Cryptocurrency returns prediction using candlestick patterns analysis and multi-layer deep LSTM neural networks

Mohammad VAHIDPOUR[1], Amir DANESHVAR[2*], Mohsen AMINI KHOUZANI[3], Mahdi HOMAYOUNFAR[4]

[1] Department of Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Department of Industrial Management, Electronic Branch, Islamic Azad University, Tehran, Iran

[3] Department of Financial Engineering, Shahr-e-Qods Branch, Islamic Azad University, Shahr-e-Qods, Tehran, Iran

[4] Department of Industrial Management, Rasht Branch, Islamic Azad University, Rasht, Iran

mohammad.vahidpour@srbiau.ac.ir, a_daneshvar@iauec.ac.ir[*], amini_k_m@yahoo.com, homayounfar@iaurasht.ac.ir

[*]**Corresponding author:** Amir DANESHVAR

a_daneshvar@iauec.ac.ir[*]

Abstract: Financial markets are characterised by their dynamic, non-linear, and fluctuating nature. Analysing financial time series in these contexts is a complex and challenging task. Candlestick patterns are recognised as among the most widely used financial tools and offer invaluable insights into market sentiment and psychology. However, manual analysis of these patterns presents significant challenges. Therefore, leveraging machine learning methods becomes a necessity for overcoming these challenges. In this study, a four-step framework was introduced in which the data preparation process is executed on the price data of the 20 cryptocurrencies. Forty-eight candlestick patterns were extracted alongside returns. Employing the long short-term memory (LSTM) neural network, structured with multiple layers, each specialising in a specific cryptocurrency, enables individualised prediction of market returns. Evaluation of model accuracy and sensitivity is conducted via the confusion matrix, and two distinct trading strategies assess the capital portfolio. The research findings underscore the profitability of the proposed model across all scenarios. Candlestick patterns serve as powerful tools for understanding market sentiments and identifying shifts in market trends. However, their standalone efficacy is limited. Integrating them with other technical analysis tools facilitates more informed decision-making and fosters a deeper understanding of market dynamics.

Keywords: Financial, Cryptocurrency, Candlestick, Long Short-Term Memory, Machine Learning.

1. Introduction

Financial markets experience price changes driven by economic forces, innovation, technology, competition, policy, regulations, news events, and more. These factors ultimately affect the balance of supply and demand. A financial time series embodies a chronological array of data intricately linked to the fiscal performance of a specific market. The purpose of financial time series analysis is for modelling to better understand market behavior and forecast financial data and future trends. Financial data have a complex and chaotic structure that is caused by the dynamics and changes of the market, so their analysis is a challenging task that has attracted the attention of researchers. Approaches to market analysis primarily adhere to two schools of thought. The fundamental analysis methodology hinges upon economic, environmental, political, and financial determinants. Conversely, the technical approach leans on statistics, leveraging past price movements and trading volumes as foundational pillars. Within financial markets, prices oscillate between upward, downward, or neutral trajectories. The crux of technical analysis rests upon three cardinal principles: i- Inclusion of all factors in price dynamics; ii- Price movements dictated by prevailing trends; iii- The recurrence of historical patterns (Nazário et al., 2017).

One of the most widely used tools for illustrating price movements in financial markets are candlestick patterns. Originating circa 1750, Munehisa Homma pioneered these patterns to dissect the Japanese rice market. The interplay of buyers' and sellers' behaviours and actions instigates shifts

https://doi.org/10.33436/v34i1y202410

in the dynamics of supply and demand, which can be visually encapsulated within candlestick patterns (Nison, 2001). These configurations adeptly capture key price metrics such as highs, lows, openings, and closings, thereby affording investors insights into market sentiments and psychology. Consequently, they serve as potent tools for short-term price prognostications and future milestones (Tharavanij et al., 2017).

Over a hundred distinct candlestick patterns exist, but the human ability to identify and interpret them all is limited, particularly due to factors like fatigue and information overload. To address these limitations, machine learning techniques, particularly deep neural networks, offer promising solutions. Deep neural networks, a type of machine learning with layered information processing, can decipher complex patterns with enhanced accuracy. This makes them invaluable in financial time series analysis, where they excel at predicting prices and future market trends with greater precision, operating at high speeds, adapting to new data, and facilitating automatic learning (Gers et al., 2000).

In recent years, cryptocurrencies, a type of digital currency using encryption technology, have captured significant attention. A wide variety of cryptocurrencies have emerged, each boasting distinct features, technologies, and objectives. Attributes such as security, transparency, speed, low fees, and global accessibility have propelled cryptocurrencies into the mainstream as innovative tools within financial systems (Morris, 2006). As the cryptocurrency markets undergo continuous evolution, the integration of advanced analytics techniques holds significant potential to augment trading strategies and decision-making processes within this swiftly changing domain.

The objective of this study is to address the existing gap in the literature concerning candlestick analysis within the cryptocurrency market. To achieve this aim, the study seeks to amalgamate and enhance existing methods through the utilisation of machine learning techniques. To this end, a four-step architecture is proposed, involving the collection and processing of data from various cryptocurrencies across multiple stages. Subsequently, different candlestick and price patterns are extracted and utilised as input for the proposed Long Short-Term Memory (LSTM) neural network. The primary goal of this proposed model is to offer solutions aimed at improving decision-making and enhancing the efficiency and profitability of portfolios through the application of deep neural network methodologies.

This study employs a hybrid approach that integrates candlestick patterns with other technical parameters in the analysis of financial markets, particularly cryptocurrency markets. This method utilises a model that combines comprehensive and diverse information, ultimately improving market trend prediction. Incorporating cutting-edge artificial intelligence techniques and neural networks for analysing extensive datasets of cryptocurrencies enhances the accuracy, reliability, and efficiency of the predictive model. This innovative hybrid approach, which simultaneously exploits candlestick patterns and LSTM deep neural networks, enables the aggregation of more extensive information and the identification of complex and concealed patterns within financial markets. Additionally, adopting a multi-layered architecture establishes a dynamic and more dependable model adaptable to varying asset numbers, effectively leveraging market information.

