Predicting Stock Market Index Movements With Machine Learning

Nazif Ayyıldız

ÖZGÜR PRESS

Predicting Stock Market Index Movements With Machine Learning

Nazif Ayyıldız

ÖZGÜR PRESS

Published by
Özgür Yayın-Dağıtım Co. Ltd.
Certificate Number: 45503

15 Temmuz Mah. 148136. Sk. No: 9 Şehitkamil/Gaziantep
+90.850 260 09 97
+90.532 289 82 15
www.ozguryayınlari.com
info@ozguryayinlari.com

Predicting Stock Market Index Movements With Machine Learning
Nazif Ayyıldız

Language: English
Publication Date: 2023
Cover design by Mehmet Çakır
Cover design and image licensed under CC BY-NC 4.0
Print and digital versions typeset by Çizgi Medya Co. Ltd.

ISBN (PDF): 978-975-447-821-1

DOI: https://doi.org/10.58830/ozgur.pub354

Suggested citation:
Ayyıldız, N. (2023). *Predicting Stock Market Index Movements With Machine Learning*. Özgür Publications. DOI: https://doi.org/10.58830/ozgur.pub354. License: CC-BY-NC 4.0

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see https://www.ozguryayinlari.com/

ÖZGÜR PRESS

This book has been derived from the doctoral thesis titled "PREDICTION OF STOCK MARKET INDEX MOVEMENTS USING MACHINE LEARNING METHODS: AN APPLICATION ON THE STOCK MARKETS OF DEVELOPED AND DEVELOPING COUNTRIES" prepared in the Department of Business Administration at Niğde Ömer HALİSDEMİR University, Institute of Social Sciences. Furthermore, the thesis was supervised by Prof. Dr. Ömer İSKENDEROĞLU and approved by the doctoral thesis committee consisting of Prof. Dr. Ömer İSKENDEROĞLU, Prof. Dr. Mutlu Başaran ÖZTÜRK, Prof. Dr. Erdinç KARADENİZ, Assoc. Prof. Dr. Ayberk Nuri BERKMAN, and Dr. Research Assistant Mehmet BEYAZGÜL.

Contents

Introduction

The future inherently contains risks and uncertainties. Effective management of the future is closely related to accurate forecasting. Therefore, the future can only be turned into an opportunity and potential gains can be realized when it is accurately predicted. Predicting stock price movements is of great interest in the financial world, as it provides the opportunity for investors to generate returns or take precautionary measures against potential losses. However, the question of whether stock prices can be predicted or not is a topic of debate in the literature. The efficient market hypothesis, as proposed by Fama, suggests that stock prices are significantly determined by new information and that prices follow a random walk model that cannot be predicted solely based on historical information. While this hypothesis found support when it was first introduced, it has been criticized and its validity has been questioned in numerous studies that demonstrate stock prices can be predicted to some extent. Furthermore, traditional or classical methods, also known as fundamental and technical analysis, are used for the purpose of forecasting stock prices in capital markets. In fundamental analysis, research is conducted on the economy, industry, and individual companies to determine the true value of a stock. If the true value of a stock is greater than its market price, a buy decision is made; otherwise, a sell decision is made. On the other hand, technical analysis, unlike fundamental analysis, does not require an in-depth analysis of the economy, industry, or individual companies. It relies solely on past price movements and trading volumes of stocks to predict the direction of future prices. However, in recent years, with advancements in technology, computer-based machine learning techniques have become widely used in stock markets. Machine learning allows for the identification of historical relationships and trends in data, leading to more accurate and faster predictions compared to traditional methods.

The objective of this study is to predict the directional movements of stock market indices in developed and developing countries using machine learning classification methods, compare the performance of these methods, and determine the best prediction method. To achieve this goal, a daily data set and technical indicators for the period from January 1, 2012, to December 31, 2021, were used to predict the directions of stock market indices using decision trees, random forests, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks.

The study consists of four sections. In the first section, information is provided about the historical development of stock exchanges, followed by an explanation of the important functions and activities of stock exchanges. Subsequently, the topic of stock market indices is explored, comparing major stock exchanges in developed and developing countries in terms of market capitalization and the number of companies. The final part discusses the random walk theory and the efficient market hypothesis, explaining the analytical methods used in stock markets. In the second section, basic concepts related to machine learning are presented, followed by an exploration of the history of machine learning. Different types of machine learning, such as supervised, unsupervised, semi-supervised, and reinforcement learning, are discussed, along with explanations of machine learning methods. The section concludes with a summary of common machine learning algorithms used in predicting stock market index directions. The third section summarizes studies that predict stock market indices in developed and developing countries using different machine learning methods. The literature findings are discussed, providing an overview of the research conducted in this area. In the final section where the analysis is performed, the purpose and significance of the study are explained, and the data set and methods used are introduced. The findings from the analysis of predicting stock market index directions are presented. Finally, the results are discussed from different perspectives.

CHAPTER 1

Stock Exchanges and Analysis in the Stock Market

The first chapter, titled 'Stock Exchanges and Analysis in the Stock Market,' is divided into five subsections. In the first subsection, the historical development of stock exchanges is explored. The second subsection explains the important functions and activities of stock exchanges. The third subsection focuses on stock market indices. In the fourth subsection, major stock exchanges in developed and developing countries are compared. The final subsection delves into the random walk theory and the efficient market hypothesis, elucidating the analytical methods used in stock markets.

1.1.THE HISTORY OF STOCK EXCHANGES

In the modern financial world, although there are many stock exchanges where various financial instruments are bought and sold, the most well-known among these exchanges is the securities exchange. To describe securities exchanges, terms such as beurs (Dutch), borse (German), bourse (French), borsa (Italian), and bolsa (Spanish and Portuguese) are used. In the literature, in some studies, the word 'borsa' is thought to have originated from the Van der Beurs family, who are believed to have opened one of the first exchanges in Antwerp in 1531. The term 'market' has its origins in Italian, derived from the word 'piazza.' The concept of 'pazar' in Turkish, which comes from the Persian word 'bazar' also means 'market.' In a financial context, the market encompasses places where financial products are bought and sold. However, the market concept also includes over-the-counter markets (Mosselaar, 2018: 45).

Looking back into history, the origins of stock exchanges can be traced back to the Roman Empire approximately 1,770 years ago. The oldest known exchange is believed to have been established around 250 AD in the ancient city of Aizaoni, located in Kütahya Çavdarhisar, Turkey. Aizaoni, which was discovered in 2012 and added to the UNESCO World Heritage List, contains inscriptions on stone blocks where Emperor Diocletian's wage controls against inflation in 301 AD are recorded. These inscriptions list the selling prices of all goods sold in imperial markets. Prices of goods that were traded at the time were fixed to combat inflation nationwide. Cross-pricing was used among the produced goods to prevent purchases at exorbitant prices (Nadirgil, 2015: 16).

The earliest stock exchanges typically emerged in small venues with the presence of brokers who gathered regularly in places near markets and fairs. Slow and costly transportation and information technology limited the scale and scope of these early exchanges. Generally, the initial exchanges were located in historically significant trading centers, such as Antwerp, Amsterdam, Genoa, and London. The Antwerp Exchange, established in 1460, is one of the oldest exchanges in Europe. Following Anvers, exchanges in Amsterdam, Paris, and London were established in the 16th century, Berlin in the 17th century, Vienna in the 18th century, and Brussels, Rome, Milan, Madrid, and Istanbul Exchanges in the 19th century (Güngör, 2016: 130). In the United States, the first exchanges were established first in New York in 1792, followed by Boston and Philadelphia. These early exchanges were located in specific street locations, such as the spot under the famous Buttonwood tree on Wall Street or other places, established for their specific purposes. The first companies that shareholders owned were companies like the Dutch East India Company, founded by Dutch merchants in 1602. The shares of the Dutch East India Company were traded on the Amsterdam Stock Exchange for about two centuries until the company went bankrupt. Since the early companies established were not known outside their regions, the shares issued were traded locally. As a result, most investors invested in shares of companies they were familiar with (Levinson, 2006: 129).

In the early exchanges with small trading scales, call auctions were often organized to match buyers and sellers. Call auctions utilized buy offers and auction methods. In the late 19th century, when increasing lists of stocks became difficult to complete at auctions in some markets, these markets were converted into continuous trading systems. The New York Stock Exchange (NYSE), which was a pioneer in this regard, switched to continuous trading in 1872 by using experts in fixed brokerage locations. Nevertheless, around the world, many exchanges continued to use the call auction format until the

20th century. As a result, the number of investors, the number of securities traded, and trading volumes in many exchanges remained quite small until the 20th century (Chambers and Dimson, 2016: 119-120).

Between 1871 and 1914, there was an increase in the number and value of stock exchanges. Just before World War I, there were mainly 89 main stock exchanges, most of which were located in Europe. During this period, in terms of stock market activities and market capitalization, major stock exchanges in the United Kingdom, France, Germany, and the United States stood out. In the second half of the 20th century, several significant developments paved the way for stock market activities. In 1975, fixed brokerage commissions were eliminated in the United States. In 1985, foreign brokerage firms began to be accepted as members of the Tokyo Stock Exchange. In 1986, fixed brokerage commissions were removed from the London Stock Exchange, and foreign brokerage firms began to be accepted as full members. This development was referred to as the "big bang" in financial circles. In 1987, in London, and in 1999, in the United States, it became legally permissible to conduct commercial banking and investment banking activities together. In the 1990s, many developing countries passed similar laws to encourage stock trading, almost eliminating countries without stock exchanges. There was a similar development in the status of stock exchanges. Starting in 1993 and continuing until 2000, most small stock exchanges worldwide transformed into profit-oriented joint-stock companies. In many platforms, stock exchange-traded funds (ETFs) were offered to the public. By the 2000s, following the developments in internet and information technologies, all major stock exchanges switched to electronic trading systems. In countries such as Argentina and Portugal, a significant portion of transactions on smaller stock exchanges shifted to larger and more liquid stock exchanges. This increased competition among stock exchanges that wanted to dominate the trading of the most active stocks. Many stock exchanges pursued the path of consolidation or acquisition to concentrate their trading activities in larger financial centers and reduce costs.

1.2.THE FUNCTIONS AND ACTIVITIES OF STOCK EXCHANGES

Stock exchanges are authorized institutions responsible for regulating, controlling, and determining and announcing market prices with the aim of providing a specific order and trust in the buying and selling of securities (Avadhani, 2011: 154). The significant functions of stock exchanges, which are of great importance for national economies, are listed below.

• In stock exchanges, savers transform their savings into investments to gain returns, while companies in need of funds receive financial support. In this context, stock exchanges play a fundamental role known as providing capital formation by supplying resources for the coordination of savings and investments in economies, ensuring the most effective and efficient coordination for the development of countries (Sreevidya, 2014: 31).

• Stock exchanges bring together all buyers and sellers, ensuring a single price for each security. In other words, stock exchanges serve as mechanisms for creating a fixed price mechanism for determining the values of securities (Darskuviene, 2010: 7).

• Stock exchanges keep a complete record of all transactions occurring in different securities daily and provide regular information about their prices and trading volumes. Determining price trends enables investors to make rapid decisions regarding the buying and selling of the securities they are interested in (Singh, 2012: 56).

• Stock exchanges establish a continuous and organized market, allowing securities to be sold at their market value at any time. Stock exchanges fulfill the function of providing liquidity, enabling investors to effectively utilize their funds (Quiry et al., 2005: 8).

• Stock exchanges play a role in making large companies' shares available to the public, assisting these companies in obtaining new sources of finance, and, at the same time, enabling the distribution of ownership of large-scale firms among numerous smaller-scale firms or individuals. The function of disseminating ownership allows economic units with small savings accumulations to become shareholders in much larger corporate structures through capital markets (Tunalı, 2013: 136).

• Stock exchanges ensure that they monitor and follow high standards to guarantee efficient, fair, and transparent operation. Additionally, listed companies are obligated to disclose all financial information related to their operations to the public. Such requirements subject company management to continuous scrutiny and evaluation, promoting corporate governance (Musonera and Safari, 2008: 64).

• Stock exchanges are among the best places to monitor the economic effects of inflation, investments, development, and growth, as well as adverse situations like war, economic crises, and political instability. The overall performance of stock exchanges reflects the economic performance of countries. Stock exchanges are often regarded as the pulse, mirror, or barometer of the economy (Deshmukh, 2017: 32).

• Transactions in stock exchanges are conducted exclusively among exchange members, following rules and regulations that ensure adequate transparency and adherence to procedures. Protecting investors is of paramount importance in stock exchanges, preventing investors from being knowingly and unfairly harmed (Abraham and Kannappan, 2018: 1028).

• Stock exchanges, through their websites and various corporate publications, as well as by organizing educational programs, aim to educate the public and encourage people to invest in securities (Shibinu, 2014: 13).

Stock exchanges perform a range of crucial activities in capital markets to ensure the reliable, transparent, efficient, stable, fair, and competitive trading of financial instruments. While the nature and scope of these activities may vary from one country to another, the standard activities of stock exchanges are listed below (Watson and Head, 2010: 96; Leria, 2011: 6; Faure, 2013: 134; Baker and Lang, 2020: 16):

- Managing the licensing process, including the listing, suspension, or cancellation of securities.
- Establishing, operating, and managing markets within the exchange.
- Transmitting and matching trade orders.
- Executing transactions in a timely and expedited manner.
- Providing information on the prices, volumes, and conditions of securities, the characteristics of registered securities, and issuers.
- Granting trading permissions on the exchange.
- Determining and collecting exchange revenues.
- Developing necessary rules for trading and other activities.
- Enforcing disciplinary actions against unethical and illegal behavior.
- Preventing potential conflicts of interest between exchange shareholders and market operators.
- Resolving disputes that arise in exchange operations and transactions.

1.3. STOCK EXCHANGE INDICES

An index is a financial indicator that measures the proportional change in one or more variables' movements. Indices are tools that allow complex and often difficult-to-understand events to be expressed with a single number, providing general information about the outcomes of changes. The measure of the proportional changes shown by values of a specific statistical event over time is referred to as an index or index number. Indices are used to provide information about the price movements of products in financial, commodity, or other markets. Financial indices are created to measure the price movements of stocks, bonds, bonds, and other types of investments. Indices that aim to capture the general behavior of stock markets are referred to as stock market indices or market indices. Stock market indices are calculated by selecting a group of stocks that represent the entire market, a specific sector, or segment of the market (Jose et al., 2014:108).

Stock market indices generally reflect the state of the national economy, and more specifically, the economic performance of companies listed on the stock exchange. In the case of economic growth, an increase in the demand for company products is expected to lead to higher sales, earnings, and dividends for those companies. This situation is anticipated to drive up stock prices and the overall market index (Cma, 2021: 10). Over time, the direction and trend of the variables included in the index are examined through economic analyses. Economists can make assessments about the performance of economies by analyzing stock market movements. When conducting economic assessments, the main index movements of developed stock markets in the United States and Europe are evaluated first, and then the indices of developing countries are interpreted. Therefore, stock market indices serve as platforms for the continuous evaluation of national economies (Karan, 2004: 57).

Stock market indices, which serve as a general indicator of the stock market, provide a broad overview of market performance based on the prices of the stocks included in the index. In the stock market, indices aggregate the prices or returns of a large number of stocks, representing all the variations in these variables with a single number. Without such an index, it would be impossible to determine the direction and extent of movement in the stock market. Stock market indices are used for monitoring the performance of the market they represent and for predicting risk. Indices that measure changes in stock prices generally fall into two categories. Some are simple indices that calculate the overall price level in this area without considering the relative importance of the stocks in the index, giving equal weight to all stock prices and using arithmetic or geometric averages. Others are complex indices that take into account the total market values of stock prices, weighted by the market capitalization of individual stocks and cover a large number of stocks. Price indices do not consider dividends paid when calculating the index and ensuring its continuity. Therefore, only gains resulting from the appreciation of shares are reflected in the index. In return indices, dividend payments are taken into account in the calculation of the index, and adjustments are made based on dividend payments. In this context, gains derived from the appreciation of shares, as well as gains from dividends, are reflected in the index (Varlık, 2017: 41; Secp, 2019: 20). Indices are primarily calculated based on the market values of the freely tradable shares of the stocks included in their composition. In other words, an investor investing in an index fund earns returns approximately equal to the market average (Karan, 2004: 271).

1.4.IMPORTANT STOCK EXCHANGES IN DEVELOPED AND DEVELOPING COUNTRIES

In the international economic and financial world, countries are grouped according to their levels of development using various criteria by certain institutions and organizations. These classifications are typically made by international organizations such as the World Bank, the International Monetary Fund (IMF), and the United Nations. In addition to international organizations, institutions like Morgan Stanley Capital International (MSCI) or PricewaterhouseCoopers (PwC) also classify countries based on their stock market performance or to create future perspectives.

According to the classification made by the World Bank, which categorizes countries by their level of development, countries with high per capita income are generally referred to as high-income economies, while countries with low to middle-level per capita income are classified as low and middle-income economies (World Bank, 2021: 1). The IMF, which uses different criteria for classification, groups countries into advanced economies, emerging markets, and developing economies (IMF, 2016: 162). Finally, according to the classification by the United Nations, countries are categorized into three groups: developed economies, transition economies, and developing economies. In this classification by the United Nations, some subgroups are also defined based on geographical location or temporary criteria. One of the most popular subgroups is the G-7, which represents the most developed seven economies. The G-7 subgroup consists of the United States, Japan, the United Kingdom, France, Canada, Germany, and Italy (WESP, 2021: 163-164).

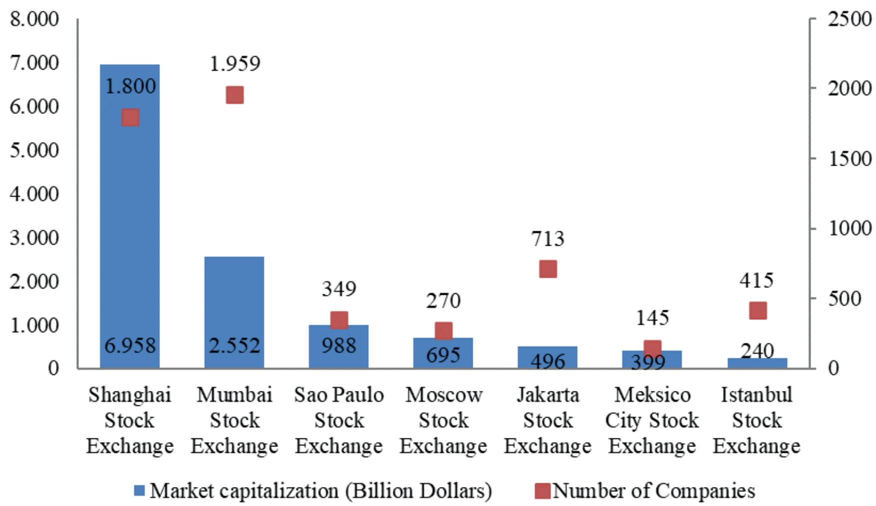
In recent years, the concept of E-7 (Emerging-7) countries, referred to as developing market economies, is frequently used in economic circles. The term E-7 countries was first used by PricewaterhouseCoopers (PwC), an international management consulting firm, in the Stern Review Report on October 30, 2006. The E-7 concept represents seven economies that are believed to have much larger economies in the future. The E-7 group consists of the People's Republic of China, India, Brazil, Mexico, Russia, Indonesia, and Turkey (Samadder et al., 2012: 11).

In the context of the study, G-7 economies are considered as developed countries, while E-7 economies are regarded as developing countries. When selecting the stock exchanges and indices in each country, the stock exchange with the highest market capitalization in each country and the most popular index on that exchange were preferred. In this context, this section of the study examines important stock exchanges in G-7 and E-7 countries, comparing them based on market capitalization and the number of listed companies within their respective development groups.

1.4.1. Some Major Stock Exchanges in Developed Countries

In recent years, various classifications have been made to categorize countries based on their wealth. One of the most well-known groups among these country groups is the G-7 countries, which are used to represent developed countries and include the seven wealthiest countries in the world. The G-7 countries consist of the United States, Japan, the United Kingdom, France, Canada, Germany, and Italy (WESP, 2021: 163-164). The most important stock exchanges in these developed countries are the New York Stock Exchange in the United States, the Tokyo Stock Exchange in Japan, the London Stock Exchange in the United Kingdom, the Paris Bourse in France, the Toronto Stock Exchange in Canada, the Frankfurt Stock Exchange in Germany, and the Milan Stock Exchange in Italy. The market capitalization and the number of listed companies for these major stock exchanges in G-7 countries in the year 2020 are examined in Figure 1.

Figure 1: Market Capitalization and Number of Companies for G-7 Country Stock Exchanges in 2020



**Reference: WFE, 2021 (Data obtained using the World Federation of Exchanges' Statistical Portal.)*

When examining the market capitalizations of major stock exchanges in developed countries for the year 2020, it can be observed that the New York Stock Exchange in the United States had the highest market value, with $26.232 trillion. Following the New York Stock Exchange, the Tokyo Stock Exchange had $6.718 trillion, the London Stock Exchange had $4.045 trillion, the Paris Bourse had $2.997 trillion, the Toronto Stock Exchange had $2.608 trillion, the Frankfurt Stock Exchange had $2.284 trillion, and the Milan Stock Exchange had $734 billion in market capitalization. In terms of the number of listed companies in 2020, the Tokyo Stock Exchange in Japan ranked first with 3,758 listed companies, while the New York Stock Exchange was second with 2,974 listings. These two stock exchanges were followed by the London Stock Exchange with 2,347 listings, the Toronto Stock Exchange with 1,660 listings, the Paris Bourse with 846 listings, the Frankfurt Stock Exchange with 485 listings, and the Milan Stock Exchange with 464 listings. When evaluating the market capitalizations and the number of listed companies for major stock exchanges in developed countries in 2020, it can be seen that the New York Stock Exchange, with the highest market value, had relatively fewer listed companies compared to the Tokyo Stock Exchange. In this section, major stock exchanges in developed countries have been further explained.

1.4.1.1.New York Stock Exchange

The New York Stock Exchange (NYSE), commonly referred to as "Wall Street," has its roots dating back to the 1800s. It is renowned for its infrastructure and corporate structure and is the largest stock exchange in the world. As of 2020, the NYSE had a market capitalization exceeding $26 trillion and listed the financial assets of a total of 2,974 companies, with 2,466 being domestic and 508 foreign. Companies listed on the NYSE benefit from private liquidity providers and a highly integrated trading platform that utilizes advanced technology. The NYSE features several popular indices, including the NYSE 100, which includes the largest 100 companies listed on the exchange, the DJIA (Dow Jones Industrial Average) representing 30 major industrial companies, and the S&P 500, which covers 500 large companies. The NYSE 100 index, which spans a wide range of sectors and includes only U.S. companies, is commonly used to assess the overall market conditions. This index is weighted based on public market value and calculated based on both price and total return. It is reviewed quarterly, and companies with an average daily trading volume of less than 100,000 shares are removed from the index.

1.4.1.2.Tokyo Stock Exchange

The Tokyo Stock Exchange, established on May 15, 1878, is Japan's largest stock exchange. It serves as the operator of Japan's main cash capital market. As of 2020, it is the third-largest stock exchange in the world by market capitalization, exceeding $6.7 trillion, and it is the largest stock exchange in Asia. A total of 3,758 companies are listed on the Tokyo Stock Exchange, including 3,754 domestic companies and 4 foreign companies. The primary index used to gauge the performance of the Tokyo Stock Exchange is the Nikkei 225 Index, which is an important benchmark for all markets and investors. This index has been calculated since 1950 and consists of the 225 stocks with the highest trading volume on the Tokyo Stock Exchange, with a price-weighted calculation. Stocks with higher prices have more weight in the index, making large companies more influential in the index's movements. Given that Japan's economy is heavily reliant on exports to the United States, the Nikkei 225 Index closely follows the movements of U.S. markets and indices.

1.4.1.3.London Stock Exchange

The London Stock Exchange (LSE), established in 1801, is one of the world's oldest stock exchanges with a history spanning over 300 years. It is also among the largest stock exchanges globally, and foreign companies often choose to dual-list by listing their shares on the London Stock Exchange. In 2007, the London Stock Exchange acquired the Italian Stock Exchange for 1.6 billion Euros.

As of 2020, the London Stock Exchange had a market capitalization exceeding $4 trillion, with a total of 2,347 listed companies, including 1,979 domestic companies and 368 foreign companies. The London Stock Exchange develops various indices that serve as benchmarks for financial products. One of the most popular indices is the FTSE 100 Index, also known as the "Footsie." The FTSE 100 Index was created in January 1984 with a base level of 1,000 to measure the performance of the largest 100 companies that have passed size and liquidity screening on the London Stock Exchange. The abbreviation FTSE stands for Financial Times and Stock Exchange. Many international investors consider FTSE indices, particularly the FTSE 100, as representative of the UK market.

1.4.1.4.Paris Stock Exchange

The Paris Bourse, established in 1724 in France, is considered one of the important historical financial centers. On September 22, 2000, it merged with the Brussels and Amsterdam stock exchanges to form Euronext, and it became known as Euronext Paris. As of 2020, Euronext Paris had a market capitalization of approximately $3 billion and listed 846 companies. The most commonly used index on the exchange is the CAC 40 Index. The CAC 40 reflects the performance of the largest and most active 40 stocks listed on Euronext Paris, with a market capitalization-weighted composition. The CAC 40 Index was first calculated on June 15, 1988, and serves as the basis for structured products, funds, exchange-traded funds, options, and futures. Two-thirds of the businesses and activities of the companies listed in the index are located outside of France, and approximately 45% of the listed shares are owned by foreign investors. The CAC 40 Index is similar to the Dow Jones Industrial Average in the United States, the DAX in Germany, and the Nikkei in Japan. Profit and loss calculations for companies in the CAC 40 are made in Euros.

1.4.1.5.Toronto Stock Exchange

The Toronto Stock Exchange (TSX), also known as "Toronto Menkul Kıymetler Borsası" in Turkish, was founded in 1861. It is one of the largest securities exchanges in North America and is located in Toronto, Canada. Trading on the Toronto Stock Exchange is conducted in the Canadian dollar currency. As of 2020, the Toronto Stock Exchange had a market capitalization exceeding $2.6 trillion and listed 1,660 companies, making it the largest stock exchange in Canada. The primary index used to gauge the performance of the Toronto Stock Exchange is the TSX Index, which covers approximately 95% of the Canadian stock market. The TSX Index has been a leading benchmark for market activity in Canadian stock markets since its inception in 1977. It serves both as a benchmark for comparison and as an investment tool. The TSX Index is designed to preserve the liquidity characteristics of narrower indices while providing a representation of a broad benchmark index. It undergoes a review by the index committee every quarter to ensure its accuracy and relevance.

