JOIV : Int. J. Inform. Visualization, 7(3-2): Empowering the Future: The Role of Information Technology in Building Resilience - November 2023 2016-2024

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Evaluation of Cryptocurrency Price Prediction Using LSTM and CNNs Models

Ng Shi Wen [a], Lew Sook Ling [a,*]

[a] *Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia*

*Corresponding author: *sllew@mmu.edu.my*

Abstract*— Cryptocurrencies created by Nakamoto in 2009 have gained significant interest due to their potential for high returns. However, the cryptocurrency market's unpredictability makes it challenging to forecast prices accurately. To tackle this issue, a deep learning model has been developed that utilizes Long Short-Term Memory (LSTM) neural networks and Convolutional Neural Networks (CNNs) to predict cryptocurrency prices. LSTMs, a type of recurrent neural network, are well-suited for analyzing time series data and have been successful in various prediction applications. Additionally, CNNs, primarily used for image analysis tasks, can be employed to extract relevant patterns and characteristics from input data in Bitcoin price prediction applications. This study contributes to the existing related works on cryptocurrency price prediction by exploring various predictive models and techniques, which involve a machine learning model, deep learning model, time series analysis, and as well as a hybrid model that combines deep learning methods to predict cryptocurrency prices as well as enhance the accuracy and reliability of the price predictions. To ensure accurate predictions in this study, a trustworthy dataset from investing.com was sought. The dataset, sourced from investing.com, consists of 1826 time series data samples. The dataset covers the time frame from January 1, 2018, to December 31, 2022, providing data for a period of 5 years. Subsequently, pre-processing was conducted on the dataset to guarantee the quality of the input. As a result of absent values and concerns regarding the dataset's obsolescence, an alternative dataset was sourced to avoid these issues. The performance of the LSTM and CNN models was evaluated using root mean squared error (RMSE), mean squared error (MSE), mean absolute error (MAE) and R-squared (R^2). It was observed that they outperformed each other to a certain degree in short-term forecasts compared to long-term predictions, where the R^2 values for LSTM range from 0.973 to 0.986, while for CNNs, they range from 0.972 to 0.988 for 1 day, 3 days and 7 days windows length. Nevertheless, the LSTM model demonstrated the most favorable performance with the lowest error rate. The RMSE values for the LSTM model ranged from 1203.97 to 1645.36, whereas the RMSE values for the CNNs model ranged from 1107.77 to 1670.93. As a result, the LSTM model exhibited a lower error rate in RMSE and achieved the highest accuracy in R^2 compared to the CNNs model. Considering these comparative outcomes, the LSTM model can be deemed as the most suitable model for this specific case.

***Keywords*— Cryptocurrency; Bitcoin; LSTM; CNNs; prediction; price.**

Manuscript received 27 Dec. 2022; revised 11 Apr. 2023; accepted 24 Jun. 2023. Date of publication 30 Nov. 2023.

I. Introduction

The digital or virtual currency known as cryptocurrency is protected by described cryptography, making it almost unlikely to fake or double-spend [1], [2]. The market for these digital currencies is very unstable, and cryptocurrencies are still a relatively new concept [3]. Cryptocurrencies tend to be uninsured and difficult to convert into real money (such as US dollars or euros) since they do not need banks or any other third party to control them. Moreover, as cryptocurrencies are intangible assets based on technology, they are vulnerable to hacking, just like any other intangible asset. Bitcoin, Ethereum, and Litecoin are three of the most well-known cryptocurrencies currently in circulation. Bitcoin and other cryptocurrencies are increasingly being used in lieu of traditional fiat currencies, which are traded on online exchanges [4], [5]. The technique of determining what the value of a cryptocurrency will be in the future is referred to as cryptocurrency price prediction. This challenging endeavor calls for an analysis of a wide range of factors, including shifts in consumer preferences, rates of product uptake, developments in relevant technologies, and existing market conditions. To get past this obstacle, a deep learning approach, the LSTM model and CNNs model, are developed and utilized to predict the price of cryptocurrency.

In recent years, there has been much interest in the regular stock market, but owing to its quick rise, the cryptocurrency market has established itself as being particularly noteworthy. Due to the industry's rapid growth, several researchers developed computer algorithms to uncover methodologies that provide accurate forecasts for the bitcoin market [6]–[16]. Predicting the behavior of financial markets has always been challenging, requiring much prior financial expertise and strong data interpretation abilities. The traditional method, especially when applied to the due to the vast quantity of data that is collected every day. Because deep learning approaches are inherently skilled at conducting in-depth data analyses, they are thus more successful than other methods at predicting the behavior of data relating to dynamics. In addition, while assessing the market, deep learning reads data directly from the dataset and actively promotes itself during each learning cycle [17], [18]. To put it another way, in this study, the performance of the LSTM model and CNNs model improves in tandem with the size of the dataset.

On the other hand, accurate cryptocurrency price forecasting is essential for investors because it can help them choose the best time to buy and sell their assets [19]. This is because the information may guide when investors should buy and sell their assets. It is essential for businesses and organizations who wish to accept cryptocurrency payments or utilize initial coin offers to raise capital because of cryptocurrencies' versatility. In light of this, analyzing the many deep learning approaches that are now in use seems to be an absolute need to develop adequate algorithms for predicting the most important cryptocurrencies. This investigation aims to evaluate the performance of a certain supervised deep learning algorithm, the LSTM and CNNs models regarding prediction making. The contribution of this study is to establish the most efficient model for trading short-term cryptocurrencies and to provide direction for future investigation.

The structure of this study is organized as an executive literature review of previous work presented in Section 2. Section 3 describes the proposed methodology. The experimental assessments are presented in Section 4. The conclusion and discussion are presented in Section 5.

II. Materials and Method

An executive of related works or literature review is a concise summary of relevant research studies, providing decision-makers with an overview of the current knowledge in the cryptocurrency prediction field. It helps to inform strategic decision-making by highlighting key findings, trends, and areas for further research. Numerous industry specialists have extensively tried to predict the cryptocurrency market by employing diverse approaches, techniques, and algorithms [6]–[16]; they aim to formulate the most effective investment strategies by conducting comprehensive analyses of various technical and emotional factors. As part of their evaluation, they have attempted to assess the impact of specific qualities on cryptocurrency prices. These specialists analyze the cryptocurrency market using a variety of instruments and methods. Technical analysis involves the examination of historical price and trading volume data, the identification of patterns, and the application of indicators to forecast future price movements. A summary of the related works is presented in Table 1.