The study structure is outlined as follows: Section 2 provides the background of the study and reviews previous research on utilising candlestick patterns for trend forecasting. Section 3 introduces the concepts and literature pertinent to the research, delineating various types of candlestick patterns and their implications. Section 4 presents the proposed model in four stages. Section 5 elaborates on the dataset, research scope, model settings, and neural network employed. Section 6 articulates the research findings, and in Section 7 the conclusions of the study and delineates avenues for future research.

2. Related works

The following describes studies related to the analysis of candlestick charts and price patterns. In a study (Chootong & Sornil, 2012), introduces a trading strategy that incorporates price action patterns, candlestick patterns, and a variety of technical trading indicators such as moving averages, exponential moving averages, Bollinger bands, and trading volume. The dataset comprised shares from the Thai stock market. The primary objective of this research was to amalgamate diverse patterns to enhance investment returns by offering more comprehensive and precise trading signals. In another study (Lu et al., 2012), explored the profitability of candlestick patterns in the Taiwan stock market. The research approach involved buying during Bullish (Bearish) patterns and holding until the occurrence of Bearish (Bullish) patterns. The findings revealed that bullish reversal patterns are statistically significant and yield profits. Additionally, the study underscored that the profitability of bullish patterns varies across different market conditions.

The following studies related to technical analysis and candlestick patterns in predicting market movements. In a study (Jain et al., 2022) examines technical analysis in cryptocurrencies and its effectiveness in using candlestick charts and technical indicators, concluding that technical analysis can lead to positive returns. Additionally, the results indicate that relying solely on candlestick patterns may not be efficient, and it is essential to use other indicators to examine trends alongside candlestick patterns. A study (Santur, 2022) delved into technical analysis and the application of candlestick charts for market trend prediction. A software framework employing 24 candlestick patterns was developed to forecast trend directions. Results indicated that employing a strategy centred on recognising candlestick patterns and taking suitable actions aligned with the trend direction resulted in higher profits. The research achieved an accuracy rate of 53.8% in trend direction detection. In a study (Tharavanij et al., 2017), the effectiveness of candlestick patterns is examined and tested through two distinct strategies. The dataset comprises the top 50 shares in the Thai market, with eight candlestick patterns analysed in the research. The methodology employed in this study incorporates various statistical tests such as adjusted t-tests and binomial tests. A study (Prado et al., 2013), examined the efficacy of candlestick patterns in forecasting stock behavior in the Brazilian market. They investigated 16 patterns, analysing their identification, approval percentage in subsequent periods, and average profit or loss per pattern. While the findings revealed a notable association between certain candlestick patterns and market trends in Brazil, the overall predictive capacity of the patterns was found to be limited. Another study (Caginalp & Laurent, 1998) delves into the predictive capabilities of candlestick patterns using S&P 500 data spanning from 1992 to 1996. Eight candlestick patterns were employed in this investigation. Statistical analyses were conducted on two datasets, employing non-parametric criteria and statistical validation to assess the predictive efficacy of these models. The research findings revealed that candlestick patterns demonstrate predictive potential.

Next, studies regarding the utilization of machine learning algorithms and deep neural networks for predicting market movements are outlined. In a study (Aycel & Santur, 2022), the exploration revolves around defining and identifying candlestick patterns and their application in learning algorithms for predicting price direction and strength. The study analysed daily data from Bitcoin, Dollar, Gold, Brent Oil, and Bist100, incorporating 21 candlestick patterns. The findings suggest that integrating candlestick patterns into algorithmic development yields profitability by leveraging historical price data to enhance trading outcomes. In a study (Karmelia et al., 2022), machine learning techniques and deep learning methods are applied to address financial data classification issues, particularly focusing on candlestick patterns. The study utilises ten candlestick patterns and employs FNN (Feedforward Neural Networks) with sampling techniques. Data from five shares in the LQ45 IDX index are utilised, and the confusion matrix serves as a performance evaluation tool. Various machine algorithm methods, including recurrent neural networks, logistic regression, support vector machines, and others, are employed to train classification models .A study (Kusuma et al., 2019) focuses on leveraging deep learning neural networks and candlestick charts for market movement prediction. Historical stock market data is transformed into candlestick data, and a convolutional neural network model is utilised to analyse patterns and forecast price movements. Ultimately, a web-based system is developed. The primary objective is to demonstrate the efficacy of employing deep learning techniques and candlestick charts for market forecasting. In a study (Goswami et al., 2009), a hybrid methodology incorporating candlestick analysis is employed to forecast short-term market price volatility. This model aims to recognise lucrative candlestick patterns using SOM and SBR techniques to detect price fluctuations. The study's findings indicate that utilising candlestick patterns as a technical indicator can prove beneficial for short-term forecasting purposes.

Reviewed articles emphasise the importance of candlestick patterns in technical analysis. These patterns offer valuable insights into market sentiment, price trends, and supply and demand dynamics, aiding traders in mitigating risks. Research suggests that candlestick patterns can be profitable for forecasting cryptocurrency prices. However, their effectiveness might be limited if used alone. Integrating them with other technical indicators or leveraging machine learning algorithms can significantly enhance prediction accuracy. Studies exploring deep learning alongside candlestick patterns demonstrate this potential. Overall, candlestick pattern analysis remains a valuable resource for cryptocurrency traders, potentially optimising the performance of their trading and investment systems.

