A New Algorithmic Trading Approach Based on Ensemble Learning and Candlestick Pattern Recognition in Financial Assets

Üzeyir AYCEL[1*], Yunus SANTUR[1]
[1] Software Engineering, Institute of Science, Firat University, Elazig, Turkey
[*1] uzeyir.aycel@hotmail.com, [2] ysantur@firat.edu.tr

Abstract: Financial assets considered as time series are chaotic in nature. The main goal of investors is to take a position at the right time and in the right direction by making predictions about the future of this chaotic series. These time series consist of the opening, low, high, and closing prices of a certain period. The approaches used to make predictions about trend direction and strength using moving averages and indicators based on them have noise and lag problems as they are obtained statistically. Candlestick charts, on the other hand, reflect the price-based psychology of bear and bull investors, and facilitate the interpretation of price movements by consolidating the said opening, closing, lowest and highest prices in a single image. It is known that it was applied to Japanese rice markets for the first time in history and there are more than 100 candle patterns. In this study, an extensible architecture software framework using factory patterns and an object-oriented approach is proposed for defining candlestick patterns and developing intelligent learning algorithms based on them. In the studies carried out for financial assets, the profit factor, which shows the portfolio gain of the strategy, is used. It is desirable that this number of wins be greater than 1. When the proposed approach is tested for 5 major financial assets, this value was obtained as greater than 1 for all assets. The proposed software framework can also be used in the development of new robotic approaches in terms of being applicable to all kinds of financial assets in every period.

Key words: Algorithmic trading, Ensemble learning, Financial forecasting, Pattern recognition, XGBoost.

Finansal Varlıklarda Topluluk Öğrenme ve Mum Grafiği Patern Tanımaya Dayalı Yeni Bir Algoritmik Ticaret Yaklaşımı

Öz: Zaman serileri olarak ele alınan finansal varlıklar yapısı gereği kaotiktir. Kaotik olan bu seri üzerinde geleceğe yönelik tahminlerde bulunarak doğru zamanda, doğru yönde pozisyon almak yatırımcıların temel hedefidir. Bu zaman serileri belirli bir periyotta ki açılış, en düşük, en yüksek ve kapanış fiyatından oluşmaktadır. Hareketli ortalamalar ve bunlara dayalı olarak göstergeleri kullanarak trend yönü ve gücü hakkında tahminler yapmak için kullanılan yaklaşımlar istatistiksel olarak elde edildiğinden gürültü ve gecikme problemlerine sahiptir. Mum grafikler ise bahsi edilen açılış, kapanış, en düşük ve en yüksek fiyatları tek bir görselde konsolide ederek, fiyat hareketlerinin yorumlanmasını kolaylaştırmalarının yanı sıra ayı ve boğa yatırımcılarının fiyat temelli psikolojilerini de yansıtırlar. Tarihte ilk defa Japon pirinç piyasalarına uygulandığı ve 100'den fazla mum patern olduğu bilinmektedir. Bu çalışmada mum grafik paternlerinin tanımlanması ve bunlara dayalı akıllı öğrenme algoritmalarının geliştirilmesi için fabrika deseni ve nesne tabanlı yaklaşım kullanan genişleyebilir mimaride bir yazılım çatısı önerilmiştir. Finansal varlıklar için gerçekleştirilen çalışmalarda stratejinin portföy kazancını gösteren kazanç katsayısı kullanılmaktadır. Bu kazanç sayısının 1'den büyük olması arzu edilmektedir. Önerilen yaklaşım 5 major finansal varlık için test edildiğinde, tüm varlıklarda bu değer 1'den büyük olarak elde edilebilmiştir. Önerilen yazılım çatısı her türlü finansal varlığa her periyotta uygulanabilir olması açısından yeni robotik yaklaşımların geliştirilmesinde de kullanılabilecektir.

Anahtar kelimeler: Algoritmik ticaret, Desen tanıma, Finansal tahmin, Topluluk öğrenimi, XGBoost.

1. Introduction

An exchange is an organized market in which tradeable securities, commodities, foreign Exchange, and futures are bought or sold. The stock market is one of the most popular investment environments that work with the principle of capital exchange due to its regular trading mechanism, up-to-date prices, and real-time trading, as well as creating a safe environment for its investors. Another tool that can be shown as an investment tool is stock market stocks [1]. Stocks and sector indices such as banking, construction and health are examples of financial time series. The future values of financial time series have a complex and chaotic structure that changes dynamically due to instant decisions of investors. Therefore, the estimation process on these time series is considered to be a challenging task based on more than one condition [2].

---

[*] Corresponding author: uzeyir.aycel@hotmail.com. ORCID Number of authors: [1] 0000-0003-0847-9418, [2] 0000-0002-8942-4605

There are two main approaches where attempting to predict stock prices and investors get information before investing. These are called technical analysis and fundamental analysis. The technical analysis develops statistical-based approaches to predict future prices using past price movements and trading volumes. Fundamental analysis, on the other hand, is to have an idea about whether to invest or not by taking into account other factors, especially economic, environmental, political, and financial factors in addition to these statistical approaches [3]. Fundamental analysis can be used to examine a company's balance sheet such as assets, liabilities and earnings, and even the company's management and competitive conditions. In technical analysis, it is generally used for short-term transactions, but also instant, minute, 5 minute, hourly, daily, weekly and monthly transactions. For this purpose, online platforms such as Tradingview and Investing and package programs called data terminals are also used [4]. In addition, stock charts containing previous price movements are used to predict future price movements of stocks [2].

Understanding how to use and interpret charts with technical analysis is of paramount importance to any trader. The starting point of technical analysis is to read the charts correctly and to choose a chart type suitable for the analysis style. These charts consist of price and date information in their most basic form. Thanks to the graphs, the trend, the resistance, and support levels that the price has difficulty in passing, the pattern formations can be determined and the decision is made to enter a buy or sell position using various analysis tools. Therefore, charts are the final point of departure in the decision mechanism for every investor.

There are basically three types of charts on trading platforms. These charts are called line, bar, and candlestick charts. The line chart, as seen in Figure 1, is the simplest form of chart used to observe price development information at a glance. These charts, which are usually drawn over the closing prices, can be drawn over the opening, lowest or highest price. However, in any case, it only refers to a single data.

Figure 1. An example line graph representation

The bar graph, also known as the bar graph, has the representation in Figure 2. It is the type of chart that shows the opening, closing, lowest and highest prices in the session and is widely used in technical analysis. In the bar chart, the top of the bar shows the highest price the share has reached throughout the day. The lowest point of the bar also indicates the lowest price that the share has gone through throughout the day. The lengthening of the bar indicates that the volatility of the share has increased, while the shortening of the neck indicates that the volatility of the share has decreased. In a bar chart, the closing price is a line to the right on the bar, and the opening is a line to the left. The fact that the lowest price of the stock and the closing price are close to each other indicates that there is selling pressure in the stock in the short term, but this indicates that the stock has technically weakened. On the contrary, the closeness of the highest price reached during the day with the closing price indicates that the stock is technically stronger in the short run [5].

Figure 2. An example bar graph representation

The use of candlestick patterns in time series has been very popular around the world lately because they are more visually appealing and generally easier to read and interpret than the bar or line charts. However, candlestick charts give investors and analysts a clearer view, as well as give a more accurate depiction of market sentiment. Many traders are able to trade using candlestick charts, exemplified in Figure 3, without any helpful indicators. Candlestick charts, in addition to showing the market movements in detail, make it easy to see the relationship between the opening (Open) and closing (Close) and the highest (High) and lowest (Low) price movements in a single visual, and it is called OHCL for short [6].

