Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 45 (2024), No. 4, pp. 1061–1072
DOI : 10.47974/JIOS-1691

A hybrid deep learning model for accurate time series forecasting of cryptocurrencies

Aditya Nagdiya *
Vivek Kapoor †
Vrinda Tokekar §
Department of Information Technology
Institute of Engineering & Technology
Devi Ahilya Vishwavidyalaya
Indore
Madhya Pradesh
India

Abstract

A growing number of people and organizations are choosing to invest in cryptocurrencies. The development of precise forecasting models for cryptocurrencies is crucial due to their very volatile market. Financial forecasting has long made use of time series analysis and prediction, but conventional time series analysis techniques have trouble capturing intricate patterns and nonlinear relationships. On the other hand, although deep learning models show promise in time series analysis, their effectiveness depends on large amounts of data, which might result in overfitting. To predict Bitcoin prices, this research presents a hybrid model that combines long short-term memory (LSTM) and convolutional neural networks (CNN), where the CNN extracts features from the time series data, while the LSTM captures persistent patterns over time.

Subject Classification: *Primary 93A30, Secondary 49K15.*

Keywords: *Deep learning, Hybrid models, Time series forecasting.*

1. Introduction

The surge in cryptocurrencies has sparked increased enthusiasm for the development of precise forecasting models tailored to these digital assets. Notably, cryptocurrencies like Bitcoin, Ethereum, and Litecoin have

* *E-mail:* `aditya22.ad@gmail.com (Corresponding Author)`
† *E-mail:* `vkapoor@ietdavv.edu.in`
§ *E-mail:* `vrindatokekar@yahoo.com`

garnered attention for their decentralized structure and the possibility for substantial returns. Despite these advantages, the crypto market is characterized by high volatility, posing a significant challenge in accurately predicting future price movements. Accurate price prediction can assist investors in making informed decisions, minimize their risks, and maximize their returns. Therefore, there is a growing interest in developing models for Cryptocurrencies price prediction [1], [2].

Traditional and conventional time series methodologies, such as ARIMA [3], exponential smoothing, and autoregressive models [4], have been used to predict the prices of cryptocurrencies. However, these models have limitations in capturing the complex patterns, dependencies and relationship in the data [5]. Conversely, deep learning models, exemplified by the Long Short-Term Memory (LSTM) neural network, have demonstrated potential in the realm of time series based analysis and prediction. [6]. These models excel in capturing intricate nonlinear relationships and complex patterns within the data, showcasing commendable performance across diverse domains, including finance. Nevertheless, it is imperative to recognize that deep learning models demand substantial data volumes, and there exists a risk of overfitting, particularly in scenarios where the dataset is limited in size. [7]. In recent time, deep learning models have demonstrated promising results in forecasting the highly volatile crypto price [8]. Deep learning models, including recurrent neural networks (RNNs) [9], convolutional neural networks (CNNs) [10], and LSTM [11], have been employed to discern and capture both the temporal and spatial patterns within the data.

In this paper, we suggest a hybrid model that merges a CNN and LSTM to forecast the prices of cryptocurrencies. The CNN is used to extract important features from the time series data, while the LSTM is used to capture prolonged dependencies within the crypto price data. The proposed model is evaluated on two datasets of daily closing prices of Bitcoin and Ethereum. Therefore, the main aim of this study is to introduce a model for cryptocurrency price prediction that combines traditional time-series-based analysis and various deep learning methodologies to capture the linear as well as non-linear relationships within the crypto price data and enhance the price forecasting accuracy.

In chapter 2, we present an overview of previous studies on time-series-based forecasting of cryptocurrencies using deep learning. In chapter 3, we describe the methodology of our hybrid CNN- LSTM model. In chapter 4, we present the model training process and result analysis of the experiments conducted on the Bitcoin dataset. Finally in chapter 5, we conclude the paper, with implications and potential applications of our model and also provide recommendations for future research.

2. Related Work

Different conventional time series based methodologies, including ARIMA, exponential smoothing, and moving averages, have been applied to cryptocurrency price prediction. For example, Poongod et al. (2020) [3] proposed an ARIMA-based model for bitcoin price prediction, and they got 49% accuracy in a simulation environment. In recent years, several studies have investigated the use of deep learning models for time-series-based prediction of cryptocurrencies [8], [12]. Many of these studies have used RNN-based models, like LSTM and gated recurrent units (GRU) [13], or genetic algorithm [14] to understand the temporal relationships within the data [15], [16], on relatively stable stock price data [17].

