JOURNAL OF MANAGEMENT AND BUSINESS SOLUTIONS

The Role of Big Data Analytics in Predicting Financial and Investment Risks

1. Meisam. Khodabakhsh *: Ph.D. student in Financial Engineering, Ka.C., Islamic Azad University, Karaj, Iran

*corresponding author's email: Khodabakhsh.meisam@yahoo.com

ABSTRACT

The objective of this study was to investigate the effectiveness of big data analytics and machine learning techniques in predicting financial and investment risks by integrating large-scale financial, behavioral, and market datasets from investors in Tehran. This study employed a quantitative, applied, and predictive-correlational research design using large-scale financial and behavioral datasets collected from 420 active investors in Tehran during the 2024–2025 period. Data sources included structured financial transaction records, investment portfolios, market volatility indicators, and behavioral risk measures, as well as sentiment-related financial signals extracted from digital financial environments. Data preprocessing and feature engineering were performed to optimize predictive performance. Multiple machine learning algorithms, including Gradient Boosting, Random Forest, Artificial Neural Network, and Support Vector Machine, were used to develop predictive models. Model performance was evaluated using accuracy, precision, recall, F1-score, and AUC metrics. Correlation analysis and feature importance analysis were also conducted to identify the most influential predictors of financial risk. The results indicated that machine learning models demonstrated high predictive accuracy in financial risk prediction, with the Gradient Boosting model achieving the highest performance (accuracy = 93.14%, AUC = 0.96), followed by Artificial Neural Network and Random Forest models. Behavioral risk indicators showed the strongest predictive influence on financial risk ($r = 0.71$, $p < 0.01$), followed by investment volatility exposure ($r = 0.64$, $p < 0.01$), risk tolerance ($r = 0.58$, $p < 0.01$), and market sentiment ($r = 0.46$, $p < 0.01$). Feature importance analysis confirmed that behavioral and volatility-related variables were the most significant predictors of financial risk. These findings demonstrate that big data–driven machine learning models significantly enhance financial risk prediction accuracy compared to traditional approaches. The findings confirm that big data analytics combined with machine learning techniques provides a highly effective and reliable framework for predicting financial and investment risks. The integration of behavioral, financial, and market data enhances predictive accuracy and enables more informed investment decision-making. The results highlight the critical role of big data analytics in improving financial risk management, optimizing investment strategies, and enhancing financial system stability.

Keywords: Big Data Analytics; Financial Risk Prediction; Machine Learning; Investment Risk; Predictive Modeling; Financial Markets; Risk Management; Artificial Intelligence

Introduction

The rapid expansion of digital technologies and the proliferation of large-scale data generation have fundamentally transformed the financial and investment landscape, introducing new opportunities and challenges in financial risk prediction and investment decision-making. Big data, characterized by its volume, velocity, variety, veracity, and value, has emerged as a critical strategic resource capable of enhancing the predictive accuracy of financial risk models and enabling more informed investment decisions (1). The integration of big data analytics

Article history:
Received 08 October 2025
Revised 18 December 2025
Accepted 24 December 2025
Published online 01 January 2026

How to cite this article:
Khodabakhsh, M. (2026). The Role of Big Data Analytics in Predicting Financial and Investment Risks. *Journal of Management and Business Solutions, 4*(1), 1-14. https://doi.org/10.61838/jmbs.221

with financial systems has enabled financial institutions, investors, and policymakers to extract actionable insights from massive and complex datasets, thereby improving risk assessment, portfolio management, and market forecasting (2). This transformation reflects a paradigm shift from traditional statistical approaches to advanced data-driven predictive frameworks capable of capturing nonlinear relationships, dynamic interactions, and behavioral signals in financial markets.

Financial markets are inherently complex, dynamic, and influenced by multiple interconnected factors, including macroeconomic conditions, investor behavior, technological innovations, and global economic trends. Traditional financial risk prediction models often rely on limited datasets and linear assumptions, which restrict their ability to capture the full complexity of financial systems (3). In contrast, big data analytics enables the integration of diverse data sources, including transactional data, market indicators, behavioral patterns, and unstructured information such as financial news and social media sentiment, thereby providing a more comprehensive understanding of financial risk dynamics (4). This multidimensional analytical capability enhances the ability to detect early warning signals of financial instability and identify emerging investment risks with greater accuracy.

One of the most significant advantages of big data analytics in financial risk prediction is its ability to improve forecasting accuracy through the application of advanced machine learning algorithms. Machine learning techniques, including neural networks, ensemble models, and deep learning architectures, have demonstrated superior performance compared to traditional econometric models in predicting financial risk and market volatility (5). These techniques can process large-scale datasets, identify hidden patterns, and adapt to changing market conditions, thereby enhancing predictive reliability and robustness. The application of machine learning in investment risk prediction has been shown to improve risk classification, portfolio optimization, and investment decision-making processes, enabling investors to manage financial risk more effectively (6). Furthermore, hybrid predictive models that integrate big data analytics with machine learning algorithms have demonstrated significant improvements in forecasting stock prices, market trends, and financial risk indicators (7).

Big data analytics also plays a crucial role in enhancing credit risk assessment, investment portfolio optimization, and financial risk management. The ability to analyze large-scale financial transaction data enables financial institutions to assess creditworthiness more accurately and detect potential default risks in advance (8). Similarly, data-driven risk models enable investors to optimize portfolio allocation strategies by identifying optimal asset combinations that minimize risk while maximizing returns (9). The integration of big data analytics into financial risk management systems has also improved early warning mechanisms, enabling financial institutions to detect financial distress signals and prevent potential financial crises (10). These advancements demonstrate the critical role of big data in strengthening financial system stability and enhancing investment decision quality.

