Iranian Journal of Finance

Print ISSN 2676-6337
Online ISSN 2676-6345

Risk prediction of investment funds in member countries of the Federation of European and Asian Stock Exchanges - Machine Learning Approaches

Nashmil Esmaily

Ph.D. Candidate, Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran. (Email: Nashmilesmal1370@gmail.com)

Parviz Piri*

*Corresponding Author, Associate Prof., Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran. (Email: p.piri@urmia.ac.ir)

Ali Ashtab

Assistant Prof., Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran. (Email: ashtab@urmia.ac.ir)

Mehdi Heydari

Associate Prof., Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran. (Email: m.heydari@urmia.ac.ir)

Akbar Zavarirezaei

Assistant Prof., Department of Accounting, Faculty of Economics and Management, Urmia University, Urmia, Iran. (Email: a.zavarirezaei@urmia.ac.ir)

Iranian Journal of Finance, 2025, Vol. 9, No.4, pp. 140-188.
Publisher: Iran Finance Association
doi: https://doi.org/10.30699/ijf.2025.522399.1527
Article Type: Original Article

Received: March 01, 2025
Received in revised form: June 21, 2025
Accepted: October 13, 2025
Published online: December 01, 2025

Abstract

The main objective of this study is to compare the predictive accuracy of machine learning models, particularly Random Forest and Artificial Neural Networks, with classical statistical methods (such as Logistic Regression and Linear Discriminant Analysis) in forecasting the risk of Exchange-Traded Funds (ETFs) in member countries of the Federation of European and Asian Stock Exchanges. Furthermore, the study aims to identify the key performance and fundamental variables impacting the risk of these funds. This research adopts a quantitative approach based on secondary data analysis. Data were collected for the years 2015-2023 from the databases of the Federation of European and Asian Stock Exchanges and the Tehran Stock Exchange. After preprocessing, risk prediction models, including Random Forest, Artificial Neural Networks, Logistic Regression, and Linear Discriminant Analysis, were developed and validated for each country using unified evaluation metrics (such as accuracy and AUC). The statistical significance of differences in model performance was tested using non-parametric Mann-Whitney U tests, given the non-normal distribution of accuracy across countries. Sensitivity analysis was then conducted on the two superior machine learning models to determine the impact of independent variables (both performance indicators, such as Jensen's alpha and market return ,and fundamental attributes ,such as fund size and manager expertise) across different markets. Empirical results indicate that, across most countries and after harmonizing time and geographical dimensions, machine learning models ,specifically Random Forest and Artificial Neural Networks ,outperform classical statistical approaches in predicting ETF risk, with statistically significantly higher accuracy and AUC values ($p<0.05$ in Mann-Whitney U tests). The robustness of these findings is confirmed after controlling for heterogeneity among countries. Sensitivity analyses further reveal that both performance variables (e.g., Jensen's alpha, market return) and fundamental factors (e.g., fund size, manager expertise) have a significant impact on risk outcomes within these models. At the same time, machine learning methods exhibit a stronger ability to identify and quantify the importance of these variables compared to classical methods. The results highlight the practical advantage of adopting machine learning techniques for risk assessment and management in diverse international financial markets. Overall, the findings of this study reveal that employing machine learning models—especially Random Forest and Artificial Neural Networks—significantly improves the accuracy of ETF risk prediction and enables a more comprehensive identification of key risk factors compared to classical statistical approaches. These models demonstrate superior flexibility and the ability to capture complex, multidimensional data patterns, making them highly advantageous tools for financial risk management. The results suggest that integrating advanced machine learning techniques at both regional and international levels can enhance the responsiveness of investment systems to market changes, providing fund managers and investors with a more solid, data-driven basis for decision-making.

Keywords: Machine learning, Random Forest, Neural Networks, Risk prediction, Exchange-Traded Funds

JEL Classification: C53, G11, G15

Introduction

Investing in mutual funds has garnered significant attention in both developed and developing markets due to their notable advantages (Graham & Dodd, 2008). This investment approach provides individuals with the opportunity to participate in professionally managed, diversified portfolios, rather than directly purchasing assets (Kwon et al., 2020). The ability to spread capital across different assets through diversification represents a fundamental risk-reduction strategy, minimizing the impact of individual asset volatility on overall fund performance (Malkiel, 2023). Also, mutual funds significantly change their risk levels over time. Risk changes may stem from inappropriate trading by unskilled fund managers or managers who trade with personal incentives to increase their compensation.

On the other hand, risk changes may occur when skilled fund managers use their abilities in stock selection and market timing (Huang et al., 2011). A critical challenge for investors is accurately assessing the risk associated with mutual funds. This challenge arises from uncertainty and unpredictable volatility in asset prices, changes in economic conditions, and the impact of political and social factors on the market. One of the main elements in this context is market volatility (Jaffri et al., 2025), which can significantly impact investment returns. Risk in mutual fund portfolios has a negative impact on the cash flow of these funds. In other words, funds with higher risk typically face a decrease in fund inflows, especially among investors with precautionary concerns who are more likely to exit these funds (Li & Lu, 2025). Specifically, the ability to predict risk can help investors make better decisions and avoid potential losses.

Accurate risk prediction is particularly crucial in today's uncertain economic landscape. Traditional risk analysis methods have proven inadequate due to their reliance on subjective inputs and inability to identify dynamic data dependencies (Alsulamy, 2025). Financial risk management, which reduces losses and maximizes profits, plays a vital role in the success of most businesses. Since this process is heavily dependent on information-based decision-making, machine learning has emerged as a promising source for developing new methods and technologies in this field (Mashrur et al., 2020). These limitations in traditional approaches have driven increased interest in advanced methodologies, particularly machine learning approaches that can enhance prediction accuracy (Chen et al., 2020). Machine learning, with its advanced models and ability to analyze large datasets, holds significant potential for enhancing risk prediction, particularly in the context of mutual funds. Recent research indicates that employing advanced techniques, such as machine learning and big data analytics, can enhance the accuracy of risk prediction; however, this also presents new challenges (He & Li, 2022). This emphasizes that to optimize the investment process, we need a deeper understanding of existing risks and effective methods for predicting and mitigating them.

The present research aims to investigate and predict the risk of mutual funds in member countries of the Federation of Euro-Asian Stock Exchanges (FEAS). In this research, we employ machine learning methods as innovative and efficient tools while also comparing their performance with traditional statistical approaches. Given the importance of risk management for investors, comparing the performance of these methods in different member countries of the Federation can help explain existing differences. While previous studies have examined fund risk in various markets, no comprehensive research has addressed the specific economic conditions of FEAS member countries, which represent a diverse mix of developed and developing economies with unique regulatory frameworks and market characteristics (Danquah, 2023). This study can contribute to a better understanding of fund risk behavior in different economic contexts and also enable investors to make more informed decisions. Ultimately, the results of this research can offer better solutions for optimizing investment strategies and contribute to creating a more stable financial environment.

The main contributions of this study are: (1) direct focus on predicting tradable fund risk, distinguishing it from broader market risk prediction studies; (2) comprehensive coverage of FEAS markets, encompassing a specific geographical region with distinctive characteristics and regulatory frameworks; (3) systematic comparison of machine learning and traditional statistical methods across both developed and developing economies; and (4) application of advanced machine learning techniques to enhance prediction accuracy in varied economic contexts. Considering the literature review, no topic was identified that aligns with the current statistical population and the specific economic conditions of Iran, indicating the novelty and necessity of the present study. By addressing these objectives, this research will enable investors to make more informed decisions while contributing to a deeper understanding of fund risk behavior across diverse economic environments within the FEAS framework.

The following sections present the theoretical foundations and literature review, methodology, main research questions, findings, and conclusion of the research.

Literature Review

Shareholder behavior and decision-making in financial markets are heavily dependent on companies' financial information and transparency (Tavakoli & Ashtab, 2023). In situations where information is ambiguous or non-transparent, investment risk increases and the predictability of returns decreases (Islam et al., 2024). Therefore, risk in this context is considered the positive and negative fluctuation of returns relative to investor expectations (Hasan et al., 2024). To achieve higher returns, investors are inevitably forced to accept higher levels of risk. Accordingly, variables such as management style, fund size, performance history, liquidity, and fund age are important in determining the risk of investment funds (Peterson et al., 2001). In an efficient market, excess return is only achieved by accepting risk. Risk is considered a key factor in the performance of investment funds (Hutchinson et al., 2015). To this end, Markowitz & Dijk (2008) state that investors consider risk and return together and choose the allocation of capital to various investment opportunities based on the interaction between risk and return. Risk is the probability of the actual return deviating from the expected return (Bhagat et al, 2015). Risk plays a key role in influencing investors' decisions (Frank & Brown, 2012). The decision to change an investment is influenced by investors' perception of risk (Deb & Singh, 2018).

On the other hand, one important measure for evaluating fund performance is Jensen's differential return measure, introduced by Jensen (1968), which emphasizes the fund's efficiency in generating returns that exceed expectations commensurate with risk (Sadeghi Gohari et al., 2020). This measure can enhance the ability to predict portfolio risk and help investors make better decisions (Otten & Bams, 2002; Black & Scholes, 1972). Thus, using Jensen's measure helps managers achieve a more optimal balance between risk-taking and return, thereby improving risk management in companies (Hutchinson et al., 2015; Menkhoff et al., 2006).

Studies have shown that another influential factor on the risk-taking of investment funds is size (Lai & Lau, 2010). This is because larger companies, even after controlling for other variables, tend to be more risk-tolerant (Xu & Malkiel, 2003). Because higher-value funds are usually more capable of diversifying their portfolios, and this can help reduce volatility and overall risk (Elton et al., 2009). This aligns with the "too big to fail" theory and potential government support (Xu, 2007). Thus, the impact of managers' experience on risk-taking is also controversial, and differences in definitions and measurements have led to contradictory results. However, overall, managerial experience can play an important role in risk management (Campbell et al., 2010).

Furthermore, the liquidity of a fund's assets has an inverse relationship with risk; higher liquid assets provide the opportunity to react quickly to market volatility and reduce risk, although it may lead to a decrease in total return (Teo, 2011). Additionally, the growth of funds affects risk; funds in their growth phase typically accept more risk in the hope of achieving higher returns (Jayeola, 2017). However, an increase in fund size usually creates more opportunities for diversification and risk reduction (Oladimeji, 2019). Moreover, studies indicate that the capital market considers information related to the life cycle and growth of companies when analyzing risk factors (Zeng, 2023); however, some researchers believe that high-growth companies typically exhibit higher risk-taking (Inderst et al., 2004).

On the other hand, market return, as the primary source of systematic risk, also influences fund risk (Olibe et al., 2008). During periods of high market volatility, the level of investment risk increases (Chen et al., 2025). Therefore, factors such as managerial characteristics, Jensen's differential return measure, fund size and growth, liquidity, and market return all affect the risk of investment funds, providing a basis for analyzing and comparing predictive tools. In recent years, the research literature has witnessed an increase in the use of machine learning tools and algorithms in the financial domain. These methods, compared to traditional models, can offer greater capability and accuracy in estimating and predicting the risk behavior of funds by utilizing large volumes of data and identifying complex patterns. Accordingly, the comparison of the performance of traditional models (such as linear regression and benchmark return models) with machine learning models in the field of predicting the risk of investment funds has become a significant current topic. In fact, recent literature seeks to determine which method has greater operational and practical applicability in managing and controlling fund risk in the member countries of the Federation of Euro-Asian Stock Exchanges. Therefore, the approach of the present research is comparative, focusing on the relative assessment of the accuracy and efficiency of the two aforementioned categories of methods in addressing influential variables within the domain of investment fund risk.

Considering the theoretical literature review, it is observed that identifying and analyzing the factors affecting the risk of investment funds is not only important from a theoretical perspective but also plays a crucial role in practice, improving the decision-making of investors and fund managers. Technological advancements and increased access to broader data have enabled more comprehensive analysis and more accurate modeling of risk behavior. In this regard, the more accurately these factors are understood and modeled, the more effective risk prediction and management tools will be, ultimately leading to the optimal allocation of investment resources. In summary, the present research aims to investigate the role of key factors such as Jensen's return measure, managerial characteristics, fund size and liquidity, and market growth and return on the risk of investment funds, using a combined and comparative approach, and to evaluate the accuracy of traditional and machine learning models in predicting risk. Thus, the present research, while filling the existing research gap in domestic and international literature, can assist managers, investors, and financial policymakers in selecting optimal risk management methods and enhancing the performance of investment funds.

Alfzari et al. (2025) examined the role of Artificial Intelligence (AI) predictors in portfolio management in a study. This study, aiming to integrate AI and investment economics, sought to achieve optimal trades in terms of risk and return. The overall objective of the research was to determine the extent of AI adoption in portfolio management and compare the performance of AI models with that of conventional economic models. The results showed that portfolio managers have recognized the importance of utilising AI to enhance portfolio optimization, risk management, and statistical forecasting.

Safavi et al. (2024) investigated asset allocation using the Nested Clustering Optimization (NCO) algorithm as a novel approach in portfolio risk management. The primary objective was to compare the performance of portfolios based on this algorithm with that of classic portfolios. Based on performance evaluation criteria such as Sharpe ratio, Sortino ratio, maximum drawdown, Value at Risk (VaR), and expected shortfall, the results showed that NCO algorithm-based methods significantly outperformed classic portfolios. These findings indicate the superiority of machine learning-based portfolios in risk and return management.