TABLE I
SUMMARY TABLE OF RELATED WORKS

Past studies	Propose Method	Dataset	Result
[6]	GRU + LSTM	Litecoin Monero	Not available
[7]	GRU + LSTM	Litecoin Zcash	• MSE Loss, Litecoin ○ 1-day: 0.02038 ○ 3-days: 0.02103 ○ 7-days: 0.02337 ○ 30-days: 0.02637 • MSE Loss, Zcash ○ 1-day: 0.00461 ○ 3-days: 0.00483 ○ 7-days: 0.00524 ○ 30-days: 0.00816
[8]	ANN + LSTM	Bitcoin, Ethereum, Ripple	Not available
[9]	LSTM vs. ARIMA	Bitcoin	• RMSE of LSTM are 198.448 (single feature) and 197.515 (multi-feature) • ARIMA model RMSE is 209.263
[10]	ANN versus LSTM	Bitcoin	• RMSE and MAE for after 1 day, 10 days, 20 days, and 30 days • RMSE: 53.30, 67.99, 91.41, 45.71 • MAE: 28.93, 39.21, 60.63, 23.43
[11]	ARIMA	Bitcoin	MAPE, 0.87 for the next one-day prediction MAPE, 5.98 for the next seven day
[12]	ARIMA vs. Linear Regression	Bitcoin	Not available
[13]	ARIMA	Bitcoin	MSE: 16000
[14]	CNNs + SGRUs	Bitcoin, Ethereum, Ripple	RMSE: 43.933 RMSE: 3.511 RMSE: 0.00128
[15]	MICDL	Bitcoin	Lag = 7; 14 MAE: 170.761; 170.147 RMSE: 257.728; 257.847 R2: 0.952 Acc: 53.04%; 51.88% GM: 30.886; 29.173 Sen: 0.698; 0.582 Spe: 0.306; 0.434
		Ethereum	Lag = 7; 14 MAE: 9.233; 9.146 RMSE: 13.551; 13.492 R2: 0.964 Acc: 50.84%; 51.11% GM: 29.582; 30.461 Sen: 0.483; 0.628 Spe: 0.526; 0.363
		Ripple	Lag = 7; 14 MAE: 0.005; 0.007 RMSE: 0.007; 0.009 R2: 0.958; 0.953 Acc: 49.07%; 49.23% GM: 25.053; 26.157 Sen: 0.366; 0.442 Spe: 0.630; 0.549
[16]	CNNs + LSTM	Bitcoin	MAE: 209.89 RMSE: 258.31 MAPE: 2.35 Precision: 0.64 Recall: 0.81 F1: 0.69

In the prior work, Gated Recurrent Unit (GRU) and LSTM were used to provide a novel method of bitcoin price prediction [6]. The authors of this research used data on Litecoin and Monero that they obtained on Investing.com, a website that offers a financial markets platform, real-time data, quotations, charts, financial tools, breaking news, and analysis. Five separate components make up these two cryptocurrency datasets: cost, open, close, high, low, and volume. Compared to the 1851 data points selected for Bitcoin from January 30, 2015, to February 23, 2020, 1279 were chosen for Litecoin from August 24, 2016 to that date. The authors of this research focused more on the suggested bitcoin pricing structure than the production.

Meanwhile, another study in the same general region at the same time [7] anticipates cryptocurrency prices using GRU and LSTM. The author used the study materials Litecoin and Zcash. This study's conclusion looked at the MSE Loss for window sizes of one day, three days, seven days, and thirty days. According to the evaluation above of the journal article, using the GRU and LSTM prediction models has both benefits and drawbacks. The first journal discovered that while GRU has gates, LSTM is equivalent. "To tackle perishable gradients in LSTM, upgrades, reset, and renewal gates exist. The renewal gate determines the total quantity of digital data that must be transported cyclically from the past to the present.

Additionally, it is mentioned that the LSTM network may suffer from gradient distortion. Additionally, the second paper demonstrated how the application of GRU and LSTM can quickly identify the interdependence between two coins by using historical data and a direction algorithm. However, the model's flaws were demonstrated after the paper, demonstrating how accuracy decreases with non-uniform data due to practicality.

Moreover, the previous research also looked at how to anticipate bitcoin prices using LSTM and Artificial Neural Network (ANN) [8]. The MSE for 7, 14, 21, 30, and 60 days has been computed in this study using Bitcoin, Ethereum, and Ripple from August 7, 2015, to June 2, 2018. The benefit of using this model is that LSTM is better at using pertinent information hidden in historical memory than ANN. However, the optimization method and model parameters are not well-tested when deciding on a suitable historical and predictive memory length. The comparison between the two models may be made without a doubt. Hence, successfully compared and contrasted the LSTM model with the Autoregressive Integrated Moving Average (ARIMA) model [9]. In his research, the author settled on Bitcoin for the years 2017–2018. Results from this comparison of LSTM reveal that the single-feature RMSE is 198.448 and the multi-feature RMSE is 197.515, whereas the ARIMA model's RMSE is 209.263. The limiting and delimitation of these two models were done after the comparison output, and the results revealed that the LSTM was better at identifying long-term dependencies. Furthermore, the LSTM model will result in vanishing gradients since it may be too simple for the system to learn, while inclinations can be regularized, making it more appropriate for forecasting larger fluctuations in time series data. Additionally, even though RNN can manage long-term dependence, it often performs poorly as a result of gradient succession and long-term reliance.

Additionally, the research described for the LSTM model conducts bitcoin prediction using the LSTM and ANN techniques [10]. The study uses Bitcoin historical data from August 7, 2015, to June 2, 2018. After one, ten, twenty, and thirty days, the effects of RMSE and MAE were evaluated. The results showed that the RMSE was 53.30, 67.99, 91.41, and 45.71, while the MAE was 28.93, 39.21, 60.63, and 23.43. The suitability of LSTM for modeling time series dynamics is discussed in this work, along with its benefits and drawbacks, as well as attempts at model optimization and computational tools. Additionally, the following table lists other studies that have used the ARIMA model to forecast cryptocurrency prices. According to the paper, the author states that historical Bitcoin data were collected from May 1, 2013, to June 7, 2019 [11]. The model MAPE value ARIMA was calculated to make predictions for the next day (0.87) and the following seven days (5.98).