1.4.1.6.Frankfurt Stock Exchange

The Frankfurt Stock Exchange, with a history dating back to 1585, is Germany's most significant stock exchange and one of the oldest in the world. It plays a central role in the German economy, which is one of the world's largest. The Frankfurt Stock Exchange hosts the trading of over half of the stocks listed on the other seven stock exchanges in Germany. As of 2020, the Frankfurt Stock Exchange had a market capitalization exceeding $2.28 trillion and listed a total of 485 companies, including 438 domestic companies and 47 foreign companies. The primary benchmark index for the exchange is the DAX 30 Index, which represents the performance of the largest 30 companies in Germany in terms of trading volumes and market capitalization. The DAX 30 Index was first calculated in 1988, and the list of companies composing the index is reviewed every three months based on their performance. The DAX 30 Index includes some of Germany's leading brands and is closely watched by investors and experts as an indicator of the overall economic situation in Germany.

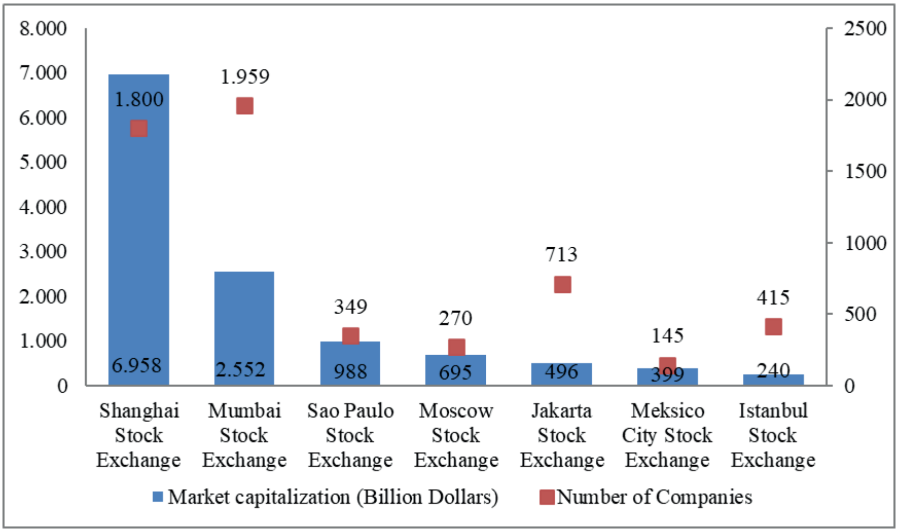
1.4.1.7. Milan Stock Exchange

The Milan Stock Exchange, founded in 1808, is Italy's most significant securities exchange. In 2007, it was acquired by the London Stock Exchange Group, becoming a part of that group. However, in 2020, the London Stock Exchange Group decided to sell the Milan Stock Exchange to Euronext, one of Europe's leading stock exchange operators. As of 2020, the Milan Stock Exchange had a market capitalization of $734 billion and listed the financial assets of 464 companies, including 326 domestic and 138 foreign companies. The primary benchmark index for the exchange is the FTSE MIB, which is designed to measure the performance of the Italian stock markets. It is a price index that includes leading companies from various sectors in Italy with high liquidity. The FTSE MIB index replaced the MIB-30 index and is essential for indicating both the performance of the Italian stock market and the country's economic direction. The FTSE MIB index covers approximately 80% of the total domestic market capitalization, making it a significant indicator for both Italian stock market performance and the country's economic outlook.

1.4.2.Some Important Stock Exchanges in Developing Countries

In recent years, the concept of "E-7 (Emerging-7) countries" is frequently used to refer to some developing nations. These countries have shown high economic performance among developing market economies, setting them apart from other developing nations. The E-7 countries, all of which are members of the G-20, are noteworthy for their rapid economic growth in recent years. The E-7 countries consist of the People's Republic of China, India, Brazil, Russia, Indonesia, Mexico, and Turkey. This section examines the market capitalization and the number of listed companies on these stock exchanges in the E-7 countries, which are progressing on the path of development and are considered emerging market economies. The market capitalization and the number of listed companies on the major stock exchanges in the E-7 countries in the year 2020 are analyzed in Figure 2.

Figure 2: Market Capitalizations and Number of Companies on E-7 Country Stock Exchanges in 2020



**Reference: WFE, 2021 (Data obtained using the World Federation of Exchanges' Statistical Portal.)*

When examining the market capitalizations of significant stock exchanges in developing countries in the year 2020, it is observed that the Shanghai Stock Exchange in the People's Republic of China had the highest market value at 6.958 trillion dollars. Following the Shanghai Stock Exchange, the Mumbai Stock Exchange in India had a market capitalization of 2.552 trillion dollars, the Sao Paulo Stock Exchange had 988 billion dollars, the Moscow Stock Exchange had 695 billion dollars, the Jakarta Stock Exchange had 496 billion dollars, the Mexico City Stock Exchange had 399 billion dollars, and the Istanbul Stock Exchange had 240 billion dollars. In 2020, when the number of registered companies is examined, it stands out that the Mumbai Stock Exchange in India had the highest number with 1,959 companies, followed by the Shanghai Stock Exchange in the People's Republic of China with 1,800 companies. These two exchanges were followed by the Jakarta Stock Exchange with 713 listings, the Istanbul Stock Exchange with 415 listings, the Sao Paulo Stock Exchange with 349 listings, the Moscow Stock Exchange with 270 listings, and the Mexico City Stock Exchange with 145 listings. When evaluating the market values and the number of listed companies on major stock exchanges in developed countries together, it is noted that although the Shanghai Stock Exchange had the highest market value, it had relatively fewer listed companies compared to the Mumbai Stock Exchange. This suggests that the Shanghai Stock Exchange may have larger, more established companies, while the Mumbai Stock Exchange has a larger number of smaller companies. This section also provides information on the important stock exchanges in countries that are making progress on the path to development and are often referred to as rising market economies.

1.4.2.1.Shanghai Stock Exchange

The Shanghai Stock Exchange, established on November 26, 1990, is the first stock exchange founded in the People's Republic of China. In 1992, a regulation was introduced to allow foreign investors to trade on the Shanghai Stock Exchange. This move was expected to contribute significantly to the development of the capital market (SSE, 2021). As of 2020, approximately 1,800 companies have their securities traded on the Shanghai Stock Exchange, with a market capitalization of around 7 trillion dollars. The main index used to reflect the market performance of the Shanghai Stock Exchange is the SSE Composite Index. The components of the SSE Composite Index cover all the stocks listed on the Shanghai Stock Exchange. The index was first calculated on July 15, 1991, with a base date of December 19, 1990. The base period is the total market value of all stocks on that day (Liang, 2010: 19-36; Bekaert and Hodrick, 2012: 402).

1.4.2.3.Mumbai Stock Exchange

The National Stock Exchange of India (NSE) was established in 1992 under the leadership of brokerage firms and banks, and it has a unique status as a private company, unlike other stock exchanges in the country. The NSE is headquartered in Mumbai and has five branches across the country (Singh, 2012, 94). As of the year 2020, the NSE has a market capitalization exceeding 2.5 trillion dollars, with a total of 1,959 companies listed, including 1 foreign company and 1,958 domestic companies. The NIFTY 50 index is the most important index calculated on the National Stock Exchange of India. It began its calculation in 1996 and is primarily based on market capitalization, consisting of 50 stocks from 24 different sectors. The index is reevaluated every six months, specifically on January 31 and July 31 (Nseindia, 2021: 1).

1.4.2.4.Sao Paulo Stock Exchange

The Sao Paulo Stock Exchange is one of the advanced stock exchanges in the Latin American region, and it was founded in the year 1890. On May 8, 2008, it merged with the Brazil Mercantile and Futures Exchange (Bolsa de Mercadorias & Futuros or BM&F) to operate under a single entity (Nyasha and Odhiambo, 2013: 6). As of the year 2020, the Sao Paulo Stock

Exchange, with a market capitalization approaching 1 trillion dollars, lists the stocks of 349 companies. The most well-known index associated with the Sao Paulo Stock Exchange is the Bovespa Index. Initially created in 1968, this index is primarily calculated based on market capitalization and is reevaluated every four months. The Bovespa Index represents the largest companies in the Brazilian capital market, accounting for approximately 80% of Brazil's capital market and is viewed as a general indicator of the country's economy (Bovespa, 2021: 1).

1.4.2.5.Moskow Stock Exchange

The RTS Stock Exchange, based in Moscow, was established in 1995 to facilitate trading on a consolidated system for different local exchanges (Taliyev, 2011: 53). As of the year 2020, the market capitalization of the exchange stands at 695 billion dollars, with a total of 270 listed companies, comprising 213 domestic and 57 foreign companies (WFE, 2021). The primary benchmark index for the Moscow Stock Exchange is the RTS Index. The RTS Index is a capitalization-weighted composite index calculated based on the prices of the largest and most liquid stocks traded on the exchange. It was first calculated on September 1, 1995, with a starting value of 100 basis points and is calculated in real-time, expressed in US dollars (Moex, 2021: 1).

1.4.2.5.Jakarta Stock Exchange

The Indonesia Stock Exchange was originally opened in 1977 as the Jakarta Stock Exchange. Initially, it operated as a government institution. In 1992, the exchange was privatized, and it transitioned into a company structure owned by brokerage firms. In 2007, the Jakarta Stock Exchange and the Surabaya Stock Exchange were merged under the umbrella of the Indonesia Stock Exchange (IDX) (Johnson and Soenen, 1996: 37-38). As of the year 2020, the market capitalization of the Indonesia Stock Exchange stands at 496 billion dollars, with 713 companies listed. The most important index of the exchange is the IDX Composite Index. This index, which started its calculation in 1983, measures the stock price performance of all companies listed on the main and development boards of the Indonesia Stock Exchange. The IDX Composite Index is viewed as an indicator of the Indonesian economy (IDX, 2021: 1).

1.4.2.6.Mexico City Stock Exchange

The Mexico Stock Exchange, headquartered in Mexico City, is the largest securities exchange in the country. It was originally established as the Mexico Commercial Exchange in 1886 and adopted its current name in 1975. The

Mexico Stock Exchange is the second-largest securities exchange in Latin America in terms of market capitalization of the companies listed (Leria, 2011: 4). As of the year 2020, the market capitalization of the Mexico Stock Exchange is approximately 400 billion dollars, with the financial assets of 145 companies being traded, consisting of 140 domestic companies and 5 foreign companies. The IPC Index (Indice de Precios y Cotizaciones) is the primary reference stock market index designed to measure the performance of the largest and most liquid stocks listed on the Mexico Stock Exchange. This index is a market capitalization-weighted index that represents stocks from various sectors of the economy that are listed on the exchange. The IPC Index started its calculation on October 30, 1978, and as of February 2009, it includes the exchange's A-group shares (S&PGlobal, 2021: 1).

1.4.2.7.Istanbul Stock Exchange

Borsa Istanbul, originally founded in Istanbul during the Ottoman Empire in 1866 under the name Dersaadet Tahvilat Borsası. In 1906, it was renamed as Esham ve Tahvilat Borsası. After the establishment of the Republic of Turkey, in 1929, it was renamed as the Istanbul Securities and Exchange Bourse. During this period, the lack of sufficient capital within the country and the existence of laws protecting the Turkish currency hindered the stock exchange from achieving its desired and targeted success. In the 1980s, efforts to develop the capital markets gained momentum, and in 1985, the first stock exchange in Istanbul was established under the name Istanbul Securities Exchange (İMKB). In 2013, Borsa İstanbul rebranded and modernized itself with a new image. As of 2020, Borsa İstanbul has a market capitalization of 237.4 billion dollars and lists the securities of 415 companies. Borsa İstanbul calculates various types of indices to allow investors to track market movements. The BIST 100 Index is the main index used for the Borsa İstanbul stock market. It is a continuation of the Composite Index that started with 40 companies' stocks in 1986 and has gradually been limited to 100 companies' stocks over time. A review is conducted every three months to determine which stocks will be added to or removed from the index (BIST, 2021: 33).

1.5.PREDICTION AND ANALYSIS IN THE STOCK MARKET

In the world of investing, there are two main analysis methods that are utilized before making investment decisions. These methods are fundamental analysis and technical analysis (Gurav and Sidnal, 2018: 2). Before delving into the topics of fundamental and technical analysis, it is essential to examine the "Random Walk Theory" and the "Efficient Market Hypothesis" proposed by Eugene Fama. The validity of these theories is a subject of debate in financial circles. These theories suggest that it may not be possible to predict market trends in a systematic way and that markets are highly efficient, making it difficult to gain an advantage through analysis or predictions.

1.5.1.Random Walk Theory and Efficient Market Hypothesis

Random Walk Theory, suggests that each security's price in the stock market follows a random walk, and the movements in one security's price are independent of the movements in another security's price. In theory, it is proposed that a stock's market price essentially forms randomly, and any attempt to predict the stock market is bound to fail. In other words, it argues that stock prices are not dependent on past prices, and there are no trends, making predictive models useless. According to the theory, security prices will change as new information about the economy, sectors, and companies emerges. Therefore, any random information that surfaces will also lead to random movements in prices. Even when information becomes predictable or expectations exist, they will be quickly reflected in security prices as they emerge (Fama, 1965: 56).

The Efficient Market Hypothesis is a long-debated topic in the finance literature. However, there is still no common consensus on it. Essentially, the hypothesis is based on the idea that price changes follow the concept of a random walk. In general, a securities market is considered efficient if the prices of securities traded in the market fully reflect all available information and if prices respond rapidly and accurately to new information (Deckman and Dale, 1986: 5). The Efficient Market Hypothesis relies on certain fundamental assumptions, including the presence of numerous buyers and sellers in the market, rationality among investors, complete information and available funds for investment, and the use of all available information to maximize benefits, and investors cannot make extra profits using the information in the market (Redhead, 2008: 483-484). Prices change randomly in response to new information, and the impact of new information on prices is immediate. Transaction costs in the market are also quite low (Barone, 2003: 1).

The Efficient Market Hypothesis classifies market efficiency into three different forms based on the types of information. In the weakest form, known as weak-form efficiency, it is assumed that investors cannot achieve above-average returns by using historical price movements. Therefore, the use of technical analysis is considered to be of no benefit. In the next level, known as semi-strong form efficiency, it is assumed that investors cannot achieve above-average returns by using publicly available information in addition to past prices. Consequently, techniques like fundamental analysis, financial statement analysis, or other publicly available information are deemed to be unhelpful in making investment decisions. Finally, in the strongest form, known as strong-form efficiency, it is assumed that even with access to private or insider information, investors cannot achieve above-average returns. This suggests that even with non-public information, no one can make abnormal profits. When the market is strongly efficient, it implies that all information, whether public or private, is rapidly reflected in prices. The Efficient Market Hypothesis posits that predicting stock price movements is impossible because stock prices already incorporate all available information. Although the hypothesis gained wide acceptance in the 1970s, it has faced increasing criticism and challenges since the 1980s. Despite years of research and numerous academic studies, there is still no consensus on whether financial markets are genuinely efficient or not (Fama, 1970: 383-384).

The Efficient Market Hypothesis (EMH) asserts that it is impossible to predict the price movement of financial securities, making it impossible for a stock market investor to outperform or "beat" the market in the long run. However, research has shown that certain techniques can provide investors with returns above the norm on various stock exchanges. While in the 1970s, the Efficient Market Hypothesis's assumptions were primarily criticized, by the 1980s, with the increasing number of opposing critical studies, the hypothesis as a whole started to be questioned. Despite years of research and thousands of journal articles, there is still no consensus on whether financial markets are efficient or not (Lo and MacKinlay, 2002: 6).

1.5.2.Analysis Methods Used in the Stock Market

Investment in the stock market requires a well-balanced approach to generate returns that can offset the risks involved. The selection of financial assets and the timing of buying and selling play a crucial role in achieving this balance. Fundamental and technical analysis methods are of great importance in determining the selection of financial assets and the timing of investments. In fundamental analysis, the true value of stocks is determined, and the aim is to identify undervalued stocks based on their market value compared to their intrinsic value for investment. In technical analysis, historical price movements of financial assets are examined in an attempt to predict the future direction of prices (Akgüç, 2013: 301). While fundamental analysis has been around for a long time, technical analysis can be considered a relatively newer analytical method compared to fundamental analysis. This section delves into the topics of fundamental and technical analysis.

1.5.2.1.Fundamental Analysis

Fundamental analysis is an analytical method used to determine the true value of a stock by taking into account economic, sector-specific, and company-related factors that can influence the stock's value. It involves finding the intrinsic value of a stock and comparing it to market prices to make buy or sell decisions. The primary goal of fundamental analysis is to discover the real value of a stock. In fundamental analysis, when a stock's market price is below its intrinsic value, a buying decision is typically recommended, and when it is above, a selling decision is considered (Suresh, 2013: 44). Also known as intrinsic value analysis, fundamental analysis involves examining a company's fundamental and financial data to determine the value of its stock and subsequently forecast the stock's future performance. Therefore, fundamental analysis encompasses evaluating the company's operations and growth prospects, the industry in which it operates, and the broader economic conditions (Quirin and Allen, 2000: 149).

In fundamental analysis, the intrinsic value of publicly traded securities is sought without necessarily relying on the prices of these securities in the capital market. Expectations of the company, disclosed financial reports, product information, market data, and the overall economic environment are all considered. Financial reports, market data, and environmental information are used to perform a valuation that can be compared to market prices. Through such a comparison, investment strategies that can yield above-average returns can be developed (Aktaş, 2008: 44).

Fundamental analysis consists of three stages: economic analysis, sector analysis, and company analysis.

- In the economic analysis, macroeconomic indicators within the economy are examined. Factors like increasing national income, low inflation, a growing gross domestic product (GDP), and lower price volatility have a positive impact on the stock market. When the economy is in a growth trend, companies benefit from this growth, leading to increases in stock values. In contrast, during an economic downturn, companies are affected, resulting in declines in stock values (Drakopoulou, 2015: 2; Baresa et al., 2013: 45).

- Sector analysis involves studying the specific sectors in which investments will be made to determine which sector or sectors to invest in. Each sector responds differently to economic trends, so classifying companies by sectors is crucial for diversifying risks. The growth potential of the chosen sector is also researched. This involves examining the sector's life cycle and determining its current position within that cycle. Understanding the life cycle stage of a sector helps in predicting the future success of companies operating in that sector. The competitive conditions within the sector are also investigated (Baresa et al., 2013: 47; Canbaş and Doğukanlı, 2007: 547).
- In company analysis, it's crucial to forecast the future performance of the chosen company and determine the intrinsic value of its stock. Company analysis comprises two main areas: qualitative and quantitative factors. Qualitative factors related to the company include characteristics of the products and services, demand conditions, customer loyalty, market share, company culture, and the quality of management. Quantitative factors, on the other hand, involve examining the company's financial statements and reports for the current and past years. The financial performance of the company is analyzed using financial analysis techniques. One of the most well-known and used methods in financial analysis is ratio analysis, which involves comparing related financial statement items to explore their relationships (Apak and Demirel, 2009: 184; Suresh, 2013: 45). In ratio analysis, different types of ratios are calculated to evaluate various aspects of a company's performance. Liquidity ratios assess the company's ability to meet short-term obligations, activity ratios measure the efficiency and productivity of the company's operations, leverage ratios examine the company's capital structure and long-term debt-paying ability, and profitability ratios assess whether the company's earnings are sufficient (Spahija and Xhaferi, 2019: 32).

1.5.2.3. Technical Analysis

Technical analysis is a technique that involves analyzing historical data for an asset, typically using charts, to make predictions about its future performance (Kirkpatrick and Dahlguist, 2011: 3). According to the principles of technical analysis, stock prices move in trends that can be predicted in advance. Technical analysis acknowledges that new information doesn't instantly reach all investors in the market. Stock prices don't reach their equilibrium instantly but rather over a certain period of time. Therefore, technical analysis directly contradicts the Efficient Market Hypothesis, which assumes that new information is instantly reflected in stock prices (Konuralp, 2001: 308).

In the formation of price trends in the market, various factors and influences come into play. Technical analysis, when examining price trends, focuses on the consequences of these factors and influences rather than their causes. Therefore, in technical analysis, the aim is not to measure the true value of a stock but to identify trends that can indicate how a stock will behave in the future (Papuccu, 2019: 247). To identify trends, charts are prepared that show the daily, weekly, and price, as well as the lowest and highest prices and trading volumes of stocks. In addition to stock charts, charts related to market and industry sectors are also examined as part of the analysis (Akgüç, 2013: 862). As a result of this analysis, buy and sell points for a stock or financial asset are determined, and attempts are made to predict the direction in which price movements will go. Uncertainty and dealing with probabilities are fundamental principles of technical analysis (Pring, 2014: 29).

Charles Dow, considered the pioneer of technical analysis, conducted his initial studies in 1900-1902, emphasizing the use of averages to determine the general trends of stocks. His work, known as Dow Theory, introduced eight assumptions and principles. These principles include the idea that averages discount everything, the market moves in trends, the primary trends, Bull and Bear markets, consist of three phases, averages must confirm each other, volume should confirm the trend, trends are assumed to continue until a clear reversal signal is given, lines can substitute for second trends, and when conducting analysis, only closing prices should be used (Myers, 2003: 207).

In technical analysis, identifying medium and long-term trends is relatively simpler, but errors in short-term trend predictions can increase. Therefore, technical indicators, known as indicators, are used to identify short and medium-term trends. Technical indicators allow analysts to see the direction in which prices are moving and measure the strength of the trend. Technical analysts use these indicators to make buy and sell decisions (Ceylan and Korkmaz, 2010: 341).

There are numerous technical indicators used in technical analysis, and many of them are widely accepted and frequently employed. In this section, some of the commonly used technical indicators in the literature have been explained.

- One of the simplest methods commonly used to predict future prices based on past price movements is the use of moving averages. Moving averages help to smooth price fluctuations, identify primary trends, and provide a more accurate means of tracking prices. In this method, the average of past prices for a specific period is calculated. Subsequently, for each new day, the average is recalculated by moving a specific number of days back. Moving averages are plotted on price charts, and signals for buying or selling are generated when the price chart crosses above or below the moving average or when moving averages with different periods intersect. The choice of periods can vary based on the investor's profile and the characteristics of the investment period (Murphy, 2000: 45).

- The Moving Average Convergence Divergence (MACD) indicator was developed by Gerald Appel in 1979. It is obtained by subtracting the 26-day moving average of prices from the 12-day moving average. An additional 9-day moving average, known as the signal line, is then plotted over the MACD. When the indicator falls below the signal line, it suggests a sell signal, and when it rises above the signal line, it indicates a buy signal (Ceylan and Korkmaz, 2010: 344). Because this indicator is derived from moving averages, it is used as a trend-following indicator. Therefore, it can be useful in trending markets. However, in sideways or ranging markets, the signals provided by the indicator can be misleading. The best results are often obtained when using weekly charts (Appel, 2005: 183).
- The Commodity Channel Index (CCI) was developed by Donald R. Lambert in 1980. It is an indicator that measures how far prices deviate from their averages, indicating the strength of the current trend. The higher the index value, the stronger the upward trend, and the lower the index value, the stronger the downward trend is perceived (Wood, 2009: 11). The index value is calculated using the simple moving average method. For the CCI, a suitable period should be about one-third of the cycle length, but any period between 5 and 25 days can be used. Typically, a 20-day period is considered the standard data basis length (Lambert, 1980: 8).
- The Relative Strength Index (RSI) was developed by Welles Wilder in 1978. It is an indicator designed to compare a stock's price to its own past price to measure the relative strength of a stock's past performance (Pring, 2014: 80). The indicator value, calculated by dividing a stock's price by an average or an index value, can range from 0 to 100 based on the movement in prices. Crossing above the reference level of 50 indicates that prices will rise, while crossing below suggests that

prices will fall. A value above 70 indicates that prices will increase significantly and then decrease shortly, while a value below 30 suggests that prices have hit a bottom and will increase shortly. Support and resistance levels are easily identifiable in the indicator. Relative Strength Indices in the range of 9-25 days are commonly used based on the stock's characteristics and the desired timeframe. It is a good indicator for showing short-term changes in the trend (Karan, 2004: 522).

- The Stochastic Oscillator, developed by George C. Lane in the late 1950s, is based on the assumption that in an upward trend, closing prices will cluster near the upper end of the period's trading range, while in a downward trend, they will cluster near the lowest price point. The Stochastic Oscillator tracks momentum rather than price, and because momentum changes direction before price, it provides signals earlier than price-following indicators (Naved and Srivastava, 2015: 926). The Stochastic Oscillator uses two different lines in its calculations. The points where these lines intersect each other represent buy or sell signals. Values on the Stochastic Oscillator range from 0 to 100. A value above 70 is considered a signal that prices will rise, while a value below 30 suggests that prices will fall. The Stochastic Oscillator is considered a successful indicator for stock buy and sell decisions (Murphy, 1999: 246).
- The Momentum indicator, developed by Welles Wilder in 1978, measures the rate and strength of a stock's upward movement. It is calculated by taking the difference between today's price and the price from a specific number of periods ago (Wilder, 1978: 53). The Momentum indicator reveals the speed, strength, and pace of price changes and trend movements. Crossing the 100 axis upward is interpreted as a buy signal, while crossing it downward is interpreted as a sell signal. The upper boundary for Momentum is typically set at 120, and the lower boundary at 80. When the indicator is outside this range, it is considered to represent extreme levels (Ceylan and Korkmaz, 2010: 344). Since it is used to track short-term movements, the selected number of days is usually less than 30. The Momentum indicator, typically created using 10-day periods, is interpreted based on its fluctuations around the zero line (Murphy, 1999: 229).

CHAPTER 2

Machine Learning

In this section, we first introduce fundamental concepts related to the discipline of machine learning. Then, we delve into the history of machine learning. Following that, we focus on the types of machine learning, including supervised, unsupervised, semi-supervised, and reinforcement learning. Next, we attempt to explain machine learning methods. In the final section, we summarize the machine learning algorithms commonly used in predicting stock market index directions.

2.1.THE DISCIPLINE OF MACHINE LEARNING

Machine learning is a subfield of artificial intelligence that enables computer systems to learn from data and make predictions about future data using the knowledge acquired through this learning process. The word "learning" implies not only gaining knowledge or understanding through reading, education, or experience but also improving performance through experience. Machine learning is a growing, evolving, and relatively new field of artificial intelligence (Mitchell, 2006: 1-7; Dulhare and Ahmad, 2020: 156). There are various definitions of machine learning in the literature. In its most general sense, machine learning can be defined as systems based on algorithms that can learn from datasets and improve their performance over time with more data (Cui, Wong, and Lui, 2006: 598). Additionally, machine learning can be described as a field of study based on developing computer algorithms that transform data into understandable actions (Bell, 2020: 2). According to the definition provided by the Presidency of the Republic of Turkey, Digital Transformation Office, machine learning is both an "automated data analysis method for creating analytical models" and a subfield of artificial intelligence based on the idea that "systems can learn from data, identify patterns, and make decisions with minimal human intervention" (Turkey National Artificial Intelligence Strategy; 2021: 92).

In the pioneering work by Mitchell (1997) in the field of machine learning, machine learning is defined as "a computer program that can improve its performance on a task T, as measured by a performance measure P, with experience E" (Mitchell, 1997: 17-18). In Mitchell's work from 2006, where the practical value of machine learning algorithms in various application domains is emphasized, a different definition is provided. In this study, machine learning is expressed as follows: "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E" (Mitchell, 2006: 1). A machine learning approach involves a set of design choices that specify the type of training experience, the target function to be learned, a representation for this target function, and an algorithm for learning the target function from training examples. A well-defined machine learning problem requires a well-defined task, performance measure, and source of training experience (Mitchell, 2006: 1).

