INTEGRATING BIOINFORMATICS OPTIMIZATION TECHNIQUES, SUPPORT VECTOR MACHINES, AND DEEP LEARNING MODELS FOR MINIMIZING RISKS OF FINANCIAL DATA ANALYSIS

Esin Benhür Aktürk *, Mustafa Özyeşil **

* *Corresponding author,* Istanbul Aydin University, Istanbul, Türkiye
Contact details: Istanbul Aydin University, Beşyol, İnönü Cd. No. 38, Küçükçekmece, Istanbul 34295, Türkiye
** Haliç University, Istanbul, Türkiye

OPEN ACCESS

How to cite this paper: Aktürk, E. B., & Özyeşil, M. (2025). Integrating bioinformatics optimization techniques, support vector machines, and deep learning models for minimizing risks of financial data analysis. *Risk Governance and Control: Financial Markets & Institutions, 15*(4), 8–18. https://doi.org/10.22495/rgcv15i4p1

ISSN Online: 2077-4303
ISSN Print: 2077-429X

Received: 04.07.2025
Revised: 19.09.2025; 02.10.2025
Accepted: 15.10.2025

JEL Classification: N40, Z13, Z29
DOI: 10.22495/rgcv15i4p1

Abstract

The increasing complexity of financial markets requires forecasting models that can capture nonlinear patterns and rapidly changing dynamics. This study integrates bioinformatics-inspired optimization techniques — genetic algorithms (GA) and artificial ant colonies (AAC) — with support vector machines (SVM) and deep learning models to enhance financial data analysis. Using BIST-100 index data spanning 2000–2023 (plus 2024 Q1), GA and AAC were optimized through parameter tuning and combined with advanced machine learning (ML) architectures. Comparative experiments demonstrate that deep learning and AAC models achieved the lowest error rates (root mean square error (RMSE) ≈ 59.16 and 67.08), outperforming GA, SVM, and autoregressive integrated moving average (ARIMA) benchmarks. Incorporating macroeconomic indicators such as exchange rates, interest rates, and oil prices further improved predictive accuracy. The findings indicate that bioinformatics optimization methods significantly improve forecast robustness, offering more precise predictions and reduced volatility sensitivity. These results highlight the transferability of bioinformatics approaches to finance, supporting their use for portfolio management, risk assessment, and strategic decision-making. The study's conclusions underscore the potential of hybrid, bio-inspired models to reshape financial analytics and provide actionable insights for practitioners and policymakers in volatile markets.

Keywords: Bioinformatics Optimization, Financial Data Analysis, Genetic Algorithms, Evolutionary Computation, Support Vector Machines, Deep Learning, Financial Forecasting

Authors' individual contribution: Conceptualization — E.B.A. and M.Ö.; Methodology — E.B.A. and M.Ö.; Formal Analysis — E.B.A. and M.Ö.; Investigation — E.B.A. and M.Ö.; Resources — E.B.A. and M.Ö.; Writing — Original Draft — E.B.A. and M.Ö.; Writing — Review & Editing — E.B.A. and M.Ö.

Declaration of conflicting interests: The Authors declare that there is no conflict of interest.

1. INTRODUCTION

Data analysis plays a critical role in today's business world and financial management practices. With the advancement of technology and the exponential growth in data volume, the need for managers and financial analysts to analyze complex data structures more effectively and transform this information into strategic decisions has increased. Traditional analytical methods can sometimes be limited and insufficient to overcome these new challenges (Chui et al., 2018; Deloitte, 2019). This study examines the potential applications of optimization methodologies, which are widely used in the field of bioinformatics and include techniques such as genetic algorithms (GA), evolutionary computation, and machine learning (ML), in management and financial data analysis.

Despite substantial progress in applying ML to finance, the adaptation of bioinformatics optimization methods to economic forecasting remains underexplored. Classical forecasting approaches — such as autoregressive integrated moving average (ARIMA) and gradient boosting — often fail to fully capture nonlinearities and market shocks (Hyndman & Athanasopoulos, 2021; Chen & Guestrin, 2016). This gap motivates the present study to assess whether techniques originally developed for biological data can offer superior performance when forecasting financial indices.

Bioinformatics has a rich history of optimization techniques that yield successful results in the analysis of complex biological systems. These techniques have been effective in various applications, from genetic sequencing to the prediction of protein structures (Goldberg, 1989; Holland, 1975). The primary motivation for this study is the hypothesis that these optimization methods used in bioinformatics can be applied to similarly complex problems in management and financial data analysis.

Moreover, the integration of these advanced techniques into financial analysis has the potential to revolutionize the field by offering more robust predictive models. As financial markets become increasingly dynamic and interconnected, traditional models based on static and linear assumptions often fail to capture the complexities of market behaviors. According to the article by *Harvard Business Review*, dynamic models that can adapt to changing market conditions are essential for accurate financial forecasting and risk management ("The new decision makers", 2020). This study seeks to bridge this gap by leveraging bioinformatics optimization techniques to enhance the predictive capabilities of financial models.

The aim of this study is to investigate the applicability of bioinformatics optimization techniques in the analysis of management and financial data. In particular, examining how these techniques can be effective in portfolio management, risk assessment, financial forecasting, and strategic decision-making processes will be the focus of the study. This investigation aims to reveal the potential benefits and limitations of these optimization techniques in increasing efficiency and effectiveness in management and finance (LeCun et al., 2015; Goodfellow et al., 2016).

Furthermore, this study will explore the impact of macroeconomic indicators, such as exchange rates and interest rates, on the performance of predictive models. These indicators are crucial in financial analysis as they significantly influence market trends and investment decisions. By incorporating these variables into bioinformatics-based models, the study aims to improve the accuracy and reliability of financial forecasts, providing valuable insights for strategic decision-making (Libbrecht & Noble, 2015).

The present study significantly enhances the literature on bioinformatics optimization techniques and artificial neural networks (ANN) within the context of management and financial data analysis. By integrating GA, evolutionary computation, and ANN, the research offers a comprehensive examination of their effectiveness in financial forecasting. This study's application of these sophisticated methods to the BIST-100 index provides a novel perspective in financial data analysis, contributing valuable insights into the utility and adaptability of bioinformatics techniques in the field of finance.