In conclusion, it can be claimed that the ARIMA model works better for short-term predictions, especially in the next two periods. However, ARIMA has its challenges in handling long-term periods of data, which may result in lower accuracy prediction. Furthermore, conducted another comparison and contrast between ARIMA and Linear Regression [12]. The work of comparison includes a forecast for the price of Bitcoin over the following four months. After comparing the results of these two methods, the author concludes that the ARIMA model performs better than the LR model and that the linear regression method yields less accurate predictions than the time series method. Similarly, to forecast Bitcoin prices, used the closing element from September 2015 to September 2018 [13]. By comparing the output from the MSE and RSS models, it can be shown that the ARIMA model fits the data better while operating with a single input, but the ARIMA model does not consider the rapid price fluctuations that occur throughout the research process.

Several researchers devised the Convolution Neural Networks approach to predict the cryptocurrency price. One of them was the authors' proposal of a novel method for predicting the price of cryptocurrencies that incorporates Convolutional Neural Networks (CNNs) and stacked gated recurrent units (SGRUs) [14]. The Bitcoin sample data was obtained from Kaggle between January 1, 2012, and March 31, 2021, while the historical closing prices of Ethereum and Ripple were obtained from the Bitstamp exchange website beginning in January 2021. This study uses the root mean squared error (RMSE) as a singular evaluation metric to evaluate the model's prediction performance. The RMSE results for each cryptocurrency were 43.933, 3.511 and 0.00128 respectively. The advantages and disadvantages indicated in this paper are that a high-level discriminative representation of the data is encoded using a one-dimensional convolutional neural network, and the stacked gated recurrent unit then captures the long-range dependencies of the representation. However, this study only utilizes one week of historical data due to limited computing resources. In addition, a new modular Multiple-Input Cryptocurrency Deep Learning Model (MICDL) was also integrated with LSTM and CNNs[15]. The authors also analyzed Bitcoin, Ethereum, and Ripple from January 1, 2017, to October 31, 2020, as well as the evaluation metrics for MAE, RMSE, and R^2 with Lag 7 and 14 days for each cryptocurrency. The authors indicated that employing the MICDL model reduces overfitting and computational cost with security, whereas there was no complete examination of cryptocurrency information, such as daily trading volume or economic and technical trading indicators. Lastly, it focuses on the combined Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks in a hybrid neural network model for Bitcoin price forecasting [16]. The model was trained on a big dataset of Bitcoin prices from December 30, 2016, to August 31, 2018. The researchers used MSE, RMSE, MAPE, Precision, Recall, and F1-score matrices to predict and measure model performance accuracy. The results were shown as MAE: 209.89, RMSE: 258.31, MAPE: 2.35, Precision: 0.64, Recall: 0.81 and F1: 0.69. The authors mentioned the study's pros and cons. CNN-LSTM neural network hybrid performs well in Bitcoin forecasting and is more suited for Bitcoin forecasting and considers external factors that involve all potential influences on Bitcoin pricing, including macroeconomic factors and investor interest. In contrast, the limitations of the proposed approach would be the authors did not indicate the corporation of the cryptocurrency's popularity on other social software.

To conclude, the literature study offered three-time series model prediction approaches: Long Short-Term Memory (LSTM), Autoregressive Integrated Moving Average (ARIMA), and Convolutional Neural Networks (CNNs) are presented in Table 1. Most researchers found that LSTM and CNNs were better than ARIMA [6]–[16]. Each algorithm has pros and cons, as expected. CNNs have shown promise in time series data prediction applications by accurately capturing temporal patterns and producing sequential data-based predictions. This study will compare CNNs and LSTMs for Bitcoin price prediction. The reason for comparing CNNs and LSTM models for Bitcoin price prediction is motivated by each model architecture's unique strengths and characteristics. The researchers have proven that CNNs are widely recognized for their effectiveness in handling visual imagery [20]. They excel at automatically extracting meaningful features from raw data, such as images, through convolutional layers and pooling operations.

On the other hand, the author mentioned that LSTM models are well-suited for capturing long-term dependencies in sequential data [21]. They are particularly adept at modeling time series data, where the order and temporal relationships of the data points are crucial for making accurate predictions. From here, two models will be compared in this paper to determine which is more reliable for forecasting bitcoin prices.

This section discusses the proposed model for predicting the price of bitcoin using the LSTM algorithm compared to CNN's algorithm. Fig 1 demonstrates the process flow of the cryptocurrency price prediction. Initially, it is necessary to compile a historical Bitcoin price data database. In addition to the prices, this information should include the compiled dates. In addition, for the dataset to be appropriate for use with an LSTM model and a CNN model, it must be cleaned and pre-processed after collection. This may involve removing missing or incorrect values, normalizing the data, and transforming the data into a format compatible with the LSTM and CNN models.

Regarding model training, both models are capable of being trained using pre-processed data. This typically involves separating the data into training and validation sets and utilizing the training set to change the internal weights and biases of the model using the backpropagation approach. After training the model, its accuracy and performance may be evaluated using the validation set. Calculating mean absolute error and root mean square error could be essential.

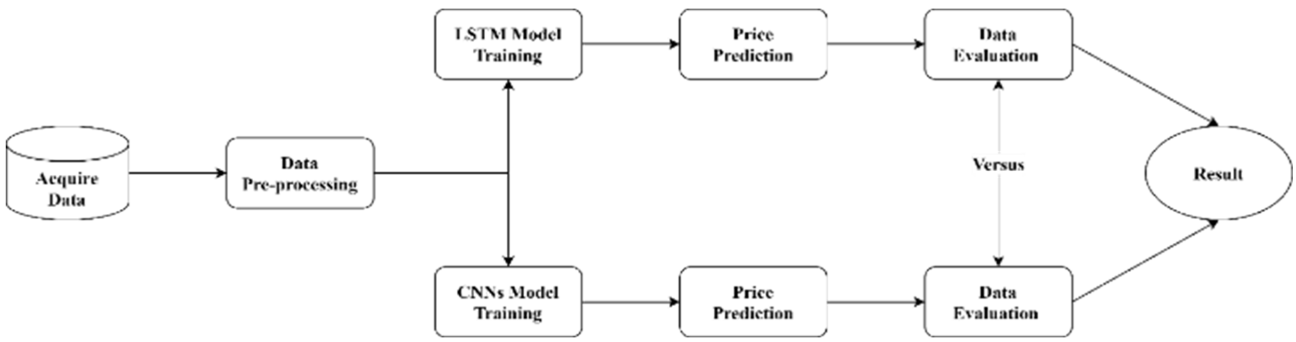


Fig. 1 Process flow cryptocurrency price prediction

A. Data Collection

This study utilized data obtained from Investing.com. It offers details on stocks, options, analysis, digital currencies, futures, and commodities. The submitted information is continually revised to ensure the desired outcome. Investing.com provides data export capabilities that make it straightforward to obtain the price history of each cryptocurrency. This study compiled Bitcoin's historical price for the five years from January 1, 2018, to December 31, 2022. The information gathered for Bitcoin comprises five characteristics: Date, Price, Open, High, Low, Vol., and change %. The historical price of Bitcoin from 2018 to 2022 is illustrated in Fig 2, which contains 1826 sample data.

