Exploring the Influence of Bahavioural Traits on Financial Risk Aversion: A Machine Learning Approach

Thi Nhu Ngoc VO

Supervised by Dr. Silvia Cirstea

A DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE MSC APPLIED DATA SCIENCE

SCHOOL OF COMPUTING AND INFORMATION SCIENCES

ANGLIA RUSKIN UNIVERSITY

SEPTEMBER 2023

Declaration

I, Thi Nhu Ngoc VO, hereby declare that to the best of my knowledge, the intellectual content of this dissertation titled “Exploring the Influence of Bahavioural Traits on Financial Risk Aversion: A Machine Learning Approach” is the product of my own work. This work has not been submitted for any degree or other purposes, neither in the same nor in a similar way and has not yet been published anywhere. All materials contained and assistance received in preparing this thesis have been explicitly sourced and acknowledged. I also declare that the information published in this dissertation has been obtained and presented in accordance with academic rules and ethical conduct.

Acknowledgement

This thesis would not have been feasible without the invaluable assistance of several individuals. I would like to use this opportunity to express my heartfelt appreciation for their encouragement, understanding, and support.

To begin, I want to express my gratitude to my thesis supervisor, Dr. Silvia Cirstea, for countless insightful remarks and ideas. She patiently listened, discussed, and gave me valuable comments and recommendations in every part of the research. I am fortunate to have her supervision and constant support, which has inspired and enabled me to finish my thesis.

In addition, a special thanks to the instructors and lecturers at Anglia Ruskin University, particularly those from the School of Computing and Information Sciences. The university has provided me with a strong knowledge foundation that would serve me well not only in this paper but also for my future life and career growth.

I would also like to thank to the financial advisors and experts I have reached out to during the course of conducting the research for invaluable suggestions on the questionnaire as well as the general reasoning of this paper. This has helped a lot in improving the quality and mutual understanding between readers and author of the thesis. Besides, I am thankful for my friends and colleagues for helping me to distribute questionnaires, and the respondents who participated in the survey.

Finally, it would be impossible to say enough about my dear family, whose understanding has helped me to overcome this challenging time to complete this paper with quality in time.

Table of Contents

Abbreviations

AGR	Agreeableness
AMOS	Analysis of Moment Structures
ANOVA	Analysis of Variance
BFI	Big Five Inventory
CART	Classification and Regression Tree
CRT	Cognitive Reflection Test
CST	Conscientiousness
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
EDA	Exploratory Data Analysis
EEA	European Economic Area
EUT	Expected Utility Theory
EXT	Extraversion
fMRI	Functional Magnetic Resonance Imaging
GDPR	General Data Protection Regulation
k-NN	k-Nearest-Neighbours
MATLAB	Matrix Laboratory (programming language)
NEO	Neuroticism - Extraversion - Openness
NEU	Neuroticism

OLS	Ordinary Least Squares
OPE	Opennes to Experience
RPR	Risk Preference
SEM	Structural Equation Modeling
SMOTE	Synthetic Minority Oversampling Technique
SPSS	Statistical Package for the Social Sciences
SVM	Support Vector Machine
UK	United Kingdom
VC	Vapnik-Chervonenkis

List of Figures

List of Tables

Abstract

This research delves into the intricate relationship between psychological factors, particularly personality traits and cognitive reflection, and their influences on risk aversion within the realm of investment decisions. Through the application of a structured questionnaire, based on the Big Five Inventory and the Cognitive Reflection Test (CRT), the study aims to discern how individual attributes might sway investment behaviors.

With a dataset consisting of 321 respondents, the research employs various machine learning models, namely Multinomial Logistics Regression, Random Forest, Decision Trees, Support Vector Machines, and k-Nearest Neighbors, to predict risk preferences.

The findings underscore neuroticism as the most influential predictor of risk aversion, aligning with prior academic assertions that individuals with heightened neuroticism display more risk-averse tendencies. The CRT scores showcased a nonlinear relationship with risk aversion, suggesting a complex interplay between cognitive abilities and financial decisions.

The insights gleaned from this study are paramount for individual investors, financial advisors, and financial institutions, advocating for a more personalized and understanding-driven approach towards investment behaviors. Ultimately, in bridging the gap between behavioral finance and traditional financial theories, this research offers a fresh perspective on the psychological underpinnings that drive financial decisions, emphasizing the need for more holistic and behavior-informed financial planning and advice.

Keywords: Behavioral finance, Machine Learning, personal characteristics, Big Five personality traits, Cognitive Reflection Test, Risk aversion

Chapter 1

Introduction

1.1. Problem Background

The domain of financial investment has seen substantial changes throughout time, influenced not just by market forces but also by the involvement of individual investors. Throughout history, financial theories have often asserted that rationality serves as the fundamental basis for making investment choices. However, when delving more into the essence of investing, a contrasting narrative emerges, whereby individual behaviours, influenced by distinct personality characteristics and cognitive processes, significantly impact the financial decisions undertaken.

Despite the rapid growth and development of the financial investment sector, characterised by its diverse range of strategies and high number of transactions, a perplexing issue continues to exist. Although market trajectories might exhibit volatility and seem to lack predictability, the fundamental aspect of this phenomenon typically resides in comprehending the behaviours and preferences of individual investors. In an age when the subjective contentment derived from investment results assumes a critical role alongside financial returns, there exists a compelling need to further explore the behavioural aspects of investing. The research paper titled "Exploring the Influence of Bahavioural Traits on Financial Risk Aversion: A Machine Learning Approach" aims to shed light on this particular aspect.

The primary objective is not alone to analyse and comprehend the complex association between behavioural characteristics and risk aversion, but rather to use these findings in order to enhance personal happiness in investment activities. This study endeavours to use machine learning techniques to comprehensively investigate the intricate interplay between investor personalities, cognitive inclinations, and their influence on risk preferences. For a significant number of individuals, the act of investing has a significance that extends beyond mere monetary benefits, including the experiential aspect of the process. It is of utmost importance to guarantee contentment throughout this endeavour, not just for individual investors but also for financial advisers and organisations that are committed to providing investing experiences that are tailored and gratifying. This research positions itself at the vanguard of reshaping the future landscape of individual investment by combining behavioural insights with advanced machine learning approaches.

1.2. Research Aim

The research aims to harness the capabilities of machine learning to investigate the intricate relationship between individual behavioral traits and financial risk aversion, with an ultimate objective of enhancing personal investment satisfaction. Through this exploration, we aspire to offer tailored investment insights that align with the unique behavioral profiles of individual investors.

1.3. Research Objectives

Working from that starting point, the specific research objectives are:

i. Using machine learning approaches, examine and clarify the relationships among the Big Five personality traits, Cognitive Reflection Test score, and their impact on risk aversion in financial decision-making.
ii. Evaluate the effectiveness and accuracy of machine learning models in predicting financial risk preferences, setting a benchmark for future enhancements in personalized financial advising.
iii. Advise financial institutions on how to modify their services in accordance with a thorough comprehension of unique behavioural qualities, so optimising customer experience and investment satisfaction.

1.4. Research Scope

Current individual investors and potential investors (both hereinafter referred to as "investors" or "individual investors") are the subjects of this research. The study was conducted and finalized in the period June to September 2023. Primary data source is a collection of 350 survey responses collected during 10^{th} July – 10^{th} Aug 2023. Secondary data sources are collected from book sections, online news, and journals cited in References.

Investors' characteristics cover two broad areas: personality and cognition. Personality traits are explored by the Big Five personality inventory. Cognition is examined by Cognitive Reflection Test. Risk preference is expressed through level of risk aversion which will be further determined in Chapter 3: Research methodology.

1.5. Abridged Methodology

The study applies qualitative analysis to explore ideas, understand concepts, and gather in-depth insights; and quantitative analysis to systematically collect, analyze data and test relationships. Machine learning approaches to predict financial risk aversion by using a dataset that includes both the Big Five personality characteristics and cognitive reflection scores. The procedure started by doing data preprocessing and feature engineering. Following this, five models, including Logistics Regression, Decision Tree, Random Forest, Support Vector Machine (SVM) and k-Nearest-Neighbours (k-NN), were trained and evaluated. The selection of the optimal model was made by considering the accuracy scores and cross-validation outcomes. The used technique facilitated the identification and analysis of patterns and importances between behavioural features and preferences in investing risk. Questionnaires were sent to respondents through social networks. The survey examined the behavior of 321 investors with consent through 17 observed questions with six-point Likert scale.

1.6. Expected Contribution

The primary motivation for research is to harness the power of machine learning in comprehending the intricate factors that determine risk behaviour, with a specific focus on examining the impact of behavioural features on investment-related risk choices.

This study employs sophisticated machine learning methodologies to provide comprehensive insights across several industries.

For individual investors: Using machine learning models, this study elucidates the nuanced investing behaviours associated with specific personality traits. This empowers investors with a predictive instrument to better interpret market dynamics, ensuring that their investment decisions align with their individual risk tolerances.

For financial organizations: Integrating machine learning and behavioural finance, this research provides brokerage firms with a dynamic predictive tool. Such a tool can effectively assess investment patterns based on behavioural characteristics, thereby enhancing the credibility of advisory services and allowing for more customised client offerings.

For the domain of behavioral finance: This study seeks to pave the way for the fusion of these two disciplines, validating the potential of machine-driven behavioural finance insights not only in mature markets but also in emerging financial ecosystems.

For the author: This endeavour gives the author hands-on experience combining the rigour of machine learning with the complexities of behavioural finance. This synthesis enhances the researcher's understanding of how cutting-edge technology can be seamlessly incorporated with evolving financial theories in order to generate more actionable insights.

1.7. Research Structure

Chapter 1: Introduction
Chapter 2: Literature Review
Chapter 3: Methodology
Chapter 4: Results & Discussion
Chapter 5: Conclusion

Chapter 2

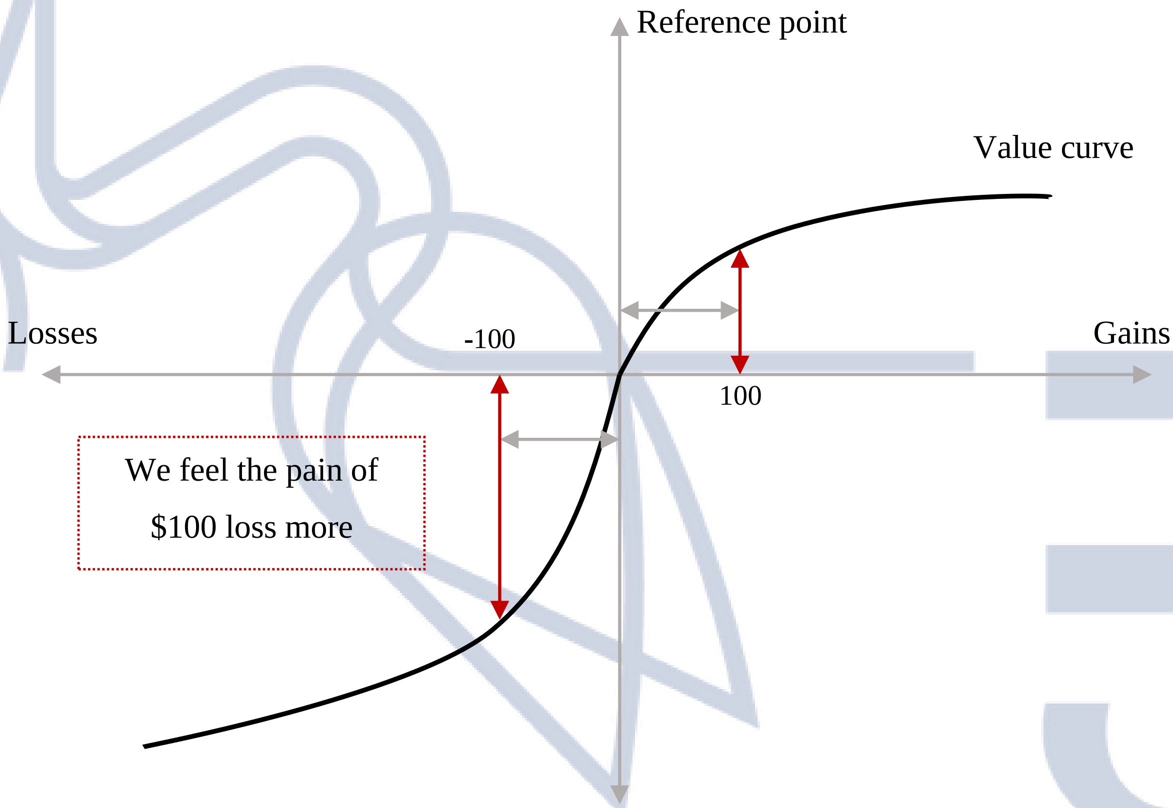
Literature Review

2.1. Standard vs. Behavioral Finance

Modern Portfolio Theory, developed by Markowitz, states that an efficient portfolio can be constructed based on the highest expected return given the amount of risk assumed (Markowitz, 1991). Efficient Market Hypothesis, proposed by Fama, suggests that financial markets are efficient, and prices always reflect available information (Fama, 1960). These theories assume rational participants and efficient markets, leading to the belief that individuals make optimal judgments and mistakes in the market do not affect prices. Both fail to explain market disruptions or abnormalities such as bubbles, overreactions, and reversals. Another dominant notion among traditional finance theories according to Bernoulli (2011), Expected Utility Theory (EUT), where economic behavior is seen as rational, failed to explain why individuals are attracted to insurance and gambling. People tend to overestimate the likelihood of certain events and behave differently depending on the context of losses or gains.

Behavioral finance introduced Prospect Theory as a critique to traditional finance theories. Kahneman and Tversky (1979; 1992) conclusively established that people's decision-making under risk contradicts the predictions of the EUT. One of their key arguments is that people do not evaluate outcomes by their objective probabilities but rather by subjectively transformed probabilities (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Kahneman was ultimately awarded the Nobel Memorial Prize in Economics in 2002 for his work on Prospect Theory. Prospect Theory focuses on subjective decision-making influenced by an individual's value system, rather than rational expectations, and proposes that individuals are risk-averse for gains and risk-seeking for losses. The value function is S-shaped, concave for gains and convex for losses, these features are illustrated in figure 2.1.

Figure 2.1: Illustration of the value function



Source: Compiled by the author

This leads to the concept of loss aversion, where individuals experience greater pain from losses than pleasure from equal gains. Loss aversion and biases impacting decision-making under risk are key concepts in behavioral finance. It aims to understand and explain actual investor behavior by incorporating behavioral evidence into decision-making models.

