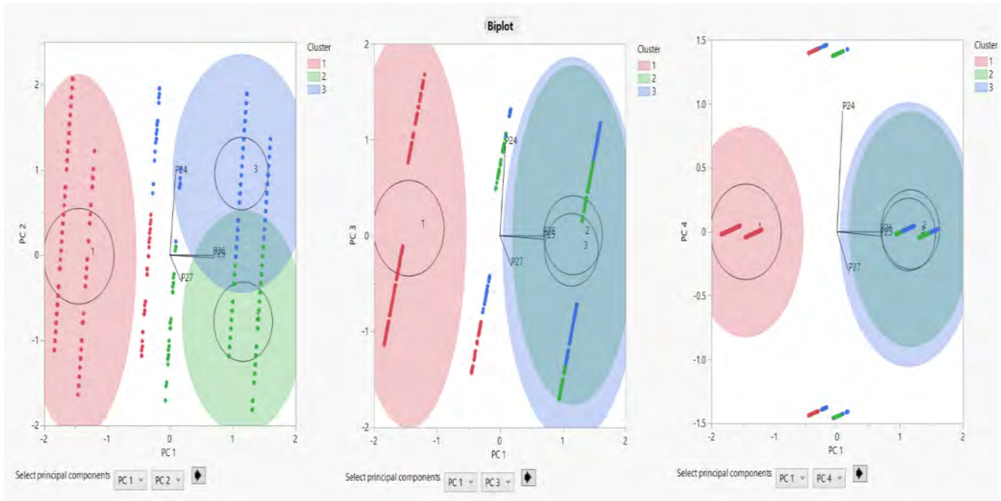


Figure 4-4. Clusters in the first two principal components of investment type data

Figure 4-4 presents a biplot of the investment type data based on the first two principal components. The biplot displays the relationships between the points and clusters in the PC1-PC2, PC1-PC3, and PC1-PC4 pairs.

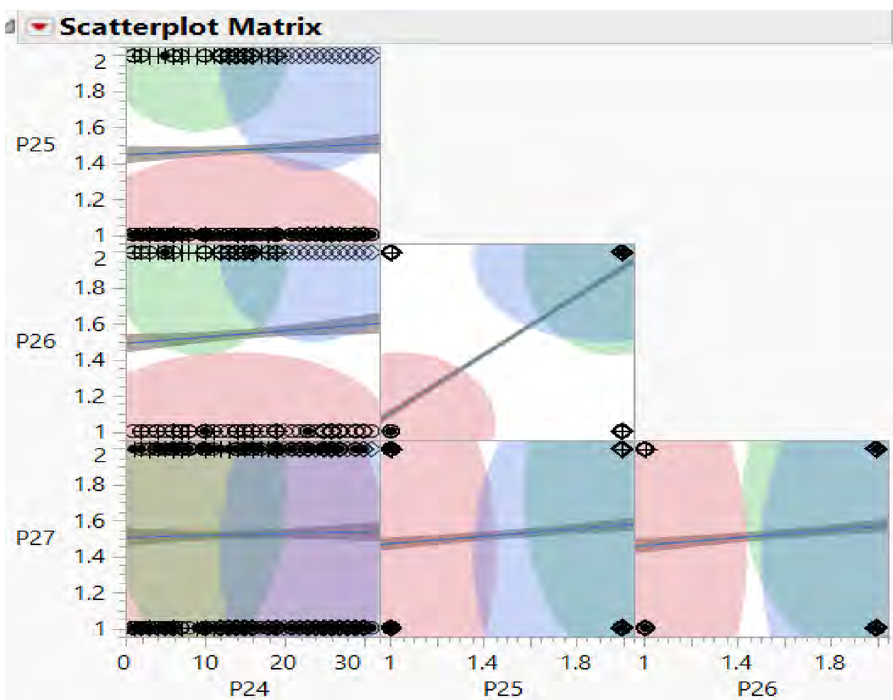


Figure 4-5. Regression line and confidence interval on the scatterplot matrix for investment type data

Figure 4-5 illustrates the regression line and confidence interval on the scatterplot matrix for the investment type data. The region inside the ellipses on the scatterplot matrix clearly demarcates the area occupied by the Y variables of the investment type data. Figure 4-6 presents a three-dimensional (3D) biplot of the points and clusters in the first three principal components of the investment type data. The eigenvalues of the four columns of data are 1.8465613, 0.9965518, 0.967521, and 0.1893659, respectively.

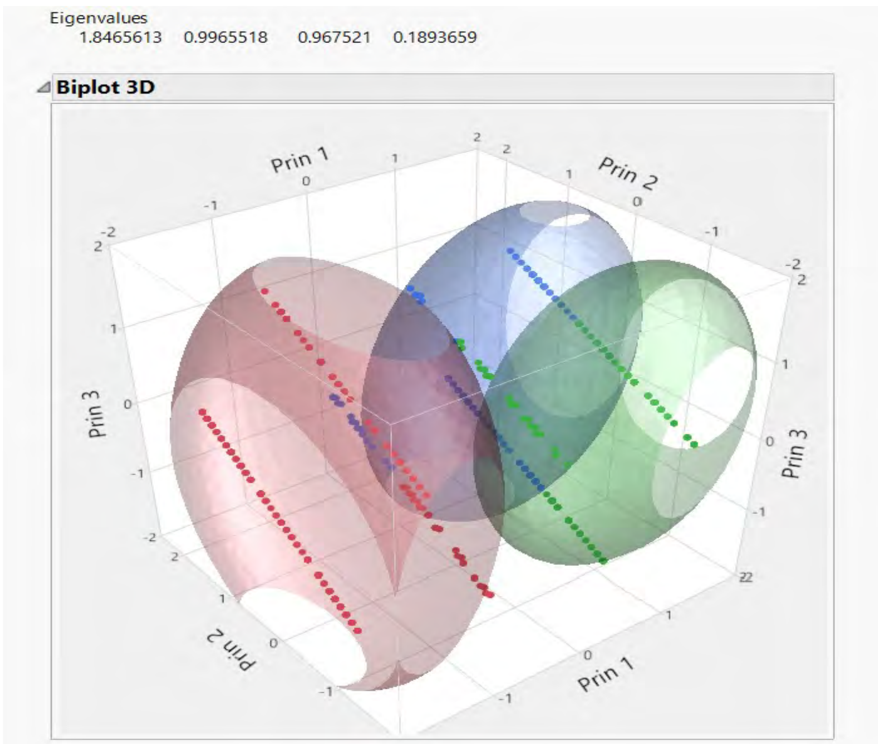


Figure 4-6. First three principal components of investment type data

Figure 4-7 displays a scatterplot matrix, along with confidence ellipses, based on the current number of clusters in the investment type data. To assess the effectiveness of the utilized clustering method with a large dataset, 10,000 simulated samples of investment type data were generated using JMP. The simulation results were then used to construct a new data table, incorporating the estimated cluster mixing probabilities, means, and standard deviations for each cluster, as depicted in Figure 4-8.

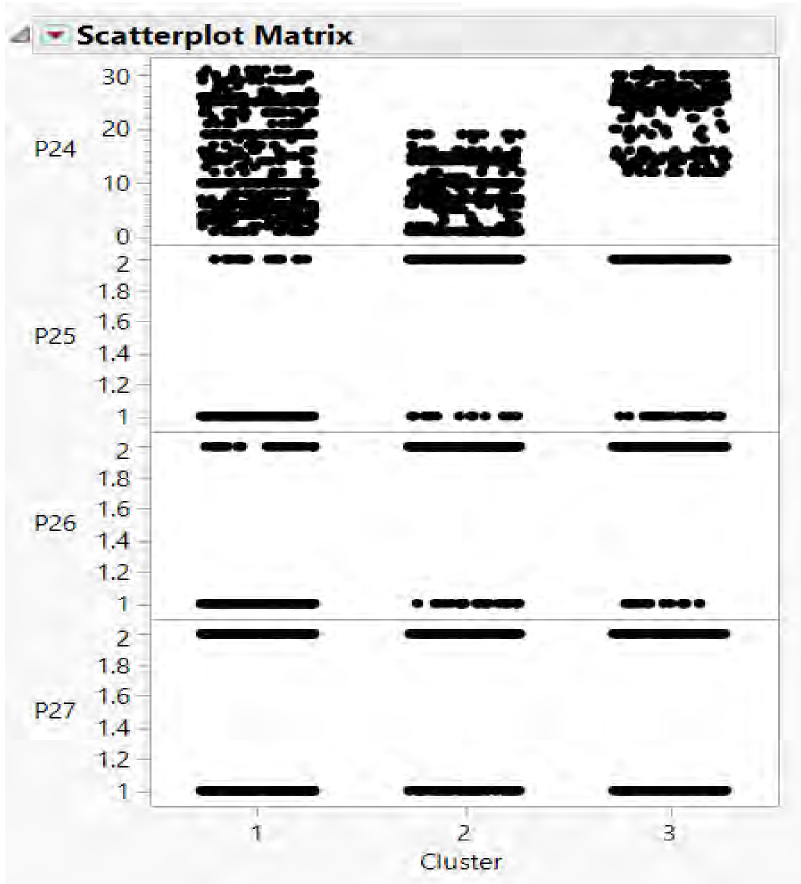


Figure 4-7. A Scatterplot matrix with confidence ellipses based on the clusters of investment type data

The description of each cluster are as follows:

Cluster 1: Suitable investment products for this group include stocks/shares on the stock market, mutual funds, and government securities. These respondents have not invested in the stock market in the last three years and do not regularly monitor stock performance. Many of them did not invest in government bonds.

Cluster 2: Suitable investment products for this group include stocks/shares on the stock exchange, mutual funds, voluntary pension funds, and government securities. These respondents have invested in the stock market in the last three years and regularly monitor stock performance. Many of them have investments in government bonds.

Cluster 3: Suitable investment products for this group include stocks/shares, voluntary pension funds, and government securities. These respondents have been investing in the stock market for the last three years and regularly monitor stock performance. Many of them did not invest in government bonds.

Simulate Clusters 10000 samples - JMP Pro

Simulate Clus...

Columns (5/0)

- P24
- P25
- P26
- P27
- Cluster

Rows

All rows	10,000
Selected	0
Excluded	0
Hidden	0
Labeled	0

	P24	P25	P26	P27	Cluster
1	18.287906602	1.5862597697	1.5156412618	1.6116394867	1
2	17.848493115	1.7344980988	1.6764332057	0.9610399885	3
3	23.976954204	2.0584888745	2.2584742178	2.3329528448	3
4	19.650301413	1.072122683	1.4123901386	1.1167468755	1
5	17.708549416	1.2159721779	0.8165702657	1.4723481249	1
6	6.0168303373	1.8315323173	1.5946110811	1.7617527796	2
7	12.796618552	2.0480838705	1.8110781659	0.9559294054	2
8	24.337651396	1.9225432642	2.0671020189	1.0409415512	3
9	29.849528611	0.9592742333	1.0669468213	1.4243387162	1
10	7.918834595	1.1682454502	1.1773674291	1.9250109194	1
11	17.939891571	2.1275926765	1.6262474187	1.7195493991	2
12	23.548408934	0.8463336537	1.1016040097	1.2030431638	1
13	6.9659331416	1.1669313802	0.8423863026	0.8827679553	1
14	8.9038906168	1.921076582	1.9839612455	1.4506485828	2
15	4.1984101168	0.8961507405	1.0104170628	0.8550215482	1
16	3.2628552519	2.0266679755	1.5866828697	2.2375936828	2
17	3.720247273	0.9446531428	0.9195116811	0.8104852529	1
18	27.431710591	2.4935360104	2.0156824075	1.3012363448	3
19	15.551739532	1.8787626184	2.0493778755	1.4884638353	2
20	12.084699385	2.0683484923	1.756770316	1.6017739108	2
21	-2.211936881	1.2145900277	0.9796638314	0.9903824225	1
22	12.335517516	2.0985295488	1.9596949818	1.4998139008	2
23	3.2000188885	2.1261208343	1.9379036773	2.1235380391	2
24	10.93821439	1.2226463826	0.8106613877	1.6320853206	1
25	5.1213465824	1.8129495779	1.6627662841	1.3518483472	2
26	28.59197483	1.7915061031	2.1373497364	1.7982000656	3
27	19.337806518	1.6203059403	1.8290379199	2.4093172836	3
28	11.930904831	0.7847535869	0.9916031275	1.9950953755	1
29	11.779351866	2.0434374107	2.1016030869	1.2139917959	2
30	18.68982151	2.0312777941	1.5107376559	1.316567647	3
31	11.350449364	0.6645149128	0.8850254154	1.5113938894	1
32	19.90168657	1.6362569466	1.9800113111	1.7909432477	3
33	18.800964488	0.7342837645	0.9191999497	1.9910568344	1

Figure 4-8. A part of the simulated investment type data table

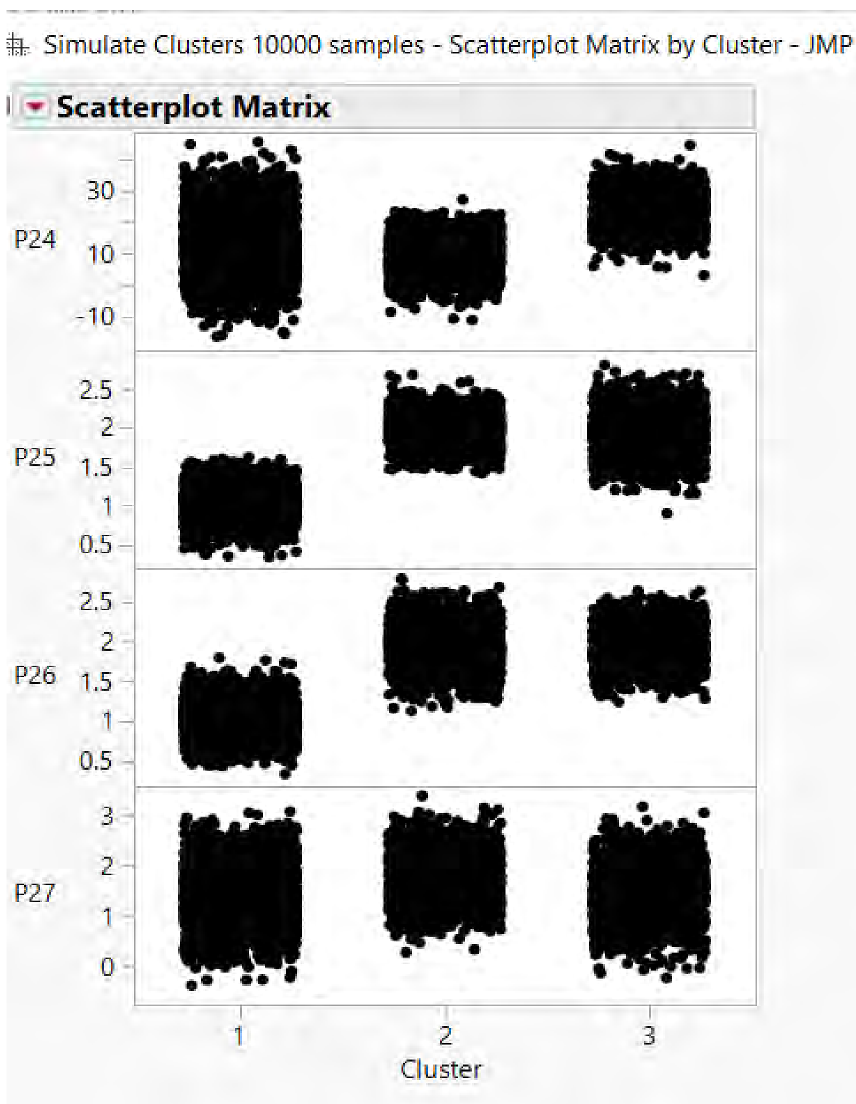


Figure 4-9. A scatterplot matrix based on the clusters of simulated investment type data

Figure 4-9 displays a scatterplot matrix with confidence ellipses, which are generated based on the current number of clusters in simulated investment data. The figure illustrates the optimization of the clusters through simulations performed on a scale of 10,000. This scatterplot matrix provides a visual representation of the clustering patterns and the distribution of the investment data.

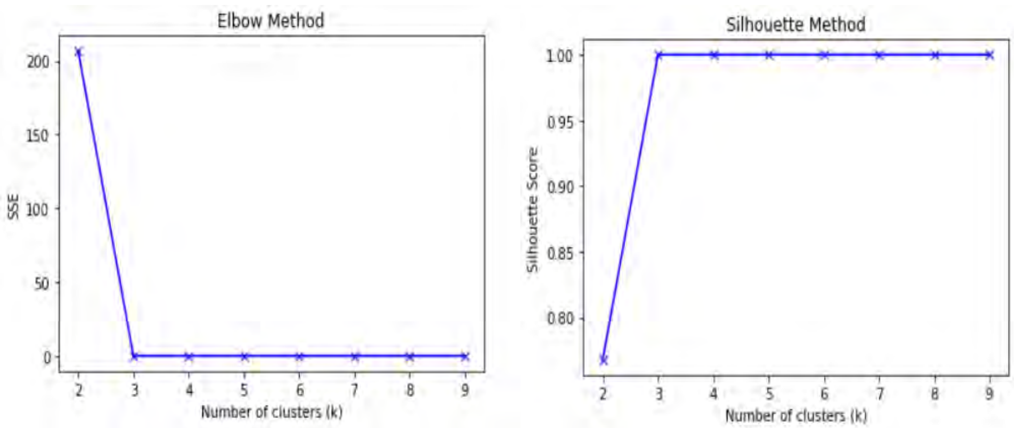


Figure 4-10. SSE and Silhouette score for each value of cluster (Elbow & Silhouette Methods)

Figure 4-10 depicts a two-dimensional plot of three mixture components that sum to a constant. The plot provides a visual representation of how the three components are mixed to form a single, overall mixture. First, the data is loaded, and then set a range of the possible numbers of clusters to try (from 2 to 9). Then, the K-Means algorithm is used to cluster the data for each value of k and calculate the SSE for each k. The Sum of Squared Errors (SSE) values are plotted against k to generate the Elbow curve and look for the "elbow point" where the SSE starts to level off. This can give us an indication of the optimal number of clusters. Also, the Silhouette score is calculated for each k, which measures how well each data point fits into its assigned cluster compared to the other clusters. The Silhouette scores are plotted against k to identify the value of k that maximizes the Silhouette score, which can also give us an indication of the optimal number of clusters. It is important that it is needed to preprocess the data for removing missing values before clustering. The script is written in Python and is designed to run in JMP. The script uses the Silhouette and Elbow methods to determine the optimal number of clusters for the investment type cluster. Specifically, the script reads an Excel file containing the data, preprocesses it as needed, and then applies the KMeans clustering algorithm with a range of possible numbers of clusters:

```
# Import libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
from sklearn.cluster import KMeans
from sklearn.metrics import silhouette_score

# Load data from Excel file
data = pd.read_excel("file path ")

# Preprocess data (if needed)
# ...

# Set range of possible number of clusters
k_range = range(2, 10)

# Calculate SSE (Sum of Squared Errors) for each k
sse = []
for k in k_range:
    kmeans = KMeans(n_clusters=k, random_state=0)
    kmeans.fit(data)
    sse.append(kmeans.inertia_)

# Plot the Elbow curve
plt.plot(k_range, sse, 'bx-')
plt.xlabel('Number of clusters (k)')
plt.ylabel('SSE')
plt.title('Elbow Method')
plt.show()

# Calculate Silhouette score for each k
sil_scores = []
for k in k_range:
    kmeans = KMeans(n_clusters=k, random_state=0)
    kmeans.fit(data)
    labels = kmeans.labels_
    sil_scores.append(silhouette_score(data, labels))

# Plot the Silhouette scores
plt.plot(k_range, sil_scores, 'bx-')
plt.xlabel('Number of clusters (k)')
plt.ylabel('Silhouette Score')
plt.title('Silhouette Method')
plt.show()
```

4.2. Proposing ANFIS-based Investment Recommender Systems

The ANFIS is a type of FIS that generates single-output FISs. In ANFIS, the system parameters are optimized using input and output training data through grid partitioning and a combination of training algorithms. The ANFIS operates using "IF_THEN" rules based on triangular MFs and the architecture of a FIS consists of three parts: fuzzy rules, MFs, and the reasoning mechanism to generate output.

Each ANFIS has several inputs and one output for the investment type, where the inputs can be either fuzzy or non-fuzzy. The inputs have a maximum of one and a minimum of zero MFs. The system was implemented using the fuzzy logic toolbox in MATLAB. The ANFIS process consists of six steps, including importing data, designing the FIS, preparing data, generating the FIS, training the FIS, and modeling the FIS.

The proposed system offers two levels of service to potential investors: the differential investment type service and the integrated investment type service. The differential investment type service includes six separate ANFISs for demographic, decision, personality traits, experiences, financial, and managerial traits data, while the integrated investment type service offers a combined investment type recommender ANFIS. Both levels of service provide customized recommendations for the type of investment.

To create the inputs for the combined investment type recommender ANFIS, data from the six categories were clustered separately. The data clustering process is explained in the subheadings of "Clustering Input Demographic Data", "Clustering Input Key Decision Factors Data", "Clustering Input Personality Traits Data", "Clustering Input Experiences Data", "Clustering Input Financial Data", and "Clustering Input Managerial Traits Data".

4.2.1. Demographics ANFIS

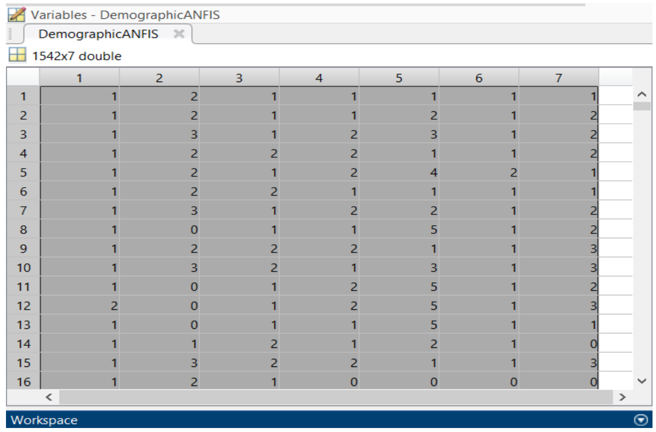
The questionnaire included several questions designed to gather demographic information from respondents. Specifically, six questions (input 1-6) were asked to obtain information on the respondents' age, gender, education level, income, occupation, and marital status. The purpose of collecting this information was to gain a deeper understanding of the relationship between demographics and investment decision-making. By analyzing the responses, we can gain insight into how demographic factors are linked to the selection of specific investment types and products.

Table 4-2. MFs of the demographics ANFIS inputs

Input1	MFs	Gender	Frequency
	MF1	Male	1307
	MF2	Female	191
Input2	MFs	Age/year	Frequency
	MF1	15-34	359
	MF2	35-54	387
	MF3	55-79	100
Input3	MFs	Location	Frequency
	MF1	Budapest	784
	MF2	Other	704
Input4	MFs	Education	Frequency
	MF1	College or university economics	564
	MF2	College or university non-economics	596
	MF3	Postgraduate	73
	MF4	Other	278
Input**5**	MFs	Job	Frequency
	MF1	Employee middle management	231
	MF2	Small-medium business	115
	MF3	Graduate freelance	69
	MF4	Employed lower manager	138
	MF5	Subordinate intellectual worker	659
	MF6	Skilled worker	51
	MF7	Employed senior management	67
	MF8	Micro or self-employed	88
	MF9	Other	80
Input6	MFs	Income /HUF	Frequency
	MF1	Under 200,000	1385
	MF2	200,004-349999	104
	MF3	Above 350,000	7

The ANFIS-based investment recommendation system utilizes six demographic inputs and one investment type output. These inputs include: 1) gender, 2) age, 3) location, 4) education, 5) job, and 6) income. The system utilizes two MFs for gender, with option 1 "male" assigned to MF1 and option 2 "female" assigned to MF2. For age, three MFs are utilized, with option 1 "15-34 years old" assigned to MF1, option 2 "35-54 years old" assigned to MF2, and option 3 "55-79 years old" assigned to MF3. The location input utilizes two MFs, with option 1 "Budapest" assigned to MF1 and option 2 "other location" assigned to MF2. The education input utilizes four MFs, with option 1 "College or university economics" assigned to MF1, option 2 "College or university non-economics" assigned to MF2, option 3 "Postgraduate" assigned to MF3, and option 4 "Other" assigned to MF4. The input 5 "job" with 9 MFs pertains to the potential investors' occupation. Option 1, "Employee middle management," is assigned to MF1; option 2, "Small medium business," is assigned to MF2; option 3, "Graduate freelance," is assigned to MF3; option 4, "Employed lower manager," is assigned to MF4; option 5, "Subordinate intellectual worker," is assigned to MF5; option 6, "Skilled worker," is assigned to MF6; option 7, "Employed senior management," is assigned to MF7; option 8, "Micro or self-employed," is assigned to MF8; and option 9, "Other," is assigned to MF9. Similarly, the input 6 "income" with 3 MFs is related to the potential investors' monthly income. Option 1, "Under 200,000 HUF," is assigned to MF1; option 2, "200,004-349999 HUF," is assigned to MF2; and option 3, "Above 350,000 HUF," is assigned to MF3 (as shown in Table 4-2).

The output of the system is defined as three clusters of investment types/products. These clusters encompass a variety of investment products, including listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The system utilizes 1542 train data pairs for inputs and output, with the implication method set to Min and aggregation method set to Max.



Variables - DemographicANFIS

DemographicANFIS

1542x7 double

	1	2	3	4	5	6	7
1	1	2	1	1	1	1	1
2	1	2	1	1	2	1	2
3	1	3	1	2	3	1	2
4	1	2	2	2	1	1	2
5	1	2	1	2	4	2	1
6	1	2	2	1	1	1	1
7	1	3	1	2	2	1	2
8	1	0	1	1	5	1	2
9	1	2	2	2	1	1	3
10	1	3	2	1	3	1	3
11	1	0	1	2	5	1	2
12	2	0	1	2	5	1	3
13	1	0	1	1	5	1	1
14	1	1	2	1	2	1	0
15	1	3	2	2	1	1	3
16	1	2	1	0	0	0	0

Workspace

Figure 4-11. A part of imported data to MATLAB to propose the demographics ANFIS

The data imported into MATLAB consisted of 7 columns, with 6 columns related to the demographics of potential investors and one column related to investment type clusters. The fuzzy function utilized the inputs and output for demographics ANFIS. The DemographicANFIS was designed using the Sugeno type as a new FIS. As shown in Figure 4-11, the imported data in MATLAB was used as inputs and output to propose the DemographicANFIS.

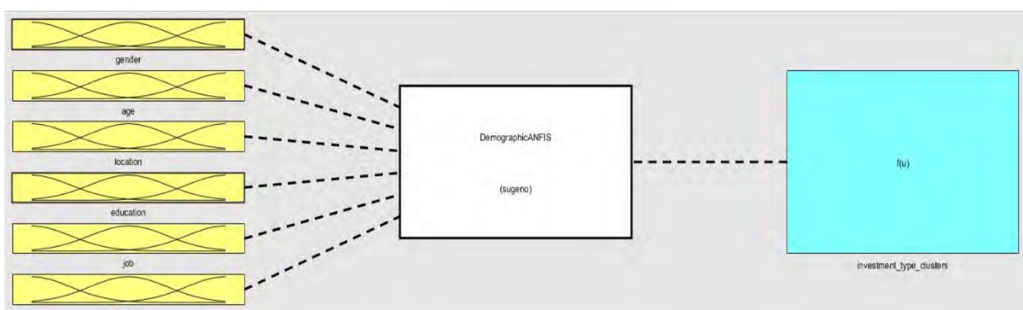


Figure 4-12. The properties of the DemographicANFIS

The proposed Demographics ANFIS system, named "DemographicANFIS," is depicted in Figure 4-12 along with its properties. The system utilizes six inputs, including gender, age, location, education, job, and income, to generate an output of investment type clusters.

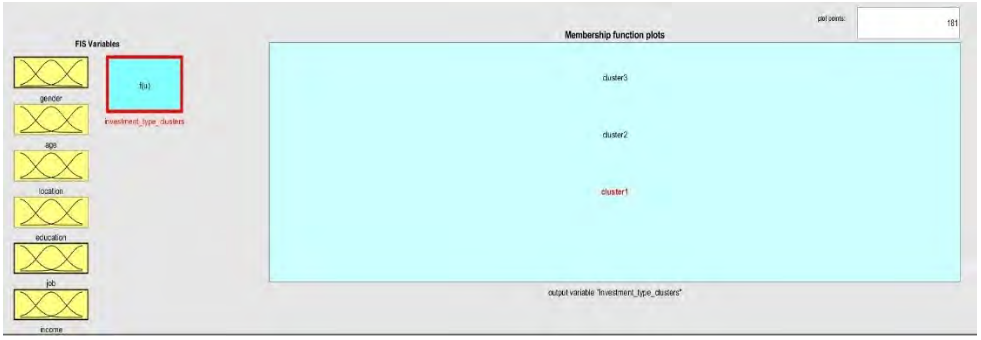


Figure 4-13. Output MFs in the DemographicANFIS

As shown in Figure 4-13, the shape of the MFs for the output in the Demographic ANFIS is presented. Constant type MFs have been utilized for the three investment types of "Cluster 1", "Cluster 2", and "Cluster 3”.

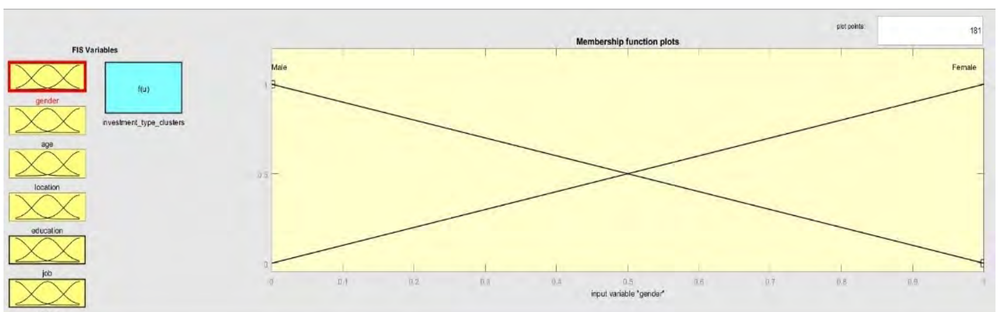


Figure 4-14. MFs shape for input 1 (gender) in the DemographicANFIS

Figure 4-14 illustrates the shape of the MFs for input 1 of the DemographicANFIS. The MF shape utilized is a trimf with two MFs, 'male' and 'female'."

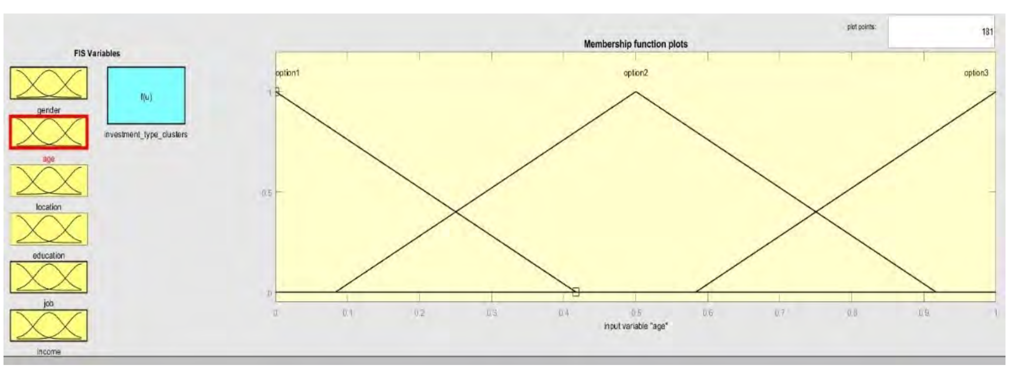


Figure 4-15. MFs shape for input 2 (age) in the DemographicANFIS

Figure 4-15 illustrates the MFs' shape for input 2 of the DemographicANFIS. The MF shape employed is trimf with 3 options, namely "option 1", "option 2", and "option 3"."

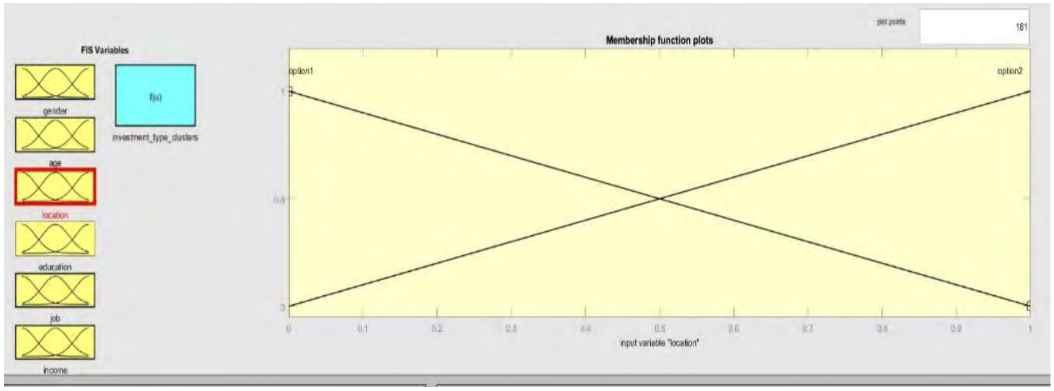


Figure 4-16. MFs shape for input 3 (location) in the DemographicANFIS

As depicted in Figure 4-16, the MFs for input 3 of the DemographicANFIS are illustrated. The shape of the MFs is triangular (trimf) with two options, namely "option 1" and "option 2."

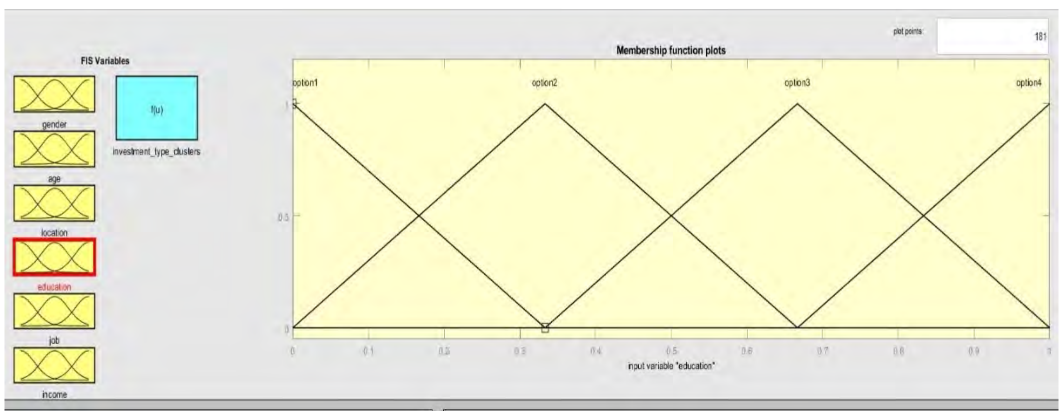


Figure 4-17. MFs shape for input 4 (education) in the DemographicANFIS

Figure 4-17 illustrates the MF shape for input 4 of the DemographicANFIS. The MF shape is a trimf with four options, namely "Option 1", "Option 2", "Option 3", and "Option 4"."

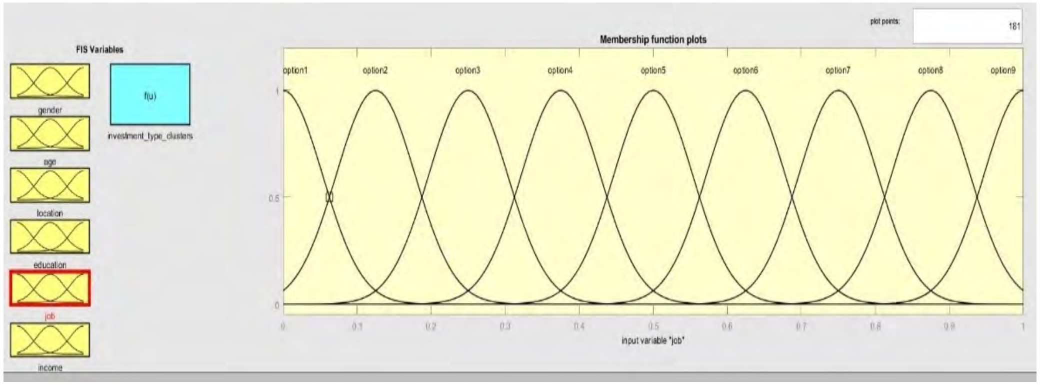


Figure 4-18. MFs shape for input 5 (job) in the DemographicANFIS

Figure 4-18 illustrates the shape of the MFs for input 5 of the DemographicANFIS model. The MFs are represented using the gaussmf with a total of nine options, including "Option 1", "Option 2", "Option 3", "Option 4", "Option 5", "Option 6", "Option 7", "Option 8", and "Option 9".

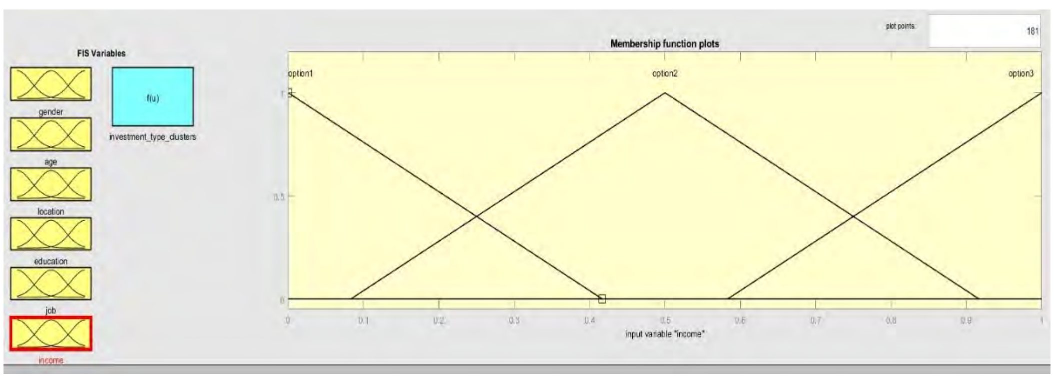


Figure 4-19. MFs shape for input 6 (income) in the DemographicANFIS

Figure 4-19 illustrates the shape of the MFs for input 6 of the DemographicANFIS. The MF shape is a trimf with three options, namely "Option 1", "Option 2", and "Option 3"."

4.2.2. Proposing DemographicANFIS

The Demographic ANFIS is a system that utilizes ANNs and fuzzy logic to make investment recommendations based on demographic information. It consists of six inputs: 1) gender, 2) age, 3) location, 4) education, 5) job, and 6) income.

These inputs are used to determine the best investment type for a particular individual or group. The system produces a single output, which is the recommended investment type, based on the analysis of the input demographics.

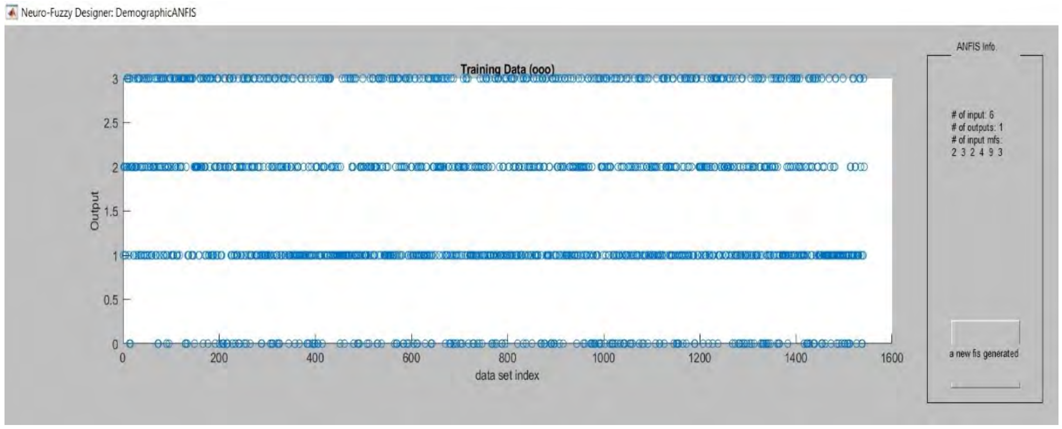


Figure 4-20. Prepared data in the DemographicANFIS

Figure 4-20 depicts the data preparation process for the DemographicANFIS model. The data is prepared for the subsequent steps of training and validation. A grid partition method is utilized, and the optimization is performed using a hybrid approach with an error tolerance of 0 and 3 epochs. This results in the generation of a new FIS (FIS) known as the DemographicANFIS. The x-axis represents the data set index, which comprises 1542 entries, while the y-axis shows the output of the prepared data across four levels (4-3). Level 0 represents the original data, while levels 1-3 display the three levels of prepared data.

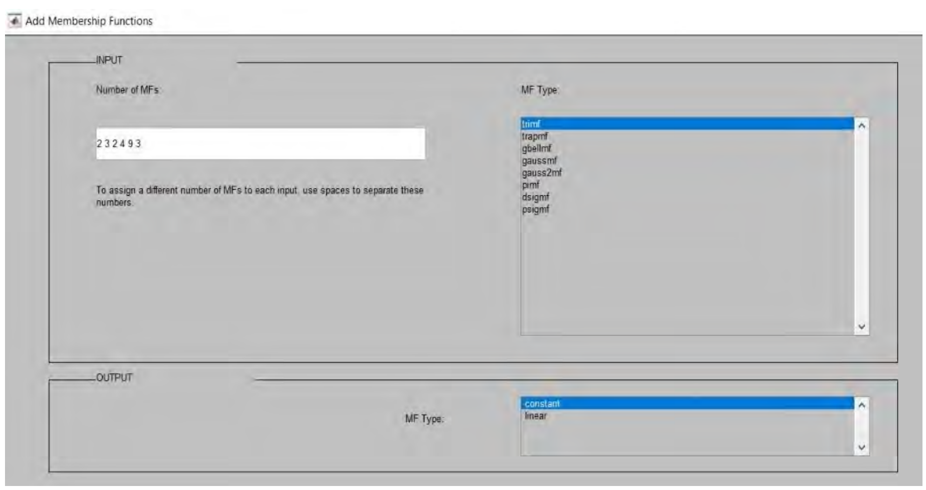


Figure 4-21. Information for generating the DemographicANFIS as a new FIS

Figure 4-21 displays a summary of the MFs for the Demographic ANFIS system. This ANFIS was created as a new FIS (FIS) with six fuzzy inputs and one non-fuzzy output. Figure 4-22 presents the error of the trained data in the Demographic ANFIS grid. The training process was a hybrid with three epochs, and the error for each epoch was approximately 0.87. A RMSE value between 0.2 and 0.5 indicates a good ability of the model to predict data, while a value greater than 0.75 suggests excellent accuracy and prediction capabilities. In this case, the RMSE value exceeds 0.75. The information regarding the training process of the Demographic ANFIS system is provided below:

ANFIS Info:

Number of nodes: 2647
Number of linear parameters: 1296
Number of nonlinear parameters: 69
Total number of parameters: 1365
Number of training data pairs: 1542
Number of checking data pairs: 0
Number of fuzzy rules: 1296

The training process for the Demographic ANFIS began and completed after two epochs, with the minimal training RMSE value of 0.86683.

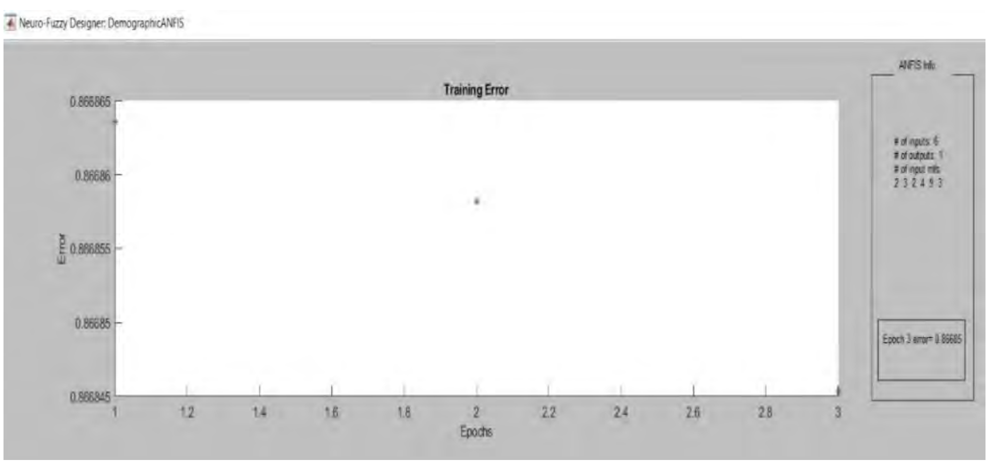


Figure 4-22. RMSE in the DemographicANFIS

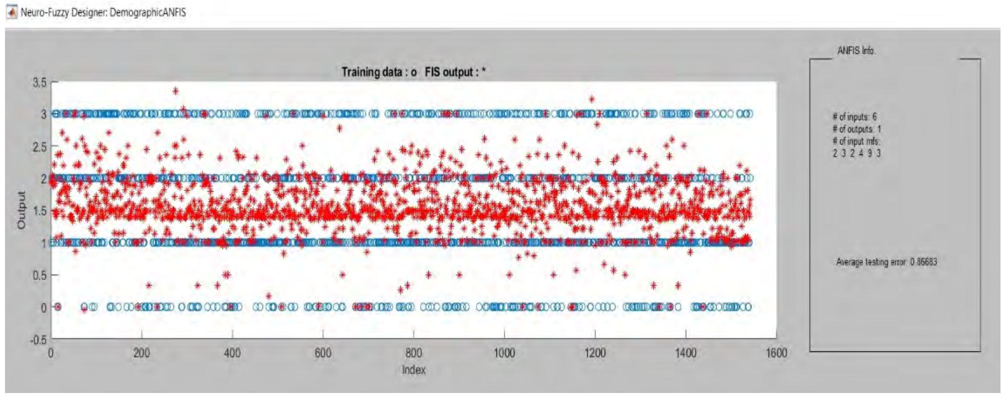


Figure 4-23. Trained Data in the DemographicANFIS

Figure 4-23 depicts the trained data in the Demographic ANFIS system, represented in red. The average training error of the system is 0.86683 and it generated 720 rules. The figure indicates that the least amount of validation error occurred in epoch 3, which suggests that the parameters of the model are well-aligned with the training data. As a result, it can be inferred that the Demographic ANFIS system demonstrates exceptional generalization performance. The training data fully showcases the characteristics of the FIS-modeled data. Figure 4-24 displays a portion of the generated rules in the proposed Demographic ANFIS system, presented in verbose format. The rules can be adjusted, added, or deleted based on expert opinions and investor feedback.

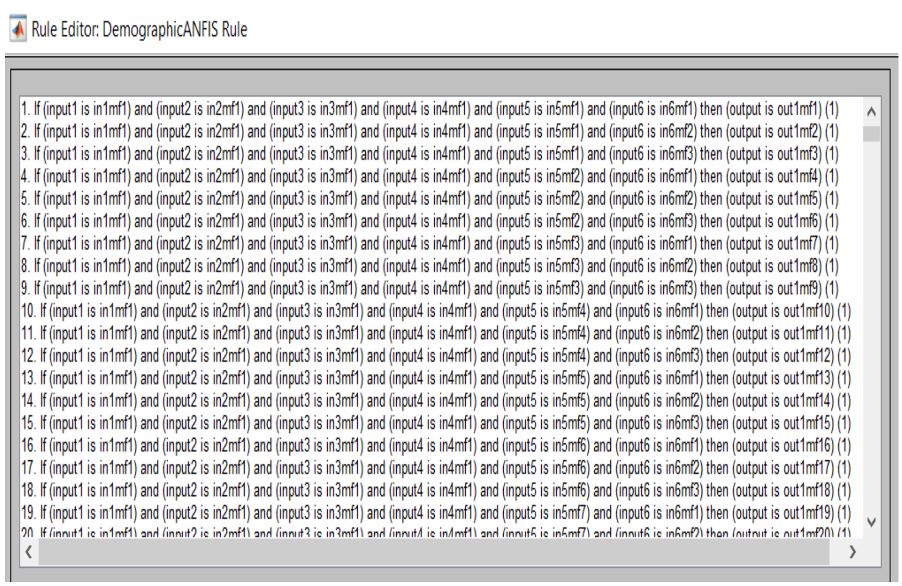


Figure 4-24. A part of the generated rules in the DemographicANFIS

Figure 4-25 showcases a portion of the rule viewer, depicting the open system of the Demographic ANFIS. The figure highlights the presence of 1296 rules and 101 plot points. As depicted, the trained Demographic ANFIS system generates numerous rules with adjusted MF parameters, utilizing the training data. These rules are related to the epoch where the training error is at its minimum value, which in this case is epoch 3. To further reduce errors, the number of epochs can be increased in instances where multiple epochs have the same minimum training error (RMSE). The red line in the figure provides an opportunity to manipulate and improve the MF status, resulting in improved system performance. Ideally, any necessary changes to the MF status can be made based on system feedback. Attachment 2 provides further information regarding the rule generation process by the Demographic ANFIS.

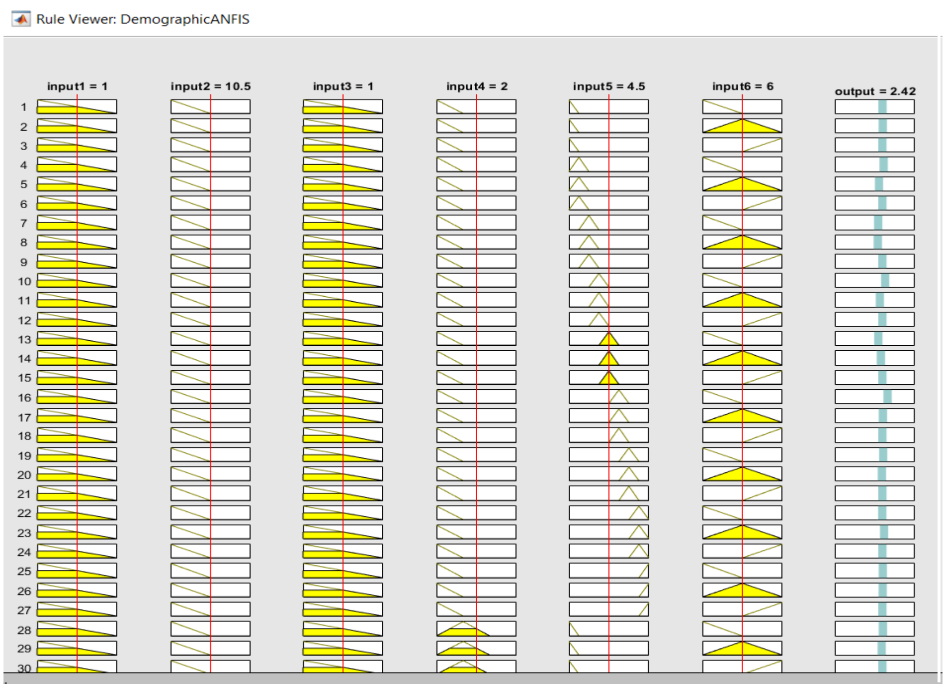


Figure 4-25. A part of the rule viewer in the DemographicANFIS

The proposed Demographic ANFIS system aims to determine the relationship between demographics and investment type selection. Figures 4-26 (a-d) present 3D graphs that depict the impact of certain input pairs on investment type. These surface graphs are nonlinear and monolithic, displaying investment type recommendations for given inputs. These figures provide visual representation of the relationship between different pairs of demographic inputs and investment type. The six sub-figures, labeled as (a) to (f), each demonstrate the impact of a specific pair of demographic inputs on investment type recommendations. These 3D graphs allow for the exploration of the nonlinear and monolithic relationships between demographic inputs and investment type. Figure 4-26a shows the relationship between gender and income on investment type shows how these two demographic factors influence investment type recommendations. For example, the graph might show that higher income individuals are more likely to choose a certain type of investment compared to lower income individuals.

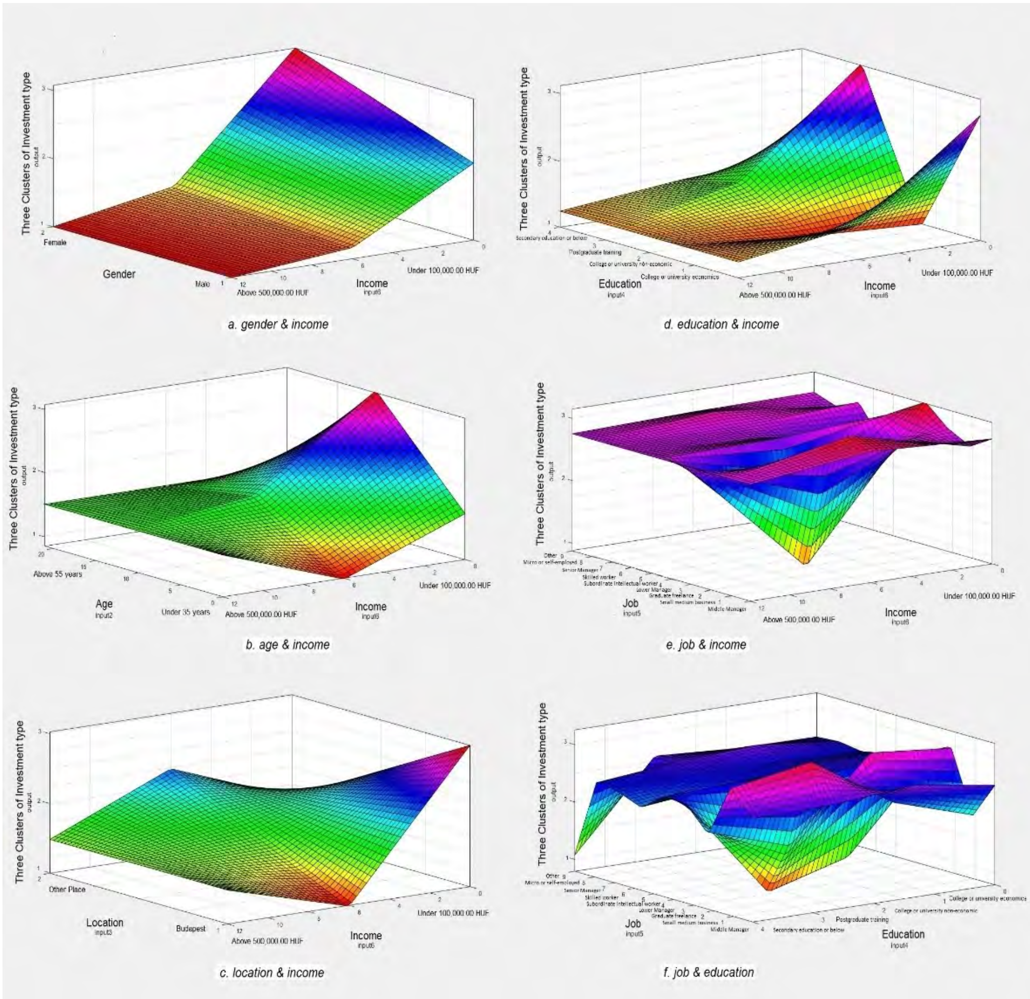


Figure 4-26. Effectiveness of the relations of each pair of demographics inputs on investment type

Figure 4-26b shows the relationship between age and income on investment type displays how these two demographic factors impact investment type recommendations. The graph may reveal a pattern where older individuals with higher income tend to choose a certain type of investment, while younger individuals with lower income prefer another type of investment. Figure 4-26c shows the relationship between location and income on investment type illustrates the relationship between these two demographic factors and investment type recommendations. The graph may indicate that individuals in certain locations with higher income tend to choose a specific type of investment, while those in other locations with lower income prefer another type of investment. Figure 4-26d shows the relationship between education and income on investment type shows how these two demographic factors affect investment type recommendations. For example, the graph might demonstrate that individuals with higher education and higher income tend to choose a certain type of investment, while those with lower education and lower income prefer another type of investment. Figure 4-26e shows the relationship between job and income on investment type displays the impact of these two demographic factors on investment type recommendations. The graph may indicate that individuals with certain jobs and higher income tend to choose a specific type of investment, while those with different jobs and lower income prefer another type of investment. Figure 4-26f shows the relationship between job and education on investment type illustrates the relationship between these two demographic factors and investment type recommendations. The graph may reveal a pattern where individuals with certain jobs and higher education tend to choose a specific type of investment, while those with different jobs and lower education prefer another type of investment. These figures provide insight into the complex relationships between demographic inputs and investment type and can be useful for making informed investment recommendations based on demographic data.

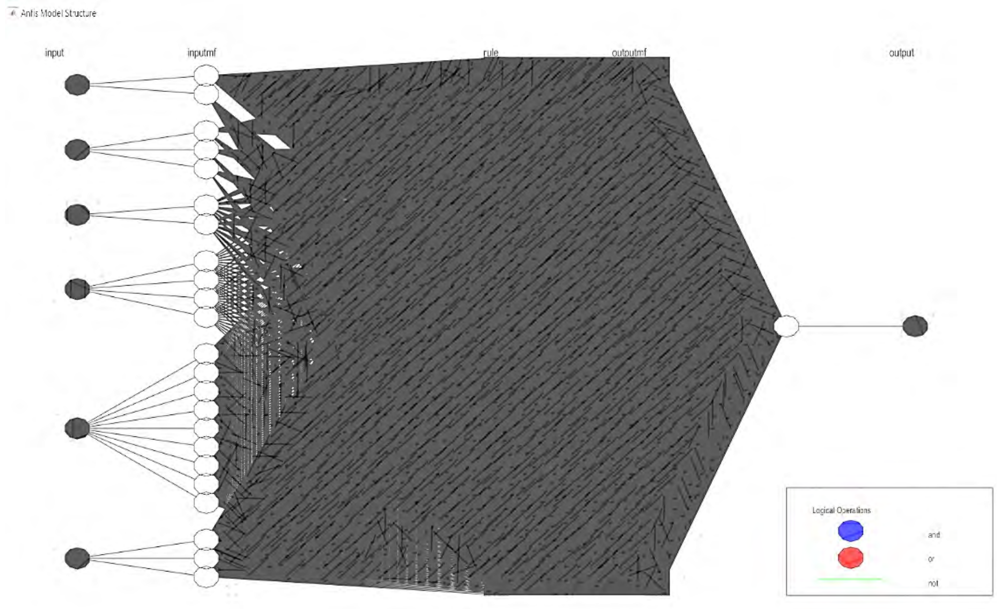


Figure 4-27. DemographicANFIS Model Structure

Figure 4-27 depicts the structure of the Demographic ANFIS model, which serves as an investment recommendation system based on demographic information. The figure showcases the inputs, MFs, various layers of the ANFIS system, and the output, which is a recommendation for investors to select an investment type based on demographic clusters. The model encompasses several key components including fuzzification, which translates the inputs into a fuzzy set, implication rules, which determine the relationships between inputs and outputs, normalization, which scales the output of the implication rules, defuzzification, which converts the fuzzy output into a crisp value, and integration, which combines the outputs of multiple MFs into a single, aggregated output MF. Based on the comparison between actual and predicted values, it can be concluded that the Demographic ANFIS structural model provides an excellent fit for the investment recommendation system considering potential investors' demographics.

4.2.3. Key Decision Factors ANFIS

The questionnaire included several questions that aimed to measure the key decision factors of potential investors. Four questions (input 1-4) were asked from the respondents to gather information about their key decision factors and to gain a deeper understanding of the relationship between these factors and their attitudes towards finance. This information was used to establish an ANFIS-based IRS model, which was designed to customize investment types based on the key decision factors of potential investors. The ANFIS layer of the model was based on the four questions that were asked to identify the effective factors in investment decision making. The study employed a survey to gather data on participants' opinions regarding environmental awareness factors, system value, and expectations for returns with socially conscious investments. Respondents were asked to indicate their level of agreement with a series of statements using a 5-point Likert scale, with options ranging from completely disagree to completely agree. Four key decision factors were analyzed in the study: environmental factors, system value, potential for higher returns with socially conscious investments, and potential for lower returns with socially conscious investments. The output of the study was investment type or product, which was divided into three clusters. The investment types analyzed in the study included listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The study employed an ANFIS to analyze the data. The ANFIS operates using "IF-THEN" rules, which are based on MFs of the inputs. The architecture of the FIS used in the study consisted of three parts: fuzzy rules, MFs, and a reasoning mechanism to generate output. The ANFIS used in the study had four inputs, each with three MFs. The inputs had gaussmf MFs with a maximum value of one and a minimum value of zero. The data was processed using the fuzzy logic toolbox in MATLAB. The proposed system generates an output for the investment type/product, which is defined as three clusters represented by three constant MFs. The data for the investment type includes a variety of investment products, such as listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The number of training data pairs used for all inputs and output is 1542, with the implication method set to Min and aggregation method set to Max.

1542x5 table

	1 EnvironmentalFactors	2 SystemValue	3 ExpectToHigherReturns	4 AgreeWithLowerReturns	5 InvestmentType	6
1	3	3	3	2	2	
2	3	3	3	3	2	
3	3	3	3	3	2	
4	4	4	4	3	2	
5	4	4		2	4	
6	3	4	2	4	2	
7	4	5	1	3	2	
8	1	1	1	1	2	
9	5	4	3	4	2	
10	4	4	4	3	2	
11	1	1	1	1	2	
12	3	1	1	2	2	
13	1	1	1	1	2	
14	5				2	
15	5	5	4	3	2	

Figure 4-28. Apart from imported data to MATLAB to propose the DecisionANFIS

The data imported into MATLAB consisted of five columns, four of which pertained to key factors in decision making for potential investors, and the fifth column related to investment type clusters. The ANFIS algorithm was utilized to incorporate the inputs and output for the key decision factors in a fuzzy function. A new FIS of the Sugeno type, named DecisionANFIS, was designed to process the imported data. Figure 4-28 illustrates a portion of the data imported into MATLAB, which serves as inputs and output for the proposed key decision factors in the ANFIS system.

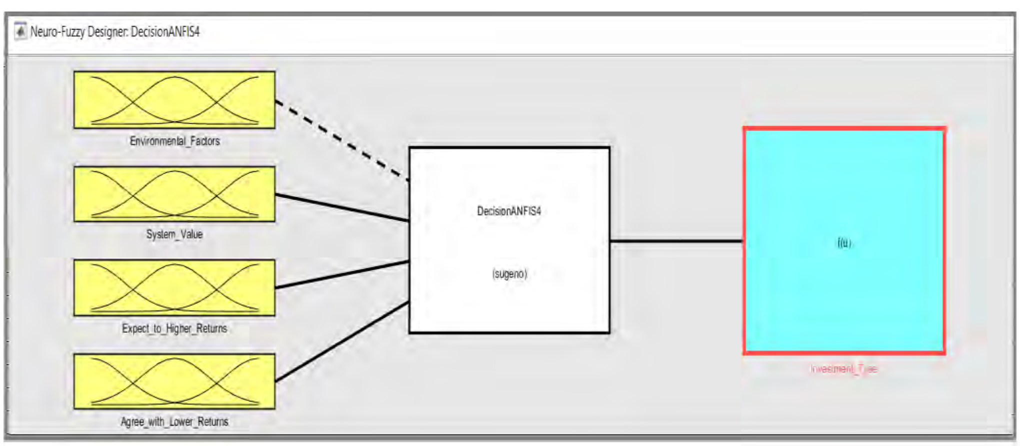


Figure 4-29. The properties of the DecisionANFIS

As depicted in Figure 4-29, the design of the "DecisionANFIS" system, utilizing ANFIS, is presented. The system incorporates four key inputs, including: 1) Environmental factors, 2) System value, 3) Expectation of higher return, and 4) Agreement with lower return. The system's output is investment type clusters.

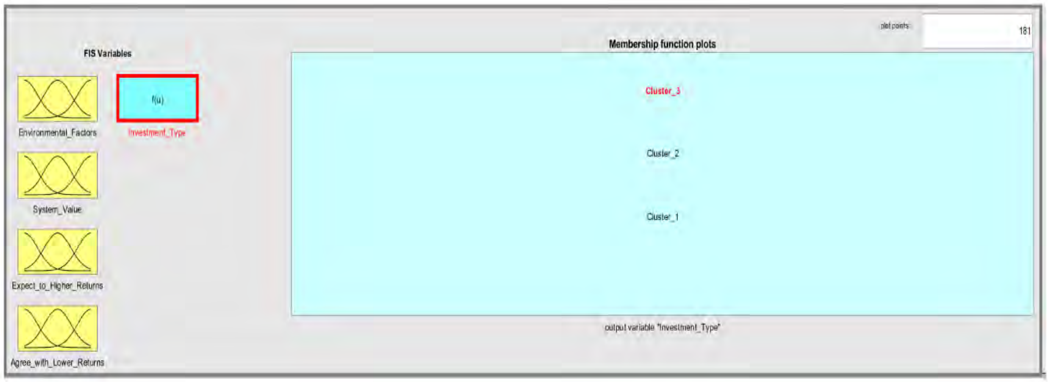


Figure 4-30. Output MFs in the DecisionANFIS

Figure 4-30 illustrates the shape of the MFs for the output in the DecisionANFIS model. The model employs a constant MF for each of the three investment types, namely "Cluster 1", "Cluster 2", and "Cluster 3".

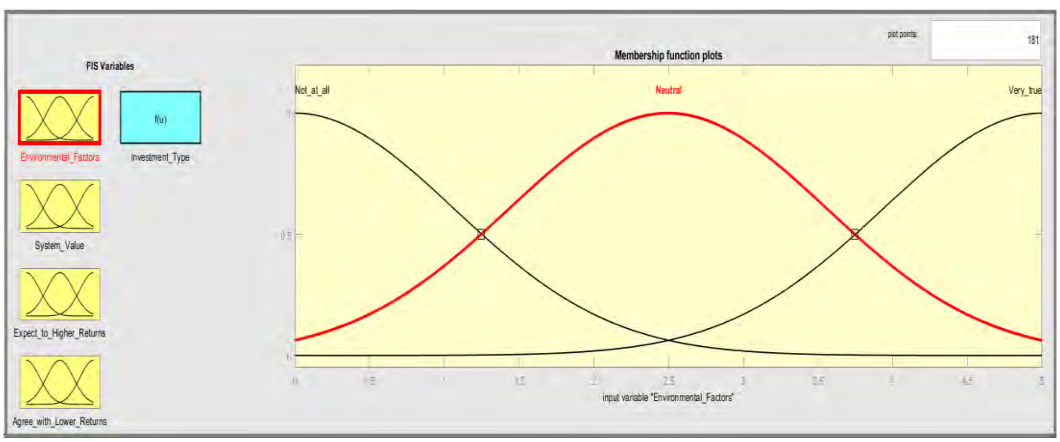


Figure 4-31. Inputs' MFs in the DecisionANFIS

As shown in Figure 4-31, the MFs for all inputs of the DecisionANFIS are illustrated. The chosen MF shape is gaussian, with three designated categories:

"Not at all," "Neutral," and "Very true." These MFs provide a clear visualization of the input membership to the corresponding fuzzy sets.

4.2.4. Proposing DecisionANFIS

The DecisionANFIS is a component of the investment recommendation system framework. It consists of four inputs that represent key decision factors that influence potential investors in their investment choices. These factors include environmental factors, system value, higher return, and lower return. The DecisionANFIS has one output, investment type, which is divided into three clusters. The data that is considered key decision factors was explained in an earlier chapter. Figure 4-19 represents the data that was prepared for the training and validation stages of the DecisionANFIS. The data was trained using a grid partition method, and the optimization process was a hybrid approach with an error tolerance of 0 and three epochs. The purpose of this component of the system is to help identify the most influential factors in investment decisions and make recommendations based on that information.

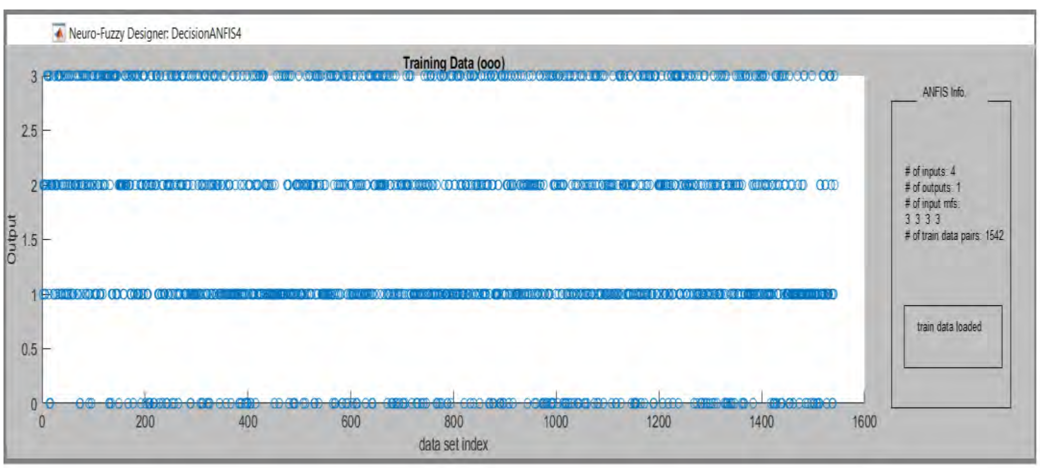


Figure 4-32. Prepared data in the DecisionANFIS

The Decision ANFIS is generated as a new FIS and Figure 4-32 provides a summary of the MFs for the Decision ANFIS. The Decision ANFIS is a new FIS that has been created based on the input and rules generated by other ANFIS models, such as the Demographic ANFIS. Figure 4-33 provides information about the MFs used in the Decision ANFIS, which are used to process the input data and make decisions based on the defined rules. The MFs help to interpret the input data and provide a more accurate and interpretable representation of the relationships between the inputs and the investment type.

