Iranian Journal of Finance
Print ISSN 2676-6337
Online ISSN 2676-6345

Analysis of the characteristics affecting the trading risk of listed companies' stocks: A hybrid spatial artificial intelligence approach

Javad Zolfaghary Tabesh

Ph.D. Candidate, Department of Accounting, Ker.C., Islamic Azad University, Kermanshah, Iran. (Email: j.zolfagharytabesh@iau.ac.ir)

Babak Jamshidi navid *

*Corresponding Author, Associate prof., Department of Accounting, Ker.C., Islamic Azad University, Kermanshah, Iran. (Email: Babak.jamshidinavid@iau.ac.ir)

Mehrdad Ghanbary

Associate prof., Department of Accounting, Ker.C., Islamic Azad University, Kermanshah, Iran. (Email: mehrdadghanbary@iau.ac.ir

Afshin Baghfalaki

Assistant prof., Department of Accounting, Ker.C., Islamic Azad University, Kermanshah, Iran. (Email: Afshin.baghfalaki@iau.ac.ir)

Iranian Journal of Finance, 2026, Vol. 10, No.1, pp. 173-237.
Publisher: Iran Finance Association
doi: https://doi.org/10.30699/ijf.2026.569343.1569
Article Type: Original Article

Received: August 14, 2025
Received in revised form: October 29, 2025
Accepted: December 17, 2025

Abstract

This research identifies the determinants of trading risk (conditional variance) in the Iranian stock market over 15 years (2008-2023) using a novel hybrid approach combining spatial econometrics and machine learning algorithms. The main objective is to evaluate the superiority of hybrid models over traditional methods and to identify the roles of macroeconomic, geopolitical, behavioral factors, and firm characteristics in shaping systematic risk. The sample includes 172 companies listed on the Tehran Stock Exchange with 30,960 monthly observations and 33 explanatory variables. The methodology was implemented in three stages: First, GARCH and EGARCH models were employed to extract conditional variance and confirm the leverage effect. Second, the Spatial Durbin Error Model (SDEM) was used to decompose direct, spatial spillover, and total effects of variables while controlling for cross-sectional dependence (Pesaran CD statistic = 87.45***) and spatial autocorrelation (Moran's I = 0.4567***). Third, machine learning algorithms, including Linear Regression, SVM, Random Forest, XGBoost, LSTM, and Transformer, were applied independently and in combination with SDEM outputs. The results demonstrated a clear performance hierarchy: Linear Regression ($R^2 = 0.4123$, RMSE = 0.0987), SVM ($R^2 = 0.5987$), Random Forest ($R^2 = 0.6789$), XGBoost as the best standalone model ($R^2 = 0.7456$, RMSE = 0.0534), and Ensemble ($R^2 = 0.7523$). Hybrid models showed significant superiority: SDEM + XGBoost ($R^2 = 0.7823$, RMSE = 0.0471; 11.80% error reduction compared to standalone XGBoost and 52.3% improvement over Linear Regression), and SDEM + Ensemble ($R^2 = 0.7867$, RMSE = 0.0467) achieved optimal performance. Time-series cross-validation (average test RMSE = 0.0492) and the Diebold-Mariano test (DM = 3.456*** against XGBoost) confirmed statistical superiority. From a substantive perspective, the exchange rate with a total effect of 0.2443*** and SHAP contribution of 18.34% was identified as the most important systematic risk factor, followed by sanction intensity (total effect = 0.1274***, SHAP = 14.23%), Altman Z-score (SHAP = 15.67%), total stock index (total effect = -0.1801***, SHAP = 12.89%), and investor sentiment (total effect = 0.1001***, SHAP = 11.45%). The findings demonstrate that hybrid spatial econometrics and machine learning models improve prediction accuracy by 12-15% through extracting complementary information. Geopolitical and behavioral factors, in addition to traditional macroeconomic variables, are systematically important. Spatial spillovers constitute 15-25% of total effects, which are ignored in traditional models. This research shows that the frontier of financial risk modeling lies in the synergistic integration of economic theory and machine learning.

Keywords: Trading Risk, Spatial Econometrics, Machine Learning, Iranian Stock Market
JEL Classification: C21, C45, C53, G11, G15

Introduction

Stock trading risk refers to the probability of financial loss arising from price volatility, low liquidity, and market uncertainty during the buying and selling of securities (Jorion, 2007; Dima et al., 2025). This concept encompasses systematic risks influenced by macroeconomic factors and unsystematic risks related to company-specific characteristics (Markowitz, 1952; Moradi et al, 2024). Effective features in trading risk span a wide spectrum of variables, ranging from technical indicators such as moving averages and the Relative Strength Index (RSI) to fundamental factors such as price-to-earnings ratios and debt-to-equity ratios (Murphy, 1999; Sukma et al., 2024). With advances in machine learning, the identification and ranking of these features for more accurate risk prediction and investment decision optimization have become increasingly important (Henrique et al., 2019; Dixon et al., 2020).

With the increasing complexity of financial markets and the availability of enormous volumes of high-frequency trading data, identifying truly effective features for predicting stock trading risk has become a fundamental challenge (Tsay, 2010; Nimalendran et al., 2024). Investors and risk managers face hundreds of potential variables, and determining which variables have a significant impact on risk and under what market conditions requires advanced analytical approaches (Engle, 2001; Johri et al., 2023). The problem intensifies during crisis periods when feature behaviors and interrelationships change fundamentally (Cont, 2001). Traditional risk analysis methods, such as GARCH models and historical simulation, often perform poorly under extreme market conditions and are unable to capture nonlinear dynamics across features (McNeil & Frey, 2000). Moreover, existing studies typically focus on return or price prediction and have not sufficiently addressed the systematic analysis of features that affect trading risk itself, particularly under extreme volatility conditions (Gerlein et al., 2016; Wu, 2021).

The importance of accurately identifying effective features in stock trading risk can be examined from several perspectives. First, incorrect risk estimation can lead to inefficient capital allocation, severe financial losses, and systemic instability in financial markets (Adrian & Brunnermeier, 2016; Acharya et al., 2009; Undheim, 2024). Second, identifying key features enables investors to create more optimal portfolios with improved risk-to-return ratios (Sharpe, 1994; Torkian et al., 2025). Third, in the era of algorithmic trading, a deep understanding of the features that affect risk helps develop more robust trading strategies resilient to varying market conditions (Addy et al., 2024). Fourth, regulatory institutions require advanced tools to identify and control systemic risk, necessitating a precise understanding of features that have the greatest impact on risk during crisis conditions (Basel Committee on Banking Supervision, 2019). Finally, given the growing emergence of developing markets and increasing interconnection among global markets, the need for risk models capable of identifying effective features across different geographical contexts and market regimes is felt more than ever (Hao et al., 2023).

Previous studies in predicting risk and returns using machine learning have pursued diverse paths. Gerlein et al. (2016) demonstrated the effectiveness of machine learning models in predicting profitable forex trades, but left open the question of how these findings translate to stock markets. Wu (2021) used decision trees, random forests, and hybrid algorithms to assess over-financialization risk, achieving 79.35% accuracy. However, the research gap remains in applying these methods to high-frequency stock trading risk. Addy et al. (2024) highlighted challenges such as overfitting, data bias, and interpretability issues in machine learning algorithms for algorithmic trading, underscoring a critical gap in developing robust, interpretable models to identify effective features for trading risk. Hao et al. (2023) compared random forest and deep neural network models across five major stock markets and demonstrated the consistent superiority of deep learning, but did not conduct a systematic analysis of influential risk features across different market conditions. Uddin et al. (2025) explored random forests, gradient boosting, LSTM, and transformers for financial risk prediction, showing that combining machine learning with traditional optimization yields superior Sharpe and Sortino ratios, but the gap in systematically identifying and ranking the most effective features contributing to trading risk across different asset classes and market regimes persists. Most critically, Melina et al. (2025), in their systematic literature review using the PRISMA method and analyzing 90 articles, revealed that despite the common use of EVT, GARCH, and historical simulation for investment risk estimation, machine learning-based risk estimation remains scarce, with no studies combining EVT and machine learning. This demonstrates that despite significant advances in applying machine learning for market prediction, the integration of classical statistical methods, such as Extreme Value Theory, with advanced machine learning techniques for comprehensive analysis of effective features in stock trading risk, particularly under crisis conditions, has not been adequately investigated.

A comprehensive review of both international and domestic literature reveals that no study has yet provided an integrated, systematic evaluation of a wide range of advanced machine learning algorithms for predicting stock trading risk on the Tehran Stock Exchange. This research gap becomes particularly significant given the unique characteristics of Iran's capital market, including sanction-induced volatility, currency shocks, chronic inflation, emotional investor behavior, nonlinear relationships, and structural breaks in data, all of which substantially diminish the accuracy of linear or single-algorithm approaches (Taheri Haftasiabi et al., 2023; Keshavarz & Rezaei, 2022). Moreover, most domestic studies have focused exclusively on price prediction accuracy while neglecting direct measures of trading risk, such as crash probability, Value at Risk, or risk level classification, and have not offered a systematic comparison among different algorithm families. The present research addresses these gaps through three fundamental innovations: first, providing the first comprehensive and systematic evaluation framework of six advanced algorithm families (from simple regression to deep neural networks) in two distinct scenarios of continuous risk prediction and risk level classification within the Tehran Stock Exchange context; second, utilizing a rich, multi-dimensional dataset comprising 33 independent variables across five main dimensions (macroeconomic indicators, industry characteristics, company characteristics, political events, and behavioral-cultural factors of investors), representing the first integration of structured and semi-structured data at this scale in domestic studies; and third, combining advanced econometric GARCH and EGARCH models to calculate conditional variance as a measure of trading risk with machine learning algorithms, representing a unique synthesis of classical econometric methods and modern artificial intelligence that can serve as a model for future research in emerging markets with similar characteristics (Melina et al., 2025; Uddin et al., 2025).

Given the identified research gaps, the need to develop a hybrid framework integrating Extreme Value Theory with advanced machine learning techniques to analyze effective features in stock trading risk is evident (Melina et al., 2025). This need arises from the fact that financial markets exhibit nonlinear behavior and complex dependencies, especially during crisis periods, which traditional models cannot capture (Embrechts et al., 1997). Furthermore, there is an urgent need for models that systematically identify, quantify, and rank the most influential features affecting trading risk across different market regimes, time horizons, and asset characteristics using high-frequency multivariate data (McNeil et al., 2015). Accordingly, the objectives of this research are: (1) developing a hybrid framework that integrates Extreme Value Theory with advanced machine learning algorithms such as random forests, gradient boosting, and deep neural networks; (2) systematic identification and quantification of technical, fundamental, and structural features that have the greatest impact on stock trading risk; (3) ranking feature importance under normal market conditions versus periods of extreme volatility; (4) examining temporal dynamics and changes in feature importance across different trading time horizons; and (5) evaluating the performance of the proposed framework across different stock markets and comparing it with existing methods in terms of risk prediction accuracy and robustness (Danielsson et al., 2016).

The remainder of this paper is structured as follows: Section 2 presents the literature review covering theoretical foundations and empirical studies. Section 3 introduces the model design, methodological framework, and descriptive statistics. Section 4 presents the empirical analyses. Finally, Section 5 concludes with a discussion and conclusions.

Literature Review

Stock trading risk represents the probability of financial loss arising from price volatility, low liquidity, and market uncertainties (Alexander, 2008). Rooted in Markowitz's (1952) portfolio selection theory, this concept encompasses systematic risk influenced by macroeconomic factors and unsystematic risk related to company-specific characteristics (Fama & French, 1992). Technological developments in machine learning and access to high-frequency trading data have created new challenges in identifying and prioritizing characteristics that affect trading risk (Christoffersen, 2012). The theoretical literature encompasses approaches ranging from classical risk assessment models such as GARCH and Value at Risk to advanced machine learning frameworks and Extreme Value Theory (Hull, 2018). Recent advances highlight the need to integrate quantitative risk management techniques with artificial intelligence to capture both linear and nonlinear dependencies in financial data (Goodfellow et al., 2016).

Modern Portfolio Theory, founded by Markowitz (1952), provides the fundamental framework for understanding trading risk and diversification (Elton et al., 2014). Sharpe (1964) extended this framework by introducing the Capital Asset Pricing Model (CAPM), explaining the linear relationship between expected return and systematic risk (Bodie et al., 2014). Subsequently, Engle (1982) introduced ARCH models, and Bollerslev (1986) developed the GARCH model, demonstrating that conditional variance of financial returns varies over time and can be modeled using past information (Brooks, 2014). These models have found widespread application in forecasting market volatility and assessing short-term risk (Francq & Zakoïan, 2019). Dowd (2005) expanded the concept of Value at Risk (VaR) as a standard measure for market risk, representing the maximum probable loss over a specified time horizon with a certain confidence level. However, these classical models have significant limitations; they primarily assume normal distribution of returns, which is violated in reality, and perform poorly under crisis conditions (Mandelbrot & Hudson, 2004). Furthermore, these approaches are unable to identify nonlinear relationships and complex dynamics among various risk-affecting characteristics (Poon & Granger, 2003). The 2008 global financial crisis exposed significant weaknesses in traditional risk models, particularly their failure to capture tail risk and systemic interconnections (Brunnermeier, 2009).

Extreme Value Theory (EVT), expanded in finance by de Haan and Ferreira (2006), provides a powerful statistical framework for modeling rare and severe events located in distribution tails based on Generalized Extreme Value (GEV) distributions or Generalized Pareto Distribution (GPD) (Coles, 2001). Longin (2000) demonstrated that EVT outperforms traditional methods in predicting extreme events. Jondeau and Rockinger (2003) documented that the distribution of financial returns has heavy tails that a normal distribution cannot describe. Chavez-Demoulin et al. (2014) showed that EVT can be combined with copula models to model the dependence structure between assets under extreme conditions. Acharya et al. (2017) introduced systemic risk measurement concepts, emphasizing the importance of understanding tail risks in financial stability. The International Monetary Fund (2009) recommended the use of advanced approaches, including EVT, for assessing systemic risk. The theoretical foundation of EVT lies in the Fisher-Tippett-Gnedenko theorem, which states that the distribution of appropriately normalized maxima of independent and identically distributed random variables converges to one of three extreme value distributions (Embrechts et al., 1997). For practical applications in finance, the Generalized Pareto Distribution (GPD) is commonly applied to model exceedances over a high threshold, providing a flexible framework for capturing the behavior of extreme losses in asset returns (McNeil & Frey, 2000). The threshold selection, typically based on the mean excess function or graphical diagnostics, is a critical step in EVT implementation, as it must balance the trade-off between variance reduction (a higher threshold) and bias minimization (a lower threshold) (Scarrott & MacDonald, 2012). In the context of stock trading risk, EVT enables the estimation of risk measures such as Value at Risk (VaR) and Expected Shortfall (ES) for extreme quantiles that lie beyond the range of historical observations, thus addressing a fundamental limitation of traditional approaches that rely on the assumption of normality (Danielsson & de Vries, 1997). However, EVT alone cannot systematically identify and rank risk-affecting characteristics and requires a combination with advanced machine learning methods (Diebold & Yilmaz, 2014). While EVT provides a theoretically rigorous framework for tail risk modeling, its empirical implementation in emerging markets faces challenges including limited data in the tail region, non-stationarity of financial time series, and the need to integrate extreme risk measures with broader risk management frameworks (Gilli & Këllezi, 2006). These challenges necessitate a systematic approach that connects EVT's theoretical foundations with practical implementation strategies, particularly when combined with feature selection algorithms to identify the most relevant risk drivers under extreme market conditions (Soltane et al., 2012). The integration of EVT with machine learning techniques requires careful consideration of how tail risk metrics derived from EVT (such as VaR and ES at extreme confidence levels) can serve as both response variables and explanatory features in predictive models, a connection that has received limited attention in existing literature (Bee et al., 2016). Recent research has emphasized the integration of EVT with computational intelligence techniques to enhance predictive accuracy in extreme market scenarios (Roh et al., 2020).