Machine learning is fundamentally focused on two related questions: "How can we build computer systems that automatically improve with experience?" and "What are the fundamental theoretical laws that govern all learning systems, regardless of whether they are implemented in humans, in computers, or in organizations?" Machine learning is a highly significant field both in addressing these fundamental scientific and engineering questions and in generating highly practical applications in various domains (Mitchell, 2017: 1). Additionally, as an interdisciplinary field, machine learning shares common ground with mathematical areas such as probability, statistics, and optimization, alongside computer science and artificial intelligence. Since the goal is to program machines to learn, machine learning is inherently a subfield of computer science. However, it is also seen as a branch of artificial intelligence because the ability to turn experience into expertise and discover meaningful patterns in complex data is fundamental to human intelligence. In contrast to traditional artificial intelligence, which attempts to mimic intelligent behavior, machine learning aims to leverage the strengths and special abilities of computers to complement human intelligence and tackle tasks that often go beyond human capabilities (Shwartz and David, 2014: 24-25).

Machine learning is built upon mathematical disciplines like probability, statistics, and optimization theory to extract patterns from data. It involves creating models to enable machines to learn from historical data and make predictions about the future. In this sense, machine learning differs from traditional software engineering methods. In traditional methods, software developers look at data sets and attempt to create patterns. Once patterns are identified, a set of rules is applied to transform input data into the desired output. These rules are explicitly coded using a programming language, resulting in software. However, in machine learning, unlike traditional programming, all the hard work is done by machine learning algorithms instead of a human (Grigorev, 2020: 1-2). With the help of statistical methods and computational power, machine learning algorithms can identify trends, detect hidden patterns and behaviors, and make future predictions (Geron, 2019: 6).

2.2.THE HISTORY OF MACHINE LEARNING

Machine learning encompasses applications that simulate human behaviors such as learning, reasoning, and problem-solving. The idea of machines exhibiting intelligent behaviors akin to humans has always been an area of interest throughout history. Over the years, machine learning research has been pursued with varying levels of intensity, employing different approaches and emphasizing different objectives. Several developments can be considered as milestones in the historical evolution of the concept of machines learning. One of these milestones is the invention of the mechanical calculator. In 1642, Blaise Pascal invented the first mechanical calculator capable of performing addition. In 1672, Gottfried Leibniz improved the previous model, creating a mechanical calculator that could perform addition, subtraction, multiplication, and division. Certainly, historical hoaxes related to machine learning also exist. Approximately a century after the invention of the mechanical calculator, in 1770, Wolfgang Von Kempelen created a robot named the "Mechanical Turk," which could play chess. This machine, designed in the form of a Turkish figure carved from wood and dressed in Ottoman attire, featured a chessboard in front of it and was placed on a wheeled cabinet. It consisted of numerous levers, pulleys, and complex mechanical systems. The mechanical chess-playing robot was presented to Empress Maria Theresa of Austria and played chess against her advisor, defeating him in 30 minutes. In subsequent periods, the robot played chess against various individuals throughout Europe, often emerging victorious, which contributed to its recognition. In 1840, Edgar Allan Poe, the robot's final owner, revealed the true nature of the machine. It was disclosed that there was a hidden compartment inside the cabinet large enough to accommodate a human, and the machine was an illusion

(Will, 2013: 1). It is believed that the possibility of machine intelligence in Kempelen's deception served as an important source of inspiration for the development of the concept of machine learning. Following Kempelen, Charles Babbage (1834) aimed to develop a mechanical machine capable of displaying truly intelligent behaviors. However, the outcome of this work concluded that a machine exhibiting behaviors as intelligent as a human could not be produced (Babbage, 1834). In 1842, Ada Lovelace created the first computer algorithm, which was a significant development enabling the creation of faster and more practical computer applications (Wilson, 2018).

In the modern world, the concept of machine learning was first introduced in Alan Turing's 1950 work titled "Computing Machinery and Intelligence." In this work, which aimed to answer the question, "Can machines think?" the idea that machines could think intelligently like humans was put forward. The study included three participants: a human, a human judge, and a machine, and it explored the notion that machines could think just like humans (Turing, 1950: 433-460). Nine years after Turing, Arthur Samuel from IBM played a significant role in popularizing the concept of machine learning by using the term itself. He introduced a self-learning checkers game program in his work "Some Studies in Machine Learning Using the Game of Checkers" (Samuel, 1959: 535-554). Another significant development in the field of machine learning occurred in 1959. Cahit Arf's work titled "Can Machines Think, and How Can They Think?" emphasized that machines could exhibit certain characteristics that might indicate they can think like a human. The study also provided examples showing that machines could be designed to perform logical and analytical tasks, such as calculation, language use, and analogy-making (Arf, 1959: 91-103). In 1962, the expert checkers player Robert Nealey played the game against an IBM 7094 computer and was defeated. This development is considered a turning point in the field of machine learning (Bell, 2020: 2).

The second half of the 20th century marked the golden age of machine learning, with many significant developments in the field. The first of these developments took place in 1957 when the first "artificial neural network" for computers was created. This neural network was developed by Frank Rosenblatt with the aim of simulating the thinking process of the human brain. Ten years after the creation of the initial neural network, in 1967, T.M. Cover and P.E. Hart introduced the "Nearest Neighbor Algorithm," which was used to solve data classification problems (Cover and Hart, 1967: 21). In 1981, Gerald Dejong developed the "Explanation-Based Learning" system, allowing a computer to analyze training data and create a general rule by discarding irrelevant data, effectively enabling machine learning (Dejong, 1981). In 1985, a notable innovation in the field of machine learning occurred with the formation of the first machine learning research team as part of AT&T (ATT, 2021). In 1992, the first automatic speech recognition system was introduced. This system was created by Padma Ramesh and Jay Wilpon using a machine learning approach called Hidden Markov Models (Ramesh and Wilpon, 1992: 381). 1995 saw the proposal of the "Support Vector Machines" algorithm by Corinna Cortes and Vladimir N. Vapnik, which revolutionized large-scale classification (Cortes and Vapnik, 1995: 273). In the same year, Haffner, LeCun, and Bengio recommended the first convolutional neural network with a large-scale application to control recognition (Haffener, LeCun, and Bengio, 1995). In 1996, the Adaboost algorithm was developed by Freund, Y. and Schapire R., which is used for processing unstructured data through decision trees (Freund and Schapire, 1996: 325).

One of the significant developments in the field of machine learning occurred in 1988 when artificial neural networks were used to predict the price movements of financial assets. With the help of the created model, IBM stock prices were predicted by identifying nonlinear movements and patterns in historical data (White, 1988: 451-458). In 1997, an IBM computer named "Deep Blue" faced off against the world chess champion, Garry Kasparov. Designed to explore 200 million possible chess positions per second, Deep Blue defeated its opponent after a six-game match. Deep Blue made history by becoming the first machine to defeat a world chess champion (IBM, 2021: 1).

As the 21st century entered its first half, machine learning started to become one of the most important fields, and significant developments continued to take place. In 2001, natural language processing systems and other machine learning technologies were used together for the first time in interactive voice response systems. In 2011, researchers working on deep neural networks discovered new algorithms that made it possible to train models on millions of examples. In 2014, both Google and Facebook elevated machine learning to become a core technology for their businesses, with machine learning researchers from AT&T playing key roles in both companies. In 2015, Interactions acquired AT&T's speech and language research team. The merger of the organizations' customer service interactions and award-winning adaptive understanding technologies led to revolutionary new approaches. In 2017, Digital Roots, acquired by Interactions, provides machine learning-based social media services to businesses, allowing brands to quickly filter, respond to, and engage with their social media followers (Wilpon et al., 2021: 5-6).

In 2019, the Organization for Economic Co-operation and Development (OECD) adopted the "Principles on Artificial Intelligence," the first international standard for the responsible management of trustworthy artificial intelligence by governments (OECD, 2019). The OECD Financial Markets Committee's 2021-2022 report presented an analysis of artificial intelligence and machine learning. The report examined how early adopters in the financial sector have been impacted by machine learning technologies and how these innovative mechanisms have transformed their business models. The report emphasized that machine learning has found widespread application in the financial sector. Machine learning is utilized in various financial areas, including asset management, risk management, algorithmic trading, credit commitment processes, blockchain-based financial services, mobile banking operations, customer chatbots, credit commitment and scoring, credit loss estimation, fraud monitoring and detection, customer services, and insurance operations (OECD, 2021: 15).

In 2019, as part of the "Breakthrough Strategic Approach for Identifying Priority Technology Areas" conducted by the Presidency of the Republic of Turkey, the following areas were identified as priority technology areas: advanced functional material technologies, engine technologies, biotechnological drugs, the Internet of Things (IoT), energy storage, robotics, mechatronics, artificial intelligence, big data, information security, broadband technologies, and micro/nano and optoelectronic technologies. Among these areas, "Artificial Intelligence and Machine Learning" were determined as having "the highest impact potential in terms of economic impact, societal benefit, and national security". As of 2021, a total of 205 AI startups were identified in Turkey. Large-scale companies dominate AI investments through in-house investments. The most important sources of funding for new ventures are venture capital and private equity, with machine learning being one of the most funded areas (Presidential Digital Transformation Office, 2021: 47).

2.3.THE TYPES OF MACHINE LEARNING

Machine learning, as a subfield of artificial intelligence, focuses on the learning aspect of computers (Mirjalili, Faris, and Aljarah, 2020: 4). The types of machine learning are generally categorized into four main groups in the literature based on their usage, applications, and algorithms. These machine learning types include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, which will be explained in this section.

2.3.1.Supervised learning

Supervised learning is a learning technique that aims to predict the relationship between inputs and outputs, attempting to find patterns in labeled datasets (Cielen, Meysman, and Ali, 2016: 66). Supervised learning is also referred to as a mapping function, as it defines hidden patterns in the data. In supervised learning, there is a labeled dataset consisting of input data pairs and a target output. The target output is a value associated with the input data. A supervised learning algorithm processes the training labeled dataset to predict similar new datasets or the mapping function. Supervised learning can perform two main tasks: classification and regression (Mohri, Rostamizadeh, and Talwalkar, 2018: 4). The process is based on regression and classification models, and the learning process continues until the desired performance and functionality are achieved. An algorithm is created to map inputs to outputs using a set of variables. Regression algorithms aim to predict the dependent variable, while classification algorithms determine which category the data belongs to. Examples of supervised learning models include logistic regression, decision trees, random forests, or multilayer perceptrons (Geron, 2019: 26-27).

2.3.2.Unsupervised Learning

In cases where data is not properly classified and there are unlabeled observations, it is not possible to perform supervised learning. Unsupervised learning models are often utilized in data mining applications that involve large volumes of unstructured input data (Park, 2020: 24). Unsupervised learning is a learning technique that focuses on commonalities rather than responding to feedback. The method aims to determine the probability of commonalities in a specific dataset and uses these commonalities to develop a model. The algorithm, in line with its objective, learns how to accomplish a task without providing a logical approach for doing something. Therefore, the unsupervised approach is more complex than the supervised process. Unsupervised learning can also be described as employing a reward-based approach to confirm the success of achieving specific goals without providing explicit instructions on how to attain those goals (Watt, Borhani, and Katsaggelos, 2020: 14). In unsupervised learning, there are no concepts of correct or incorrect outputs; instead, similarities, differences, and patterns are mathematically represented. Unsupervised learning algorithms are frequently used for applications such as clustering, association, and anomaly detection (Geron, 2019: 25).

2.3.3.Semi-Supervised Learning

Semi-supervised learning is employed when there is a small amount of labeled data and a vast amount of unlabeled data. In semi-supervised learning, a combination of both labeled and unlabeled data is used. In other words, classified and unclassified training data are used together to make predictions for unknown points (Mohammed, Khan, & Bashier, 2017: 10). Semi-supervised learning is often preferred in cases where easy access to unclassifiable data is provided, and obtaining labeled data is costly (Neil, 2021: 12). When comparing semi-supervised learning and supervised learning techniques, it can be observed that in supervised learning, there is a large number of labeled data and a small number of data to be predicted, while in semi-supervised learning, the opposite is true (Kızılkaya and Oğuzlar, 2018: 92). In this context, semi-supervised learning can be said to be situated between supervised learning and unsupervised learning (Brownlee, 2016: 17).

2.3.4.Reinforcement learning

Reinforcement learning is a type of learning that proves beneficial when precise models are not realistic. The method focuses on how machines should operate to maximize certain aspects of cumulative reward and minimize risk (Mohammed, Mohsen, & Bahsier, 2017: 30). Reinforcement learning combines elements from supervised learning and dynamic programming (Harmon and Harmon, 1997: 3). Reinforcement learning algorithms advance the learning process through feedback. Continuous trials are conducted towards a goal, and the actions taken are classified as successful or unsuccessful until the goal is reached. Reinforcement learning algorithms ensure that they find the best possible outcome by not repeating erroneous actions based on their learned experiences. Compared to supervised and unsupervised learning types, it can be said that reinforcement learning is more complex and challenging to implement (Berry, 2020: 12).

2.4.MACHINE LEARNING METHODS

In the field of machine learning, there are classification, clustering, and regression methods available.

In classification methods, a unit's membership in a group is determined based on its characteristics (Harrington, 2012: 7). The algorithms written to solve classification problems are referred to as classifiers. The goal of the algorithm is to find a classifier that can label unprocessed examples with the highest accuracy and make predictions about new inputs that the system has not seen before. To achieve this goal, the algorithm maps input variables to targets, labels, or categories (Sarker, 2021: 5). The prominent classification algorithms include decision trees, random forests, k-nearest neighbors, support vector machines, and the Naive Bayes algorithm.

In clustering methods, data points are assigned to different clusters without knowing the structure of the data points (Neapolitan and Jiang, 2018: 320). Algorithms use a similarity measure to identify clusters in the data. The similarity measure can be expressed as Euclidean distance or probabilistic distance. Various algorithms can be used in the data clustering process, including density-based, hierarchical, partitioning, or grid-based algorithms (Geron, 2019: 25). Clustering algorithms are applied in various fields, including pattern recognition, data analysis, image processing, and business analysis. A clustering algorithm can have several desired features, such as scalability, the ability to extract clusters in any shape, handling different data types, not requiring input parameters, being insensitive to data input, the ability to perform noise-based clustering, and speed (Dulhare, Ahmad, and Ahmad, 2020: 204-205).

In regression methods, the effect of the variation in independent variables on the mean value of the dependent variable is estimated. More specifically, regression algorithms measure the impact of changing the value of an independent variable on the value of the dependent variable when other independent variables are held constant (Mohri et al., 2012: 2). The difference between regression and classification methods is that in regression, the predicted values are continuous, whereas in classification, they are discrete. In regression methods, unlike classification methods, the idea of proximity between different categories stands out (Harrington, 2012: 177). In the field of machine learning, linear and logistic regression algorithms are widely used (Campesato, 2020: 52).

2.5.MACHINE LEARNING ALGORITHMS

Algorithms are a set of instructions on how a computer should interact with data, how data should be manipulated, and transformed. An algorithm can be as simple as adding a column of numbers or as complex as recognizing a person's face in an image. Machine learning algorithms are different from other algorithms. In most algorithms, a programmer attempts to build the model by entering the algorithm. However, the process is different in machine learning. In machine learning, the data itself builds the model. As more data is added to the algorithm, the algorithm becomes more complex. A machine learning algorithm can perform better as it is exposed to more data (Hurwitz and Kirsch, 2018: 28).

In the field of machine learning, numerous different algorithms have been developed for use in classification, clustering, or regression methods. This section examines some of the popular classification algorithms widely used in various application areas, especially in predicting stock and market index directions.

2.5.1.Decision Trees Algorithm

Decision trees are one of the most commonly used machine learning algorithms for classification and prediction purposes (Williams, 2011: 205). In the study by Morgan and Sonquist (1963), the regression tree algorithm was proposed, while in the study by Morgan and Messenger (1973), the classification tree algorithm was recommended (Morgan and Sonquist: 1963; Morgan and Messenger: 1973). The decision trees algorithm builds the model by hierarchically and recursively partitioning the data according to an appropriate statistical method. The data to be classified, waiting at the root node of the tree, progresses along the flowchart on the tree based on the values of their own features. The goal is to maximize the impurity decrease value with the other nodes in the tree and minimize the impurity decrease value at each node. When the placement process in the corresponding class's leaves is completed, the classification process is completed with the decision tree model created (Witten, Frank, and Hall, 2011: 99). Decision trees provide an advantage in terms of ease of interpretation and comprehensibility for decision-makers since they represent a flowchart (Chien and Chen, 2008: 34)

2.5.2.Random Forest Algorithm

The random forest algorithm, as described by Breiman (2001), is defined as a combination of tree predictors where each tree is constructed independently based on a randomly sampled vector of values, and all trees in the forest have the same distribution (Breiman, 2001: 5). A generalization of decision trees, the random forest algorithm performs classification across multiple trees. Multiple decision trees of the same type are combined to create a forest of decision trees. When a numerical prediction is needed from the data, the average of tree predictions is calculated. When a categorical prediction is needed from the data, the mode of tree predictions is determined. There are two important differences between decision trees and the random forest algorithm. Firstly, in the decision trees algorithm, the entire sample is used, while in the random forest algorithm, only a subset of the sample is used. In other words, for each tree, a random and fixed number of observations are taken. This process is also called bootstrapping. Secondly, in the decision trees algorithm, all features in the model are optimized, while in the random forest algorithm, a random subset of features is chosen at each split (Hull, 2021; 136). In the field of machine learning, it is well-known that random forest algorithms have a high level of predictive accuracy. These algorithms are commonly used for solving both regression and classification problems (Breiman, 2001:29).

2.5.3.K-Nearest Neighbors Algorithm

The K-Nearest Neighbors Algorithm, proposed by Cover and Hart in 1967, is based on the principle of assigning an unclassified sample point to the class closest to it among a previously classified set of points (Cover and Hart, 1967: 21). In other words, in this algorithm, the classification of a data point is done by analogy based on its similarity to previous data points. To determine the class, each test example is processed individually with each data point in the training set. For instance, to determine its class, k data points in the training set that are closest to that specific example are selected. Among the selected examples, the class with the most instances is assigned to the test example (Han, Kamber, and Pei, 2012: 423). The K-Nearest Neighbors Algorithm is commonly used in regression and classification problems (Zhang et al., 2017: 43).

2.5.4.Naive Bayes Algorithm

The Naive Bayes algorithm is a machine learning algorithm built on the principles of conditional probability rules and is based on Bayes' theorem discovered by Thomas Bayes in 1812 (Berrar, 2018: 1). Minsky's work in 1961 is one of the pioneering studies that used Bayes' theorem as a classifier in algorithms (Minsky, 1961: 406). The Naive Bayes classification algorithm relies on assumptions such as independence between variables, normality, and the equal importance of all attributes. With the Naive Bayes classification algorithm, the class to which an example belongs is predicted based on the highest probability calculation. The algorithm learns which classes examples belong to from the data set allocated for training and performs classification predictions for observations in the test data based on previously observed probabilities. Bayes' theorem is used for probability estimation in this process (Barber, 2012: 242). This algorithm can achieve classification performance that is equal to or better than results obtained from more complex algorithms in terms of computation, ease of use, and requiring less time for implementation (Cichosz, 2015: 118).

2.5.5.Logistic Regression Algorithm

Due to its attributes, the logistic regression algorithm has become one of the most preferred algorithms in recent years for machine learning classification problems (Tuffery, 2011: 437). One of the pioneering studies, Berkson's work in 1944, used it for the analysis of biological experiments (Berkson, 1944: 357). Logistic regression works with multiple independent variables and includes a sigmoid function to calculate probabilities (Prajapati, 2013: 157). In logistic regression, the goal is to find the most suitable model to describe the relationship between two or more independent variables and dependent variables. There are two significant differences between linear regression and logistic regression. The first difference is that in linear regression analysis, the dependent variable to be predicted is continuous, while in logistic regression analysis, the dependent variable takes on a discrete (categorical) value. Second, in linear regression analysis, the value of the dependent variable is predicted, while in logistic regression, the probability of one of the possible values of the dependent variable occurring is estimated. With logistic regression, especially in data sets with multiple dimensions and, therefore, many features, the binary dependent variable is used to compare the selected class with all the other classes (Hackeling, 2014: 98).

2.5.6.Support Vector Machines Algorithm

Support Vector Machines, which were recommended in the study of Cortes and Vapnik (1995) and are one of the most popular supervised learning algorithms in machine learning, are especially used for solving classification and regression problems (Cortes and Vapnik, 1995; 273-299). Support Vector Machines are a kind of classifier learning method based on statistical learning theory. They are based on the principle of structural risk minimization, aiming to minimize the upper bound of the generalization error. As a result, they provide better generalization performance in solving machine learning problems compared to traditional neural networks. Support Vector Machines create a hyperplane in the pattern space to maximize the margin of separation between positive and negative examples by assigning points to two separate half-spaces in the pattern space. Support vectors are data points close to the hyperplane and affect the position and orientation of the hyperplane. Using these support vectors, the margin of the classifier is maximized (Hua and Zhang, 2006: 1037). This algorithm, especially in large and high-dimensional datasets, provides better classification performance compared to other machine learning algorithms (Marsland, 2015: 169)

2.5.7.Artificial Neural Networks Algorithm

Artificial Neural Networks, inspired by biological neural networks as seen in the study by McCulloch and Pitts (1943), were developed with the aim of self-generating new knowledge and discoveries through learning, similar to the learning abilities of the human brain, without any external assistance (McCulloch and Pitts, 1943: 115). Artificial learning algorithms based on the principle of learning from experiences aim to produce an output from multiple inputs. Artificial neural cells, or neurons, within the network have the ability to make predictions about similar examples they have never encountered before, either with supervised or unsupervised learning, based on the data. In addition to single-layer networks consisting of only input and output layers, networks composed of an input layer, one or more hidden layers, and an output layer are also used. The goal of the system is to find the output that is closest to the actual value by recognizing patterns in the data from the input layer (Tektaş and Karataş, 2004: 338). Artificial Neural Networks algorithms that can acquire learning abilities without requiring a mathematical model and a rule-based structure enable fast and cost-effective classification and predictions (Çınaroğlu and Avcı, 2020, 9). Artificial Neural Networks are often used when data is unlabeled or unstructured (Hurwitz and Kirsch, 2018: 31).

CHAPTER 3

Literature Review on the use of Machine Learning Algorithms in Predicting Stock Market Movements

In the literature, there are numerous studies that employ different evaluation criteria and methods for predicting the direction of stock market indices. In recent years, parallel to advancements in computer technologies, many machine learning algorithms have been developed and continue to be developed. Machine learning methods are becoming more widespread in their use due to their flexibility, generalization, risk assessment, and prediction capabilities, as they operate with fewer assumptions compared to traditional methods (Filiz et al., 2017: 232).

The literature reveals that various machine learning methods have been used, developed, and proposed for predicting the price movements of stock market indices. Artificial neural networks, support vector machines, decision trees, random forests, k-nearest neighbors, logistic regression, and naive Bayes prominently stand out as the most frequently used methods for predicting the direction of stock market indices or stock prices. These methods are commonly utilized for predicting the direction of stock market indices in various countries.

It is generally accepted in financial circles that the development level of a country is closely related to its financial development, and developed countries have advanced financial markets. Within the scope of this study, the aim is to compare the performance of machine learning classification algorithms in predicting the direction of stock market indices of both developed and developing countries. Accordingly, this section of the study is divided into four subsections in line with this aim. The first subsection summarizes studies where machine learning methods are employed to predict the direction of stock market indices in developed countries, while the second subsection does the same for developing countries. The third subsection focuses on studies that simultaneously predict the direction of stock market indices in both developed and developing countries using different machine learning methods. Finally, the last subsection discusses the findings obtained from the literature.

3.1.STUDIES ON THE DIRECTIONAL PREDICTION OF STOCK MARKET INDICES IN DEVELOPED COUNTRIES

In advanced capital markets, factors such as total assets, market liquidity, the number of individual and institutional investors, financial product diversity, and high market liquidity often lead to low transaction costs. Developed countries' capital markets are commonly referred to as developed markets in financial circles (Teker and Özer, 2012: 2). Within the scope of this study, the aim is to compare the performance of machine learning classification algorithms in predicting the directional movement of stock market indices in developed and developing countries. In this section, studies related to the prediction of stock market indices in developed countries using various machine learning methods are presented. These studies typically have a common feature: predicting the direction of the indices using machine learning methods and often presenting multiple models comparatively based on their method performance.

One of the pioneering studies in predicting stock market indices using machine learning methods is the study by Kimoto et al. (1990). This study aimed to predict the direction of the Japan Tokyo Stock Exchange TOPIX index. To achieve this goal, a total of 33 months of data from 1987 to 1989 were used, and predictions were made using multiple regression and artificial neural networks. According to the results of the study, it was found that artificial neural networks provided higher prediction accuracy

In the study by Cao and Tay (2001), the performance of a multi-layer perceptron trained by support vector machines and backpropagation algorithms in predicting the directional movement of the S&P 500 index was compared. The daily price index data from the period of 01.04.1993 to 31.12.1995 were used as the dataset. It was determined that the support vector machines model provided more successful predictions based on criteria such as normalized mean squared error, mean absolute error, directional symmetry, correct upward trend, and correct downward trend. Additionally, it was emphasized that applying the support vector machines method for predicting financial time series is advantageous.

In the study by Enke and Thawornwong (2005), an attempt was made to predict the directional movement of the S&P 500 index using artificial neural networks, generalized regression neural networks, probabilistic neural networks, and linear regression methods based on daily data from the period of 1976 to 1999. According to the results obtained from the study, it was observed that neural network classification model-based strategies were more successful compared to strategies using other models.

Huang et al. (2005) investigated the predictability of the weekly directional movement of the NIKKEI 225 index using the support vector machines method. A total of 676 observations from the period of 01.01.1990 to 31.12.2002 were used in the research. To evaluate prediction success, the performance of the support vector machines algorithm was compared to the performance of linear discriminant analysis, quadratic discriminant analysis, and feedforward neural networks. The analysis results showed that support vector machines outperformed other classification methods.

In the study by Shen et al. (2012), the aim was to predict the next-day directional movement of the NASDAQ, S&P500, and DJIA indices using the support vector machines method. Daily data from the period of 04.01.2000 to 25.10.2012 were used for the analysis. Due to holidays that vary from country to country and result in market closures, NASDAQ was primarily used for data alignment, and missing data from other sources were replaced using linear interpolation. The analysis results indicated that the support vector machines algorithm achieved a prediction success rate of 74.4% for NASDAQ, 76% for S&P500, and 77.6% for DJIA.

Adebiyi et al. (2014) used artificial neural networks and autoregressive integrated moving average methods to predict the direction of the NYSE index. The study utilized daily data from the period of 1988 to 2011 to predict the next day's direction of the NYSE index. The results showed that both models were successful in predicting the index's direction, with the artificial neural networks model outperforming.

