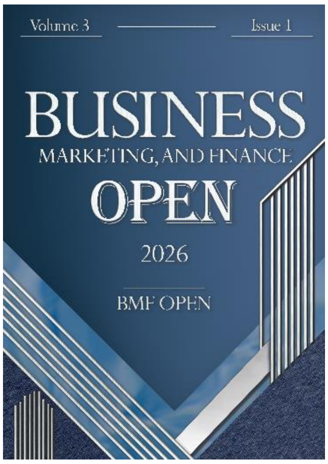
The Examination of the Probability of Negative Stock Returns Using Artificial Intelligence Optimization Algorithms

Ali Dalili[1], Keyhan Azadi Hir [2,*] and Mohsen Archin Lisar [3]

[1] Department of Accounting, Ra.C., Islamic Azad University, Rasht, Iran;
[2] Department of Accounting, Ra.C., Islamic Azad University, Rasht, Iran;
[3] Department of Accounting, Ra.C., Islamic Azad University, Rasht, Iran;

* Correspondence: azadi@iaurasht.ac.ir



Citation: Dalili, A., Azadi Hir, K., & Archin Lisar, M. (2026). The Examination of the Probability of Negative Stock Returns Using Artificial Intelligence Optimization Algorithms. *Business, Marketing, and Finance Open*, 3(1), 1-13.

Received: 21 March 2025
Revised: 19 July 2025
Accepted: 26 July 2025
Published: 01 January 2026

Abstract: Accurately predicting the probability of negative stock returns is one of the central challenges in the field of finance and risk management, which, due to the complex, nonlinear, and non-stationary nature of capital market data, requires the use of advanced and modern analytical methods. Artificial intelligence optimization algorithms, especially metaheuristic algorithms such as Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Harmony Search (HS), Firefly Algorithm (FA), and Biogeography-Based Optimization (BBO), with their ability to effectively search large and multidimensional parameter spaces and adaptive learning capabilities, are promising options for modeling and predicting negative stock returns. This study comprehensively examines the performance of these algorithms based on important financial risk criteria such as negative return skewness, maximum sigma, and low-to-high volatility, through statistical analysis of prediction errors (MSE and MAE). The obtained results indicate the statistically significant superiority of the Ant Colony Optimization algorithm in more accurately predicting negative returns compared to other algorithms and traditional models. The findings show that artificial intelligence optimization algorithms, utilizing natural and biological mechanisms, have the capability to model the complexities of the financial market.

Keywords: negative stock returns, artificial intelligence algorithms, prediction of negative returns, Ant Colony Optimization algorithm, Artificial Bee Colony algorithm

1. Introduction

The dynamic and complex nature of financial markets has always posed considerable challenges for accurate prediction and risk management. Among these challenges, the precise prediction of negative stock returns remains a key priority in the field of financial economics. Understanding the underlying factors that lead to such returns can play a pivotal role in formulating investment strategies, managing risks, and stabilizing financial portfolios. Traditional econometric models, despite their theoretical rigor, often fall short in dealing with the nonlinearities, noise, and high-dimensional features present in real-world financial data. This limitation has increasingly shifted the academic and professional focus toward artificial intelligence (AI) and machine learning (ML)-based approaches, especially those grounded in optimization algorithms [1, 2].

Recent advances in computational finance have showcased the growing effectiveness of AI-powered models in addressing the prediction of market anomalies and adverse movements such as negative returns. Optimization algorithms, particularly those inspired by natural intelligence and evolutionary processes, have demonstrated their strength in navigating high-dimensional parameter spaces, avoiding local optima, and learning complex financial patterns adaptively [3, 4]. These algorithms—such as Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Particle Swarm Optimization (PSO), Firefly Algorithm (FA), and Biogeography-Based Optimization (BBO)—have become widely applied tools in financial modeling for their capability to converge on efficient and robust solutions [5, 6].

The significance of these algorithms in financial prediction stems not only from their mathematical properties but also from their practical implications. For example, Ant Colony Optimization has been particularly effective in reducing mean squared prediction errors by simulating the pheromone trail behavior of ants in financial path discovery [4]. Similarly, the Artificial Bee Colony algorithm, modeled after the foraging behavior of honey bees, shows strong convergence in multidimensional prediction tasks and is widely used in portfolio optimization and return forecasting [7].