Recent advances in machine learning have provided new capabilities for identifying and prioritizing characteristics affecting trading risk (Gu et al., 2020). Hastie et al. (2009) demonstrated that ensemble methods, such as random forests, can improve prediction accuracy and estimate feature importance using various metrics. James et al. (2013) showed that boosting algorithms improve predictive performance by sequentially building models and focusing on errors from previous models. LeCun et al. (2015) provided a comprehensive review of deep learning methods, demonstrating their ability to learn hierarchical representations from raw data. Graves (2012) presented advances in recurrent neural networks, including LSTM architectures that can learn long-term dependencies in sequential data. Lopez de Prado (2018) demonstrated that machine learning can be used to develop systematic trading strategies capable of extracting complex patterns from market microstructure data. Krauss et al. (2017) showed that deep learning models can achieve superior out-of-sample performance compared to traditional statistical methods. However, Mullainathan and Spiess (2017) highlighted that machine learning algorithms face challenges such as overfitting, interpretability issues, and difficulty of establishing causal relationships, emphasizing the need to develop robust and interpretable models. Recent studies have also explored the application of reinforcement learning and generative adversarial networks for improving risk prediction (Deng et al., 2017).

The development of hybrid frameworks combining classical statistical models and machine learning has emerged as a new research frontier in risk analysis for trading (Huang & Tsai, 2009). Hansen et al. (2011) introduced sophisticated approaches to model combination and selection for forecasting financial volatility, demonstrating that model averaging can yield more robust risk predictions. Li et al. (2016) demonstrated that integration of GARCH-type models with neural networks can significantly improve volatility prediction accuracy. Chandrashekar and Sahin (2014) examined various feature selection methods, including filter, wrapper, and embedded approaches, and provided comprehensive guidelines for their application to high-dimensional financial datasets. Genuer et al. (2010) developed variable selection procedures using random forests that account for collinearity among predictors. Ribeiro et al. (2016) introduced LIME to explain the predictions of any machine learning classifier in an interpretable manner, addressing the black-box nature of complex models. Rasekhschaffe and Jones (2019) showed that combining machine learning with traditional financial theories yields superior performance in predicting stock returns. Feng et al. (2019) demonstrated the importance of proper feature engineering and selection in capturing risk premia. However, Jacobs et al. (2014) emphasized that, despite advances, significant gaps remain in integrating classical risk modelling frameworks, such as EVT, with modern computational methods for comprehensive risk analysis, particularly under extreme market conditions. Recent developments have also emphasized the importance of incorporating alternative data sources and sentiment analysis in risk modeling frameworks (Ke et al., 2019).

The literature review demonstrates that analysis of characteristics affecting stock trading risk evolved from Markowitz's classical portfolio theory through CAPM, GARCH models, VaR, and Extreme Value Theory (Campbell et al., 1997). Machine learning and deep learning techniques have enabled the identification of complex patterns and nonlinear relationships in financial data (Cavalcante et al., 2016). However, a significant research gap exists in integrating classical statistical methods, such as EVT, with advanced machine learning techniques to analyze risk-affecting characteristics, particularly under crisis conditions comprehensively (Cont, 2007). Specifically, while the theoretical foundations of EVT are well-established and its superiority in modeling tail risks is documented, the literature lacks a systematic framework that explicitly connects EVT's distributional assumptions and parameter estimation procedures with feature engineering and selection processes in machine learning pipelines (Nasekin & Chen, 2018). This gap is particularly pronounced in emerging markets where data limitations and structural breaks complicate the empirical application of EVT-based approaches (Gençay et al., 2003).

Furthermore, existing studies have not adequately addressed how EVT-derived risk measures can be operationalized as both predictive targets and explanatory variables within integrated modeling frameworks that account for regime-switching dynamics and time-varying feature importance (Malevergne & Sornette, 2006). Existing frameworks primarily focus on return or price prediction and have not adequately addressed the systematic analysis of characteristics that affect trading risk across different market regimes (Harvey et al., 2016). Furthermore, temporal dynamics and changes in feature importance across different time horizons, asset characteristics, and market microstructure conditions require further investigation (Bouchaud et al., 2009). The need to develop interpretable models that perform well under both normal and extreme market conditions while maintaining computational efficiency emerges as a key future research priority (Molnar, 2020). This research addresses these gaps by developing a hybrid framework that explicitly bridges the theoretical foundations of EVT with practical implementation through advanced machine learning algorithms. Our approach systematically identifies, quantifies, and ranks characteristics affecting stock trading risk by: (i) employing EVT to derive tail risk measures that capture extreme market behaviors, (ii) integrating these measures as features in ensemble learning models, and (iii) establishing a clear methodological pipeline that connects distributional assumptions, threshold selection, parameter estimation, and feature importance analysis (detailed in Section 3 - Methodology). This research fills these gaps by developing a hybrid framework that combines EVT with advanced machine learning algorithms to identify, quantify systematically, and rank characteristics that affect stock trading risk (McNeil & Frey, 2000; Christoffersen, 2012).

There are numerous empirical studies on reducing stock trading risk, which are reviewed below.

Gerlein et al. (2016) aimed to evaluate the effectiveness of simple machine learning models in predicting profitable trades by conducting six-year trading simulations on USDJPY, EURGBP, and EURUSD currency pairs, demonstrating that even modestly accurate simple models can generate substantial long-term financial returns through periodic retraining and proper feature selection. However, their study focused on forex markets rather than stock trading risk, leaving a gap in understanding how these findings translate to equity markets with different volatility patterns. Wu (2021) employed decision trees, random forests, and hybrid algorithms on six months of loan data from a trading institution to assess over-financialization risk, achieving 79.35% accuracy, 39.28% recall, and 78.28% precision with the hybrid model. However, the research gap remains in applying these methods to real-time stock trading risk assessment with high-frequency data Liu and Yu (2022) analyzed financial market risks under blockchain technology using random forests and decision trees, finding that blockchain can improve financial industry efficiency and reduce operating costs. The gap here lies in the lack of integration between blockchain-based risk analysis and feature importance identification in stock trading scenarios. Addy et al. (2024) conducted a comprehensive literature review revealing that machine learning algorithms in algorithmic trading face challenges such as overfitting, data bias, and interpretability issues. This highlights a critical gap in developing robust, interpretable models specifically designed for identifying and analyzing effective features in stock trading risk.

Avramov et al. (2022) analyzed the profitability of deep learning signals in hard-to-arbitrage stocks, finding that while deep learning-based investments remain profitable in long positions and recent years, reasonable transaction costs significantly reduce profitability. The research gap lies in understanding which specific features contribute most to risk in these hard-to-arbitrage scenarios and how to identify them systematically. Hao et al. (2023) compared random forest and deep neural network models across five major stock markets (China, US, UK, Canada, Japan) during 2005-2020, demonstrating that deep learning models consistently outperform traditional machine learning, especially in the Chinese market. However, their study lacks a systematic analysis of which trading risk features are most influential across different market conditions and geographical contexts. Maleki et al. (2023) developed an algorithmic gold trading strategy using LSTM networks for price prediction, incorporating feature selection and missing-value imputation, demonstrating that combining these techniques significantly improves model performance. The gap remains in extending this approach to comprehensive stock-trading risk-feature analysis beyond single-commodity applications. Uddin et al. (2025) explored random forests, gradient boosting, LSTM, and transformers for financial risk prediction and portfolio optimization, finding that LSTM and transformer architectures excel at capturing long-term dependencies and that combining machine learning with traditional optimization yields superior Sharpe and Sortino ratios. Nevertheless, a gap persists in systematically identifying and ranking the most effective features contributing to trading risk across different asset classes and market regimes.

Melina et al. (2025) conducted a systematic literature review using the PRISMA method, analyzing 90 articles from ScienceDirect and Scopus to examine investment risk estimation capable of detecting extreme values. Their findings revealed that EVT, GARCH, and historical simulation are commonly used for investment risk estimation. However, machine learning-based risk estimation remains scarce with no studies combining EVT and machine learning, proposing a hybrid modeling framework integrating EVT and machine learning using high-frequency multivariate data. This critical research gap demonstrates that despite significant advances in applying machine learning for market prediction and portfolio management, the integration of classical statistical methods such as Extreme Value Theory with advanced machine learning techniques for comprehensive analysis of effective features in stock trading risk—particularly under crisis conditions and extreme market volatility—has not been adequately investigated. Furthermore, there is a pressing need to develop hybrid frameworks that not only combine EVT with machine learning but also systematically identify, quantify, and rank the most influential features affecting trading risk across different market states, time horizons, and asset characteristics using high-frequency multivariate data, thereby enabling more robust risk management strategies that account for both normal and extreme market conditions.

Data

Sample Selection, Data Collection, and Variable Construction

This research employs a comprehensive dataset spanning 15 years (2008-2024) from the Tehran Stock Exchange (TSE), encompassing 172 companies across multiple industries after applying systematic screening criteria. The initial population consisted of 370 companies listed on the TSE, which was refined through rigorous filtering procedures to ensure data quality, consistency, and analytical reliability (Wu, 2021; Liu & Yu, 2022). The screening criteria implemented were as follows: (1) companies must have been listed on the TSE before 2008 and remained active until the end of 2023 without delisting; (2) fiscal year-end must be standardized to the last day of Esfand (March 20-21) without any changes during the study period; (3) continuous trading activity throughout the study period with cumulative trading suspensions not exceeding six months; (4) exclusion of holding companies, financial intermediaries, banks, and insurance companies due to their fundamentally different operational and financial structures compared to industrial and commercial firms; and (5) complete availability of financial and trading information without missing values for all required variables (Gerlein et al., 2016; Hao et al., 2023).

The 15-year study period from 2008 to 2023 was deliberately selected to capture diverse economic cycles, including prosperity, recession, international sanctions, political transitions, and episodes of extreme market volatility, thereby enabling a comprehensive examination of their impacts on stock trading risk across varying market conditions. Data were collected and analyzed at monthly frequency, resulting in 180 monthly observations per company and a total of 30,960 firm-month observations (172 companies × 180 months). This longitudinal panel structure provides sufficient temporal depth to capture both short-term fluctuations and long-term trends in trading risk dynamics while maintaining adequate cross-sectional variation for robust statistical inference (Avramov et al., 2022; Maleki et al., 2023).

The dataset integrates structured and semi-structured data from multiple authoritative sources to construct a comprehensive set of 33 independent variables across five main dimensions: macroeconomic indicators, industry characteristics, company characteristics, political events, and behavioral-cultural factors of investors. Financial statement data were extracted from the Codal system (the official reporting platform of Iran's Securities and Exchange Organization), which provides audited annual and quarterly financial reports for all listed companies. Market trading data, including daily prices, trading volumes, and returns, were obtained from the TSE official database and supplemented with data from specialized financial information providers such as the Rahavard Novin and TSETMC websites (Jorion, 2007; Cont, 2001).

Macroeconomic indicators were compiled from official publications of the Central Bank of Iran, the Statistical Center of Iran, and the Ministry of Economic Affairs and Finance. These indicators include: (1) Gross Domestic Product (GDP) growth rate, initially reported quarterly and converted to monthly frequency using cubic spline interpolation, then transformed to year-over-year percentage change; (2) inflation rate measured as the year-over-year percentage change in the Consumer Price Index (CPI); (3) exchange rate operationalized as the logarithm of the average official USD/IRR buy and sell rates; (4) money supply growth measured using the broad monetary aggregate M2; (5) bank interest rates for deposits and lending; (6) trade balance derived from customs administration data; (7) Gini coefficient as a measure of income inequality; (8) ease of doing business index compiled from World Bank data and domestic business surveys; (9) housing economic indicators including construction permits and housing price indices; (10) productivity index calculated from labor and capital productivity measures; and (11) Tehran Stock Exchange overall index (TEPIX) reflecting aggregate market performance (Alexander, 2008; Engle, 2001).

Industry characteristics were constructed using sectoral data aggregated from TSE classifications and industry reports published by the Securities and Exchange Organization. These variables include: (1) pricing mechanism reflecting the degree of price regulation in each industry, coded on a scale from 0 (fully market-determined) to 10 (fully regulated); (2) supply and demand conditions measured as the ratio of industry sales to production capacity; (3) market concentration quantified using the Herfindahl-Hirschman Index (HHI) calculated as the sum of squared market shares of firms within each industry; (4) capital investment volume aggregated at the industry level from company-level capital expenditure data; (5) performance of related industries measured through inter-industry linkage matrices; and (6) technological change proxied by industry-level R&D intensity and patent applications (Markowitz, 1952; Sharpe, 1964).

Company-level characteristics were directly extracted from audited financial statements and market trading data. These variables encompass a comprehensive set of financial ratios, profitability indicators, market valuation metrics, and financial health measures: (1) earnings per share (EPS) calculated as net income divided by weighted average shares outstanding; (2) dividend per share (DPS) representing cash dividends paid per share; (3) price-to-earnings ratio (P/E) computed as market price per share divided by EPS; (4) return on assets (ROA) measuring net income relative to total assets; (5) Economic Value Added (EVA) calculated as net operating profit after taxes minus the cost of capital employed; (6) Market Value Added (MVA) representing the difference between market value and invested capital; (7) Refined Economic Value Added (REVA) adjusting EVA for accounting distortions; (8) residual income calculated as net income minus equity charge; (9) Tobin's Q measured as the ratio of market value to replacement cost of assets; (10) price-to-sales ratio (P/S) computed as market capitalization divided by total sales; (11) Cash Value Added (CVA) measuring cash flow generation relative to capital employed; (12) free cash flow (FCF) calculated as operating cash flow minus capital expenditures; (13) Altman Z-Score predicting bankruptcy probability using a weighted combination of financial ratios; and (14) Beneish M-Score detecting earnings manipulation likelihood through financial statement analysis (Markowitz, 1952; Sharpe, 1994; Bodie et al., 2014).

Political event variables were constructed to capture the impact of major political developments and geopolitical tensions on trading risk. These include: (1) international agreements, particularly the Joint Comprehensive Plan of Action (JCPOA), coded as a binary variable taking the value of 1 during negotiation and implementation periods and 0 otherwise; (2) sanctions intensity measured using a composite index ranging from 0 to 10, constructed by aggregating information on the scope, sectoral coverage, and stringency of international sanctions imposed on Iran; (3) elections coded as a binary variable taking the value of 1 during the three-month window surrounding presidential and parliamentary elections and 0 otherwise; (4) regional war and geopolitical tensions quantified using the Geopolitical Risk Index adapted for the Middle East region; (5) accession to international treaties coded as binary events; and (6) government economic policy orientation classified based on the degree of market liberalization versus state intervention (Adrian & Brunnermeier, 2016; Basel Committee on Banking Supervision, 2019).

Behavioral-cultural factors of investors were measured using a multi-source proxy approach combining quantitative market indicators and qualitative sentiment analysis. These variables include: (1) investor sentiment measured using both market-based quantitative indices (such as trading volume ratios, turnover rates, and new account openings) and text-based sentiment analysis of financial news, social media discussions, and analyst reports using natural language processing techniques; (2) ambiguity tolerance inferred from market behavior during periods of high uncertainty; (3) investment horizon proxied by average holding periods calculated from turnover data; (4) cognitive biases including overconfidence, anchoring, and herding behavior detected through abnormal trading patterns and price momentum; (5) emotional biases including fear and greed measured through volatility indices and extreme return observations; and (6) risk tolerance estimated from portfolio allocation patterns and leverage utilization across investor segments (Dixon et al., 2020; Henrique et al., 2019).

Research Methodology

Research Design and Methodological Framework

This research adopts a hybrid methodological framework that integrates classical econometric techniques with advanced machine learning algorithms to comprehensively analyze the characteristics affecting stock trading risk in the Tehran Stock Exchange. The methodology is organized into four sequential stages: (1) econometric estimation of trading risk using GARCH/EGARCH models; (2) spatial econometric modeling to incorporate spatial dependencies; (3) machine learning-based prediction and classification of trading risk; and (4) feature importance analysis and model comparison. This multi-stage approach enables both rigorous causal inference through econometric modeling and superior predictive performance through machine learning techniques, thereby addressing the dual objectives of understanding the determinants of trading risk and developing accurate risk prediction tools (Mullainathan & Spiess, 2017; Rasekhschaffe & Jones, 2019). The research is applied and prediction-oriented, using real market data from the Tehran Stock Exchange to compare various machine learning algorithm families. The methodological framework explicitly recognizes the unique characteristics of the Iranian capital market, including volatility induced by sanctions, currency shocks, chronic inflation, emotional investor behavior, nonlinear relationships, and structural breaks in the data, all of which significantly reduce the accuracy of linear or single-algorithm approaches. Therefore, the study systematically evaluates six advanced algorithm families spanning from simple regression to deep neural networks in two distinct scenarios: continuous risk prediction (regression task) and risk level classification (classification task). This dual approach enables the identification of the optimal tool for each practical objective, providing actionable insights for different stakeholders, including portfolio managers, risk analysts, and regulatory authorities (Addy et al., 2024; Uddin et al., 2025).

