Financial risk tolerance profiling from text

Information Processing and Management 61 (2024) 103704

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

Financial risk tolerance profiling from text

Frank Xing

School of Computing, National University of Singapore, Singapore, Singapore

ARTICLE INFO

Keywords:
Artificial intelligence in finance
Risk tolerance
Risk profiling
Text mining
Convolutional neural network

ABSTRACT

Traditionally, individual financial risk tolerance information is gathered via questionnaires or similar structured psychometric tools. Our abundant digital footprint, as an unstructured alternative, is less investigated. Leveraging such information can potentially support large-scale and cost-efficient financial services. Therefore, I explore the possibility of building a computational model that distills risk tolerance information from user texts in this study, and discuss the design principles discovered from empirical results and their implications. Specifically, a new quaternary classification task is defined for text mining-based risk profiling. Experiments show that pre-trained large language models set a baseline micro-F1 of circa 0.34. Using a convolutional neural network (CNN), the reported system achieves a micro-F1 of circa 0.51, which significantly outperforms the baselines, and is a circa 4% further improvement over the standard CNN configurations (micro-F1 of circa 0.47). Textual feature richness and supervised learning are found to be the key contributors to model performances, while other machine learning strategies suggested by previous research (data augmentation and multi-tasking) are less effective. The findings confirm user texts to be a useful risk profiling resource and provide several insights on this task.

1. Introduction

Risk has been a central topic in finance from the very beginning (Markowitz, 1952; Sharpe, 1964) and is still a critical concept in many financial decision-making and modeling processes today. For example, the capital asset pricing model (CAPM) calculates market risk premium at an aggregated level and uses it to explain different expected returns from different financial assets; the asset allocation models use the risk aversion at an individual level to decide the optimal portfolio holding weights (Thavaneswaran et al., 2021; Xing et al., 2019b); banks use companies' auditing and fraudulent risk information and platforms use individuals' credit risk inferred from their self-disclosures to make lending decisions (Saha et al., 2016; Siering, 2023). The digitalization trend of the financial market, the increasing diversity of financial products, and the rising influence of retail investors together add more uncertainty to investment. As a result, investors may suffer the risk of loss of income or even loss of principal when investing. In such a context, investors need to choose investment projects based on their investment goals and risk preferences, and many investors will consult with financial advisors before investing, despite the accompanying cost.

There is a pressing need to leverage the information available and transform financial planning into a more economic and inclusive process. Although risk tolerance is an important factor in financial planning and consulting, a formal definition of it is challenging (Hemrajani et al., 2023). Only in recent years have practitioners clearly realized the differences between many risk-related concepts, including *risk appetite/need, risk perception, risk preference, risk attitude, risk tolerance, risk aversion, risk capacity, risk-taking behavior, risk profile, and more*. Grable (2018) defines risk tolerance as the willingness to engage in risky behavior in which possible outcomes can be negative. Therefore, investors with high risk tolerance are more likely to engage in more

E-mail address: xing@nus.edu.sg.

https://doi.org/10.1016/j.ipm.2024.103704
Received 20 November 2023; Received in revised form 29 February 2024; Accepted 2 March 2024
Available online 5 March 2024

high-risk investments, while investors with low risk tolerance tend to be more conservative. Understanding such investors' risk tolerance information helps financial institutions gauge customers' comfort level with investment risk and provide customers with personalized information. In order to help customers make better financial decisions, financial institutions need to provide customers with appropriate guidance. Before investing, many institutions require customers to answer questions in a survey, then complete investment portfolios and provide customers with services and suggestions based on the survey results. This remains a standard procedure for digital services, e.g., robo-advisory. Investors are classified into different categories according to the amount of loss they can tolerate. A common practice by financial institutions is to divide users' risk tolerance into several categories, e.g., radical, moderate, and conservative. DBS Bank, e.g., currently customizes each of its themed portfolios into Slow n' Steady (Risk Level 2), Comfy Cruisin' (Risk Level 3), and Fast n' Furious (Risk Level 4) for the "digiPortfolio" investment product.[1]

Numerous previous studies argue that risk tolerance is closely related to other personal traits at an individual level, and an investor's behavior and goal can be understood through risk tolerance together with cognitive and emotional biases, as well as investors' sentiment (Lengkeek et al., 2023; Xing et al., 2020; Yekrangi & Abdolvand, 2021). For instance, Pompian and Longo (2004) suggested that investment advisors consider client gender and personality to assess risk tolerance before executing an investment program according to the following four-step method: (1) Ask your client to take a personality type test; (2) Evaluate responses to determine personality type; (3) Assess risk tolerance using the "Type and Gender-Based Risk tolerance Scales"; (4) Execute investment program. Nobre and Grable (2015) advised to better understand clients' risk-taking behavior via evaluating their risk tolerance, which was influenced by their risk profile, risk perception, and risk need.

The risk tolerance information is also important at a macro level. With knowledge of investors' risk attitudes and psychology, behavioral finance reveals and explains some irrational behaviors of investors in the financial market. Investors and financial planners may have cognitive and emotional biases when making important investment decisions (Athota et al., 2023; Yekrangi & Abdolvand, 2015). An example of emotional bias is that investors' overconfidence may make them more inclined to receive news that enhances their self-confidence but ignore information that differs from their opinions. When suffering a loss, the feeling of pain caused by the loss may make investors continue to hold these assets because they want to avoid the feeling of pain, which may lead to continued loss of assets. The more confident people are, the more frequently they will trade, and the more likely they will receive low returns. People with low risk tolerance may experience opportunity losses from not investing in stocks, while people with high risk tolerance in short-term investing may cause unnecessary losses in wealth (Yao & Hanna, 2005). The survey by Ainia and Lutfi (2019) shows that risk tolerance had a significant and positive effect on investment decision-making: the higher a person's risk tolerance level, the higher the person's opportunity to allocate funds to high-risk assets. An understanding of risk tolerance was one necessary factor for a person to be able to make optimal portfolio choices in terms of risk-reward trade-offs, and choosing a portfolio not consistent with risk tolerance may cause investor disappointment and inferior utility (Moreschi, 2005).