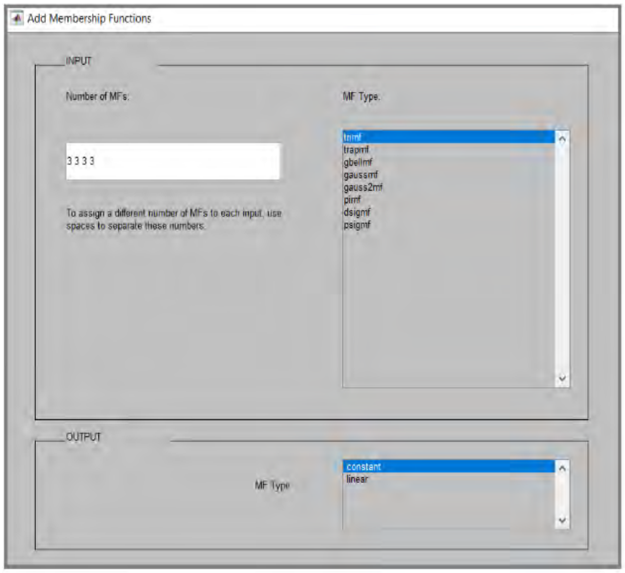


Figure 4-33. Information for generating the DecisionANFIS as a new FIS

The DecisionANFIS is a newly generated FIS that consists of 4 inputs and 1 output. The data set used for this system has 1542 entries, with the x-axis representing the index and the y-axis representing the distribution of the output, based on investment-type clusters. Figure 4-34 illustrates the training error, expressed as the RMSE, in the DecisionANFIS grid. During the training process, a Gaussmf was utilized for each of the four inputs. The system was trained as a hybrid with 3 epochs, resulting in an error of approximately 0.93 for each epoch. The DecisionANFIS system was designed to predict investment type clusters based on the inputs provided. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model.

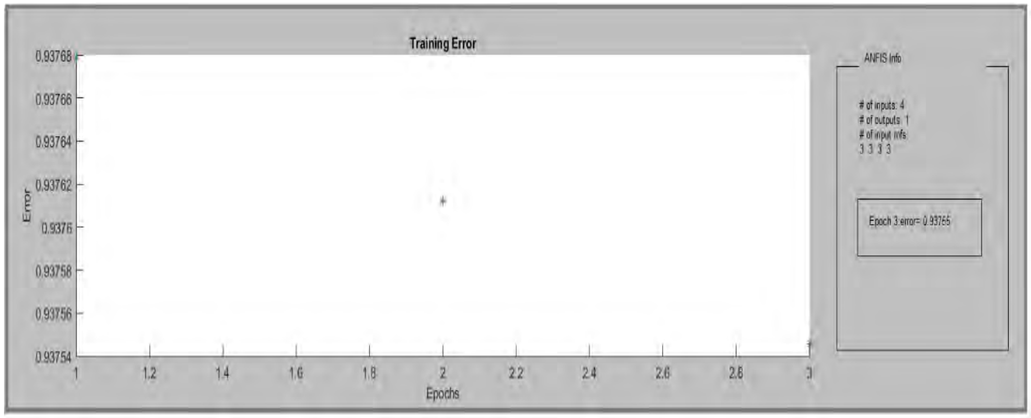


Figure 4-34. Training RMSE in the DecisionANFIS

The information provided is about the training process of the DecisionANFIS. According to the information, the DecisionANFIS system has 193 nodes, with 81 linear parameters and 36 nonlinear parameters, making for a total of 117 parameters. The training process was performed using 1542 training data pairs and 0 checking data pairs. The DecisionANFIS system uses 81 fuzzy rules. The training process was initiated and completed at epoch 2, where the training error reached its minimum value of 0.93748. The value is expressed in terms of the RMSE, which measures the difference between the predicted values and the actual values in the training data. The lower the RMSE value, the better the fit of the model to the training data. The training was stopped when the designated epoch number was reached, and the training was completed with a minimum training RMSE of 0.93748.

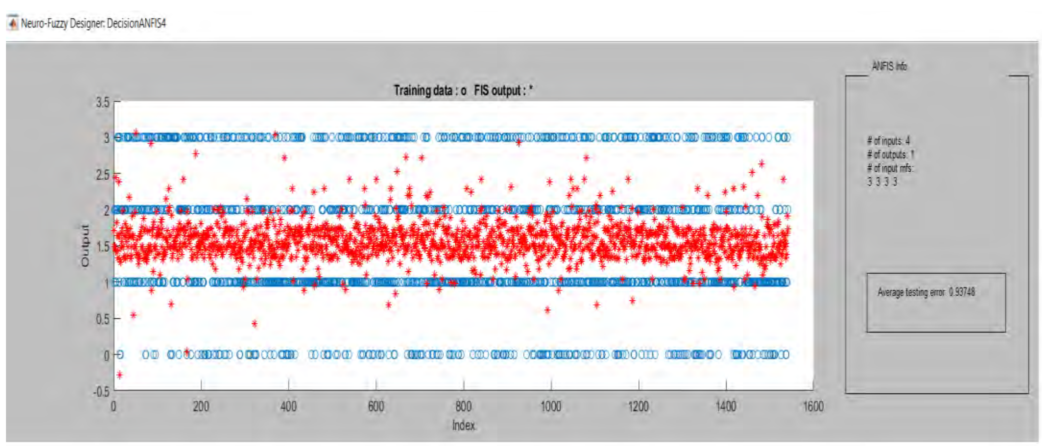


Figure 4-35. Trained data in the DecisionANFIS

Figure 4-35 represents the trained Decision ANFIS system. The average training error, as indicated in the figure, is 0.93748. This figure indicates that the Decision ANFIS system has generated 81 rules. Figure 4-36 is a portion of the rule viewer and depicts a part of the open system of the Decision ANFIS. The figure highlights the presence of 81 rules and 181 plot points. The rule viewer provides an overview of the rules generated by the Decision ANFIS system, which are based on the inputs and training data used to train the system. The plot points in the figure help to visualize the relationship between the inputs and outputs, as well as the performance of the system in making decisions based on the rules generated.

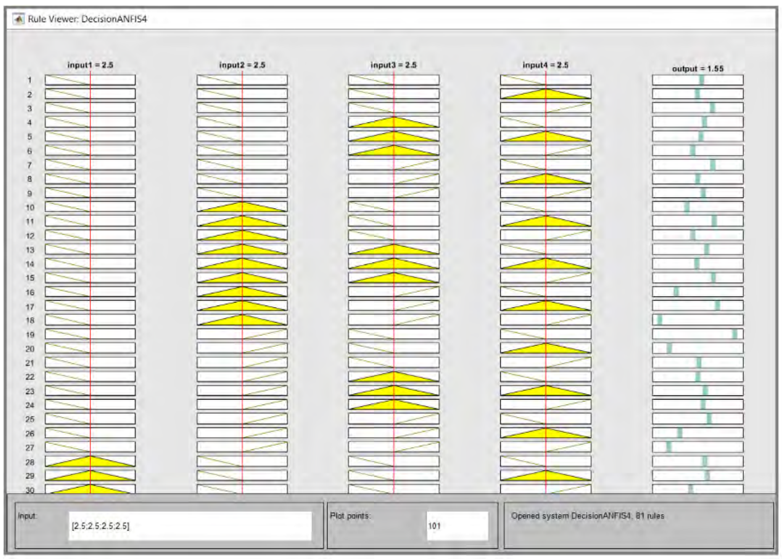


Figure 4-36. A part of the rule viewer in the DecisionANFIS

Figure 4-37 depicts all the generated rules in the proposed Decision ANFIS system in a verbose format. This means that the figure presents all the rules in a clear and detailed manner, making it easy to understand. The verbose format also allows for modifications to be made to the rules, such as adding, changing, or deleting rules. These modifications can be based on expert opinions and investor feedback. This feature makes the Decision ANFIS system flexible and customizable, allowing for personalized recommendations to be made based on specific needs and preferences.

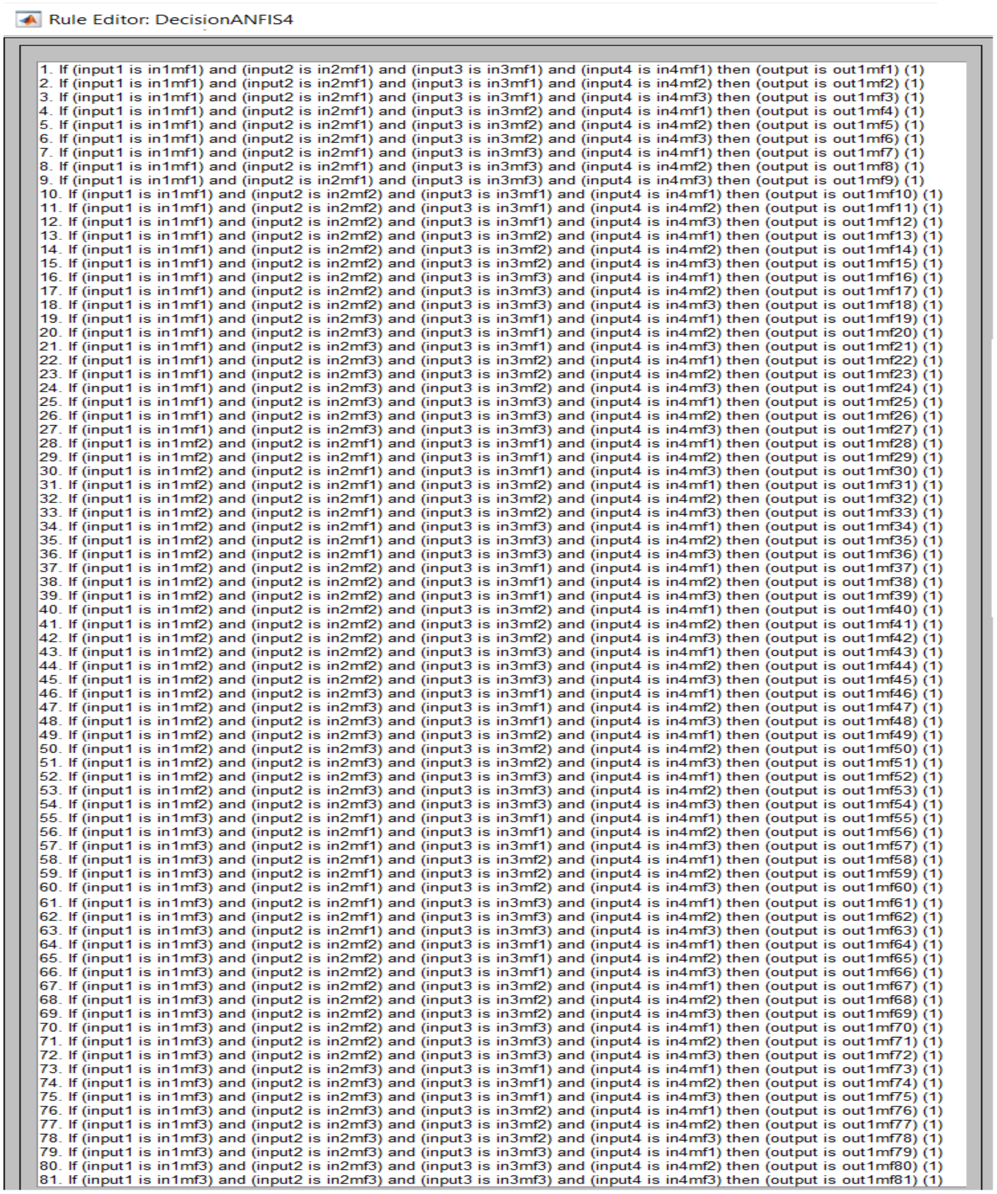


Figure 4-37. All generated rules in the DemographicANFIS

Here is a configuration file for a Decision-based ANFIS (DecisionANFIS) with four inputs and one output. The type of the system is Sugeno and the version of the system is 2.0. The system has 81 rules, and the methods for AND, OR, implication, aggregation, and defuzzification are "prod", "probor", "prod", "sum", and "wtaver", respectively. The inputs, named "input1", "input2", "input3", and "input4", each have a range of [0, 5]. Each input has 3 MFs defined by trimf. The MFs for each input are described in detail in the respective sections of the configuration file. The output, named "output", has a range of [0, 3] and 81 MFs defined as constant MFs. The values for each of the 81 constant MFs are specified in the configuration file:

```
[System]
Name='DecisionANFIS4'
Type='sugeno'
Version=2.0
NumInputs=4
NumOutputs=1
NumRules=81
AndMethod='prod'
OrMethod='probor'
ImpMethod='prod'
AggMethod='sum'
DefuzzMethod='wtaver'

[Input1]
Name='input1'
Range=[0 5]
NumMFs=3
MF1='in1mf1':'trimf',[-2.50000000000211 0.000273256633754278 2.5043873649622]
MF2='in1mf2':'trimf',[-0.000270415661509097 2.50078035535276 5.00021628442632]
MF3='in1mf3':'trimf',[2.49793365855891 5.00050671449916 7.5]

[Input2]
Name='input2'
Range=[0 5]
NumMFs=3
MF1='in2mf1':'trimf',[-2.5 -0.000771830685955874 2.49713252280317]
MF2='in2mf2':'trimf',[6.76704703404358e-05 2.50001118437064 4.99975717761454]
MF3='in2mf3':'trimf',[2.50359229904918 5.00078307851985 7.5]

[Input3]
Name='input3'
Range=[0 5]
NumMFs=3
MF1='in3mf1':'trimf',[-2.49999999431746 0.005034674582037 2.51645973512364]
MF2='in3mf2':'trimf',[-0.00503031860999279 2.50847630177763 4.99891399645676]
MF3='in3mf3':'trimf',[2.51605323132303 5.0034539387042 7.5]

[Input4]
Name='input4'
Range=[0 5]
NumMFs=3
MF1='in4mf1':'trimf',[-2.49999999993703 0.00190170550136344 2.51228576095747]
MF2='in4mf2':'trimf',[-0.00190270047511544 2.50030897395295 5.00159339313379]
MF3='in4mf3':'trimf',[2.49539968257195 4.99840471583401 7.4999999998001]

[Output1]
Name='output'
Range=[0 3]
NumMFs=81
MF1='out1mf1':'constant',[1.90992356871414]
MF2='out1mf2':'constant',[-0.702485966820823]
MF3='out1mf3':'constant',[8.51901389188576]
MF4='out1mf4':'constant',[3.64368113675292]
MF5='out1mf5':'constant',[1.55600103265047]
MF6='out1mf6':'constant',[-3.40811015857288]
MF7='out1mf7':'constant',[8.75725769931706]
MF8='out1mf8':'constant',[-0.5417342374968]
```

MF9='out1mf9':'constant',[2.91508282270922]
MF10='out1mf10':'constant',[-6.98964005134042]
MF11='out1mf11':'constant',[9.57672797539354]
MF12='out1mf12':'constant',[-3.42303444212422]
MF13='out1mf13':'constant',[5.03334742891396]
MF14='out1mf14':'constant',[-1.09166387920432]
MF15='out1mf15':'constant',[9.09753123157907]
MF16='out1mf16':'constant',[-13.7483956139737]
MF17='out1mf17':'constant',[11.6412909004446]
MF18='out1mf18':'constant',[-23.7701060547523]
MF19='out1mf19':'constant',[22.2233738266717]
MF20='out1mf20':'constant',[-17.8657661430125]
MF21='out1mf21':'constant',[0.238612456497012]
MF22='out1mf22':'constant',[-0.41084741798542]
MF23='out1mf23':'constant',[3.9997387624487]
MF24='out1mf24':'constant',[2.78717131951822]
MF25='out1mf25':'constant',[6.47672215490544]
MF26='out1mf26':'constant',[-11.416450566028]
MF27='out1mf27':'constant',[-18.3211290586829]
MF28='out1mf28':'constant',[3.81140413739008]
MF29='out1mf29':'constant',[5.43067276224834]
MF30='out1mf30':'constant',[-4.89085673658399]
MF31='out1mf31':'constant',[-3.16860881533702]
MF32='out1mf32':'constant',[2.70090484215443]
MF33='out1mf33':'constant',[3.5757783376]
MF34='out1mf34':'constant',[6.7391765841315]
MF35='out1mf35':'constant',[-4.97954979609414]
MF36='out1mf36':'constant',[22.8330719877315]
MF37='out1mf37':'constant',[1.74599976855572]
MF38='out1mf38':'constant',[1.4681384492677]
MF39='out1mf39':'constant',[2.13570165184749]
MF40='out1mf40':'constant',[2.36135575184674]
MF41='out1mf41':'constant',[1.55920307607364]
MF42='out1mf42':'constant',[-0.206187636817465]
MF43='out1mf43':'constant',[1.12301906642153]
MF44='out1mf44':'constant',[2.36254155671928]
MF45='out1mf45':'constant',[1.07991907944234]
MF46='out1mf46':'constant',[5.52289182398196]
MF47='out1mf47':'constant',[2.11788921226246]
MF48='out1mf48':'constant',[-0.296028833269121]
MF49='out1mf49':'constant',[-3.21585633876496]
MF50='out1mf50':'constant',[2.46748875415298]
MF51='out1mf51':'constant',[1.98646675824672]
MF52='out1mf52':'constant',[4.99611112555997]
MF53='out1mf53':'constant',[1.4002561076866]
MF54='out1mf54':'constant',[0.712278197133587]
MF55='out1mf55':'constant',[-0.324660349345829]
MF56='out1mf56':'constant',[-5.1037215191204]
MF57='out1mf57':'constant',[-0.0475028993733485]
MF58='out1mf58':'constant',[22.5392226737821]
MF59='out1mf59':'constant',[-0.998564989896525]
MF60='out1mf60':'constant',[2.46400079375895]
MF61='out1mf61':'constant',[-15.6761137828228]
MF62='out1mf62':'constant',[4.79679599030644]
MF63='out1mf63':'constant',[-15.2466520687227]
MF64='out1mf64':'constant',[-0.692158593909971]
MF65='out1mf65':'constant',[4.51135213974123]
MF66='out1mf66':'constant',[1.15236600437226]
MF67='out1mf67':'constant',[3.15517575336455]
MF68='out1mf68':'constant',[1.45094579268146]
MF69='out1mf69':'constant',[1.19117916650417]
MF70='out1mf70':'constant',[1.61045364569496]

MF71='out1mf71':'constant',[0.629468530885751]
MF72='out1mf72':'constant',[2.11046489457963]
MF73='out1mf73':'constant',[1.06525846682759]
MF74='out1mf74':'constant',[2.16643573363876]
MF75='out1mf75':'constant',[1.36806205977103]
MF76='out1mf76':'constant',[3.81785008162497]
MF77='out1mf77':'constant',[1.34119861732613]
MF78='out1mf78':'constant',[1.33094626022074]
MF79='out1mf79':'constant',[1.21882248191825]
MF80='out1mf80':'constant',[1.29075202134171]
MF81='out1mf81':'constant',[1.33624903318031]

The expression "T[Rules]" refers to the rules that are present in a FIS. Each line in the given expression represents a separate rule. A FIS rule consists of a set of antecedents and a corresponding conclusion. The antecedents define conditions under which the rule is applied, and the conclusion represents the output of the system when the antecedents are met. Each line in the expression has the following format:

<Antecedent 1>, <Antecedent 2>, ..., <Antecedent n>, <Conclusion> (Membership degree) : Rule weight

For example, in the first rule:

1 1 1 1, 1 (1) : 1

- The antecedents are 1 1 1 1 which represent the activation levels of the inputs, where 1 represents full activation of that input.
- The conclusion is 1 and the membership degree is (1), indicating that the conclusion is fully activated.
- The rule weight is 1. The rule weight is a value that indicates the importance of the rule.

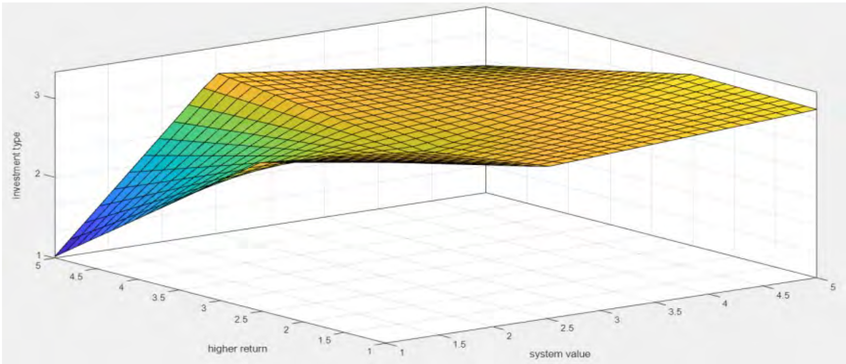
In a similar manner, the other rules can be interpreted. The rules with antecedents 1 1 1 2, 1 1 1 3, 1 1 2 1, 1 1 2 2, and 1 1 2 3 have similar structures, with different antecedents and conclusions

[Rules]
1 1 1 1, 1 (1) : 1
1 1 1 2, 2 (1) : 1
1 1 1 3, 3 (1) : 1
1 1 2 1, 4 (1) : 1
1 1 2 2, 5 (1) : 1
1 1 2 3, 6 (1) : 1
1 1 3 1, 7 (1) : 1
1 1 3 2, 8 (1) : 1
1 1 3 3, 9 (1) : 1
1 2 1 1, 10 (1) : 1

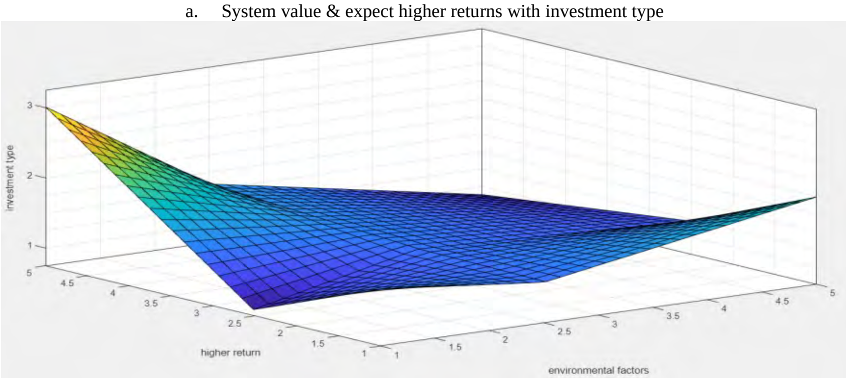
1 2 1 2, 11 (1) : 1
1 2 1 3, 12 (1) : 1
1 2 2 1, 13 (1) : 1
1 2 2 2, 14 (1) : 1
1 2 2 3, 15 (1) : 1
1 2 3 1, 16 (1) : 1
1 2 3 2, 17 (1) : 1
1 2 3 3, 18 (1) : 1
1 3 1 1, 19 (1) : 1
1 3 1 2, 20 (1) : 1
1 3 1 3, 21 (1) : 1
1 3 2 1, 22 (1) : 1
1 3 2 2, 23 (1) : 1
1 3 2 3, 24 (1) : 1
1 3 3 1, 25 (1) : 1
1 3 3 2, 26 (1) : 1
1 3 3 3, 27 (1) : 1
2 1 1 1, 28 (1) : 1
2 1 1 2, 29 (1) : 1
2 1 1 3, 30 (1) : 1
2 1 2 1, 31 (1) : 1
2 1 2 2, 32 (1) : 1
2 1 2 3, 33 (1) : 1
2 1 3 1, 34 (1) : 1
2 1 3 2, 35 (1) : 1
2 1 3 3, 36 (1) : 1
2 2 1 1, 37 (1) : 1
2 2 1 2, 38 (1) : 1
2 2 1 3, 39 (1) : 1
2 2 2 1, 40 (1) : 1
2 2 2 2, 41 (1) : 1
2 2 2 3, 42 (1) : 1
2 2 3 1, 43 (1) : 1
2 2 3 2, 44 (1) : 1
2 2 3 3, 45 (1) : 1
2 3 1 1, 46 (1) : 1
2 3 1 2, 47 (1) : 1
2 3 1 3, 48 (1) : 1
2 3 2 1, 49 (1) : 1
2 3 2 2, 50 (1) : 1
2 3 2 3, 51 (1) : 1
2 3 3 1, 52 (1) : 1
2 3 3 2, 53 (1) : 1
2 3 3 3, 54 (1) : 1
3 1 1 1, 55 (1) : 1
3 1 1 2, 56 (1) : 1
3 1 1 3, 57 (1) : 1
3 1 2 1, 58 (1) : 1
3 1 2 2, 59 (1) : 1
3 1 2 3, 60 (1) : 1
3 1 3 1, 61 (1) : 1
3 1 3 2, 62 (1) : 1
3 1 3 3, 63 (1) : 1
3 2 1 1, 64 (1) : 1
3 2 1 2, 65 (1) : 1
3 2 1 3, 66 (1) : 1
3 2 2 1, 67 (1) : 1
3 2 2 2, 68 (1) : 1
3 2 2 3, 69 (1) : 1
3 2 3 1, 70 (1) : 1
3 2 3 2, 71 (1) : 1
3 2 3 3, 72 (1) : 1

3 3 1 1, 73 (1) : 1
3 3 1 2, 74 (1) : 1
3 3 1 3, 75 (1) : 1
3 3 2 1, 76 (1) : 1
3 3 2 2, 77 (1) : 1
3 3 2 3, 78 (1) : 1
3 3 3 1, 79 (1) : 1
3 3 3 2, 80 (1) : 1
3 3 3 3, 81 (1) : 1

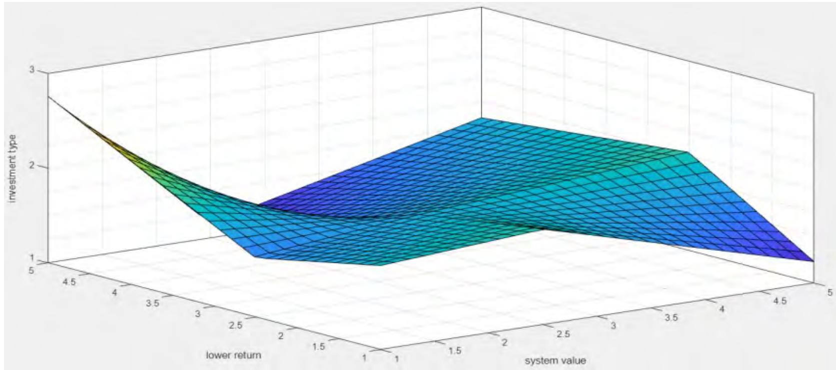
The proposed Decision ANFIS system is essentially a model that seeks to understand the relationship between key decision factors and the type of investment. It does this by analyzing four key decision factors pairs, and the effect each pair has on investment type.



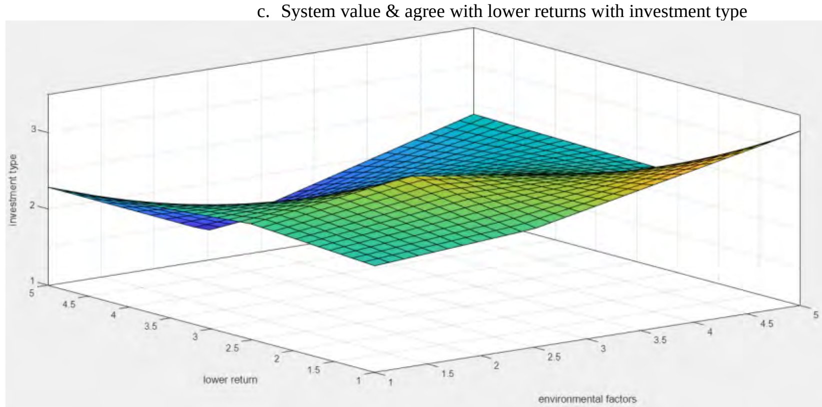
a. System value & expect higher returns with investment type



b. Environmental awareness factors & expect higher returns with investment type



c. System value & agree with lower returns with investment type



d. Environmental awareness factors & agree with lower returns with investment type

Figure 4-38. Effectiveness of the relations of each pair of key decision factors inputs on investment type

These four key decision factor pairs are: Figure 4-38a shows the system value & expectation of higher returns with socially conscious investments - This pair of factors examines the relationship between a potential investor's evaluation of the investment system's value and their expectations of receiving higher returns from socially responsible investments. Figure 4-38b shows the environmental awareness factors & expectation of higher returns with socially conscious investments - This pair of factors analyzes the relationship between a potential investor's level of environmental awareness and their expectation of receiving higher returns from socially responsible investments. Figure 4-38c shows the system value & agreement with lower returns with socially conscious investments - This pair of factors examines the relationship between a potential investor's evaluation of the investment system's value and their agreement to accept lower returns from socially responsible investments. Figure 4-38d shows the environmental awareness factors & agreement with lower returns with socially conscious investments - This pair of factors analyzes the relationship between a potential investor's level of environmental awareness and their agreement to accept lower returns from socially responsible investments. In summary, figures 4-38 visualize the relationship between each pair of key decision factors and their impact on investment type. This figure is a collection of four 3D graphs (a-d), each graph representing one of the key decision factor pairs. The graphs are nonlinear and monolithic, meaning that they display the investment type recommendations for a given set of inputs. By examining the relationship between these key decision factors and investment type, the Decision ANFIS system provides valuable insight into the factors that influence an individual's investment choices.

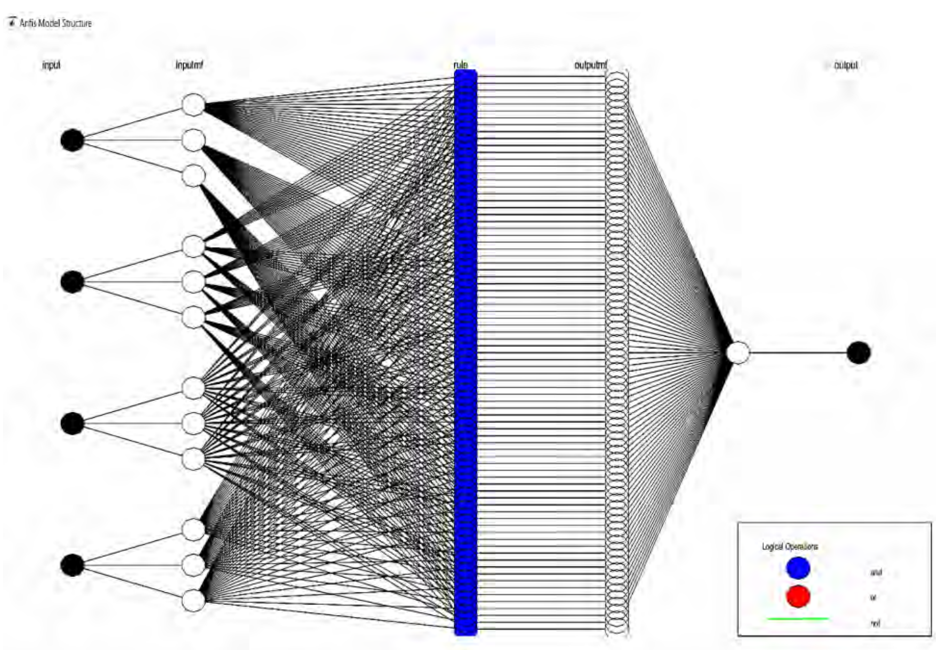


Figure 4-39. DecisionANFIS Model Structure

Figure 4-39 displays the structure of the Decision ANFIS Model. The model represents the inputs, MFs, and various layers of the ANFIS, culminating in an investment type recommendation for the investor based on cluster analysis. The model's structure includes the following components:

- Fuzzification: This layer maps the input values to MF values.
- Implication rules: In this layer, the system evaluates the rules and calculates the consequent MF values.

- Normalization: This layer normalizes the outputs of the rules.
- Defuzzification: This layer converts the MF values into a single crisp value, representing the final investment type recommendation.
- Integration: This layer aggregates the output MFs to generate the final output value.

4.2.5. Personality Traits ANFIS

The questionnaire used in this study included questions to measure the financial awareness and risk appetite of potential investors. This information is used to determine the appropriate investment products for the individual. Six questions were asked to assess financial awareness and risk appetite, and six inputs were designed in the ANFIS system based on these questions.

The first input, "safety," consisted of 5 MFs and was related to the potential investors' feelings of safety regarding different types of assets. These included securities (MF1), real estate (MF2), cash (MF3), gold (MF4), and bank deposits (MF5).

The second input, "excess money," also consisted of 5 MFs and focused on how potential investors use money beyond their usual monthly expenses. The five options for this question included setting the money aside for plans, spending it on entertainment, saving it for unexpected events, spending it on hobbies, or having no excess money.

The third input, "computational awareness," consisted of 3 MFs and measured the potential investors' understanding of mathematical concepts. Respondents were given the choice between three options: receiving two million forints with the decision left to the individual, receiving one million forints unconditionally, or receiving ten million forints if they could correctly predict a randomly chosen number.

The fourth input, "investment fund," consisted of 2 MFs and assessed whether the potential investors had an investment fund or not.

The fifth input, "saving factors," related to the potential investors' perception of factors that influence long-term savings, with four options provided and multiple answers allowed.

The responses to a question about investment options were grouped into different categories (as outlined in Table 4-3). These options included considerations such as low cost, low risk, potential for high returns, and government support. A total of 1539 potential investors selected one or more of these options in their answer. Based on this input, the system considered 10 MFs for analysis.

Table 4-3. MFs groups for Input 5 of Personality Trait ANFIS based on the responses

MFs	Grouping Multiple Choices (Input 5)	Code	Frequency/ responses
MF1	State Aid-Opportunity for High Returns-Low Risk-Low Cost	SOLrLc	359
MF2	State Aid-Opportunity for High Returns-Low Cost	SOLc	296
MF3	Opportunity for High Returns-Low Risk	OLr	280
MF4	Opportunity for High Returns-Low Risk-Low Cost	OLrLc	140
MF5	Low Risk	Lr	138
MF6	Low Risk-Low Cost	LrLc	131
MF7	State Aid-Opportunity for High Returns	SO	77
MF8	State Aid-Low Risk	SLr	53
MF9	State Aid-Low Risk-Low Cost	SLrLc	43
MF10	Opportunity for High Returns-Low Risk-State Aid	OLrS	22
Total of responses			**1539**

The final input for the ANFIS-based investment recommendation system was accounting knowledge. This input was like the third input and sought to gauge the potential investor's understanding of accounting principles. Participants were presented with a scenario involving the purchase of a one-year government bond for 100,000 forints, with an annual interest rate of 3% and an initial account management fee of 1% of the annual opening balance. They were then asked to calculate the amount of money that would be in their account after one year.

Table 4-4. Responses to Q 211 (Input 6)

Response to Q. 211	Frequency	Percentage
A	8	0.52%
B (Correct)	1377	89.3%
C	2	0.13%
D	72	4.67%
E	16	1.04%
F	12	0.78%
G	4	0.26%
H	-	-
I	1	0.06%
Total answers	**1492**	**96.76%**

There were ten possible answers to this question, with the second option (B) being the correct answer. The accounting knowledge of the potential investors was measured based on their response to this question. As shown in Table 4-4, more than 89% of respondents chose the correct answer. The system considered two MFs for this output, with MF1 being the correct answer and MF2 being the incorrect answer.

The proposed IRS framework utilizes a combination of six inputs, including safety, excess money, computational awareness, investment fund, saving factors, and accounting knowledge, to generate one output, which is divided into three clusters of investment types/products. These investment types include listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The system is trained using 1542 data pairs, with a minimum implication and maximum aggregation. The data is imported into MATLAB, which includes 7 columns, with 6 columns representing the input personality traits and one column representing the output investment type clusters. The fuzzy function of the system utilizes the ANFIS technique, with a new FIS designed in the Sugeno type, referred to as the PersonalityTraitsANFIS. Figure 4-40 illustrates a portion of the imported data used as inputs and output in the proposed PersonalityTraitsANFIS.

Variables - PersonalityTraitsANFIS

PersonalityTraitsANFIS

1542x7 double

	1	2	3	4	5	6	7
1	1	1	1	2	3	1	1
2	2	1	1	2	7	1	2
3	5	1	1	1	3	1	2
4	5	2	2	2	2	1	2
5	5	1	1	2	1	2	1
6	2	2	2	2	5	1	1
7	2	1	2	1	5	1	2
8	3	2	2	1	2	1	2
9	2	1	2	2	3	1	3
10	3	3	2	2	2	1	3
11	1	1	2	1	3	1	2
12	2	3	2	1	3	1	3
13	1	1	2	2	1	1	1
14	2	1	2	2	5	1	0
15	1	3	2	1	4	1	3

Figure 4-40. A part of imported data to MATLAB to propose the PersonalityTraitsANFIS

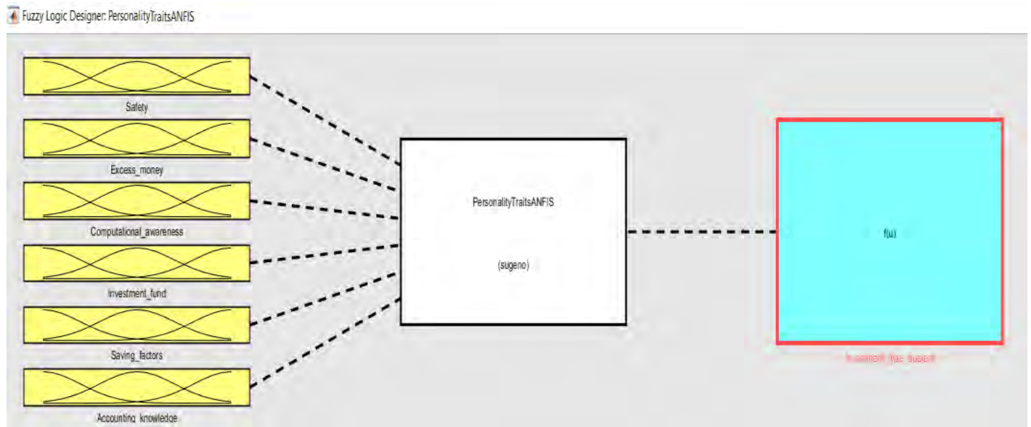


Figure 4-41. The properties of the PersonalityTraitsANFIS

Figure 4-41 illustrates the design of the "PersonalityTraitsANFIS" system, which incorporates six inputs: 1) safety, 2) excess money, 3) computational awareness, 4) investment fund, 5) saving factors, and 6) accounting knowledge. The system also includes one output, which is an investment type cluster.

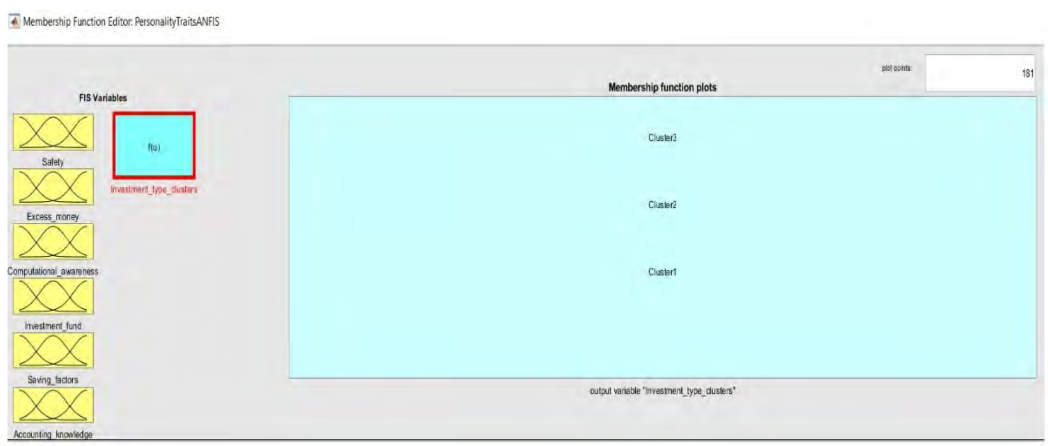


Figure 4-42. Output MFs in the PersonalityTraitsANFIS

As shown in Figure 4-42, the shape of the MFs for the output in the Personality Traits ANFIS model is presented. Constant kind MFs have been utilized for the three investment types, namely "Cluster 1", "Cluster 2", and "Cluster 3".

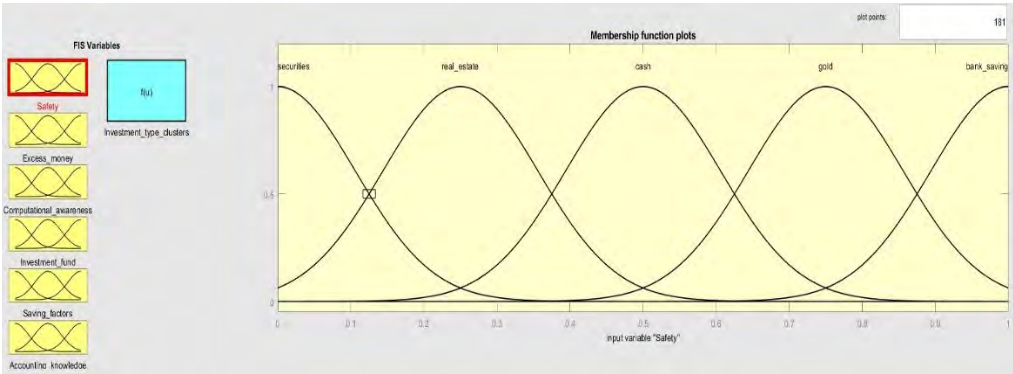


Figure 4-43. MFs shape for input 1 (safety) in the PersonalityTraitsANFIS

Figure 4-43 illustrates the shape of the MFs for input 1 in the Personality Traits ANIS. The MF shape employed is gaussian with five distinct categories: 'securities', 'real estate', 'cash', 'gold', and 'bank savings'.

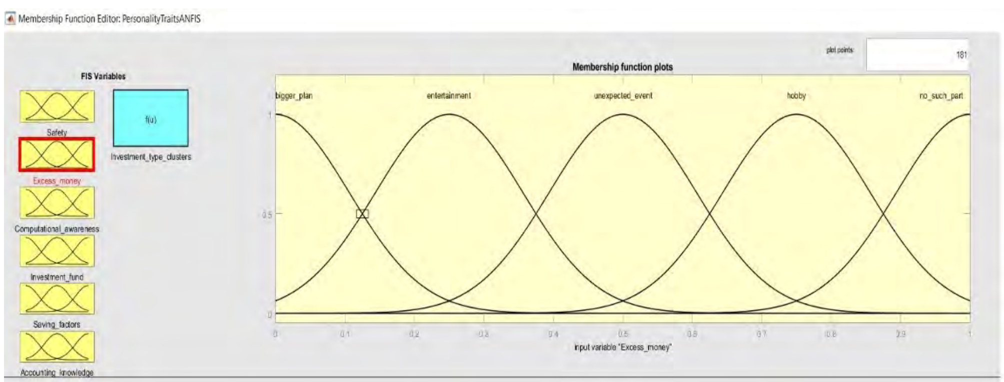


Figure 4-44. MFs shape for input 2 (excess money) in the PersonalityTraitsANFIS

Figure 4-44 illustrates the shape of the MFs for input 2 of the PersonalityTraitsANFIS system. The MFs are represented in the form of a gaussian function with five different categories, namely "bigger plan", "entertainment", "unexpected event", "hobby", and "no such part".

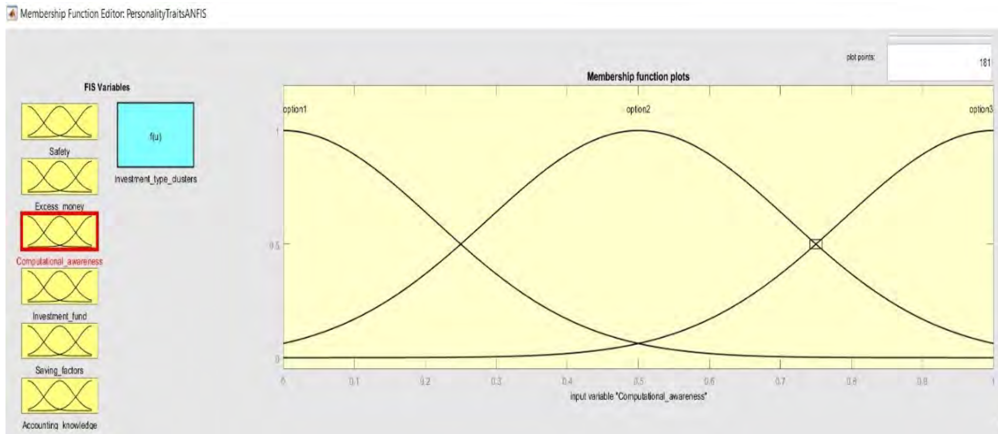


Figure 4-45. MFs shape for input 3 (computational awareness) in the PersonalityTraitsANFIS

Figure 4-45 illustrates the shape of the MFs for the third input in the Personality Traits ANFIS model. The MF shape utilized is gaussian, with three options represented: 'Option 1', 'Option 2', and 'Option 3'.

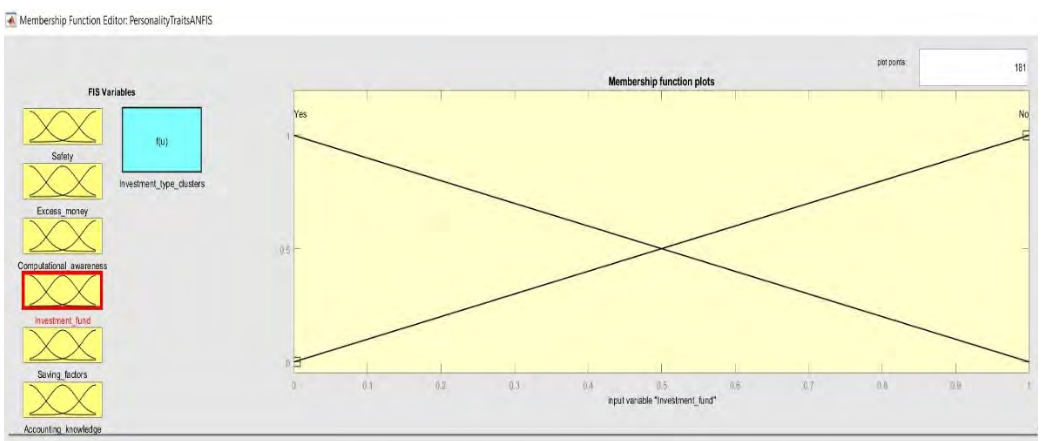


Figure 4-46. MFs shape for input 4 (investment fund) in the PersonalityTraitsANFIS

Figure 4-46 illustrates the shape of the MFs for input 4 in the PersonalityTraitsANFIS model. The MF shape is triangular, with two MFs designated as "yes" and "no".

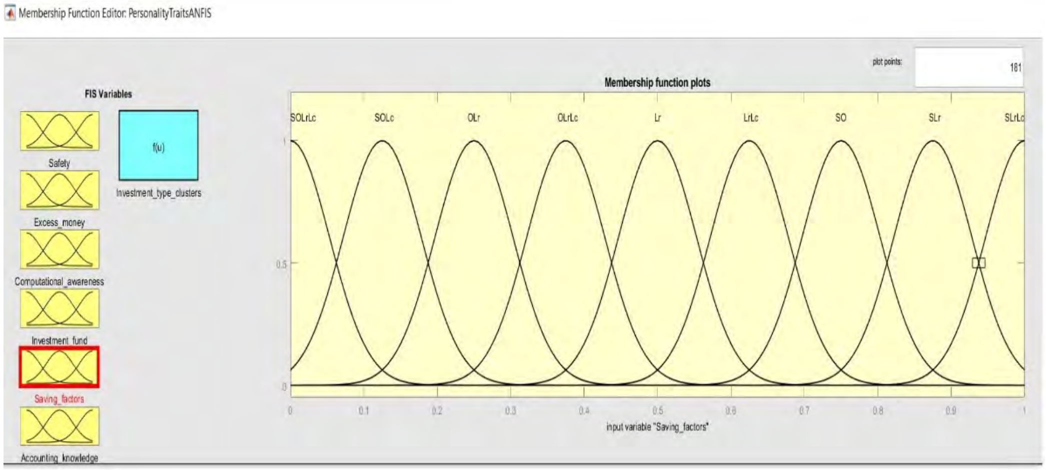


Figure 4-47. MFs shape for input 5 (saving factors) in the PersonalityTraitsANFIS

Figure 4-47 illustrates the MF shapes for input 5 of the PersonalityTraitsANFIS. The MFs are gaussian in nature, with 9 abbreviated codes as listed in Table 4-3 (SOLrLc, SOLc, OLr, OLrLc, Lr, LrLc, SO, SLr, and SLrLc). The MATLAB fuzzy toolbox allows for a maximum of 9 MFs per input. During the analysis, there were ten groups of data for this input, however, the tenth group was omitted as it consisted of only 22 potential investors out of a total of 1542 and ranked last. This exclusion does not impact the research results.

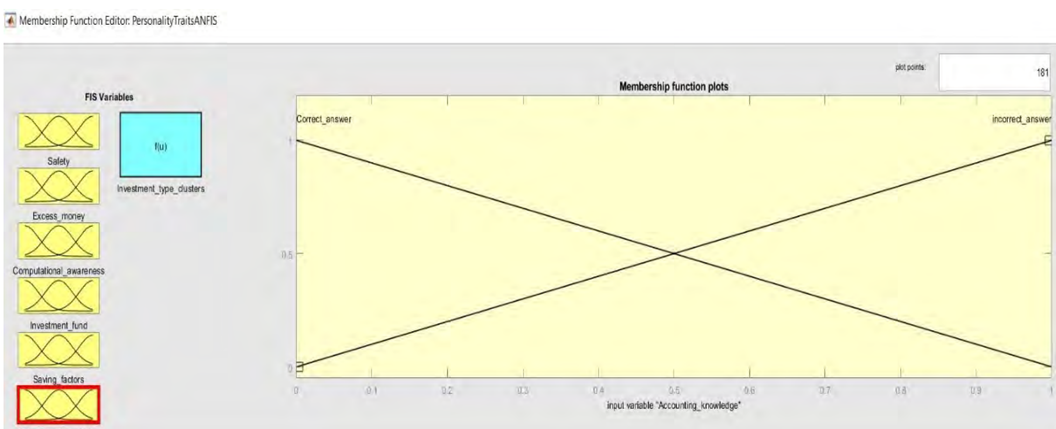


Figure 4-48. MFs shape for input 6 (accounting knowledge) in the PersonalityTraitsANFIS

Figure 4-48 displays the MF shape for input 6 in the PersonalityTraitsANFIS system. The MF shape is represented using a trimf with two distinct MFs: 'Correct Answer' and 'Incorrect Answer.'"

4.2.6. Proposing PersonalityTraitsANFIS

The proposed IRS considers six personality traits as inputs, including safety, excess money, computational awareness, investment fund, saving factors, and accounting knowledge. These inputs are used to determine the appropriate investment type/product for the individual. The output of the system consists of three clusters, each representing a different investment type or product. "Safety" refers to the individual's preference for secure investments with low risks. "Excess money" refers to the individual's disposable income or surplus funds available for investment. "Computational awareness" refers to the individual's understanding and familiarity with computers and technology. "Investment fund" refers to the individual's investment portfolio or fund size. "Saving factors" refers to the individual's savings habits and behavior. "Accounting knowledge" refers to the individual's understanding of financial accounting principles and practices. The combined ANFIS framework uses these six inputs to generate investment recommendations, grouping the recommendations into three clusters based on the characteristics of the investment type or product. The recommended investment type/product will vary depending on the individual's personality traits and investment goals.

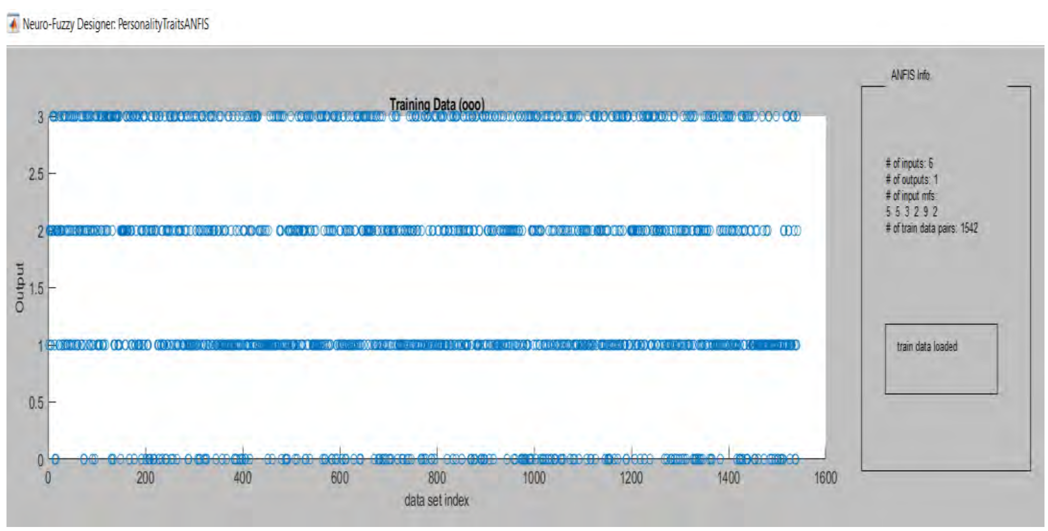


Figure 4-49. Prepared data in the PersonalityTraitsANFIS

Figure 4-49 presents the data that was prepared for the training and validation steps of the PersonalityTraitsANFIS. The grid partition method was utilized to train the new FIS and a hybrid optimization approach was used, with an error tolerance of 0 and 3 epochs. The result of this process was the generation of the PersonalityTraitsANFIS as a new FIS. Figure 4-50 provides a summary of the MFs for the PersonalityTraitsANFIS system. The system comprises 6 inputs and 1 output and is named PersonalityTraitsANFIS. The data set index is shown on the x-axis, while the y-axis displays the distribution of the output based on investment-type clusters. The number of data points in the dataset is 1542.

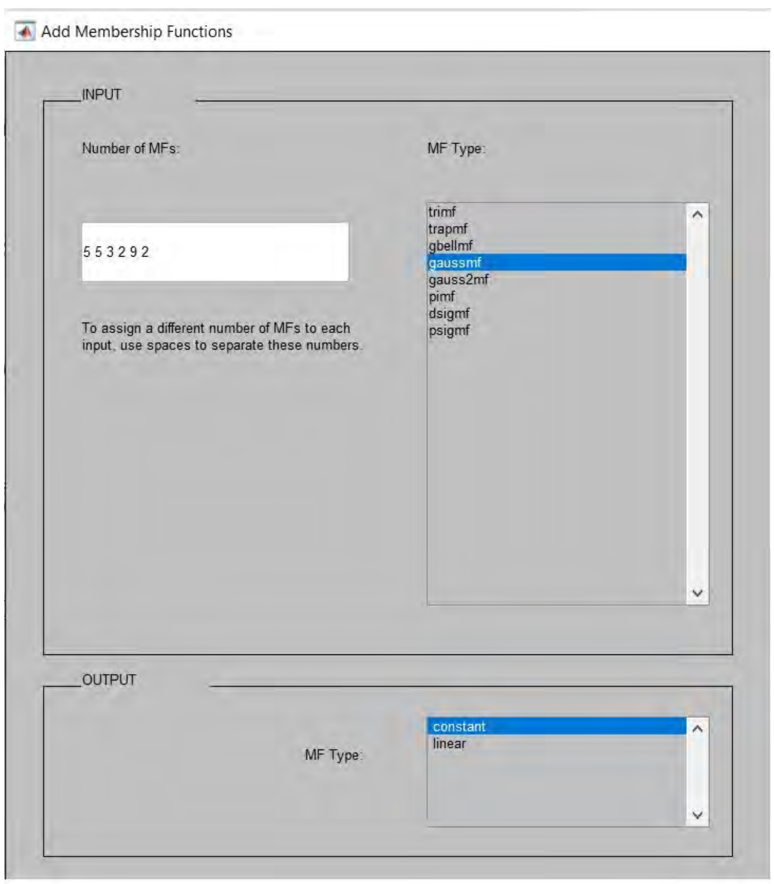


Figure 4-50. Information for generating the PersonalityTraitsANFIS as a new FIS

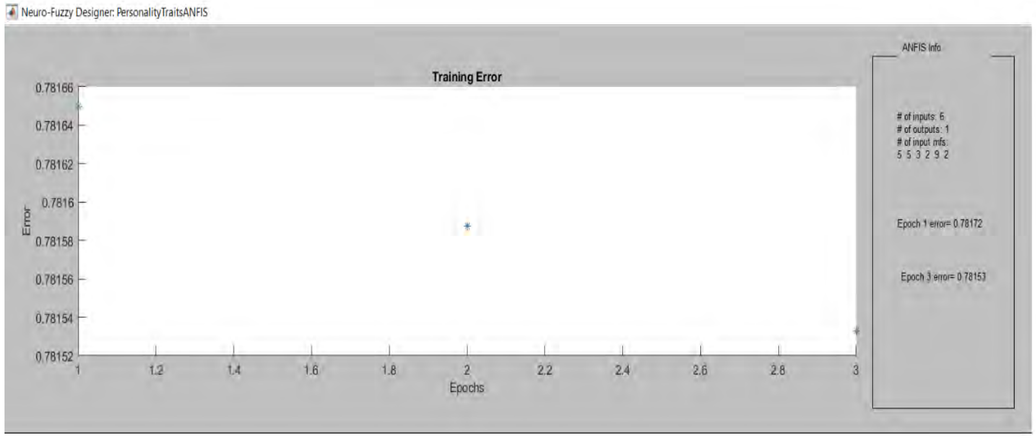


Figure 4-51. RMSE in the PersonalityTraitsANFIS

Figure 4-51 displays the trained PersonalityTraits ANFIS grid. It has six inputs and one output, which represents the investment type clusters. The training of the FIS was performed using a hybrid method with 3 epochs. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model. The error for each epoch is around 0.78. The information provided in the ANFIS info section indicates that the PersonalityTraits ANFIS has 5461 nodes, 2700 linear parameters, 52 nonlinear parameters, and a total of 2752 parameters. It was trained using 1542 training data pairs, and there were no checking data pairs used. The system has 2700 fuzzy rules.

The training process of the PersonalityTraits ANFIS is indicated by the "Start training ANFIS" message. The process took three epochs to complete, with the minimal training RMSE reaching 0.78153 at epoch 3. This value indicates the level of error in the model's prediction, with lower values indicating a more accurate model.

Figure 4-52 illustrates the trained Personality Traits ANFIS system. The average training error, as indicated in the figure, is 0.78148. This suggests that the system has a relatively high accuracy in its predictions, although there is still some room for improvement.

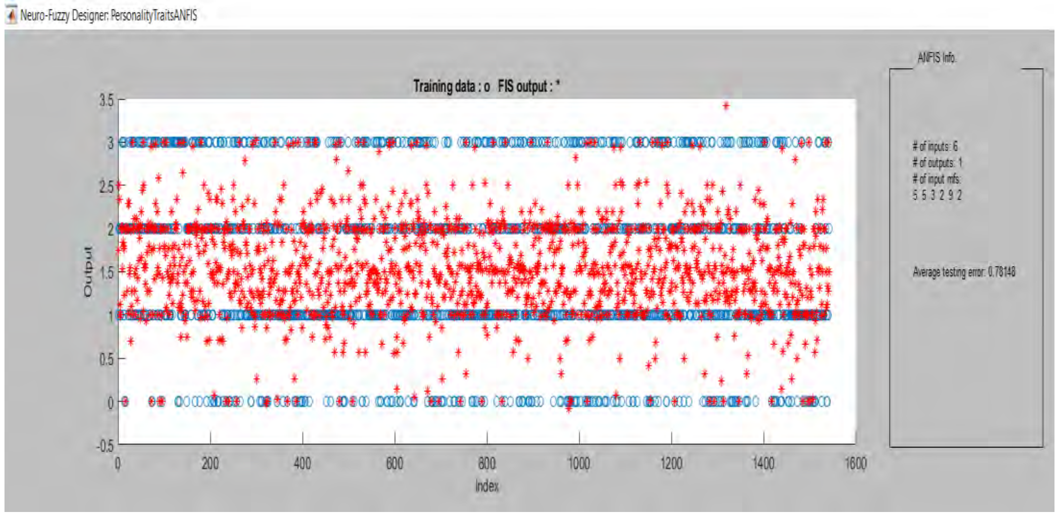
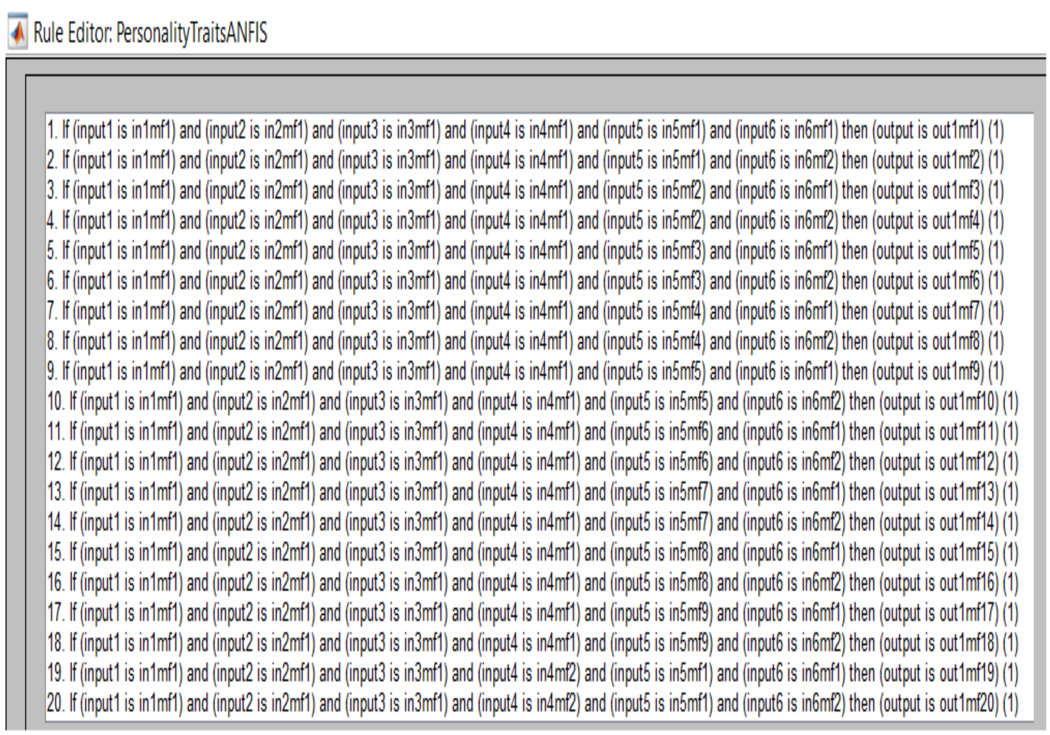


Figure 4-52. Trained data in the PersonalityTraitsANFIS

The Personality Traits ANFIS system has generated 2,700 rules, which are used to make investment recommendations based on the investor's personality traits. Figure 4-53 shows a portion of these generated rules in verbose format, allowing for a more detailed analysis of the system's workings.


Rule Editor: PersonalityTraitsANFIS

1. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf1) and (input6 is in6mf1) then (output is out1mf1) (1)
2. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf1) and (input6 is in6mf2) then (output is out1mf2) (1)
3. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf2) and (input6 is in6mf1) then (output is out1mf3) (1)
4. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf2) and (input6 is in6mf2) then (output is out1mf4) (1)
5. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf3) and (input6 is in6mf1) then (output is out1mf5) (1)
6. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf3) and (input6 is in6mf2) then (output is out1mf6) (1)
7. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf4) and (input6 is in6mf1) then (output is out1mf7) (1)
8. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf4) and (input6 is in6mf2) then (output is out1mf8) (1)
9. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf5) and (input6 is in6mf1) then (output is out1mf9) (1)
10. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf5) and (input6 is in6mf2) then (output is out1mf10) (1)
11. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf6) and (input6 is in6mf1) then (output is out1mf11) (1)
12. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf6) and (input6 is in6mf2) then (output is out1mf12) (1)
13. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf7) and (input6 is in6mf1) then (output is out1mf13) (1)
14. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf7) and (input6 is in6mf2) then (output is out1mf14) (1)
15. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf8) and (input6 is in6mf1) then (output is out1mf15) (1)
16. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf8) and (input6 is in6mf2) then (output is out1mf16) (1)
17. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf9) and (input6 is in6mf1) then (output is out1mf17) (1)
18. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf1) and (input5 is in5mf9) and (input6 is in6mf2) then (output is out1mf18) (1)
19. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf2) and (input5 is in5mf1) and (input6 is in6mf1) then (output is out1mf19) (1)
20. If (input1 is in1mf1) and (input2 is in2mf1) and (input3 is in3mf1) and (input4 is in4mf2) and (input5 is in5mf1) and (input6 is in6mf2) then (output is out1mf20) (1)

Figure 4-53. A part of the generated rules in the PersonalityTraitsANFIS

For example, the experts can create two rules based on the feedback of potential investors. These rules may not be created by the system.

Expert's Rule 1:
If (safety is in securities) and (excess cash is set aside for bigger plans) and (computational awareness is 2 million forints conditionally) and (investment fund is Yes) and (saving factors is State Aid-Opportunity for High Returns-Low Cost) and (for accounting knowledge, they answered the accounting question correctly then (output is Suitable investment products for this group include stocks/shares on the stock exchange, mutual funds, voluntary pension funds, and government securities. These respondents have invested in the stock market in the last three years and regularly monitor stock performance. Many of them have investments in government bonds.)

Expert's Rule 2:
If (safety is in securities) and (excess cash is set aside for bigger plans) and (computational awareness is 2 million forints conditionally) and (investment fund is Yes) and (saving factors is Opportunity for High Returns-Low Risk) and (for accounting knowledge, they answered the accounting question correctly then (output is Suitable investment products for this group include stocks/shares, voluntary pension funds, and government securities. These respondents have been investing in the stock market for the last three years and regularly monitor stock performance. Many of them did not invest in government bonds.)

Rule 1 by the system states that if the respondent's safety preference is in securities, they have excess cash set aside for bigger plans, their computational awareness is at least 2 million forints, they have invested in an investment fund, and their saving factors are State Aid-Opportunity for High Returns-Low Risk-Low Cost, and they answered the accounting question correctly, then the suitable investment products for this group include stocks/shares on the stock market, mutual funds, and government securities. However, many of them did not invest in government bonds, and they have not invested in the stock market in the last three years. Rule 1 by the expert is similar to rule 1 by the system, but it states that the respondents have invested in the stock market in the last three years and regularly monitor stock performance. Additionally, many of them have investments in government bonds. Therefore, the recommended investment products for this group include stocks/shares on the stock exchange, mutual funds, voluntary pension funds, and government securities. Rule 2 by the expert is also similar to system's rule 1 and expert's rule 2, but the saving factors are Opportunity for High Returns-Low Risk, and the respondents have been investing in the stock market for the last three years and regularly monitor stock performance. Additionally, many of them did not invest in government bonds. Therefore, the suitable investment products for this group include stocks/shares, voluntary pension funds, and government securities.

It is possible to make modifications to these rules based on expert opinions and investor feedback. This could involve adding, changing, or deleting rules to ensure that the system remains up-to-date and relevant and that its recommendations are in line with current market trends and the needs of the investors. Overall, the figures demonstrate the functionality and versatility of the Personality Traits ANFIS system in making investment recommendations.

Figure 4-54 illustrates a portion of the rule viewer, which displays the open system of the PersonalityTraits ANFIS. The figure showcases the presence of 2700 rules and 101 plot points. The PersonalityTraits ANFIS is a system that utilizes artificial intelligence and fuzzy logic to make investment recommendations based on individual personality traits. Attachment 3 provides further information about the rule generation process used by the PersonalityTraits ANFIS. This process involves using training data to generate rules that reflect the relationships between personality traits and investment type. The number of rules and plot points can be used as indicators of the complexity of the system and the amount of data used for training.

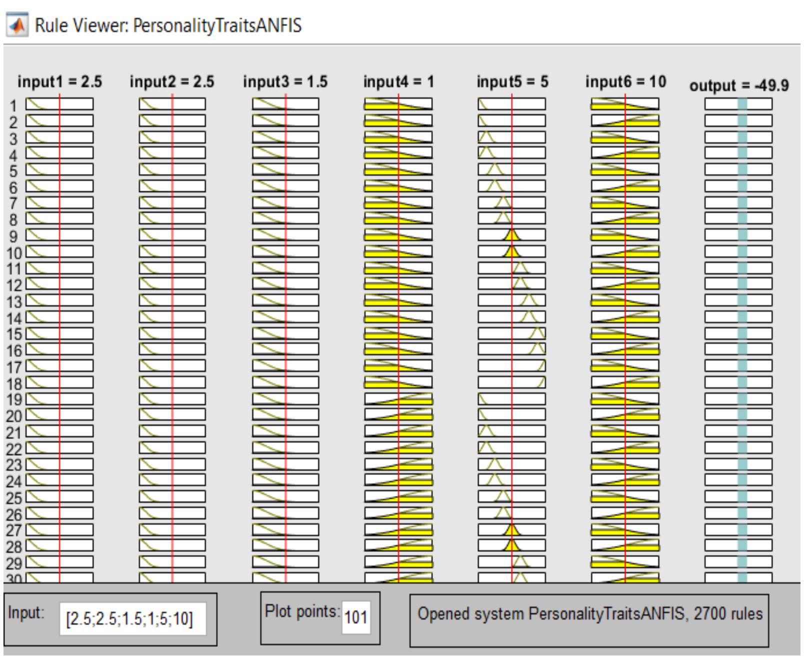
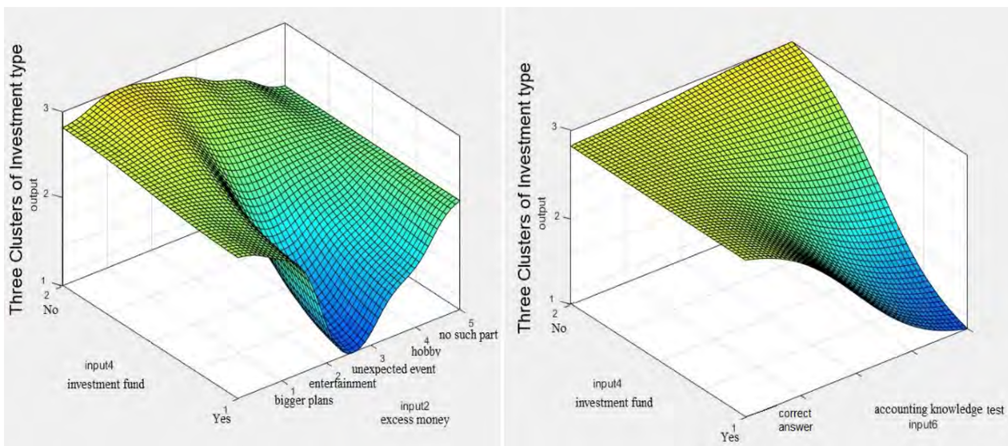


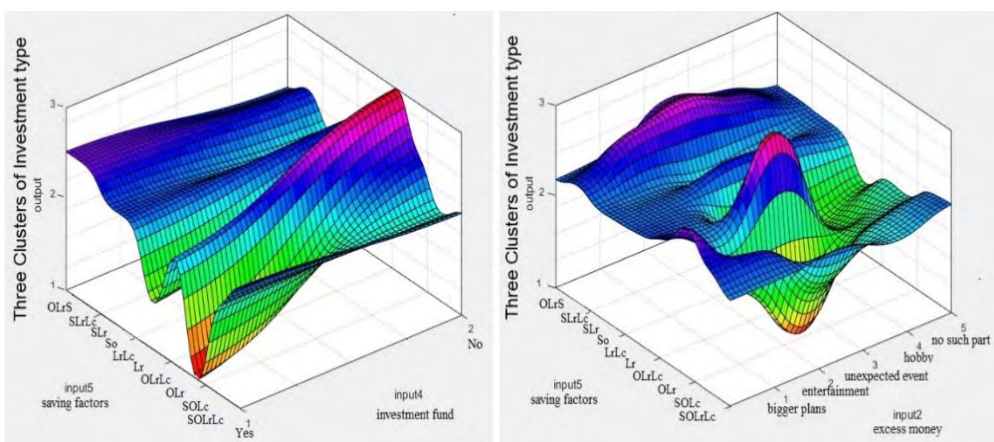
Figure 4-54. A part of the rule viewer in the PersonalityTraitsANFIS

The PersonalityTraitsANFIS is a proposed system that aims to determine the relationship between an individual's personality traits and their investment type preferences. Figures 4-55 (a-e) are 3D graphs that demonstrate the effect of different input pairs (i.e., personality traits) on the investment type. These surface graphs are nonlinear, meaning that they don't follow a straight line or simple pattern, and monolithic, meaning they are complete and integrated, presenting investment type recommendations for specific inputs. These graphs visually represent the complex relationship between an individual's personality traits and their investment preferences and provide insight into the system's decision-making process.



