DIGITAL TRANSFORMATION AND ADMINISTRATION INNOVATION

Citation: Dalili, A., Azadi Hir, K., & Archin Lisar, M. (2025). Comparison of the Efficiency of Traditional Multivariate Regression and Modern AI-Based Optimization Algorithms in Predicting the Probability of Negative Stock Returns. *Digital Transformation and Administration Innovation*, 3(3), 1-12.

Received date: 2025-02-11 Revised date: 2025-06-13 Accepted date: 2025-06-22 Published date: 2025-07-01

Comparison of the Efficiency of Traditional Multivariate Regression and Modern AI-Based Optimization Algorithms in Predicting the Probability of Negative Stock Returns

Ali Dalili[1], Keyhan Azadi Hir[1]*, Mohsen Archin Lisar[1]

1. Department of Accounting, Ra.C., Islamic Azad University, Rasht, Iran

*Correspondence: azadi@iaurasht.ac.ir

Abstract

Timely identification and prediction of negative stock returns represent a critical challenge in financial analysis and risk management, significantly influencing the effectiveness of investment decision-making. This study evaluates the traditional multivariate regression approach alongside the performance of advanced artificial intelligence (AI) optimization algorithms, including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO), in modeling the probability of negative stock returns. A comprehensive financial dataset was compiled from companies listed on the Tehran Stock Exchange. After conducting preprocessing and feature selection procedures, the models were evaluated based on performance indicators such as accuracy, recall, F1-score, and generalizability. Focusing on Iran's capital market, the research analyzed data from 101 publicly listed companies across two time periods (2010–2015 and 2023–2024). The objective was to assess the accuracy, stability, and generalizability of AI-based models in comparison to the classical regression model. The findings indicate that AI algorithms demonstrate superior performance over multivariate regression due to their ability to model complex, nonlinear relationships between financial variables. These algorithms showed higher predictive accuracy in detecting negative returns. Moreover, their capacity to manage large datasets and reduce statistical noise contributed to enhanced model stability when dealing with irregular and outlier data. These results suggest that employing AI-based optimization methods can serve as an effective tool for market risk analysis and improving decision-making processes under conditions of uncertainty. The present study offers a novel perspective on integrating traditional and intelligent techniques to enhance predictive accuracy in financial markets and may serve as a practical reference for financial analysts and policymakers.

Keywords: Artificial Intelligence Optimization Algorithms, Negative Stock Returns, Multivariate Regression, Financial Prediction, Conservative Reporting.

1. Introduction

In the dynamic and highly volatile landscape of financial markets, the ability to accurately predict stock behavior—especially the occurrence of negative returns—has long been a critical challenge for investors, analysts, and policymakers. Traditional econometric models, including linear regressions and multivariate statistical techniques, have historically been the dominant tools in financial prediction. However, the increasing complexity, volume, and velocity of financial data have made these methods increasingly inadequate. This inadequacy has paved the way for the integration of artificial intelligence (AI) and machine learning (ML) technologies into the domain of stock market prediction, marking a profound paradigm shift in financial analytics and decision-making processes (Ahmed et al., 2022; Subha, 2025).

Artificial intelligence, with its subfields of machine learning, deep learning, and natural language processing, offers enhanced capabilities in identifying nonlinear patterns, learning from massive data streams, and adapting to changing market dynamics. These advantages are crucial in financial environments characterized by noise, volatility, and rapid fluctuations (Ashtiani & Raahemi, 2023; Shaikh et al., 2025). The increasing adoption of these technologies in financial contexts underscores their growing importance and effectiveness. As noted by Subha (2025), AI is no longer merely an experimental innovation in stock trading; it has evolved into a core strategic asset for firms seeking data-driven insights and competitive advantage (Subha, 2025).

Numerous studies have demonstrated the efficacy of AI-based models over traditional statistical tools. For instance, machine learning algorithms such as Random Forest, Support Vector Machines, Gradient Boosting, and deep neural networks have shown superior performance in stock price forecasting, particularly under volatile or uncertain market conditions (Khandagale, 2023; Yang, 2023). These models are capable of capturing complex interactions among variables and can dynamically update their internal parameters to adapt to new market signals (Musale, 2024). As emphasized by Lin and Marques (2024), AI systems do not merely replicate human decision-making—they extend and enhance it, creating new possibilities for forecasting, portfolio optimization, and automated trading (Lin & Marques, 2024).

Furthermore, advanced applications of AI in finance go beyond price forecasting. They now encompass sentiment analysis using textual data, pattern recognition from trading signals, fraud detection, credit scoring, and risk management (Chen et al., 2023; Dash, 2023; Singh, 2022). According to Ashtiani and Raahemi (2023), the inclusion of unstructured data—such as financial news and social media sentiment—into prediction models via AI techniques has significantly enriched the forecasting process and helped capture market psychology more effectively (Ashtiani & Raahemi, 2023). These capabilities are particularly valuable in markets where irrational behaviors and speculative bubbles often dominate.

Despite these promising developments, the implementation of AI in financial contexts is not without its challenges. Data quality, overfitting, model interpretability, and ethical concerns remain key issues that require rigorous attention (Ahmed et al., 2022; Ghallabi et al., 2025). As AI models often operate as "black boxes," the transparency of decision-making processes becomes critical, especially when used in high-stakes environments like trading or credit assessment. Moreover, issues such as data snooping bias and lack of generalizability to different market regimes must be carefully addressed (Ayyıldız, 2023; Lin & Marques, 2024).