- Date: The specific time period's date.
- Price: The final cost for the specified time period.
- Open: The price that is open for that specific time period.
- High: The highest price for that specific time period.
- Low: The lowest cost for the specified time period.
- Volume: The total number of transactions within the specified time period.

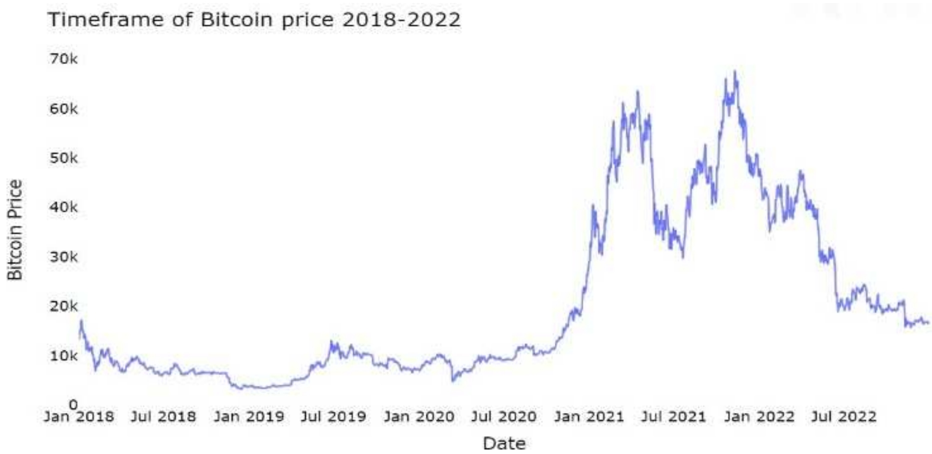


Fig. 2 The Bitcoin historical price for 2018 – 2022

B. Pre-processing

Several pre-processing procedures are used to tidy up the historical Bitcoin data, such as feature selection, timestamp and data type conversion, missing value check, filtering attribute, train-test split, and normalization of the min-max scaling. Due to the large number of features involved in the Bitcoin data at this stage, the Date and Price features were chosen to undertake price predictions. Follow by the timestamp conversion was performed where the timestamp in DD Mon, YYYY format was converted into YY-MM- DD date format. Meanwhile, the data typed conversion was carried out to convert the object data type to float data type, the affected characteristics are Price, Open, High and Low. After verifying, there are no missing values or NaN values in the Bitcoin dataset. The characteristics (Open, High, Low, Vol, and Change%) are filtered out by dropping the dataset column. For evaluation purpose, a training set and testing set were created using the historical bitcoin price indicated for 1823 sample data, where the training set was raging as 80% which are 1460 samples data, testing set was 20%, which is 366 samples data to examine the evaluation for 1 day, 3 days and 7 days windows length. The window length refers to the number of previous time steps or data points that are considered as input to the model. It represents the historical information or sequence of data that the model uses to make predictions.

As a result of outliers and fluctuations, the proposed model cannot be applied directly to the original data values. Normalization is performed during the pre-processing phase in order to eliminate chaotic data and enhance precision. Price was chosen as the primary data during the pre-processing phase so that MinMaxScaler could be applied. The objective of MinMaxScaler is to transform the Price data, so it is scaled between 0 and 1. MinMax Scaler has been used as a normalization method, as presented in Equation (1). A summary of the training and testing data is presented in Table 2.

$$x_{scaled} = \frac{x - x_{min}}{y_{max} - y_{min}} \quad (1)$$

TABLE II
TRAINING AND TESTING DATA

Method	Windows Length (Days)	Training Set (80%)	Testing Set (20%)
LSTM CNNs	1, 3, and 7	1460	366

C. Long Short-Term Memory (LSTM)

Hochreiter invented the LSTMnd has since been enhanced and used by several researchers [22], [23]. The LSTM, on the other hand, is regarded as an improved RNN model. An RNN contains connections between its hidden layers as opposed to a typical neural network. Thus, in addition to the input from the input layer, the input of the hidden layer also includes the output of the hidden layer from the previous iteration. Figure 3 demonstrates how a recurrent neural network is used. Although LSTM and RNN have comparable extended structures, there are differences in the hidden layer's memory cell structure. According to the author, the forget, inputand output gate have all been added to the LSTM's hidden layer memory cell based on the RNN's structure [24]–[26]. The design of these three distinct gate structures effectively overcomes the vanishing gradient issue, making it particularly appropriate for resolving issues with long-term dependency.

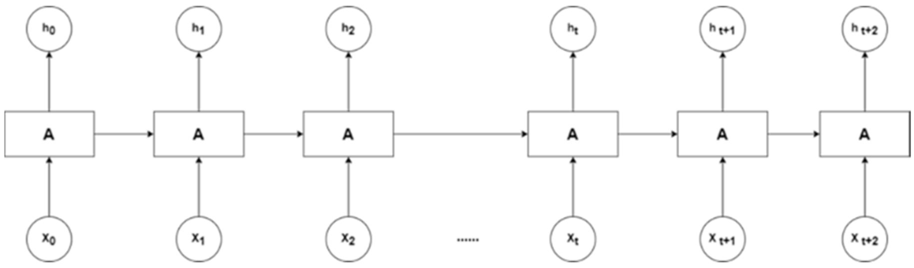


Fig. 3 The Expended Recurrent Neural Network Structure [26]

Figure 3 starts by eliminating X_0 from the input list before creating h0, which functions as an input along with X_1, for the subsequent step. The input for the next step is thus h0 and X_1. The input with X_2 for the subsequent phase is h1 from the previous step, and so on. In this way, it retains context-specific memory while learning. The Equation (2) applicable to the current situation is.

$$h_t = f\ (h_{t-1}, X_t) \tag{2}$$

$$h_t = tanh\ (W_{hh}\ h_{t-1} + W_{xh} x_t) \tag{3}$$

$$y_t = W_{hy}\ h_t \tag{4}$$

W stands for the weight, h stands for the single hidden vector, W_{hh} denotes as the weight from the previous hidden state, W_{hx} symbolizes the weight from the current input state, and $tanh$ is the activation function that implements a nonlinearity that constricts the activations to the range. [- 1.1]. Equation (3) illustrates how an RNN's activation function is used. Weight in the output state is W_{hy} and the output equation (4) represents the out state as y_t.