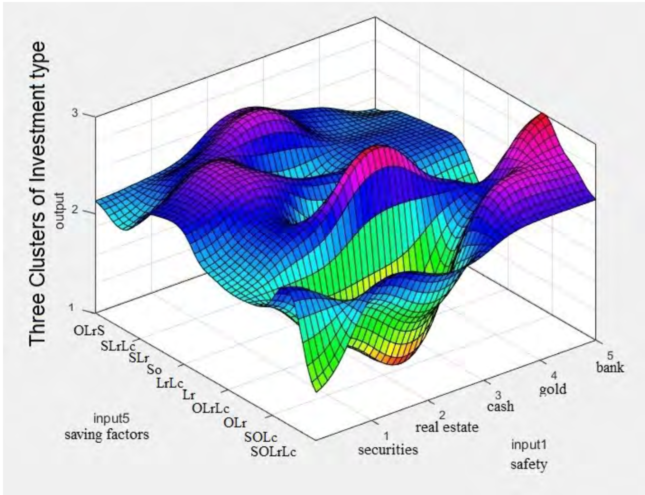
a. investment fund & excess cash

b. investment fund & accounting knowledge



c. saving factors & investment fund

d. saving factors & excess cash



e. saving factors & safety

Figure 4-55. Effectiveness of the relations of each pair of personality traits inputs on investment type

Figure 4-55 presents the effectiveness of the relationships between different pairs of personality trait inputs and investment type. The figure likely depicts the impact that each pair of personality traits has on the recommended investment type. For example: Figure 4-55a shows the investment fund & excess money - this pair of inputs could represent the relationship between the investment fund an individual has available and the amount of extra money they have. The figure might show how the presence of excess money affects the recommended investment type when an individual has a certain level of investment fund. Figure 4-55b shows the investment fund & accounting knowledge - this pair of inputs could represent the relationship between an individual's investment fund and their level of accounting knowledge. The figure might show how an individual's accounting knowledge affects the recommended investment type when they have a certain level of investment fund. Figure 4-55c shows the saving factors & investment fund - this pair of inputs could represent the relationship between an individual's saving habits and their investment fund. The figure might show how an individual's saving habits affect the recommended investment type when they have a certain level of investment fund. Figure 4-55d shows the saving factors & excess money - this pair of inputs could represent the relationship between an individual's saving habits and the amount of excess money they have. The figure might show how an individual's saving habits affect the recommended investment type when they have a certain amount of excess money. Figure 4-55e shows the saving factors & safety - this pair of inputs could represent the relationship between an individual's saving habits and their level of safety concerns. The figure might show how an individual's safety concerns affect the recommended investment type when they have certain saving habits. Figure 4-56 illustrates the Personality Traits ANFIS Model Structure, including the inputs, MFs, different layers of the ANFIS system, and the output recommendation to investors regarding investment type selection. The model includes the layers of fuzzification, implication rules, normalization, defuzzification, and integration, which result in the aggregated output MF. The MFs are likely categorized into different clusters, and the ANFIS system utilizes these clusters to make its investment type recommendations.

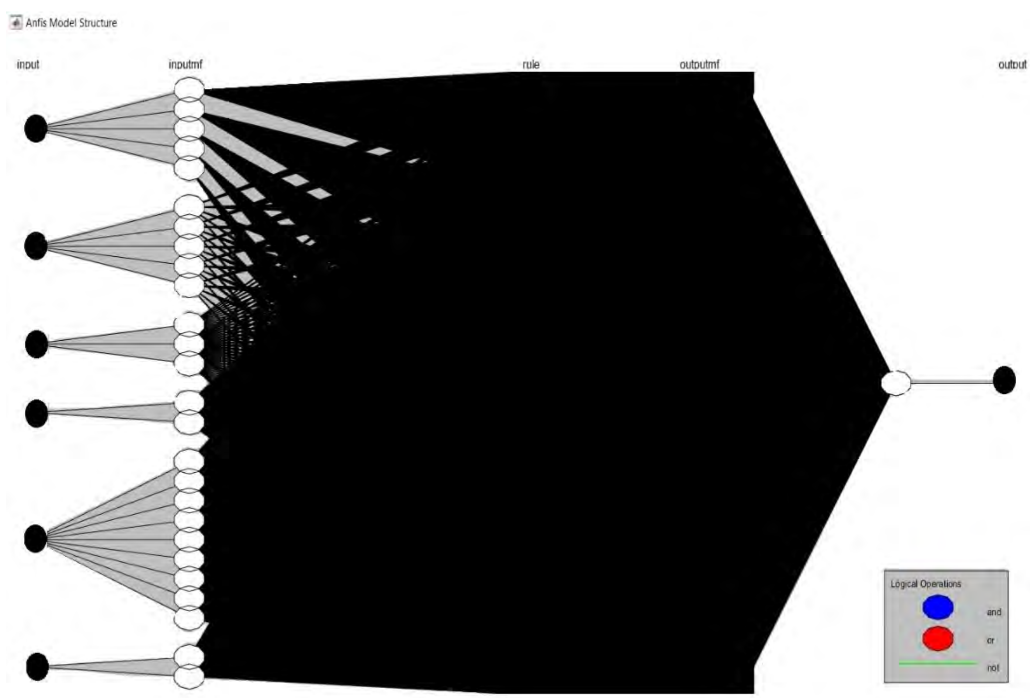


Figure 4-56. PersonalityTraitsANFIS Model Structure

4.2.4. Experiences ANFIS

The potential investors' experiences were measured through a questionnaire. Two questions (input 1-2) aimed to gather information about the respondents' experiences with digital solutions and their utilization in investment. Five sections of questions (input 3-7) evaluated the respondents' satisfaction with their current bank and openness to new savings trends and services.

Based on these questions, seven inputs were designed for the ANFIS-based IRS framework:

Input 1 (online shopping) with 2 MF evaluates the respondents' online shopping behavior in the past three months. "Yes" is assigned to MF1 and "No" to MF2.

Input 2 (online service) with 2 MFs gauges the respondents' use of online services like Spotify, Netflix, etc. "Yes" is assigned to MF1 and "No" to MF2.

Input 3 (bank accounts) with 3 MFs assesses the number of bank accounts the respondents use, with ranges of 4-1 assigned to MF1, 2-3 to MF2, and more than 3 to MF3.

Input 4 (bank status) with 3 MFs examines the respondents' bank account status, with MF1 assigned to premium customers, MF2 to private customers, and MF3 to regular customers.

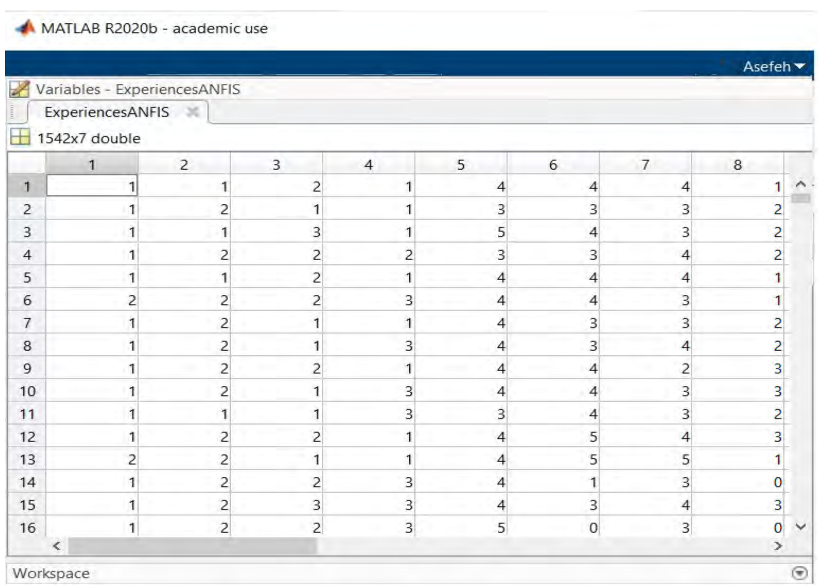
Input 5 (bank satisfaction) with 5 MFs evaluates the respondents' satisfaction with their bank based on answers to five questions on a five-point scale. The averages of the answers were rounded and assigned to MF1 (strongly disagree), MF2 (disagree), MF3 (neither agree nor disagree), MF4 (agree), and MF5 (strongly agree).

Input 6 (investment expectations) with 5 MFs assesses the respondents' investment expectations from their bank based on answers to four questions on a five-point scale. The averages of the answers were rounded and assigned to MF1 (strongly disagree), MF2 (disagree), MF3 (neither agree nor disagree), MF4 (agree), and MF5 (strongly agree).

Input 7 (security) with 5 MFs evaluates the respondents' feeling of safety and security with their bank based on answers to four questions on a five-point scale.

The averages of the answers were rounded and assigned to MF1 (strongly disagree), MF2 (disagree), MF3 (neither agree nor disagree), MF4 (agree), and MF5 (strongly agree).

The proposed framework employs a combined IRS using ANFIS that takes seven inputs into consideration, including online shopping, online service, bank accounts, bank status, bank satisfaction, investment expectations, and security. The system has a single output, which is the investment type/product, divided into three clusters. The investment types involve various financial products, such as listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The training data set consists of 1542 pairs of inputs and outputs, with a minimum implication and a maximum aggregation. The data was imported into MATLAB, consisting of 8 columns, with 7 columns related to the potential investors' experiences and 1 column related to the investment type clusters. In the fuzzy function, the inputs and outputs are designated for the ANFIS experiences. A new FIS was designed in the Sugeno type, referred to as ExperiencesANFIS. Figure 4-57 illustrates the imported data into MATLAB as inputs and outputs for the proposed ExperiencesANFIS.



	1	2	3	4	5	6	7	8
1	1	1	2	1	4	4	4	1
2	1	2	1	1	3	3	3	2
3	1	1	3	1	5	4	3	2
4	1	2	2	2	3	3	4	2
5	1	1	2	1	4	4	4	1
6	2	2	2	3	4	4	3	1
7	1	2	1	1	4	3	3	2
8	1	2	1	3	4	3	4	2
9	1	2	2	1	4	4	2	3
10	1	2	1	3	4	4	3	3
11	1	1	1	3	3	4	3	2
12	1	2	2	1	4	5	4	3
13	2	2	1	1	4	5	5	1
14	1	2	2	3	4	1	3	0
15	1	2	3	3	4	3	4	3
16	1	2	2	3	5	0	3	0

Figure 4-57. Imported data to MATLAB to propose the ExperiencesANFIS

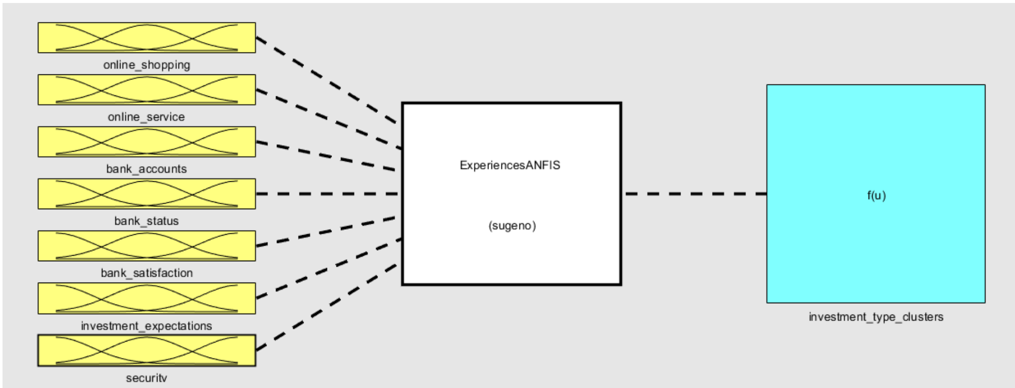


Figure 4-58. The properties of the ExperiencesANFIS

Figure 3-58 displays the proposed 'ExperiencesANFIS' system and its specifications. Seven inputs were integrated into the system, including: 1) Online shopping experience, 2) Online service experience, 3) Bank account usage, 4) Bank status, 5) Bank satisfaction, 6) Investment expectations, and 7) Security. The system generates a single output, which is the cluster type of investment.

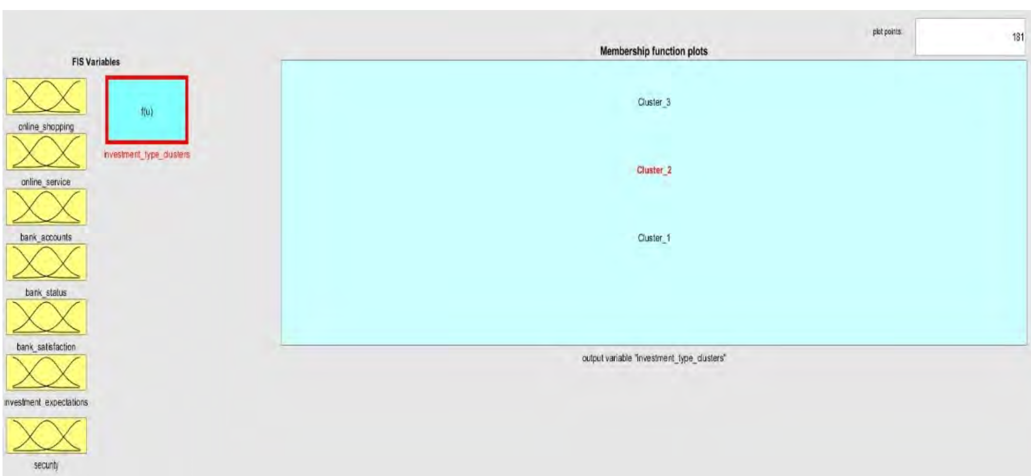


Figure 4-59. Output MFs in the ExperiencesANFIS

Figure 4-59 displays the MF shape for the output in the ExperiencesANFIS model. Three MFs of the constant type are utilized for the investment categories "Cluster 1," "Cluster 2," and "Cluster 3."

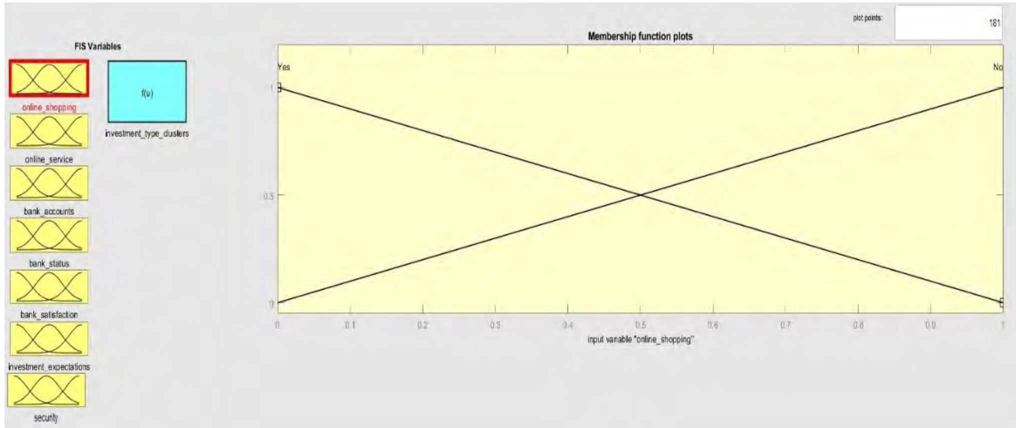


Figure 4-60. MFs shape for input 1 (online shopping) in the ExperiencesANFIS

Figure 4-60 displays the shape of the MF for input 1 of the ExperienceANFIS. The MF shape is triangular, with two MFs designated as 'Yes' and 'No'."

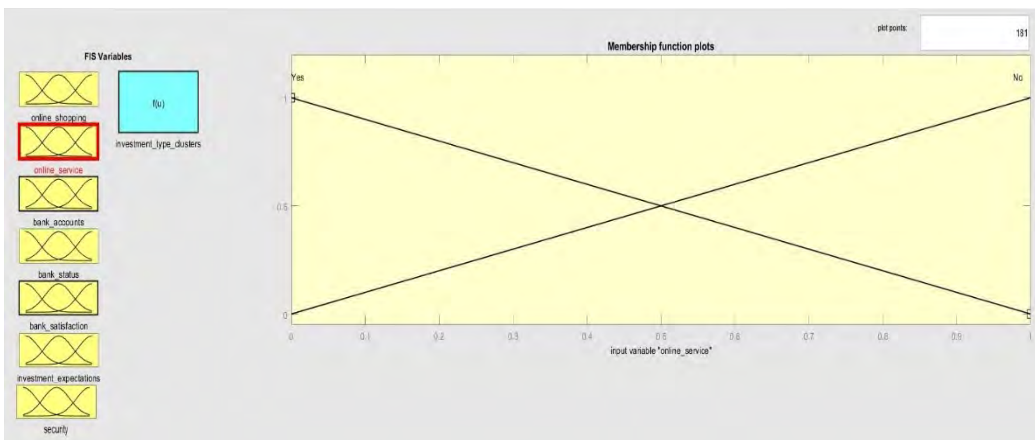


Figure 4-61. MFs shape for input 2 (online service) in the ExperiencesANFIS

Figure 4-61 depicts the shape of the MFs for input 2 in the ExperiencesANFIS model. The MFs are represented using triangular MFs with two categories, "Yes" and "No".

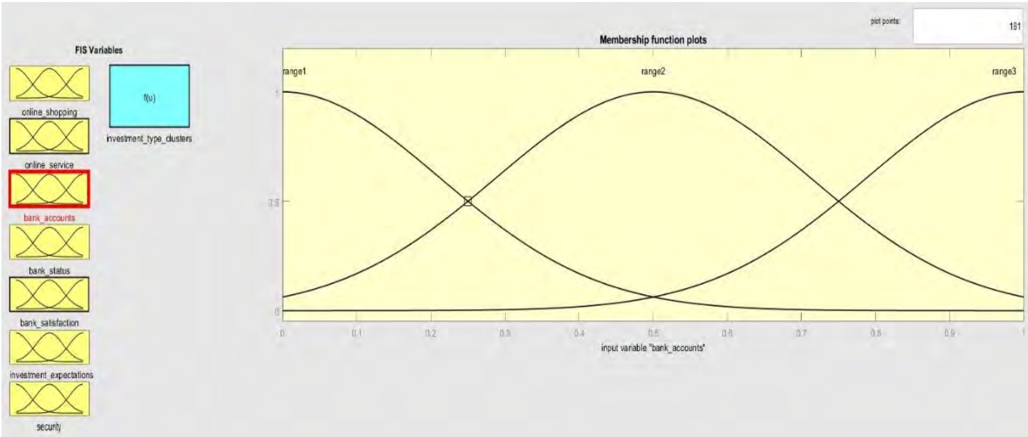


Figure 4-62. MFs shape for input 3 (bank accounts) in the ExperiencesANFIS

Figure 4-62 displays the MF shape for input 3 in the ExperiencesANFIS model. The MF shape is represented using gaussmf, with three ranges defined as "Range 1", "Range 2", and "Range 3".

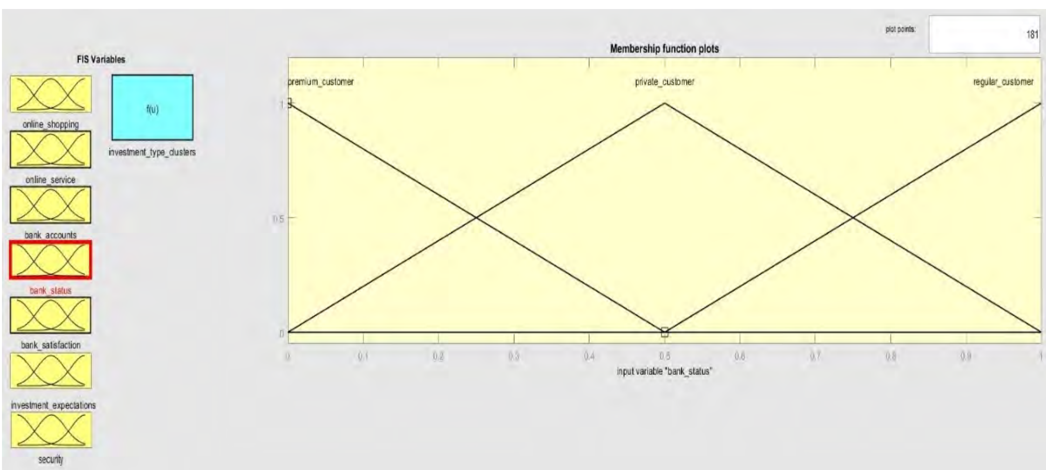


Figure 4-63. MFs shape for input 4 (bank status) in the ExperiencesANFIS

Figure 4-63 illustrates the MF shapes for input 4 of the ExperiencesANFIS. The MF shape is triangular, represented by three MFs, including "Premium Customer," "Private Customer," and "Regular Customer.

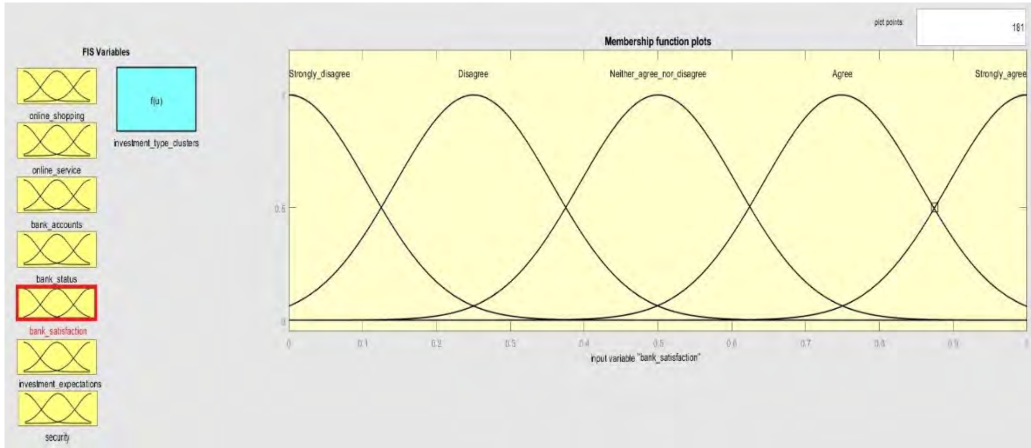


Figure 4-64. MFs shape for input 5 (bank satisfaction) in the ExperiencesANFIS

Figure 4-64 illustrates the MF shape for the input 5 of the ExperiencesANFIS. The MF shape is represented as gaussmf, with five MFs on a five-point scale ranging from "strongly disagree" to "strongly agree."

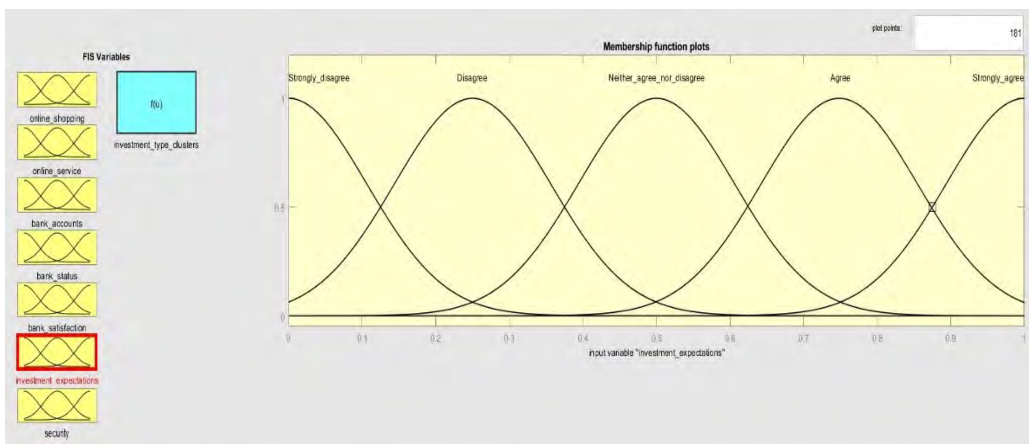


Figure 4-65. MFs shape for input 6 (investment expectations) in the ExperiencesANFIS

Figure 4-65 illustrates the shape of the MFs for input 6 of the ExperiencesANFIS. The MFs are represented by a gaussian distribution with five MFs on a five-point scale ranging from "strongly disagree" to "strongly agree."

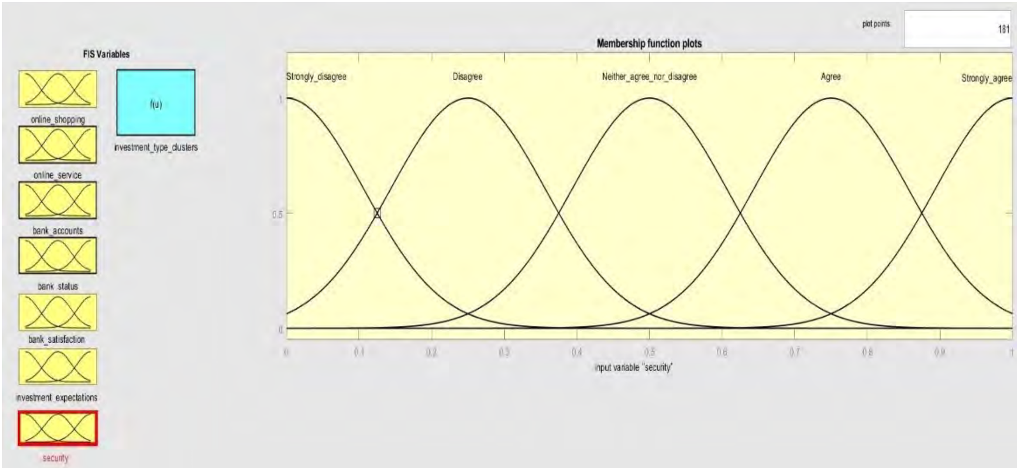


Figure 4-66. MFs shape for input 7 (security) in the ExperiencesANFIS

Figure 4-66 illustrates the MF shape for the input 7 in the ExperiencesANFIS system. The MF is represented by a gaussian shape, with 5 functions distributed along a five-point scale ranging from "strongly disagree" to "strongly agree”.

4.2.5. Proposing ExperiencesANFIS

The system described has seven inputs that relate to the individual's experience in online shopping, online service, bank accounts, bank status, bank satisfaction, investment expectations, and security. These inputs are used to determine the individual's investment type, which is the system's output. The investment type output is categorized into three clusters. The presence of seven inputs indicates that the system takes into consideration various aspects of the individual's experience when making investment recommendations. The three clusters for investment type suggest that the system classifies the investment recommendations into three distinct categories, potentially providing a more nuanced and detailed approach to investment decision-making.

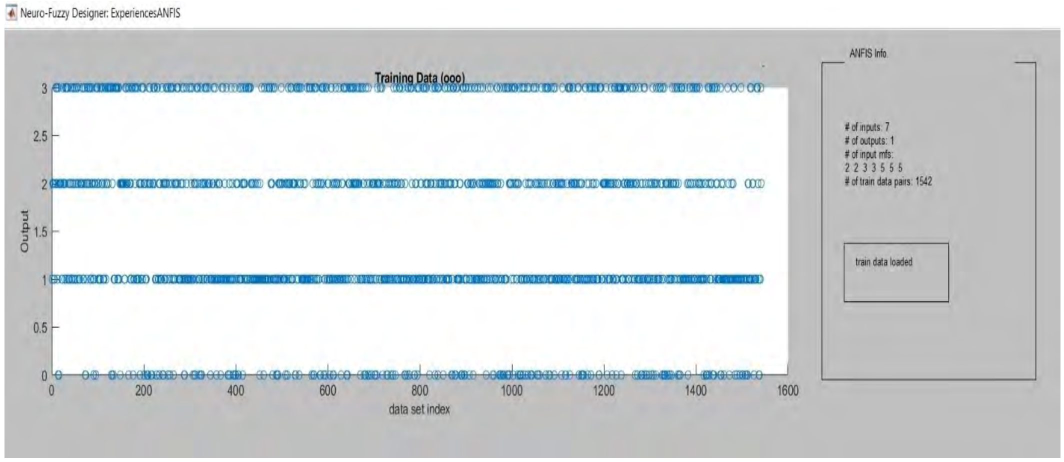


Figure 4-67. Prepared data in the ExperiencesANFIS

Figure 4-67 illustrates the processed data that will be utilized for the training and validation phases of the Experiences ANFIS. To train a new FIS, a grid partition method was employed, and the optimization was conducted using a hybrid approach with an error tolerance of 0 and 3 epochs. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model. This resulted in the creation of the Experiences ANFIS as a new FIS. Figure 4-68 presents a summary of the MFs for the Experiences ANFIS.

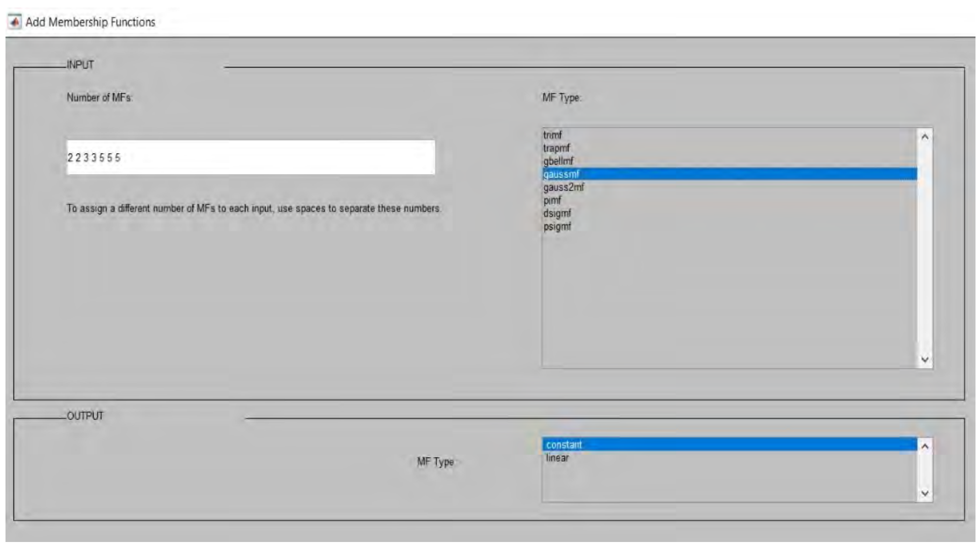


Figure 4-68. Information for generating the ExperiencesANFIS as a new FIS

The ExperiencesANFIS is a newly generated FIS that consists of 7 inputs and 1 output. The data set index is represented by 1542 on the x-axis, while the y-axis displays the distribution of the output based on investment-type clusters. The information about the training process of the ExperiencesANFIS is provided as follows:

ANFIS Information:

- The system has a total of 9060 nodes.
- There are 4500 linear parameters and 50 nonlinear parameters, making the total number of parameters 4550.
- The system was trained using 1542 training data pairs and 0 checking data pairs.
- The ExperiencesANFIS has 4500 fuzzy rules.

The training of the ExperiencesANFIS was performed using ANFIS, and the following details were noted:

- The training process was started by ANFIS and completed at epoch 2, as the designated epoch number was reached.
- The minimal training RMSE was recorded as 0.788496, which indicates the level of accuracy achieved by the system during the training process.

The proposed ExperiencesANFIS system was developed at a stage where the number of nodes was 9060 and the number of fuzzy rules was 4500. Due to the heavy data processing requirements of this system, a suitable server was not available. To address this, a sample of 500 rows of data was selected and used to propose the ExperiencesANFIS system.

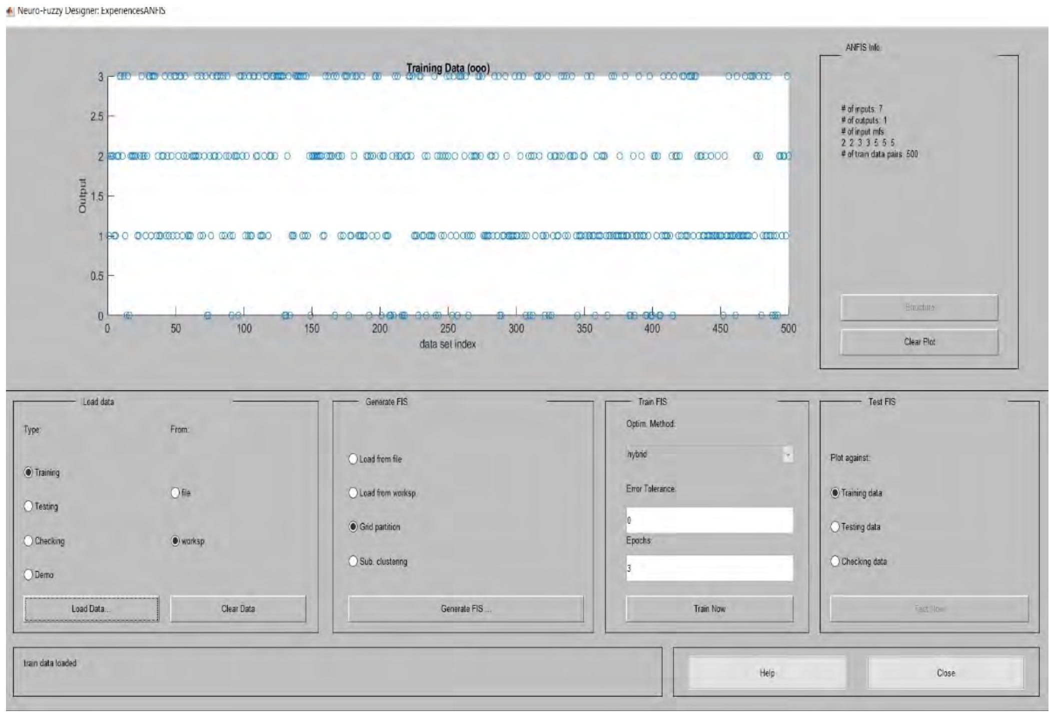


Figure 4-69. Prepared sample data in the ExperiencesANFIS

Figure 4-69 presents sample data, consisting of 500 rows, for the purpose of training and validation in the Experiences ANFIS system. To prepare the data for training, a grid partition method was used and the optimization process was conducted using a hybrid method with an error tolerance of 0 and 3 epochs. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model. This resulted in the creation of a new FIS, known as the Experiences ANFIS, which consists of 7 inputs and 1 output. The data set index, displayed on the x-axis, consists of 500 rows, while the y-axis represents the distribution of the output based on investment type clusters. In other words, the y-axis displays how the output is grouped based on different types of investments.

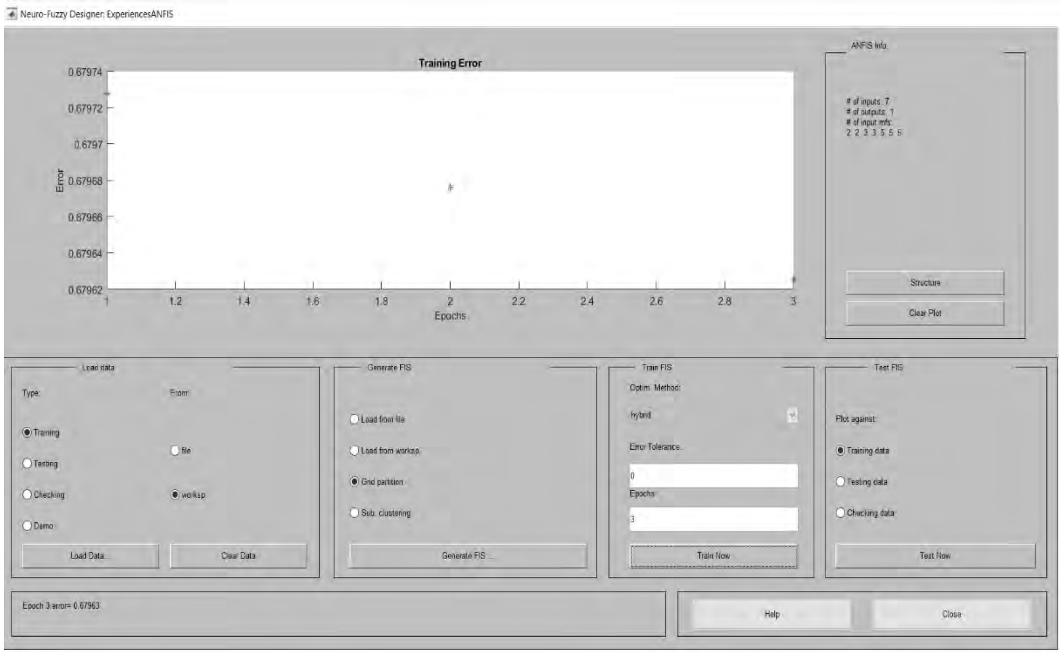


Figure 4-70. The trained ExperiencesANFIS with sample data

Figure 4-70 displays the trained grid for the Experiences ANFIS system. This system has seven inputs and one output, which categorizes investment types into clusters. The system is a hybrid type and has undergone 3 training epochs. The error for each of these epochs is around 0.67. The ANFIS information provided in the text shows that the system has 9060 nodes, 4500 linear parameters, 50 nonlinear parameters, and a total of 4550 parameters. The system has been trained using 500 training data pairs and 0 checking data pairs and has generated 4500 fuzzy rules. The training process of the Experiences ANFIS is described as follows: the system starts training, and after the designated epoch number (2) is reached, the ANFIS training process is completed. The minimal training RMSE is 0.679575.

Figure 4-71 presents the results of the trained Experiences ANFIS, based on the sample data used for training. The average training error is 0.67958, which represents the average difference between the predicted and actual values during training. This system has generated 4,500 rules based on the sample data.

Figure 4-71. Trained data in the ExperiencesANFIS based on sample

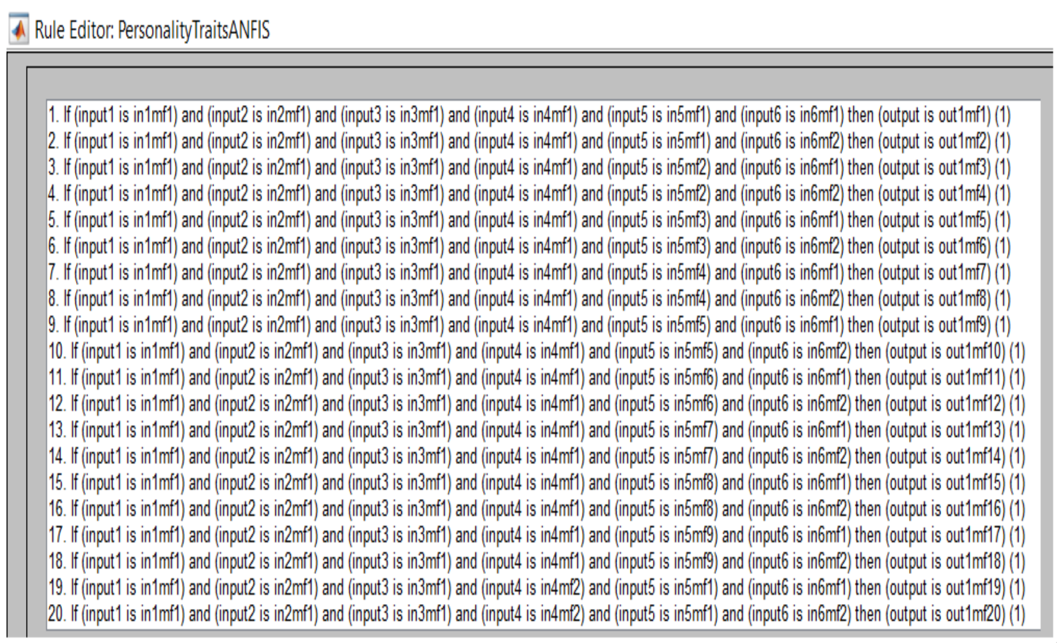
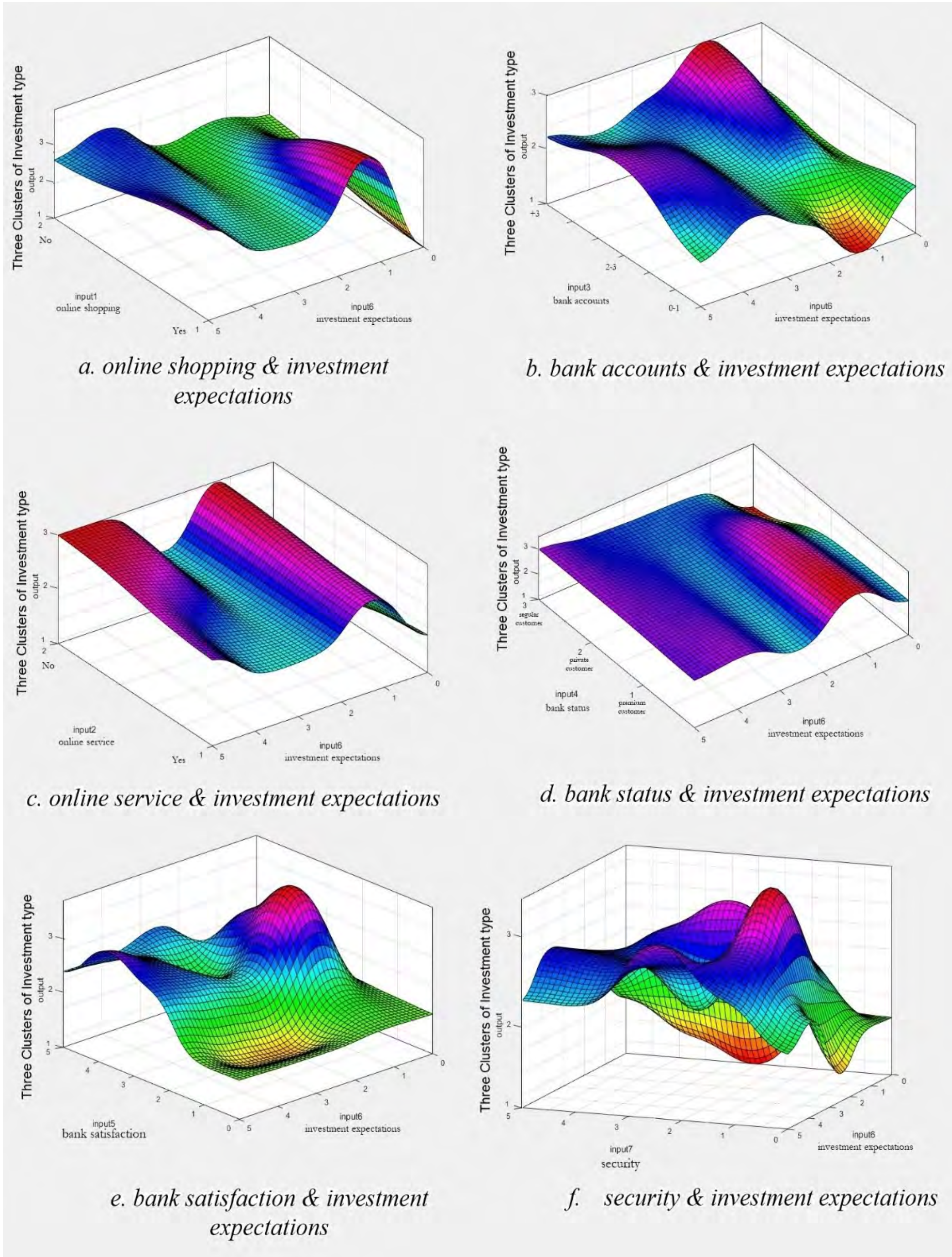


Figure 4-72. A part of the generated rules in the ExperiencesANFIS

Figure 4-72 highlights a portion of the generated rules in a detailed format, referred to as the "verbose format". The verbose format allows for easy inspection of the rules and their characteristics. These rules can be modified, added, or deleted based on expert opinions and feedback from investors. This provides a level of customization and adaptability to the Experiences ANFIS, ensuring that it remains relevant and effective in generating investment recommendations. The Experiences ANFIS system proposed is designed to investigate the relationship between experiences and the type of investment. Figures 4-73 (a-f) present a 3D graphical representation of the impact of certain input pairs on investment type. These graphs are nonlinear and monolithic, meaning they show investment type recommendations for given inputs as a single, uninterrupted surface. The effect of some specific input pairs on investment type is described through these graphs. The goal of the Experiences ANFIS is to help better understand how experiences influence investment decisions, and the information provided by the 3D graphs can assist in this understanding. These figures are meant to demonstrate the effectiveness of the relationship between different pairs of experience inputs and investment type. The figure is comprised of six sub-figures (4-73a through 4-73f) that show the relationship between various input pairs and investment expectations. Figure 4-73a and 4-73b depict the relationship between online shopping experience and investment expectations, as well as between bank account experience and investment expectations, respectively. Figure 4-73c and 4-73d depict the relationship between the combination of online shopping experience and bank account experience and investment expectations. Figure 4-73e and 4-73f illustrate the relationship between bank satisfaction and investment expectations, as well as between security and investment expectations, respectively. The aim of these sub-figures is to give an understanding of how different experience inputs affect investment expectations, and how they can be used to make informed investment decisions.



a. online shopping & investment expectations

b. bank accounts & investment expectations

c. online service & investment expectations

d. bank status & investment expectations

e. bank satisfaction & investment expectations

f. security & investment expectations

Figure 4-73. Effectiveness of the relations of each pair of Experiences inputs on investment type

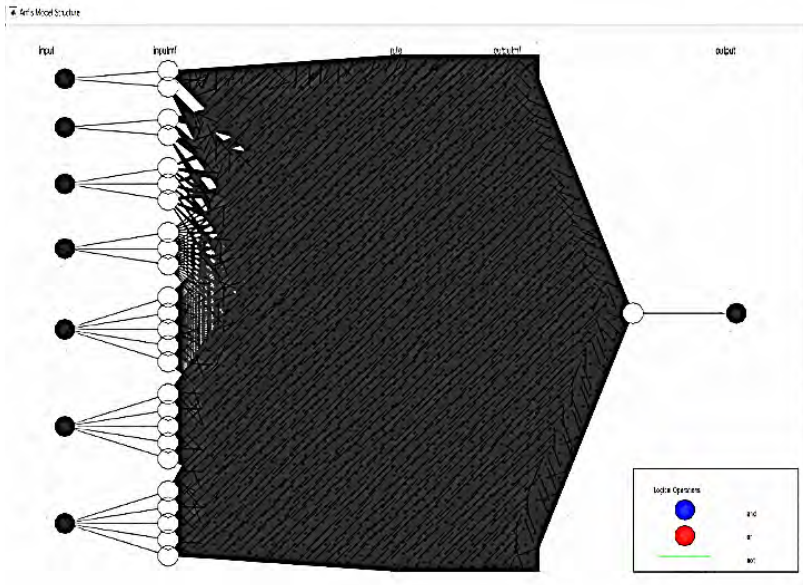


Figure 4-74. ExperiencesANFIS Model Structure

Figure 4-74 depicts the structure of the Experience ANFIS model, which serves as a recommendation system for investment type selection based on the inputs and cluster formations. The model illustrates the various components of the ANFIS system, including inputs, MFs, and multiple layers of the ANFIS. The fuzzification layer is responsible for converting the inputs into fuzzy sets, allowing for a more nuanced representation of the data. The implication rules layer then applies the relevant rules to the fuzzified inputs, generating intermediate results. The normalization layer ensures that the results are properly scaled and weighted, while the defuzzification layer converts the intermediate results back into crisp values. Finally, the integration or aggregated output MF layer integrates the results from each rule, producing the final output, which serves as a recommendation for investment type selection.

4.2.6. Financial ANFIS

The questionnaire contained questions that assessed the financial status of potential investors. Four questions (inputs 1-4) were asked about their current savings and financial situation. This information was gathered to understand the potential impact of an unexpected event or significant expense on the investors. The ANFIS model was designed with four inputs based on the questionnaire responses, as follows:

Input 1 (current savings) was evaluated with 6 MF based on the potential investors' savings. Option 1, "I could buy a new car", was assigned to MF1; option 2, "I could buy a 4-5-year-old car", was assigned to MF2; option 3, "I could buy a property from my savings", was assigned to MF3; option 4, "I could pay for a foreign vacation", was assigned to MF4; option 5, "I can buy a mid-range smartphone anytime", was assigned to MF5; and option 6, "There are no savings", was assigned to MF6. (See Table 4-5 for details).

Table 4-5. potential investors' statement to have savings

	Options	**Frequency**
MF1	I could buy a new car for myself	410
MF2	I could buy a 4-5-year-old car for myself	345
MF3	I could buy property from my savings	396
MF4	I could pay for a foreign vacation	276
MF5	I can buy a mid-range smartphone anytime	83
MF6	There are no	29

The input of monthly expenses with 6 MF relates to the financial status of potential investors, specifically their average monthly living expenses. The first option, "We regularly set aside savings" is assigned to MF1, the second option "We allocate additional funds for entertainment and shopping, as well as holiday expenses" is assigned to MF2, the third option "We occasionally spend on fun purchases such as new clothing" is assigned to MF3, the fourth option "We allocate a regular budget for entertainment and shopping" is assigned to MF4, the fifth option "Our monthly income only covers necessary expenses such as housing and food" is assigned to MF5, and the sixth option "We do not have any expendable income" is assigned to MF6 (refer to Table 4-6).

Table 4-6. Average monthly living expenses of potential investors

	Options	Frequency
MF1	We can regularly set aside savings	1086
MF2	In addition to our regular spending (entertainment, shopping), we can set you aside for the holidays	244
MF3	We just go out a little every month to buy new clothes for fun	90
MF4	We can spend it regularly on entertainment and shopping	104
MF5	Our monthly income provides our basic livelihood (housing, food)	13
MF6	We do not come out of our monthly revenue	4

The current financial situation of potential investors is assessed by inputting 5 different mutual funds (MF) based on the investor's description of their financial statement. Option 1, "I have no daily problems but in the long run," is assigned to MF1; option 2, "Everything's okay," is assigned to MF2; option 3, "I am confident in my financial stability, but concerned about the future for my children," is assigned to MF3; option 4, "I am struggling but managing to get by," is assigned to MF4; and option 5, "hopeless," is assigned to MF5. This information is reflected in Table 4-7.

Table 4-7. Description of the current financial statement by potential investors

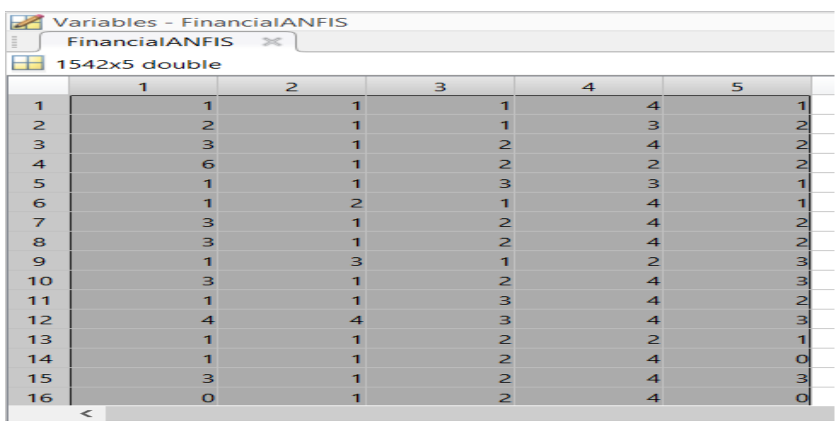
	Options	Frequency
MF1	I have no daily problems but no overall	471
MF2	Everything is okay	573
MF3	I am calm about myself, but the future of the children is uncertain	470
MF4	It is hard, but I live	23
MF5	hopeless	5

The spending plan for savings, input 4, is designed to assess the expected usage of savings by potential investors. The options provided for this input have been assigned to four different MFs, as follows: Option 1, "I plan to use my savings in one or two weeks or one month," is assigned to MF1; Option 2, "Within one year or 2-3 years," is assigned to MF2; Option 3, "Within 4-5 years or 5-8 years," is assigned to MF3; and Option 4, "Over 8 years or I do not plan to use my savings," is assigned to MF4, as indicated in Table 4-8.

Table 4-8. Expectation to spend savings by the potential investors

	Options	Frequency
MF1	I plan to use my savings in one or two weeks or one month	38
MF2	Within one year or 2-3 years	240
MF3	Within 4-5 years or 5-8 years	143
MF4	Over 8 years or I do not plan to use my savings	1119

The output of the investment type clusters is defined as three clusters. The data for the investment type involves various financial products, including listed stocks, mutual funds, voluntary pension funds, government securities and bonds, among others. To train the system, a total of 1542 data pairs were utilized for both the inputs and outputs. The aggregation method used was Max, while the implication method applied was Min.


Variables - FinancialANFIS

FinancialANFIS

1542x5 double

	1	2	3	4	5
1	1	1	1	4	1
2	2	1	1	3	2
3	3	1	2	4	2
4	6	1	2	2	2
5	1	1	3	3	1
6	1	2	1	4	1
7	3	1	2	4	2
8	3	1	2	4	2
9	1	3	1	2	3
10	3	1	2	4	3
11	1	1	3	4	2
12	4	4	3	4	3
13	1	1	2	2	1
14	1	1	2	4	0
15	3	1	2	4	3
16	0	1	2	4	0

Figure 4-75. A part of imported data to MATLAB to propose the FinancialANFIS

The data imported into MATLAB consisted of five columns, with four columns related to the financial information of potential investors and one column indicating the investment type clusters. In the fuzzy function, the inputs and outputs were designated for the financial ANFIS. A new FIS, FinancialANFIS, was designed using the Sugeno type. Figure 4-75 illustrates the imported data in MATLAB, which served as inputs and outputs to develop the FinancialANFIS.

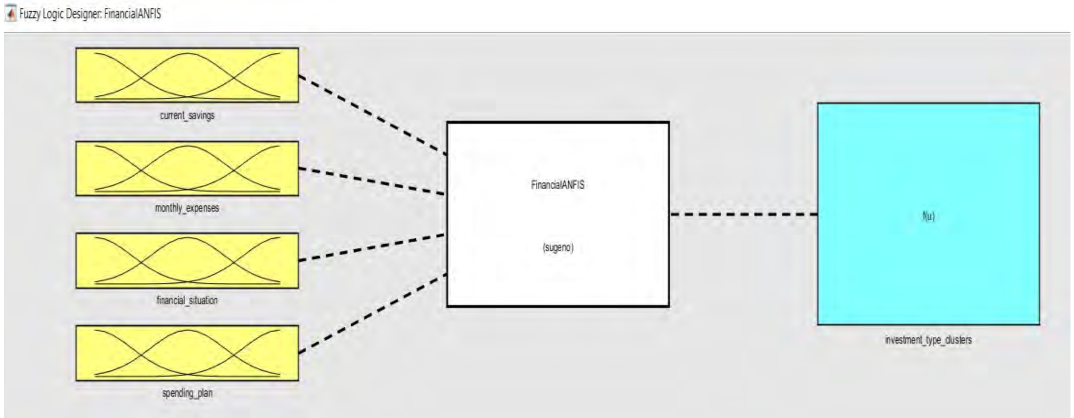


Figure 4-76. The properties of the FinancialANFIS

Figure 4-76 illustrates the designed Financial ANFIS system, referred to as "FinancialANFIS," and its properties. The system has four inputs, namely, current savings, monthly expenses, financial situation, and spending plan for savings, and one output, the investment type cluster.

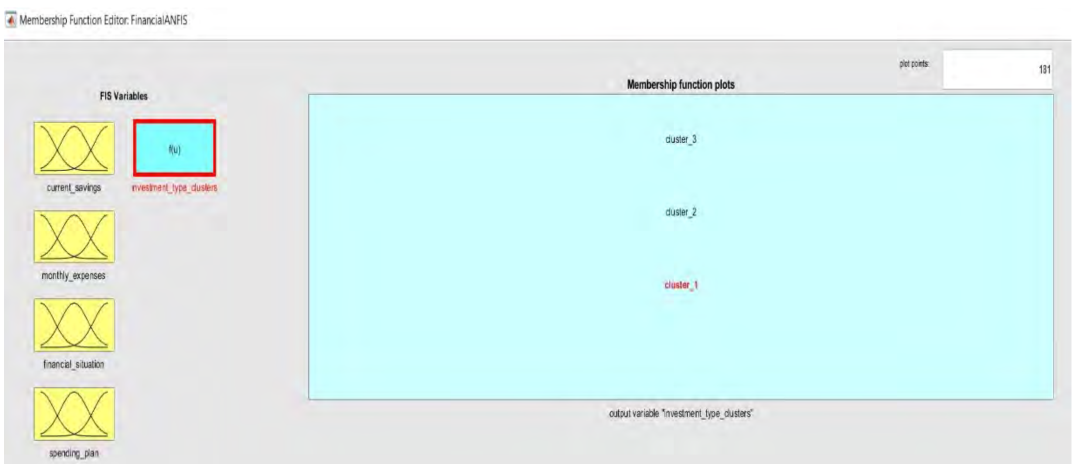


Figure 4-77. Output MFs in the FinancialANFIS

Figure 4-77 illustrates the shape of the MFs for the output in FinancialANFIS. The output is divided into three constant-type MFs, corresponding to the investment types 'Cluster 1', 'Cluster 2', and 'Cluster 3.

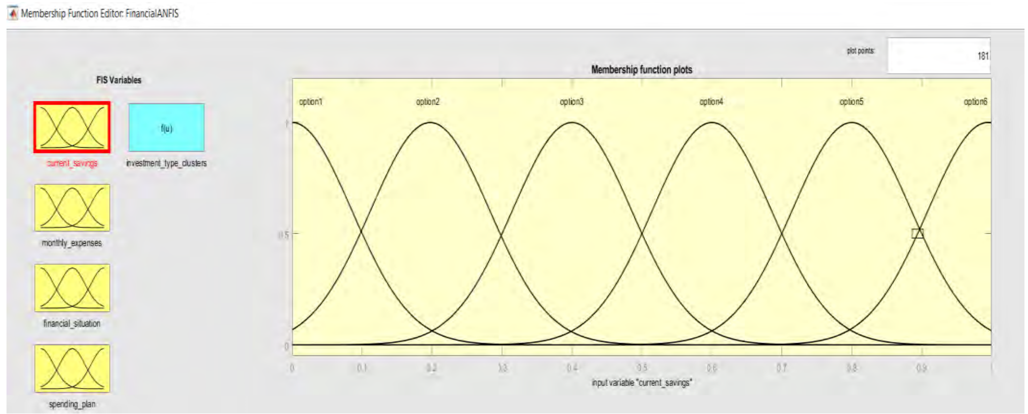


Figure 4-78. MFs shape for input 1 (current saving) in the FinancialANFIS

Figure 4-78 illustrates the shape of the MFs for input 1 in the FinancialANFIS. The shape of the MFs is gaussian and is composed of 6 options: "Option 1", "Option 2", "Option 3", "Option 4", "Option 5", and "Option 6".

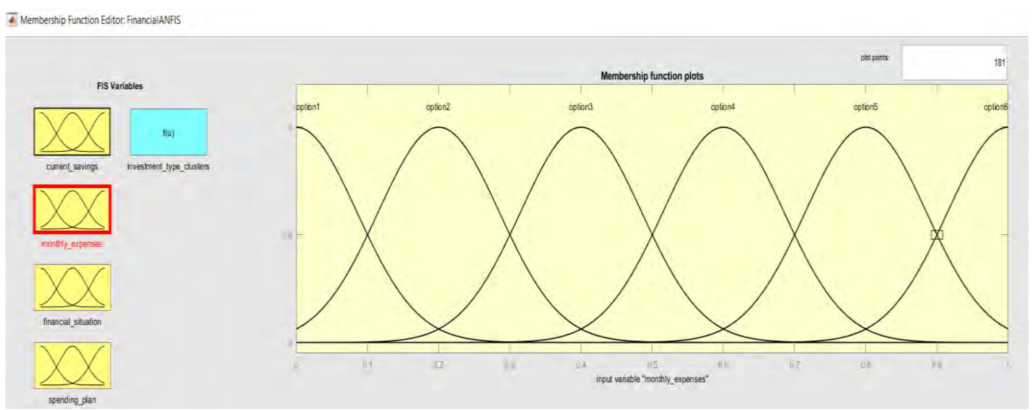


Figure 4-79. MFs shape for input 2 (monthly expenses) in the FinancialANFIS

Figure 4-79 displays the shape of the MFs for the second input of the FinancialANFIS. The MFs have a gaussian shape with six options, namely "Option 1", "Option 2", "Option 3", "Option 4", "Option 5", and "Option 6".

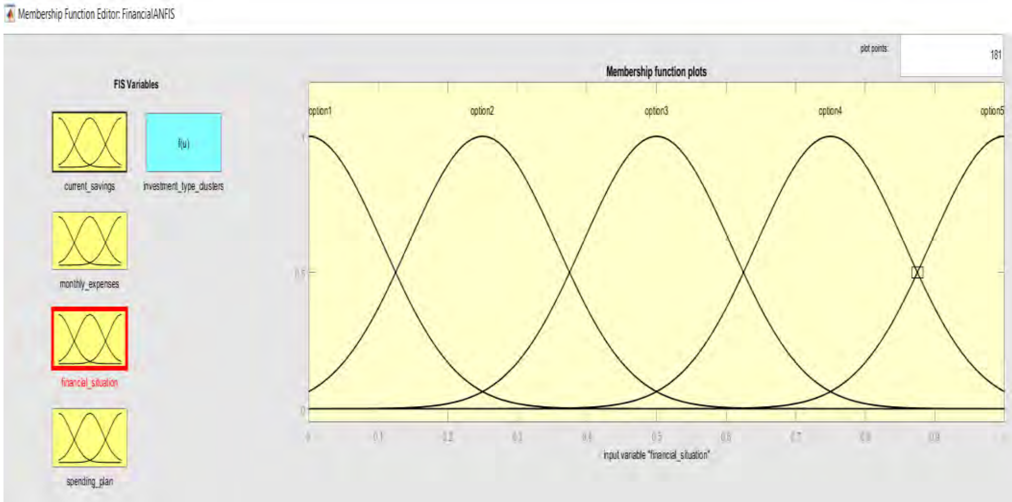


Figure 4-80. MFs shape for input 3 (financial situation) in the FinancialANFIS

Figure 4-80 illustrates the shape of the MFs for input 3 in the FinancialANFIS. The MFs are modeled using gaussmf and are represented by five options, namely "Option 1", "Option 2", "Option 3", "Option 4", and "Option 5".

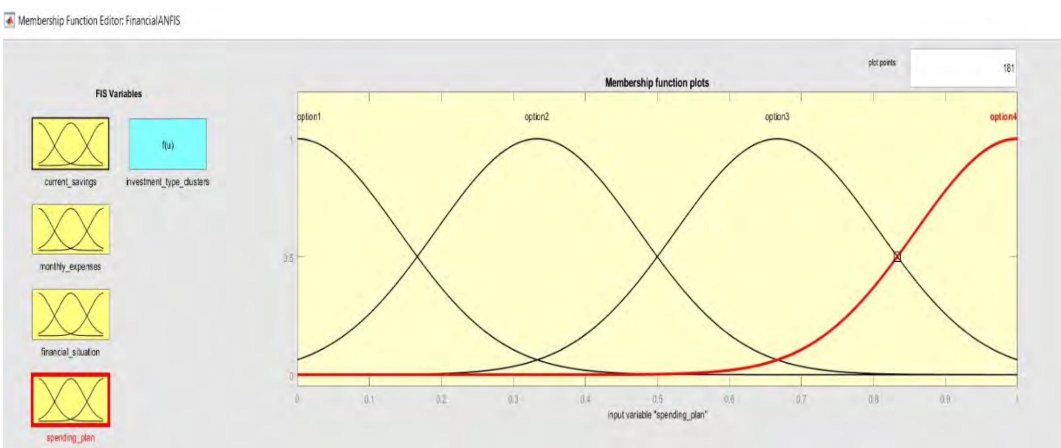


Figure 4-81. MFs shape for input 4 (spending plan for savings) in the FinancialANFIS

Figure 4-81 illustrates the shape of the MFs for the fourth input of FinancialANFIS. The MF shape is a triangular MF with four options, namely "Option 1", "Option 2", "Option 3", and "Option 4".

4.2.7. Proposing FinancialANFIS

The Financial ANFIS system was designed to help evaluate the financial situation of potential investors. To gather relevant information, four questions were asked to determine the current savings and financial situation of the individual. Based on the responses to these questions, four inputs were created for the Financial ANFIS system. The purpose of these inputs is to provide an understanding of the individual's financial status, which would be used to make recommendations for investments. The Financial ANFIS system uses these inputs, along with other relevant data, to make informed investment suggestions based on an individual's financial situation.

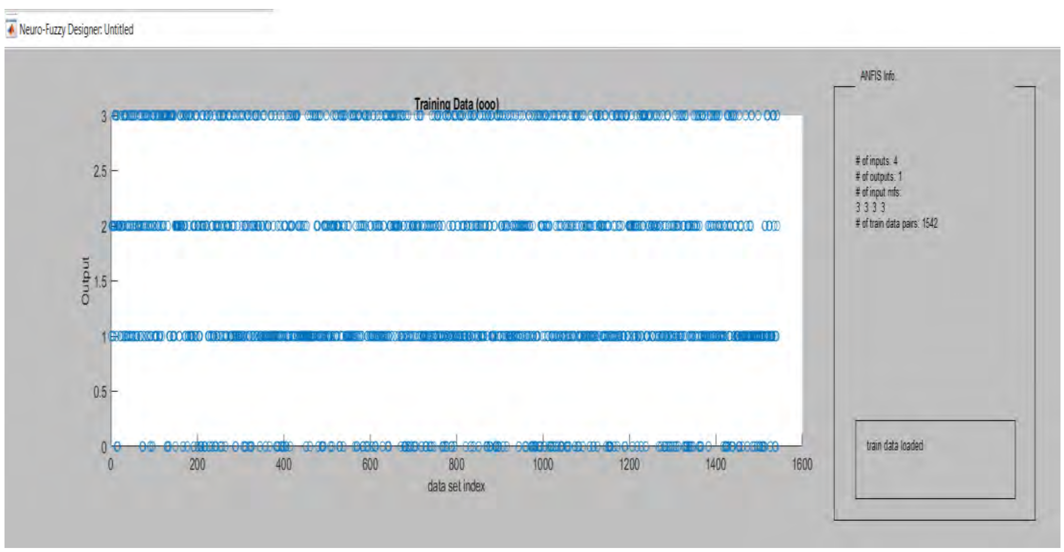


Figure 4-82. Prepared data in the FinancialANFIS

Figure 4-82 represents the pre-processed data for the next stages of training and validation in the Financial ANFIS system. To train the new FIS, a grid partition approach was employed, and the optimization method was hybridized with an error tolerance of 0 and a total of 3 epochs. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model. This resulted in the generation of the Financial ANFIS as a new FIS. Figure 4-83 displays a summary of the MFs for the Financial ANFIS system.

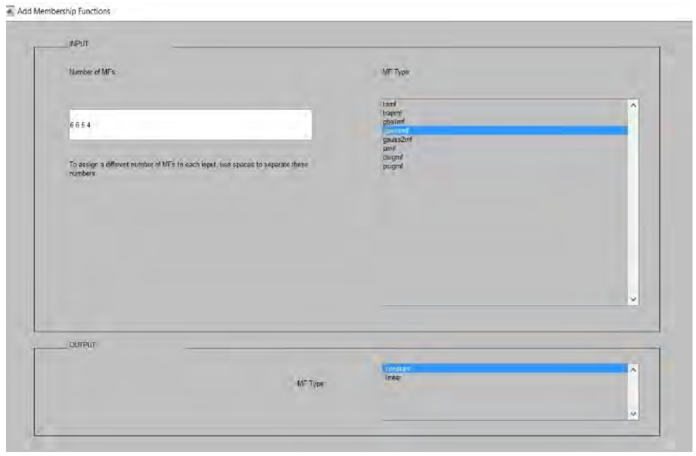


Figure 4-83. Information for generating the FinancialANFIS as a new FIS

The Financial ANFIS system is a newly generated FIS that contains 4 inputs and 1 output. The data set used in this system has 1542 records, with the x-axis representing the index of the data set and the y-axis showing the distribution of the output based on investment type clusters. In other words, the y-axis displays how the output (e.g., a predicted investment type) is distributed based on different investment-type groups or clusters.