Figure 3. An example candlestick chart

Candlestick charts were learned and used by the whole world when Steve Nison went to Japan and did research [6]. The traditional Japanese candlestick interpretation technique has the potential to be a unique approach that can be used in the development of financial forecasting models. Traders use candlesticks as in Figure 3 and can analyze the OHLC data, which includes all price movements, in a single visual [7]. As seen in Figure 4, the green candle indicates an upward trend in the market. Since the closing price is higher than the opening price, it is considered positive for the uptrend and indicates the fact that the bulls are in control. The longer the candle body, the stronger the buying interest. The red candle indicates that the market has a downward trend. This indicates that there is more selling than buying. It indicates the fact that the downtrend is also under the control of the bears as the closing price is lower than the opening price. The longer the candle body, the stronger the supply-demand balance between bears and bulls is assumed. The lines above and below the candle body are called "Shadow" or "Wick". The upper shadow represents the tested but failed high point of the price, while the lower shadow represents the tested but failed low point of the price. There are more than 104 candle formations that support candlestick charts. Japanese candlesticks have been affected by noisy stock prices [6].
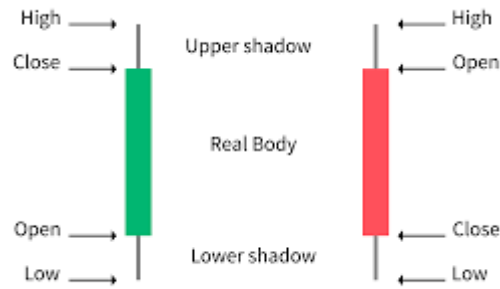
Figure 4. Candlestick description

The work will proceed as follows: Chapter 2 provides a summary of the work in the field of financial forecasting. Chapter 3 presents the architecture of the proposed model and the proposed system. In Chapter 4, the experimental study of the proposed approach is mentioned. The results of the study and future studies are given in Chapter 5.

2. Background

Prediction of financial time series is of primary importance for investors. In recent years, there has been a noticeable increase in the publications published in this field. In this context, different algorithms and methods have been developed. Research in this area will be outlined below.

Kusuma et al. (2019) investigated the predictability of the stock market using deep convolutional network and candlestick charts. In this study, different neural networks such as convolutional neural network, residual network and visual geometry group network are used. Consequently, it has been used to design a decision support framework that can be used by traders to provide suggested indicators of future stock price direction [8].

Hung et al. (2021), taking a new approach, proposed a deep predictive (DPP) method for price action by using candlestick charts in stock historical data. This method consists of three steps: 1. parsing a particular candlestick chart into sub-charts; 2. Using the CNNautoencoder to get the best representation of subcharts; 3. Application of RNN to predict price movements from a collection of sub-chart representations [9].

Sadeghi and Farid (2021) aim to design a stock market prediction based on candle patterns and use fuzzy logic to model market rules and candles. It explores how a trading algorithm should be designed using trading information analysis and signaling techniques and how they should be combined with the capabilities of fuzzy logic. In this study, the short- and medium-term benefits of the proposed method have been demonstrated in various markets. Again, this study enables investors to use private information to make safer investments. Also, in this study, fuzzy logic is used to implement trading systems based on candlestick patterns [10].

Yee et al. (2021) prefer to examine the states of the Ichimoku Kinko Hyo and Japanese Candlestick indicators to generate strong bullish and strong bearish signals. This model using real data determined from Bursa Malaysia, three counters NOTION, ADVENTA and RANHILL have been selected with a strong fundamental background through fundamental analysis and technical analysis in a short-term trade. Comparative work has been done on the performance of this combination analysis for 8 months of daily data market price action. It was determined that the experimental results revealed that the combination algorithm was successfully predicted for the short term. Therefore, this study has been a promising method for traders and investors who are interested in estimating the stock price to make a profit [11].

Lin et al (2021) proposed the PRML model, a new candlestick pattern recognition model that uses machine learning methods to make stock trading decisions even better. Four popular machine learning methods and 11 different feature types were applied to all possible cases of everyday patterns to initiate the pattern recognition scheme. The PRML model was applied to forecast all Chinese market stocks from January 1, 2000 to October 30, 2020 [12].

Ardianti et al. (2021), in their study, performed the estimation of the trading strategy by analyzing the candlestick pattern using Artificial Neural Network (ANN). Technical indicator tools and candlestick pattern were created as features and label data during the modeling process. The method has been applied to four stocks from IDX for a certain period of time through their technical indicators. In the 28-day period, the model was found to

produce the highest accuracy, reaching 85.96%. In addition, K-Fold Cross-Validation was used to evaluate the result of the model performance [13].

To highlight its capabilities, Chen and Tsai (2020) proposed a two-step approach to automatically recognize candlestick patterns. As a first step, Gramian Angular Field (GAF) was used to encode the time series as different image types. The second step was to use a Convolutional Neural Network (CNN) with GAF images to learn the eight critical types of candlestick patterns. In this paper, his approach is named GAF-CNN [14].

Yassini et al. (2019) analyzed the existence of different candlestick patterns in this study. It can predict trends in the foreign exchange market. The first set of minor hypotheses includes whether a candle with an inverted color is advancing with these patterns. The second group analyzes the trading profit significantly profitable transaction after various patterns. In this study, the three "open" prices of the First candles are considered as different possible stop points. These three different points have different consequences, so we have divided them into three different scenarios and discussed about them separately. Based on the small hypotheses in the end, it is concluded that there is no evidence to predict the strength of candlestick patterns in trend recognition.

Gökül (2021), in his work, focuses on the prediction of stock price using a technical analysis tool (candlestick chart). In this study, one-year stock price data of 30 companies, which are the leading companies in the individual sector, on the NSE (National Stock Exchange) are used to predict share prices. Example charts are from moneycontrol.com. In this study, the stock price is estimated using previous price changes in the stock market. Comparing these with a real price, it was found that how many were true and how many were false [16].

Ho and Huang (2021), on the other hand, proposed a multi-channel collaboration network by combining candlestick charts and social media data for stock trend forecasts. First, social media sentiment features were extracted from Twitter using the Natural Language Toolkit and sentiment analysis data. Next, the stock's historical time series data is converted into a candlestick chart to illuminate patterns in the stock's movement. Finally, the stock's sentiment features and candlestick chart are integrated to predict stock price action over 4-, 6-, 8-, and 10-day timeframes. We evaluated our model for five high-demand stocks (Apple, Tesla, IBM, Amazon, and Google) and found that our collaborative network achieved promising results, comparing favorably with single-network models using sentiment data or candlesticks alone [17].

Lin et al (2021), a simple eight-trigram feature engineering schema approach of inter-day candlestick patterns, a new approach to daily stock pattern prediction, which combines traditional candlestick charting with current artificial intelligence, is an ensemble machine learning framework. Various machine learning methods, including deep learning methods, are applied to stock data to predict the closing price direction [18].

Ananthi and Vijayakumar (2021) make predictions on the stock prices of any company determined by users in their proposed system for the next few days. It estimates the general condition of the stock by using estimated stock price and data sets obtained from different data regarding a particular stock. It is done by regression and candlestick pattern detection to make stock price prediction. In this proposed system, it generates signals to predict the market on the candlestick chart [19].

Aycel and Santur (2022), in their study, developed a new hybrid approach that can be used as an alternative to moving averages such as SMA, WMA and EMA used in the literature. Their performance was compared with the PSAR, BB, MACD, RSI and Stoch RSI indicators, which are widely used in the literature. The proposed approach has been validated by backtesting 30 stocks in BIST30 [20].