The diversity in algorithmic approaches further enriches the landscape. For example, optimization-based algorithms such as Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Biogeography-Based Optimization (BBO), Firefly Algorithm (FA), and Harmony Search (HS) have gained traction for their ability to handle multidimensional, nonlinear, and chaotic systems—conditions that characterize financial markets. These algorithms, by mimicking natural phenomena, allow for multi-point search and adaptive learning, making them well-suited to forecasting rare events like negative stock returns (S. & Sornalakshmi, 2024; Shaikh et al., 2025). Studies by Ghallabi et al. (2025) and Salisu et al. (2024) affirm that these heuristic models often outperform regression-based approaches, especially in high-volatility environments and during periods of external economic shocks (Ghallabi et al., 2025; Salisu et al., 2024).

The specific use of AI to predict negative stock returns—a subset of financial prediction research—has garnered significant interest due to its direct implications for risk management and portfolio protection. Negative return prediction models help investors prepare for market downturns, mitigate losses, and rebalance their portfolios accordingly. According to Shaghaghi Shahri (2024), such models are particularly vital in emerging markets like Iran, where political instability, regulatory shifts, and currency fluctuations introduce additional layers of uncertainty (Shaghaghi Shahri, 2024). AI models, when trained on comprehensive datasets incorporating both macroeconomic and firm-level variables, can effectively anticipate these downturns and support informed decision-making.

A critical evaluation by Mintarya et al. (2023) demonstrates that the integration of ensemble methods and hybrid architectures yields significant improvements in prediction accuracy over single-model approaches (Mintarya et al., 2023). Similarly, Musale (2024) underscores that algorithmic diversity, including evolutionary algorithms and swarm intelligence, provides the flexibility needed to model chaotic systems like capital markets (Musale, 2024). In parallel, studies such as those by Chen et al. (2023) and Ayyıldız (2023) have shown how machine learning tools can be fine-tuned using performance metrics like Mean Absolute Error (MAE) and Mean Squared Error (MSE) to evaluate predictive reliability (Ayyıldız, 2023; Chen et al., 2023).

In recent years, comparative studies have assessed the efficacy of AI algorithms against conventional methods. For instance, Lombardo et al. (2022) analyzed the ability of ML classifiers to predict bankruptcy in U.S. markets and found that non-linear classifiers significantly outperformed traditional models in early warning detection (Lombardo et al., 2022). Similarly, Syukur and Istiawan (2021) examined LQ45 index prediction in Indonesia and concluded that ensemble AI methods consistently yielded more accurate results than linear regression-based counterparts (Syukur & Istiawan, 2021). These findings suggest that AI not only enhances precision but also facilitates early detection of financial risks, including those related to stock declines.

The shift toward AI-driven models is also reflective of broader changes in data availability and computational power. With the proliferation of high-frequency trading data, cloud computing, and real-time analytics, AI models are now more scalable and adaptable than ever before (Lin & Marques, 2024; Shaikh et al., 2025). Moreover, as emphasized by S. and Sornalakshmi (2024), the democratization of AI tools and platforms enables even small investors to access sophisticated forecasting mechanisms that were once limited to institutional actors (S. & Sornalakshmi, 2024).

Given the potential of these technologies, the present study seeks to evaluate and compare the effectiveness of traditional multivariate regression and contemporary AI optimization algorithms in predicting the likelihood of negative stock returns.

2. Methods and Materials

The statistical population of the present study includes all companies listed on the Tehran Stock Exchange during two time periods: 2010–2015 and 2023–2024. These two intervals were selected to evaluate and analyze the stability and generalizability of predictive models for the risk of negative stock returns under differing economic and capital market conditions. Sample selection was based on criteria such as trading volume, stock liquidity, continuity of company operations in both periods, and availability of relevant financial data. Ultimately, 101 companies were selected as the study sample, providing a suitable representation of the entire population and enabling more precise analysis. The sampling method was purposive, based on access to complete and high-quality data. This sampling strategy was adopted to ensure the accuracy and validity of the data and to exclude companies with incomplete or unstable data from the predictive models. Given the importance of precision and stability in financial predictive models, ensuring data quality and selecting a representative sample population are fundamental steps in research design, enabling the extraction of valid and generalizable results.

The data used in this study were collected from official and credible sources within Iran's capital market. Financial information for companies was obtained from platforms such as the Tehran Stock Exchange Information Dissemination System, the Codal system, and associated financial databases. These data included annual and interim financial statements, stock prices, and other information relevant to the variables required for predictive modeling. To ensure the accuracy and completeness of the data, a preprocessing stage was conducted, involving data cleaning, the removal of incomplete or invalid entries, and the correction of outliers. Additionally, the data were normalized and standardized to ensure temporal consistency and enable comparisons across periods. The data collection process was designed to encompass all variables influencing the risk of negative stock returns, as well as macroeconomic factors to be analyzed in subsequent stages. This approach significantly enhanced the accuracy and credibility of both predictive models and statistical analyses.