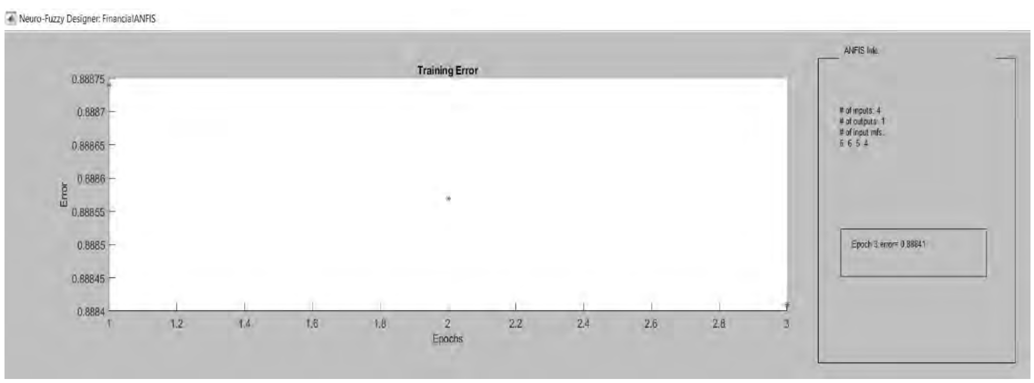


Figure 4-84. The trained FinancialANFIS

Figure 4-84 depicts the trained grid of the Financial ANFIS system. The system takes into account four inputs and one output related to investment type clusters. The training process for the FIS is hybrid and involves three epochs. The error for each epoch is around 0.89. The following information relates to the training data process of the Financial ANFIS.

The system has 1489 nodes, 720 linear parameters, 42 nonlinear parameters, and a total of 762 parameters. The training data consists of 1542 pairs and there are no checking data pairs. The system has 720 fuzzy rules. The training process begins with three epochs and the error for each epoch is recorded, with the final error being 0.888411. The minimum training RMSE is 0.888411. The training process is completed after the designated epoch number (3) is reached, indicating that the Financial ANFIS training is complete.

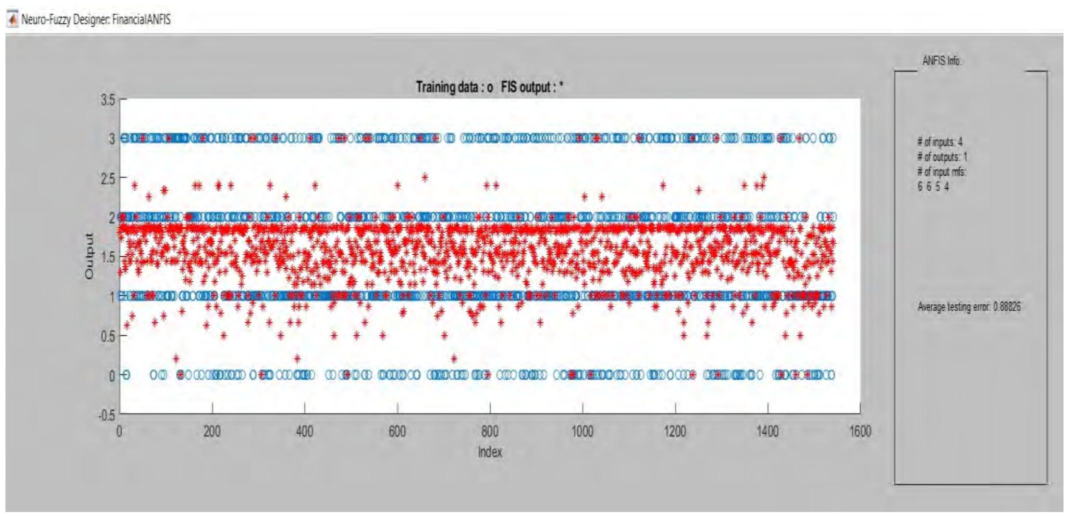


Figure 4-85. Trained data in the FinancialANFIS

Figure 4-85 represents the trained Financial ANFIS system. The figure displays the average training error, which is 0.88826. This suggests that the system has been trained to a certain level of accuracy, but there may still be room for improvement. Figure 4-47 illustrates a portion of the 720 rules generated by the proposed Financial ANFIS system, presented in a verbose format. This format provides detailed information about each rule, making it easier to understand the system's decision-making process. It is noteworthy that the generated rules can be altered based on expert opinions and feedback from investors. This allows for customization and refinement of the system to better suit specific needs and preferences.

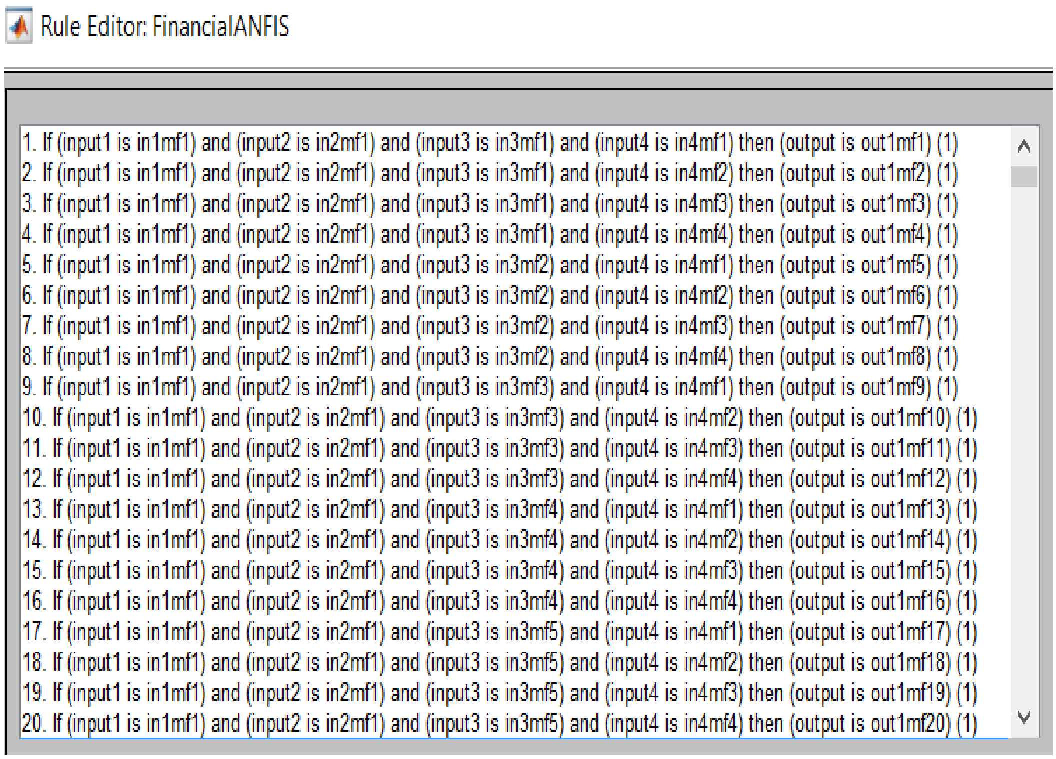
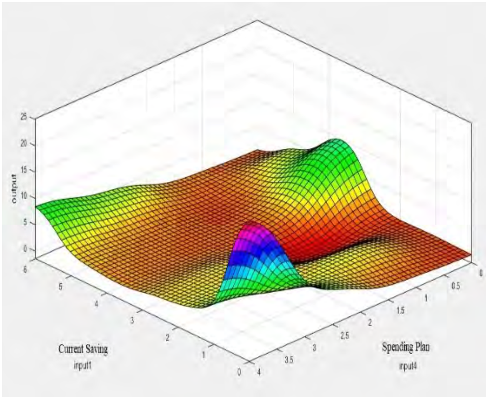


Figure 4-86. A part of the generated rules in the FinancialANFIS

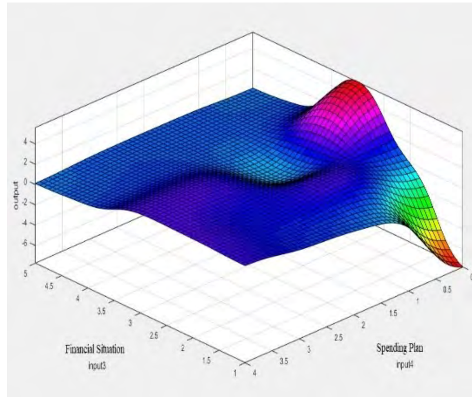
Figure 4-86 is a part of the generated rules in the Financial ANFIS system. Figure 4-48 depicts a portion of the rule viewer for the Financial ANFIS system, showcasing its open system. The figure highlights the presence of 720 rules and 101 plot points within the system. The rules generated in the Financial ANFIS system are used to make recommendations for financial investments. The number of rules, 720, represents the various conditions and scenarios that the system takes into consideration when making recommendations. The plot points, 101, are used to visualize the relationships between the inputs and outputs of the system. Overall, Figure 4-86 and Figure 4-87 provide a glimpse into the workings of the Financial ANFIS system and how it uses rules and plot points to make investment recommendations based on financial data.

Figure 4-87. A part of the rule viewer in the FinancialANFIS

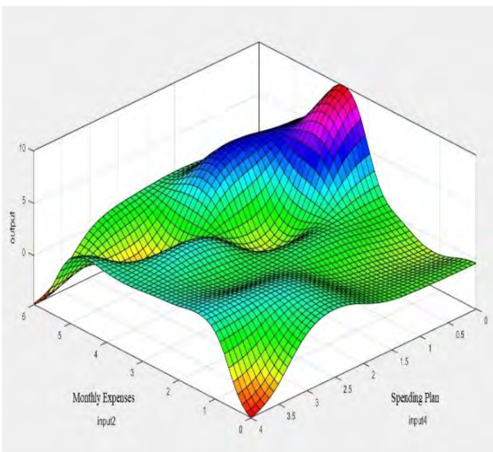
Attachment 4 provides information about the rule generation process of the Financial ANFIS. The Financial ANFIS system is designed to identify the relationship between financial factors and investment type. The graphs in Figures 4-88 (a-d) are 3D visual representations that illustrate the impact of certain input pairs on investment type. These graphs are nonlinear and show the investment type recommendations for a given set of inputs. The surface graphs are monolithic, meaning they are composed of a single piece and do not have any discontinuities or gaps. In other words, the graphs provide a comprehensive and seamless representation of the investment type recommendations based on the financial inputs.



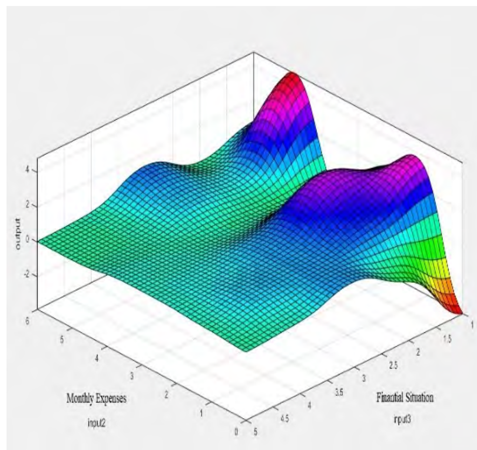
a. current savings & spending plan



c. financial situation & spending plan



b. monthly expenses & spending plan



d. monthly expenses & financial situation

Figure 4-88. Effectiveness of the relations of each pair of financial inputs on investment type

These figures demonstrate the effectiveness of the relationships between four financial input pairs and investment type. These input pairs are 4-88a. current savings and spending plan - this figure shows the impact of an individual's current savings balance and spending plan on their investment type selection. 4-88b. monthly expenses and spending plan - this figure displays the effect of an individual's monthly expenses and spending plan on their investment type choice. 4-88c. financial situation and spending plan - this figure highlights the impact of an individual's overall financial situation and spending plan on their investment type preference. 4-88d. monthly expenses and financial situation - this figure presents the effect of an individual's monthly expenses and overall financial situation on their investment type decision. In general, the figure aims to demonstrate the effectiveness of the relationships between these financial inputs and investment type, providing insights into how these factors influence an individual's investment choices. Figure 4-89 represents the structure of the Financial ANFIS Model. The model displays the inputs, MFs, various layers of the ANFIS system, and the output, which is a recommendation to investors regarding the selection of an investment type based on the defined clusters. The model highlights the layers of the system, including fuzzification, implication rules, normalization, defuzzification, and integration, which results in an aggregated output MF. The fuzzification layer assigns the inputs to their respective MFs, transforming them into a fuzzy representation. The implication rules layer uses these fuzzy inputs to determine the best investment type based on a set of predefined rules. The normalization layer ensures that the rules' strengths sum up to 1. The defuzzification layer converts the fuzzy outputs into crisp outputs, providing a clear investment recommendation. Finally, the integration layer aggregates the MFs into a single output MF, resulting in a comprehensive investment recommendation for the investors.

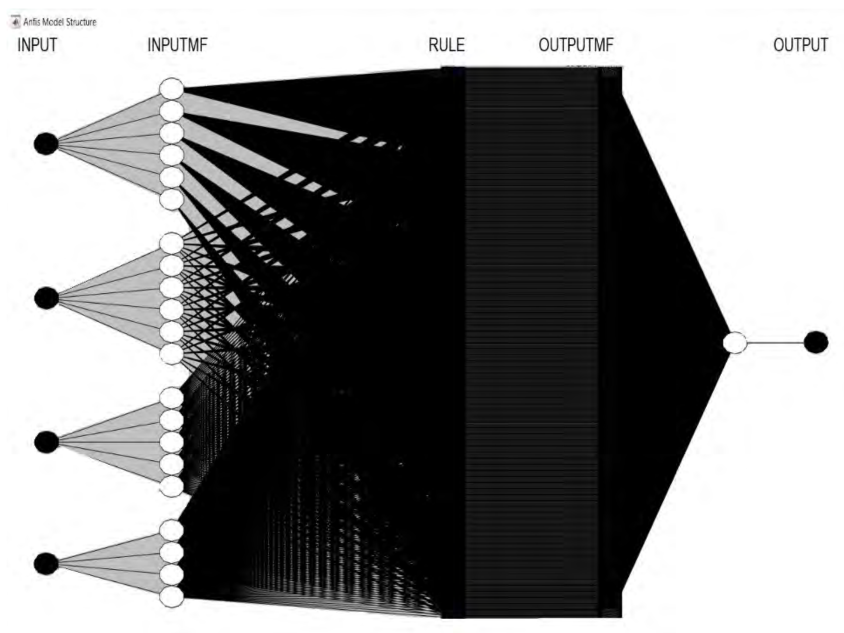


Figure 4-89. FinancialANFIS Model Structure

4.2.8. Managerial Traits ANFIS

The questionnaire used in the study measured the personal characteristics and managerial traits of potential investors to better understand their decision-making processes and investment choices. Nine questions were asked about the personal characteristics, which aimed to assess the respondents' personality traits in regard to finance. This information was used to understand the relationship between personal habits and attitudes towards finance and investment decisions.

Nine inputs were designed for potential investors' managerial traits using ANFIS based on the questionnaire responses. Input 1, related to planning and purposefulness, was based on five multiple-choice answers, with a five-degree scale: strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree. The questions were designed to assess the respondents' planning and purposefulness in life, including their goal-setting abilities, flexibility in their plans, and their ability to adapt to changing circumstances. The average of all answers was calculated and rounded to prepare the data for input 1, with MF1=strongly disagree, MF2=disagree, MF3=neither agree nor disagree, MF4=agree, and MF5=strongly agree. The frequency of each option was recorded in Table 4-9.

Table 4-9. MFs of the Managerial traits for input 1 (planning)

	Options	Frequency
MF1	Strongly disagree	4
MF2	Disagree	28
MF3	Neither agree nor disagree	814
MF4	Agree	690
MF5	Strongly agree	11

Input 2, labeled "stress," has two MFs related to decision making anxiety. The first option, "I get nervous after decision making," is assigned to MF1, while the second option, "I worry about whether my decision is right," is assigned to MF2.

Input 3, labeled "pace," has two MFs related to the speed at which potential investors work. The first option, "I work rather rushing," is assigned to MF1, while the second option, "I work more comfortably than a little rush," is assigned to MF2.

Input 4, labeled "influential," has two MFs related to the potential investor's personality. The first option, "I try to be influential," is assigned to MF1, and the second option, "I let things happen around me and I adjust to them," is assigned to MF2.

Input 5, labeled "daily schedule," has three MFs related to the potential investor's daily routine. The first option, "I plan it," is assigned to MF1, the second option, "It is dictated by my family," is assigned to MF2, and the third option, "My workplace dictates," is assigned to MF3.

Input 6, labeled "strategy," has two MFs related to the potential investor's planning style. The first option, "I can see my options to choose," is assigned to MF1, and the second option, "I limit it to my possibilities," is assigned to MF2.

Input 7, labeled "attachment," has two MFs related to the potential investor's work dedication. The first option, "I am dynamic in work with clear assignment of tasks," is assigned to MF1, and the second option, "I am working on disorganized, deconstructive, or passive," is assigned to MF2.

Table 4-9. MFs of the Managerial traits for input 2-9

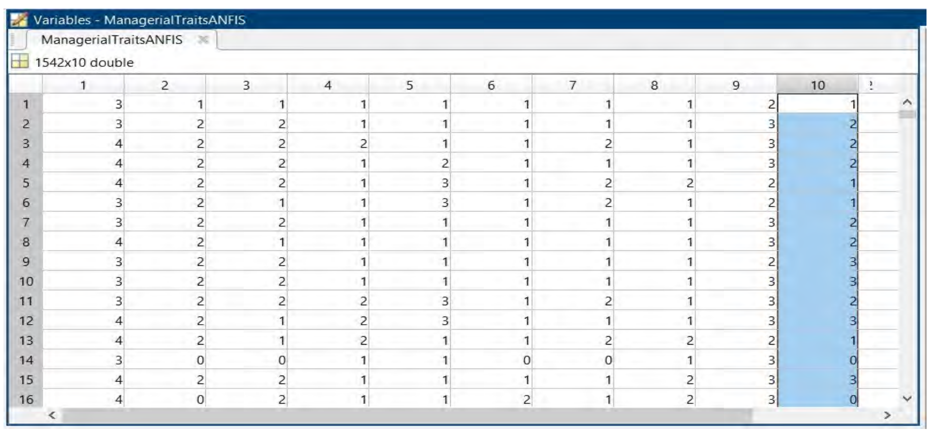
Inputs	**MFs**	**Options**	**Frequency**
Input2 stress	MF1	I get nervous after decision making	279
	MF2	I worry about whether my decision is right	1253
Input3 pace	MF1	I work rushing	591
	MF2	I work more comfortably than a little rush	949
Input4 influential	MF1	I try to be influential	1245
	MF2	I let things happen around me and I adjust to them	295
Input5 daily schedule	MF1	I plan it	820
	MF2	It is dictated by my family	111
	MF3	My workplace dictates	610
Input6 strategy	MF1	I can see my options to choose from	1380
	MF2	I limit it to my possibilities	159
Input7 attachment	MF1	I am dynamic in work with clear assignment of tasks	910
	MF2	I am working on disorganized, deconstructive, or passive	618
Input8 satisfaction	MF1	If the work, I am doing seems pointless	1142
	MF2	If the work, I am doing does not satisfy me mentally	398
Input9 planning time	MF1	1-3 weeks	104
	MF2	4-8 weeks	488
	MF3	more than 8 weeks or I do not usually plan my holidays in advance	949

Input 8, labeled "satisfaction," has two MFs related to the potential investor's job satisfaction. The first option, "If the work, I am doing seems pointless," is assigned to MF1, and the second option, "If the work, I am doing doesn't satisfy me mentally," is assigned to MF2.

Input 9, labeled "planning time," has three MFs related to the potential investor's vacation planning habits. The first option, "1-3 weeks," is assigned to MF1, the second option, "4-8 weeks," is assigned to MF2, and the third option, "more than 8 weeks or I do not usually plan my vacations," is assigned to MF3).

The frequency of each option for MFs of the Managerial traits for input 2-9 was recorded in Table 4-10.

The output of the investment type clustering is defined as three clusters. The investment type data includes various financial products such as listed stock mutual funds, voluntary pension funds, government securities/bonds, and other investment options. For all input and output data, 1542 training data pairs were used, with the Min-Max normalization technique applied for aggregation purposes.


Variables - ManagerialTraitsANFIS

ManagerialTraitsANFIS

1542x10 double

	1	2	3	4	5	6	7	8	9	10
1	3	1	1	1	1	1	1	1	2	1
2	3	2	2	1	1	1	1	1	3	2
3	4	2	2	2	1	1	2	1	3	2
4	4	2	2	1	2	1	1	1	3	2
5	4	2	2	1	3	1	2	2	2	1
6	3	2	1	1	3	1	2	1	2	1
7	3	2	2	1	1	1	1	1	3	2
8	4	2	1	1	1	1	1	1	3	2
9	3	2	2	1	1	1	1	1	2	3
10	3	2	2	1	1	1	1	1	3	3
11	3	2	2	2	3	1	2	1	3	2
12	4	2	1	2	3	1	1	1	3	3
13	4	2	1	2	1	1	2	2	2	1
14	3	0	0	1	1	0	0	1	3	0
15	4	2	2	1	1	1	1	2	3	3
16	4	0	2	1	1	2	1	2	3	0

Figure 4-90. A part of imported data to MATLAB to propose the ManagerialTraitsANFIS

The data imported into MATLAB consisted of 10 columns, with 9 columns related to the managerial traits of potential investors and 1 column related to investment type clusters, as shown in Figure 4-90. In the fuzzy function, the inputs and outputs for the managerial traits were specified for ANFIS. A new FIS was designed in the Sugeno type, called ManagerialTraitsANFIS. The imported data in MATLAB was used as inputs and outputs for the proposed ManagerialTraitsANFIS.

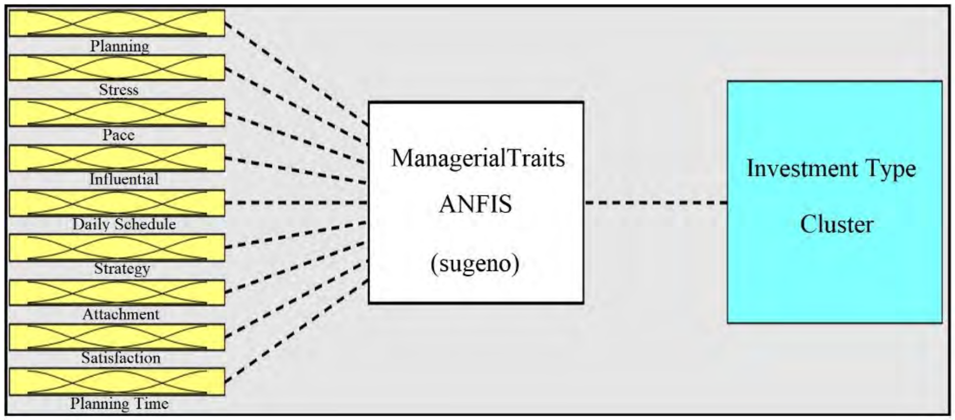


Figure 4-91. The properties of the ManagerialTraitsANFIS

Figure 4-91 depicts the designed Managerial Traits ANFIS system, referred to as "ManagerialTraitsANFIS", and its properties. Nine inputs were incorporated, including: 1) Planning, 2) Stress, 3) Pace, 4) Influential, 5) Daily Schedule, 6) Strategy, 7) Attachment, 8) Satisfaction, and 9) Planning Time. The system outputs one result, an investment type cluster.

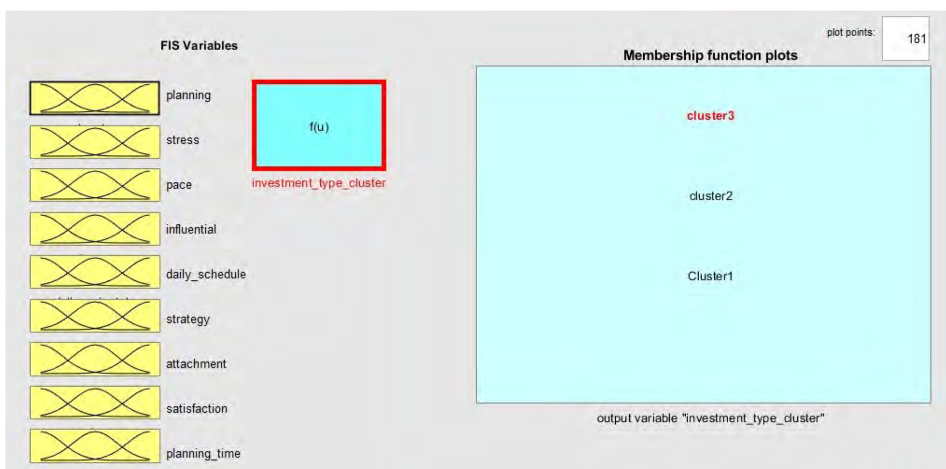


Figure 4-92. Output MFs in the ManagerialTraitsANFIS

Figure 4-92 illustrates the shape of the MFs for the output in the Managerial Traits ANFIS model. Three MFs of the constant type are utilized for the investment types, including "Cluster 1", "Cluster 2", and "Cluster 3."

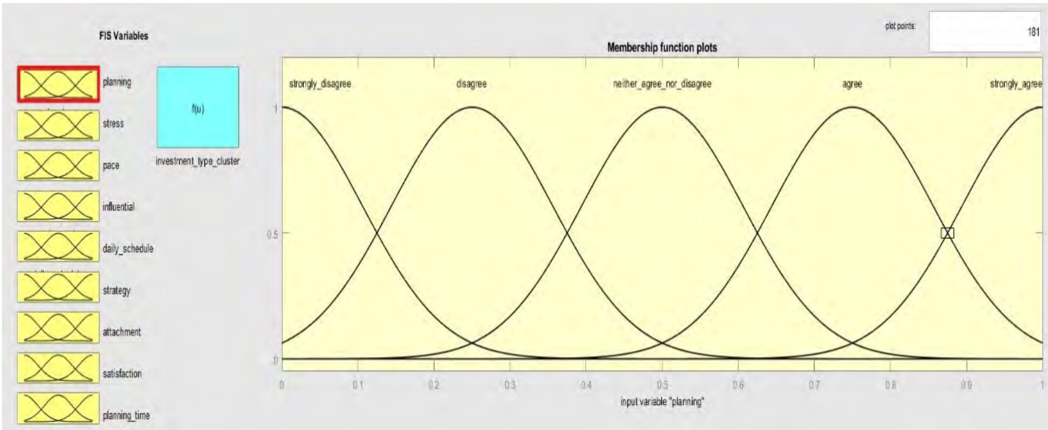


Figure 4-93. MFs shape for input 1 (planning) in the ManagerialTraitsANFIS

Figure 4-93 displays the MFs for the first input of the Managerial Traits ANFIS model. The MFs are represented using gaussian curves with five functions on a five-point scale, ranging from "strongly disagree" to "strongly agree".

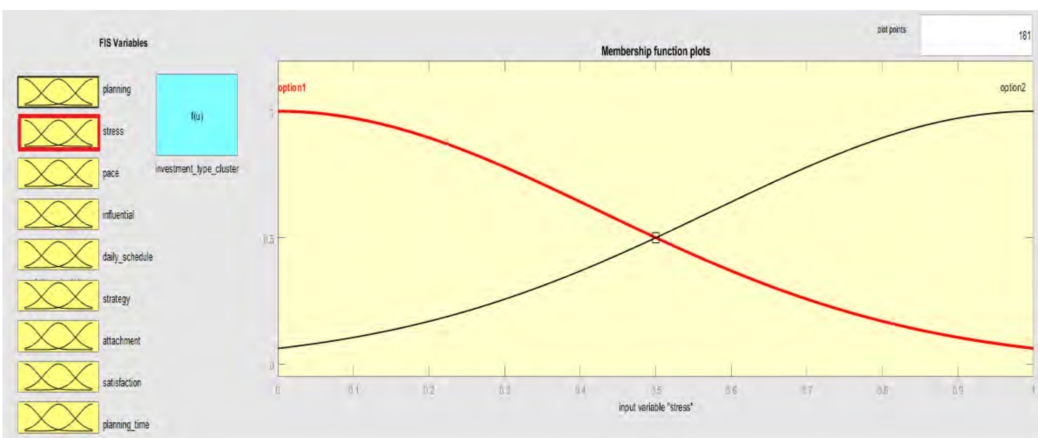


Figure 4-94. MFs shape for input 2 (stress) in the ManagerialTraitsANFIS

Figure 4-94 showcases the shape of the MFs for input 2 in the Managerial Traits ANFIS model. The shape of the MFs is represented using Gaussmf with two options, "Option 1" and "Option 2."

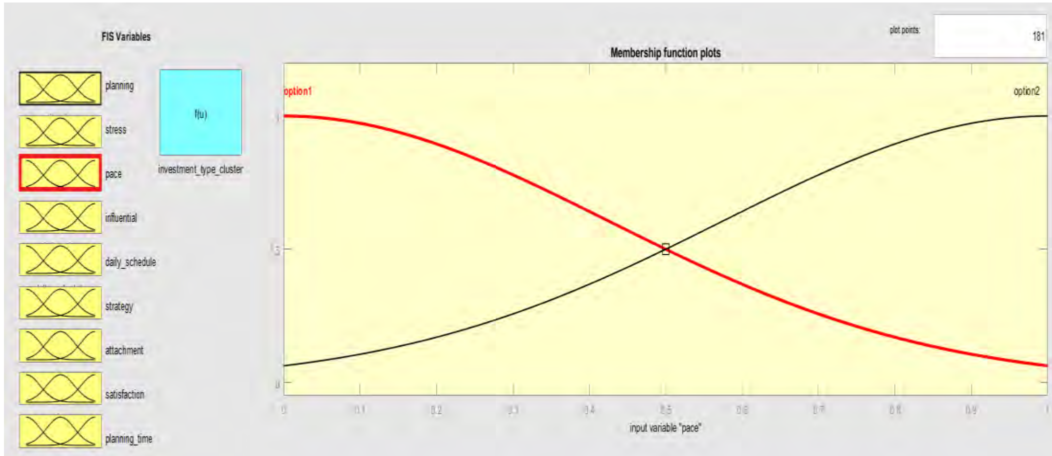


Figure 4 95. MFs shape for input 3 (pace) in the ManagerialTraitsANFIS

Figure 4-95 shows the MF shape for input 3 in the ManagerialTraitsANFIS model. The MF shape used is Gaussian, with two options, "Option 1" and "Option 2."

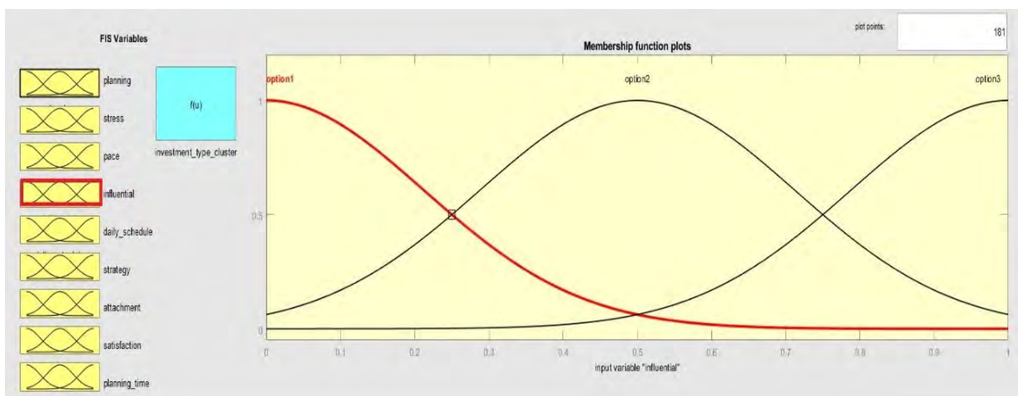


Figure 4-96. MFs shape for input 4 (influential) in the ManagerialTraitsANFIS

Figure 4-96 illustrates the MF shapes for input 4 of the Managerial Traits ANFIS model. The MF shape utilized is Gaussmf and there are three options available, namely "Option 1", "Option 2", and "Option 3".

Figure 4-97. MFs shape for input 5 (daily schedule) in the ManagerialTraitsANFIS

Figure 4-97 displays the shape of the MFs for input 5 in the ManagerialTraitsANFIS model. The MF shape is represented by the Gaussmf, with two options, "Option 1" and "Option 2".

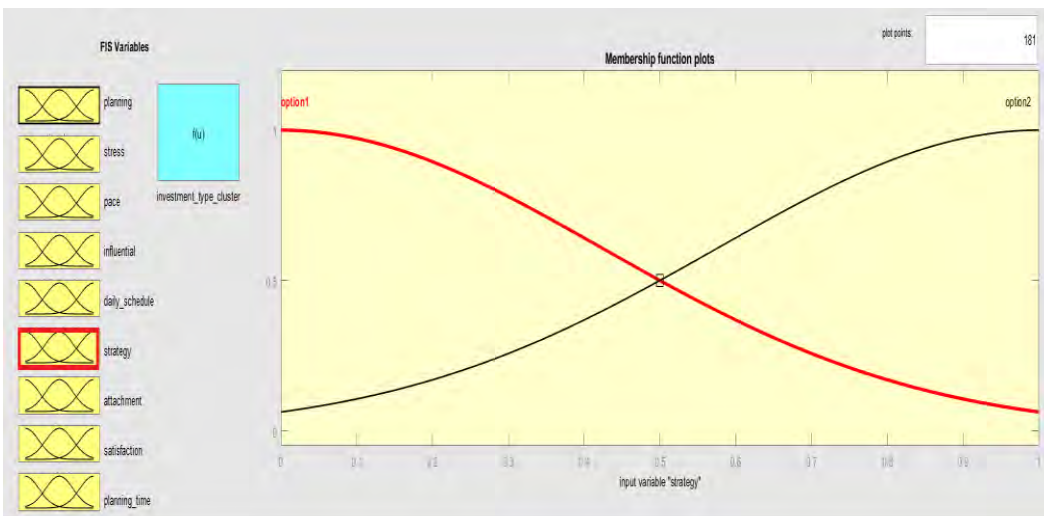


Figure 4-98. MFs shape for input 6 (strategy) in the ManagerialTraitsANFIS

Figure 4-98 presents the MF shapes for input 6 of the Managerial Traits ANFIS. The MF shape adopted is Gaussian, with two options, "Option 1" and "Option 2".

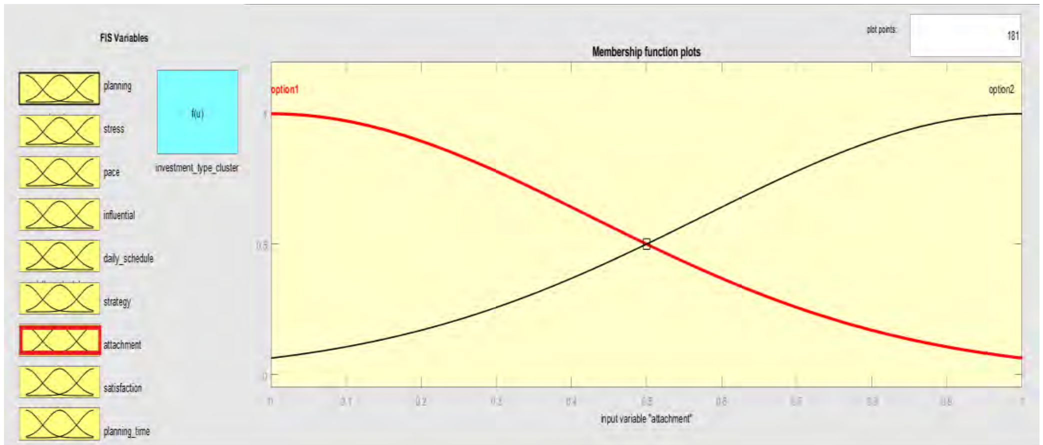


Figure 4-99. MFs shape for input 7 (attachment) in the ManagerialTraitsANFIS

Figure 4-99 shows the MFs for input 7 in the ManagerialTraitsANFIS. The MFs are represented using the gaussian shape, with two options: "Option 1" and "Option 2".

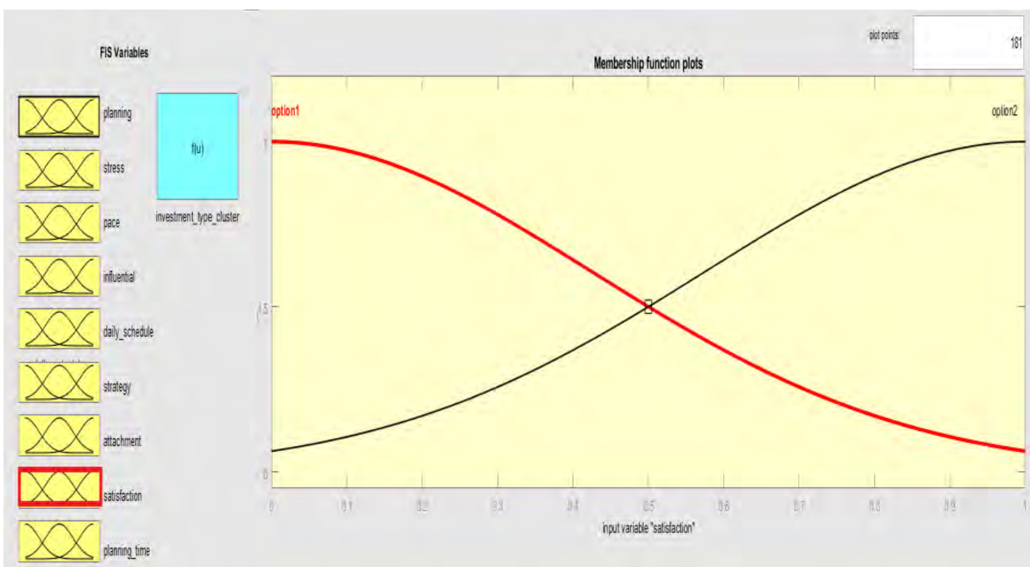


Figure 4-100. MFs shape for input 8 (satisfaction) in the ManagerialTraitsANFIS

Figure 4-100 illustrates the shape of the MFs for the input 8 of the ManagerialTraitsANFIS. The MF shape utilized is Gaussian, with two options present: "Option 1" and "Option 2" .

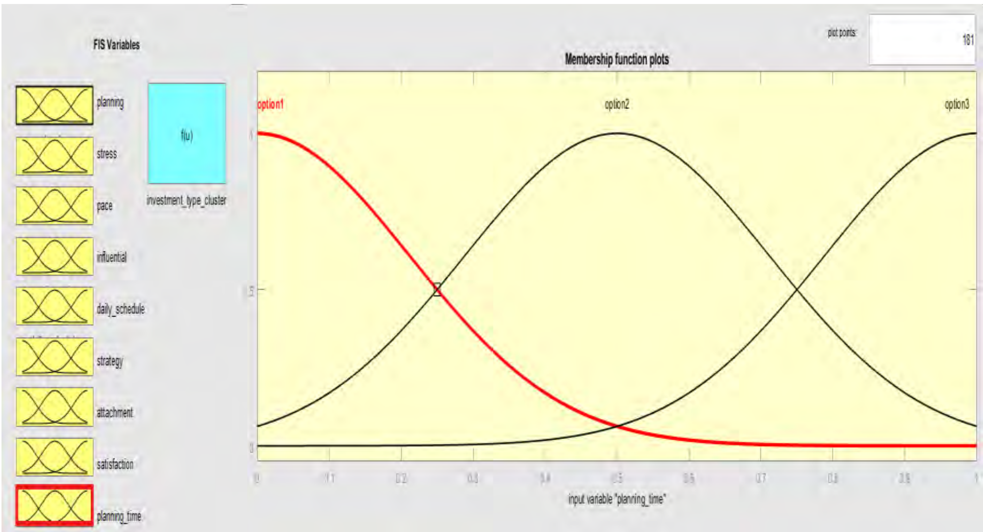


Figure 4-101. MFs shape for input 9 (planning time) in the ManagerialTraitsANFIS

Figure 4-101 illustrates the shape of the MFs for the 9th input of the Managerial Traits ANFIS. The MF shape is represented using gaussian functions with three options, "Option 1", "Option 2", and "Option 3".

4.2.9. Proposing ManagerialTraitsANFIS

The ManagerialTraitsANFIS is a system that uses ANFIS to make predictions. It consists of nine inputs and one output. The input parameters are used to make predictions about the output, which in this case is managerial traits. The system uses triangular MFs to make these predictions. The use of three MFs indicates that the inputs are grouped into three distinct categories, each with its own MF. These categories and the associated MFs are used to make predictions about the output, in this case, managerial traits. The ANFIS system is trained using historical data and makes predictions based on that training.

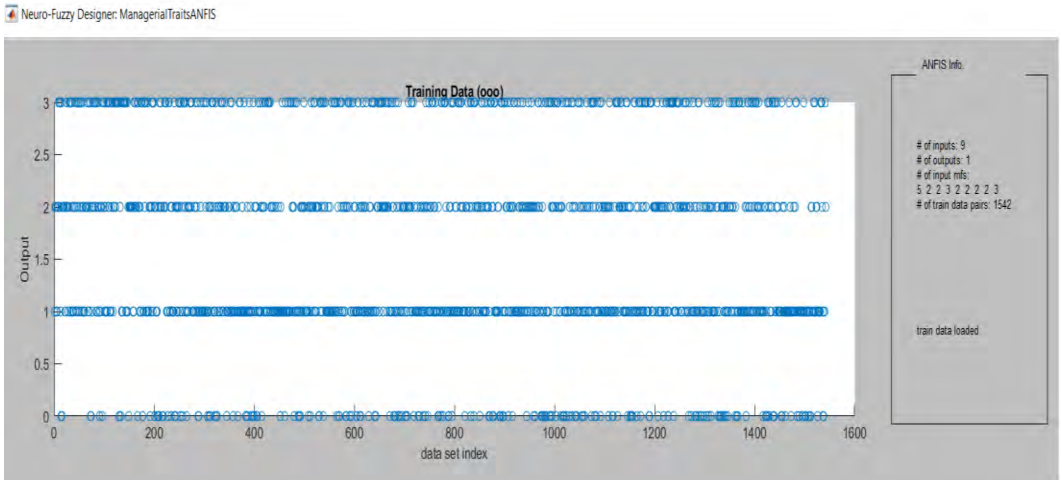


Figure 4-102. Prepared data in the ManagerialTraitsANFIS

Figure 4-102 represents the data preparation for the Managerial Traits ANFIS system. This data is used for the training and validation steps in the system. To prepare the data, a grid partition method was employed, and the optimization process was performed using a hybrid method with an error tolerance of 0 and 3 epochs. This resulted in the generation of the Managerial Traits ANFIS as a new FIS. Figure 4-103 showcases the summary information about the MFs of the Managerial Traits ANFIS system. The MFs represent the shape and characteristics of the input data used by the system to make its predictions.

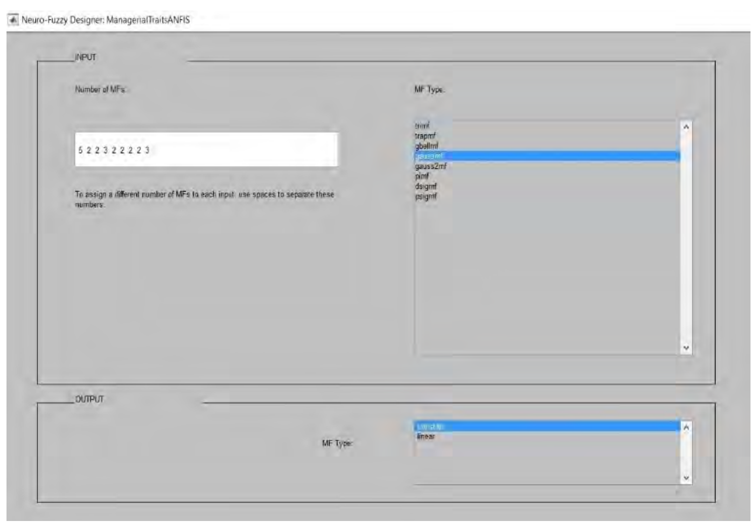


Figure 4-103. Information for generating the ManagerialTraitsANFIS as a new FIS

The ManagerialTraitsANFIS is a newly generated FIS that consists of 9 inputs and 1 output. The data set used to develop this system contains 1542 data points and is plotted on the x-axis. The y-axis displays the distribution of the output, based on the investment-type clusters. In other words, the data set provides information on the relationship between the 9 inputs and the investment type, and the y-axis shows how the output is spread across different investment-type clusters.

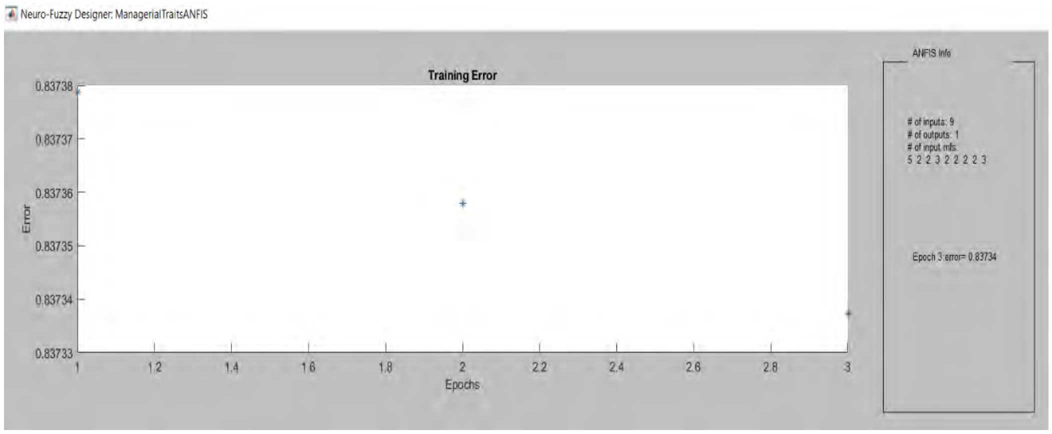


Figure 4-104. The trained ManagerialTraitsANFIS

The figure 4-104 presents the trained ManagerialTraitsANFIS network. The network has 4 inputs and one output, which represents investment type clusters. The training process of the FIS was performed using a hybrid approach. The error for each epoch is around 0.84, as shown in the process. The ANFIS information section provides details on the number of nodes (5818), linear parameters (2880), nonlinear parameters (46), total parameters (2926), training data pairs (1542), and fuzzy rules (2880) used in the network. The training process of the ManagerialTraitsANFIS was initiated, and the results are displayed after each epoch. The training process was completed after 3 epochs, and the final minimal training RMSE is 0.837341. This value represents the accuracy of the trained network in making predictions. The selection of three epochs for the training process is determined by the system based on factors such as the size and complexity of the data set, the available computational resources, and the desired level of accuracy for the model.

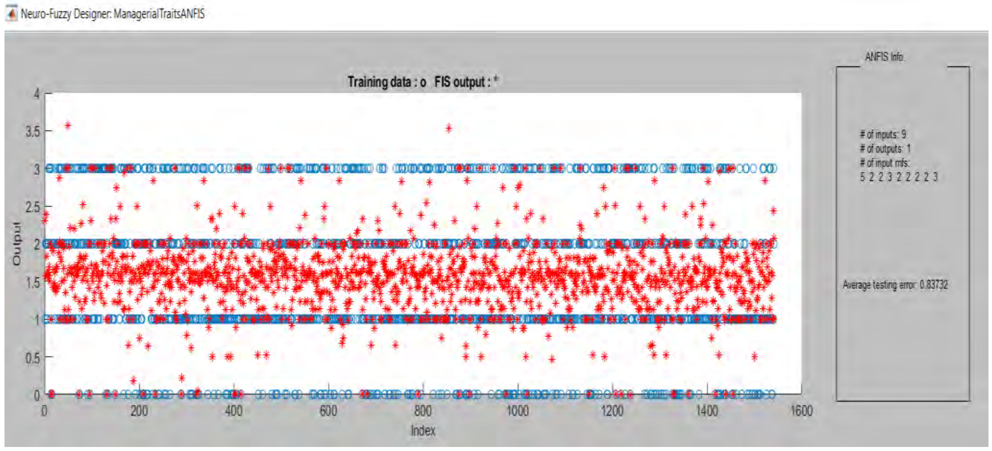


Figure 4-105. Trained data in the ManagerialTraitsANFIS

Figure 4-105 depicts the trained Managerial Traits ANFIS system. The average training error of the system is 0.83732, indicating that the system has achieved a certain level of accuracy in its training. The Managerial Traits ANFIS system has generated 2,880 rules. Figure 4-106 showcases a portion of the generated rules in the proposed Managerial Traits ANFIS system, displayed in a detailed format. This provides the flexibility to add, modify, or remove rules based on expert opinions and feedback from investors.

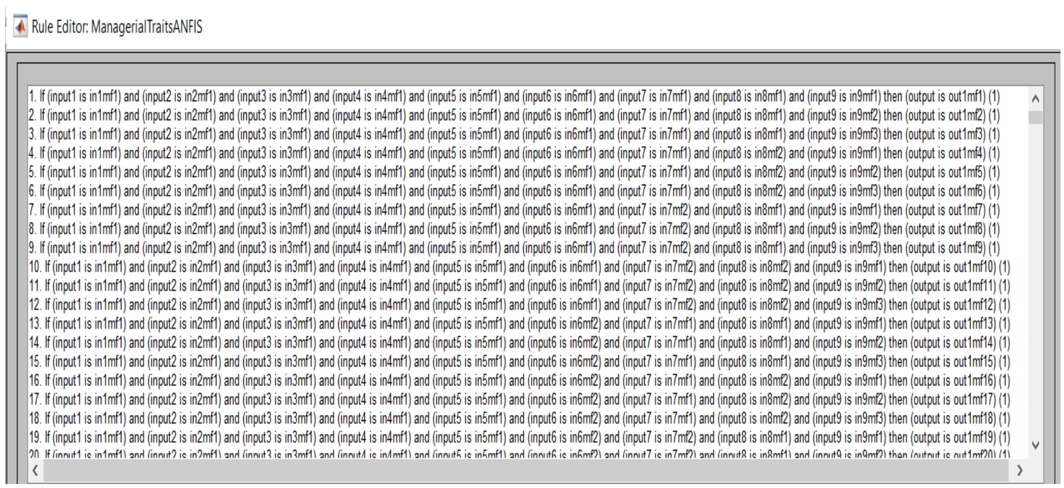
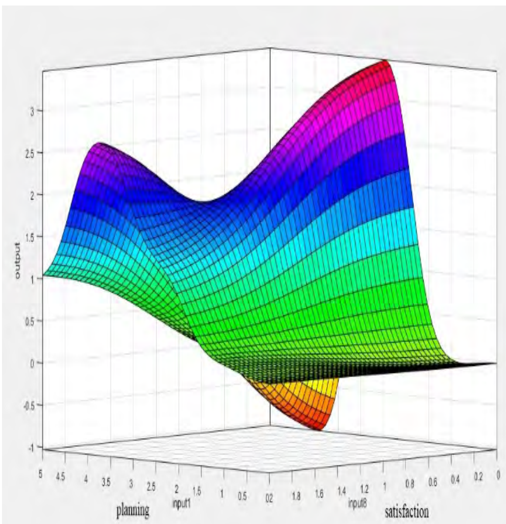
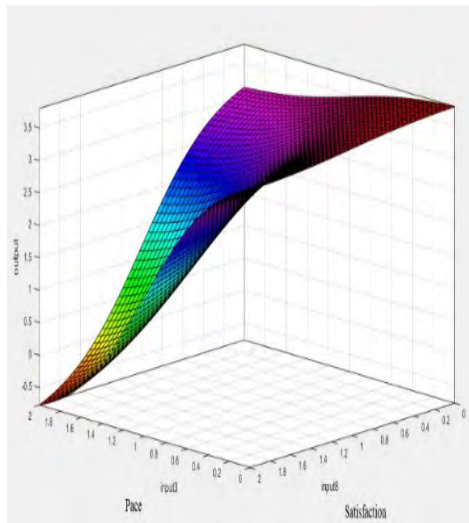


Figure 4-106. A part of the generated rules in the ManagerialTraitsANFIS

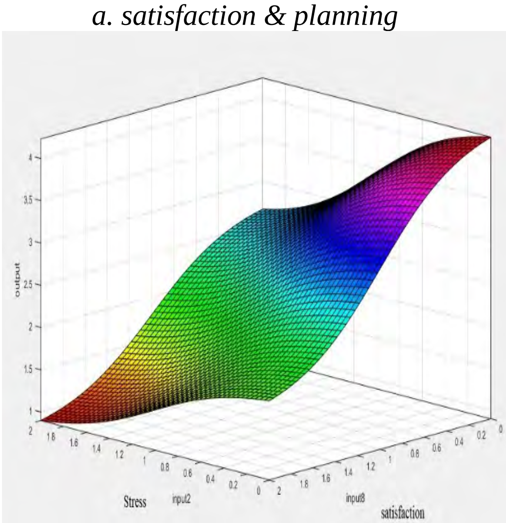
Attachment 5 provides information on the rule generation process by the Managerial Traits ANFIS. The Managerial Traits ANFIS system is designed to determine the relationship between various managerial traits and their impact on investment type selection. Figures 4-107 (a-h) depict the 3D graphs that showcase the effect of certain input pairs on investment type. These surface graphs are nonlinear and monolithic, displaying the recommended investment type for specific inputs. These graphs provide a visual representation of the relationship between managerial traits and investment type, making it easier to understand the system's recommendations.



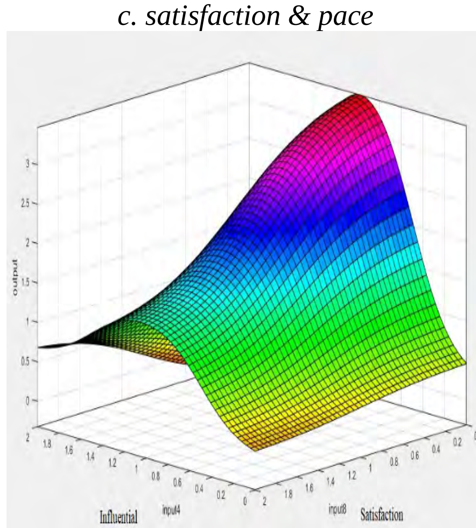
a. satisfaction & planning



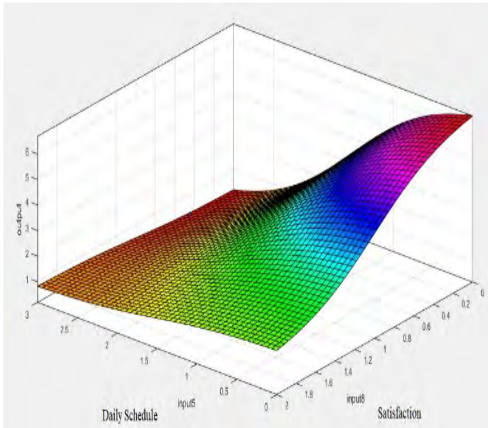
c. satisfaction & pace



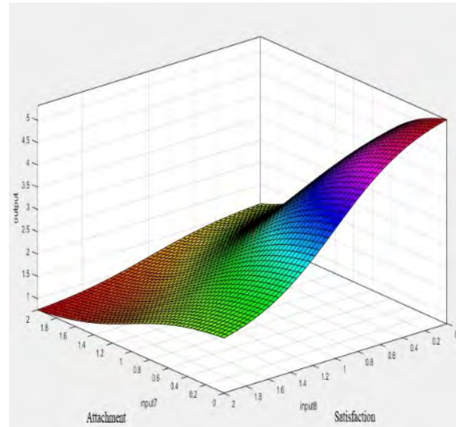
b. satisfaction & stress



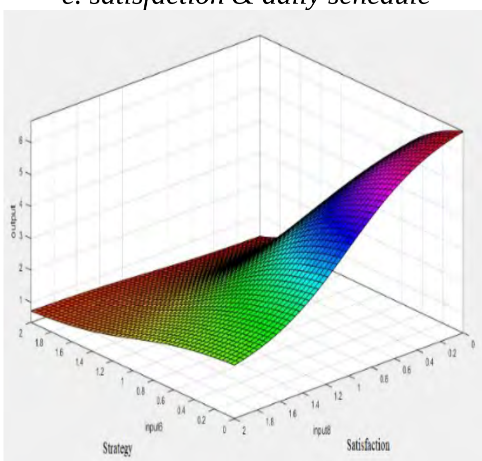
d. satisfaction & influential



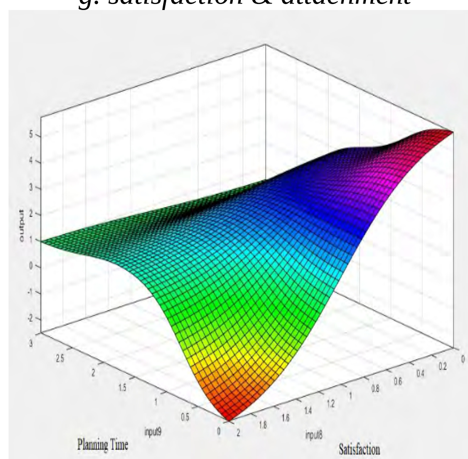
e. satisfaction & daily schedule



g. satisfaction & attachment



f. satisfaction & strategy



h. satisfaction & planning time

Figure 4-107. Effectiveness of the relations of each pair of managerial traits inputs on investment type

Figure 4-107 illustrates the effectiveness of the relationships between pairs of managerial traits inputs on investment type. The relationships being analyzed include Figure 4-107a shows the satisfaction and Planning. This relationship looks at the impact of an individual's level of satisfaction and their ability to plan on their investment type. Figure 4-107b shows the satisfaction and Stress: This relationship assesses the effect of an individual's satisfaction level and their stress level on their investment type. Figure 4-107c shows the satisfaction and Pace. This relationship evaluates the influence of an individual's satisfaction level and their preferred pace of work on their investment type. Figure 4-107d shows the satisfaction and Influential. This relationship examines the relationship between an individual's level of satisfaction and their level of influence on their investment type. Figure 4-107e shows the satisfaction and Daily Schedule. This relationship explores the impact of an individual's satisfaction level and their daily schedule on their investment type. Figure 4-107f shows the satisfaction and Strategy. This relationship analyzes the relationship between an individual's satisfaction level and their preferred strategy on their investment type. Figure 4-107g shows the satisfaction and Attachment. This relationship examines the effect of an individual's satisfaction level and their attachment to certain things or people on their investment type. Figure 4-107h shows the satisfaction and Planning Time. This relationship evaluates the influence of an individual's satisfaction level and their preferred time for planning on their investment type. The purpose of these figures is to provide insight into the different relationships between managerial traits and investment type, and to determine the effectiveness of these relationships in influencing investment decisions.

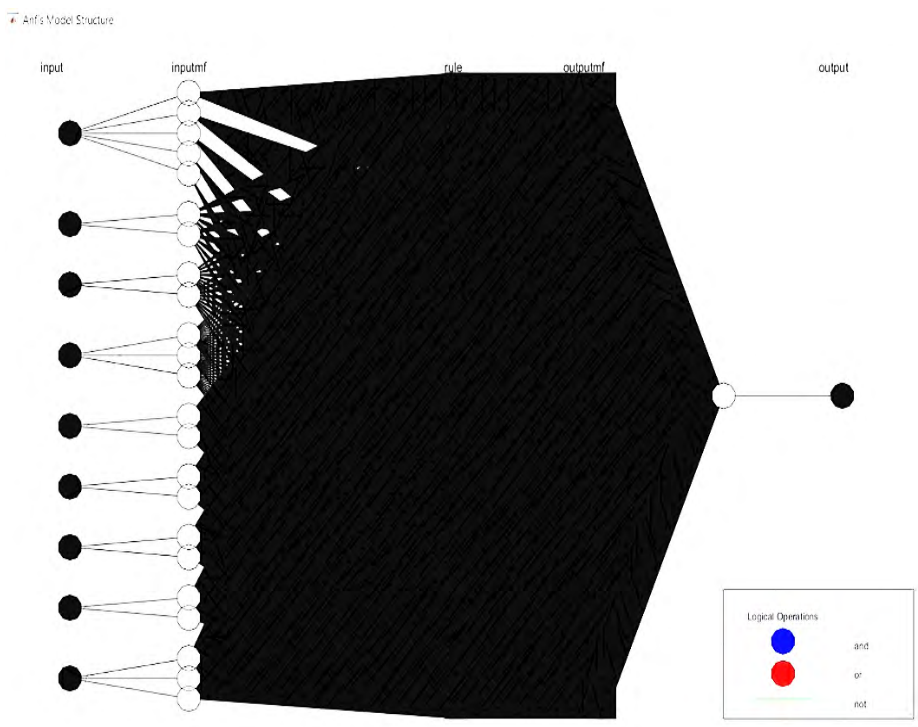


Figure 4-108. ManagerialTraitsANFIS Model Structure

Figure 4-108 shows the structure of the ManagerialTraitsANFIS model, which is used to make investment-type recom mendations to investors based on their managerial traits. The model includes nine inputs, each with its own set of MFs that map the inputs to fuzzy sets. The inputs represent various managerial traits, such as

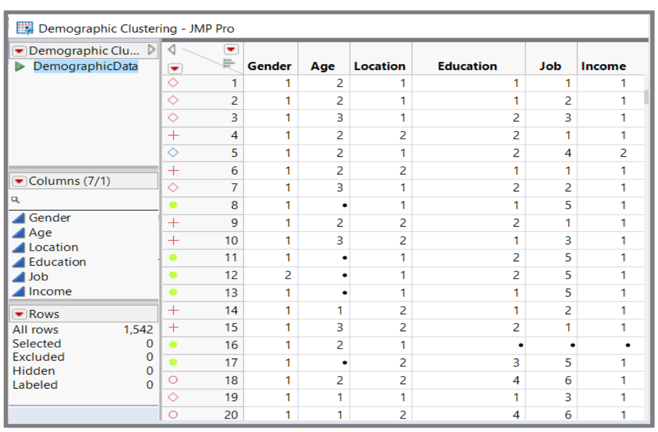
Planning, Stress, Pace, Influential, Daily Schedule, Strategy, Attachment, Satisfaction, and Planning Time. The model also includes several layers, including fuzzification, implication rules, normalization, defuzzification, and integration. The fuzzification layer maps the crisp inputs to fuzzy sets based on the MF. The implication rules layer uses fuzzy sets to generate a set of IF-THEN rules that describe the relationship between the inputs and the output. The normalization layer normalizes the output of the implication rules to ensure that the output MF is a valid probability distribution. The defuzzification layer maps the output MF to a crisp value, which represents the investment type recommendation. The integration layer aggregates the output of the defuzzification layer to generate a final output. The figure is a graphical representation of the ANFIS architecture, it shows how the inputs are passing through different layers and finally getting an output which is the recommended investment type for the investors. It also shows how different MFs are used in the fuzzification layer and how the rules are generated based on the inputs and MFs. Overall, the figure illustrates the process by which the ManagerialTraitsANFIS model generates investment-type recommendations for investors based on their managerial traits.

4.3. Inputs for Combined Investment Type Recommender ANFIS

In the second step of proposing the system, six categories of data were used as inputs for the ANFIS system, including "respondents' demographics," "key factors in investment decision making by respondents," "personality traits, knowledge, and ability of the respondents," "respondent's experiences," "respondents' financial situation," and "managerial traits of the respondents." The JMP software was utilized to cluster each category of data and create inputs for the combined ANFIS, employing the K-Means and SOM methods (JMP Documentation, 2015). The combined method of K-Means and SOM was used to cluster the demographic data, as the result of the combined clustering method was superior to using K-Means alone for this group of data with multiple features. SOM is an unsupervised machine learning technique that can be used to cluster data with many features. In this method, in addition to clustering the data, they are mapped to a two-dimensional map, making it easier to visualize the clusters. The JMP software considers the center of the clusters selected by K-Means as a point and calculates the probability of the presence of that point in each group (JMP Documentation, 2015). The process of clustering is repeated in two steps based on the expectation-maximization (EM) algorithm. In the expectation step, the probability of the presence of each point in a cluster is calculated, and in the second step, a new center is identified for each cluster based on the probability of presence. This process is repeated until the stability of the clusters is achieved (Clustering Methods for Unsupervised Machine Learning, 2019). The JMP software uses the CCC to select the optimal number of clusters, which fits the data best, based on the highest CCC value (SAS Help Center: Cubic Clustering Criterion, 2015). The performance of the CCC is evaluated using Monte Carlo methods.

4.3.1. Clustering Input demographics data

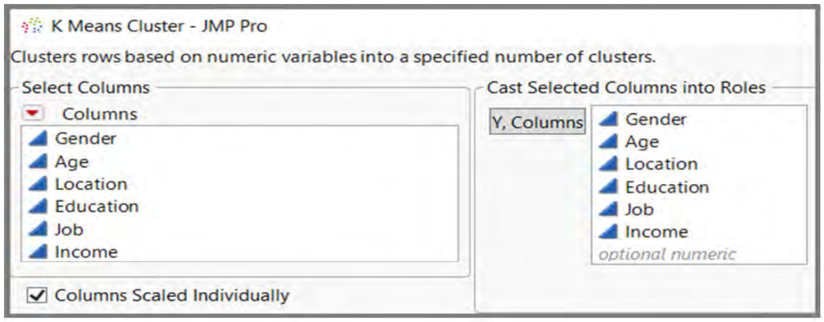
In the initial step, demographic data was brought into JMP, consisting of 1542 rows and 6 columns. Figure 4-109 depicts the imported demographic data within JMP, which was then subjected to further processing. The data was prepared using the K-Means and SOM methods within JMP.



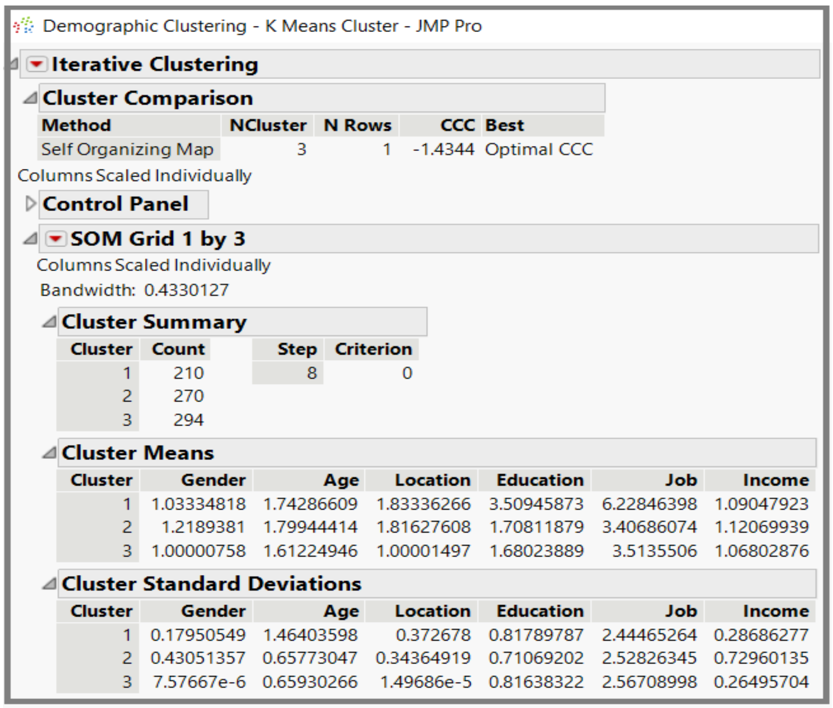
	Gender	Age	Location	Education	Job	Income
1	1	2	1	1	1	1
2	1	2	1	1	2	1
3	1	3	1	2	3	1
4	1	2	2	2	1	1
5	1	2	1	2	4	2
6	1	2	2	1	1	1
7	1	3	1	2	2	1
8	1	•	1	1	5	1
9	1	2	2	2	1	1
10	1	3	2	1	3	1
11	1	•	1	2	5	1
12	2	•	1	2	5	1
13	1	•	1	1	5	1
14	1	1	2	1	2	1
15	1	3	2	2	1	1
16	1	2	1	•	•	•
17	1	•	2	3	5	1
18	1	2	2	4	6	1
19	1	1	1	1	3	1
20	1	1	2	4	6	1

Figure 4-109. A part of imported demographic data in JMP

In the next step, the data rows are grouped into a specified number of clusters based on the numeric variables. The columns in the data include information about the gender, age, location, education, job, and income of the respondents. This process is visualized in Figure 4-110.