3. Material and Method

In this study, a software framework was created to develop algorithmic strategies supported by statistical pattern analysis. As shown in Figure 5, the proposed approach has four stages. In the first stage, the software framework was created with object-based coding for the 18 wax patterns used in the study. At this stage, object-oriented programming and factory design cover were used in order to increase the patterns to be used in the future with minimum code cost for future studies. In the second stage, one-hot encoding was done to determine the candle type. Thus, in the third stage, a data set labeled with candle types was created. In the fourth stage, which is the last stage, algorithmic buy/sell strategies based on patterns and/or statistical buy/sell strategies supported by candlestick patterns will be developed and portfolio return and buy&wait strategies will be compared.
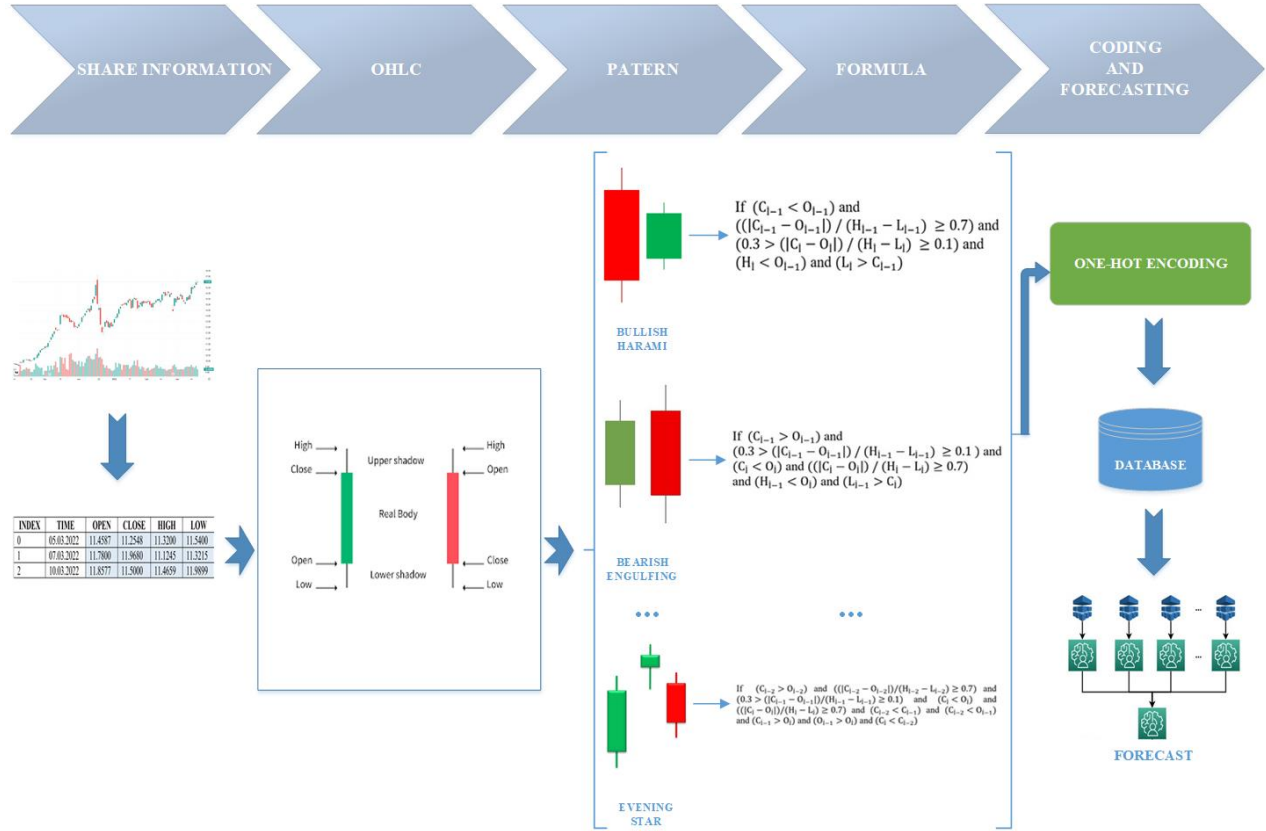
Figure 5. Proposed approach

3.1. Candlestick math

$$v_i = [O_i, C_i, H_i, L_i]$$

$$\Delta v_i = C_i - O_i$$

$$max_{v_i} = max(C_i, O_i)$$

$$min_{v_i} = min(C_i, O_i)$$

3.1.1. Doji

As in Figure 6a, candlesticks formed when the opening and closing prices are very close to each other or are exactly equal, are called doji. If the current stock has the same opening and closing prices, a doji occurs. Doji looks like a straight line in appearance. It is a term used to describe the equal power of buyers and sellers in the market. The formulation of the Doji candlestick pattern is shown below.

$$Doji(v_i) = \begin{cases} 1, & if & \left[\begin{matrix} \left(\dfrac{|\Delta v_i|}{H_i - L_i} < 0.1 \right) & and \\ \left(\dfrac{H_i - max_{v_i}}{3 * |\Delta v_i|} > 1 \right) & and \\ \left(\dfrac{min_{v_i} - L_i}{3 * |\Delta v_i|} > 1 \right) \end{matrix} \right] \\ 0, & otherwise \end{cases} \tag{1}$$

3.1.2. Doji star

As seen in Figure 6b, it is an undecided candle whose lower and upper wicks are equal and the opening and closing values are close to each other. Like the doji formation, the doji star formation is similar in appearance to a straight line.

The formulation of the Doji star candlestick pattern is shown below.

$$DojiStar(v_i, v_{i-1}) = \begin{cases} 1, & if \begin{bmatrix} C_{i-1} > O_{i-1} & and & \dfrac{|\Delta v_{i-1}|}{H_i - L_i} \geq 0.7 & and \\ \dfrac{|\Delta v_i|}{H_i - L_i} < 0.1 & and & C_{i-1} < C_i & and \\ C_{i-1} < O_i & and & \dfrac{H_i - max_{v_i}}{3 * |\Delta v_i|} > 1 & and \\ \dfrac{min_{v_i} - L_i}{3 * |\Delta v_i|} > 1 \end{bmatrix} \\ 0, & otherwise \end{cases}$$

(2)

a) Doji patern   b) Doji star patern
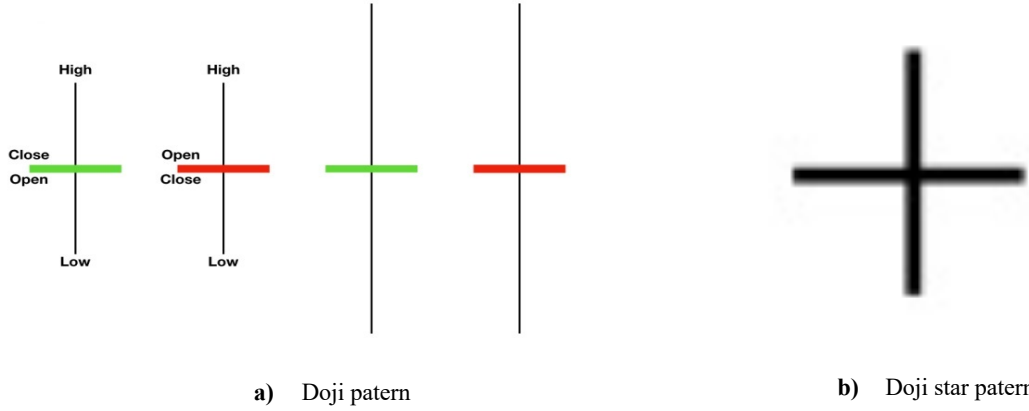
Figure 6. Doji and doji star candlestick pattern description

3.1.3. Dragonfly doji

The dragonfly doji, also known as the dragonfly doji, is a particular Doji where both the open and closed price is at or relatively close to the high of a particular day. It is a bearish or bullish (depending on context) candle with a long lower wick and open/close near the upper end as shown in Figure 7a.