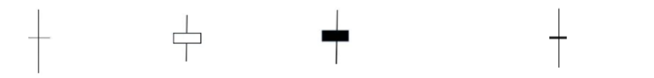
3. Candlestick patterns

Using candlestick pattern analysis allows traders to improve their strategy. A candlestick chart usually consists of bodies and shadows. The body of the candlestick represents the difference between the open and closed price. A full body indicates the closing price is lower than the open price, while an empty body indicates the closing price is higher than the open price. Shadows, thin lines at the top and bottom of the body, indicate the highest and lowest prices. Candlestick patterns are generally divided into two main categories: trend continuation patterns, indicating the continuation of the price direction, and reversal patterns, indicating a change in the price direction. Price trends are typically divided into three states: bullish, where prices are rising; bearish, where prices are falling; and neutral or stagnant, where there is no significant change in price direction (Xie et al., 2021). Table 1 describes 48 famous patterns along with their shape.

Table 1. Candlestick patterns (Avasaram, 2012; Morris, 2006; Prado et al., 2013)

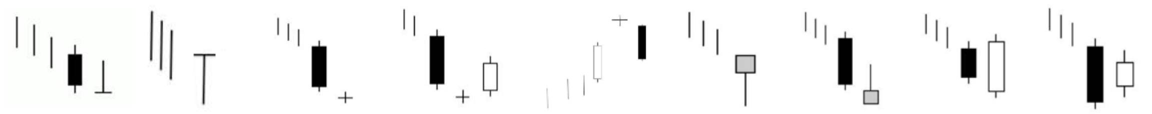
Neutral candlestick patterns

Doji	Spinning Doji	Long-Legged Doji

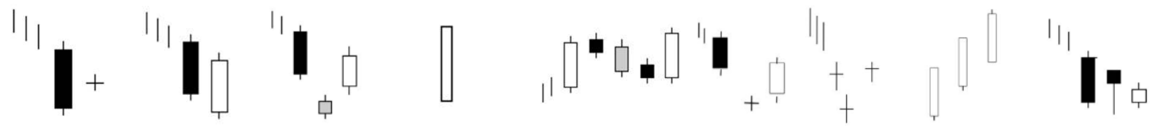


Bullish candlestick patterns

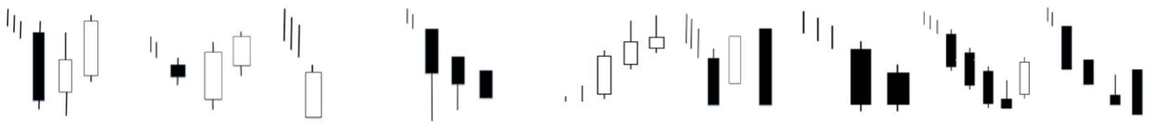
Gravestone Doji	Dragonfly Doji	Doji Star	Morning Doji Star	Evening Doji Star	Hammer	Inverted Hammer	Engulfing Pattern	Harami Pattern



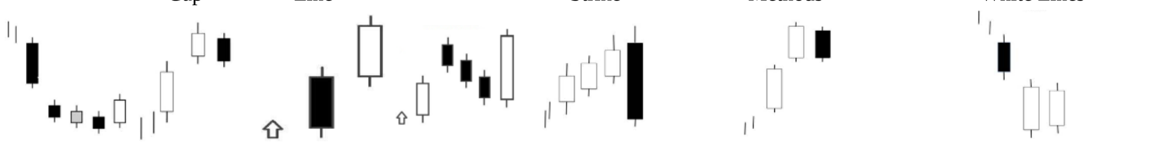
Harami Cross Pattern	Piercing Pattern	Morning Star	Marubozu	Rising Three Methods	Abandone d Baby	Tristar Pattern	Three Advancing White Soldier	Unique Three River



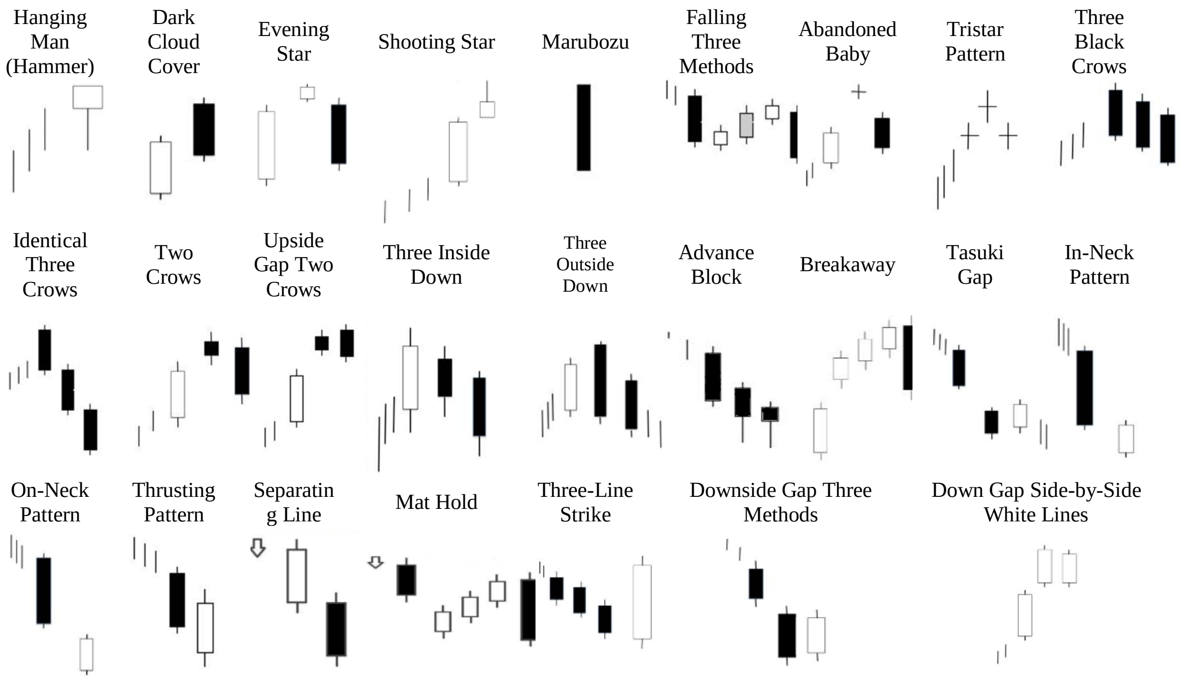
Three Inside Up	Three Outside Up	Belt-Hold	Three Stars In The South	Advance Block	Stick Sandwich	Matching Low	Ladder Bottom	Concealing Baby Swallow



Breakaway	Tasuki Gap	Separating Line	Mat Hold	Three-Line Strike	Upside Gap Three Methods	Up Gap Side-by-Side White Lines



Bearish candlestick patterns



4. Proposed method

In the following, the proposed model is described in four stages.