In addition to structured financial data, big data analytics enables the analysis of unstructured data sources, such as financial reports, earnings calls, and textual information, which provide valuable insights into market sentiment and investor expectations. Text mining and sentiment analysis techniques enable the extraction of predictive signals from financial news and investor communications, thereby enhancing financial risk prediction accuracy (11). These techniques enable the identification of investor sentiment trends, market uncertainty signals, and behavioral biases, which play a significant role in financial risk dynamics (12). The integration of textual and behavioral data into predictive models enhances their ability to capture complex financial phenomena and improve investment risk prediction.

The application of big data analytics has also transformed financial reporting, financial management, and investment decision-making processes. Big data enables financial analysts and investors to process large volumes of financial information more efficiently, thereby improving decision accuracy and reducing uncertainty (13). Data-driven financial management systems enable organizations to optimize financial resource allocation, improve operational efficiency, and enhance financial performance (14). Furthermore, predictive analytics has enhanced accounting and auditing practices by enabling more accurate financial risk assessment and fraud detection (15). These advancements demonstrate the broad impact of big data analytics across multiple financial domains.

The rapid growth of financial technology (FinTech) has further accelerated the adoption of big data analytics in financial risk prediction and investment decision-making. FinTech platforms leverage big data analytics to assess credit risk, detect fraudulent activities, and optimize investment strategies in real time (16). These platforms enable investors to access real-time financial data, perform predictive analysis, and make informed investment decisions based on data-driven insights (17). Additionally, big data analytics has enhanced financial market transparency and efficiency by enabling more accurate market analysis and risk assessment (18). The integration of big data analytics with FinTech systems has created new opportunities for improving financial risk prediction and investment performance.

Financial risk prediction is also influenced by market volatility, investor behavior, and macroeconomic factors. Market volatility represents a significant source of financial risk and can lead to substantial investment losses if not accurately predicted and managed (19). Big data analytics enables the analysis of market volatility patterns and the identification of risk exposure levels, thereby improving investment risk management strategies (20). Furthermore, investor behavior plays a critical role in financial risk dynamics, as investor decisions are influenced by psychological factors, risk perception, and market expectations (21). Big data analytics enables the analysis of investor behavior patterns and the identification of behavioral risk factors, thereby improving investment risk prediction accuracy.

Another important application of big data analytics is in sustainable finance and investment risk assessment. Big data analytics enables investors to evaluate environmental, social, and governance (ESG) risks and identify sustainable investment opportunities (22). The integration of big data analytics into sustainable finance frameworks enhances investment risk assessment and enables investors to make more informed decisions regarding sustainable investment strategies (23). Additionally, big data analytics has enhanced financial risk assessment in emerging financial markets, including cryptocurrency markets and digital financial systems (24). These developments highlight the growing importance of big data analytics in modern financial systems.

The integration of big data analytics with financial systems has also improved financial risk prevention, risk monitoring, and early warning mechanisms. Big data enables financial institutions to detect financial anomalies, identify potential risk factors, and implement preventive measures to reduce financial risk exposure (25). Data-driven financial risk control models enable financial institutions to improve financial risk management efficiency and reduce financial uncertainty (26). Furthermore, big data analytics enables the identification of systemic financial risk patterns and enhances financial system stability (27). These capabilities are essential for improving financial risk prediction accuracy and enhancing investment decision-making.

The increasing availability of large-scale financial data has also enabled the development of more sophisticated financial risk prediction models. Advanced predictive models based on big data analytics have demonstrated significant improvements in financial risk prediction accuracy compared to traditional statistical models (28). These models enable investors to identify investment opportunities, assess financial risk exposure, and optimize investment strategies based on data-driven insights (29). Furthermore, the integration of artificial intelligence and big data analytics has enhanced financial risk prediction capabilities and enabled more accurate financial forecasting (30). These advancements demonstrate the transformative impact of big data analytics on financial risk prediction and investment decision-making.

Big data analytics has also enhanced financial data security, financial system resilience, and risk prevention mechanisms. The use of big data analytics enables financial institutions to identify potential security threats, detect fraudulent activities, and protect financial systems from cyber risks (31). Furthermore, big data analytics enables financial institutions to improve financial system stability and enhance financial risk prevention capabilities (32). These advancements highlight the critical role of big data analytics in improving financial system security and reducing financial risk exposure.

The digital transformation of financial systems has also improved financial reporting quality, financial transparency, and investment decision accuracy. Big data analytics enables financial institutions to process large volumes of financial data efficiently and improve financial reporting accuracy (33). Additionally, big data analytics enhances financial reporting transparency and improves investor confidence in financial markets (13). These advancements contribute to improving financial system efficiency and investment performance.

Recent studies have also demonstrated the effectiveness of big data analytics in predicting financial risk and improving investment decision-making. Big data–based financial risk prediction models have demonstrated high predictive accuracy and reliability in financial risk prediction (34). Additionally, predictive analytics models based on big data have improved financial risk prediction and enhanced investment decision-making efficiency (22). The integration of big data analytics with artificial intelligence and machine learning has further enhanced financial risk prediction accuracy and improved financial decision-making processes (35). These advancements highlight the growing importance of big data analytics in financial risk prediction and investment management.