The following eight equations may be used to illustrate how the LSTM processes information based on Fig 4 below.

$$i_t = \sigma(W_i x_t + H_i h_{t-1} + b_i) \tag{5}$$

$$f_t = \sigma(W_f x_f + H_f h_{t-1} + b_f) \tag{6}$$

$$o_t = \sigma(W_o x_f + H_o h_{t-1} + b_o) \tag{7}$$

$$c_{\tilde{t}} = tanh(W_c x_t + H_c h_{t-1} + b_c) \tag{8}$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \tag{9}$$

$$h_t = o_t * tanh\ tanh\ (c_t) \tag{10}$$

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{11}$$

$$\tanh(x) = \frac{a^x - a^{-x}}{e^x + e^{-x}} \tag{12}$$

Each symbol has a distinct representational meaning when seen from both sides of the Equation. The input gate, denoted as i_t, plays a crucial role in selecting which data within the cell needs to be modified. On the other hand, the forget gate, symbolized as f_t determines which information should be discarded from the cell. The output gate, represented by o_t , regulates the amount of information to be output. The symbol c_t represents the candidate value for the memory cell's state at a given time, t. The current state of the memory cell, denoted as c_t, is calculated by combining i_t and c_t using element-wise multiplication (*) and f_t and c_{t-1} , the previous state of the memory cell. This combination allows for the updating of the memory cell's state. The symbol h_t corresponds to the output value that is filtered by the output gate o_t. This filtered output reflects the relevant information from the memory cell. To ensure appropriate transformations, the sigmoid function, denoted by σ, is employed to squash values between 0 and 1. Additionally, the hyperbolic tangent function, tanh, is used to limit values between -1 and 1. The symbol x_t represents the input provided to the memory cell at time t, which is subject to the gating operations described above. Weight matrices, specifically $W_i, W_f, W_o, W_c, H_i, H_f, H_o,$ *and* H_c are used to multiply the respective inputs and gate values, while bias vectors, b_i, b_f, b_o, *and* b_c, are added to the resulting calculations.

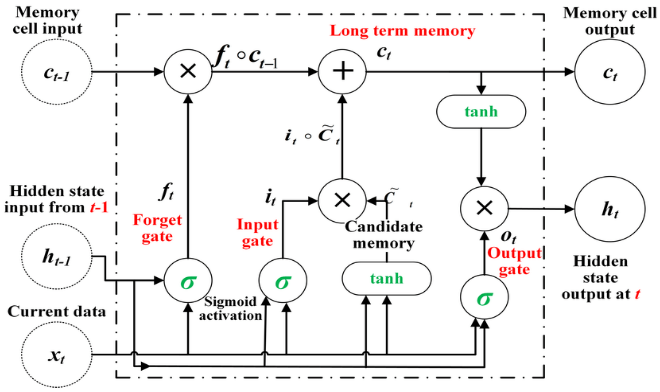


Fig. 4 The hidden layer LSTM memory cell structure [24]

D. Convolutional Neural Networks

It is well known that this model has been used to recognize patterns and features in photos, videos, and audio signals and that it did so with surprisingly excellent results [28]. In this study, 1D CNN models presented in Fig 5, where the data and kernels are one-dimensional vectors, were taken into consideration.

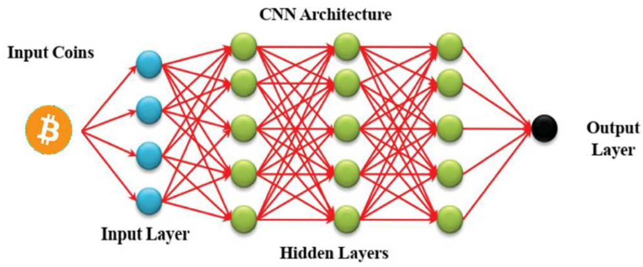


Fig. 5 Basic Convolutional Neural Networks Architecture [15]

In order to learn hierarchical representations of the input data, CNNs make use of a number of essential components [27]. Beginning with one or more convolutional layers, CNNs conduct convolution to extract local features by sliding filters or kernels over the input data [28]. After each convolutional layer, activation functions, often ReLUs (Rectified Linear Units), induce nonlinearity. These layers aid in the network's learning of spatial pattern hierarchies by collecting pertinent elements at various levels of abstraction. Following the convolutional layers, regularization is a process that makes minor adjustments to the algorithm to improve the generalization of the model. This subsequently enhances the model's efficacy on untested data. This method facilitates translation invariance and reduces computational complexity.

The layers of keras.Conv1D accepts multiple parameters, including filters, kernel_size, activation, kernel_regularizer, whereas keras.layers accepts only one parameter. The specification list at below:

- Filters: The number of filters (or kernels) that will be utilized in the convolutional layer.
- input_shape: represents the number of time steps in the input data.
- kernel_size: Convolution filters size.
- activation: Activation function used on the convolution layer's output.

- kernel_regularizer: Regularizer to penalize the kernel of the layer.

Table 3 presents the parameters set to the above values for the five instances of the convolutional layers.

TABLE III
CNNs PARAMETER

Filters	Kernel_Size	Activation	Input Shape	Regularization
64	1	relu	(tiem_step, 1)	Dropout (0.2) L2 = 0.01

The resultant feature maps are then converted to a 1-dimensional vector and transmitted through layers with complete connectivity. Connecting every neuron from the preceding layer to every neuron in the current layer enables intricate transformations and the acquisition of representations at a higher level. These layers perform computations on the flattened features, resulting in more abstract and meaningful representations. The final layer with full connectivity is the output layer, which generates the network's predictions. The activation function used in the output layer is dependent on the task. A linear activation function may be used for regression tasks, whereas softmax is frequently used for classification tasks to generate probability scores. CNNs learn optimal weights through backpropagation during training. This procedure entails calculating the gradient of the loss function concerning the network weights and adjusting the weights to minimize the loss. Optimizers such as Adam are employed in this optimization procedure.