The formulation of the Dragonfly doji candlestick pattern is shown below.

$$DragonFly(v_i) = \begin{cases} 1, & if \begin{bmatrix} \left(\dfrac{|\Delta v_i|}{H_i - L_i} < 0.1\right) & and \\ \left(\dfrac{H_i - max_{v_i}}{|\Delta v_i|} < 1\right) & and \\ \left(\dfrac{min_{v_i} - L_i}{3 * |\Delta v_i|} > 1\right) \end{bmatrix} \\ 0, & otherwise \end{cases}$$
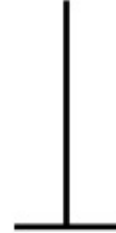
(3)

3.1.4. Gravestone doji

173

The Bears version of the Dragonfly Doji is the Gravestone Doji (Tombstone Doji). It is a bearish reversal candle with a long upper wick and an open/close near the lower end as shown in Figure 7b.

The formulation of the Gravestone doji candlestick pattern is shown below.

$$GraveStone(v_i) = \begin{cases} 1, & if & \left[\begin{array}{ll} \left(\dfrac{|\Delta v_i|}{H_i - L_i} < 0.1 \right) & and \\ \left(\dfrac{H_i - max_{v_i}}{3 * |\Delta v_i|} > 1 \right) & and \\ \left(\dfrac{min_{v_i} - L_i}{|\Delta v_i|} \leq 1 \right) \end{array} \right] \\ 0, & otherwise \end{cases} \tag{4}$$
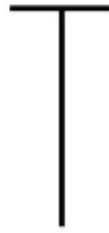
a) Dragonfly doji patern   **b)** Gravestone doji patern

Figure 7. Dragonfly doji and gravestone doji candle pattern description

3.1.5. Hammer

This pattern is formed at the bottom of a trend or during a downtrend. As it can be seen in Figure 8a, since it shows a bottom-beating uptrend, this formation is called a hammer. The upper shadow is almost absent. The lower shadow of the candlestick must be at least twice the length of the body.

The formulation of the Hammer candle mold is shown below.

$$Hammer(v_i) = \begin{cases} 1, & if & \left[\begin{array}{ll} \left(\dfrac{H_i - L_i}{3 * |\Delta v_i|} < 1 \right) & and \\ \left(\dfrac{C_i - L_i}{0.001 + H_i - L_i} > 0.6 \right) & and \\ \left(\dfrac{O_i - L_i}{0.001 + H_i - L_i} > 0.6 \right) \end{array} \right] \\ 0, & otherwise \end{cases} \tag{5}$$

3.1.6. Hanging man

Hanging Man acts as a reversal signal of an uptrend. As seen in Figure 8b, it has the same structure as the hammer. The difference between them is seen at the end of the hammer downtrend; hanging man is seen at the end of an uptrend. Both are a reversal candlestick pattern.

The formulation of the Hanging man candlestick is shown below.

$$\tag{6}$$

$$HangingMan(v_i, v_{i-1}, v_{i-2}) = \begin{cases} 1, & if \quad \begin{bmatrix} \left(\dfrac{H_i - L_i}{-4 * \Delta v_i} > 1\right) & and \\ \left(\dfrac{C_i - L_i}{0.001 + H_i - L_i} > 0.75\right) & and \\ \left(\dfrac{O_i - L_i}{0.001 + H_i - L_i} > 0.75\right) & and \\ (H_{i-1} < O_i) \; and & (H_{i-2} < O_i) \end{bmatrix} \\ 0, & otherwise \end{cases}$$

a) Hammer patern   **b)** Hanging man patern

Figure 8. Hammer and hanging man candle pattern description

3.1.7. Inverted hammer

The Inverted Hammer pattern is formed at the end of a downtrend and gives a bullish signal. Their bodies are below, their shadows are above, and the drop shadow is almost absent as in figure 9a.

The formulation of the Inverted hammer candle mold is shown below.

$$InvHammer(v_i) = \begin{cases} 1, & if \quad \begin{bmatrix} \left(\dfrac{H_i - L_i}{-3 * \Delta v_i} > 1\right) & and \\ \left(\dfrac{H_i - C_i}{0.001 + H_i - L_i} > 0.6\right) & and \\ \left(\dfrac{H_i - O_i}{0.001 + H_i - L_i} > 0.6\right) \end{bmatrix} \\ 0, & otherwise \end{cases} \tag{7}$$

3.1.8. Shooting star

The Shooting Star has the opposite view of the Hammer pattern. It is seen at the end of an uptrend. In the Shooting Star candle, the opening and closing are very close to each other, as in figure 9b. It is a candle formation at the end of an uptrend that will move the trend down.

The formulation of the Shooting star candlestick pattern is shown below.

$$ShootingStar(v_i, v_{i-1}) = \begin{cases} 1, & if \quad \begin{bmatrix} (O_{i-1} < C_{i-1} < O_i) & and \\ \left(\dfrac{H_i - max_{v_i}}{|\Delta v_i|} \geq 3\right) & and \\ \left(\dfrac{max_{v_i} - L_i}{|\Delta v_i|} \leq 1\right) \end{bmatrix} \\ 0, & otherwise \end{cases} \tag{8}$$

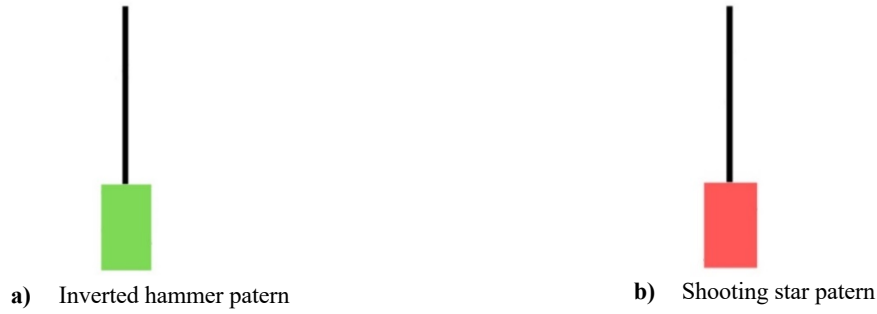**a)** Inverted hammer patern   **b)** Shooting star patern

Figure 9. Inverted hammer and shooting star candle pattern description

3.1.9. Bullish harami

The Bullish harami (Pregnant Bull) formation is formed by first a red body as in figure 10a, and a green body that stays within the boundaries of this red body and then emerges completely inside the red body. It occurs during a downtrend in the market. The second small green body formed indicates that the bears' strength is decreasing.

The formulation of the Bullish harami candle pattern is shown below.

$$BullishHarami(v_i, v_{i-1}) = \begin{cases} 1, & if & \begin{bmatrix} (C_{i-1} < O_{i-1}) & and & \left(\frac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.7\right) & and \\ \left(0.3 > \frac{|\Delta v_i|}{H_i - L_i} \geq 0.1\right) & and & (H_i < O_{i-1}) & and \\ (L_i > C_{i-1}) \end{bmatrix} \\ 0, & & otherwise \end{cases} \quad (9)$$

3.1.10. Bearish harami

The Bearish harami (Pregnant Bear) pattern consists of a first green body, as in figure 10b, and a red body that falls within the boundaries of this green body and then occurs. It occurs in an uptrend in the market. The formation basically resembles a pregnant woman.