However, there is limited research on the models that combine CNN with LSTM for forecasting cryptocurrency prices. One of the earliest studies that used LSTM for btc price prediction was conducted by karakoyun et al. (2018) [18]. Here the btc closing prices were predicted using LSTM and achieved better results contrasted with conventional models such as ARIMA and exponential smoothing. In another study, politis et al. (2021) [19] used LSTM and GRU for price of Ethereum. They found that the LSTM-GRU model surpassed the other models, achieving an accuracy of up to 78.9% in forecasting the direction of price movements. More recently, Sridhar et al. (2021) [20] proposed a model that uses self-attention to forecast the prices of dogecoin. They evaluated their model on the hourly dogecoin price variation and model gave an impressive accuracy of 98.46%. While the aforementioned studies have shown promising results using RNN-based models, there is limited research on the use of combined CNN-LSTM models for forecasting cryptocurrency prices. A few studies have used CNNs to extract features from the time-series-based data before feeding it to the LSTM for further processing. For instance, Li et al. (2020) [21] used a CNN-LSTM model to forecast the prices of Bitcoin and achieved better results in comparison to different other deep learning models. In general, existing literature indicates that deep learning models have the capability to surpass traditional and conventional time-series-based forecasting models in predicting cryptocurrency prices. However, there is a requirement for more research on combined CNN-LSTM models for this task.

3. Hybrid CNN-LSTM model

The proposed model for forecasting cryptocurrency time series, employing both CNN and LSTM, consists of two primary components: a CNN layer and an LSTM layer. The CNN is tasked with extracting characteristics from the time series based crypto data, while the LSTM focuses on capturing long-term dependencies within the data.

In the CNN layer, a sequence of historical prices serves as input, undergoing convolutional filtering to extract relevant features. These filters systematically scan the sequential input of historical prices to identify and extract distinctive patterns or features. Each filter focuses on recognizing specific temporal patterns, capturing relevant information such as trends, fluctuations, or other key characteristics present in the price data. The convolutional operation involves sliding the filters across the input sequence, performing localized computations at each step. Through this process, the CNN effectively distills essential features from the time based crypto data, creating a transformed representation that emphasizes important aspects for LSTM layer. The output from the CNN layer is then directed into the LSTM layer, which is designed to capture enduring dependencies within this hitorical price data which is in time-series-based format. The LSTM's ability to selectively retain and discard information, coupled with its recurrent connections, enables it to capture and remember long-term dependencies in the time series data. By processing the features extracted by the CNN layer through its memory cell and gates, the LSTM can learn and retain important patterns and relationships that persist over extended periods in the cryptocurrency price data. This combination of CNN with LSTM allows the model to address both short-term and long-term dynamics in the time series. As a countermeasure against overfitting, dropout is implemented on the LSTM layer. This involves randomly setting a portion of the LSTM units to zero during training, compelling the remaining units to acquire more resilient and generalizable representations. The output of the LSTM layer is subsequently fed into a fully connected layer with a linear activation function, which generates the forecasted price for the next time step. In essence, the suggested model architecture amalgamates the advantages of CNN and LSTM to extract significant features from the crypto data and capture long-term dependencies, respectively, while preventing overfitting with dropout. Equations that explain the above explained structure of hybrid CNN-LSTM model:

$$X = [x1, x2, ..., xn] \tag{1}$$

$$Ci = \mathrm{Conv1D}(X, Wi) \tag{2}$$

$$Pi = \mathrm{MaxPooling1D}(Ci) \tag{3}$$

$$H = \mathrm{LSTM}(P1, P2, ..., Pk) \tag{4}$$

$$Y = \mathrm{Dense}(H, Wy) \tag{5}$$

Here in equation (1) X is the input time series of cryptocurrency prices, where n is the length of the time series. In equation (2) Ci is the result of applying the i-th convolution filter Wi to the input time series X using the 1D convolution operation (Conv1D). In equation (3) Pi is the result of applying the max pooling operation (MaxPooling1D) to Ci to reduce its dimensionality. In equation (4) H is the output of the LSTM layer (LSTM), which takes as input the pooled outputs $P1, P2, ..., Pk$ from the convolutional layer. In equation (5) Y is the final output of the model, which is obtained by applying a dense layer (Dense) to the output of the LSTM layer H using weight matrix Wy.