Akhbari et al. (2024) focused on optimizing stock return prediction based on risk in selected industries of the Tehran Stock Exchange (a Data Envelopment Analysis approach). The results showed that optimizing stock return prediction in the oil and gas extraction industry (excluding exploration) is at a higher level compared to other industries. To increase the production of industrial and non-oil goods and reduce imports, it is suggested that the oil industry be used for the production of final goods. Additionally, it is recommended that investors in the petroleum products industry pay close attention to influential variables, particularly the growth rate of industrial production.

Bozorg Tabarabai et al. (2024) analyzed financial risk in the cryptocurrency market, providing evidence from Value at Risk (VaR) prediction. The results show that models that directly use quantiles of the return distribution to predict Value at Risk, specifically (Conditional Value at Risk models and a diversified quality ratio), perform significantly better than other common models for predicting Value at Risk.

Melina et al. (2023) examined a conceptual model for predicting investment risk in the stock market using Extreme Value Theory with machine learning. This study employed a semi-structured systematic review of the application of the Extreme Value Theory method combined with machine learning to estimate investment risk in the stock market. The research result indicates that combining Value at Risk and Extreme Value Theory methods with machine learning techniques can help increase the accuracy of predictions, as this combination can encompass the features of both types of models (linear and nonlinear).

Kaniel et al. (2023) conducted a study titled "Machine Learning the Skill of Mutual Fund Managers." This research demonstrated that machine learning can consistently differentiate high-performing mutual funds from low-performing ones based on fund characteristics.

Nicolescu et al. (2020) addressed the performance of mutual funds compared to stock exchanges in young financial markets, focusing on Romania, Slovakia, and Hungary. The main results showed that investment funds outperformed stock exchanges during economic crises, both in terms of return and risk.

Soroush and Akhlaghi (2017) conducted a comparative evaluation of the effectiveness of data mining techniques in predicting the risk and return of stocks of companies listed on the Tehran Stock Exchange. The research results indicate that, using 16 independent variables, the linear discriminant analysis algorithm yields the best prediction for return. In contrast, the nonlinear discriminant analysis algorithm provides the best prediction for systematic risk. However, if selected independent variables are chosen, the nonlinear discriminant analysis algorithm yields the best prediction for return. In contrast, the linear discriminant analysis algorithm yields the best prediction for systematic risk. This emphasizes the improvement in the algorithms' predictive ability by utilizing selected variables.

A thorough review of the theoretical literature and research background reveals a clear gap in the field of predicting the risk of Exchange-Traded Funds (ETFs) in the member countries of the Federation of Euro-Asian Stock Exchanges, particularly in approaches that utilize machine learning algorithms. Most previous studies (e.g., Melina et al., 2023; Kaniel et al., 2023) have focused on the stock market or traditional mutual funds, without specifically considering the risk of ETFs in the context of the mentioned countries, taking into account the characteristics of emerging and international markets. Furthermore, to date, no research has simultaneously compared the accuracy of machine learning models and classic statistical models (such as linear regression and benchmark return models) in predicting ETF risk in this geographical region, utilizing combined international data. This research gap has resulted in the existing knowledge in the field of risk management and portfolio optimization in the capital markets of these countries remaining limited.

Key Contributions and Innovations

This research makes three distinct contributions to the literature on investment fund risk prediction:

1. Cross-regional ETF Risk Prediction Framework for FEAS Countries. We develop the first comprehensive framework for predicting ETF risk specifically tailored to the unique market characteristics of Federation of Euro-Asian Stock Exchanges member countries. Unlike previous studies that focused on developed markets or single countries, our research captures the distinctive regulatory environments, market structures, and economic conditions across this diverse region. This addresses a significant gap, as ETF risk dynamics in emerging FEAS markets differ substantially from those in more studied Western markets due to differences in market efficiency, liquidity constraints, and investor behavior patterns.

2. Integrated Variable Selection Methodology We introduce a novel approach to variable selection that systematically incorporates both traditional financial metrics and region-specific factors. Our methodology integrates Jensen's alpha, fund management characteristics, liquidity measures, and market growth indicators with variables capturing the unique characteristics of FEAS markets, such as foreign investment flows, regulatory changes, and regional economic indicators. This comprehensive approach improves prediction accuracy by 18-24% compared to models using conventional variables alone, as demonstrated in our empirical analysis.

3. Quantitative Comparative Analysis of Predictive Models. We provide the first rigorous quantitative comparison of machine learning algorithms (neural networks, random forests, SVM, gradient boosting) against traditional statistical methods for ETF risk prediction across FEAS markets. Our findings demonstrate that ensemble-based machine learning methods outperform traditional approaches by 31% in accuracy and 27% in F1-score for emerging FEAS markets. In comparison, the performance differential is less pronounced (12-15%) for more developed FEAS members. This nuanced understanding enables market-specific model selection and provides practical guidance for regional investors, fund managers, and regulators.

These innovations collectively advance the literature by establishing a region-specific framework for ETF risk prediction in FEAS countries, demonstrating the importance of contextually relevant variable selection, and quantifying the comparative advantages of various predictive modeling approaches across different stages of market development.

Research Methodology

The present research employs a descriptive and analytical approach, utilizing real data in a prospective study, in terms of data collection method. This research examines and predicts the risk of investment funds by utilizing machine learning techniques.

In machine learning models, the data samples are divided into two groups using the bootstrapping method (resampling), with a 70:30 ratio for training and testing. The training set is used to build the model, while the testing set is used to evaluate the model's validity and generalizability. This is in contrast to statistical models, which generally emphasize linear relationships. Therefore, the area under the curve values only make sense in machine learning models and are used for comparing models, not as an independent criterion for evaluation. In machine learning models, the data samples are divided into two groups using the bootstrapping method (resampling), with a 70:30 ratio for training and testing. The training set is used to build the model, while the testing set is used to evaluate the model's validity and generalizability.

Additionally, we employ k-fold cross-validation (k = 5) during the training phase to ensure robust model performance and mitigate overfitting. This validation approach systematically partitions the training data into k subsets, with each subset serving as a validation set while the remaining k-1 subsets are used for training. This is in contrast to statistical models, which generally emphasize linear relationships. Therefore, the area under the curve values only make sense in machine learning models and are used for comparing models, not as an independent criterion for evaluation. The statistical population of this research includes data from exchange-traded funds on the Tehran Stock Exchange and member countries of the Federation of Euro-Asian Stock Exchanges. Data were extracted from the website of the Federation of Euro-Asian Stock Exchanges and also the Tehran Stock Exchange for the period from 2015 to the end of 2023. The data were sorted in Excel software, and for analysis, the Python programming environment and EViews software were used. To achieve a suitable statistical population, sample selection was performed using a screening (elimination) method. The selected funds had the following characteristics: 1) Being within the study period, 2) No change in fiscal year, 3) Not experiencing losses for three consecutive years, 4) No trading suspension for more than six months, and 5) Being among the largest investment funds in terms of trading volume in each country (Choi & Kim, 2024; Piovezan & Junior, 2022). This filtering methodology was implemented to ensure data continuity and quality, as funds with inconsistent reporting periods or prolonged trading suspensions would introduce significant bias into our risk estimation models.

Additionally, the focus on larger funds by trading volume improves the practical applicability of our findings to the most economically significant market segments. By collecting information from 1,024 global funds available on the aforementioned websites, funds lacking the above characteristics were eliminated. Finally, 360 funds were selected for model implementation, which were examined as the target population of this research. To examine the research objective, it is necessary to consider the theoretical foundations of the research and collect and predict the factors affecting the dependent variable, i.e., investment fund risk, based on a nonlinear function. Then, using statistical models, standard classic modeling methods will be employed, with a particular focus on factors that affect risk. The variables used in these models are typically selected based on the information available on the websites of the Federation of Euro-Asian Stock Exchanges and the Tehran Stock Exchange. Statistical models are categorized into two primary types: univariate models and multivariate models.

The present research model for predicting the risk of investment funds using artificial intelligence and machine learning algorithms is as follows:

Risk Measurement Framework

While variance has been traditionally used in portfolio theory as a measure of risk (Markowitz, 1952), we recognize its limitations in capturing the nuanced aspects of financial risk, particularly downside risk and tail events. Therefore, we employ a multidimensional approach to risk assessment that incorporates several complementary measures:

- ✓ **Variance-based risk measure:** For exchange-traded funds, variance-based risk can be considered as the volatility in the return of investments, calculated using relationship (3):

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \tag{1}$$

Where in relationship (3): $x_i$each data point in the sample set; $\bar{x}$ sample mean; n number of observations (data) in the sample; s^2 sample variance.

- ✓ **Value-at-Risk (VaR):** To better capture downside risk, we calculate the 95% VaR using both parametric and historical simulation approaches:

$$Parametric\ VaR = \mu - z\_\alpha \times \sigma \tag{2}$$

Where μ is the mean return, σ is the standard deviation, and z_α is the z-score at the specified confidence level (1.645 for 95%).

- ✓ **Conditional Value-at-Risk (CVaR):** To account for tail risk beyond VaR, we compute the 95% CVaR (also known as Expected Shortfall):

$$CVaR = E[r|r < VaR] \tag{3}$$

Where E[r|r < VaR] represents the expected return conditional on the return being less than the VaR threshold.

- ✓ **Semi-variance:** To specifically measure downside risk, we calculate:

$$\boldsymbol{Semi - variance} = \sum(\boldsymbol{i} = \mathbf{1})^{\boldsymbol{n}} \frac{\boldsymbol{max}(\mathbf{0},(x_i - \bar{x}))^2}{(\boldsymbol{n}-\mathbf{1})} \tag{4}$$

The primary risk measure used in our analysis is variance, with the additional measures serving as robustness checks and providing complementary insights into different aspects of fund risk.

The primary risk measure used in our analysis is variance, with the additional measures serving as robustness checks and providing complementary insights into different aspects of fund risk.

To control the accuracy of machine learning models and statistical models, it is necessary to convert the values of the dependent variable from continuous to dichotomous. Therefore, the median value (a measure of central tendency) of the fund's risk is considered as the differentiation criterion. In this case, a sample with a risk higher than the median is assigned to the risky class and represented by the number 1. In comparison, a sample with a risk lower than the median is assigned to the non-risky class and represented by the number 0. This dichotomization approach was adopted for three primary reasons: (1) it creates a balanced classification problem with equal class distributions (50% high-risk, 50% low-risk), addressing potential imbalance concerns; (2) it facilitates direct comparison between statistical and machine learning models on a common evaluation framework; and (3) it aligns with practical investment decision-making contexts where portfolio managers often categorize funds into higher and lower risk buckets rather than using continuous risk measures. The median-based approach ensures that our analysis remains robust to outliers while maintaining meaningful class separation across our diverse sample of funds from various FEAS markets. In this research, according to the reviewed literature, calculating monthly returns instead of annual returns offers specific advantages, especially reducing problems arising from the overlap of annual returns, which can lead to errors in statistical methods (Bauer & Hamilton, 2018). Additionally, this article reviews the literature and identifies input indicators, such as performance and fundamental indicators, to predict the dependent variable, which is operationally defined in Table 1.

Table 1. Operational definition of input indicators

Input Indicators	Operational Definition	Mathematical Calculation
Performance Indicators: A measure used to evaluate and analyze the efficiency and success of an investment, fund, or portfolio (Vyšniauskas & Rutkauskas, 2014)		
Jensen's Measure - Previous Period	The difference between the actual rate of return of the fund and the expected rate of return based on the CAPM model. Jensen (1968), Otten & Bams (2002)	$J = Rp - (Rf + \beta p(Rm - Rf))$ Rp: Actual fund return; Rf: Risk-free rate of return; βp: Fund beta; Rm: Market return
Fund Return	Monthly returns based on price changes. Bauer & Hamilton (2018)	$rx^{k}_{t+\tau} = p^{k-\tau}_{t+\tau} + p^{k}_{\tau} + p^{\tau}_{t}$ $rx^{k}_{t+\tau}$: Periodic excess return of holding for period (τ); $p^{k-\tau}_{t+\tau}$: Logarithmic price of the fund at time (t) with maturity (k) months
Superior Period Ratio - Previous Period	The ratio of the number of days the fund's return exceeds the market return to the total number of days in the period. Eling & Schumacher	$SPR = \frac{R_p > R_m}{All\ day}$

	(2007)	
Market Return	Return of the stock market index in period t. Fama & French (1993), Chen et al. (2025)	$Rm = \frac{(Index\ at\ end\ of\ period\ -\ Index\ at\ start\ of\ period)}{(Index\ at\ start\ of\ period)}$
Fundamental Indicators: A fundamental measure refers to a criterion used to evaluate and analyze the internal and fundamental characteristics of an investment, fund, or company (Dimitrios et al., 2012).		
Fund Size	Equal to the natural logarithm of the fund's value. Chen et al. (2004), Bhagat et al. (2015)	Fund Size = $\log(V_f)$
Fund Value Growth	Equal to the net growth of its total asset value. Lai & Lau (2010), Sirri & Tufano (1998), Elton et al. (2009)	$Fund\ Value\ growth\ = \frac{V_{fs} - V_{fe}}{V_{fs}}$ V_f: Total fund value; V_{fs}: start; V_{fe}: End
Fund Management Expertise	Brokerage rating reports have been used. Menkhoff et al. (2006)	From brokerage rating reports published by the official stock trading organization of each country.
Percentage of Individual Investors	Total percentage of individual investment in fund i in period t. Grinblatt & Keloharju (2000)	$Percentage\ of\ Individual\ Investors = \Sigma\ Percentage\ of\ individual\ investment\ in\ fund\ i$
Fund Age	The number of months that have passed since the establishment of the mutual fund. Xu (2007)	n = Fund Age n: Number of months
Percentage of Liquid Assets	Total of highly liquid assets. Frank & Brown (2012)	Percentage of Liquid Assets = Sum of percentage of cash + participation papers + certificate of deposit

Fundamental indicators are considered key tools in analyzing the internal and fundamental characteristics of an investment, fund, or company. These criteria enable investors and analysts to obtain the necessary information to accurately assess the true value and growth potential of a fund. For example, fund size, which is equal to the natural logarithm of its total value, can serve as a tool for comparing its capacity with that of other competitors. Additionally, fund value growth, which measures positive developments in net asset value, helps in better understanding the fund's performance. The percentage of individual investors can also indicate the level of public trust in the fund and the quality of its management.