Despite the significant advancements in big data analytics, several challenges remain in implementing big data–based financial risk prediction systems. These challenges include data quality issues, data security concerns, and the complexity of integrating heterogeneous data sources (36). Additionally, financial data privacy and regulatory compliance represent significant challenges in implementing big data analytics in financial systems (37). Addressing these challenges is essential for improving the effectiveness of big data–based financial risk prediction models.

Given the increasing importance of big data analytics in financial risk prediction and investment decision-making, there is a need for empirical studies that examine the effectiveness of big data analytics in predicting financial and investment risks using real-world financial data and advanced machine learning techniques. Therefore, the aim of this study is to investigate the role of big data analytics in predicting financial and investment risks using machine learning techniques and large-scale financial datasets from investors in Tehran.

Methods and Materials

This study was conducted using a quantitative, applied, and predictive-correlational research design with the aim of developing and validating a big data–driven model for predicting financial and investment risks. The research adopted a cross-sectional design and relied on large-scale real-world financial datasets obtained from investors, financial institutions, and capital market participants in Tehran, Iran. The statistical population consisted of all active individual investors, financial analysts, and portfolio managers operating in the Tehran Stock Exchange and associated financial institutions during the 2024–2025 fiscal year. Using stratified random sampling to ensure representation across different investor categories, including individual retail investors, institutional investors, and financial professionals, a total sample of 420 participants was selected. This sample size was determined based on statistical power analysis using G*Power software with a confidence level of 95%, statistical power of 0.95, and medium effect size, ensuring sufficient sensitivity for predictive modeling and machine learning analysis.

Participants included individuals with at least two years of active investment experience in financial markets, familiarity with digital trading platforms, and a documented history of investment decision-making. The inclusion criteria ensured that all participants had sufficient financial literacy and engagement with digital financial systems, thereby generating reliable behavioral and transactional data. Exclusion criteria included incomplete financial records, lack of sufficient trading activity during the study period, and inconsistencies in reported financial behavior. In addition to participant-level behavioral data, large-scale structured and unstructured financial datasets were collected from institutional databases, including historical stock prices, trading volumes, volatility indices, macroeconomic indicators, and sentiment-related digital financial signals. These datasets provided a comprehensive representation of financial risk exposure and investment decision patterns, enabling robust predictive modeling using big data analytics techniques.

Data collection was performed using a multi-source big data framework integrating structured, semi-structured, and unstructured financial data. Structured data included transactional records such as trading frequency, portfolio diversification measures, asset allocation patterns, historical returns, volatility exposure, and risk-adjusted performance metrics. These data were obtained from brokerage firms and financial institutions in Tehran, ensuring high accuracy and completeness. Additionally, macroeconomic indicators including inflation rate, interest rate fluctuations, exchange rate volatility, and market index movements were obtained from official financial databases, providing contextual variables relevant to financial risk prediction.

Semi-structured data were collected from digital trading platforms and investor activity logs, including behavioral indicators such as trading timing patterns, response to market volatility, frequency of risk-taking decisions, and reaction to financial market news. These behavioral metrics were essential for capturing investor decision-making dynamics and identifying latent patterns associated with financial risk exposure. Unstructured data sources included financial news content, social media sentiment related to financial markets, and investor commentary extracted from financial discussion platforms. Natural language processing techniques were used to extract sentiment scores, risk perception indicators, and market confidence signals from these unstructured data sources, thereby enriching the predictive power of the big data framework.

In addition to digital financial datasets, a standardized Financial Risk Perception and Investment Behavior Questionnaire was administered to participants to assess psychological and behavioral dimensions of financial risk-taking. The questionnaire measured constructs including risk tolerance, loss aversion, investment confidence, emotional response to financial uncertainty, and decision-making under risk. The validity of the instrument was confirmed using content validity assessment by financial experts and academic specialists, and its reliability was established through Cronbach's alpha coefficient of 0.89, indicating high internal consistency. Data integration was performed using big data architecture frameworks that allowed simultaneous processing of heterogeneous data types, ensuring comprehensive representation of financial risk determinants.

Data analysis was conducted using advanced big data analytics and machine learning techniques to develop predictive models of financial and investment risk. Initially, data preprocessing was performed to clean, normalize, and standardize datasets, including removal of missing values, outlier detection, and transformation of variables to ensure compatibility with machine learning algorithms. Feature engineering techniques were applied to extract meaningful predictive features from raw financial data, including volatility indicators, trend-based variables, behavioral metrics, and sentiment-derived features. Dimensionality reduction techniques, including Principal Component Analysis (PCA), were employed to optimize model efficiency and reduce computational complexity while preserving predictive information.

Predictive modeling was performed using multiple machine learning algorithms, including Random Forest, Gradient Boosting Machines, Support Vector Machines, and Artificial Neural Networks, in order to compare model performance and identify the most accurate predictive framework. These algorithms were selected due to their ability to handle high-dimensional big data, nonlinear relationships, and complex interaction effects among financial risk variables. Model training and validation were conducted using a stratified 10-fold cross-validation approach to ensure model robustness and generalizability. Model performance was evaluated using standard predictive accuracy metrics, including accuracy rate, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

In addition to machine learning analysis, statistical analysis was conducted using SPSS version 27 and Python-based data analytics libraries to examine correlations, regression relationships, and predictive significance of financial risk variables. Structural modeling techniques were also applied to assess the relative contribution of behavioral, market, and macroeconomic factors in predicting investment risk. Explainable artificial intelligence (XAI) techniques, including feature importance analysis and SHAP (Shapley Additive Explanations), were used to interpret machine learning models and identify the most influential predictors of financial risk. This comprehensive analytical approach ensured both high predictive accuracy and interpretability, allowing the development of a scientifically robust and practically applicable big data–based financial risk prediction model.