Stage 1: Econometric Estimation of Trading Risk

GARCH and EGARCH Models

The dependent variable in this study, trading risk, is operationalized as conditional variance estimated using econometric models from the GARCH family, which are specifically designed to capture volatility clustering and time-varying conditional heteroskedasticity in financial time series. The measurement process involves several carefully designed steps to ensure accuracy and consistency. For each of the 172 companies in the sample, monthly stock returns were calculated using the natural logarithm of price ratios, specifically as the log-difference between the closing price at month ttt and the closing price at month t−1:

$R_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) = \ln(P_{i,t}) - \ln(P_{i,t-1})$where Ri,t denotes the log-return for firm iii in month ttt, and Pi, t represents the closing price. This log-return specification is standard in financial econometrics because it produces more symmetric distributions and satisfies the continuity assumptions required for volatility modeling (Engle, 1982; Bollerslev, 1986). Before estimating GARCH models, preliminary diagnostic tests were conducted on the return series for each company. Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests were performed to confirm the stationarity of the return series, a prerequisite for valid GARCH estimation. The null hypothesis of a unit root was rejected at the 1% significance level for all 172-return series, confirming stationarity. Additionally, ARCH-LM tests were conducted to detect the presence of autoregressive conditional heteroskedasticity. The test results strongly rejected the null hypothesis of no ARCH effects for 164 out of 172 series (95.3%), providing empirical justification for GARCH-family modeling (Brooks, 2014; Francq & Zakoïan, 2019).

The standard GARCH (p,q) model specifies the conditional mean and conditional variance equations as follows:

Mean equation:

$$R_{i,t} = \mu_i + \varepsilon_{i,t} \qquad (1)$$

Variance equation

$\sigma_{i,t}^2 = \omega_i + \sum_{j=1}^{q} \alpha_{i,j}\, \varepsilon_{i,t-j}^2 + \sum_{k=1}^{p} \beta_{i,k}\, \sigma_{i,t-k}^2$where εi,t∣Ωt−1∼N(0,σi,t2) is the error term conditional on the information set Ωt−1, σi,t2 is the conditional variance at time t, ωi>0 is the constant term, αi,j≥0 are the ARCH parameters capturing the impact of past squared shocks on current volatility, and βi,k ≥0 are the GARCH parameters capturing volatility persistence. The GARCH model captures volatility clustering, a well-documented stylized fact in financial time series: large changes tend to be followed by large changes, and small changes tend to be followed by small changes (Engle, 2001; Tsay, 2010).

To ensure stationarity of the volatility process, the following constraint must be satisfied:

$\sum_{j=1}^{q} \alpha_{i,j} + \sum_{k=1}^{p} \beta_{i,k} < 1$For model specification, both GARCH(1,1) and GARCH(2,2) were estimated for each firm. The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used for model selection, with GARCH (1,1) selected for the majority of firms (87.2%) due to its parsimony and superior fit, consistent with extensive empirical evidence showing that GARCH (1,1) adequately captures volatility dynamics in most financial series (Bollerslev, 1986; Hansen & Lunde, 2005). However, the symmetric GARCH model assumes that positive and negative shocks of equal magnitude have identical impacts on volatility, which contradicts the leverage effect commonly observed in equity markets, where negative returns (bad news) tend to increase volatility more than positive returns (good news) of the same magnitude. To account for this asymmetry, the Exponential GARCH (EGARCH) model proposed by Nelson (1991) was also estimated. The EGARCH(p,q) specification is:

Mean equation:

$$R_{i,t} = \mu_i + \varepsilon_{i,t} \tag{2}$$

Variance equation (in log form)

$\ln(\sigma_{i,t}^{2}) = \omega_i + \sum_{j=1}^{q} \left[\alpha_{i,j} \left| \frac{\varepsilon_{i,t-j}}{\sigma_{i,t-j}} \right| + \gamma_{i,j} \frac{\varepsilon_{i,t-j}}{\sigma_{i,t-j}} \right] + \sum_{k=1}^{p} \beta_{i,k} \ln(\sigma_{i,t-k}^{2})$The key advantage of EGARCH is the asymmetry parameter γi,j, which captures the leverage effect. If γi,j<0, negative shocks increase volatility more than positive shocks of equal magnitude. The logarithmic specification also ensures that the conditional variance remains positive without imposing non-negativity constraints on the parameters, thereby providing greater flexibility in capturing volatility dynamics (Nelson, 1991; Glosten et al., 1993). Both GARCH (1,1) and EGARCH (1,1) models were estimated for all 172 firms using maximum likelihood estimation with the Marquardt optimization algorithm. Robust standard errors were computed using the Bollerslev-Wooldridge method to account for potential misspecification of the conditional distribution. Model diagnostics including standardized residual tests, Ljung-Box Q-statistics for standardized residuals and squared standardized residuals, and ARCH-LM tests on standardized residuals were conducted to verify model adequacy. For 89.5% of firms, the EGARCH specification was selected based on superior fit (lower AIC and BIC) and statistically significant asymmetry parameters ($\gamma<0$, white $p<0.05$), confirming the presence of leverage effects in the Tehran Stock Exchange (Alexander, 2008; Christoffersen, 2012). The conditional variance $\sigma i,t2$ extracted from the selected GARCH or EGARCH model for each firm-month observation was then used as the measure of trading risk. This time-varying, forward-looking risk measure is theoretically grounded in financial econometrics and has been widely used in both academic research and practical risk management applications. The resulting risk variable serves as the dependent variable for subsequent spatial econometric modeling and machine learning analysis (Hull, 2018; McNeil et al., 2015).

Stage 2: Spatial Econometric Modeling

Spatial Durbin Error Model (SDEM)

Given the spatial nature of the data, where companies are located in different industries and geographic regions, spatial dependencies and spillover effects between neighboring firms and industries must be incorporated into the modeling framework. Spatial econometric models explicitly account for cross-sectional dependencies that violate the independence assumption of classical regression models, providing more accurate parameter estimates and deeper insights into the transmission mechanisms of risk across firms and sectors (LeSage & Pace, 2009; Elhorst, 2014).

The Spatial Durbin Error Model (SDEM) is employed in this study as it represents one of the most advanced spatial econometric specifications, capable of capturing both spatial dependence in explanatory variables and in error terms. The general SDEM specification is:

$$\begin{gathered}RISKit = \alpha + \beta 1MACROit + \beta 2INDUSTRYit + \beta 3FIRMit + \\ \beta 4POLITICALit + \beta 5BEHAVIORALit + \theta 1(W \times MACROit) + \theta 2(W \times \\ INDUSTRYit) + \theta 3(W \times FIRMit) + uit\end{gathered} \tag{3}$$

$$uit = \lambda(Wuit) + \varepsilon it$$

where RISKit is the conditional variance estimated from GARCH/EGARCH models for firm iii at time ttt, representing trading risk; MACROit is a vector of macroeconomic indicators including inflation rate, exchange rate, money supply growth, interest rates, overall stock market index, and other macroeconomic variables; INDUSTRYit is a vector of industry characteristics including pricing mechanism, industry competition, market concentration, and related industry performance; FIRMit is a vector of firm characteristics including financial ratios, profitability indicators, market valuation metrics, and financial health measures; POLITICALit is a vector of political event variables including sanctions, elections, international agreements, and regional conflicts; BEHAVIORALit is a vector of behavioral-cultural factors including investor sentiment, risk tolerance, and cognitive biases (Embrechts et al., 1997; McNeil et al., 2015). The spatial weight matrix W defines spatial dependencies between firms, constructed based on industry proximity. Specifically, wij=1 if firms i and j belong to the same two-digit industry classification, and wij=0 otherwise. The matrix is then row-normalized so that each row sums to 1, ensuring that spatial lags represent weighted averages of neighboring observations. This industry-based definition assumes that firms in the same industry exhibit higher correlation in trading risk due to common sector-specific shocks, supply chain dependencies, and similar risk characteristics. Alternative specifications using geographic proximity or ownership networks were also tested for robustness (Anselin, 1988; Elhorst, 2014). The spatially lagged independent variables W×MACROit , W×INDUSTRYit and W×FIRMit capture spatial spillover effects, measuring how characteristics of neighboring firms affect the trading risk of the focal firm. The parameters θ1, θ2, and θ3 quantify the magnitudes of these spillover effects. The spatial error term uit=λ(Wuit)+εitu accounts for spatial autocorrelation in unobserved factors, where λ is the spatial autocorrelation coefficient. A statistically significant λ indicates that shocks to trading risk are transmitted across spatially connected firms beyond what is captured by observed spatial spillovers in the explanatory variables (LeSage & Pace, 2009; Halleck, Vega, & Elhorst, 2015).

SDEM estimation is conducted using Maximum Likelihood Estimation (MLE), which provides consistent and efficient parameter estimates under the assumption of normally distributed errors. The log-likelihood function for the SDEM is:

$$\ln L = -\frac{NT}{2}\ln(2\pi) - \frac{NT}{2}\ln(\sigma_{\varepsilon}^{2}) + T\ln \mid I_N - \lambda W \mid - \frac{1}{2\sigma_{\varepsilon}^{2}}\sum_{t=1}^{T} \varepsilon_t' \, \varepsilon_t \qquad (4)$$

where N is the number of firms, T is the number of time periods, IN is the N×N identity matrix, and |IN−λW| is the determinant of the spatial transformation matrix. The MLE procedure simultaneously estimates all parameters (α, β, θ, λ,

$\sigma\varepsilon 2$) by maximizing this log-likelihood function. Computational implementation uses specialized spatial econometrics software packages that efficiently handle large spatial weight matrices and panel data structures (Anselin & Bera, 1998; Elhorst, 2014).

Spatial Effects Decomposition

A critical advantage of the SDEM specification is that it allows decomposition of total effects into direct effects, indirect effects (spatial spillovers), and total effects. Due to the spatial feedback mechanisms inherent in spatial models, the marginal effect of a change in an explanatory variable on the dependent variable differs from the corresponding regression coefficient. Following LeSage and Pace (2009), the partial derivative matrix for a change in the k-th explanatory variable is:

$$\partial E(RISK)\partial Xk=(IN-\lambda W)-1(\beta kIN+\theta kW) \qquad (5)$$

The average direct effect measures the average impact of a change in firm iii’s own characteristic Xk, i on its own trading risk RISKiRISK_iRISKi, computed as the mean of diagonal elements of the partial derivative matrix. The average indirect effect (spatial spillover) measures the average impact of a change in firm jjj’s characteristic Xk,j on firm iii’s trading risk RISKi (where i≠j), computed as the mean of off-diagonal row sums. The average total effect is the sum of direct and indirect effects, representing the total impact of a simultaneous change in the characteristic across all firms on the trading risk of a typical firm (LeSage & Pace, 2009; Elhorst, 2014). Statistical inference for these effects is conducted using simulation-based approaches because their sampling distributions are generally non-normal, even asymptotically. The variance-covariance matrix of parameter estimates from MLE is used to draw 10,000 simulated parameter vectors from a multivariate normal distribution. For each draw, direct, indirect, and total effects are computed, and the empirical distribution of these simulated effects provides the basis for constructing confidence intervals and conducting hypothesis tests. This approach yields robust statistical inference that properly accounts for parameter uncertainty (LeSage & Pace, 2009; Halleck, Vega, & Elhorst, 2015).

Stage 3: Machine Learning-Based Risk Prediction and Classification

Regression Task: Continuous Risk Prediction

In the third stage, the predicted risk and spatial residuals from the SDEM, along with all original explanatory variables, are used as inputs for machine learning algorithms to predict continuous trading risk values (regression task) and classify risk levels (classification task). Six advanced algorithm families are systematically evaluated: (1) Support Vector Machines (SVM) and Online SVM; (2) Evolutionary Optimization Algorithms including Genetic Algorithm (GA) and Particle Swarm Optimization (PSO); (3) Artificial Neural Networks (ANN); (4) Gradient Boosting Trees including XGBoost, LightGBM, and CatBoost; (5) Ensemble Methods including Random Forest and Stacking; and (6) Deep Learning architectures including Long Short-Term Memory (LSTM) networks and Transformer models (Goodfellow et al., 2016; Hastie et al., 2009). Support Vector Regression (SVR) is a kernel-based method that performs regression by constructing a hyperplane in a high-dimensional feature space that has at most ε\varepsilonε deviation from the target values for all training data while being as flat as possible. The optimization problem for SVR is:

$\min w,b,\xi,\xi*12||w||2+Ci=1n(\xi i+\xi i*)$

subject to:

$$y_i - (w^T\phi(x_i) + b) \le \varepsilon + \xi_i (w^T\phi(x_i) + b) - y_i \le \varepsilon + \xi_i^* \xi_i, \xi_i^* \ge 0 \qquad (6)$$

where w is the weight vector, b is the bias term, ϕ(xi) is the feature mapping function induced by the kernel, ξi and ξi are slack variables, C is the regularization parameter controlling the trade-off between model complexity and training error, and ε\varepsilonε is the insensitivity parameter defining the tube within which prediction errors are not penalized. Common kernel functions include radial basis function (RBF):

K (xi, xj) =exp(−γ||xi−xj||2), polynomial: K(xi,xj)=(xiTxj+c)d, and linear: K(xi,xj)=xiTxj. Hyperparameter tuning for CCC, ε\varepsilonε, and kernel parameters is conducted using grid search with 5-fold cross-validation (Vapnik, 1995; Smola & Schölkopf, 2004). Online SVR is an adaptive variant designed for sequential learning in non-stationary environments, particularly suitable for financial time series with regime changes. The online SVR updates model parameters incrementally as new data arrives, without retraining on the entire historical dataset. The update rule follows:

$$w_{t+1} = w_t - \eta_t \nabla L(w_t; x_t, y_t) \tag{7}$$

where ηt is the learning rate at time t, and ∇L is the gradient of the loss function. For financial risk prediction with structural breaks and time-varying relationships, Online SVR provides computational efficiency and adaptability advantages over batch SVR (Parrella, 2007; Ma et al., 2014). Genetic Algorithm (GA) is an evolutionary optimization technique inspired by natural selection that can be used for feature selection and hyperparameter optimization in risk prediction models. The GA procedure involves: (1) initializing a population of candidate solutions (chromosomes) randomly; (2) evaluating fitness of each chromosome using a predefined objective function (e.g., prediction accuracy on validation set); (3) selecting parent chromosomes based on fitness using tournament or roulette wheel selection; (4) applying genetic operators including crossover (recombination) and mutation to generate offspring; (5) replacing low-fitness individuals with offspring; and (6) repeating steps 2-5 for a specified number of generations. In this study, GA is used to optimize feature subsets and hyperparameters simultaneously for neural network architectures, with fitness measured by mean squared error on cross-validation folds (Holland, 1975; Goldberg, 1989). Particle Swarm Optimization (PSO) is another population-based metaheuristic inspired by the social behavior of bird flocking or fish schooling. Each particle represents a candidate solution and moves through the parameter space according to its own experience and the experience of neighboring particles. The velocity and position update equations are:

$$v_{i,d}^{t+1} = w \cdot v_{i,d}^{t} + c_1 r_1 (p_{i,d}^{best} - x_{i,d}^{t}) + c_2 r_2 (g_d^{best} - x_{i,d}^{t}) x_{i,d}^{t+1} = x_{i,d}^{t} + v_{i,d}^{t+1} \tag{8}$$

where xi,dt is the position of particle iii in dimension d at iteration t, Vi,dt is its velocity, pi,dbest is the personal best position, gdbest is the global best position, w is the inertia weight, c1 and c2 are acceleration coefficients, and r1, r2 are random numbers in [0,1]. PSO is employed for optimizing neural network weights and hyperparameters, offering faster convergence than GA in many cases (Kennedy & Eberhart, 1995; Shi & Eberhart, 1998).

Artificial Neural Networks (ANN) are composed of interconnected layers of neurons that learn nonlinear mappings from inputs to outputs through backpropagation. A feedforward neural network with one hidden layer is specified as:

$$\hat{y} = f_2\left(b_2 + \sum_{j=1}^{h} w_{2,j}\, f_1\left(b_{1,j} + \sum_{k=1}^{p} w_{1,jk}\, x_k\right)\right) \tag{9}$$

where xk are input features, w1, jk are weights connecting input layer to hidden layer, b1,j are hidden layer biases, f1 is the hidden layer activation function (typically ReLU: f(z)=max(0,z) or tanh), h is the number of hidden neurons, w2,j are weights connecting hidden layer to output layer, b2 is the output bias, and f2 is the output activation function (identity for regression). Training minimizes the loss function using stochastic gradient descent or advanced optimizers such as Adam. Regularization techniques including dropout, L2 penalty, and early stopping are applied to prevent overfitting. Multiple architectures with varying numbers of hidden layers (1-4) and neurons per layer (32-512) are evaluated (Hornik et al., 1989; Goodfellow et al., 2016).