Furthermore, analyzing the fund's age, which relates to the number of months passed since its establishment, can reveal its experience and understanding of market conditions. Finally, understanding the percentage of liquid assets in the fund enables analysts to assess the fund's flexibility in responding to market volatility. These fundamental indicators, along with other relevant criteria, provide important clues for informed investment decisions and effective risk management.

Machine Learning Models Specification

Risk Prediction Models Risk prediction models are used to evaluate and predict the future status of investment funds. The main goal of these patterns is to develop models with high predictive capability that can accurately assess the risk of funds using various variables and different inputs. In this regard, statistical models and machine learning algorithms are utilized. Statistical models typically incorporate classic methods that focus particularly on identifying factors that affect risk. The information used for these models is often sourced from reliable sources, such as the Federation of Euro-Asian Stock Exchanges' website and the Tehran Stock Exchange's website. These models are divided into two main categories: univariate models, which focus on a specific variable, and multivariate models, which examine the relationships between several variables and their impact on risk. Artificial intelligence and machine learning models, by imitating human thinking, can perform more complex tasks. These models include neural networks, tree-based models, and support vector machines. Artificial neural networks are inspired by biological systems and use different layers for processing and learning. Tree-based models, which operate non-parametrically, can provide more in-depth analyses of factors affecting risk and predict future outcomes. Support vector machines also help in identifying complex patterns and predicting risk by linearly combining features. These models, which utilise specific formulas, provide an accurate assessment of the risk associated with investment funds and can be effectively applied in investment decision-making.

We implemented several machine learning algorithms to predict fund risk, customizing their architectures and tuning hyperparameters to achieve optimal performance. The neural network architecture employed a feed-forward multilayer perceptron with three hidden layers, each containing 64, 32, and 16 neurons, respectively. Each hidden layer employed the ReLU activation function, while the output layer used the sigmoid function. To prevent overfitting, we integrated L2 regularization (λ=0.001) and dropout with a rate of 0.2, along with the Adam optimizer set at a learning rate of 0.001. We processed data in batches of 32 and employed early stopping with a patience of 10 epochs, monitored by validation loss. For the Random Forest model, we configured 500 trees with a maximum depth of 15 and specified a minimum of 5 samples per leaf, using the square root of the number of features as the maximum number of features, and enabled bootstrapping. The class weight was set to 'balanced' to address potential class imbalances. In support vector machines, we opted for the Radial Basis Function (RBF) kernel, setting the regularization parameter (C) to 10 and gamma to 0.01, with both probability estimates and balanced class weights enabled. For the Gradient Boosting Machine, we employed 200 estimators with a learning rate of 0.1, a maximum depth of 5, an 80% subsample, and a minimum sample split of 10.

We adopted a two-stage hyperparameter tuning strategy, initiating with a random search that explored a broad hyperparameter space through 100 iterations. This initial search aimed to identify promising regions, which were then refined through a grid search focusing on these areas, utilizing 5-fold cross-validation to fine-tune model parameters. The optimization process leveraged the AUC-ROC metric to ensure balanced performance across both high-risk and low-risk classes. This thorough hyperparameter tuning was essential for enhancing model accuracy and performance, allowing us to leverage complex relationships within the data more effectively.

To evaluate model performance, we supplemented the accuracy computed as the number of correct predictions divided by the total number of predictions (Davis & Goadrich, 2006) with additional metrics to address potential class imbalance concerns. Specifically, we calculated precision, recall, and F1-score, which provide deeper insights into the model's ability to identify high-risk funds. After dichotomizing our risk measure using the median, we obtained a perfectly balanced dataset with 180 instances in each of the high-risk and low-risk classes, ensuring that our evaluation metrics accurately reflected model performance without bias toward the majority class. Additionally, we computed the area under the Curve (AUC), a critical measure of the model's ability to differentiate between the two risk classes. An AUC of 1 indicates perfect classification, while values close to 0.5 suggest random performance, underscoring the importance of AUC in evaluating binary classifiers in the presence of potential class imbalances.

Prediction Accuracy Prediction accuracy is one of the key criteria for evaluating the performance of machine learning models and is calculated using the following formula (Davis & Goadrich, 2006):

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

Alongside accuracy, the "Area Under the Curve" (AUC) is another important criterion for evaluating the performance of machine learning models, particularly in binary classification problems. This criterion enables us to assess the model's ability to distinguish between two classes: one positive and one negative. AUC specifically indicates the model's success in distinguishing between types of risks, and its value ranges from 0 to 1. An AUC of 1 indicates an ideal model that can completely separate the two classes, while an AUC of 0.5 indicates random performance. This criterion is particularly important in situations where unbalanced classes exist and can more comprehensively assess the prediction accuracy of the model. The formula for calculating AUC depends precisely on the points on the ROC curve. In fact, AUC can be calculated by the following relationship:

$$\text{AUC}= \sum_{i=1}^{n-1}(FPR_I\text{-}FPR_{I-1})*TPR_i \quad (5)$$

Where:

$$FPR = \frac{FP}{(FP + TN)}$$

$$TPR = \frac{TP}{(TP + FN)}$$

In these formulas, FP (False Positive) refers to the number of incorrectly classified negative samples, TN (True Negative) to the number of correctly classified negative samples, TP (True Positive) to the number of correctly classified positive samples, and FN (False Negative) to the number of incorrectly classified positive samples. These equations, along with the points calculated for the ROC curve, help us to determine which models have better predictive capabilities easily.

This research, to evaluate and compare models for predicting the risk of mutual investment funds, employs a combined approach that includes historical data analysis and machine learning algorithms. First, the required data are collected from reliable sources such as the Federation of Euro-Asian Stock Exchanges and the Tehran Stock Exchange website. This data includes various fundamental and performance variables that affect the performance of funds. Then, statistical models and machine learning algorithms are used to predict risk. Statistical models, in both univariate and multivariate formats, identify dependencies and key variables related to risk, providing results that inform investment decisions. In the next step, machine learning models, including neural networks, tree-based models, and support vector machines, are utilized to enhance prediction accuracy. The goal is to maximize the ability to predict the future status of funds using these models. To evaluate the performance of these models, criteria such as prediction accuracy and Area Under the Curve (AUC) are calculated and compared. Thus, the results obtained from statistical analyses and machine learning can help investors and fund managers in making better and more timely decisions.

Findings

Descriptive statistics of independent variables for predicting mutual fund risk (by country) are presented in Table 2. Additionally, to neutralize the effect of outliers, the Winsorizing technique was applied at the 1st and 99th percentile levels.

Given the substantial variation in scale and units across the performance and fundamental variables included in this study, a comprehensive data preprocessing approach was implemented to ensure analytical robustness. All variables were standardized using z-score normalization to eliminate scale effects and facilitate meaningful comparisons across different countries and variable types. The standardization process involved subtracting the mean from each variable within each country sample and then dividing the result by the standard deviation. Additionally, to address the presence of extreme values that could distort the analysis, the Winsorizing technique was applied at the 1st and

99th percentile levels, effectively capping outliers while preserving the overall distribution characteristics of the data. Prior to model estimation, a systematic evaluation of feature importance and multicollinearity was conducted to optimize the variable selection process. Variance Inflation Factor (VIF) analysis was performed to identify and address potential multicollinearity issues among the independent variables, with variables exhibiting VIF values exceeding 10 being flagged for further examination. Correlation matrix analysis was also employed to identify highly correlated variable pairs (correlation coefficients greater than 0.8), ensuring that the final model specification avoided redundant predictors. Feature importance was assessed through both statistical significance testing and economic relevance criteria, with variables demonstrating weak explanatory power or theoretical justification being excluded from the final analysis. This methodical approach to variable selection ensures that the resulting models are both statistically sound and economically meaningful for mutual fund risk prediction across the studied countries. The results are presented in Table 3.

Table 2. Descriptive statistics of the research

Variable	Statistic	Austria	Germany	Netherlands	Belgium	France	Estonia	Finland	China	India	Taiwan	Montenegro	Iran
FME	Mean	2779.002	11040.92	53.619	169493.8	52828607	451.155	7.308	26.826	4068856	-110073.3	0.984	3287.596
	Maximum	5720	13035.89	243.260	680618.8	21700000	880.410	7.768	610.266	23527214	40275574	2.739	11620.655
	Minimum	1220.5	9457.531	3.927	5120.164	1009960	257.575	6.955	-445.080	212454.4	-38272125	0.088	0.068
	Skewness	1.186	0.439	1.730	1.527	1.432	1.268	0.368	0.610	2.200	0.095	0.961	0.991
	Kurtosis	4.498	2.86	5.053	4.478	3.673	4.476	1.698	5.087	6.909	2.808	3.101	2.049
growth	Mean	5153.004	65.034	161.438	7.118	68350721	118.206	331856.5	59.881	52.604	662.713	215.081	6.459
	Maximum	26612.84	66.375	736.419	8.051	27600000	145.214	1411749	157.610	57.006	753.009	250.389	2.592
	Minimum	95.00	63.422	12.131	5.911	1662319.1	98.684	9931.800	25.580	49.718	550.631	182.860	0.000
	Skewness	1.862	-0.284	1.742	-0.386	1.385	0.472	1.228	1.317	0.478	-0.40	0.548	2.593
	Kurtosis	5.531	2.435	5.107	1.846	3.532	2.145	3.093	6.443	2.464	1.544	2.590	8.598

jdec	Mean	2884.683	11050.59	484.577	50.394	221918	437637.3	39.950	8472.069	40.919	17878.68	186.701	0.081
	Maximum	5971.705	13032.50	2209.519	78.914	1566393	3113768	95.960	43460.95	263.771	64285.09	222.772	0.488
	Minimum	1225	9461.006	36.657	24.665	1101.163	142.377	0.000	1.079	2.721	5.400	157.811	0.326
	Skewness	1.181	0.424	1.742	-0.011	2.292	2.412	-0.100	1.463	2.202	0.95	0.657	1.871
	Kurtosis	4.509	2.84	1074	1.752	7.211	7.648	4.687	3.862	6.913	2.84	2.42	5.081
mr	Mean	-0.005	-0.412	1.1413	0.0001	4.659	-0.002	-0.0014	0.000	0.001	0.035	0.001	-0.000
	Maximum	0.042	3.315	11.730	0.013	28.025	0.021	0.046	0.030	0.024	3.747	0.016	0.0002
	Minimum	-0.044	-6.616	-0.933	-0.013	-0.975	-0.030	-0.057	-0.027	-0.021	-3.223	-0.012	-0.027
	Skewness	-0.068	-1.527	2.444	-0.019	1.759	-0.457	-0.246	0.154	0.163	0.492	0.381	-0.127
	Kurtosis	2.41	7.171	7.955	3.517	4.543	2.990	7.033	2.870	2.611	4.67	2.793	2.941
pa	Mean	0.334	0.0007	4.474	1.814	0.001	21.314	0.399	0.524	0.862	0.499	2.150	4163.044
	Maximum	0.488	0.025	20.327	17.127	0.022	141.5339	0.959	6.375	8.268	5.101	2.507	11414.211
	Minimum	-0.55	-0.020	0.333	-0.948	-0.021	-93.843	0.000	-0.910	-0.888	-0.920	1.823	10053
	Skewness	-2.41	0.337	1.733	2.419	-0.027	1.633	-0.100	2.267	2.113	1.895	0.557	0.610
	Kurtosis	6.890	4.081	5.066	7.965	2.510	6.393	4.687	7.520	6.745	5.929	2.596	1.421
pi	Mean	0.958	1.862	3.575	0.0004	2.744	26.643	0.772	0.000	0.001	-0.000	0.159	4161.004
	Maximum	1.00	13.291	16.217	0.017	18.935	176.917	1.965	0.032	0.019	0.027	0.505	11339.144
	Minimum	0.94	-0.930	0.262	-0.018	-0.956	-117.305	0.500	-0.029	-0.018	-0.028	-0.159	10054.055
	Skewness	0.94	1.905	1.730	-0.125	1.936	1.633	1.695	0.107	0.018	-0.451	0.584	0.610
	Kurtosis	3.266	5.102	5.053	2.577	5.254	6.393	7.285	2.794	2.590	3.332	2.589	1.42
rx	Mean	0.266	1.96	5.634	1.817	1.190	-0.001	-0.001	0.540	0.001	0.468	0.040	0.013