E. Evaluation Metrics

There are two important characteristics which involve accuracy of the value prediction and direction predictions. These two characteristics would reflect the various neural networks to predict the Bitcoin price. In this study, root mean squared error (RMSE), mean absolute error (MAE), R-squared (R^2), and mean squared error (MSE) are adopted to serve as the evaluation metrics of the cryptocurrency prediction models. All evaluation metric defined as

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\tilde{p}_i - p_i)^2 \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{p}_i - p_i)^2} \quad (14)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\tilde{p}_i - p_i| \quad (15)$$

$$R^2 = 1 - \frac{\Sigma(p_i - \hat{p}_i)^2}{\Sigma(p_i - \bar{p})^2} \quad (16)$$

where N represents the total number of predictions, p represents the actual price, and $\tilde{p}$ represents the predicted price. The hyperparameter values with the lowest MAE, MSE, RMSE are used as the best settings, the prediction's accuracy will be greater.

III. RESULT AND DISCUSSION

This section presents the results of comparing the two proposed LSTM and CNNs models. These two models are trained on the same Bitcoin historical data with the same timestamp. At the same time, we likewise examine the prediction performance of the model. The cryptocurrency was trained utilizing three different Windows Length which are 1 day, 3 Days, and 7 Days. In the following Table 4 is presented the comparison result of the models on the Bitcoin dataset. The evaluation metrics, MSE, MAE, RMSE are implemented to evaluate the model performance. All models exhibited similar performance, regarding the performance metrics RMSE, MSE, MAE, and R^2.

TABLE IV
EXPERIMENTAL RESULT FOR BITCOIN HISTORICAL DATASET

Method	Windows Length (Days)	RMSE	MSE	MAE	R^2
LSTM	1	1203.97	1449536.48	875.06	0.986
	3	1311.71	1720574.33	926.96	0.980
	7	1645.36	2707216.78	1191.05	0.973
CNNs	1	1107.77	1227161.09	801.50	0.988
	3	1699.56	2888501.37	1363.85	0.972
	7	1670.93	2792011.33	1197.38	0.972

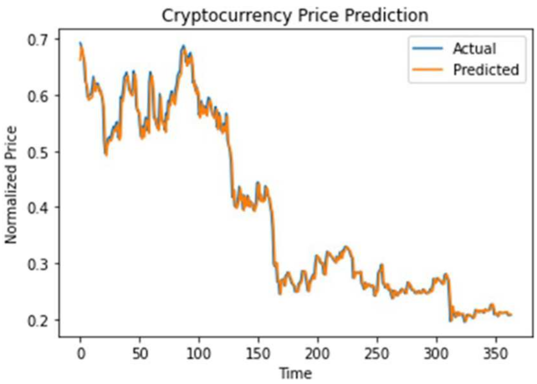
When looking at the RMSE, MSE, and MAE values for the LSTM approach, we observe that when the window length goes from 1 day to 7 days, there is a slow but steady rise in all three of these values. The LSTM model reported with a 1-day window length achieves an RMSE of 1203.97, MSE of 1,449,536.48, MAE of 875.06, and an R^2 value of 0.986. The LSTM model's performance marginally deteriorates over a 3-day window length in comparison to a 1-day window length. The results show an RMSE of 1311.71, MSE of 1,720,574.33, MAE of 926.96, and an R^2 value of 0.980. Among the tested window lengths, the LSTM model with a 7-day window length shows the largest prediction errors. It obtains a RMSE of 1645.36, an MSE of 2,707,216.78, an MAE of 1191.05, and R^2 value of 0.973.This suggests that the 3 days and 7 days window length could indicate more noise or irrelevant information into the model, leading to somewhat less accurate predictions. Nevertheless, even with a window length of 7 days, the model still achieves reasonably low prediction errors. This is an essential point to bring up, and it bears repeating.

Overall, the proposed model of CNNs where the windows length with 1 days indicates the lowest result among the all the output in CNNs, which are RMSE consists of 1107.77, MSE is 1227161.09 and MAE is 801.50, whereas the proposed model of CNNs with 3 day's windows length has the highest value in RMSE (1699.56), MSE (2888501.37) and MAE (1363.85). With a 7 days window length, the CNNs performance is slightly decreased compared with the 3-days windows length, it achieves an RMSE of 1670.93, MSE of 2792011.33, MAE of 1197.38, and an R^2 value of 0.972. More specifically, the extended window lengths are able to bring forth more complications that the CNNs model finds difficult to adequately describe.

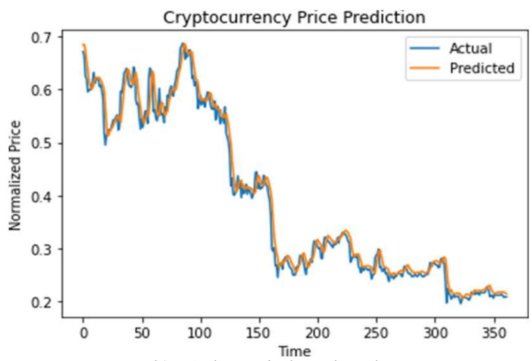
Despite this, when comparing the LSTM and CNNs model, we notice that the LSTM model generally outperforms the CNNs model in terms of prediction accuracy. The LSTM method consistently achieves lower RMSE, MSE, and MAE values across all window lengths, indicating better overall performance. However, to be mentioned that it is worth nothing that the differences in performance between LSTM and CNNs model are relatively small, where the RMSE values for the LSTM model range from 1203.97 to 1645.36, while the CNNs model ranges from 1107.77 to 1670.93.

By examining the R^2 value, it interprets which model had the best fit performance, R^2 was used as the overall model performance metric. To compare these two models, the R2 values for LSTM range from 0.973 to 0.986, while for CNNs, they range from 0.972 to 0.988. In addition, the R^2 results for both models indicate a score that is substantially higher than 0.90 and close to 1. From here, we observed that despite the fact that the R^2 statistic shows that all models have similar metrics value, these similarities suggest that both methods are capable of providing accurate forecasts and perform a strong fit to the Bitcoin historical data, with LSTM having a slight advantage.