Figures 4-110. Y columns to cluster demographics data


Demographic Clustering - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	N Rows	CCC	Best
Self Organizing Map	3	1	-1.4344	Optimal CCC

Columns Scaled Individually

Control Panel

SOM Grid 1 by 3

Columns Scaled Individually

Bandwidth: 0.4330127

Cluster Summary

Cluster	Count
1	210
2	270
3	294

Step	Criterion
8	0

Cluster Means

Cluster	Gender	Age	Location	Education	Job	Income
1	1.03334818	1.74286609	1.83336266	3.50945873	6.22846398	1.09047923
2	1.2189381	1.79944414	1.81627608	1.70811879	3.40686074	1.12069939
3	1.00000758	1.61224946	1.00001497	1.68023889	3.5135506	1.06802876

Cluster Standard Deviations

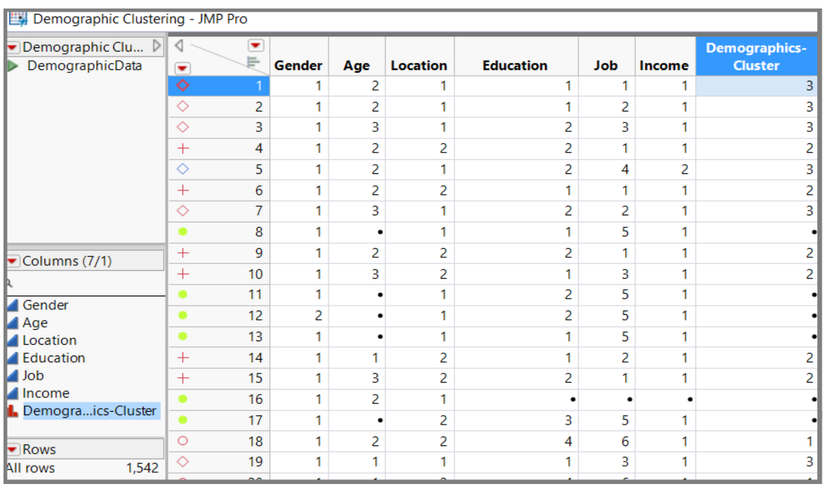
Cluster	Gender	Age	Location	Education	Job	Income
1	0.17950549	1.46403598	0.372678	0.81789787	2.44465264	0.28686277
2	0.43051357	0.65773047	0.34364919	0.71069202	2.52826345	0.72960135
3	7.57667e-6	0.65930266	1.49686e-5	0.81638322	2.56708998	0.26495704

Figure 4-111. Iterative Clustering of demographic data by K Means & SOM method in JMP

The clustering indicated by JMP is determined using the CCC method. Figure 4-111 represents the results of using a combined K-Means and SOM method to group demographics into three clusters. The clustering process is iterative and is divided into 8 steps, with the final count of the first cluster being 210, the second cluster 270, and the third cluster 294. The means and standard deviations for each cluster are also provided. The script for the clustering is written in Python and uses the K Means Cluster function. The script specifies the variables to be clustered

(Gender, Age, Location, Education, Job, Income) and sets the number of clusters to be 3. The script also defines the use of SOM and the standard deviations of each cluster to be displayed in a report. The script (supports Python) for the clustering is the following:

```
K Means Cluster(
        Y( :Gender, :Age, :Location, :Education, :Job, :Income ),
        {SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
        Number of Clusters( 3 ), SOM, Go},
        SendToReport(
                Dispatch( {}, "Cluster Comparison", OutlineBox, {Close( 1 )} ),
                Dispatch(
                        {"SOM Grid 1 by 3"},
                        "Cluster Standard Deviations",
                        OutlineBox,
                        {Close( 0 )}
                )
        )
)
```



	Gender	Age	Location	Education	Job	Income	Demographics-Cluster
1	1	2	1	1	1	1	3
2	1	2	1	1	2	1	3
3	1	3	1	2	3	1	3
4	1	2	2	2	1	1	2
5	1	2	1	2	4	2	3
6	1	2	2	1	1	1	2
7	1	3	1	2	2	1	3
8	1	•	1	1	5	1	•
9	1	2	2	2	1	1	2
10	1	3	2	1	3	1	2
11	1	•	1	2	5	1	•
12	2	•	1	2	5	1	•
13	1	•	1	1	5	1	•
14	1	1	2	1	2	1	2
15	1	3	2	2	1	1	2
16	1	2	1	•	•	•	•
17	1	•	2	3	5	1	•
18	1	2	2	4	6	1	1
19	1	1	1	1	3	1	3

Figure 4-112. A part of clusters for each row of demographic data

Figure 4-112 illustrates the addition of a new column to the data table, which holds the cluster assignments for each row. The cluster assignment is based on numeric variables in the demographic data, and the number of clusters is specified beforehand. The column displays the cluster number assigned to each row in the data table.

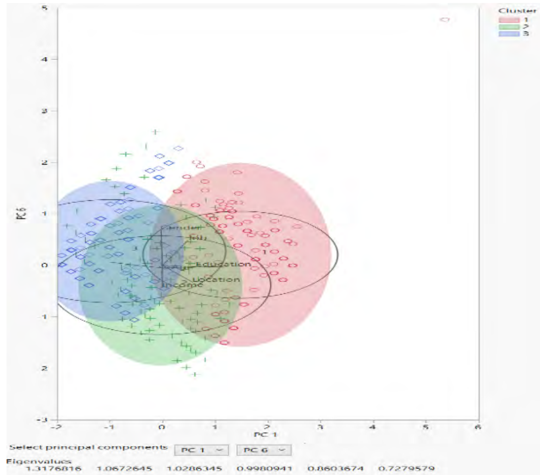


Figure 4-113. Example of Biplot for Demographic Data Clusters SOM

Figure 4-113 presents a visual representation of the demographic data clustering results using SOMs. The biplot displays the points and clusters in the first two principal components, specifically PC1 and PC6. The three clusters of demographic data are illustrated, and there is overlap between the clusters, with some points belonging to two or even all three clusters. This visualization provides a graphical representation of the clustering results and offers insights into the relationships and similarities between the different demographic groups.

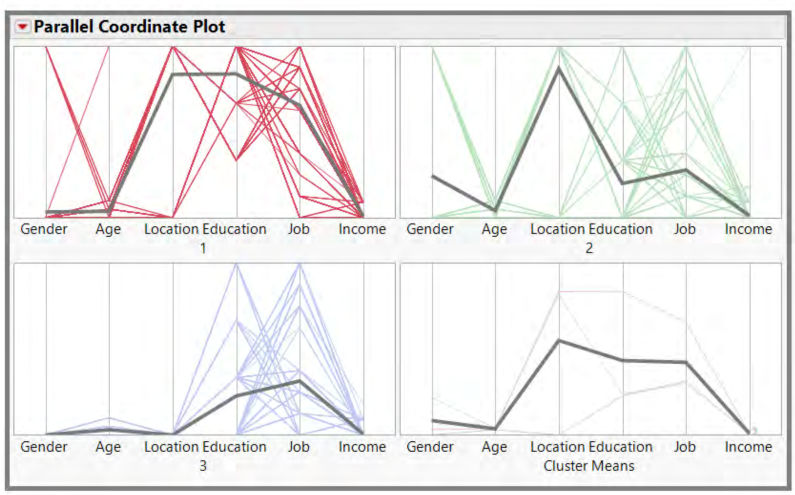


Figure 4-114. Parallel Coordinate Plot for Demographic Data

Figure 4-114 depicts a graphical representation of demographic data separated into clusters. The plot displays connected line segments that correspond to each row in the data table. The figure suggests that the clusters share similar features, which can make it challenging to differentiate between them.

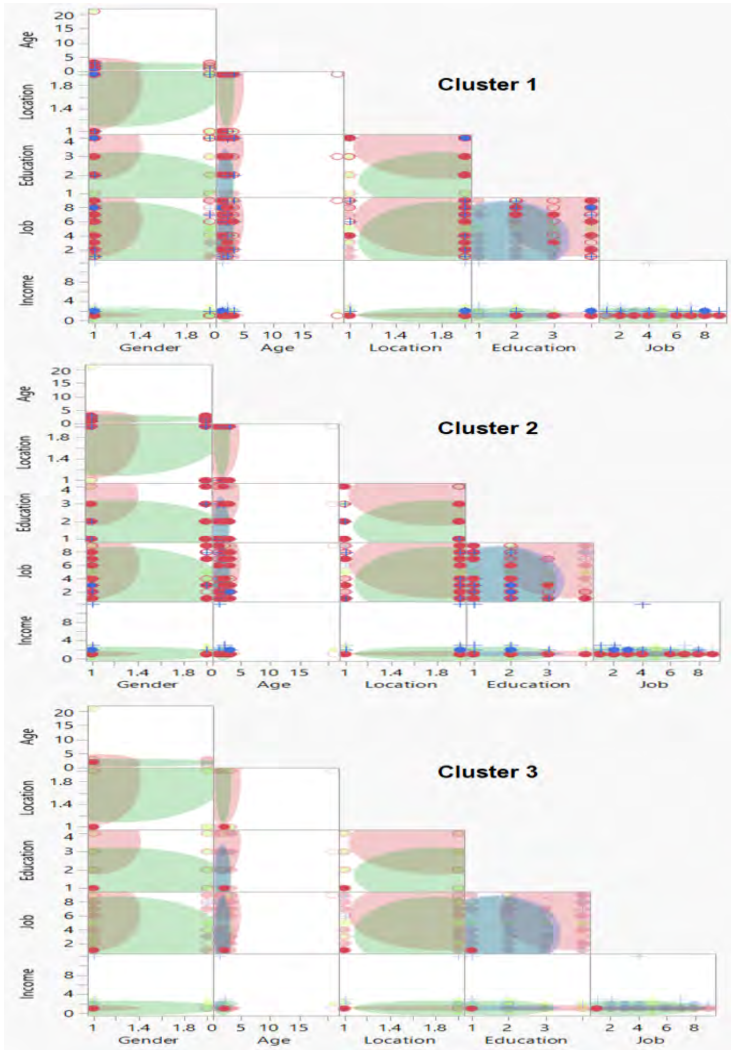
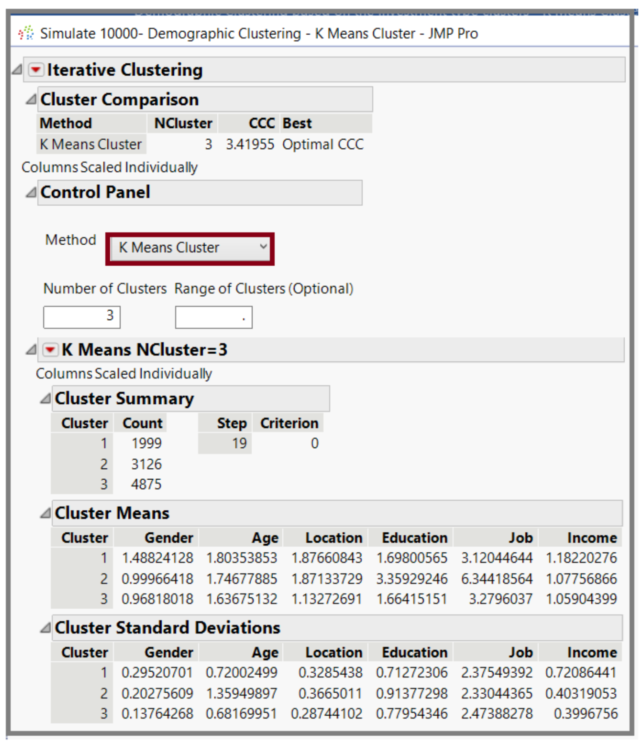


Figure 4-115. Scatterplot matrix for demographic data

However, the use of SOM helps to effectively distinguish between the three clusters. The gray line represents the mean of the data. Figure 4-115, on the other hand, presents a regression line and confidence interval on a scatterplot matrix for the demographic data separated by clusters. The clear shaded region inside the ellipses in the scatterplot matrix represents the area between each Y variable of the demographic data. The matrix also contains ellipses, points, and a lower triangular scatter matrix for the covariates. The ellipses with different overlays indicate the different levels of the categorical variable X. The linear discriminant method used in this matrix is based on the pooled covariance matrix.

The performance of the clustering method applied to a large volume of data was evaluated by simulating demographic data in JMP, with a scale of 10,000 samples. The estimated cluster mixing probabilities, means, and standard deviations for each cluster were used to create a new data table with simulated demographic data. The number of clusters indicated by JMP was determined based on the CCC. Figures 4-116 and 4-117 present the three clusters obtained from the simulated demographic data using the K-Means method. The figures display the iterative clustering process and the cluster summary in 19 steps, including the count of the first cluster (1999), the second cluster (3126), and the third cluster (4875). The mean and standard deviation values for each cluster are also indicated.


Simulate 10000- Demographic Clustering - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	3.41955	Optimal CCC

Columns Scaled Individually

Control Panel

Method: K Means Cluster

Number of Clusters: 3 Range of Clusters (Optional): .

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count
1	1999
2	3126
3	4875

Step	Criterion
19	0

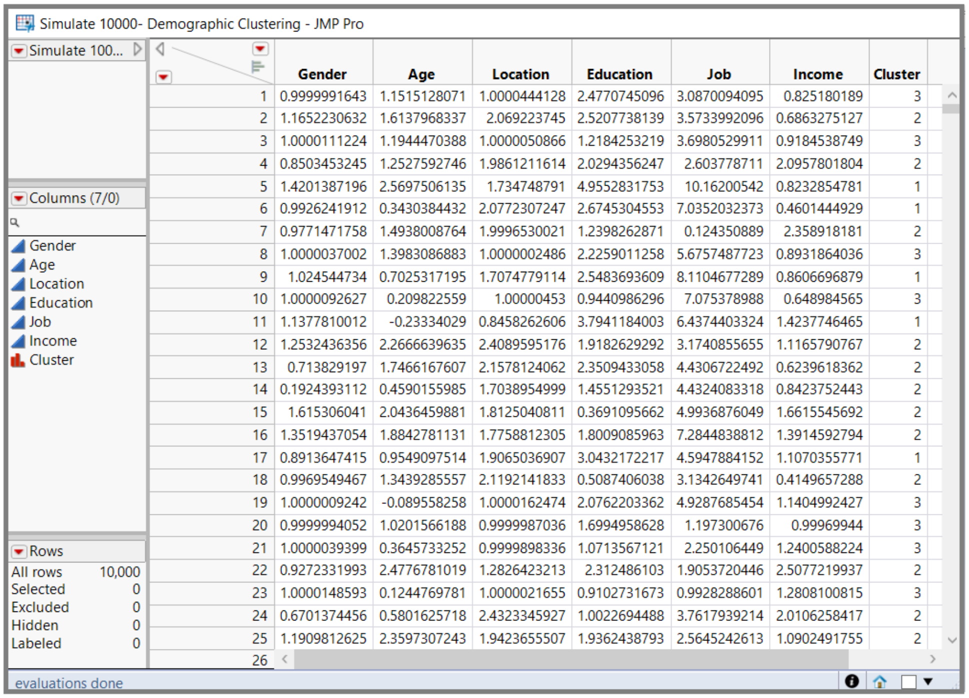
Cluster Means

Cluster	Gender	Age	Location	Education	Job	Income
1	1.48824128	1.80353853	1.87660843	1.69800565	3.12044644	1.18220276
2	0.99966418	1.74677885	1.87133729	3.35929246	6.34418564	1.07756866
3	0.96818018	1.63675132	1.13272691	1.66415151	3.2796037	1.05904399

Cluster Standard Deviations

Cluster	Gender	Age	Location	Education	Job	Income
1	0.29520701	0.72002499	0.3285438	0.71272306	2.37549392	0.72086441
2	0.20275609	1.35949897	0.3665011	0.91377298	2.33044365	0.40319053
3	0.13764268	0.68169951	0.28744102	0.77954346	2.47388278	0.3996756

Figure 4-116. Iterative Clustering of simulated sample demographic data by the K-Means method



Simulate 10000- Demographic Clustering - JMP Pro

	Gender	Age	Location	Education	Job	Income	Cluster
1	0.9999991643	1.1515128071	1.0000444128	2.4770745096	3.0870094095	0.825180189	3
2	1.1652230632	1.6137968337	2.069223745	2.5207738139	3.5733992096	0.6863275127	2
3	1.0000111224	1.1944470388	1.0000050866	1.2184253219	3.6980529911	0.9184538749	3
4	0.8503453245	1.2527592746	1.9861211614	2.0294356247	2.603778711	2.0957801804	2
5	1.4201387196	2.5697506135	1.734748791	4.9552831753	10.16200542	0.8232854781	1
6	0.9926241912	0.3430384432	2.0772307247	2.6745304553	7.0352032373	0.4601444929	1
7	0.9771471758	1.4938008764	1.9996530021	1.2398262871	0.124350889	2.358918181	2
8	1.0000037002	1.3983086883	1.0000002486	2.2259011258	5.6757487723	0.8931864036	3
9	1.024544734	0.7025317195	1.7074779114	2.5483693609	8.1104677289	0.8606696879	1
10	1.0000092627	0.209822559	1.00000453	0.9440986296	7.075378988	0.648984565	3
11	1.1377810012	-0.23334029	0.8458262606	3.7941184003	6.4374403324	1.4237746465	1
12	1.2532436356	2.2666639635	2.4089595176	1.9182629292	3.1740855655	1.1165790767	2
13	0.713829197	1.7466167607	2.1578124062	2.3509433058	4.4306722492	0.6239618362	2
14	0.1924393112	0.4590155985	1.7038954999	1.4551293521	4.4324083318	0.8423752443	2
15	1.615306041	2.0436459881	1.8125040811	0.3691095662	4.9936876049	1.6615545692	2
16	1.3519437054	1.8842781131	1.7758812305	1.8009085963	7.2844838812	1.3914592794	2
17	0.8913647415	0.9549097514	1.9065036907	3.0432172217	4.5947884152	1.1070355771	1
18	0.9969549467	1.3439285557	2.1192141833	0.5087406038	3.1342649741	0.4149657288	2
19	1.0000009242	-0.089558258	1.0000162474	2.0762203362	4.9287685454	1.1404992427	3
20	0.9999994052	1.0201566188	0.9999987036	1.6994958628	1.197300676	0.99969944	3
21	1.0000039399	0.3645733252	0.9999898336	1.0713567121	2.250106449	1.2400588224	3
22	0.9272331993	2.4776781019	1.2826423213	2.312486103	1.9053720446	2.5077219937	2
23	1.0000148593	0.1244769781	1.0000021655	0.9102731673	0.9928288601	1.2808100815	3
24	0.6701374456	0.5801625718	2.4323345927	1.0022694488	3.7617939214	2.0106258417	2
25	1.1909812625	2.3597307243	1.9423655507	1.9362438793	2.5645242613	1.0902491755	2
26							

Figure 4-117. A part of the simulated demographic data table

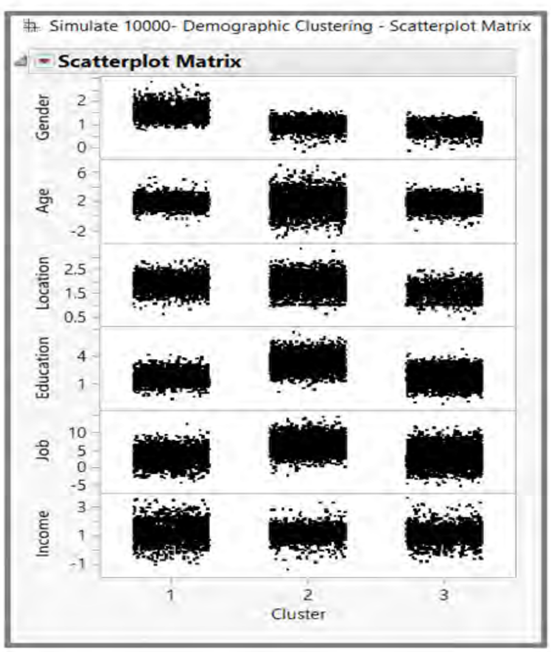


Figure 4-118. A scatterplot matrix based on the clusters of simulated demographic data

Figure 4-118 represents a scatterplot matrix that was generated based on the clusters obtained from the simulated demographic data. The figure displays the confidence ellipses, which are used to quantify the degree of uncertainty surrounding the clusters. This scatterplot matrix displays the results of a simulation run on a scale of 10,000, which aimed to optimize the number of clusters present in the demographic data. The scatterplot matrix and the confidence ellipses together provide a visual representation of the distribution and grouping of the demographic data, based on the results of the simulation.

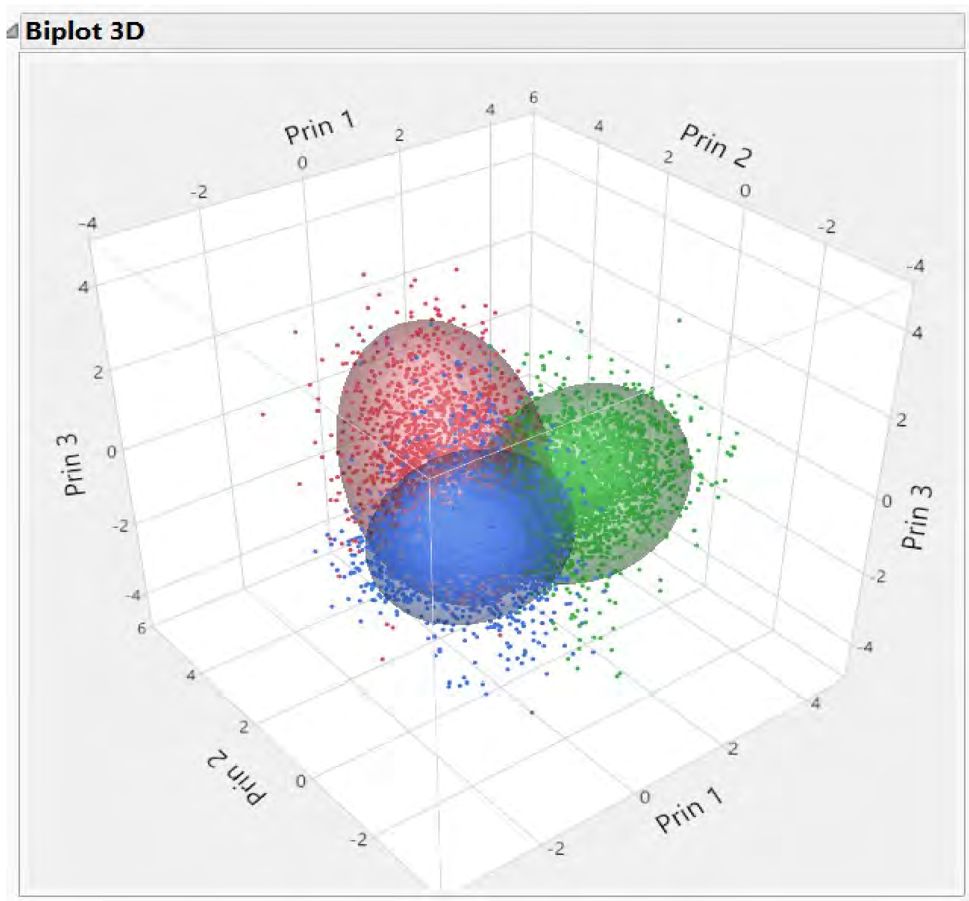
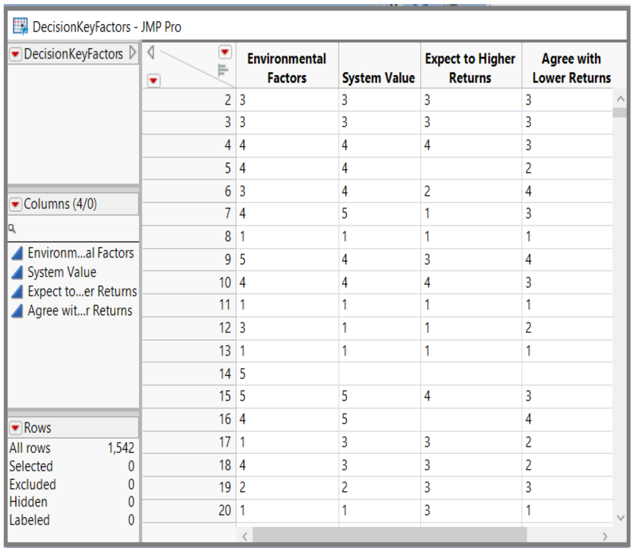


Figure 4-119. First three principal components of demographic data

Figure 4-119 represents the biplot visualization of the first three principal components of the demographic data. It displays the distribution of the data points and the clusters they form in a 3D space. The plot visualizes the relationship between the first three principal components and the demographic data, allowing for a deeper understanding of the underlying patterns and trends in the data.

4.3.2. Clustering Input Key decision factors data

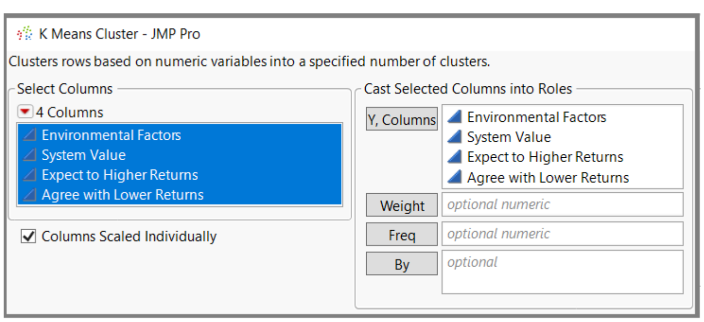
In the process of clustering data, the key decision factors data is first imported into JMP software, with four columns and 1542 rows of information. As demonstrated in Figure 4-120, the imported data is then visually represented in JMP. To further prepare the data, the K-Means and SOM methods are utilized within JMP to cluster the key decision factors data.



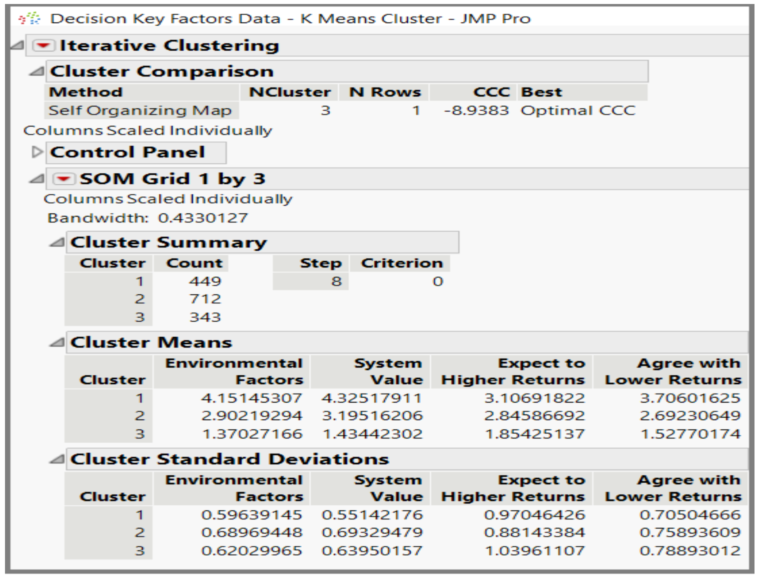
	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns
2	3	3	3	3
3	3	3	3	3
4	4	4	4	3
5	4	4		2
6	3	4	2	4
7	4	5	1	3
8	1	1	1	1
9	5	4	3	4
10	4	4	4	3
11	1	1	1	1
12	3	1	1	2
13	1	1	1	1
14	5			
15	5	5	4	3
16	4	5		4
17	1	3	3	2
18	4	3	3	2
19	2	2	3	3
20	1	1	3	1

Figure 4-120. A part of imported key decision factors data in JMP

The next step in the process involves grouping the rows of data based on the numerical variables into a specified number of clusters. This is shown in Figure 4-121, which displays the results of the clustering process. The columns in the figure represent information regarding the level of agreement among respondents with regards to the key affective factors that impact their investment decisions.



Figures 4-121. Y columns to cluster key decision factors data



Decision Key Factors Data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	N Rows	CCC	Best
Self Organizing Map	3	1	-8.9383	Optimal CCC

Columns Scaled Individually

Control Panel

SOM Grid 1 by 3

Columns Scaled Individually

Bandwidth: 0.4330127

Cluster Summary

Cluster	Count	Step	Criterion
1	449	8	0
2	712		
3	343		

Cluster Means

Cluster	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns
1	4.15145307	4.32517911	3.10691822	3.70601625
2	2.90219294	3.19516206	2.84586692	2.69230649
3	1.37027166	1.43442302	1.85425137	1.52770174

Cluster Standard Deviations

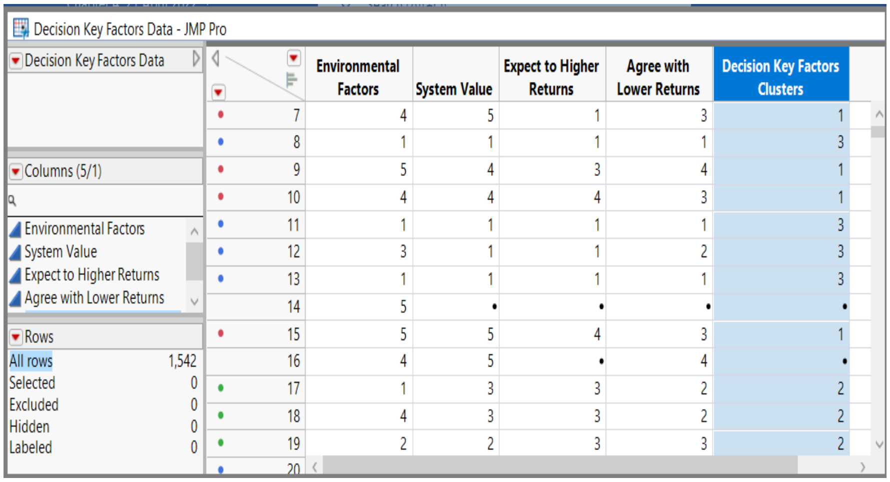
Cluster	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns
1	0.59639145	0.55142176	0.97046426	0.70504666
2	0.68969448	0.69329479	0.88143384	0.75893609
3	0.62029965	0.63950157	1.03961107	0.78893012

Figure 4-122. Iterative Clustering of key decision factors data by K Means & SOM

Figure 4-122 depicts the result of the iterative clustering of key decision factor data using a combination of K-Means and SOM methods. The number of clusters was determined based on the CCC criteria and three clusters were indicated. The figure displays the cluster summary in 8 steps, including the count for each cluster, as well as the mean and standard deviation for each cluster. The script used for the clustering process is written in Python and includes parameters such as the number of clusters (3), the number of rows in the SOM (1), and the bandwidth of the SOM (0.433012701892219). The script also includes instructions for visualizing the results using various plots, including biplot and parallel coordinate plots. The following is the script (supports Python) for the clustering:

```
K Means Cluster(
	Y(
		:Environmental Factors, :System Value, :Expect to Higher Returns,
		:Agree with Lower Returns
	),
	{SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
	Number of Clusters( 3 ), SOM, Go(
		Show Biplot Rays( [0, 0, 1] ),
		Parallel Coord Plots,
		Biplot( 1 ),
		Biplot 3D( 1 )
	)},
	SendToReport(
		Dispatch( {}, "Control Panel", OutlineBox, {Close( 1 )} ),
		Dispatch(
			{"SOM Grid 1 by 3"},
			"Cluster Standard Deviations",
```

```
                    OutlineBox,
                    {Close( 0 )}
            )
        )
)
```



Decision Key Factors Data - JMP Pro

Decision Key Factors Data

Columns (5/1)

Environmental Factors
System Value
Expect to Higher Returns
Agree with Lower Returns

Rows

Rows	
All rows	1,542
Selected	0
Excluded	0
Hidden	0
Labeled	0

	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns	Decision Key Factors Clusters
7	4	5	1	3	1
8	1	1	1	1	3
9	5	4	3	4	1
10	4	4	4	3	1
11	1	1	1	1	3
12	3	1	1	2	3
13	1	1	1	1	3
14	5	•	•	•	•
15	5	5	4	3	1
16	4	5	•	4	•
17	1	3	3	2	2
18	4	3	3	2	2
19	2	2	3	3	2
20					

Figure 4-123. A part of clusters for each row of key decision factors data

Figure 4-123 represents a portion of the clustering process performed on the key decision factor data. The figure illustrates the addition of a new column in the data table, which displays the cluster assigned to each row. The clustering process involves grouping the rows of the data table based on numerical variables into a defined number of clusters. In other words, the figure shows how each row of the key decision factor data has been assigned to a cluster based on the values of the numerical variables.

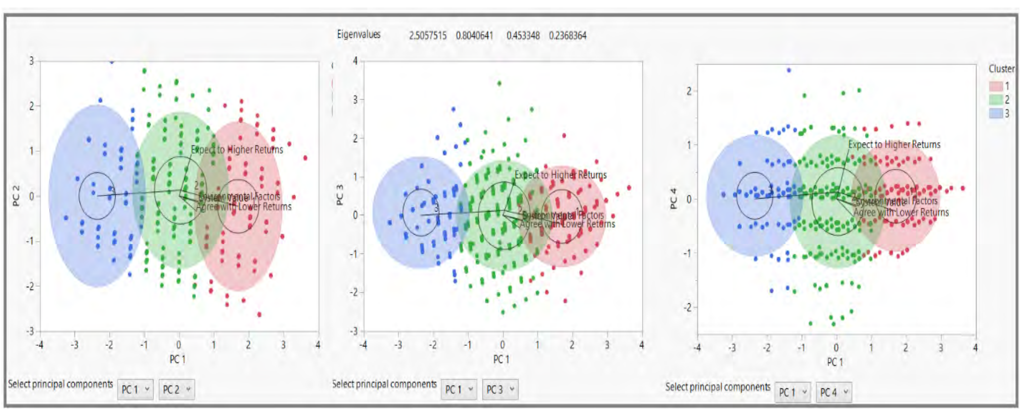


Figure 4-124. Biplots for Key decision factors Data Clusters SOM

Figure 4-124 displays biplots of the key decision factor data clusters obtained from the SOM. The biplots present the distribution of the data points in the first two principal components of the data. There are three biplots in total, each showing the relationship between different pairs of principal components such as "PC1 and PC2", "PC1 and PC3", and "PC1 and PC4". This figure provides an illustration of all three clusters of the key decision factor data and demonstrates how some of the clusters overlap with each other.

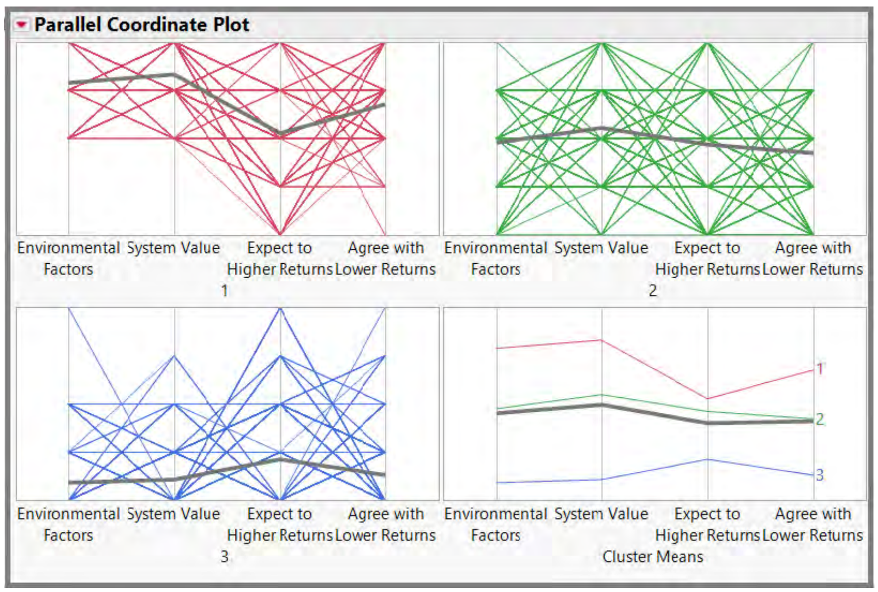


Figure 4-125. Parallel Coordinate Plot for Key decision factors Data

Figure 4-125 represents a graphical depiction of the key decision factor data for each cluster. The line segments connecting the data points in each cluster suggest that the features of the clusters are like each other, making it challenging to differentiate between them. However, the use of SOM effectively highlights the distinctions between the three clusters. The gray line in the figure represents the mean value of the data. Figure 4-126, on the other hand, is a scatterplot matrix that represents the regression line and confidence interval for the key decision factor data and each cluster. The clear region inside the ellipses on the scatterplot matrix highlights the relationship between the Y variable of the key decision factor data and the categorical variable X. The scatterplot matrix includes ellipses, points, and a lower triangular scatter matrix for the covariates, with different ellipses representing different levels of the categorical variable X. The linear discriminant method in the matrix is based on the pooled covariance matrix.

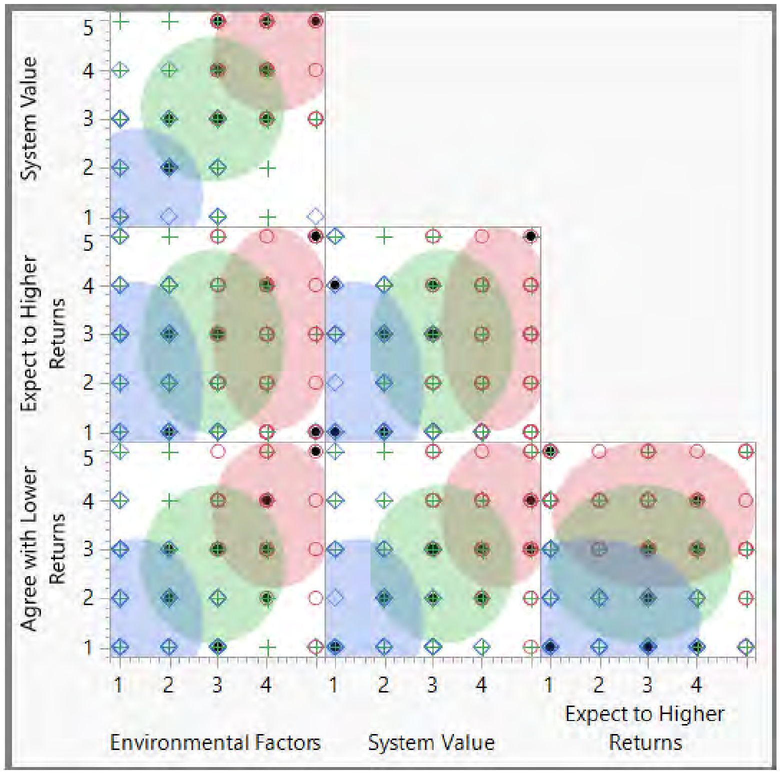
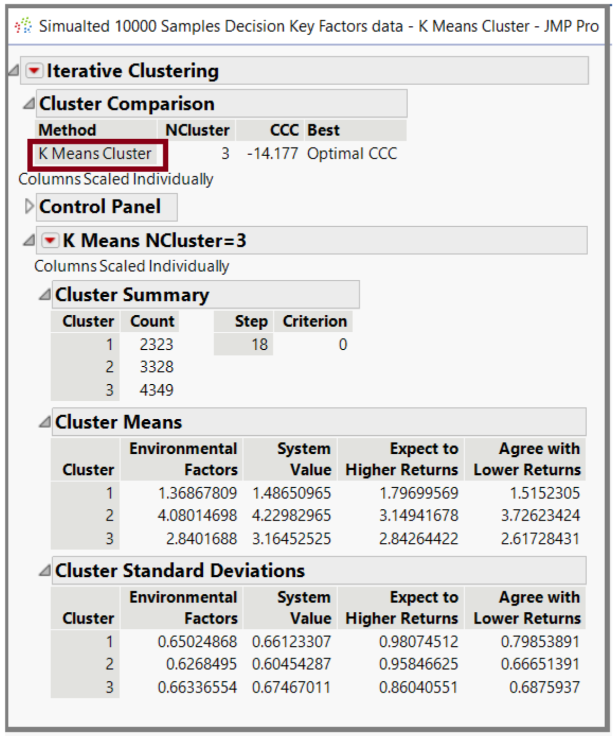


Figure 4-126. Scatterplot matrix for key decision factors data

The performance of the clustering method used in the system was evaluated using a large volume of simulated key decision factor data. The data was generated in JMP software with a sample size of 10,000. A new data table was created based on the estimated cluster mixing probabilities, means, and standard deviations for each cluster. The number of clusters was determined by the CCC criteria in JMP and was found to be three.

Figures 4-127 and 4-128 illustrate the results of the K-Means method used for clustering the simulated key decision factor data. The figures show the iterative process of clustering and a summary of the clusters after 18 steps. The summary includes the count of each cluster with the first cluster having 2323 samples, the second cluster having 3328 samples, and the third cluster having 4349 samples. Additionally, the means and standard deviations are provided for each cluster.



Simualted 10000 Samples Decision Key Factors data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	-14.177	Optimal CCC

Columns Scaled Individually

Control Panel

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count
1	2323
2	3328
3	4349

Step	Criterion
18	0

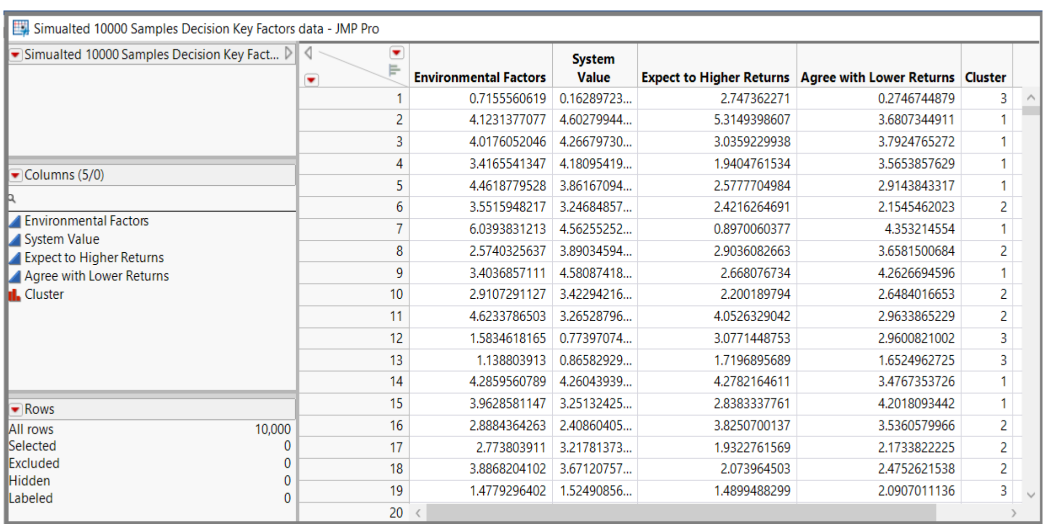
Cluster Means

Cluster	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns
1	1.36867809	1.48650965	1.79699569	1.5152305
2	4.08014698	4.22982965	3.14941678	3.72623424
3	2.8401688	3.16452525	2.84264422	2.61728431

Cluster Standard Deviations

Cluster	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns
1	0.65024868	0.66123307	0.98074512	0.79853891
2	0.6268495	0.60454287	0.95846625	0.66651391
3	0.66336554	0.67467011	0.86040551	0.6875937

Figure 4-127. Iterative Clustering of simulated sample key decision factors data by K-Means method



Simualted 10000 Samples Decision Key Factors data - JMP Pro

Simualted 10000 Samples Decision Key Fact...

Columns (5/0)

- Environmental Factors
- System Value
- Expect to Higher Returns
- Agree with Lower Returns
- Cluster

Rows

All rows	10,000
Selected	0
Excluded	0
Hidden	0
Labeled	0

	Environmental Factors	System Value	Expect to Higher Returns	Agree with Lower Returns	Cluster
1	0.7155560619	0.16289723...	2.747362271	0.2746744879	3
2	4.1231377077	4.60279944...	5.3149398607	3.6807344911	1
3	4.0176052046	4.26679730...	3.0359229938	3.7924765272	1
4	3.4165541347	4.18095419...	1.9404761534	3.5653857629	1
5	4.4618779528	3.86167094...	2.5777704984	2.9143843317	1
6	3.5515948217	3.24684857...	2.4216264691	2.1545462023	2
7	6.0393831213	4.56255252...	0.8970060377	4.353214554	1
8	2.5740325637	3.89034594...	2.9036082663	3.6581500684	2
9	3.4036857111	4.58087418...	2.668076734	4.2626694596	1
10	2.9107291127	3.42294216...	2.200189794	2.6484016653	2
11	4.6233786503	3.26528796...	4.0526329042	2.9633865229	2
12	1.5834618165	0.77397074...	3.0771448753	2.9600821002	3
13	1.138803913	0.86582929...	1.7196895689	1.6524962725	3
14	4.2859560789	4.26043939...	4.2782164611	3.4767353726	1
15	3.9628581147	3.25132425...	2.8383337761	4.2018093442	1
16	2.8884364263	2.40860405...	3.8250700137	3.5360579966	2
17	2.773803911	3.21781373...	1.9322761569	2.1733822225	2
18	3.8868204102	3.67120757...	2.073964503	2.4752621538	2
19	1.4779296402	1.52490856...	1.4899488299	2.0907011136	3
20					

Figure 4-128. A part of the simulated key decision factors data table

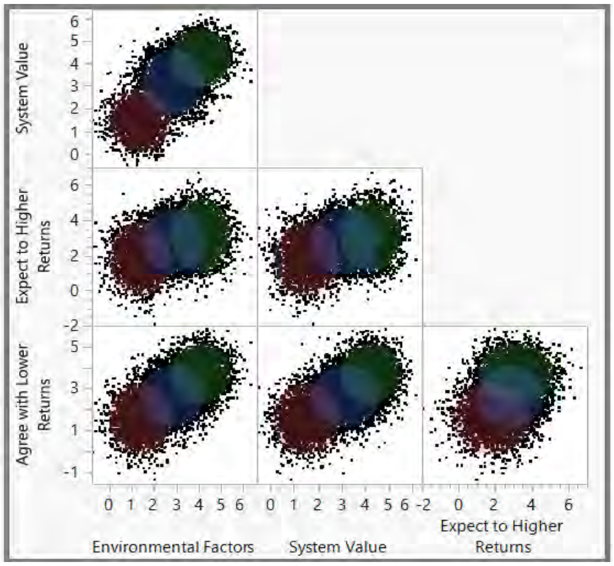


Figure 4-129. A scatterplot matrix based on the clusters of simulated key decision factors data

Figure 4-129 is a visual representation of the clustering of key decision factors data. The figure presents a scatterplot matrix, with each panel showing the relationship between two decision factors. Confidence ellipses are added to the scatterplot to provide a visual representation of the uncertainty of the data. The scatterplot matrix is based on simulated data, meaning that the data was generated by running simulations on a 10,000 scale. The purpose of this figure is to show how the optimization of the clusters, or grouping of data points, can be achieved through simulations. This figure can help to provide insight into how the key decision factors are related to each other, and how they are affected by the optimization process.

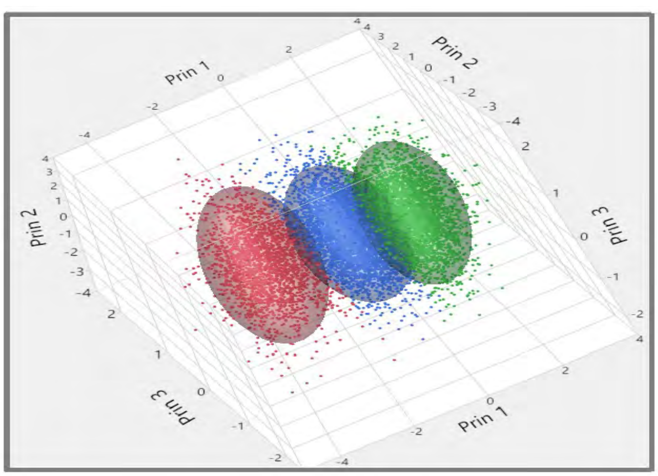
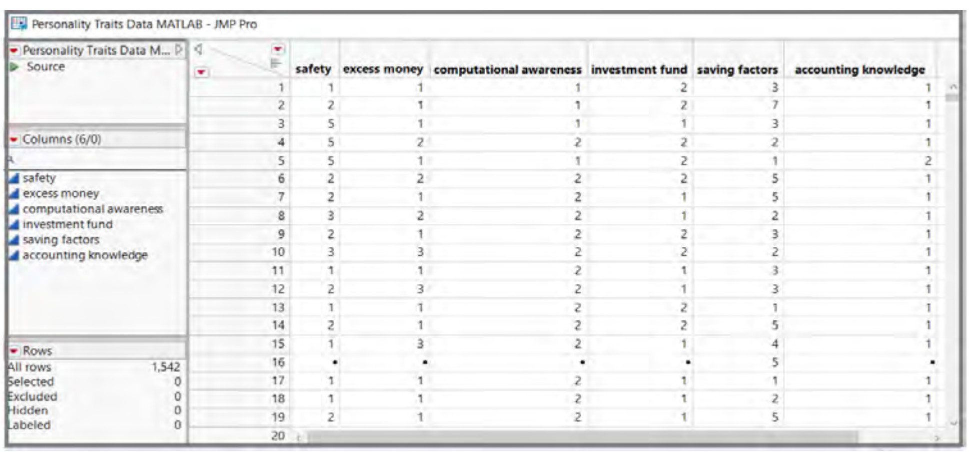


Figure 4-130. The first three principal components of the key decision factor data

Figure 4-130 represents a biplot in three dimensions, showing the distribution of points and clusters in the first three principal components of the key decision factors data. A biplot is a graphical representation of two or more variables in which points are plotted to reflect their relationships with each other. The use of the first three principal components as axes in this biplot suggests that they capture most of the variance in the key decision factors data and that they are important in understanding the distribution of the points and clusters. The biplot allows us to visualize the clustering patterns of the data and the relationships between the data points and the principal components, providing valuable insights into the structure of the data.

4.3.3. Clustering Input Personality Traits Data

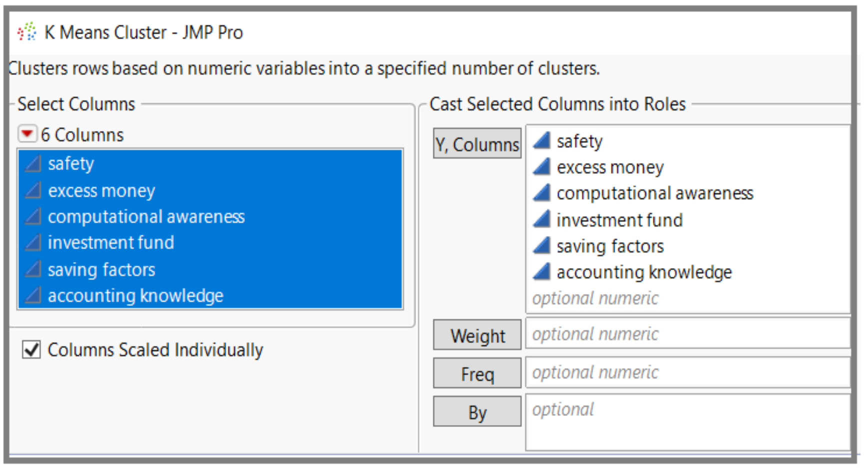
In the first step of data clustering, the personality traits data was imported into JMP software, which consisted of 1542 rows and four columns. Figure 4-131 represents the visualization of the imported personality traits data in JMP. Afterwards, the data was pre-processed using two techniques, K-Means and SOM, within JMP to prepare it for the clustering process. The purpose of this preparation was to group similar data points together, making it easier to analyze and interpret the results.



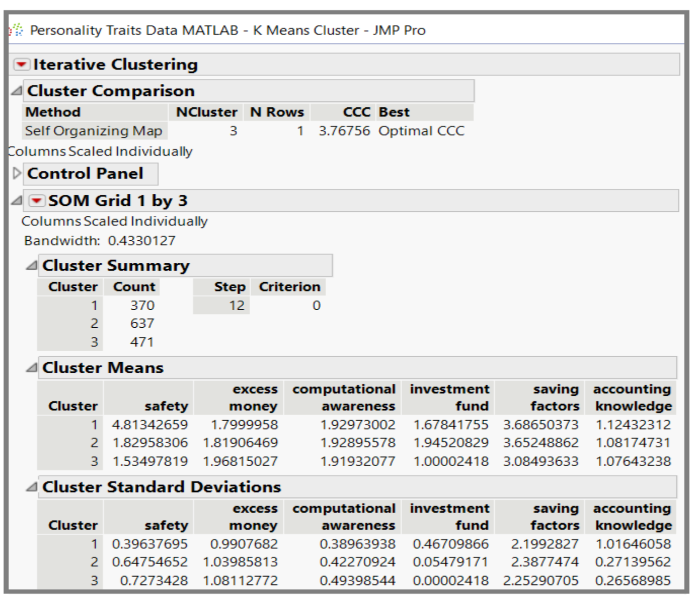
	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge
1	1	1	1	2	3	1
2	2	1	1	2	7	1
3	5	1	1	1	3	1
4	5	2	2	2	2	1
5	5	1	1	2	1	2
6	2	2	2	2	5	1
7	2	1	2	1	5	1
8	3	2	2	1	2	1
9	2	1	2	2	3	1
10	3	3	2	2	2	1
11	1	1	2	1	3	1
12	2	3	2	1	3	1
13	1	1	2	2	1	1
14	2	1	2	2	5	1
15	1	3	2	1	4	1
16	•	•	•	•	5	•
17	1	1	2	1	1	1
18	1	1	2	1	2	1
19	2	1	2	1	5	1
20						

Figure 4-131. A part of imported personality traits data in JMP

The cluster rows in the data represent the grouping of numeric variables into a defined number of clusters. In this instance, the columns represent the personality traits of respondents in various subjects, including safety, excess money, computational awareness, investment fund, saving factors, and accounting knowledge, as shown in Figure 4-132. In other words, the data is divided into clusters based on the values of the numeric variables. The columns of the data contain information about the personality traits of respondents in different areas related to finance and investment. This information is depicted in this figure.



Figures 4-132. Y columns to cluster personality traits data



Personality Traits Data MATLAB - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	N Rows	CCC	Best
Self Organizing Map	3	1	3.76756	Optimal CCC

Columns Scaled Individually

Control Panel

SOM Grid 1 by 3

Columns Scaled Individually

Bandwidth: 0.4330127

Cluster Summary

Cluster	Count
1	370
2	637
3	471

Step	Criterion
12	0

Cluster Means

Cluster	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge
1	4.81342659	1.7999958	1.92973002	1.67841755	3.68650373	1.12432312
2	1.82958306	1.81906469	1.92895578	1.94520829	3.65248862	1.08174731
3	1.53497819	1.96815027	1.91932077	1.00002418	3.08493633	1.07643238

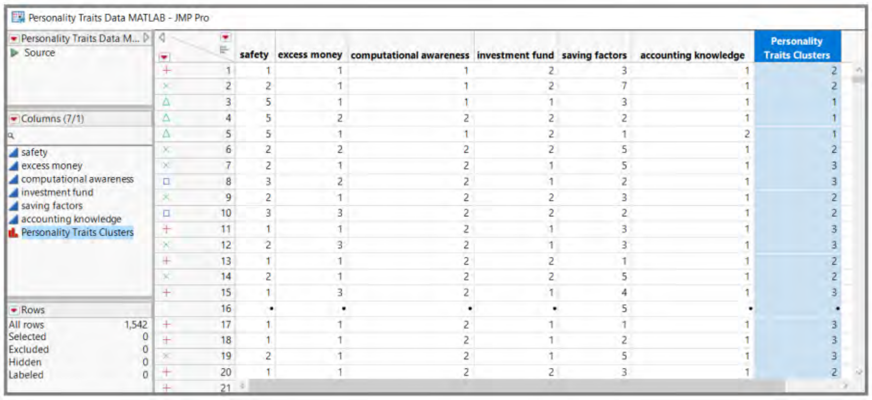
Cluster Standard Deviations

Cluster	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge
1	0.39637695	0.9907682	0.38963938	0.46709866	2.1992827	1.01646058
2	0.64754652	1.03985813	0.42270924	0.05479171	2.3877474	0.27139562
3	0.7273428	1.08112772	0.49398544	0.00002418	2.25290705	0.26568985

Figure 4-133. Iterative Clustering of personality traits data by K Means & SOM

The number of clusters shown in Figure 4-133 is determined by the CCC (Cluster Characteristic Criteria) method. The figure represents the result of clustering personality traits using a combined method of K-Means and SOM. The clustering process is performed iteratively in 12 steps, resulting in three clusters with counts of 370, 637, and 471 individuals in each cluster respectively. Additionally, the figure provides the mean and standard deviation values for each cluster. The script for the clustering process, written in Python, performs the K-Means clustering on the following personality traits: safety, excess money, computational awareness, investment fund, saving factors, and accounting knowledge. The script includes parameters for the SOM algorithm, such as the number of rows in the SOM grid, bandwidth, and the number of clusters. The output is then sent to a report, which displays the cluster standard deviations in an outline box. The script (supports Python) for the clustering is in the following:

```
K Means Cluster(
	Y(
		:safety, :excess money, :computational awareness, :investment fund,
		:saving factors, :accounting knowledge
	),
	{SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
	Number of Clusters( 3 ), SOM, Go},
	SendToReport(
		Dispatch( {}, "Control Panel", OutlineBox, {Close( 1 )} ),
		Dispatch(
			{"SOM Grid 1 by 3"},
			"Cluster Standard Deviations",
			OutlineBox,
			{Close( 0 )}
		)
	)
)
```



	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge	Personality Traits Clusters
1	1	1	1	2	3	1	2
2	2	1	1	2	7	1	2
3	5	1	1	1	3	1	1
4	5	2	2	2	2	1	1
5	5	1	1	2	1	2	1
6	2	2	2	2	5	1	2
7	2	1	2	1	5	1	3
8	3	2	2	1	2	1	3
9	2	1	2	2	3	1	2
10	3	3	2	2	2	1	2
11	1	1	2	1	3	1	3
12	2	3	2	1	3	1	3
13	1	1	2	2	1	1	2
14	2	1	2	2	5	1	2
15	1	3	2	1	4	1	3
16	•	•	•	•	5	•	•
17	1	1	2	1	1	1	3
18	1	1	2	1	2	1	3
19	2	1	2	1	5	1	3
20	1	1	2	2	3	1	2
21							

Figure 4-134. A part of clusters for each row of personality traits data

Figure 4-134 illustrates the allocation of clusters for each row of personality trait data. It depicts a portion of the data table with an additional column added, which represents the assigned cluster for each row of data. The cluster assignment is based on the grouping of numeric variables into a specified number of clusters. Essentially, the figure demonstrates how the personality trait data has been divided into several clusters, each having a unique cluster number assigned to it.

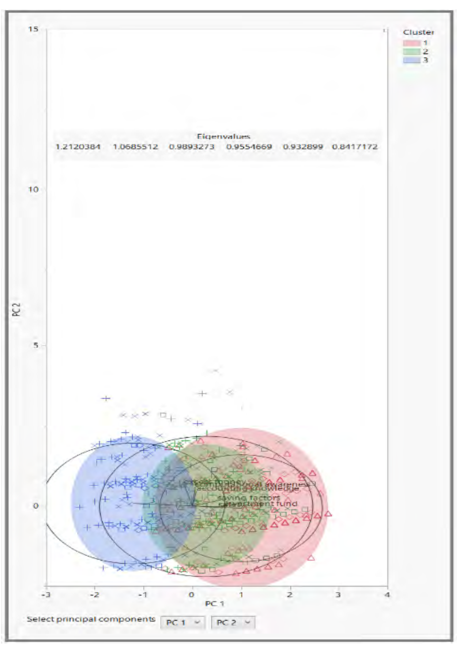


Figure 4-135. Example Biplot for Personality traits Data Clusters SOM

Figure 4-135 presents a biplot representation of the personality traits data in the first two principal components, labeled as "PC1" and "PC2". The figure displays the distribution of the three clusters in the data, with some overlap between the two clusters visible in certain areas. Additionally, there is a section where all three clusters overlap completely. A biplot is a type of data visualization that displays the relationships between the points in a dataset and the principal components, which are new variables constructed from the original data that explain most of its variance. In this case, the biplot shows the distribution of the personality traits data in the first two principal components, "PC1" and "PC2". The three clusters represent different groupings of the data points based on their similarity in terms of the personality traits. The overlapping of the two and three clusters suggests that there is some degree of overlap in the traits represented by each cluster.

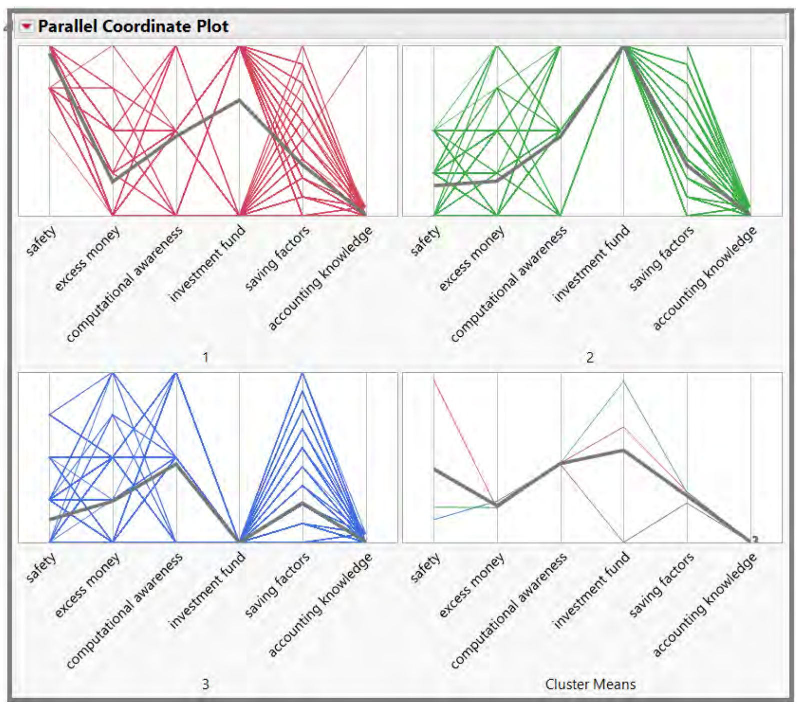


Figure 4-136. Parallel Coordinate Plot for Personality traits Data

Figure 4-136 displays the plot for each cluster of personality traits data, where the connected line segments represent each row of the data table. This figure highlights the similarities in features among the clusters, making it challenging to distinguish between them. However, the application of SOM effectively resolves this issue, as the distinction between the three clusters is depicted clearly. The gray line in the figure represents the mean of the data. Figure 4-137 presents the regression line and confidence interval on the scatterplot matrix, where the data is separated into each cluster of the personality traits. The clear shaded region inside the ellipses on the scatterplot matrix represents the relationship between each Y variable of the personality traits data. This matrix contains ellipses, points, and a lower triangular scatter matrix for the covariates, where ellipses with different overlays are shown for each level of the categorical variable X. The linear discriminant method used in this matrix is based on the pooled covariance matrix.

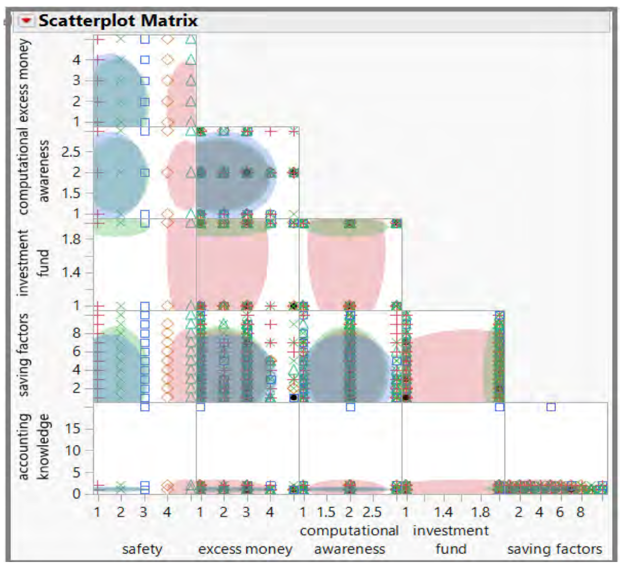
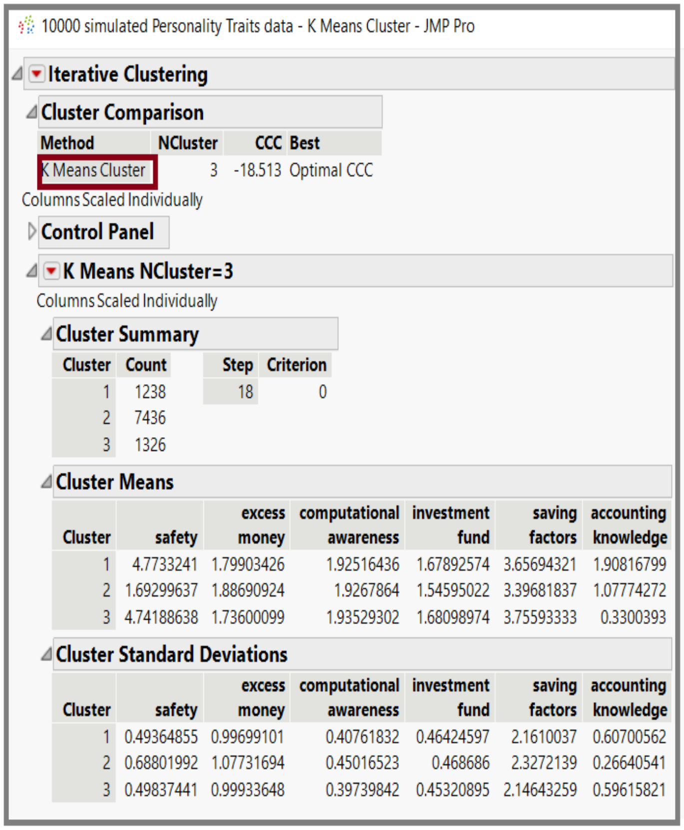


Figure 4-137. Scatterplot matrix for personality traits data

The used clustering method was evaluated with a large volume of simulated personality trait data in JMP, with a sample size of 10,000. A new data table was created using the estimated mixing probabilities, means, and standard deviations for each cluster. The number of clusters indicated by JMP was determined using the CCC method and resulted in three clusters. Figures 4-138 and 4-139 present the results of the K-Means method applied to the simulated personality trait data, showing an iterative clustering process in 18 steps. The count of data points in each cluster is displayed, with 1238 in the first cluster, 7436 in the second cluster, and

1326 in the third cluster. The means and standard deviations for each cluster are also indicated.



10000 simulated Personality Traits data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	-18.513	Optimal CCC

Columns Scaled Individually

Control Panel

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count	Step	Criterion
1	1238	18	0
2	7436		
3	1326		

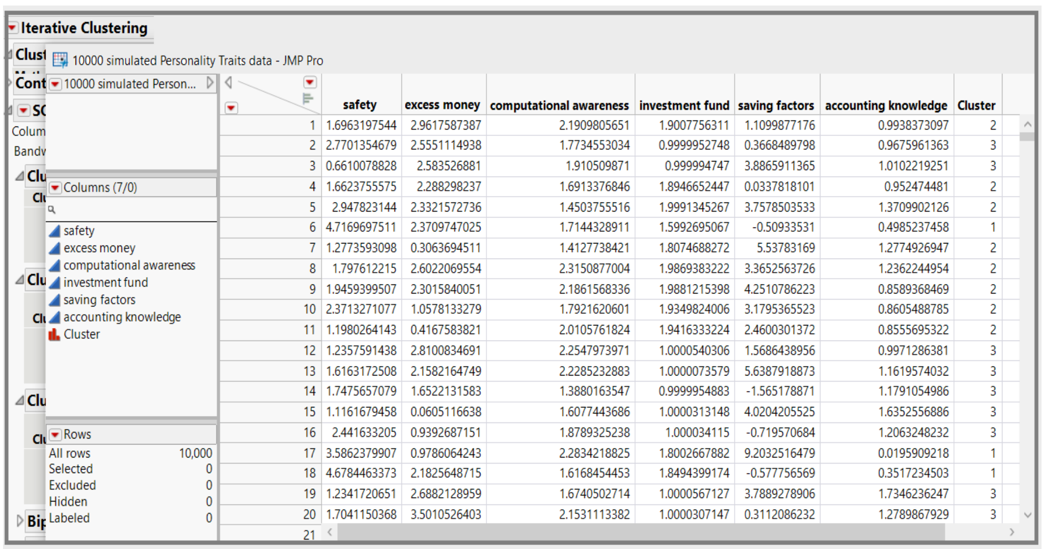
Cluster Means

Cluster	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge
1	4.7733241	1.79903426	1.92516436	1.67892574	3.65694321	1.90816799
2	1.69299637	1.88690924	1.9267864	1.54595022	3.39681837	1.07774272
3	4.74188638	1.73600099	1.93529302	1.68098974	3.75593333	0.3300393

Cluster Standard Deviations

Cluster	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge
1	0.49364855	0.99699101	0.40761832	0.46424597	2.1610037	0.60700562
2	0.68801992	1.07731694	0.45016523	0.468686	2.3272139	0.26640541
3	0.49837441	0.99933648	0.39739842	0.45320895	2.14643259	0.59615821

Figure 4-138. Iterative Clustering of simulated sample personality traits data by the K-Means method



10000 simulated Personality Traits data - JMP Pro

Columns (7/0): safety, excess money, computational awareness, investment fund, saving factors, accounting knowledge, Cluster

Rows: All rows 10,000; Selected 0; Excluded 0; Hidden 0; Labeled 0

	safety	excess money	computational awareness	investment fund	saving factors	accounting knowledge	Cluster
1	1.6963197544	2.9617587387	2.1909805651	1.9007756311	1.1099877176	0.9938373097	2
2	2.7701354679	2.5551114938	1.7734553034	0.9999952748	0.3668489798	0.9675961363	3
3	0.6610078828	2.583526881	1.910509871	0.999994747	3.8865911365	1.0102219251	3
4	1.6623755575	2.288298237	1.6913376846	1.8946652447	0.0337818101	0.952474481	2
5	2.947823144	2.3321572736	1.4503755516	1.9991345267	3.7578503533	1.3709902126	2
6	4.7169697511	2.3709747025	1.7144328911	1.5992695067	-0.50933531	0.4985237458	1
7	1.2773593098	0.3063694511	1.4127738421	1.8074688272	5.53783169	1.2774926947	2
8	1.797612215	2.6022069554	2.3150877004	1.9869383222	3.3652563726	1.2362244954	2
9	1.9459399507	2.3015840051	2.1861568336	1.9881215398	4.2510786223	0.8589368469	2
10	2.3713271077	1.0578133279	1.7921620601	1.9349824006	3.1795365523	0.8605488785	2
11	1.1980264143	0.4167583821	2.0105761824	1.9416333224	2.4600301372	0.8555695322	2
12	1.2357591438	2.8100834691	2.2547973971	1.0000540306	1.5686438956	0.9971286381	3
13	1.6163172508	2.1582164749	2.2285232883	1.0000073579	5.6387918873	1.1619574032	3
14	1.7475657079	1.6522131583	1.3880163547	0.9999954883	-1.565178871	1.1791054986	3
15	1.1161679458	0.0605116638	1.6077443686	1.0000313148	4.0204205525	1.6352556886	3
16	2.441633205	0.9392687151	1.8789325238	1.000034115	-0.719570684	1.2063248232	3
17	3.5862379907	0.9786064243	2.2834218825	1.8002667882	9.2032516479	0.0195909218	1
18	4.6784463373	2.1825648715	1.6168454453	1.8494399174	-0.577756569	0.3517234503	1
19	1.2341720651	2.6882128959	1.6740502714	1.0000567127	3.7889278906	1.7346236247	3
20	1.7041150368	3.5010526403	2.1531113382	1.0000307147	0.3112086232	1.2789867929	3
21							

Figure 4-139. A part of the simulated personality traits data table

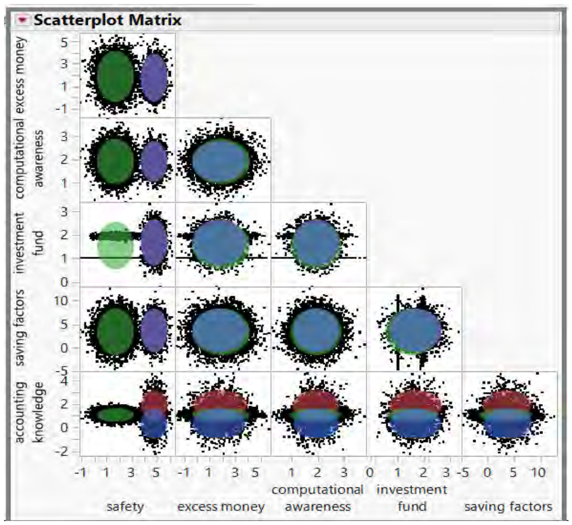


Figure 4-140. A scatterplot matrix based on the clusters of simulated personality traits data

Figure 4-140 represents a scatterplot matrix that displays the distribution of simulated personality traits data. The matrix is accompanied by confidence ellipses that provide a visual representation of the clusters present in the data. The figure is created based on the current number of clusters in the data, which has been optimized through simulations on a scale of 10,000. This figure highlights the optimization process of the clusters, providing insights into how the data is grouped and how the confidence ellipses change in size and shape based on the number of clusters. The figure serves as a valuable tool for visualizing the results of the simulations and understanding the distribution of the personality traits data.