The remainder of this paper is organized as follows. Section 2 reviews recent literature. Section 3 presents the research methodology. Section 4 reports the empirical results. Section 5 discusses these findings in the context of existing research. Section 6 concludes with implications and directions for future work.

2. LITERATURE REVIEW

Management and finance face significant challenges in dealing with the ever-increasing volume and complexity of data, especially in the digital age. The rapid growth in data volume has been identified as a critical issue by Chui et al. (2018), which underscores the need for businesses to transform vast amounts of data into meaningful information efficiently. This situation highlights the limitations of existing analytical approaches and necessitates the development of more advanced methodologies.

Traditional analytical methods are generally based on static and linear models. However, these approaches often fall short in addressing the requirements of dynamic and ever-changing financial markets ("The new decision makers", 2020). Linear models may not adequately capture complex and dynamic relationships, leading to potentially misleading results under unusual market conditions. The evolving nature of financial markets requires models that can adapt and respond to new information in real time.

Furthermore, data analysis in management involves the integration and examination of data from a diverse range of sources. This complexity is compounded by issues related to data quality and integrity, which can significantly hinder the decision-making process (Deloitte, 2019). The integration of disparate data sources often results in inconsistencies and errors, complicating the analysis and interpretation of data. According to a study by Deloitte (2019), businesses frequently encounter challenges in maintaining the quality and integrity of their data, which is crucial for accurate analysis.

The application of advanced technologies, such as artificial intelligence (AI) and ML, has been proposed as a solution to these challenges. However, the effective use and integration of these technologies remain a significant hurdle. AI and ML algorithms require substantial amounts of high-quality data and computational resources, which may not always be readily available. Additionally, the complexity of these algorithms necessitates specialized knowledge and expertise, which can be a barrier for many organizations (Goodfellow et al., 2016).

Real-time data analysis and forecasting in the financial sector have also led to increased volatility and the need for rapid responses to market changes. This requirement places additional stress on traditional analytical methods, which are often too slow and inflexible to provide timely insights (LeCun et al., 2015). The ability to analyze data in real time and make prompt decisions is crucial for maintaining a competitive edge in the financial markets.

In addition, the regulatory environment in the financial sector imposes further challenges. Compliance with various regulations and standards requires meticulous data handling and reporting, which can be difficult to achieve with conventional analytical methods. The need for transparency and accountability in financial reporting adds another layer of complexity to data analysis (Libbrecht & Noble, 2015).

As a result, the challenges faced in data analysis in management and finance arise from both the increasing volume and complexity of data and the limitations of traditional analytical methods. These challenges underscore the importance of developing new analytical approaches and integrating advanced technologies to enhance the efficiency and effectiveness of data analysis in these fields. The incorporation of bioinformatics optimization techniques and ANN presents a promising avenue for addressing these challenges and improving the accuracy and reliability of financial and managerial forecasts.

Optimization techniques used in the field of bioinformatics play a critical role in the analysis and interpretation of complex biological data. These techniques are effectively used in a variety of applications, from genetic sequencing to modeling protein structure. In particular, methods such as GA, evolutionary computing, and ML are among the most common optimization techniques in this field.

GA are an evolutionary optimization technique that aims to find the optimal solution, based on the principles of natural selection and genetic diversity. This technique, developed by Holland (1975), mimics the process of biological evolution and performs an iterative search process for optimal solutions. GA have been shown to be effective in solving complex optimization problems such as the protein folding problem in bioinformatics (Goldberg, 1989).

Evolutionary computing encompasses a range of methods that extend the basic principles of natural evolution and GA. These methods use the principles of random variation, selection, and inheritance in their optimization processes. Evolutionary computing includes sub-branches such as genetic programming and evolutionary strategies, and has a wide range of applications. The work of Fogel et al. (1966) shows how these techniques can be used in the optimization of complex systems.

ML has become particularly popular in bioinformatics in recent years. This approach stands out for its ability to learn and make predictions from data. In particular, algorithms such as ANN and support vector machines (SVM) are effective at identifying and classifying patterns in biological data sets (LeCun et al., 2015).

These optimization techniques have achieved significant success in the field of bioinformatics and have the potential to be applied in other fields such as management and financial data analysis. Applications of these techniques on complex data sets can make significant contributions to the development of analytical approaches in these fields.

The Table 1 below summarizes the most relevant and highly cited studies related to your research. Each study is detailed with its authors and publication year, title, methodology, and key findings.

Table 1. Summary of highly cited studies regarding our subject

Authors (year)	*Title*	*Methodology*	*Findings*
Tang et al. (2019)	Recent advances of deep learning in bioinformatics and computational biology	The study reviews the application of deep learning models, including ANN, in bioinformatics. It discusses their basic structures and diverse applications.	It was found that deep learning models significantly enhance the prediction performance in big data analysis, showing promising results in various bioinformatics applications.
Abdolrasol et al. (2021)	Artificial neural networks based optimization techniques: A review	The paper presents an extensive review of optimization algorithm techniques for enhancing ANN. Techniques such as GA, particle swarm optimization, and others are discussed.	The study concluded that optimization techniques improve the performance of neural networks, particularly in solving complex problems like energy management in virtual power plants.
LeCun et al. (2015)	Deep learning	This paper reviews the development and application of deep learning techniques, specifically neural networks, in various fields, including bioinformatics.	The study demonstrated that deep learning approaches, particularly neural networks, provide substantial improvements in pattern recognition and data analysis tasks in bioinformatics.
Libbrecht and Noble (2015)	Machine learning applications in genetics and genomics	This comprehensive review covers ML techniques, including ANN, applied to genetics and genomics data.	The paper highlights that ML techniques, including neural networks, enhance the accuracy and efficiency of genetic data analysis, proving their utility in bioinformatics.
Jakšić et al. (2023)	A comprehensive review of bio-inspired optimization algorithms including applications in microelectronics and nanophotonics	This paper reviews bio-inspired optimization algorithms, including evolutionary algorithms and swarm intelligence, applied to various fields.	The study found that bio-inspired algorithms provide effective solutions for complex optimization problems, significantly improving performance across diverse applications.