With the accumulating digital footprints on social media and advances in natural language processing (NLP) comes the opportunity to know your customer (including risk tolerance, behavioral biases, personality, and many other associated aspects) through analyzing the online user generated content (UGC). In fact, the literature on personality detection or risk profiling for corporate entities from text is abundant (Vinciarelli & Mohammadi, 2014; Yin et al., 2020). It is also reported that text-derived personality traits effectively depict and predict consumer perceptive behaviors in financial and health contexts (Yang et al., 2023). However, there was scant previous research that attempted to profile users' financial risk tolerance directly from the UGC to the best of my knowledge. The most relevant studies in this thread are those that measure patients' personality and subjective risk tolerance through questionnaire surveys and used regression methods to establish the relationship between personality traits and risk tolerance. These studies are key to the major challenge in this research task: the lack of risk tolerance labels for existing text corpora. In this research, I summarize the results of these studies and calculate user risk tolerance labels via personality traits. This way, a convolutional neural network (CNN) model that directly infers the financial risk tolerance of users from UGC has been trained. Since this study aims to test the effectiveness of UGC features on the new task rather than optimizing system performances, the CNN architecture is chosen over transformer-based models for its simpler architecture, better interoperability, and a rich past literature to compare with when it is used as a test bed. The presented method can help financial service providers better understand customers' risk preferences in a fast and cost-efficient manner, thus promoting financial inclusion. From the client's perspective, providing this model helps them choose appropriate financial products according to their personal investment preferences, thus reducing possible losses in investment. As a result, clients are more satisfied with the service and will be more willing to continue investing with the institution.

This study attempts to address two main *research objectives*. The *first* objective is to test whether financial risk profiling can directly benefit from user generated texts. Previous research documented that (1) financial risk tolerance is associated with personality traits, and (2) personality traits can be modeled from user texts. However, it is unclear to what extent the useful information can be preserved. The *second* objective is to develop modeling guidelines via experimenting with effective techniques on personality detection, including recurrent CNN (Nasir & Malik, 2024), data augmentation (Yang et al., 2023), and multi-tasking (Li et al., 2022).

To preview the main result, it has been discovered that individuals' digital footprint is an effective source of information for financial risk tolerance profiling. Rich text representation features (pre-trained word embeddings from various language models) benefit the model performance more than machine learning tricks, e.g., sentence augmentation and multitasking. Specifically, this study makes three major *contributions*:

[1] https://www.dbs.com.sg/personal/investments/other-investments/dbs-digiportfolio

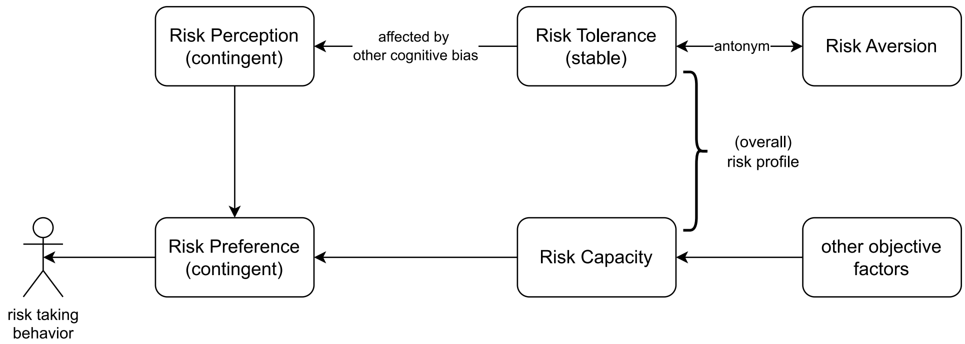

Fig. 1. Risk-related terminologies and their relations.

1. It formally proposes the financial risk tolerance profiling task as a quaternary classification problem and summarizes a proxy risk labeling method via personality from previous studies;
2. A first-of-its-kind dataset for the above-mentioned task is synthesized and made available for research purposes upon reasonable requests;
3. A computational model based on the CNN architecture is trained and it shows significant improvement over strong training-free baselines.

The remainder of this article provides more details on the research objectives, the concept of financial risk tolerance, and its relation to personality traits (Section 2); Section 3 elaborates on a meta-analysis of risk tolerance calculation, synthesis of datasets, and the model that predicts risk tolerance from text; Sections 4 and 5 present the experimental results; Section 6 analyzes and discusses the experimental results; Finally, future works of this study are discussed in Section 7.

2. Literature review

2.1. The concept of risk tolerance

Previous studies have discussed multiple risk tolerance-related concepts, including risk attitude, risk aversion, risk preference, risk appetite, risk capacity, etc. A brief exhibition of such concepts is provided in Grable (2018). Due to the nebulous nature of those concepts, there are no widely agreed precise definitions yet. However, I try to distinguish them primarily based on Grable (2018) to create clarity for terminologies used in this article: the construction is illustrated in Fig. 1.

The overall risk profile is used as the umbrella term that considers both the investor's psychological state and other objective factors, such as his/her principal amount, income, life cycle, and many more. Despite the complexity and interdependence between the subjective and objective factors as reported by Piovesan and Willadsen (2021) and Prinz et al. (2014), risk tolerance is used to summarize the effect of subjective factors. The objective factors, on the other hand, determine risk capacity, which evaluates an individual's financial ability to withstand financial losses. Risk aversion is treated as the antonym of risk tolerance. It is theorized that risk tolerance is further influenced by other contextual cognitive biases, and finally forms the risk perception. Risk perception and risk capacity together contribute to risk preference, which is represented in economic analysis as a utility function and refers to the general feeling that one choice is better than another. This risk preference explains the risk-taking behavior of a rational agent. In the construction of Fig. 1, it is clear that an investor's high risk preference does not necessarily mean that the investor's risk tolerance is high, but may also be attributed to a low risk capacity or other cognitive biases.

The review by Hertwig et al. (2019) concluded that what is called risk tolerance here was a moderately stable psychological trait with both general and domain-specific components when measured through self-reports but not behavioral tests. Sahm (2012) pointed to the relatively stable risk preference according to a panel of 12,003 individuals over a decade. More previous studies show that risk tolerance was a stable personality trait and was unlikely to change substantially over life (Van de Venter et al., 2012), which supported the theory of Nicoletta Marinelli and Palmucci (2017) that risk tolerance was a genetic, predispositional, and stable personality trait. To summarize, it is reasonable to model and predict risk tolerance at an individual level since it does not change drastically over time.

2.2. Risk tolerance and personality traits

The study on the correlation between risk tolerance and personality traits requires a well-defined theory of personality. Cattell (1943) pioneered the computational study of personality by factor analysis and cluster analysis, leading to the identification of the 16PF (personality factor) structure. Five repeated factors in experiments of self-ratings, staff ratings, and teammate ratings were later discovered from the 22 variables in Cattell's work. In another research (Norman, 1967), four experts refined these factors through word selection criteria, semantic analysis, and classification, giving rise to five broad personality dimensions (McCrae & John, 1992), named as Extroversion (EXT), Neuroticism (NEU), Agreeableness (AGR), Conscientiousness (CON) and Openness (OPN). This theory is known as the Big Five personality traits today and remains popular in human–computer interactions and computational social science studies, e.g., Lee and Wu (2022). Subsequent research has employed vocabulary and questionnaire methods to validate the structure of these dimensions.