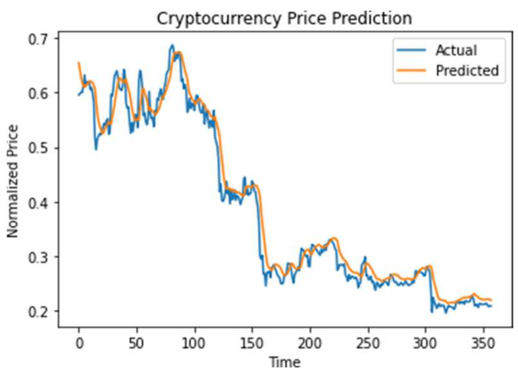
Fig 7 and 8 illustrate the graph output for the Bitcoin historical price prediction. The blue line represents the actual Bitcoin historical price, while the orange line represents the predicted Bitcoin price. By looking at Fig 7 and 8, it is hard to differentiate between the two model prediction results because the difference between the actual and predicted values is very minute. Nonetheless, based on our observation, the 1-day window length for the LSTM and CNNs model is able to perform an efficient prediction with higher accuracy, whereas the 3-day and 7-day window lengths might be predicted the trend, but likely differ from the actual.



a) 1 day windows length

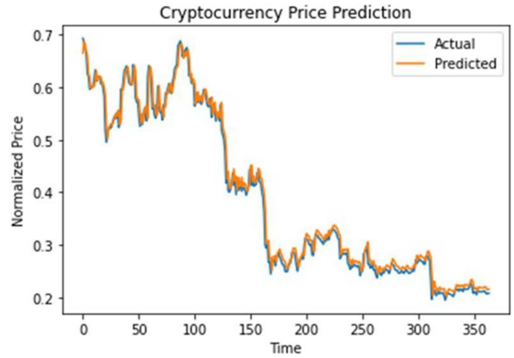


b) 3 days windows length

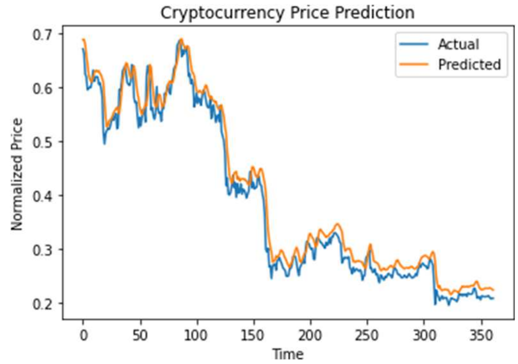


c) 7 days windows length

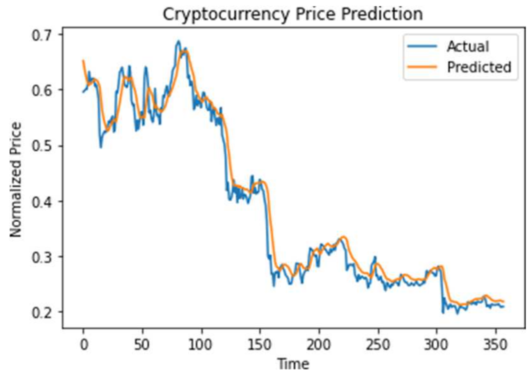
Fig. 7 Predicted results of LSTM with different windows length



a) 1 day windows length



b) 3 days windows length



c) 7 days windows length

Fig. 8 Predicted results of CNNs with different windows length

Overall, based on the analysis above, it is possible to observe that both the LSTM and CNN models outperform one another somewhat more in short-term forecasts than in long-term predictions. The pattern of rising prediction errors (RMSE, MSE, MAE) with increasing window lengths is consistent with this fact.

IV. Conclusion

In this paper, deep learning techniques were applied to predict the price of cryptocurrency with the employment of LSTM and CNNs models. The historical price of cryptocurrency is acquired, namely Bitcoin. The collected data are converted, pre-processed, and normalized to filter the unnecessary data. The LSTM and CNNs models are then loaded with the pre-processed data. LSTM models transform the time series data into input-output sequences appropriate for LSTM model training. After that, CNN's architecture will handle the input representation successfully, and depending on experimentation and the task difficulty, additional hyperparameters, like the number of layers and their sizes, may need to be changed. To improve the model, we have implemented the parameter regularization techniques to reduce overfitting to quite an extent. By comparing the RMSE, MSE, and MAE, the implementation highlights that the LSTM has slightly better performance than CNNs, where the result is shown in Table 4. Overall, the results indicate that both the LSTM and CNNs methods are effective for forecasting, with the LSTM method performing marginally better, supported by the result analysis above.

The limitation conducted in this study is due to the high volatility of Bitcoin data, which gives rise to the complexity of performing the consistency of the evaluation metrics. Where been experiment in this study is the values of RMSE, MSE, and MAE contribute too large values compared to previous research [11], [15]. The highest value of evaluation metrics might indicate issues with forecasting large average errors compared to the actual values. On the other hand, since how the implementation was carried out was different, and the dataset that was chosen was different, it may carry out various outcomes. The applicability of LSTM and CNN models for short-term or long-term predictions can vary based on the dataset, the specific forecasting assignment, and other variables such as the architecture and hyperparameters.

In future work, except the price of the cryptocurrency, other features such as external factors, law and regulation, trading volume per day, and seasonality trends can serve as the input features for the prediction to train more accurately patterns to gain more rich and valuable insight into the dynamics between cryptos, at the meantime also improve the forecast accuracy of the prediction model. Additionally, we might consider combining multiple models or predictions using ensemble methods[29]–[31]. Ensemble techniques, such as tagging or boosting, can reduce the risk of overfitting and enhance overall performance by capitalizing on the strengths of multiple models.

REFERENCES

[1] A. Lewis, *The Basics of Bitcoins and Blockchains. An Introduction to Cryptocurrencies and the Technology that Powers Them*. Mango Media Inc., 2021.