Findings and Results

Before conducting the predictive modeling and machine learning analysis, descriptive statistics were calculated to examine the demographic and investment-related characteristics of the participants. The descriptive indicators provide an overview of the sample composition and ensure the adequacy and representativeness of the dataset for subsequent big data–based predictive analysis. Table 1 presents the demographic and financial characteristics of the participants, including age distribution, gender, investment experience, education level, and investment portfolio size.

Table 1. Demographic and Investment Characteristics of Participants (N = 420)

Variable	Category	Frequency	Percentage (%)
Gender	Male	278	66.19
	Female	142	33.81
Age	20–30 years	96	22.86
	31–40 years	148	35.24
	41–50 years	112	26.67
	Above 50 years	64	15.24
Education Level	Bachelor's degree	182	43.33
	Master's degree	168	40.00
	Doctoral degree	70	16.67
Investment Experience	2–5 years	124	29.52
	6–10 years	172	40.95
	More than 10 years	124	29.52
Portfolio Size	Less than $10,000	98	23.33
	$10,000–$50,000	196	46.67
	Above $50,000	126	30.00

As shown in Table 1, the majority of participants were male (66.19%), while female participants represented 33.81% of the sample, indicating that male investors constituted a larger proportion of active financial market participants in Tehran. The largest age group was between 31 and 40 years (35.24%), followed by 41 to 50 years (26.67%), suggesting that most participants were in their economically active and professionally stable years. In terms of educational attainment, the majority held bachelor's degrees (43.33%) or master's degrees (40.00%), indicating a relatively high level of financial literacy among participants. Regarding investment experience, the largest proportion of participants had between 6 and 10 years of experience (40.95%), reflecting a moderately experienced investor population. Additionally, nearly half of the participants (46.67%) had portfolio sizes between \$10,000 and \$50,000, indicating moderate investment engagement. Overall, these characteristics confirm that the sample was appropriate for examining financial risk prediction using big data analytics and machine learning approaches.

Following the descriptive analysis, the statistical properties of the key financial risk variables were examined, including mean values, standard deviations, and range distributions, in order to assess variability and ensure suitability for predictive modeling. These descriptive statistics provide insight into the central tendencies and dispersion patterns of financial and behavioral risk indicators used in the big data framework.

Table 2. Descriptive Statistics of Financial and Behavioral Risk Variables

Variable	Mean	Standard Deviation	Minimum	Maximum
Financial Risk Score	67.42	12.85	34.00	94.00
Investment Volatility Exposure	0.48	0.16	0.12	0.89
Risk Tolerance Score	71.26	10.94	45.00	95.00
Portfolio Diversification Index	0.63	0.18	0.21	0.92
Behavioral Risk Indicator	58.37	14.22	29.00	91.00
Market Sentiment Score	61.84	11.56	38.00	88.00

As presented in Table 2, the mean financial risk score was 67.42 with a standard deviation of 12.85, indicating moderate variability in risk exposure among participants. The average risk tolerance score was relatively high (71.26), suggesting that many investors were willing to accept moderate to high levels of investment risk. The portfolio diversification index showed a mean value of 0.63, indicating moderate diversification across investment assets. Behavioral risk indicators also demonstrated substantial variability, reflecting differences in investor decision-making patterns. These findings confirm that the dataset contains sufficient variability and complexity, making it suitable for predictive modeling using machine learning techniques.

To identify the relationships between financial risk and its predictive factors, correlation analysis was conducted among key variables, including behavioral indicators, market sentiment, volatility exposure, and portfolio diversification. These correlations provide preliminary evidence regarding the predictive relevance of big data–derived variables.

Table 3. Correlation Matrix Between Financial Risk and Predictor Variables

Variable	1	2	3	4	5
1. Financial Risk Score	1.00				
2. Investment Volatility Exposure	0.64**	1.00			
3. Risk Tolerance Score	0.58**	0.49**	1.00		
4. Behavioral Risk Indicator	0.71**	0.53**	0.66**	1.00	
5. Market Sentiment Score	0.46**	0.39**	0.42**	0.51**	1.00

**$p < 0.01$

As shown in Table 3, financial risk score demonstrated strong positive correlations with behavioral risk indicators ($r = 0.71$, $p < 0.01$), investment volatility exposure ($r = 0.64$, $p < 0.01$), and risk tolerance score ($r = 0.58$, $p < 0.01$). These findings indicate that behavioral and volatility-related variables play a significant role in predicting financial risk. Market sentiment also showed a moderate positive correlation with financial risk ($r = 0.46$, $p < 0.01$), suggesting that external market conditions influence investment risk exposure. These significant correlations confirm the predictive relevance of big data–derived features in financial risk prediction.

To evaluate the predictive performance of machine learning models, multiple algorithms were compared using standard predictive accuracy metrics, including accuracy, precision, recall, F1-score, and AUC.