The formulation of the Bearish harami candle pattern is shown below.

$$BearishHarami(v_i, v_{i-1}) = \begin{cases} 1, & if & \begin{bmatrix} (C_{i-1} > O_{i-1}) & and & \left(\frac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.7\right) & and \\ \left(0.3 > \frac{|\Delta v_i|}{H_i - L_i} \geq 0.1\right) & and & (H_i < C_{i-1}) & and \\ (L_i > O_{i-1}) \end{bmatrix} \\ 0, & & otherwise \end{cases} \quad (10)$$

a) Bullish harami patern   **b)** Bearish harami patern

Figure 10. Bullish harami and bearish harami candle pattern views

3.1.11. Bullish engulfing

The most important factor in the formation of Bullish engulfing (Bull Engulfing) formation is that the red candle formed on the first day remains inside the green candle formed on the second day, as in figure 11a. The formation of the formation in a falling or correcting market gives us a bearish signal in this formation.

The formulation of the Bullish engulfing candle pattern is shown below.

$$BullishEngulfing(v_i, v_{i-1}) = \begin{cases} 1, & if \begin{bmatrix} (C_{i-1} < O_{i-1}) & and & \left(\frac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and \\ \left(0.3 > \frac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.1\right) & and & (C_i > O_i) & and \\ (H_{i-1} < C_i) & and & (L_{i-1} > O_i) \end{bmatrix} \\ 0, & otherwise \end{cases} \quad (11)$$

3.1.12. Bearish engulfing

The Bearish engulfing pattern is the combination in which the body boundaries of the red candle are larger than the body boundaries of the green candle, as seen in figure 11b. The red candle is formed after the green and swallows the green completely. This indicates that the asset price has fallen more than its previous rise. Generally, this pattern indicates an uptrend end and a downtrend start.

The formulation of the Bearish engulfing candle pattern is shown below.

$$BearishEngulfing(v_i, v_{i-1}) = \begin{cases} 1, & if \begin{bmatrix} (C_{i-1} > O_{i-1}) & and & \left(\frac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and \\ \left(0.3 > \frac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.1\right) & and & (C_i < O_i) & and \\ (H_{i-1} < O_i) & and & (L_{i-1} > C_i) \end{bmatrix} \\ 0, & otherwise \end{cases} \quad (12)$$

a) Bullish engulfing patern   **b)** Bearish engulfing patern

Figure 11. Bullish engulfing and bearish engulfing candle pattern description

3.1.13. Dark cloud cover

The dark cloud cover pattern is a top return pattern. When the market is in an uptrend, a green candlestick is observed on the first day of the pattern as seen in figure 12a. The second day gaps up and opens higher, but closes lower than the middle of the body of the previous candlestick, creating a strong red candlestick. The second day does not close under the body of the first day.

The formulation of the dark cloud cover candle mold is shown below.

$$DarkCloudCover(v_i, v_{i-1}) = \begin{cases} 1, & if & \begin{bmatrix} (C_{i-1} > O_{i-1}) & and & \left(\dfrac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.7\right) & and \\ \left(\dfrac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and & (C_i < O_i) & and \\ (O_i \geq C_{i-1}) & and & \left(O_{i-1} < C_i < \dfrac{O_{i-1} + C_{i-1}}{2}\right) \end{bmatrix} \\ 0, & & otherwise \end{cases}$$ (13)

3.1.14. Piercing patern

Piercing Pattern (Piercing Candles) pattern has an opposite structure to the Dark Cloud Cover pattern. Piercing pattern is formed at the end of a downtrend. It is a bottom reversal pattern. As in Figure 12b, a red candlestick appears first. A green candlestick gapes down and closes in the upper half of the body of the first candlestick. The second candle does not close above the body of the first candle.
The formulation of the piercing pattern wax pattern is shown below.

$$DarkCloudCover(v_i, v_{i-1}) = \begin{cases} 1, & if & \begin{bmatrix} (C_{i-1} < O_{i-1}) & and & \left(\dfrac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.7\right) & and \\ \left(\dfrac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and & (C_i > O_i) & and \\ (O_i \leq C_{i-1}) & and & \left(C_i < \dfrac{O_{i-1} + C_{i-1}}{2}\right) \end{bmatrix} \\ 0, & & otherwise \end{cases}$$ (14)

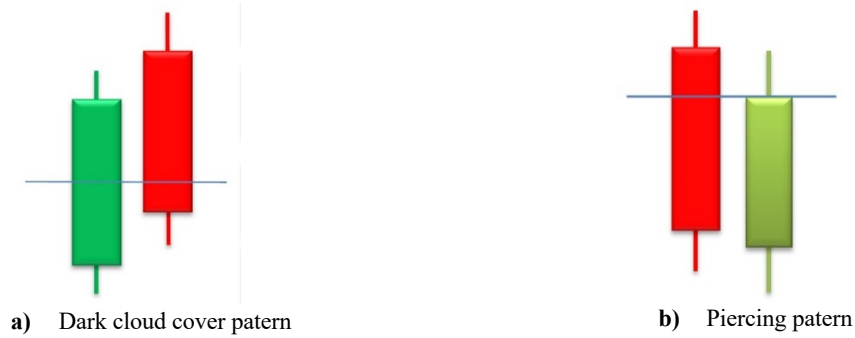**a)** Dark cloud cover patern   **b)** Piercing patern

Figure 12. Dark cloud cover and piercing wax pattern description

3.1.15. Evening star

Evening star is a candle formation that indicates the possible reversal of an uptrend. As seen in Figure 13a, a short candlestick comes after the first green candle. Finally, a red candlestick appears.
The formulation of the Evening star candlestick pattern is shown below.

$$EveningStar(v_i, v_{i-1}, v_{i-2}) = \begin{cases} 1, & if & \begin{bmatrix} (C_{i-2} > O_{i-2}) & and & \left(\dfrac{|\Delta v_{i-2}|}{H_{i-2} - L_{i-2}} \geq 0.7\right) & and \\ \left(0.3 > \dfrac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.1\right) & and & (C_i < O_i) & and \\ \left(\dfrac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and & (C_{i-2} < C_{i-1}) & and \\ (C_{i-2} < O_{i-1}) & and & (C_{i-1} > O_i) & and \\ (O_{i-1} > O_i) & and & (C_i < C_{i-2}) \end{bmatrix} \\ 0, & & otherwise \end{cases}$$ (15)

178

3.1.16. Evening doji star

Evening doji star means the end of the uptrend and is a sell signal for traders trading in the forex market. It consists of a doji following a green bullish candle as in Figure 13b. The doji following the green candle is formed by creating a gap in the upward direction. On the third day, a red bearish candle is formed, which is inside the green bullish candle formed on the first day.

The formulation of the Evening doji star candlestick pattern is shown below.

$$
EveningDojiStar(v_i, v_{i-1}, v_{i-2}) = \begin{cases} 1, & if \begin{bmatrix} (C_{i-2} > O_{i-2}) & and & \left(\dfrac{|\Delta v_{i-2}|}{H_{i-2} - L_{i-2}} \geq 0.7\right) & and \\ \left(\dfrac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} < 0.1\right) & and & \left(\dfrac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and \\ (C_{i-2} < C_{i-1}) & and & (C_{i-2} < O_{i-1}) & and \\ (C_{i-1} > O_i) & and & (O_{i-1} > O_i) & and \\ \left(\dfrac{H_{i-1} - max_{v_{i-1}}}{3 * |\Delta v_{i-1}|} > 1\right) & and & \left(\dfrac{min_{v_{i-1}} - L_{i-1}}{3 * |\Delta v_{i-1}|} > 1\right) & and \\ (C_i < C_{i-2}) & and \\ \end{bmatrix} \\ 0, & otherwise \end{cases}
$$

(16)