Ricciardi and Simon (2001) categorize behavioral finance as an interdisciplinary field that incorporates knowledge from various scientific and business disciplines to elucidate the decision-making process of investors. A crucial component of this field, investor behavior, is dedicated to understanding and predicting the systematic influences of psychological decision-making processes on financial markets. As Fromlet (2001) suggests, behavioral finance is a domain where individual conduct and market phenomena are closely intertwined, drawing insights from both psychological and financial literature. Alternatively, Pompian (2016) puts it:

> Behavioral finance attempts to understand and explain *actual* investor behavior, in contrast to theorizing about investor behavior. It differs from traditional (or standard) finance, which is based on assumptions of how investors and markets *should* behave. Behavioral finance is about understanding how people make decisions, both individually and collectively. By understanding how investors and markets behave, it may be possible to modify or adapt to these behaviors in order to improve economic outcomes.

2.2. Behavioral Characteristics

In the financial industry, risk is a contentious concept, with the majority of perspectives asserting that it is unexpected and associated with uncertainty (Ton and Dao, 2014). Traditional finance focuses on quantitative risk metrics such as variance, standard deviation, and beta and implies that everyone has a consistent risk perspective. However, behavioural finance takes into account qualitative elements where risk can be observed in both personality and cognition (Olsen, 2007; 2008). Using a "two-system" approach, Thaler and Sheffrin (1981) developed a framework that emphasises the interaction between emotions and cognition in decision-making. The emotive psychology is thought to be risk-averse and myopic, whereas the cognitive system is thought to be risk-neutral and to view intertemporal situations from a longer-term perspective. A two-system decision process is an additional potential mechanism for comprehending the relationship between personality and cognition and risk preferences.

2.2.1. Personality theories and The Big Five Personality Traits

2.2.1.a. The impact of personality on various scopes and specifically on the economy

Personality, defined as a pattern of characteristic thoughts, feelings, and behaviors that persist over time and situations, has been recognized as a significant motivator of human behavior (Phares, 1988; Durand, Newby and Sanghani, 2008). Researchers have demonstrated that various personality characteristics can predict a wide range of human actions, preferences, and outcomes, including job performance (Tisu et al., 2020), career success (Semeijn, van der Heijden and De Beuckelaer, 2020), attitudes towards materialism, money and well-being (Górnik-Durose, 2020), and academic achievement (Herrera, Al-Lal and Mohamed, 2020).

In the realm of economics, the relevance of personality to the economy has been explored (Borghans et al., 2008). In the financial domain, researchers have found connections between personality traits and investment term decisions (Mayfield, Perdue and Wooten, 2008; Jadlow and Mowen, 2010), risk-taking behaviors (Mishra, Lalumière and Williams, 2010), and investment portfolio choices and outcomes (Gambetti and Giusberti, 2019; Akhtar and Das, 2020). Empirical evidence has connected personality factors and individual preferences to home tenure features and real estate investment behaviors (Ben-Shahar and Golan, 2014). Studies have also revealed that certain personality traits, such as confidence, optimism, emotional intelligence, and adaptability, influence mutual fund investors' behavior (Chang, Chen and Fang, 2016). Particular personality attributes have been shown to have a signigicant influence on perceptional errors (Sadi et al., 2011), risk aversion and cognitive biases, as well as socially responsible investment (Nga and Ken Yien, 2013).

While the examination of investors' personality traits presents a novel and intriguing approach to understanding investment decision-making, there remains a scarcity of research on this subject. The predictive power of major personality traits suggests the potential for similar studies on risk preference, an aspect currently lacking in both psychological and economic research. This research aims to address this gap and contribute to the current empirical knowledge on risk preferences by employing a machine learning approach to examine the influence of behavioral traits on financial risk aversion.

2.2.1.b. Measuring personality with the Big Five personality traits

The assessment of personality traits often relies on the widely accepted five-factor model, known as the Big Five. This model, proposed by McCrae & Costa in 1987, comprises five major domains: Agreeableness, Conscientiousness, Extraversion, Neuroticism, and Openness to experience. It has been extensively utilized in personality research due to its comprehensive nature and ability to capture commonalities among existing systems of personality traits (Soto and John, 2017).

The Big Five approach has gained popularity in personality research, showing agreement on the five personality traits across various cultures and age groups (De Fruyt et al., 2009). The dimensions have been found to remain relatively stable throughout adulthood and have a genetic basis (Digman, 1989; Vukasović and Bratko, 2015). Additionally, the Big Five traits have been linked to various outcomes, such as job performance, occupational status, and academic achievement (Roberts and Robins, 2000; Tett et al., 2013).

To assess personality traits, researchers have developed tools like the BFI-10, a concise questionnaire comprising ten statements to evaluate each personality dimension (Bortoli et al., 2019; Rammstedt et al., 2020). These tools enable the creation of individual profiles based on the prominence of each trait, allowing researchers to analyze the unique blend of personality traits in individuals.

Table 2.1 summarizes the key characteristics and measurement of each of the five personality traits.

Table 2.1: Five dimensions of measuring personality traits

Factor	High Scorer Characteristics	Factor Measures…	Low Scorer Characteristics
(N) Neuroticism	Worrying, nervous, emotional, insecure.	adjustment versus emotional instability. Identifies individual prone to psychological distress, unrealistic ideas, and maladaptive coping mechanisms.	Calm, relaxed, unemotional, hearty, self-satisfied
(E) Extraversion	Sociable, active, talkative, person-oriented, optimistic, fun-loving.	the quantity and intensity of interpersonal interaction, activity level, need for stimulation, and capacity for joy.	Reserved, sober, non-exuberant, aloof, task-oriented, quiet.

(O) Openness to experience	Curious, broad interests, creative, original.	amount of proactive-seeking, and appreciation of experience for its own sake, toleration for and exploration of the unfamiliar.	Conventional, down-to-earth, narrow interests, unartistic, unanalytical.
(A) Agreeableness	Soft-hearted, good-natured, trusting, helpful, gullible, straightforward.	the quality of interpersonal orientation along a continuum for compassion to antagonism in thoughts, feelings and actions.	Cynical, rude, suspicious, uncooperative, ruthless, irritable, and manipulative.
(C) Conscientious-ness	Organized, reliable, hard-working, punctual, neat, and persevering.	the degree of organization persistence, and motivation and goal-directed behavior. Contrasts dependable, fastidious people with those who are lackadaisical.	Aimless, unreliable, lazy, careless, negligent, weak-willed.

Source: The NEO Personality Inventory Manual (Costa and McCrae, 2008).

2.2.1.c. The Big Five personality traits in the financial sector

The association between Big Five personality traits and investment behavior has garnered significant attention in behavioral finance literature. Researchers have explored the link between individual personality characteristics and investment decisions, shedding light on the role of traits in shaping financial preferences. Hunter and Kemp (2004) examined the personality differences between investors in risky and "normal" firms, finding that individuals investing in e-commerce businesses displayed higher openness to experience, influencing their investment preferences. Similarly, Mayfield et al. (2008) identified that extroverted and conscientious individuals showed a greater inclination for short-term investing, while those with neurotic tendencies and risk aversion avoided short-term instruments.

Investigating the Big Five personality model, Oehler et al. (2018) provided experimental evidence that extroversion and neuroticism significantly influenced individuals' behavior, with extroverted participants paying more for financial assets and neurotic individuals retaining lower-risk assets. Chen, Ho, and Liu (2019) observed that investors with conscientious, amiable, extraverted, and open traits outperformed those with high neuroticism in the long run. Durand et al. (2008) noted that higher neuroticism and openness to experience corresponded to greater risk tolerance. Rustichini et al. (2012) revealed that neuroticism negatively impacted risk-taking in the domain of gains but had a diminished effect in the domain of losses, while conscientiousness influenced risk-taking attitudes. Bortoli (2019) also reported a positive relationship between risk propensity and openness to experience.

These studies collectively demonstrate the significance of personality traits in shaping risk preferences and investment decisions, providing valuable insights into the interplay of cognition and risk aversion in financial choices.

2.2.2. Cognitive theories and Cognitive Reflection Test

2.2.2.a. Cognitive ability and investment risk preference implications

Various studies have explored the relationship between risk preference and cognitive ability, indicating that cognitive capacity may play a role in shaping investment decisions. The "two-system" model proposes that risk preference can be influenced by both emotional personality-related responses (system-one) and intentional, calculating decision-making processes (system-two), making risk preference partially endogenous to cognitive capacity (Benjamin, Brown and Shapiro, 2013).

Moreover, research has shown that cognitive ability, including general intelligence and specific cognitive capacities, significantly affects decision-making (Lubinski and Humphreys, 1997; Frederick, 2005). Cognitive ability may impact investment choices through three main pathways. Firstly, a higher cognitive ability allows for better comprehension of financial concepts, reducing the barriers to entry in financial markets (Christelis, Jappelli and Padula, 2010). Secondly, cognitive capacity may influence risk preferences and the curvature of the value function, leading to variations in risk-taking behavior (Frederick, 2005; Benjamin, Brown and Shapiro, 2013; Dohmen et al., 2018).

Thirdly, individuals with lower cognitive abilities may be more susceptible to overconfidence, leading them to underestimate financial risks (Barber and Odean, 2001; Spaniol and Bayen, 2005).

The relationship between cognitive ability and risk aversion has significant implications for economic research and practical applications. Studies have shown that those with lower cognitive skills tend to be more risk-averse, impacting their investment decisions (Park, 2016, p.201; Dohmen et al., 2018). This knowledge is valuable for constructing screening models and designing contracts based on observable proxies for cognitive capacity. Additionally, cognitive ability's inclusion in reduced-form models and structural estimates can provide insights into the impact of risk aversion and impatience on economic decisions (Cawley, Heckman and Vytlacil, 2001).

In summary, the connection between cognitive ability and risk preferences in investment decisions offers valuable insights for understanding human behavior in financial contexts. The studies presented here emphasize the relevance of cognitive capacity in shaping risk preferences and decision-making processes, contributing to the growing field of behavioral finance.

2.2.2.b. Measuring cognitive capacity with Cognitive Reflection Test

The Cognitive Reflection Test (CRT) was developed by Frederick in 2005 to assess an individual's cognitive ability, particularly the capacity to suppress an instinctive incorrect answer in favor of a more reflective, accurate one. This ability is a vital component of decision-making, a concept validated in recent studies, such as those by Cueva et al. (2016), and used alongside the Big Five personality framework (Corgnet, Espín and Hernán-González, 2015).

Consisting of three items, the CRT measures the propensity to resist the initial impulsive response to provide a more thoughtful answer (Frederick, 2005). Despite its simplicity, CRT's impact is profound, as it underlines the importance of cognitive control in steering away from instinctive errors to find accurate solutions, thus providing invaluable insight into a person's cognitive mechanisms and decision-making competencies (Thoma et al., 2015).

Table 2.2: The Cognitive Reflection Test

(CRT1) A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost? £__________
(CRT2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? __________ minutes
(CRT3) In a lake, there is a patch of lily pads. Everyday, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? __________ days

Source: Cognitive Reflection and Decision Making (Frederick, 2005)

2.3. Investor Risk Profiles and Risk Preferences

Risk perception reflects an individual's subjective judgment of risk characteristics and severity, which is influenced by factors such as information source reliance, personal knowledge, market volatility, and regulatory measures (Brighetti et al., 2011; Linden, 2015). Each person's investment behaviour is governed by their risk profile or risk appetite, which signifies the level of risk an individual is willing to accept in pursuit of a potential return (Fehr-Duda and Epper, 2012; Pompian, 2016). Those with a high-risk preference, or risk-seeking individuals, are drawn to the potential for high rewards, accepting the associated risks. On the other hand, individuals with a low-risk appetite or risk-averse individuals prioritise stability and capital protection over high returns (Guiso, Sapienza and Zingales, 2018).

In certain circumstances, an investor may display risk neutrality, showing indifference towards uncertain outcomes and focusing on returns, making high-risk investments more appealing if they offer higher returns. This behaviour often occurs when the investment represents a minor portion of their wealth (Conlin, Lynn and O'Donoghue, 2003).

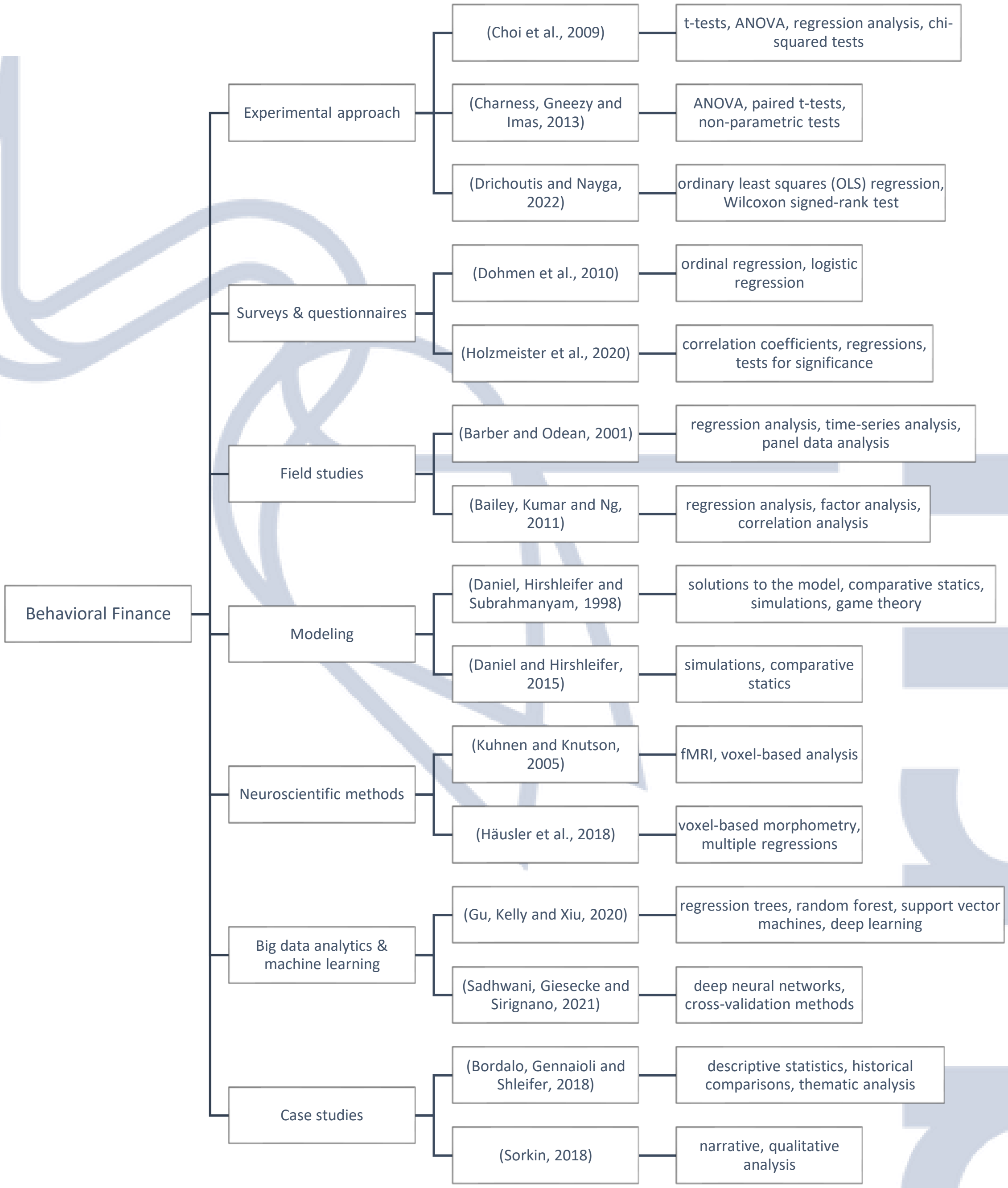
Risk aversion, the focus of this paper, prompts individuals to consistently avoid risky investment choices, even at the cost of experiencing potential losses. Risk aversion and materialism are found to influence the relationship between financial literacy and financial decision outcomes, potentially leading to premature divestment from profitable assets and unwarranted retention of losing ones (Hastings, Madrian and Skimmyhorn, 2013).