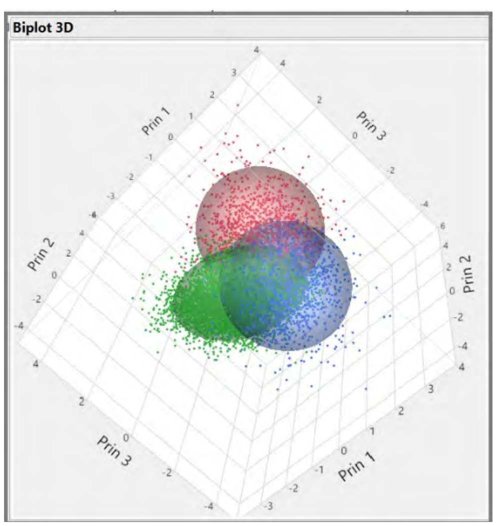
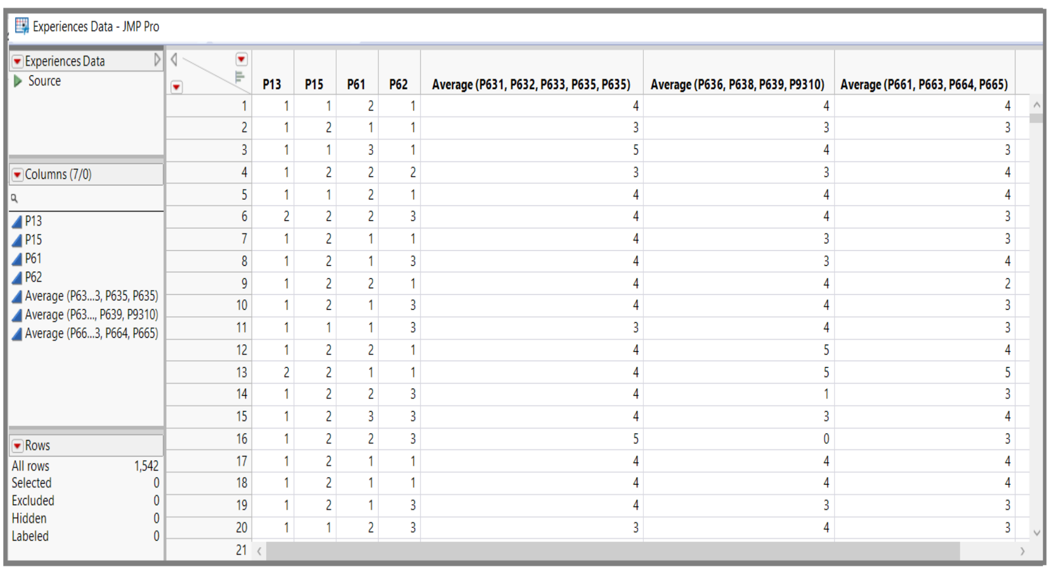


Figure 4-141. First three principal components of personality traits data

Figure 4-141 presents a 3D biplot that displays the points and clusters in the data of the first three principal components of personality traits. The biplot provides a graphical representation of the relationships between the points and the principal components, helping to visualize the distribution of the data in a 3D space. The clusters in the biplot indicate groups of similar points, allowing for a better understanding of the patterns and trends in the personality traits data.

4.3.4. Clustering Input Experiences Data

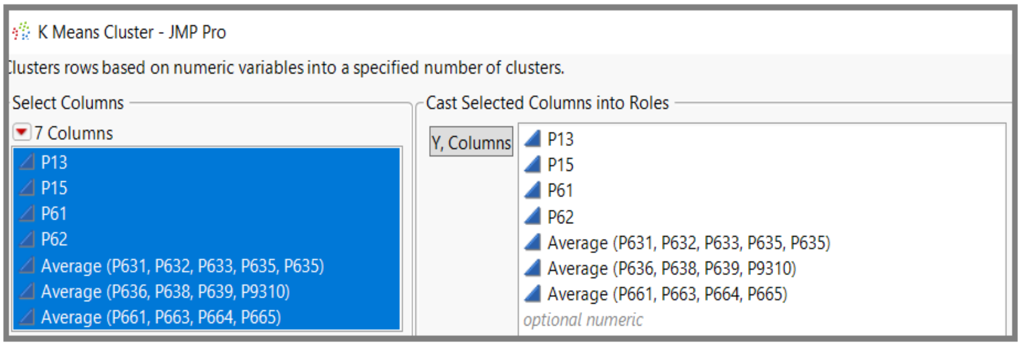
The process of creating a combined IRS using ANFIS involved several steps. In the first step, the data used for clustering was imported into JMP software in four columns with 1542 rows. Figure 4-142 visualizes the imported data in JMP and demonstrates that the experiences data was successfully imported. After importing the data, it was prepared for further analysis using two methods, K-Means and SOM, within JMP. These methods were used to cluster the experiences data and aid in the development of the ANFIS-based IRS.



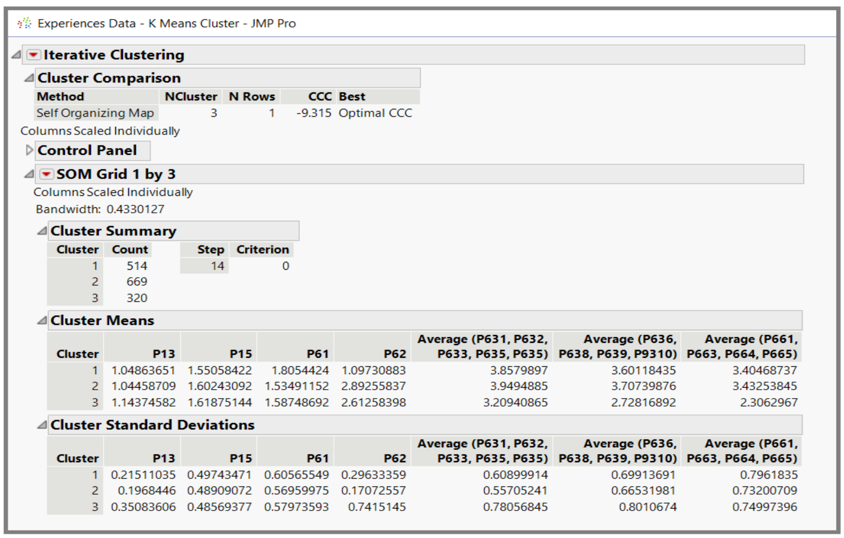
	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)
1	1	1	2	1	4	4	4
2	1	2	1	1	3	3	3
3	1	1	3	1	5	4	3
4	1	2	2	2	3	3	4
5	1	1	2	1	4	4	4
6	2	2	2	3	4	4	3
7	1	2	1	1	4	3	3
8	1	2	1	3	4	3	4
9	1	2	2	1	4	4	2
10	1	2	1	3	4	4	3
11	1	1	1	3	3	4	3
12	1	2	2	1	4	5	4
13	2	2	1	1	4	5	5
14	1	2	2	3	4	1	3
15	1	2	3	3	4	3	4
16	1	2	2	3	5	0	3
17	1	2	1	1	4	4	4
18	1	2	1	1	4	4	4
19	1	2	1	3	4	3	3
20	1	1	2	3	3	4	3
21							

Figure 4-142. A part of imported experiences data in JMP

In the next step, the rows are grouped into clusters based on their numerical values. The number of clusters is determined beforehand. The columns in this process contain information about the experiences of survey participants on various subjects. This is depicted in Figure 4-143. Essentially, the data collected from the survey respondents is divided into a set number of clusters based on their numerical values. These clusters provide insights into the experiences of the participants on different subjects, which are depicted in this figure.



Figures 4-143. Y columns to cluster experience data



Experiences Data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	N Rows	CCC	Best
Self Organizing Map	3	1	-9.315	Optimal CCC

Columns Scaled Individually

Control Panel

SOM Grid 1 by 3

Columns Scaled Individually

Bandwidth: 0.4330127

Cluster Summary

Cluster	Count
1	514
2	669
3	320

Step	Criterion
14	0

Cluster Means

Cluster	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)
1	1.04863651	1.55058422	1.8054424	1.09730883	3.8579897	3.60118435	3.40468737
2	1.04458709	1.60243092	1.53491152	2.89255837	3.9494885	3.70739876	3.43253845
3	1.14374582	1.61875144	1.58748692	2.61258398	3.20940865	2.72816892	2.3062967

Cluster Standard Deviations

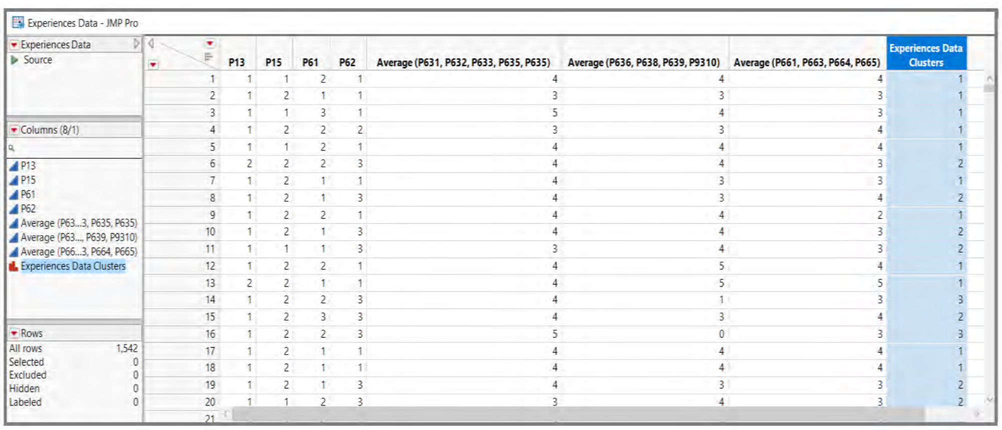
Cluster	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)
1	0.21511035	0.49743471	0.60565549	0.29633359	0.60899914	0.69913691	0.7961835
2	0.1968446	0.48909072	0.56959975	0.17072557	0.55705241	0.66531981	0.73200709
3	0.35083606	0.48569377	0.57973593	0.7415145	0.78056845	0.8010674	0.74997396

Figure 4-144. Iterative Clustering of experiences data by K Means & SOM

The figure 4-144 depicts the results of a clustering algorithm that uses a combination of K-Means and SOM methods. The number of clusters is determined by considering the CCC. The results indicate three clusters for experiences. The clustering process involves 14 steps of iterative clustering, which are summarized in the figure, including the count of each cluster: 514 for the first cluster, 669 for the second cluster, and 320 for the third cluster. Additionally, the means and standard deviations are provided for each cluster. The script for the clustering algorithm is written in Python and is provided in the text. The script applies the K-Means clustering method to the variables P13, P15, P61, P62, and the average of P631 to P635, among others. The SOM method is used with a specified number of clusters (3) and a bandwidth of 0.433. The script also includes instructions to send the results to a report, which includes a control panel and a table of cluster standard deviations.

The script (supports Python) for the clustering is in the following:

```
K Means Cluster(
	Y(
		:P13, :P15, :P61, :P62, :"Average (P631, P632, P633, P635, P635)"n,
		:"Average (P636, P638, P639, P9310)"n, :"Average (P661, P663, P664, P665)"n
	),
	{SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
	Number of Clusters( 3 ), SOM, Go},
	SendToReport(
		Dispatch( {}, "Control Panel", OutlineBox, {Close( 1 )} ),
		Dispatch(
			{"SOM Grid 1 by 3"},
			"Cluster Standard Deviations",
			OutlineBox,
			{Close( 0 )}
		)
	)
)
```



	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)	Experiences Data Clusters
1	1	1	2	1	4	4	4	1
2	1	2	1	1	3	3	3	1
3	1	1	3	1	5	4	3	1
4	1	2	2	2	3	3	4	1
5	1	1	2	1	4	4	4	1
6	2	2	2	3	4	4	3	2
7	1	2	1	1	4	3	3	1
8	1	2	1	3	4	3	4	2
9	1	2	2	1	4	4	2	1
10	1	2	1	3	4	4	3	2
11	1	1	1	3	3	4	3	2
12	1	2	2	1	4	5	4	1
13	2	2	1	1	4	5	5	1
14	1	2	2	3	4	1	3	3
15	1	2	3	3	4	3	4	2
16	1	2	2	3	5	0	3	3
17	1	2	1	1	4	4	4	1
18	1	2	1	1	4	4	4	1
19	1	2	1	3	4	3	3	2
20	1	1	2	3	3	4	3	2
21								

Figure 4-145. A part of clusters for each row of experiences data

Figure 4-145 illustrates the saved column in the data table that stores the assigned cluster for each row. The new column represents the clustering of the data based on numeric variables, where each row is assigned to one of a specified number of clusters. This information is useful for analyzing and categorizing the experience data based on similarities in the numeric variables.

Figure 4-146. Example Biplot for Experiences Data Clusters SOM

Figure 4-146 is a representation of the first two principal components of the experience data using a biplot. The pair of principal components, labeled as "PC1 and PC2", have been plotted to visualize the clusters in the experience data. The biplot shows the distribution of the data points in three clusters, which are labeled as "Cluster 1", "Cluster 2", and "Cluster 3". It is observed that there is some overlap between the two clusters and in a particular region, all three clusters overlap, indicating that there is a mixture of data points in that area.

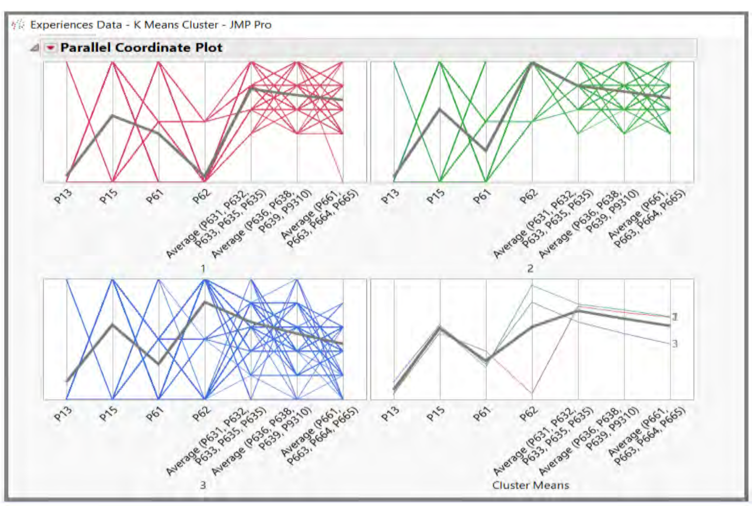


Figure 4-147. Parallel Coordinate Plot for Experiences Data

Figure 4-147 displays a plot for each cluster of the experience data, where the line segments connecting the points represent the rows of the data table. The plot suggests that the clusters share similar features, which may make it challenging to differentiate between them. However, the SOM algorithm effectively distinguishes between the three clusters, as indicated by the gray line representing the mean.

Figure 4-148 presents the regression line and confidence interval on a scatterplot matrix for the experience data and each cluster separately. The clear shaded region inside the ellipses on the scatterplot matrix represents the relationship between the Y variable of the experience data. The matrix also features ellipses, points, and a lower triangular scatter matrix for the covariates, with different ellipses being displayed for each level of the categorical variable X. The linear discriminant method depicted in the matrix is based on the pooled covariance matrix.

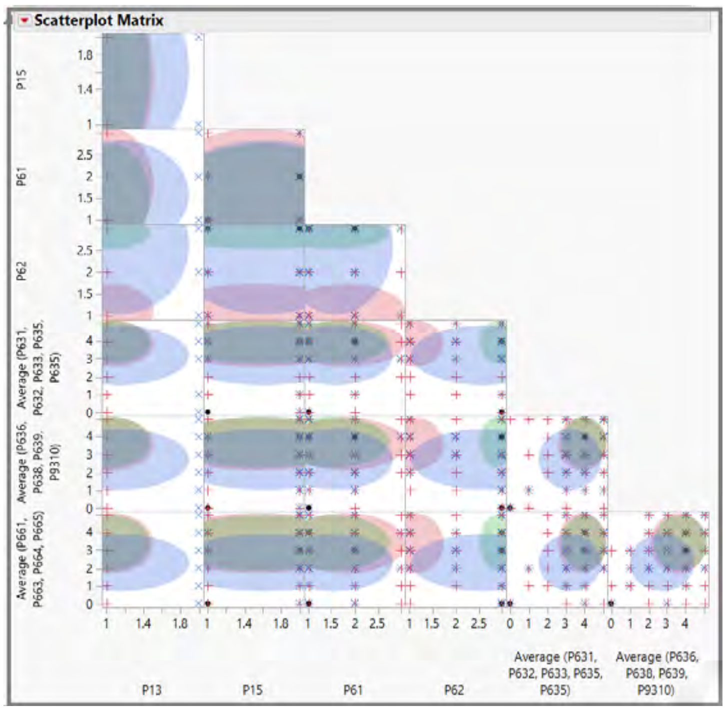
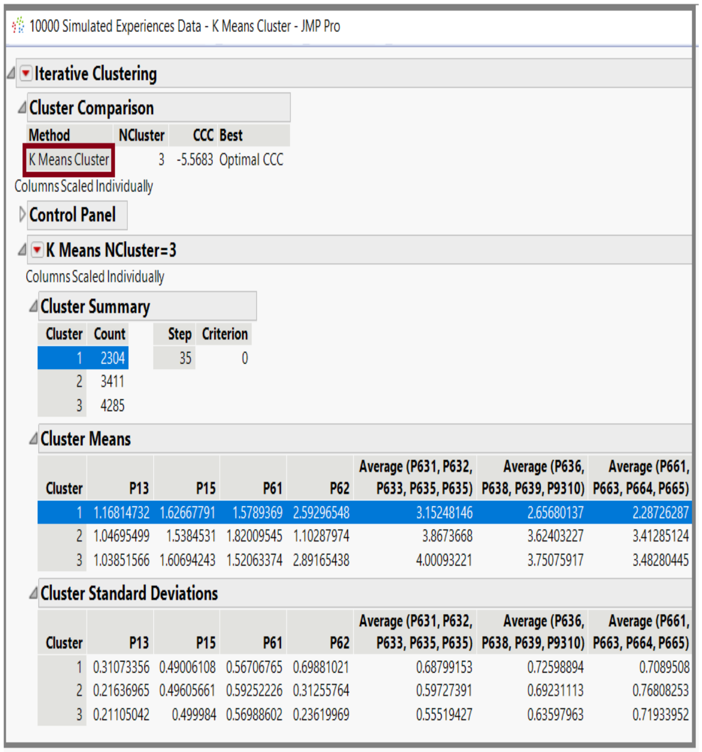


Figure 4-148. Scatterplot matrix for experiences data

The effectiveness of the clustering method applied to a large dataset was evaluated through simulation of experience data in JMP. A data table was generated using 10,000 samples and the estimated cluster mixing probabilities, means, and standard deviations for each cluster. The number of clusters was determined based on the criteria of the CCC in JMP. Figures 4-149 and 4-150 present the results of the K-Means method applied to the simulated experience data, which resulted in three clusters. The iteration process involved 35 steps, and the count of samples in each cluster was recorded as 2304 for the first cluster, 3411 for the second cluster, and

4285 for the third cluster. Additionally, the means and standard deviations for each cluster are also displayed.



10000 Simulated Experiences Data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	-5.5683	Optimal CCC

Columns Scaled Individually

Control Panel

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count	Step	Criterion
1	2304	35	0
2	3411		
3	4285		

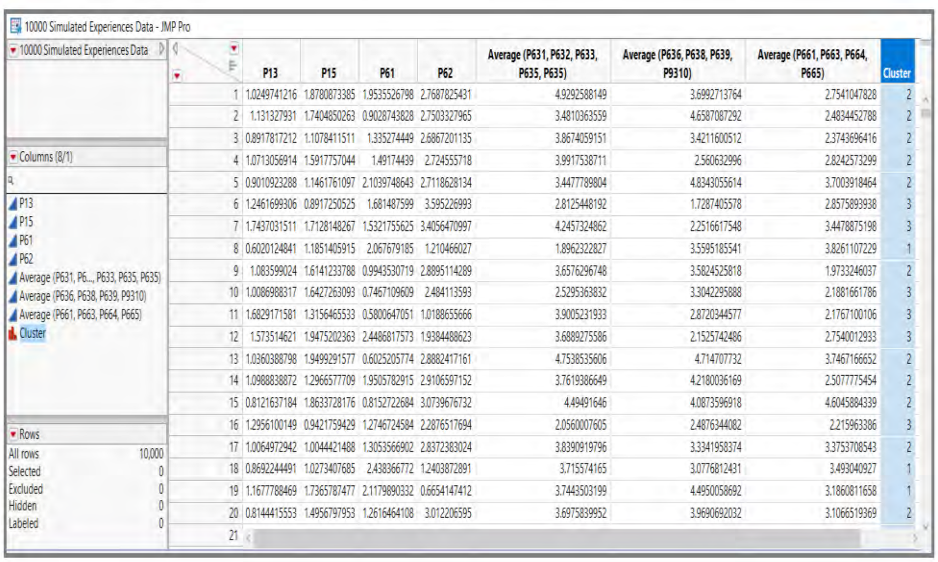
Cluster Means

Cluster	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)
1	1.16814732	1.62667791	1.5789369	2.59296548	3.15248146	2.65680137	2.28726287
2	1.04695499	1.5384531	1.82009545	1.10287974	3.8673668	3.62403227	3.41285124
3	1.03851566	1.60694243	1.52063374	2.89165438	4.00093221	3.75075917	3.48280445

Cluster Standard Deviations

Cluster	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)
1	0.31073356	0.49006108	0.56706765	0.69881021	0.68799153	0.72598894	0.7089508
2	0.21636965	0.49605661	0.59252226	0.31255764	0.59727391	0.69231113	0.76808253
3	0.21105042	0.499984	0.56988602	0.23619969	0.55519427	0.63597963	0.71933952

Figure 4-149. Iterative Clustering of simulated sample experiences data by the K-Means method



10000 Simulated Experiences Data - JMP Pro

10000 Simulated Experiences Data

Columns (8/1)

P13
P15
P61
P62
Average (P631, P6..., P633, P635, P635)
Average (P636, P638, P639, P9310)
Average (P661, P663, P664, P665)
Cluster

Rows

All rows	10,000
Selected	0
Excluded	0
Hidden	0
Labeled	0

	P13	P15	P61	P62	Average (P631, P632, P633, P635, P635)	Average (P636, P638, P639, P9310)	Average (P661, P663, P664, P665)	Cluster
1	1.0249741216	1.8780873385	1.9535526798	2.7687825431	4.9292588149	3.6992713764	2.7541047828	2
2	1.131327931	1.7404850263	0.9028743828	2.7503327965	3.4810363559	4.6587087292	2.4834452788	2
3	0.8917817212	1.1078411511	1.335274449	2.6867201135	3.8674059151	3.4211600512	2.3743696416	2
4	1.0713056914	1.5917757044	1.49174439	2.724555718	3.9917538711	2.560632996	2.8242573299	2
5	0.9010923288	1.1461761097	2.1039748643	2.7118628134	3.4477789804	4.8343055614	3.7003918464	2
6	1.2461699306	0.8917250525	1.681487599	3.595226993	2.8125448192	1.7287405578	2.8575893938	3
7	1.7437031511	1.7128148267	1.5321755625	3.4056470997	4.2457324862	2.2516617548	3.4478875198	3
8	0.6020124841	1.1851405915	2.067679185	1.210466027	1.8962322827	3.5595185541	3.8261107229	1
9	1.083599024	1.6141233788	0.9943530719	2.8895114289	3.6576296748	3.5824525818	1.9733246037	2
10	1.0086988317	1.6427263093	0.7467109609	2.484113593	2.5295363832	3.3042295888	2.1881661786	3
11	1.6829171581	1.3156465533	0.5800647051	1.0188655666	3.9005231933	2.8720344577	2.1767100106	3
12	1.573514621	1.9475202363	2.4486817573	1.9384488623	3.6889275586	2.1525742486	2.7540012933	3
13	1.0360388798	1.9499291577	0.6025205774	2.8882417161	4.7538535606	4.714707732	3.7467166652	2
14	1.0988838872	1.2966577709	1.9505782915	2.9106597152	3.7619386649	4.2180036169	2.5077775454	2
15	0.8121637184	1.8633728176	0.8152722684	3.0739676732	4.49491646	4.0873596918	4.6045884339	2
16	1.2956100149	0.9421759429	1.2746724584	2.2876517694	2.0560007605	2.4876344082	2.215963386	3
17	1.0064972942	1.0044421488	1.3053566902	2.8372383024	3.8390919796	3.3341958374	3.3753708543	2
18	0.8692244491	1.0273407685	2.438366772	1.2403872891	3.715574165	3.0776812431	3.493040927	1
19	1.1677788469	1.7365787477	2.1179890332	0.6654147412	3.7443503199	4.4950058692	3.1860811658	1
20	0.8144415553	1.4956797953	1.2616464108	3.012206595	3.6975839952	3.9690692032	3.1066519369	2
21								

Figure 4-150. A part of the simulated experiences data table

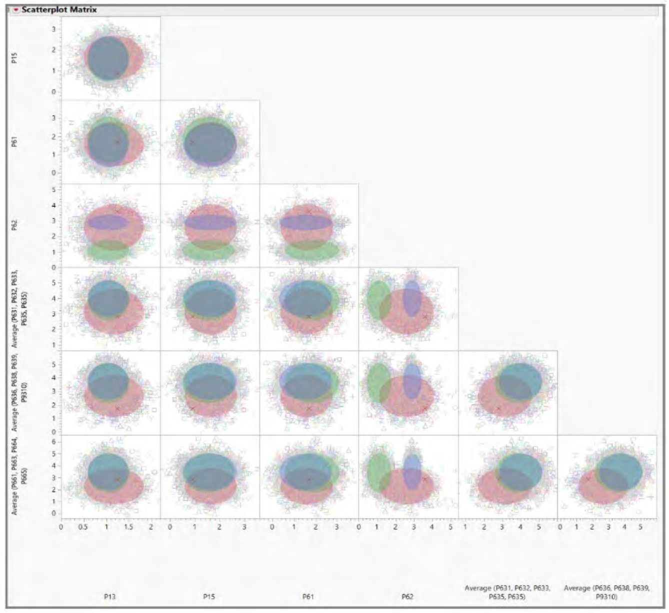


Figure 4-151. A scatterplot matrix based on the clusters of simulated experiences data

Figure 4-151 represents a scatterplot matrix that displays confidence ellipses based on the current number of clusters in simulated experience data. The scatterplot matrix demonstrates how the optimization of clusters is achieved through simulations performed on a scale of 10,000. The confidence ellipses provide an estimation of the variance within each cluster, visually representing the grouping of data points with similar characteristics.

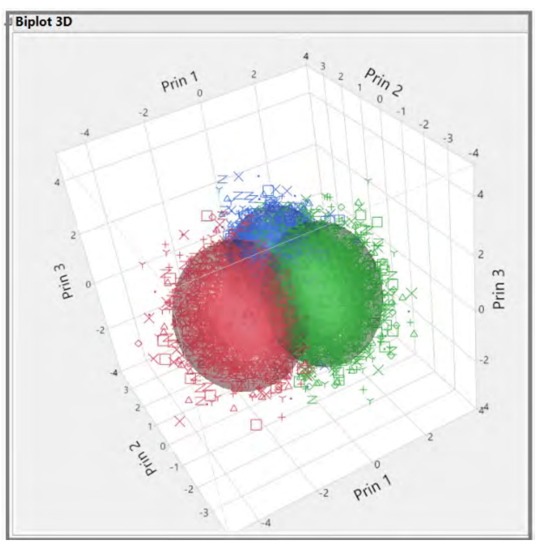
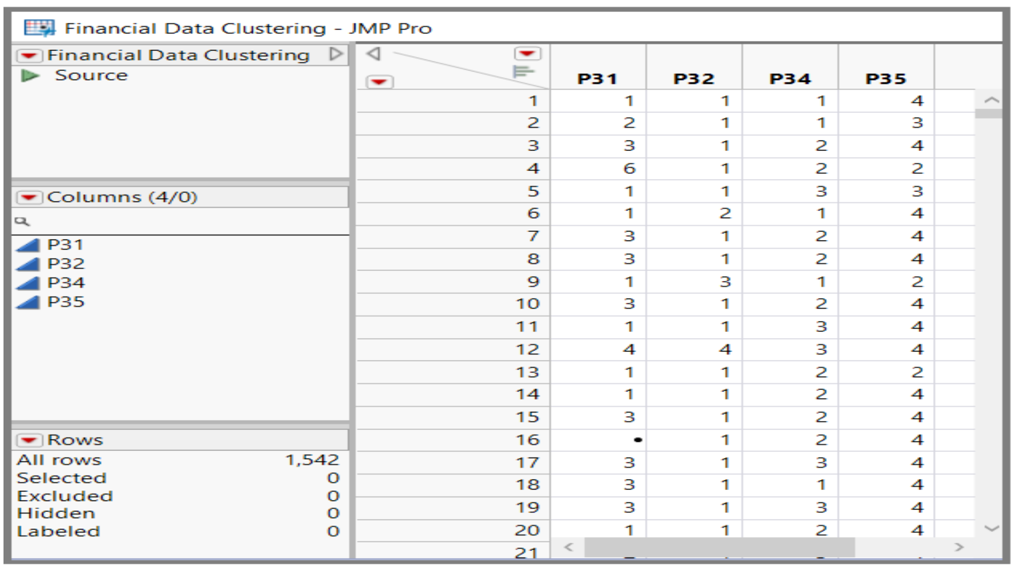


Figure 4-152. First three principal components of experiences data

Figure 4-152 presents a 3D biplot representation of the experience data, displayed through the first three principal components. The biplot provides a visual representation of the relationship between the points and clusters in the data, allowing for easy interpretation and analysis of the underlying patterns and structures. This visualization helps to understand the distribution of the experience data and how it is associated with different clusters, providing valuable insights for further analysis.

4.3.5. Clustering Input Financial Data

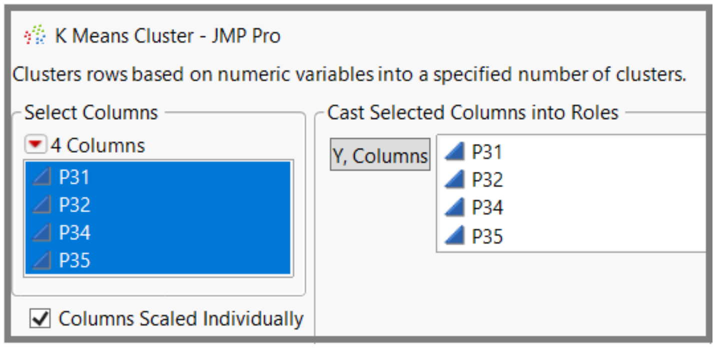
In the initial step of clustering the financial data, it is imported into JMP software in the form of four columns with 1542 rows of data. This is shown in Figure 4-153. The data is then preprocessed using K-Means and SOM methods within JMP to prepare it for further analysis. The purpose of this preprocessing is to group similar data points together into meaningful clusters, which will later be used to make investment recommendations. The K-Means and SOM methods are commonly used techniques for data clustering and play a crucial role in organizing the financial data for the Combined ANFIS system.



	P31	P32	P34	P35
1	1	1	1	4
2	2	1	1	3
3	3	1	2	4
4	6	1	2	2
5	1	1	3	3
6	1	2	1	4
7	3	1	2	4
8	3	1	2	4
9	1	3	1	2
10	3	1	2	4
11	1	1	3	4
12	4	4	3	4
13	1	1	2	2
14	1	1	2	4
15	3	1	2	4
16	•	1	2	4
17	3	1	3	4
18	3	1	1	4
19	3	1	3	4
20	1	1	2	4
21				

Figure 4-153. A part of imported financial data in JMP

The process of clustering the data involves grouping the rows of numeric variables into a specified number of clusters. In this case, the data is related to four questions about the respondents' current savings and financial situation. These questions serve as a measure of the impact of an unexpected event or major release on the potential investors. The results of this clustering are presented in Figure 4-154.



Figures 4-154. Y columns to cluster financial data

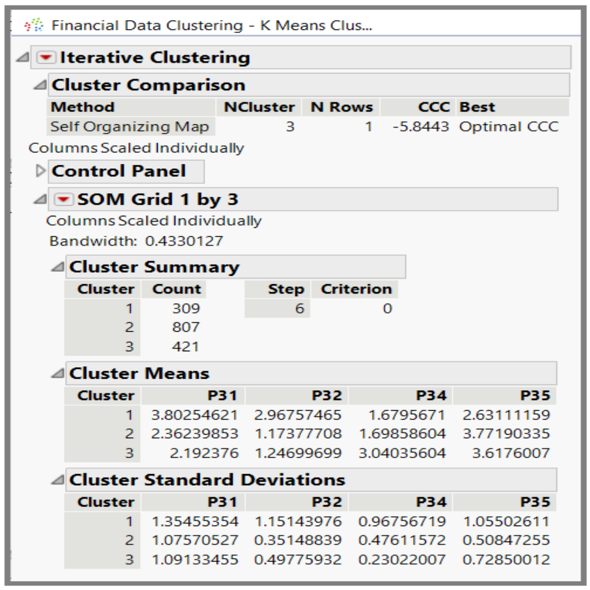
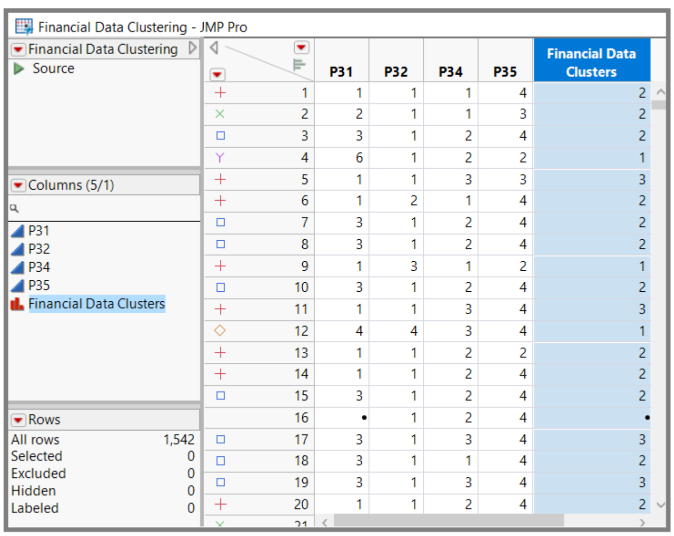


Figure 4-155. Iterative Clustering of financial data by K Means & SOM

The figure 4-155 shows the results of a clustering analysis performed on financial data using a combination of K-Means and SOM method. The number of clusters, three in this case, is determined based on the CCC (Cluster Characteristics Consistency) criterion. The figure displays the results of six iterative clustering steps, including the count of data points in each cluster (309, 807, and 421) and the means and standard deviations of each cluster. The script used for the clustering is written in Python and utilizes the K-Means Cluster function to cluster the data based on the variables P31, P32, P34, and P35. The script sets the number of clusters to 3, sets the SOM method for clustering, and outputs the standard deviations of each cluster in an OutlineBox. The output of the script can be sent to a report for further analysis. The script (supports Python) for the clustering is the following:

```
K Means Cluster(
        Y( :P31, :P32, :P34, :P35 ),
        {SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
        Number of Clusters( 3 ), SOM, Go},
        SendToReport(
                Dispatch( {}, "Control Panel", OutlineBox, {Close( 1 )} ),
                Dispatch(
                        {"SOM Grid 1 by 3"},
                        "Cluster Standard Deviations",
                        OutlineBox,
                        {Close( 0 )}
                )
        )
)
```



	P31	P32	P34	P35	Financial Data Clusters
1	1	1	1	4	2
2	2	1	1	3	2
3	3	1	2	4	2
4	6	1	2	2	1
5	1	1	3	3	3
6	1	2	1	4	2
7	3	1	2	4	2
8	3	1	2	4	2
9	1	3	1	2	1
10	3	1	2	4	2
11	1	1	3	4	3
12	4	4	3	4	1
13	1	1	2	2	2
14	1	1	2	4	2
15	3	1	2	4	2
16	•	1	2	4	•
17	3	1	3	4	3
18	3	1	1	4	2
19	3	1	3	4	3
20	1	1	2	4	2

Figure 4-156. A part of clusters for each row of financial data

Figure 4-156 depicts the addition of a new column to the data table, which stores the cluster assignment for each row. The clustering of the rows is based on numerical variables in the financial data, and the number of clusters is specified beforehand. The new column in the data table holds the cluster label for each corresponding row, providing an organized representation of the financial data based on its similarity within the specified number of clusters.

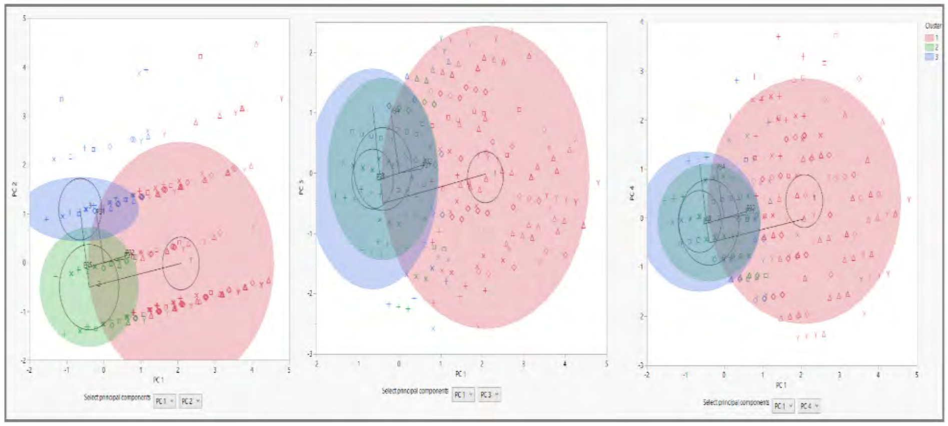


Figure 4-157. Example Biplots for Financial Data Clusters SOM

Figure 4-157 illustrates three biplots of the points and clusters in the first two principal components of the financial data. The biplots showcase the relationships between the pair of principle components, specifically "PC1 and PC2", "PC1 and PC3", and "PC1 and PC4". The figure displays all the data points of the three clusters present in the financial data. It can be observed that some of the points belong to two or even all three clusters, suggesting overlapping between the clusters. This highlights the presence of interdependence and complex relationships within the financial data, as different data points can belong to multiple clusters.

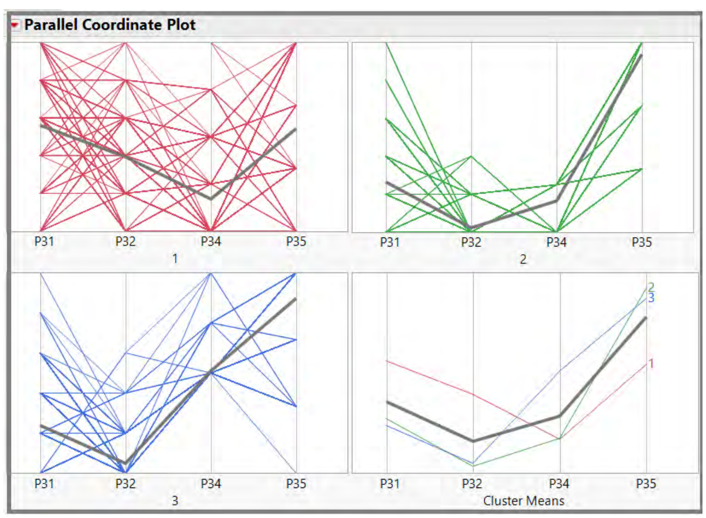


Figure 4-158. Parallel Coordinate Plot for Financial Data

Figure 4-158 presents a graphical representation of financial data, separated into individual clusters, where connected line segments denote each row in the data table. The figure highlights that the clusters can possess similar characteristics, making it challenging to differentiate between them. However, the SOM algorithm is effective in clearly separating the three clusters. The gray line represents the mean value. Figure 4-159 presents a scatterplot matrix, displaying the regression line and the confidence interval for the financial data, separated by cluster. The clear region inside the ellipses on the scatterplot matrix indicates the relationship between each dependent variable of the financial data. The matrix encompasses ellipses, points, and a lower triangular scatter matrix of the covariates. The ellipses with different overlays signify different levels of the categorical independent variable X. The linear discriminant analysis in this matrix is based on a combined covariance matrix.

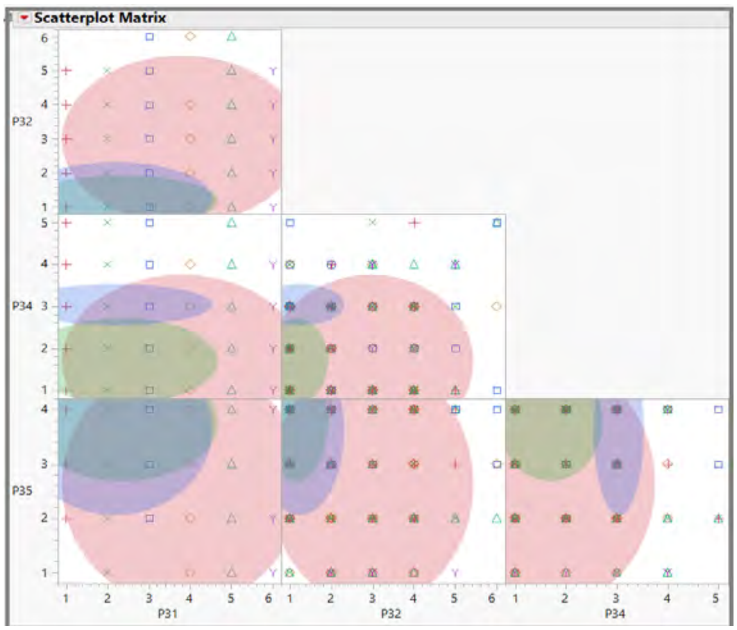
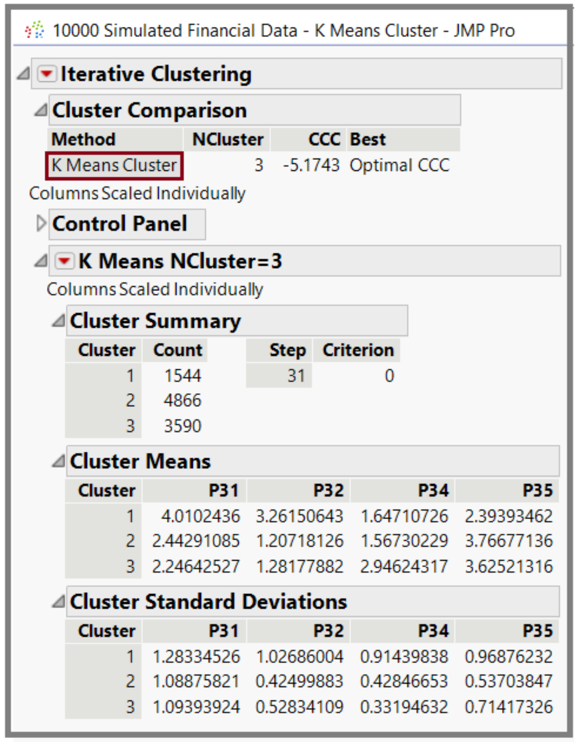


Figure 4-159. Scatterplot matrix for financial data

In order to assess the performance of the chosen clustering method with a large dataset, financial data was simulated using JMP with a sample size of 10,000. A new data table was created with the simulated financial data based on the estimated cluster mixing probabilities, means, and standard deviations for each cluster. The number of clusters determined by JMP is based on the CCC. Figures 4-160 and 4-161 present the results of the simulation, which depict three clusters for the simulated financial data generated through the K-Means method. The figures illustrate the iterative clustering process and cluster summary in 31 steps, with the count of the first cluster being 1544, the second cluster 4866, and the third cluster 3590. The figures also show the means and standard deviations for each cluster.



10000 Simulated Financial Data - K Means Cluster - JMP Pro

Iterative Clustering

Cluster Comparison

Method	NCluster	CCC	Best
K Means Cluster	3	-5.1743	Optimal CCC

Columns Scaled Individually

Control Panel

K Means NCluster=3

Columns Scaled Individually

Cluster Summary

Cluster	Count
1	1544
2	4866
3	3590

Step	Criterion
31	0

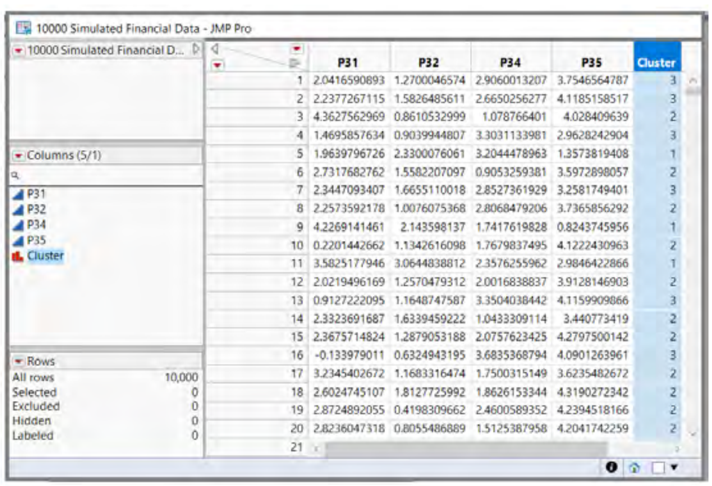
Cluster Means

Cluster	P31	P32	P34	P35
1	4.0102436	3.26150643	1.64710726	2.39393462
2	2.44291085	1.20718126	1.56730229	3.76677136
3	2.24642527	1.28177882	2.94624317	3.62521316

Cluster Standard Deviations

Cluster	P31	P32	P34	P35
1	1.28334526	1.02686004	0.91439838	0.96876232
2	1.08875821	0.42499883	0.42846653	0.53703847
3	1.09393924	0.52834109	0.33194632	0.71417326

Figure 4-160. Iterative Clustering of simulated sample financial data by the K-Means method



10000 Simulated Financial Data - JMP Pro

10000 Simulated Financial D...

Columns (5/1)

P31
P32
P34
P35
Cluster

Rows

All rows	10,000
Selected	0
Excluded	0
Hidden	0
Labeled	0

	P31	P32	P34	P35	Cluster
1	2.0416590893	1.2700046574	2.9060013207	3.7546564787	3
2	2.2377267115	1.5826485611	2.6650256277	4.1185158517	3
3	4.3627562969	0.8610532999	1.078766401	4.028409639	2
4	1.4695857634	0.9039944807	3.3031133981	2.9628242904	3
5	1.9639796726	2.3300076061	3.2044478963	1.3573819408	1
6	2.7317682762	1.5582207097	0.9053259381	3.5972898057	2
7	2.3447093407	1.6655110018	2.8527361929	3.2581749401	3
8	2.2573592178	1.0076075368	2.8068479206	3.7365856292	2
9	4.2269141461	2.143598137	1.7417619828	0.8243745956	1
10	0.2201442662	1.1342616098	1.7679837495	4.1222430963	2
11	3.5825177946	3.0644838812	2.3576255962	2.9846422866	1
12	2.0219496169	1.2570479312	2.0016838837	3.9128146903	2
13	0.9127222095	1.1648747587	3.3504038442	4.1159909866	3
14	2.3323691687	1.6339459222	1.0433309114	3.440773419	2
15	2.3675714824	1.2879053188	2.0757623425	4.2797500142	2
16	-0.133979011	0.6324943195	3.6835368794	4.0901263961	3
17	3.2345402672	1.1683316474	1.7500315149	3.6235482672	2
18	2.6024745107	1.8127725992	1.8626153344	4.3190272342	2
19	2.8724892055	0.4198309662	2.4600589352	4.2394518166	2
20	2.8236047318	0.8055486889	1.5125387958	4.2041742259	2
21					

Figure 4-161. A part of the simulated financial data table

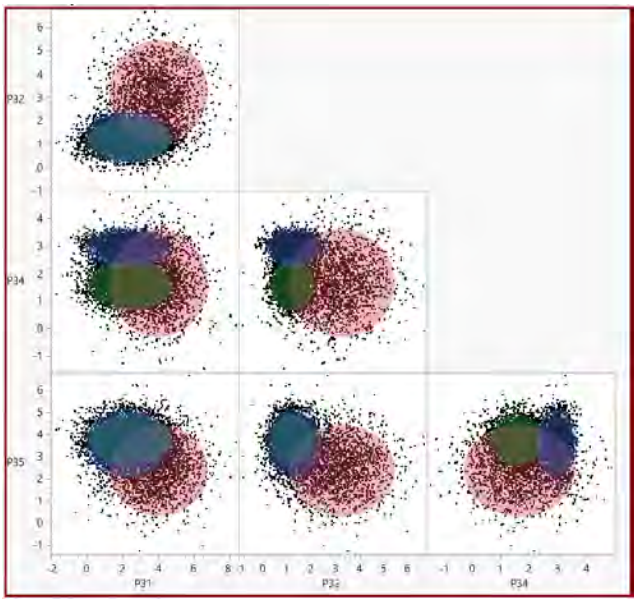


Figure 4-162. A scatterplot matrix based on the clusters of simulated financial data

Figure 4-162 presents a scatterplot matrix that displays the optimized clusters of simulated financial data. The scatterplot matrix includes confidence ellipses to visualize the distribution of data points within each cluster. The figure demonstrates the optimization of clusters by simulations conducted on a large scale of 10,000 data points. This information helps to understand the relationships between the financial variables and how they are grouped into clusters.

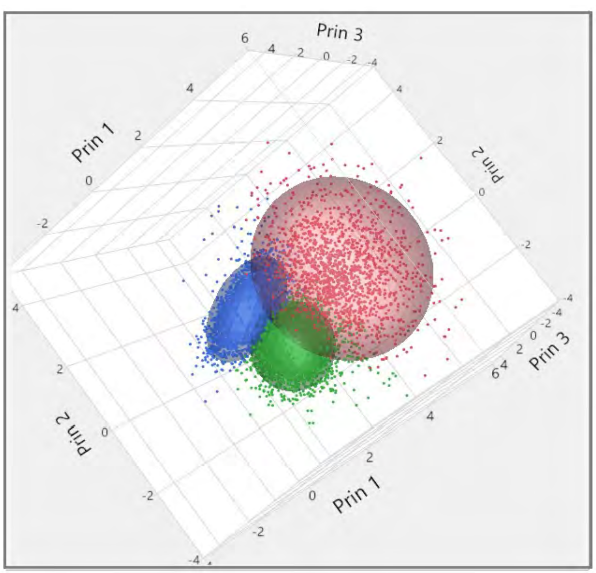
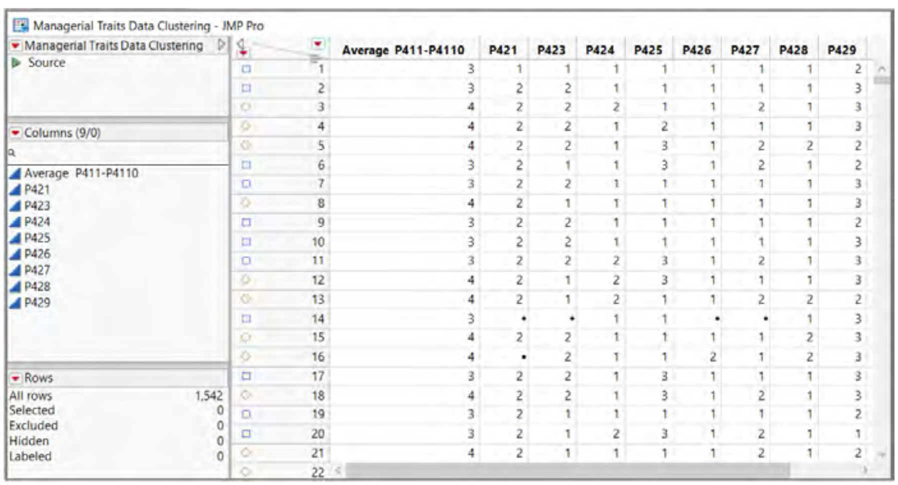


Figure 4-163. First three principal components of financial data

Figure 4-163 displays a 3D biplot representation of the financial data points and their respective clusters in the first three principal components. The biplot is a graphical representation that projects the data points onto a 3D plane, allowing for visualization of both the relationships between the data points and the structure of the underlying clusters. The first three principal components have been selected as they capture the most significant variations in the financial data. By visualizing the data in this manner, it is possible to gain insights into the underlying patterns and relationships within the financial data, helping to inform investment decision-making.

4.3.6. Clustering Input Managerial Traits Data

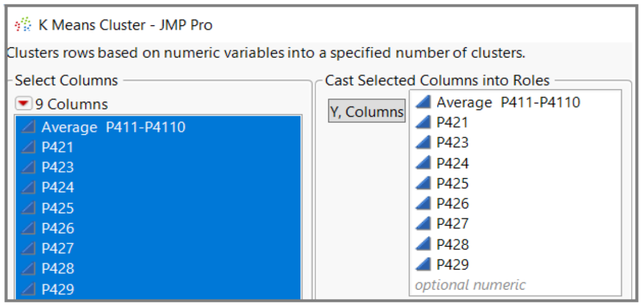
The initial step in the clustering process involved importing managerial trait data into JMP software, which is presented in four columns with a total of 1542 rows. The imported data can be visualized in Figure 4-164. After importing the data, the K-Means and SOM methods were employed in JMP to prepare the data for further analysis. The purpose of using these methods is to group similar data points together into clusters, allowing for a more comprehensive understanding of the characteristics of the managerial trait data.



	Average P411-P4110	P421	P423	P424	P425	P426	P427	P428	P429
1	3	1	1	1	1	1	1	1	2
2	3	2	2	1	1	1	1	1	3
3	4	2	2	2	1	1	2	1	3
4	4	2	2	1	2	1	1	1	3
5	4	2	2	1	3	1	2	2	2
6	3	2	1	1	3	1	2	1	2
7	3	2	2	1	1	1	1	1	3
8	4	2	1	1	1	1	1	1	3
9	3	2	2	1	1	1	1	1	2
10	3	2	2	1	1	1	1	1	3
11	3	2	2	2	3	1	2	1	3
12	4	2	1	2	3	1	1	1	3
13	4	2	1	2	1	1	2	2	2
14	3	•	•	1	1	•	•	1	3
15	4	2	2	1	1	1	1	2	3
16	4	•	2	1	1	2	1	2	3
17	3	2	2	1	3	1	1	1	3
18	4	2	2	1	3	1	2	1	3
19	3	2	1	1	1	1	1	1	2
20	3	2	1	2	3	1	2	1	1
21	4	2	1	1	1	1	2	1	2
22									

Figure 4-164. A part of imported managerial traits data in JMP

In the next step, the data collected from the respondents is divided into different clusters based on numeric variables. This grouping is done to categorize the data into a specific number of clusters for better analysis. The data includes information related to four questions about the current savings and management traits of the respondents. These questions are used to determine the level of difficulty an unexpected event or major release may pose for potential investors. This information is presented in Figure 4-165.



Figures 4-165. Y columns to cluster managerial traits data

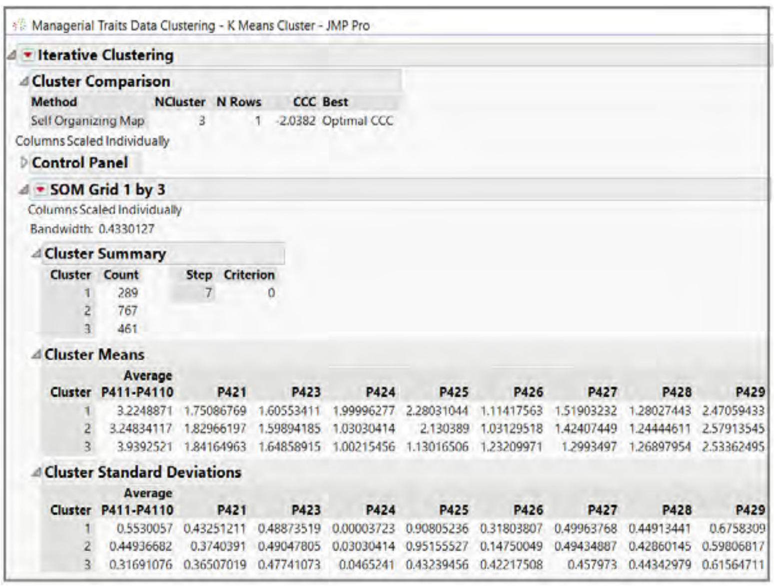
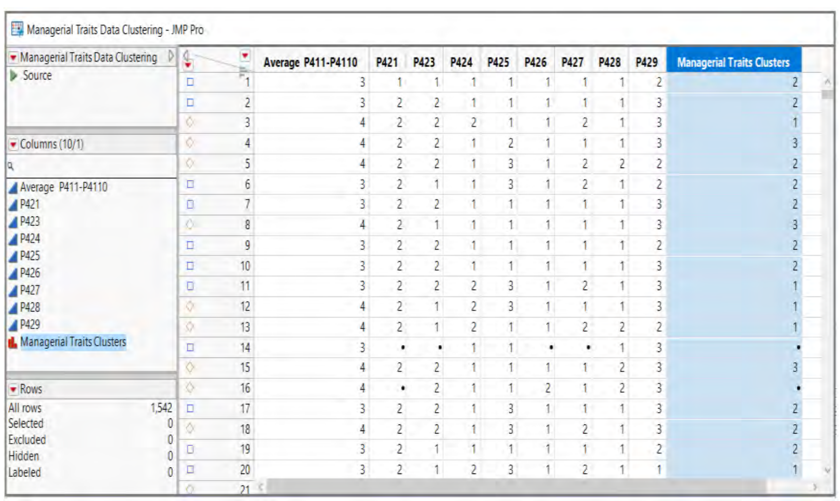


Figure 4-166. Iterative Clustering of managerial traits data by K Means & SOM

This paragraph describes the results of a clustering analysis conducted using the K-Means and SOM combined method in order to determine the number of clusters for managerial traits. The number of clusters was determined based on the CCC. The results of the analysis are shown in Figure 4-166, which indicates that three clusters were identified. The number of observations in each cluster are summarized, with Cluster 1 having 289 observations, Cluster 2 having 767, and Cluster 3 having 461. The mean and standard deviation for each cluster is also indicated. The script used for the clustering is provided, which was written in Python and executed using JMP software. The script uses the K-Means algorithm and SOM to cluster the data based on various managerial traits (such as "Average P411-P4110", "P421", "P423", "P424", "P425", "P426", "P427", "P428", and "P429"). The script also generates a report that includes the standard deviations for each cluster. The script (supports Python) for the clustering is in the following:

```
K Means Cluster(
	Y(
		:"Average  P411-P4110"n, :P421, :P423, :P424, :P425, :P426, :P427, :P428,
		:P429
	),
	{SOM N Rows( 1 ), SOM Bandwidth( 0.433012701892219 ), Single Step( 0 ),
	Number of Clusters( 3 ), SOM, Go},
	SendToReport(
		Dispatch( {}, "Control Panel", OutlineBox, {Close( 1 )} ),
		Dispatch(
			{"SOM Grid 1 by 3"},
			"Cluster Standard Deviations",
			OutlineBox,
			{Close( 0 )}
		)
	)
)
```



	Average P411-P4110	P421	P423	P424	P425	P426	P427	P428	P429	Managerial Traits Clusters
1	3	1	1	1	1	1	1	1	2	2
2	3	2	2	1	1	1	1	1	3	2
3	4	2	2	2	1	1	2	1	3	1
4	4	2	2	1	2	1	1	1	3	3
5	4	2	2	1	3	1	2	2	2	2
6	3	2	1	1	3	1	2	1	2	2
7	3	2	2	1	1	1	1	1	3	2
8	4	2	1	1	1	1	1	1	3	3
9	3	2	2	1	1	1	1	1	2	2
10	3	2	2	1	1	1	1	1	3	2
11	3	2	2	2	3	1	2	1	3	1
12	4	2	1	2	3	1	1	1	3	1
13	4	2	1	2	1	1	2	2	2	1
14	3	•	•	1	1	•	•	1	3	•
15	4	2	2	1	1	1	1	2	3	3
16	4	•	2	1	1	2	1	2	3	•
17	3	2	2	1	3	1	1	1	3	2
18	4	2	2	1	3	1	2	1	3	2
19	3	2	1	1	1	1	1	1	2	2
20	3	2	1	2	3	1	2	1	1	1

Figure 4-167. A part of clusters for each row of managerial traits data

Figure 4-167 illustrates the addition of a new column in a data table, which stores the cluster assignment for each row. The cluster assignments are based on the numeric variables present in the data related to managerial traits and are grouped into a specified number of clusters. The new column in the data table provides a clear and organized representation of the clustering of the managerial trait data.

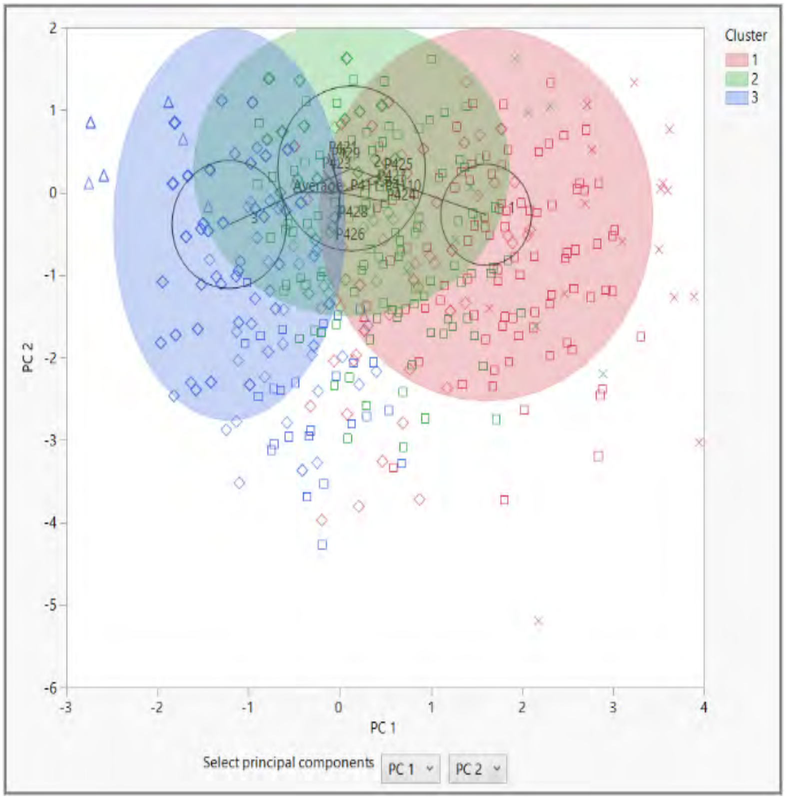


Figure 4-168. Example Biplot for Managerial traits Data Clusters SOM

Figure 4-168 is a biplot representation of the managerial traits data in the first two principal components, labeled as "PC1" and "PC2". The plot depicts the distribution of the rows of the managerial traits data into three clusters. The figure reveals that there is some overlap between the two clusters, and in one section, all three clusters overlap. This biplot provides a visual representation of the relationships and similarities between the different clusters of managerial traits data in the first two principal components.

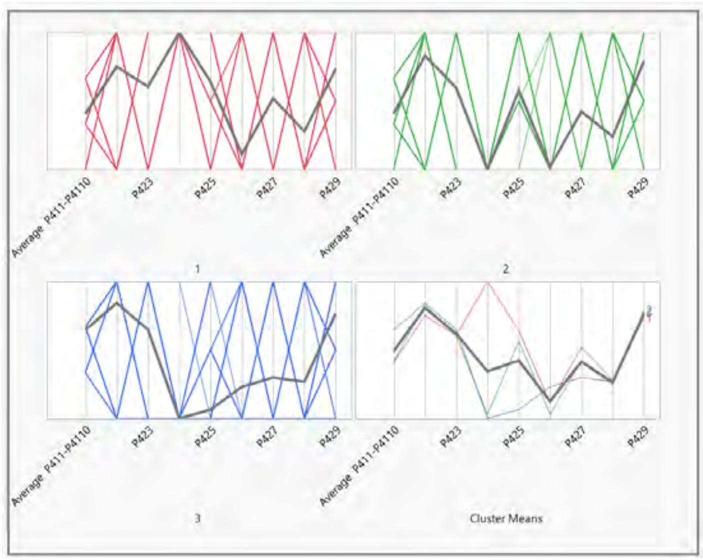


Figure 4-169. Parallel Coordinate Plot for Managerial traits Data

The figure 4-169 and 4-170 provide a visual representation of the relationship between managerial traits data and the clusters that are formed from that data. Figure 4-169 displays the plot for each cluster of managerial traits data separately, where the line segments connect each data point, revealing that the clusters can have similar features. However, the use of SOM is effective in highlighting the differences between these clusters, as demonstrated by the gray line, which represents the mean. Figure 4-170 displays a scatterplot matrix that contains a regression line and confidence interval for each cluster of managerial traits data. The clear shaded region inside the ellipses on the scatterplot matrix indicates the relationship between the Y variable of the managerial traits data. The scatterplot matrix also includes ellipses, points, and a lower triangular scatter matrix for the covariates. The different overlays of the ellipses are based on the categorical variable X. The linear discriminant method used in this matrix is calculated using the pooled covariance matrix.

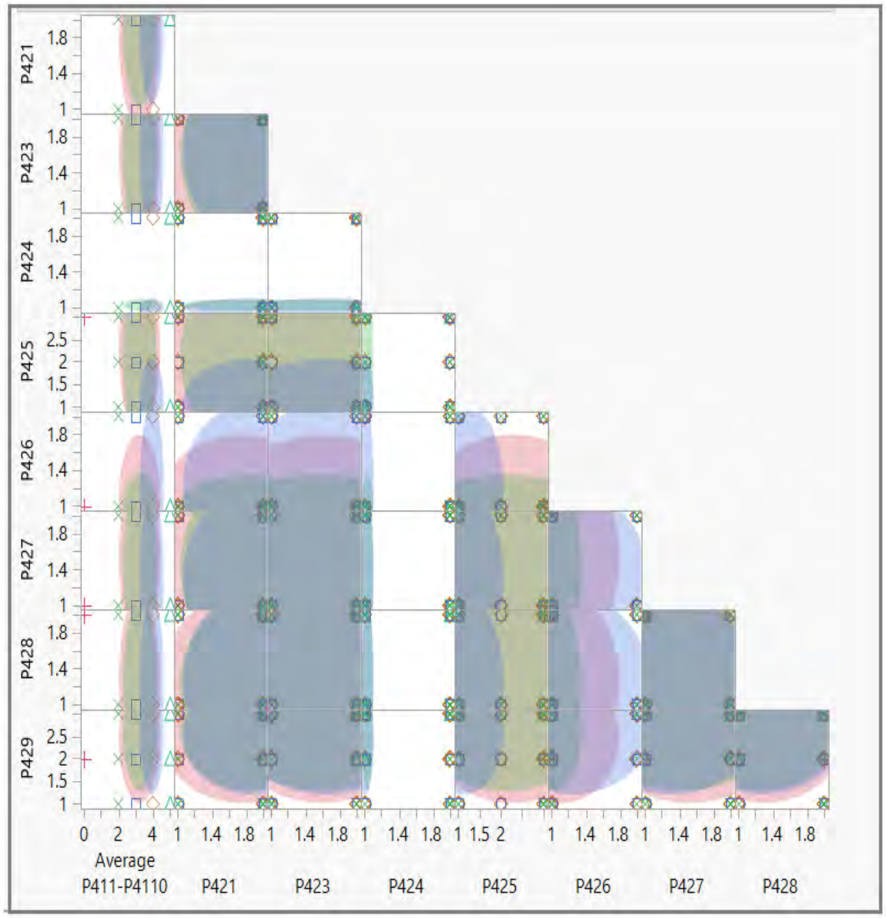


Figure 4-170. Scatterplot matrix for managerial traits data

The performance of the used clustering method was evaluated using a large volume of simulated managerial traits data. The data was generated in JMP with a sample size of 10,000. A new data table was constructed using the estimated cluster mixing probabilities, means, and standard deviations for each cluster. The number of clusters was determined by JMP using the CCC. Figures 4-171 and 4-172 display the results of the K-Means method, showing the simulation of the managerial traits data into three clusters. The iteration process and cluster summary in 41 steps are also shown, including the count of each cluster, with 3895 samples in Cluster 1, 4215 samples in Cluster 2, and 1890 samples in Cluster 3. Additionally, the means and standard deviations for each cluster are presented.