Using the construction of Big Five traits, Epstein and Garfield (1992) classified investors into different personality types and concluded that only when users invest in stocks that are consistent with their personality types can they receive income. Later, Lauriola and Levin (2001) showed that personality traits can predict preferences for gains and losses. People with high openness scores can tolerate higher risks, while investors with high neuroticism scores are more inclined to avoid risk. Durand et al. (2008) examined relationships between Big Five personality traits and investment decisions according to portfolios of 21 Australian investors, which showed that individuals who had more openness were more able to withstand investment portfolios with high risk. Lee et al. (2010) found that individuals with high agreeableness, high intelligence scores, and low rigorous scores can accept more losses. A 2014 survey (Prinz et al., 2014) showed that agreeableness and openness modestly affected students' financial decision-making. Ozer and Mutlu (2019) found that conscientiousness, agreeableness, and openness have significant effects on financial behavior. Most recently, Exley et al. (2021) and Rodrigues and Gopalakrishna (2023) reported that the significance of different personality traits may be unstable and different across generations: the uncontrolled demographic feature of samples may be a reason for discrepancies in research findings. Gambetti and Giusberti (2019) discovered that anxious individuals were likely to save money and avoid investments, perceiving high risks with low control and returns, while people with high extroversion, self-control, and independence would make more investments. Lai (2019) concluded that perceived behavioral control of individuals regarding stock investment is influenced by personality traits of agreeableness, extraversion, conscientiousness, and openness.

Personality has also been associated with more complicated behavioral finance variables other than risk tolerance, such as investor prejudices, sentiment, overconfidence, and herding. A review article reported that conscientiousness had a positive relationship with overconfidence. Baddeley et al. (2010) conducted a simulated task for a functional magnetic resonance imaging (f-MRI) analysis and revealed that herding tendencies were negatively related to sociability (including extraversion and empathy), while positively related to risk-taking (including impulsivity and venturesomeness).

Based on the abundant empirical evidence elaborated above, I hypothesize that personality information is closely related to risk tolerance. If personality information can be detected from texts, the same source may also contain important clues for the individual's risk tolerance.

2.3. Personality detection from text

Unlike the financial risk tolerance profiling task, personality detection from text is a well-studied area. Many machine learning models, including Support Vector Machine (SVM) and Naive Bayes classifier, are applied to use linguistic features for personality detection, such as the Mairesse feature (Mairesse et al., 2007), Medical Research Council (MRC) dictionary (Wilson, 1988), and Linguistic Inquiry and Word Count (LIWC) (Tausczik & Pennebaker, 2010). Deep learning models that have been described for personality detection are mainly variants of CNNs and RNNs (recurrent neural networks, e.g., bidirectional LSTM and GRU) or a combination of them. For instance, Majumder et al. (2017) applied CNN to process textual features. Sun et al. (2018) proposed a model that combined LSTM and CNN, and tried to capture the number of sentence vectors that were closely connected in some coordinates. They also concluded that persons with the same traits were likely to express sentiments in similar ways. Rahman et al. (2019) compared several activation functions including $sigmoid(\cdot)$, $tanh(\cdot)$ and leaky $ReLU(\cdot)$ for personality detection from text, and found that the overall performance using $tanh(\cdot)$ was better than the other two activation functions. Ren et al. (2021) employed text sentiment analysis and BERT to generate sentence-level embedding: this technique improved detection performance on both the Myers–Briggs Type Indicator (MBTI) labeled and Big Five labeled datasets. Yang et al. (2023) designed a CNN-LSTM with a word-layer-person hierarchical attention network (wlpHAN) and a fine-tuning module for personality detection. Ablation analysis suggested that the correct attention mechanism, data augmentation, and fine-tuning are useful for this task.

Based on the wide acceptance of CNN as an effective text feature extractor and classifier especially for short and social media texts (Kim, 2014), the risk tolerance profiling model in this article uses the CNN architecture in a similar manner as described in Majumder et al. (2017).

3. Methodology

3.1. Deriving risk tolerance labels from personality traits

One major challenge in the proposed risk profiling task is the lack of high-quality and aligned risk tolerance labels. In this study, a meta-analysis is conducted to summarize a linear regression model from the literature to infer risk tolerance levels based on the Big Five model.

Three studies by Pak and Mahmood (2015), Pinjisakikool (2018), and Wong and Carducci (2013) are compared because they all used linear regression methods to establish the relation between risk tolerance and personality scores, though different scales were used in the original questionnaires. In order to agglomerate the results from different studies, I first transform different scales into a 5-point scale system. Subsequently, these risk tolerance levels will be used as the supervision and ground truth for model evaluation.

When the dependent variable and the independent variable in the regression equations have a linear relation, the dependent variable and the independent variable can be respectively normalized. If we set X as a function of the independent variable x and its scale in the original questionnaire, and let the minimum value and maximum value of the original scale be a and b, then the normalization is:

$$X = \frac{x-a}{b-a}. \tag{1}$$

Set Y as the new dependent variable whose desired minimum value in the new scale system is A and the maximum value is B, then,

$$Y = (B-A) \times X + A. \tag{2}$$

Substituting formula (1) in formula (2), the transformation becomes:

$$Y = (B-A) \times \frac{x-a}{b-a} + A. \tag{3}$$

In Pinjisakikool (2018), the regression equation is:

$$\begin{aligned} risk\_tol_7 = 2.936 &+ 0.125EXT_5 + 0.121OPN_5 \\ &- 0.176AGR_5 - 0.096CON_5 - 0.112NEU_5, \end{aligned} \tag{4}$$

where the personality scale is a 5-point scale, and the risk tolerance is a 7-point scale. Therefore, risk tolerance needs to be re-scaled to 5-point using formula (3) as follows:

$$risk\_tol_7 = (risk\_tol_5 - 1) \times \frac{7-1}{5-1} + 1 = \frac{3}{2}\ risk\_tol_5 - \frac{1}{2}. \tag{5}$$

Substitute $risk\_tol_7$ with formula (5), we will have:

$$\begin{aligned} risk\_tol_5 = 2.29 &+ 0.083EXT_5 + 0.08OPN_5 \\ &- 0.117AGR_5 - 0.064CON_5 - 0.075NEU_5. \end{aligned} \tag{6}$$

Similarly, both personality and risk tolerance in the research of Pak and Mahmood (2015) are 6-point scales, and the regression equation is as follows:

$$risk\_tol_6 = 4.037 - 0.187AGR_6 + 0.317OPN_6. \tag{7}$$

By transforming the independent variables and the dependent variable into the 5-point scale respectively, a model aligned with the one from Pinjisakikool (2018) is obtained as below.

$$\begin{aligned} 1.25\ risk\_tol_5 - 0.25 = 4.037 &- 0.187 \times (1.25AGR_5 - 0.25) \\ &+ 0.317 \times (1.25OPN_5 - 0.25). \end{aligned} \tag{8}$$