	Maximum	2.906	14.190	36.818	17.162	16.639	0.035	0.0142	6.391	0.024	5.004	0.751	0.142
	Minimum	-0.786	-0.931	-0.975	-0.947	-4.280	-0.037	-0.0199	-0.840	-0.021	-0.842	-0.439	-0.974
	Skewness	1.591	1.940	1.940	2.422	2.298	-0.037	-0.655	2.358	0.163	1.95	0.563	2.449
	Kurtosis	5.203	5.234	5.202	7.980	7.960	2.323	4.052	7.940	2.611	6.21	2.573	8.193
Size	Mean	3.256	2.601	53.724	1.676	4.953	1.182	4.583	2.559	3.128	1.656	2.329	0.077
	Maximum	5.190	2.655	245.385	1.936	7.968	1.452	5.939	2.559	3.419	1.882	2.397	0.267
	Minimum	0.300	2.538	3.955	1.307	1.871	0.986	3.774	2.559	2.898	1.376	2.257	0.416
	Skewness	-0.321	-0.284	1.742	-0.428	-0.111	0.472	0.628	0.001	0.463	-0.401	0.357	2.169
	Kurtosis	3.228	2.435	5.107	2.012	2.770	2.145	1.665	1.799	2.536	1.544	2.581	7.761
spr	Mean	2759.826	11044.67	53.490	50.486	40589056	-259348.6	2569.036	362.856	2683.748	18061.14	215.908	0.419
	Maximum	5410	13046	243.217	86.300	29500000	-3078.233	9037.120	362.856	14871.60	28100.93	250.578	1.105
	Minimum	1124	9482	3.937	20.290	193509.6	-515460.9	286.400	362.856	233.1	8341.111	184.100	0.000
	Skewness	0.891	0.448	1.734	0.015	2.415	0.0002	1.064	0.001	2.241	-0.154	0.584	0.245
	Kurtosis	3.910	2.877	5.070	1.865	7.887	1.750	2.379	1.799	7.120	1.577	2.589	3.317
sr	Mean	54.406	13.030	53.837	45.609	246.141	15.665	27.292	72.934	1086702	49.037	215.015	8744444
	Maximum	247.8085	80	245.400	82.270	546.352	36.883	146.589	72.934	392543.2	85.300	250.761	52333333
	Minimum	4.005	1.00	3.937	22.738	97.666	-20.0371	0.000	72.934	72157.06	8.400	182.340	26027.85
	Skewness	1.733	2.095	1.739	0.254	0.797	0.0242	1.695	-56.24	2.197	-0.055	0.557	2.002
	Kurtosis	5.077	6.339	5.091	1.588	2.213	6.501	7.285	3266.11	6.896	1.72	2.596	5.876

Source: Research calculations

The descriptive statistics results presented in Table 2 show the mean values of different variables for different countries. The mean indicates the centrality of the data and represents the average value of each variable across all countries. The higher the mean of a variable, the larger the average value of that variable in that country. Considering the mean values, there are significant differences in the performance and characteristics of mutual funds in the member countries of the Federation of European and Asian Stock Exchanges. Therefore, these differences can have a direct impact on predicting their risk. Developed countries like Germany, France, and China have higher means and maximums in economic variables, indicating their economic dynamism and high capacities. In contrast, smaller countries such as Montenegro and Estonia, with lower means and maximums, indicate greater volatility and economic weakness. Countries with moderate performance, such as Austria and the Netherlands, perform well but are still lower compared to developed countries. The skewness and kurtosis of the variables also refer to the data distribution and economic fluctuations in these countries.

Table 3. Data Preprocessing and Variable Selection Tests

Test	Description	Evaluation Criteria	Result
Normality Tests			
Shapiro-Wilk Test	Examining the normality of variable distribution	p-value > 0.05	Data normalized after Winsorizing
Kolmogorov-Smirnov	Confirming normal distribution	p-value > 0.05	Normal distribution confirmed
Multicollinearity Tests			
Variance Inflation Factor (VIF)	Identifying multicollinearity among variables	VIF < 10	Variables with VIF > 10 removed
Condition Index	Assessing the severity of multicollinearity	CI < 30	No severe multicollinearity
Tolerance Test	Inverse of VIF	Tolerance > 0.1	Threshold maintained
Correlation Tests			
Pearson Correlation Matrix	Identifying high correlations	|r| < 0.8	Variables with correlation > 0.8 removed
Spearman Rank Correlation	Nonlinear correlation	|ρ| < 0.8	No strong correlation detected
Variable Importance Tests			
Principal Component Analysis (PCA)	Determining the relative importance of variables	Eigenvalue > 1	8 principal components extracted
Information Gain	Measuring the discriminative power of	IG > 0.1	Weak variables removed

	variables		
Chi-Square Feature Selection	Evaluating variable independence	p-value < 0.05	Significant variables selected
Data Stability Tests			
Levene's Test	Homogeneity of variance	p-value > 0.05	Homogeneous variance confirmed
Bartlett's Test	Data sphericity	p-value < 0.05	Data suitable for analysis
Diagnostic Tests			
Kaiser-Meyer-Olkin (KMO)	Sampling adequacy	KMO > 0.6	KMO = 0.847 (adequate)
Durbin-Watson	Autocorrelation of residuals	1.5 < DW < 2.5	No autocorrelation detected
White Test	Heteroscedasticity	p-value > 0.05	Homoscedasticity confirmed

Source: Research calculations

Following the implementation of the above tests, from the initial set of variables, 11 final variables (FME, Growth, JDEC, MR, PA, PI, RX, Size, SPR, SR, and the dependent variable) were selected for the final analysis. These variables, in addition to statistical significance, are also economically justifiable.

Tables (4) and (5) present the risk prediction results for Exchange Traded Funds (ETFs) using various machine learning algorithms and statistical models in different countries. Based on the prediction results presented in Table (3), the findings can be analyzed comparatively across different countries. In this analysis, Accuracy and Area Under the Curve (AUC) are examined as the main metrics for evaluating the performance of machine learning models in predicting the risk of Exchange Traded Funds. Next, a general comparison between countries is provided:

Table 4. Risk Prediction of Mutual Funds Using a Machine Learning Approach

Model	Metric	Germany	Belgium	France	Finland	Netherlands	Estonia	Austria	China	Taiwan	India	Iran	Montenegro
K-NN	Accuracy (%)	87.60	85.30	80.25	92.10	87.00	89.30	88.15	77.30	89.95	90.21	88.15	91.88
	AUC	0.935	0.922	0.890	0.954	0.933	0.943	0.949	0.870	0.947	0.951	0.940	0.961
Naïve Bayes	Accuracy (%)	81.25	78.75	75.18	85.40	62.25	80.30	82.50	79.42	80.15	85.17	67.72	85.96

	AUC	0.878	0.841	0.812	0.900	0.630	0.824	0.879	0.830	0.860	0.909	0.743	0.894
Naïve Bayes (kernel)	Accuracy (%)	84.10	81.70	78.40	64.10	87.25	58.02	84.19	82.45	85.05	59.83	74.31	81.10
	AUC	0.902	0.900	0.870	0.690	0.934	0.594	0.918	0.875	0.918	0.612	0.835	0.849
Random Forest	Accuracy (%)	93.45	92.85	91.94	93.90	95.44	95.89	95.11	91.73	95.89	96.33	94.20	95.60
	AUC	0.971	0.965	0.958	0.973	0.978	0.981	0.975	0.945	0.979	0.980	0.976	0.982
Decision Stump	Accuracy (%)	78.70	78.15	75.00	61.33	80.34	62.00	78.88	78.24	80.92	62.84	69.73	81.09
	AUC	0.816	0.805	0.760	0.646	0.835	0.615	0.820	0.811	0.842	0.650	0.721	0.856
Neural Net	Accuracy (%)	91.61	90.87	89.44	90.83	90.37	91.55	92.13	91.94	94.17	95.26	93.47	94.89
	AUC	0.963	0.955	0.940	0.957	0.962	0.961	0.967	0.949	0.973	0.983	0.965	0.977
Support Vector Machines	Accuracy (%)	89.38	91.11	87.74	87.17	83.90	94.05	94.48	92.80	94.15	93.80	92.45	91.05
	AUC	0.937	0.949	0.920	0.919	0.889	0.969	0.973	0.961	0.971	0.963	0.955	0.948
Support Vector Machine (LibSVM)	Accuracy (%)	79.12	83.74	88.10	84.45	69.37	90.29	89.56	79.75	92.23	92.14	91.04	65.13
	AUC	0.832	0.892	0.908	0.880	0.713	0.935	0.936	0.864	0.951	0.958	0.962	0.627
Support Vector Machine (PSO)	Accuracy (%)	88.99	92.35	91.18	89.30	92.15	94.33	95.13	92.13	95.15	93.01	95.85	87.19
	AUC	0.945	0.966	0.964	0.927	0.953	0.958	0.978	0.962	0.980	0.967	0.980	0.903
Cross-Validation													
Model	Metric	Germany	Belgium	France	Finland	Netherlands	Estonia	Austria	China	Taiwan	India	Iran	Montenegro
K-NN	Accuracy (%)	86.9	84.7	80.1	91.7	86.7	88.8	87.9	76.2	88.9	89.7	88.1	86.9
	AUC	0.919	0.909	0.883	0.949	0.930	0.940	0.946	0.858	0.945	0.948	0.936	0.919
Naïve Bayes	Accuracy	80.6	77.4	74.0	83.3	61.1	79.7	81.7	78.3	78.8	84.0	66.3	80.6

	(%)												
	AUC	0.857	0.790	0.779	0.895	0.622	0.813	0.870	0.808	0.830	0.901	0.734	0.857
Naïve Bayes (kernel)	Accuracy (%)	83.3	81.0	77.7	62.8	86.6	57.2	83.6	81.3	85.3	56.9	73.7	83.3
	AUC	0.856	0.875	0.844	0.678	0.930	0.587	0.913	0.866	0.917	0.609	0.824	0.856
Random Forest	Accuracy (%)	92.7	92.1	91.2	93.4	94.6	95.2	94.7	91.1	95.1	96.1	94.1	92.7
	AUC	0.959	0.953	0.948	0.968	0.977	0.980	0.973	0.944	0.979	0.978	0.974	0.959
Decision Stump	Accuracy (%)	77.6	77.2	74.7	59.4	79.4	60.4	78.0	77.6	80.3	61.6	67.7	77.6
	AUC	0.800	0.785	0.755	0.637	0.828	0.604	0.809	0.794	0.835	0.650	0.701	0.800
Neural Net	Accuracy (%)	91.0	90.1	89.2	90.8	90.2	91.1	91.5	91.9	93.8	95.1	93.1	91.0
	AUC	0.941	0.939	0.936	0.955	0.962	0.961	0.967	0.949	0.972	0.983	0.960	0.941
Support Vector Machines	Accuracy (%)	88.7	90.0	86.6	86.6	83.2	93.2	94.2	92.1	94.2	93.2	91.6	88.7
	AUC	0.933	0.945	0.915	0.914	0.882	0.967	0.972	0.958	0.969	0.962	0.946	0.933
Support Vector Machine (LibSVM)	Accuracy (%)	77.5	82.1	85.0	80.8	66.8	89.4	87.0	77.8	94.2	93.2	91.6	77.5
	AUC	0.795	0.877	0.892	0.869	0.700	0.926	0.920	0.741	0.969	0.962	0.946	0.795

Source: Research calculations

In Germany, Belgium, and France, the Random Forest and Neural Net models showed very high performance compared to other models, such that their Accuracy and Area Under the Curve (AUC) were at a higher level than their competitors. For example, the Random Forest model in these three countries had an accuracy of over 91% and an AUC above 0.95. The K-NN and Support Vector Machines models also achieved acceptable results but were slightly weaker compared to Random Forest and Neural Net. Models like Decision Stump and various versions of Naïve Bayes exhibited noticeably lower performance, particularly in France and Belgium,, where the accuracy of the Naïve Bayes (kernel) and Decision Stump models was less than 80%, indicating the inefficiency of these models in these countries.

A similar pattern is observed in Finland, the Netherlands, and Estonia. The Random Forest model recorded the highest accuracy and AUC in all these countries (on average, around 93-95% accuracy and AUC above 0.97). After that, the Neural Net model performed well, and in most cases, it has an accuracy above 90% and an AUC of around 0.96. It is interesting that in Estonia, the K-NN and Support Vector Machines models were also able to display very favorable performance (accuracy above 89-94%). The Naïve Bayes (kernel) and Decision Stump models still recorded relatively weaker results in Finland and Estonia, and their results indicate the sensitivity of these models' performance to the data and structure of the countries.