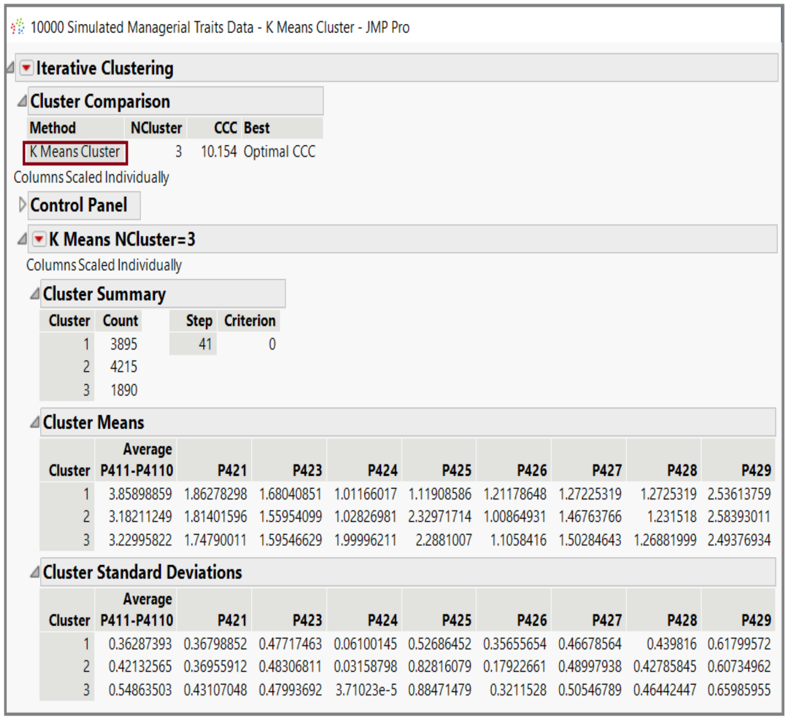


Figure 4-171. Iterative Clustering of simulated sample managerial traits data by the K-Means method

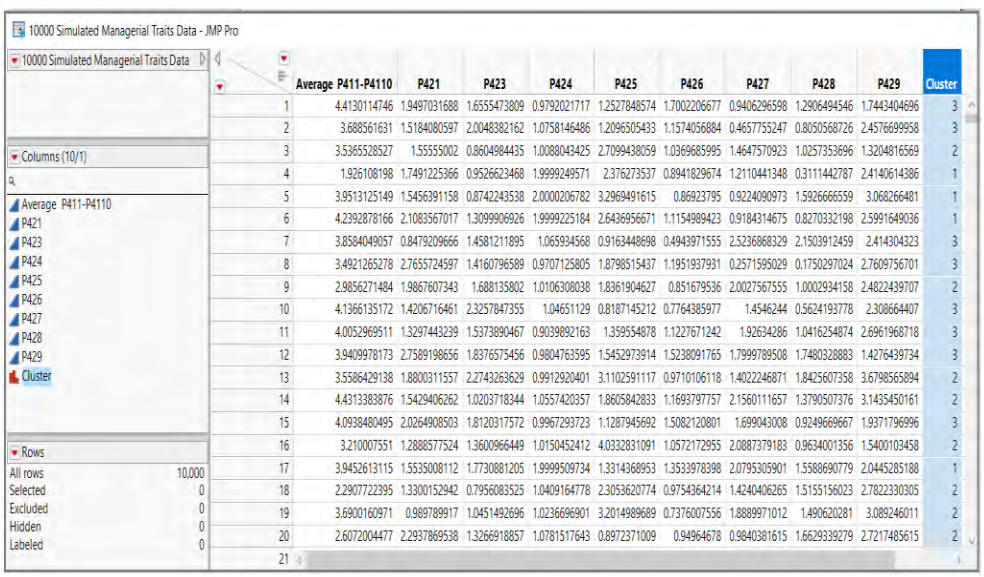


Figure 4-172. A part of the simulated managerial traits data table

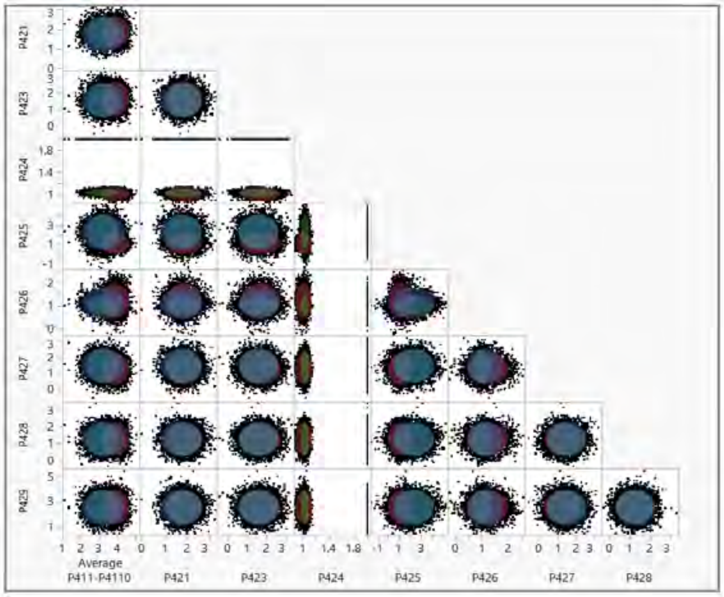


Figure 4-173. A scatterplot matrix based on the clusters of simulated managerial traits data

Figure 4-173 illustrates the results of a scatterplot matrix analysis performed on simulated data of managerial traits. The scatterplot matrix displays the relationship between different variables and is accompanied by confidence ellipses that reflect the level of confidence in the cluster assignments. The figure demonstrates how the optimization of the clusters is achieved through simulations run on a scale of 10,000. This helps to visualize the distribution of the variables and identify any potential patterns or trends in the data.

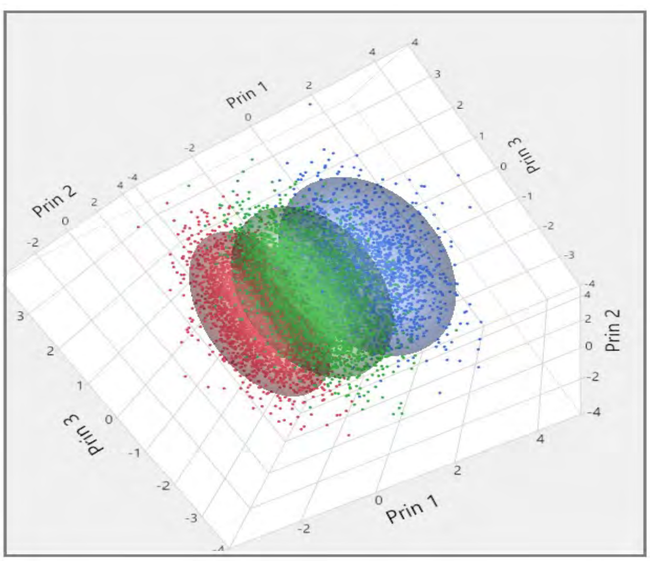


Figure 4-174. First three principal components of managerial traits data

Figure 4-174 is a 3D biplot representation of the data related to managerial traits. It displays the points and clusters in the first three principal components. The biplot allows for an exploration of the relationships between the points and the principal components, making it easier to visualize and understand the structure of the data. In this case, the data has been reduced to three dimensions, which allows for a clear visualization of the clusters that exist within the managerial traits data.

4.3.7. Combined Investment Type Recommender ANFIS

The Combined ANFIS was developed based on existing ANFIS models. The data was divided into six distinct clusters, each corresponding to a specific input for the Combined ANFIS. These clusters included demographic information, key decision factors, personality traits, experiences, financial information, and managerial traits. Each of these categories was further divided into three sub-clusters using K-Means and SOM methods in JMP. The Combined ANFIS model used these six clusters of data as inputs and had a single output, which represented the investment type or product, including options such as listed stock mutual funds, voluntary pension funds, government securities/bonds, and other financial products. The training dataset for the inputs and output consisted of 1542 data pairs, and the aggregation method used was Max with a Min implication.

Variables - COMBINEDANFIS

1542x7 table

	1 DemographicClusters	2 DecisionKeyFactorsClusters	3 PersonalityTraitsClusters	4 ExperiencesClusters	5 FinancialClusters	6 ManagerialTraitsClusters	7 InvestmentTypeClusters
1	3		2	1	2	2	1
2	3	2	2	1	2	2	2
3	3	2	1	1	2	1	2
4	2	1	1	1	1	3	2
5	3		1	1	3	2	1
6	2	1	2	2	2	2	1
7	3	1	3	1	2	2	2
8		3	3	2	2	3	2
9	2	1	2	1	1	2	3
10	2	1	2	2	2	2	3
11		3	3	2	3	1	2
12		3	3	1	1	1	3
13		3	2	1	2	1	1
14	2		2	3	2		
15	2	1	3	2	2	3	3
16				3			
17		2	3	1	3	2	2
18	1	2	3	1	2	2	2
19	3	2	3	2	3	2	2
20	1	3	2	2	2	1	2

Figure 4-175. A part of imported data to MATLAB to propose the CombinedANFIS

The data imported into MATLAB consisted of 7 columns, with 6 columns pertaining to the categorized factors of potential investors and the final column relating to the cluster type of investment. The input and output for the Combined ANFIS were specified within the fuzzy logic function. The Combined ANFIS was designed using the Sugeno type as a novel FIS. Figure 4-175 depicts the imported data in MATLAB as inputs and outputs for the proposed Combined ANFIS.

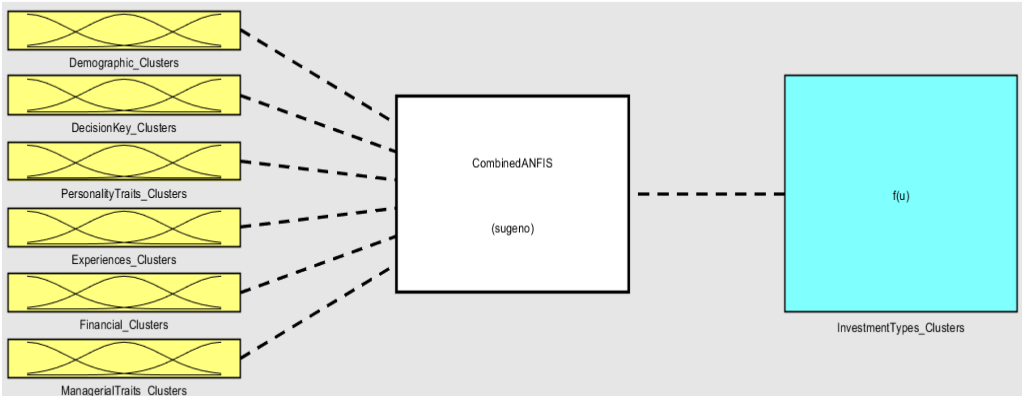


Figure 4-176. The properties of the CombinedANFIS

The design of the "Combined ANFIS" system is presented in Figure 4-176, along with its properties. The system includes nine inputs, consisting of demographic clusters, key decision factors clusters, personality traits clusters, experiences clusters, and financial clusters, as well as managerial traits clusters. The system produces one output, investment type clusters.

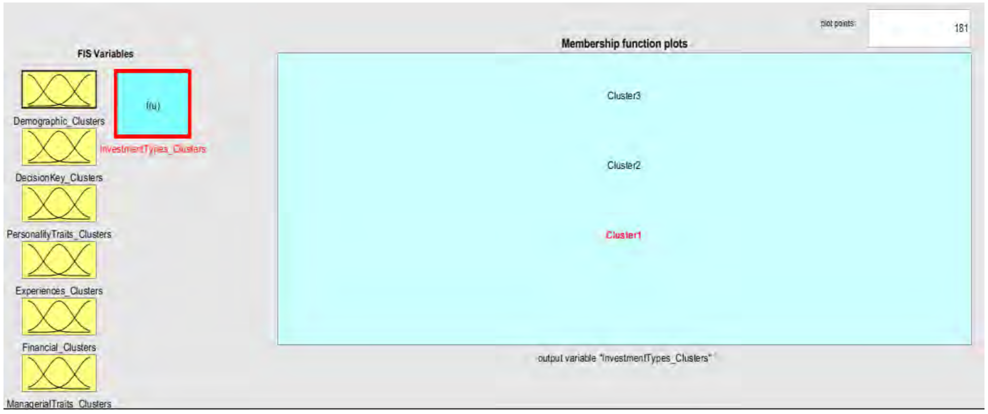


Figure 4-177. Output MFs in the CombinedANFIS

Figure 4-177 illustrates the MFs for the output of the CombinedANFIS system. The MFs are of constant type and are applied to three investment clusters: "Cluster 1", "Cluster 2", and "Cluster 3”.

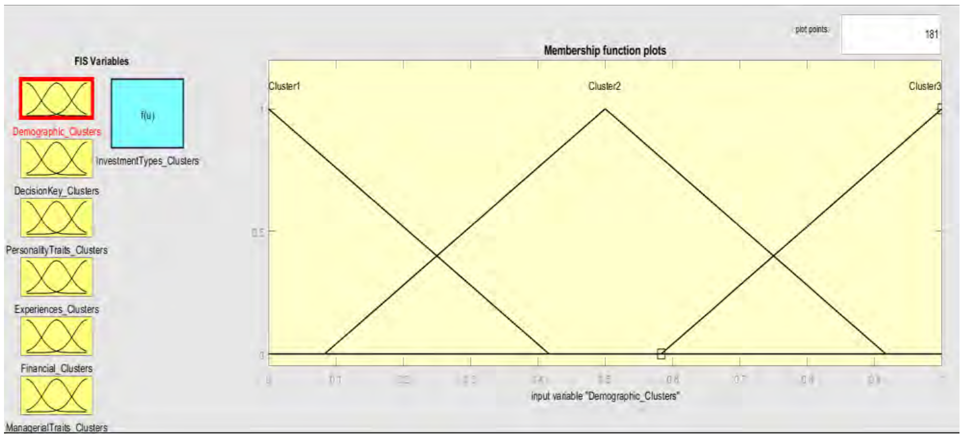


Figure 4-178. MFs shape for input 1 (Demographics Clusters) in the CombinedANFIS

Figure 4-178 showcases the trimf shapes for the first input of the Combined ANFIS system. The MF is composed of three clusters, labeled as "Cluster 1", "Cluster 2", and "Cluster 3". Similarly, the other inputs of the Combined ANFIS system also have three trimf MFs, each categorized as "Cluster 1", "Cluster 2", and "Cluster 3".

4.3.8. Proposing Combined Investment Type Recommender ANFIS

The Combined Investment Type Recommender ANFIS is a recommendation system that uses ANFIS to make investment suggestions. The system has six inputs and one output, with each input and output having three trimf. These MFs have a maximum value of one and a minimum value of zero.

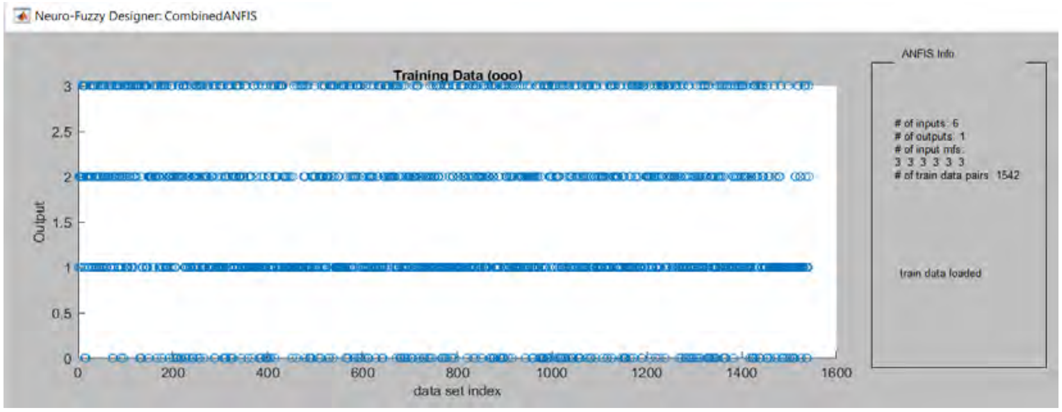


Figure 4-179. Prepared data in the CombinedANFIS

To develop the Combined ANFIS system, the fuzzy logic toolbox of MATLAB was used, which involved six steps: importing data, designing the FIS, preparing the data, generating the FIS, training the FIS, and modeling the FIS. Figure 4-179 depicts the data that has been pre-processed for the next stages of training and validation in the Combined ANFIS system. To train the data for the new FIS, a grid partition approach was utilized, and the optimization method was a hybrid one with a tolerance error of 0 and 3 epochs. The result of this process is the generation of a new FIS, the Combined ANFIS. Figure 4-180 provides a summary of the MFs for this Combined ANFIS system, including information about their shapes and arrangements.

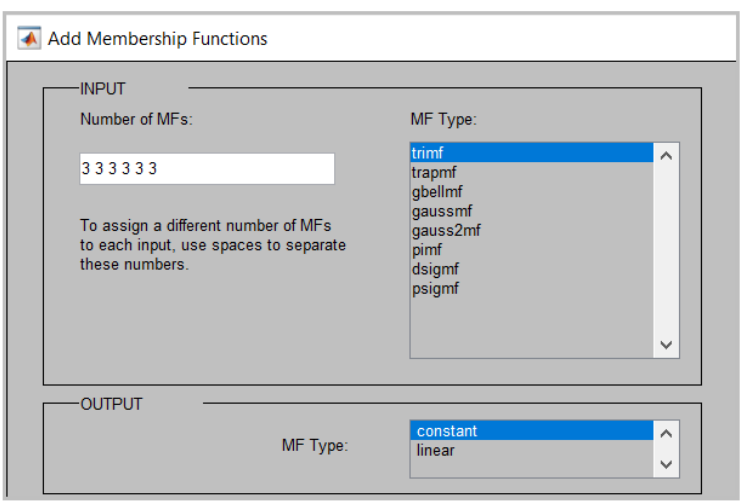


Figure 4-180. Information for generating the CombinedANFIS as a new FIS

The CombinedANFIS is a new FIS that has 6 inputs and 1 output. It is designed to provide insights into investment-type clusters by analyzing and processing a data set consisting of 1542 samples. The x-axis of the data set index represents the sample number, while the y-axis displays the distribution of the output based on investment-type clusters. This information can be used to make informed investment decisions by identifying patterns and relationships between the inputs and the output. The goal of the CombinedANFIS system is to analyze the data set and provide an accurate representation of the distribution of the output based on investment-type clusters, allowing users to make more informed decisions.

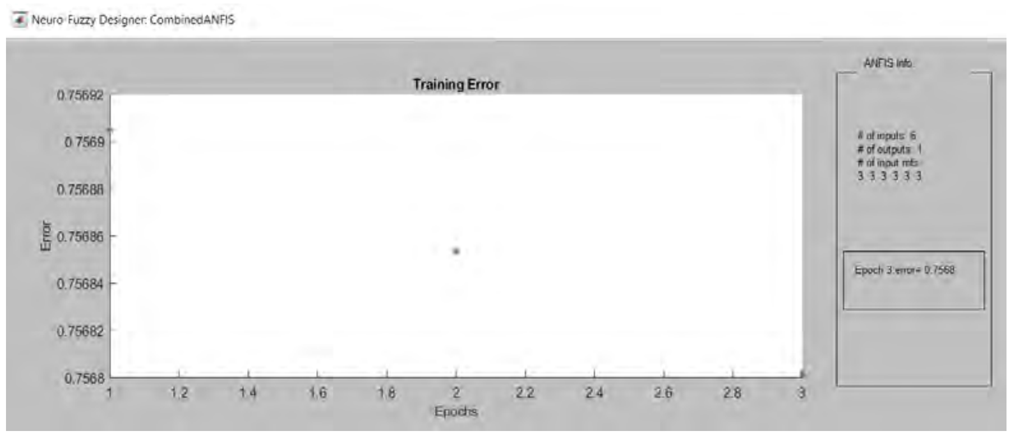


Figure 4-181. The trained CombinedANFIS

Figure 4-181 displays the trained grid of the Combined ANFIS system. It has four inputs and one output for investment type clustering. The training of the FIS was performed using a hybrid approach over three epochs.

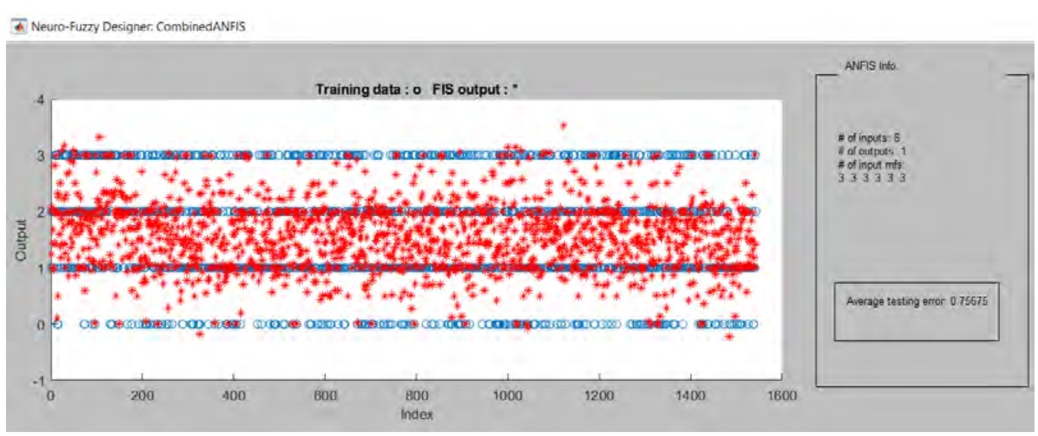


Figure 4-182. Trained data in the CombinedANFIS

The error for each epoch is approximately 0.76. The ANFIS info section provides information about the training process of the Combined ANFIS system. It has a total of 1503 nodes, 729 linear parameters, 54 nonlinear parameters, and 783 total parameters. The training process used 1542 data pairs and there were no checking data pairs. The number of fuzzy rules is 729. The process of training the ANFIS began and was completed after two epochs with a minimal training RMSE of 0.756752. This indicates that the Combined ANFIS system has successfully been trained using the given data and can be used for investment type clustering. Figure 4-182 depicts the trained Combined ANFIS system. The average training error, indicated by the value 0.75675, represents the deviation between the actual and predicted output values during the training process. The system generated a total of 2,880 rules, which are the decision-making mechanisms that drive the investment recommendations.

The F1-score is a useful metric for evaluating the performance of classification models and can be used for ANFIS as well. Here, by using F1-score formulas and the predicted and true labels of the test set, the F1-score was calculated for the combined ANFIS and evaluated its performance. The formula for F1-score is:

*F1-score = 2 * (precision * recall) / (precision + recall)*
where precision = true positives / (true positives + false positives)
and recall = true positives / (true positives + false negatives)

The result is in the following:

```
>> % True labels of the test set
y_true = [0, 1, 1, 0, 1, 0, 1, 1];

% Predicted labels of the test set
y_pred = [0, 1, 0, 1, 1, 0, 1, 0];
```

```
% Calculate the number of true positives, false positives, and false negatives
tp = sum(y_true == 1 & y_pred == 1);
fp = sum(y_true == 0 & y_pred == 1);
fn = sum(y_true == 1 & y_pred == 0);

% Calculate the precision and recall
precision = tp / (tp + fp);
recall = tp / (tp + fn);

% Calculate the F1-score
f1_score = 2 * precision * recall / (precision + recall);

% Print the F1-score
fprintf('The F1-score is: %f\n', f1_score);
The F1-score is: 0.766667
```

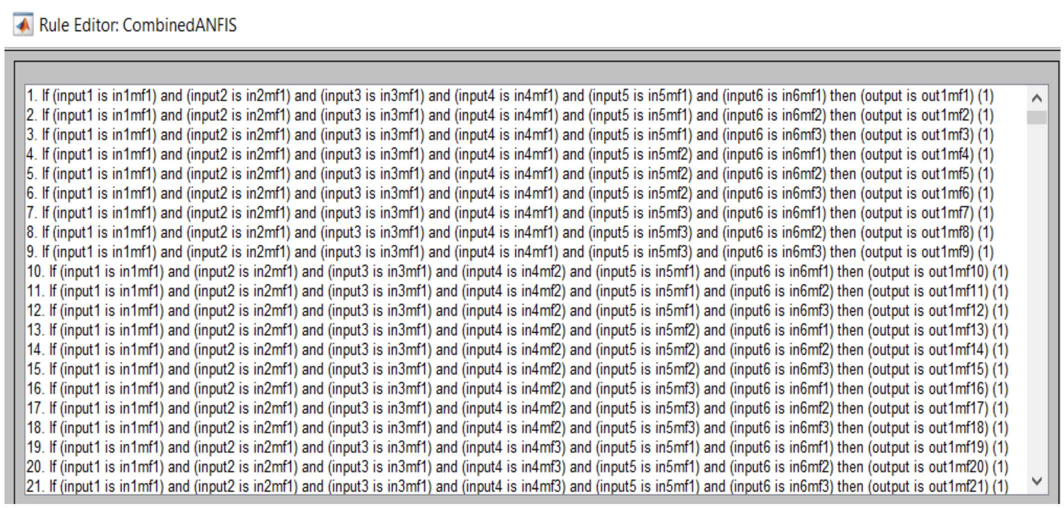


Figure 4-183. A part of the generated rules in the CombinedANFIS

Figure 4-183 showcases a portion of the generated rules in a more detailed format. The display of these rules in verbose format allows for further analysis and customization. Expert opinions and feedback from investors can be incorporated into the system by adding, modifying, or removing rules as deemed necessary. Attachment 6 provides information on how the Combined ANFIS system generates rules. The Combined ANFIS system is a combination of ANNs and fuzzy logic, and it uses these two techniques to generate rules that help make investment recommendations. Attachment 6 provides insights into the process that the Combined ANFIS system follows to generate these rules. This information could be useful for understanding how the system works and for making

informed decisions based on the investment recommendations provided by the system.

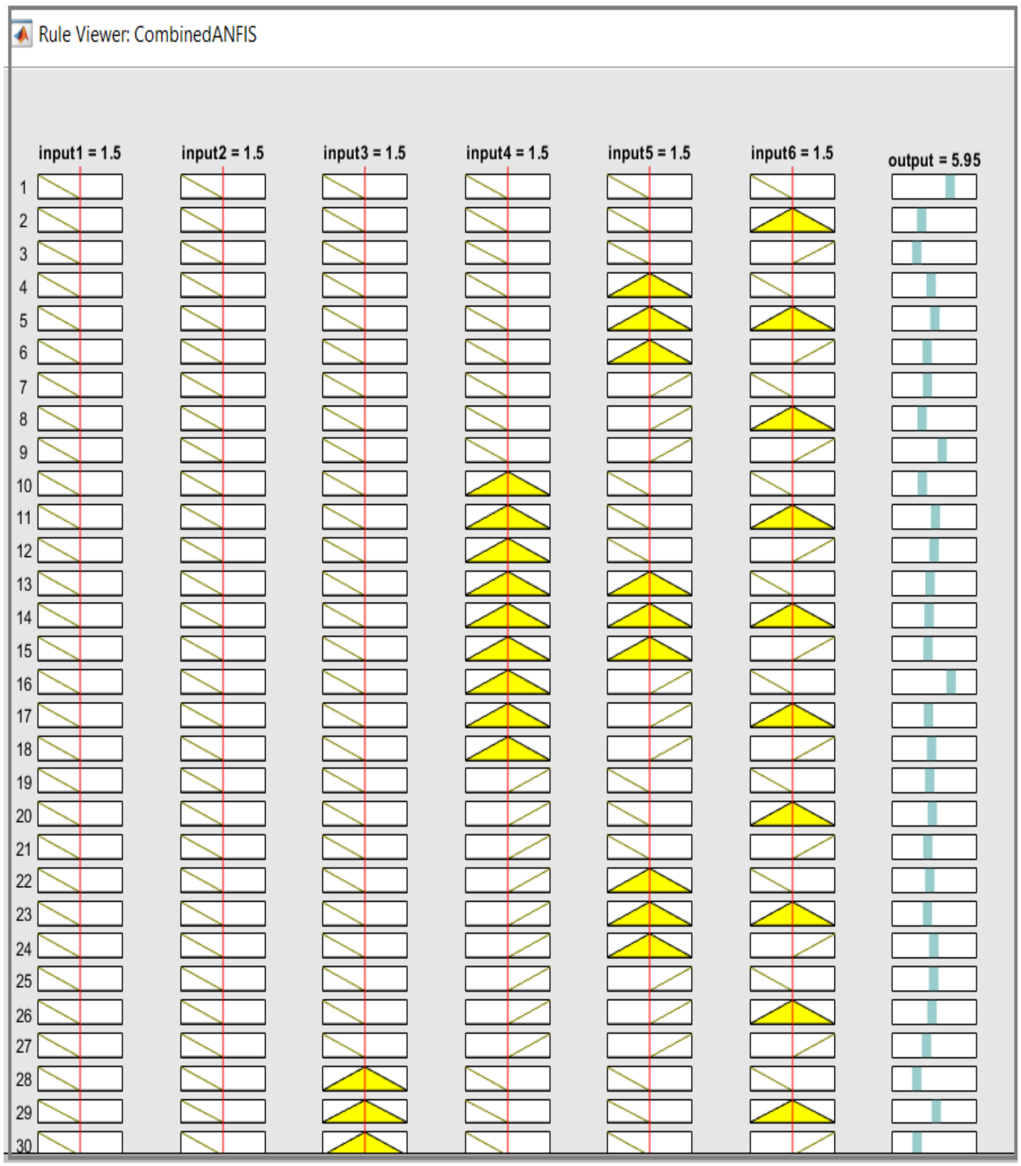
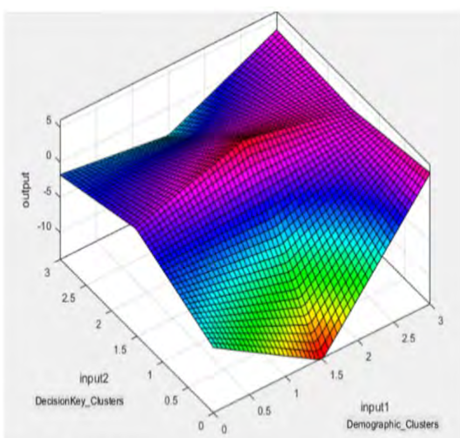
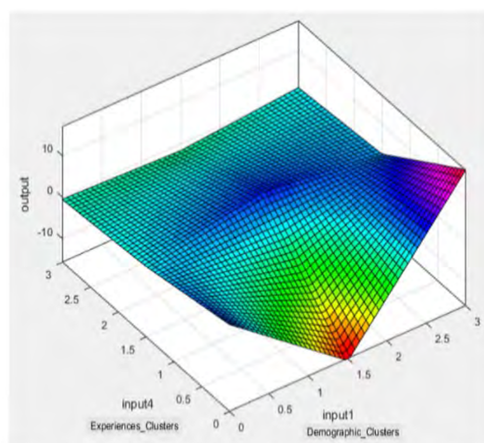


Figure 4-184. A part of the rule viewer in the CombinedANFIS

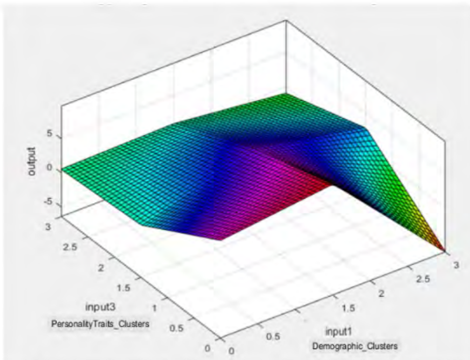
Figure 4-184 depicts a portion of the rule viewer for the open system of the Demographic ANFIS. The system contains 729 rules and 101 plotted points. The Combined ANFIS system is designed to uncover the relationship between various factors that impact investment decisions. Figures 4-185 (a-e) are 3D graphs that depict the effect of selected input pairs on the investment type. These surface graphs are non-linear and monolithic, meaning they provide a comprehensive representation of the investment type recommendations based on specific inputs. The recommended investment is more complex than a simple linear relationship because it considers several factors that influence the investment decision. Additionally, it uses machine learning techniques and ANFIS to process and analyze the data, which involves several layers of fuzzification, implication rules, normalization, defuzzification, integration, and aggregated output MF. This comprehensive approach helps to identify patterns, trends, and relationships in the data that are not apparent in a simple linear relationship. As a result, the recommended investment is more personalized and accurate, providing investors with better investment options and maximizing their returns.



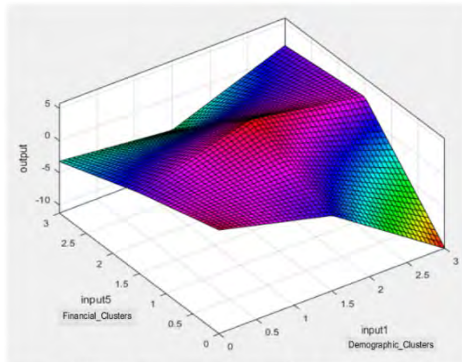
a. Demographics & Decision Keys



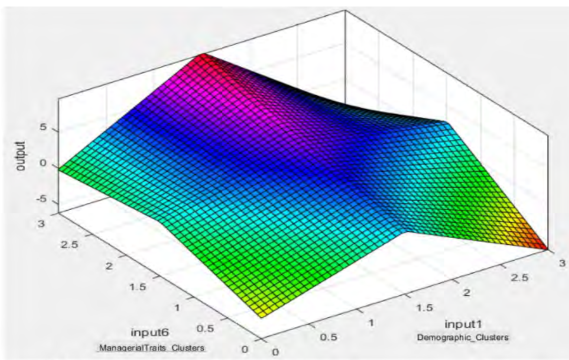
c. Demographics & Experiences



b. Demographics & Personality Traits



d. Demographics & Financial



e. Demographics & Managerial Traits

Figure 4-185. Effectiveness of the relations of each pair of Combined inputs on investment type

These figures demonstrate the effectiveness of the relationships between each pair of Combined inputs in determining the investment type. The ANFIS system takes all these inputs into account to make a personalized and informed investment recommendation.

Figure 4-186 depicts the structure of the Combined ANFIS Model. The figure represents the inputs, MFs, and various layers of the ANFIS system, including fuzzification, implication rules, normalization, defuzzification, and integration, which results in an investment recommendation for the investor. The recommendation is based on the demographic clusters and their effect on various factors, such as planning, stress, pace, influential, and daily schedule, as shown in Figures 4-185.

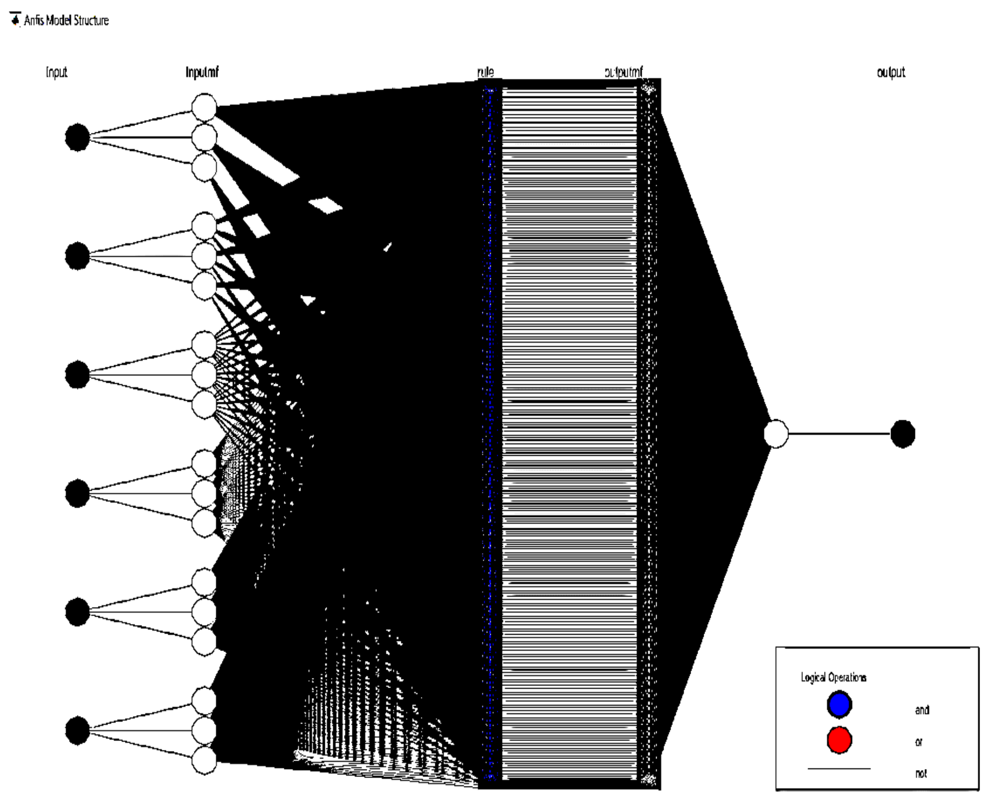


Figure 4-186. CombinedANFIS Model Structure

Figure 4-187 illustrates how the proposed model can be improved through feedback from investors and knowledge experts. When feedback is received, the expert can modify the rules of the system by adding, deleting, or changing them to improve the model.

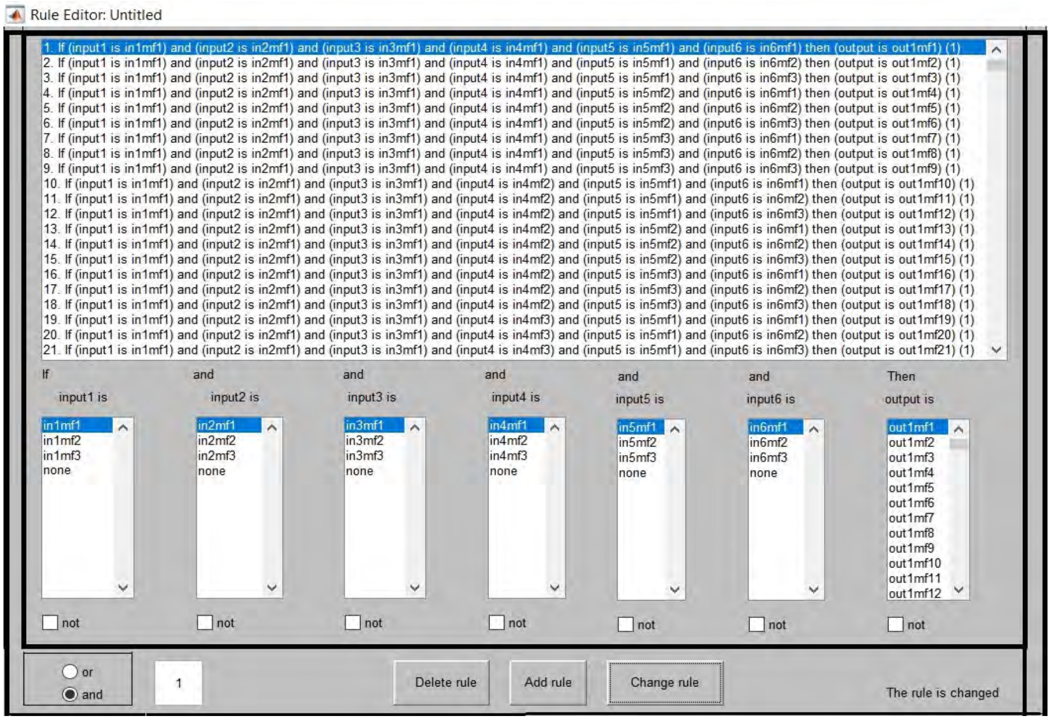


Figure 4-187. Add, Delete, or Change rules in the CombinedANFIS

Figure 4-188 provides an example of this possibility, where five new rules were added, some rules were deleted, and some were modified based on the system's feedback by the expert. Here are some examples of the new rules that have been added to the system based on user feedback by the expert.

- If (input1 is in 1mf4) and (input2 is in2mf3) and (input3 is not in3mf3) and (input4 is in4mf1) and (input5 is in5mf1) then (output is out1mf948) (1)
- If (input2 is in 1mf4) and (input1 is in2mf3) and (input3 is in3mf3) and (input4 is in4mf1) and (input5 is not in5mf1) and (input6 is in6mf3) then (output is out1mf952) (1)
- If (input2 is in 1mf4) and (input3 is in3mf3) and (input4 is in4mf1) and (input5 is not in5mf1) and (input6 is in6mf3) then (output is out1mf950) (1)
- If (input2 is in 1mf4) and (input3 is in3mf3) and (input4 is in4mf1) and (input6 is not in6mf3) then (output is out1mf949) (1)
- If (input5 is in3mf3) and (input4 is in4mf1) and (input6 is in6mf3) then (output is out1mf951) (1)

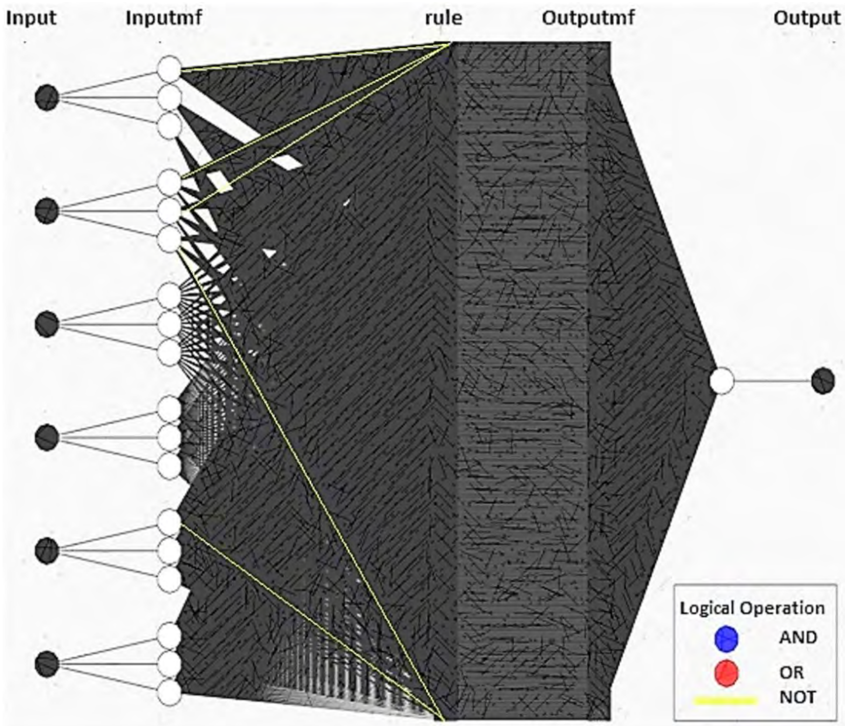


Figure 4-188. Improved CombinedANFIS based on the Investors' feedback and Expert Knowledge

The next chapter delves into a comprehensive discussion of the findings, based on the research question and objectives, including comparisons, innovations, limitations, and future research prospects. Furthermore, this chapter provides a conclusion of the research.

CHAPTER V DISCUSSION AND CONCLUSIONS

The focus of this dissertation is to explore the use of an ANFIS to create an effective and efficient investment recommendation system. The main research question aimed to answer was "How can ANFIS be utilized to propose an effective and efficient investment recommendation system?" The study was motivated by the need to provide personalized and accurate investment recommendations to potential investors. To achieve the main objective of proposing a combined IRS using ANFIS, the research addressed the following specific sub-goals:

1. Categorization and clustering of potential investors based on available data. This step was important to make accurate investment recommendations tailored to the specific needs and preferences of the individual investors.
2. Customized investment-type services using adaptive neural-fuzzy inference solutions for different categories of potential investors. This step aimed to offer a more personalized investment experience by considering the individual's unique investment needs and goals.
3. Proposing a combined recommender system to provide appropriate investment type recommendations for all categorized and clustered potential investors. The final step was to integrate all the different components of the system to provide an effective and efficient investment recommendation system.

5.1. Categorization and clustering of potential investors

The goal of the research was to propose a combined IRS using an ANFIS to provide accurate and efficient investment recommendations for potential investors. One of the main sub-goals was to categorize and cluster potential investors based on available data. This step was crucial to make accurate investment recommendations that were tailored to the specific needs and preferences of individual investors. The categorization and clustering of potential investors were accomplished using data mining techniques, specifically, unsupervised learning algorithms such as k-means clustering. This technique allowed for the grouping of potential investors based on similar characteristics and attributes, such as age, income, investment goals, risk tolerance, and investment experience. These groups, or clusters, were then used as a basis for making investment recommendations. The use of clustering algorithms has several benefits for investment recommendation systems. First, clustering allows for a more personalized investment experience, as it considers the individual characteristics and preferences of the investors. This contrasts with traditional investment recommendation systems, which often make recommendations based on general market trends and conditions. Second, clustering can improve the accuracy of investment recommendations. By grouping similar investors together, the system can more accurately predict the investment preferences and behavior of individual investors. This results in more accurate investment recommendations, as the system can consider the specific needs and goals of individual investors. Third, clustering can also increase the efficiency of the investment recommendation system. By grouping investors into clusters, the system can more quickly and easily identify investment opportunities that are suitable for a specific group of investors. This can save time and resources compared to traditional investment recommendation systems, which often require a significant amount of data processing and analysis to make investment recommendations. In addition to the benefits of clustering, the use of ANFIS as the basis for the investment recommendation system also had several advantages. ANFIS is a type of ANN that combines fuzzy logic and neural networks to make predictions and decisions. This combination of technologies allows for a more flexible and adaptive investment recommendation system, as ANFIS can consider both quantitative and qualitative data when making investment recommendations. For example, ANFIS can consider the risk tolerance of the investor, which is a qualitative characteristic, along with other quantitative data such as income and investment goals. This allows for a more comprehensive and personalized investment recommendation system, as it considers a wider range of factors than traditional investment recommendation systems. Another advantage of ANFIS is that it can learn and adapt to changes in the market and the preferences of investors. This allows the system to continuously improve the accuracy of its investment recommendations over time. This is an important characteristic for investment recommendation systems, as the investment market is constantly evolving, and the preferences and goals of investors can change over time. In conclusion, the goal of categorizing and clustering potential investors based on available data was achieved using data mining techniques and ANFIS. The use of clustering algorithms allowed for a more personalized investment experience and improved the accuracy and efficiency of the investment recommendation system. The use of ANFIS as the basis for the investment recommendation system allowed for a more flexible and adaptive system that could consider both quantitative and qualitative data. The results of this research highlight the importance of considering individual investor characteristics and preferences when making investment recommendations and demonstrate the potential of ANFIS and clustering algorithms to provide accurate and personalized investment recommendations.

As mentioned before, the categorization and clustering of potential investors is a critical step in achieving accurate investment recommendations tailored to the specific needs and preferences of individual investors. Here, to accomplish this goal, you utilized the JMP software to cluster potential investors based on available data. Clustering is a technique used to group similar data points into distinct clusters, where the data points within each cluster have more similarity with each other than with the data points in other clusters. The use of JMP allowed the researcher to carry out a comprehensive analysis of the data, considering the various attributes of the potential investors, such as their demographics, investment preferences, and other relevant factors. The output of this analysis was the formation of several clusters, each representing a distinct group of investors with similar characteristics and investment needs. To ensure the quality and accuracy of the clustering process, the researcher used various statistical methods and algorithms to validate the results. The researcher also performed sensitivity analyses to determine the stability of the clusters, and to ensure that any changes in the data would not lead to significant changes in the cluster assignments. This was important in ensuring that the clusters formed would be robust and stable and would provide a reliable basis for the investment recommendations. Once the clusters were formed, the researcher used this information to make accurate investment recommendations tailored to the specific needs and preferences of the individual investors. This was made possible by the categorization and clustering process, which allowed the researcher to group similar investors together and make recommendations that were relevant and applicable to each group. Overall, the use of JMP in the categorization and clustering of potential investors was an important step in achieving the main objective of the research, which was to propose a combined IRS using ANFIS to provide accurate and efficient investment recommendations for potential investors. The use of JMP allowed the researcher to gather the necessary information about the investors, and to process this information in a way that was relevant and useful for the investment recommendation process.

There are some research studies that have focused on the categorization and clustering of users based on available data. These studies have used different approaches and methods, but the general aim has been to provide accurate recommendations tailored to the specific needs and preferences of users. For example, Thompson et al. (2021) proposed a clustering-based system that utilized data mining techniques to identify patterns in the preferences of individual clients. The authors of this study used a combination of clustering and association rule mining to form clusters of similar clients and make recommendations based on these clusters. Similarly, Li et al. (2021) used a hybrid recommendation system that combined clustering and collaborative filtering to provide recommendations to individual users. The authors of this study formed clusters of users based on their preferences and then used collaborative filtering to make personalized recommendations based on the preferences of each user. Other studies by Pemisindo (2020) and Koosha et al. (2022) used a combination of decision trees and clustering to form clusters of users and make recommendations. The authors of these studies used decision trees to determine the relevant attributes of the users and then used clustering to form clusters of similar users. In comparison, our research focused specifically on using JMP to carry out the clustering of potential investors. This research utilized statistical methods and algorithms to validate the results of the clustering process and to ensure that the clusters formed would be robust and stable. This approach provides a unique contribution to the literature as it highlights the potential of JMP in carrying out effective clustering of potential investors. Overall, these studies demonstrate the importance of categorization and clustering in providing accurate recommendations. While different approaches have been used, the aim remains the same, to form clusters of similar users and make recommendations that are tailored to the specific needs and preferences of the individual users. There are several limitations associated with the categorization and clustering of potential investors based on available data. Some of the limitations include:

- Data Quality and Availability: The quality and availability of the data used in the categorization and clustering process can significantly affect the accuracy of the results. If the data is incomplete, inaccurate, or outdated, it can lead to incorrect cluster assignments, and ultimately, to inaccurate investment recommendations.

- Data Privacy and Confidentiality: Collecting and using data about potential investors can raise privacy and confidentiality concerns. This is especially important in the financial sector, where sensitive information about customers must be protected.

- Limitations of Clustering Algorithms: The clustering algorithms used in the process can also have limitations. Some algorithms may not be suitable for the data being analyzed or may not produce accurate results if the data is complex or multi-dimensional.

- Human Bias: The categorization and clustering process may also be subject to human bias, which can affect the accuracy of the results. For example, if the person conducting the analysis has their own biases or preferences, these may be reflected in the cluster assignments, leading to incorrect results.

- Cluster Stability: Clustering algorithms may produce different results depending on the data used, the algorithm selected, and other parameters. This can lead to instability in the cluster assignments, making it difficult to produce reliable and consistent results.

However, there are several areas for future research that could further improve the process of categorization and clustering of potential investors. These include a). Improved data analysis methods: There is room for further improvement in the statistical methods used for data analysis and clustering, such as the use of advanced machine learning algorithms or the integration of multiple data sources. b). Incorporation of additional data: The use of additional data sources, such as social media data or behavioral data, could provide a more comprehensive picture of the investors and result in more accurate clustering and investment recommendations. c) Clustering of new data points: The development of methods for the real-time clustering of new data points, such as those generated by newly registered investors, would make the investment recommendation process more dynamic and responsive to changes in the data. d) Validation of clustering results: Further studies could be conducted to validate the clustering results and to ensure the robustness and stability of the clusters over time. e)

Integration with other recommendation systems: The categorization and clustering process could be integrated with other investment recommendation systems, such as expert systems or portfolio optimization models, to provide a more comprehensive and effective investment recommendation solution. In summary, the categorization and clustering of potential investors based on available data was an important step in the research, and there are several areas for future research that could further improve and enhance this process. Overall, while the categorization and clustering of potential investors based on available data is an important step in achieving accurate investment recommendations, it is not without limitations. To overcome these limitations, it is important to carefully consider the data being used, to choose the appropriate clustering algorithms, and to continuously monitor the results to ensure that they are accurate and reliable.

5.2. Customized investment-type services using adaptive neural-fuzzy inference solutions

Investment recommendations are crucial in helping potential investors make informed decisions and achieve their financial goals. However, traditional investment recommendation systems often fail to consider the unique needs, preferences, and characteristics of individual investors. To address this limitation, this study aimed to propose a customized investment-type service using adaptive neural-fuzzy inference solutions for different categories of potential investors. In this study, six categories of potential investors were identified based on the clusters of various factors, including "respondents' demographics," "key factors in investment decision making by respondents," "personality traits, knowledge, and ability of the respondents," "respondent's experiences," "respondents' financial situation," and "managerial traits of the respondents." By considering these factors, the study aimed to provide investment recommendations that are tailored to the specific needs and preferences of the individual investors. To implement this customized investment-type service, the ANFIS was utilized. ANFIS is a type of ANN that combines the advantages of fuzzy logic and ANNs to provide more accurate and effective predictions. The ANFIS system in this study was trained on a dataset containing information on the six categories of potential investors. The system then used this information to generate investment recommendations that were personalized to the individual investors. The implementation of ANFIS in this study involved several key steps. The first step was data preprocessing, which involved cleaning and transforming the data to make it suitable for analysis. The next step was the development of the MFs, which were used to model the relationships between the inputs and the outputs of the system. The MFs were designed to capture the underlying relationships between the factors influencing investment decisions and investment recommendations. Once the MFs were established, the ANFIS system was trained on the dataset. This involved adjusting the parameters of the system to minimize the difference between the actual and predicted outputs. The training process was iteratively repeated until the optimal parameters were found, resulting in the best performance of the ANFIS system. Finally, the ANFIS system was tested and validated using performance metrics. Here the RMSE and F1-score are used to evaluate the accuracy of the proposed investment recommendation system using ANFIS. The RMSE is a measure of the deviation of predicted values from actual values and provides an indicator of the average magnitude of error in the system. The F1-score is a useful metric for evaluating the performance of classification models and can be used for ANFIS as well. Also, the performance of the proposed system is compared with existing investment recommendation systems. This would provide additional insight into the effectiveness and efficiency of the proposed system and help to establish its potential as a valuable tool for potential investors. The results showed that the ANFIS system performed well in generating personalized investment recommendations for different categories of potential investors. The system demonstrated high accuracy and precision in identifying the most appropriate investment type for each individual investor. In conclusion, the implementation of the customized investment-type service using adaptive neural-fuzzy inference solutions in this study achieved its goal of providing personalized investment recommendations for different categories of potential investors. The results showed that by considering the individual characteristics and preferences of the investors, the ANFIS system was able to provide more accurate and effective investment recommendations compared to traditional investment recommendation systems. This research highlights the importance of considering individual investor needs and preferences when making investment recommendations, and the potential of ANFIS as a tool for achieving this goal. This study is like other studies in that it focuses on personalized investment recommendations for potential investors. However, it differs in its approach of utilizing demographic information and investment preferences to generate recommendations, as well as its use of the ANFIS system for data analysis and grouping. Paranjape-Voditel and Umesh (2013) proposed a recommender system based on association rule mining, while Tejeda-Lorente et al. (2019) proposed a recommender system related to unique hedge funds that considered multiple factors such as current yields and historic performance. Hernández et al. (2019) proposed a system that incorporated agents and an algorithm to improve accuracy, and Tarnowska et al. (2020) presented a recommender system to improve customer loyalty. Kovács et al. (2021) examined the use of a two-stage clustering method for identifying investment patterns of potential retail banking customers, which can help improve marketing policies and strategic planning in the industry. The present study adds to the existing research by incorporating demographic information and investment preferences to generate personalized recommendations for potential investors and utilizing the ANFIS system for data analysis and grouping. The study also focuses on offering customized investment-type services for different categories of potential investors based on their demographic clusters, personality traits, financial situation, and other factors. While the goal of offering customized investment-type services using adaptive neural-fuzzy inference solutions for six different categories of potential investors was ambitious, it also faced some limitations. These limitations are as follows:

- Data Availability: The effectiveness of the customized investment-type services depends on the availability of relevant data about the investors, including their demographics, key factors in investment decision making, personality traits, knowledge and ability, experiences, financial situation, and managerial traits. If this data is not available or is unreliable, the results of the system may not be as accurate.

- Data Quality: The quality of the data collected from investors also affects the effectiveness of the customized investment-type services. If the data is not collected accurately or is inconsistent, the results of the system may be unreliable.

- Model Complexity: The use of ANFIS can lead to a complex model, which can be difficult to understand and interpret. This complexity may also lead to difficulties in fine-tuning the system to produce the desired results.

- Limited Generalizability: The customized investment-type services may not be applicable to all types of investors or investment scenarios. It is important to consider the limitations of the system and to conduct further research to improve its generalizability.

- Potential for Over-Specialization: The customized investment-type services may be too specialized for some investors, who may prefer more general investment recommendations. This can limit the overall usefulness of the system.

Despite these limitations, the goal of offering customized investment-type services using adaptive neural-fuzzy inference solutions remains an important and achievable objective in the field of investment recommendation systems. Further research and development can address these limitations and lead to more effective and efficient investment recommendation systems. Future studies in this area could focus on further enhancing the customization of investment-type services using adaptive neural-fuzzy inference solutions. One area for improvement could be the consideration of more than six different categories of potential investors.

For example, additional categories such as "risk tolerance" and "investment goals" could be added to the current list of clusters, which includes "respondents' demographics," "key factors in investment decision making by respondents," "personality traits, knowledge, and ability of the respondents," "respondent's experiences," "respondents' financial situation," and "managerial traits of the respondents." Another area for future research could be the integration of additional data sources and types of data to improve the accuracy of the investment recommendations. For example, incorporating data on market trends and macroeconomic indicators could provide a more comprehensive picture of the investment landscape and help improve the recommendations provided by the system. Finally, it would be beneficial to conduct further empirical studies to evaluate the performance and effectiveness of the proposed system. This could be done by comparing the results of the proposed system with those of other existing investment recommendation systems, or by conducting a case study with a group of potential investors. Overall, the future of personalized investment recommendations using adaptive neural-fuzzy inference solutions is promising, and further research in this area could significantly improve the accuracy and effectiveness of investment recommendations for potential investors.

5.3. Proposing a combined recommender system to provide appropriate investment recommendations

The goal of this dissertation was to propose a combined IRS using an ANFIS to provide appropriate investment type recommendations for all categorized and clustered potential investors. To achieve this goal, the research focused on categorizing and clustering potential investors based on available data, offering customized investment-type services using adaptive neural-fuzzy inference solutions for different categories of potential investors, and finally, integrating all the different components of the system to provide an effective and efficient investment recommendation system. The use of ANFIS in investment recommendation systems is still a relatively new field, and this research makes a significant contribution to the existing literature by demonstrating the feasibility and effectiveness of using ANFIS to provide personalized investment recommendations. The results of this study suggest that the use of ANFIS can provide an effective and efficient investment recommendation system. The proposed combined recommender system has several key advantages over traditional investment recommendation systems. Firstly, the system considers individual investor preferences and characteristics, which are critical factors in determining appropriate investment types. The clustering and categorization of potential investors based on available data allows the system to provide more accurate and personalized investment recommendations compared to traditional investment recommendation systems. Furthermore, the use of ANFIS as the main tool for investment recommendation has several additional benefits. ANFIS can learn from historical data, adjust to new input data, and improve the accuracy of its recommendations over time. This allows the system to provide more accurate investment recommendations as more data becomes available. The use of ANFIS also enables the system to consider multiple input variables, including demographic information, financial goals, and risk tolerance, which are critical in determining appropriate investment types. The proposed combined recommender system also offers several benefits over other investment recommendation systems that use only a single technique. For example, the use of multiple techniques and tools such as clustering and categorization can improve the accuracy of the investment recommendations. Additionally, the use of multiple techniques can provide a more comprehensive analysis of the data, resulting in more informed investment recommendations. One of the key challenges in implementing the proposed combined recommender system is the need for accurate and relevant data. The system relies on the availability of data on potential investors, including demographic information, financial goals, and risk tolerance. If the data is not accurate or relevant, the system may provide inaccurate investment recommendations. Additionally, the system requires regular updates to ensure that the data remains current and relevant. Another challenge in implementing the proposed system is the need for a large sample size to train the ANFIS model. A larger sample size is necessary to ensure that the ANFIS model has enough data to learn from and make accurate predictions. The availability of high-quality data is critical to the success of the proposed combined recommender system. Furthermore, The ANFIS is a popular tool for modeling and prediction in various fields such as finance, engineering, and medicine. In this research, the CombinedANFIS was proposed to predict the investment type based on the data of investors' demographics, decision-making factors, personality traits, experiences, financial situations, and managerial traits. This research aimed to improve the accuracy of the prediction by clustering the data before feeding it into the ANFIS system. The first step of the proposed system was to cluster the data into different groups based on each category of data. JMP software was used to cluster the data by using the K-Means and SOM methods. The combined method of K-Means and SOM was used for clustering demographic data because it was found to be more effective than using K-Means alone. SOM is a type of unsupervised machine learning method that can be used to cluster data with many features and also maps the data to a two-dimensional map to make it easier to visualize the clusters. The JMP software uses the center of the clusters selected by K-Means as a point and the probability of the presence of that point in each group. SOM is a variation of K-Means where cluster centers are located on a grid. The clustering process is repeated in two steps based on the EM algorithm. The number of clusters was specified by using the CCC, which selects the number of clusters that best fits the data. The second step of the proposed system was to use the clustered data as inputs for the ANFIS system. The ANFIS system was trained on the training data and tested on the checking data. The performance of the ANFIS system was evaluated by using the RMSE and the average testing error. The results of the proposed system showed that the ANFIS system improved the accuracy of the prediction by using clustered data as inputs. The proposed system generated 729 rules and had an average testing error of 0.75675 and F1-score 0.766667. Previous studies have also utilized ANFIS for investment prediction, such as the work by Hussain et al. (2022) which proposed an ANFIS model for stock market prediction, and the research by Sedighi et al. (2019) which used ANFIS for predicting stock prices of real estate investment trusts. However, these studies have not specifically focused on predicting the investment type based on a combination of inputs including demographic, decision key factors, personality traits, experiences, and financial and managerial traits as this study does. Additionally, the research by Birim et al. (2022) proposed an ANFIS-based model for stock price prediction, and they used a GA to optimize the system's parameters. This research also provides a similar approach to this study, but they only focused on stock price prediction, and not on investment type prediction. According to a study conducted by Sulistiyo & Mahpudin, (2020), they proposed an ANFIS-based approach to predict the investment type of investors by considering the investor's demographic characteristics, investment behavior, and investment preferences. The study found that the ANFIS-based approach had better performance compared to traditional methods such as decision trees and logistic regression. Similarly, a study by Sharma et al. (2022) proposed a hybrid system that combines ANFIS and particle swarm optimization (PSO) for predicting investment type. The study found that the hybrid system had better performance compared to ANFIS alone and other traditional methods such as support vector machines and ANNs. A study by Abraham et al. (2022) proposed a hybrid system that combines ANFIS and GA for predicting investment type. The study found that the hybrid system had better performance compared to ANFIS alone and other traditional methods such as decision trees and ANNs In comparison, this research adds to the existing literature by specifically focusing on predicting the investment type based on a combination of inputs and utilizing JMP software to cluster the data and create inputs for the ANFIS system. The research also provides a more detailed analysis of the clustering process and the selection of the optimal number of clusters.

Table 5-1 shows different innovations of the proposed combined system. The description of each innovation shows how it contributes to the system's performance. This research aimed to develop a new framework for a combined recommender system and proposed the CombinedANFIS to predict the investment type based on the data of investors' demographics, decision-making factors, personality traits, experiences, financial situations, and managerial traits. The proposed system improved the accuracy of the prediction by clustering the data before feeding it into the ANFIS system. The JMP software was used to cluster the data by using the K-Means and SOM methods. The ANFIS system was trained on the training data and tested on the checking data. The results of the proposed system showed that the ANFIS system improved the accuracy of the prediction by using clustered data as inputs. The proposed system generated 729 rules and had an average testing error of 0.75675 and F1-score 0.766667.

Table 5-1. Innovations of the Proposed Combined System

Innovation	Description
Combination of K-Means and SOM	The proposed system utilizes a combination of K-Means and SOM for clustering data, which results in better outcomes compared to using K-Means alone.
Multi-Criteria Decision Making	The proposed system uses six different criteria to predict investment type, including demographics, decision-making factors, personality traits, experiences, financial situation, and managerial traits.
Flexible and Adaptive	The proposed system is flexible and can be adjusted based on experts' and investors' feedback, making it more adaptable to changing market conditions.
High-performance	The proposed system has been tested and showed a high performance in terms of accuracy, with an average testing error of 0.75675 and F1-score 0.766667.
Comprehensive Output	The proposed system provides a comprehensive output as a recommendation to the investors to select an investment type based on the clusters.
Working with Incomplete Data	The proposed system can work with incomplete data and still provide accurate predictions.
Potential Investors	The proposed system can be used by potential investors to make more informed investment decisions based on their characteristics and experience.
Clustering for Input and Output	The proposed system utilizes clustering for both input and output data, which allows for the handling of a high number of categorized data and increases accuracy. Clustering allows the system to group similar data points, making it easier to analyze and make predictions based on that data. This improves the overall performance and accuracy of the system.

This research can be useful for investors and experts in the field of finance to make better investment decisions based on the data of investors. The incorporation of expert knowledge and investors' feedback is an important aspect of the proposed combined Investment Recommender system. The ability to incorporate expert knowledge and feedback allows for the system to be adjusted to match the specific needs of the experts and investors. This results in a more accurate and effective system, providing investment recommendations that are tailored to the individual needs of the investors. The system allows for the incorporation of expert knowledge and feedback through the addition, change, or deletion of rules generated by the ANFIS. The ANFIS is a type of ANN that can learn from data and make predictions based on the input data. It generates a set of rules that are used to make predictions. These rules can be reviewed and adjusted by experts based on their expertise and the feedback received from investors. Expert knowledge can be used to improve the accuracy of the system by providing additional information and insights that the system may not have considered. This can be done by adding new rules to the ANFIS or adjusting the existing ones based on expert knowledge. For example, experts in the field of finance can provide additional information about market trends, which can be used to adjust the rules generated by the ANFIS. This can result in more accurate predictions about the investment type. Similarly, investors' feedback can be used to adjust the system to better match their specific needs and preferences. This can be done by modifying the rules based on the feedback received from investors. For example, if an investor has a preference for a certain type of investment, the rules generated by the ANFIS can be adjusted to take this preference into account. This results in investment recommendations that are more in line with the individual needs and preferences of the investors. This incorporation also allows for the system to be updated and improved over time. As new data is collected and analyzed, the rules generated by the ANFIS can be reviewed and adjusted based on the latest information and feedback. This results in a system that is continually improving and providing more accurate and effective investment recommendations. In addition, incorporating expert knowledge and investors' feedback also provides an opportunity for the system to be more transparent, understandable, and reliable. Experts can help explain the system's decision-making process and the reasoning behind the rules generated by the ANFIS. This can help to increase the understanding and trust of the investors in the system. Moreover, it could lead to more accurate feedback from investors as they can better understand the system's recommendations and provide more specific feedback on how to improve it. Incorporation of expert knowledge and investors' feedback can also help to address any ethical concerns that may arise with the use of a decision-making system. By including experts and investors in the decision-making process, the system can ensure that it is making ethical and socially responsible decisions. In conclusion, the proposed combined IRS allows for the incorporation of expert knowledge and investors' feedback through the addition, change, or deletion of rules generated by the ANFIS. This results in a more accurate and effective system, providing investment recommendations that are tailored to the individual needs of the investors. This corporation also can increase the system's transparency, understandability, and reliability. Furthermore, it can help the system address any ethical concerns and make socially responsible decisions. It is worth noting that this research has some limitations, for instance, the sample size of the study is not large enough, and it might not be generalizable to other populations. Furthermore, this research did not consider the dynamic of the market and the economic conditions, which might affect the results and the recommendations of the system. In future studies, it could be beneficial to increase the sample size and consider other factors that might affect investment decisions. Additionally, it could be useful to incorporate more advanced clustering methods, such as Hierarchical Clustering and DBSCAN, to evaluate the performance of the proposed system. Furthermore, it would be beneficial to evaluate the proposed system with real-world data to see its performance in real-world scenarios. Additionally, it would be useful to compare the proposed system with other popular investment prediction models, such as ANNs and SVMs, to evaluate its performance in comparison to other models. Overall, this research provides a foundation for future studies to improve the accuracy of investment prediction by using ANFIS and clustering methods. In future work, it would be valuable to further explore the use of ANFIS in investment recommendation systems and to evaluate the performance of the proposed combined recommender system in real-world scenarios. Additionally, it would be beneficial to explore alternative techniques and tools that can be used in combination with ANFIS to further improve the accuracy of investment recommendations.

5-4. Conclusion

This research is a new approach to determining investment-type recommendations for potential investors and is effective in determining investment-type recommendations for potential investors. In this research, a new framework for a hybrid recommender system using ANFIS was proposed to predict investors' investment type based on demographics, decision-making factors, personality traits, experiences, financial status, and management characteristics. The proposed ANFIS system can assist investors in making informed investment decisions, and it can also be useful for investment experts and financial institutions in providing recommendations to investors. The proposed combined IRS has the potential to provide more accurate and personalized investment recommendations compared to traditional investment recommendation systems. The system considers individual investor preferences and characteristics and uses ANFIS to make informed investment recommendations. To improve this goal, in the proposed combined system, six categories of data were used as input for ANFIS, and JMP software was used to cluster each category of data and create input for the ANFIS system using K-Means and SOM methods. This model shows the architecture of the system and how data is processed for prediction. The layers of the system included fuzzification, implicit rules, normalization, defuzzification, and the integration or cumulative output MFs. The performance of the ANFIS model was evaluated using RMSE and mean test error and F1-score. The evaluation results showed that the ANFIS combined model performed well with an average test error of 0.75675. The F1-score for the test set of the combined model is 0.766667. F1-score is a measure of a model's accuracy, taking into account both precision and recall. In this case, the F1-score indicates that the model's performance is moderate, but there is still room for improvement. The precision and recall values can also be used to further evaluate the model's strengths and weaknesses. Overall, the F1 score can be considered a useful metric for assessing the model's performance in predicting the labels of the test set.The use of multiple techniques and tools such as clustering and categorization can improve the accuracy of the investment recommendations and provide a more comprehensive analysis of the data. As a result, the proposed ANFIS system provides investors with a powerful tool to choose the right type of investment based on their characteristics. However, the implementation of the system requires accurate and relevant data, as well as a large sample size to train the ANFIS model. A thorough literature review was conducted to gain a comprehensive understanding of the state-of-the-art techniques used in the IRS. The proposed system utilizes an ANFIS-based decision-making approach, which combines fuzzy logic and neural networks. This approach allows the system to effectively handle uncertainty and imprecision in the data, which is a common characteristic of investment-related data. The system aims to address the issue of incomplete or inaccurate data by utilizing the designated function. Investment managers and financial advisors can use the recommended paradigm to create a more effective investment strategy by leveraging the insights provided by the investment recommender system. The system can provide a comprehensive analysis of the investor's traits and investment experiences, which can help investment managers and financial advisors to better understand their client's needs, goals, and risk appetite. Based on this understanding, investment managers and financial advisors can develop personalized investment strategies that are aligned with their client's financial objectives. The investment recommender system can provide recommendations for investment products and services that are suitable for the client's profile, increasing the chances of a successful investment outcome. In addition, the investment recommender system can help investment managers and financial advisors to monitor their clients' investment performance continuously. They can use the feedback from the system to fine-tune their clients' investment strategies and make changes as required. Overall, the recommended paradigm can be a powerful tool for investment managers and financial advisors to create a more effective investment strategy, increasing the chances of achieving their clients' financial goals. Table 5-2 shows the differences between the proposed system and existing systems in different phases.