This can be further simplified as:

$$risk\_tol_5 = 4.2545 - 0.187AGR_5 + 0.317OPN_5. \tag{9}$$

Similarly, both personality and risk tolerance in the research of Wong and Carducci (2013) are 9-point scales, and the regression equation is as follows:

$$risk\_tol_9 = 4.44 + 0.02EXT_9 + 0.18OPN_9 - 0.13AGR_9 - 0.15CON_9. \tag{10}$$

By transforming the independent variables and the dependent variable into 5-point scales, we can get:

$$risk\_tol_5 = 2.67 + 0.2EXT_5 + 0.18OPN_5 - 0.13AGR_5 - 0.15CON_5. \tag{11}$$

By summarizing the regressive results from the three studies, that are, formula (6) (9) and (11), we will have:

$$\begin{aligned} risk\_tol_5 = 3.0715 &+ 0.094EXT_5 + 0.192OPN_5 \\ &- 0.145AGR_5 - 0.071CON_5 - 0.025NEU_5. \end{aligned} \tag{12}$$

Formula (12) suggests that Openness and Agreeableness (coef. > 0.1) are the two most prominent personality traits that influence the individual's risk tolerance level. This interpretation is also consistent among the studies by Pak and Mahmood (2015), Pinjisakikool (2018), and Wong and Carducci (2013). The corresponding 5-point average and median risk tolerance scores in different studies are subsequently transformed and presented as in Table 1, showing the heterogeneous populations these studies are conducted on. It can be observed that the research of Pinjisakikool (2018) pooled a conservative population (claimed to be representative of the Dutch population), whereas the research of Pak and Mahmood (2015) accessed a higher risk tolerance group (potential private investors in a post-Soviet transition country, i.e., Kazakhstan).

Table 1
Descriptive statistics of reported risk tolerance scores after transformation.

	Mean	Median	Min	Max
Pinjisakikool (2018)	1.9	1.89	–	–
Pak and Mahmood (2015)	3.736	3.896	–	–
Wong and Carducci (2013)	2.75	–	–	–

Table 2
Descriptive statistics of inferred risk tolerance scores on personality datasets.

risk_tol/dataset	Source	#users	Mean	Median	Min	Max
MyPersonality (Markovikj et al., 2021)	Facebook	250	3.34	3.36	2.74	3.69
Essay (Pennebaker & King, 1999)	Students	2479	3.18	3.18	2.53	3.84
PAN15 (Pardo et al., 2015)	Twitter	334	3.32	3.29	2.93	3.62

Table 3
Distribution of risk tolerance levels among surveyed population.

risk_tol	Our targeted percentage	Actual number of users
gambler	10	273
willing after research	40	1067
cautious	40	887
risk avoider	10	240

Table 4
Data samples from the synthesized corpus.

User ID	Text	Big Five labels
02002056707	"this is my first writing assignment of college" "it does not seem like it could be so bad" "in fact , college itself is not so bad yet" … …	ynynn
64e929be3ff0	"found out that Jolly Pirate Donuts near her house Awesome" "is feeling a little subbydub today" "has a new baby sister Little Baby NoName" … …	nyyny
… …	… …	… …

3.2. Synthesizing a risk tolerance corpus

Because the major challenge of this study was the lack of risk labels for texts, an essential requirement is for the textual data to have labeled features that has been established to associate with risk tolerance. Grable (2016) listed 11 highly relevant factors (p.25, Table 2.1), where personality information is more often collected than other demographic information in NLP research. Therefore, three representative datasets for personality research, i.e., MyPersonality (Markovikj et al., 2021), Essay (Pennebaker & King, 1999), and PAN-15 (Pardo et al., 2015) are used to synthesize a corpus for risk profiling. At the data pre-processing step, I converted all letters to lowercase letters and removed all non-ASCII characters. For Twitter (X) data, I replaced hashtags with the plain text of the tags, and removed @ tags and URLs. Long sentences are divided into several short sentences, and the last short sentence may be shorter than the max length and padded. In the experiments, the max length is set to 20 words.

The fields left in this combined dataset include user ID, content, and Big Five personality. Among them, the PAN-15 dataset includes the Twitter content of 334 Twitter users (152 in English). The texts published by the same user are first combined into one piece of long text, and in the subsequent data pre-processing step again divided according to their length. The value of users' Big Five personality in the PAN-15 dataset is from [−0.5, 0.5], where the value is proportionally mapped to [0, 5] in order to calculate the risk tolerance of each user. The value range of Big Five personality for the MyPersonality dataset is already [0,5]. The personality traits of the dataset Essays have only binary values 'y' and 'n', which are mapped to 3.75 and 1.25 respectively, to fit into the interval of [0,5]. Then, the user's risk tolerance scores are calculated according to formula (12). The results, shown in Table 2, illustrate the high distributional consistency among all three component datasets. The last dataset preparation step is to categorize continuous risk tolerance scores. To achieve this, I refer to survey results of demographic distributions from previous research (Kim et al., 2021), and rank and divide the user's risk tolerance scores proportionally (see Table 3). The dataset size is considered appropriate when referred to other psychometric research, e.g., Manolika (2023) and Zhu et al. (2022). Data samples from this corpus are exhibited in Table 4.

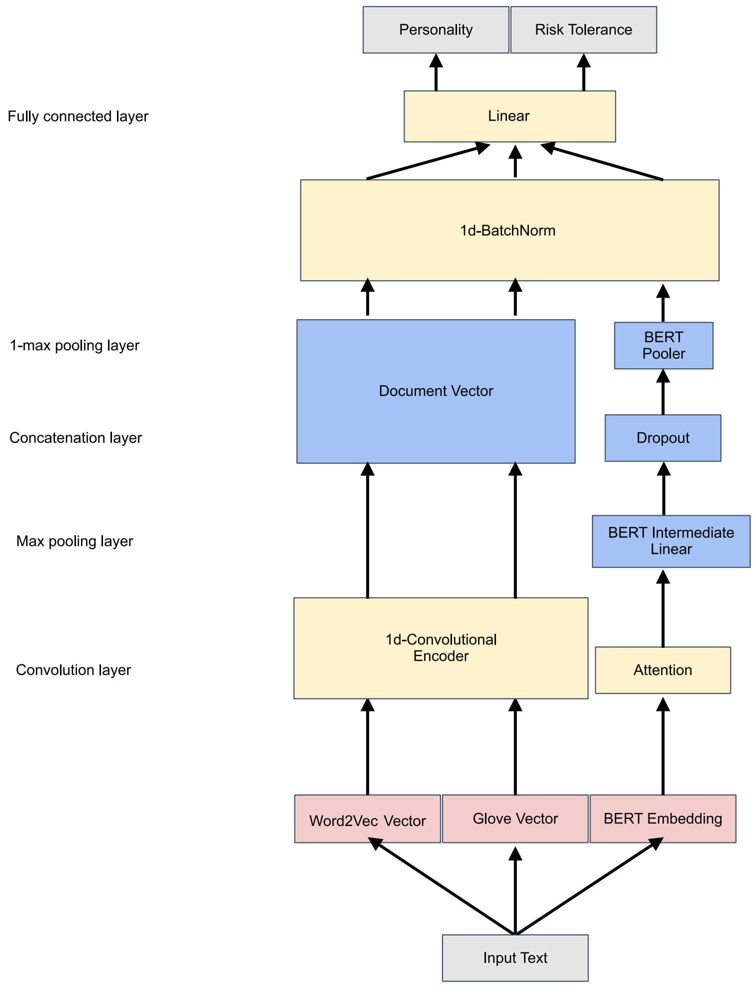

Fig. 2. A CNN model for text-based financial risk tolerance profiling.

