

A Review of Machine Learning Models in Finance: Evidence from Academic Research in Turkey

Asst. Prof. Dr. Naime İrem Duran  [0000-0002-8953-2171](https://orcid.org/0000-0002-8953-2171)

Abstract

In recent years, the application of machine learning models in finance has attracted great attention due to its potential to improve decision-making processes and risk management strategies. The aim of this article is to present a comprehensive review of academic research conducted in Turkey on the use of machine learning models in finance. It aims to identify machine learning techniques commonly used in Turkish finance studies, evaluate their effectiveness, and provide insights into successful applications. The findings reveal that regression analysis is widely used in predicting financial variables such as stock prices and exchange rates. Clustering techniques have been effective in customer segmentation and market basket analysis. Decision trees are frequently used in credit scoring and fraud detection tasks due to their interpretability and ease of implementation. Moreover, artificial neural networks, especially deep learning algorithms; It has shown promising results in complex financial tasks such as sentiment analysis, anomaly detection, and algorithmic trading. In conclusion, this review underlines the significant potential of machine learning models in finance in Turkey. A few suggestions can be made regarding machine learning in finance in Turkey to identify future research areas. These may include developing customized machine learning models for specific financial applications that require more in-depth analysis, improving the quality and size of datasets, and investigating new techniques outside of existing models. There is also a need for more studies to provide practical guidance on how machine learning techniques are applied by financial institutions and how these applications can be improved.

1 Introduction

Machine learning is defined as a tool that provides benefits beyond human potential by identifying patterns through learning from data, which would be difficult or nearly impossible for the human mechanism to perform. Machine learning enables computers to learn similar to humans by developing algorithms. With the rapid advancements in technology in recent years, it is observed that the applications of machine learning in the field of finance are rapidly evolving.

The aim of this study is to examine academic research conducted in the field of finance in Turkey, determine which techniques are used for which topics, and provide researchers with a general evaluation, as well as evaluate existing and missing topics. To achieve this goal, the first part of the study will include a definition of machine learning and the techniques used, followed by a summary of the literature review, and finally, an evaluation will be made in the conclusion section.

2 The Definition and Methodologies of Machine Learning

Machine learning is generally defined as a subset of artificial intelligence. Prediction is made based on the information obtained from data using machine learning techniques. With machine learning techniques, models that learn the relationship between variables enable predictions to be made on computers without any programming. Nowadays, the scope of machine learning is expanding rapidly, and it is increasingly prevalent in all of our lives. In summary, machine learning is a scientific study where algorithms and statistical models learn from data without the need for computer programming, with minimal human intervention in the results.

Machine learning is particularly used in finance for price prediction, risk assessment, and financial market analysis.

2.1 Supervised Learning Algorithms

Supervised learning algorithms consist of algorithms that learn from the available data during the training phase and utilize this data during the prediction phase. Supervised learning algorithms include methods such as Linear Regression, Logistic Regression, Support Vector Machines, Decision Trees, and Random Forests.

2.2 Unsupervised Learning Algorithms

In contrast to supervised learning algorithms, in unsupervised learning algorithms, the aim is to discover hidden patterns among the data. Therefore, unsupervised learning algorithms are used to tackle more complex problems. Unsupervised learning algorithms include K-Means Clustering, Hierarchical Clustering, and PCA (Principal Component Analysis).

2.3 Reinforcement Learning Algorithms

Reinforcement Learning is a type of learning that involves training an agent, such as a robot or a vehicle, based on the responses it receives from the environment, without the presence of training data. In this algorithm, the agent aims to achieve the best outcome through trial and error.

2.4 Deep Learning

Deep Learning is a subfield of machine learning that utilizes artificial neural networks to create algorithms. Deep learning techniques include algorithms such as Classic Neural Networks, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Restricted Boltzmann Machines (RBM), and Deep Reinforcement Learning, Transfer Learning and Self Organizing Maps.