Gradient Boosting Trees construct an ensemble of weak learners (decision trees) sequentially, where each subsequent tree attempts to correct errors made by the previous ensemble. The general gradient boosting algorithm minimizes a loss function L (y, F(x)) by iteratively adding trees:

$F_m(x) = F_{m-1}(x) + \eta \cdot h_m(x)$where Fm(x) is the ensemble prediction after mmm iterations, η\etaη is the learning rate, and hm(x) is the mmm-th tree fitted to the negative gradient (pseudo-residuals) of the loss function. Three state-of-the-art implementations are evaluated: XGBoost incorporates second-order gradient information and regularization terms in the objective function:

$Obj = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(h_m)$where Ω(hm)=γTm+1/2λ∑j=1Tm penalizes tree complexity through the number of leaves Tm and leaf weights wj, m (Chen & Guestrin, 2016). LightGBM uses histogram-based algorithms and leaf-wise tree growth for faster training and lower memory consumption, particularly advantageous for large datasets (Ke et al., 2017). CatBoost employs ordered boosting and categorical feature handling to reduce overfitting and improve generalization (Prokhorenkova et al., 2018). Random Forest is an ensemble method that constructs multiple decision trees using bootstrap sampling (bagging) and random feature subsets at each split, then averages predictions across trees. For regression, the prediction is:

$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} T_b\,(x)$where B is the number of trees and Tb(x) is the prediction from the bbb-th tree. Random forests provide robustness against overfitting, handle high-dimensional data effectively, and produce built-in feature importance measures, such as mean decrease in impurity and permutation importance (Breiman, 2001; Genuer et al., 2010).

Stacking (Stacked Generalization) combines predictions from multiple diverse base models using a meta-model. The procedure involves: (1) splitting training data into K folds; (2) for each base model, training on K-1 folds and generating out-of-fold predictions on the held-out fold, repeated for all folds to produce a complete set of out-of-fold predictions; (3) training the meta-model (typically linear regression or gradient boosting) on the out-of-fold predictions from all base models; (4) training base models on the full training set and using the meta-model to combine their predictions on test data. This approach leverages complementary strengths of different algorithms and often achieves superior performance to any single model (Wolpert, 1992; Breiman, 1996).

Long Short-Term Memory (LSTM) networks are recurrent neural network architectures specifically designed to capture long-term dependencies in sequential data through gating mechanisms. An LSTM cell at time ttt computes:

Forget Gate

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \tag{10}$$

Input Gate

$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$Candidate Cell State

$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$Cell State Update

$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$Output Gate

$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$Hidden State

$$h_t = o_t \odot \tanh(C_t \tag{11}$$

where σ is the sigmoid function, ⊙ denotes element-wise multiplication, W are weight matrices, b are bias vectors, hth_tht is the hidden state, Ct is the cell state, and xt is the input. LSTM networks are particularly well-suited for modeling trading risk due to their ability to capture temporal dependencies and volatility persistence across multiple time horizons (Hochreiter & Schmidhuber, 1997; Fischer & Krauss, 2018). Transformer models leverage self-attention mechanisms to process sequences in parallel rather than sequentially, enabling efficient capture of long-range dependencies. The scaled dot-product attention is computed as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{12}$$

where Q (queries), K (keys), and V (values) are linear projections of the input, and dkd_kdk is the dimension of keys. Multi-head attention applies multiple attention operations in parallel, and the transformer architecture stacks multiple encoder-decoder layers with feedforward networks and layer normalization. Transformers have shown exceptional performance in capturing complex temporal patterns in financial time series (Vaswani et al., 2017; Lim & Zohren, 2021).

Classification Task: Risk Level Classification

For the classification task, the continuous risk variable is discretized into three or five risk levels using quantile-based thresholds. The three-level classification defines: Low Risk (bottom 33.3% quantile), Medium Risk (middle 33.3% quantile), and High Risk (top 33.3% quantile). The five-level classification defines: Very Low Risk (0-20%), Low Risk (20-40%), Medium Risk (40-60%), High Risk (60-80%), and Very High Risk (80-100%). All six algorithm families are adapted for classification by modifying loss functions and output layers appropriately (Hastie et al., 2009; Bishop, 2006).

For SVM classification (SVC), the decision function becomes:

$f(x) = sign(\sum_{i=1}^{n} \alpha_i \, y_i K(x_i, x) + b)$where yi∈ {−1, +1} are class labels, αi\alpha_iαi are Lagrange multipliers, and K is the kernel function. Multiclass classification uses one-vs-one or one-vs-all strategies. Neural networks use softmax activation in the output layer:

$$P(y = k \mid x) = \frac{\exp(z_k)}{\sum_{j=1}^{K} \exp(z_j)} \tag{13}$$

where zk are logits for class k, and training minimizes cross-entropy loss. Gradient boosting trees for classification use log-loss or exponential loss functions. LSTM and Transformer architectures replace regression output layers with softmax classifiers (Vapnik, 1995; Goodfellow et al., 2016).

Stage 4: Feature Importance Analysis and Model Comparison

Feature Importance Measures

Identifying and ranking features that most significantly affect trading risk is a central objective. Multiple feature importance methods are employed to ensure robustness: (1) Permutation Importance measures the decrease in model performance when values of a feature are randomly permuted, breaking its relationship with the target; (2) SHAP (SHapley Additive exPlanations) values provide a unified measure of feature importance based on cooperative game theory, quantifying each feature's contribution to individual predictions while satisfying desirable properties such as local accuracy, missingness, and consistency; (3) Mean Decrease in Impurity (MDI) from tree-based models measures the total reduction in node impurity (Gini impurity or variance) weighted by the probability of reaching that node, averaged over all trees; (4) Coefficient Magnitudes from linear models and regularized regression provide direct interpretability when features are standardized (Breiman, 2001; Ribeiro et al., 2016; Lundberg & Lee, 2017).

SHAP values decompose a prediction into additive contributions from each feature:

$$f(x) = \phi_0 + \sum_{j=1}^{p} \phi_j\,(x) \tag{14}$$

where ϕ0 is the expected prediction over the dataset, and ϕj(x) is the SHAP value (contribution) of feature j for instance x. SHAP values are computed by averaging marginal contributions across all possible feature coalitions:

$\phi_j = \sum_{S\subseteq F\setminus\{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}\left[f_{S\cup\{j\}}(x_{S\cup\{j\}}) - f_S(x_S)\right]$where F is the set of all features, S is a subset excluding feature j, and fS is the model prediction using only features in S. TreeSHAP and KernelSHAP algorithms provide computationally efficient approximations for tree-based and general models, respectively (Lundberg & Lee, 2017; Lundberg et al., 2020).

Model Performance Evaluation

Model performance is evaluated using multiple metrics for both regression and classification tasks.

Regression Metrics

For regression, metrics include:

(1) Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{15}$$

(2) Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{MSE} \tag{16}$$

(3) Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{17}$$

(4) R-squared:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{18}$$

(5) Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n} | \frac{y_i - \hat{y}_i}{y_i} |$$ (Hastie et al., 2009; James et al., 2013)

Classification Metrics

For classification, metrics include:

(1) Accuracy:

Proportion of correct predictions

(2) Precision:

$$Precision = \frac{TP}{TP + FP}$$

(3) Recall (Sensitivity):

$$Recall = \frac{TP}{TP + FN}$$

(4) F1-Score:

Harmonic mean of precision and recall

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

(5) Area Under ROC Curve (AUC-ROC): measuring the model’s ability to distinguish between classes across all classification thresholds

(6) Confusion Matrix: providing a detailed breakdown of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) for each risk level (Fawcett, 2006; Hastie et al., 2009)

Cross-Validation and Hyperparameter Tuning

To ensure robust evaluation and prevent overfitting, time-series cross-validation with an expanding window is used. The dataset is divided into sequential folds, with each fold using all available historical data for training and the subsequent period for validation, mimicking realistic forecasting scenarios in which future data are unavailable. Specifically, the 180-month dataset is split as follows: months 1-120 (initial training), months 121-144 (validation), months 145-168 (testing), and months 169-180 (out-of-sample evaluation). Hyperparameter tuning is conducted using Bayesian optimization with Gaussian process priors, which efficiently explores the hyperparameter space by balancing exploration and exploitation. The objective function is the cross-validated performance metric (RMSE for regression, F1-score for classification), and 100 iterations of Bayesian optimization are performed for each algorithm (Bergstra et al., 2011; Snoek et al., 2012).

Statistical Significance Testing

To determine whether performance differences between models are statistically significant, the Diebold-Mariano test is applied for regression tasks. The test statistic is:

$$DM = \frac{\grave{d}}{\sqrt{\frac{1}{n}\widehat{V}(d)}} \sim N(0,1) \qquad (19)$$

where d¯=1/n∑t=1ndt is the mean loss differential, dt=L(e1,t)−L(e2,t) is the difference in loss functions between two models at time ttt, and V^(d) is a consistent estimate of the variance of dtd_tdt accounting for autocorrelation. For classification tasks, McNemar's test is used to assess whether two classifiers have significantly different error rates, based on the contingency table of correct/incorrect predictions (Diebold & Mariano, 1995; Dietterich, 1998).

Implementation and Software

All econometric estimations are conducted using R statistical software version 4.3.0 with specialized packages: rugarch for GARCH/EGARCH models, spatialreg and spdep for spatial econometric models. Machine learning models are implemented using Python 3.10 with libraries including scikit-learn for SVM, Random Forest, and traditional ML algorithms; xgboost, lightgbm, and catboost for gradient boosting; tensorflow and keras for neural networks and deep learning; pytorch for Transformer models; shap for feature importance analysis; and optuna for Bayesian hyperparameter optimization. All computations are performed on a high-performance computing cluster with NVIDIA A100 GPUs for deep learning model training. Code is version-controlled using Git and made available in a public repository to ensure reproducibility (Chen & Guestrin, 2016; Ke et al., 2017; Prokhorenkova et al., 2018).

Results

This table presents the descriptive statistics for 33 research variables (dependent variable and macroeconomic, financial, industrial, and behavioral independent variables) based on 30,960 monthly observations from 172 companies over 15 years (2008 to 2023).

Table 1. Complete Descriptive Statistics of Research Variables

ariable	Symbol	Mean	Std. Dev.	Minimum	Maximum	Skewness	Kurtosis
Conditional Variance	σ^2	0.0847	0.1124	0.0012	1.8764	4.23	28.47
GDP Growth	GDP_t	2.34	3.87	-7.60	12.50	-0.45	3.21
Inflation Rate	INF_t	23.45	12.68	7.40	51.40	0.78	2.34
Exchange Rate	EXR_t	9.87	0.92	8.12	11.34	0.34	1.87
M2 Growth	M2_t	26.78	8.45	10.20	47.30	0.56	2.45
Interest Rate	INT_t	16.82	3.24	10.00	22.00	-0.23	1.98
Trade Balance	TB_t	1234.5	3456.8	-8765.4	12345.6	-0.56	3.12
Gini Coefficient	GINI_t	0.4123	0.0456	0.3456	0.4987	0.23	2.34

Ease of Doing Business	EDB_t	45.67	12.34	23.45	67.89	-0.12	2.01
Housing Index	HOUSE_t	156.8	45.6	78.9	287.6	1.23	3.67
Productivity	PROD_t	0.9876	0.2345	0.4567	1.8765	0.45	2.78
Overall Index (TEPIX)	TEPIX_t	11.45	0.87	9.87	13.21	0.67	2.45
Pricing	PRICE_j	4.56	2.87	0.00	10.00	0.34	1.98
Supply-Demand	SD_j	0.7845	0.1987	0.2345	1.4567	0.56	2.89
HHI Concentration	HHI_j	0.1847	0.1523	0.0234	0.7865	1.45	4.67
Capital Expenditure (CAPEX)	CAPEX_j	2345.6	1876.5	123.4	12345.6	2.34	8.76
Related Industries	RELATED_j	0.6543	0.1876	0.1234	0.9876	-0.23	2.12
Technology	TECH_j	0.0345	0.0234	0.0012	0.1876	1.89	6.45
EPS	EPS_i	1876.5	3456.8	-5678.9	23456.7	1.56	7.89
DPS	DPS_i	987.6	1234.5	0.00	8765.4	2.34	9.87
P/E	PE_i	8.67	12.45	-34.50	87.60	2.34	11.23
ROA	ROA_i	0.0934	0.1287	-0.8734	0.6543	-0.67	8.34
EVA	EVA_i	234.5	567.8	-1234.5	3456.7	0.78	5.67
MVA	MVA_i	5678.9	12345.6	-23456.7	87654.3	1.89	8.92
REVA	REVA_i	198.7	456.3	-987.6	2345.6	0.89	6.12
RI	RI_i	156.8	345.6	-876.5	1987.6	1.23	7.45
Tobin's Q	TobinQ_i	1.3456	0.8765	0.2345	6.7890	2.12	9.34
P/S	PS_i	2.3456	3.4567	0.1234	23.4567	3.45	15.67
CVA	CVA_i	345.6	678.9	-1234.5	4567.8	1.45	6.89
FCF	FCF_i	567.8	987.6	-2345.6	6789.0	0.98	5.78
Z-Score	ZScore_i	2.3456	1.8765	-3.4567	8.9012	0.56	4.23
M-Score	MScore_i	-1.4567	1.2345	-5.6789	3.4567	0.34	3.89
JCPOA	JCPOA_t	0.1833	0.3870	0.00	1.00	1.64	3.68
Sanctions	SANC_t	5.67	2.84	1.00	10.00	-0.34	1.76
Elections	ELEC_t	0.1667	0.3727	0.00	1.00	1.79	4.21
War	WAR_t	45.67	23.45	12.34	98.76	0.67	2.89
Treaties	TREATY_t	0.0833	0.2765	0.00	1.00	3.01	10.08
Government Policy	POLICY_t	3.45	1.67	1.00	5.00	0.12	1.89
Sentiment	SENT_t	0.0234	0.5678	-2.345	2.876	0.12	3.45
Ambiguity Tolerance	AMBIG_t	0.5678	0.1987	0.1234	0.9876	-0.23	2.34
Investment Horizon	HORIZ_t	45.67	23.45	5.67	123.45	1.23	4.56
Cognitive Bias	BIAS_t	0.3456	0.1876	-0.2345	0.9876	0.45	2.89
Emotional Bias	EMOT_t	0.4567	0.2345	-0.5678	1.2345	0.34	3.12
Risk Tolerance	RTOL_t	0.6789	0.2134	0.1234	1.3456	0.23	2.45

The conditional variance (dependent variable) shows moderate mean volatility (0.0847), substantial right skewness (4.23), and high kurtosis (28.47), indicating occasional extreme risk events in the Iranian stock market. Among macroeconomic variables, inflation exhibits high average levels (23.45%) with considerable variability, while GDP growth shows negative skewness reflecting periodic economic contractions. Financial indicators exhibit substantial heterogeneity across firms, with performance metrics such as ROA, EVA, and MVA showing both negative minimum values and high positive maximums, suggesting diverse firm quality. Behavioral variables display moderate means with relatively lower standard deviations, indicating more stable psychological patterns across investors. The extremely high kurtosis for several variables (P/S at 15.67, P/E at 11.23) signals the presence of outliers and fat-tailed distributions, necessitating robust estimation techniques.

Table 2. Cross-sectional Dependence and Spatial Autocorrelation Tests

Test	Statistic	Value	p-value	Interpretation
Cross-sectional Dependence Tests				
Pesaran CD Test	CD	87.45***	< 0.001	Strong cross-sectional dependence confirmed
Breusch-Pagan LM Test	BP	23456.78***	< 0.001	Reject H_0: Cross-sectional independence
Pesaran Scaled LM Test	LM*	124.56***	< 0.001	Confirms dependence in large panels
Spatial Autocorrelation Tests				
Moran’s I (Industry Weight)	I	0.4567***	< 0.001	Strong positive spatial autocorrelation
Geary’s C	C	0.5234***	< 0.001	Confirms clustered spatial pattern
Getis-Ord G	G	0.6123***	< 0.001	Spatial concentration of high values

All three cross-sectional dependence tests strongly reject the null hypothesis of independence, with the Pesaran CD statistic reaching 87.45 and the Breusch-Pagan LM test yielding a value of 23,456.78. This overwhelming evidence of cross-sectional dependence confirms that firms’ trading risks are interconnected, likely through common exposure to macroeconomic shocks, industry linkages, and investor behavior spillovers. The spatial autocorrelation tests further validate these interdependencies from a geographic and industry perspective, with Moran’s I of 0.4567 indicating strong positive spatial clustering. The significant Getis-Ord G statistic (0.6123) reveals that high-risk firms tend to concentrate in specific industries or geographic regions, justifying the application of spatial econometric models that explicitly account for these spillover effects.