3.3. Model architecture and implementation details

A CNN model is built based on the architecture described by Majumder et al. (2017) and several useful model features are experimented with to test for their effectiveness. Fig. 2 illustrates the model architecture. In detail, the following features may improve the model performance according to the literature:

1. Richness of representations: Using multiple text representations is a key factor that influences the model performance. Recent studies, e.g., Yang et al. (2023) have shown that psychologically inspired lexicons and middle layers from large language models provide additional useful information to the network input. The network input in Fig. 2 is a concatenation from sentence embeddings, including Word2Vec (Mikolov et al., 2013), Glove (Pennington et al., 2014), and BERT (Devlin et al., 2019), to preserve semantic information as much as possible.
2. Text augmentation: This is often useful when the model training phase underfits or overfits because of limited data size. Yang et al. (2023) reported SPDFiT (Self-Taught Personality Detection Fine-Tuning), which uses Bayesian learning to assign possible pseudo labels for new texts. In this study, the textaugment Python library[2] is used to substitute words and create semantic equivalents of existing texts. Synonymous substitution is a common method in NLP, which increases the amount of data in the dataset. The method is dedicated to providing more training data, thus improving the classification effect of short texts through global augmentation methods.
3. Multi-task learning: Previous studies documented that personality detection may be learned with closely related tasks, such as internet use behaviors (Mark & Ganzach, 2014) and emotion detection (Li et al., 2022). The multi-task fashion is thus

[2] https://github.com/dsfsi/textaugment

experimented, i.e., combines the 5 personality traits and risk tolerance as outputs for the same network, so that parameters can be shared between the two tasks. Cross entropy loss function is used, where personality traits remain in 2 categories ('y' and 'n'), and risk tolerance was divided into 4 categories.

For the BERT embeddings, "bert-base-uncased"[3] with 10% dropout is used. Each contributing representation has an output dimension of 100 after batch normalization. These together with the Mairesse features form a final in-feature size of $3 \times 100 + 84 = 384$ for the fully connected layer (see Fig. 2). The representations are not frozen and will also be trained. Model parameters are empirically set: training batch size = 16, and maximum epoch = 4. A standard Adam optimizer (learning rate = 0.001 and weight decay = 0) from the PyTorch package is used.

3.4. Linguistic features

This study uses linguistic features from Mairesse et al. (2007) and applies the author's original Java program to extract features. In particular, the feature set includes some features of the Medical Research Council (MRC) Psycholinguistic Database and Linguistic Inquiry and Word Count (LIWC). The MRC machine-usable dictionary contains both linguistic and psycholinguistic attributes for 150,837 words (Wilson, 1988). The LIWC dictionary (Tausczik & Pennebaker, 2010) contained attributes that reflect different emotions, thinking styles, social concerns, and even parts of speech. The MRC database of Oxford Text Archive (Wilson, 1988) is used for calculating linguistic features. Finally, a total of 84 features were extracted, including 70 features of LIWC and 14 features of MRC.

For the sake of coverage, three models, i.e., Word2Vec, Glove, and BERT, are used to produce sentence embeddings. Word2Vec was developed by simply training a neural network for the next word prediction task (Mikolov et al., 2013), which aimed to obtain a vectorized representation of the word through the context of the word. Glove (Pennington et al., 2014) applied a co-occurrence matrix, and considered both local and global information. This study used pre-trained Word2Vec and Glove vectors. Bidirectional Encoder Representations from Transformers (BERT) is a larger model of pre-training language representations developed by Google. Unlike the fixed word representations for Word2Vec and Glove, BERT representations are at the sentence level and jointly produced from a neural network. BERT (Devlin et al., 2019) included pre-training and fine-tuning on various specific tasks. BERT was unsupervised and could use only plain text corpus for training.

In this research, out-of-vocabulary words are counted for their frequencies of occurrence. If the frequency is greater than or equal to the threshold (=1 in our case), a separate word vector for this word will be created with the randomized values of each dimension between $[-0.25, 0.25)$ to match the pre-trained embeddings. The dimensions of the word/sentence representations in this study are 300 for Word2Vec and Glove, and 768 for BERT.

4. Experiment

To make better use of our size-limited data for training, 10-fold cross-validation has been implemented. Cross-validation also provides more information about the performance metrics stability of the experimented model and enables robustness testing. Cross-validation randomly samples the corpus into 10 portions. Only one portion is left as the test set each time, and the remaining nine portions are used as the training set. Subsequently, performance metrics were calculated on each test set and averaged to obtain the final result as reported in Table 5. Beside data, the variances introduced by models are minimal. Experiments show that performance metrics will converge with different initialization manual seeds. The dispersion information is also used to show the significance of performance differences in Table 6.

Table 5 enables ablation analysis for the introduction of each new feature as well as comparisons to several training-free baseline metrics reported in the first three rows. Strategic guess assumes that the risk tolerance level distribution information (Table 3) is available and generates classification labels according to those probabilities. The recent generative language models[4] GPT-3.5 and GPT-4 are prompted using the below template to classify the user texts into different risk tolerance levels. When the response does not contain a classification or refuses to answer, the strategic guess results are used. Except for those ill-answered cases, the GPT models are not prompted with knowledge of the probability distribution.

```
completion = openai.ChatCompletion.create(
 model="gpt-model-name",
 messages=[
  {"role": "system", "content": "You are a financial advisor,
  skilled in understanding and judging the financial risk tolerance level of a client through conversations.
  You will rate the client's risk tolerance level from 0 to 3.
  0 means low tolerance and 3 means high tolerance."},
  {"role": "user", "content": "[example content 1]"},
  {"role": "assistant", "content": "1"},
  {"role": "user", "content": "You are doing a great job."},
  {"role": "user", "content": "Here is another client [example content 2]"} ]
)
```

[3] https://huggingface.co/bert-base-uncased

[4] https://platform.openai.com/docs/models

Table 5
Experimental results with different model settings on the synthesized corpus.