3 The Method

The study focused on articles retrieved from Google Scholar, Dergipark, and YÖK databases spanning the period from 2020 to 2024. The search was conducted using specific keywords, with "machine learning and finance" being the primary search term due to time constraints. The identified studies were categorized into two main groups: theoretical and empirical. The analysis aimed to investigate the applications of artificial intelligence technologies in finance and determine the most commonly utilized methods

4 Review of Studies Conducted in Machine Learning and Finance

In the study conducted by Söylemez (2020), artificial neural networks algorithm was employed to model gold price predictions between 2014 and 2019. In the prediction model, variables such as fear index, oil prices, dollar index, and Dow Jones index were used. The analysis concluded by selecting the model with the best performance achieved using artificial neural networks. Karadağ, Bölükbaş, and Ünal (2020) utilized customer data from a bank to predict suitable financial data for customers using K-NN, Random Forest, and Naive-Bayes algorithms. The study determined that the random forest model was the best algorithm.

In Aksoy's study (2021), data from five manufacturing companies for the period 2010-2020 were used to predict stock prices using artificial neural networks, regression, and K-Nearest Neighbor algorithms. It was observed that the artificial neural networks algorithm performed the best, although all algorithms used could be utilized by investors. Karacan and Kırdar (2021) provided information on academic studies related to the prediction of stock prices and portfolio management using machine learning methods. They emphasized the increasing potential of the field. Filiz et al. (2021) predicted the BIST 100 index using major world indices, exchange rates (dollar and euro), and gold variables. Various algorithms such as Naive Bayes, logistic regression, artificial neural networks, support vector, and random forest decision tree were used, with the support vector machine algorithm achieving the best performance. In the study by Usta et al. (2021), the BIST-30 index and financial ratios for the years 2010-2019 were utilized to predict stock prices using XG-Boost, Random Forest, and Artificial Neural Networks algorithms. The XG-Boost algorithm yielded the best results. Salur (2021) used artificial neural networks algorithm to predict financial failure using data from 144 companies traded on the Borsa Istanbul between 2013-2018. The study concluded that the artificial neural networks algorithm provided effective results. Yıldız and Yıldız (2021) used artificial neural networks algorithm to predict investment alternatives with variables such as exchange rates, BIST 100 index, gold prices, and TL deposit interest rates for the period 2015-2020. It was noted that the model performances were high. Taş et al. (2021) utilized SP 500 index data from 2000 to 2020 to make predictions using LSMP and MLP methods. It was observed that the performance of both methods was found to be close to each other.

Doğan and Türe (2022) employed machine learning techniques to predict the financial risk of countries using data from 75 countries between 2015 and 2019. Logistic regression, support vector machines, K-nearest neighbors, and decision tree algorithms were utilized for modeling, with the KNN algorithm demonstrating the best performance. Kırca et al. (2022) predicted financial risk for 2600 agents for the period 2017-2021. Among linear regression, random forest, Gradient Boosted, and XGBoost methods, they found that Gradient Boosting performed the best. Sarı (2022) utilized artificial neural networks and wavelet transform artificial neural networks methods to predict weekly financial stress index, interest rates, yield spread, and other variables affecting financial stress for 30 years. The study concluded that the wavelet transforms artificial neural networks method outperformed others. Uyar and Yavuz (2022) modeled the relationship between stock prices and intellectual capital variables using data from 124 manufacturing companies operating in the Istanbul Stock Exchange between December 31, 2010, and 2018. They found Support Vector Machines (SVM) to be the most effective model and emphasized the importance of including intangible fixed assets in financial reporting. Yıldız (2022) examined Turkish studies utilizing artificial intelligence technology in finance between 2015 and 2021 and observed a rapid increase in studies conducted after 2020. Altan and Demirci (2022) categorized 133 firms into successful and unsuccessful categories for three periods and used their cash flow ratios as independent variables. They experimented with XGBoost, Gradient Boosting, and Artificial Neural Networks methods and preferred the XGBoost method as the suitable algorithm.