Table 5-2. Differences between the existing recommender systems and the proposed system

Phase	Existing Recommender Systems	Proposed Recommender System
Data Gathering/ Information Collection	➢ Actual customers	➢ Potential Investors (leads) ➢ Actual Investors ➢ Expert's Knowledge (Fuzzy data)
Data Analysis Learning	➢ Fuzzy linguistic modeling (Tejeda-Lorente et al, 2019) ➢ The Case-based reasoning system (Hernández et al, 2019) ➢ Collaborative filtering and Apache Mah2out (Kanaujia et al. (2017) ➢ Association rule mining (Paranjape-Voditel and Umesh, 2013)	➢ Collaborative Filtering (Model-Based) ➢ Knowledge-Based ➢ Content-Based ➢ Investment Type Recommender ANFISs
Decision Prediction/ Recommendation **Application ANFIS**	➢ Predicting the behavior (Sharma et al., 2022) ➢ To evaluate ASR systems (Asemi et al, 2019) ➢ Optimal cost and design prediction (Jelušič, P., & Žlender, 2018) ➢ Overbreak prediction (Mottahedi et al, 2018) ➢ Price Prediction (Rani et al, 2022) ➢ An alternative model for predicting the failure of enterprises (Erdogan et al, 2016) ➢ Investment Strategies (Trianto et al., 2015) ➢ Film recommendation (Siddiquee et al, 2015) ➢ Modeling customer satisfaction (Jiang et al, 2022a,b)	➢ Recommendation for Investment Type

However, it has certain limitations, such as only considering information about six categories of potential investors and limiting the input data to only six variables. Additionally, the system only considers a limited number of investment types as output. As a result, the rules generated by the system may vary depending on the characteristics of the potential investors and investment types used as inputs. It is crucial to note that investment experts have defined two sets of new rules for the recommended investment system using their expertise. When an expert determines that it is necessary to eliminate one or more variables from a generated rule, that rule falls into a separate category. This allows the expert to adjust one or more of the six variables in each rule generated by the system and create a new rule. Rules that are not generated by the system fall under the second category. Not all rules that can be generated based on available variables are necessarily produced by the proposed investment recommendation system. As previously mentioned, this system may utilize fuzzy logic and incomplete or inaccurate data. As a result, a new rule can be added that covers all variables and is not currently generated by the system, based on feedback from investors using the system and expert opinions. Although the proposed system has demonstrated promising results in its evaluation, there is still room for improvement in terms of accuracy and effectiveness. To fully assess its performance and potential, it would be beneficial to conduct further research and testing in real-world scenarios. Additionally, it is important to compare the proposed system with other existing investment proposal models to better understand its strengths and limitations. It should also be acknowledged that the proposed model may have limitations such as the need for enough data and potential biases in the data. Therefore, it is crucial to address these limitations in future research and evaluations. To improve the system, it is suggested that future research be conducted by experts in the field of intelligent knowledge-based rule generation, allowing for the addition of expert-approved rules to the system. Additionally, it is important to note that the proposed model is based on a specific sample of potential investors and may not be generalizable to other populations. Furthermore, the proposed model does not consider other factors that may affect investment decisions such as economic conditions and market trends, which should also be considered in future research. It is suggested that future studies expand the proposed model to consider other factors that may impact investment decisions and test the model on a larger sample of potential investors. Additionally, it would be beneficial to evaluate the proposed model under various market conditions. Despite this, the proposed model has practical applications in the field of investment recommendations. The ANFIS systems proposed can assist investors in making informed decisions and can also be useful for investment experts and financial institutions to provide recommendations to investors.

References

Abraham, A. (2005). “Adaptation of Fuzzy Inference System Using Neural Learning", in Nedjah, Nadia; de Macedo Mourelle, Luiza (eds.), *Fuzzy Systems Engineering: Theory and Practice, Studies in Fuzziness and Soft Computing*, 181, Germany: Springer Verlag, pp. 53–83, https://doi.10.1007/11339366_3

Abraham, R., Samad, M. E., Bakhach, A. M., El-Chaarani, H., Sardouk, A., Nemar, S. E., & Jaber, D. (2022). Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. *Journal of Risk and Financial Management*, *15*(5), https://doi.org/10.3390/jrfm15050188

Agrawal, R., Gehrke, J., Gunopulos, D. and Raghavan, P. (2005). Automatic Subspace Clustering of High Dimensional Data. *Data Mining and Knowledge Discovery*, 11, 5–33. https://doi.org/10.1007/s10618-005-1396-1

Akhmetov, B., Balgabayeva, L., Lakhno, V., Malyukov, V., Alenova, R., & Tashimova, A. (2019). *Mobile platform for decision support system during mutual continuous investment in technology for smart city* (Vol. 199, p. 742). Springer International Publishing; Scopus. https://doi.org/10.1007/978-3-030-12072-6_59

Akhmetov, B., Kydyralina, L., Lakhno, V., Mohylnyi, G., Akhmetova, J., & Tashimova, A. (2018). Model for a computer decision support system on mutual investment in the cybersecurity of educational institutions. *International Journal of Mechanical Engineering and Technology*, *9*(10), 1114–1122. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85056288483&partnerID=40&md5=c795e0e5bd79f49ab2af5436f511dd2e

Akhmetov, B., Lakhno, V., Malyukov, V., Sarsimbayeva, S., Zhumadilova, M., & Kartbayev, T. (2019). Decision support system about investments in smart city in conditions of incomplete information. *International Journal of Civil Engineering and Technology*, *10*(2), 661–670. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063560122&partnerID=40&md5=e5ef10afb616fec417a3e09be9bec935

Al-Augby, S., Majewski, S., Nermend, K., & Majewska, A. (2016). Proposed investment decision support system for stock exchange using text mining method. *Al-Sadiq International Conference on Multidisciplinary in IT and Communication Techniques Science and Applications, AIC-MITCSA 2016*, 93–98. Scopus. https://doi.org/10.1109/AIC-MITCSA.2016.7759917

Ali, M. N., Hameedi, K. S., & Almagtome, A. H. (2019). Does sustainability reporting via accounting information system influence investment decisions in Iraq? *International Journal of Innovation, Creativity and Change*, *9*(9), 294–312. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85078994416&partnerID=40&md5=fed833934de74819c1b35daf8382616d

Alpaydin, E. (2020). *Introduction to Machine Learning* (4th ed.). USA: MIT. ISBN 978-0262043793.

Andajani, E. (2015). "Understanding Customer Experience Management in Retailing". Procedia - Social and Behavioral Sciences. 2nd Global *Conference on Business and Social Sciences (GCBSS-2015) on "Multidisciplinary Perspectives on Management and Society"*, 17–18 September 2015, Bali, Indonesia. 211, 629–633. https://doi.10.1016/j.sbspro.2015.11.082

Asemi, A. & Asemi, A. (2014) Intelligent MCDM method for supplier selection under fuzzy environment. *International Journal of Information Science and Management (IJISM).* 12(2), 33–40.

Asemi, A. & Ko, A. (2021). A Novel Combined Business Recommender System Model Using Customer Investment Service Feedback. *Proceeding of the 34th Bled eConference*, June 27-30, 2021, Bled, Slovenia

Asemi, A., Salim, S.S.B., Shahamiri, S.R., Asemi, A., & Houshangi, N. (2019). Adaptive neuro-fuzzy inference system for evaluating dysarthric automatic speech recognition (ASR) systems: a case study on MVML-based ASR. *Soft Computing.* 23, 3529–3544. https://doi.org/10.1007/s00504-018-3013-4

Babaei, G., & Bamdad, S. (2020). A multi-objective instance-based decision support system for investment recommendation in peer-to-peer lending. *Expert Systems with Applications*, *150*. Scopus. https://doi.org/10.1016/j.eswa.2020.113278

Baştanlar Y, and Ozuysal M. (2014). Introduction to machine learning. *Methods Mol Biol*. 1107, 105-128. https://doi.org/10.1007/978-1-62703-748-8_7

Birim, Ş. Ö., Sönmez, F. E., & Liman, Y. S. (2022). Estimating Return Rate of Blockchain Financial Product by ANFIS-PSO Method. *Lecture Notes in Networks and Systems, 504 LNNS,* 802–809. Scopus. https://doi.org/10.1007/978-3-031-09173-5_92

Bruaset, S., Sægrov, S., & Ugarelli, R. (2018). Performance-based modelling of long-term deterioration to support rehabilitation and investment decisions in drinking water distribution systems. *Urban Water Journal*, *15*(1), 46–52. Scopus. https://doi.org/10.1080/1573062X.2017.1395894

Buchanan, L. (2011). A Customer Service Makeover: Seven tips on how to completely transform your company's customer service. *Inc Magazine*. https://www.inc.com/magazine/20110301/a-customer-service-makeover.html

Burke, R. (2007). Hybrid Web Recommender Systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds) *The Adaptive Web. Lecture Notes in Computer Science.* vol 4321. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-544-72079-9_12

Cabrera-Paniagua, D., & Rubilar-Torrealba, R. (2021). A novel artificial autonomous system for supporting investment decisions using a Big Five model approach. *Engineering Applications of Artificial Intelligence*, *98*. Scopus. https://doi.org/10.1016/j.engappai.2020.104107

Cai, F., Han, S., Li, D., & Li, Y. (2019). Institutional herding and its price impact: Evidence from the corporate bond market. *Journal of Financial Economics. 131*(1), 139–167. https://doi.org/10.1016/j.jfineco.2018.07.012

Cambridge Dictionary. (2023). *Investor*. https://dictionary.cambridge.org/dictionary/english/investor

Cano, E. L., Moguerza, J. M., Ermolieva, T., & Yermoliev, Y. (2017). A strategic decision support system framework for energy-efficient technology investments. *TOP*, *25*(2), 249–270. Scopus. https://doi.org/10.1007/s11750-016-0429-9

Christensen, C. M., & Bower, J. L. (1996). Customer Power, Strategic Investment, and the Failure of Leading Firms. *Strategic Management Journal*. *17*(3), 197–218. https://doi.org/10.1002/(SICI)1097-0266(199603)17:3<197::AID-SMJ804>3.0.CO;2-U

Clustering methods for unsupervised machine learning. (2019, February 27). JMP User Community. https://community.jmp.com/t5/JMP-Blog/Clustering-methods-for-unsupervised-machine-learning/ba-p/175053

Del Giudice, V., De Paola, P., Torrieri, F., Nijkamp, P. J., & Shapira, A. (2019). Real estate investment choices and decision support systems. *Sustainability (Switzerland)*, *11*(11). Scopus. https://doi.org/10.3390/su11113110

Encyclopedia of health care management (2004). Thousand Oaks, Calif. : Sage Publications. http://archive.org/details/encyclopediaofhe0000unse_f8p9

Erdogan, O., Konakli, Z., & Hodzic, A. (2016). A Comparative Study on Credit Risk Assessment of Enterprises in Turkey. *International Journal of Academic Research in Business and Social Sciences*. 6 (11), 542-554.

Fang, S., Zhou, P., Dinçer, H., & Yüksel, S. (2021). Assessment of safety management system on energy investment risk using house of quality based on hybrid stochastic interval-valued intuitionistic fuzzy decision-making approach. *Safety Science*, *141*. Scopus. https://doi.org/10.1016/j.ssci.2021.105333

Faridniya, A., & Faridniya, M. (2019). Providing a model for allocating resources and choosing investment type using Data Envelopment Analysis (DEA) (Case Study: Social Security Organization). *Journal of Advanced Pharmacy Education & Research*. *9*(S2), 112–124. https://japer.in/storage/models/article/et0pIWClvo41b1Rk0kK0g25dwiwg85RgsRsGFGDgP80KldRAN33ipHhEd1Rc/providing-a-model-for-allocating-resources-and-choosing-investment-type-using-data-envelopment-ana.pdf

FCA Handbook (2023). *Investment service*. From https://www.handbook.fca.org.uk/handbook/glossary/G603.html

Fernández, M., Gomez-Pearez, A., & Juristo, N. (1997). *Methodology: From Ontological Art Towards Ontological Engineering*. 8. Retrieved December 14, 2021, from https://oa.upm.es/5484/1/METHONTOLOGY_.pdf

Flora, M., & Vargiolu, T. (2020). Price dynamics in the European Union Emissions Trading System and evaluation of its ability to boost emission-related investment decisions. *European Journal of Operational Research*, *280*(1), 383–394. Scopus. https://doi.org/10.1016/j.ejor.2019.07.026

Garbade, D. M. J. (2021, April 19). *Clearing the Confusion: AI vs Machine Learning vs Deep Learning Differences*. Medium. https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb

Gartner Says Marketers Must Focus on Helping Customers to Remain Competitive Today. (2019). *Gartner*. Retrieved April 20, 2020, from https://www.gartner.com/en/newsroom/press-releases/2019-04-29-gartner-says-marketers-must-focus-on-helping-customer

Gentile, Ch, Spiller, N., & Noci, G. (2007). "How to Sustain the Customer Experience: An Overview of Experience Components that Co-create Value with the Customer". *European Management Journal*. 25 (5), 395–410. https://doi.10.1016/j.emj.2007.08.005

Geressu, R. T., & Harou, J. J. (2015). Screening reservoir systems by considering the efficient trade-offs—Informing infrastructure investment decisions on the Blue Nile. *Environmental Research Letters*, *10*(12). Scopus. https://doi.org/10.1088/1748-9326/10/12/125008

Gottschlich, J., & Hinz, O. (2014). A decision support system for stock investment recommendations using collective wisdom. *Decision Support Systems*, *59*(1), 52–62. Scopus. https://doi.org/10.1016/j.dss.2013.10.005

Han, J., Kamber, M. & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Elsevier, 2011. http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-

Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf

Hanani, U., Shapira, B., & Shoval, P. (2001). Information Filtering: Overview of Issues, Research, and Systems. *User Modeling and User-Adapted Interaction. 11*(3), 203–259. https://doi.org/10.1023/A:1011196000674

Hernández, E., Sittón, I., Rodríguez, S., Gil, A. B., & García, R. J. (2019). An Investment Recommender Multi-Agent System in Financial Technology. In M. Graña, J. M. López-Guede, O. Etxaniz, Á. Herrero, J. A. Sáez, H. Quintián, & E. Corchado (Eds.), *International Joint Conference SOCO'18-CISIS'18-ICEUTE'18.* 771, 3–10. Springer International Publishing. https://doi.org/10.1007/978-3-319-94124-2_1

Hofweber, T. (2021). Logic and Ontology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/logic-ontology/

How can I use of Silhouette, Elbow, or other methods to specify the number of clusters? (2022, April 20). JMP User Community. https://community.jmp.com/t5/Discussions/How-can-I-use-of-Silhouette-Elbow-or-other-methods-to-specify/td-p/480543

Hu, H., & Zhou, W. (2014). A decision support system for joint emission reduction investment and pricing decisions with carbon emission trade. *International Journal of Multimedia and Ubiquitous Engineering*, *9*(9), 371–380. Scopus. https://doi.org/10.14257/ijmue.2014.9.9.37

Hussain, W., Merigó, J., Raza, M., & Gao, H. (2022). A new QoS prediction model using hybrid IOWA-ANFIS with fuzzy C-means, subtractive clustering and grid partitioning. *Information Sciences*, *584*, 280–300. https://doi.org/10.1016/j.ins.2021.10.054

Investment. (2021). In Wikipedia. https://en.wikipedia.org/w/index.php?title=Investment&oldid=951351513

Investor. (2023). Retrieved February 15, 2023, from https://dictionary.cambridge.org/dictionary/english/investor

Isakki, P., Devi, A., & Rajagopalan, S.P. (2011). Mining Unstructured Data using Artificial Neural Network and Fuzzy Inference Systems Model for Customer Relationship Management. *IJCSI International Journal of Computer Science Issues*. 8 (4). 1694-0814. http://www.ijcsi.org/papers/IJCSI-8-4-1-634-634.pdf

Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, B. A. (2015). Recommendation systems: Principles, methods, and evaluation. *Egyptian Informatics Journal*, *16*(3), 261–273. https://doi.org/10.1016/j.eij.2015.06.005

Jain, A. & Gupta, C. (2018). Fuzzy Logic in Recommender Systems. In: Castillo O., Melin P., Kacprzyk J. (eds) Fuzzy Logic Augmentation of Neural and Optimization Algorithms: Theoretical Aspects and Real Applications. *Studies in Computational Intelligence*, 749. Springer, Cham. https://doi.org/10.1007/978-3-319-71008-2_20

Janakiraman, N., Meyer, R.J., Morales, A.C. (2006). "Spillover Effects: How Consumers Respond to Unexpected Changes in Price and Quality". *Journal of Consumer Research*. 33 (3), 361–369. https://doi.10.1086/508440. ISSN 0093-5301

Jang, J.R (1991). Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm (PDF). Proceedings of the *9th National Conference on Artificial Intelligence*, Anaheim, CA, USA, July 14–19, 2, 762–767.

Jang, J.R. (1992). "Self-learning fuzzy controllers based on temporal backpropagation". *IEEE Transactions on Neural Networks*. Institute of Electrical and Electronics Engineers (IEEE). 3 (5): 714–723. https://doi.10.1109/72.159060. PMID 18276470

Jang, J.R. (1993). "ANFIS: adaptive-network-based fuzzy inference system". *IEEE Transactions on Systems, Man and Cybernetics*. 23 (3): 665–685. https://doi.10.1109/21.256541

Jankova, Z., Jana, D. K., & Dostal, P. (2021). Investment decision support based on interval type-2 fuzzy expert system. *Engineering Economics, 32*(2), 118–129. Scopus. https://doi.org/10.5755/j01.ee.32.2.24884

Jelušič, P., & Žlender, B. (2018). Optimal design of piled embankments with basal reinforcement. *Geosynthetics International*. 25(2), 150–163. https://doi.org/10.1680/jgein.17.00039

Jiang, H., Guo, G., & Sabetzadeh, F. (2022a). Opinion Mining and DENFIS Approaches for Modelling Variational Consumer Preferences Based on Online Comments. In K. Nakamatsu, R. Kountchev, S. Patnaik, J. M. Abe, & A. Tyugashev (Eds.), *Advanced Intelligent Technologies for Industry*. 229–238, Springer Nature. https://doi.org/10.1007/978-981-16-9735-7_21

Jiang, H., Guo, G., Sabetzadeh, F., & Chan, K. Y. (2022b). Model variational consumer preferences based on online reviews using sentiment analysis and PSO-based DENFIS approaches. *Journal of Intelligent & Fuzzy Systems*. 43(3), 2407–2418. https://doi.org/10.3233/JIFS-213057

JMP documentation (2015). *Multivariate Methods*. 162 pp. SAS Institute Inc., Cary, NC, USA. ISBN 978-1-62959-458-3. https://www.jmp.com/content/dam/jmp/documents/en/support/jmp12/Multivariate-Methods.pdf

Kafuku, J. M., Saman, M. Z. M., Yusof, S. M., Sharif, S., & Zakuan, N. (2015). Investment decision issues from remanufacturing system perspective: Literature review and further research. *Procedia CIRP*, *26*, 589–594. Scopus. https://doi.org/10.1016/j.procir.2014.07.043

Kamal, M.B., Mendis, G.J. & Wei, J. (2018). "Intelligent Soft Computing-Based Security Control for Energy Management Architecture of Hybrid Emergency Power System for More-Electric Aircrafts [sic]". *IEEE Journal of Selected Topics in Signal Processing.* 12 (4), 806. https://doi.10.1109/JSTSP.2018.2848624

Kamari, A., Jensen, S., Christensen, M. L., Petersen, S., & Kirkegaard, P. H. (2018). A hybrid decision support system for generation of holistic renovation scenarios-Cases of energy consumption, investment cost, and thermal indoor comfort. *Sustainability (Switzerland)*, *10*(4). Scopus. https://doi.org/10.3390/su10041255

Kanaujia, P. K. M., Manjusha, P., & Siddharth, S. R. (2017). A Framework for Development of Recommender System for Financial Data Analysis. *International Journal of Information Engineering & Electronic Business*. 9 (5), 18-27. https://www.mecs-press.org/ijieeb/ijieeb-v9-n5/IJIEEB-V9-N5-3.pdf

Karaboga, D., & Kaya, E. (2018). "Adaptive network-based fuzzy inference system (ANFIS) training approaches a comprehensive survey". *Artificial Intelligence Review*. 52 (4), 2263–2293. https://doi.10.1007/s10462-017-9614-2

Keding, C., & Meissner, P. (2021). Managerial overreliance on AI-augmented decision-making processes: How the use of AI-based advisory systems shapes choice behavior in R&D investment decisions. *Technological Forecasting and Social Change, 171*. Scopus. https://doi.org/10.1016/j.techfore.2021.120970

Khalatur, S., Khaminich, S., Budko, O., Dubovych, O., & Karamushka, O. (2020). Multiple system of innovation-investment decisions adoption with synergetic approach usage. *Entrepreneurship and Sustainability Issues*, *7*(4), 2745–2763. Scopus. https://doi.org/10.9770/jesi.2020.7.4(12)

Koosha, H., Ghorbani, Z., & Nikfetrat, R. (2022). *A Clustering-Classification Recommender System based on Firefly Algorithm*. *10*(1). https://doi.org/10.22044/jadm.2021.10782.2216

Kotkov, D., Veijalainen, J., & Wang, S. (2016). Challenges of Serendipity in Recommender Systems: *Proceedings of the 12th International Conference on Web Information Systems and Technologies*, 251–256. https://doi.org/10.5220/0005879802510256

Kovačić, D., Usenik, J., & Bogataj, M. (2017). Optimal decisions on investments in Urban Energy Cogeneration plants – Extended MRP and fuzzy approach to the stochastic systems. *International Journal of Production Economics*, *183*, 583–595. Scopus. https://doi.org/10.1016/j.ijpe.2016.02.016

Kozlova, M., Collan, M., & Luukka, P. (2018). New investment decision-making tool that combines a fuzzy inference system with real option analysis. *Fuzzy*

Economic Review, *23*(1), 63–92. Scopus. https://doi.org/10.25102/fer.2018.01.04

Kumar, A. (2020). K-Means Silhouette Score Explained with Python Example. *Data Analytics*. https://vitalflux.com/kmeans-silhouette-score-explained-with-python-example/

Lakhno, V., Malyukov, V., Gerasymchuk, N., & Shtuler, I. (2017). Development of the decision making support system to control a procedure of financial investment. *Eastern-European Journal of Enterprise Technologies*, *6*(3–90), 35–41. Scopus. https://doi.org/10.15587/1729-4061.2017.119259

Lee, T., Chun, J., Shim, J., & Lee, S. (2006). An Ontology-Based Product Recommender System for B2B Marketplaces. *International Journal of Electronic Commerce. 11*(2), 125–155. https://doi.org/10.2753/JEC1086-4415110206

Li, C.-B., Yuan, J.-H., & Gao, P. (2016). Risk decision-making based on Mahalanobis-Taguchi system and grey cumulative prospect theory for enterprise information investment. *Intelligent Decision Technologies*, *10*(1), 49–58. Scopus. https://doi.org/10.3233/IDT-150236

Li, L., Cao, X., & Zhang, S. (2022). Shared energy storage system for prosumers in a community: Investment decision, economic operation, and benefits allocation under a cost-effective way. *Journal of Energy Storage*, *50*. Scopus. https://doi.org/10.1016/j.est.2022.104710

Li, L., Zhang, Z., & Zhang, S. (2021). Hybrid Algorithm Based on Content and Collaborative Filtering in Recommendation System Optimization and Simulation. *Scientific Programming*, e7427409. https://doi.org/10.1155/2021/7427409

Liang, T. P. (2008). Recommendation systems for decision support: An editorial introduction. *Decision Support Systems*. 45 (3), 385-386. https://doi.org/10.1016/j.dss.2007.05.003.

Luo, W. (2020). Application of improved clustering algorithm in investment recommendation in embedded system. *Microprocessors and Microsystems*, *75*. Scopus. https://doi.org/10.1016/j.micpro.2020.103066

Masoud, E. Y. (2013). The Effect of Perceived Risk on Online Shopping in Jordan. *European Journal of Business and Management.* 5(6), 76–87. https://www.iiste.org/Journals/index.php/EJBM/article/view/4689

Meyer, A, S. (2007). Understanding Customer Experience. *Harvard Business Review*. https://hbr.org/2007/02/understanding-customer-experience

Mo, J.-L., Schleich, J., Zhu, L., & Fan, Y. (2015). Delaying the introduction of emissions trading systems-Implications for power plant investment and operation from a multi-stage decision model. *Energy Economics*, *52*, 255–264. Scopus. https://doi.org/10.1016/j.eneco.2015.11.009

Mottahedi, A., Sereshki, F., & Ataei, M. (2018). Overbreak prediction in underground excavations using hybrid ANFIS-PSO model. *Tunnelling and Underground Space Technology*, 80, 1–9. https://doi.org/10.1016/j.tust.2018.05.023

Mutanov, G., Milosz, M., Saxenbayeva, Z., & Kozhanova, A. (2018). *Investments Decision Making on the Basis of System Dynamics* (Vol. 769, p. 303). Springer Verlag; Scopus. https://doi.org/10.1007/978-3-319-76081-0_25

Naranjo, R., & Santos, M. (2019). A fuzzy decision system for money investment in stock markets based on fuzzy candlesticks pattern recognition. *Expert Systems with Applications*, *133*, 34–48. Scopus. https://doi.org/10.1016/j.eswa.2019.05.012

Niu, D., Song, Z., Wang, M., & Xiao, X. (2017). Improved TOPSIS method for power distribution network investment decision-making based on benefit evaluation indicator system. *International Journal of Energy Sector Management*, *11*(4), 595–608. Scopus. https://doi.org/10.1108/IJESM-05-2017-0005

Orciuoli, F., & Parente, M. (2017). An ontology-driven context-aware recommender system for indoor shopping based on cellular automata. *Journal of Ambient Intelligence and Humanized Computing*. *8*(6), 937–955. https://doi.org/10.1007/s12652-016-0411-2

Ortner, J., Velthuis, L., & Wollscheid, D. (2017). Incentive systems for risky investment decisions under unknown preferences. *Management Accounting Research*, *36*, 43–50. Scopus. https://doi.org/10.1016/j.mar.2016.09.001

Papapostolou, C. M., Kondili, E. M., & Tzanes, G. (2018). *Optimisation of water supply systems in the water – energy nexus: Model development and implementation to support decision making in investment planning* (Vol. 43, p. 1218). Elsevier B.V.; Scopus. https://doi.org/10.1016/B978-0-444-64235-6.50211-4

Paranjape-Voditel, P. and Umesh, D. (2013). A stock market portfolio recommender system based on association rule mining. *Applied Soft Computing*. 13 (2). 1055-1063.

Pemisindo, F. (2020). Hybrid Recommendation System with Clustering and Classification Method Based on Case of Clinic XYZ. *International Journal of Advanced Trends in Computer Science and Engineering. 9*(2), 1352–1357. https://doi.org/10.30534/ijatcse/2020/69922020

Petković, D., & Ćojbašić, Ž. (2012). Adaptive neuro-fuzzy estimation of autonomic nervous system parameters effect on heart rate variability. *Neural Computing and Applications*. 21(8), 2065–2070.

Petković, D., Issa, M., Pavlović, ND., Zentner, L., & Ćojbašić, Ž. (2012). Adaptive neuro-fuzzy controller for the adaptive compliant robotic gripper. *Expert System and Applications*. 39(18), 13295–13304.

Portfolio investment. (2019). In Wikipedia. https://en.wikipedia.org/w/index.php?title=Portfolio_investment&oldid=918362764

Quitoras, M. R., Cabrera, P., Campana, P. E., Rowley, P., & Crawford, C. (2021). Towards robust investment decisions and policies in integrated energy systems planning: Evaluating trade-offs and risk hedging strategies for remote communities. *Energy Conversion and Management*, *229*. Scopus. https://doi.org/10.1016/j.enconman.2020.113748

Rajab, Sh. & Sharma, V. (2017). Neuro-fuzzy approach in business applications: a decade survey. *International Journal of Business Information Systems* (IJBIS). 26 (4). http://www.inderscience.com/offer.php?id=87748

Rani, P., Shokeen, J., Singh, A., Singh, A., Kumar, S., & Raghuvanshi, N. (2022). Stock Price Prediction Using Reinforcement Learning. In A. Khanna, D. Gupta, S. Bhattacharyya, A. E. Hassanien, S. Anand, & A. Jaiswal (Eds.), *International Conference on Innovative Computing and Communication*s (pp. 69–76). Springer. https://doi.org/10.1007/978-981-16-2597-8_6

Ren, K., & Malik, A. (2019). Investment recommendation system for low-liquidity online peer to peer lending (P2PL) marketplaces. *WSDM 2019 - Proceedings of the 12th ACM International Conference on Web Search and Data Mining*, 510–518. Scopus. https://doi.org/10.1145/3289600.3290959

Renna, P. (2017). A Decision Investment Model to Design Manufacturing Systems based on a genetic algorithm and Monte-Carlo simulation. *International Journal of Computer Integrated Manufacturing*, *30*(6), 590–605. Scopus. https://doi.org/10.1080/0951192X.2016.1187299

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: Grouplens (1994). Open architecture for collaborative filtering of netnews. In: *Proceedings ACM Conference on Computer-Supported Cooperative Work*, pp. 175–186

Resnick, P., Varian, H.R. (1997). Recommender systems. *Communications of the ACM*. 40(3), 56–58.

Ribas, J. R., & da Silva Rocha, M. (2015). A decision support system for prioritizing investments in an energy efficiency program in favelas in the city of Rio de Janeiro. *Journal of Multi-Criteria Decision Analysis*, *22*(1–2), 89–99. Scopus. https://doi.org/10.1002/mcda.1524

Rühr, A., Streich, D., Berger, B., & Hess, T. (2019). A classification of decision automation and delegation in digital investment management systems. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2019-January*, 1435–1444. Scopus. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85099474028&partnerID=40&md5=4b540866ec7ddd81afed3d00d76a58dc

Salge, T. O., Kohli, R., & Barrett, M. (2015). Investing in information systems: On the behavioral and institutional search mechanisms underpinning hospitals' is investment decisions. *MIS Quarterly: Management Information Systems*, *39*(1), 61–89. Scopus. https://doi.org/10.25300/MISQ/2015/39.1.04

Sarkar, D., Bali, R., Sharma, T. (2018) Machine Learning Basics. In: *Practical Machine Learning with Python*. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-3207-14

SAS Help Center: Cubic Clustering Criterion. (2015). Retrieved April 20, 2022, from https://documentation.sas.com/doc/en/emref/14.3/n1dm4owbc3ka5jn11yjkod7ov1va.htm

Scaparra, M. P., Starita, S., & Sterle, C. (2015). *Optimizing investment decisions for railway systems protection* (Vol. 27, p. 233). Springer Netherlands; Scopus. https://doi.org/10.1007/978-3-319-04426-2_11

Scheinbaum, A.C. (2017). *Online Consumer Behavior: Theory and Research in Social Media, Advertising and E-tail.* Routledge. https://www.worldcat.org/title/online-consumer-behavior-theory-and-research-in-social-media-advertising-and-e-tail/oclc/754734005/viewport

Sedighi, M., Jahangirnia, H., Gharakhani, M., & Farahani Fard, S. (2019). A Novel Hybrid Model for Stock Price Forecasting Based on Metaheuristics and Support Vector Machine. *Data*. *4*(2), Article 2. https://doi.org/10.3390/data4020075

Selden, P.L. (1998). *Sales Process Engineering: An Emerging Quality Application*. Quality Progress: 59–63.

Senecal, S., & Nantel, J. (2004). The influence of online product recommendations on consumers' online choices. *Journal of Retailing*. *80*(2), 159–169. https://doi.org/10.1016/j.jretai.2004.04.001

Sharma, S., Rana, V., & Malhotra, M. (2022). Automatic recommendation system based on hybrid filtering algorithm. *Education and Information Technologies*. 27(2), 1523–1538. https://doi.org/10.1007/s10639-021-10643-8

Siddiquee, MN., Haider, N. and Rahman, MR. (2015). Movie Recommendation System Based on Fuzzy Inference System and Adaptive Neuro-Fuzzy Inference System. *International Journal of Fuzzy System Applications*. 4(4), 31-69. https://doi.10.4018/IJFSA.2015100103

Siejka, M. (2017). THE ROLE OF SPATIAL INFORMATION SYSTEMS IN DECISION-MAKING PROCESSES REGARDING INVESTMENT SITE SELECTION. *Real Estate Management and Valuation*, *25*(3), 62–72. Scopus. https://doi.org/10.1515/remav-2017-0023

Slovic, P. (1972). Psychological Study of Human Judgment: Implications for Investment Decision Making. *The Journal of Finance. 27*(4), 779–799. https://doi.org/10.2307/2978668

Srivastava, V. K., Goel, R. K., & Bharti, P. K. (2013). Application Of Fuzzy Logic in Decision Making & Grid Technology. *International Journal of Engineering Research.* 2(8), 5, 761-765.

Starita, S., & Scaparra, M. P. (2016). Optimizing dynamic investment decisions for railway systems protection. *European Journal of Operational Research*, *248*(2), 543–557. Scopus. https://doi.org/10.1016/j.ejor.2015.07.025

Sulistiyo, H., & Mahpudin, E. (2020). Demographic analysis for the selection of an investment type for amateur golfers. In *Advances in Business, Management, and Entrepreneurship*. CRC (Communications Regulatory Commission) Press.

Sun, J., Wang, H., & Chen, J. (2020). Decision-making of port enterprise safety investment based on system dynamics. *Processes*, *8*(10), 1–17. Scopus. https://doi.org/10.3390/pr8101235

Szafranko, E., Srokosz, P. E., Jurczak, M., & Śmieja, M. (2022). Application of ANFIS in the preparation of expert opinions and evaluation of building design variants in the context of processing large amounts of data. *Automation in Construction*. 133, 104045. https://doi.org/10.1016/j.autcon.2021.104045

Tahmasebi, P. (2010). Comparison of optimized neural network with fuzzy logic for ore grade estimation. *Australian Journal of Basic and Applied Sciences*. 4, 764–772.

Tao, Z., Moncada, J. A., Poncelet, K., & Delarue, E. (2021). Review and analysis of investment decision making algorithms in long-term agent-based electric power system simulation models. *Renewable and Sustainable Energy Reviews*, *136*. Scopus. https://doi.org/10.1016/j.rser.2020.110405

Tarnowska, K., Ras, Z. W., & Daniel, L. (2020). *Recommender System for Improving Customer Loyalty*. Springer International Publishing. https://doi.org/10.1007/978-3-034-13438-9

Tejeda-Lorente, Á., Bernabé-Moreno, J., Herce-Zelaya, J., Porcel, C., & Herrera-Viedma, E. (2019). A risk-aware fuzzy linguistic knowledge-based recommender system for hedge funds. *Procedia Computer Science.* 162, 916.

Teotónio, I., Cabral, M., Cruz, C. O., & Silva, C. M. (2020). Decision support system for green roofs investments in residential buildings. *Journal of Cleaner Production*, *249*. Scopus. https://doi.org/10.1016/j.jclepro.2019.119365

Teresa Swartz, D.I. (2002). *Handbook of Services Marketing and Management*. Thousand Oaks, CA: Sage.

The Code of Hammurabi (2008). The Avalon project: *The Code of Hammurabi*. Translated by L. W. King. https://avalon.law.yale.edu/ancient/hamframe.asp

Thomas, E. B., Dolecheck, K. A., Mark, T. B., Eastwood, C. R., Dela Rue, B. T., & Bewley, J. M. (2019). A decision-support tool for investment analysis of automated oestrus detection technologies in a seasonal dairy production system. *Animal Production Science*, *59*(12), 2280–2287. Scopus. https://doi.org/10.1071/AN17730

Thompson, J. R. J., Feng, L., Reesor, R. M., & Grace, C. (2021). Know Your Clients' Behaviours: A Cluster Analysis of Financial Transactions. *Journal of Risk and Financial Management.* *14*(2). https://doi.org/10.3390/jrfm14020050

TOP15: A legolvasottabb oldalak között a Portfolio. (2018). *Portfoliocsoport.hu*. Retrieved April 18, 2020. https://portfoliocsoport.hu/hirek/2018/04/10/top15_a_legolvasottabb_oldalak_kozott_a_portfolio

Trianto, A. B., Hadi, I. M., Liong, T. H., & Purqon, A. (2015). Analysis Prediction of Indonesian Banks (BCA, BNI, MANDIRI) using Adaptive Neuro-Fuzzy Inference System (ANFIS) and Investment Strategies. In A. Purqon, T. Hidayat, R. J. R. Hwu, & H. Arimura (Eds.), *5th International Conference on Mathematics and Natural Sciences (icmns 2014)* (Vol. 1677, p. 080007). Amer Inst Physics. https://doi.org/10.1063/1.4930738

Ullah, F., & Sepasgozar, S. M. E. (2020). Key factors influencing purchase or rent decisions in smart real estate investments: A system dynamics approach using online forum thread data. *Sustainability (Switzerland)*, *12*(11). Scopus. https://doi.org/10.3390/su12114382

van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., & Verhoef, P. C. (2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research. 13*(3), 253–266. https://doi.org/10.1177/1094670510375599

Van Looy, A. (2016). Social Customer Relationship Management. In A. Van Looy (Ed.), *Social Media Management: Technologies and Strategies for Creating Business Value* (pp. 87–111). Springer International Publishing. https://doi.org/10.1007/978-3-319-21994-5_5

Verhoef, P.C.; Lemon, K.N., Parasuraman, A.; Roggeveen, A., Tsiros, M., Schlesinger, L.A. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies. *Journal of Retailing. Enhancing the Retail Customer Experience*. 85 (1), 31–41. https://doi.10.1016/j.jretai.2008.11.001

von Appen, J., & Braun, M. (2018). Interdependencies between self-sufficiency preferences, techno-economic drivers for investment decisions and grid integration of residential PV storage systems. *Applied Energy*, *229*, 1140–1151. Scopus. https://doi.org/10.1016/j.apenergy.2018.08.003

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly.* 26(2). xiii-xxiii. https://www.sis.uta.fi/~pt/TIEA5_Thesis_Course/Session_08_2013_02_11/

webster02-AnalyzingPastToPrepareForFuture-WritingLiteratureReview_0.pdf

Wei, Q., Li, S., Gou, X., & Huo, B. (2019). Joint optimal decision of the shared distribution system through revenue-sharing and cooperative investment contracts. *Industrial Management and Data Systems*, *119*(3), 578–612. Scopus. https://doi.org/10.1108/IMDS-07-2018-0285

Woodcock, N., Green, A., Starkey, M., & The Customer FrameworkTM. (2011). Social CRM as a business strategy. *Journal of Database Marketing & Customer Strategy Management.* 18(1), 50–64. https://doi.org/10.1057/dbm.2011.7

World Bank (2018). Portfolio Investment. Retrieved April 18, 2020, from https://data.worldbank.org/indicator/BN.KLT.PTXL.CD

Xue, J., Ye, J., Tao, Q., & Wang, D. (2019). Multi-scenarios based operation mode and investment decision of source-storage-load system in business park. *Dianli Zidonghua Shebei/Electric Power Automation Equipment*, *39*(2), 78–83 and 92. Scopus. https://doi.org/10.16081/j.issn.1006-6047.2019.02.012

Yan, Y., Hong, L., He, X., Ouyang, M., Peeta, S., & Chen, X. (2017). Pre-disaster investment decisions for strengthening the Chinese railway system under earthquakes. *Transportation Research Part E: Logistics and Transportation Review*, *105*, 39–59. Scopus. https://doi.org/10.1016/j.tre.2017.07.001

Yera, R. and Martin, L. (2017). Fuzzy Tools in Recommender Systems: A Survey. *International Journal of Computational Intelligence Systems.* 10 (2017) 776–803.

Zahin, S., Latif, H.H., Paul, S.K., & Azeem, A. (2013). A comparative analysis of power demand forecasting with artificial intelligence and traditional approach. *International Journal of Business Information Systems* (IJBIS). 3(3). http://www.inderscience.com/offer.php?id=54469

Zhou, D., Yu, Y., Wang, Q., & Zha, D. (2019). Effects of a generalized dual-credit system on green technology investments and pricing decisions in a supply chain. *Journal of Environmental Management*, *247*, 269–280. Scopus. https://doi.org/10.1016/j.jenvman.2019.06.058

ATTACHMENT 1: INVESTMENT QUESTIONNAIRE

Portfolio, Corvinus University of Budapest, and Dorsum, one of the region's leading providers of innovative investment software, are launching joint research. The purpose of our research is to find out how conscious our readers are about their finances. In our research, we are curious about our readers' savings, spending habits, their use of digital financial solutions, or their view of the state of the economy in the years to come. Become our partner, fill out our questionnaire: it is only 8 minutes!

IN PARTNERSHIP WITH | CORVINUS | Dorsum IT SOLUTIONS

PAGE 1/7

We would like you to answer the questions below to determine whether the statements you make are true. Your answers will help us understand how open you are to digital solutions and how you are using them consciously.

1-1 Do you have a smartphone?

Yes No

1-2 Do you have a mobile internet subscription?

Yes No

1-3 Have you bought online in the last three months? (Any product or service or even a ticket, etc.)

Yes No

1-4 Do you use a password management application to store your online passwords? (Note that when the browser automatically saves your passwords, only apps specifically used for this purpose)

Yes No

1-5 Do you have a regular subscription to use any online service? (For example, Spotify, Apple Music, Netflix, Dropbox, OneDrive, etc.)

Yes No

1-6 Have you changed any of your visibility settings on social networks in the past year? (For example, who can see the post on Facebook)

Yes No

IN PARTNERSHIP WITH CORVINUS Dorsum IT SOLUTIONS

Here are questions that measure your financial awareness and risk appetite. Knowing these will help you determine which savings or investment products are right for you.

2-1 **When do you feel most safe from the following?**

If my property is in real estate

If my money is deposited in a bank account

If I keep my money in cash

If my assets are in gold

If my assets are insecurities

2-2 **What do you do most with the part of your monthly income that you do not spend on overheads and food?**

There is no such part

I spend it on entertainment and stuff

I will put it aside to have spare time for an unexpected event

Set aside for my bigger plans (e.g., vacation, purchase a car, a house)

I spend time on my hobby

2-3 **Which of the following would you choose?**

One million forints unconditionally

Two million forints, but it is up to you to decide whether to get it

Ten million forints, but only if you find out in advance what number a machine will randomly choose from one to ten

2-4 **Which of the following investment products do you think is right for you? (Multiple answers can be marked)**

Listed stock

Mutual fund

Voluntary pension fund

Government security

Other financial products

2-5 **Have you had a stock market investment in the last 3 years?** Yes No

2-6 **If so, do you regularly monitor/follow the performance of the stock?** Yes No

2-7 **Do you have a government bond investment?** Yes No

2-8 **Do you have an investment fund?** Yes No

2-9 **If so, do you regularly monitor the fund's performance?** Yes No

2-10 **Which of the following four factors are most important to you in your long-term savings? (Multiple answers can be marked)**

State aid   Opportunity for high returns   Low-risk   Low cost

2-11 **At the beginning of the year, he buys a one-year government bond for $ 100,000, which pays 3% interest. How much money will your account have in one year if the account management fee is 1% of the annual opening balance? Enter the amount in thousands of forints:**

…………………………

The following questions relate to your current savings and financial situation. These are used to gauge how difficult an unexpected event or major release is for you.

3-1 Do you have savings?

There are no
I can buy a mid-range smartphone anytime
I could pay for an abroad vacation
I could buy a 4-5-year-old car for myself
I could buy a new car for myself
I could buy property from my savings

3-2 Which statement is true for their household in terms of monthly living?

We do not come out of our monthly revenue
Our monthly income provides our basic livelihood (housing, food)
We just go out a little every month to buy new clothes for fun
We can spend it regularly on entertainment and shopping
In addition to our regular spending (entertainment, shopping), we can set you aside for the holidays
We can regularly set aside savings

3-3 What investment are your current savings? Choose from the following options (Multiple answers can be marked)

There is no such/ in cash/ bank deposits government securities /Mutual fund /Investment insurance/ Corporate bond in shares/ Owned by my own business/ Real Estate (flat, house, land, farmland) /In art treasures, collections (paintings, coins, old-timer cars, etc.) /precious metals/ In a derivative financial product (options, futures, etc.)/ Equity settled/ Trusts/ other Investment Funds or Investment

3-4 How would you describe your current financial situation?

Hopeless / It is hard, but I live / I have no daily problems, but not overall, I am worried about the security of my future/ I am calm about myself, but the future of the children is uncertain / Everything's okay

3-5 I expect my savings today to be:

In one or two weeks/ Within a month / In one year/ Within 2-3 years/ Within 4-5 years/Within 5-8 years/ Over 8 years/ I do not plan to use my savings

3-6 Do you or any members of your household living with you have the following? (Multiple answers can be marked)

Health fund savings/ Membership of pension fund guaranteeing adequate pension, pension insurance/ accident insurance/ Comprehensive property insurance for all properties you own/ Property insurance covering all valuable valuables (e.g., CASCO)/ Liability Insurance for Professional and Other Liability (e.g., Damage to Others Due to Breaking of Own Pipe)/ Sufficient security money/ A relative, friend, acquaintance who can count on nursing care, school/doctor referral, homework

The following questions aim to assess your characteristics. Your answers will give you a better understanding of the relationship between human habits and attitudes to finance and how these factors are linked to choosing the right savings and investment product.

4-1 How are these statements true to you?

4-1-1 I do not plan my future, I prefer drifting with events, I plan flexibly

1 Not at all 2 3 Neutral 4 5 Very true

4-1-2 When I set a goal for myself, I usually plan the steps to get there

1 Not at all 2 3 Neutral 4 5 Very true

4-1-3 If I feel like my job is getting too risky, I do not waste my time on it

1 Not at all 2 3 Neutral 4 5 Very true

4-1-4 My destiny is in my own hands

1 Not at all 2 3 Neutral 4 5 Very true

4-1-5 It is up to me how I reach my goals

1 Not at all 2 3 Neutral 4 5 Very true

4-1-6 If my plan does not go as I expected, I will let it go

1 Not at all 2 3 Neutral 4 5 Very true

4-1-7 The factors that ensure my success are in my hands

1 Not at all 2 3 Neutral 4 5 Very true

4-1-8 I keep a detailed list of my plans

1 Not at all 2 3 Neutral 4 5 Very true

4-1-9 I like working in teams and getting help and assistance from the right professionals

1 Not at all 2 3 Neutral 4 5 Very true

4-1-10 When I reach my goal, I reward myself

1 Not at all 2 3 Neutral 4 5 Very true

4-2 Choose from the following options

4-2-1 When I decide

I tend to be nervous afterward if I have made the right decision

Instead, he worries if I am going to make the right decision

4-2-2 When I plan my day

Rather, I focus on the tasks I see given in the day and organize my other activities around them

I would rather imagine what my day should be like and shape my business

4-2-3 When I do a task

I work more hastily than comfortably

I work more comfortably than a little rush

4-2-4 What is your characteristic?

I try to influence the course of things

I let things happen around me and I adjust to them

4-2-5 My daily schedule
My workplace dictates
It is dictated by my family
I plan it
4-2-6 When I plan
Only my imagination limits my possibilities
I can see what options I can choose from
4-2-7 During my work
I find myself disorganized, deconstructive, passive
Rather, I see dynamic work, clear assignment of tasks
4-2-8 It bothers me more
If the work, I am doing seems pointless
If the work, I am doing does not satisfy me mentally
4-2-9 How many weeks in advance do you usually plan your vacation?
1 week
2-3 weeks
4-6 weeks
6-8 weeks
More than 8 weeks
I do not usually plan my holidays in advance

IN PARTNERSHIP WITH CORVINUS Dorsum IT SOLUTIONS

PAGE 5/7

Below, we would like you to answer two elaborate questions. Answering is optional, but it can help us understand how much you are planning your finances and how conscious you are. **In a few sentences, tell me that**

5-1 how you collected the money for your most recent major investment (monthly savings, cash flow, the larger amount received, etc.) The answer is optional.

……………………………………..

5-2 Is your monthly budget planned in your household, and if so, to what extent? Answering is optional.

……………………………………..

At the end of our survey, we would like to evaluate how satisfied you are with the services of your current bank and how open you are to the latest savings trends and emerging services. Your answer will give you a clearer picture of the weaknesses of the Hungarian banking sector and outline the innovation directions that will benefit you and the Hungarian public in the financial field in the medium and long term.

6-1 Please enter the number of banks you are currently using: …………

6-2 Bankomnál

I am a regular customer

I am a premium customer

I am a private banking client

6-3 Please indicate to what extent you agree with the following statements

6-3-1 Overall, I am satisfied with my bank

1 Not at all 2 3 Neutral 4 5 Very true

6-3-2 My account statement is completely understandable to me

1 Not at all 2 3 Neutral 4 5 Very true

6-3-3 I am satisfied with the online services provided by my bank (internet banking, mobile application, etc.)

1 Not at all 2 3 Neutral 4 5 Very true

6-3-4 I would love to be a customer of a bank that is accessible only online and telephone clerks

1 Not at all 2 3 Neutral 4 5 Very true

6-3-5 I am satisfied with the range of services provided by my bank

1 Not at all 2 3 Neutral 4 5 Very true

6-3-6 I would love to make an investment decision based on a robot's recommendation

1 Not at all 2 3 Neutral 4 5 Very true

6-3-7 I would be willing to switch banks for better digital services

1 Not at all 2 3 Neutral 4 5 Very true

6-3-8 I would accept an investment proposal from a robot if I could discuss the details with my advisor first

1 Not at all 2 3 Neutral 4 5 Very true

6-3-9 I would love to vote online at the general meetings of the companies behind the shares I hold

1 Not at all 2 3 Neutral 4 5 Very true

6-3-10 I would love to use a platform that makes all the information about the securities of my choice accessible

1 Not at all 2 3 Neutral 4 5 Very true

6-4 When making your investment decision, who / what do you prefer to give your word? (List the first place that matters most when making your decision, the last one that least matters) Drag the answers to the right place.

- Bank algorithm
- Social media recommendation
- Family
- friends
- Banking consultant
- Choices of people like me (my peers, people in similar jobs, etc.)
- I rely on myself

6-5 Please indicate how much you agree with the statements below

6-5-1 Environmental awareness factors also play a role in my investment decisions (such as how polluting a company is)

1 Not at all 2 3 Neutral 4 5 Very true

6-5-2 It is important for me that my investment reflects the value system I represent (for example, a company does not exploit its suppliers in developing countries)

1 Not at all 2 3 Neutral 4 5 Very true

6-5-3 I can expect higher returns in exchange for socially conscious investments

1 Not at all 2 3 Neutral 4 5 Very true

6-5-4 I am willing to make lower returns if I can invest my money socially consciously

1 Not at all 2 3 Neutral 4 5 Very true

6-6 Please indicate how much you agree with the statements below

6-6-1 My bank keeps my information confidential

1 Not at all 2 3 Neutral 4 5 Very true

6-6-2 Google keeps my information confidential

1 Not at all 2 3 Neutral 4 5 Very true

6-6-3 In exchange for a better service, I would be happy to share my financial status information with my bank (such as my payment habits, information about my money/securities registered with the bank, etc.)

1 Not at all 2 3 Neutral 4 5 Very true

6-6-4 In exchange for a better service, I would be happy to share my financial status information with a new non-bank player (such as my payment habits, my money/securities registered with other banks, etc.)

1 Not at all 2 3 Neutral 4 5 Very true

6-6-5 It would help me manage all my banking matters through a single platform

1 Not at all 2 3 Neutral 4 5 Very true

6-7 Please indicate which platforms you have heard of (You can mark more than one answer)

PayPal/ Transfer/Wise/ Revolut /Plus500/ eToro /Simple

6-8 Please indicate which of the following platforms you have used at least once (Multiple answers are possible)

PayPal /Transfer/Wise/ Revolut /Plus500/ eToro/ Simple

IN PARTNERSHIP WITH | BUDAPESTI CORVINUS EGYETEM | Dorsum IT SOLUTIONS

At the end of our survey, please answer some demographic questions. They do not contain personal information, they do not identify you based on their answers, and they are used for analytical purposes only.

7-1 Your gender

Male Female

7-2 It is your age ....

7-3 Where does he live?

Budapest County town City Village

7-4 What is your education?

Primary schools /High school, secondary education /College, or university - economic /College or university - not economic/ Postgraduate training/ I am currently studying

7-5 What is your job?

Small and/or medium business / large Contractor /Graduate freelance /Micro and/or sole trader Farmer, farm worker or Agricultural/ Employee senior manager /Employee middle management/ Employed lower manager, team leader, supervisor/ Subordinate Intellectuals (Employee)/another intellectual, service, trader (employee) /Skilled worker (employee) /Trained or auxiliary (employee) manual worker, seasonal worker (employee)

ATTACHMENT 2: GENERATING RULES PROCESS FOR DEMOGRAPHICANFIS

[System]
Name='DemographicANFIS'
Type='sugeno'
Version=2.0
NumInputs=6
NumOutputs=1
NumRules=1296
AndMethod='prod'
OrMethod='probor'
ImpMethod='prod'
AggMethod='sum'
DefuzzMethod='wtaver'

[Input1]
Name='input1'
Range=[0 2]
NumMFs=2
MF1='in1mf1':'trimf',[-2 0.000333047420963386 1.99981415378178]
MF2='in1mf2':'trimf',[0.000442453779770493 2.0003330110654 4]

[Input2]
Name='input2'
Range=[0 21]
NumMFs=3
MF1='in2mf1':'trimf',[-10.5 -0.00243999954277444 10.5009016647353]
MF2='in2mf2':'trimf',[-0.00426113799422104 10.4975598614319 21.0000001229014]
MF3='in2mf3':'trimf',[10.5 21 31.5]

[Input3]
Name='input3'
Range=[0 2]
NumMFs=2
MF1='in3mf1':'trimf',[-1.99999999514709 0.00624622576903484 2.00002374077672]
MF2='in3mf2':'trimf',[-0.00193069552001991 2.00622273541233 4]

[Input4]
Name='input4'
Range=[0 4]
NumMFs=4
MF1='in4mf1':'trimf',[-1.33333333137059 0.00293171412369073 1.34204420573551]
MF2='in4mf2':'trimf',[-0.00200680495691898 1.33725182401531 2.66765409107521]
MF3='in4mf3':'trimf',[1.33432343993649 2.66834094973322 3.99885323149771]

MF4='in4mf4':'trimf',[2.66872624331244 4.00068503268348 5.33333333333333]

[Input5]
Name='input5'
Range=[0 9]
NumMFs=9
MF1='in5mf1':'trimf',[-1.125 0.000129835998789197 1.12603591835356]
MF2='in5mf2':'trimf',[8.33373538882956e-05 1.1241002468386 2.2463782754096]
MF3='in5mf3':'trimf',[1.12470673176409 2.24781007956832 3.37267224716138]
MF4='in5mf4':'trimf',[2.2494192826356 3.37722967512196 4.50422036500721]
MF5='in5mf5':'trimf',[3.37772553584201 4.50703482738974 5.62790095552386]
MF6='in5mf6':'trimf',[4.50457684557055 5.62802266062106 6.74968949973325]
MF7='in5mf7':'trimf',[5.62375818688295 6.74822748983581 7.87467050988712]
MF8='in5mf8':'trimf',[6.74599701793529 7.87383117113062 8.99926799292603]
MF9='in5mf9':'trimf',[7.87484961347295 8.99998119438774 10.125]

[Input6]
Name='input6'
Range=[0 12]
NumMFs=3
MF1='in6mf1':'trimf',[-6 0.000309741878187414 5.99738707320999]
MF2='in6mf2':'trimf',[0.0140390606608012 6.0003097862919 12]
MF3='in6mf3':'trimf',[6 12 18]

[Output1]
Name='output'
Range=[0 3]
NumMFs=1296

ATTACHMENT 3: GENERATING RULES PROCESS FOR PERSONALITYTRAITSANFIS

[System]
Name='PersonalityTraitsANFIS'
Type='sugeno'
Version=2.0
NumInputs=6
NumOutputs=1
NumRules=2700
AndMethod='prod'
OrMethod='probor'
ImpMethod='prod'
AggMethod='sum'
DefuzzMethod='wtaver'
[Input1]
Name='input1'
Range=[0 5]
NumMFs=5
MF1='in1mf1':'gaussmf',[0.549885272788741 0.0107057738877879]
MF2='in1mf2':'gaussmf',[0.536265624145132 1.25834667281187]
MF3='in1mf3':'gaussmf',[0.528427243195111 2.50701153998811]
MF4='in1mf4':'gaussmf',[0.522146428621981 3.75476072549451]
MF5='in1mf5':'gaussmf',[0.52706107972212 5.00199527787233]

[Input2]
Name='input2'
Range=[0 5]
NumMFs=5
MF1='in2mf1':'gaussmf',[0.526775281705263 -0.00211378995020233]
MF2='in2mf2':'gaussmf',[0.528486497772131 1.2477209132315]
MF3='in2mf3':'gaussmf',[0.530533974073317 2.50025412230114]
MF4='in2mf4':'gaussmf',[0.527829452340464 3.75118787918883]
MF5='in2mf5':'gaussmf',[0.531022588611375 4.99984736430129]

[Input3]
Name='input3'
Range=[0 3]
NumMFs=3
MF1='in3mf1':'gaussmf',[0.6390421443634710 0.00203308373918616]
MF2='in3mf2':'gaussmf',[0.63677349506533 1.50252315914775]
MF3='in3mf3':'gaussmf',[0.6323976605473 3.00266536560544]

[Input4]
Name='input4'
Range=[0 2]
NumMFs=2
MF1='in4mf1':'gaussmf',[0.853821880708378 0.00823448346805835]
MF2='in4mf2':'gaussmf',[0.836436788609586 2.0118556425019]

[Input5]
Name='input5'
Range=[0 10]
NumMFs=9
MF1='in5mf1':'gaussmf',[0.544487556737433 0.00761971509636061]
MF2='in5mf2':'gaussmf',[0.532760315277465 1.25438666606257]
MF3='in5mf3':'gaussmf',[0.527396656391524 2.50132622182475]
MF4='in5mf4':'gaussmf',[0.530263880094273 3.74845904913642]
MF5='in5mf5':'gaussmf',[0.536000610371835 4.9966597761886]
MF6='in5mf6':'gaussmf',[0.529820668942144 6.24982766098038]
MF7='in5mf7':'gaussmf',[0.529886536717095 7.49888372702663]
MF8='in5mf8':'gaussmf',[0.532683149632252 8.74828142426306]
MF9='in5mf9':'gaussmf',[0.533762626807301 9.9983923336356]

[Input6]
Name='input6'
Range=[0 20]
NumMFs=2
MF1='in6mf1':'gaussmf',[8.49302320942183 -0.000835618437160805]
MF2='in6mf2':'gaussmf',[8.49719656968019 19.9975574922863]

[Output1]
Name='output'
Range=[0 3]
NumMFs=2700

ATTACHMENT 4: GENERATING RULES PROCESS FOR FINANCIALANFIS

```
[System]
Name='FinancialANFIS Ruls'
Type='sugeno'
Version=2.0
NumInputs=4
NumOutputs=1
NumRules=720
AndMethod='prod'
OrMethod='probor'
ImpMethod='prod'
AggMethod='sum'
DefuzzMethod='wtaver'

[Input1]
Name='input1'
Range=[0 6]
NumMFs=6
MF1='in1mf1':'gaussmf',[0.505924599435783 -0.00189439683556451]
MF2='in1mf2':'gaussmf',[0.50428216819331 1.19648451077156]
MF3='in1mf3':'gaussmf',[0.50365659837494 2.39255827828617]
MF4='in1mf4':'gaussmf',[0.512200648939718 3.58800496023208]
MF5='in1mf5':'gaussmf',[0.52941473878969 4.78640770722068]
MF6='in1mf6':'gaussmf',[0.513868815314422 5.9978167195858]

[Input2]
Name='input2'
Range=[0 6]
NumMFs=6
MF1='in2mf1':'gaussmf',[0.509630866505576 1.97976824547824e-05]
MF2='in2mf2':'gaussmf',[0.509606393516228 1.20001404785888]
MF3='in2mf3':'gaussmf',[0.509578643022572 2.40000052522814]
MF4='in2mf4':'gaussmf',[0.509583852803805 3.59997953441616]
MF5='in2mf5':'gaussmf',[0.509636419631854 4.79996463518774]
MF6='in2mf6':'gaussmf',[0.50964076741018 5.99997512062908]

[Input3]
Name='input3'
Range=[1 5]
NumMFs=5
MF1='in3mf1':'gaussmf',[0.424659313579378 0.999999328009502]
MF2='in3mf2':'gaussmf',[0.424633377799767 2.00000656309624]
MF3='in3mf3':'gaussmf',[0.424659188130285 3.00000070153398]
MF4='in3mf4':'gaussmf',[0.424653712042049 4.0000030314382]
MF5='in3mf5':'gaussmf',[0.424660803292909 5.00000003952626]

[Input4]
Name='input4'
```

Range=[0 4]
NumMFs=4
MF1='in4mf1':'gaussmf',[0.566259232288328 2.64744897299502e-05]
MF2='in4mf2':'gaussmf',[0.566216490513653 1.33335137308839]
MF3='in4mf3':'gaussmf',[0.566192952572373 2.66667649845446]
MF4='in4mf4':'gaussmf',[0.566192779653757 4.00001216833845]

[Output1]
Name='output'
Range=[0 3]
NumMFs=720

ATTACHMENT 5: GENERATING RULES PROCESS FOR COMBINEDANFIS

[System]
Name='CombinedANFISRules'
Type='sugeno'
Version=2.0
NumInputs=6
NumOutputs=1
NumRules=729
AndMethod='prod'
OrMethod='probor'
ImpMethod='prod'
AggMethod='sum'
DefuzzMethod='wtaver'

[Input1]
Name='input1'
Range=[0 3]
NumMFs=3
MF1='in1mf1':'trimf',[-1.5 0.00140695604436714 1.50280652185293]
MF2='in1mf2':'trimf',[0.000747550787495075 1.50308515227308 3.00007862112638]
MF3='in1mf3':'trimf',[1.50336593459829 3.00167766380987 4.5]

[Input2]
Name='input2'
Range=[0 3]
NumMFs=3
MF1='in2mf1':'trimf',[-1.49999997570307 0.00650474541903476 1.51283496518699]
MF2='in2mf2':'trimf',[-0.00479830458586012 1.51679367948954 3.00114025887823]
MF3='in2mf3':'trimf',[1.52102940665605 3.01030906678536 4.5]

[Input3]
Name='input3'
Range=[0 3]
NumMFs=3
MF1='in3mf1':'trimf',[-1.5 -0.00448550565766976 1.49095253320211]
MF2='in3mf2':'trimf',[-0.000303141665806183 1.49742644788497
2.99993927017783]
MF3='in3mf3':'trimf',[1.50384024683207 3.00191322099505 4.5]

[Input4]
Name='input4'
Range=[0 3]
NumMFs=3
MF1='in4mf1':'trimf',[-1.5 0.00174791252420499 1.50348443193079]
MF2='in4mf2':'trimf',[0.000843226558188782 1.50415722521298 3.00022644987206]
MF3='in4mf3':'trimf',[1.50483885260968 3.00240847145672 4.5]

[Input5]

Name='input5'
Range=[0 3]
NumMFs=3
MF1='in5mf1':'trimf',[-1.5 0.00233327716726784 1.50464628171306]
MF2='in5mf2':'trimf',[0.000891919522777258 1.5031977592608 2.99988949908588]
MF3='in5mf3':'trimf',[1.50173079395181 3.00086399378097 4.5]

[Input6]
Name='input6'
Range=[0 3]
NumMFs=3
MF1='in6mf1':'trimf',[-1.5 -0.00308970585223127 1.49378446182136]
MF2='in6mf2':'trimf',[6.68606838579428e-05 1.49278360013402 3.00321700397349]
MF3='in6mf3':'trimf',[1.49179742621458 2.99586282460635 4.49999999318986]

[Output1]
Name='output'
Range=[0 3]
NumMFs=729

ATTACHMENT 6: PUBLICATIONS RELATED TO THE RESEARCH

Journal Articles:

Asemi, A., Asemi, A. & Ko, A. (2023). Customizing Investment Recommendations Using ANFIS and Potential Investor's Financial Situation in Retail Banking. *Journal of Big Data*. Under Review. Q1

Asemi, A., Asemi, A. & Ko, A. (2023). ANFIS-based model for investment recommender system using Financial Management Traits. *IEEE Transactions on Fuzzy Systems*. Under Review. Q1

Asemi, A., Asemi, A. & Ko, A. (2023). Investment Intelligence: A Combined Neuro-Fuzzy Inference-Based Recommender System for Personalized Investment Strategies. *Expert Systems with Applications*. Under Review. Q1

Fatahi Nafchi, N., Asemi, A., & Asemi, A. (2023). A Fuzzy Delphi-based Inference System for Detecting and Controlling Rice Weeds. *IEEE Robotics & Automation Magazine*. Under Review. Q1

Asemi, A., Asemi, A. & Ko, A. (2023). Developing an ANFIS-based Investment Recommender System using Multimodal Neural Network Pretraining. IEEE Intelligent Systems. Under Review. Q1

Asemi, A., Asemi, A. & Ko, A. (2023). Investment Recommendation Using ANFIS and Potential Investors' Experiences. *IEEE Intelligent Systems*. Under Publish. Q1

Asemi, A., Asemi, A. & Ko, A. (2023). A model for the investment recommender system using ANFIS based on the Potential Investors' Decision Key Factors (PIDKFs). *Big Data*. Under publish. Q2

Asemi, A., Asemi, A. & Ko, A. (2023). Unveiling the Impact of Managerial Traits on Investor Decision Prediction: ANFIS Approach. *Soft Computing*. Under publish. Q2

Asemi, A., Asemi, A. & Ko, A. (2023). A Model for Investment Type Recommender System based on the Potential Investors' Demographic and feedback using ANFIS. *Journal of Big Data*. Under publish. Q1

Asemi, A. and Asemi, A. (2022). A Judgment-Based Model for Usability Evaluating of Interactive Systems Using Fuzzy Multi Factors Evaluation (MFE). *Applied Soft Computing*. https://doi.org/10.1016/j.asoc.2022.108411. Q1

Asemi, A. and Asemi, A. (2022). "Data for Usability Evaluating of Interactive Systems based on the Judgment-Based Model". *Data in Brief*. Under publish. Q2

Asemi, A. and Asemi, A., Ko, A., & Alibeigi, A. (2022). An integrated model to evaluate big data properties for analytical methods in recommender systems. *Journal of Big Data*. https://doi.org/10.1186/s40537-022-00564-z. Q1

Asemi, A. and Asemi, A. (2022). Non-Empirical ISO 9241-210:2019-based Usability Evaluation Using Fuzzy Inference Analyser. *Library Hi Tech*. *Special Issue on*

"*Social Robots: Services and Applications*." https://doi.org/10.1108/LHT-02-2022-0091. Q1

Kovács, T., Ko, A., and Asemi, A. (2021). Exploration of the investment patterns of potential retail banking customers using two-stage cluster analysis. *Journal of Big Data*. 8:141. https://doi.org/10.1186/s40537-021-00529-4. Q1

Asemi, A., Ko, A., and Asemi. A. (2021). Infoecology of Deep Learning & Smart Manufacturing: Thematic & Concept Interactions. *Library Hi Tech*. https://doi.org/10.1108/LHT-08-2021-0252. Q1

Asemi, A., Ko, A. & Nowkarizi, M. (2020). Intelligent libraries: A review on Expert Systems, Artificial Intelligence, and Robot. *Library High Tech*. 39 (2). 412-434. https://doi.org/10.1108/LHT-02-2020-0038 Q1

Asemi, A. & Ko, A (2020). A Bibliometrics Literature Review on Cryptocurrency. *Library Philosophy and Practice*. https://digitalcommons.unl.edu/libphilprac/3714/ Q2

Asemi, A. & Ko, A. (2020). The investigation on Infoecology in the field of Smart Manufacturing. *Library High Tech*. https://doi.org/10.1108/LHT-03-2024-0057. Q1

Asemi, A., Salim, S.S.B., Shahamiri, S.R., Asemi, A., and Houshangi, N. (2019). Adaptive neuro-fuzzy inference system for evaluating dysarthric automatic speech recognition (ASR) systems: a case study on MVML-based ASR. *Soft Computing*. 23 (10). 3529-3544. https://doi.org/10.1007/s00504-018-3013-4. Q1

Conference Papers:

Asemi, A., Asemi, A. & Ko, Andrea (2023). Systematic Review and Propose an ANFIS-Based Investment Type Recommender System using Investors' Demographic. *ICICT 2023: 8th International Congress on Information and Communication Technology,* London, 24-23 Feb.

Asemi, A. and Ko, A. (2021). A Novel Combined Business Recommender System Model Using Customer Investment Service Feedback. Proceeding of the 34th Bled eConference, June 27-30, 2021, Bled, Slovenia.

Asemi, A. & Ko, Andrea (2019). Infoecology of Smart Manufacturing. *Conference: OGIK2019,* Budapest, Hungary 2019.11.08. - 2019.11.09. Budapest: Milton Friedman University. **Best Presentation Award.** https://njszt.hu/hu/news/2019-12-06/ogik2019-konferencia-beszamolo