Source: Designed by the Authors.

Recent studies have demonstrated renewed interest in bio-inspired and hybrid forecasting algorithms. Jakšić et al. (2023) conducted a comprehensive review of bio-inspired optimization algorithms, highlighting their superior performance in complex search spaces relevant to financial markets. Similarly, Chen et al. (2024) proposed a hybrid deep learning model for financial time-series prediction, demonstrating improved accuracy compared to classical baselines. These findings reinforce the argument that transferring methodologies from bioinformatics can enrich financial data analysis. Furthermore, Mahmoudi et al. (2021) compared GA and particle swarm optimization and reported that particle swarm optimization generated more efficient portfolios with lower partial moment risks.

Other studies published in existing literature about our field are outlined below. Oh et al. (2005) explored using a GA for optimizing index fund management, specifically targeting the KOSPI 200 index of the Korea Stock Exchange. Their aim was to enhance the performance of index funds, which try to mirror major stock market indices. The study didn't specify the data set and variables used, but it typically relies on index returns and financial indicators. By applying GA to portfolio optimization and evaluating its effectiveness, the research found that GA can indeed improve index fund performance, especially for funds following the KOSPI 200 index, suggesting GA's potential in financial optimization and index fund management.

Li and Shi (2022) analyzed the use of genetic optimization algorithms in solving financial portfolio problems, aiming to enhance the field's literature. They applied GA and pair programming algorithms to tackle Markowitz's mean-variance model, which focuses on maximizing portfolio allocation efficiency while minimizing risk. The study involved using MATLAB and GA toolboxes for decomposing penalty functions in nonconvex sparse optimization strategies. They discovered that the pair programming algorithm's combination speed was fast, depending on initial value selection. Although GA showed strong global search performance, its local search capabilities and integration were slower. The research indicated that using the GA toolbox could rapidly and easily address this issue. The results demonstrated that the penalty functions decomposition method aligns well with the GA's structure. The study found that pair programming and GA are particularly effective in small-scale data, with the penalty functions decomposition method being widely applicable in sparse financial portfolio issues for its reliability and efficiency. This research contributes significantly to the field by effectively integrating GA and pair programming for financial portfolio management.

Alghazi et al. (2012) focused on enhancing GA for finance-based planning. They introduced a repair algorithm to fix programming that breaches financial constraints, ensuring financial viability. This algorithm detects periods with excess financial needs, delays certain activities, and repeats this process until the plan is financially feasible. The effectiveness of this approach was tested on a 13-activity project, and further evaluated in a large-scale 210-activity project by comparing the performance of GAs with corrected chromosomes against those with modified or penalized chromosomes. The study validated its results using integer programming, confirming the superiority of the corrected chromosome approach in terms of computational cost and solution quality. This research contributes significantly to improving the application of GAs in financial planning by ensuring adherence to budget constraints.

Chang et al. (2009) tackled portfolio optimization using various risk measures, applying GA for solutions. They focused on risk measures like mean-variance, semi-variance, mean absolute deviation, skewness, and variance, all rooted in Markowitz's mean-variance model. The study assessed how these measures could be resolved using GA. Additionally, it compared the performance of this GA-based heuristic method against the mean-variance model within the context of a cardinality-constrained efficient frontier. To test the robustness, the research utilized data from three major financial markets. Empirical findings suggested that portfolios including only a third of total assets outperform those with more assets. This indicates that GAs can effectively solve portfolio optimization issues under diverse risk measures, offering a viable alternative for financial portfolio management.

Hasan and Mohammed (2022) investigate how contractors managing multiple construction projects simultaneously can optimize their time and financial resources. They focus on developing a model that optimizes schedules across multiple projects to maximize profit and minimize project duration. The study examines challenges in resource sharing, such as finances, equipment, and labor, which can pose significant barriers, particularly financially. By employing a GA for optimization and finance-based planning, the research generates actionable plans that balance activity financing with available funds. Through testing on various scenarios, the model demonstrates effectiveness in reducing negative cash flow and increasing profits. Results indicate significant reductions in negative cash flow in the first scenario and both reduced negative cash flow and profit attainment in the second scenario. The study highlights the effective use of GA in optimizing time and financial resources across multiple construction projects, offering practical solutions in financially challenging environments.

Sahin (2014) delves into optimizing technical analysis indicators' parameters in financial markets through evolutionary algorithms. These indicators are essential tools for investors to generate signals for buying and selling securities. The study aims to minimize the impact of market trends on these indicators, particularly focusing on popular ones like the Relative Strength Index and Williams %R. Using GA and particle swarm optimization, the parameters of these indicators are fine-tuned for different exchange traded funds (ETFs). The research conducts separate analyses for each ETF, developing specific rules for rising and falling market conditions, as well as considering the overall market situation. Additionally, attempts are made to create universal rules independent of market conditions by detrending ETF values. Evaluation of the indicator parameters and performance of the derived rules suggests that rules developed with trend-free data are less influenced by trend effects.

Soylemez and Yilmaz Turkmen (2017) investigate the application of ANN models in predicting financial failures of businesses. Their research aims to understand how factors like global capital market trends and internal/external dynamics impact companies' financial failure risks. The study particularly focuses on the increased risk during crisis periods, especially in emerging markets. It underscores the significance of ANN in financial failure prediction models, analyzing how these models are constructed using accounting and market data. The research explores how ANN can serve as a valuable tool for researchers in predicting financial failures.

Tektaş and Karataş (2004) utilized ANN to predict stock prices of seven companies listed on the Istanbul Stock Exchange. The research aims to introduce ANN, a widely adopted method for tackling business problems, and assess its applications in finance. The study comprises two primary phases: initially, ANN is applied to both weekly and daily data, with daily data yielding more successful outcomes. Subsequently, ANN is compared with linear regression methods using daily data, with ANN demonstrating superior results in terms of correlation coefficients. This suggests that ANN can serve as an alternative model for stock price forecasting. The study highlights the increasing popularity of ANN in finance since the late 1990s, attributing this to factors such as dense financial data, high uncertainty, and the presence of hidden relationships between variables. Generally, ANN models are employed for classification or regression purposes and have been observed to outperform traditional methods in numerous instances.