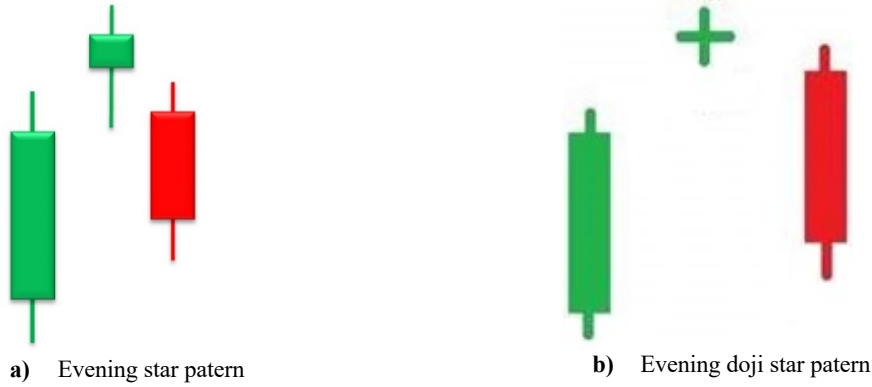**a)** Evening star patern   **b)** Evening doji star patern

Figure 13. Evening star and evening doji star candle pattern description

3.1.17. Morning star

This candle formation is the downtrend counterpart of the Evening Star candlestick formation in an uptrend. For the Morning Star formation to occur, the characteristics of the three candles should be as follows; The first candle should be a long bearish or bearish candle, the second candle should be a small body bullish or doji that closes below the first candle, and the third candle should be a bullish or bullish candle that opens above the second candle.

The formulation of the Morning star candlestick pattern is shown below.

$$
MorningStar(v_i, v_{i-1}, v_{i-2}) = \begin{cases} 1, & if \begin{bmatrix} (C_{i-2} > O_{i-2}) & and & \left(\dfrac{|\Delta v_{i-2}|}{H_{i-2} - L_{i-2}} \geq 0.7\right) & and \\ \left(0.3 > \dfrac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.1\right) & and & (C_i > O_i) & and \\ \left(\dfrac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and & (C_{i-2} > C_{i-1}) & and \\ (C_{i-2} > O_{i-1}) & and & (C_{i-1} < O_i) & and \\ (O_{i-1} < O_i) & and & (C_i > C_{i-2}) \\ \end{bmatrix} \\ 0, & otherwise \end{cases}
$$

(17)

3.1.18. Morning doji star

The Morning Doji Star is a bullish reversal pattern seen in a downtrend very similar to the Morning Star. The only difference is that the Morning Doji Star must have a doji candle in the second line. As can be seen in Figure 14b, it is an inverse pattern of the Evening Doji Star pattern. The first candle is a downtrend candle and the second candle is a doji below the body of the first candle. The last candle should close between the body of the first candle.

The formulation of the Morning doji star candlestick pattern is shown below.

$$MorningDojiStar(v_i, v_{i-1}, v_{i-2}) = \begin{cases} 1, & if \begin{bmatrix} (C_{i-2} < O_{i-2}) & and & \left(\frac{|\Delta v_{i-2}|}{H_{i-2} - L_{i-2}} \geq 0.7\right) & and \\ \left(\frac{|\Delta v_{i-1}|}{H_{i-1} - L_{i-1}} \geq 0.1\right) & and & (C_i > O_i) & and \\ \left(\frac{|\Delta v_i|}{H_i - L_i} \geq 0.7\right) & and & (C_{i-2} > C_{i-1}) & and \\ (C_{i-2} > O_{i-1}) & and & (C_{i-1} < O_i) & and \\ (O_{i-1} < O_i) & and & (C_i > C_{i-2}) & and \\ \left(\frac{H_{i-1} - max_{v_{i-1}}}{3 * |\Delta v_{i-1}|} > 1\right) & and & \left(\frac{min_{v_{i-1}} - L_{i-1}}{3 * |\Delta v_{i-1}|} > 1\right) \end{bmatrix} \\ 0, & otherwise \end{cases} \quad (18)$$

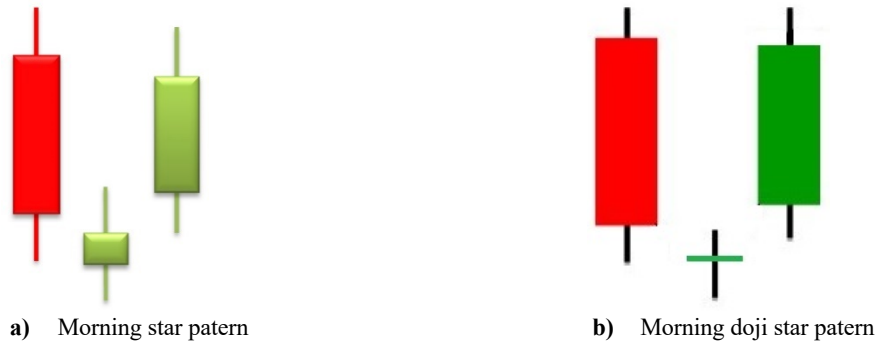a)   Morning star patern   b)   Morning doji star patern

Figure 14. Morning star and morning doji star candlestick patterns

4. Experimental Results

To validate the proposed approach, gold, brent oil, bist100 index, bitcoin (BTC), and usd were used. In the experimental studies, the first stage involves labeling the trend (1: increase, 0: decrease) and the candle pattern type of the daily bar with one hot encoding, using the daily closing values of the relevant financial assets. At this stage, the raw data set is transformed and used as in Figure 15.
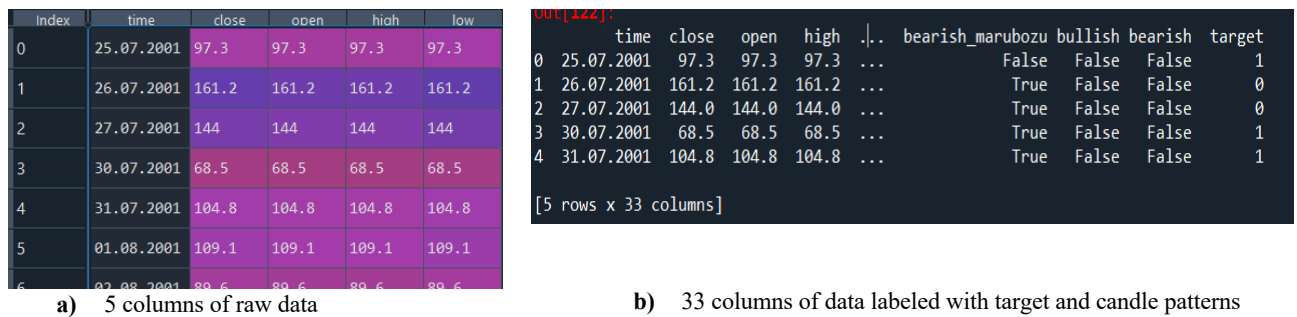

a)   5 columns of raw data   b)   33 columns of data labeled with target and candle patterns

Figure 15. Preparation of the data set

In the second stage, the data set was divided into training and testing and trained with the XGBoost algorithm. In the training phase, the system learns the rise and fall patterns by using the labeled data set. The XGBoost algorithm used at this stage is a learning algorithm that uses collective learning and is widely used in data science problems. Written in C++ language, XGBoost C++ can work with high performance on big data such as financial data because it works with the parallel programming principle, and it is more convenient than other collective learning methods against overfitting/underfitting problems. The tree-based XGBoost algorithm is typically as in Figure 16.
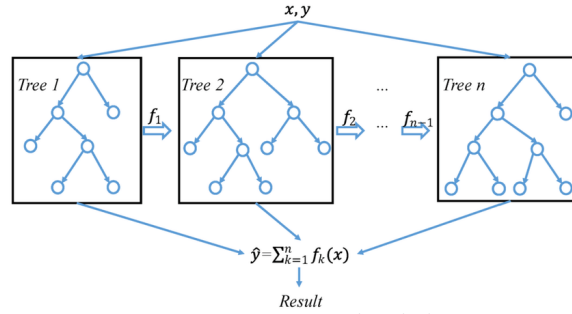

Figure 16. XGBoost description

Table 1. Hyperparameters selected for XGBoost

MAX_DEPTH	ETA	ERROR FUNCTION	GAMMA	ESTIMATORS
5	0.1	rmse	0.05	100

In the test phase, the bars in this period that were not used before were subjected to Backtest using the previously learned system. Backtest is expressed as making buy/sell transactions according to the rules determined in manual or algorithmic transactions in financial markets. Backtesting typically calculates PF (Profit Factor - coefficient of gain) by obtaining GF (Gross Profit – Gross Profit), GL (Gross Loses – Gross Loss) values on historical data (19). In the ideal Backtest, it aims to have a PF value greater than 1, a low number of transactions (for paying less commission), a low slippage value (price difference between trading), and a linear yield curve instead of a fluctuating one. In this study, commission and slippage are ignored since a robot that trades with daily bars and stays in the trend direction is considered instead of a high-frequency algorithmic robot.

$$PF = {GP}/{GL} \tag{19}$$

In all charts, the GP yield curve is shown in green and the GL curve in red.