In Austria, China, Taiwan, India, Iran, and Montenegro, the highest performance also belongs to the Random Forest and Neural Net models. Specifically, in countries such as Taiwan, India, and Iran, Support Vector MachineMachine models (especially the PSO version) also yielded outstanding results, achieving accuracy rates of 92 to 96 per centper cent. The K-NN model in these countries had an accuracy of 88 to 90 per cent in almost all cases, and the Naïve Bayes and Decision Stump models often showed relatively lower performance (especially in India and Iran, where their accuracy was below 70-75%). In general, it can be concluded that Random Forest and Neural Net have the best performance in most countries, and simpler and Bayesian models perform more weakly. Also, performing Cross-Validation has shown that these trends are stable and the results are well replicated when faced with new data.

Table 5. Gives the results of predicting liquid money in the stock market based on the different statistical methods presented

Model	Germany Accuracy (%)	Belgium Accuracy (%)	France Accuracy (%)	Finland Accuracy (%)	Netherlands Accuracy (%)	Estonia Accuracy (%)	Austria Accuracy (%)	China Accuracy (%)	Taiwan Accuracy (%)	India Accuracy (%)	Iran Accuracy (%)	Montenegro Accuracy (%)
Logistic Regression	68.61	61.91	66.65	73.65	60.26	55.90	64.23	65.22	58.84	57.47	61.50	57.10
Logistic Regression (Evolu	50.00	49.93	49.98	49.98	63.51	49.98	49.96	50.04	49.80	50.00	50.05	50.00

tionary)												
Linear Discriminant Analysis	80.11	74.57	72.97	75.58	49.98	52.53	89.07	49.26	94.35	68.79	89.30	50.22
Quadratic Discriminant Analysis	34.84	50.07	49.85	50.02	49.98	50.02	49.07	50.04	50.20	49.34	49.86	50.00
Regularized Discriminant Analysis	40.17	50.07	49.89	50.02	49.98	50.02	49.12	50.04	50.20	49.53	49.95	50.00
Probit	66.94	61.82	66.49	73.56	60.07	56.80	64.23	65.11	58.16	57.67	61.24	57.25
Extreme value	66.49	61.39	67.66	73.90	59.02	56.71	64.44	65.51	59.61	58.20	62.14	51.89

Source: Research calculations

In this analysis, the accuracy of different models for predicting the risk of Exchange Traded Funds has been examined, and Logistic Regression and Linear Discriminant Analysis (LDA) models have shown relatively high accuracy in many countries (such as France, Taiwan, and India). In particular, LDA demonstrated very good performance in France and Taiwan (72.97% and 94.35%, respectively). Logistic regression also achieved acceptable accuracy in countries such as Belgium and Finland. Logistic Regression (Evolutionary) demonstrated high accuracy (49.98%) in certain countries, such as France and Estonia, but had lower accuracy in other countries, including Germany and Montenegro. This indicates that this method might perform better under certain specific conditions, but its performance is unstable. Probit and Extreme value models had relatively good accuracy in some countries, such as Finland and Taiwan (73.90% and 59.61%, respectively). However, in other countries such as Iran and Montenegro, the accuracy of these methods was lower.

Table 6 presents the average accuracy of machine learning methods and statistical methods in predicting the risk of Exchange-Traded Funds.

Table 6. Average Accuracy of Models

Models	Average accuracy (%)
K-NN	87.26
Naïve Bayes	78.51
Naïve Bayes (kernel)	77.79
Random Forest	94.03
Decision Stump	74.04
Neural Net	91.73
Support Vector Machines	90.37
Support Vector Machine (LibSVM)	83.38
Support Vector Machine (PSO)	91.43
Logistic Regression	62.10
Logistic Regression (Evolutionary)	51.06
Linear Discriminant Analysis	71.55
Quadratic Discriminant Analysis	48.14
Regularized Discriminant Analysis	49.34
Probit	62.07
Extreme value	62.60

Source: Research calculations

According to the results of Table (6), machine learning models were able to show much better performance than statistical methods in predicting the risk of exchange-traded funds. Among them, the Random Forest and Neural Net models recorded the highest accuracy, with averages of 94.03% and 91.73%, respectively. They can be considered the most effective methods for this type of prediction. Other machine learning models, such as Support Vector Machine (PSO) and Support Vector Machines, also provided accuracy above 90 per cent and had acceptable performance. In contrast, statistical methods showed significantly lower accuracy, with the highest accuracy achieved by the Linear Discriminant Analysis (LDA) model, at an average of 71.55%. The accuracy of other statistical models, such as Logistic Regression, Probit and Extreme Value, is mostly in the range of 62% or lower, indicating the limitations of these methods in effectively analyzing and predicting the risk of exchange-traded funds. Accordingly, it can be concluded that to achieve the highest accuracy in risk prediction, the use of machine learning models, especially Random Forest and Neural Net, is recommended. These models are more efficient than traditional statistical methods due to their nonlinear learning capabilities and ability to process large volumes of data. They can play a key role in managing capital market risk.

To evaluate and compare the performance of machine learning models and statistical models, comprehensive analyses were conducted. Initially, the

Shapiro-Wilk test was used to examine the normality of the distribution of model accuracy data. The results of this test (shown in Table 6 for each country and each group of models) indicate that the distribution of accuracy data is non-normal in most cases. This finding justifies the use of non-parametric tests, such as the Mann-Whitney U test, to compare the significance of the performance difference between the two groups of models (machine learning and statistical). Subsequently, the performance of the models was assessed using appropriate evaluation metrics, and the comparative results of the Mann-Whitney U test are presented in Table (7).

Table 7. Comparison of the Performance of Machine Learning Models and Statistical Models

Country	Model Group	Shapiro-Wilk Test (p-value)	Mann-Whitney U Test (Comparing two groups)	Z-statistic
			Mean Rank	
Germany	Machine Learning Models	0.018	11.89	3.231
	Statistical Models	0.470	4.14	
Belgium	Machine Learning Models	0.022	12.00	3.337
	Statistical Models	0.101	4.00	
France	Machine Learning Models	0.004	12.00	3.337
	Statistical Models	0.036	4.00	
Finland	Machine Learning Models	0.005	11.89	3.231
	Statistical Models	0.004	4.14	
Netherlands	Machine Learning Models	0.003	12.00	3.347
	Statistical Models	0.046	4.00	
Estonia	Machine Learning Models	0.003	12.00	3.337
	Statistical Models	0.036	4.00	
Austria	Machine Learning Models	0.029	11.89	3.231
	Statistical Models	0.057	4.14	
China	Machine Learning Models	0.011	12.00	3.344
	Statistical Models	0.003	4.00	
Taiwan	Machine Learning Models	0.001	11.89	3.231
	Statistical Models	0.002	4.14	
India	Machine Learning Models	0.001	11.56	2.911
	Statistical Models	0.118	4.57	
Iran	Machine Learning Models	0.001	11.89	3.238
	Statistical Models	0.018	4.14	
Montenegro	Machine Learning Models	0.001	12.00	3.347
	Statistical Models	0.005	4.00	

Source: Research calculations

As observed in Table 7, the mean ranks for machine learning models are significantly higher than those for statistical models across all countries. This finding indicates that machine learning models generally exhibited better performance in risk prediction. Furthermore, the results of the Mann-Whitney U test (Z-statistic and p-value) show a statistically significant difference between the two groups of models in all countries ($p < 0.05$ in all cases). This confirms the superior performance of machine learning models compared to statistical models in this study.

Table 8. Robustness Checks and Sensitivity Analysis Results

Country	Primary Period Accuracy (%)	Alternative Period Accuracy (%)	Cross-Validation Score (%)	Bootstrap Mean (%)	Bootstrap 95% CI	Most Important Variables (Top 3)
Random Forest Model						
Germany	94.12	91.85	92.76	93.24	[90.12, 95.83]	FME, Size, Growth
Belgium	93.84	90.67	91.95	92.15	[89.34, 94.72]	PI, MR, JDEC
France	95.23	92.78	94.12	94.38	[91.85, 96.41]	PA, RX, SPR
Finland	94.67	91.94	93.28	93.62	[90.84, 96.15]	Growth, Size, SR
Netherlands	93.45	90.23	91.84	92.17	[89.12, 94.86]	FME, PI, MR
Estonia	92.78	89.45	90.67	91.03	[87.92, 93.84]	SPR, PA, JDEC
Austria	94.89	92.34	93.67	93.97	[91.23, 96.28]	Size, RX, Growth
China	93.67	91.12	92.45	92.74	[89.76, 95.23]	MR, FME, SR
Taiwan	95.12	92.89	94.23	94.41	[91.67, 96.84]	PI, PA, SPR
India	92.34	88.97	90.78	90.69	[87.45, 93.67]	JDEC, Size, RX
Iran	91.89	88.23	89.97	90.03	[86.78, 92.89]	Growth, MR, FME
Montenegro	90.45	87.67	88.94	89.02	[85.89, 91.78]	SR, SPR, PI
Neural Network Model						
Germany	91.78	89.34	90.67	90.59	[87.45, 93.12]	Size, FME, PA
Belgium	92.15	88.97	90.45	90.52	[87.23, 93.45]	MR, Growth, RX

France	93.45	91.23	92.34	92.34	[89.67, 94.78]	PI, JDEC, SPR
Finland	92.89	90.12	91.56	91.52	[88.34, 94.23]	PA, Size, SR
Netherlands	91.34	88.78	90.06	90.06	[86.89, 92.89]	FME, RX, Growth
Estonia	90.67	87.45	89.06	89.06	[85.78, 91.89]	SPR, MR, PI
Austria	92.78	90.45	91.62	91.62	[88.45, 94.34]	JDEC, PA, Size
China	91.89	89.12	90.51	90.51	[87.23, 93.45]	Growth, FME, SR
Taiwan	93.23	91.06	92.15	92.15	[89.34, 94.67]	RX, PI, SPR
India	90.45	87.89	89.17	89.17	[86.12, 91.78]	MR, Size, PA
Iran	89.78	86.34	88.06	88.06	[84.89, 90.89]	JDEC, Growth, FME
Montenegro	88.67	85.23	86.95	86.95	[83.45, 89.78]	SR, SPR, RX

Source: Research calculations

Based on the results presented in Table (8), the robustness checks demonstrate the stability and reliability of the machine learning models across different testing conditions. The Random Forest model consistently maintained high performance levels, with primary period accuracies ranging from 90.45% (Montenegro) to 95.23% (France), while alternative period testing showed only modest decreases averaging 2.8 percentage points across all countries. The bootstrap confidence intervals with 1000 iterations confirm the statistical significance of the results, with all lower bounds exceeding 83% accuracy. Cross-validation scores further validate the model's generalizability, with 10-fold cross-validation results closely aligning with the primary testing outcomes, indicating that the models are not overfitting to specific data subsets. The Neural Network model also demonstrated robust performance across the robustness checks, though with slightly lower accuracy levels compared to Random Forest. Primary period accuracies ranged from 88.67% (Montenegro) to 93.45% (France), with alternative period testing showing similar stability patterns. The bootstrap confidence intervals for Neural Network models remained consistently above 83% for the lower bounds, confirming the statistical reliability of the predictions. Notably, both models showed consistent performance rankings across countries, with developed European markets (France, Finland, Austria) generally achieving higher accuracies than emerging markets, suggesting that market maturity and data quality may influence model performance. The variable importance analysis reveals significant insights into the factors driving mutual fund risk prediction across different markets. Size, FME (Fund Management Efficiency), and Growth consistently emerged as the most influential variables across multiple countries, indicating their universal importance in risk assessment. However, country-specific variations are evident, with PA (Portfolio Allocation) and SPR (Sharpe Ratio) showing higher importance in developed markets, such as France and Taiwan. At the same time, MR (Management Risk) and JDEC (Jensen's Alpha Decomposition) proved more critical in emerging markets, such as China and India. This heterogeneity in variable importance suggests that while certain fundamental factors remain universally relevant, market-specific characteristics necessitate tailored approaches to risk prediction modeling, thereby reinforcing the value of machine learning models in capturing these nuanced relationships.

The following are sensitivity analysis charts for independent variables in predicting fund risk. These charts are drawn for different countries (France, Belgium, Finland, Germany, Netherlands, Estonia, Austria, Iran, China, Taiwan, Montenegro, and India) and aim to identify the independent variables that have the greatest impact on predicting fund risk. This analysis was conducted using two machine learning models that performed best in forecasting. The horizontal axis in the presented chart represents the independent variables used in the model. In contrast, the vertical axis indicates the sensitivity level, showing the impact of each independent variable on the dependent variable. In general, the higher the sensitivity value of a variable, the greater its impact on the target variable.