Karlsson and Nordberg (2015) investigated the location dependence of artificial neural networks in predicting index movements. They predicted the direction of five different developed stock market indices: the US S&P500, Germany's DAX, Japan's Nikkei, Denmark's OMX30, and Sweden's OMX30, using the artificial neural networks method. Data from January 1, 2014, to June 1, 2014, were used as the training dataset, and data from June

2, 2014, to December 1, 2014, were used as the test dataset. The results of the study indicated that the artificial neural networks method was successful in predicting all the indices and its success did not significantly differ from one country to another.

In the study by Türkmen and Cemgil (2015), selected stocks traded on NASDAQ between 01.01.2010 and 30.06.2014 were used to predict the directional movement. They employed support vector machines, multi-layer perceptron, random forests, and stacked autoencoders, a deep learning algorithm, for prediction. The prediction results were evaluated based on precision and sensitivity measures, and support vector machines were found to be the most successful model according to both criteria.

Milosević (2016) aimed to predict the direction of 1,739 selected stocks from indices such as S&P1000, FTSE 100, and S&P Europe 350. For each stock, price data at the end of each quarter from the first quarter of 2012 to the end of 2015 were used as the dataset. They applied support vector machines, decision trees, random forests, logistic regression, naive Bayes, and Bayesian neural networks to make separate directional predictions for each stock. The results showed that random forests achieved the most successful prediction performance with an accuracy rate of 75.1%. Random forests were followed by decision trees with 66.8%, logistic regression with 64.3%, support vector machines with 63.9%, Bayesian neural networks with 62.5%, and naive Bayes with 54.5% accuracy.

In the study by Aktaş (2019), the directional movement of DJIA and NDX indices in the United States was predicted using the random forests and support vector machines methods. The analysis used daily data from the period of 2012 to 2018, with data from 2012 to 2017 as the training dataset and data from 2018 as the test dataset. The results of the analysis showed that the random forest algorithm performed better for the DJIA index, while the support vector machines algorithm yielded better results for the NDX index.

In the study by Subasi et al. (2021), predictions were made for the directional movement of four different stock indices: NASDAQ, NYSE, NIKKEI, and FTSE. The study used daily data from March 24, 2010, to March 24, 2020, and compared the prediction accuracy of various machine learning algorithms. According to the results of the study, the random forest algorithm achieved an accuracy of 93%, the decision tree algorithm had an accuracy of 79%, artificial neural networks achieved an accuracy of 75%, k-nearest neighbors had an accuracy of 71%, and support vector machines achieved an accuracy of 62%. The random forest algorithm was found to be the most successful in terms of prediction performance.

3.2.STUDIES ON THE DIRECTIONAL PREDICTION OF STOCK MARKET INDICES IN DEVELOPING COUNTRIES

Within the scope of this study, the performance of machine learning classification algorithms in predicting the direction of stock indices in both developed and developing countries is intended to be compared. In this context, national and international studies predicting stock indices of developing countries using different machine learning methods are examined in this section. The common feature of these studies is that they predict the direction of indices using machine learning methods and typically present a comparative analysis by using multiple methods.

In the pioneering study of Bao et al. (2005), one of the first studies in which machine learning methods were used to predict stock indices in developing countries, the Shanghai Stock Exchange SCI index's daily closing prices were predicted using the support vector machine method. A total of 140 data points from January 21, 2002, to September 27, 2002, were used, with 100 as training data and the remaining 40 as test data. The analysis showed that the support vector machine method achieved an accuracy performance of 60.72% in predicting the direction of the SCI index.

In the study by Kumar and Thenmozhi (2006), the aim was to predict the direction of India's National Stock Exchange's S&P CNX NIFTY market index. The study used a total of 1,360 daily data points from January 1, 2000, to May 31, 2005. Machine learning methods, including artificial neural networks, random forests, and support vector machines, were used to predict the direction of the index. The results of the study indicate that support vector machines are the most successful method in predicting the direction of the stock market. After support vector machines, the random forest method was found to be the second most successful classification method.

In the study by Ou and Wang (2009), data from the period of January 3, 2000, to December 29, 2006, was used to predict the movement of the Hang Seng Index of the Hong Kong Stock Exchange. Various classification algorithms, including k-nearest neighbors, naive Bayes, logistic regression, decision trees, artificial neural networks, and support vector machines, were employed in the study. The results showed that all algorithms achieved an accuracy rate of over 80%, with support vector machines being the most successful algorithm.

Cao, Parry, and Leggio (2011) aimed to predict the price movements of Class A shares traded on the Shanghai Stock Exchange (SHSE) between 1999 and 2008. They utilized single-variable linear models, multivariable linear models, single-variable artificial neural networks, and multivariable artificial neural networks. The multivariable artificial neural networks were found to provide the highest accuracy in predicting the direction of the SHSE index.

In the study by Kara et al. (2011), the goal was to predict the movement of the Istanbul Stock Exchange National 100 Index using data from January 2, 1997, to December 31, 2007. They compared the performance of artificial neural networks and support vector machines in making predictions. The analysis results indicated that the artificial neural networks algorithm had an average performance of 75.74%, while the support vector machines algorithm had an average performance of 71.52%. It was concluded that the artificial neural networks algorithm had relatively higher prediction accuracy compared to support vector machines.

In the study by Subha and Nambi (2012), the aim was to predict the daily movement direction of India's BSE-SENSEX and NSE-NIFTY stock market indices. The study used data from January 2006 to May 2011, and the k-nearest neighbors and logistic regression algorithms were employed to predict the direction of the indices. The results showed that the k-nearest neighbors algorithm achieved a prediction accuracy of 79.65%, while the logistic regression algorithm had a lower accuracy of 54.11%. Therefore, the k-nearest neighbors algorithm exhibited a more successful classification performance compared to the logistic regression algorithm.

Dunis et al. (2013) focused on predicting the weekly changes in the Madrid stock market's IBEX35 stock index using support vector machines and artificial neural networks. The data covered the period from October 18, 1990, to October 29, 2010. Technical indicators, such as the Relative Strength Index (RSI) and Moving Average Convergence Divergence (MACD), were utilized. The analysis revealed that better results were achieved when shorter training periods were used in comparison to longer ones, and analyzing recent data provided better outcomes. Consequently, it was determined that for the chosen period and market, support vector machines outperformed artificial neural networks.

In Wang's study (2013), the direction of South Korea's KOSPI 200 index and Hong Kong's HSI index was predicted using a hybrid method of artificial neural networks and decision trees. The study, based on daily data from 2002 to 2011, resulted in a prediction accuracy rate of 77%.

Imandoust and Bolandraftar (2014) focused on predicting the movement direction of the Tehran Stock Exchange (TSE) index using daily data from 2007 to 2012. Three models, decision trees, random forests, and naive Bayes, were used for prediction, and their performances were compared. The analysis indicated that the decision trees model achieved a prediction accuracy of 80.08%, the random forest model had an accuracy of 78.81%, and the naive Bayes algorithm achieved an accuracy of 73.84%. The decision trees algorithm was highlighted as the most successful in terms of prediction performance.

In Karagül's study (2014), the direction of stock prices of 42 companies operating in the food, textile, and cement sectors within the BIST 100 index in the 2006-2011 period was predicted using the support vector machines method. Single-fold cross-validation was employed to test the learning algorithm's performance and classification. The study's results showed that the method predicted the next day's direction of the index with an accuracy of 97.6%.

In the study by Ballings et al. (2015), data from 5,767 publicly traded European companies were used to predict stock price directions. Random forests, artificial neural networks, logistic regression, support vector machines, and k-nearest neighbors were employed as machine learning methods for stock price direction prediction. The results indicated that the random forests algorithm provided the most successful outcomes, followed by support vector machines, artificial neural networks, k-nearest neighbors, and logistic regression algorithms.

Nou et al. (2015) aimed to predict returns and volatilities in the OMX Baltic Benchmark price index of the Baltic Stock Exchange. Daily closing data from September 4, 2001, to March 1, 2021, were used, and predictions for the index's direction were made using random forests, support vector machines, and k-nearest neighbors algorithms. The results showed that support vector machines and k-nearest neighbors algorithms demonstrated more successful performance.

Patel et al. (2015) focused on predicting the movement direction of the CNX Nifty and BSE price indices in Indian stock markets. They used data from 2003 to 2012 and employed naive Bayes, artificial neural networks, support vector machines, and random forest models. The results indicated that the most successful model was naive Bayes with a success rate of 90.19%. Following that, random forest achieved an accuracy rate of 89.98%, support vector machines 89.33%, and artificial neural networks 86.69%.

Zahedi and Rounaghi (2015) used 20 financial ratios calculated for the Tehran Stock Exchange from 2006 to 2012 to predict stock price directions. They employed artificial neural networks and principal component analysis methods. The analysis results suggested that the principal component analysis algorithm was more successful than artificial neural networks for predicting stock prices in the Tehran Stock Exchange.

In Udomsak's study (2015), the performance of support vector machines and naive Bayes classifiers was compared for predicting the SET100 index of the Thailand Stock Exchange. Data from January 1, 2010, to February 1, 2010, were used in the analysis. The mean squared error and the ratio of correctly classified examples were used to compare the two algorithms. The results showed that naive Bayes achieved a 66% accuracy rate, while support vector machines had a 56% prediction accuracy. Therefore, naive Bayes outperformed support vector machines in predicting the Thailand Stock Exchange, although both algorithms were considered not sufficiently advanced to accurately model the stock market.

In the study by Pehlivanlı et al. (2016), they aimed to predict the direction of the BIST 100 index and investigated the machine learning model that best predicts the index. Daily data from the BIST 100 index for the period from April 1, 2007, to January 31, 2013, was used, and support vector machines, multi-layer perceptrons, and random forest models were compared. The model with the highest accuracy rate in predicting the index direction was determined to be the support vector machines model.

Filiz et al. (2017) focused on predicting the direction of the BIST-50 index using daily data covering 2,591 trading days from 2006 to 2016. They employed machine learning methods including k-nearest neighbors, naive Bayes, and artificial neural networks. The results of the analysis indicated that naive Bayes achieved a 92.28% accuracy rate, artificial neural networks achieved 91.66%, and k-nearest neighbors achieved 89.58% in predicting the direction of the BIST-50 index.

Tekin and Çanakoğlu (2018) used data from 25 companies traded on Borsa Istanbul from 2010 to 2017. They utilized 1,928 daily data points and 12 technical indicators to predict price movement directions. Naive Bayes, multi-layer perceptrons, random forest, and support vector machines were used as the prediction methods. The findings revealed that the naive Bayes algorithm had the highest accuracy rate at 53.17%.

Aktaş (2019) conducted a study where the movement directions of the DJIA and NDX indices in the U.S. were predicted using daily data and technical indicators from the period 2012-2018. The model used data from 2012 to 2017 as training data, 2018 as test data, and 2019 for validation. The results showed that the most successful method for the DJIA index was random forests, while for the NDX index, the support vector machines method yielded better results.

In Kara's study (2019), the direction of the BIST 100 index was predicted using data from the period 1995-2018. They employed various machine learning methods, including artificial neural networks, support vector machines, decision trees, naive Bayes, k-nearest neighbors, logistic regression, and linear discriminant analysis. The classification methods achieved accuracy rates of 83.83%, 78.43%, 65.04%, 61.74%, 55.48%, 76.70%, and 76.87%, respectively. According to the results, the best classification method that could be used to predict the direction of the BIST 100 index was artificial neural networks.

In the study by Papuccu (2019), artificial neural networks, support vector machines, and naive Bayes algorithms were used to predict the movement direction of the BIST 100 index. The dataset included daily closing values from 2009 to 2018. Technical indicators were used as input for the models predicting stock index movements. The analysis results showed that all three models could be used to capture stock index movements, with artificial neural networks having the most successful performance.

Ismail et al. (2020) aimed to predict the direction of three indices: the Kuala Lumpur Composite Index, Kuala Lumpur Industrial Index, and Kuala Lumpur Technology Index. Logistic regression, artificial neural networks, support vector machines, and random forest methods were employed to predict the next day's movement of these indices. The results indicated that the support vector machine was the most successful method for this purpose. The study suggested that using a combination of machine learning methods might be a more accurate approach to predict stock price movements.

In Aksoy's study (2021), data from five manufacturing companies in the BIST 30 and BIST Corporate Governance Index were used to predict the direction of stock prices for the period from 2010/3 to 2020/3. The dataset included quarterly financial statement data for these companies and an average of five macroeconomic variables. Prediction methods included k-nearest neighbors, classification and regression trees, and artificial neural networks. The analysis results showed that artificial neural networks achieved a general classification accuracy of 98.05%, classification and regression trees achieved 96.10%, and k-nearest neighbors achieved 92.20% in predicting stock price direction.

Chaengkham and Wianwiwat (2021) aimed to predict the movement of leading stock indices in four Southeast Asian countries: Indonesia, Malaysia, Singapore, and Thailand. They used the support vector machines method with data spanning from January 2002 to December 2019. The results showed that support vector machines achieved prediction accuracy ranging from 58.14% to 65.12% for these countries' stock markets. The highest prediction accuracy was obtained for the Indonesian Stock Exchange (IDX) at 65.12%, followed by the Stock Exchange of Thailand (SET) at 60.47%, while the Singapore Exchange (SGX) and Bursa Malaysia (BM) achieved 58.14% accuracy in their predictions.

In the study by Filiz et al. (2021), the goal was to predict the direction of change of the BIST-100 index using logistic regression, artificial neural networks, naive Bayes, support vector machines, and decision trees. Input variables included major world indices and macroeconomic indicators such as gold and the dollar. The dataset covered the period from January 1, 2006, to December 1, 2020. The analysis results showed that the support vector machines algorithm achieved a classification accuracy of 71.9%, followed by logistic regression at 70.6%, naive Bayes at 70.4%, decision trees at 70.3%, and artificial neural networks at 70.2%. Therefore, it was determined that support vector machines were the most successful method for predicting the BIST 100 index.

Kemalbay and Alkış (2021) aimed to predict the daily upward or downward movement direction of the BIST 100 index using machine learning algorithms. For this purpose, logistic regression and k-nearest neighbors algorithms were applied using independent variables that had a statistically significant impact on the direction of the BIST 100 index. The analysis was conducted using daily return data from January 11, 2010, to October 13, 2016, comprising 1700 trading days. The results of the analysis showed that the logistic regression algorithm achieved an accuracy rate of 81%, outperforming the k-nearest neighbors algorithm, which had an accuracy rate of 78%.

3.3.STUDIES ON THE DIRECTIONAL PREDICTION OF STOCK MARKET INDICES IN BOTH DEVELOPED AND DEVELOPING COUNTRIES

In the literature, there are very few studies that jointly examine stock market indices of both developed and developing countries using machine learning methods. These studies commonly involve predicting the direction of indices using different machine learning techniques and comparing the performance of these methods.

One of the pioneering studies in the literature that used machine learning methods to predict stock market indices of both developed and developing countries is the study by Phua et al. (2003). In this study, the direction of Germany's DAX index and the US DJIA, FTSE-100, NASDAQ indices, as well as China's HSI index, was predicted using artificial neural networks. The price data of all the stocks that served as components of these indices during the period from January 4, 1994, to September 30, 2002, were used as the dataset. The results showed that the artificial neural networks algorithm achieved an average accuracy rate of over 60% in predicting the direction of the indices for all five markets. The HSI index had the highest accuracy, with a 74% success rate, while the DAX, DJIA, FTSE, and NASDAQ indices had accuracy rates of 68%, 70%, 73%, and 73%, respectively.

In the study by Liao and Wang (2010), the direction of the US DJIA, NASDAQ, and S&P 500 indices, as well as China's Shanghai and Shenzhen Stock Exchange SAI, SBI, and HIS indices, was predicted using artificial neural networks. The study used daily data from December 19, 1990, to June 7, 2008, covering an 18-year period. The study evaluated the method's success using various parameters, and it concluded that artificial neural networks were a successful method for predicting these indices.

In the study by Dash and Dash (2016), the direction of the S&P 500 index from the developed country category, which represents the United States, and the BSE SENSEX index from the developing country category, representing India, was predicted using artificial neural networks, support vector machines, naive bayes, k-nearest neighbors, and decision trees. The study utilized a dataset from January 4, 2010, to December 31, 2014. Additionally, a classification model using artificial neural networks was suggested in the study, and the model outputs were converted into simple trading signals of buy, hold, and sell using appropriate rules. The results of the study indicated that the proposed model outperformed classifiers such as support vector machines, naive bayes, k-nearest neighbors, and decision trees in terms of generating higher profit percentages.

In the study by Özer et al. (2018), the movement direction of stock market indices from developed countries, including the US NASDAQ, UK FTSE 100, Germany DAX, France CAC 40, as well as indices from developing countries such as China's SHANGHAI, India's NIFTY 50, Mexico's IPC, and Turkey's BIST 100, was attempted to be predicted. The study used weekly data from the period of 2012 to 2016 and applied artificial neural networks and fuzzy logic models for prediction. The analysis results compared the predictive accuracy of both methods and showed that both methods were successful in predicting the indices, providing similar prediction results.

In the study by Ali et al. (2021), the direction movements of Japan's Tokyo Stock Exchange Nikkei 225 index from developed countries and the indices of China's SZSE Composite, Pakistan's KSE-100, and South Korea's KOSPI Composite from developing countries were predicted using artificial neural networks and support vector machines. The study covered a 10-year period from January 1, 2011, to September 27, 2020, using daily data derived from stock trading, including open, close, high, low prices, and various technical indicators. The success performance of the models was evaluated based on accuracy, precision, and sensitivity criteria, and the results indicated that artificial neural networks outperformed support vector machines in predicting the daily directional movements of the KSE-100, KOSPI, Nikkei 225, and SZSE indices.

In the study by Domingo et al. (2021), the goal was to predict market trends in both developed and developing countries. The markets analyzed included the advanced markets represented by the US S&P 500 and the UK FTSE 100 indices, as well as the developing markets represented by South Africa's ALSI and Brazil's BOVESPA indices. Market direction was predicted using support vector machines, k-nearest neighbors, decision trees, and random forest methods. The study employed a dataset from July 3, 1995, to August 28, 2018, spanning a total of 6042 daily data points. The data was distributed into training and testing sets with 83% and 17% of the total data, respectively. The models that best predicted market direction were random forests and decision trees for the S&P 500, FTSE 100, and ALSI indices, while for the BOVESPA index, the random forest model performed the best. Consequently, it was determined that random forests were the best models for predicting market direction in both developed and developing markets.

3.4.LITERATURE FINDINGS

In the literature, there are numerous studies that focus on predicting the directional movements of stock market indices in both developed and developing countries using machine learning methods. These studies often employ various machine learning techniques and emphasize method performance. Among the commonly used methods for predicting stock market indices or stock price movements, artificial neural networks, support vector machines, decision trees, random forests, k-nearest neighbors, logistic regression, and naive Bayes stand out. Many of the studies aiming to predict the direction of stock market indices use two or more machine learning methods to make predictions and compare these methods based on their predictive accuracy. The determination of which method is successful generally follows a similar process. In this process, historical data related to the index are divided into training and test datasets. Subsequently, a model is trained using the training dataset. Then, an attempt is made to predict values in the test dataset. Finally, the predicted values are compared with the actual values to calculate the model's predictive accuracy. As a result, the method with higher predictive accuracy is considered more successful. The studies examined within the scope of this research and the most successful methods found in these studies are presented in Table 1.

Table 1: Applied Studies on Index Direction Prediction and the Most Successful Methods

Machine Learning Methods	Developed Cauntries	Developing Cauntries	Developed and Developing Cauntries	Number of Studies
Support Vector Machines	Cao ve Tay (2001), Huang vd. (2005), Shen vd. (2012), Türkmen ve Cemgil (2015),	Bao vd. (2005) Kumar ve Thenmozhi (2006), Ou ve Wang (2009), Kara vd. (2011), Dunis vd.(2013), Karagül (2014), Pehlivanlı vd. (2016), Ismail vd.(2020), Chaengkham ve Wianwiwat (2021)		13
Artificial Neural Networks	Kimoto vd. (1990), Enke ve Thawornwong (2005) Adebiyi (2014), Karlsson ve Nordberg (2015)	Cao, Parry ve Leggio (2011), Kara (2019), Papuccu(2019), Aksoy (2021)	Phua vd (2003), Liao ve Wang (2010), Dash ve Dash (2016), Özer vd.(2018), Ali vd. (2021)	13
Random Forest	Miloseviç (2016), Aktaş (2019), Subasi vd. (2021),	Ballings vd. (2015), Aktaş (2019)	Dimingo vd. (2021)	6
Naive Bayes		Patel vd. (2015), Udomsak (2015), Filiz vd. (2017), Tekin ve Çanakoğlu(2018)		4

Machine Learning Methods	Developed Cauntries	Developing Cauntries	Developed and Developing Cauntries	Number of Studies
K-Nearest Neighbors		Subha ve Nambi (2012), Nou vd. (2015)		2
Logistic Regression		Filiz vd. (2021), Kemalbay ve Alkış (2021)		2
Decision Trees		Imandoust ve Bolandraftar (2014)		1
Number of Studies	11	24	6	41

When Table 1 is examined in the context of both developed and developing countries, it is observed that more studies have been conducted on the stock market indices of developing countries. Although the number of countries in the developed category is relatively less, their stock market indices have been examined in a considerable number of studies. Furthermore, there is limited literature that explores stock index prediction for both developed and developing countries together. In this context, this study is believed to make a significant contribution to the literature.

When Table 1 is examined in the context of the most successful methods, it can be observed that the most successful methods are Support Vector Machines and Artificial Neural Networks. Following these methods, in order, come Random Forest, Naive Bayes, K-Nearest Neighbors, Logistic Regression, and Decision Trees.

CHAPTER 4

Prediction of the Movements of Developed and Developing Country Stock Market Indices Using Machine Learning Methods

In this section, the purpose and significance of the research were initially explained. Following that, the dataset and methods used in the research were introduced. Then, the findings obtained by predicting the direction of stock market indices using machine learning methods were presented. In the final section, the findings from the analysis were discussed. Within the scope of this study, the daily dataset and technical indicators for developed and developing countries' stock market indices for the period between January 1, 2012, and December 31, 2021, were used to predict the directional movements using machine learning methods, including decision trees, random forests, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks.

4.1.PURPOSE AND SIGNIFICANCE OF THE RESEARCH

The main objective of this study is to utilize machine learning classification methods to predict the movements of stock market indices in both developed and developing countries. Furthermore, it aims to compare the performance of these methods and identify the most effective prediction approach. To achieve this goal, daily data spanning from January 1, 2012, to December 31, 2021, and various technical indicators were employed to forecast the trajectories of stock market indices. The machine learning algorithms employed in this study include decision trees, random forests, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks.

Predicting the movements of stock prices is a critical aspect of decision-making in investment. However, stock markets, characterized by their high volatility, dynamism, and complexity, are influenced by a multitude of factors, including macroeconomic variables, global events, and human behavior. This continuous exposure to various influences has made the predictability of stock returns a contentious topic in the academic literature. The Efficient Market Hypothesis (EMH), proposed by Fama in 1970, posited that stock prices fully incorporate all available information (Fama, 1970: 386). However, the validity of the Efficient Market Hypothesis has faced ongoing scrutiny and debate in empirical research since its inception. While the belief in market efficiency was supported in the 1960s and 1970s, the subsequent decades have witnessed numerous studies arguing that asset prices are not independent of past information, thereby challenging the notion of market efficiency (Nevasalmi, 2020: 86).

In stock markets, an effective prediction model or technique must be capable of providing accurate forecasts regarding market direction to minimize investment risks and uncertainties (Rodriguez and Rodriguez, 2015: 2). Two distinct analysis methods are commonly employed for predicting market prices in the realm of stock trading: fundamental analysis and technical analysis. Fundamental analysis involves forecasting stock prices based on the financial assessment of companies or sectors. The objective is to identify the gap between the real value of an asset and its market price, enabling the purchase of the asset at a low value and its subsequent sale at a higher price. On the other hand, technical analysis assumes that stock prices are determined by market forces and relies on the belief that history tends to repeat itself. It employs historical securities data to predict future prices (Spahija and Xhaferi, 2019: 8-9).

Recent years have witnessed a growing recognition that traditional methods are inadequate for analyzing and predicting the highly volatile movements of stock prices. In this context, machine learning techniques based on artificial intelligence have emerged as more effective alternatives compared to traditional approaches (Rothman, 2021: 6). Machine learning methods, capable of processing substantial volumes of data to unveil intricate relationships within the datasets, empower the generation of quicker and more accurate predictions compared to traditional methodologies (Nou et al., 2021: 7).

The predictability of stock index movements can bring various benefits to stock markets, investors, market regulators, and business managers. Predictability in the market will primarily lead to increased investment flows into the market. On the other hand, it will provide significant benefits for investors such as protecting their savings, avoiding transaction costs, seizing investment opportunities, and anticipating risks that may arise in extraordinary situations. Furthermore, it will be beneficial for market regulators in terms of taking appropriate decisions and corrective measures. Additionally, it can offer business managers the opportunity to act correctly in maximizing corporate values (Mallikarjuna and Rao, 2019: 1-2). In this context, predicting stock index movements with new and different methods is considered an important subject.

4.2.DATA SET AND METHODOLOGY OF THE RESEARCH

This research focuses on G-7 economies as developed countries and E-7 economies as developing countries. To select stock market indices in each country, the primary stock index of the stock exchange with the highest market capitalization in each country was given priority. Consequently, the stock market indices for developed countries were identified as NYSE 100 (USA), NIKKEI 225 (Japan), FTSE 100 (UK), CAC 40 (France), DAX 30 (Germany), FTSE MIB (Italy), and TSX (Canada). As for the stock market indices representing emerging markets, they were determined as SSE (China), BOVESPA (Brazil), RTS (Russia), NIFTY 50 (India), IDX (Indonesia), IPC (Mexico), and BIST 100 (Turkey).

The main goal of this study is to forecast the directional movements of the selected stock market indices by employing machine learning classification techniques, including decision trees, random forests, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks. When it comes to making predictions using machine learning, it is crucial to use an adequately sized dataset for modeling. Nevertheless, the exact definition of what constitutes an adequate dataset size remains uncertain and context-dependent. For instance, decision trees, support vector machines, and naive Bayes methods are advised for use with small to medium-sized datasets, while k-nearest neighbors, random forests, and artificial neural networks are recommended for dealing with large-scale datasets. Therefore, the selection of the research period was initially determined by reviewing related literature. It was observed that similar studies in the literature employed datasets of varying sizes. Subsequently, based on studies by Huang et al. (2005), Ali et al. (2021), and Subaşı et al. (2021), the appropriate research period was chosen as a ten-year period from January 1, 2012, to December 31, 2021.

The historical daily data for the selected indices, which includes the opening price, closing price, highest price, and lowest price, throughout the defined research period, was sourced from the https://tr.investing.com website. Following this, an investigation into technical indicators that have been commonly used as input variables in similar studies, with the potential to enhance the prediction performance of the methods, was conducted. These technical indicators, frequently employed in analogous studies concentrating on predicting the direction of stock market indices, such as Kumar and Thenmozhi (2006), Kara et al. (2011), Patel et al. (2015), Qiu and Song (2016), Kara and Ecer (2018), and Papuccu (2019), were computed and integrated into the analysis as input variables. The technical indicators utilized as input variables and their respective calculation methods are provided in Table 2.