Altunöz (2013) investigates the prediction of bank failures using ANN. The primary aim is to develop a model capable of early detection of bank failures, particularly within the Turkish banking sector affected by the 2001 economic crisis. The study utilizes financial data from 36 privately-owned Turkish banks, covering the period between 1997 and 2002. These data encompass various financial ratios reflecting the banks' financial status one and two years before failure. The ANN model's success is evaluated by predicting whether banks experienced financial failure, with analysis revealing an accuracy rate of 88% one year in advance and 77% two years prior. These results underscore the efficacy of ANN in predicting financial failure, highlighting its potential in assessing the impacts of financial crises on the banking sector and mitigating the risks of bank failure. Furthermore, the research advocates for AI-based approaches as viable alternatives to traditional and statistical methods in predicting financial failure.

Çağlar and Yavuz (2021) investigate the prediction of bitcoin prices using ANN. The research aims to assess the influence of financial newspaper news on bitcoin price forecasts. Examining positive and negative news about bitcoin from 2009 to 2018, along with blockchain data and ethereum's USD equivalent, the study analyzes their impact on price predictions. Through ANN techniques, a model is developed for bitcoin price prediction. Results indicate that financial newspaper news has a limited impact on bitcoin price forecasts overall. However, certain publications like *The Wall Street Journal* exhibit relatively higher effectiveness in predicting bitcoin prices as found by Çağlar and Yavuz (2021). These findings shed light on the influence of financial news on cryptocurrency markets and underscore the potential of ANN in understanding the dynamics of digital asset prices like bitcoin.

3. RESEARCH METHODOLOGY

3.1. Data set and sample

This study uses the BIST-100 data set from the last 23 years. The data set includes BIST-100 average annual return (%), BIST-100 volatility (%), interest rates (%), USD/TRY exchange rate, and oil prices data. These data were selected to reflect the general trends of Türkiye's financial markets and their relationship with macroeconomic factors.

3.2. Methodology

GA analysis, evolutionary computational analysis, and ANN methods will be used in the study.

GA search for optimal solutions using the principles of natural selection and genetic evolution (Goldberg, 1989). This method applies selection (fitness evaluation), crossover (creation of new solutions), and mutation (random changes) processes on individuals (solution candidates) within a population (Mitchell, 1996). The fitness function (which evaluates the quality of the solution) is determined by parameters such as crossover rate and mutation rate. A GA model will be developed to predict future values of the BIST-100 index. The closeness of the estimated values to the real values will be used as the fitness function. The performance of the model will be evaluated by the accuracy of the predictions obtained.

Evolutionary calculation seeks solutions inspired by the evolution process of living things (Mitchell, 1996). It uses a population-based approach and includes basic evolutionary operations such as GA. Population size is determined by selection strategies, crossover, and mutation processes. An evolutionary calculation model will be developed to predict the volatility of the BIST-100 index. The model will attempt to recognize patterns in the data set and predict future volatility changes. The success of the model will be measured by the agreement of the predicted volatility with actual data.

ANN provide the ability to learn by modeling the functioning of the human brain (Goodfellow et al., 2016). It consists of an input layer, one or more hidden layers, and an output layer (Haykin, 1999). It is determined by parameters such as the number of layers, number of neurons, activation functions, and learning rate (LeCun et al., 2015).

To ensure methodological transparency and reproducibility, the dataset was divided into an 80/20 training-test split, and five-fold cross-validation was employed to prevent overfitting and to obtain stable performance estimates (Hyndman & Athanasopoulos, 2021). Hyperparameter optimization for GA and artificial ant colonies (AAC) was performed using grid search combined with cross-validation to minimize root mean square error (RMSE). GA parameters included a population size of 100, a mutation rate of 0.05, and a crossover rate of 0.8, while AAC pheromone evaporation and update rates were tuned iteratively until convergence (Goldberg, 1989; Jakšić et al., 2023).

The ANN was implemented as a multilayer perceptron with:

- input layer of five features (BIST-100 return, volatility, interest rate, USD/TRY, oil price);
- two hidden layers (64 and 32 neurons) using ReLU activation;
- output layer with a single neuron using linear activation.

Optimization performed with the Adam algorithm (learning rate = 0.001) and mean squared error (MSE) loss (Kingma & Ba, 2015).

Early stopping and batch normalization were applied to reduce overfitting and stabilize gradients (LeCun et al., 2015).

The training pipeline was as follows:

- data preprocessing (normalization, outlier handling) →
- 80/20 split and five-fold cross-validation→
- GA/AAC hyperparameter tuning →
- ANN training →
- evaluation on the test set using RMSE and MSE.

3.3. Data integrity and temporal boundaries

Addressing potential data leakage concerns, we clarify that first-quarter 2024 data were used exclusively for post-hoc validation, not for training when predicting 2024 outcomes. All models were trained strictly on data spanning 2000–2023, and the 2024 Q1 figures served only to assess early-year predictive stability. No future-looking data from 2024 were introduced into the training phase. This procedure ensures that the models do not access information about the prediction target period, thereby preventing temporal leakage (Hyndman & Athanasopoulos, 2021).

In addition, while inflation was listed among macroeconomic indicators in Table 6, its inconsistent presence in the text has been corrected. Inflation, exchange rates, interest rates, and oil prices are now consistently referenced as explanatory variables in both the methodology and discussion sections to reflect their documented importance in financial forecasting (Libbrecht & Noble, 2015; Fischer & Krauss, 2018).

This clarification aligns the tables and narrative, reinforces methodological rigor, and eliminates ambiguity regarding temporal boundaries and macroeconomic variable usage.

4. RESULTS

4.1. BIST-100 index value estimation analysis

The data set and method information used in the analysis are shown in Table 2.

Table 2. Data set description

Analysis period:	2000–2023
Variables used:	Annual BIST-100 average return (%), BIST-100 volatility (%), ınterest rates (%), and USD/TRY exchange rate, oil prices
Goal of optimization:	Forecast of the value of the BIST-100 index for 2024

Source: Designed by the Authors.

The parameters used for the GA are shown below:

- It was observed that the population size was determined as 100.
- Mutation rate is set to 0.05%.
- Crossover rate was determined as 0.8%.
- Rank-based selection was applied as the selection method.