Model settings	Macro-precision	Macro-recall	Macro-F1	Micro-precision	Micro-recall	Micro-F1
Strategic guess	0.2500	0.2500	0.2500	0.3400	0.3400	0.3400
gpt-3.5-turbo	0.2484	0.2424	0.2221	0.3489	0.3489	0.3489
gpt-4-1106-preview	0.2512	0.2506	0.2222	0.3587	0.2590	0.2842
CNN (W)	0.2391	0.2896	0.2538	0.4711	0.4711	0.4711
CNN-aug (W)	0.2367	0.2854	0.2540	0.4750	0.4750	0.4750
CNN (G)	0.2445	0.2996	0.2621	0.4938	0.4938	0.4938
CNN-MT (G)	0.2416	0.3035	0.2690	0.4830	0.4830	0.4830
CNN-MT (W+G+B)	**0.2569**	**0.3086**	**0.2774**	**0.5066**	**0.5066**	**0.5066**

Table 6
Descriptive statistics and robustness test results (micro-F1).

	Strategic guess	CNN (W)	CNN-MT (W+G+B)
Sample mean	0.3244	0.4711	0.5066
Standard deviation	0.0351	0.0094	0.0179
Sample size	3	10	10
		Welch's t-value	p-value
Strategic guess/CNN (W)		7.1624	0.0095***
CNN (W)/CNN-MT (W+G+B)		5.5525	0.0001***

5. Results and robustness tests

The experimental results in Table 5 show that training or fine-tuning is very important to the risk tolerance profiling task. It is important to note that the CNN-based results in Table 5 are using a fixed manual seed (seed = 0) for generating random numbers, therefore do not reflect the universal or the best performances. Although GPT is believed to be a model of basic reasoning capability and commonsense knowledge, it does not significantly outperform the strategic guess. This may indicate that a large amount of useful (risk-related) textual features are not covered in those large language models yet. By using simple training, i.e., exposing the predictive model to textual features, the CNN (W) model already shows significant improvement from zero-shot learning without text information in terms of the micro-F1 metric (Table 6). CNN (W) is the model described by Majumder et al. (2017): it used just the Word2Vec embeddings and changed the target output from personality traits to the risk tolerance level. The CNN-MT (W+G+B) model is an improved version with multi-tasking and rich textual embedding inputs. By testing whether the average performance metrics are significantly different with two unknown unequal standard deviation samples (Zimmerman, 2012), Table 6 shows that, even based on the small sample sizes, leveraging the textual features and constructing an appropriate architecture are useful for this new task.

6. Discussion and implications

In this section, the implications of the experimental results are further discussed. In terms of large language models, it is interesting to observe that GPT-4 is not much superior to GPT-3.5 and optimizes precision over recall. A closer investigation reveals that GPT-4 refrains from answering more often, probably due to safety tuning, so the metrics are inclined to those of strategic guess. When comparing CNN-based models, there are observable improvements when using richer embeddings: the additional Glove representation improves CNN by over 0.02, and the additional Glove and BERT representations improve CNN-MT by over 0.02 in terms of micro-F1 scores. The expansion of embeddings seems a major source of model improvement other than training or fine-tuning. A possible reason is that risk tolerance (the target in the task) information largely resides in the language context.

Text augmentation is experimented on the CNN (W) model. Marivate and Sefara (2020) studied the effect of different approaches to text augmentation, and found that augmentation reduced the possibility of over-fitting. After performing synonym replacement of the training set, the number of records in the new dataset was twice that of the original dataset. The number of records in the test set remained unchanged. The results showed that text augmentation, again, only has minimal effect on the model performance metrics. Therefore, this feature is abandoned from the final CNN-MT(W+G+B) model. In fact, combining different sources of data, instead of text augmentation, seems to be more effective. This is evidenced by comparing with model settings where only the Essays (Pennebaker & King, 1999) data is used.

A common belief is that multi-tasking improves closely related tasks. For instance, Li et al. (2022) designed a multi-task model framework to predict personality traits and emotional behaviors simultaneously, which performed better than a single CNN model, especially in the measurement of recall. The experimental results here, however, show that multi-tasking with personality is not so effective, especially in the case of financial risk profiling. CNN-MT (G) only achieves a comparable macro-F1 to CNN (G) and its micro-F1 is even slightly lower (0.4830 < 0.4938). These results indicate that the new task does not tend to overfit to the data, and is not complimentary to the personality detection task.

Based on the above discussions on comparing different model variants, the final model is set as using all the Word2Vec, Glove, and BERT representations, predicting personality traits and risk tolerance types together based on the synthesized dataset. This final model achieves the best results across all the metrics, including accuracy, precision, recall, and F1 score. It is observed that improving micro-metrics is easier. This is because the risk tolerance classes are skewed: accurately predicting the “gambler” and “risk avoider” types is difficult. The macro-metrics are significantly affected by averaging with the low precision and recall components. It is also observed that the improvement in micro-metrics is more balanced, whereas macro-precision remains similar across the models in Table 5: the improvement in macro-metrics mainly comes from the higher recalls.

This study has two important *theoretical implications* for the information science and information management field. First, it adds knowledge to the recent hype that large language models are good at every professional task. The experimental results show that GPT models’ performance is only comparable to a strategic guess for financial risk profiling. Indeed, in many cases the outputs are “Based on the provided text, it is difficult to assess your risk tolerance level. Could you please share more information about your financial goals, investment preferences, and attitude towards financial risks?” or “You seem to have a mix of cautiousness and determination, which suggests a moderate risk tolerance”. The outputs do not use the Big Five personality categories and only show a superficial understanding of risk tolerance related concepts. The study indicates training to be important for this task, which echoes the recent findings that domain adaptation (Suzuki et al., 2023) and descriptive prompting (Wen et al., 2023) are needed for financial analysis and personality detection. Second, the study proves user generated texts to be a useful information source for financial planning (Heo et al., 2022). With a carefully built deep learning model, micro-F1 can be significantly improved from strong baselines (circa 0.34) to circa 0.50. Given the unbalanced data distribution, this means the binary classification problem (“will-to-take-risk” and “more-cautious”) is basically solved. However, it seems more difficult to identify the more extremely risk-taking or risk-averse investors. This indicates that the risk profiling process as a whole may still need some human intervention.

This study also has *practical implications* for information systems researchers and algorithm engineers. The risk tolerance profiling task needs knowledge of applied psychology. Consequently, the richness of embeddings (especially including LIWC, etc.) is a primary influence factor on the model performance. It is also empirically tested that other techniques from personality detection, such as text augmentation and multi-task learning, are less effective for the risk tolerance profiling task. The model can be integrated into the risk profiling practices, which are required for customer knowledge assessment, investment product recommendation, etc. The model result may replace a formal questionnaire in low-stake situations, and be used as an assistive tool to remind financial advisors when there is a significant discrepancy in the risk profiles created from multiple channels (Xing et al., 2019a).