This research is applied and quantitative in nature, utilizing a modeling and forecasting approach to compare the efficiency of traditional multivariate regression with modern methods based on artificial intelligence (AI) optimization algorithms in predicting the probability of negative stock returns. In the first step, financial data of companies were collected and preprocessed. Subsequently, two categories of predictive models were developed:

- **Traditional Multivariate Regression Model**: Multivariate regression, as a classical statistical method, was employed to model linear relationships between multiple independent variables and the dependent variable. This model has long been used in financial analyses and is capable of explaining the impact of various variables on the risk of negative stock returns.
- **Modern Models Based on AI Optimization Algorithms**: Five selected AI optimization algorithms—Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Biogeography-Based Optimization (BBO), Firefly Algorithm (FA), and Harmony Search (HS)—were developed into predictive models. These algorithms, known for their high capability in exploring large optimization spaces, can identify complex and nonlinear relationships in financial data.

The model execution process began by setting the initial population of each algorithm to 100 samples. Each algorithm then proceeded with the optimization process by defining an appropriate objective function to minimize error metrics such as Mean Squared Error (MSE) and Mean Absolute Error (MAE). Model performance was evaluated and compared based on criteria including prediction accuracy, stability, and generalizability. To validate the results, statistical analyses such as the independent t-test and Levene's test were conducted to assess significant differences between models. This method, while adhering to a rigorous scientific framework, facilitated the use of advanced AI capabilities in the financial domain and opened new horizons for enhancing the analytical tools of conservative financial reporting.

Artificial Intelligence Optimization Algorithms (Metaheuristic Algorithms) are a class of optimization methods inspired by natural and social processes, capable of solving complex problems involving large and multidimensional search spaces. These algorithms are typically used to find near-optimal solutions within a reasonable time and, due to their evolutionary and adaptive structures, can effectively address nonlinear, multi-objective, and unstable problems. In this study, five selected AI optimization algorithms were applied, which are described below:

- **Ant Colony Optimization (ACO)**: This algorithm is based on the social behavior of ants in searching for optimal paths to food sources. In ACO, a number of agents (ants) traverse different paths in the search space in parallel. Using pheromone parameters, better paths are reinforced, increasing the likelihood of their selection in subsequent generations. ACO is particularly effective in solving combinatorial and continuous optimization problems and is widely used in finance, especially in feature selection and parameter optimization.
- **Artificial Bee Colony (ABC)**: This algorithm mimics the collective behavior of honey bees in searching for optimal food sources. Worker, observer, and scout bees operate in coordination to explore and exploit the search space. Due to its simplicity, global search capability, and fast convergence, ABC is highly efficient for optimizing continuous problems and financial data.
- **Biogeography-Based Optimization (BBO)**: Inspired by biogeography theory, which explains how biological species migrate between habitats, this algorithm considers problem solutions as habitats. The migration of information between habitats improves the quality of solutions. BBO is suitable for solving multivariate and nonlinear optimization problems and is effective in identifying complex relationships in financial applications.
- **Firefly Algorithm (FA)**: FA is based on the behavior of fireflies, where the intensity of their light is used to attract other fireflies. Each firefly moves toward brighter ones, gradually converging to optimal regions of the search space. FA is appropriate for optimizing nonlinear, multi-objective problems with complex distributions of optimal points.
- **Harmony Search (HS)**: Inspired by the process of improvisation in musical performance, in HS, each "harmony" represents a solution to the problem. Through the use of a harmony memory, random variations, and combination strategies, the search space is explored and improved. HS is widely applied in financial optimization problems due to its simple structure and strong exploratory power.

3. Findings and Results

To obtain an initial and more comprehensive understanding of the statistical characteristics of the dataset used in this study, descriptive indicators related to the research variables are presented in this section. These indicators include the mean, median, standard deviation, minimum, and maximum values for the dependent variables as well as selected independent variables after optimization using the PCO algorithm. The purpose of presenting these descriptive statistics is to identify the overall features of the data, such as central tendency and dispersion, which play a critical role in the preliminary understanding of the data structure. Additionally, these indicators provide a strong foundation for conducting more detailed statistical analyses and hypothesis testing in later stages of the research. The table below presents a sample of descriptive statistics for three of the study's dependent variables. It is worth noting that the data pertain to companies listed on the Tehran Stock Exchange during the time period from 2010 to 2015. This time span was chosen to examine market changes within a defined economic cycle.

Table 1: Descriptive Statistics of Research Variables

Variable	Mean	Std. Deviation	Minimum	Maximum
Maximum Sigma	1.50	1.28	0.20	4.00
Bottom-to-Top Volatility	0.2901	0.2988	-0.13	2.99
Negative Skewness of Stock Return	0.3605	0.6107	-6.52	6.81
Firm Size	6.85	0.6925	4.98	8.22
Return on Equity (ROE)	0.5348	0.3042	-0.85	0.90
Return on Assets (ROA)	0.6054	0.1303	-0.24	0.98
Tax	0.1626	0.1124	0.10	0.21
Financial Leverage	0.6099	0.2039	0.13	0.93
Market-to-Book Ratio	0.5706	0.3020	0.23	3.02
Working Capital Management	0.2646	0.2062	0.10	0.99
Board Independence	0.6705	0.1540	0.20	0.85
Past Stock Returns	0.5526	0.2051	-0.54	1.23
Trading Volume	2.57	0.2329	1.15	3.99
Conservatism	0.1404	0.3342	-2.01	2.97
Opacity	0.1124	0.1941	0.04	1.93
Investor Heterogeneity	0.033	0.055	-0.46	0.50
Information Asymmetry	0.2170	0.2072	0.04	0.99
Institutional Ownership	0.7337	0.1535	0.27	0.94