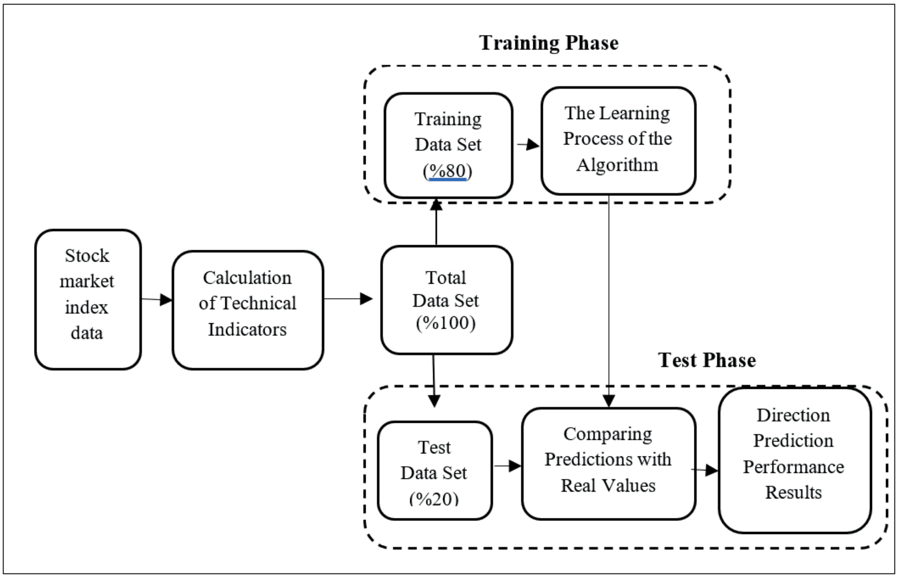
Table 2: Technical Indicators and Calculation Methods

Technical Indicators	**Calculation**
Simple Moving Average (MA)	$\frac{C_t + C_{t-1} + \cdots + C_{t-30}}{n}$
Weighted Moving Average (WMA)	$\frac{((n) * C_t + (n-1) * C_{t-1} + \cdots + C_{t-14}}{(n + (n-1) + \cdots + 1)}$
Exponential Moving Average (EMA)	$EMA(k)_t = EMA(k)_{t-1} + a * (C_t - EMA(k)_{t-1}$
Momentum (MOM)	$C_t - C_{t-n}$
Stochastic %K	$\frac{C_t - LL_{t-n}}{HH_{t-n} - LL_{t-n}} * 100$
Stochastic %D	$\frac{\sum_{i=0}^{n-1} K_{t-i}\%}{n}$
Relative Strength Index (RSI)	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n)/(\sum_{i=0}^{n-1} Dw_{t-i}/n)}$
Moving Average Convergence Divergence (MACD)	$MACD(n)_{t-i} + \frac{2}{n+1} * DIFF_t - MACD(n)_{t-1}$
Larry Williams %R (LW)	$\frac{H_n - C_t}{H_n - L_n} * 100$
Commodity Channel Index (CCI)	$\frac{M_t - SM_t}{0{,}015 D_t}$
* C_t:Closing Price L_t:Lowest Price H_t:Highest Price $DIFF_t = EMA(12)_t - EMA(26)_t$ a:Correction Factor LL_t:Minimum of minimums for last t days	HH_t:The maximum of the maximums for the last t days $M_t = (H_t + L_t + C_t)/3$ $SM_t = \sum_{i=0}^{n} M_{t-i+1}/n)$ $D_t = \sum_{i=1}^{n} \| M_{t-i+1} - SM_t \|/n$ UP_t:upward price at time t DV_t:downward price at time t

Reference: (Kumar and Thenmozhi, 2006: 15).

The following formulas for Simple Moving Average, Weighted Moving Average, Exponential Moving Average, Momentum, Stochastic K, Stochastic D, Relative Strength Index, Moving Averages Convergence Divergence, Larry William's R, and Commodity Channel Index were used as input variables in measuring the performance of all methods employed in the study. The application process of machine learning methods used for predicting stock index directions is shown in Figure 3.

Figure 2: Machine Learning Application Stages



Machine learning algorithms were used for stock index direction prediction, utilizing technical indicators calculated from closing prices, highest prices, and lowest prices in addition to stock index data. The next day's movement direction, classified as either "fall" or "rise," based on closing prices of stock indices, was used as the output data. Before proceeding to the training and testing stages, the dataset underwent a thorough review to ensure there were no erroneous or missing data. In the machine learning application process, the dataset is divided into two parts: the training dataset and the testing dataset. The machine learns from the training dataset, and the learned model's predictions are then compared with the testing dataset. The results obtained are evaluated (Compesato, 2020: 42).

In similar studies found in the literature, it was observed that the dataset was divided into different percentage ratios such as 90:10, 80:20, 83:17, 70:30, and 60:40. Although there is no general consensus on the specific ratio for data splitting, it is suggested that an 80:20 ratio is used for classification tasks involving large datasets (Racz, Bajusz, and Heberger, 2021: 11). Afterward, the dataset was divided into these ratios to determine the ratio that produced the highest performance through trial and error. It was noted that the prediction performance of the algorithms improved when the size of the training dataset was increased from 50% to 80%. However, it was observed that when the training size was increased from 80% to 90%, the prediction performance of most algorithms decreased. Therefore, taking into account both the trial and error method employed and the recommendations in the literature, the dataset was divided into 80% for training and 20% for testing.

After the training stage, the predictions made by the algorithms were compared with the testing dataset to evaluate the prediction performance of the methods. Seven different machine learning algorithms, including decision trees, random forests, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks, were used to predict the "fall" or "rise" movement directions of developed and developing countries stock indices.

The decision tree algorithm is used to predict the direction of stock market indices. This algorithm creates a tree structure to classify the upward or downward trends of the index based on past data. Subsequently, the values of decision boundaries in this tree structure are learned from the training data, and experimental predictions are made. When decision trees are applied to continuous data, a logical test is conducted at each internal node, where X_i represents a feature in the data space, and C is a threshold value for X_i within the observed range: $X_i > C$. Various criteria, such as maximizing differences or minimizing similarity in child nodes, are used to determine the threshold value C. Suppose a dataset represents class values T, with n classes such as $C_1, C_2 \dots, C_n$. In this case, the probability for each class is calculated as $P_i = C_i/T$, and entropies are computed for the classes. When the T class values are divided using feature B in the dataset, the gain ratio resulting from this split is calculated. The gain ratio is used as a criterion to repeatedly split the training dataset in each node of the tree so as to maximize the gain ratio. Each leaf node is processed to contain only observation values belonging to a single class. To calculate the gain ratio, the operations described in equations (1), (2), (3), and (4) are sequentially applied (Quinlan, 1993: 302; Han, Kamber, and Pei, 2012: 332-337).

$$\text{Entropy}(T) = -\sum_{i=1}^{n}(p_i \log_2(P_i) \quad (1)$$

$$\mathrm{Gain}(B,T) = \mathrm{Entropy}(T) - \sum_{i=1}^{n} \frac{|T_i|}{|T|} \left(\mathrm{Entropy}(T_i)\right) \tag{2}$$

$$\mathrm{Split\ Information}(B) = -\sum_{i=1}^{k} \frac{|T_i|}{|T|} \mathrm{Log}_2\left(\frac{|T_i|}{|T|}\right) \tag{3}$$

$$\mathrm{Gain\ Ratio} = \frac{\mathrm{Gain}(B,T)}{\mathrm{Split\ Information}(B)} \tag{4}$$

The random forest algorithm combines multiple decision trees to predict the directional movements of stock market indices. In other words, it creates multiple decision trees by using subsets of the dataset and combines the predictions of these trees to obtain the final result. In the random forest algorithm, the number of decision trees (n) to be created based on the features of the dataset is first determined. Then, at each node of the created decision trees, a random selection of m variables is made, and calculations are performed using the Gini index to determine which branch is the best. Subsequently, the best branch is split into two sub-branches. This process continues until a single class remains at each node, meaning the Gini index is reduced to zero. Finally, the class with the most votes among the predictions made by the individual decision trees is selected as the predicted class. In the calculation of the Gini index, as shown in equation (5), T represents the entire dataset, p_j represents the square of the division of the number of elements in the dataset that are smaller and larger than the data point in consideration, and c represents the selected data point (Breiman et al., 1984: 121; Breiman, 2001: 1).

$$I_G = \sum_{j=1}^{e} p_j^2 \tag{5}$$

The k-nearest neighbors (k-NN) algorithm is one of the classification methods used to predict the upward and downward movements of stock market indices. In the k-nearest neighbors algorithm, example-based classification is performed based on the distance measure of examples in the known class category in the training dataset (North, 2016: 94). Each example in n data instances defined by p attributes, $x_1^t = (x_1, \dots, x_{1p}), \dots, x_1^t = (x_1, \dots, x_{np})$, represents a point in a p-dimensional vector space. The distance between sample points i and j is denoted as $d(x_i, x_j)$, and when all attributes are numeric, the algorithm calculates it using the Euclidean distance, as shown in equation (6) (Mitchell, 1997: 232).

$$d(x_i * x_j) = \sqrt{\sum_{k=1}^{p} (x_{ik} - y_{jk})^2} \tag{6}$$

To determine which class a new instance belongs to, the algorithm calculates the distances from this calculated point to all points in the training dataset, and the calculated distances are sorted from smallest to largest. Taking into account the predefined parameter, the number of nearest neighbors (k), k data points with the closest distances are selected. To determine the class of the new observation, majority voting or weighted voting methods are commonly used. The votes of the k nearest neighbors in the training dataset, $x_m; m = 1, \dots, k$, are inversely proportional to the distance to the new sample point x_z. This relationship is expressed in equation (7) (Cover and Hart, 1967: 21-22).

$$\text{vote } (x_m) = \begin{cases} \infty, & d(x_m, x_z) = 0 \\ \frac{1}{d(x_m, x_z)}, & \text{other} \end{cases} \tag{7}$$

One of the classification methods used to predict the direction of stock market indices is Naive Bayes. The Naive Bayes algorithm learns which class examples belong to by using the training dataset, and it makes class predictions for observations in the test data. In this process, Bayes' theorem is used for probability estimation. Bayes' theorem, expressed as P(C/X), gives the probability of event C occurring when event X has happened. This theorem is provided in equation (8) (Minsky, 1961: 406; Jurafsky and Martin, 2020: 57).

$$P\left(\frac{C}{X}\right) = \frac{P(C) * \left(\frac{C}{X}\right)}{P(X)} \tag{8}$$

Logistic regression is commonly used for predicting the direction of stock market indices. In this method, a logistic regression model is constructed based on historical data to estimate the probability of the index going up or down. Logistic regression is a classification method used when the dependent variable belongs to two categorical classes. When predicting the direction of the index, the dependent variable is typically divided into two categories, such as "up" or "down." In the algorithm, assuming that Y is the dependent variable, and $X^T = (X_1, \dots . X_p)$ is a vector of independent variables: the two-category dependent variable is usually encoded as Y = 0 and Y = 1. When the value of X vector, $X^T = (X_1, \ldots . X_p)$ is known, the probability of Y taking the value of 1, $\pi(X) = P(Y = 1|X = x)$ is represented. The multiple logistic regression model is provided in equation (9) (Hosmer et al., 2013: 50; Hilbe, 2015: 5).

$$\pi(X) = P(Y = 1|X = x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\ldots+\beta_p x_p)}} \tag{9}$$

When the logit transformation is applied to the multiple logistic regression model, which is represented as $lojit[\pi(x)] = g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right.$, it transforms into the linear model as expressed in equation (10).

$$g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right. = \beta_0 + \beta_1 x_1 + \ldots . + \beta_p x_p \tag{10}$$

The parameter $\beta^T = \beta_0 + \beta_1 + \ldots . + \beta_p$ can be estimated using the maximum likelihood method based on the likelihood function $L(\beta) = \prod_{i=1}^{n} (x_i)^{y_i}\ (1 - \pi(x_i)^{1-y_i}$ where g(x) can be continuous and take values between-∞ and +∞, depending on the range of x's values (Hosmer et al., 2013: 50; Hilbe, 2015:5).

Support Vector Machines, commonly used algorithms in predicting the direction of stock market indices, involve creating a differentiating hyperplane based on historical data. This hyperplane is utilized to predict whether the index will rise or fall. The generated hyperplane ensures maximum margin separation to effectively classify data points. When considering a dataset consisting of l elements for training, where $\{x_i, y_i\}, i = 1,2, \ldots, l$, with w as the weight vector, b as a constant term, $y_i \in \{-1,1\}$ as label values, and $x_i \in \Re^d$ as the feature vector, the support vectors, which are training points, are represented by equations (11), (12), and (13) (Cortes and Vapnik, 1995: 273; Marsland, 2009: 119).

$$y_i = 1 \ ; \ w * x_i + b \geq 1 \tag{11}$$

$$y_i = -1 \ ; \ w * x_i + b \geq -1 \tag{12}$$

$$y_i = w * x_i + b \geq 1 \qquad \forall i \tag{13}$$

Artificial neural networks are commonly used methods in binary classification problems such as predicting the direction of stock market indices. When designing an artificial neural network model, a typical three-layer structure is used, consisting of an input layer, one or more hidden layers, and an output layer. The input layer includes nodes corresponding to each feature in the dataset. Hidden layers are used to capture complex patterns in the data. The output layer typically consists of a single node in the case of binary classification, providing class predictions. The network's training is a process where the model learns to improve its performance on the dataset. During this process, the network learns from a portion of the data (the training set). The learning process aims to minimize the error function between the predicted classes and the actual classes. Artificial neural networks often use the backpropagation algorithm for error backpropagation. In the artificial neural networks algorithm, a fitness function is defined by summing the squares of error signals for each neuron in the output layer. In the k-th iteration of training, if $y_i$represents the output value of the i-th neuron in the output layer of the artificial neural network and d_i is the desired value to be output by that neuron, the error signal for the i-th neuron is calculated as $e_i = d_i - y_i(k)$. The fitness function, denoted by E, representing the sum of squared error signals for each neuron in the output layer of the artificial neural network, is expressed in equation (14) (Rojas, 1996: 157).

$$E = \frac{1}{2}\ {}_{i}e_i^2\ k = \frac{1}{2}\ {}_{i}(d_i - y_i\ k)^2 \tag{14}$$

The backpropagation algorithm aims to minimize the fitness function. Since the fitness function depends on the weight values of the artificial neural network, the algorithm computes the amount of change in the network's weights optimally using the gradient descent method. Denoted by , the learning rate in the gradient descent method is expressed in equation (15) (Haykin, 2009: 605; Li and Huang, 2021: 3).

$$\Delta w_{ij} = -\eta \frac{\partial E(w)}{\partial w_{ij}} \tag{15}$$

The operation of the backpropagation algorithm involves two main phases: the forward and backward passes. In the forward pass, the network's outputs are calculated and compared with the desired or target outputs. Subsequently, the error or loss is computed by comparing the desired and actual outputs. In the backward pass, the network's weights are adjusted based on the errors computed during the forward pass. The forward and backward pass phases are repeated iteratively until the error reaches an acceptably low level (Cilimkovic, 2010: 4).

4.3.RESEARCH FINDINGS

In this section of the study, metrics used to evaluate the performance of machine learning classification methods are explained. Subsequently, the findings from the analyses are presented. When machine learning algorithms are used for binary classification problems, the evaluation of classification is defined based on the confusion matrix. The confusion matrix used for binary classification problems is presented in Table 3 (Han, Kamber and Pei, 2012: 364).

Table 3: Confusion Matrix for Binary Classification

	True Positive Value	**True Negative Value**
Predicted Positive Value	True Positive	False Negative
Predicted Negative Value	False Positive	True Negative

The rows of the confusion matrix table, used for comparing the predicted and actual values in classification problems, represent the predicted class, while the columns represent the actual class. In the confusion matrix, True Positives and True Negatives indicate the number of correctly classified positive and negative examples, respectively, while False Negatives and False Positives represent the number of misclassified negative and positive examples, respectively. Evaluating the performance of classifiers plays a critical role in the creation and selection of classification models. While the most common measure used to evaluate machine learning classification performance is accuracy, various other performance metrics have been proposed. However, there is no general guideline among practitioners regarding which metric to choose when evaluating a classifier's performance (Liu et al. 2014: 20). Commonly used metrics for assessing the classification performance of machine learning algorithms are presented in Table 4.

Table 4: Classification Performance Metrics

Metrics	Formula
Accuracy	$\frac{\text{True Positives} + \text{True Negatives}}{\text{Positives} + \text{Negatives}}$
Error Rate	$\frac{\text{False Positives} + \text{False Negatives}}{\text{Positives} + \text{Negatives}}$
Precision	$\frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$
Recall (Sensitivity)	$\frac{\text{True Positives}}{\text{Positives}}$
Specificity	$\frac{\text{True Negatives}}{\text{Negatives}}$
F-Score (F-Measure)	$2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$

Reference:(Hossin and Sulaiman, 2015: 4).

The most commonly used metric for evaluating classification performance is the accuracy metric, which measures the ratio of correctly classified examples to the total data (Sokolova and Lapalme, 2009: 429-430). The error rate metric, used to evaluate classification performance, is the inverse of accuracy and measures the ratio of incorrectly classified examples. The precision metric measures the ratio of examples predicted as positive to the actual positive examples. Sensitivity measures the ratio of correctly classified positive examples to the actual positives, while specificity measures the ratio of correctly classified negative examples. The F-Score represents the harmonic mean between sensitivity and precision. These metrics, which have values ranging from 0 to 1, indicate the quality of performance. A herror rate close to zero and the other metrics close to 1 suggest high performance (Gong, 2021:4).

In the context of predicting the next day's direction of stock market indices, it has been observed in the literature that the accuracy metric is widely used. In this study, the accuracy metric was used to evaluate the prediction performance of machine learning methods. In this section, the results of the prediction performance of decision trees, random forests, k-nearest neighbors, naive bayes, logistic regression, support vector machines, and artificial neural networks for stock market indices of developed and developing countries are presented.

4.3.1.Decision Trees Algorithm Prediction Results

In the study, the directional movements of stock indices of developed countries, such as NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX indices, and developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100 indices, were predicted using the decision trees algorithm for machine learning classification methods. The confusion matrices and accuracy rates for the prediction of the directional movements of stock indices of developed countries using the decision trees algorithm are presented in Table 5.

Table 5:Confusion Matrix and Accuracy Rates for the Prediction of Stock Index Directions of Developed Countries Using the Decision Trees Algorithm

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	73	6	79	53,37%
	Projected Decline	229	196	425	
	Total	302	202	504	
NIKKEI 225	Projected Rise	28	5	33	54,14%
	Projected Decline	222	240	462	
	Total	250	245	495	
FTSE 100	Projected Rise	261	201	462	58,89%
	Projected Decline	7	37	44	
	Total	268	238	506	
CAC 40	Projected Rise	79	3	82	59,57%
	Projected Decline	204	226	430	
	Total	283	229	512	
DAX 30	Projected Rise	42	3	45	54,74%
	Projected Decline	226	235	461	
	Total	268	238	506	
FTSE MIB	Projected Rise	173	108	281	56,56%
	Projected Decline	114	116	230	
	Total	287	224	511	
TSX	Projected Rise	156	68	224	59,16%
	Projected Decline	137	141	278	
	Total	293	209	502	

When examining Table 5, it is observed that the decision trees algorithm accurately predicted the daily directions of the NYSE 100 index over 504 days with 73 increases and 196 decreases. For the NIKKEI 225 index, it predicted 28 increases and 240 decreases for its 495 daily directions. In the case of the FTSE 100 index, the algorithm predicted 261 increases and 37 decreases for its 506 daily directions. Similarly, for the CAC 40 index, it predicted 79 increases and 226 decreases over 512 days. As for the DAX 30 index, it predicted 42 increases and 235 decreases for its 506 daily directions. For the FTSE MIB index, the algorithm predicted 173 increases and 116 decreases for its 511 daily directions. Finally, for the

TSX index, it correctly predicted 156 increases and 141 decreases for its 505 daily directions. When examining the calculated accuracy rates, it can be observed that the decision trees algorithm achieved the best prediction with a 59.57% accuracy rate for the CAC 40 among the stock indices of developed countries. The lowest accuracy rate was achieved for the NYSE 100 with a 53.37% accuracy rate.

The confusion matrices and accuracy rates for the decision trees algorithm regarding the prediction of the directions of the stock indices of developing countries can be examined in Table 6.

Table 6: Confusion Matrix and Accuracy Rates for the Decision Trees Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	255	216	471	55.35%
	Projected Decline	1	14	15	
	Total	256	230	486	
BOVESPA	Projected Rise	167	104	271	60.41%
	Projected Decline	90	129	219	
	Total	257	233	490	
RTS	Projected Rise	251	161	412	61,90%
	Projected Decline	31	61	92	
	Total	282	222	504	
NFTY 50	Projected Rise	273	179	452	61,49%
	Projected Decline	12	32	44	
	Total	285	211	496	
IDX	Projected Rise	242	184	426	59.18%
	Projected Decline	16	48	64	
	Total	258	232	490	
IPC	Projected Rise	68	3	71	61.43%
	Projected Decline	191	241	432	
	Total	259	244	503	
BIST 100	Projected Rise	296	180	476	64.14%
	Projected Decline	0	26	26	
	Total	296	206	502	

When Table 6 is examined, it is observed that the decision tree algorithm correctly predicted the direction of movement for various stock market indices. The SSE index's 486-day movement direction had 255 upward and 14 downward movements correctly predicted. The BOVESPA index's 490-day movement direction had 167 upward and 129 downward movements correctly predicted. The RTS index's 504-day movement direction had 251 upward and 61 downward movements correctly predicted. The NFTY 50 index's 496-day movement direction had 273 upward and 32 downward movements correctly predicted. The IDX index's 490-day movement direction had 242 upward and 48 downward movements correctly predicted. The IPC index's 503-day movement direction had 68 upward and 241 downward movements correctly predicted. And the BIST 100 index's 502-day movement direction had 296 upward and 26 downward movements correctly predicted. When examining the accuracy rates calculated based on these prediction results, it is observed that among the stock market indices of developing countries, the decision tree algorithm provided the highest accuracy for predicting the BIST 100 index with a 64.14% accuracy rate, while the SSE index had the lowest accuracy with a 55.35% accuracy rate.

4.3.2.Random Forest Algorithm Prediction Results

In the study, the movement directions of stock market indices in developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX, and stock market indices in developing countries, namely SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, were predicted using machine learning classification methods, specifically the random forest algorithm. The confusion matrices and accuracy rates for the random forest algorithm's predictions of the movement directions of stock market indices in developed countries can be examined in Table 7.

Table 7: Confusion Matrix and Accuracy Rates for the Random Forest Algorithm's Prediction of Stock Indices' Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	291	170	461	64.09%
	Projected Decline	11	32	43	
	Total	302	202	504	
NIKKEI 225	Projected Rise	38	0	38	57.17%
	Projected Decline	212	245	457	
	Total	250	245	495	
FTSE 100	Projected Rise	260	198	458	59.29%
	Projected Decline	8	40	48	
	Total	268	238	506	
CAC 40	Projected Rise	105	13	118	62.70%
	Projected Decline	178	216	394	
	Total	283	229	512	
DAX 30	Projected Rise	18	0	18	50.59%
	Projected Decline	250	238	488	
	Total	268	238	506	
FTSE MIB	Projected Rise	182	85	267	62.82%
	Projected Decline	105	139	244	
	Total	287	224	511	
TSX	Projected Rise	78	4	82	56.37%
	Projected Decline	215	205	420	
	Total	293	209	502	

When Table 7 is examined, it can be observed that the random forest algorithm correctly predicted 291 upward and 32 downward movements in forecasting the 504-day direction of the NYSE 100 index, 38 upward and 245 downward movements in forecasting the 495-day direction of the NIKKEI 225 index, 260 upward and 40 downward movements in forecasting the 506-day direction of the FTSE 100 index, 105 upward and 216 downward movements in forecasting the 512-day direction of the CAC 40 index, 18 upward and 238 downward movements in forecasting the

506-day direction of the DAX 30 index, 182 upward and 139 downward movements in forecasting the 511-day direction of the FTSE MIB index, and 78 upward and 205 downward movements in forecasting the 502-day direction of the TSX index.

When examining the accuracy rates in Table 7, it can be observed that among the stock indices of developed countries, the index with the highest predicted accuracy by the random forest algorithm is NYSE 100 with a 64.09% accuracy rate, while the index with the lowest percentage of 50.59% accuracy is DAX 30.

Confusion matrices and accuracy rates for the prediction of the movement direction of stock indices in developing countries by the random forest algorithm can be examined in Table 8.

Table 8: Confusion Matrix and Accuracy Rates for the Random Forest Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	65	4	69	59.88%
	Projected Decline	191	226	417	
	Total	256	230	486	
BOVESPA	Projected Rise	47	6	53	55.92%
	Projected Decline	210	227	437	
	Total	257	233	490	
RTS	Projected Rise	88	6	94	60.32%
	Projected Decline	194	216	410	
	Total	282	222	504	
NFTY 50	Projected Rise	80	3	83	58.06%
	Projected Decline	205	208	413	
	Total	285	211	496	
IDX	Projected Rise	204	139	343	60.61%
	Projected Decline	54	93	147	
	Total	258	232	490	
IPC	Projected Rise	236	173	409	61.03%
	Projected Decline	23	71	94	
	Total	259	244	503	
BIST 100	Projected Rise	195	74	269	65.14%
	Projected Decline	101	132	233	
	Total	296	206	502	

When examining Table 8, it can be observed that the random forest algorithm correctly predicted 65 upward and 226 downward movements in forecasting the 486-day direction of the SSE index, 47 upward and 227 downward movements in forecasting the 490-day direction of the BOVESPA index, 88 upward and 216 downward movements in forecasting the 504-day direction of the RTS index, 80 upward and 208 downward movements in forecasting the 496-day direction of the NIFTY 50 index, 204 upward and 93 downward movements in forecasting the 490-day direction of the IDX index, 236 upward and 71 downward movements in forecasting the 503-day direction of the IPC index, and 195 upward and 132 downward movements in forecasting the 502-day direction of the BIST 100 index. Upon examining the accuracy rates calculated based on these prediction results, it can be seen that among the stock indices of developing countries, the index with the highest predicted accuracy by the random forest algorithm is NYSE 100 with a 65.14% accuracy rate, while the index with the lowest percentage of 55.92% accuracy is BOVESPA.

4.3.3.K-Nearest Neighbors Algorithm Prediction Results

In the scope of the study, the directional movements of stock indices of developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX indices, and stock indices of developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, were predicted using the k-nearest neighbors algorithm, a machine learning classification method. Confusion matrices and accuracy rates for the k-nearest neighbors algorithm's prediction of the movement direction of stock indices in developed countries can be examined in Table 9.