Stage 1: dataset preparation

The dataset preparation steps are illustrated schematically in Figure 1. Initially, asset prices are obtained via API at regular intervals, specifically in hourly and daily time frames, and stored in the MYSQL database. Subsequently, the data stored in the database undergoes processing through two distinct paths. In the first path, the simple returns are computed for the designated time period. The formula in Eq. (1) captures the percentage change in price between two time periods, offering insights into the growth or decline of an investment.

$$r_t = \frac{p_t}{p_{t-1}} - 1 \tag{1}$$

In the second path, various types of candlestick patterns are identified and extracted. Following this, the data from both paths are merged, thereby preparing the dataset for the specified number of assets under examination. The dataset is then partitioned into two segments: the training portion is utilised for neural network training, while the test portion is employed to assess the proposed model against various criteria.

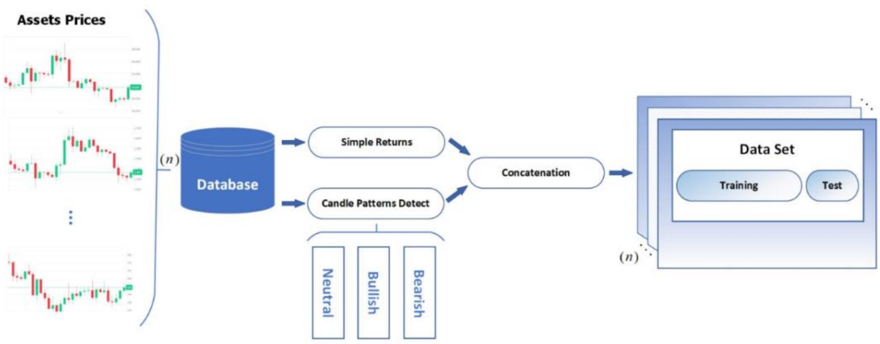


Figure 1. Schematic of data preparation

In the proposed model, the return value for each cryptocurrency is calculated and considered in the period T which is expressed in Eq. (2). In this study, T is set to 30, representing the size of the sliding window, which aims to capture the recent price changes in the model's analysis.

$$Returns = \{ r_t \quad r_{t-1} \quad \cdots \quad r_{t-T} \} \tag{2}$$

Furthermore, pre-processing is conducted on the open, closed, high, and low prices to enable the utilization of candlestick patterns within the model. Various patterns are extracted, as outlined in Table 1, which encompasses 48 distinct candlestick patterns. The corresponding index value is extracted for each of the different formed patterns. If the candlestick pattern is established, its status value is set to 1; otherwise, if it is not established, the status value is set to zero. The status of all 48 different candlestick patterns is arranged in an array as shown in Eq. (3). The variable c_i represents the $i-th$ candlestick pattern from the first to the 48th candlestick pattern.

$$Patterns = \{ c_1 \quad c_2 \quad \cdots \quad c_{48} \} \tag{3}$$

To construct the final dataset, a history of the latest returns, denoted by the variable r, within the time interval from 0 to T periods prior, along with 48 candlestick patterns extracted with the symbol c in the last period t, are merged together. These patterns are then encapsulated into the final feature represented by Eq. (4), serving as the input for the neural network. In this equation, the attributes of each commodity are denoted by the index i and the time period t. The set of 48 distinct candlestick patterns encompasses the following.

Candlestick patterns: Doji, Spinning Doji, Gravestone Doji, Dragonfly Doji, Doji Star, Long-Legged Doji, Morning Doji Star, Evening Doji Star, Hammer, Hanging Man (Hammer), Inverted Hammer, Engulfing Pattern, Harami Pattern, Harami Cross Pattern, Piercing Pattern, Dark Cloud Cover, Morning Star, Evening Star, Shooting Star, Marubozu, Rising/Falling Three Methods, Abandoned Baby, Tristar Pattern, Three Advancing White Soldiers, Three Black Crows, Identical Three Crows, Two Crows, Upside Gap Two Crows, Unique Three River, Three Inside Up/Down, Three Outside Up/Down, Belt-Hold, Three Stars In The South, Advance Block, Stick Sandwich, Matching Low, Ladder Bottom, Concealing Baby Swallow, Breakaway, Tasuki Gap, In-Neck Pattern, On-Neck Pattern, Thrusting Pattern, Up/Down-gap Side-by-Side White Lines, Separating Lines, Mat Hold, Three-Line Strike, Upside/Downside Gap Three Methods.

$$\begin{aligned} Features_{i,t} &= Returns_{i,t} + Patterns_{i,t} \\ &= \{ r_{i,t} \quad r_{i,t-1} \quad \cdots \quad r_{i,t-T} \quad c_{i,t,1} \quad c_{i,t,2} \quad \cdots \quad c_{i,t,48} \} \end{aligned} \tag{4}$$

Stage 2: long short-term memory model

Figure 2 illustrates the architectural schematic of the deep neural network model designed for predicting the return value of assets. The proposed architecture is designed in a multi-layered manner, consisting of n parallel layers, where n equals the number of assets used in the study. This architecture allows for dynamic adjustment of the layers based on the number of assets, enabling each layer to be assigned to a specific asset. The segregation of each layer simulates the network independently for each asset, facilitating separate calculations for each asset without interference from others. This segregation enhances the accuracy of the output.