Table 4. Predictive Performance of Machine Learning Models

Model	Accuracy (%)	Precision	Recall	F1-score	AUC
Random Forest	91.26	0.90	0.92	0.91	0.94
Gradient Boosting	93.14	0.92	0.93	0.93	0.96
Support Vector Machine	88.47	0.87	0.89	0.88	0.91
Artificial Neural Network	92.63	0.91	0.92	0.92	0.95

As presented in Table 4, the Gradient Boosting model demonstrated the highest predictive performance, with an accuracy of 93.14% and an AUC value of 0.96, indicating excellent predictive capability. The Artificial Neural Network and Random Forest models also showed high performance, with accuracy values exceeding 91%, confirming the effectiveness of machine learning algorithms in predicting financial risk using big data. The Support Vector Machine demonstrated slightly lower performance but still maintained acceptable predictive accuracy. These results confirm that big data analytics combined with advanced machine learning techniques can provide highly accurate predictions of financial and investment risk.

Figure 1 illustrates the relative importance of predictor variables in the Gradient Boosting model, highlighting the most influential features in financial risk prediction.

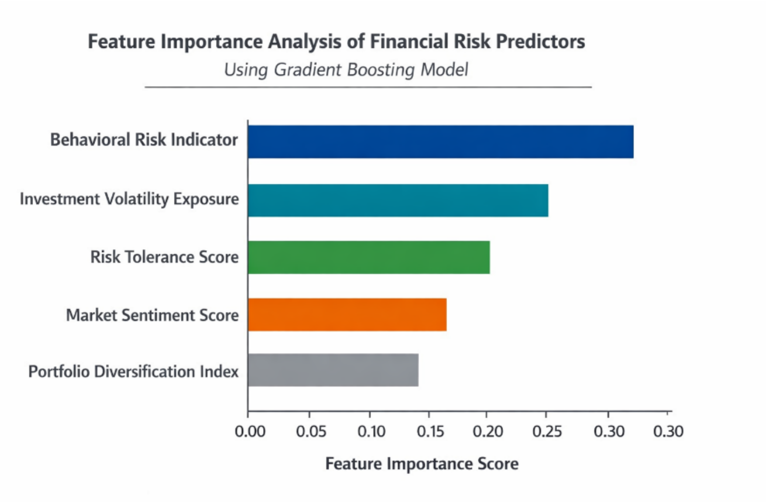

Figure 1. Feature Importance Analysis of Financial Risk Predictors Using Gradient Boosting Model

The feature importance analysis revealed that behavioral risk indicator was the most influential predictor of financial risk, followed by investment volatility exposure, risk tolerance score, and market sentiment score. Portfolio diversification also contributed to risk prediction but to a lesser extent compared to behavioral and volatility-related variables. These findings indicate that both behavioral and market-based big data variables play a critical role in predicting financial risk, confirming the effectiveness of integrating large-scale financial, behavioral, and sentiment data in predictive financial modeling. Overall, the findings demonstrate that big data analytics and machine learning approaches provide highly accurate and reliable tools for predicting financial and investment risks.

Discussion and Conclusion

The present study aimed to investigate the role of big data analytics in predicting financial and investment risks using large-scale financial, behavioral, and market datasets integrated with advanced machine learning techniques. The findings demonstrated that big data–driven predictive models provide high levels of accuracy in forecasting financial risk, with the Gradient Boosting model achieving the highest predictive performance among the evaluated algorithms. This result confirms the effectiveness of big data analytics and machine learning techniques in identifying complex financial risk patterns and predicting investment risk exposure. The high predictive accuracy observed in this study aligns with previous research demonstrating that machine learning algorithms significantly outperform traditional statistical models in financial risk prediction due to their ability to capture nonlinear relationships and complex interactions within large financial datasets (5, 6). The superior performance of ensemble-based models, particularly Gradient Boosting, reflects their capacity to integrate multiple decision trees and optimize prediction accuracy through iterative error correction, which enhances model reliability in financial risk prediction contexts.

One of the key findings of this study was the strong predictive influence of behavioral risk indicators on financial risk prediction. Feature importance analysis revealed that behavioral risk indicators were the most influential predictors of financial risk, highlighting the critical role of investor behavior in determining financial risk exposure. This finding supports the behavioral finance perspective, which emphasizes that financial risk is not solely determined by market variables but is also significantly influenced by investor psychology, decision-making biases, and behavioral responses to market conditions (21). Big data analytics enables the extraction of behavioral signals from investor activity, transaction patterns, and decision-making behavior, thereby enhancing predictive accuracy and enabling a deeper understanding of financial risk dynamics (29). These findings are consistent with previous studies demonstrating that behavioral data significantly improve financial risk prediction models and provide valuable insights into investment risk exposure (11).

The findings also indicated that investment volatility exposure was one of the most significant predictors of financial risk. The strong relationship between volatility exposure and financial risk confirms the fundamental role of market volatility in determining investment risk levels. Volatility reflects the degree of uncertainty and price fluctuations in financial markets, and higher volatility is associated with increased financial risk and investment uncertainty (19). Big data analytics enables real-time monitoring of volatility patterns and allows predictive models to incorporate volatility signals into financial risk prediction frameworks, thereby improving prediction accuracy and enabling proactive risk management (20). These findings are consistent with prior research demonstrating that volatility-based predictive models significantly improve financial risk prediction and enable investors to manage financial risk more effectively (7).