Table 9: Confusion Matrix and Accuracy Rates for the K-Nearest Neighbors Algorithm's Prediction of Stock Indices' Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	167	170	337	52.98%
	Projected Decline	135	32	167	
	Total	302	202	504	
NIKKEI 225	Projected Rise	95	100	195	48.48%
	Projected Decline	155	145	300	
	Total	250	245	495	
FTSE 100	Projected Rise	119	98	217	51.19%
	Projected Decline	149	140	289	
	Total	268	238	506	
CAC 40	Projected Rise	51	19	70	50.98%
	Projected Decline	232	210	442	
	Total	283	229	512	
DAX 30	Projected Rise	13	0	13	49.60%
	Projected Decline	255	238	493	
	Total	268	238	506	
FTSE MIB	Projected Rise	182	159	341	48.34%
	Projected Decline	105	65	170	
	Total	287	224	511	
TSX	Projected Rise	78	14	92	54.38%
	Projected Decline	215	195	410	
	Total	293	209	502	

When examining Table 9, it can be observed that the k-nearest neighbors algorithm correctly predicted 167 upward and 32 downward movements in forecasting the 504-day direction of the NYSE 100 index, 95 upward and 145 downward movements in forecasting the 495-day direction of the NIKKEI 225 index, 119 upward and 140 downward movements in forecasting the 506-day direction of the FTSE 100 index, 51 upward and 210 downward movements in forecasting the 512-day direction of the CAC 40 index, 13 upward and 238 downward movements in forecasting the 506-day direction of the DAX 30 index, 182 upward and 65 downward movements in forecasting the 511-day direction of the FTSE MIB index, and 78 upward and 195 downward movements in forecasting the 502-day direction of the TSX index.

Upon examining the accuracy rates calculated based on these prediction results, it can be seen that among the stock indices of developed countries, the index with the highest predicted accuracy by the k-nearest neighbors algorithm is TSX with a 54.38% accuracy rate, while the index with the lowest percentage of 48.34% accuracy is FTSE MIB.

Confusion matrices and accuracy rates for the k-nearest neighbors algorithm's prediction of the movement direction of stock indices in developing countries can be examined in Table 10.

Table 10: Confusion Matrix and Accuracy Rates for the K-Nearest Neighbors Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	10	4	14	48.46%
	Projected Decline	246	226	572	
	Total	256	230	486	
BOVESPA	Projected Rise	47	6	53	49.39%
	Projected Decline	210	227	437	
	Total	257	233	490	
RTS	Projected Rise	55	6	61	53.77%
	Projected Decline	227	216	443	
	Total	282	222	504	
NFTY 50	Projected Rise	51	3	54	52.22%
	Projected Decline	234	208	442	
	Total	285	211	496	
IDX	Projected Rise	204	211	415	45.92%
	Projected Decline	54	21	75	
	Total	258	232	490	
IPC	Projected Rise	191	173	364	52.09%
	Projected Decline	68	71	139	
	Total	259	244	503	
BIST 100	Projected Rise	195	144	339	51.29%
	Projected Decline	101	62	163	
	Total	296	206	502	

When examining Table 10, it can be observed that the k-nearest neighbors algorithm correctly predicted 10 upward and 226 downward movements in forecasting the 486-day direction of the SSE index, 47 upward and 227 downward movements in forecasting the 490-day direction of the BOVESPA index, 55 upward and 216 downward movements in forecasting the 504-day direction of the RTS index, 51 upward and 208 downward movements in forecasting the 496-day direction of the NIFTY 50 index, 204 upward and 21 downward movements in forecasting the 490-day direction of the IDX index, 191 upward and 71 downward movements in forecasting the 503-day direction of the IPC index, and 195 upward and 62 downward movements in forecasting the 503-day direction of the BIST 100 index.

Upon examining the accuracy rates in Table 10, it can be seen that among the stock indices of developing countries, the index with the highest predicted accuracy by the k-nearest neighbors algorithm is RTS with a 53.77% accuracy rate, while the index with the lowest percentage of 45.92% accuracy is IDX

4.3.4.Naive Bayes Algorithm Prediction Results

In the scope of the study, the directional movements of stock indices of developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX indices, and stock indices of developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, were predicted using the Naive Bayes algorithm, a machine learning classification method. Confusion matrices and accuracy rates for the Naive Bayes algorithm's prediction of the movement direction of stock indices in developed countries can be examined in Table 11.

Table 11: Confusion Matrix and Accuracy Rates for the Naive Bayes Algorithm's Prediction of Stock Indices' Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	244	123	367	64.09%
	Projected Decline	58	79	137	
	Total	302	202	504	
NIKKEI 225	Projected Rise	180	133	313	58.99%
	Projected Decline	70	112	182	
	Total	250	245	495	
FTSE 100	Projected Rise	194	125	319	60.67%
	Projected Decline	74	113	187	
	Total	268	238	506	
CAC 40	Projected Rise	222	123	345	64.06%
	Projected Decline	61	106	167	
	Total	283	229	512	
DAX 30	Projected Rise	194	117	311	62.25%
	Projected Decline	74	121	195	
	Total	268	238	506	
FTSE MIB	Projected Rise	214	112	326	63.80%
	Projected Decline	73	112	185	
	Total	287	224	511	
TSX	Projected Rise	232	118	350	64.34%
	Projected Decline	61	91	152	
	Total	293	209	502	

When examining Table 11, it can be observed that the Naive Bayes algorithm correctly predicted 244 upward and 79 downward movements in forecasting the 504-day direction of the NYSE 100 index, 180 upward and 112 downward movements in forecasting the 495-day direction of the NIKKEI 225 index, 194 upward and 113 downward movements in forecasting the 506-day direction of the FTSE 100 index, 222 upward and 106 downward movements in forecasting the 512-day direction of the CAC 40 index, 194 upward and 121 downward movements in forecasting the 506-day direction of the DAX 30 index, 214 upward and 112 downward movements in forecasting the 511-day direction of the FTSE MIB index, and 232 upward and 91 downward movements in forecasting the 502-day direction of the TSX index correctly. Upon examining the accuracy rates obtained from these prediction results, it can be seen that among the stock indices of developed countries, the index with the highest predicted accuracy by the Naive Bayes algorithm is TSX with a 64.34% accuracy rate, while the index with the lowest percentage of 58.99% accuracy is NIKKEI 225.

Confusion matrices and accuracy rates for the Naive Bayes algorithm's prediction of the movement direction of stock indices in developing countries can be examined in Table 12.

Table 12: Confusion Matrix and Accuracy Rates for the Naive Bayes Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	181	124	305	59.05%
	Projected Decline	75	106	181	
	Total	256	230	486	
BOVESPA	Projected Rise	210	134	344	63.06%
	Projected Decline	47	99	146	
	Total	257	233	490	
RTS	Projected Rise	188	100	288	61.51%
	Projected Decline	94	122	216	
	Total	282	222	504	
NFTY 50	Projected Rise	223	109	332	65.52%
	Projected Decline	62	102	164	
	Total	285	211	496	
IDX	Projected Rise	168	97	265	61.84%
	Projected Decline	90	135	225	
	Total	258	232	490	
IPC	Projected Rise	180	118	298	60.83%
	Projected Decline	79	126	205	
	Total	259	244	503	
BIST 100	Projected Rise	216	90	306	66.14%
	Projected Decline	80	116	196	
	Total	296	206	502	

When examining Table 12, it can be observed that the Naive Bayes algorithm correctly predicted 181 upward and 106 downward movements in forecasting the 486-day direction of the SSE index, 210 upward and 99 downward movements in forecasting the 490-day direction of the BOVESPA index, 188 upward and 122 downward movements in forecasting the 504-day direction of the RTS index, 223 upward and 102 downward movements in forecasting the 496-day direction of the NIFTY 50 index, 168 upward and 135 downward movements in forecasting the 490-day direction of the IDX index, 180 upward and 126 downward movements in forecasting the 503-day direction of the IPC index, and 216 upward and 116 downward movements in forecasting the 502-day direction of the BIST 100 index correctly. Upon examining the accuracy rates obtained from these prediction results, it can be seen that among the stock indices of developing countries, the index with the highest predicted accuracy by the Naive Bayes algorithm is BIST 100 with a 66.14% accuracy rate, while the index with the lowest percentage of 59.05% accuracy is SSE.

4.3.5.Logistic Regression Algorithm Prediction Results

In the scope of the study, the directional movements of stock indices of developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX indices, and stock indices of developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, were predicted using the Logistic Regression algorithm, a machine learning classification method. Confusion matrices and accuracy rates for the Logistic Regression algorithm's prediction of the movement direction of stock indices in developed countries can be examined in Table 13.

Table 13: Confusion Matrix and Accuracy Rates for the Logistic Regression Algorithm's Prediction of Stock Indices' Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	269	57	326	82.14%
	Projected Decline	33	145	178	
	Total	302	202	504	
NIKKEI 225	Projected Rise	232	73	305	81.62%
	Projected Decline	18	172	190	
	Total	250	245	495	
FTSE 100	Projected Rise	259	67	326	84.98%
	Projected Decline	9	171	180	
	Total	268	238	506	
CAC 40	Projected Rise	253	54	307	83.59%
	Projected Decline	30	175	205	
	Total	283	229	512	
DAX 30	Projected Rise	242	71	313	80.83%
	Projected Decline	26	167	193	
	Total	268	238	506	
FTSE MIB	Projected Rise	264	68	332	82.19%
	Projected Decline	23	156	179	
	Total	287	224	511	
TSX	Projected Rise	270	45	315	86.45%
	Projected Decline	23	164	187	
	Total	293	209	502	

When examining Table 13, it can be observed that the Logistic Regression algorithm correctly predicted 269 upward and 145 downward movements in forecasting the 504-day direction of the NYSE 100 index, 232 upward and 172 downward movements in forecasting the 495-day direction of the NIKKEI 225 index, 259 upward and 171 downward movements in forecasting the 506-day direction of the FTSE 100 index, 253 upward and 175 downward movements in forecasting the 512-day direction of the CAC 40 index, 242 upward and 167 downward movements in forecasting the 506-day direction of the DAX 30 index, 264 upward and 156 downward movements in forecasting the 511-day direction of the FTSE MIB index, and 270 upward and 164 downward movements in forecasting the 502-day direction of the TSX index correctly. Upon examining the accuracy rates obtained from these prediction results, it can be seen that the Logistic Regression algorithm achieved an accuracy rate of over 80% in predicting the stock indices of developed countries. Furthermore, the TSX index stands out as the best predicted index among the stock indices of developed countries with an accuracy rate of 86.45%. The index with the lowest percentage, with an accuracy rate of 80.83%, is DAX 30.

Confusion matrices and accuracy rates for the Logistic Regression algorithm's prediction of the movement direction of stock indices in developing countries can be examined in Table 14.

Table 14: Confusion Matrix and Accuracy Rates for the Logistic Regression Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	57	33	90	81.48%
	Projected Decline	199	197	396	
	Total	256	230	486	
BOVESPA	Projected Rise	234	58	292	83.47%
	Projected Decline	23	175	198	
	Total	257	233	490	
RTS	Projected Rise	251	49	300	84.13%
	Projected Decline	31	173	204	
	Total	282	222	504	
NFTY 50	Projected Rise	259	56	315	83.47%
	Projected Decline	26	155	181	
	Total	285	211	496	
IDX	Projected Rise	225	48	273	83.47%
	Projected Decline	33	184	217	
	Total	258	232	490	
IPC	Projected Rise	242	55	297	85.69%
	Projected Decline	17	189	206	
	Total	259	244	503	
BIST 100	Projected Rise	258	52	310	82.07%
	Projected Decline	38	154	192	
	Total	296	206	502	

When examining Table 14, it can be observed that the Logistic Regression algorithm correctly predicted 57 upward and 197 downward movements in forecasting the 486-day direction of the SSE index, 234 upward and 175 downward movements in forecasting the 490-day direction of the BOVESPA index, 251 upward and 173 downward movements in forecasting the 504-day direction of the RTS index, 259 upward and 155 downward movements in forecasting the 496-day direction of the NIFTY 50 index, 225 upward and 184 downward movements in forecasting the 490-day direction of the IDX index, 242 upward and 189 downward movements in forecasting the 503-day direction of the IPC index, and 258 upward and 154 downward movements in forecasting the 502-day direction of the BIST 100 index correctly.

Upon examining the accuracy rates in Table 14, it can be seen that the Logistic Regression algorithm achieved an accuracy rate of over 80% in predicting the stock indices of developing countries. Furthermore, the IPC index stands out as the best predicted index among the stock indices of developing countries with an accuracy rate of 85.69%. The index with the lowest percentage, with an accuracy rate of 81.48%, is SSE.

4.3.6.Support Vector Machines Algorithm Prediction Results

In the scope of the study, the directional movements of stock indices of developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX indices, and stock indices of developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, were predicted using the Support Vector Machines algorithm, a machine learning classification method. Confusion matrices and accuracy rates for the Support Vector Machines algorithm's prediction of the movement direction of stock indices in developed countries can be examined in Table 15.

Table 15: Confusion Matrix and Accuracy Rates for the Support Vector Machines Algorithm's Prediction of Stock Indices' Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	295	133	428	72.22%
	Projected Decline	7	69	76	
	Total	302	202	504	
NIKKEI 225	Projected Rise	210	59	269	80.00%
	Projected Decline	40	186	226	
	Total	250	245	495	
FTSE 100	Projected Rise	239	64	303	81.62%
	Projected Decline	29	174	203	
	Total	268	238	506	
CAC 40	Projected Rise	235	49	284	81.05%
	Projected Decline	48	180	228	
	Total	283	229	512	
DAX 30	Projected Rise	228	53	281	81.62%
	Projected Decline	40	185	225	
	Total	268	238	506	
FTSE MIB	Projected Rise	245	60	305	80.04%
	Projected Decline	42	164	206	
	Total	287	224	511	
TSX	Projected Rise	246	30	276	84.66%
	Projected Decline	47	179	226	
	Total	293	209	502	

When examining Table 15, it can be observed that the Support Vector Machines algorithm correctly predicted 295 upward and 69 downward movements in forecasting the 504-day direction of the NYSE 100 index, 210 upward and 186 downward movements in forecasting the 495-day direction of the NIKKEI 225 index, 239 upward and 174 downward movements in forecasting the 506-day direction of the FTSE 100 index, 235 upward and 180 downward movements in forecasting the 512-day direction of the CAC 40 index, 228 upward and 185 downward movements in forecasting the 506-day direction of the DAX 30 index, 245 upward and 164 downward movements in forecasting the 511-day direction of the FTSE MIB index, and 246 upward and 179 downward movements in forecasting the 502-day direction of the TSX index correctly. When examining the accuracy rates, it can be seen that the Support Vector Machines algorithm achieved an accuracy rate of 84.66% in predicting the stock indices of developed countries. Furthermore, the TSX index stands out as the best predicted index among the stock indices of developed countries with an accuracy rate of 84.66%. The index with the lowest percentage, with an accuracy rate of 72.22%, is NYSE 100.

Confusion matrices and accuracy rates for the Support Vector Machines algorithm's prediction of the movement direction of stock indices in developing countries can be examined in Table 16.

Table 16: Confusion Matrix and Accuracy Rates for the Support Vector Machines Algorithm's Prediction of Stock Indices' Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	233	60	293	82.92%
	Projected Decline	23	170	193	
	Total	256	230	486	
BOVESPA	Projected Rise	220	53	273	81.63%
	Projected Decline	37	180	217	
	Total	257	233	490	
RTS	Projected Rise	240	47	287	82.34%
	Projected Decline	42	175	217	
	Total	282	222	504	
NFTY 50	Projected Rise	233	40	273	81.45%
	Projected Decline	52	171	223	
	Total	285	211	496	
IDX	Projected Rise	236	78	314	79.59%
	Projected Decline	22	154	176	
	Total	258	232	490	
IPC	Projected Rise	222	47	269	83.30%
	Projected Decline	37	197	234	
	Total	259	244	503	
BIST 100	Projected Rise	236	47	283	78.69%
	Projected Decline	60	159	219	
	Total	296	206	502	

In Table 16, when examined, it can be seen that the support vector machines algorithm accurately predicted the direction of the SSE index for 233 increases and 170 decreases over 486 daily movements, for the BOVESPA index, 220 increases and 180 decreases over 490 daily movements, for the RTS index, 240 increases and 175 decreases over 504 daily movements, for the NIFTY 50 index, 233 increases and 171 decreases over 496 daily movements, for the IDX index, 236 increases and 154 decreases over 490 daily movements, for the IPC index, 222 increases and 197 decreases over 503 daily movements, and for the BIST 100 index, 236 increases and 159 decreases over 502 daily movements. When examining the accuracy rates obtained from the prediction results, it can be seen that the support vector machines algorithm accurately predicted the direction of gthe rowing countries' stock indices with an accuracy of over 80%. Among these indices, the IPC index has the highest accuracy rate of 83.30%, while the BIST 100 index has the lowest accuracy rate of 78.69%.

4.3.7. Artificial Neural Networks Algorithm Prediction Results

The study involves predicting the direction of stock indices for developed countries, including NYSE 100, NIKKEI 225, FTSE 100, CAC 40, DAX 30, FTSE MIB, and TSX, as well as developing countries, including SSE, BOVESPA, RTS, NIFTY 50, IDX, IPC, and BIST 100, using machine learning classification methods, specifically the Artificial Neural Networks algorithm. The confusion matrices and accuracy rates for predicting the movement direction of developed country stock indices using the Artificial Neural Networks algorithm can be examined in Table 17.

Table 17: Confusion Matrix and Accuracy Rates for the Artificial Neural Networks Algorithm's Prediction of Stock Indices Direction in Developed Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
NYSE 100	Projected Rise	263	43	306	83.73%
	Projected Decline	39	159	198	
	Total	302	202	504	
NIKKEI 225	Projected Rise	190	34	224	81.01%
	Projected Decline	60	211	271	
	Total	250	245	495	
FTSE 100	Projected Rise	244	9	253	93.48%
	Projected Decline	24	229	253	
	Total	268	238	506	
CAC 40	Projected Rise	243	47	290	83.01%
	Projected Decline	40	182	222	
	Total	283	229	512	
DAX 30	Projected Rise	197	21	218	81.82%
	Projected Decline	71	217	288	
	Total	268	238	506	
FTSE MIB	Projected Rise	259	48	307	85.13%
	Projected Decline	28	176	204	
	Total	287	224	511	
TSX	Projected Rise	243	21	264	85.86%
	Projected Decline	50	188	238	
	Total	293	209	502	

When examining Table 17, it can be seen that the artificial neural networks algorithm correctly predicted the direction of the NYSE 100 index's daily movement in 263 instances of increase and 159 instances of decrease. For the NIKKEI 225 index, it correctly predicted the direction in 190 instances of increase and 211 instances of decrease in its 495 daily movements. In the case of the FTSE 100 index, it made accurate predictions in 244 instances of increase and 229 instances of decrease in its 506 daily movements. As for the CAC 40 index, the algorithm correctly predicted the direction in 243 instances of increase and 182 instances of decrease in its 512 daily movements. For the DAX 30 index, it achieved accurate predictions in 197 instances of increase and 217 instances of decrease in its 506 daily movements. Regarding the FTSE MIB index, it successfully predicted the direction in 259 instances of increase and 176 instances of decrease in its 511 daily movements. In the case of the TSX index, it made correct predictions in 243 instances of increase and 188 instances of decrease in its 502 daily movements.

Looking at the accuracy rates presented in Table 17, the artificial neural networks algorithm achieved the highest accuracy in predicting the FTSE 100 index with a rate of 93.48%, while the lowest accuracy was observed in predicting the NIKKEI 225 index with a rate of 81.01%. Additionally, the algorithm consistently achieved accuracy rates of over 80% for all developed country stock indices, indicating its strong predictive performance.

The confusion matrix and accuracy rates for predicting the direction of developing country stock indices using the artificial neural networks algorithm can be found in Table 18.

Table 18: Confusion Matrix and Accuracy Rates for the Artificial Neural Networks Algorithm's Prediction of Stock Indices Direction in Developing Countries

Index	Movement Direction	Real Rise	Real Decline	Total	Accuracy
SSE	Projected Rise	217	40	257	83.74%
	Projected Decline	39	190	229	
	Total	256	230	486	
BOVESPA	Projected Rise	187	16	203	82.45%
	Projected Decline	70	217	287	
	Total	257	233	490	
RTS	Projected Rise	255	43	298	86.11%
	Projected Decline	27	179	206	
	Total	282	222	504	
NFTY 50	Projected Rise	262	123	385	70.56%
	Projected Decline	23	88	111	
	Total	285	211	496	
IDX	Projected Rise	230	50	280	84.08%
	Projected Decline	28	182	210	
	Total	258	232	490	
IPC	Projected Rise	219	14	233	89.26%
	Projected Decline	40	230	270	
	Total	259	244	503	
BIST 100	Projected Rise	281	52	333	86.65%
	Projected Decline	15	154	169	
	Total	296	206	502	

When examining Table 18, it can be observed that the artificial neural networks algorithm accurately predicted the direction of the following developing country stock indices: SSE with 217 correct upward and 190 correct downward predictions, BOVESPA with 187 correct upward and 217 correct downward predictions, RTS with 255 correct upward and 179 correct downward predictions, NIFTY 50 with 262 correct upward and 88 correct downward predictions, IDX with 230 correct upward and 182 correct downward predictions, IPC with 219 correct upward and 230 correct downward predictions, and BIST 100 with 281 correct upward and 154 correct downward predictions.

Regarding the accuracy rates, it is evident that the artificial neural networks algorithm achieved the highest accuracy of 89.26% in predicting the IPC index's direction among developing country stock indices, while the lowest accuracy of 70.56% was observed in predicting the NIFTY 50 index's direction.

4.4.DISCUSSION OF FINDINGS

The primary objective of this study is to employ machine learning classification methods for predicting the trends of stock market indices in both developed and developing countries. Furthermore, the study aims to assess the performance of these methods and identify the most effective prediction approach. To accomplish this objective, we utilized daily data and technical indicators spanning from January 1, 2012, to December 31, 2021. Our analysis involved an evaluation of the performance of various machine learning algorithms, including Decision Trees (DT), Random Forest (RF), k-Nearest Neighbors (k-NN), Naive Bayes (NB), Logistic Regression (LR), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) for the prediction of index movements. The prediction accuracy rates of these machine learning algorithms, pertaining to stock market index directions in both developed and developing countries, are detailed in Table 19.

Table 19: Prediction Accuracy Rates of Machine Learning Algorithms

STOCK MARKET INDICES OF DEVELOPED COUNTRIES							
INDICES	**DT**	**RF**	**k-NN**	**NB**	**LR**	**SVM**	**ANN**
NYSE 100	0,5337	0,6409	0,5298	0,6409	0,8214	0,7222	*0,8373
NIKKEI 225	0,5414	0,5717	0,4848	0,5899	*0,8162	0,8000	0,8101
FTSE 100	0,5889	0,5929	0,5119	0,6067	0,8498	0,8162	*0,9348
CAC 40	0,5957	0,6270	0,5098	0,6406	*0,8359	0,8105	0,8301
DAX 30	0,5474	0,5059	0,4960	0,6225	0,8083	0,8162	*0,8182
FTSE MIB	0,5656	0,6282	0,4834	0,6380	0,8219	0,8004	*0,8513
TSX	0,5916	0,5637	0,5438	0,6434	*0,8645	0,8466	0,8586
AVERAGE	0,5663	0,5900	0,5026	0,6260	0,8256	0,7943	*0,8343
STOCK MARKET INDICES OF DEVELOPING COUNTRIES							
INDICES	**DT**	**RF**	**k-NN**	**NB**	**LR**	**SVM**	**ANN**
SSE	0,5535	0,5637	0,4846	0,5905	0,8148	0,8292	*0,8374
BOVESPA	0,6041	0,5592	0,4939	0,6306	*0,8347	0,8163	0,8245
RTS	0,6190	0,6032	0,5377	0,6151	0,8413	0,8234	*0,8611
NIFTY 50	0,6149	0,5806	0,5222	0,6552	*0,8347	0,8145	0,7056
IDX	0,5918	0,6061	0,4592	0,6184	0,8347	0,7959	*0,8408
IPC	0,6143	0,6103	0,5209	0,6083	0,8569	0,8330	*0,8926
BIST 100	0,6414	0,6514	0,5129	0,6614	0,8207	0,7869	*0,8665
AVERAGE	0,5996	0,5872	0,5045	0,6197	*0,8340	0,8142	0,8326
OVERALL AVERAGE	**0,5830**	**0,5886**	**0,5036**	**0,6228**	**0,8298**	**0,8042**	***0,8335**

* *The accuracy rate of the method that predicts the index's direction with the highest accuracy.*

When considering Table 19, which provides insights into the accuracy rates of machine learning algorithms for predicting stock market index movements in both developed and developing countries, we find that artificial neural networks emerge as the most successful method with an overall average accuracy rate of 83.35%. The artificial neural networks algorithm exhibits an average accuracy of 83.43% for predicting the movement direction of stock market indices in developed countries and an average accuracy of 83.26% for those in developing countries. Notably, the FTSE 100 index, which the algorithm predicts with an accuracy rate of 93.48%, boasts the highest accuracy, setting a new benchmark for practical prediction accuracy. On the other hand, the NIFTY 50 index, predicted with an accuracy of 70.56%, is the lowest in terms of accuracy by the artificial neural networks algorithm. This leads us to conclude that the artificial neural networks algorithm effectively predicts stock market index movements in both developed and developing countries with an impressive performance rate exceeding 70%. This outcome aligns with the results of studies such as Dash and Dash (2016) and Ali et al. (2021), where different machine learning methods were compared in predicting stock market indices of developed and developing countries. However, it diverges from the findings of the Dimingo et al. (2021) study. In Dash and Dash (2016), various machine learning methods were used to predict the direction of the USA's S&P 500 in the developed country category and India's BSE SENSEX in the developing country category, with artificial neural networks developing as the most successful method. Similarly, the Ali et al. (2021) study predicted the direction of Japan's Tokyo Stock Exchange Nikkei 225 index from developed countries and China's Shanghai Stock Exchange SZSE Composite index, Pakistan's Karachi Stock Exchange KSE-100 index, and South Korea's Korea Stock Exchange KOSPI Composite index from developing countries using different machine learning algorithms, with artificial neural networks proving to be the most successful method. In contrast, the Dimingo et al. (2021) study reached a different conclusion by predicting the direction of the USA's S&P 500 in developed countries and the UK's FTSE 100 index, along with South Africa's ALSI and Brazil's BOVESPA indices in developing countries. They used different machine learning algorithms and identified random forest as the most successful method. Logistic regression, the second-ranking method with an overall average accuracy rate of 82.98%, demonstrates its competence. This algorithm predicts the direction of stock market indices in developed countries with an average accuracy of 82.56% and those in developing countries with an average accuracy of 83.40%. It's essential to highlight that it achieves an 86.45% accuracy rate in predicting the FTSE 100 index, the highest among all, while its lowest prediction accuracy is for the DAX 30 index, standing at 80.83%. Logistic regression excels at forecasting the movement directions of stock market indices in both developed and developing countries, surpassing an average accuracy rate of over 80%. Support vector machines take the third spot based on the overall average prediction accuracy rates, amounting to 80.42%. This algorithm predicts the direction of stock market indices in developed countries with an average accuracy of 79.43% and those in developing countries with an average accuracy of 81.42%. The TSX index, with an accuracy of 84.66%, is the index most accurately predicted by the support vector machines algorithm. The NYSE 100 index, with an accuracy of 72.22%, receives the lowest accuracy prediction. The support vector machines algorithm effectively forecasts the movement directions of stock market indices in both developed and developing countries, showcasing consistent and commendable average accuracy rates. In summary, considering the overall average prediction accuracy rates, artificial neural networks, logistic regression, and support vector machines take the lead, followed by Naive Bayes with an accuracy of 62.28%, Random Forest with an accuracy of 58.86%, Decision Trees with an accuracy of 58.30%, and k-Nearest Neighbors with an accuracy of 50.36%.