In each layer, the input features consist of a history of return values over a period of size T, which is assumed to be 30 in this study. Additionally, the status of 48 different candlestick patterns is represented as binary values at time t, which are merged together. The input vector of the neural network is the sum of the number of features of candlestick patterns and the return values. The LSTM neural network is comprised of an input layer, two intermediate layers, and an output layer designed linearly. This design converts the vector resulting from the intermediate layer into a vector of possible values, thereby producing the output equivalent to the return prediction for time t+1. By utilising LSTM, this neural network is capable of learning long-term dependencies between input and output information.

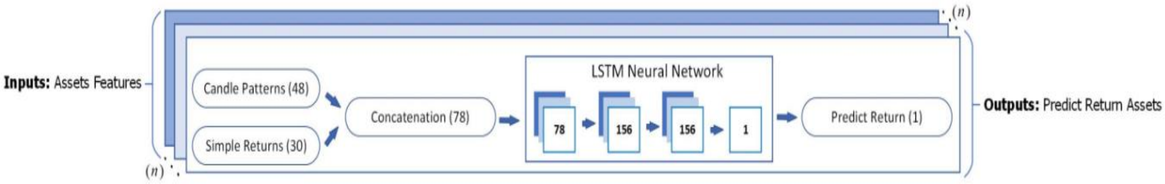


Figure 2. LSTM neural network architecture schematic

Stage 3: prediction accuracy

In artificial intelligence, the confusion matrix functions as a table illustrating the performance of supervised learning algorithms. Each column represents predicted values, and each row corresponds to actual samples. Confusion matrix is used to determine the value of evaluation indicators such as accuracy and sensitivity, whose relationship is expressed in Eq. (5). TP counts correctly classified positive samples. TN counts correctly classified negative samples. FP represents incorrectly classified positive samples, and FN represents incorrectly classified negative samples. Accuracy indicates how well the model performs overall in detecting true and false positives, while sensitivity indicates the ability of the model to detect true positives (Stehman, 1997*)*.

Output results from the proposed LSTM model at time t are compared with the return value of each asset at time t+1. If both values are positive or negative, it indicates a correct prediction; otherwise, it is considered incorrect. The number of correct and incorrect classifications for all cryptocurrencies is recorded daily and hourly. Accuracy and sensitivity metrics are utilized to assess model performance. Consequently, the count of correct and incorrect classifications for all assets in both daily and hourly time frames during the testing period is tallied, and then the mentioned metrics are calculated.

$$Accurancy = \frac{TP+TN}{TP+TN+FN+FP}$$
$$Sensitivity = \frac{TP}{TP+FN} \tag{5}$$

Stage 4: portfolio strategy

Two strategies have been employed to simulate the portfolio:

- Buy and Hold Strategy: This approach is straightforward and widely used in capital market investment. It entails purchasing assets and holding them for a specified period of time t.
- Suggested strategy: The confusion matrix generated in the preceding step is used, where if the forecasted value exceeds the threshold, a purchase is initiated; conversely, if the forecasted value falls below the negative threshold, a sale is executed. Otherwise, there is no alteration in the portfolio size, as indicated in Eq. (6).

$$Action = \begin{cases} Buy & predict \geq Threshold \\ Sell & predict \leq Threshold \\ Hold & otherwise \end{cases} \tag{6}$$

5. Experimental settings

In this section, details on data collection, the historical period used, programming libraries, and neural network settings are provided.

5.1. Cryptocurency market data

In this study, data from the years 2022 and 2023 pertaining to the top 20 cryptocurrencies by market cap, as per Binance records, have been utilised. The data is collected at daily and hourly intervals via the Binance API. The cryptocurrencies considered in this study include BTC, ETH, BNB, SOL, ADA, DOGE, AVAX, TRX, DOT, LINK, MATIC, SHIB, ICP, LTC, BCH, ATOM, UNI, ETC, XLM, XMR. Binance stands as one of the most prominent and widely used cryptocurrency exchanges globally. Binance has quickly gained popularity for its extensive range of supported cryptocurrencies and advanced trading features. As a comprehensive cryptocurrency exchange platform, Binance facilitates the buying, selling, and trading of various digital assets. In this paper data was collected including unixtime, high, low, open, and close. This information is gathered on an hourly and daily basis, and all data is stored in a MySQL database.

One of the prominent PHP libraries for technical analysis is known as "Technical Analysis for Traders." This library, developed by a programmer named "Tommic," can be explored on GitHub This library provides support for various technical analysis tools, including moving averages, price-volume indicators, candlestick patterns, and more. To extract candlestick patterns, the trader section of the mathematical extensions plugin is utilised.

5.2. Algorithm configuration

In this section, the basic settings of the neural network, such as the number of layers and the size of each layer, the evaluation settings and the number of training steps are stated in Table 2. The Python programming language, along with the PyTorch library, has been employed to implement the neural network.

Table 2. Hyperparameters of Long Short-Term Memory

Hyperparameters	Value
Depth	4
Input Layer	78
Hidden layers	2x156
Output layer	1
Evaluation criteria	Mean Squared Error
Learning rate	0.001
Periods (Epochs)	50

6. Results

In this section, the results obtained from the proposed model are examined. For neural network training, 21 months of data, equivalent to 87.5%, have been utilised, while the remaining 12.5% consisting of the last three months of data have been used for testing. The time range of the training and testing periods is illustrated in Table 3. For the daily time frame, a total of 8461 patterns were identified within this range. Moreover, 118,674 patterns were identified in the hourly time frame for this range. The results are detailed in Table 4.