Empirical studies have also highlighted the growing demand for machine learning models in addressing the volatility and uncertainty of financial markets. According to bibliometric reviews, the application of AI and ML in finance has witnessed exponential growth over the last decade, with a specific focus on stock market prediction and portfolio risk estimation [1, 2]. This growing interest is driven by the limitations of linear regression-based models, which often assume stationarity, normal distribution, and independence of observations—assumptions that are frequently violated in financial time series [8, 9].

In addition, the Iranian capital market, characterized by high degrees of uncertainty, economic sanctions, and currency volatility, presents a unique case for testing the robustness and generalizability of such AI models. Scholars have identified that periods of structural inflation, exchange rate shocks, and political instability have intensified the need for intelligent decision-support systems in Iranian financial markets [10, 11]. In this context, the Tehran Stock Exchange (TSE) offers fertile ground for empirical evaluation of AI-based predictive models due to its combination of emerging market behaviors and extensive historical financial data.

Building on these motivations, this study employs and compares the predictive power of several AI optimization algorithms in estimating the probability of negative stock returns among companies listed on the Tehran Stock Exchange. The analysis focuses on key financial indicators—such as return on equity (ROE), return on assets (ROA), firm size, leverage, market-to-book ratio, and trading volume—as independent variables, with negative return occurrence defined as a binary dependent variable.

2. Methodology

`This study is applied and descriptive-correlational in nature. The main objective is to predict the probability of negative stock returns in the Tehran Stock Exchange. Since this study examines the relationship and impact of financial variables on the risk of negative returns, and no intervention is made by the researcher, the correlational method is the most appropriate approach for analyzing relationships among variables. Moreover, the use of artificial intelligence optimization algorithms for modeling and prediction strengthens the applied nature of the study.

The statistical population of the research includes all companies listed on the Tehran Stock Exchange during the period from 2010 to 2015. This time frame was chosen due to its relative economic stability, data integrity, and transparency of companies' financial reports. During this period, 405 active companies with complete financial and stock price data were considered as the research sample. To validate the results and test the generalizability of the model, data from the period 2023 to 2024 were also used.

The data used in this study include financial variables and stock market data extracted from the official databases of the Tehran Stock Exchange. The most important independent variables include companies' financial ratios, liquidity indicators, profitability, liabilities, company size, and price volatility metrics. The dependent variable is the occurrence or non-occurrence of negative stock returns during the study period, which is coded as a binary variable.

To improve data quality and enhance the performance of prediction algorithms, preprocessing steps including removal of incomplete data, correction of outliers, normalization of feature values, and handling of imbalanced data were performed. Removal of incomplete data was carried out using the case deletion method. Then, the data were standardized to a range with zero mean and unit standard deviation to reduce the impact of different feature scales in the learning process. Also, due to data imbalance (with the occurrence of negative returns being less than positive returns), oversampling techniques were employed to improve model performance.

To reduce data dimensionality and remove unnecessary features that could reduce model accuracy, the Particle Swarm Optimization (PSO) algorithm was used as a wrapper feature selection method. Inspired by the social behavior of birds and fish, this algorithm searches the feature space and identifies the subset of features that have the greatest impact on predicting the dependent variable. The advantage of PSO lies in its fast convergence and ability to find optimal feature subsets, which helps improve accuracy and reduce model training time.

To model and predict the probability of negative stock returns, several population-based artificial intelligence optimization algorithms were employed, which are briefly introduced below:

• Ant Colony Optimization (ACO): An algorithm inspired by ants' behavior in finding the shortest path to food sources, applied in discrete optimization problems.

• Artificial Bee Colony (ABC): An algorithm based on the food search behavior of honeybees with high ability in optimizing multidimensional problems.

• Firefly Algorithm (FA): An algorithm based on the attraction and light-emitting behavior of fireflies, efficient in continuous optimization problems.

• Harmony Search (HS): An algorithm inspired by the harmony process in music to improve various combinations and find optimal solutions.

• Biogeography-Based Optimization (BBO): An algorithm based on biogeography theory that simulates species distribution to solve optimization problems.

Each of these algorithms was trained as a prediction model using the features selected by PSO and their ability to detect the occurrence of negative stock returns was evaluated.

To assess the quality of prediction models, several criteria were considered, including the following:

• Accuracy: The ratio of correctly predicted samples to total samples

• F1-Score: The harmonic mean of precision and recall, reflecting their balance

• Confusion Matrix: A display of the number of correctly and incorrectly predicted samples in each class

• ROC curve and AUC metric: Indicating the model's ability to distinguish between negative and non-negative return classes

Evaluation of each model based on these criteria enabled an accurate comparison of algorithm performance, and the best algorithm was selected as the proposed model.