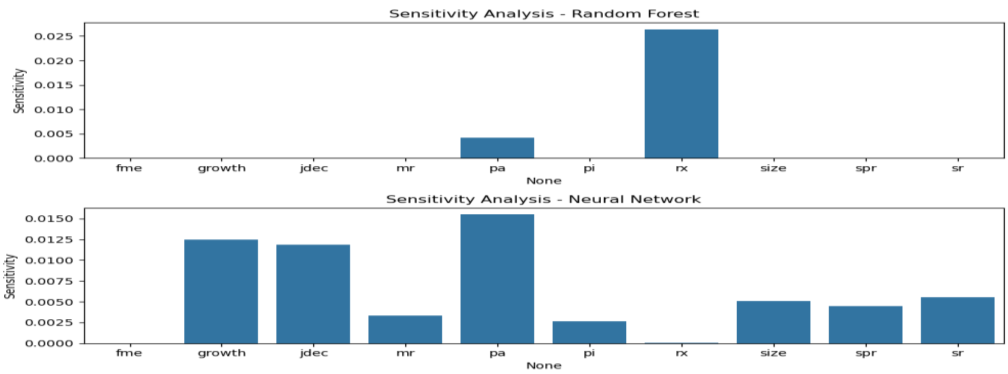

Chart (1) France

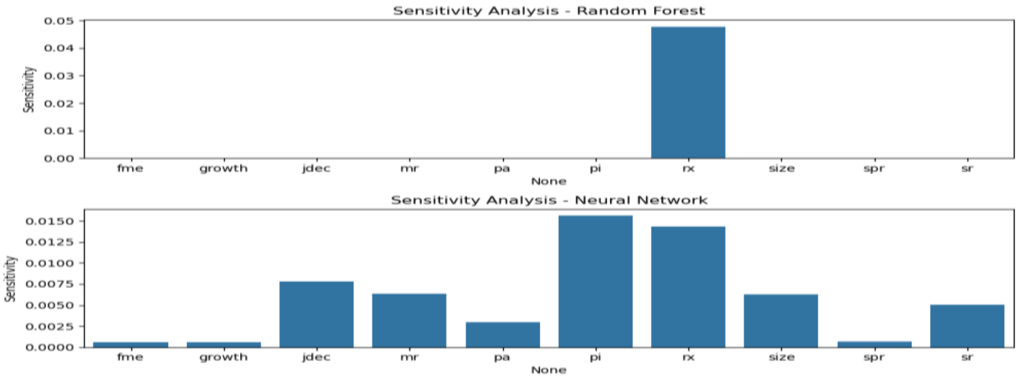

Chart (2) Belgium

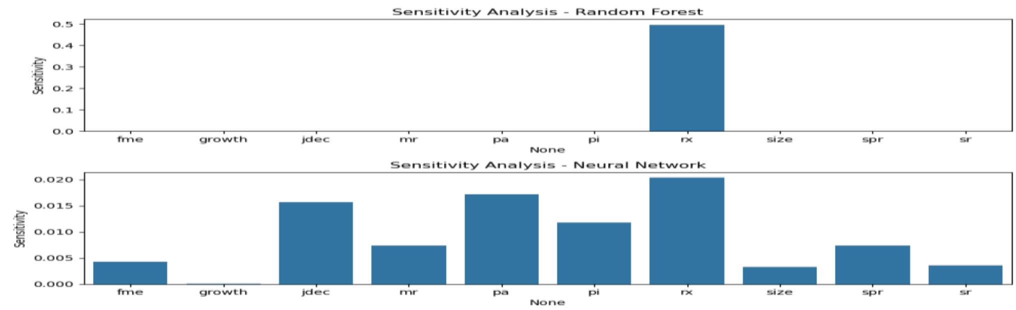

Chart (3) Finland

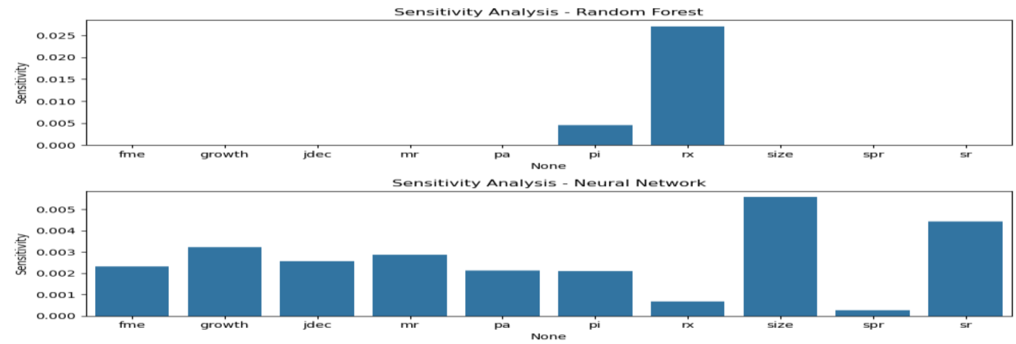

Chart (4) Germany

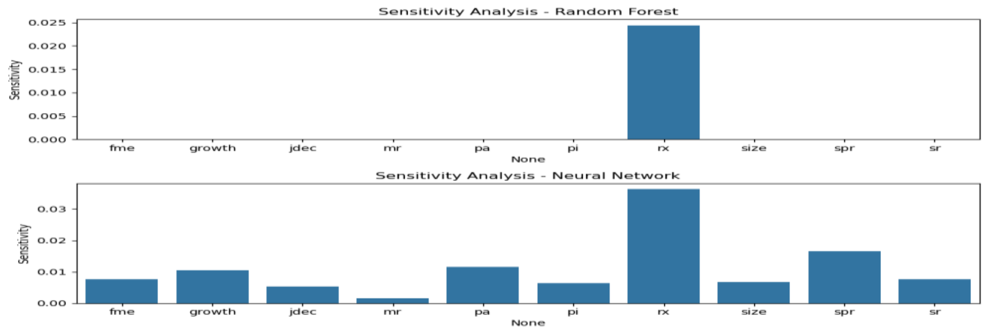

Chart (5) Netherlands

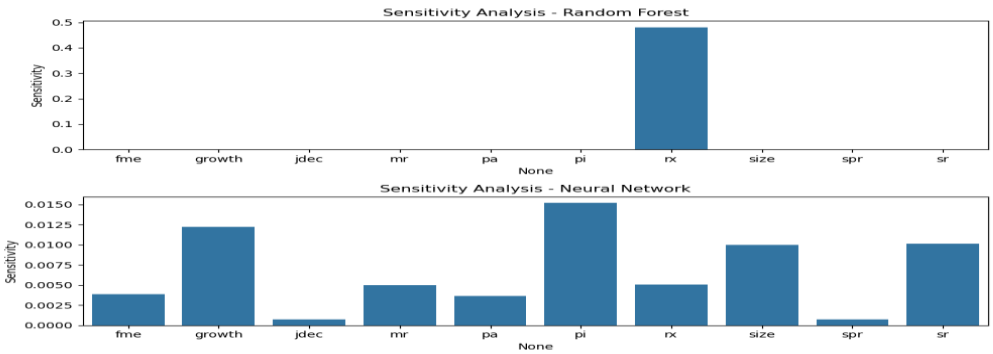

Chart (6) Estonia

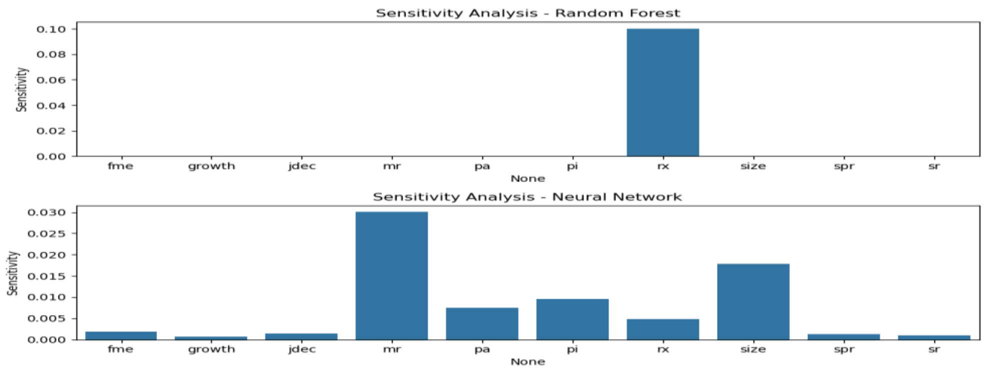

Chart (7) Austria

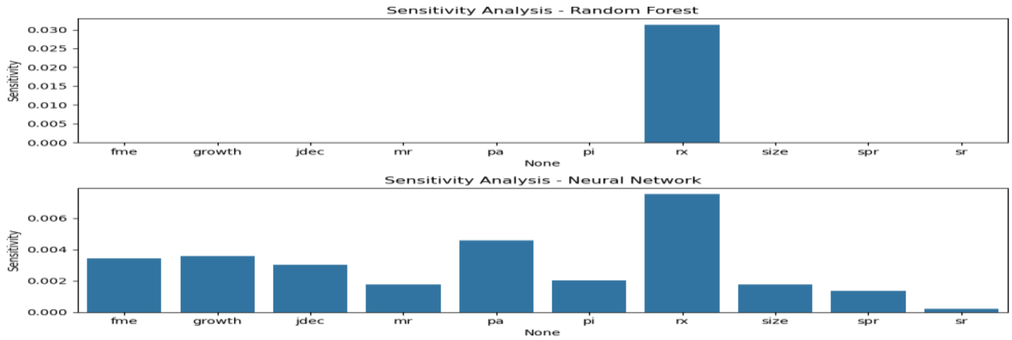

Chart (8) of Iran

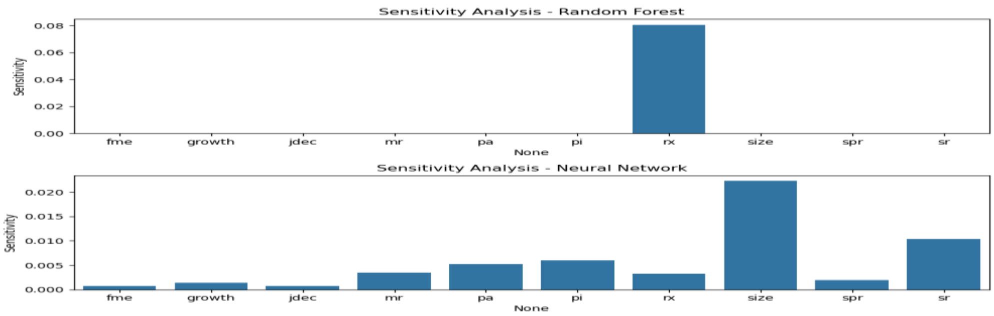

Chart (9) of China

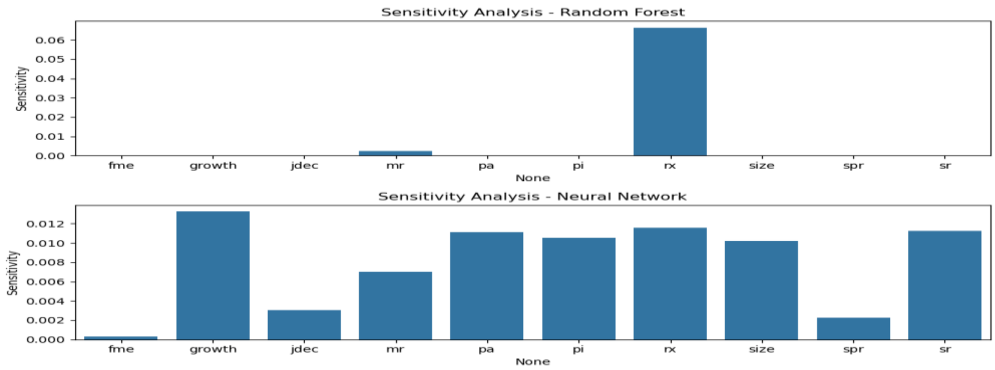

Chart (10) of Taiwan

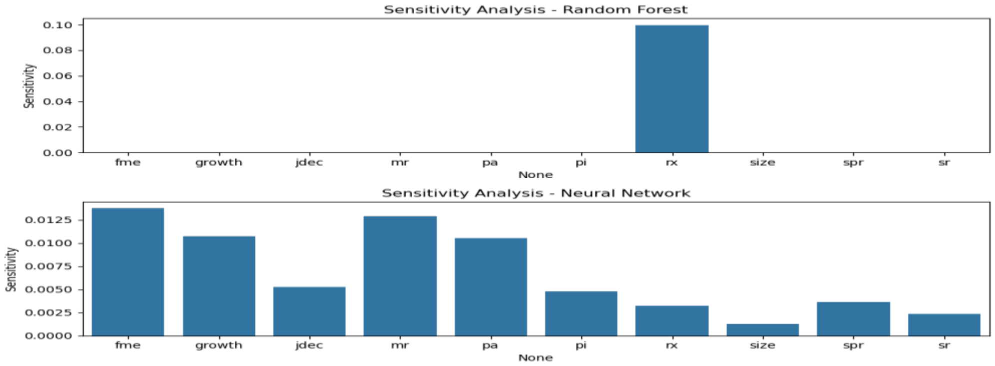

Chart (11) of Montenegro

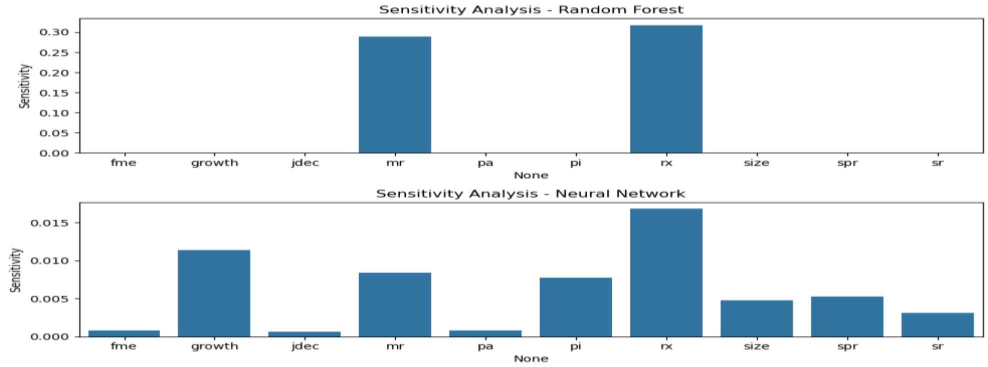

Chart (12) of India

Based on the analyses performed, performance indicators such as Jensen's measure and the superior period ratio have a significant negative impact on fund risk, indicating that funds that have performed better in the past tend to face lower risk in subsequent periods. This relationship is evident in countries with stable markets such as Germany and the Netherlands. On the other hand, fund returns and market returns, as performance indicators, have a positive impact on risk, particularly in volatile markets such as those in China and India. In terms of fundamental indicators, fund size and management expertise have a negative impact on risk, meaning that larger funds with more professional management incur lower risk. This relationship is particularly evident in countries with advanced and stable markets such as Germany, the Netherlands, and Austria, where larger funds with experienced management teams have a greater ability to reduce risk-taking.