2.4. Applications of Machine Learning in Behavioral Finance Research

In behavioral finance research, some of the commonly used software to analyze data, perform statistical analyses, and model financial behaviors in behavioral finance include: SPSS, R (a free and open-source programming language and software environment for statistical computing and graphics), Stata, MATLAB (a programming language and computing environment), AMOS (a specialized software used for structural equation modeling (SEM)), etc. A bibliographic review on 348 articles published in 2011–2021 from journals indexed in the Scopus Q1 and Q2 finance journal database revealed an upward trajectory in the publication trend beginning in 2015 and identified the application of Artificial Intelligence and Machine Learning in bankruptcy prediction, stock price prediction, portfolio management, oil price prediction, anti-money laundering, behavioural finance, big data analytics, and blockchain (Ahmed et al., 2022). In behavioural finance, the use of machine learning and artificial intelligence would enable more sophisticated data analysis, personalised financial services, enhanced risk management, and improved decision support. These techniques can considerably contribute to a deeper understanding of the influence of personality traits on financial decision-making and assist investors in making better-informed and more efficient investment decisions.

Figure 2.2 on the next page summarises the behavioral finance taxonomy. The experimental approach and field studies have been foundational. While the former, conducted in controlled environments, helps test specific financial behavior hypotheses through games or simulations, the latter observes these behaviors in natural, real-world settings, such as the trading patterns on stock markets or consumer purchasing habits. Surveys and questionnaires, on the other hand, tap directly into individuals' experiences, beliefs, and attitudes, providing a snapshot of how people navigate financial decisions in diverse settings.

Figure 2.2: Behavioral finance taxonomy



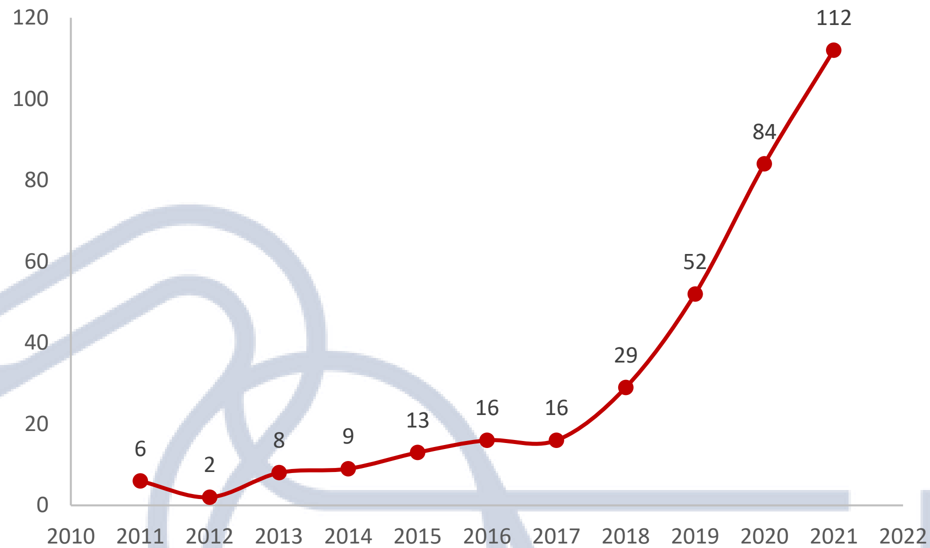
Delving deeper, we have modeling and neuroscientific methods. Models in behavioral finance aim to capture the real essence of decision-making, blending traditional economic theories with psychological elements to explain observed behavioral biases. Neurofinance, a relatively newer avenue, explores the neural underpinnings of these decisions. Techniques like fMRI scans delve into the brain's processes during financial decisions, marking the intersection of finance with neurobiology. Meanwhile, case studies offer granular insights, presenting deep dives into specific financial behaviors, institutions, or phenomena, enriching the broader understanding of the field.

Yet, amidst these traditional methodologies, there's a discernible gap: the application of machine learning and big data analytics. With the explosion of data in today's digital age, there's an unprecedented opportunity to delve deeper. While traditional methods often work with limited datasets, machine learning algorithms can process vast amounts of data, identifying intricate patterns that might be invisible to the human eye. This not only can provide macro-level insights but can also personalize financial strategies for individuals based on historical data. Despite its potential, the use of machine learning in behavioral finance is still in its nascent stages, presenting a golden opportunity for future research.

Traditional statistical methods, such as regression analysis, time series analysis, econometric models and hypothesis testing, have long been prevalent in behavioral finance research. Researchers might be more familiar with these methods and prefer their interpretability over complex machine learning algorithms. These methods have a long history in finance and economics and are well-understood by researchers and practitioners.

However, as machine learning has gained popularity in various domains, including finance, there has been growing interest in applying machine learning algorithms to behavioral finance problems. Machine learning algorithms offer the potential to uncover complex patterns, relationships, and nonlinearities in large datasets, which may not be easily identifiable using traditional methods.

Figure 2.3: Artificial Intelligence & Machine Learning in finance literature growth



Source: Artificial intelligence and machine learning in finance: A bibliometric review (Ahmed et al., 2022)

There is vast potential of machine learning to analyze complex financial data and reveal insights into the investor behavior. As datasets become more accessible and researchers gain proficiency in these methods, we can expect a growing interest in leveraging machine learning algorithms to gain deeper insights into investor behavior, market dynamics, and financial decision-making processes. The use of machine learning and artificial intelligence in behavioral finance allows for more sophisticated data analysis, personalized financial services, improved risk management, and enhanced decision support. These techniques can significantly contribute to a deeper understanding of the influence of personality traits on financial decision-making and help investors make more satisfactory investment choices.

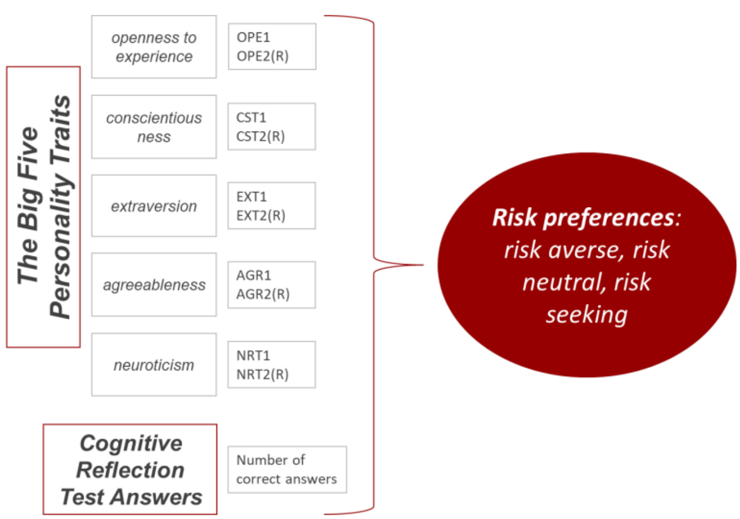
Chapter 3

Methodology

3.1. Research design and rationale

The primary aim of this study is to elucidate the association between individual attributes, namely, personality traits and cognitive reflection, and their influence on risk aversion within investment decisions. Figure 3.1 delineates the research methodology employed in this study. Out of the 11 independent variables, 10 belong to the Big Five Inventory. These traits are assessed through a structured questionnaire, with each trait represented by two items: one directly scored and the other reverse-scored. This approach ensures a thorough and precise evaluation of each respective trait. The remaining independent variable is derived from the Cognitive Reflection Test (CRT). The metric for this variable is derived from the aggregate of accurate responses on the CRT, which signifies the respondent's degree of cognitive reflection.

Figure 3.1. Research model



Source: Compiled by the author

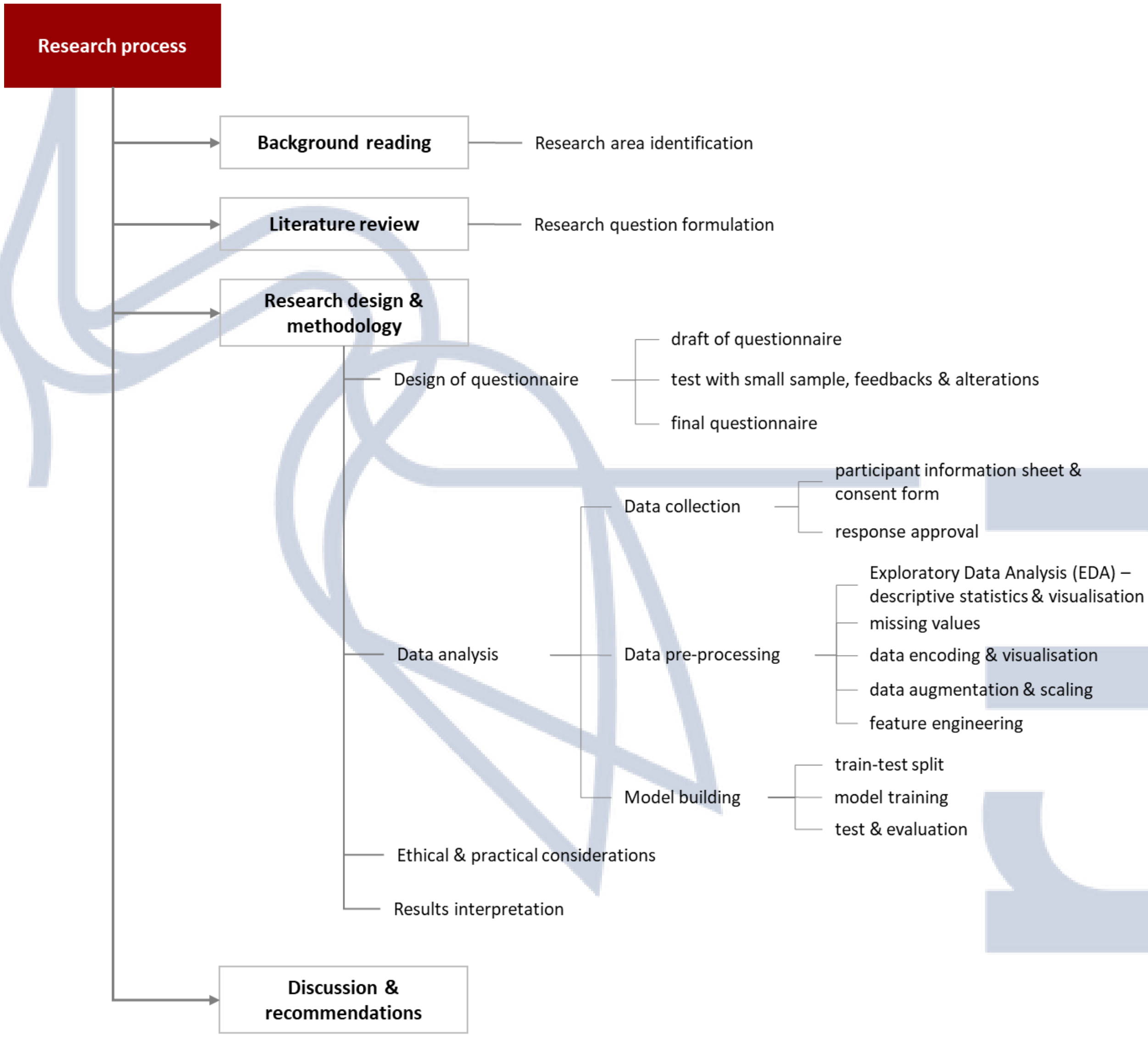
The dependent variable measures an individual's propensity for risk aversion in investment contexts. This study seeks to ascertain the extent to which the independent variables influence this specific trait.

The methodological emphasis on both personality and cognitive reflection is predicated on the postulation that these elements significantly inform investment behavior, as corroborated in section 2 of the literature review. Utilizing a combination of directly scored and reverse-scored items within the questionnaire is designed to bolster the response reliability and attenuate potential response bias. Through examining the interrelationships between these independent variables and risk aversion, this investigation endeavors to furnish a comprehensive insight into the determinants underpinning investment decision-making.

3.2. Research process

Figure 3.2 provides an overview of the research journey in sequential order. It begins with Background Reading, an exploratory phase that examines existing literature to determine the primary research topic. The Literature Review that follows refines previously obtained insights and facilitates the formulation of primary research questions. Research Design & Methodology constitute the core of the study. This encompasses the detailed blueprint of the *Design of Questionnaire* for primary data collection, as outlined in Section 3.2.1. The subsequent breakdown under *Data Analysis* sheds light on the methods and sources of *Data Collection*, definitions and metrics of *Variables & Measurement*, and the architectural framework that structures data flow, *Data Pipeline Framework*. The research employs specific computational models, elaborated in *Machine Learning Models to be Applied*, and the performance of these models will be gauged based on the *Evaluation Metrics* detailed in Section 3.2.6. The process culminates with a contemplative reflection on *Ethical and Practical Considerations* in Section 3.3, which delves into the ethical nuances and potential real-world challenges of the study. Chapter 4, "Results Interpretation," delves into the evaluation metrics of the applied machine learning models to understand their efficacy and accuracy. Chapter 5, "Discussion & Recommendation," synthesizes findings to offer actionable insights for stakeholders, accompanied by recommendations to harness optimal investment behaviors.

Figure 3.2: Research process diagram



Source: Compiled by the author

3.2.1. Design of questionnaire

In this study, the selected variables encompassing individual behavioral traits and risk aversion have been extensively cited and recognized within the realms of managerial and organizational psychology scholarship. The formulation of the questions was underpinned by a meticulous review of pertinent literature, coupled with a comprehensive comprehension of investor psychology.

The structured questionnaire was delineated into four distinct sections. The inaugural section was dedicated to components associated with the Big Five personality traits. Due to the potential advantages of reducing research expenditures, enhancing participation rates, and streamlining the survey administration process, the BFI-10 scale, as proposed by Gosling, Rentfrow, and Swann (2003), was adopted.