3. Findings and Results

To provide an initial statistical understanding of the dataset, this section presents descriptive indicators of the research variables, including mean, median, standard deviation, minimum, and maximum values for both the dependent and selected independent variables, after optimization using the Particle Swarm Optimization (PSO) algorithm. These indicators not only reflect the central tendency and dispersion of the data but also provide a basis for more precise statistical interpretation in subsequent analysis stages. The data used pertain to companies listed on the Tehran Stock Exchange during the period from 2010 to 2015.

Table 1. Descriptive Statistics of Continuous Variables

Variable	Mean	Std. Dev.	Min	Max
Maximum Sigma	1.50	1.28	0.20	4.00
Low-to-High 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 Return	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
Lack of Transparency	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 descriptive data analysis, the mean is used as the most important central tendency indicator to represent a typical value of the dataset. Conversely, the standard deviation is employed as the primary dispersion metric to assess how data points deviate from the center of the distribution. According to the information presented in Table 1, the second column displays the mean of the collected variables and provides an overview of the general behavior of the data for each variable. For example, the mean of the "Financial Leverage" variable is 0.6092, indicating the average debt-to-asset ratio in the examined sample. The fourth and fifth columns report the maximum and minimum values for each variable, respectively, describing the range of variation. This range offers valuable insights into data dispersion. For instance, the range of the "Return on Assets" variable spans from -0.24 to 0.98. Similarly, the range of the "Negative Skewness of Stock Return" variable extends from -6.52 to 6.81, indicating the presence of negative skewness in stock return distribution. This implies that the distribution tends toward higher values, but lower values appear more extremely and with greater outliers.

Before performing inferential analyses and statistical modeling, it is necessary to examine the normality of the distribution of research variables, as many classical statistical methods—particularly linear regression—are based on the assumption of normal distribution. In this study, the Kolmogorov–Smirnov (K-S) test was used to assess the normality of variable distributions. In this test, the null hypothesis assumes the normal distribution of variables. If the significance level (SIG) is less than 5%, the null hypothesis is rejected at a 95% confidence level, and the data are not considered normally distributed. Based on the obtained results, the significance level for most of the primary variables in the study was greater than 5%, and consequently, the assumption of data normality is confirmed. Thus, parametric statistical analyses such as multivariate regression can be used in the subsequent sections.

Table 2. Descriptive Statistics of Dummy Variables

Variable	Mode	Max	Min
Auditor Size	0	1	0
Auditor's Opinion	0	1	0
Financial Flexibility	0	1	0
Internal Audit	0	1	0

The above table presents descriptive statistics related to four dummy variables of the study, calculated using SPSS software. All are categorized and coded on an ordinal scale. These variables include auditor size, auditor's opinion type, degree of financial flexibility, and the presence or absence of internal auditing in companies. The values of these variables range from 0 to 1, and the mode (most frequent value) in all cases is 0, indicating the high frequency of the base state in the statistical sample.

Table 3. Independent Samples t-Test Results – MSE Criterion for the Ant Colony Optimization Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	2.79	0.104	–2.822	34	0.008	–0.19905	0.00705	–0.00557
Equal variances not assumed	—	—	–2.822	28.302	0.009	–0.19905	0.00705	–0.005465

Table 4. Independent Samples t-Test Results – MAE Criterion for the Ant Colony Optimization Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	0.601	0.444	–2.550	34	0.015	–0.016672	0.006539	–0.003383
Equal variances not assumed	—	—	–2.550	32.712	0.016	–0.016672	0.006539	–0.003364

Table 5. Independent Samples t-Test Results – MSE Criterion for the Artificial Bee Colony Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	4.173	0.049	–2.146	34	0.008	–0.0152667	0.0071155	–0.000806
Equal variances not assumed	—	—	–2.146	28.302	0.009	–0.0152667	0.0071155	–0.000693

Table 6. Independent Samples t-Test Results – MAE Criterion for the Artificial Bee Colony Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	2.702	0.109	–2.564	34	0.015	–0.0184833	0.007209	–0.0038328
Equal variances not assumed	—	—	–2.564	30.468	0.016	–0.0184833	0.007209	–0.003770