In the descriptive analysis of the data, the mean serves as the principal measure of central tendency, used to present a representative overall value for the dataset. This indicator plays a crucial role in demonstrating the average status of the variables and contributes to a better understanding of the general behavior of the data. Conversely, the standard deviation is employed as a key dispersion metric to measure the extent to which data points deviate from the central value, thereby indicating variability and volatility. According to the data in Table 1, the second column represents the mean values of the collected variables and reflects the overall data behavior for each variable. For example, the mean value of "Financial Leverage" is 0.6092, indicating the average debt-to-asset ratio in the studied sample. The fourth and fifth columns show the maximum and minimum values for each variable, which reflect the range of variation and provide valuable information about the spread of the data. For instance, the range of "Return on Assets" spans from -0.24 to 0.98. Additionally, the variable "Negative Skewness of Stock Returns" ranges from -6.52 to 6.81, indicating a negatively skewed distribution of returns. This suggests that while the data are skewed toward higher values, lower values are more extreme and dispersed, highlighting a pattern of asymmetric risk. These statistical details help improve understanding of data distribution characteristics and enable more precise analyses.

Table 2. Descriptive Statistics of the Study's Dummy Variables

Variable	Mode	Maximum	Minimum
Auditor Size	0	1	0
Auditor Opinion	0	1	0
Financial Flexibility	0	1	0
Internal Auditing	0	1	0

The above table presents descriptive statistics for four dummy variables in the study, calculated using SPSS software. All these variables were classified and coded on an ordinal scale, with values ranging from 0 to 1. These variables include auditor size, auditor opinion type, level of financial flexibility, and the presence or absence of internal auditing in the firms. The mode, or the most frequently occurring value, for all these variables is 0, indicating a high frequency of the baseline condition in the sampled companies. These results suggest that in the majority of sampled firms, the baseline scenario was dominant, and there were no substantial differences across these dimensions. Analyzing these dummy variables provides a clearer understanding of the role of auditing characteristics and financial structure in modeling and forecasting the risk of negative stock returns.

Table 3. Performance of the Ant Colony Optimization (ACO) Algorithm Based on MSE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	ANT COLONY	6	31	5.17	0.05
	REGRESSION	6	48.78	8.23	
	TOTAL	12	79.78	13.4	
Sigma	ANT COLONY	6	27	4.5	0.045
	REGRESSION	6	51	8.5	
	TOTAL	12	78	13	
Volatility	ANT COLONY	6	34	5.67	0.42
	REGRESSION	6	44	7.33	
	TOTAL	12	78	13	

In the "bottom-to-top volatility" criterion, the significance level was above 0.05, indicating no statistically significant difference between the Ant Colony Optimization algorithm and multivariate regression. Although ACO had lower prediction error, this was not enough to conclude its superiority. However, for the criteria of "negative skewness of stock returns" and "maximum sigma," the Mann–Whitney test showed that ACO significantly outperformed regression. The mean errors of ACO in these indices were substantially lower than those of regression. This suggests that the Ant Colony Optimization algorithm has greater ability in predicting indicators with nonlinear and unstable behavior, such as negative skewness and extreme sigma values. On more linear metrics such as volatility, however, no significant difference was found. The likely reason for ACO's superior performance lies in its strength in optimization across complex and high-risk spaces.

Table 4. Performance of the Ant Colony Optimization (ACO) Algorithm Based on MAE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	ANT COLONY	6	28.98	4.83	0.05
	REGRESSION	6	42.6	7.10	
	TOTAL	12	71.04	11.93	
Sigma	ANT COLONY	6	24	4	0.016
	REGRESSION	6	54	9	
	TOTAL	12	78	13	
Volatility	ANT COLONY	6	39	6.5	1
	REGRESSION	6	39	6.5	
	TOTAL	12	78	13	

According to Table 4, the nonparametric Mann–Whitney test was used to compare the performance of the Ant Colony Optimization algorithm and multivariate regression using MAE as the error metric for the three main indicators related to the probability of negative stock returns. For the volatility index, the significance level was above 0.05, leading to the acceptance of the null hypothesis (no significant difference in means). This means no meaningful statistical difference was observed between the two models, despite ACO's slightly higher average MAE. This indicates that in low-volatility and relatively stable markets, both models perform similarly. For negative skewness, the Mann–Whitney test revealed a significant difference between the models, rejecting the null hypothesis. ACO, with a mean MAE of 4.83, performed considerably better than regression (MAE = 7.10), indicating superior ability to detect asymmetrical behaviors and outliers in returns—common during crises or liquidity shortages. For maximum sigma, ACO again outperformed regression (MAE = 4 vs. 9), confirming its advantage in detecting extreme return divergence. These results, consistent with the MSE findings, reinforce that ACO excels in more complex and nonlinear indices such as negative skewness and extreme sigma, while showing no advantage in simpler, more linear indicators. These findings highlight that AI optimization methods like ACO—through multi-point search and stepwise optimization—have greater potential for modeling and predicting complex and unstable stock market behaviors in Iran.