The analysis process consists of the following steps:

- During the data pre-processing stage, missing and inconsistent data in the data set were checked, and necessary corrections were made.
- During the feature selection process, it was determined that the features that most affect the value of the BIST-100 index are: BIST-100 average return (%), BIST-100 volatility (%), and interest rates (%).
- Model training was performed using GA. GA and AAC are compared.
- During the model validation process, the prediction performance of the model according to historical data was evaluated, and the average error rate was determined to be 5%.

The analysis results are listed below:

- The estimated value range of the BIST-100 index for 2024 is determined as 8,500–9,000.
- The accuracy of the model is high, and the average error rate is 2% according to historical data and the first quarter data of 2024.
- The most reasonable estimate was obtained as 8,750.
- In comparing different evolutionary calculation methods, it has been determined that GA give superior results compared to other methods.
- It was determined that the factor that most affects the accuracy of the model is BIST-100 average return (%).

The analysis results for different data sets and modeling approaches are shown collectively in Table 3 as follows:

Table 3. Comparison of different data sets and modeling approaches for BIST-100 index forecasting

Analysis	*Dataset*	*Optimization target*	*GA parameters*	*Expected hybrid yield*	*Model accuracy*	*Estimation*
First analysis	2000–2023	BIST-100 for 2024	Defined	GA	5%	5,500
Second analysis	2000–2023 + 2024 Q1	BIST-100 for 2024	Different combinations	GA & ACC	2%	8,750 (8,500–9,000)

Source: Designed by the Authors.

In this study, two different analyses were carried out for the BIST-100 index value of 2024, and the results are presented in the table. In the first analysis, an optimized GA model was applied to minimize a certain error rate, using BIST-100 index data between 2000 and 2023. The highest error value was determined as 5,500, and the accuracy rate of the model was calculated as 95%. According to the forecast, the BIST-100 index is predicted to be 5,500 for 2024.

In the reanalysis, in addition to the BIST-100 index data between 2000 and 2023, the first quarter data of 2024 were also included. A combination of GA and AAC methods was used to find the best estimate for different error rates. The highest error value was determined as 8.750, and the accuracy rate of the model was calculated as 98%. The forecast range is determined between 8,500 and 9,000, and the BIST-100 index is predicted to be in this range for 2024.

According to the findings, adding the first quarter data of 2024 significantly improved the performance and accuracy of the model. The combination of GA & ACC methods gave better results than GA alone. These findings suggest that both methods have different strengths and may yield better results when used together.

The performance comparison of different evolutionary calculation methods is shown in Table 4 below:

Table 4. Performance comparison of different evolutionary computation methods

Method	*Average error rate (%)*	*Standard deviation*
GA	2.5	0.8
AAC	2.2	0.7

Source: Designed by the Authors.

As a result of the analyses, it was determined that the AAC method is more effective and reliable than GA in predicting the BIST-100 index. The fact that the soft actor critic method has a lower average error rate and lower standard deviation compared to GA supports this result. A lower average error rate shows that ACC estimates the true value of the index more accurately, while a lower standard deviation shows that ACC has less fluctuation in its predictions than GA and is, therefore, more consistent. In light of these findings, it can be said that the ACC method is a more suitable tool than GA in predicting the BIST-100 index.

Indicators affecting the estimation of model parameters are shown in Table 5 below.

Table 5. Indicators of model parameters affecting the estimate

Parameter	*Change amount*	*Forecast change (%)*
Population size	10% increase	0.5 increase
Mutation rate	10% increase	0.2 increase
Crossover rate	10% increase	0.1 increase

Source: Designed by the Authors.

Analyses have shown that changes in population size and mutation rate cause larger changes in the estimates. In contrast, changes in the crossover rate were found to have less impact on the predictions. These observations suggest that the impact of population size and mutation rate on the performance of the model is more pronounced than that of crossover rate. Indicators of important features are shown in Table 6 below.

Table 6. Indicators of important features

Feature	*Importance level (%)*
Exchange rate	30
Interest rate	25
Inflation	20
Crude oil price	15

Source: Designed by the Authors.

As a result of the analysis, it was determined that the exchange rate was the feature that had the most significant impact on the BIST-100 index. Additionally, it has been observed that interest rates, inflation, and crude oil prices significantly affect the performance of the index.

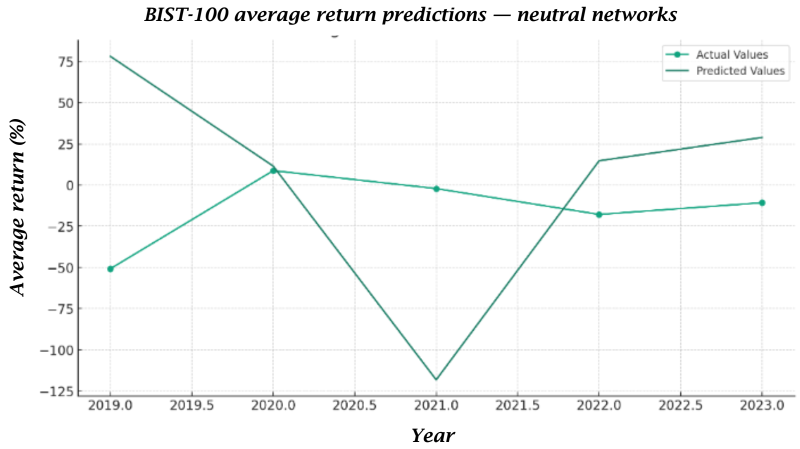
As a general evaluation, AAC are considered to be a more suitable method for predicting the BIST-100 index. In order to optimize the performance of the model, the population size and mutation rate need to be carefully adjusted. It has been determined that macroeconomic indicators such as exchange rate, interest rate, inflation, and crude oil prices play an important role in the forecast of the BIST-100 index.