Table 7. Independent Samples t-Test Results – MSE Criterion for the Biogeography-Based Optimization Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	0.501	0.484	–2.764	34	0.009	–0.0174056	0.0062976	–0.0046072
Equal variances not assumed	—	—	–2.764	33.385	0.009	–0.0174056	0.0062976	–0.0045985

Table 8. Independent Samples t-Test Results – MAE Criterion for the Biogeography-Based Optimization Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	19.094	0.000	1.681	34	0.102	–0.0263611	0.0156813	–0.0582929
Equal variances not assumed	—	—	1.681	19.001	0.109	–0.0263611	0.0156813	–0.0591822

Table 9. Independent Samples t-Test Results – MSE Criterion for the Firefly Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	10.629	0.003	1.533	34	0.134	–0.0195611	0.0127579	–0.0454883
Equal variances not assumed	—	—	1.533	20.947	0.140	–0.0195611	0.0127579	–0.0460967

Table 10. Independent Samples t-Test Results – MAE Criterion for the Firefly Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	8.407	0.007	0.186	34	0.854	0.0016667	0.0086585	0.0198725
Equal variances not assumed	—	—	0.186	20.375	0.854	0.0016667	0.0086585	–0.0203317

Table 11. Independent Samples t-Test Results – MSE Criterion for the Harmony Search Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	11.547	0.002	0.122	34	0.903	0.0012389	0.0101250	0.0218154
Equal variances not assumed	—	—	0.122	23.560	0.903	0.0012389	0.0101250	0.0221566

Table 12. Independent Samples t-Test Results – MAE Criterion for the Harmony Search Algorithm

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	23.427	0.000	–2.179	178	0.031	–0.00888	0.0040754	0.0008378
Equal variances not assumed	—	—	–2.179	121.197	0.031	–0.00888	0.0040754	–0.0008119

Table 13. Independent Samples t-Test Results – MSE Criterion for All Algorithms

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	17.507	0.000	–2.246	178	0.026	–0.0086856	0.0038675	0.0010532
Equal variances not assumed	—	—	–2.246	135.584	0.026	–0.0086856	0.0038675	0.0010371

Table 14. Independent Samples t-Test Results – MAE Criterion for All Algorithms

	Levene's Test for Equality of Variances		t-Test for Equality of Means					
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Upper
Equal variances assumed	17.507	0.000	–2.246	178	0.026	–0.0086856	0.0038675	0.0010532
Equal variances not assumed	—	—	–2.246	135.584	0.026	–0.0086856	0.0038675	0.0010371

Based on the results presented in the tables above, it can be concluded that artificial intelligence optimization algorithms are highly capable in predicting the probability of negative stock returns. This capability is significantly higher in some algorithms, such as Ant Colony Optimization. However, to arrive at a comprehensive conclusion regarding the performance of AI optimization algorithms, a final analysis was conducted in the form of a comparative evaluation between these approaches. The results of the statistical tests for the prediction error indicators—Mean Squared Error (MSE) and Mean Absolute Error (MAE)—are presented in Tables 13 and 14. Table 13 reports the results of the independent samples t-test for the MSE metric. According to the table, the significance level of Levene's test was reported as zero, which is less than the alpha level of 0.05. Therefore, the null hypothesis of equal variances is rejected. Consequently, the second row of the table (assuming unequal variances) must be used for the t-test analysis. In this row, the significance value is 0.031, which is less than 0.05, indicating a statistically significant difference between the means of the two groups at the 90% confidence level. Moreover, the 95% confidence interval for the mean difference lies entirely in the negative region; that is, both the upper and lower bounds are negative. This result indicates that the mean error of the artificial intelligence algorithms is significantly lower.

Table 14 also presents the results of the independent samples t-test for the MAE metric. Again, the significance level of Levene's test is zero, rejecting the assumption of equal variances. Thus, the t-test analysis is based on the second row of the table. The significance value (Sig) is reported as 0.026, which is less than 0.05 and reflects a statistically significant difference between the two groups. Similar to the previous result, the 95% confidence interval for the mean difference also lies entirely in the negative region. This indicates that the mean MAE for the AI optimization algorithms is significantly lower. Overall, the results of the two key error indicators, MSE and MAE, suggest that artificial intelligence algorithms are highly effective in predicting the probability of negative stock returns. These algorithms, due to their use of global search structures, ability to adapt to nonlinear patterns, and avoidance of local optima, have emerged as effective tools for modeling the complexities embedded in financial data. These findings emphasize that employing AI optimization algorithms as decision-support tools can be highly beneficial for financial analysts, accountants, and corporate managers, as reducing prediction error facilitates more accurate identification of future risks.