4.1. Experimental studies for Bist100 index, BTC, gold, Dollar and Brent oil

In recent years, there has been an increase in the number of approaches based on machine learning for popular financial assets such as the Bist100 index, BTC, gram gold, dollar, and brent oil [21-29].

The Bist100 index represents the 100 most valuable companies in the Borsa Istanbul index. The data set obtained for Bist100 consists of 4792 pieces of data covering daily bars in the range of 11.04.2003 – 10.05.2022. When the training and testing ratio is selected as 50%, PF=1.43 as seen in Figure 17a, 162 of the transactions resulted a profit, and 74 of them resulted in loss. An approach that can be used to reduce the amount of loss is the stop loss strategy (stop loss when the threshold value is reached) when entering the position, but the main motivation of this study is to prove the performance of algorithmic transactions that will only be performed with pattern recognition, so stop loss and technical / fundamental indicators are not used.

The data set for btc, the most popular blockchain-based cryptocurrency, was used between 18.07.2010 – 10.05.2022 with 4315-day bar closing. Since the data set is relatively small, the training:test partitioning was used as 0.85:0.15. As can be seen in Figure 17b, PF=35 could be obtained in experimental studies, 143 of the transactions were closed with profit and only 13 with loss. With the high PF value obtained, commission and slippage have also become insignificant. It is also effective that cryptocurrencies are in a bullish trend and are very popular in the high PF value and linear yield curve. Cohen (2021), the PF value in his study is much lower than

this study, but in addition to the difference in the approach, it differs from each other in the selected periods for training and testing. In this context, no one-to-one comparison has been made [30].

For gram gold, one of the financial assets that are considered an important investment tool and can be considered the least volatile, 2780-day bar closing data between 11.04.2013 and 10.05.2022 was used, and the data set was divided equally. As can be seen in Figure 17c, PF= 33.44 was obtained, and only 37 lossy transactions were realized against 316 profitable transactions.

The 5000-day bar closing between 11.04.2000 and 13.06.2019 was used for the dollar, which is one of the important investment instruments in our country and can be extremely volatile from time to time due to the socio-economic structure. PF=12.1 was obtained when a short test period was selected. As can be seen in Figure 17d, when training:test is divided as 50% in experimental studies, more than 90% profitable transactions are obtained in the Backtest phase.

Sudden changes in brent oil, which is an important indicator in international markets, is one of the most important commodities that has the power to deeply affect especially emerging markets. In the study, the data set was chosen as 600-day bar closing between 13.04.2020 – 10.05.2022. When the training and testing rate was chosen as 50%, PF=2.01 as seen in figure 17e. Despite 28 profitable transactions, 16 lossy transactions were carried out.
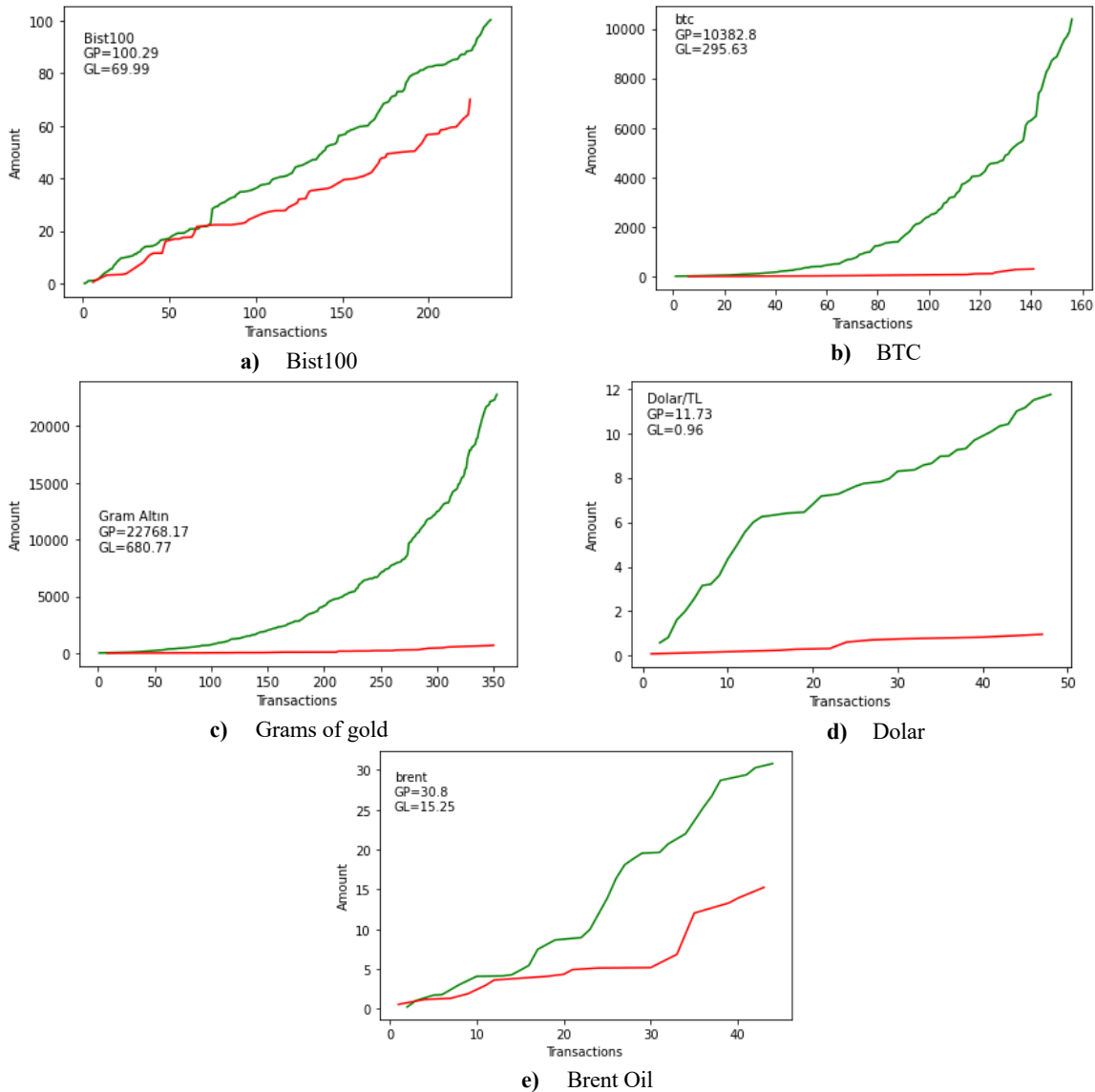

a) Bist100

b) BTC

c) Grams of gold

d) Dolar

e) Brent Oil

Figure 17. GP and GL values of datasets

5. Discussion, Conclusion and Future Works

This study, which was carried out to identify only candlestick patterns for both the bist100 index and global financial assets, revealed the success of an algorithmic trading strategy based solely on candlestick charts. Candlestick charts can be used in almost any period, from minute to daily bars. Although only 21 of the most popular patterns were used in the proposed study, the earnings coefficient could be obtained above 1 in all major financial assets and the profit/loss ratio could be above 90%.