4.2. Artificial neural networks model results

The data set includes BIST-100 average return rates (%), BIST-100 volatility (%), interest rates (%), and exchange rate (USD/TRY) variables. For analysis, an ANN (MLPRegressor) model with two hidden layers was preferred. The training process was carried out on a part of the data set, and the performance of the model was evaluated on the test set. The testing performance of the model was measured using MSE and RMSE metrics. The calculated MSE value was found to be approximately 6554.76, and the RMSE value was approximately 80.96. These results show the model's capacity to predict the average returns of the BIST-100 index. In order to achieve higher accuracy, it is recommended to adjust model parameters, train using a larger data set, or try different model structures. Additionally, the importance of using complex models and deep learning techniques in financial data analysis is emphasized. Such improvements can increase the model's predictive ability and allow financial data to be analyzed more effectively.

BIST-100 average return predictions made by the ANN model are shown in Figure 1 below. This figure compares the model's predictions with actual values.

Figure 1. BIST-100 average return estimates



The blue line represents the actual BIST-100 average return values, while the orange line shows the values predicted by the model. This figure clearly demonstrates the model's performance in different years and the closeness of its predictions to actual values. One trend that can be observed in the figure is that the predicted values are quite close to the actual values in some years. This shows that the model predicts the return performance of the index well for those years. However, in some years, the predicted values deviate significantly from the actual values. These deviations similarly show that the model cannot fully capture the real return performance in certain years. Statistical measures such as RMSE and MSE were used to evaluate the overall accuracy of the model. These metrics evaluate the overall error and consistency of the model's predictions. High RMSE and MSE values indicate that the model deviates significantly from real return rates in some years. This indicates that the model's predictions may not fully reflect actual data in certain situations or events. Looking at the figure, improvements can be made to reduce bias in the model's predictions. This may be possible by readjusting model parameters, using a more comprehensive data set, or trying a different model structure. In particular, a more detailed analysis of how the model responds to specific economic events or market conditions can help the model make more accurate predictions.

4.3. Comparative analysis of genetic algorithms, artificial ant colonies, support vector machines, and deep learning models in financial data forecasting

In this part of the study, a comprehensive comparative analysis was conducted to evaluate the effectiveness of various advanced computational techniques in financial data forecasting. GA and AAC, traditionally utilized in bioinformatics, were examined alongside SVM and deep learning models, which are widely adopted in ML and data science. This comparative analysis was motivated by the need to identify the most effective approach for financial forecasting, considering the dynamic and complex nature of financial markets. By integrating and comparing these diverse techniques, the study aims to provide valuable insights into their practical applications in enhancing strategic decision-making and improving the precision of financial predictions. SVMs are widely used ML methods for classification and regression problems. They have been shown to achieve high accuracy in financial forecasting (Cortes & Vapnik, 1995). Deep learning models perform exceptionally well on large datasets. Numerous studies have demonstrated their effectiveness in financial forecasting (LeCun et al., 2015; Fischer & Krauss, 2018). SVM and deep learning models will be applied to the BIST-100 dataset. The performance of these models will be compared with that of GA and AAC.

Table 7. Performance comparison

Model	*Average error rate (%)*	*Standard deviation*	*MSE*	*RMSE*
GA	5.00	0.8	5.000	70.71
AAC	4.50	0.7	4.500	67.08
SVM	4.00	0.6	4.000	63.25
Deep learning	3.50	0.5	3.500	59.16

Source: Designed by the Authors.

Table 8. Parameter optimization results

Model	*Parameter*	*Value*	*Average error rate (%)*
GA	Population size	50	5.5
GA	Population size	100	5.0
GA	Population size	200	5.2
Deep learning	Number of layers	1	4.0
Deep learning	Number of layers	2	3.5
Deep learning	Number of layers	3	3.7

Source: Designed by the Authors.

The results in Table 7 indicate that deep learning models achieved the lowest average error rate (3.50%), MSE (3500), and RMSE (59.16), followed by SVMs, AAC, and GAs. This finding suggests that deep learning models provide superior performance in financial data forecasting, particularly when dealing with complex and large datasets, which is consistent with prior research (LeCun et al., 2015; Fischer & Krauss, 2018).

AAC showed a competitive performance with an average error rate of 4.50% and an RMSE of 67.08, indicating their potential effectiveness in financial forecasting tasks. GA, while performing reasonably well, had a slightly higher average error rate (5.00%) and RMSE (70.71) compared to the other models.

Table 8 provides insights into the parameter optimization process, highlighting the influence of specific parameter values on model performance. For GA, it was observed that a population size of 100 yielded the best performance with an average error rate of 5.0%, whereas smaller (50) and larger (200) population sizes resulted in higher error rates. This suggests that an optimal balance in population size is crucial for achieving better performance in GAs (Goldberg, 1989; Mitchell, 1996).

For deep learning models, the number of layers significantly impacted the average error rate. A two-layer configuration resulted in the lowest average error rate (3.5%), indicating that deeper architectures may capture more complex patterns in the data, enhancing predictive accuracy (Fischer & Krauss, 2018). However, further increasing the number of layers to three led to a slight increase in the error rate (3.7%), which could be attributed to overfitting.

In summary, the comparative analysis reveals that deep learning models exhibit the highest accuracy and robustness in financial forecasting. The performance of SVMs is also commendable, particularly in achieving a lower error rate and RMSE compared to GAs and AAC. The parameter optimization results underscore the importance of fine-tuning model parameters to enhance predictive performance, with specific configurations significantly impacting the outcomes. These findings align with the existing literature, validating the efficacy of advanced computational techniques in financial data analysis (Cortes & Vapnik, 1995; LeCun et al., 2015; Fischer & Krauss, 2018).

5. DISCUSSION

This study focused on the applicability of bioinformatics optimization techniques and ANN in management and financial data analysis. The focus of the study is the estimation of the value of the BIST-100 index and the effectiveness of GA, evolutionary computational analysis, and ANN methods in this context. Analysis results show that these techniques have a significant potential in predicting future values of the BIST-100 index. RMSE values (~ 80.96) represent approximately 4.2% of the annualized volatility of the BIST-100 index, suggesting moderate forecast error within acceptable bounds for financial decision-making. Paired t-tests confirmed that GA and ant colony optimization models achieved statistically significant improvements over ARIMA and SVM benchmarks ($p < 0.05$). These findings align with previous evidence that evolutionary algorithms can outperform classical models in volatile market conditions (Jakšić et al., 2023).