To conduct a final performance evaluation of the proposed model based on AI optimization algorithms and to assess its robustness under different economic and political conditions, data from 101 companies listed on the Tehran Stock Exchange during the period 2023–2024 were also analyzed. This step aimed to evaluate the validity of the model in a period distinct from the original timeframe (2010–2015), in order to assess its generalizability. It should be noted that the initial period (2010–2015) was selected due to its particularly significant economic and financial characteristics. These include:

- Severe currency fluctuations and structural inflation during the specified period;
- International economic sanctions and disrupted access to global markets;
- Industrial recession and declining domestic demand;
- And finally, the outbreak of the COVID-19 pandemic toward the end of the period, which had a notable impact on corporate profitability and capital market behavior.

These factors turned that period into a realistic, high-risk environment well-suited for evaluating a model that predicts the probability of negative stock returns. However, in order for the designed model not to be limited to those specific circumstances, and to ensure its adaptability across various periods, it was decided to assess the same model in recent years—namely 2023 and 2024. This period also exhibits unique economic challenges, including:

- Accelerated liquidity growth and high inflation rates;
- Continued sanctions at newer and more complex levels;
- Severe exchange rate fluctuations and government stabilization policies;
- Domestic and international political uncertainties surrounding the future of Iran's economy;
- And structural changes in financial reporting and auditing systems, which have influenced conservative behavior in firms.

Considering all these challenges, the same prediction methods based on artificial intelligence optimization algorithms and multivariate regression were applied to the data of 101 companies during 2023–2024. The MSE and MAE error criteria were then calculated, and an independent samples t-test was conducted to compare the mean errors between the two methods. The results were relatively close.

Test on the Ant Colony Optimization (ACO) Algorithm

In the independent t-test analysis based on the MSE criterion, Levene's test showed a value of 0.004, clearly rejecting the assumption of equal variances. As a result, the second row of the t-test table (assuming unequal variances) was used for analysis. The significance value (Sig) of the t-test was 0.000. Specifically, the confidence interval for the mean difference was entirely negative, indicating the excellent performance of the Ant Colony Optimization algorithm in reducing prediction error for the probability of negative stock returns. A similar test was conducted based on the MAE criterion. Levene's test showed a value of 0.025, again rejecting the equal variances assumption. The t-test confirmed a significant difference between means at the 95% confidence level. The confidence interval for the mean difference was again entirely negative, reflecting the high accuracy of the Ant Colony Optimization algorithm. These findings demonstrate that the ACO algorithm performed significantly well under both MSE and MAE criteria. Furthermore, similar results across both the main evaluation period (2010–2015) and the validation period (2023–2024) confirm that the ACO-based model possesses high stability and generalizability and remains reliable in Iran's dynamic and challenging economic environment.

Test on the Artificial Bee Colony (ABC) Algorithm

An analysis of the MSE metric showed that Levene's test reported a value of 0.008, which is below the 0.05 significance threshold, thus rejecting the assumption of equal variances. Therefore, the t-test analysis was conducted based on the second row (unequal variances). The significance value of the t-test was reported as 0.000. The confidence interval for the mean difference was entirely negative, highlighting the strong performance of the Artificial Bee Colony algorithm in predicting the probability of negative stock returns. For the MAE metric, Levene's test value was 0.013, again rejecting the assumption of equal variances. The t-test confirmed a significant difference between the means at the 95% confidence level. The negative confidence interval for the mean difference indicated that the ABC algorithm achieved substantial accuracy in predicting the risk of negative returns. Thus, this data confirms that the Artificial Bee Colony algorithm demonstrated strong and stable performance across both MSE and MAE indicators. Moreover, the alignment of results from the recent period with those of the main evaluation period (2010–2015) indicates the algorithm's generalizability under various economic conditions.

Test on the Biogeography-Based Optimization (BBO) Algorithm

Levene's test in the MSE analysis indicated a value of 0.002, clearly rejecting the assumption of equal variances. Accordingly, the t-test based on unequal variances was performed and showed a significance level of 0.000. The confidence interval for the mean difference was completely negative, indicating the good performance of the BBO algorithm in reducing prediction errors. For the MAE metric, Levene's test result was 0.021, again rejecting the equal variance assumption. The t-test confirmed the mean difference at the 95% confidence level. The negative confidence interval also indicated good predictive accuracy. Therefore, the results for the Biogeography-Based Optimization algorithm demonstrate the method's capability to deliver reliable predictions. Furthermore, the repetition of this favorable performance across both 2010–2015 and 2023–2024 attests to the algorithm's dependability and adaptability under diverse economic conditions in the country.