Table 3.1: Big Five Inventory questionnaire

Code	Item	Agreement level					
Big Five Inventory: I see myself as…							
EXT1	Extraverted, enthusiastic.	1	2	3	4	5	6
AGR1	Critical, quarrelsome.	1	2	3	4	5	6
CST1	Dependable, self-disciplined.	1	2	3	4	5	6
NRT1	Anxious, easily upset.	1	2	3	4	5	6
OPE1	Open to new experiences, complex.	1	2	3	4	5	6
EXT2(R)	Reserved, quiet.	1	2	3	4	5	6
AGR2(R)	Sympathetic, warm.	1	2	3	4	5	6
CST2(R)	Disorganised, careless.	1	2	3	4	5	6
NRT2(R)	Calm, emotionally stable.	1	2	3	4	5	6
OPE2(R)	Conventional, uncreative.	1	2	3	4	5	6

"R" denotes reverse-scored items.

Source: Rammstedt and John (2007)

The Cognitive Reflection Test (CRT) was then presented, followed by a question regarding the participant's previous interaction with the test. Afterwards, a section comprised of four queries targeted risk aversion specifically. Adapted from metrics created by Gomez-Mejia and Balkin (1989) and Mayfield et al. (2008), these questions are germane to financial decisions. The final section of the questionnaire requested respondents' demographic information.

To enhance the validity of the questionnaire, preliminary in-depth interviews were conducted with five seasoned investors, each with more than five years of investment experience. Their feedback provided crucial insights into prospective survey question modifications. After modifications, the questionnaire was pilot-tested on a sample of fifty participants. This pilot phase ensured precise data processing, facilitated the accumulation of feedback, and allowed for further modifications to ensure comprehension. The finalised survey was then prepared for distribution to the primary sample group.

According to Fisher (2010), the six-point Likert scale, a renowned instrument for assessing respondents' attitudes and perceptions, was a major component of the survey. Individual investors were invited to indicate their level of agreement with respect to behavioural traits, risk propensities, and investment goals using this scale. The scale's increments ranged from 1 to 6, representing a continuum from "totally disagree" to "totally agree."

3.2.2. Data collection

According to Durand *et al.*, personality is a "motivator" of human conduct. A growing body of evidence supports the claim that individual investment decisions are intricately linked to personality traits. Existing literature examines in depth the relationship between personality and the behaviour of established investors, but the effects of these traits on individuals outside of the investment industry remain unexplored. Consequently, this research also surveys prospective investors, such as those who are unfamiliar with financial markets or who possess basic knowledge but refrain from active participation.

There is little research on the impacts of personality factors on behavioral intentions through cognitive inflection. According to Mayfield *et al.*, behavioral intentions, which represent an individual's propensity towards a particular behaviour, are predominately cognitive in nature. Given the importance of CRT scores to this study and cognizant of cognitive reflection as an innate disposition emphasising restrained and reflective decision-making, respondents with prior CRT exposure were excluded from the sample pool.

To collect relevant data, the study employed a carefully crafted online questionnaire, as described in section 3.2.1. The survey was distributed via multiple channels, such as social media, investor groups, undergraduate cohorts, and professional networks. Participation was completely optional. Beginning on 10 July 2023 and concluding a month later on 10 August 2023, 350 participants with consent responded to the survey. After filtering, which included the exclusion of 29 respondents familiar with the CRT, a coherent dataset consisting of 321 respondents was curated for subsequent analysis.

3.2.3. Variables and measurements

Table 3.2: Description of study variables

Factor name & sign	Observed variables	Range	Description
Neuroticism (NEU)	NEU1, NEU2R	1 – 6	High scores indicate tenseness, moodiness, anxiety, and insecurity
Extraversion (EXT)	EXT1, EXT2R	1 – 6	High scores indicate assertiveness, sociability, talkativeness, optimism, and being upbeat and energetic
Openness to experience (OPE)	OPE1, OPE2R	1 – 6	High scores indicate an active imagination, aesthetic sensitivity, a preference for variety, intellectual curiosity, and broad cultural interest
Agreeableness (AGR)	AGR1, AGR2R	1 – 6	High scores indicate altruism, personal warmth, sympathy towards others, helpfulness, and cooperation
Conscientious ness (CST)	CST1, CST2R	1 – 6	High scores indicate purposefulness, being strong-willed, determination, organization, reliability, and punctuality

CRT score (CRT)	CRTRUNre	0 – 3	The number of correct answers in the CRT
Risk aversion (RPR)	RPR1, RPR2, RPR3, RPR4	1 – 6	High scores indicate high risk aversion/avoidance and low risk-seeking attitude towards financial investing

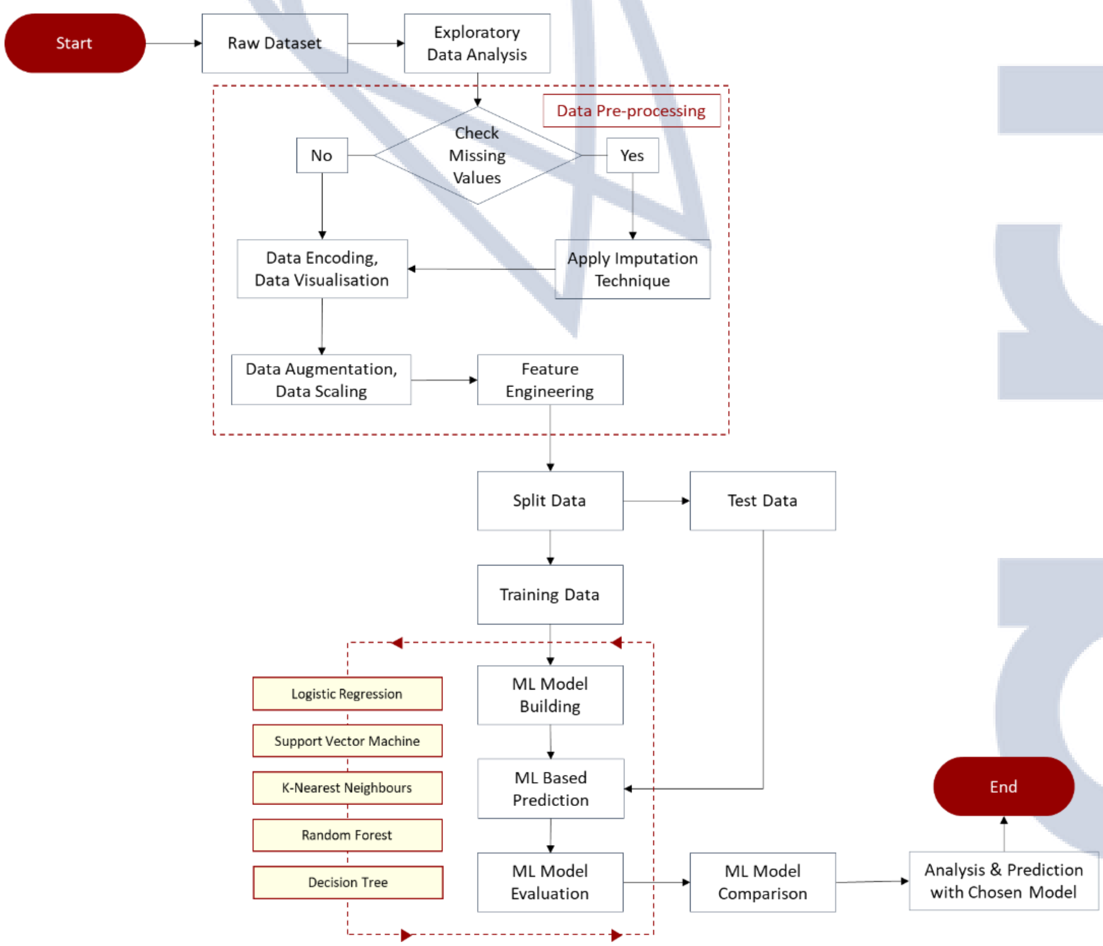
Source: Compiled by the author

Table 3.2 provides a comprehensive breakdown of the variables of the study, elaborating on the nuances and characteristics of each. Under the column Factor Name & Sign, each variable's designated name is listed alongside its associated sign. The Observed Variables column describes the specific manifestations of these factors as documented by the research data, thereby providing insight into the observable metrics being evaluated. For example, NEU1 and NEU2R represent questionnaire items that portrait Neuroticism. The Range column delineates the minimum and maximum values or categories that each variable can assume. Lastly, the Description column functions as an explanatory guide by providing context-specific explanations for each variable, thereby ensuring that its indications of variable scores are made clear. In essence, this table serves as a map, navigating readers through the landscape of study variables and their characteristics. The fundamental level of the research model comprises the variables NEU, EXT, OPE, AGR, CST, and CRTRUN, which are examined for their effect on RPR.

3.2.4. Data pipeline framework

This procedure, illustrated in figure 3.3, begins with Exploratory Data Analysis (EDA), a preliminary foray into the unprocessed data that reveals its structural complexities, anomalies, recurring patterns, and salient characteristics. Following this preparatory work, the data goes on a preprocessing regimen. This juncture encompasses a plethora of tasks, ranging from addressing missing data to ensure its completeness to data encoding, which translates categorical variables into machine-readable formats, to data scaling, which normalises the range of independent variables or features, and feature engineering, which creatively reshapes or augments features to boost their predictive efficacy. Additionally, data augmentation techniques such as resampling, employing both oversampling via SMOTE and bootstrapping combined with noise injection, have been incorporated to balance the dataset and improve its robustness. The data is then divided into training and testing subsets using the train-test split mechanism, thereby preserving an unexplored segment for model validation. Within the parameters of this study, five distinct classification paradigms listed in Section 3.2.5 are developed and elaborated upon. Each model is then analysed through the lens of particular evaluation metrics, as described in section 3.2.6. The pipeline concludes with the model comparative evaluation, which contrasts each model's efficacy to determine the archetype that best resonates with the dataset's nuances. This data infrastructure underpins a holistic approach to data management and examination, refining the path from raw data to insightful predictions. The code corresponding to these steps can be found in Appendix B of this paper.

Figure 3.3: Data pipeline framework flow chart



Source: Compiled by the author

3.2.5. Machine learning models to be applied

Embarking on an analytical journey, the task is to scrutinize the influence of behavioral traits on financial risk aversion using a machine learning approach. Given the dataset's complexities with 11 independent variables, the quintessential objective is to classify data into distinct categories, ensuring that the patterns within the data are captured and translated into actionable insights. This mandates the application of potent machine learning algorithms tailored for multiclass classification tasks. The methodology aligns with cutting-edge research, leveraging proven techniques adapted for optimal performance on the current dataset.

(i) Multinomial Logistic Regression

Extending binary logistic regression, this technique thrives in multiclass classification scenarios. Due to its simplicity and efficacy, this linear model is frequently an excellent starting point for moderate datasets (Hosmer Jr., Lemeshow, and Sturdivant, 2013). L1 (Lasso) regularisation techniques will be utilised to prevent overfitting, particularly when a dataset is relatively small. Cross-validation, specifically k-fold cross-validation, can be employed to tune hyperparameters like the regularization strength, ensuring the model's robustness and generalization capability (Zhang, Oles, and Vera, 2020).

(ii) Decision Trees

Classification and Regression Trees (CART) algorithms are adept at providing clear decision criteria. Their versatility is showcased through their ability to handle both categorical and numeric data seamlessly. The essence of these algorithms lies in recursively partitioning the dataset based on features to attain the highest classification purity. Notably, CART trees are inherently interpretable: the underlying logic of each decision can be explicitly understood, offering insights into the importance of each variable in decision-making processes. Decision trees are interpretable, which means that the logic underlying each classification decision can be observed, making them useful for understanding the importance of variables (Quinlan, 2014). Pruning mechanisms, which play a crucial role in reducing tree complexity and enhancing predictive accuracy, strengthen their dependability (Quinlan, 2018).

(iii) Random Forest

As an ensemble of decision trees, Random Forest constructs an ensemble of trees, aggregating their predictions to improve performance and reduce overfitting. This ensemble method handles overfitting better than a single decision tree, which can be especially beneficial when dealing with limited data, as it facilitates generalisation to unseen data. Random Forests also produces a measure of "feature importance," which can be useful for determining which variables have the most impact on the prediction (Breiman, 2001). Hyperparameters, such as the maximum depth of the trees and the minimum samples required to divide, can be fine-tuned using cross-validation to ensure robustness (Probst, Wright, & Boulesteix, 2019).

(iv) Support Vector Machines (SVM)

Support Vector Machines are appropriate for multiclass classification, specifically when using the 'one-vs-one' or 'one-vs-all' techniques. SVMs identify the hyperplanes in a high-dimensional space that determine optimal boundaries and demarcate distinct classes most effectively. Even though SVMs have a firm theoretical foundation based on the Vapnik-Chervonenkis (VC) theory, as demonstrated by numerous empirical studies (Cortes and Vapnik, 1995), they are also robust in practise. Regularisation in SVMs is intrinsically determined by the selection of the C parameter, which determines the trade-off between maximisation of the margin and minimization of classification errors. Cross-validation will be used to determine the optimal C value and kernel selection (Boser, Guyon, and Vapnik, 2019). The significance of hyperparameter optimisation has been emphasised in recent research (Albon, 2020), particularly in terms of kernel selection and the C parameter.

(v) k-Nearest Neighbors (k-NN)

k-NN is a straightforward, non-parametric, and "lazy" multiclass classification algorithm. A data point is classified based on how its neighbours are classified. Despite its simplicity, k-NN can be remarkably accurate, particularly when the number of variables (11 in this case) is not excessively large, thereby avoiding the curse of dimensionality. The efficacy of the k-NN algorithm relies heavily on the underlying distance metric. As emphasized in recent studies, normalization of features and hyperparameter optimization, especially the choice of k, is paramount for accuracy (James, Witten, Hastie, and Tibshirani, 2021). Cross-validation, as suggested by Alpaydin (2020), can optimize these hyperparameters.

In conclusion, the selection of these finely tuned algorithms offers a balanced combination of linear and non-linear methods, interpretable and black-box models, and parametric and non-parametric approaches. All of these classification algorithms have been extensively acknowledged in academic literature for their ability to handle classification tasks of varying complexity.

3.2.6. Evaluation metrics

In any classification assignment involving machine learning, it is crucial to objectively assess the predictive model's performance. Even though absolute accuracy is essential, it does not always provide a complete picture, particularly when classes are unbalanced or when various categories of misclassification have varying consequences. To gain a comprehensive comprehension of the efficacy and development opportunities of the applied models, a suite of evaluation metrics is employed. These metrics cast light on various aspects of the model's performance, ensuring that its predictions are both robust and reliable. To evaluate the predicted outcomes, the following evaluation metrics will be specified:

- Confusion matrix: a table presenting the actual vs. predicted classifications, used to evaluate the performance of a classification model.