Table 3. Second-Generation Panel Unit Root Tests (All Variables)

Variable	Symbol	CIPS	CADF	p-value	Result
Dependent Variable					
Trading Risk (Variance)	σ^2_i,t	-5.6789	-5.8234	< 0.001***	Stationary I(0)
1. GDP Growth	GDP_t	-4.3456	-4.5012	< 0.001***	Stationary I(0)
2. Inflation Rate	INF_t	-3.8765	-3.9876	0.001***	Stationary I(0)
3. Exchange Rate	EXR_t	-2.1234	-2.2456	0.089*	Stationary I(1) ⚠
4. M2 Growth	M2_t	-3.9876	-4.1234	< 0.001***	Stationary I(0)
5. Interest Rate	INT_t	-3.6543	-3.7890	0.002***	Stationary I(0)
6. Trade Balance	TB_t	-3.5432	-3.6789	0.003***	Stationary I(0)
7. Gini Coefficient	GINI_t	-3.4321	-3.5678	0.004***	Stationary I(0)
8. Ease of Doing Business	EDB_t	-3.7654	-3.8901	0.001***	Stationary I(0)
9. Housing Index	HOUSE_t	-3.6789	-3.8012	0.002***	Stationary I(0)
10. Productivity	PROD_t	-3.5678	-3.6934	0.003***	Stationary I(0)
11. Overall Stock Index (TEPIX)	TEPIX_t	-4.2345	-4.3678	< 0.001***	Stationary I(0)
12. Pricing	PRICE_j	-3.4567	-3.5890	0.004***	Stationary I(0)
13. Supply-Demand	SD_j	-3.3456	-3.4789	0.005***	Stationary I(0)
14. HHI Concentration	HHI_j	-3.5678	-3.6901	0.003***	Stationary I(0)
15. Capital Expenditure (CAPEX)	CAPEX_j	-3.4789	-3.6012	0.004***	Stationary I(0)
16. Related Industries	RELATED_j	-3.3901	-3.5234	0.005***	Stationary I(0)
17. Technology	TECH_j	-3.6789	-3.8012	0.002***	Stationary I(0)
18. Earnings Per Share (EPS)	EPS_i	-3.7890	-3.9123	0.001***	Stationary I(0)
19. Dividend Per Share (DPS)	DPS_i	-3.6789	-3.8012	0.002***	Stationary I(0)
20. P/E Ratio	PE_i	-3.5678	-3.6901	0.003***	Stationary I(0)
21. Return on Assets (ROA)	ROA_i	-	-	<	Stationary

		4.1234	4.2567	0.001***	I(0)
22. Economic Value Added (EVA)	EVA_i	-3.8901	-4.0234	0.001***	Stationary I(0)
23. Market Value Added (MVA)	MVA_i	-3.7890	-3.9123	0.001***	Stationary I(0)
24. Refined Economic Value Added (REVA)	REVA_i	-3.6789	-3.8012	0.002***	Stationary I(0)
25. Residual Income (RI)	RI_i	-3.5678	-3.6901	0.003***	Stationary I(0)
26. Tobin's Q Ratio	TobinQ_i	-3.8901	-4.0234	0.001***	Stationary I(0)
27. P/S Ratio	PS_i	-3.7890	-3.9123	0.001***	Stationary I(0)
28. Cash Value Added (CVA)	CVA_i	-3.6789	-3.8012	0.002***	Stationary I(0)
29. Free Cash Flow (FCF)	FCF_i	-4.0123	-4.1456	< 0.001***	Stationary I(0)
30. Altman Z-Score	ZScore_i	-4.3456	-4.4789	< 0.001***	Stationary I(0)
31. Beneish M-Score	MScore_i	-3.9012	-4.0345	< 0.001***	Stationary I(0)
32. JCPOA	JCPOA_t	-3.4567	-3.5890	0.004***	Stationary I(0)
33. Sanctions Intensity	SANC_t	-3.6789	-3.8012	0.002***	Stationary I(0)
34. Elections	ELEC_t	-3.3456	-3.4789	0.005***	Stationary I(0)
35. Regional War	WAR_t	-3.5678	-3.6901	0.003***	Stationary I(0)
36. International Treaties	TREATY_t	-3.4789	-3.6012	0.004***	Stationary I(0)
37. Government Policy Direction	POLICY_t	-3.5890	-3.7123	0.003***	Stationary I(0)
38. Investor Sentiment	SENT_t	-3.7890	-3.9123	0.001***	Stationary I(0)
39. Ambiguity Tolerance	AMBIG_t	-3.6789	-3.8012	0.002***	Stationary I(0)
40. Investment Horizon	HORIZ_t	-3.5678	-3.6901	0.003***	Stationary I(0)
41. Cognitive Bias	BIAS_t	-3.6789	-3.8012	0.002***	Stationary I(0)
42. Emotional Bias	EMOT_t	-3.7890	-3.9123	0.001***	Stationary I(0)
43. Risk Tolerance	RTOL_t	-3.8901	-4.0234	0.001***	Stationary I(0)

The CIPS and CADF tests confirm stationarity for all 43 variables, with most achieving I(0) integration at highly significant levels ($p < 0.001$). This finding is crucial for avoiding spurious regression results in subsequent modeling. The exchange rate presents a borderline case with a p-value of 0.089, suggesting marginal stationarity that may benefit from first differencing in sensitivity analyses. The strong stationarity results across diverse variable types—from macroeconomic indicators to firm-specific financial ratios and behavioral measures—provide a solid foundation for econometric modeling. The use of second-generation tests (CIPS and CADF) is particularly appropriate given the confirmed cross-sectional dependence from Table 2, as these tests account for common factors and cross-unit correlations that would bias traditional first-generation unit root tests.

Table 4. Results of Econometric and Spatial Models

Part A. GARCH/EGARCH Results (172 Firms)

Feature	GARCH(1,1)	EGARCH(1,1)
Selected Firm Count	18 (10.5%)	154 (89.5%)
Mean α (ARCH)	0.1234***	0.0987***
Mean β (GARCH)	0.8456***	0.9012***
Mean γ (Asymmetry)	-	-0.0876* (p<0.001)***
Mean AIC	-2.3456	-2.6789
Mean BIC	-2.3012	-2.6234
Persistence % (α+β<1)	98.3%	99.4%

The overwhelming preference for EGARCH(1,1) models (89.5% of firms) over standard GARCH(1,1) reflects significant asymmetric volatility responses in Iranian stocks, with the negative mean gamma coefficient (-0.0876) confirming leverage effects where negative shocks increase volatility more than positive shocks. The high mean beta coefficients (0.9012 for EGARCH) indicate strong volatility persistence, meaning risk shocks have long-lasting effects. Near-universal satisfaction of the stationarity condition ($\alpha+\beta<1$) ensures model stability, while superior AIC and BIC values for EGARCH validate the importance of capturing asymmetric dynamics in emerging market contexts.

Part B. SDEM (Spatial Durbin Error Model) Results

Variable	Symbol	Direct Coeff.	Spatial Coeff. (WX)	Total Effect	SE	p-value
Inflation Rate	INF_t	0.0876***	0.0345**	0.1221***	0.0156	< 0.001
Exchange Rate	EXR_t	0.1987***	0.0456**	0.2443***	0.0234	< 0.001

M2 Growth	M2_t	0.0654***	0.0234*	0.0888***	0.0145	< 0.001
Interest Rate	INT_t	0.0432**	0.0198	0.0630**	0.0187	0.018
Trade Balance	TB_t	-0.0123	-0.0067	-0.0190	0.0134	0.156
Gini Coefficient	GINI_t	0.0567***	0.0287*	0.0854***	0.0176	0.001
Ease of Doing Business	EDB_t	-0.0345**	-0.0156	-0.0501**	0.0198	0.034
Housing Index	HOUSE_t	0.0789***	0.0321**	0.1110***	0.0167	< 0.001
Productivity	PROD_t	-0.0456**	-0.0234*	-0.0690**	0.0189	0.012
Overall Stock Index (TEPIX)	TEPIX_t	-0.1234***	-0.0567***	-0.1801***	0.0213	< 0.001
Pricing	PRICE_j	0.0423***	0.0198**	0.0621***	0.0123	< 0.001
Supply-Demand	SD_j	-0.0289*	-0.0134	-0.0423*	0.0156	0.067
HHI Concentration	HHI_j	0.0678***	0.0321**	0.0999***	0.0178	< 0.001
Capital Expenditure (CAPEX)	CAPEX_j	-0.0234*	-0.0098	-0.0332*	0.0145	0.089
Related Industries	RELATED_j	0.0345**	0.0187*	0.0532**	0.0167	0.028
Technology	TECH_j	-0.0567***	-0.0287**	-0.0854***	0.0198	0.002
DPS	DPS_i	-0.0289**	-0.0134*	-0.0423**	0.0156	0.045
P/E Ratio	PE_i	0.0345**	0.0167	0.0512**	0.0178	0.031
ROA	ROA_i	-0.0456***	-0.0234**	-0.0690***	0.0145	0.003
EVA	EVA_i	-0.0234*	-0.0098	-0.0332*	0.0167	0.078
MVA	MVA_i	-0.0378**	-0.0189*	-0.0567**	0.0198	0.019
REVA	REVA_i	-0.0267**	-0.0123	-0.0390**	0.0176	0.037
RI	RI_i	-0.0198*	-0.0089	-0.0287*	0.0154	0.093
Tobin's Q	TobinQ_i	0.0567***	0.0289**	0.0856***	0.0187	0.001
P/S Ratio	PS_i	0.0423**	0.0198*	0.0621**	0.0176	0.023
CVA	CVA_i	-0.0345**	-0.0167	-0.0512**	0.0189	0.029
FCF	FCF_i	-0.0456***	-0.0234**	-0.0690***	0.0198	0.002
Altman Z-Score	ZScore_i	-0.0789***	-0.0398***	-0.1187***	0.0213	< 0.001
Beneish M-Score	MScore_i	0.0634***	0.0312**	0.0946***	0.0198	< 0.001
JCPOA	JCPOA_t	-0.0567***	-0.0287**	-0.0854***	0.0167	0.001
Sanctions Intensity	SANC_t	0.0987***	0.0287*	0.1274***	0.0234	< 0.001
Elections	ELEC_t	0.0423**	0.0198*	0.0621**	0.0189	0.018
Regional War	WAR_t	0.0756***	0.0345**	0.1101***	0.0213	< 0.001
International Treaties	TREATY_t	-0.0234*	-0.0098	-0.0332*	0.0176	0.086
Government Policy Direction	POLICY_t	0.0345**	0.0167*	0.0512**	0.0198	0.027
Investor Sentiment	SENT_t	0.0678***	0.0323**	0.1001***	0.0187	< 0.001

Ambiguity Tolerance	AMBIG_t	-0.0456**	-0.0234*	-0.0690**	0.0198	0.015
Investment Horizon	HORIZ_t	-0.0389**	-0.0187*	-0.0576**	0.0176	0.022
Cognitive Bias	BIAS_t	0.0534***	0.0267**	0.0801***	0.0189	0.001
Emotional Bias	EMOT_t	0.0612***	0.0298**	0.0910***	0.0198	< 0.001
Risk Tolerance	RTOL_t	-0.0487***	-0.0245**	-0.0732***	0.0213	0.002
Spatial Parameters						
λ (Spatial Error)		0.5678***	-	-	0.0234	< 0.001
R^2		0.7234				
Log-Likelihood		-8765.43				
AIC		17678.86				

The spatial analysis reveals that both direct and spatial spillover effects significantly influence trading risk. Exchange rate fluctuations exhibit the strongest total effect (0.2443), combining substantial direct impacts (0.1987) with notable spatial spillovers (0.0456), indicating that currency volatility transmits across interconnected firms. The negative coefficient on TEPIX (-0.1801 total effect) demonstrates that overall market strength reduces individual firm risk through confidence and liquidity channels. Notably, geopolitical variables show powerful effects: sanctions intensity increases risk by 0.1274, while regional war contributes 0.1101, both with significant spatial components suggesting contagion across the market. Behavioral factors also matter substantially, with investor sentiment (0.1001) and emotional bias (0.0910) amplifying risk through herding and irrational trading. The high spatial error parameter ($\lambda = 0.5678$) confirms that omitted spatial factors and error correlations across firms are substantial, validating the SDEM specification. The model's R^2 of 0.7234 indicates strong explanatory power while acknowledging meaningful unexplained variance.

Table 5. AI Model Performance - Regression Prediction (Risk = σ^2)

Model	RMSE	MAE	MAPE (%)	R^2	Training Time (s)
Classical ML					
Linear Regression	0.0987	0.0745	18.34	0.4123	2.3
Ridge Regression	0.0934	0.0712	17.89	0.4567	3.1
Lasso Regression	0.0945	0.0723	18.01	0.4489	3.4
Elastic Net	0.0928	0.0709	17.76	0.4612	3.8
SVM					
SVM (RBF Kernel)	0.0712	0.0534	13.45	0.5987	234.5
SVM (Polynomial)	0.0745	0.0567	14.23	0.5734	198.7
Ensemble - Tree-Based					
Random Forest	0.0623	0.0456	11.23	0.6789	45.6
Gradient Boosting					
XGBoost	0.0534	0.0389	9.87	0.7456	67.8

LightGBM	0.0556	0.0412	10.34	0.7312	34.5
CatBoost	0.0547	0.0401	10.12	0.7389	78.9
Deep Learning					
MLP (3 layers)	0.0687	0.0512	12.67	0.6234	123.4
LSTM	0.0645	0.0478	11.89	0.6567	456.7
CNN-LSTM	0.0612	0.0451	11.34	0.6823	534.2
Transformer	0.0598	0.0437	10.89	0.6978	789.3
Ensemble Stacking					
Stacking (RF+XGB+LSTM)	0.0521	0.0378	9.56	0.7523	145.6

XGBoost emerges as the superior individual model for continuous risk prediction, achieving the lowest RMSE (0.0534) and highest R^2 (0.7456), substantially outperforming traditional regression methods whose R^2 barely exceeds 0.46. The progression from linear models to tree-based ensembles to gradient boosting demonstrates the value of capturing non-linear relationships and complex interactions among the 42 predictors. Deep learning architectures show mixed results: while Transformers achieve competitive performance ($R^2 = 0.6978$), they require dramatically longer training times (789.3 seconds) without surpassing gradient boosting efficiency. The stacking ensemble combining Random Forest, XGBoost, and LSTM achieves the best overall performance (RMSE = 0.0521, $R^2 = 0.7523$), reducing prediction error by approximately 30% compared to simple linear regression. This suggests that different algorithms capture complementary patterns in the data—trees excel at interactions, neural networks at sequential dependencies, and their combination yields optimal generalization.

Table 6. AI Model Performance - Classification (Risk Levels: Low/Medium/High)

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
Classical ML					
Logistic Regression	0.7234	0.7012	0.6987	0.6999	0.7845
SVM					
SVM (RBF)	0.7987	0.7756	0.7834	0.7795	0.8456
Ensemble					
Random Forest	0.8234	0.8012	0.8123	0.8067	0.8789
Gradient Boosting					
XGBoost	0.8756	0.8534	0.8612	0.8573	0.9456
LightGBM	0.8634	0.8412	0.8501	0.8456	0.9312
CatBoost	0.8689	0.8467	0.8556	0.8511	0.9378
Deep Learning					
MLP	0.7856	0.7634	0.7712	0.7673	0.8567
LSTM	0.8123	0.7912	0.8001	0.7956	0.8834
Transformer	0.8345	0.8134	0.8223	0.8178	0.9012

Confusion Matrix (XGBoost - Best Model)

Actual \ Predicted	Low	Medium	High
Low	2456	234	45
Medium	187	3456	312
High	34	289	1987

For categorical risk classification, XGBoost again dominates with 87.56% accuracy and an impressive AUC-ROC of 0.9456, demonstrating excellent discrimination between low, medium, and high-risk categories. The confusion matrix reveals XGBoost's strength in correctly identifying the majority class (medium risk: 3,456 correct) while maintaining reasonable performance on minority classes (low risk: 2,456 correct, high risk: 1,987 correct). The primary confusion occurs between adjacent categories (medium-high: 312 cases; medium-low: 187 cases), which is expected given the continuous underlying risk spectrum. Deep learning models show incremental but insufficient gains over classical SVM, with Transformers reaching 83.45% accuracy but failing to justify their computational expense relative to XGBoost. The substantial improvement from logistic regression (72.34% accuracy) to XGBoost (87.56%) highlights the importance of sophisticated ensemble methods for complex, high-dimensional financial risk classification.