Table 3. Historical scope of the training and testing period

Milestones	Status	Date
Start	Train	01.01.2022
Finish	Train	09.30.2023
Start	Test	10.01.2023
Finish	Test	12.31.2023

Table 4. Summary of patterns analysis in daily and hourly time frames

Parameter	Daily	Hourly
Total detected patterns	8461	118674
Average patterns on each asset	423	5934

6.1. Neural network training error

To evaluate the training of the neural network model, the mean square error (MSE) evaluation criterion has been employed. In Figure 3, the total MSE error for all assets across 50 training iterations is depicted. Figure 3a of the neural network input comprises candlestick and simple return features, while Figure 3b includes only the return feature, and Figure 3c solely encompasses the candlestick feature. Notably, the overall error rate experienced an approximate 82% reduction. The objective of this comparison chart is to assess the predictability of candlestick patterns and returns independently versus their combined utilisation with other auxiliary features. Encouragingly, the outcomes indicate that Figure 3a exhibits lower noise levels and yields superior results compared to the other configurations.

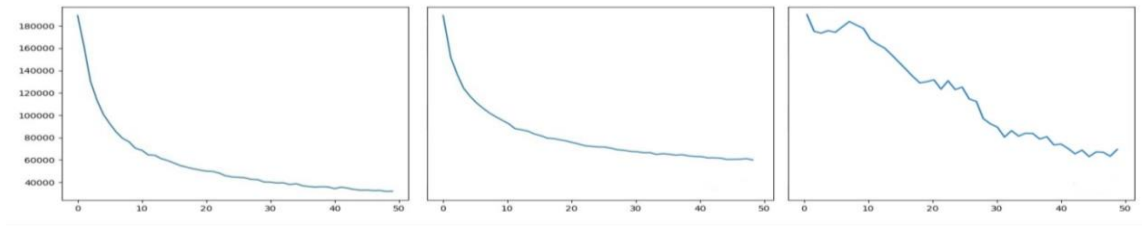


Figure 3a. Candlesticks with returns futures

Figure 3b. Returns futures

Figure 3c. Candlesticks futures

Figure 3. Neural network training accumulates errors for: 3a, 3b, and 3c

6.2. Accuracy and sensitivity evaluation

The confusion matrix serves as an important tool for evaluating the performance of the proposed model. The calculation results for a daily time frame within the defined test interval, as delineated in Table 5, are presented. In this table, the green cells signify the correct samples. The accuracy rates for different scenarios varied as follows: when both return and candlestick patterns were employed, the accuracy stood at 0.61, while using return alone yielded a slightly lower accuracy of 0.51. Similarly, utilising candlestick patterns in isolation also resulted in an accuracy of 0.51. In the combined mode, where both return and candlestick patterns were considered, the sensitivity rate was equal to 0.69.

Confusion matrix for the hourly time frame during the specified test period is presented in Table 6. The accuracy values varied across different modes: for the return and candle mode, it equalled 0.52, while for the candle mode alone, it was 0.51, and for the return mode alone, it stood at 0.5. Notably, in the combined mode, where both return and candlestick patterns were considered, the sensitivity value reached 0.76.

Table 5. Confusion matrix for daily time frame

	Actual Values (Candles+Returns)		Actual Values (Returns)		Actual Values (Candles)	
Predicted Values	621 TP	277 FN	264 TP	256 FN	526 TP	372 FN
	413 FP	465 TN	622 FP	633 TN	491 FP	387 TN

Table 6. Confusion matrix for hourly time frame

	Actual Values (Candles+Returns)		Actual Values (Returns)		Actual Values (Candles)	
Predicted Values	13550 TP	4250 FN	2610 TP	2402 FN	12703 TP	5097 FN
	11702 FP	4239 TN	14349 FP	14380 TN	11281 FP	4660 TN

6.3. Investigating buy/hold strategy based on initial capital

An initial capital of value 100 has been allocated for each of the studied assets. The evaluation was conducted using the buy-and-hold strategy for the test period within the daily time frame. Changes in the capital amounts for each currency are illustrated in Figure 4. The results indicate that Avax yielded the highest profit at 134.2, followed by Link at 129.4 and Sol at 128.9. Conversely, BCH currency demonstrated the lowest profit value at 95.8 within the portfolio.

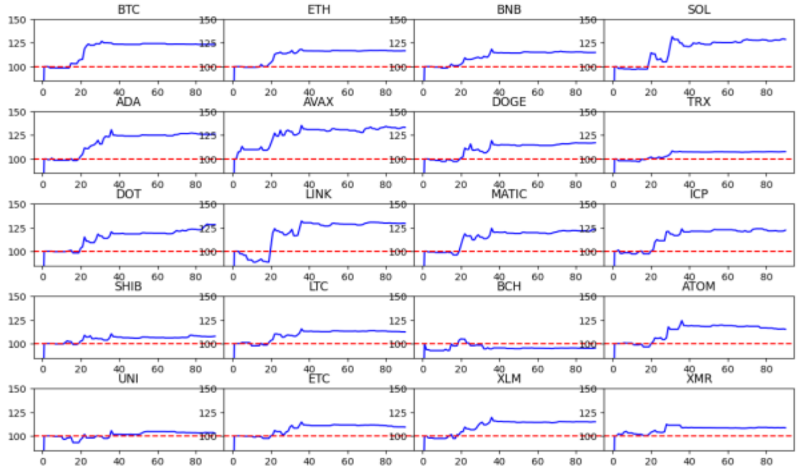


Figure 4. Profitability comparison of top 20 cryptocurrencies with buy and hold strategy

6.4. Investigating proposed strategy based on initial capital

In the implementation of the proposed model on 20 cryptocurrencies with an initial capital of 100, three possible modes may occur for each asset buy, sell, or hold. In Figure 5a, the blue lines represent the state of buying or holding the currency, indicating its presence in the portfolio. Empty spaces denote the sale of the currency and its absence in the portfolio during the test trial period. The results in Figure 5b reveal that SOL demonstrated the highest value at 195.9, followed by LINK at 169.1 and SHIB at 135.6. Conversely, the BCH currency exhibited the lowest value at 104.4.