Table 3.3: Confusion matrix for 3-class classification

		True Class		
		C1	C2	C3
Predicted Class	C1	C11	C12	C13
	C2	C21	C22	C23
	C3	C31	C32	C33

Source: Compiled by the author

Where:

- C11, C22, C33 are the number of true predictions for each class.
- C12, C13 are the number of Class 1 instances wrongly predicted as Class 2 and Class 3 respectively. Similarly for other off-diagonal elements.

- Accuracy score: the ratio of correctly predicted instances to the total number of instances. In a 3-class system, it's the sum of the diagonal elements of the confusion matrix divided by the total number of samples. Accuracy score provides a general measure of how well the classifier predicts the correct categories.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$$

- Precision: the fraction of correctly identified cases for each class out of all cases predicted to be that class. High precision means that an algorithm returned substantially more relevant results than irrelevant.

$$Precision\ (for\ a\ specific\ class) = \frac{Number\ of\ True\ Positives\ for\ the\ class}{Total\ Number\ Predicted\ to\ be\ that\ class}$$

- Recall (sensitivity): the fraction of actual positives that the model correctly identified for each class out of total actual instances of that class. It measures the model's ability to identify all positive cases. High recall indicates the model can detect most of the positive cases.

$$Recall\ (for\ a\ specific\ class) = \frac{Number\ of\ True\ Positives\ for\ the\ class}{Total\ Actual\ Instances\ of\ that\ class}$$

- F1-score: the harmonic mean between precision & recall. An F1-score is in the range [0, 1]. It is a way to summarize the model's accuracy and robustness in a single number. An F1-score closer to 1 indicates better classification performance.

$$F1 - score = \frac{2}{\frac{1}{precision} + \frac{1}{recall}}$$

A combination of evaluation metrics is a particularly accurate method for determining a machine learning model's efficacy. Each metric provides a unique perspective on the performance of the model, from precision to nuanced measurements such as the F1-score. By utilising these metrics, a comprehensive comprehension of model strengths and areas for improvement is attained, ensuring that predictions in practical applications are robust and reliable.

3.3. Ethical and practical considerations

Maintaining the ethical integrity of this research is paramount, especially in the realm of participant data protection and privacy. In alignment with the General Data Protection Regulation (GDPR), several procedures have been instituted to uphold participants' rights and privacy:

- Pseudoanonymisation: Personal identifiers are substituted with unique codes, ensuring participant confidentiality while permitting required data linkage by the researcher.
- Data Usage: Only essential personal data will be harnessed to meet research objectives.
- Data Processing: All data operations are executed within UK boundaries, adhering strictly to its legal framework. For participants from the European Economic Area (EEA), it's notable that UK's data protection standards mirror EEA's provisions.
- Security Protocols: Data will be stored on password-protected systems or encrypted devices, accessible solely to authorized personnel. Secure channels, such as encrypted transfers, will be employed for data transmissions.
- Data Sharing: On occasions, pseudoanonymized data might be shared with key stakeholders like project supervisors to ensure research rigour.
- Anonymity: Rigorous efforts will be exerted to uphold participant anonymity in all research publications.
- Data Review: Participants may be granted the chance to vet and verify certain data, like interview transcripts, before dissemination.
- Online Surveys: Transparency is maintained through clear links to the privacy policies of both the survey platform and the academic institution.

Every procedure has been meticulously crafted to prioritize participants' rights, trust, and comfort. This rigorous ethical adherence not only underpins the integrity of the research but also ensures its acceptance and trustworthiness in the academic community.

Chapter 4

Results & Discussion

4.1. Description of sample respondents

321 questionnaire respondents' descriptive data is presented in table 4.1.

Table 4.1: Descriptive statistics of respondents

	Frequency	Percent
Gender		
Male	213	66.4
Female	108	33.6
Age		
18-25 years	143	44.5
26-35 years	86	26.8
35-50 years	71	22.1
Above 50 years	21	6.5
Academic level		
High school diploma	107	33.3
Bachelor's degree	94	29.3
Master's degree	84	26.2
Doctoral degree	30	9.3
Professional degree	6	1.9
Annual income		
Under £20,000	68	21.2
£20,000 – £40,000	90	28.0
£40,000 – £60,000	102	31.8

£60,000 – £80,000	37	11.5
Above £80,000	24	7.5
Knowledge on financial investing		
Never heard about financial investment	23	7.2
Aware of financial investment but have no intention to participate	45	14.0
Aware of financial investment and have intention to invest in the future	113	35.2
Have participated in financial investment	140	43.6
Total	321	100.0

Source: Compiled by the author

The profile of the survey participants above shows that the sample was diverse in terms of demographics, academic level, income, as well as financial investing knowledge and intention. The majority of the respondents were male (66.4%), relatively consistent with the current financial investing market's gender distribution (Anon., 2021). The respondents are of all ages above 18, with the highest proportion accounted by 143 people from the 18-25 age group (44.6%), which somewhat parallels the high number of undergraduates (107 people). Only 6.5% were above middle age, and the rest are adults between 25 and 50 years old. Almost 30% of the respondents had a degree of Bachelor; one thirds of all respondents had studied higher education. The predominant income group is £40,000 – £60,000 per annum, which was owned by more than 30% of respondents. 28% earned from £20,000 to £40,000 annually, 21% under £20,000, and the rest earned £60,000. The respondents' knowledge and experience regarding financial investing also varies: they either had knowledge and/or intention or not, and either have actively invested or not. This satisfies the audience target stated in section 3.2.2. of this paper. The majority (43.6%) participated in active investing. Among the rest, more than 35% were potential investors.

4.2. Evaluation of model performances

The analysis presented in this section provides a meticulous evaluation of the performances of five distinct machine learning algorithms, namely, Multinomial Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and k-

Nearest Neighbors (k-NN). A granular examination of key metrics such as accuracy, precision, recall, and the F1-score will be delivered for each method, leading to a holistic comparative analysis.

4.2.1. Multinomial Logistic Regression

Performance Metrics: Utilizing a hyperparameter value of **C=0.1**, the Multinomial Logistic Regression yielded a training accuracy of 0.74 and an impressive testing accuracy of 0.78. The model's precision, recall, and F1-score metrics across the three classes unveiled valuable insights. Specifically, class 3.0 demonstrated superior precision and recall, highlighting the model's adeptness in recognizing this category.

Interpretation: The relatively balanced precision and recall metrics across the three classes underscore the model's ability to impartially distinguish among the categories, effectively minimizing biases. This balance suggests the algorithm's robust capacity to handle the intrinsic complexities of the dataset and its nuances. The success in classification of class 3.0 might be attributed to distinct patterns or features that make it more discernible.

4.2.2. Decision Tree

Performance Metrics: The Decision Tree algorithm, when evaluated through cross-validation, exhibited an average accuracy of 0.71. When applied to the dataset, the training and testing accuracies were noted as 0.74 and 0.75, respectively.

Interpretation: The Decision Tree model is inherently interpretable, delineating clear decision boundaries based on the features. While the recorded accuracy isn't the pinnacle amongst the evaluated algorithms, the model's transparency in decision-making stands out. Moreover, the balance across precision and recall for each class indicates its ability to treat each class with equal sensitivity.

4.2.3. Random Forest

Performance Metrics: Among ensemble techniques, Random Forests have gained prominence due to their efficacy, and this study reaffirms this. Exhibiting a mean cross-validation accuracy of 0.80, and subsequent training and testing accuracies at 0.85 and 0.86, the model's performance is commendable.

Interpretation: Random Forest's superiority stems from its design, which aggregates multiple decision trees, reducing the chances of overfitting. The stellar testing accuracy alongside the training accuracy suggests the model's generalized nature. Especially noteworthy is the model's efficacy in classifying instances of class 1.0, suggesting possible distinct features within this class facilitating more accurate classifications.

4.2.4. Support Vector Machine (SVM)

Performance Metrics: SVM, with its inherent ability to determine optimal hyperplanes, achieved a mean cross-validation accuracy of 0.74. This manifested as training and testing accuracies of 0.74 and 0.78, respectively, in our evaluations.

Interpretation: The SVM showcased a particular proclivity towards class 3.0, achieving stellar precision and recall metrics for this class. Its holistic performance, while commendable, suggests there might be room for further enhancements, possibly through refined feature engineering or optimal kernel selection. SVM's mathematical rigor in distinguishing classes, especially in high-dimensional spaces, reaffirms its potential for such classification tasks.

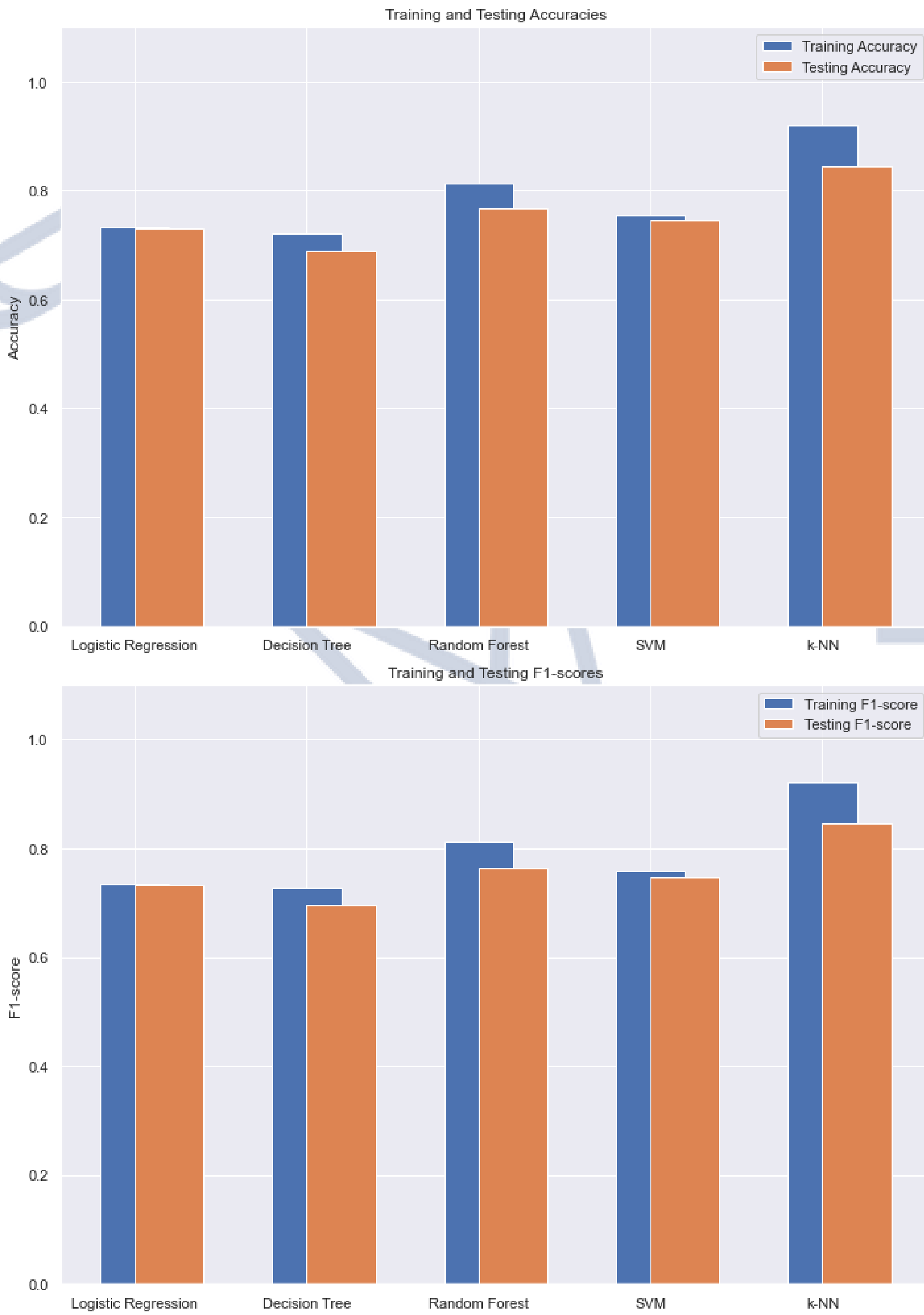
4.2.5. k-Nearest Neighbors (k-NN):

Performance Metrics: Emerging as one of the top contenders in this analysis, k-NN demonstrated a mean cross-validation accuracy of 0.86. Additionally, the algorithm achieved remarkable training and testing accuracies of 0.93 and 0.88, respectively.

Interpretation: The inherent simplicity of k-NN, which classifies based on proximity to neighbors, coupled with its unparalleled accuracy, makes it an exceptional model for this dataset. The impressive precision and recall metrics across classes indicate the model's capability to classify instances with high fidelity. However, while its performance is praiseworthy, k-NN doesn't inherently provide insights into feature importance or inter-relations.

4.2.6. Comparative Analysis:

Figure 4.1: Comparison of evaluation metrics



Source: Generated by Python, as attached in Appendix B

When juxtaposing the performance metrics of the models:

- **Accuracy**: k-NN stands out as the most accurate model, closely rivaled by the Random Forest classifier. The consistent training and testing accuracies of both models hint at their robustness and reduced overfitting tendencies.
- **Precision & Recall**: While SVM and Logistic Regression exhibited certain biases towards specific classes, k-NN and Random Forest displayed well-rounded metrics, suggesting their adaptability and generalized nature for varied datasets.
- **Cross-Validation and Robustness**: Random Forest and k-NN, with their consistent scores, reflect reliability and robustness. The minimal standard deviations further accentuate their stability across diverse data splits.
- **Complexity vs. Interpretability**: While Random Forest offers a profound understanding of feature importance, it's more intricate due to its ensemble nature. On the contrary, simpler models like the Decision Tree and k-NN, while potent, differ in their interpretative capabilities.

The comprehensive metrics and analysis of five distinct machine learning algorithms applied to the dataset has provided valuable insights into their respective strengths and weaknesses. Recommending a suitable model for predictive purposes would require weighing these aspects while also acknowledging the limitations and potential areas of improvement.