When Table 19 is examined based on the best method for predicting stock market indices, it becomes apparent that the artificial neural networks, while the best method, may not be universally superior for all indices. In other words, it doesn't exhibit the highest accuracy performance for all indices. As a result, artificial neural networks are the best prediction method for NYSE 100, FTSE 100, DAX 30, FTSE MIB in developed countries, as well as for SSE, RTS, IDX, IPC, and BIST 100 in developing countries. On the other hand, logistic regression emerges as the best prediction method for NIKKEI 225, CAC 40, and TSX in developed countries, as well as for BOVESPA and NIFTY 50 in developing countries. This outcome aligns with studies such as Phua et al. (2003), Liao and Wang (2010), and Özer et al. (2018), which investigated the position dependence of artificial neural networks. However, it contradicts Karlsson and Nordberg (2015). In the studies mentioned, Phua et al. (2003) focused on the direction of Germany's DAX, the USA's DJIA, FTSE-100, NASDAQ indices in developed countries, as well as China's HSI index in developing countries. Liao and Wang (2010) investigated the direction of the USA's DJIA, NASDAQ, and S&P500 indices, as well as China's Shanghai and Shenzhen Stock Exchange SAI, SBI, and HIS indices. Özer et al. (2018) examined developed country stock indices such as the USA's NASDAQ, the UK's FTSE100, Germany's DAX, France's CAC40, as well as developing country indices like China's SHANGAİ, India's NIFTY50, Mexico's IPC, and Istanbul's BIST100. In all these studies, the direction of indices was predicted using artificial neural networks, and the method proved to be successful in predicting indices. Karlsson and Nordberg (2015), on the other hand, studied developed country stock indices such as S&P500 (USA), DAX (Germany), Nikkei 225 (Japan), OMX30 (Denmark), and OMX30 (Sweden), and artificial neural networks were found to be a successful method for direction prediction. However, it was noted that its success did not significantly vary from country to country.

When Table 19 is examined, focusing on the developed category, artificial neural networks with an average of 83.43% emerge as the best prediction method in developed countries. In developing countries, logistic regression with an average of 83.40% and artificial neural networks with an average of 83.26% stand out as the leading methods. However, it is observed that the prediction performance of these methods varies among countries and/or indices, but does not significantly differ based on the development categories of countries. No studies were found in the literature that addressed this topic in the context of developed and developing countries. In this context, it is believed that the study can contribute to the literature.

When Table 19 is specifically examined based on developed countries alone, artificial neural networks with an average of 83.43% emerge as the best prediction method in developed countries. This result aligns with studies that focus on developed countries' stock indices, such as Kimoto et al. (1990) and Enke and Thawornwong (2005). In the Kimoto et al. (1990) study, the direction of Japan's Tokyo Stock Exchange TOPIX index was predicted, and in the Enke and Thawornwong (2005) study, the best method for predicting the movement direction of the USA's S&P 500 index was found to be artificial neural networks.

When Table 19 is specifically examined based on developing countries alone, logistic regression with an average of 83.40% and artificial neural networks with an average of 83.26% stand out as the most successful methods in predicting the movement direction of developing countries' stock market indices. This result aligns with studies comparing machine learning methods' performance on the stock market indices of developing countries, where artificial neural networks were identified as the most successful method in Cao, Parry, and Leggio (2011), Kara (2019) and Papuccu (2019) and support vector machines were found to be the most successful method in Filiz et al. (2021) and Kemalbay and Alkış (2021) studies. From these studies, it was determined that artificial neural networks were the best method for predicting the movement direction of China's SHSE in the Cao, Parry, and Leggio (2011) study, and Turkey's BIST 100 index in the Kara (2019) and Papuccu (2019) studies. However, in the studies by Filiz et al. (2021) and Kemalbay and Alkış (2021), logistic regression was identified as the best method for predicting the movement direction of the BIST 100 index.

When Table 19 is examined based on the worst-performing methods, it is observed that the method with the lowest prediction average is k-Nearest Neighbors with a general accuracy rate of 50.36%. This result aligns with studies comparing the prediction performance of different machine learning methods on indices, similar to Filiz et al. (2017) and Kara (2019) studies.

In these studies, k-Nearest Neighbors emerged as the worst method for predicting the direction of BIST 50 in Filiz et al. (2017) and BIST 100 in Kara (2019).

Conclusion

Stock prices are influenced by various factors such as general economic conditions, company performance, political events, and global trends. Investors who can accurately predict the future movements of stock prices, which are dynamic and complex, have the opportunity to seize significant gains and, at the same time, minimize potential losses. However, the prediction of stock prices is a topic that has been widely discussed in the literature. The Efficient Market Hypothesis forms the basis of this controversial subject. According to this hypothesis put forth by Fama (1970), stock prices quickly and accurately reflect all publicly available information. Therefore, the rapid incorporation of new information into market prices makes it impossible to predict prices based on past information. The Efficient Market Hypothesis received strong support in its early days. However, especially since the 1980s, many empirical studies have questioned the validity of this hypothesis, and opposing views asserting that market prices are to some extent predictable have been put forward.

Technical analysis is one of the prediction methods actively used in capital markets. In technical analysis, future prices are predicted based solely on the past price movements and trading volume data of stocks. However, in recent years, parallel to technological advancements, machine learning methods based on artificial intelligence are widely used in stock markets. Machine learning detects past relationships and trends in data, allowing for more accurate, fast, and parametric predictions compared to other forecasting methods. Therefore, in recent years, machine learning has emerged as a significant tool in the financial world for predicting stock prices, ushering in a new era.

This research aims to predict the directional movements of developed and developing stock market indices using machine learning classification methods, comparing their performance, and determining the best prediction method. In this context, developed countries representing G-7 economies and developing countries representing E-7 economies have been considered. The main stock market indices of each country, based on the largest market capitalization, have been selected. For developed countries, the selected stock market indices are: NYSE 100 (USA), NIKKEI 225 (Japan), FTSE 100 (UK), CAC 40 (France), DAX 30 (Germany), FTSE MIB (Italy), and TSX (Canada). For developing countries, the chosen stock market indices are: SSE (China), BOVESPA (Brazil), RTS (Russia), NIFTY 50 (India), IDX (Indonesia), IPC (Mexico), and BIST 100 (Turkey). A ten-year research period from 01.01.2012 to 31.12.2021 has been selected as the analysis period. Daily data from this period were used to calculate technical indicators such as Simple Moving Average, Weighted Moving Average, Exponential Moving Average, Momentum, Stochastic K, Stochastic D, Relative Strength Index, Moving Average Convergence Divergence, Larry Williams' R, and Commodity Channel Index. These technical indicators were used as input data for the analysis. Using the created prediction model, the success performance of machine learning classification methods such as decision trees, random forest, k-nearest neighbors, naive Bayes, logistic regression, support vector machines, and artificial neural networks in predicting the directional movements of stock market indices was examined.

Artificial neural networks achieved an accuracy of 83.35%, logistic regression 82.98%, support vector machines 80.42%, naive Bayes 62.28%, random forest 58.86%, decision trees 58.30%, and k-nearest neighbors 50.36% in predicting the directional movements of developed and developing countries stock market indices. According to the results obtained, during the analyzed period, artificial neural networks were determined as the best method for predicting the directional movements of stock market indices in both developed and developing countries. Along with artificial neural networks, logistic regression and support vector machines were also found to predict the directional movements of all indices with an accuracy rate exceeding 70%. Furthermore, it was concluded that the best method, artificial neural networks, was not universally applicable to all indices, meaning it was not the method with the highest accuracy performance for all indices. Accordingly, artificial neural networks were found to be the best method for predicting the NYSE 100, FTSE 100, DAX 30, FTSE MIB indices in developed countries, as well as the SSE, RTS, IDX, IPC, BIST 100 indices in developing countries. For the NIKKEI 225, CAC 40 and TSX indices in developed countries, and the BOVESPA and NIFTY 50 indices in developing countries, logistic regression was determined to be the best prediction method. Additionally, it was noted that the prediction performance of machine learning methods varied across countries and indices but did not significantly differ based on the development status of countries.

Throughout the period under scrutiny, artificial neural networks, logistic regression, and support vector machines have consistently demonstrated their prowess in accurately anticipating the fluctuations in stock market indices. These findings have the potential to serve as invaluable guiding stars for future financial prognostications. Consequently, they present an opportunity to offer specific recommendations for investors, portfolio managers, corporations, and economic policymakers. For investors, the utilization of artificial neural networks, logistic regression, and support vector machines can be a means to fine-tune their investment strategies in preparation for future periods, empowering them to make astute investment decisions. Portfolio managers, too, can enhance the quality of their services by seamlessly incorporating these methodologies into their portfolio management processes. Moreover, the provision of market trend projections alongside stock-specific forecasts can present a more comprehensive and insightful approach. Corporations can effectively employ these predictive techniques to gain profound insights into stock performance and trends, assess future growth potential, and meticulously evaluate potential risks. Lastly, economic policymakers can harness the identified successful methodologies to meticulously assess the efficacy of economic policies or to adeptly chart the course of the economy.While the results of this study offer valuable insights, it's crucial to acknowledge the presence of several limitations. To begin with, the exclusion of algorithms like deep learning, gradient boosting, multi-layer perceptron, and ensemble learning methods, which are commonly employed for classification tasks, has constrained the diversity of methods compared and the overall scope of the analysis. Moreover, the study's confinement to a ten-year timeframe might introduce potential variations in prediction outcomes that could manifest differently in shorter or longer periods. Furthermore, the absence of macroeconomic data in the analysis is a notable limitation. Financial market performance is often intricately connected to economic indicators, interest rates, and various macroeconomic factors. The exclusion of these influential factors from the analysis may inherently limit the predictive performance. The number and selection of technical indicators employed in the study also represent a limitation. Incomplete coverage of specific technical indicators may result in an inadequate representation of the dataset in the analysis. All these constraints should be considered when interpreting the study's results in their current form and when planning future research. In light of these limitations, further research exploring alternative methods and input variables for predicting stock market index movements holds the potential to enhance predictive accuracy. This, in turn, will guide the direction of future studies.

References

Abraham, S. and Kannappan, M. (2018),” The Role of Stock Exchanges in Corporate Governance”, International Journal of Pure and Applied Mathematics, Vol: 119 No.17, ISSN: 1314-3395, ss 1023-1033.

Adebiyi A.A, Adewumi A.O, and Ayo C.K. (2014) Comparison of ARIMA and artificial neural networks modelsfor stock price prediction. J Appl Math 2014:9–11. https://doi.org/10.1155/2014/614342

Aksoy, A. and Tanrıöven, C. (2007). Sermaye Piyasası Yatırım Araçları ve Analizi. Ankara: Gazi Kitapevi

Aksoy, B. (2021). Pay senedi fiyat yönünün makine öğrenmesi yöntemleri ile tahmini: Borsa İstanbul örneği. Business and Economics Research Journal, 12(1), 89-110. http://dx.doi.org/10.20409/berj.2021.312

Aktaş, M. (2008), Hisse Senedi Seçimi ve Hisse Senedi Seçiminde Etkili Olan Finansal Oranların Belirlenmesine Yönelik İMKB Şirketleri Üzerinde Bir Uygulama. Doctoral Thesis. Department of Finance. Social Sciences Institute. Istanbul University. Istanbul

Aktaş, S. (2019), Dow Jones Endüstri Ortalaması ve Nasdaq 100 Endekslerinin Makine Öğrenmesi Algoritmaları ile Tahmini, Istanbul Technical University. Social Sciences Institute. Master’s Thesis. Istanbul.

Ali,M, Khan , D. M. Aamir M, Ali A. and Ahmad Z. (2021).Predicting the Direction Movement of Financial Time Series Using Artificial Neural Network and Support Vector Machine. Hindawi Complexity. ID 2906463, https://doi.org/10.1155/2021/2906463

Apak, S. and Demirel, E. (2009). Finansal Yönetim: Sermaye Piyasaları. (1. Cilt) İstanbul: Papatya Yayıncılık.

Appel, G. (2015), “Technical Analysis Power Tools For Active Investors”, Pearson Education Inc, USA, 2005.

Arı, A. and Berberler, M. E. (2017). Yapay Sinir Ağları ile Tahmin ve Sınıflandırma Problemlerinin Çözümü İçin Arayüz Tasarımı . Acta Infologica , 1 (2) , 55-73. https://dergipark.org.tr/tr/pub/acin/issue/33868/335553

Arf, C. (1959) "Makine Düşünebilir Mi ve Nasıl Düşünebilir?", Atatürk Üniversitesi–Üniversite Çalışmalarını Muhite Yayma ve Halk Eğitimi Yayınları Konferanslar Serisi No: 1, 1959, Erzurum, s. 91-103

ATT (2021), AT&T Labs | Our Legacy, Web: Accession Date: 17.05.2021

Avadhani, V.A. (2011), Investment Management. Himalaya Publishing House, Eighth Revised Edition.

Babbage, C. (1834), "Statement Addressed to the Duke of Wellington respecting the Calculating Engine", London. Reprinted in The Works of Charles Babbage: Volume 3, ed. M.Campbell-Kelly.

Baker, M., Lang, W. (2020), "The Basic of Saving And İnvesting: Investor Education", Investor Production Institute, Investor Production Trust (IPT), ss.16,https://www.tn.gov/content/dam/tn/commerce/documents/securities/posts/The-Basics-of-Savings-and-Investing.pdf

Ballings, M, Poel, DV, Hespeel, N. and Grypb, R. (2015). Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications.Volume 42, Issue 20, 15 November 2015, Pages 7046-7056. https://doi.org/10.1016/j.eswa.2015.05.013.

Bao, Y, Kun, L., Guo, W. (2005) "Forecasting stock composite index by fuzzy support vector machines regression." Machine Learning and Cybernetics. Proceedings of 2005 International Conference on. Vol. 6. IEEE, 2005, http://www.sci-int.com/pdf/636327843900768486.pdf

Barber, D. (2012). Bayesian Reasoning and Machine Learning, Cambridge University Press, Cambridge.

Baresa, S., Bogdan, S. and Ivanovic, Z. (2013)," Strategy of Stock Valuatıon By Fundamental Analysis", University of Tourism and Management, Skopje, Vol. 4, Iss. 1, ISSN 1857-6982, pp. 45-51

Barone, R. (2003). From Efficient Markets to Behavioral Finance, University of Lecce Economics Working Paper No. 46/24, December, (pp.1-27) http://papers.ssrn.com/_id=493545, 17.01.2011.

Bekaert, G. and Hodrick, R. J. (2012), International Financial Management, 2nd edition, ISBN-13: 978-0-13-216276-0

Bell, J. (2020), Machine Learning:Hands-On for Developers and Technical Professionals, Second Edition, John Wiley & Sons, Inc., Indianapolis, Indiana ISBN: 978-1-119-64214-5

Berkson, J. (1944). Application of the Logistic Function to Bio-Assay, Journal of the American Statistical Association, 39(227), 357-365.

Berrar, D. (2018). Bayes' Theorem and Naive Bayes Classifier, Reference Module in Life Sciences, DOI:10.1016/B978-0-12-809633-8.20473-1

Berry, Ryan (2020), Machine Learning, The Ultimate Guide to Understand AI Big Data Analytics and the Machine Learning's Building Block

Application in Modern Life, http://longfiles.com/0a2smkas08eu/B086Y5JCF2.epub.html

BIST (2021), Cevaplarla Borsa ve Sermaye Piyasası, Borsa İstanbul, adresinden 15.05.2021 tarihinde alınmıştır.

BIST (2021b),Borsa İstanbul Borsacılık Faaliyetlerine İlişkin Esaslar Yönetmeliği, adresinden 10.05.2021 tarihinde alınmıştır.

Boisvert, S. and Gaa, C. (2010) "Innovation and Competition in Canadian Equity Markets", Bank Of Canada Revıew, Financial Markets Department. Summer 2010. Ss.16.

Borsaitaliana (2021), Borsa Italiana: sito ufficiale della Borsa di Milano -Borsa Italiana, Web: Accession Date: 26.05.2021

Breiman, L. (2001). Random Forests, Machine Learning, 45 (1): 1-32.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone.(1984) Classification and regression trees. Wadsworth and Brooks/Cole, Monterey, California, USA.

Brownlee, J. (2016), Master Machine Learning Algorithms: Discover How They Work and Implement Them From Scratch, Machine Learning Mastery.

Bovespa (2021) Ibovespa-B3, Web: adresinden 20.05.2021 adresinden 21.05.2021 tarihinde alınmıştır.

Cao, L. and Tay, F.E.H.(2001). Financial Forecasting Using Support Vector Machines, Neural Computing and Applications, 2001)10:184–192.

Cao, R., Parry, M. and Leggio, K. (2011). The three-factor model and artificial neural networks: Predicting stock price movement in China. Annals OR. 185. 25-44. 10.1007/s10479-009-0618-0.

Campesato, O. (2020), Artificial Intelligence, Machine Learning And Deep Learning, ISBN: 9781683924654, https://www.ebooks.com/en-us/book/209937164/artificial-intelligence-machine-learning-and-deep-learning/oswald-campesato/

Canbaş, S., Doğukanlı, H. (2007). Finansal Pazarlar Finansal Kurumlar ve Sermaye Pazarı Analizleri. Adana: Karahan.

Ceylan, A. and Korkmaz, T. (2010), Sermaye Piyasası ve Menkul Değer Analizi, Ekin Kitabevi, Bursa.

Chambers, D. and Dimson, E. (2016), "Fınancıal Market History: Reflections On The Past For Investors Today", CFA Institute Research Foundation, ISBN 978-1-944960-13-1 https://www.cfainstitute.org/-/media/documents/book/rf-publication/2016/financial-market-history-full-book.ashx

Chaengkham, S. and Wianwiwat S. (2021) Stock market index prediction using machine learning: evidence from leading Southeast Asian countries, Thailand and The World Economy | Vol. 39, No.2. 56-64.

Chien, C. F., and Chen, L. F. 2008. "Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-Technology Industry," Expert Systems with Applications, vol. 34, p. 280-290

Cichosz, P. (2015). Data Mining Algorithms Explained R, John Wiley & Sons Press, UK.

Cielen, D., Meysman, A. D. B., Ali, M. (2016), Introducing Data Science: Bıg Data, Machıne Learnıng, And More, Usıng Python Tools, Printed in the United States of America, ISBN: 9781633430037.

Cilimkovic, M. (2010), Neural Networks and Back Propagation Algorithm, CMA (2021), Financial Investments and Stock Markets, Web: adresinden 11.05.2021 tarihinde alınmıştır.

Cortes, C., and Vapnik, V.N. (1995). Support-Vector Networks. Machine Learning, 20, 273-297.

Cover, T.M. and Hart, P.E. (1967) "Nearest neighbor pattern classification". IEEE Transactions on Information Theory, IT-13(1):21–27 (1967).

Çınaroğlu, E. and Avcı, T., (2020), "THY Hisse Senedi Değerinin Yapay Sinir Ağları ile Kestirimi", Atatürk University Journal of Economics and Administrative Sciences, 34(1): 1-20, DOI: 10.16951/atauniiibd.530322

Darskuviene, V. (2010), Financial Markets, Vytautas Magnus University, Leonardo da Vinci Programme Project, Education and Culture, Lifelong Learning Programme.

Dash, R. and Dash, P. K. (2016). A Hybrid Stock Trading Framework Integrating Technical Analysis with Machine Learning Techniques. The Journal of Finance and Data Science, 2(1), 42-57.

Dayı, F. (2017),"Birden Fazla Borsada İşlem Gören Hisse Senetlerinin Değerlemesi", Doctoral Thesis, Department of Business Administration, Institute of Social Sciences, Gazi University, Ankara

Deckman, R.T. and Dale, M. (1986). Efficient Capital Markets and Accounting: A Critical Analysis, Second Edition, Prentice-Hall.

Dennis, K. (2004), Blaise Pascal and the First Calculator.Tech Directions. Nov2004, 64 (4), 10.

Deshmukh, S. Salgaonkar, A. and Harichandan, D. (2017), "Financial Markets", PGDFM - Post Graduate Diploma in Financial Management, University of Mumbai.

Deutsche Börse. (2021). About Deutsche Börse. Erişim Adresi: https://www.deutsche-boerse.com/dbg-en/investor-relations

Dimingo, R., Miwamba, J.W. M. and Bonga L.B.(2021) Predıctıon Of Stock Market Dırectıon: Applıcatıon Of Machıne Learnıng Models, NTERNATIONAL ECONOMICS, Volume 74, Issue 4 – Novmber, 499-536,

Doğukanlı, H. (2017), "Uluslararsı Finans", Gözden Geçirilmiş Beşinci Baskı, Karahan Kitabevi, Adana

Drakopoulou, V. (2015),"A Review of Fundamental and Technical Stock Analysis Techniques", Journal of Stock & Forex Trading, 5 (1) DOI: 10.4172/2168-9458.1000163

Dulhare, U.N. Ahmad, K. and Ahmad, K.A.B. (2020), Machine Learning and Big Data: Concepts, Algorithms,Tools and Applications, First edition, John Wiley & Sons, Inc., USA, ISBN 9781119654742

Dunis,, R., de la Fuente P. (2013) Forecasting IBEX-35 moves using support vector machines. Neural Comput & Applic 23, 229–236 (2013).

Enke, D., and Thawornwong, S. (2005). The Use of Data Mining and Neural Networks for Forecasting Stock Market Returns. Expert Systems with Applications, 29(4), 927-940.

Erdoğan, M. (2017), "Sıfırıncı Yasa", Journal of Accounting Science World, 19(3); s751.

Euronext. (2021). About Euronext. Erişim Adresi: https://www.euronext.com/about-euronext

Fama, Eugene F. (1965), "Random Walks in Stock Market Prices", Financial Analysts Journal, Vol. 21, No. 5, 55-59.

Fama, Eugene F. (1970), "Efficient capital markets: A review of theory and empirical work", The journal of Finance, 25(2), 383-417.

Faure, Ap (2013), "Equity Market: An Introduction",1st Edition, Quoin Institute, Limited & bookboon.com, ISBN 978-87-403-0594-4, s.134

FTSE (2020), FTSE 100 FTSE overview - London Stock Exchange, Web: / http://www.ftse.com/Analytics/FactSheets/temp/47432326-b750-42d5-97fc-180f60818666.pdf Access Date: 25.05.2021

Filiz, E., Karaboğa, H. A. and Akoğul, S. (2017). BIST-50 Endeksi Değişim Değerlerinin Sınıflandırılmasında Makine Öğrenmesi Yöntemleri ve Yapay Sinir Ağları Kullanımı. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 26(1), 231-241.

Filiz, E., Akogul, S. and Karaboğa, H. A. (2021). Büyük Dünya Endeksleri Kullanılarak BIST-100 Endeksi Değişim Yönünün Makine Öğrenmesi Algoritmaları ile Sınıflandırılması. Bitlis Eren University Journal of Science, 10 (2) , 432-441 . DOI: 10.17798/bitlisfen.889007

Freund, Y.and Schapire, R. (1996). Game theory, on-line prediction and boosting. In Proceedings COLT'86: Conf. on Comput. Learning Theory (pp. 325–332). New York, NY: ACM Press.

Geron, A. (2019) Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems, Second Edition,

Gibson, G. L. (1998), The Stock Exchanges of London, Paris and New-York, A Comparison, New York & London. ss. 45

Gong, M. (2021). A Novel Performance Measure for Machine Learning Classification. International Journal of Managing Information Technology, 13(1), 11–19.

Gurav, U. and Sidnal, N. (2018). Predict Stock Market Behavior: Role of Machine Learning Algorithms. DOI: 10.1007/978-981-10-7245-1_38.

Güngör, O.C. (2016),"İstanbul Menkul Kıymetler Borsası Özel Dili Üzerine", Journal of Turkish Language and Literature,Volume:2, Issue:3, Doi Number:10.20322

Grigorev, A. (2020). Machine Learning Bookcamp MEAP V06 (A. Books (ed.)). Copyright. Manning Publications. Erişim adresi: https://www.manning.com/books/machine-learning-bookcamp

Hackeling, G. (2014). Mastering Machine Learning with Scikit-Learn, Packt Publishing, Birmingham.

Haffener P. LeCun,Y and Bengio Y., (1995), The handbook of brain theory and neural networks Convolutional networks for images, speech, and time series, 3361 (10).

Han, J., Kamber, M., and Pei, J. (2012). "Data Mining: Concepts and Techniques". Third Edition, Elsevier Science & Technology. USA.

Harmon, M. E. and Harmon, S. S. (1997). Reinforcement Learning: A Tutorial. Wright Lab Wright-Patterson Afb Oh.

Harrington, P. (2012), Machine Learning In Action, NY, USA, Manning Publications Co.

Haykin, S. (2009), Neural Networks and Learning Machines, Pearson Education, Third Edition, USA,

Hilbe, J.M. (2015), Practical Guide to Logistic Regression, CRC press, 1st Edition, International Standard Book Number-13: 978-1-4987-0958-3. ISBN 9781498709576,

Hosmer JR, David W, Lemeshow S, and Sturdivant RX (2013), Applied Logistic Regression.3rd ed. New Jersey, USA, John Wiley & Sons, s.50.

Hua, Z. and Zhang, B. (2006), "A Hybrid Support Vector Machines And Logistics Regression Approach For Forecasting Intermittent Demand Of Spare Parts", Applied Mathematics and Computation, C: 181, 2006, pp.1035-1048

Huang, W.N, Yoshiteru, Wang, Shou-Yang (2005), Forecasting Stock Market Movement Direction with Support Vector Machine", Computers&Operations Research 32, 2513–2522

Hull, Isaiah. (2021). Machine Learning for Economics and Finance in TensorFlow 2: Deep Learning Models for Research and Industry, ISBN: 978-1-4842-6372-3, DOİ: 10.1007/978-1-4842-6373-0.

Hurwitz, J and Kirsch, D. 2018. Dummies için Makine Öğrenimi: IBM Limited Edition. New Jersey (NJ): John Wiley & Sons, Inc.

Hossin, M. and M.N, Sulaiman (2015), A Review on Evaluation Metrics for Data Classification Evaluations, International Journal of Data Mining & Knowledge Management Process, 5, 1-11, doi = 10.5121/ijdkp.2015.5201

Imandoust, S. B., and Bolandraftar, M. (2014). Forecasting the direction of stock market index movement using three data mining techniques: The case of Tehran Stock Exchange. Int. Journal of Engineering Research and Applications, 4(6), 106-117.