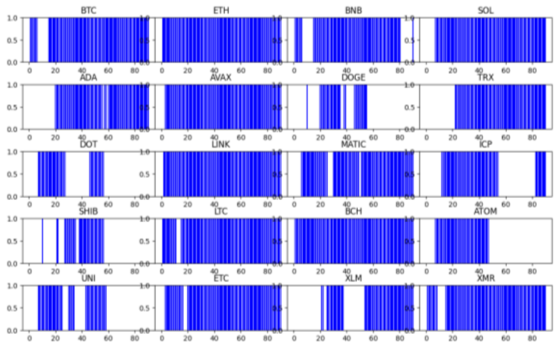


Figure 5a. Portfolio Changes

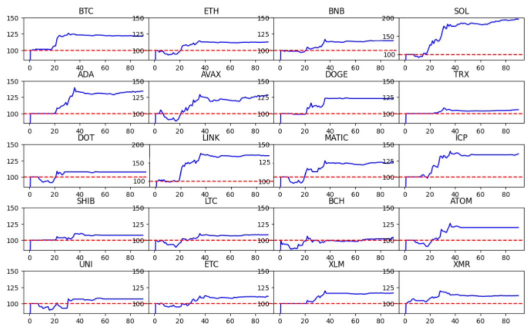


Figure 5b. Profits

Figure 5. Portfolio Changes and Profits in the Proposed Strategy using 20 Cryptocurrencies (see in 5a and 5b)

Figure 6 illustrates the max, min, and average cumulative profits of the capital portfolio generated by both the proposed model and the buy-and-hold strategy for 20 cryptocurrencies during the test period. At the conclusion of the test, the portfolio yielded the highest value of 195.9 for the proposed model and 133.3 for the buy-and-hold strategy. Conversely, in the lowest scenario, the proposed model yielded 105.9, while the buy-and-hold strategy returned 95.2. The average value for the proposed model stood at 122.9, compared to 116.8 for the buy-and-hold strategy.

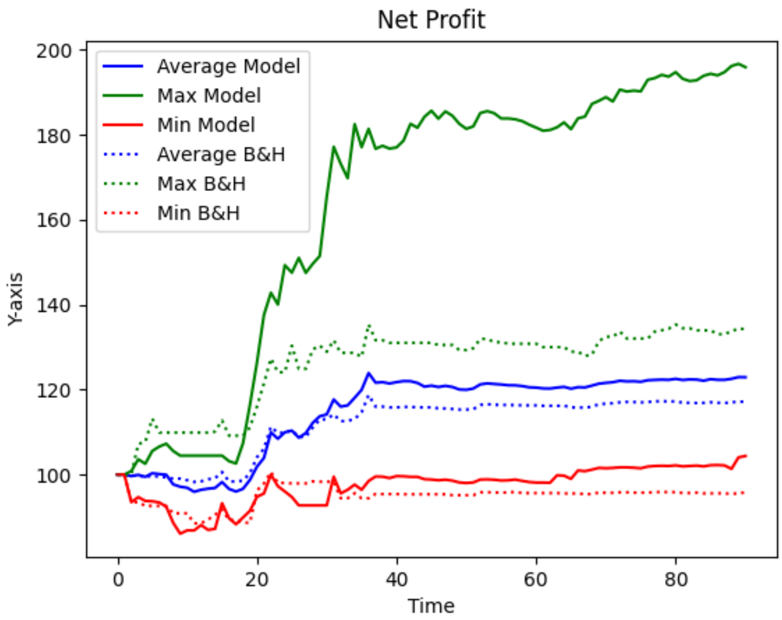


Figure 6. Portfolio cumulative profit for top 20 cryptocurrencies

The comparison between the graphs of the two strategies reveals that the proposed model was consistently more profitable compared to the buy and hold strategy in all cases. Furthermore, even in the lowest-performing scenario, the buy and hold strategy incurred losses, whereas the proposed strategy remained profitable.

7. Conclusions and discussion

In this study, a model designed to predict returns in the cryptocurrency market was introduced. The model utilizes LSTM neural networks trained on extracted candlestick pattern data and return values to discern market trends. Twenty cryptocurrencies were examined in this study, with an average of 423 and 5934 candlestick patterns identified per cryptocurrency in the daily and hourly time frames, respectively. While including more historical data and a larger number of cryptocurrencies may improve the model's performance, it also leads to increased complexity and training time.

The neural network's training error significantly decreased by 82%, indicating effective learning and convergence towards an optimal solution. This suggests the satisfactory quality of the input data, which includes candlestick patterns alongside return values. While the network's error for candlestick features alone exhibited some noise, integrating them with other indicators like returns led to superior outcomes. These findings underscore that candlestick patterns offer a robust tool for technical analysis and discerning market sentiments, as well as the potential for detecting shifts in market trends. Nevertheless, relying solely on candlestick patterns yields subpar results and proves insufficient. By amalgamating them with other technical indicators, more precise outcomes can be attained. (Aycel & Santur, 2022; Jain et al., 2022; Prado et el., 2013; Santur, 2022).

The proposed model achieved an accuracy of 0.61 in the daily time frame and 0.52 in the hourly time frame. These results indicate that the model effectively learned and identified trends in both daily and hourly intervals. Notably, the daily timeframe exhibited a 9% improvement in accuracy compared to the hourly data. This suggests that the higher trading volume in daily data leads to lower volatility and more stable prices. Consequently, identifying long-term trends becomes more precise with less influence from short-term emotional fluctuations often seen in hourly data.

The proposed model demonstrably outperformed the buy-and-hold strategy in forecasting and capital management for cryptocurrency markets. It achieved a profitability rate of 195.9, which was 73.0% higher than buy-and-hold in the optimal scenario. Even in the least favourable scenario, where the buy-and-hold strategy incurred losses, the proposed strategy remained profitable. The disparity between the two strategies in the least favorable scenario was approximately 10.7%, indicating the superior performance of the proposed strategy over the buy and hold approach.