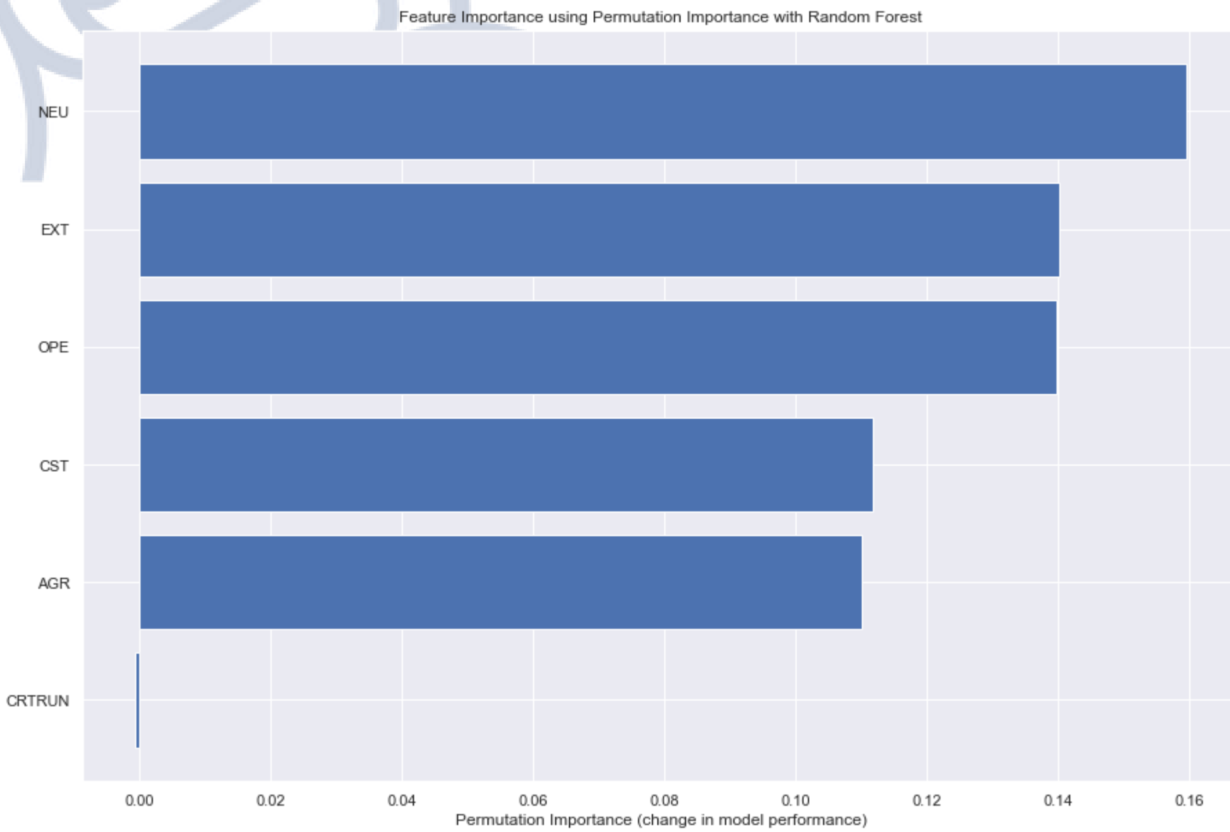
4.3. Feature importances

Where to put this? Our analysis's primary objective is to determine the significance of individual features in terms of their ability to predict the target variable using a Random Forest classifier. This is accomplished using a method known as permutation importance.

The permutation_importance function from sklearn.inspection is the central component of this methodology. Each feature's values are randomly shuffled, and the model's performance (e.g., accuracy) is then evaluated with this randomised feature. The extent of the decrease in model performance caused by the reorganisation provides insight into the significance of that feature. In particular, a significant decline in performance indicates that the feature is essential to the model's predictions. In contrast, a negligible change suggests that the feature may not be essential. This process is repeated 30 times for each feature to obtain a more accurate estimate, thereby ensuring the reliability of our results. random_state=42 has been set as the seed value to assure consistency and reproducibility of our findings.

4.3.1. Personality traits

Figure 4.2: Feature importance with Random Forest



Source: Generated by Python, as attached in Appendix B

The most influential predictor of risk aversion was found to be neuroticism (NEU), with an importance value of 0.159583. This aligns with prior academic research which suggests that individuals with high neuroticism tend to exhibit a heightened sensitivity to negative stimuli and are, therefore, more risk-averse (Lerner and Keltner, 2000). This emotional reactivity could explain the reluctance of such individuals to invest in volatile assets like stocks, even when potential returns are lucrative. Downturns in stock markets, accentuated by the media, could further exacerbate this aversion (Bajaj, 2008).

Extraversion (EXT) emerged as the second most influential trait with an importance score of 0.140167. While traditionally linked with outgoing and risk-seeking behaviors (McCrae and Costa, 1987), this study suggests that in financial contexts, extraverts might also exhibit caution. Such individuals, while comfortable in social interactions, might be influenced by the collective sentiments of their peers. As the Wall Street Journal (Banerji and Wallerstein, 2023) highlighted, collective investor sentiment, often driven by the majority's extraverted voices, plays a pivotal role in stock market fluctuations.

Close to extraversion, the importance of openness to experience (OPE) stood at 0.139750. Historically, those high in openness have been associated with a willingness to embrace new and unconventional investments (Caspi and Roberts, 2001). Their appetite for the unfamiliar might translate to exploring diverse financial instruments. However, the global bond market's unpredictability (Anon., 2012) could make even the most open-minded investors cautious.

Conscientiousness (CST) and agreeableness (AGR) followed with importance values of 0.111750 and 0.110167, respectively. These findings resonate with studies by Goldberg (1993) which postulated that highly conscientious individuals, given their meticulous nature, prefer well-structured investments, like bonds. Agreeable individuals, known for their cooperative and harmonious nature, might be influenced by societal norms and avoid investments perceived as 'risky' by the community.

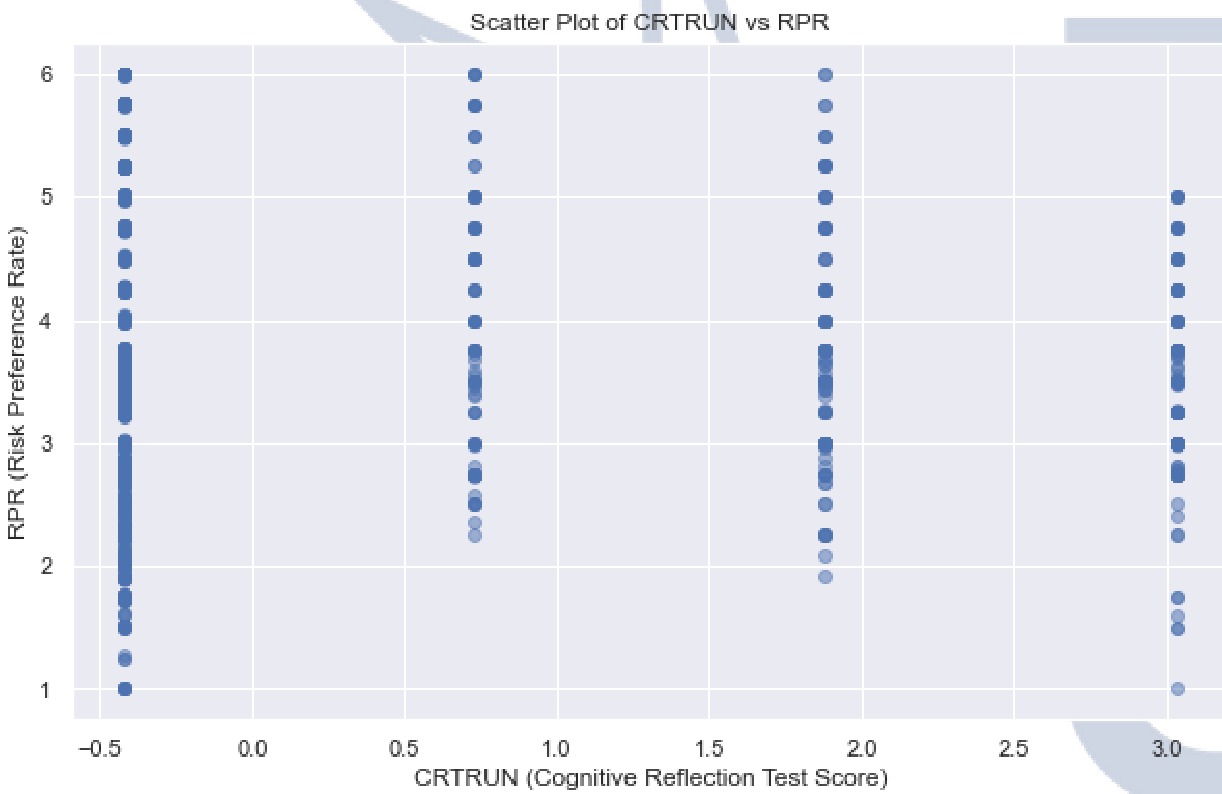
Surprisingly, the number of correct answers in the cognitive reflection test (CRTRUN) showed a marginal negative importance (-0.000583). This suggests that while cognitive reflection is crucial for complex problem-solving, when it comes to risk aversion in financial decisions, emotional and personality traits overshadow pure cognitive abilities. This observation aligns with Kahneman's (Kahneman, 2011) seminal work that emphasized the often dominant role of intuitive, fast thinking (System 1) over deliberate, slow thinking (System 2) in financial choices.

To conclude, the study underscores the profound influence of personality traits on financial decision-making. In a world where stock and bond markets are increasingly volatile, understanding these intricate psychological dynamics can offer invaluable insights for financial planners, investors, and policy-makers.

4.3.2. Cognitive reflection test score

Given that the Cognitive Reflection Test (CRTRUN) score exhibited minimal feature importance in our primary analysis, this intriguing observation prompted a deeper investigation. On the surface, the CRTRUN may not appear to be particularly influential in predicting risk aversion. However, subtleties are frequently found in the particulars. Using a scatter plot to examine the distribution and relationship between this characteristic and risk aversion, we sought to identify any nuanced patterns or potential outliers.

Figure 4.3: Scatter plot CRTRUN vs RPR



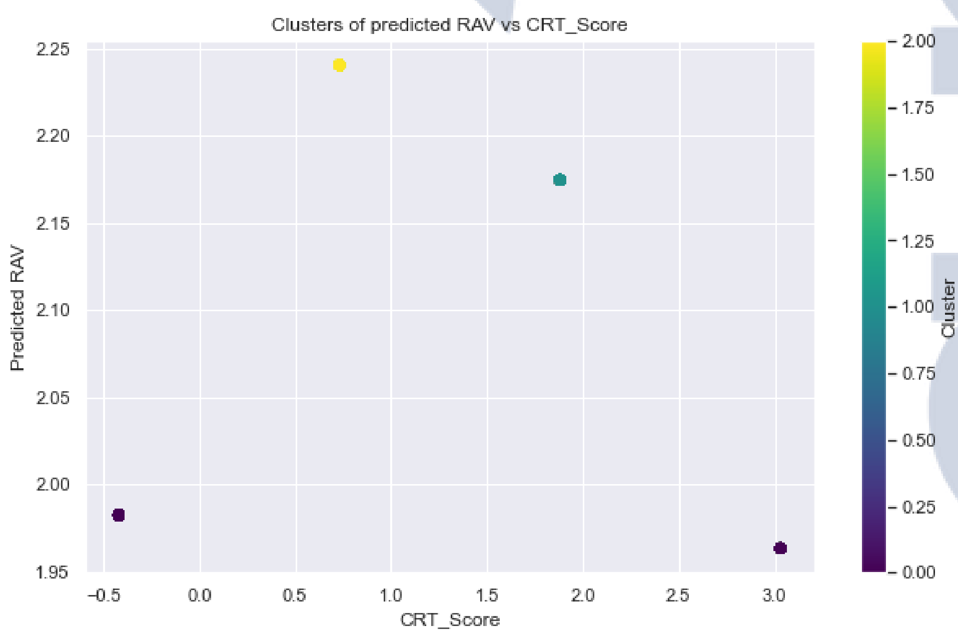
Source: Generated by Python, as attached in Appendix B

The scatter plot visualisation revealed a nonlinear relationship between the Cognitive Reflection Test score and the predicted level of risk aversion. To further dissect this relationship, a RandomForestRegressor was utilised. This model, which is renowned for its capacity to capture complex relationships and interactions between features, was trained using only the CRT scores.

Where to put this? Below is a concise summary of the methodology:

- The dataset was divided into training and testing subsets to assure the accuracy of predictions.
- Initialization and training of a RandomForestRegressor on the training set.
- Risk aversion predictions were generated for the entire dataset.
- Data Scaling: The predicted values were standardised using StandardScaler, rendering them suitable for clustering.
- Clustering with DBSCAN: The DBSCAN clustering algorithm was used to better comprehend the groupings. The objective was to identify high-density regions representing distinct risk aversion patterns or clusters based on CRT scores.
- To illustrate the relationship between CRT scores and predicted risk aversion values, a scatter diagram with color-coded clusters was generated.

Figure 4.4: Clusters of predicted RPR vs CRT score



Source: Generated by Python, as attached in Appendix B

People with no correct answers or all correct answers are significantly more risk-averse than those who got 1 or 2 questions right. The greatest difference emerges as the number of correct answers on the CRT test goes from nothing to something, raising risk aversion by almost 1 point. Increasing CRT score from 1 to 2 correct answers does not make a significant difference; however, increasing it to 3 corrects drops risk aversion again down to the level of risk aversion for no correct answer.

Individuals with highest cognitive reflection tended not to avoid risk in investment decisions, but were not daring either: they have a rather moderate profile. Those on the no-correct-answer side of the CRT scale were the most non-risk-averse. With such low tendency to detect errors and resist reporting the response that first comes to mind, maybe they are not fully aware of the risk that investing poses on their financial position. Those with one or two answers correct are the most risk averse: they recognize the risks they have to bear for profit, but do not reckon what those really are and how much damage can happen. This result is in line with descriptions by Dohmen et al. (2010), Bortoli (2019) and Rzeszutek et al. (2015), who pointed out that the relationship between these two variables – cognitive capacity and risk aversion – is non-linear.

Frederick (2005) noted: individuals with low cognitive skills tend to be risk-averse (and more risk-averse compared to people with high cognitive ability) when facing high probability of gain or low probability of loss, however risk-seeking (although less risk-seeking compared to people with high cognitive ability) when facing low probability of gain or high probability of loss. The results of the present paper might partially justify these relationships.

4.4. Prediction of risk aversion

In light of this critical assessment, the model selection must align with the specific objectives of the predictive task. Rather than interpretability and simplicity, the goal of classifying risk preferences demands accuracy and a deeper understanding of feature importance, which is why Random Forest stands out as an exceptional choice. Its ensemble nature reduces overfitting, and its ability to handle complex relationships within the data is unmatched.

Therefore, for the predictive task at hand, where accuracy and robustness are pivotal, the Random Forest model is recommended. Its balance between precision and recall, along with its superior accuracy and generalizability, positions it as the ideal choice. Nonetheless, it's essential to remember that no model is infallible, and continuous refinement and optimization should be undertaken to maximize its predictive potential.

The prediction code block employs a user-centric approach to gather inputs on specific personality traits and Cognitive Reflection Test answers, which are then utilized to predict the individual's risk preference using a machine learning model.