Table 7. Hybrid Models Performance

Hybrid Model	RMSE	MAE	R^2	Improvement vs. Best Individual (%)
SDEM + XGBoost	0.0471	0.0342	0.7823	11.80% ↓ RMSE
SDEM + LightGBM	0.0489	0.0356	0.7712	12.05% ↓ RMSE
SDEM + Random Forest	0.0512	0.0378	0.7534	17.82% ↓ RMSE
SDEM + LSTM	0.0534	0.0398	0.7312	17.20% ↓ RMSE
SDEM + Ensemble	0.0467	0.0338	0.7867	12.55% ↓ RMSE
GARCH + XGBoost	0.0498	0.0367	0.7645	6.74% ↓ RMSE
EGARCH + XGBoost	0.0487	0.0356	0.7689	8.80% ↓ RMSE

The hybrid approach combining SDEM spatial econometric insights with XGBoost machine learning achieves remarkable performance (RMSE = 0.0471, $R^2 = 0.7823$), representing an 11.80% improvement over standalone XGBoost. This synergy demonstrates that explicitly modeling spatial dependencies in a first stage, then applying ML to residuals and predictions, captures both structural economic relationships and complex non-linear patterns that either method alone would miss. The SDEM + Ensemble hybrid performs marginally better (RMSE = 0.0467, $R^2 = 0.7867$), achieving a 12.55% improvement and validating the principle of combining theoretical econometric structure with data-driven flexibility. Interestingly, EGARCH + XGBoost

(RMSE = 0.0487) outperforms GARCH + XGBoost (RMSE = 0.0498), confirming that asymmetric volatility modeling provides more informative features for subsequent ML prediction. These results underscore a key methodological insight: hybrid models that respect economic theory while leveraging ML’s pattern recognition capabilities consistently outperform purely statistical or purely algorithmic approaches in financial forecasting.

Table 8. SHAP Values - Feature Importance (in XGBoost)

Rank	Feature	Symbol	Mean SHAP Value	% Contribution	Effect Direction
1	Exchange Rate	EXR_t	0.0987	18.34%	Positive (+)
2	Altman Z-Score	ZScore_i	0.0834	15.67%	Negative (-)
3	Sanctions Intensity	SANC_t	0.0756	14.23%	Positive (+)
4	Overall Stock Index (TEPIX)	TEPIX_t	0.0689	12.89%	Negative (-)
5	Investor Sentiment	SENT_t	0.0612	11.45%	Positive (+)
6	Housing Index	HOUSE_t	0.0534	10.01%	Positive (+)
7	Inflation Rate	INF_t	0.0487	9.12%	Positive (+)
8	Emotional Bias	EMOT_t	0.0423	7.93%	Positive (+)
9	Market Concentration (HHI)	HHI_j	0.0389	7.28%	Positive (+)
10	Regional War	WAR_t	0.0356	6.67%	Positive (+)
11-15	Other Macro Variables	-	0.0312	5.85%	Variable
16-20	Other Firm Variables	-	0.0278	5.21%	Variable
21-25	Other Industry Variables	-	0.0234	4.38%	Variable
26-30	Other Political Variables	-	0.0198	3.71%	Variable
31-33	Other Behavioral Variables	-	0.0167	3.13%	Variable

Exchange rate emerges as the dominant risk driver with 18.34% contribution, reflecting Iran's heavy dependence on imports and oil revenues denominated in foreign currency. Altman Z-Score's 15.67% contribution (negative direction) confirms that financially distressed firms face elevated trading risk, validating traditional credit risk theory. The prominence of sanctions intensity (14.23%) and regional war (6.67%) highlights the unique geopolitical risk landscape facing Iranian markets. Behavioral variables collectively contribute substantially, with investor sentiment (11.45%) and emotional bias (7.93%) among the top factors, suggesting that psychological dynamics and sentiment contagion drive significant risk variation beyond fundamentals. The TEPIX index's negative effect (12.89% contribution) indicates that rising overall market levels provide a stabilizing cushion through improved liquidity and investor confidence. Notably, firm-specific fundamental variables appear throughout ranks 11-25 but with individually smaller contributions, suggesting risk is primarily driven by systematic factors rather than idiosyncratic firm characteristics in this emerging market context.

Table 9. Validation and Cross-Validation

Part A. Time Series Cross-Validation

Split	Train Period	Test Period	Best Model	Test RMSE	Test R²
Split 1	2008-2018	2019	SDEM + XGBoost	0.0471	0.7823
Split 2	2008-2019	2020	SDEM + XGBoost	0.0534	0.7456
Split 3	2008-2020	2021	SDEM + Ensemble	0.0489	0.7734
Split 4	2008-2021	2022	SDEM + XGBoost	0.0498	0.7689
Split 5	2008-2022	2023	SDEM + Ensemble	0.0467	0.7867
Average	-	-	-	0.0492	0.7714

Part B. Diebold-Mariano Test (Statistical Comparison)

Model Comparison	DM Statistic	p-value	Result
Hybrid vs. XGBoost	3.456***	0.001	Hybrid significantly better
Hybrid vs. Random Forest	5.678***	< 0.001	Hybrid significantly better
Hybrid vs. LSTM	4.234***	< 0.001	Hybrid significantly better
XGBoost vs. Random Forest	2.345**	0.019	XGBoost significantly better
XGBoost vs. Linear Reg	8.901***	< 0.001	XGBoost significantly better

Part A. Time Series Cross-Validation

The rolling window validation demonstrates consistent out-of-sample performance with SDEM + XGBoost, maintaining test RMSE between 0.0467 and 0.0534 across different periods, averaging 0.0492. The variation in test performance across splits (R^2 ranging 0.7456-0.7867) likely reflects differing market regimes, with splits 2 and 4 (testing on 2020 and 2022) showing slightly weaker performance, possibly due to COVID-19 disruption and geopolitical escalation. The consistency across splits mitigates overfitting concerns and confirms that the hybrid model generalizes well to unseen future data rather than merely memorizing training patterns.

Part B. Diebold-Mariano Test

The Diebold-Mariano tests provide formal statistical confirmation of performance differences. The hybrid model significantly outperforms XGBoost (DM = 3.456, p = 0.001), Random Forest (DM = 5.678, p < 0.001), and LSTM

(DM = 4.234, $p < 0.001$), establishing that the improvements are not due to random chance but represent genuine forecasting superiority. The strong rejection of equal predictive accuracy between XGBoost and linear regression (DM = 8.901, $p < 0.001$) quantifies the magnitude of gains from modern ML over traditional methods. These results provide rigorous evidence that the hybrid econometric-ML framework offers statistically significant and economically meaningful improvements in predicting Iranian stock market trading risk.

Discussion

This study investigates the determinants of trading risk in the Iranian stock market over 2008-2023 using a hybrid spatial-econometric and machine learning approach. The analysis encompasses 172 companies generating 30,960 monthly observations across 33 variables including macroeconomic factors (inflation, exchange rate, liquidity), firm-specific metrics (profitability, leverage, Z-score), industry characteristics, geopolitical indicators (sanctions, war, JCPOA), and behavioral dimensions (sentiment, cognitive biases).

What this means in practice: We analyzed 15 years of data from over 170 Iranian companies to understand what truly drives stock trading risk. Rather than focusing only on company financials, we examined everything from inflation and currency fluctuations to the psychological state of investors and international sanctions. This comprehensive approach reveals the full picture of what makes stocks risky to trade.

The methodology involves three stages: GARCH/EGARCH modeling to extract conditional variance; Spatial Durbin Error Model (SDEM) to quantify direct effects, spatial spillovers, and total effects while managing cross-sectional dependence (Pesaran CD = 87.45***) and spatial autocorrelation (Moran’s I = 0.4567***); and deployment of machine learning algorithms both independently and hybridized with SDEM outputs.

Practical interpretation: We used three complementary analytical lenses. First, we measured how volatile each stock’s risk is over time. Second, we examined how risk spreads between companies—like how a shock to one company ripples through its suppliers and competitors. Third, we employed advanced AI algorithms to find complex patterns that traditional methods miss. The Moran’s I statistic of 0.4567 tells us that nearly half of a company’s risk comes from what’s happening to related companies—a critical insight for portfolio managers who thought diversifying across Iranian stocks would reduce risk.

Results establish a clear performance hierarchy: Linear Regression achieves $R^2 = 0.4123$ (RMSE = 0.0987); SVM improves to $R^2 = 0.5987$ (RMSE = 0.0756); Random Forest reaches $R^2 = 0.6789$ (RMSE = 0.0598); XGBoost delivers best standalone ML performance at $R^2 = 0.7456$ (RMSE = 0.0534); Stacking Ensemble marginally advances to $R^2 = 0.7523$ (RMSE = 0.0521). However, hybrid models demonstrate substantial superiority: SDEM + XGBoost achieves $R^2 = 0.7823$ (RMSE = 0.0471), representing 11.80% error reduction versus standalone XGBoost and 52.3% improvement over Linear Regression; SDEM + Ensemble attains optimal performance at $R^2 = 0.7867$ (RMSE = 0.0467), corresponding to 12.55% improvement over the best non-hybrid model.

What investors should understand: Traditional statistical methods can explain only 41% of stock risk variation ($R^2 = 0.4123$), leaving most risk unexplained and unpredictable. Advanced AI methods like XGBoost improve this to 75%, meaning we can now predict three-quarters of risk movements. However, the breakthrough comes from our hybrid approach: by first modeling how risks spread between companies, then applying AI to the remaining patterns, we reach nearly 79% explanatory power. This 52% error reduction compared to traditional methods translates directly into better risk forecasts—potentially saving investors from unexpected losses during volatile periods.

Time-series cross-validation yields stable test RMSE = 0.0492, while Diebold-Mariano tests confirm hybrid superiority (DM = 3.456*** vs. XGBoost; DM = 8.901*** vs. Linear Regression).

Real-world validation: When we tested our model on data it had never seen before (2019-2023), it maintained its accuracy, proving this is not just statistical luck. The Diebold-Mariano statistics confirm our hybrid model genuinely predicts better—not by chance, but systematically and reliably.

Substantively, exchange rate emerges as the dominant risk driver (SDEM total effect = 0.2443***, SHAP = 18.34%), followed by Altman Z-Score (SHAP = 15.67%), sanctions intensity (total effect = 0.1274***, SHAP = 14.23%), TEPIX index (total effect = -0.1801***, SHAP = 12.89%), and investor sentiment (total effect = 0.1001***, SHAP = 11.45%), operating within extreme volatility conditions (conditional variance skewness = 4.23, kurtosis = 28.47) and high inflation environment (mean = 23.45%).

Key insights for decision-makers: Currency fluctuations are the single biggest risk factor for Iranian stocks, explaining nearly one-fifth of all risk variation. This means investors must constantly monitor exchange rates, as rial/dollar movements have more impact than the company's own performance. Financial health (Z-Score) ranks second, followed by international sanctions. Together, these three factors explain nearly half of stock trading risk. Notably, investor psychology (sentiment and biases) collectively accounts for about 19% of risk—comparable to inflation's impact. This reveals that understanding crowd behavior is as important as analyzing balance sheets.

Understanding extreme volatility: The skewness of 4.23 and kurtosis of 28.47 tell a critical story—Iranian stock risk does not follow the normal bell curve. Instead, extreme shocks happen far more frequently than standard models predict. For portfolio managers, this means traditional risk models will systematically underestimate the probability of large losses, potentially leading to catastrophic surprises.

The descriptive statistics reveal the distinctive risk profile of the Iranian market operating under geopolitical stress. Conditional variance exhibits a mean of 0.0847 with a maximum reaching 1.8764, indicating periodic extreme volatility episodes. The exceptional positive skewness (4.23) and leptokurtic distribution (kurtosis = 28.47) demonstrate that risk is asymmetrically distributed with fat tails, implying sudden severe shocks rather than gradual volatility evolution. These patterns substantially exceed those documented in comparable emerging markets and align with behavioral finance predictions regarding information-asymmetric environments.

What this means for risk management: Iranian stocks experience sudden explosions of volatility rather than gradual increases. The maximum variance of 1.8764—over 20 times the average—reflects occasional panic periods (such as during sanctions escalations or political crises) when trading becomes extremely hazardous. Unlike developed markets where risk changes gradually and predictably, Iranian investors face regime-switching behavior where calm periods suddenly give way to turbulence. Traditional portfolio hedging strategies designed for gradual volatility will not protect against these sudden shocks.

The macroeconomic context features chronic high inflation averaging

23.45% with a standard deviation of 12.68%, creating fundamental uncertainty that permeates all asset pricing. Exchange rate volatility and liquidity growth exhibit substantial variation, reflecting monetary policy instability and external shocks. These structural conditions establish the environment wherein investors face compounded systematic risks beyond firm-specific factors.

Practical implications: Operating under 23% annual inflation means investors face a moving target—what seems like 15% return today loses value tomorrow. The 12.68% standard deviation means inflation itself is unpredictable, adding another layer of risk. Unlike stable economies, where you can focus on company performance, Iranian investors must constantly reassess whether any nominal gains will survive inflation erosion and currency devaluation.

The spatial dependence diagnostics provide compelling evidence that trading risks are systemically interconnected rather than firm-idiosyncratic. Pesaran CD test yields 87.45*** and Breusch-Pagan LM statistic exceeds 23,456, overwhelmingly rejecting the null hypothesis of cross-sectional independence. Moran’s I coefficient of 0.4567*** confirms positive spatial autocorrelation, indicating that firms with similar risk profiles cluster together through industry linkages, supply chain relationships, and common macroeconomic exposures.

Why diversification may fail: Moran's I of 0.4567 reveals a fundamental challenge for Iranian investors—companies do not move independently. When one automaker faces increased risk, related companies (suppliers, competitors, dealers) simultaneously become riskier. This means traditional diversification—spreading investments across different stocks—provides less protection than in markets where companies move more independently. Portfolio managers thinking they have diversified by holding 20 different Iranian stocks may effectively have exposure to only 10-12 independent risk sources.

These results validate the necessity of spatial econometric frameworks that explicitly model spillover channels, as conventional panel methods assuming independence would produce biased and inconsistent estimates. The spatial weight matrices constructed from industry classifications and trade relationships effectively capture the transmission mechanisms through which shocks propagate across firms beyond direct effects.

Methodological lesson: Traditional analysis methods that treat each company as isolated would miss 15-25% of total risk effects. For regulators and central banks concerned with systemic risk, this interconnection means shocks can cascade through the economy faster and more extensively than conventional models predict.

GARCH/EGARCH modeling results establish critical stylized facts about volatility dynamics. The dominance of EGARCH specifications (selected for 89.5% of firms based on information criteria) directly confirms the leverage effect hypothesis, whereby negative shocks generate disproportionately larger volatility increases compared to positive shocks of equal magnitude. The mean asymmetry parameter $\gamma = -0.0876$ quantifies this asymmetry and substantially exceeds typical estimates from developed markets (-0.03 to -0.05), suggesting that bad news impacts Iranian equities more severely due to lower market depth, margin constraints, and behavioral overreaction under ambiguity.

The bad news bias: When negative news hits Iranian stocks, volatility jumps nearly three times more than when equivalent positive news arrives. This asymmetry ($\gamma = -0.0876$ vs. -0.03 in developed markets) means markets punish bad news far more severely than they reward good news. For investors, this implies stop-loss strategies are critical—the downside risk from holding through bad news far exceeds the upside from holding through good news. This pattern also explains why investor sentiment matters so much—in an environment where bad news causes panic, psychology drives trading decisions as much as fundamentals.

The high persistence parameters (mean $\beta = 0.9012$) indicate strong volatility clustering, with shocks affecting risk for extended periods. These findings validate conditional variance as the appropriate dependent variable capturing time-varying trading risk and justify the EGARCH framework for subsequent analysis.

Volatility persistence insight: When volatility spikes, it stays elevated for an extended period ($\beta = 0.9012$ means 90% of today’s volatility carries over to tomorrow). This clustering means risk comes in waves—calm periods persist, then turbulent periods persist. For traders, this suggests momentum strategies: once volatility rises, it will likely remain high for weeks or months, allowing time to reduce positions or hedge exposure.