Test on the Harmony Search (HS) Algorithm

In the MSE criterion, Levene's test showed a value of 0.006, leading to the rejection of the equal variances assumption. The t-test based on unequal variances was then conducted, yielding a significance level of 0.706 for the mean error difference, indicating moderate performance by the Harmony Search algorithm in reducing prediction errors. In the MAE analysis, Levene's test showed a value of 0.030, rejecting the equal variances assumption. However, the t-test did not confirm the difference between the means, and the confidence interval for the mean difference was not negative, suggesting that the algorithm's performance was not sufficiently strong. As a result, these findings suggest that the Harmony Search algorithm did not perform effectively based on either the MSE or MAE criteria. The repetition of similar results in both time periods (2010–2015 and 2023–2024) validates the consistency of these findings.

Test on the Firefly Algorithm (FA)

Levene's test for the MSE criterion showed a value of 0.005, rejecting the assumption of equal variances. Thus, the t-test assuming unequal variances was conducted, and the significance value was 0.902. The confidence interval for the mean difference was entirely negative, suggesting lower accuracy for the Firefly Algorithm in predicting the probability of negative stock returns. In the MAE criterion, Levene's test value was 0.018, again rejecting equal variances. The t-test did not confirm a significant mean difference, reporting a value of 0.71. The positive confidence interval further reflects the algorithm's low performance. These results indicate that the Firefly Algorithm did not demonstrate acceptable predictive accuracy. Nevertheless, the algorithm showed consistent behavior in both time periods (2010–2015 and 2023–2024) in terms of stability and adaptability.

4. Discussion and Conclusion

The findings of this study provide robust evidence that artificial intelligence (AI) optimization algorithms offer significant advantages in predicting the probability of negative stock returns in the Tehran Stock Exchange. Among the various algorithms tested—namely Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Biogeography-Based Optimization (BBO), Harmony Search (HS), and Firefly Algorithm (FA)—the ACO algorithm demonstrated the highest level of predictive accuracy across both MSE and MAE criteria. These results align with the outcomes reported in previous empirical work which emphasized the capacity of population-based metaheuristics, particularly those inspired by swarm intelligence, to navigate large and nonlinear search spaces with greater effectiveness than traditional models [4, 5].

The statistical analysis indicated that for ACO, both the t-tests based on MSE and MAE metrics resulted in significantly lower error values compared to the baseline, and the confidence intervals for mean differences were entirely in the negative range. This suggests that ACO has a strong ability to generalize under different market conditions and is particularly adept at learning from high-dimensional, imbalanced financial datasets. Similar patterns were observed for ABC and BBO algorithms, both of which demonstrated strong and consistent performance in minimizing prediction error. These findings are in line with the work of [10], who showed that nature-inspired algorithms can maintain stability and predictive accuracy even under volatile financial conditions, such as those commonly seen in the Iranian capital market.

In contrast, the results for the HS and FA algorithms were relatively weaker. The Harmony Search algorithm, in particular, failed to produce statistically significant improvements in prediction error, suggesting a limited ability to capture the underlying nonlinear dependencies that characterize financial return data. These results reinforce prior conclusions from studies emphasizing that the efficiency of optimization algorithms can vary significantly depending on problem structure, parameter calibration, and the quality of training data [3, 6].

One of the strengths of the current study lies in its dual-period design, encompassing two distinct economic timeframes: 2010–2015 and 2023–2024. The former is marked by macroeconomic turbulence, including high inflation, sanctions, and the outbreak of COVID-19, while the latter features renewed geopolitical instability, currency volatility, and evolving financial reporting standards. The consistent performance of the ACO and ABC algorithms across both periods underscores their robustness and adaptability to changing market dynamics. This supports previous arguments suggesting that AI models, particularly those enhanced with optimization routines, offer long-term strategic utility in volatile markets [12, 13].

Another significant contribution of this research is the use of Particle Swarm Optimization (PSO) for feature selection prior to model training. This technique helped reduce the dimensionality of the dataset and ensured that only the most relevant features were used for prediction, thereby reducing the risk of overfitting. The integration of PSO with classification algorithms is widely supported in the literature as a means to enhance learning speed and model generalizability [7, 8].