8. Future works

In this study, the impact of candlestick patterns on profitability, in conjunction with the return parameter, was examined. However, future research could explore other parameters of technical analysis and technical indicators. This could involve the utilisation of moving average indicators (such as SMA, EMA, etc.), volatility indicators (such as Bollinger Band, ATR, Keltner, etc.), volume indicators (such as Volume, MFI, etc.), trend indicators (such as RSI, MACD, Ichimoku, etc.), and more.

To enhance profitability and mitigate risk within the investment portfolio while attaining predefined objectives, one can consider integrating the research concept with innovative diversification methods and investment portfolio management utilising machine learning. Additionally, to manage transactional risks within the portfolio, implementing profit limits to safeguard earned profits or loss limits to mitigate capital depletion can be beneficial strategies.

The current study employed the LSTM neural network to analyse data on the financial system time series. However, alternative machine learning methods, including other neural network architectures (such as CNN, GRU, etc.), reinforcement learning techniques, and tree-based algorithms, can be explored and compared to assess their performance.

REFERENCES

Avasaram (2012) *Candle sticks tutorials.* avasaram.com/tutorials/candlesticks/tutorialLauncherCandleSticks.jsp, Arizona, USA.

Aycel, Ü. & Santur, Y. (2022) A new algorithmic trading approach based on ensemble learning and candlestick pattern recognition in financial assets. *Turkish Journal of Science and Technology*. 17(2), 167-184. doi: 10.55525/tjst.1124256.

Caginalp, G. & Laurent, H. (1998) The predictive power of price patterns. *Applied Mathematical Finance*. 5, 181-206. doi: 10.1080/135048698334637.

Chootong, C. & Sornil, O. (2012) Trading signal generation using a combination of chart patterns and indicators. *International Journal of Computer Science Issues*. 9(6), 202-209.

Gers, F. A., Schmidhuber, J. & Cummins, F. (2000) Learning to forget: Continual prediction with LSTM. *Neural computation*. 12(10), 2451-2471. doi: 10.1162/089976600300015015.

Goswami, M. M., Bhensdadia, C. K. & Ganatra, A. P. (2009) Candlestick analysis based short term prediction of stock price fluctuation using SOM-CBR. In: *2009 IEEE International Advance Computing Conference*. IEEE. pp. 1448-1452. doi: 10.1109/IADCC.2009.4809230.

Jain, R., Bhardwaj, P. & Soni, P. (2022). Can The Market of Cryptocurrency Be Followed with the Technical Analysis. *International Journal for Research in Applied Science & Engineering Technology*. 10(IV), 2425-2445. doi: 10.22214/ijraset.2022.41760.

Karmelia, M. E., Widjaja, M. & Hansun, S. (2022) Candlestick Pattern Classification Using Feedforward Neural Network. *International Journal of Advance Soft Computing & Its Applications.* 14(2), 79-95. doi: 10.15849/IJASCA.220720.06.

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.

Lu, T. H., Shiu, Y. M. & Liu, T. C. (2012) Profitable candlestick trading strategies-The evidence from a new perspective. *Review of Financial Economics*. 21(2), 63-68. doi: 10.1016/j.rfe.2012.02.001.

Morris, G. L. (2006) *Candlestick charting explained: Timeless techniques for trading stocks and futures* (3rd ed.). New York, NY: McGraw-Hill.

Nazário, R. T. F., e Silva, J. L., Sobreiro, V. A. & Kimura, H. (2017) A literature review of technical analysis on stock markets. *The Quarterly Review of Economics and Finance*. 66(C), 115-126. doi: 10.1016/j.qref.2017.01.014.

Nison, S. (2001) *Japanese candlestick charting techniques: a contemporary guide to the ancient investment techniques of the Far East*. Penguin Publishing Group.

Prado, H. A., Ferneda, E., Morais, L. C., Luiz, A. J. & Matsura, E. (2013) On the effectiveness of candlestick chart analysis for the Brazilian stock market. *Procedia Computer Science*. 22, 1136-1145. doi: 10.1016/j.procs.2013.09.200.

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. doi: 10.14744/sigma.2022.00039.

Stehman, S. V. (1997) Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*. 62(1), 77-89. doi: 10.1016/S0034-4257(97)00083-7.

Tharavanij, P., Siraprapasiri, V. & Rajchamaha, K. (2017) Profitability of candlestick charting patterns in the stock exchange of Thailand. *SAGE Open*. 7(4). doi: 10.1177/2158244017736799.

Xie, H., Fan, K. & Wang, S. (2021) *Candlestick forecasting for investments: applications, models and properties*. New York. Routledge.

Mohammad VAHIDPOUR, Ph.D. Candidate in Information Technology Management, Science and Research Branch, Islamic Azad University, Tehran, Iran. His current research interests include blockchain, cryptocurrencies, software engineering, artificial intelligence, machine learning, modelling and time series analysis.

Amir DANEHSVAR, Assistant Professor, Department of Industrial Management, Faculty of Management, Electronic Branch, Islamic Azad University, Tehran, Iran. His current research interests include data envelopment analysis, multi criteria decision making, structural equation, modelling, regression analysis, time series analysis, system dynamics and meta heuristic algorithms.

Mohsen AMINI KHOUZANI, Assistant Professor, Department of Financial Engineering, Shahr-e-Qods Branch, Islamic Azad University, Shahr-e-Qods, Tehran, Iran. His current research interests include financial analysis and economics, management, planning, e-commerce and social media, strategic management, financial planning and budgeting.

Mahdi HOMAYOUNFAR, Assistant Professor, Department of Industrial Management, Rasht Branch, Islamic Azad University, Rasht, Iran. His current research interests includes operation management, optimisation, prediction planning, production/operation management, SCM, production planning, linear programming, industrial management, simulation and modelling.