[2] J. Krusling, "Research Guides: Blockchain and Cryptocurrency: Cryptocurrency." Accessed: October 12, 2023. [Online]. Available: https://ggu.libguides.com/c.php?g=1157511&p=8946880

[3] S. Cruickshank, "The future of finance? A look at bitcoin's boom and how cryptocurrency works." Accessed: October 12, 2023. [Online]. Available: https://hub.jhu.edu/2018/01/08/what-is-bitcoin-and-how-does-it-work/

[4] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." Accessed: October 11, 2023. [Online]. Available: https://bitcoin.org/en/bitcoin-paper

[5] Z. Li, Q. Lu, S. Chen, Y. Liu, and X. Xu, "A Landscape of Cryptocurrencies," IEEE, Oct. 2023. doi: 10.1109/bloc.2019.8751469.

[6] S. Biswas, M. Pawar, S. Badole, N. Galande, and S. Rathod, "Cryptocurrency Price Prediction Using Neural Networks and Deep Learning," IEEE, Oct. 2023. doi: 10.1109/icaccs51430.2021.9441872.

[7] S. Tanwar, N. P. Patel, S. N. Patel, J. R. Patel, G. Sharma, and I. E. Davidson, "Deep Learning-Based Cryptocurrency Price Prediction Scheme With Inter-Dependent Relations," *IEEE Access*, vol. 9, pp. 138633–138646, Oct. 2023, doi: 10.1109/access.2021.3117848.

[8] W. Yiying and Z. Yeze, "Cryptocurrency Price Analysis with Artificial Intelligence," IEEE, Oct. 2023. doi: 10.1109/infoman.2019.8714700.

[9] S. M. Raju and A. M. Tarif, "Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis," Jun. 2020. Accessed: October 19, 2023. [Online]. Available: https://arxiv.org/abs/2006.14473

[10] W. Zhengyang, L. Xingzhou, R. Jinjin, and K. Jiaqing, "Prediction of Cryptocurrency Price Dynamics with Multiple Machine Learning Techniques," ACM, Oct. 2023. doi: 10.1145/3340997.3341008.

[11] I. M. Wirawan, T. Widiyaningtyas, and M. M. Hasan, "Short Term Prediction on Bitcoin Price Using ARIMA Method," IEEE, Oct. 2023. doi: 10.1109/isemantic.2019.8884257.

[12] S. D. I. of E. and S. H. Department of CSE India, G. Vidyulatha, M. Mounika, and N. Arpitha, "Crypto Currency Prediction Model using ARIMA," *Turkish Journal of Computerand Mathematics Education*, vol. 11, no. 3, pp. 1654–1660, 2020, [Online]. Available: https://turcomat.org/index.php/turkbilmat/article/view/11754/8594

[13] A. Azari, "Bitcoin Price Prediction: An ARIMA Approach," Apr. 2019, [Online]. Available: http://arxiv.org/abs/1904.05315

[14] C. Y. Kang, C. P. Lee, and K. M. Lim, "Cryptocurrency Price Prediction with Convolutional Neural Network and Stacked Gated Recurrent Unit," *Data (Basel)*, vol. 7, no. 11, p. 149, Oct. 2023, doi: 10.3390/data7110149.

[15] I. E. Livieris, N. Kiriakidou, S. Stavroyiannis, and P. Pintelas, "An Advanced CNN-LSTM Model for Cryptocurrency Forecasting," *Electronics (Basel)*, vol. 10, no. 3, p. 287, Oct. 2023, doi: 10.3390/electronics10030287.

[16] Y. Li and W. Dai, "Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model," *The Journal of Engineering*, vol. 2020, no. 13, pp. 344–347, Oct. 2023, doi: 10.1049/joe.2019.1203.

[17] K. Singh and A. Mohd Syuhaidi Abu Bakar, "The Role Of Higher Learning Institutions in Developing B40 & M40 Communities Living Near Higher Education Institutions," 2021. [Online]. Available: https://www.researchgate.net/publication/358137709

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, Oct. 2023, doi: 10.1038/nature14539.

[19] Y.-C. Wen and A. Hongchindaket, "Interviewing Young Adults about Cryptocurrency Literacy and Investment Intention," IEEE, Oct. 2023. doi: 10.1109/icbir57571.2023.10147709.

[20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol. 60, no. 6, pp. 84–90, Oct. 2023, doi: 10.1145/3065386.

[21] J. Brownlee, "How to Develop LSTM Models for Time Series Forecasting," MachineLearningMastery.com. Accessed: October 19, 2023. [Online]. Available: https://machinelearningmastery.com/how-to-develop-lstm-models-for-time-series-forecasting/

[22] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, Oct. 2023, doi: 10.1162/neco.1997.9.8.1735.

[23] F. Qian and X. Chen, "Stock Prediction Based on LSTM under Different Stability," IEEE, Oct. 2023. doi: 10.1109/icccbda.2019.8725709.

[24] T. Zebin, M. Sperrin, N. Peek, and A. J. Casson, "Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks," IEEE, Oct. 2023. doi: 10.1109/embc.2018.8513115.

[25] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, "A New Forecasting Framework for Bitcoin Price with LSTM," IEEE, Oct. 2023. doi: 10.1109/icdmw.2018.00032.

[26] M. Wildan Putra Aldi, Jondri, and A. Annisa, "Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin," vol. 5, no. 2, pp. 3548–3555, 2018.

[27] S. Kazeminia, H. Sajedi, and M. Arjmand, "Real-Time Bitcoin Price Prediction Using Hybrid 2D-CNN LSTM Model," IEEE, Oct. 2023. doi: 10.1109/icwr57742.2023.10139275.

[28] A. Naseer, E. Nava Baro, S. Daud Khan, Y. Vila, and J. Doyle, "A Novel Cryptocurrency Prediction Method Using Optimum CNN," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1051–1063, Oct. 2023, doi: 10.32604/cmc.2022.020823.

[29] P. C. Sekhar, M. Padmaja, B. Sarangi, and Aditya, "Prediction of Cryptocurrency using LSTM and XGBoost," IEEE, Oct. 2023. doi: 10.1109/icbds53701.2022.9935871.

[30] W. F. Hidayat, M. F. Julianto, Y. Malau, A. Setiadi, and Sriyadi, "Implementation of LSTM and Adam Optimization as a Cryptocurrency Polygon Price Predictor," IEEE, Oct. 2023. doi: 10.1109/icitri59340.2023.10249571.

[31] D. G. Taslim and I. M. Murwantara, "A Comparative Study of ARIMA and LSTM in Forecasting Time Series Data," IEEE, Oct. 2023. doi: 10.1109/icitacee55701.2022.9924148.