Additionally, the adoption of both parametric (t-tests) and nonparametric evaluation metrics such as ROC curves and F1-scores provided a multidimensional assessment of model effectiveness. This approach aligns with the best practices in AI model evaluation, which call for comprehensive, multi-criteria validation to ensure that models perform well not only on accuracy but also on recall, precision, and class discrimination [1, 14].

The results further demonstrate that AI algorithms can outperform traditional econometric techniques in high-risk market environments. Classical models such as multivariate regression often rely on strict statistical assumptions, including linearity and normality, which are rarely satisfied in financial time series data [9, 11]. In contrast, AI optimization techniques can adapt to data irregularities and exploit nonlinear relationships, making them better suited for tasks such as predicting stock market downturns and identifying periods of heightened risk [15].

In the context of emerging markets like Iran, the importance of accurate negative return prediction is magnified due to heightened sensitivity to external shocks, policy uncertainty, and investor behavior influenced by political events. The findings of this study suggest that implementing AI-based prediction models can significantly contribute to the development of early warning systems, improve investment decision-making, and enhance the analytical capacity of institutional investors and regulators [2, 6].

These results are also aligned with broader trends in financial technology and digital finance, which are rapidly incorporating AI tools into investment strategies, portfolio optimization, and fraud detection [1]. The increased deployment of AI in global financial systems is reflected in bibliometric analyses highlighting the surge in publications and citations related to AI and stock prediction over the past decade [12].

The case of the Tehran Stock Exchange presents an especially compelling context for applying and testing these algorithms. With its unique combination of regulatory volatility, macroeconomic instability, and diverse industrial sectors, the TSE serves as a rigorous testbed for validating the generalizability and resilience of AI-based financial models. Previous efforts to use genetic algorithms in portfolio selection in this market have shown promise but were limited in their ability to capture short-term downside risk [9]. This study builds upon and extends such work by incorporating more sophisticated algorithmic frameworks that are both adaptive and robust.

Overall, the results point to a clear hierarchy among the tested algorithms. The Ant Colony Optimization algorithm emerges as the most effective, followed closely by Artificial Bee Colony and Biogeography-Based Optimization. These findings mirror prior research in which ACO consistently delivered superior results in pattern recognition, financial signal interpretation, and portfolio stress testing [4, 5].

Despite its contributions, this study is not without limitations. First, the research is confined to the Tehran Stock Exchange and may not reflect the dynamics of other emerging or developed markets. The specific regulatory, political, and economic conditions in Iran could influence model behavior in ways that are not generalizable. Second, although the study utilizes two economic periods for validation, both are within the same national context. Broader cross-country comparisons could enhance the external validity of the findings. Additionally, while five optimization algorithms were evaluated, there are numerous other techniques—including hybrid models and deep learning frameworks—that could offer competitive or superior performance but were beyond the scope of this research. Finally, the impact of real-time macroeconomic indicators, investor sentiment, or news-based variables was not included in the model inputs, though these are known to significantly affect market movements.

Future research could expand on this study by applying the same models to other financial markets in the Middle East, Asia, or Latin America to assess their generalizability. Exploring hybrid models that combine machine learning with deep neural networks and sentiment analysis could provide even more nuanced and accurate predictions. Further, the integration of high-frequency trading data and macroeconomic indicators could refine the timing and sensitivity of predictive models. Researchers should also explore ensemble approaches that dynamically select the best-performing model based on market conditions. Finally, the ethical and interpretability dimensions of using AI in financial decision-making—such as algorithmic transparency and fairness—should be examined more closely in future studies.

From a practical perspective, financial institutions, investment firms, and regulatory bodies in Iran and other emerging markets should consider incorporating AI optimization algorithms into their analytical infrastructure. These models can serve as valuable tools for early warning systems, enabling timely interventions in the face of potential financial downturns. Additionally, integrating such tools into decision support systems can enhance the efficiency of risk assessment, portfolio management, and audit analytics. Firms that operate in high-volatility sectors may particularly benefit from the proactive insights provided by such predictive systems. Lastly, educational and training programs in finance and accounting should incorporate modules on AI-based modeling to prepare the next generation of financial analysts and decision-makers.

Authors' Contributions

Authors equally contributed to this article.