Table 5. Performance of the Artificial Bee Colony (ABC) Algorithm Based on MSE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	ABC COLONY	6	34.05	4.75	0.047
	REGRESSION	6	43.5	7.25	
	TOTAL	12	77.1	12	
Sigma	ABC COLONY	6	27	4.5	0.045
	REGRESSION	6	51	8.5	

	TOTAL	12	78	13	
Volatility	ABC COLONY	6	36	6	0.63
	REGRESSION	6	42	7	
	TOTAL	12	78	13	

The results showed that for the volatility index, the error difference between the two methods was not statistically significant, indicating similar performance. However, for negative skewness of stock returns, the ABC algorithm had a lower MSE (4.75 vs. 7.25), indicating higher accuracy in predicting sudden drops and asymmetric market behavior. In the sigma index as well, ABC achieved better identification of extreme deviations in returns (MSE = 4.5 vs. 8.5). Overall, the Artificial Bee Colony algorithm demonstrated greater accuracy under volatile and complex market conditions compared to multivariate regression, while showing no significant difference under low-volatility scenarios.

Table 6. Performance of the Artificial Bee Colony (ABC) Algorithm Based on MAE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	ABC COLONY	6	24	4.12	0.061
	REGRESSION	6	54	6.35	
	TOTAL	12	78	10.47	
Sigma	ABC COLONY	6	24	4	0.016
	REGRESSION	6	54	9	
	TOTAL	12	78	13	
Volatility	ABC COLONY	6	39	6.5	1
	REGRESSION	6	39	6.5	
	TOTAL	12	78	13	

To assess the accuracy of the ABC algorithm in predicting negative stock returns, the Mann–Whitney test was used to compare MAE values between ABC and multivariate regression. For negative skewness, a significant difference was observed: ABC had a lower MAE (4.12 vs. 6.35), reflecting superior accuracy in forecasting declining and asymmetric behaviors. In the maximum sigma index, ABC also outperformed regression (MAE = 4 vs. 9), confirming its advantage in identifying turbulent and high-risk market conditions. These results indicate that the ABC algorithm delivers more accurate and reliable performance on critical market indicators than regression, especially in nonlinear and unstable capital market environments like Iran.

Table 7. Performance of the Biogeography-Based Optimization (BBO) Algorithm Based on MSE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	BBO ALGORITHM	6	57	4.5	0.75
	REGRESSION	6	19.5	3.25	
	TOTAL	12	76.5	7.75	
Sigma	BBO ALGORITHM	6	30	5	0.04
	REGRESSION	6	48	8	
	TOTAL	12	78	13	
Volatility	BBO ALGORITHM	6	25.5	6.1	0.05
	REGRESSION	6	46.5	6.9	
	TOTAL	12	78	13	

Table 7 compares the performance of the Biogeography-Based Optimization (BBO) algorithm and multivariate regression in terms of MSE error. The Mann–Whitney test results indicate that for the indices "bottom-to-top volatility" and "negative skewness of returns," the differences between the two methods were not statistically significant, although BBO showed lower errors. However, in the "maximum sigma" index, the difference was statistically significant, with BBO performing better. Thus, BBO is more precise in identifying extreme risk behaviors compared to regression, but its statistical superiority is not definitive across all indicators.

Table 8. Performance of the Biogeography-Based Optimization (BBO) Algorithm Based on MAE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	BBO ALGORITHM	6	31.2	5.20	0.63
	REGRESSION	6	21	3.50	
	TOTAL	12	52.2	8.70	
Sigma	BBO ALGORITHM	6	23	3.83	0.01
	REGRESSION	6	55	9.17	

	TOTAL	12	78	13	
Volatility	BBO ALGORITHM	6	29.4	4.90	0.05
	REGRESSION	6	46.5	7.75	
	TOTAL	12	75.9	12.65	

According to Table 8, the Mann–Whitney test comparing the BBO algorithm and multivariate regression using the MAE error metric shows that BBO performed better on two indicators—"bottom-to-top volatility" and "maximum sigma"—with statistically significant differences. This indicates that BBO has higher accuracy in detecting extreme fluctuations and outlier behaviors. However, for the "negative skewness of returns" index, the difference was not statistically significant, and no definitive superiority of BBO can be asserted. Overall, BBO shows better performance under high-risk market conditions compared to regression, though no substantial differences were observed under other conditions.

Table 9. Performance Analysis of the Firefly Algorithm Based on MSE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	FF COLONY	6	45	7.50	0.04
	REGRESSION	6	21	3.50	
	TOTAL	12	66	11	
Sigma	FF COLONY	6	30	5.00	0.05
	REGRESSION	6	51	8.50	
	TOTAL	12	81	13.5	
Volatility	FF COLONY	6	39	6.50	0.083
	REGRESSION	6	33	5.50	
	TOTAL	12	72	12	

As shown in Table 9, the Firefly Algorithm (FA) and multivariate regression were compared using MSE error. For negative skewness, regression performed significantly better. However, for maximum sigma, FA demonstrated superior accuracy with a statistically significant difference. For bottom-to-top volatility, no significant difference was found between the two methods. Overall, the Firefly Algorithm performed better under volatile market conditions but did not exhibit consistent superiority in more stable or asymmetrical indicators.