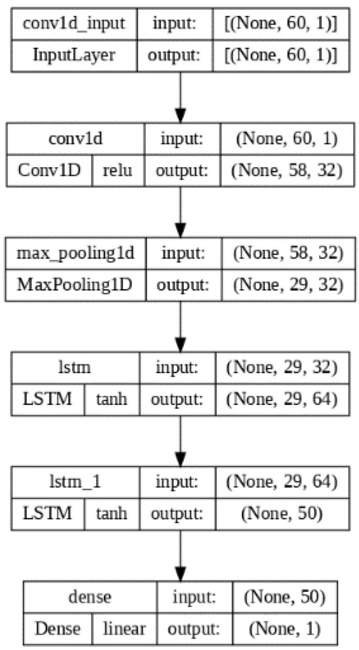


Figure 1

Architecture of the hybrid CNN-LSTM model proposed

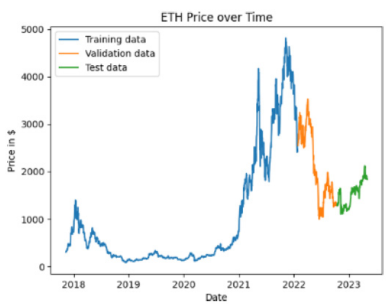
The proposed CNN-LSTM architecture combines the capabilities of CNNs and LSTMs to capture both local and global temporal patterns in the input time series. The convolutional layer applies filters with small receptive fields to capture local patterns in the time-series-based data, while the LSTM layer uses its memory cells to capture longterm dependencies in it.

The proposed model is a fusion of deep learning techniques, incorporating both CNN and LSTM. The CNN layer comprises 32 filters with a kernel size of 3 and a stride of 1, then next to it a max pooling layer with a pool size of 2. The resulting output from the max pooling layer is then inputted into the LSTM layer, featuring 64 units and a dropout rate of 0.2. Following this LSTM layer, there is another LSTM layer which consists of 50 units and a dropout rate of 0.2. Subsequently, lastly the dense layer, equipped with a single output neuron and a linear activation function, ultimately generating the predicted price for the subsequent day. The detailed architecture is visually presented in "Figure 1".

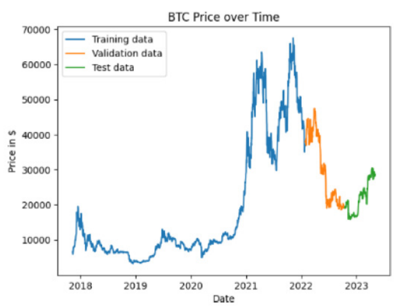
4. Model Training and Evaluation

The first step in building any time series model is data preprocessing. In this study, we used a datasets of daily closing prices of Bitcoin and Ethereum, obtained from Yahoo Finance. Both the datasets consist of 2000 observations from November-2017 to May-2023. The btc data is illustrated in "Figure 2a". The eth data is illustrated in "Figure 2b".

We pre-processed the data by first normalizing the prices using the MinMaxScaler function from the scikit-learn library. This ensured that all the values remained between 0 and 1, which is important for the CNN



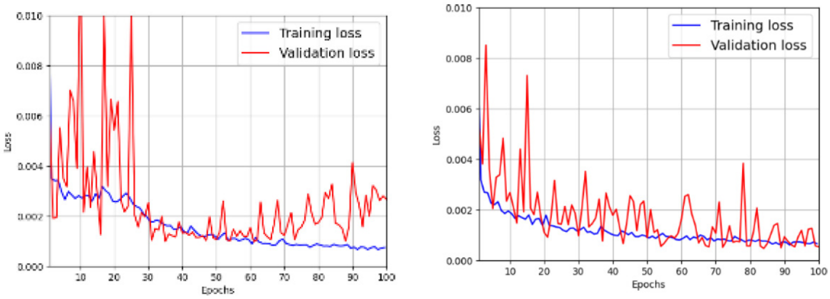
(a) Bitcoin data split



(b) Ethereum data split

Figure 2

Training, validation and test data split



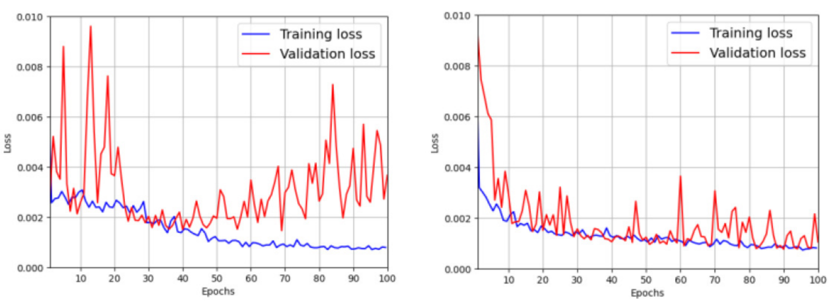
(a) LSTM Model (b) Proposed Model

Figure 3

Training and validation losses on bitcoin price

model to extract meaningful features. The dataset was partitioned into training, validation, and testing sets, maintaining 77:13:10 ratio, respectively.