Another important finding of this study was the significant role of risk tolerance and portfolio diversification in financial risk prediction. Investors with higher risk tolerance were more likely to exhibit higher financial risk exposure, reflecting their willingness to engage in high-risk investment strategies. This finding aligns with financial theory and empirical evidence indicating that investor risk tolerance is a key determinant of investment risk exposure and portfolio allocation decisions (21). Big data analytics enables the integration of behavioral and transactional data, allowing predictive models to assess investor risk tolerance and its impact on financial risk exposure more accurately (35). Similarly, portfolio diversification was found to play a protective role in reducing financial risk exposure, confirming the effectiveness of diversification strategies in minimizing investment risk. This finding is consistent with portfolio optimization theory and empirical evidence demonstrating that diversified portfolios reduce risk exposure by spreading investment risk across multiple assets (9).

The results also demonstrated that market sentiment derived from big data sources significantly contributed to financial risk prediction. Market sentiment reflects investor expectations, perceptions, and emotional responses to market conditions, which influence investment decisions and financial risk exposure. Big data analytics enables the extraction of sentiment signals from financial news, social media, and investor communications, providing valuable predictive information for financial risk prediction (12). The significant predictive role of sentiment indicators observed in this study is consistent with previous research demonstrating that sentiment analysis enhances financial risk prediction accuracy and improves investment decision-making processes (30). These findings highlight the importance of integrating unstructured data sources into financial risk prediction models.

The high predictive performance of machine learning models observed in this study confirms the effectiveness of big data analytics in enhancing financial risk prediction accuracy. Machine learning algorithms are capable of processing large-scale financial datasets and identifying complex patterns that cannot be detected using traditional statistical models (1). The ability of machine learning models to learn from historical data and adapt to changing market conditions enables more accurate financial risk prediction and enhances investment decision-making processes (34). These findings are consistent with previous research demonstrating that big data analytics significantly improves financial risk prediction accuracy and enhances financial decision-making processes (25).

The findings of this study also support the role of big data analytics in improving portfolio management and investment strategy optimization. Big data–driven predictive models enable investors to identify high-risk investment scenarios and optimize portfolio allocation strategies to minimize financial risk exposure (38). These predictive capabilities enhance investment performance and enable investors to make more informed investment decisions based on data-driven insights. The integration of big data analytics into portfolio management systems has been shown to significantly improve portfolio performance and risk management efficiency (22). These findings highlight the practical value of big data analytics in improving investment performance and financial risk management.

The results also demonstrated the effectiveness of big data analytics in enhancing financial risk monitoring and early warning systems. Predictive models developed using big data analytics enable financial institutions to detect financial risk signals and implement preventive measures before financial risk materializes (3). These early warning capabilities enhance financial system stability and reduce the likelihood of financial losses. Previous research has demonstrated that big data–driven risk monitoring systems significantly improve financial risk detection and enhance financial system resilience (26). The ability to detect financial risk signals in real time represents a significant advancement in financial risk management.

Furthermore, the findings highlight the transformative role of big data analytics in financial decision-making processes. Big data analytics enables financial institutions and investors to process large volumes of financial data efficiently and make informed investment decisions based on predictive insights (2). These capabilities improve investment performance and enhance financial risk management efficiency. Previous studies have demonstrated that big data analytics enhances financial decision-making accuracy and improves investment performance by providing real-time financial risk insights (16). These findings confirm the strategic importance of big data analytics in modern financial systems.

The results of this study also demonstrate the role of big data analytics in improving financial system stability and risk prevention. Big data analytics enables financial institutions to identify financial risk patterns, monitor financial system performance, and implement risk prevention strategies (32). These capabilities enhance financial system resilience and reduce financial risk exposure. The integration of big data analytics with financial systems has been shown to significantly improve financial risk prevention and financial system stability (27). These findings highlight the critical importance of big data analytics in improving financial system resilience.

Overall, the findings of this study confirm that big data analytics provides a powerful and effective framework for predicting financial and investment risks. The integration of big data analytics with machine learning techniques enhances financial risk prediction accuracy, improves investment decision-making, and strengthens financial risk management capabilities. These findings are consistent with prior research demonstrating that big data analytics significantly improves financial risk prediction and investment performance (36). The results also highlight the importance of integrating behavioral, market, and sentiment data into financial risk prediction models to improve predictive accuracy and financial risk management effectiveness.

Despite the significant contributions of this study, several limitations should be acknowledged. First, the study was conducted using financial data from investors in Tehran, which may limit the generalizability of the findings to other financial markets with different economic structures, regulatory environments, and investment behaviors. Financial markets vary significantly across countries, and investor behavior, market volatility, and financial system characteristics may influence financial risk prediction models differently. Second, although this study utilized large-scale financial datasets, the availability and quality of financial data may affect predictive model performance. Data incompleteness, reporting biases, and data integration challenges may introduce potential limitations in predictive accuracy. Third, this study employed cross-sectional data, which limits the ability to capture long-term temporal changes in financial risk dynamics. Longitudinal data analysis would provide a more comprehensive understanding of financial risk evolution over time. Additionally, while advanced machine learning models were used, model performance may be influenced by hyperparameter selection, data preprocessing methods, and feature engineering processes.