Table 10. Performance Analysis of the Firefly Algorithm Based on MAE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	FF COLONY	6	57	7.50	0.04
	REGRESSION	6	25.5	4.25	
	TOTAL	12	82.5	11.75	
Sigma	FF COLONY	6	27.48	4.58	0.05
	REGRESSION	6	44.52	7.42	
	TOTAL	12	72	12	
Volatility	FF COLONY	6	48	6.27	0.088
	REGRESSION	6	30	5.14	
	TOTAL	12	78	11.41	

In the skewness criterion, regression outperformed the Firefly Algorithm with a statistically significant difference. Conversely, in the maximum sigma indicator, the Firefly Algorithm performed better with lower error, and the difference was also significant. For bottom-to-top volatility, no meaningful difference was observed, and both methods performed similarly. As a result, each method showed relative advantages in different indicators.

Table 11. Performance Analysis of the Harmony Search Algorithm Based on MSE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	HS COLONY	6	53	8.83	0.025
	REGRESSION	6	31	5.17	
	TOTAL	12	84	14	
Sigma	HS COLONY	6	29	4.83	0.049
	REGRESSION	6	49	8.00	
	TOTAL	12	78	12.83	
Volatility	HS COLONY	6	45	7.50	0.33
	REGRESSION	6	33	5.50	
	TOTAL	12	78	13	

In the skewness indicator, regression significantly outperformed the Harmony Search (HS) algorithm. In contrast, the HS algorithm showed significantly better performance in the maximum sigma index. For the volatility metric, no significant difference was found, and both models showed comparable accuracy.

Table 12. Performance Analysis of the Harmony Search Algorithm Based on MAE Error

Model	Group	N	Sum of Ranks	Mean Rank	Sig
Skewness	HS COLONY	6	44.16	7.36	0.025
	REGRESSION	6	21	3.50	
	TOTAL	12	65.16	10.86	
Sigma	HS COLONY	6	25.56	4.26	0.049
	REGRESSION	6	43.98	7.33	
	TOTAL	12	69.54	11.59	
Volatility	HS COLONY	6	47	6.83	0.33
	REGRESSION	6	31	5.17	
	TOTAL	12	78	12	

In the evaluation of skewness based on MAE, multivariate regression significantly outperformed the Harmony Search algorithm. However, in the maximum sigma index, the HS algorithm achieved better performance with a statistically significant difference. For bottom-to-top volatility, no meaningful difference was found, and both models exhibited similar accuracy. To ensure accuracy, Mann–Whitney tests were used for pairwise comparisons and the Kruskal–Wallis test for overall ranking. These tests showed that in certain indicators, AI-based algorithms significantly outperformed the traditional method and can enhance the prediction of stock crash risk in financial markets.

Based on the results obtained, the Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), and Biogeography-Based Optimization (BBO) algorithms significantly outperformed multivariate regression across both error metrics: Mean Squared Error (MSE) and Mean Absolute Error (MAE). The related hypotheses for these three algorithms were confirmed. In contrast, the Firefly Algorithm (FA) and Harmony Search (HS) did not achieve statistically significant improvements over regression and thus their associated hypotheses were rejected. Overall, AI-based optimization algorithms demonstrated superior performance compared to traditional multivariate regression in modeling stock market behavior and predicting crash risk.

Comparisons across the three key risk identification indicators—negative skewness of returns, bottom-to-top volatility, and maximum sigma—revealed varied results:

- The Ant Colony Optimization algorithm was superior to regression in skewness and sigma indicators, but not significantly different in volatility.
- The Artificial Bee Colony algorithm showed a similar pattern, with notable advantage in the sigma indicator.
- While the Biogeography-Based Optimization algorithm occasionally produced lower errors, the differences with regression were not consistently statistically significant, indicating a relatively stable and equal performance.
- The Harmony Search algorithm underperformed compared to regression in skewness but performed better in the sigma index.
- The Firefly Algorithm failed to show significant differences across most indicators and recorded the weakest performance overall.

The comparison between the algorithms themselves also revealed statistically significant differences only for certain indicators, such as skewness. Overall, ACO was identified as the top-performing algorithm across both MSE and MAE metrics—especially in critical indicators like skewness and sigma—followed closely by the ABC algorithm.

4. Discussion and Conclusion

The findings of this study provide robust evidence that artificial intelligence (AI)-based optimization algorithms significantly outperform traditional multivariate regression in predicting the probability of negative stock returns. Across multiple metrics—including Mean Squared Error (MSE) and Mean Absolute Error (MAE)—the Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), and Biogeography-Based Optimization (BBO) algorithms demonstrated superior predictive performance, particularly in high-volatility and nonlinear market conditions. The models showed statistically significant accuracy advantages in detecting key indicators such as negative skewness and maximum sigma values. These results align with a growing body of literature emphasizing the power of AI in financial forecasting, particularly when modeling complex, chaotic, or asymmetric data structures (Ghallabi et al., 2025; Shaikh et al., 2025; Subha, 2025).

The strength of the ACO algorithm in modeling extreme skewness and sigma levels supports previous findings that swarm intelligence and nature-inspired algorithms offer greater flexibility in handling noisy and unstructured financial data (Musale, 2024; Singh, 2022). Its superiority over regression in these indicators reinforces the view that heuristic optimization methods are better suited to environments characterized by rapid shifts, extreme values, and multi-modal distributions. These findings also parallel the results of studies where AI methods were shown to capture nonlinear patterns more effectively than statistical models, especially in detecting downturns in asset prices and crisis-level signals (Dash, 2023; Mintarya et al., 2023). As expected, no significant difference was observed in the bottom-to-top volatility metric—a more linear and stable indicator—demonstrating that in less volatile conditions, traditional regression models may still perform comparably (Syukur & Istiawan, 2021).