Erden (2023) predicted daily Ereğli stock data between 2020 and 2022 using deep learning algorithms such as LSTM, GRU, and RNN, as well as the ARIMA model. The study concluded that the RNN method performed the best. Tütüncü and Gürsakal (2023) built models for predicting credit default risk using customer credit and demographic information. They found Gradient Boosting to be the most effective model among Logistic

Regression, Random Forest, Decision Tree, Support Vector Machines, K-Nearest Neighbor, and WOE transformations for Logistic Regression algorithms. Faydalı and Solak (2023) examined applications in accounting audit. They provided detailed literature on auditor selection problems, risk control and valuation, business continuity, and audit opinion topics based on annual reports published by four major audit firms, providing insights into the use of machine learning in audit purposes. Öngün (2023) conducted time series analysis using stock data of a Turkish company and inflation, interest rates, exchange rates, oil, and gold prices data. They used regression, support vector machine, decision tree, random forest, and LSTM methods, concluding that LSTM was the most effective prediction method. Ceyhan (2023) predicted the financial ratios of 14 companies traded on BIST for 2022 using Linear Regression, Naive Bayes, Random Forest, and Random Tree algorithms. The study found that Naive Bayes and Random Forest models performed the best.

Şeyranlıoğlu and Karavardar (2024) used Artificial Neural Networks, Support Vector Machines, Decision Trees, and Random Forest techniques to calculate the values of 12 companies traded on the Borsa Istanbul Investment and Holding Index. They reported that artificial neural networks model showed the best performance. Demirel, Hazar, and Babuşçu (2024) used Feedforward Artificial Neural Networks method to predict daily CDS premium values for Turkey between 2008 and 2022. The study concluded that the model performance obtained with artificial neural networks was effective compared to traditional forecasting methods

5 Conclusion

This study examines national academic research on machine learning, a subset of artificial intelligence, in the finance domain from 2020 to the present. The analysis shows that almost all studies compare the performance of machine learning algorithms. Most studies focus on stock price forecasting and related factors, with some also exploring models for predicting customer and credit risk in banking. During the detailed investigation of the relevant period, it was noticed that articles derived from these were encountered. Therefore, the review continued by examining the articles authored by the researchers instead of the theses.

Given the ease of data availability, more research in this area is expected. The variability of journals where these studies are published, covering both engineering and finance topics, highlights the breadth of machine learning and algorithmic applications. Furthermore, machine learning techniques in empirical studies offer rapid results without requiring assumptions and can outperform traditional methods, suggesting a growing adoption of these approaches. To enhance this study, expanding the number of key concepts, examining foreign literature, making finer subject-based differentiations, and conducting more detailed comparisons could be considered. Additionally, conducting such studies for specific periods is academically significant for assessing the stage of artificial intelligence utilization

References

- Aksoy, 2021. "Pay Senedi Fiyat Yönünün Makine Öğrenmesi Yöntemleri ile Tahmini: Borsa İstanbul Örneği", *Business and Economics Research Journal*, **12**, pp:89-110.
- Altan ve Demirci 2022. " Makine Öğrenmesi ile Nakit Akış Tablosu Üzerinden Kredi Skorlaması: XGBoost Yaklaşımı", *İktisat Politikası Araştırmaları Dergisi*, **9**, pp. 397-423.
- Ceyhan, 2023. "İşletmelerde Finansal Başarısızlık Öngörüsünde Makine Öğrenmesinin Kullanımı: Bist Uygulaması", *Finans Ekonomi ve Sosyal Araştırmalar Dergisi*, **8**, pp:565-580.
- Demirel vd, 2024. "CDS Primi Tahmini: Türkiye Üzerine Bir Uygulama", *İşletme Araştırmaları Dergisi*, **16**, pp:512-520.
- Doğan ve Büyükkör, 2022. "Makine Öğrenmesi ile Finansal Zaman Serisi Tahminleme", *Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, **24**, p. 1205.
- Doğan ve Türe, 2022. "Makine Öğrenmesi Teknikleri İle Ülke Riski Tahmini", *Fiscaoeconomia*, **6**, pp:1126-1151.
- Erden, 2023. "Derin Öğrenme ve ARIMA Yöntemlerinin Tahmin Performanslarının Kıyaslanması: Bir Borsa İstanbul Hissesi Örneği", *Yönetim ve Ekonomi*, **30**, pp:419-438.
- Filiz vd, 2021. "Büyük Dünya Endeksleri Kullanılarak BIST-100 Endeksi Değişim Yönünün Makine Öğrenmesi Algoritmaları ile Sınıflandırılması", *Beü Fen Bilimleri Dergisi*, **10**, pp:432-441.
- Faydalı ve Solak, 2023. "Bağımsız Denetimde Makine Öğrenmesi: Teorik Bir Yaklaşım", *Journal of Academic Value Studies*, **9**, 95-114.
- Kaynar ve Yiğit, 2021. "Öznitelik Mühendisliği ile Makine Öğrenmesi Yöntemleri Kullanılarak BIST 100 Endeksi Değişiminin Tahminine Yönelik Bir Yaklaşım", *Journal of Yasar University*, **16**, pp:1741-1762.
- Karacan ve Kırdar, 2021. " Hisse Senedi Fiyat Tahmininde Makine Öğrenmesi ve Yapay Zeka Kullanımı", *Journal of International Social Research*, **14**, pp:666-674.