Future research should expand the scope of financial risk prediction studies by incorporating larger and more diverse financial datasets from multiple countries and financial markets. Comparative cross-national studies would provide valuable insights into the effectiveness of big data analytics in different financial environments. Future research should also explore the integration of additional data sources, including real-time financial transaction data, macroeconomic indicators, and alternative data sources such as social media and digital financial platforms. Longitudinal studies should be conducted to examine the long-term effectiveness of big data analytics in predicting financial risk and capturing dynamic financial risk patterns. Additionally, future research should investigate the use of advanced deep learning techniques, including recurrent neural networks and transformer-based models, to improve financial risk prediction accuracy. The integration of explainable artificial intelligence techniques should also be explored to enhance the interpretability and transparency of financial risk prediction models.

The findings of this study provide important practical implications for financial institutions, investors, and policymakers. Financial institutions should integrate big data analytics and machine learning techniques into their financial risk management systems to enhance risk prediction accuracy and improve financial decision-making processes. Investors should utilize big data–driven predictive tools to assess financial risk exposure and optimize investment strategies. Financial regulators and policymakers should promote the adoption of big data analytics in financial risk monitoring systems to enhance financial system stability and prevent financial crises. Investment firms should also invest in advanced data analytics infrastructure and develop predictive analytics capabilities to enhance financial performance and reduce investment risk.

Acknowledgments

We would like to express our appreciation and gratitude to all those who helped us carrying out this study.

Authors' Contributions

All authors equally contributed to this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

All ethical principles were adheried in conducting and writing this article.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

References

1. Rath S, Pandey M, Rautaray SS. Synergistic Review of Automation Impact of Big Data, AI, and ML in Current Data Transformative Era. F1000research. 2025;14:253. doi: 10.12688/f1000research.161477.2.
2. Siddiqui AH, Swetha V, Chowdhary H, Krishna R, Elangovan M, Maguluri LP. Harnessing Big Data: Strategic Insights for IT Management. International Journal of Advanced Computer Science and Applications. 2024;15(7). doi: 10.14569/ijacsa.2024.0150790.
3. Chen L, Jiang R. Internet Financial Risk Model Evaluation and Control Decision Based on Big Data. Wireless Communications and Mobile Computing. 2022;2022(1). doi: 10.1155/2022/8606624.
4. Dong S. Research on the Application of Big Data Technology. Applied and Computational Engineering. 2024;106(1):186-91. doi: 10.54254/2755-2721/106/20241335.
5. Din R, Ahmed S, Khan S, Albanyan A, Hoxha J, Alkhamees BF. A Novel Decision Ensemble Framework: Attention-Customized BiLSTM and XGBoost for Speculative Stock Price Forecasting. Plos One. 2025;20(4):e0320089. doi: 10.1371/journal.pone.0320089.
6. Diao Z. Research on Investment Project Risk Prediction and Management Based on Machine Learning. Applied and Computational Engineering. 2024;87(1):142-7. doi: 10.54254/2755-2721/87/20241561.
7. Toromade AS, Chiekezie NR. Forecasting Stock Prices and Market Trends Using Historical Data to Aid Investment Decisions. Finance & Accounting Research Journal. 2024;6(8):1472-84. doi: 10.51594/farj.v6i8.1434.