7. Conclusion and future works

In this study, a new task of financial risk tolerance profiling from the textual data produced by users is defined. A CNN model similar to those used for personality detection is developed, and experimented with several features. The final model uses Word2Vec, Glove, and BERT representations, predicts personality traits together with risk tolerance, and combines training data synthesized from three different sources. This model achieves a micro-F1 score of 0.5066 for the 4-category classification problem, which is circa 4% improvement from the simple CNN (W) model and significantly superior to strong training-free baselines.

The biggest limitation of this study is that the risk tolerance labels are derived through the synthesis of multiple datasets created for personality detection studies and meta-analyses. It becomes implausible to contact the anonymous patients and survey them for the risk tolerance ground truth or to further validate the labels. Nevertheless, several important findings are reported. First, the relation between personality traits and risk tolerance level is better understood quantitatively. Second, fine-tuning is the most important component of the financial risk profiling task, and richer psycho-linguistic features are more important than text augmentation or multi-tasking. Third, it has been proved that user-generated texts (both from a more controlled lab environment and online digital footprints) are useful information for risk tolerance profiling.

Future works would include investigations on what are the useful risk-related textual patterns; explorations on the possibility of integrating non-textual features from other risk profiling tools, such as demographic data and structured questionnaires, into CNN; and data collection that aligns personality traits and risk tolerance using individual identifications.

CRediT authorship contribution statement

Frank Xing: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization.

Data availability

Data will be made available on request.

AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to improve the readability of certain sentences. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgment

The author would like to thank Xiuyu Chen for helping with data collation and software development.

References

Ainia, N. S. N., & Lutfi, L. (2019). The influence of risk perception, risk tolerance, overconfidence, and loss aversion towards investment decision making. *Journal of Economics, Business, & Accountancy Ventura*, *21*(3), 401–413.
Athota, V. S., Pereira, V., Hasan, Z., Vaz, D., Laker, B., & Reppas, D. (2023). Overcoming financial planners' cognitive biases through digitalization: A qualitative study. *Journal of Business Research*, *154*, Article 113291.
Baddeley, M., Burke, C., Schultz, W., & Tobler, T. (2010). Impacts of personality on herding in financial decision-making. *Cambridge Working Papers in Economics*, *1006*, 1–36.
Cattell, R. B. (1943). The description of personality: basic traits resolved into clusters. *Journal of Abnormal and Social Psychology*, *38*(4), 476–506.
Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT* (pp. 4171–4186).
Durand, R. B., Newby, R., & Sanghani, J. (2008). An intimate portrait of the individual investor. *Journal of Behavioral Finance*, *8*(3), 193–208.
Epstein, I., & Garfield, D. (1992). *The psychology of smart investing: Meeting the 6 mental challenges*. John Wiley & Sons, ISBN: 978-0-471-55071-6.
Exley, J., Doyle, P., Snell, M., & Campbell, W. K. (2021). OCEAN: How does personality predict financial success? *Journal of Financial Planning*, *34*(10), 68–86.
Gambetti, E., & Giusberti, F. (2019). Personality, decision-making styles and investments. *Journal of Behavioral and Experimental Economics*, *80*, 14–24.
Grable, J. E. (2016). Financial risk tolerance. In *Handbook of consumer finance research* (pp. 19–31). Springer, ISBN: 9783319288871.
Grable, J. E. (2018). *Financial risk tolerance: A psychometric review*. CFA Institute Research Foundation, ISBN: 978-1-944-96020-9.
Hemrajani, P., Rajni, Khan, M., & Dhiman, R. (2023). Financial risk tolerance: A review and research agenda. *European Management Journal*, *41*(6), 1119–1133.
Heo, W., Kwak, E. J., & Grable, J. E. (2022). The role of big data research methodologies in describing investor risk attitudes and predicting stock market performance. In *Handbook of research on new challenges and global outlooks in financial risk management* (pp. 293–315). IGI Global.
Hertwig, R., Wulff, D. U., & Mata, R. (2019). Three gaps and what they may mean for risk preference. *Philosophical Transactions of the Royal Society B*, *374*(1766), Article 20180140.
Kim, Y. (2014). Convolutional neural networks for sentence classification. In *Proceedings of EMNLP* (pp. 1746–1751).
Kim, K., Hanna, S. D., & Ying, D. (2021). The risk tolerance measure in the 2016 survey of consumer finances: New, but is it improved? *Journal of Financial Counseling and Planning*, *32*(1), 86–103.
Lai, C.-P. (2019). Personality traits and stock investment of individuals. *Sustainability*, *11*(19), 5474.
Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. *Personality and Individual Differences*, *31*(2), 215–226.
Lee, K., Kraeussl, R., & Paas, L. (2010). *Personality and investment: Personality differences affect investors' adaptation to losses*: *Technical report 7*, (pp. 1–19). Faculteit der Economische Wetenschappen en Bedrijfskunde.
Lee, P.-J., & Wu, T.-Y. (2022). Mining relations between personality traits and learning styles. *Information Processing & Management*, *59*(5), Article 103045.
Lengkeek, M., Finn, v. d. K., & Frasincar, F. (2023). Leveraging hierarchical language models for aspect-based sentiment analysis on financial data. *Information Processing & Management*, *60*(5), Article 103435.
Li, Y., Kazemeini, A., Mehta, Y., & Cambria, E. (2022). Multitask learning for emotion and personality traits detection. *Neurocomputing*, *493*, 340–350.
Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, *30*, 457–500.
Majumder, N., Poria, S., Gelbukh, A. F., & Cambria, E. (2017). Deep learning-based document modeling for personality detection from text. *IEEE Intelligent Systems*, *32*(2), 74–79.
Manolika, M. (2023). The big five and beyond: Which personality traits do predict movie and reading preferences? *Psychology of Popular Media*, *12*(2), 197–206.
Marivate, V., & Sefara, T. (2020). Improving short text classification through global augmentation methods. In *Lecture notes in computer science*, (pp. 385–399).
Mark, G., & Ganzach, Y. (2014). Personality and internet usage: A large-scale representative study of young adults. *Computers in Human Behavior*, *36*, 274–281.
Markovikj, D., Gievska, S., Kosinski, M., & Stillwell, D. (2021). Mining facebook data for predictive personality modeling. *Vol. 7*, In *Proceedings of the international AAAI conference on web and social media* (pp. 23–26).
Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*, 77–91.
McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215.
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In *International conference on learning representations, workshop track proceedings* (pp. 1–12).
Moreschi, R. W. (2005). An analysis of the ability of individuals to predict their own risk tolerance. *Journal of Business & Economics Research*, *3*(2), 39–48.
Nasir, T., & Malik, M. K. (2024). Efficient CRNN: Towards end-to-end low resource urdu text recognition using depthwise separable convolutions and gated recurrent units. *Information Processing & Management*, *61*(1), Article 103544.
Nicoletta Marinelli, C. M., & Palmucci, F. (2017). Mind the gap: Inconsistencies between subjective and objective financial risk tolerance. *Journal of Behavioral Finance*, *18*(2), 219–230.
Nobre, L. H., & Grable, J. E. (2015). The role of risk profiles and risk tolerance in shaping client investment decisions. *Journal of Financial Service Professionals*, *69*(3), 18–21.
Norman, W. T. (1967). *2800 personality trait descriptors: normative operating characteristics for a university population*. Ann Arbor: University of Michigan.
Ozer, G., & Mutlu, U. (2019). The effects of personality traits on financial behaviour. *Journal of Business, Economics and Finance*, *8*(3), 155–164.
Pak, O., & Mahmood, M. (2015). Impact of personality on risk tolerance and investment decisions: A study on potential investors of Kazakhstan. *International Journal of Commerce and Management*, *25*(4), 370–384.
Pardo, F. M. R., Celli, F., Rosso, P., Potthast, M., Stein, B., & Daelemans, W. (2015). Overview of the 3rd author profiling task at PAN 2015. In *CEUR workshop proceedings*: *Vol. 1391*, *Working notes of CLEF 2015 - conference and labs of the evaluation forum, toulouse, France, September 8-11, 2015* (pp. 1–40).
Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296–1312.
Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of EMNLP* (pp. 1532–1543).
Pinjisakikool, T. (2018). The influence of personality traits on households' financial risk tolerance and financial behaviour. *Journal of Interdisciplinary Economics*, *30*(1), 32–54.
Piovesan, M., & Willadsen, H. (2021). Risk preferences and personality traits in children and adolescents. *Journal of Economic Behaviour and Organization*, *186*, 523–532.
Pompian, M. M., & Longo, J. M. (2004). A new paradigm for practical application of behavioral finance. *Journal of Wealth Management*, *7*(2), 127–146.
Prinz, S., Grunder, G., Hilgers, R., Holtemoller, O., & Vernaleken, I. (2014). Impact of personal economic environment and personality factors on individual financial decision making. *Frontiers in Psychology*, *5*, 1–11.
Rahman, M. A., Al Faisal, A., Khanam, T., Amjad, M., & Siddik, M. S. (2019). Personality detection from text using convolutional neural network. In *International conference on advances in science, engineering and robotics technology* (pp. 1–6).
Ren, Z., Shen, Q., Diao, X., & Xu, H. (2021). A sentiment-aware deep learning approach for personality detection from text. *Information Processing & Management*, *58*(3), Article 102532.