IBM (2021), IBM100- Deep Blue, Web: Access Date:14.05.2021

IDX (2021), Stock Index – IDX, Web: Access Date: 26.05.2021

International Monetary Fund-IMF (2014) "World Economic Outlook: Legacies, Clouds, Uncertainties". Washington (October). Ss. 162, https://www.imf.org/en/Publications/WEO/Issues/2016/12/31/Legacies-Clouds-Uncertainties#:~:text=Despite%20setbacks%2C%20an%20uneven%20global,World%20Economic%20Outlook%20(WEO).

Ismail, M.S., Noorani, M.S. Md, Ismail, M., Razak, F. Abdul, A. and Mohd A. (2020), Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange, Applied Soft Computing, Volume 93, 106422, https://doi.org/10.1016/j.asoc.2020.106422.

Johnson, R. and Soenen, L.A. (1996), "The Jakarta Stock Exchange: Risk/Return Characteristics", The Journal of Investing Spring, 5 (1) 37-46; ss.37-38, DOI:https://doi.org/10.3905/joi.5.1.37

JPX (2021), Corporate News Relaeses, Web: adresinden 16.05.2021 tarihinde alınmıştır.

Jose, C, Rajan, K., M.T, Shihabudheen (2014) "Quantitative Methods for Economıc Analysis", BA Economics, School Of Distance Education, University Of Calicut ss.66.

Joseph V.R. (2022), Optimal Ratio for Data Splitting, Jurafsky, D. and Martin, J.(2020), "Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition" Kara, İ. and Ecer, F. (2018), Bist Endeks Hareket Yönünün Tahmininde Siniflandirma Yöntemlerinin Performanslarinin Karşılaştırılması. Academic Journal of Social Research, 6(83), DOI: , s. 514-524

Kara, İ. (2019), Borsa Endeks, Hareket Yönünün Tahmininde Sınıflandırma Yöntemlerinin Performanslarının Karşılaştırılması: BIST 100 Örneği, Doktora Tezi, İşletme Anabilim Dalı, Sosyal Bilimler Enstitüsü, Afyon Kocatepe Üniversitesi, Afyon.

Kara, Y., Boyacioglu, M.A. and Baykan, Ö.K. (2011), Predicting Direction of Stock Price Index Movement Using Artificial Neural Networks And Support Vector Machines: The sample of the Istanbul Stock Exchange, Expert Systems with Applications, Volume 38, Issue 5, 5311-5319,

Karabıyık, L. and Anbar, A. (2010). Sermaye Piyasası ve Yatırım Analizi. Bursa: Ekin Yayınevi

Karagül, K. (2014). İstanbul Menkul Kıymetler Borsası'nda İşlem Gören Firmaların Destek Vektör Makineleri Kullanılarak Sınıflandırılması. Pamukkale University Journal of Engineering Sciences. 20(6), 174-178.

Karan, M.B, (2011), "Yatırım Analizi ve Portföy Yönetimi", , Gazi Kitabevi, 3.Baskı, Ankara

Karapınar, A. and Zaif, F.A. (2009). Finansal Analiz. Ankara: Gazi Kitapevi.

Karlsson S. and Nordberg, M. (2015) Stock Market Index Prediction Using Artificial Neural Networks Trained On Foreign Markets And How They Compare To A Domestıc Artıfıcıal Neural Network, KTH Royal Instıtute of Technology School of Computer Scıence and Communıcatıon, Degree Project, Fırst Level Computer Scıence Stockholm, Sweden.

Karslı, M. (2004). Sermaye piyasası borsa menkul kıymetler (Genişletilmiş 5.Baskı). İstanbul: Alfa Press

Kemalbay, G. and Alkış, B.N. (2021), Borsa endeks hareket yönünün çoklu lojistik regresyon ve k-en yakın komşu algoritması ile tahmini. Pamukkale University Journal of Engineering Sciences, 27(4), 556-569, doi: 10.5505/pajes.2020.57383.

Kızılkaya Y.M., Oğuzlar A. (2019) "Bazı Denetimli Öğrenme Algoritmalarının R Programlama Dili ile Kıyaslanması", Black Sea International Scientific Journal, 37, ss.92.

Kimoto, T., Asakawa, K., Yoda, M., ve Takeoka, M. (1990). Stock Market Prediction System with Modular Neural Networks. In 1990 IJCNN International Joint Conference on Neural Networks. 1–6.

Kirkpatrick, C.D. and Dahlguist, J.R.(2011),"Technical Analysis: The Complete Resource for Financial Market Technicians", Second Edition, Nee Jersey: FT Press, ss.3.

Koç Ustalı, N., Tosun, N., Tosun, Ö. (2021). "Makine Öğrenmesi Teknikleri ile Hisse Senedi Fiyat Tahmini", Eskişehir Osmangazi University Faculty of Economics and Administrative Sciences Journal, 16(1), 1 – 16. Doi: 10.17153/oguiibf.636017

Konuralp, G. (2001). Sermaye Piyasaları Analizler, Kuramlar ve Portföy Yönetimi. Alfa Yayınevi, İstanbul.

Kumar, M. and Thenmozhi, M., (2006). Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Forest. Indian Institute of Capital Markets 9th Capital Markets Conference Paper, Available at SSRN: https://ssrn.com/abstract=876544 or http://dx.doi.org/10.2139/ssrn.876544

Kumbure, M. M, Lohrmann C, Luukka P and Porras J (2022), Machine learning techniques and data for stock market forecasting: A literature review,, Expert Systems with Applications, Volume 197, .

Küçük H., Tepe C., and Eminoğlu I. (2013): "Classification of EMG Signals by k-Nearest Neighbor Algorithm and Support Vector Machine Methods", In 2013 21st Signal Processing and Communications Applications Conference (SIU) IEEE, s.1–4.

Lambert, D.R. (1980), Commodity Channel Index: Tool For Trading Cyclic Trends, Commodities Magazine.

Lantz, B. (2013). Machine Learning With R, Packt Publishing, Birmingham. ISBN 978-1-78216-214-8.

Leria, J.J.M. (2011), "El Mercado Y La Bolsa Mexicana De Valores", s.6. https://www.academia.edu/download/38120235/El_Mercado_y_la_Bolsa_Mexicana_de_Valores.pdf

Levinson, M. (2006), "Guide to Financial Markets", The Economıst In Assocıatıon Wıth Profile Books Ltd, This fourth edition, https://web.gccaz.edu/~bri2097936/Economics%20211%20Fall%202011/The%20Economist%20-%20Guide%20to%20the%20Financial%20Markets.pdf

Li, H. and Huang, S. (2021). Research on the Prediction Method of Stock Price Based on RBF Neural Network Optimization Algorithm. E3S Web of Conferences. DOI:10.1051/e3sconf/202123503088.

Liang, J. (2010). Market segmentation and dual-listed stock price premium: An Empirical Investigation of the Chinese Stock Market, A Thesis Submitted for the Degreeof PhD, University of St. Adrews, U.K.

Liao, Z., and Wang, J. (2010). Forecasting model of global stock index by stochastic time effective neural network. Expert Systems with Applications, 37(1), 834-841.

Liu, Y. and Zhou, Y. Wen, S. and Tang, C. (2014), A Strategy on Selecting Performance Metrics for Classifier Evaluation, International Journal of Mobile Computing and Multimedia Communication,, vol:6, doi:10.4018/IJMCMC.2014100102, s20-35

Lo, A. W. and MacKinlay, A.C.: A Non-Random Walk Down Wall Street, New Jersey, Princeton University Press, 2002.

London Stock Exchange. (2021). About Us. Erişim Adresi: https://www.londonstockexchange.com/about-us

Mallikarjuna, M. and Rao, R.P. (2019), Evaluation of forecasting methods from selected stock market returns. Financ Innov 5, 40 . https://doi.org/10.1186/s40854-019-0157-x

Marsland, S., (2015), Machine Learning an Algorithmic Perspective, USA, Second Edution, A Chapman & Hall Book CRC Press.

McCulloch, W. S., and Pitts, W. A. (1943). A logical calculus of the ideas immanent in nervous activity. Buttetin of Mathematics and Biophysics, 5, 115-133.

Meyers, Thomas (2003), The Technical Analysis Course, Third Edition, McGraw-Hill Boks, USA, 2003, s. 207.

Mijwil, Maad. (2015). History of Artificial Intelligence. 3. 1-8. DOİ: 10.13140/RG.2.2.16418.15046. s.2

Miller, J.D. (2017), Statistics for Data Science: Leverage the power of statistics for Data Analysis, Classification, Regression, Machine Learning, and Neural Networks, ISBN 978-1-78829-067-8, .

Miloseviç, N. (2016), Equity Forecast: Predicting Long Term Stock Price Movement using Machine Learning, Journal of Economics Library, Vol:3 No:2. http://www.kspjournals.org/index.php/JEL/article/view/750

Minsky, M. (1961). Steps toward artificial intelligence. In Computers and Thought, pages 406–450. McGraw-Hill,

Mirjalili, S., Faris, H. and Aljarah, I. (2020), Evolutionary Machine Learning Techniques: Algorithms and Applications, Springer Nature Singapore Pte Ltd. ISBN 978-981-32-9990-0 (eBook), https://doi.org/10.1007/978-981-32-9990-0

Mitchell, T.M. (1997), Machine Learning, McGraw Hill.

Mitchell, T. (2006), The Discipline of Machine Learning, School of Computer Science Carnegie Mellon University, Web: Access Date:27.05.2022

Mitchell, T. (2017), Machine Learning, McGraw Hill.

Moex (2021), Moscow Exchange Indices (MOEX Russia Index and RTS Index), Web: Access Date: 21.05.2021

Mohammed, M., Khan, M.B., Bashier, E.B.M. (2017), Machine Learning: Algorithms and Applications, CRC Press, Taylor & Francis Group, ISBN: 13: 978-1-4987-0538-7

Mohri, M., Rostamizadeh, A. and Talwalkar, A.(2012), Foundations Of Machine Learning, NY, USA, The MIT Press.

Mohri, M., Rostamizadeh, A. and Talwalkar, A. (2018), Foundations of Machine Learning, The MIT Press, Second edition, Cambridge, London, England. The MIT Press,, ISBN 9780262039406

Monadjemi, M. and Lodewijks, J. (2019) "Financial Markets and the Global Economy", 1st edition, ISBN 978-87-403-3188-2 https://bookboon.com/premium/reader/financial-markets-and-the-global-economy

Morgan, J.N. and Sonquist, J. A. (1963) "Problems in the analysis of survey data, and a proposal", J.Amer. Statist. Assoc., vol.58, 1963, pp.415-434.

Morgan J. N. and Messenger R. C.(1973), THAID: a sequential search program for the analysis of nominal scale dependent variables. Technical report, Institute for Social Research, Univ. of Michigan, Ann Arbor, MI.

Mosselaar, J.S. (2018), "A Concise Financial History of Europe", Robeco, Rotterdam, ss.45, https://www.robeco.com/docm/docu-summer-read-a-concise-financial-history-of-europe.pdf

Murphy, J.J.. (1999) Technical Analysis of The Financial Markets, New York Institute of Finance, USA.

Murphy, J.J. (2000), Charting Made Easy, Marketplace Books, USA, (Charting) s.45

Musonera, E. and Safari, V. (2008) "Establishing a Stock Exchange in Emerging Economies: Challenge and Opportunities", The Journal of International Management Studies, Vol. 3, No. 2, ss.64, Available at SSRN:

Myers,S.C. (2003), Financing of Corporations, Handbook of the Economics of Finance,1A, p.205-253,https://doi.org/10.1016/S1574-0102(03)01008-2.

Nadirgil, O. (2015), "Avrupa Birliği ve Türkiye'deki Ticari Ürün Borsaları ve Karşılaştırmalı Analizi", EU Specialization Thesis, Ministry of Food, Agriculture and Livestock, General Directorate of EU and Foreign Relations, Ankara.

Naved, M. and Srivastava, P. (2015), "Profitability of Oscillators Used in Technical Analysis for Financial Market", Advances in Economics and Business Management, ISSN: 2394-1545; Volume 2, Number 9; pp. 925-931

Neapolitan, R.E. and Jiang, X, (2018), Artificial Intelligence: With an Introduction to Machine Learning, Second Edition, CRC Press, ISBN:13: 978-1-138-50238-3.

Neil, C. (2021), Machine Learning. How It Works and How is Correlated to Artificial Intelligence and Deep Learning.

Nevasalmi. (2020).Forecasting Multinormal Stock Returns Using Machine Learning, The Journal of Finance and Data Science.6. , pp.87, 88, 92

North, M. (2016)," Data Mining for the Masses", CreateSpace Independent Publishing Platform, Isbn:9781523321438}, , s.94.

Nou, A., Lapitskaya. D., Eratalay, M.H. and Sharma, R. (2015) Predicting stock return and volatility with machine learning and econometric models: A comparative case study of the Baltic stock market. November 2021SSRN Electronic Journal.DOI:10.2139/ssrn.3974770

Nseindia (2021), Overview-NSE Indices, web: Access Date:26.05.2021

Nyasha, S. and Odhiambo, N.M. (2013), "The Brazilian Stock Market Development: A Critical Analysıs of Progress and Prospects During the Past 50 Years", Risk Governance & Control: Financial Markets & Institutions / Volume 3, Issue 3 ss.6.

NYSE (2021a) American Stock Exchange Temeline, New York Stock Exchange, Access Date: 11.05.2021

NYSE (2021b), New York Stock Exchange Index Series Methodology, Web: adresinden 16.05.2021 tarihinde alınmıştır.

OECD (2019), Artificial Intelligence in Society, OECD Publishing, Paris, https://dx.doi.org/10.1787/eedfee77-en.

OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities,Challenges,andImplicationsforPolicyMakers,https://www.oecd.org/finance/artificialintelligencemachinelearningbigdatainfinance.html.

Ou, P. and Wang, H. (2009). Prediction of stock market index movement by ten data mining techniques. Modern Applied Science, 3(12), 28-42.

Özer, A. Sarı, S.S. and Başakın, E.E. (2018) Bulanık Mantık ve Yapay Sinir Ağları ile Borsa Endeks Tahmini: Gelişmiş ve Gelişmekte Olan Ülkeler Örneği, Hitit University Social Sciences Institute Journal, 1(1),s.99-124.7

Qiu M, and Song Y (2016) Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Mode, PLoS ONE 11(5): e0155133. Quinlan J.R., 1993, "C4.5: Programs for Machine Learning", Morgan Kaufmann, San Mateo, CA, s.302

Papuçcu, H. (2019). "Borsa Endeksi Hareketlerinin Tahmini: Trend Belirleyici Veri", Journal of Vocational School of Social Sciences, 22(1) (E-Issn: 2564-7458) Ss. 246-256

Parameswaran, S. (2011), "Fundamentals of Financial Instruments, An Introduction to Stocks, Bonds, Foreign Exchange, and Derivatives ", John Wiley & Sons (Asia) Pte. Ltd. ISBN 978-0-470-82490-0

Park, A. (2020), Machine Learning: The Most Complete Guide for Beginners to Mastering Deep Learning, Artificial Intelligence and Data Science with Python (Data Science Mastery), ISBN 10: 1914167554 ISBN 13: 9781914167553, https://www.goodreads.com/book/show/50650863-machine-learning.

Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015). Predicting Stock and Stock Price Index Movement Using Trend Deterministic Data Preparation and Machine Learning Techniques. Expert Systems with Applications, 42(1), 259-268.

Pehlivanli, A. C., Asikgil, B. and Gulay, G. (2016) Indicator Selection with Committee Decision of Filter Methods for Stock Market Price Trend in ISE, Applied Soft Computing, 49, 792-800. DOI: 10.1016/jaSoC.2016.09.004

Phua, P. K. H., Zhu, X. and Koh, C. H., 2003. Forecasting stock index increments using neural networks with trust region methods. s.l., IEEE, pp. 260-265.

Prajapati, V. (2013). Big Data Analysis with R and Hadoop, Packt Publishing, Birmingham.

Presidential Digital Transformation Office (2021), National Artificial Intelligence Strategy 2021-2025, Republic of Turkey, https://www.cbddo.gov.tr/UYZS ve https://cbddo.gov.tr/sss/yapay-zeka/

Pring, M.J. (2014). Technical Analysis Explained, McGraw-Hill Education, Fifth Edition.

Quirin, J.J and Allen, A. (2000), "The effect of earnings permanence on fundamental information analysis", The Mid - Atlantic Journal of Business, South Orange, Dec 2000, Vol.36, Iss. 4; pg. 149.

Quiry, P., Dallocchio, M.F., Yann L. and Salvi, A. (2005), "Corporate Finance Theory and Practice", sixth edition, John Wiley & Sons Ltd, ISBN-13 978-0-470-09225-5, s.8

Racz, A., Bajusz, D., and Heberger, K .(2021) Effect of Dataset Size and Train/Test Split Ratios in QSAR/QSPR Multiclass Classification. Molecules, 26, 1111. https://doi.org/10.3390/ molecules26041111

Ramesh, P. and Wilpon, J. G. (1992). Modeling state durations in hidden markov models for automatic speech recognition. In Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on, volume 1, pages 381–384. IEEE.

RapidMiner (2022a), RapidMiner Documentation, Decision Treee, Access Date:17.06.2022

RapidMiner (2022b), RapidMiner Documentation, Neural Network, Web: Access Date: 17.06.2022

Redhead, K., (2008), Personal Finance and Investments A behavioural finance perspective, Routledge, NewYork.

Rodriguez, P. and Rodriguez, A. (2015). Predicting Stock Market Indices Movements. Capital Markets: Market Efficiency eJournal.

Rojas, R., (1996) "Neural Networks A Systematic Introduction", New York, Springer Verlag.

Rothman, T. (2021) Machine Learning Versus Fundamental Investment Analysis: A Meta-Analysis. .

Samadder, S., Ghosh, K. and Basu, T. (2012), "Scaling Analysis of Prime Stock Exchange Indices of the Emerging Seven (E7) Countries" International Journal of Applied Computational Science and Mathematics, ISSN 2249-3042 Vol:2, No:1, ss. 11-22.

Samuel, A. (1959), "Some Studies in Machine Learning Using the Game of Checkers", IBM Journal of Research and Development 3(3), 535-554.

Sarker, I.H. (2021) "Machine Learning: Algorithms, Real-World Applications and Research Directions". SN Computer Science. 2, 160.. .

SEC (2021), Ishares Trust, Web: adresinden 27.05.2021 tarihinde alınmıştır.

Securities and Exchange Commission of Pakistan-SECP (2019), Saving and Capital Markets, Web: Access Date: 05.05.2021

Sertkaya, B., (2018), "Gelişmiş ve Yükselen Piyasa Ekonomilerinde Hisse Senedi Endeksleri İle Makro İktisadi Değişkenler Arasındaki İlişki", Doctoral Thesis, Department of Economics, Institute of Social Sciences, Gazi University, Ankara

Shen, S., Jiang, H., and Zhang, T. (2012). Stock Market Forecasting Using Machine Learning Algorithms. http://cs229.stanford.edu/proj2012/ShenJiangZhang-StockMarketForecastingusingMachineLearningAlgorithms.pdf

Shibinu, S. (2014),"Economıcs Of Capıtal Market", Ba Economics, University Of Calicut, School of Distance Education, School of Dıstance Educatıon, India.

Shwartz, S.S.and David, S.B. (2014), "Understanding Machine Learning: From Theory to Algorithms", Cambridge University, United States of America, ISBN 978-1-107-05713-5.

Singh, R.R. (2012), "Stock Market Operations", Excel Books Private Limitet, Funjab University, New Delhi-India

Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management. 45. 427-437. DOI:10.1016/j.ipm.2009.03.002

Spahija, D. and Xhaferi, S. (2019), Fundamental and Technıcal Analysis of The Stock Price, International Scientific Journal Monte 1(1) DOI:10.33807/monte.1.201904160,

S&P Global (2021), The S&P/BMV IPC Turns, Access Date: 25.05.2021

Sreevidya, U. (2014), "Financial Markets and Services", University Of Calicut School Of Distance Education, Calicut University, http://14.139.185.6/website/SDE/sde168.pdf

SSE (2021), History and Development-Shanghai Stock Exchange, Web: adresinden 27.05.2021 tarihinde alınmıştır.

Subasi, A. Amir,F. Bagedo, K. Shams,A.and Sarirete, A. (2021). Stock Market Prediction Using Machine Learning, Procedia Computer Science.Volume 194, 2021, Pages 173-179. https://doi.org/10.1016/j.procs.2021.10.071

Subha MV, and Nambi ST. (2012) "Classification of stock index movement using k-nearest neighbours (k-nn) algorithm". WSEAS Transactions on Information Science and Applications, 9(9), 261-270.

Suresh, A.S (2013),"A Study on Fundamental and Technical Analysis", International Journal of Marketing, Financial Services & Management Research, Vol.2, No. 5, ISSN 2277- 3622, ss.44

Taliyev, R. (2011),"Kamuyu Aydınlatma ve Firma Değeri; Rusya Menkul Kıymetler Borsasında Bir Uygulama", Doctoral Thesis, Business Administration, Institute of Social Sciences, Ankara University, Ankara.

Teker S. and Özer B. (2012), "Sermaye Piyasaları Yapısal Karşılaştırması: Gelişmiş, Gelişmekte Olan Ülkeler ve Türkiye", Journal of Accounting and Finance July/2012,

Tekin S. and Çanakoğlu E., (2018). Prediction of Stock Returns in Istanbul Stock Exchange Using Machine Learning Methods," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2018, pp. 1-4, doi: 10.1109/SIU.2018.8404607.

Tektaş, A. and Karataş, A. (2004) "Yapay Sinir Ağları ve Finans Alanına Uygulanması: Hisse Senedi Fiyat Tahminlemesi", Atatürk University Journal of Economics and Administrative Sciences, 18(3-4). 337-349.

Thaler, R. and Mullaınathan, S.. (2000). Behavioral Economics, Massachusetts Institute of Technology Department of Economics Working Paper Series, Working Paper: 00-27, September.

TSPB (2021a) Türkiye Sermaye Piyasası Aracı Kuruluşlar Birliği, Publications and Reports, Access Date: 25.05.2021

TSPB (2021b) Türkiye Sermaye Piyasası Aracı Kuruluşlar Birliği, Publications and Reports,https://www.tspb.org.tr/wpcontent/uploads/2020/07/TurkiyeSermayePiyasasi2019r.pdf#page=16 Access Date:25.05.2021

TSX (2021) Overview of Trading on Toronto Stock Exchange and TSX Venture Exchange, Web: adresinden 18.05.2021 tarihinde alınmıştır.

Tuffery, S. (2011). Data Mining and Statistics for Decision Making, John Wiley & Sons Press, UK.

Tunalı, H. (2013). Finansal Piyasalar ve Kurumlar, Istanbul University Faculty of Open and Distance Education Economics Undergraduate Program, Http://Auzefkitap.İstanbul.Edu.Tr/Kitap/İktisat_Ao/Finansalpiyasalarvekur.Pdf, Access Date:04.04.2021

Turing, A. M. (1950) Computing Machinery and Intelligence. Mind 49: 433-460.

Türkiye National Artificial Intelligence Strategy 2021-2025. Digital Transformation Office of the Presidency of the Republic of Turkey and the Ministry of Industry and Technology, Access Date: 27.05.2022

Türkmen A.C. and Cemgil A.T. (2015) "An application of deep learning for trade signal prediction in financial markets". IEEE 2015 23nd Signal Processing and Communications Applications Conference, Malatya, Turkey, 16-19 May 2015

Udomsak, N. (2015), "How do the naive Bayes classifier and the Support Vector Machine compare in their ability to forecast the Stock Exchange of Thailand?", https://doi.org/10.48550/arxiv.1511.08987.

UNESCO (2021), Aizanoi Antique City - UNESCO World Heritage Centre, Web: Access Date:27.05.2021

Varlık, B. (2017), Spectral Analysis of Stock Prices in Borsa Istanbul, Phd Thesis, Department of Business Administration, Institute of Social Sciences, Adnan Menderes University, Aydın.

Vegas, G. (2011), "Dall'Unità ai Giorni Nostri: 150 Anni di Borsa in Italia", Pubblicazione a cura di Consob, ss.79.

Wang, A. (2013). KOSPI 200 and HSI Index Performance During the Period 2002-2011. Journal of Financial Markets, 15(3), 189-205.

Watson, D. and Head, A. (2010), "Corporate Finance: Principles and Practice", Financial Times Prentice Hall, Fifth Edition, ISBN:0273725343, 9780273725343, ss.34

Watt, J., Borhani, R. and Katsaggelos, A.K. (2020), Machine Learning Refined Foundations, Algorithms, and Applications, University of Cambridge, title:www.cambridge.org/9781108480727, DOI:10.1017/9781108690935

WFE (World Federation of Exchanges). (2021). Annual Statistics Guide. Veriler , Access Date:20.05.2022

White, H. (1988). Economic prediction using neural networks: The case of IBM daily stock returns. In IEEE 1988 international conference on neural networks, San Diego, CA, USA, 2, p451–458). https://doi.org/10.1109/ICNN.1988.23959.

Williams, G. (2011). Data Mining with Rattle and R, Springer Science+Business Media, New York, USA.

Will, K. (2013). Autonomous Chess-Playing, Department of Computer Science TUDarmstadt, https://www.semanticscholar.org/paper/Autonomous-Chess-Playing-Will/aa5c9a1fd94d9cf421255cdd6c45f2d8e5f922d8 Access Date: 25.10.2022

Wilson, A. (2018), Ensuring Data and AI Work for People and Society, The Ada Lovelace Institute, Nuffield Foundation, https://www.

adalovelaceinstitute.org/wp content/uploads/2018/11/Ada_Lovelace_Institute_prospectus-WEB.pdf

Wilpon, J. Thomson, D., Bangalore, S., Haffner P. and Johnston M. (2017) The Fundamentals of Machine Learning, Interactions, https://www.interactions.com/wp-content/uploads/2017/06/machine_learning_wp-5.pdf.

Witten, I.H., Frank, E. ve Hall, M.A. (2011). Data Mining: Practical Machine Learning Tools and Techniques, 3. Baskı, Morgan Kaufman Publications, USA.

Wood, K.W.. (2009) Trade The Patterns, The Revolutionary Way of Trading The CCI, W & A Publishing,

World Bank (2021), World Bank Country and Lending Groups, Web: Access Date:17.05.2021

World Economic Situation and Prospects (WESP), United Nations Department of Economic and Social Affairs, ISBN (PDF): 9789210547178, https://doi.org/10.18356/005f3c1b-en

Zahedi, J. and Rounaghi, M.M. (2015), Application of Artificial Neural Network Models and Principal Component Analysis Method in Predicting Stock Prices on Tehran Stock Exchange, Physica A: Statistical Mechanics and Its Applications, 438: 178-187.

Predicting Stock Market Index Movements With Machine Learning

Nazif Ayyıldız

ÖZGÜR PRESS

ISBN 978-975-447-821-1