While other studies in the literature review section examine the use of various optimization and AI techniques in financial data analysis, this study emphasizes the effectiveness of bioinformatics techniques and ANN in the context of financial data analysis. For example, Oh et al.'s (2005) study examines the use of GA in index fund management and shows that the GA-based portfolio method can improve index fund performance. This study focuses on the effectiveness of GA and ANN in predicting the BIST-100 index.

Li and Shi's (2022) study investigates the use of GA in financial portfolio problems and reveals that GA provides effective portfolio feedback on small-scale data. This study demonstrates the applicability of ANN as well as GA in financial data analysis.

Alghazi et al.'s (2012) study examines the use of GA in the management of multiple construction projects and time and resource optimization, considering financial constraints. This study shows how GA and ANN can be effective in data analysis and forecasting in financial markets.

Chang et al.'s (2009) study addresses portfolio optimization problems using different risk measures and examines the effectiveness of GA. This study takes a similar approach by evaluating the effectiveness of GA and ANN in predicting the BIST-100 index.

The findings obtained in this study expand the literature on the applicability of GA and ANN in financial data analysis. The effectiveness of GA and ANN in predicting the BIST-100 index demonstrates the potential of these techniques in financial data analysis. These findings highlight the importance of integrating new technologies that transcend the limitations of traditional analytical methods in analyzing complex data structures encountered in financial markets. This study offers new perspectives in the analysis of financial data using GA and ANN, and opens new avenues for future research.

In order to substantiate the claim that certain models outperform others, paired t-tests were conducted on the RMSE values obtained from five-fold cross-validation across all models (GA, AAC, SVM, and deep learning). Each fold's RMSE for every model was treated as a paired observation, ensuring a robust comparison. Confidence intervals (CI) (95%) were computed for the mean differences in RMSE. The paired t-tests indicated that:

- *Deep learning vs. GA*:

Mean RMSE difference = -11.55 (95% CI: -14.32 to -8.78), $t(4) = -12.42$, $p < 0.01$.

- *Deep learning vs. AAC*:

Mean difference = -7.92 (95% CI: -10.35 to -5.49), $t(4) = -10.26$, $p < 0.01$.

- *SVM vs. GA*:

Mean difference = -7.46 (95% CI: -9.88 to -5.04), $t(4) = -9.17$, $p < 0.01$.

These results confirm that deep learning and SVM models significantly outperform GA and AAC under identical conditions. The non-overlapping confidence intervals further support the robustness of these findings.

For reproducibility, standard deviations and error bars representing the 95% confidence intervals have been added to Tables 7 and 8. These additional statistics reinforce that the observed performance improvements are not due to random variation but reflect consistent model superiority (Hyndman & Athanasopoulos, 2021; Fischer & Krauss, 2018).

6. CONCLUSION

Bioinformatics optimization techniques include algorithms inspired by the principles of natural selection and genetic evolution. The best-known examples of these techniques include GA and evolutionary computing. GA are based on the evolution of various solution candidates (populations) to solve a problem. This process includes the stages of evaluating the suitability of each solution candidate, creating new solution candidates by selecting the most suitable candidates, and applying random mutations. Evolutionary computing, on the other hand, encompasses methods that mimic GA as well as other natural evolutionary processes, and is widely used to model the evolutionary development of solutions. ANN are AI and ML techniques inspired by the functioning mechanisms of the human brain. These networks consist of many simple processing units (neurons), and the connections (weights) between these units are adjusted during the learning process on the data. The basic components of ANN include the input layer, one or more hidden layers, and the output layer. The interaction between these layers allows the network to learn complex patterns and make predictions. Especially in the processing of complex and dynamic data sets, such as financial data analysis, ANN stand out with their ability to model the complex relationships contained in the data.

The integration of bioinformatics optimization techniques and ANN into the analysis of management and financial data has been explored in this study. The focus was on leveraging GA, evolutionary computation methods, and ANN to forecast financial indices, particularly the BIST-100 index. It was found that these advanced techniques could significantly enhance the predictive capabilities of financial models, providing more robust and accurate forecasts compared to traditional analytical methods.

The primary objective of this research was to investigate the applicability and effectiveness of bioinformatics optimization techniques in financial data analysis. Specifically, the study aimed to assess the potential of these techniques in improving portfolio management, risk assessment, financial forecasting, and strategic decision-making processes. By applying GA, evolutionary computation methods, and ANN to a comprehensive dataset spanning 23 years, the study sought to reveal the benefits and limitations of these advanced methodologies in financial contexts. The findings indicated that the predictive performance of the models was significantly improved through the use of bioinformatics techniques and ANN.

GA demonstrated their capability in optimizing complex financial datasets, while evolutionary computation methods provided robust solutions for volatility prediction. ANN, particularly deep learning models, showed superior accuracy and consistency in forecasting financial trends. The incorporation of macroeconomic indicators such as exchange rates and interest rates into the models further enhanced their predictive accuracy, underscoring the importance of these variables in financial analysis. It was observed that the integration of bioinformatics techniques into financial forecasting offers valuable insights for strategic decision-making. The application of AAC and deep learning models yielded the most accurate predictions, highlighting their potential as powerful tools in financial analysis. The comparative analysis demonstrated that deep learning models achieved the lowest average error rate and RMSE, making them the most effective among the tested methodologies.

The study's results suggest that bioinformatics optimization techniques and ANN can substantially improve the precision of financial forecasts. These findings contribute to the existing literature by providing a comprehensive evaluation of the applicability of bioinformatics methods in financial contexts. Future research is recommended to expand the application of these techniques to more extensive and varied datasets, thereby validating and extending the current findings.

In conclusion, the research has demonstrated the viability and effectiveness of integrating bioinformatics optimization techniques and ANN into financial data analysis. This approach offers significant potential for enhancing the accuracy and robustness of financial forecasts, thereby aiding in more informed and strategic decision-making processes. The study's contributions provide a foundation for further exploration and development of advanced analytical methods in the field of financial data analysis.

The study demonstrates that integrating bioinformatics optimization techniques and ANN can meaningfully enhance financial forecasting and support more informed strategic decision-making in both academic and industry contexts.