- Karadağ vd., 2020 . "Makine Öğrenmesi ile Bireysel Müşteriler için Finansman Ürün Önerilmesi, *Academic Perspective Procedia*, **3**, pp:438-444.
- Kırca vd, 2022. "Makine Öğrenmesi Yaklaşımı ile Acente Kredi Riski Hesaplama, *Bilişim Teknolojileri Dergisi*, **15**, pp:159-166.
- Öngün, 2023. "Çok Değişkenli Finansal Zaman Serisi Analizinde Yapay Zekanın Kullanımı, *YBS Ansiklopedisi*, **11**, pp:13-23.
- Salur, 2021. "Finansal Başarısızlık Tahmininde Yapay Sinir Ağları Modelinin Kullanımı: Borsa İstanbul'da Bir Uygulama". *Journal of Economics Finance and Accounting*, **8**, pp:17-30.
- Sarı, 2022. "Finansal Stres Endeksinin Dalgacık Dönüşümlü Yapay Sinir Ağları Kullanılarak Tahmin Edilmesi", *The Journal of International Scientific Research*, **7**, pp:282-296
- Söylemez, 2020. "Çok Katmanlı Yapay Sinir Ağları Yöntemi ile Altın Fiyatlarının Tahmini". *Sosyoekonomi*, **28**, pp: 271-291.
- Şeyranlıoğlu ve Karavardar, 2024. "Şirket Değerlemesinde Makine Öğrenmesi Algoritmalarının Kullanımına Yönelik bir Uygulama: Holding Şirketleri Örneği, *Uluslararası Ekonomi ve Yenilik Dergisi*, **10**, pp: 149-175
- Taş vd., 2021. " Finansal Piyasalarda Hisse Fiyatlarının Derin Öğrenme ve Yapay Sinir Ağı Yöntemleri ile Tahmin Edilmesi; S&P 500 Endeksi Örneği". *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, **9**, 446-460.
- Tütüncü, 2022. "Makine Öğrenmesi Algoritmaları ile Kredi Temerrüt Riskini Tahmin Etme", Bursa Uludağ Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri AnaBilim Dalı, İstatistik Bilim Dalı, Yüksek Lisans Tezi
- Tütüncü ve Gürsakal, 2023. "Kredi Temerrüt Riskini Tahmin Etmede Makine Öğrenme Algoritmalarının Karşılaştırılması". , *European Journal of Science and Technology* ,**50**, pp. 14-22.
- Uyar ve Yavuz, 2022. "Entelektüel sermaye ve hisse fiyatı arasındaki ilişkinin makine öğrenmesi ile incelenmesi", *Gazi İktisat ve İşletme Dergisi*, **8**, pp: 214-228.
- Ustali vd, 2021. "Makine Öğrenmesi Teknikleri ile Hisse Senedi Fiyat Tahmini", *Eskişehir Osmangazi Üniversitesi İİBF Dergisi*, **16**, pp:1-16.
- Yıldız, 2022. " Finans Alanında Yapay Zekâ Teknolojisinin Kullanımı: Sistemik Literatür İncelemesi", *Pamukkale Sosyal Bilimler Enstitüsü Dergisi*, **52**, pp:47-66.
- Yıldız ve Yıldız.,2021. " Prediction of Investment Alternatives with Artificial Neural Network. " *İşletme Araştırmaları Dergisi*, **13**, pp:3103-3118.