Figure 4.5: Example of risk preference prediction

```
On a scale from 1 to 6 (1 being strongly disagree and 6 being strongly agree),
rate the following statements:
Extraverted, enthusiastic: 1
Critical, quarrelsome: 2
Dependable, self-disciplined: 3
Anxious, easily upset: 4
Open to new experiences, complex: 5
Reserved, quiet: 6
Sympathetic, warm: 5
Disorganised, careless: 4
Calm, emotionally stable: 3
Conventional, uncreative: 2

Answer the following questions:
A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. H
ow much does the ball cost? £ 1
If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100
machines to make 100 widgets? Minutes: 5
In a lake, there is a patch of lily pads. Every day, the patch doubles in size.
If it takes 48 days for the patch to cover the entire lake, how long would it t
ake for the patch to cover half of the lake? Days: 47

The predicted risk preference is: Seeking
```

Source: Generated by Python, as attached in Appendix B

Chapter 5

Recommendations & Conclusion

5.1. Recommendations

In this research, personality traits, cognitive reflection, and risk behavior were all shown to have a role in investing decision. The study's findings have ramifications for individual investors, financial advisors, and financial organizations.

5.1.1. Individual investors' perspectives

Warren Buffet aptly mentioned, "Success in investing does not correlate with IQ. Once you have ordinary intelligence, what you need is the temperament to control the urges that get other people into trouble in investing." At the crux of this sentiment lies the understanding that individual investors, while influenced by global market dynamics, predominantly navigate their financial decisions influenced by their inherent personality traits. Our study, harnessing advanced machine learning methodologies, illuminated how such personal attributes, combined with cognitive reflection scores, play an influential role in risk aversion and decision-making within the financial arena.

The findings presented provide a roadmap for individual investors, guiding them to introspectively analyze their inherent personality characteristics and consequently refine their investment strategies. For instance, an investor who leans heavily towards neuroticism might be more inclined to anchor themselves in the relative security of government bonds or stable bank deposits as opposed to volatile stock market ventures. Contrarily, those who exhibit high levels of openness might diversify their portfolios, exploring unconventional investment avenues. Additionally, the significance of the Cognitive Reflection Test underscores the value of critical thinking in financial decisions. Investors with lower scores might benefit from seeking external financial advisory to ensure well-informed and strategic investment choices.

A new investor, or an investor not feeling satisfied about the current investing strategies, can assess his or her own personal characteristics following this research's results, and act accordingly. For example, an investor with high score in neuroticism might want to invest in low-risk assets such as government bonds or bank deposits rather than the fluctuant stock market. Meanwhile, an identified-open-to-experience investor may seek more investing alternatives to satisfy his or her risk appetite. Someone scoring low on the CRT might want to re-assess the level of risk he or she is bearing, maybe via outside advisory for a more comprehensive and cognitive viewpoint; while those getting one or two questions wrong should learn more about the investment plan or assets.

5.1.2. Financial Advisors and Institutional Insights

Financial institutions and advisors stand to gain immensely from this research's revelations. Recognizing the intricate interplay between an individual's psychological makeup and their financial decisions can empower financial advisors to offer personalized guidance. Given the nature of financial information - deeply personal and often closely guarded - the insights from this study offer an alternative avenue to understand client needs without prying into their financial secrets. The personal attributes, more observable and often less intrusive, can provide advisors a lens to gauge potential risk attitudes and investment intentions.

It is imperative for financial planners to recognize and understand the complexity and influence of these behavioral attributes. This understanding can guide the creation of tailored financial products, ensuring alignment with the client's personality, risk tolerance, and financial objectives. Furthermore, in a landscape where client-advisor relationships are paramount, behavioral finance paves the way for creating robust, trust-driven bonds. An advisor, equipped with behavioral finance insights, can offer not just financial advice but also emotional reassurance, particularly crucial in turbulent market conditions.

Moreover, financial organizations can strategically leverage these findings for targeted marketing campaigns. Recognizing the inherent behavioral traits of potential clients can lead to more customized promotional events, drawing in the right clientele. For instance, hosting events that cater to the extroverted, or community-focused events to resonate with the agreeable. Additionally, in the realm of banking, understanding dominant personality features can aid in discerning customer reliability, potentially refining the loan or credit card approval processes. The study thus heralds an era where the synergy between behavioral psychology and financial decision-making could redefine the paradigms of personalized financial services.

5.2. Conclusion

Psychology factors always exist in individuals naturally and can affect investors' behaviors. Studies on psychological factors help financial professionals as well as investors control these complex factors, which is contributed by the conduct of this research. Behavioral finance is a relatively new theory but is growing rapidly and helpful for traditional finance. By connecting between psychological theory and practice of finance, behavioral finance has helped to explain complicated and sophisticated issues which cannot be explained clearly by traditional economic or finance theories. The present study's results have proved the common existence of behavioral factors, namely personality traits and cognitive reflection, and their influences on risk aversion. In response to the objectives outlined in the introductory chapter of this study, the following findings were ascertained:

i. The examination of the Big Five personality traits and Cognitive Reflection Test scores revealed profound insights. Neuroticism emerged as the most influential predictor of risk aversion, followed closely by extraversion and openness to experience. Additionally, the relationship between cognitive reflection scores and risk aversion was found to be non-linear, highlighting its nuanced role in influencing investment behavior.

ii. Among the machine learning models assessed, the Random Forest classifier distinctly stood out in terms of accuracy and robustness, showcasing its superiority in capturing the intricacies of the dataset and predicting financial risk preferences.

iii. For financial institutions, the findings underscore the significance of offering tailored services. By understanding and integrating the behavioral nuances of clients, institutions can deliver more personalized and efficient advisory experiences, driving both customer satisfaction and optimal investment outcomes.

The analysis presented in this paper follows a new path to research agenda of the future – machine learning in behavioral finance. Such research topic helps financial investing practitioners gain a deeper understanding of financial markets through a deeper understanding of the individuals whose behavior creates and sustains them.

5.3. Limitations of the study

The research, while groundbreaking in its integrative approach of psychological aspects and financial decision-making, encountered certain methodological limitations. Chief among these was the incorporation of data augmentation techniques, specifically resampling and the use of SMOTE, aimed at enhancing dataset robustness. While these techniques were indispensable in bolstering the data's richness, they come with a caveat: the potential introduction of artificial patterns. Such manufactured patterns might not authentically mirror real-world scenarios, potentially skewing the findings. Another notable constraint was the inherent assumption of the static nature of behavioral traits. Relying on the premise that an individual's personality and cognitive reflection remain unchanged can be somewhat reductionist, given the dynamic nature of human psychology shaped by life experiences and ever-evolving financial landscapes.

Furthermore, our dependency on self-reported data, particularly through questionnaires, could inadvertently inject biases. The subjective nature of such data collection means often grappling with inaccuracies arising from various factors, from social desirability to plain misunderstanding of posed questions. While our machine learning approach, epitomized by the use of the Random Forest classifier, offered promising insights, it's paramount to acknowledge that algorithms, however advanced, might still not fully encapsulate the intricate nuances of human behavior.

Recognizing these limitations not only underscores the rigor of this research but also underscores the avenues available for future exploration. Future scholars in this domain might delve deeper by diversifying participant profiles, embracing longitudinal designs, or integrating a broader set of influencing factors. Such endeavors can refine our understanding and, in doing so, bridge the gap between the realms of human psychology and financial decision-making even further.

Thaler (1981, p.198) states, “rational models tend to be simple and elegant with precise predictions, while behavioral models tend to be complicated, and messy, with much vaguer predictions”. This study did not have the goal of developing a formal behavioral saving model, but only to accumulate empirical evidence that may be used to construct relationships among factors.

References

Ahmed, S., Alshater, M.M., Ammari, A.E. and Hammami, H., 2022. Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance*, 61, p.101646. https://doi.org/10.1016/j.ribaf.2022.101646.

Akhtar, F. and Das, N., 2020. Investor personality and investment performance: from the perspective of psychological traits. *Qualitative Research in Financial Markets*, 12(3), pp.333–352. https://doi.org/10.1108/QRFM-11-2018-0116.

Anon. 2012. In uncharted territory, many turn to the chartists. *Reuters*. [online] 17 Jan. Available at: <https://www.reuters.com/article/markets-bonds-technicals-idINDEE80G0EX20120117> [Accessed 4 September 2023].

Anon. 2021. *Survey: Less Than Half of Women in U.S. Invest in the Stock Market*. [online] NerdWallet. Available at: <https://www.nerdwallet.com/article/investing/survey-less-than-half-of-women-in-u-s-invest-in-the-stock-market> [Accessed 4 September 2023].

Bailey, W., Kumar, A. and Ng, D., 2011. Behavioral biases of mutual fund investors. *Journal of Financial Economics*, 102(1), pp.1–27. https://doi.org/10.1016/j.jfineco.2011.05.002.

Bajaj, V., 2008. Forget Logic; Fear Appears to Have Edge. *The New York Times*. [online] 8 Oct. Available at: <https://www.nytimes.com/2008/10/08/business/08fear.html> [Accessed 4 September 2023].

Banerji, G. and Wallerstein, E., 2023. The Stock Market Isn't as Calm as It Seems. *Wall Street Journal*. [online] 27 Jun. Available at: <https://www.wsj.com/articles/the-stock-market-isnt-as-calm-as-it-seems-3713180c> [Accessed 4 September 2023].

Barber, B.M. and Odean, T., 2001. Boys will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics*, 116(1), pp.261–292.

Benjamin, D.J., Brown, S.A. and Shapiro, J.M., 2013. Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences. *Journal of the European Economic Association*, 11(6), pp.1231–1255. https://doi.org/10.1111/jeea.12055.

Ben-Shahar, D. and Golan, R., 2014. Real estate and personality. *Journal of Behavioral and Experimental Economics*, 53, pp.111–119. https://doi.org/10.1016/j.socec.2014.08.008.

Bernoulli, D., 2011. Exposition of a new theory on the measurement of risk. In: *The Kelly Capital Growth Investment Criterion*, World Scientific Handbook in Financial Economics Series. [online] WORLD SCIENTIFIC. pp.11–24. https://doi.org/10.1142/9789814293501_0002.

Bordalo, P., Gennaioli, N. and Shleifer, A., 2018. Diagnostic Expectations and Credit Cycles. *The Journal of Finance*, 73(1), pp.199–227. https://doi.org/10.1111/jofi.12586.

Borghans, L., Duckworth, A.L., Heckman, J.J. and Weel, B. ter, 2008. The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43(4), pp.972–1059. https://doi.org/10.3368/jhr.43.4.972.

Bortoli, D.D., Jr, N. da C., Goulart, M. and Campara, J., 2019. Personality traits and investor profile analysis: A behavioral finance study. *PLOS ONE*, 14(3), p.e0214062. https://doi.org/10.1371/journal.pone.0214062.

Brighetti, G., Ottaviani, C., Nucifora, V. and Borlimi, R., 2011. Decision Making: Psychological Perspective. In: C. Lucarelli and G. Brighetti, eds. *Risk Tolerance in Financial Decision Making*, Palgrave Macmillan Studies in Banking and Financial Institutions. [online] London: Palgrave Macmillan UK. pp.133–152. https://doi.org/10.1057/9780230303829_6.

Caspi, A. and Roberts, B.W., 2001. Personality Development Across the Life Course:The Argument for Change and Continuity. *Psychological Inquiry*, 12(2), pp.49–66. https://doi.org/10.1207/S15327965PLI1202_01.

Cawley, J., Heckman, J. and Vytlacil, E., 2001. Three observations on wages and measured cognitive ability. *Labour Economics*, 8(4), pp.419–442. https://doi.org/10.1016/S0927-5371(01)00039-2.

Chang, H.Y., Chen, K.C. and Fang, H.K., 2016. The relationship study of mutual fund investor personality traits and investment behaviour. *Journal of Chinese Economic and Business Studies*, 14, pp.51–73.

Charness, G., Gneezy, U. and Imas, A., 2013. Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, 87, pp.43–51. https://doi.org/10.1016/j.jebo.2012.12.023.

Chen, T.-H., Ho, R.-J. and Liu, Y.-W., 2019. Investor personality predicts investment performance? A statistics and machine learning model investigation. *Computers in Human Behavior*, 101, pp.409–416. https://doi.org/10.1016/j.chb.2018.09.027.

Choi, J.J., Laibson, D., Madrian, B.C. and Metrick, A., 2009. Reinforcement Learning and Savings Behavior. *The Journal of Finance*, 64(6), pp.2515–2534.

Christelis, D., Jappelli, T. and Padula, M., 2010. Cognitive abilities and portfolio choice. *European Economic Review*, 54(1), pp.18–38. https://doi.org/10.1016/j.euroecorev.2009.04.001.

Conlin, M., Lynn, M. and O'Donoghue, T., 2003. The norm of restaurant tipping. *Journal of Economic Behavior & Organization*, 52(3), pp.297–321. https://doi.org/10.1016/S0167-2681(03)00030-1.

Corgnet, B., Espín, A. and Hernán-González, R., 2015. The Cognitive Basis of Social Behavior: Cognitive Reflection Overrides Antisocial but Not Always Prosocial Motives. *Economics Faculty Articles and Research*. [online] https://doi.org/10.3389/fnbeh.2015.00287.

Costa, P.T. and McCrae, R.R., 2008. The Revised NEO Personality Inventory (NEO-PI-R). In: *The SAGE handbook of personality theory and assessment, Vol 2: Personality measurement and testing*. Thousand Oaks, CA, US: Sage Publications, Inc. pp.179–198. https://doi.org/10.4135/9781849200479.n9.

Cueva, C., Iturbe-Ormaetxe, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., Yu, H. and Zhukova, V., 2016. Cognitive (ir)reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics*, 64, pp.81–93. https://doi.org/10.1016/j.socec.2015.09.002.

Daniel, K. and Hirshleifer, D., 2015. Overconfident Investors, Predictable Returns, and Excessive Trading. *Journal of Economic Perspectives*, 29(4), pp.61–88. https://doi.org/10.1257/jep.29.4.61.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A., 1998. Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, 53(6), pp.1839–1885. https://doi.org/10.1111/0022-1082.00077.

De Fruyt, F., De Bolle, M., McCrae, R.R., Terracciano, A. and Costa, P.T., 2009. Assessing the Universal Structure of Personality in Early Adolescence: The NEO-PI-R and NEO-PI-3 in 24 Cultures. *Assessment*, 16(3), pp.301–311. https://doi.org/10.1177/1073191109333760.

Digman, J.M., 1989. Five Robust Trait Dimensions: Development, Stability, and Utility. *Journal of Personality*, 57(2), pp.195–214. https://doi.org/10.1111/j.1467-6494.1989.tb00480.x.

Dohmen, T., Falk, A., Huffman, D. and Sunde, U., 2010. Are Risk Aversion and Impatience Related to Cognitive Ability? *American Economic Review*, 100(3), pp.1238–1260. https://doi.org/10.1257/aer.100.3.1238.