It is aimed to expand this study in two main ways in the future and thus to increase its originality. The first of these is to increase the number of candlestick patterns. In fact, only 21 of the patterns known to be over 100 were used and typically 40-60% of all bars were recognized in the study. Despite this, the results obtained are quite satisfactory, and it is predicted that the performance will be increased by increasing the number of patterns. In this study, only buy/sell strategies based on candlestick patterns were produced. In time series, the number and quality of input attributes can be increased by obtaining technical indicators, basic ratios of assets, visual formations, and trends in different time periods. In this context, the approach proposed in this study and the other approaches mentioned will be combined in a hybrid way. Thus, when commission and slippage losses are combined with take profit and loss situation strategies, a more effective algorithmic robot system will be brought to the literature.

Acknowledgement

References

[1] Filiz, E., Karaboğa, H. A., 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.

[2] Pabuçcu, H. (2019). Borsa endeksi hareketlerinin tahmini: trend belirleyici veri, Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, 22(1), 246-256.

[3] Şişmanoğlu, G., Koçer, F., Önde, M. A., Sahingöz, O. K. (2020). Derin Öğrenme yöntemleri ile borsada fiyat tahmini, Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 9(1), 434-445.

[4] Santur, Y. Deep learning based regression approach for algorithmic stock trading: A case study of the Bist30, Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 10(4), 1195-1211.

[5] Budak, C. (2019). Teknik analiz indikatörlerinin performans karşılaştırması üzerine bir araştırma (Doctoral dissertation), Marmara Universitesi

[6] Bozkurt, Y. (2021). Piyasa performans oranlarına göre oluşturulmuş portföylerin getiri oranlarının değerlendirilmesi: BİST 100 endeksi firmaları üzerine bir uygulama (Master's thesis), Aydın Adnan Menderes Üniversitesi, Sosyal Bilimler Enstitüsü.

[7] Madbouly, M. M., Elkholy, M., Gharib, Y. M., Darwish, S. M. (2020, April). Predicting stock market trends for japanese candlestick using cloud model, In The International Conference on Artificial Intelligence and Computer Vision (pp. 628-645), Springer, Cham.

[8] Kusuma, R. M. I., Ho, T. T., Kao, W. C., Ou, Y. Y., Hua, K. L. (2019). Using deep learning neural networks and candlestick chart representation to predict stock market, arXiv preprint arXiv:1903.12258.

[9] Hung, C. C., Chen, Y. J. (2021). DPP: Deep predictor for price movement from candlestick charts, Plos one, 16(6), e0252404.

[10] Sadeghi, M., Farid, D. (2021). Investigating candlestick patterns using fuzzy logic in the stock trading system, Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(13), 7786-7806.

[11] Yee, L. L., Mei, H. L., Isharuddin, L. (2021). Ichimoku cloud and japanese candlestick prediction combination pattern approached: the case study of malaysia stock market, Multidisciplinary Applied Research and Innovation, 2(2), 190-196.

[12] Lin, Y., Liu, S., Yang, H., Wu, H., Jiang, B. (2021). Improving stock trading decisions based on pattern recognition using machine learning technology, PloS one, 16(8), e0255558.

[13] Ardiyanti, N. P. W., Palupi, I., Indwiarti, I. (2021). Trading strategy on market stock by analyzing candlestick pattern using artificial neural network (ann) method, Jurnal Media Informatika Budidarma, 5(4), 1273-1282.

[14] Chen, J. H., Tsai, Y. C. (2020). Encoding candlesticks as images for pattern classification using convolutional neural networks, Financial Innovation, 6(1), 1-19.

[15] Yassini, S. B., Rahnamay Roodposhti, F., Fallahshams, M. (2019). Analyzing the effectiveness of candlestick technical trading strategies in foreign exchange market, International Journal of Finance & Managerial Accounting, 4(15), 25-41.

[16] Gökül, U. (2021). Forecast share price using technical analysis tool, Pacific International Journal, 4(1), 01-06.

[17] Ho, T. T., Huang, Y. (2021). Stock price movement prediction using sentiment analysis and candlestick chart representation, Sensors, 21(23), 7957.

[18] Lin, Y., Liu, S., Yang, H., Wu, H. (2021). Stock trend prediction using candlestick charting and ensemble machine learning techniques with a novelty feature engineering scheme, IEEE Access, 9, 101433-101446.

[19] Ananthi, M., Vijayakumar, K. (2021). Stock market analysis using candlestick regression and market trend prediction (CKRM), Journal of Ambient Intelligence and Humanized Computing, 12(5), 4819-4826.

[20] Aycel, Ü., Santur, Y. (2022). A new moving average approach to predict the direction of stock movements in algorithmic trading, Journal of New Results in Science, 11(1), 13-25.

[21] Kaynar, T., Yiğit, Ö. E. (2021). Öznitelik mühendisliği ile makine öğrenmesi yöntemleri kullanılarak bıst 100 endeksi değişiminin tahminine yönelik bir yaklaşım, Yaşar Üniversitesi E-Dergisi, 16(64), 1741-1762.

[22] Koç, Y. (2021). Makine öğrenmesi ile çok terimli hisse senedi yönlü tahmini; BIST100 örneği/Multinomial direction forecast with machine learning algorithms; BIST100 example (Doctoral dissertation), Kadir Has Üniversitesi.

[23] Santur, Y. (2022). Candlestick chart based trading system using ensemble learning for financial assets, Sigma Journal of Engineering and Natural Sciences, 40(2), 370-379.

[24] Aksoy, B. (2021). Pay senedi fiyat yönünün makine öğrenmesi yöntemleri ile tahmini: Borsa İstanbul örneği, Business and Economics Research Journal, 12(1), 89-110.

[25] Akdağ, M., Bozma, G. (2021). Stok akış modeli ve facebook prophet algoritması ile bitcoin fiyatı tahmini/Prediction of bitcoin price with stock to flow model and facebook prophet algorithm, Uluslararası Ekonomi İşletme ve Politika Dergisi, 5(1), 16-30.

[26] Demirel, A. C., Hazar, A. (2021). Kripto para değerlerine dayanılarak bist 100 endeks hareketi tahmininde destek vektör makineleri uygulaması, Başkent Üniversitesi Ticari Bilimler Fakültesi Dergisi, 5(1), 27-35.

[27] Altunbaş, C. (2021). Derin öğrenme ile hisse senedi piyasası tahmini (Master's thesis), Aydın Adnan Menderes Üniversitesi Sosyal Bilimler Enstitüsü.

[28] Ustali, N. K., Tosun, N., Tosun, Ö. (2021). Makine öğrenmesi teknikleri ile hisse senedi fiyat tahmini, Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi, 16(1), 1-16.

[29] Tanışman, S., Karcıoğlu, A. A., Aybars, U. G. U. R., Bulut, H. (2021). LSTM sinir ağı ve arıma zaman serisi modelleri kullanılarak bitcoin fiyatının tahminlenmesi ve yöntemlerin karşılaştırılması, Avrupa Bilim ve Teknoloji Dergisi, (32), 514-520.

[30] Cohen, G. (2021). Optimizing candlesticks patterns for Bitcoin's trading systems, Review of Quantitative Finance and Accounting, 57(3), 1155-1167.