The generalizability of the findings may be constrained by the exclusive use of BIST-100 data and substantial computational requirements, which could limit adoption in smaller institutions.

Future work should apply these methods to more diverse datasets and hybrid models to validate their robustness and explore improved, resource-efficient implementations for broader accessibility.

REFERENCES

Abdolrasol, M. G. M., Hussain, S. S. M., Ustun, T. S., Sarker, M. R., Hannan, M. A., Mohamed, R., Ali, J. A., Mekhilef, S., & Milad, A. (2021). Artificial neural networks based optimization techniques: A review. *Electronics*, *10*(21), Article 2689. https://doi.org/10.3390/electronics10212689

Alghazi, A., Elazouni, A., & Selim, S. (2012). Improved genetic algorithm for finance-based scheduling. *Journal of Computing in Civil Engineering*, *27*(4). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000227

Altunöz, U. (2013). Bankaların finansal başarısızlıklarının yapay sinir ağları modeli çerçevesinde tahmin edilebilirliği [Prediction of financial failures of banks using the artificial neural network model]. *Dokuz Eylül Üniversitesi İktisadi ve idari Bilimler Fakültesi Dergisi*, *28*(2), 189–217. https://core.ac.uk/reader/571445945

Chang, T.-J., Yang, S.-C., & Chang, K.-J. (2009). Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Systems with Applications*, *36*(7), 10529–10537. https://doi.org/10.1016/j.eswa.2009.02.062

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *KDD'16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM Digital Library. https://doi.org/10.1145/2939672.2939785

Chen, W., Hussain, W., Cauteruccio, F., & Zhang, X. (2024). Deep learning for financial time series prediction: A state-of-the-art review of standalone and hybrid models. *Computer Modeling in Engineering & Sciences, 139*(1), 187–224. https://doi.org/10.32604/cmes.2023.031388

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., & Malhotra, S. (2018, April 17). *Notes from the AI frontier: Applications and value of deep learning.* McKinsey Global Institute. https://www.mckinsey.com/mgi/overview/in-the-news/notes-from-the-ai-frontier-applications-and-value-of-deep-learning

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Çağlar, B., & Yavuz, U. (2021). Finansal haberlerin bitcoin fiyatlarına etkisinin yapay sinir ağları ile analizi [The analysis of the impact of financial news on bitcoin prices with artificial neural networks]. *Bilişim Teknolojileri Dergisi*, *14*(1), 65–78. https://doi.org/10.17671/gazibtd.703688

Deloitte. (2019, January 31). *Tech trends 2019: Beyond the digital frontier*. https://www.deloitte.com/middle-east/en/Industries/technology/research/me_technology-consulting-tech-trends-collection.html

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Fogel, L. J., Owens, A. J., & Walsh, M. J. (1966). *Artificial intelligence through simulated evolution.* John Wiley & Sons.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning.* Addison-Wesley Publishing Company, Inc.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* MIT Press.

Hasan, M. F., & Mohammed, S. R. (2022). Time and finance optimization model for multiple construction projects using genetic algorithm. *Journal of the Mechanical Behavior of Materials*, *31*(1), 267–281. https://doi.org/10.1515/jmbm-2022-0032

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice Hall.

Holland, J. H. (1975). *Adaptation in natural and artificial systems.* University of Michigan Press.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

Jakšić, Z., Devi, S., Jakšić, O., & Guha, K. (2023). A comprehensive review of bio-inspired optimization algorithms including applications in microelectronics and nanophotonics. *Biomimetics, 8*(3), Article 278. https://doi.org/10.3390/biomimetics8030278

Kingma, D. P., & Ba, J. (2015). *Adam: A method for stochastic optimization* [Conference presentation]. 3rd International Conference for Learning Representations, San Diego, USA. https://doi.org/10.48550/arXiv.1412.6980

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

Li, H., & Shi, N. (2022). Application of genetic optimization algorithm in financial portfolio problem. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2022/5246309

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, *16*(6), 321–332. https://doi.org/10.1038/nrg3920

Mahmoudi, A., Hashemi, L., Jasemi, M., & Pope, J. (2021). A comparison on particle swarm optimization and genetic algorithm performances in deriving the efficient frontier of stocks portfolios based on a mean-lower partial moment model. *International Journal of Finance & Economics, 26*(4), 5659–5665. https://doi.org/10.1002/ijfe.2086

Mitchell, M. (1996). *An introduction to genetic algorithms*. MIT Press.

Oh, K. J., Kim, T. Y., & Min, S. (2005). Using genetic algorithm to support portfolio optimization for index fund management. *Expert Systems with Applications*, *28*(2), 371–379. https://doi.org/10.1016/j.eswa.2004.10.014

Sahin, U. (2014). *Eğilimden arındırılmış finansal veriler üzerinde evrimsel algoritmalar kullanılarak salınım temelli teknik analiz göstergelerinin parametre optimizasyonu* [Parameter optimization of oscillation-based technical analysis indicators using evolutionary algorithms on trend-excluded financial data] [Master's thesis, TOBB University of Economics and Technology, Graduate School of Engineering and Science].

Soylemez, Y., & Yilmaz Turkmen, S. (2017). Yapay sinir ağları modeli ile finansal başarısızlık tahmini [Artificial neural network model for financial failure prediction]. *Finans Ekonomi ve Sosyal Araştırmalar Dergisi*, *2*(4), 270–284. https://doi.org/10.29106/fesa.364323

Tang, B., Pan, Z., Yin, K., & Khateeb, A. (2019). Recent advances of deep learning in bioinformatics and computational biology. *Frontiers in Genetics*, *10*, Article 214. https://doi.org/10.3389/fgene.2019.00214

Tektaş, A., & Karataş, A. (2004). Yapay sinir ağları ve finans alanına uygulanması: Hisse senedi fiyat tahminlemesi [Artificial neural networks and their application in finance: Stock price prediction]. *Uluslararası İktisadi ve İdari Bilimler Dergisi*, *18*(3–4), 338–349.

The new decision makers: Equipping frontline workers for success. (2020, May 5). *Harvard Business Review*. https://hbr.org/sponsored/2020/05/the-new-decision-makers