8. Peng H. Research on Credit Risk Identification of Internet Financial Enterprises Based on Big Data. Mobile Information Systems. 2021;2021:1-8. doi: 10.1155/2021/1034803.
9. Bowala S, Singh J. Optimizing Portfolio Risk of Cryptocurrencies Using Data-Driven Risk Measures. Journal of Risk and Financial Management. 2022;15(10):427. doi: 10.3390/jrfm15100427.
10. Wang C, Liu S. Innovative Risk Early Warning Model Based on Internet of Things Under Big Data Technology. Ieee Access. 2021;9:100606-14. doi: 10.1109/access.2021.3095503.
11. Yang Y, Zhang K, Fan Y. Analyzing Firm Reports for Volatility Prediction: A Knowledge-Driven Text-Embedding Approach. Informs Journal on Computing. 2022;34(1):522-40. doi: 10.1287/ijoc.2020.1046.
12. Sawhney R, Aggarwal A, Shah RR. An Empirical Investigation of Bias in the Multimodal Analysis of Financial Earnings Calls. 2021. doi: 10.18653/v1/2021.naacl-main.294.
13. Saleh I, Marei Y, Ayoush M, Afifa MA. Big Data Analytics and Financial Reporting Quality: Qualitative Evidence From Canada. Journal of Financial Reporting and Accounting. 2022;21(1):83-104. doi: 10.1108/jfra-12-2021-0489.
14. Ren S. Optimization of Enterprise Financial Management and Decision-Making Systems Based on Big Data. Journal of Mathematics. 2022;2022(1). doi: 10.1155/2022/1708506.
15. Adebiyi OO. Exploring the Impact of Predictive Analytics on Accounting and Auditing Expertise: A Regression Analysis of LinkedIn Survey Data. Asian Journal of Economics Business and Accounting. 2023;23(22):286-305. doi: 10.9734/ajeba/2023/v23i221153.
16. Parthiban M. Big Data Analytics in Fintech: A Review of Credit Risk Assessment and Fraud Detection. Eatp. 2024:3676-84. doi: 10.53555/kuey.v30i5.3514.
17. Olabanji SO, Oladoyinbo TO, Asonze CU, Adigwe CS, Okunleye OJ, Olaniyi OO. Leveraging FinTech Compliance to Mitigate Cryptocurrency Volatility for Secure US Employee Retirement Benefits: Bitcoin ETF Case Study. Asian Journal of Economics Business and Accounting. 2024;24(4):147-67. doi: 10.9734/ajeba/2024/v24i41270.
18. Adeniran IA, Efunniyi CP, Osundare OS, Abhulimen AO. Leveraging Big Data Analytics for Enhanced Market Analysis and Competitive Strategy in the Oil and Gas Industry. International Journal of Management & Entrepreneurship Research. 2024;6(8):2849-65. doi: 10.51594/ijmer.v6i8.1470.
19. Bhatti RK. Financial Market Volatility and Risk Management Strategies. Pijr. 2024:46-9. doi: 10.36106/paripex/4304085.
20. Liu J. A Hybrid Model Integrating LSTM With Multiple GARCH-Type Models for Volatility and Var Forecast. 2023. doi: 10.4108/eai.6-1-2023.2330313.
21. Prasad S, Kiran R, Sharma RK. Influence of Financial Literacy on Retail Investors' Decisions in Relation to Return, Risk and Market Analysis. International Journal of Finance & Economics. 2020;26(2):2548-59. doi: 10.1002/ijfe.1920.
22. Adeoye OB, Ani EC, Ninduwesuor-Ehiobu N, Montero DJP, Usman FO, Olu-lawal KA. A Conceptual Framework for Data-Driven Sustainable Finance in Green Energy Transition. World Journal of Advanced Research and Reviews. 2024;21(2):1791-801. doi: 10.30574/wjarr.2024.21.2.0620.
23. Ofodile OC, Oyewole AT, Ugochukwu CE, Addy WA, Adeoye OB, Okoye CC. Predictive Analytics in Climate Finance: Assessing Risks and Opportunities for Investors. GSC Advanced Research and Reviews. 2024;18(2):423-33. doi: 10.30574/gscarr.2024.18.2.0076.
24. Zhu C, Chu J. The Impact of Chinese Big Tech on the Traditional Financial Market: Evidence From Ant Group. Electronic Commerce Research. 2023;25(2):879-905. doi: 10.1007/s10660-023-09694-5.
25. Liang Q. Financial Risk Prediction and Prevention Based on Big Data Technology. Financial Engineering and Risk Management. 2023;6(11). doi: 10.23977/ferm.2023.061104.
26. Wang Y. Research on Risk Control Model of Internet Financial Based on Data Envelopment Analysis. 2023. doi: 10.1117/12.2671911.
27. Wu C, Liu J, Zhang H. Data Ecology and Accurate Portrait: Optimization of Credit Risk System for SMEs in Supply Chain Finance Based on Big Data Technology. Journal of Risk Analysis and Crisis Response. 2022;11(4). doi: 10.54560/jracr.v11i4.310.
28. Kernchen R. Risk Forecasting in the Light of Big Data. Journal of Risk Analysis and Crisis Response. 2021;10(4):160. doi: 10.2991/jracr.k.201230.001.
29. Wu J. Data Analytics and Data Mining Techniques in Financial Investment Risk Management. 2024. doi: 10.3233/faia241100.
30. Zhao Y. Empowering Sustainable Finance: The Convergence of AI, Blockchain, and Big Data Analytics. Advances in Economics Management and Political Sciences. 2024;85(1):267-73. doi: 10.54254/2754-1169/85/20240925.

31. Wang Y, Wang L. Internet Financial Data Security and Economic Risk Prevention for Android Application Privacy Leakage Detection. Computational Intelligence and Neuroscience. 2022;2022:1-12. doi: 10.1155/2022/6782281.
32. Gao L, Xiao J. Big Data Credit Report in Credit Risk Management of Consumer Finance. Wireless Communications and Mobile Computing. 2021;2021(1). doi: 10.1155/2021/4811086.
33. Bisht D, Singh R, Gehlot A, Akram SV, Singh A, Montero EC, et al. Imperative Role of Integrating Digitalization in the Firms Finance: A Technological Perspective. Electronics. 2022;11(19):3252. doi: 10.3390/electronics11193252.
34. Li X. A Study of Stock Market Risk Assessment Using Big Data. Advances in Economics Management and Political Sciences. 2024;86(1):164-9. doi: 10.54254/2754-1169/86/20240957.
35. Zhang R, Sidik MHJ. Big Data, Artificial Intelligence, and Financial Literacy: Exploring Their Combined Influence on Investment Behavior Among Chinese Household. Journal of Information Systems Engineering & Management. 2024;9(1):24446. doi: 10.55267/iadt.07.14651.
36. Aderemi S, Olutimehin DO, Nnaomah UI, Orieno OH, Edunjobi TE, Babatunde SO. Big Data Analytics in the Financial Services Industry: Trends, Challenges, and Future Prospects: A Review. International Journal of Science and Technology Research Archive. 2024;6(1):147-66. doi: 10.53771/ijstra.2024.6.1.0036.
37. Eze C, Igwama GT, Nwankwo EI, Emeihe EV. The Role of Big Data in Transforming Financial Management in U.S. Healthcare: A Conceptual Framework. Finance & Accounting Research Journal. 2024;6(11):2093-107. doi: 10.51594/farj.v6i11.1698.
38. Odeyemi O, Mhlongo NZ, Nwankwo EE, Scholatica UC, Okoye CC. Big Data Applications in Portfolio Management: A Review of Techniques and Strategies. International Journal of Science and Research Archive. 2024;11(1):1984-99. doi: 10.30574/ijsra.2024.11.1.0268.