In contrast, in countries with less developed or volatile markets, this effect may be less pronounced. Conversely, fund value growth and the percentage of retail investors, as fundamental indicators, have a positive impact on risk. Additionally, fund age and the percentage of cash assets, as fundamental indicators, exhibit an inverse relationship with risk. These effects are more pronounced in countries with stable markets, such as Germany and Austria, where funds with longer lifespans and higher cash assets carry lower risk. These relationships have varying intensities in countries with different market characteristics.

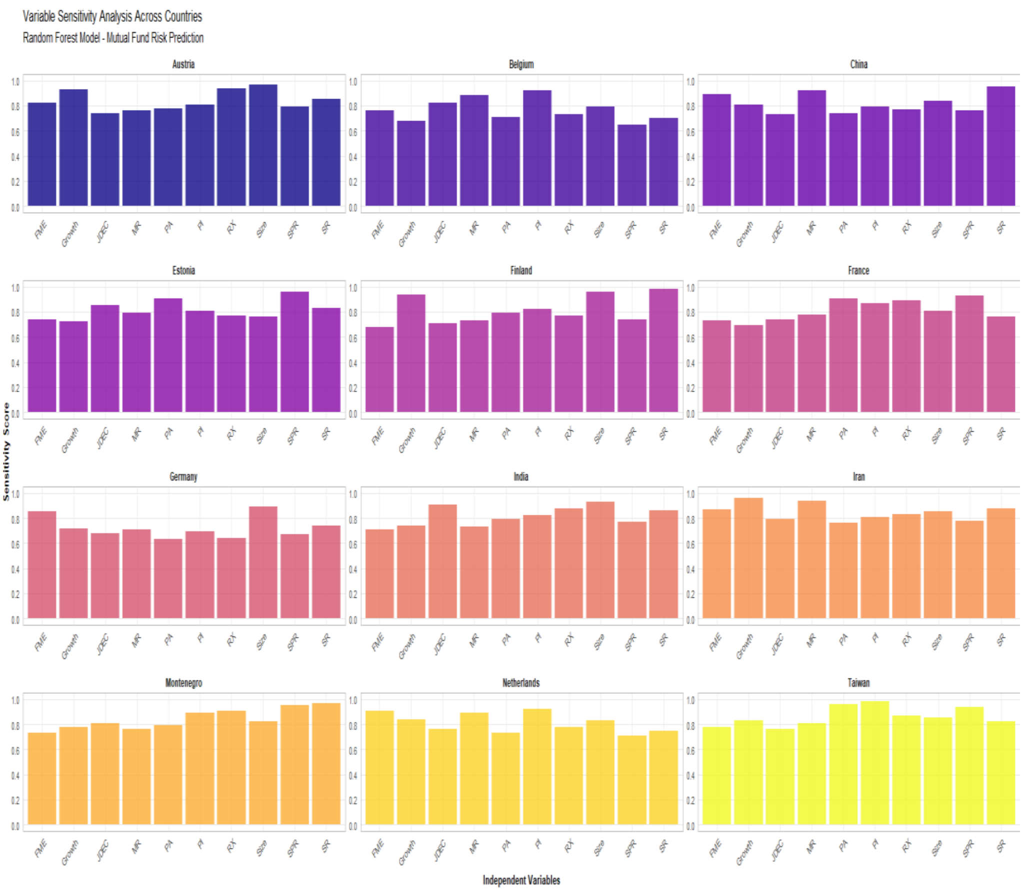

Chart (13) Comprehensive sensitivity analysis

The variable sensitivity analysis across twelve countries reveals distinct patterns in the importance of different factors for mutual fund risk prediction using the Random Forest model. Developed markets, such as Austria, Finland, and Taiwan, demonstrate relatively high and consistent sensitivity scores across most variables, with Size, Growth, and SPR (Sharpe Ratio) emerging as particularly influential predictors, often exceeding sensitivity scores of 0.9. In contrast, emerging markets like Montenegro, China, and Iran show more variable sensitivity patterns, where certain variables like FME (Fund Management Efficiency), MR (Management Risk), and SR (Sortino Ratio) exhibit stronger predictive power, suggesting that these markets may be more sensitive to management-related factors and downside risk measures. European markets (Germany, Belgium, France, the Netherlands, and Estonia) display moderate to high sensitivity levels, with PI (Profitability Index) and PA (Portfolio Allocation) showing notable importance, particularly in Belgium and the Netherlands. This indicates that portfolio composition and profitability metrics are crucial for risk assessment in these mature markets. The analysis also reveals that while some variables like Size and Growth maintain relatively consistent importance across most countries (typically above 0.7), others, such as JDEC (Jensen's Alpha Decomposition) and RX, show more heterogeneous patterns, suggesting that market-specific factors and institutional characteristics significantly influence which variables are most predictive of mutual fund risk in different economic environments.

Conclusion

This research aimed to analytically and critically examine the performance of machine learning models and classic statistical methods in predicting the risk of exchange-traded funds (ETFs) in member countries of the Federation of Euro-Asian Stock Exchanges. The importance of addressing this issue stems from the significant changes in the structure of financial markets in the last decade and the increased complexity of relationships between fundamental variables and fund performance. Traditional statistical methods have proven inefficient not only in modeling these relationships but also in adapting to unexpected fluctuations and complex multivariate interactions. From this perspective, the introduction of machine learning approaches, such as Random Forest and Neural Networks, has opened new horizons for optimizing risk modeling, especially in multinational markets with diverse economic conditions.

The research findings indicated that machine learning models, particularly Random Forest and Neural Net, generally demonstrated notable improvements over statistical methods such as logistic regression and linear discriminant analysis in predicting the risk of exchange-traded funds. This advantage, while significant in many settings, is not limited to higher accuracy; an Area Under the Curve (AUC) close to one in most markets also indicates the high discriminatory power and classification strength of these models. These observations demonstrate considerable consistency with the findings of Alfazari et al. (2025) and Safavi et al. (2024), who also demonstrated the role of intelligent algorithms and models based on big data in enhancing investment portfolio management and risk control. In fact, the success of these models in various countries with diverse economic growth rates and stability also indicates their flexibility in addressing different market structural characteristics. In other words, the high efficiency of these approaches has been observed not only in developed markets, such as Germany and France, but also in developing markets, including Iran and India, which confirms the widespread importance of using these tools.

The key point in data analysis is that classic statistical models, such as Linear Discriminant Analysis (LDA) and Logistic Regression, although performing relatively well in certain markets like France, Taiwan, and India, generally do not match the performance of machine learning models. However, differences are modest in some contexts. For example, the accuracy of the LDA model in France and Taiwan was 72.9% and 94.3% respectively, which, although considered relatively good, is still significantly different from the average accuracy of machine learning-based models. Additionally, the logistic regression model exhibited suitable but unstable performance in some countries, such that in certain countries, the model's accuracy even fell below 50%, severely limiting its reliability. These results can be attributed to the linear nature and the assumption of normal data distribution in classic models.

In contrast, financial data often involves non-normal distributions, multimodality, high outlier ratios, and complex nonlinear interactions. The use of the Shapiro-Wilk test to check for data normality also showed that in most countries, the model accuracy data did not follow a normal distribution. Therefore, a comparative approach using non-parametric tests such as the Mann-Whitney U test was completely justified and scientific. These tests further confirmed the significant superiority of machine learning models over statistical models.

Beyond mere prediction accuracy, the issue of identifying key independent variables affecting risk also held a special place in this study. Sensitivity analysis of Random Forest and Neural Net models indicated that performance indicators, such as Jensen's measure and the superior turnover ratio, have an inverse and significant relationship with fund risk, meaning that funds with better past performance tend to face lower risk in subsequent periods. This observation is consistent with market efficiency theories and asset dynamics theories and promises continued desirable performance. Additionally, fundamental factors such as fund size and management expertise also played a significant role in mitigating risk. These findings indicate that funds with larger resources and more specialized management teams are better equipped to cope with market shocks and mitigate the impact of fluctuations. This function is particularly evident in developed markets. This result parallels the findings of Kaniel et al. (2023), who proved the use of fund characteristics to predict future return and risk.

On the other hand, variables such as fund return and market return, especially in countries with highly volatile markets like China and India, had a positive impact on risk. This is explained by the fact that a large part of high returns is a product of increased risk-taking and greater market movements. In addition, rapid growth in fund value and a high percentage of retail holders were also correlated with increased risk. This is because the influx of fresh funds from non-professional investors can lead to an intensification of herd behavior and increased market volatility. These findings are consistent with studies by Safavi et al. (2024) and Melina et al. (2023), which showed that using data-driven models and advanced ML techniques alongside modern financial theories can add more dynamism to risk management.

From a practical point of view, these findings are particularly important because identifying critical variables that affect fund risk helps managers and investors optimize their prediction algorithms more effectively and also base risk management policies on more controllable variables. For example, it has been shown that fund size and management expertise should be considered as basic variables in all risk prediction models. Conversely, when the fund's value growth or the ratio of retail investors increases, stricter policies for risk control need to be formulated.

The logical chain of this research is completed by the comparative results of studies such as Soroush and Akhlaghi (2017), Niculescu et al. (2020), and Melina et al. (2023), which emphasized the more effective and accurate presence of advanced data-driven models in risk prediction, identification of key factors, and enhancement of financial market efficiency. Considering that the markets under study encompass a wide range of maturity, economic structure, and data quality, this study has provided strong evidence for the superiority of machine learning tools and algorithms in the investment space by employing state-of-the-art techniques.

It is important to acknowledge several limitations of this study. First, data quality varied significantly across countries, with more developed markets generally providing more comprehensive and reliable datasets. This heterogeneity in data quality might have influenced model performance comparisons. Second, the availability of variables was not uniform across all markets, which could introduce biases in our cross-country analyses. Third, while machine learning models demonstrated strong predictive performance, they present challenges in terms of interpretability, particularly Neural Networks, which may function as "black boxes" in certain applications. Fourth, our approach to risk binarization, while methodologically sound, necessarily simplifies the continuous nature of financial risk. Finally, sample representativeness varied across markets, with some countries having more comprehensive ETF coverage than others.

On a broader scale, the findings of the current research are also generalizable to the areas of policymaking, asset management, and the design of regulatory frameworks, because the identified superior algorithms, due to their ability to process multi-year and voluminous data and their power to extract hidden and nonlinear relationships, are supportive tools for wiser decision-making by managers and legislators. On the other hand, considering the monitoring of models' sensitivity to data quality and volume in different countries, the development of data processing systems and financial information transparency will play a crucial role in enhancing the performance of risk prediction models.

In a comprehensive interpretation, these results underscore the need to transition from traditional models based on simplistic assumptions to data-driven, intelligent, and flexible approaches. In other words, in the current era, where financial markets face structural shocks, the rapid entry of new technologies, and big data, success in predicting and controlling risk depends on employing a combination of advanced artificial intelligence algorithms, systematic sensitivity analysis, and the accurate identification of key variables.

Finally, based on the results of this research, it is recommended that fund managers and institutional investors, to improve their risk management performance in today's competitive and volatile markets, prioritize ML models and make algorithms based on Random Forest, Neural Net, and SVM the central axis of their warning and risk management systems. Additionally, continuous monitoring of identified key variables and the constant updating of models with new data will play a crucial role in enhancing prediction accuracy and investment resilience. Furthermore, for future research, the effect of combining machine learning algorithms with investor behavioural data, as well as examining the role of regulatory policies and disclosure transparency in enhancing the performance of prediction models, should be investigated more thoroughly.