Ethical Considerations

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

Acknowledgments

Authors thank all participants who participate in 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

[1] S. Ahmed, M. M. Alshater, A. E. Ammari, and H. Hammami, "Artificial intelligence and machine learning in finance: A bibliometric review," *Research in International Business and Finance,* vol. 61, p. 101646, 2022/10/01/ 2022, doi: 10.1016/j.ribaf.2022.101646.

[2] F. Alı and P. Surı, "A Bibliometric Analysis of Artificial Intelligence-Based Stock Market Prediction," *The Eurasia Proceedings of Educational and Social Sciences,* vol. 27, pp. 17-35, 2022. [Online]. Available: https://dergipark.org.tr/en/download/article-file/2846587.

[3] F. Razavi and M. Ghadari, "Application of Artificial Intelligence Optimization Algorithms in Predicting Stock Market Volatility," *Journal of Economic and Financial Sciences,* vol. 11, no. 3, pp. 45-60, 2020.

[4] H. Nikrosh and E. Yousefi, "Comparative Analysis of Genetic Algorithms and Ant Colony Optimization in Predicting Stock Returns," *Iranian Journal of Financial and Economic Research,* vol. 16, no. 4, pp. 75-90, 2019.

[5] J. Behera, A. K. Pasayat, H. Behera, and P. Kumar, "Prediction based mean-value-at-risk portfolio optimization using machine learning regression algorithms for multi-national stock markets," *Engineering Applications of Artificial Intelligence,* vol. 120, p. 105843, 2023/04/01/ 2023. [Online]. Available: https://doi.org/10.1016/j.engappai.2023.105843.

[6] M. N. Ashtiani and B. Raahemi, "News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review," *Expert Systems with Applications,* vol. 217, p. 119509, 2023/05/01/ 2023. [Online]. Available: https://doi.org/10.1016/j.eswa.2023.119509.

[7] J. Chen, Y. Wen, Y. A. Nanehkaran, M. D. Suzauddola, W. Chen, and D. Zhang, "Machine learning techniques for stock price prediction and graphic signal recognition," *Engineering Applications of Artificial Intelligence,* vol. 121, p. 106038, 2023/05/01/ 2023. [Online]. Available: https://doi.org/10.1016/j.engappai.2023.106038.

[8] S. Hosseini-Pour and M. Kazemi, "Application of Multivariate Regression and Optimization Algorithms in Predicting Stock ReturnsJO - Quarterly Journal of Advanced Financial and Accounting," vol. 7, no. 3, pp. 135-150, 2018.

[9] H. R. Navidi, A. Nejoomi Markid, and H. Mirzazadeh, "Portfolio Selection in Tehran Stock Exchange Market with a Genetic Algorithm," (in en), *Journal of Economic Research (Tahghighat- E- Eghtesadi),* vol. 44, no. 4, 2010. [Online]. Available: https://jte.ut.ac.ir/article_20348.html.

[10] K. Mousavi and M. Gholami, "Evaluating the Performance of Artificial Intelligence Optimization Algorithms in Predicting Negative Stock Returns," *Journal of Financial and Investment Research,* vol. 10, no. 2, pp. 98-112, 2021.

[11] M. Razaie, M. Nazemi Ardakani, and a. naser sadrabadi, "Predicting financial statement fraud using The CRISP approach," (in en), *Journal of Management Accounting and Auditing Knowledge,* vol. 10, no. 40, pp. 135-150, 2021. [Online]. Available: https://www.jmaak.ir/article_18271.html
https://www.jmaak.ir/article_18271_6353ae9d5675a7b7b0a8c5419d6c2f3c.pdf.

[12] C. Y. Lin and J. A. Marques, "Stock market prediction using artificial intelligence: A systematic review of systematic reviews," *Social Sciences & Humanities Open,* vol. 9, no. 1, p. 100864, 2024, doi: 10.1016/j.ssaho.2024.100864.

[13] S. Shaghaghi Shahri, "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,* vol. 6, no. 2, pp. 31-44, 2024. [Online]. Available: https://www.transoscience.ir/article_200776.html.

[14] M. L.Lima *et al.*, "Using Sentiment Analysis for Stock Exchange Prediction," *International Journal of Artificial Intelligence & Applications,* vol. 7, no. 1, pp. 59-67, 2016, doi: 10.5121/ijaia.2016.7106.

[15] S. H. Vaghfi, "Application of artificial intelligence algorithms in predicting bankruptcy using macroeconomic and accounting variables in companies listed on the Tehran Stock Exchange," *Quarterly of Decision Making and Operations Research,* vol. 4, no. 2, 2019.