Similarly, the ABC algorithm showed a strong ability to model both negative skewness and maximum sigma, though it failed to show significant superiority in more stable indicators. This is in line with the findings of Khandagale (2023), who showed that ensemble and tree-based AI models excel when tasked with identifying outlier behaviors and discontinuities in financial time series (Khandagale, 2023). The observed results also support Lombardo et al. (2022), who concluded that AI classifiers outperform regression models in risk classification scenarios, especially under volatile economic conditions (Lombardo et al., 2022). Meanwhile, the BBO algorithm provided a more moderate performance: while it did not exhibit statistically significant superiority in all areas, it showed reliable and consistent results. This stability is critical in risk-sensitive domains and validates the findings of previous studies that have recommended BBO for financial systems requiring generalizability and reduced variance in predictions (Lin & Marques, 2024; Yang, 2023).

In contrast, the Firefly Algorithm (FA) and Harmony Search (HS) displayed weaker performance, with neither algorithm outperforming regression in most indicators. While they showed potential in predicting maximum sigma, their inconsistent results in skewness and volatility metrics suggest that their heuristic search capabilities may not be as well-suited to the intricacies of stock return prediction as other algorithms. This echoes concerns in the literature regarding the tendency of certain metaheuristics to overfit or become trapped in local optima when faced with high-dimensional or highly stochastic financial data (Ayyıldız, 2023; Chen et al., 2023). These findings demonstrate that not all AI methods are equally effective in financial forecasting, and algorithm selection must be tailored to the specific nature of the prediction task.

Moreover, this study reinforces the assertion that machine learning and AI models, when applied appropriately, can provide early-warning signals for financial downturns and improve risk-adjusted decision-making in capital markets (Ahmed et al., 2022; Shaghaghi Shahri, 2024). The consistently strong performance of optimization-based algorithms in identifying negative skewness—a key indicator of downside risk—suggests a tangible opportunity for asset managers and institutional investors to preemptively adjust their portfolios. This is particularly valuable in emerging markets like Iran, where geopolitical and economic instability increases the frequency and severity of sudden stock devaluations (S. & Sornalakshmi, 2024). In such environments, the capacity to forecast not only trend but also risk asymmetry becomes a core competency, and AI models provide an edge that traditional methods cannot match.

In addition to performance evaluation, this research also contributes to methodological innovation. By applying evolutionary and nature-inspired algorithms to the prediction of negative returns, the study extends the utility of AI in financial forecasting beyond mainstream supervised learning models such as decision trees and neural networks (Ashtiani & Raahemi, 2023; Mintarya et al., 2023). It also empirically supports the claim that hybrid and ensemble systems may not always be necessary when simpler, well-tuned metaheuristics can deliver comparable or superior outcomes in specific forecasting tasks. The model validation using both MSE and MAE further ensures robustness and mitigates the risk of error metric bias—a problem that has been flagged in earlier comparative analyses (Dash, 2023; Yang, 2023).

From a theoretical standpoint, the findings corroborate the growing view that financial markets operate as complex adaptive systems. In this light, models capable of mimicking biological adaptation and swarm behavior are not merely convenient analogs but structurally appropriate frameworks for financial modeling. This supports claims by Ghallabi et al. (2025) and Salisu et al. (2024) that AI tools can better internalize feedback loops, emergent volatility, and time-varying dependencies—key features of modern financial markets (Ghallabi et al., 2025; Salisu et al., 2024). The results also align with Chen et al. (2023), who emphasized that real-world financial data requires algorithmic strategies that go beyond fixed assumptions of normality and linearity (Chen et al., 2023).

Despite its contributions, this study is not without limitations. First, the sample size was limited to 101 firms listed on the Tehran Stock Exchange across two specific periods. While this selection provided valuable insights into emerging market behavior, the findings may not be directly generalizable to more mature or highly liquid markets such as the NYSE or LSE. Second, the study only examined a limited set of optimization algorithms. Although ACO, ABC, BBO, FA, and HS represent a broad class of metaheuristics, other promising models—such as genetic algorithms, particle swarm optimization, or hybrid deep learning architectures—were excluded due to scope constraints. Third, while performance was measured using MAE and MSE, other important evaluation criteria such as precision-recall trade-offs, robustness under different market regimes, and interpretability of model output were not explored in detail. Furthermore, the computational cost and scalability of these algorithms were not fully analyzed.

Future studies should consider extending the dataset to include more companies across multiple international stock exchanges and timeframes to enhance the generalizability of results. Moreover, incorporating high-frequency trading data, macroeconomic indicators, and unstructured data sources such as financial news and social media sentiment could yield even richer and more responsive models. Comparative analysis of hybrid models combining both statistical and heuristic approaches would also be beneficial, particularly to explore if ensemble frameworks might improve performance in markets with mixed volatility profiles. Additionally, studies could explore the integration of explainable AI (XAI) tools to enhance model transparency and build investor trust. Another promising area lies in testing the resilience of models under simulated market shocks or adversarial attacks to assess their robustness in extreme conditions.