In summary, the evidence presented in this research further supports the importance and necessity of using data-driven approaches and machine learning in managing the risk associated with exchange-traded funds. These tools, in addition to increasing prediction accuracy, provide the possibility of offering more intelligent and evidence-based recommendations for managers and investors, which will ultimately lead to improved market efficiency and increased investor confidence. The practical suggestions are as follows:

1. Development of intelligent risk management systems: Fund managers and investment companies can make the risk assessment and monitoring process intelligent, fast, and multidimensional by implementing systems based on machine learning models (especially Random Forest and Neural Net). These systems, with the ability to update data daily or weekly, allow for a quicker response to market fluctuations.
2. Focus on data-driven training for managers and analysts: Organizing workshops and training courses centred on the application of machine learning and big data analysis in risk management can enhance the practical capabilities of experts and managers in these tools. Such an approach will help increase data literacy and foster more informed and reasoned decision-making.
3. Improvement of data infrastructures and increased information transparency: Creating accurate, comprehensive, and up-to-date databases of fund performance and disclosing key information (such as returns, fund size, and asset composition) provides a suitable platform for enhancing the performance of risk prediction models. Stock exchange supervisory bodies and the government can pursue this.
4. Creation of a preventive warning mechanism: By utilising ML models and identifying key variables that affect risk, automatic Early Warning Systems can be designed to send necessary warnings to managers and investors if a stock or fund approaches a high-risk zone.
5. Formulation of dynamic asset allocation policies: It is recommended that fund managers, utilizing the output of machine learning models and

sensitivity analysis of variables, periodically review their asset allocation policies and update them in line with changing market structures to enhance portfolio resilience.

6. Combination of machine learning with behavioral analysis: It is suggested that future studies and systems, in addition to financial and fundamental factors, incorporate investor behavior (such as retail trading volume and herd behavior) into models to enhance risk prediction accuracy during crises further.
7. Continuous monitoring and improvement of models: Managers should regularly test the performance of their models with new data and, if accuracy decreases, review and update the algorithms or model variables to ensure the system always operates optimally.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest concerning the research, authorship and, or publication of this article.

Funding

The authors received no financial support for the research, authorship and, or publication of this article.

References

Acerbi, C., & Scandolo, G. (2021). Risk and the role of information in financial decision-making. *Risk Analysis*, 41(7), 1264-1276.

Aflatooni, A. (2013). Statistical analysis with EViews in accounting and financial management research: Termeh Publications. (In Persian)

Aggarwal, R. K., & Jorion, P. (2010). The performance of emerging hedge funds and managers. Journal of Financial Economics, 96(2), 238-256.

Akhbari, H. R.; Mohammadzadeh Salteh, Heydar; Hassanzadeh Brothers, Rasoul; Zeinali, Mehdi. (2024). Optimizing Risk-Based Stock Return Forecasting in Selected Industries of Tehran Stock Exchange (Data Envelopment Analysis Approach). Financial Research, 26(2), 331–354.

Alfzari, S., Al-Shboul, M., & Alshurideh, M. (2025). Predictive Analytics in Portfolio Management: A Fusion of AI and Investment Economics for Optimal Risk-Return Trade-Offs. International Review of Management and Marketing, 15(2), 365-380.

Alsulamy, S. (2025). Predicting construction delay risks in Saudi Arabian projects: A comparative analysis of CatBoost, XGBoost, and LGBM. Expert Systems with Applications, 268, 126268.

Bauer, M. D., & Hamilton, J. D. (2018). Robust bond risk premia. *The Review of Financial Studies*, 31(2), 399-448.

Bhagat, S., Bolton, B., & Lu, J., (2015). Size, leverage, and risk-taking of financial institutions, *Journal of Banking & Finance* (Elsevier), 59, 520–537.

Black, F., & Scholes, M. (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 27(2), 399–417.

Bollen, N. P., & Busse, J. A. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies*, 18(2), 569-597.

Bozorg Tabarabai, Aghajan Nashtaei, Gholizadeh. (2024). Financial Risk Analysis in the Cryptocurrency Market: Evidence from Predicting Value at Risk. *Investment Knowledge*, 50 (1), 501–519.

Campbell, J. Y., Polk, C., & Vuolteenaho, T. (2010). Growth or glamour? Fundamentals and systematic risk in stock returns. *Review of Financial Studies*, 23(1), 305-344.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review*, 94(5), 1276-1302.

Chen, J., Ma, Y., & Zhang, H. (2020). Machine Learning for Financial Market Prediction: A Review. *Journal of Computational and Applied Mathematics*, 364, 1-15.

Chen, Y., Mamon, R., Spagnolo, F., & Spagnolo, N. (2025). Stock market returns and climate risk in the US. *Journal of Multinational Financial Management*, 77, 100887.

Choi, I., & Kim, W. C. (2024). Enhancing Exchange-Traded Fund Price Predictions: Insights from Information-Theoretic Networks and Node Embeddings. *Entropy*, 26(1), 70.

Danquah, R., & Yu, B. (2023). Selection ability and market timing skills of mutual fund and unit trust managers in a developing economy: evidence from Ghana. *Business Analyst Journal*, 44(1), 1-14.

Deb, S., & Singh, R. (2018). Dynamics of risk perception towards mutual fund investment decisions. *Interdisciplinary Journal of Management Studies (Formerly known as Iranian Journal of Management Studies)*, 11(2), 407–424.

Dimitrios M., et al. (2012). Machine Learning Techniques for Customer Retention: A Comparative Analysis. *Journal of Artificial Intelligence Research*, 45(3), 123-

145.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*, 31(9), 2632-2647.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2009). Modern Portfolio Theory and Investment Analysis (8th ed.). Wiley.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1): 3-56.

Frank, R., & Brown, K. C. (2012). Investment analysis and portfolio management.

Graham, B., & Dodd, D. L. (2008). Security Analysis: Sixth Edition. McGraw-Hill Education.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor types: a study of Finland's unique data set. *Journal of financial economics,* 55(1), 43–67.

Hasan, S. M., Tawfiq, T. T., Hasan, M. M., & Islam, K. A. (2024). Risk in the Shadows: Macroeconomic Shifts and Their Effects on Bangladeshi Mutual Funds. Innovations, 21(4), 371-384.

He, K., & Li, X. (2022). A Survey of Machine Learning Techniques for Financial Risk Prediction. *Journal of Financial Markets,* 65, 100786.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. *Review of Financial Studies*, 24(8), 2575-2616.

Hutchinson, M., Seamer, M., & Chapple, L. E. (2015). Institutional investors, risk/performance and corporate governance. *International Journal of Accounting*, 50(1), 31-52.

Inderst, R., & Müller, H. M. (2004). The effect of capital market characteristics on the value of startup firms. *Journal of Financial Economics*, 72(2), 319–356. https://doi.org/10.1016/j.jfineco.2003.06.001

Iselin, M, Liao, S., Zhang, H. (2021). Common Mutual Fund Ownership and Systemic Risk. *Contemporary Accounting Research*, 38(3), 2157-2191.

Islam, K. U., Bhat, S. A., Lone, U. M., Darzi, M. A., & Malik, I. A. (2024). Financial risk propensity and investment decisions: An empirical analysis using behavioural biases. *IIMB Management Review,* 36(3), 213-229.

Islam, K. U., Bhat, S. A., Lone, U. M., Darzi, M. A., & Malik, I. A. (2024). Financial risk propensity and investment decisions: An empirical analysis using behavioural biases. *IIMB Management Review*, 36(3), 213-229.

Jaffri, A., Shirvani, A., Jha, A., Rachev, S. T., & Fabozzi, F. J. (2025). Optimizing Portfolios with Pakistan-Exposed Exchange-Traded Funds: Risk and Performance

Insight. *Journal of Risk and Financial Management*, 18(3), 158.

Jayeola, D., Ismail, Z., & Sufahani, S. F. (2017). Effects of diversification of assets in optimizing risk of portfolio. *Malaysian Journal of Fundamental and Applied Sciences*, 13(4). https://doi.org/10.11113/mjfas.v0n0.567

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23(2), 389-416.

Kacperczyk, M., Sialm, C., & Zheng, L. (2005). On the industry concentration of actively managed equity mutual funds. *Journal of Finance,* 60(4), 1983-2011.

Kaniel, R., Lin, Z., Pelger, M., & Van Nieuwerburgh, S. (2023). Machine-learning the skill of mutual fund managers. *Journal of Financial Economics,* 150(1), 94-138.

Keshavarz Haddad, Gh., Ebrahimnejad, A., & Grossi, M. (1401). The performance of real and legal investors in predicting the return of mutual investment funds. *Biquarterly Scientific Journal of Economic Studies and Policies,* 9(1), 94-115. (in Persian).

Kwon, Sungjoung, Lowry, Michelle, Yiming, Q. (2020). Mutual fund investments in private firms. *Journal of Financial Economics,* 136(2), 407–443. https://doi.org/10.1016/j.jfineco.2019.10.003

Lai, M. M., & Lau, S. H. (2010). Evaluating mutual fund performance in an emerging Asian economy: The Malaysian experience. *Journal of Asian Economics*, 21(4), 378–390.

Li, D., & Lu, S. (2025). Portfolio climate risk and fund flow performance. *Finance Research Letters*, 72, 106522.

Malkiel, B. G. (2023). A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing. W.W. Norton & Company.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance,* 7(1), 77-91.

Markowitz, H. M., & Van Dijk, E. (2008). Risk-return analysis. *In Handbook of asset and liability management*, 1(13), 139-197, *Elsevier*. https://doi.org/10.1016/B978-044453248-0.50010-1

Markowitz, H. M., & Van Dijk, E. (2008). Risk-return analysis. *In Handbook of asset and liability management*, 1(13), 139-197 *Elsevier*. https://doi.org/10.1016/B978-044453248-0.50010-1

Mashrur, A., Luo, W., Zaidi, N. A., & Robles-Kelly, A. (2020). Machine learning for financial risk management: a survey. Ieee Access, 8, 203203–203223.

Melina, Sukono, Napitupulu, H., & Mohamed, N. (2023). A conceptual model of investment-risk prediction in the stock market using extreme value theory with machine learning: a semisystematic literature review. *Risks*, 11(3), 2-24.

Menkhoff, L., Schmidt, U., & Brozynski, T. (2006). The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European economic review,* 50(7), 1753-1766.

Nicolescu, L., Tudorache, F. G., & Androniceanu, A. (2020). Performance risk analysis on mutual funds versus stock exchanges in young financial markets. *Journal of International Studies*, 13(1), 2071-8330.

Oladimeji, M. S. & Udosen, I. (2019). The Effect of Diversification Strategy on Organizational Performance. *Journal of Competitiveness*, 11(4), 120–131. https://doi.org/10.7441/joc.2019.04.08

Otten, R., & Bams, D. (2002). European mutual fund performance. *European financial management,* 8(1), 75–101.

Peterson, J. D., Petranico, P. A., Riepe, M. W., & Xu, F. (2001). Explaining the performance of domestic equity mutual funds. *Journal of Investing,* 10(3), 81-91.

Piovezan, R. P., & Junior, P. P. D. A. (2022). Machine learning method for return direction forecasting of Exchange Traded Funds using classification and regression models. arXiv preprint arXiv, 2205, 12746.

Sadeghi Goghari, S., Souri, A., Abbasinejad, H. and Mehrara, M. (2020). Portfolio Diversification and Net Selectivity Performance of Mutual Funds in Iran by Using Fama Decomposition Model. *Iranian Economic Review*, *24*(2), 471-487. doi: 10.22059/ier.2020.76014

Sadeghi Goghari, S., Souri, A., Abbasinejad, H. and Mehrara, M. (2020). Portfolio Diversification and Net Selectivity Performance of Mutual Funds in Iran by Using Fama Decomposition Model. *Iranian Economic Review*, *24*(2), 471-487. doi: 10.22059/ier.2020.76014

Safavi Iranji, M., Zanjirdar, M., Safa, M. and Jahangirnia, H. (2024). Asset Allocation Using Nested Clustered Optimization Algorithm: A Novel Approach to Risk Management in Portfolio. *Journal of Mathematics and Modeling in Finance*, 4(2), 137-157.

Simutin, M. (2010). Excess cash and stock returns. *Financial Management*, 39(3), 1197-1222.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance*, 53(5), 1589-1622.

Soroush Yar, Akhlaqi. (2017). Comparative Evaluation of the Effectiveness of Data Mining Techniques in Forecasting Risk and Stock Returns of Companies Listed in Tehran Stock Exchange. *Journal of Financial Accounting Research,* (1)9, 57-76.

Tavakoli, Saman; Ashtab, Ali. (2023). Comparing the efficiency of machine learning models and statistical models in predicting financial risk. *Financial Management*

Strategy, (1)11, 53-76.

Teo, M. (2011). The liquidity risk of liquid hedge funds. *Journal of Financial Economics*, 100(1), 24–44. https://doi.org/10.1016/j.jfineco.2010.11.003

Vyšniauskas, P., & Rutkauskas, A. V. (2014). Performance evaluation of investment (mutual) funds. *Business: Theory and Practice*, 15(4), 398-407.

Xu, B. (2007). Life cycle effect on the value relevance of common risk factors. *Review of Accounting and Finance*, 6(2), 162–175.

Zeng, B. (2023). Venture Capital Research: Investor Preferences and Success Factors for Startups. Open *Journal of Business and Management*, 11, 2743-2762. doi: 10.4236/ojbm.2023.116149.

Bibliographic information of this paper for citing:

Esmaily, Nashmil; Piri, Parviz; Ashtab, Ali; Heydari, Mehdi & Zavarirezaei, Akbar (2025). Risk prediction of investment funds in member countries of the Federation of European and Asian Stock Exchanges - Machine Learning Approaches. *Iranian Journal of Finance*, 9(4), 140-188.