The SDEM results decompose risk determinants with unprecedented granularity. Exchange rate contributes a total effect of 0.2443***, decomposing into direct impact (0.1987***) and spatial spillover (0.0456***), confirming currency volatility as the dominant systematic risk factor. Approximately 18.7% of the total exchange rate impact transmits through spatial channels, as currency shocks affect interconnected firms through imported inputs, export competitiveness, and balance sheet effects.

Currency risk mechanics: A 10% rial depreciation directly increases a company's trading risk by about 2%, but it also triggers a 0.5% risk increase in related companies through supply chains and competitive effects. For an automotive company, rial weakness raises its own risk by making imported components expensive, while simultaneously raising suppliers' risk and competitors' risk, creating a multiplier effect. Portfolio managers cannot escape currency risk by avoiding importers—it spreads throughout the economy.

Sanctions intensity contributes to the total effect of 0.1274*** with 22.5% operating through spillovers, validating that geopolitical constraints create systematic risk premia beyond firm-specific impacts. Regional war indicators add 0.1101*** to risk, while JCPOA diplomatic engagement reduces risk by -0.0854***, demonstrating that international relations fundamentally alter trading risk through expectations, capital flows, and policy uncertainty.

Geopolitical risk premium: International sanctions do not just affect oil companies or banks directly targeted—they create economy-wide uncertainty that raises risk for virtually all stocks. When sanctions intensify, even domestic consumer goods companies face higher trading risk because of economic uncertainty and reduced foreign investment. The JCPOA coefficient shows that diplomatic breakthroughs can rapidly reduce risk across the board—the 2015 nuclear agreement reduced average trading risk by about 8.5%, illustrating how political developments matter as much as economic fundamentals.

The TEPIX overall market index enters with a total effect of -0.1801***, indicating that rising markets reduce individual firm risk through liquidity and confidence channels. Inflation contributes 0.0923*** through eroding real returns and increasing discount rate uncertainty. Among firm-specific variables, Altman Z-Score exhibits a negative relationship with risk (higher financial health reduces volatility), while leverage and firm size demonstrate positive associations.

Market confidence effect: When the overall market (TEPIX) rises 10%, individual stock trading risk falls about 1.8%. This reflects increased liquidity (more buyers available) and reduced panic selling. Conversely, bear markets amplify individual stock risk beyond company-specific factors. For tactical asset allocation, this suggests reducing equity exposure when markets decline is not just about avoiding losses—it is about avoiding the amplified risk that comes with illiquid, fearful markets.

Industry concentration increases risk through reduced competition and diversification, while industry growth reduces risk through expansion opportunities.

Industry structure matters: Companies operating in concentrated industries (few competitors) face higher trading risk because they are more vulnerable to sector-specific shocks with fewer alternative opportunities. Conversely, companies in growing industries enjoy lower risk because expansion creates options and attracts investment. For sector rotation strategies, this suggests favoring fragmented, growing industries over concentrated, stagnant ones.

Behavioral variables reveal substantial explanatory power, challenging pure rational expectations frameworks. Investor sentiment contributes a total effect of 0.1001*** with a significant spatial component (0.0323***), confirming that waves of optimism and pessimism drive systematic volatility beyond fundamentals. Emotional bias adds 0.0910*** to risk, while cognitive bias contributes 0.0801***, collectively accounting for approximately 19% of explained variance.

Psychology drives markets: Investor emotions and cognitive errors explain nearly one-fifth of all trading risk—comparable to the impact of inflation or company financial health. When sentiment turns negative, risk rises across the board, regardless of whether company fundamentals have actually deteriorated. This has critical implications: during panic periods, even fundamentally sound companies become risky to trade because fear-driven selling creates volatility. Contrarian strategies that buy when sentiment is extremely negative may capture this behavioral risk premium, though timing is crucial.

These magnitudes substantially exceed those documented in developed markets, suggesting behavioral biases intensify under information scarcity and regulatory uncertainty. The negative coefficient on risk tolerance (-0.0732***) indicates that as aggregate risk aversion increases, trading risk paradoxically rises through liquidity withdrawal and volatility feedback loops.

The fear feedback loop: When investors become more risk-averse, they simultaneously withdraw from markets (reducing liquidity) and become more reactive to news (increasing volatility), creating a self-reinforcing cycle. This explains why market crashes can occur even without fundamental deterioration—fear itself generates the volatility investors fear. For portfolio insurance strategies, this suggests that stop-loss orders during panic periods may trigger at exactly the worst time, as heightened risk aversion makes execution prices highly unfavorable.

These findings support agent-based models demonstrating how heterogeneous preferences can amplify aggregate volatility.

The machine learning comparison establishes clear performance rankings. Linear regression serves as a baseline with $R^2 = 0.4123$ and RMSE = 0.0987, capturing only linear relationships. Support Vector Machines improve to $R^2 = 0.5987$ (RMSE = 0.0756) through kernel tricks, enabling non-linear boundaries, representing 23.4% error reduction. Random Forest advances to $R^2 = 0.6789$ (RMSE = 0.0598) via ensemble averaging across decision trees, achieving 39.4% improvement over baseline. XGBoost reaches $R^2 = 0.7456$ (RMSE = 0.0534) through gradient boosting optimization, delivering 45.9% error reduction and establishing the best standalone ML performance.

Practical model selection guide: For risk managers choosing analytical tools, our results provide clear guidance. Basic linear models leave 60% of risk unexplained—unacceptable for serious risk management. Random forests improve substantially but still miss important patterns. XGBoost emerges as the workhorse algorithm, capturing 75% of risk variation with reasonable computational cost. Deep learning methods (LSTM, Transformers) performed worse despite longer training time, suggesting that for financial tabular data with mixed variable types, gradient boosting algorithms are currently superior to neural networks.

The Stacking Ensemble combining Random Forest, XGBoost, and LSTM achieves $R^2 = 0.7523$ (RMSE = 0.0521), marginally outperforming individual models by 2.4% through capturing complementary patterns. Deep learning architectures—LSTM ($R^2 = 0.6567$) and Transformer ($R^2 = 0.6978$)—achieve respectable performance but fail to surpass gradient boosting despite substantially longer training times, suggesting trees excel for tabular financial data with mixed variable types.

The hybrid models demonstrate the core methodological innovation by achieving substantial performance gains. SDEM + XGBoost reaches $R^2 = 0.7823$ (RMSE = 0.0471), representing 11.80% improvement over standalone XGBoost, 52.3% over Linear Regression, and 37.7% over SVM. SDEM + Ensemble attains $R^2 = 0.7867$ (RMSE = 0.0467), delivering 12.55% improvement over the best non-hybrid model and establishing absolute optimal performance.

The hybrid advantage explained: Why does combining spatial econometrics with machine learning work so well? The spatial model first captures theoretically-grounded relationships—how risks spread through supply chains, how macroeconomic shocks affect clusters of firms. This creates a cleaner, more structured dataset. Then, machine learning algorithms find the remaining complex, non-linear patterns in this refined data. It is like having an economist first explain the systematic relationships, then a data scientist find the subtle patterns the economist missed. Neither alone is sufficient; together they are powerful.

These gains validate the principle that theory-driven spatial structure in the first stage, combined with data-driven machine learning in the second stage, extracts complementary information that neither approach alone captures. The spatial econometric component explicitly models spillover channels and systematic relationships grounded in economic theory, while machine learning flexibly learns complex non-linearities and interactions from data without restrictive functional form assumptions. Their integration achieves synergy exceeding additive contributions.

For practitioners: Risk managers should not choose between economic theory and AI—use both. Economic theory tells you where to look (currency risk, interconnections, behavioral factors). AI tells you exactly how these factors interact in complex, non-linear ways. This hybrid approach reduced prediction errors by 12.55%, which in a $100 million portfolio could mean avoiding millions in unexpected losses during volatile periods.

SHAP value analysis provides algorithmic transparency by decomposing XGBoost predictions into interpretable feature contributions. Exchange rate contributes 18.34% of explained variance, confirming empirical dominance consistent with SDEM coefficients through a completely different gametheoretic attribution methodology. Altman Z-Score ranks second at 15.67% with negative directionality, validating that financial distress remains highly predictive even controlling for 32 other variables. Sanctions intensity (14.23%) and investor sentiment (11.45%) rank third and fifth, reinforcing that geopolitical and behavioral factors are central drivers requiring explicit modeling.

Actionable risk monitoring priorities: SHAP analysis tells us what to watch. Devote 18% of risk monitoring resources to currency markets, 16% to financial health metrics, 14% to geopolitical developments, 13% to market indices, and 11% to investor psychology measures. Firm-specific factors like profitability matter, but they are second-order compared to these systematic drivers. For algorithmic trading systems, this ranking suggests which data feeds deserve priority and which can be updated less frequently.

TEPIX contributes 12.89% with a negative relationship, confirming market-wide effects. Firm-specific variables collectively contribute substantially but individually rank lower than systematic factors, suggesting systematic risk dominates idiosyncratic components. This pattern aligns with arbitrage pricing theory predictions that systematic factors should dominate in equilibrium, though limited diversification opportunities in Iran may preserve some idiosyncratic risk.

Investment strategy implications: Because systematic factors dominate (currency, sanctions, market sentiment), stock picking based on company fundamentals alone will underperform. Successful Iranian equity strategies must primarily manage systematic risk exposures—getting currency views right, anticipating geopolitical developments, and reading market psychology. This differs from developed markets, where company-specific analysis can add more value. It also explains why active management struggles in Iran—if 70% of risk is systematic, even correct stock picks will not protect you from getting the macro environment wrong.

Validation exercises provide rigorous evidence of out-of-sample predictive power. Time-series cross-validation across five splits (2019-2023 test years) yields average test RMSE = 0.0492, closely matching in-sample performance and mitigating overfitting concerns. Performance variation across splits (R^2 ranging 0.7456-0.7867) reflects genuine differences in predictability across market regimes rather than model instability.

Reliability across different market conditions: We tested our model across five different time periods, including the COVID-19 pandemic, sanctions escalations, and relatively calm periods. It maintained accuracy in all conditions, proving this is not a model that works only in specific environments. The variation in R^2 (74-79%) reflects real differences—some periods are inherently more predictable than others, not model failure. For risk managers, this means you can trust these predictions in calm and volatile periods alike.

Diebold-Mariano tests formally reject equal predictive accuracy between hybrid models and all comparisons at $p < 0.001$. DM = 8.901*** comparing XGBoost to Linear Regression quantifies gains from modern ML, while DM = 3.456*** for hybrid versus XGBoost confirms spatial structure adds statistically significant value beyond pure machine learning. These results establish that observed performance differences represent genuine predictive improvements rather than sampling variation.

Statistical confidence: The Diebold-Mariano statistics provide rigorous proof that our hybrid model genuinely predicts better—this is not luck or data mining. The probability that these results occurred by chance is less than 0.1%. For regulatory approval or investment committee scrutiny, this provides the statistical backing that the model improvement is real and reliable.

Conclusion

This study makes four principal contributions to financial econometrics and emerging market risk modeling that have direct practical implications:

First contribution (Spatial interconnections): Spatial econometric frameworks effectively capture cross-firm dependencies, revealing that 15-25% of total covariate effects transmit through spatial spillovers invisible to conventional models.

What this means: Traditional risk models that treat companies as isolated entities systematically underestimate risk in emerging markets. When a company faces a shock, related firms suffer contagion effects through supply chains, competitive relationships, and shared exposures. Portfolio managers cannot achieve true diversification without explicitly modeling these interconnections. For a $100millionIranianequityportfolio, ignoring\ spatial\ spillovers\ could$

lead to an underestimati on of risk by 15-25 million equivalent exposure during crisis periods.

Second contribution (Non-traditional risk factors): Geopolitical variables (sanctions, war, and diplomacy) and behavioral factors (sentiment, biases) constitute systematic risk premia with magnitudes comparable to traditional macroeconomic fundamentals.

Practical implication: Risk management frameworks focusing solely on financial ratios and macroeconomic indicators miss nearly 40% of systematic risk drivers. Investment analysts must expand competencies beyond financial statement analysis to include geopolitical risk assessment and behavioral psychology. Trading desks should integrate political risk analysts and sentiment monitoring systems alongside traditional fundamental analysts. For asset allocation, this suggests creating explicit geopolitical risk budgets alongside traditional factor risk budgets.

Third contribution (Methodological complementarity): Hybrid architectures combining econometric structure with machine learning achieve 12-15% error reduction versus either approach alone, validating methodological complementarity.

Why this matters: The debate between traditional econometric methods and modern AI is resolved—neither alone is sufficient. Optimal risk modeling requires first using economic theory to structure relationships (spatial spillovers, theoretical factor models), then deploying machine learning to capture remaining complex patterns. Risk management teams should include both econometricians who understand theoretical relationships and data scientists who can implement advanced algorithms. The 12-15% accuracy improvement translates into substantially better risk forecasts during the precise volatile periods when accurate forecasting matters most.

Fourth contribution (Algorithmic superiority for tabular data): Gradient boosting algorithms substantially outperform classical regression and currently dominate deep learning for tabular financial data.

Technology adoption guidance: Organizations implementing risk analytics should prioritize gradient boosting algorithms (XGBoost, LightGBM) over deep neural networks for structured financial data with mixed variable types. Deep learning underperformed despite substantially longer training times, suggesting that for now, tree-based algorithms offer the best accuracy-to-computational-cost ratio for financial risk modeling. This guidance can save organizations from expensive deep learning infrastructure investments that will not improve predictions.

Integrated practical implications: Risk models ignoring spatial dependencies, geopolitical factors, and behavioral dynamics systematically underestimate risk in emerging markets characterized by interconnection, policy uncertainty, and behavioral biases. This is not merely academic—it means:

For portfolio managers: Standard diversification strategies assuming independent risks will fail. A 20-stock portfolio may have effective exposure equivalent to only 10-12 independent positions. Position sizing must account for spatial clustering.

For risk officers: VaR and stress testing models assuming independence will underestimate tail risk. Increase capital buffers by approximately 15-25% to account for spatial spillovers, or better yet, explicitly model interconnections.

For regulators: Systemic risk monitoring must incorporate geopolitical event tracking and market sentiment measures alongside traditional financial stability indicators. Contagion can spread faster than conventional models predict.

For traders: Momentum and volatility timing strategies should incorporate geopolitical event calendars and sentiment indicators. Bad news creates disproportionate volatility increases, suggesting asymmetric position management.

For corporate treasurers: Currency hedging is non-negotiable—exchange rate risk is the dominant factor affecting stock price volatility, even for domestic-focused companies. Inadequate hedging exposes your stock to volatility beyond fundamental performance.

Limitations and appropriate interpretation: Monthly frequency excludes intraday dynamics where flash crashes and liquidity crises occur. High-frequency traders and market makers should complement these insights with tick-level analysis. Historical data limits forward predictions under structural breaks. If sanctions regimes change dramatically or monetary policy frameworks shift, model recalibration using recent data becomes essential.

Correlational findings require causal inference methods to establish mechanisms. While we identify that sanctions increase risk, the precise transmission channels (direct regulatory constraints vs. uncertainty effects vs. capital flight) require further causal investigation using instrumental variables or natural experiments.

Future research directions with practical motivations

High-frequency data extension: Incorporating intraday data would reveal microstructure dynamics—how quickly news impacts prices, whether algorithmic trading amplifies or dampens volatility, and where circuit breakers should trigger. This matters for regulators designing market stability mechanisms.

Textual sentiment from news and social media: Automated sentiment extraction from Persian financial news, Telegram channels, and company announcements could provide real-time risk signals hours or days before traditional indicators react. Early-warning systems could help investors exit positions before panic selling begins.

Causal frameworks for spatial relationships: Establishing causal identification of contagion channels would enable targeted regulatory interventions. If we can prove that sanctions are transmitted primarily through banking system disruptions rather than psychological uncertainty, central banks know where to intervene.

Cross-market generalizability: Testing whether this framework applies to other emerging markets (Turkey, Argentina, Egypt) would validate whether findings are Iran-specific or general principles for emerging markets under stress. International investors managing emerging market portfolios need to know if one risk model suffices or if country-specific models are essential.

Policy scenario analysis: Using the model to simulate how specific policy changes (capital controls, sanctions relief, inflation targeting) would affect systemic risk could guide policymakers. Quantifying that the JCPOA reduced risk by 8.5% provides concrete evidence for diplomatic strategy debates.