Practitioners, particularly portfolio managers and institutional investors, should consider integrating AI-based optimization models—especially ACO and ABC—into their risk management and early warning systems for downturn prediction. These models are particularly suited for volatile or unpredictable market environments. Financial analysts should also be trained in the practical deployment and tuning of these algorithms, ensuring they understand both the capabilities and limitations of AI tools. Finally, policymakers and regulators should encourage the adoption of AI technologies in financial sectors while ensuring ethical standards, data privacy, and model accountability are upheld. The insights from such models could also support macroprudential oversight and improve systemic risk detection in emerging financial ecosystems.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all who helped us through this study.

Conflict of Interest

The authors report no conflict of interest.

Funding/Financial Support

According to the authors, this article has no financial support.

References

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

Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications*, *217*, 119509. https://doi.org/10.1016/j.eswa.2023.119509
Ayyıldız, N. (2023). Prediction of Stock Market Index Movements With Machine Learning. https://doi.org/10.58830/ozgur.pub354
Chen, J., Wen, Y., Nanehkaran, Y. A., Suzauddola, M. D., Chen, W., & Zhang, D. (2023). Machine learning techniques for stock price prediction and graphic signal recognition. *Engineering Applications of Artificial Intelligence*, *121*, 106038. https://doi.org/10.1016/j.engappai.2023.106038
Dash, A. S. (2023). Sentiment Analysis Using Machine Learning for Forecasting Indian Stock Trend: A Brief Survey. *Finance Theory and Practice*, *27*(6), 136-147. https://doi.org/10.26794/2587-5671-2023-27-6-136-147
Ghallabi, F., Souissi, B., Du, A. M., & Ali, S. (2025). ESG stock markets and clean energy prices prediction: Insights from advanced machine learning. *International Review of Financial Analysis*, *97*(1), 103889. https://doi.org/10.1016/j.irfa.2024.103889
Khandagale, H. P. (2023). Predicting Stock Prices With Machine Learning Using Comparative Analysis of Random Forest Algorithm. *Ijeast*, *8*(6), 60-68. https://doi.org/10.33564/ijeast.2023.v08i06.008
Lin, C. Y., & Marques, J. A. (2024). Stock market prediction using artificial intelligence: A systematic review of systematic reviews. *Social Sciences & Humanities Open*, *9*(1), 100864. https://doi.org/10.1016/j.ssaho.2024.100864
Lombardo, G., Pellegrino, M., Adosoglou, G., Cagnoni, S., Pardalos, P. M., & Poggi, A. (2022). Machine Learning for Bankruptcy Prediction in the American Stock Market: Dataset and Benchmarks. *Future Internet*, *14*(8), 244. https://doi.org/10.3390/fi14080244
Mintarya, L. N., Halim, J. N. M., Angie, C., Achmad, S., & Kurniawan, A. (2023). Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science*, *216*, 96-102. https://doi.org/10.1016/j.procs.2022.12.115
Musale, D. S. (2024). Enhancing Stock Market Predictions Through Artificial Intelligence. *International Journal of Advanced Research in Science Communication and Technology*, 556-566. https://doi.org/10.48175/ijarsct-15991
S., S. S., & Sornalakshmi. (2024). A Critical Study on Harnessing the Power of Artificial Intelligence in Stock Market Trading. *International Journal for Multidisciplinary Research*, *6*(3). https://doi.org/10.36948/ijfmr.2024.v06i03.22761
Salisu, A. A., Demirer, R., Gupta, R., Sangeetha, J. M., & Alfia, K. J. (2024). Technological shocks and stock market volatility over a century Financial stock market forecast using evaluated linear regression based machine learning technique. *Journal of Empirical Finance*, *79*, 101561. https://doi.org/10.1016/j.measen.2023.100950
Shaghaghi Shahri, S. (2024). Evaluation and Comparison of Classification Model Performance in Predicting Corporate Credit Ratings Using Artificial Intelligence: A Case Study of the Tehran Stock Exchange. *Transactions on Data Analysis in Social Science*, *6*(2), 31-44. https://www.transoscience.ir/article_200776.html
Shaikh, S. Z., Khan, K. R., Sherwani, F. K., & Khan, M. (2025). Smart Trading: Unlocking Artificial Intelligence in Stock Market. *The Business & Management Review*, *15*(03). https://doi.org/10.24052/bmr/v15nu03/art-17
Singh, G. (2022). Machine Learning Models in Stock Market Prediction. *International Journal of Innovative Technology and Exploring Engineering*, *11*(3), 18-28. https://doi.org/10.35940/ijitee.c9733.0111322
Subha, S. (2025). Role of Artificial Intelligence in Stock Trading. *Thiagarajar College of Preceptors Edu Spectra*, *7*(S1), 44-47. https://doi.org/10.34293/eduspectra.v7is1-feb25.005
Syukur, A., & Istiawan, D. (2021). Prediction of LQ45 Index in Indonesia Stock Exchange: A Comparative Study of Machine Learning Techniques. *International Journal of Intelligent Engineering and Systems*, *14*(1), 453-463. https://doi.org/10.22266/ijies2021.0228.42
Yang, M. (2023). Predicting the Direction of Stock Price Movement With Machine Learning Algorithms. *Advances in Economics Management and Political Sciences*, *52*(1), 283-291. https://doi.org/10.54254/2754-1169/52/20230758