Dohmen, T., Falk, A., Huffman, D. and Sunde, U., 2018. On the Relationship between Cognitive Ability and Risk Preference. *Journal of Economic Perspectives*, 32(2), pp.115–134. https://doi.org/10.1257/jep.32.2.115.

Drichoutis, A.C. and Nayga, R.M., 2022. On the stability of risk and time preferences amid the COVID-19 pandemic. *Experimental Economics*, 25(3), pp.759–794. https://doi.org/10.1007/s10683-021-09727-6.

Durand, R.B., Newby, R. and Sanghani, J., 2008. An Intimate Portrait of the Individual Investor. *Journal of Behavioral Finance*, 9(4), pp.193–208. https://doi.org/10.1080/15427560802341020.

Fama, E.F., 1960. *Efficient Markets Hypothesis*. Diss. PhD Thesis, Ph. D. dissertation. University of Chicago.

Fehr-Duda, H. and Epper, T., 2012. Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences. *Annual Review of Economics*, 4(1), pp.567–593. https://doi.org/10.1146/annurev-economics-080511-110950.

Frederick, S., 2005. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), pp.25–42. https://doi.org/10.1257/089533005775196732.

Fromlet, H., 2001. Behavioral Finance - Theory and Practical Application: Systematic Analysis of Departures from The Homo Oeconomicus Paradigm Are Essential for Realistic Financial Research and Analysis. *Business Economics*, 36(3), pp.63–69.

Gambetti, E. and Giusberti, F., 2019. Personality, decision-making styles and investments. *Journal of Behavioral and Experimental Economics*, 80, pp.14–24. https://doi.org/10.1016/j.socec.2019.03.002.

Goldberg, L.R., 1993. The structure of phenotypic personality traits. *American Psychologist*, 48(1), pp.26–34. https://doi.org/10.1037/0003-066X.48.1.26.

Górnik-Durose, M.E., 2020. Materialism and Well-Being Revisited: The Impact of Personality. *Journal of Happiness Studies*, 21(1), pp.305–326. https://doi.org/10.1007/s10902-019-00089-8.

Gu, S., Kelly, B. and Xiu, D., 2020. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), pp.2223–2273. https://doi.org/10.1093/rfs/hhaa009.

Guiso, L., Sapienza, P. and Zingales, L., 2018. Time varying risk aversion. *Journal of Financial Economics*, 128(3), pp.403–421. https://doi.org/10.1016/j.jfineco.2018.02.007.

Hastings, J.S., Madrian, B.C. and Skimmyhorn, W.L., 2013. Financial Literacy, Financial Education, and Economic Outcomes. *Annual Review of Economics*, 5(1), pp.347–373. https://doi.org/10.1146/annurev-economics-082312-125807.

Häusler, A.N., Kuhnen, C.M., Rudorf, S. and Weber, B., 2018. Preferences and beliefs about financial risk taking mediate the association between anterior insula activation and self-reported real-life stock trading. *Scientific Reports*, 8(1), p.11207. https://doi.org/10.1038/s41598-018-29670-6.

Herrera, L., Al-Lal, M. and Mohamed, L., 2020. Academic Achievement, Self-Concept, Personality and Emotional Intelligence in Primary Education. Analysis by Gender and Cultural Group. *Frontiers in Psychology*, [online] 10. Available at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.03075> [Accessed 22 July 2023].

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U. and Zeisberger, S., 2020. What Drives Risk Perception? A Global Survey with Financial Professionals and

Laypeople. *Management Science*, 66(9), pp.3799–4358, iii–iv. https://doi.org/10.1287/mnsc.2019.3526.

Hunter, K. and Kemp, S., 2004. The personality of e-commerce investors. *Journal of Economic Psychology*, 25(4), pp.529–537. https://doi.org/10.1016/S0167-4870(03)00050-3.

Jadlow, J.W. and Mowen, J.C., 2010. Comparing the Traits of Stock Market Investors and Gamblers. *Journal of Behavioral Finance*, 11(2), pp.67–81. https://doi.org/10.1080/15427560.2010.481978.

Kahneman, D., 2011. *Thinking, fast and slow*. macmillan.

Kahneman, D. and Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. In: *Econometrica*. [online] The Econometric Society. pp.263–292. Available at: <https://doi.org/10.2307/1914185> [Accessed 29 May 2021].

Kuhnen, C.M. and Knutson, B., 2005. The Neural Basis of Financial Risk Taking. *Neuron*, 47(5), pp.763–770. https://doi.org/10.1016/j.neuron.2005.08.008.

Lerner, J.S. and Keltner, D., 2000. Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion*, 14(4), pp.473–493. https://doi.org/10.1080/026999300402763.

Linden, S. van der, 2015. The social-psychological determinants of climate change risk perceptions: Towards a comprehensive model. *Journal of Environmental Psychology*, 41, pp.112–124. https://doi.org/10.1016/j.jenvp.2014.11.012.

Lubinski, D. and Humphreys, L.G., 1997. Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, 24(1), pp.159–201. https://doi.org/10.1016/S0160-2896(97)90016-7.

Markowitz, H.M., 1991. Foundations of Portfolio Theory. *The Journal of Finance*, 46(2), pp.469–477. https://doi.org/10.2307/2328831.

Mayfield, C., Perdue, G. and Wooten, K., 2008. Investment management and personality type. *Financial Services Review*, 17.

McCrae, R.R. and Costa, P.T., 1987. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), pp.81–90. https://doi.org/10.1037/0022-3514.52.1.81.

Mishra, S., Lalumière, M.L. and Williams, R.J., 2010. Gambling as a form of risk-taking: Individual differences in personality, risk-accepting attitudes, and behavioral preferences for risk. *Personality and Individual Differences*, 49(6), pp.616–621. https://doi.org/10.1016/j.paid.2010.05.032.

Nga, J.K.H. and Ken Yien, L., 2013. The influence of personality trait and demographics on financial decision making among Generation Y. *Young Consumers*, 14(3), pp.230–243. https://doi.org/10.1108/YC-11-2012-00325.

Oehler, A., Wendt, S., Wedlich, F. and Horn, M., 2018. Investors' Personality Influences Investment Decisions: Experimental Evidence on Extraversion and Neuroticism. *Journal of Behavioral Finance*, 19(1), pp.30–48. https://doi.org/10.1080/15427560.2017.1366495.

Olsen, R.A., 2007. Investors' Predisposition for Annuities: A Psychological Perspective. *Journal of Financial Service Professionals*, 61(5), pp.51–57.

Olsen, R.A., 2008. Cognitive Dissonance: The Problem Facing Behavioral Finance. *Journal of Behavioral Finance*, 9(1), pp.1–4. https://doi.org/10.1080/15427560801896552.

Park, N.Y., 2016. Domain-specific risk preference and cognitive ability. *Economics Letters*, 141, pp.1–4. https://doi.org/10.1016/j.econlet.2016.01.008.

Phares, E.J., 1988. *Introduction to personality, 2nd ed.* Introduction to personality. Glenview, IL, US: Scott, Foresman & Co. pp.xxiv, 656.

Pompian, M.M., 2016. *Risk Profiling through a Behavioral Finance Lens*. [online] CFA Institute. Available at: </en/research/foundation/2016/risk-profiling-through-a-behavioral-finance-lens> [Accessed 28 April 2021].

Rammstedt, B., Danner, D., Soto, C.J. and John, O.P., 2020. Validation of the Short and Extra-Short Forms of the Big Five Inventory-2 (BFI-2) and Their German Adaptations. *European Journal of Psychological Assessment*, 36(1), pp.149–161. https://doi.org/10.1027/1015-5759/a000481.

Ricciardi, V. and Simon, H.K., 2001. *What is Behavioral Finance?* [SSRN Scholarly Paper] Rochester, NY: Social Science Research Network. Available at: <https://papers.ssrn.com/abstract=256754> [Accessed 17 April 2023].

Roberts, B.W. and Robins, R.W., 2000. Broad Dispositions, Broad Aspirations: The Intersection of Personality Traits and Major Life Goals. *Personality and Social Psychology Bulletin*, 26(10), pp.1284–1296. https://doi.org/10.1177/0146167200262009.

Rustichini, A., DeYoung, C.G., Anderson, J. and Burks, S.V., 2012. *Toward the integration of personality theory and decision theory in the explanation of economic and health behavior*. [Working Paper] IZA Discussion Papers. Available at: <https://www.econstor.eu/handle/10419/62547> [Accessed 5 June 2021].

Rzeszutek, M., Partyka, M. and Gołąb, A., 2015. Temperament traits, social support, and secondary traumatic stress disorder symptoms in a sample of trauma therapists. *Professional Psychology: Research and Practice*, 46(4), pp.213–220. https://doi.org/10.1037/pro0000024.

Sadhwani, A., Giesecke, K. and Sirignano, J., 2021. Deep Learning for Mortgage Risk*. *Journal of Financial Econometrics*, 19(2), pp.313–368. https://doi.org/10.1093/jjfinec/nbaa025.

Sadi, R., Asl, H.G., Rostami, M.R., Gholipour, A. and Gholipour, F., 2011. Behavioral Finance: The Explanation of Investors' Personality and Perceptual Biases Effects on Financial Decisions. *International Journal of Economics and Finance*, 3(5), p.p234. https://doi.org/10.5539/ijef.v3n5p234.

Semeijn, J. h., van der Heijden, B. i. j. m. and De Beuckelaer, A., 2020. Personality Traits and Types in Relation to Career Success: An Empirical Comparison Using the Big Five. *Applied Psychology*, 69(2), pp.538–556. https://doi.org/10.1111/apps.12174.

Sorkin, I., 2018. Ranking Firms Using Revealed Preference*. *The Quarterly Journal of Economics*, 133(3), pp.1331–1393. https://doi.org/10.1093/qje/qjy001.

Soto, C.J. and John, O.P., 2017. The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113(1), pp.117–143. https://doi.org/10.1037/pspp0000096.

Spaniol, J. and Bayen, U.J., 2005. Performance in Goal-Directed Continuous-Choice Tasks: New Approaches for Studying Individual and Age-Related Differences. *Developmental Psychology*, 41(5), pp.731–745.

Tett, R.P., Simonet, D.V., Walser, B. and Brown, C., 2013. Trait Activation Theory: Applications, Developments, and Implications for Person-Workplace Fit. In: *Handbook of Personality at Work*. Routledge.

Thaler, R.H. and Shefrin, H.M., 1981. An Economic Theory of Self-Control. *Journal of Political Economy*, 89(2), pp.392–406. https://doi.org/10.1086/260971.

Thoma, V., White, E., Panigrahi, A., Strowger, V. and Anderson, I., 2015. Good thinking or gut feeling? Cognitive reflection and intuition in traders, bankers and financial non-experts. *PLoS ONE*, 10(4). https://doi.org/10.1371/journal.pone.0123202.

Tisu, L., Lupșa, D., Vîrgă, D. and Rusu, A., 2020. Personality characteristics, job performance and mental health: the mediating role of work engagement. *Personality and Individual Differences*, 153, p.109644. https://doi.org/10.1016/j.paid.2019.109644.

Ton, H.T.H. and Dao, T.K., 2014. The Effects of Psychology on Individual Investors' Behaviors: Evidence from the Vietnam Stock Exchange. *Journal of Management and Sustainability*, 4, p.125.

Tversky, A. and Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), pp.297–323. https://doi.org/10.1007/BF00122574.

Vukasović, T. and Bratko, D., 2015. Heritability of personality: A meta-analysis of behavior genetic studies. *Psychological Bulletin*, 141(4), pp.769–785. https://doi.org/10.1037/bul0000017.

Appendix A

Questionnaire

The following list consists of personality attributes that may or may not apply to you. Please indicate your level of agreement or disagreement with each statement by marking a number next to it. The higher the level of agreement, the higher the score you give, where:

1	*Totally disagree*	*2*	*Mostly disagree*	*3*	*Partially disagree*
4	*Partially agree*	*5*	*Mostly agree*	*6*	*Totally agree*

You should rate the extent to which the pair of characteristics applies to you, even if one trait applies more significantly.

Code	Item	Agreement level						Source
Big Five Inventory: I see myself as...								Rammstedt, B. and
EXT1	Extraverted, enthusiastic.	1	2	3	4	5	6	John, O.P., 2007.
AGR1	Critical, quarrelsome.	1	2	3	4	5	6	Measuring
CST1	Dependable, self-disciplined.	1	2	3	4	5	6	personality in one
NRT1	Anxious, easily upset.	1	2	3	4	5	6	minute or less: A 10-
OPE1	Open to new experiences, complex.	1	2	3	4	5	6	item short version of the Big Five
EXT2(R)	Reserved, quiet.	1	2	3	4	5	6	Inventory in English
AGR2(R)	Sympathetic, warm.	1	2	3	4	5	6	and
CST2(R)	Disorganised, careless.	1	2	3	4	5	6	German. *Journal of*
NRT2(R)	Calm, emotionally stable.	1	2	3	4	5	6	*research in*
OPE2(R)	Conventional, uncreative.	1	2	3	4	5	6	*Personality*, *41*(1),
"R" denotes reverse-scored items.								pp.203-212.

Cognitive Reflection Test: Question 1: A bat and a ball cost £1.10 in total. The bat costs £1.00 more than the ball. How much does the ball cost? Q1A: £_______. Question 2: If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? Q2A: _______ minutes. Question 3: In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? Q3A: _______ days.		Frederick, S., 2005. Cognitive reflection and decision making. *Journal of Economic perspectives*, *19*(4), pp.25-42.

Risk preference regarding financial investments:

RPR1	I am not willing to take risk when choosing a stock or investment.	1	2	3	4	5	6	Mayfield, C., Perdue, G. and Wooten, K., 2008. Investment management and personality type. *Financial services review*, *17*(3), pp.219-236.
RPR2	I prefer a low risk/high return investment with a steady performance over an investment that offers higher risk/higher return.	1	2	3	4	5	6	
RPR3	I prefer to remain with an investment strategy that has known problems than take the risk trying a new investment strategy that has unknown problems, even if the new investment strategy has great returns.	1	2	3	4	5	6	
RPR4	I view risk in investment as a situation to be avoided at all cost.	1	2	3	4	5	6	

Survey respondents' demographic information (optional):

Name:	Gender: 1. Male; 2. Female; 3. Other	
Level of financial investment knowledge:	1. I have never heard about financial investment. 2. I am aware of financial investment but have no intention to participate. 3. I am aware of financial investment and have intention to invest in the future. 4. I have participated in financial investment (before).	
Age group:	18-25 26-35	36-50 50+
Annual income:	Under £20,000 £20,000 – £40,000 £40,000 – £60,000 £60,000 – £80,000 Above £80,000	

Appendix B

Code Block

Please click here to be directed to GitHub repository with a working code.