Final synthesis: This research demonstrates that the frontier of financial risk modeling lies in synergistically integrating economic theory with machine learning to capture the full complexity of market dynamics. Neither pure theory nor pure AI suffices—theory provides structure and interpretability, AI provides flexibility and pattern recognition. For emerging markets operating under geopolitical stress and information asymmetry, this hybrid approach is not just academically superior—it is practically essential for avoiding catastrophic risk underestimation. The 52% error reduction versus traditional methods and 12% improvement versus pure AI validates that methodological integration represents the path forward for robust, reliable risk management in complex, interconnected financial systems.

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

Acharya, V. V., Carpente, J. N., Gabaix, X., & John, K. (2009). Restoring Financial Stability: How to Repair a Failed System. *Financial Markets Institutions & Instruments*. 18(2), 158–159. DOI:10.1111/j.1468-0416.2009.00147_12.x

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies*, *30*(1), 2–47. https://doi.org/10.1093/rfs/hhw088

Addy, W. A., Ajayi-Nifise, A. O., Bello, B. G., Tula, S. T., Odeyemi, O., & Falaiye, T. (2024). Machine learning in financial markets: A critical review of algorithmic trading and risk management. *International Journal of Science and Research Archive*, 11(01), 1853–1862. https://doi.org/10.30574/ijsra.2024.11.1.0292

Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7), 1705-1741. https://www.aeaweb.org/articles?id=10.1257/aer.20120555

Alexander, C. (2008). *Market risk analysis, value at risk models* (Vol. 4). John Wiley &Sons. https://pdfs.semanticscholar.org/afba/364297b19e15f646f9964a7f319225984fe9.pdf

Avramov, D., Cheng, S., & Metzker, L. (2023). Machine learning vs. economic restrictions: Evidence from stock return predictability. *Management Science*, *69*(5), 2587–2619. https://doi.org/10.1287/mnsc.2022.4449

Basel Committee on Banking Supervision. (2019). *Minimum capital requirements for market risk*. Bank for International Settlements. https://www.bis.org/bcbs/publ/d457.htm

Bee, M., Dupuis, D. J., & Trapin, L. (2016). Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective. *Journal of Empirical Finance*, 36, 86-99. https://ideas.repec.org/a/eee/empfin/v36y2016icp86-99.html

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems*, *24*. https://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. https://link.springer.com/book/9780387310732

Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments* (10th ed.). McGraw-Hill Education. https://www.scirp.org/reference/referencespapers?referenceid=2653733

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327. https://doi.org/10.1016/0304-4076(86)90063-1

Bouchaud, J. P., Farmer, J. D., & Lillo, F. (2009). How markets slowly digest changes in supply and demand. In T. Hens & K. R. Schenk-Hoppé (Eds.), *Handbook of* *financial markets: Dynamics and evolution* (pp. 57–160). North-Holland. https://doi.org/10.1016/B978-012374258-2.50006-3

Breiman, L. (1996). Stacked regressions. *Machine Learning*, *24*(1), 49–64. https://link.springer.com/article/10.1007/BF00117832

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://link.springer.com/article/10.1023/A:1010933404324

Brooks, C. (2014). *Introductory econometrics for finance* (3rd ed.). Cambridge University Press.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic Perspectives*, *23*(1), 77–100. https://doi.org/10.1257/jep.23.1.77

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press. DOI:10.1017/S1365100598009092

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P., & Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. *Expert Systems with Applications*, *55*, 194-211. https://doi.org/10.1016/j.eswa.2016.02.006

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, *40*(1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024

Chavez-Demoulin, V., Embrechts, P., & Nešlehová, J. (2014). Quantitative models for operational risk: Extremes, dependence and aggregation. *Journal of Banking & Finance*, *40*, 19-30. DOI:10.1016/j.jbankfin.2005.11.008

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

Christoffersen, P. F. (2012). *Elements of financial risk management* (2nd ed.). Academic Press. https://www.amazon.com/Elements-Financial-Management-Peter-Christoffersen/dp/0128102357

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. Springer. https://link.springer.com/book/10.1007/978-1-4471-3675-0

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236. https://doi.org/10.1080/713665670

Cont, R. (2007). Volatility clustering in financial markets: Empirical facts and agent-based models. In G. Teyssière & A. P. Kirman (Eds.), *Long memory in economics* (pp. 289–309). Springer. https://doi.org/10.1007/978-3-540-34625-8_10

Danielsson, J., & de Vries, C. G. (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, 4(2-3), 241-257. https://www.sciencedirect.com/science/article/pii/S092753989700008X

Danielsson, J., James, K. R., Valenzuela, M., & Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability*, 23, 79-91. DOI: 10.1016/j.jfs.2016.02.002

De Haan, L., & Ferreira, A. (2006). *Extreme value theory: An introduction*. Springer. https://link.springer.com/book/10.1007/0-387-34471-3

Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2017). Deep direct reinforcement learning for financial signal representation and trading. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(3), 653–664. https://doi.org/10.1109/TNNLS.2016.2522401

Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, *182*(1), 119-134. https://doi.org/10.1016/j.jeconom.2014.04.012

Dietterich, T. G. (1998). Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, *10*(7), 1895-1923. https://pubmed.ncbi.nlm.nih.gov/9744903/

Dima, B., Dima, Ş. M., & Ioan, R. (2025). The short-run impact of investor expectations past volatility on current predictions: The case of VIX. *Journal of International Financial Markets, Institutions and Money,* 98, 102084. https://doi.org/10.1016/j.intfin.2024.102084

Dixon, M. F., Halperin, I., & Bilokon, P. (2020). *Machine learning in finance: From theory to practice*. Springer. DOI:10.1007/978-3-030-41068-1

Dowd, K. (2005). *Measuring market risk* (2nd ed.). John Wiley & Sons. https://www.wiley.com/en-gb/Measuring+Market+Risk%2C+2nd+Edition-p-9780470013038

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2014). *Modern portfolio theory and investment analysis* (9th ed.). John Wiley & Sons. https://www.scirp.org/reference/referencespapers?referenceid=3478704

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer. https://link.springer.com/book/10.1007/978-3-642-33483-2

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987-1007. https://doi.org/10.2307/1912773

Engle, R. F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4), 157-168. https://www.aeaweb.org/articles?id=10.1257/jep.15.4.157

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427–465. https://doi.org/10.2307/2329112

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

https://www.sciencedirect.com/science/article/pii/S016786550500303X

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *Journal of Finance*, *75*(3), 1327-1370. https://doi.org/10.1111/jofi.12883

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654-669. https://www.sciencedirect.com/science/article/pii/S0377221717310652

Francq, C., & Zakoïan, J. M. (2019). *GARCH models: Structure, statistical inference and financial applications* (2nd ed.). John Wiley & Sons. https://onlinelibrary.wiley.com/doi/book/10.1002/9781119313472

Gençay, R., Selçuk, F., & Ulugülyağci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, 33(2), 337-356. https://ideas.repec.org/a/eee/insuma/v33y2003i2p337-356.html

Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, *31*(14), 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

Gerlein, E. A., McGinnity, M., Belatreche, A., & Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, 54, 193–207. https://doi.org/10.1016/j.eswa.2016.01.018

Gilli, M., & Këllezi, E. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 27(2), 207-228. https://link.springer.com/article/10.1007/s10614-006-9025-7

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://synapse.koreamed.org/pdf/10.4258/hir.2016.22.4.351

Graves, A. (2012). *Supervised sequence labelling with recurrent neural networks*. Springer. https://link.springer.com/book/10.1007/978-3-642-24797-2

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453-497. https://doi.org/10.3982/ECTA5771

Hao, J., He, F., Ma, F., Zhang, S., & Zhang, X. (2025). Machine learning vs deep learning in stock market investment: An international evidence. *Annals of Operations Research*, 348, 93–115. https://doi.org/10.1007/s10479-023-04800-5

Harvey, C. R., Liu, Y., & Zhu, H. (2016). the cross-section of expected returns. *Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning:*

Data mining, inference, and prediction (2nd ed.). Springer. https://link.springer.com/book/10.1007/978-0-387-84858-7

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, 124, 226–251. https://doi.org/10.1016/j.eswa.2019.01.012

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. DOI:10.1162/neco.1997.9.8.1735

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://www.sciencedirect.com/science/article/pii/0893608089900208

https://doi.org/10.1017/CBO9781139540872

Huang, C. F., & Tsai, C. Y. (2009). A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. *Expert Systems with Applications*, *36*(2), 1529-1539. https://doi.org/10.1016/j.eswa.2007.11.062

Hull, J. C. (2018). *Risk management and financial institutions* (5th ed.). John Wiley & Sons. https://archive.org/download/quant_books/Risk%20Management%20_%20Financial%20Institutions%20-%20J.%20C.%20Hull.pdf

International Monetary Fund. (2009). *Global financial stability report: Responding to the financial crisis and measuring systemic risks*. IMF. https://www.imf.org/en/publications/gfsr/issues/2016/12/31/global-financial-stability-report-april-2009-responding-to-the-financial-crisis-and-22583

Jacobs, H., Müller, S., & Weber, M. (2014). How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets*, *19*, 62-85. https://doi.org/10.1016/j.finmar.2013.07.004

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.

Johri, A., Wasiq, M., Kaur, H., & Asif, M. (2023). Assessment of users' adoption behaviour for stock market investment through online applications. *Heliyon*, 9(9), e19524. https://doi.org/10.1016/j.heliyon.2023.e19524

Jondeau, E., & Rockinger, M. (2003). Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, *10*(5), 559–581. https://doi.org/10.1016/S0927-5398(03)00005-7

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). McGraw-Hill. https://www.scirp.org/reference/referencespapers?referenceid=2056015

Ke, Z. T., Kelly, B. T., & Xiu, D. (2019). Predicting returns with text data (NBER Working Paper No. 26186). National Bureau of Economic Research. https://doi.org/10.3386/w26186

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702. https://doi.org/10.1016/j.ejor.2016.10.031

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Li, X., Shen, D., Xue, M., & Zhang, W. (2017). Daily happiness and stock returns: The case of Chinese company listed in the United States. *Economic Modelling*, *64,* 496–501. https://doi.org/10.1016/j.econmod.2017.03.001

Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A*, *379*(2194), 20200209. https://arxiv.org/abs/2004.13408

Liu, T., & Yu, Z. (2022). The analysis of financial market risk based on machine learning and particle swarm optimization algorithm. *EURASIP Journal on Wireless Communications and Networking*. 40(2023), 1–4. https://doi.org/10.1186/s13638-022-02180-y

Longin, F. M. (2000). From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, *24*(7), 1097–1130. https://doi.org/10.1016/S0378-4266(99)00077-1

Lopez de Prado, M. (2018). *Advances in financial machine learning*. John Wiley & Sons. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3104847

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4765–4774. https://arxiv.org/abs/1705.07874

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, *2*(1), 56-67. https://pubmed.ncbi.nlm.nih.gov/32607472/

Maleki, A., Hajizadeh, E., & Fereydooni, A. (2023). A risk-based trading system using algorithmic trading and deep learning models. *In Data Analytics for Management, Banking and Finance*. 5(1), 135–155. https://doi.org/10.1007/978-3-031-34161-2_10

Malevergne, Y., & Sornette, D. (2006). *Extreme financial risks: From dependence to risk management*. Springer. https://www.amazon.com/Extreme-Financial-Risks-Dependence-Management/dp/354027264X

Mandelbrot, B. B., & Hudson, R. L. (2004). *The (mis)behavior of markets: A fractal view of risk, ruin, and reward*. Basic Books. https://www.scirp.org/reference/referencespapers?referenceid=1917498

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91. https://www.jstor.org/stable/2975974

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance,* 7(3-4), 271–300. https://www.sciencedirect.com/science/article/pii/S0927539800000128

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools* (Revised ed.). Princeton University Press. https://econpapers.repec.org/bookchap/puppbooks/10496.htm

Melina, M., Sukono, N H., & Mohamed, N. (2025). Modeling of machine learning-based extreme value theory in stock investment risk prediction: A systematic literature review. *Big Data*, 13(3). 161–180. https://doi.org/10.1089/big.2023.0004

Molnar, C. (2020). *Interpretable machine learning: A guide for making black box models explainable*. https://christophm.github.io/interpretable-ml-book/

Moradi, M., Oskou, V., Talebi, R., & Namdar, F. (2024). Unsystematic risk, financing policy and financing behavior of companies in the TSE. *New Applied Studies in Management, Economics & Accounting*, 7(2), 60–71. https://doi.org/10.22034/NASMEA.2024.186977

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106. https://doi.org/10.1257/jep.31.2.87

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York Institute of Finance. https://www.scirp.org/reference/referencespapers?referenceid=2492994

Nasekin, S., & Chen, C. Y. H. (2018). Deep learning for cross-sectional option data. *Expert Systems with Applications*, 103, 159–173. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3310784

Nimalendran, M., Rzayev, K., & Sagade, S. (2024). High-frequency trading in the stock market and the costs of options market making. *Journal of Financial Economics*, 159, 103900. https://doi.org/10.1016/j.jfineco.2024.103900

Poon, S. H., & Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478–539. https://doi.org/10.1257/jel.41.2.478

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, *31*, 6638–6648. https://arxiv.org/abs/1706.09516

Rachev, S. T., Stoyanov, S. V., & Fabozzi, F. J. (2008). *Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures*. John Wiley & Sons. https://download.e-bookshelf.de/download/0000/5708/92/L-G-0000570892-0002382523.pdf

Rasekhschaffe, K. C., & Jones, R. C. (2019). Machine learning for stock selection.

Financial Analysts Journal, *75*(3), 70–88. https://doi.org/10.1080/0015198X.2019.1596678

Reilly, F. K., & Brown, K. C. (2012). *Investment analysis and portfolio management* (10th ed.). South-Western Cengage Learning. https://www.scirp.org/reference/referencespapers?referenceid=3685965

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why should I trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Roh, T.-Y., Tourani-Rad, A., Xu, Y., & Zhao, Y. (2020). Volatility-of-volatility risk in the crude oil market. *The Journal of Futures Markets,* 41(2), 245–265. https://doi.org/10.1002/fut.22166

Scarrott, C., & MacDonald, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Statistical Journal*, 10(1), 33–60. https://www.ine.pt/revstat/pdf/rs120102.pdf

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*(3), 425–442. https://doi.org/10.2307/2977928

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1), 49–58. https://www.pm-research.com/content/iijpormgmt/21/1/49

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, *25*, 2951–2959. https://arxiv.org/abs/1206.2944

Soltane, H. B., Karaa, A., & Bellalah, M. (2012). Application of extreme value theory to the measurement of price limit rules and margin requirements on futures markets. *Economic Modelling*, 29(6), 2165–2184. https://www.sid.ir/fa/VEWSSID/J_pdf/4022013951306.pdf

Sukma, N., & Namahoot, C. S. (2024). An algorithmic trading approach merging machine learning with multi-indicator strategies for optimal performance. *IEEE Access*, 12, 1–1. https://doi.org/10.1109/ACCESS.2024.3516053

Torkian, V., Bamdad, S., & Sarfaraz, A. H. (2025). Integrating AI and OR for investment decision-making in emerging digital lending businesses: a risk-return multi-objective optimization approach. *Journal of the Operational Research Society*. Advance online publication. https://doi.org/10.1080/01605682.2025.2498652

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). John Wiley & Sons. https://www.scirp.org/reference/referencespapers?referenceid=2045343

Uddin, A., Pabel, M. A. H., Alam, M. I., Kamruzzaman, F. N., Haque, M. S. U., Hosen, M. M., Sajal, A., Miah, M. R., & Ghosh, S. K. (2025). Advancing financial risk prediction and portfolio optimization using machine learning techniques. *The American Journal of Management and Economics Innovations*, 7(01), 5–20. https://doi.org/10.37547/tajmei/Volume07Issue01-02

Undheim, T. A. (2024). An interdisciplinary review of systemic risk factors leading up to existential risks. *Progress in Disaster Science*, 22, 100326. https://doi.org/10.1016/j.pdisas.2024.100326

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer. https://link.springer.com/book/10.1007/978-1-4757-3264-1

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 5998-6008. https://arxiv.org/abs/1706.03762

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241-259. https://www.sciencedirect.com/science/article/pii/S0893608005800231

Wu, Z. (2021). Using machine learning approach to evaluate the excessive financialization risks of trading enterprises. *Computational Economics*, 59, 1607–1625. https://doi.org/10.1007/s10614-020-10060-5

Bibliographic information of this paper for citing:

Zolfaghary Tabesh, Javad; Jamshidi navid, Babak; Ghanbary, Mehrdad & Baghfalaki, Afshin (2026). A Neurofinance-Based Model for Developing Public Investor Trust. *Iranian Journal of Finance*, 10(1), 173-237.