Rodrigues, C. G., & Gopalakrishna, B. (2023). Financial risk tolerance of individuals from the lens of big five personality traits – a multigenerational perspective. *Studies in Economics and Finance*.

Saha, P., Bose, I., & Mahanti, A. (2016). A knowledge based scheme for risk assessment in loan processing by banks. *Decision Support Systems*, *84*, 78–88.

Sahm, C. R. (2012). How much does risk tolerance change? *Quarterly Journal of Finance*, *2*(4), Article 1250020.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 429–442.

Siering, M. (2023). Peer-to-peer (P2P) lending risk management: Assessing credit risk on social lending platforms using textual factors. *ACM Transactions on Management Information Systems*, *14*(3), 25:1–25:19.

Sun, X., Liu, B., Cao, J., Luo, J., & Shen, X. (2018). Who am i? Personality detection based on deep learning for texts. In *IEEE international conference on communications* (pp. 1–6).

Suzuki, M., Sakaji, H., Hirano, M., & Izumi, K. (2023). Constructing and analyzing domain-specific language model for financial text mining. *Information Processing & Management*, *60*(2), Article 103194.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54.

Thavaneswaran, A., Liang, Y., Paseka, A., Hoque, M. E., & Thulasiram, R. K. (2021). A novel data driven machine learning algorithm for fuzzy estimates of optimal portfolio weights and risk tolerance coefficient. In *30th IEEE international conference on fuzzy systems, FUZZ-iEEE 2021, Luxembourg, July 11-14, 2021* (pp. 1–6).

Van de Venter, G., Michayluk, D., & Davey, G. (2012). A longitudinal study of financial risk tolerance. *Journal of Economic Psychology*, *33*(4), 794–800.

Vinciarelli, A., & Mohammadi, G. (2014). A survey of personality computing. *IEEE Transactions on Affective Computing*, *5*(3), 273–291.

Wen, Z., Cao, J., Yang, Y., Wang, H., Yang, R., & Liu, S. (2023). DesPrompt: Personality-descriptive prompt tuning for few-shot personality recognition. *Information Processing & Management*, *60*(5), Article 103422.

Wilson, M. (1988). MRC psycholinguistic database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, *20*, 6–10.

Wong, A., & Carducci, B. J. (2013). Does personality affect personal financial risk tolerance behavior? *The IUP Journal of Applied Finance*, *19*(3), 7–18.

Xing, F., Cambria, E., & Welsch, R. (2019a). *Robo-Advisory* (pp. 113–122). Springer, ISBN: 9783030302634.

Xing, F., Cambria, E., & Welsch, R. E. (2019b). Growing semantic vines for robust asset allocation. *Knowledge-Based Systems*, *165*, 297–305.

Xing, F., Malandri, L., Zhang, Y., & Cambria, E. (2020). Financial sentiment analysis: An investigation into common mistakes and silver bullets. In *Proceedings of COLING'20* (pp. 978–987).

Yang, K., Lau, R., & Abbasi, A. (2023). Deep learning personality measurement from text. *Information Systems Research*, *34*(1), 194–222.

Yao, R., & Hanna, S. D. (2005). The effect of gender and marital status on financial risk tolerance. *Journal of Personal Finance*, *4*(1), 66–85.

Yekrangi, M., & Abdolvand, N. (2015). Are individual stock investors overconfident? Evidence from an emerging market. *Journal of Behavioral and Experimental Finance*, *5*, 35–45.

Yekrangi, M., & Abdolvand, N. (2021). Financial markets sentiment analysis: developing a specialized lexicon. *Journal of Intelligent Information Systems*, *57*(1), 127–146.

Yin, C., Jiang, C., Jain, H., & Wang, Z. (2020). Evaluating the credit risk of SMEs using legal judgments. *Decision Support Systems*, *136*, Article 113364.

Zhu, Y., Hu, L., Ge, X., Peng, W., & Wu, B. (2022). Contrastive graph transformer network for personality detection. In *Proceedings of iJCAI'22*.

Zimmerman, D. W. (2012). Heterogeneity of variance and biased hypothesis tests. *Journal of Applied Statistics*, *40*(1), 169–193.