To evaluate and contrast the performance of our model, we constructed an additional LSTM model with identical last two LSTM layers as those in the our above described model. In this LSTM model, the initial LSTM layer comprises 64 units with a dropout rate of 0.2. Following this LSTM layer, there is another LSTM layer with 50 units and a dropout rate of 0.2. Subsequently, a dense layer follows, featuring a single output neuron and a linear activation function, generating the predicted price for the subsequent day. Both LSTM and the proposed models are separately trained on the preprocessed bitcoin data and ethereum data. The models underwent training using the Adam optimizer, employing a learning rate of 0.001 and a batch size of 32. The training spanned 100 epochs, and the



(a) LSTM Model (b) Proposed Model

Figure 4

Training and validation losses on ethereum price

optimal model was determined by selecting the one with the lowest mean squared error (MSE) on the validation set.

After training LSTM model, on bitcoin data we get the learning curves illustrated in “Figure 3a” and on ethereum data we get the learning curves illustrated in “Figure 4a”. Both learning curve plots illustrate that the training loss steadily decreases with increasing experience. However, the validation loss plot follows a pattern where it initially decreases but eventually starts to rise. The point at which the validation loss started increasing again is considered an inflection point, signaling a potential onset of overfitting. Training could be halted at or around this inflection point, as further experience beyond it may exhibit overfitting dynamics. After training our model, on bitcoin data we get the learning curves illustrated in “Figure 3b” and on ethereum data we get the learning curves illustrated in “Figure 4b”. Both the plots of learning curves show a good fit, as the plot of training loss and the plot of validation loss decreases to a point of stability. In both the plots, the validation curves are near to the training curves, which means that there is not too much overfitting.

Apart from these two deep learning models, we have also considered the traditional time series analysis by useing ARIMA model which captures the linear relationships in the data. The ARIMA model stands out as a widely used method in time series analysis, frequently employed in previous studies for predicting Bitcoin prices. This model consists of three crucial parameters: Autoregression (AR), Integration (I), and Moving Average (MA). Autoregression captures the linear relationship between the current value and its historical values, Integration involves differencing the data to achieve stationarity, and Moving Average accounts for the linear relationship between the current value and past errors. The selection of optimal ARIMA model parameters is guided by the Akaike Information Criterion.

In contrast, the our model undergoes evaluation on the testing set, and its outcomes are compared with those of both the ARIMA and LSTM models. Evaluation metrics encompass the Mean Squared Error, Root Mean Squared Error, Mean Absolute Error, and Mean Absolute Percentage Error. These metrics provide a measure of the model’s accuracy in predicting the bitcoin and ethereum prices. Tabel I presents the evaluation metrics for the three models using bitcoin price and Table II using ethereum price.

Table I

Model comparison on bitcoin

Model	MSE	RMSE	MAE	MAPE
ARIMA	49860746.2	7061.2	5980.5	23.29%
LSTM	7436897.1	2727.0	2451.4	10.1%
Hybrid	1017254.1	1008.5	769.5	3.3%

Table II

Model comparison on ethereum

Model	MSE	RMSE	MAE	MAPE
ARIMA	148658.1	385.5	335.2	19.51%
LSTM	46914.5	216.5	202.8	12.297%
Hybrid	7563.7	86.9	70.9	4.49%

The Table-I shows the results on bitcoin price, and it describes how our model is performing better than both the traditional ARIMA model and LSTM model, in all the evaluation metrics. The MSE of our hybrid model is 1017254.1, which is lower than the other two tradition models. The RMSE of our model is 1008.5, which is also lower than the traditional models. The MAE of our model is 769.5, which is also lower than both the traditional models. The MAPE of our model is 3.3%, which is also lower than the two traditional models. Similarly in Table II, on ethereum price also, our model is performing better than both the traditional models in terms of all evaluation metrics.

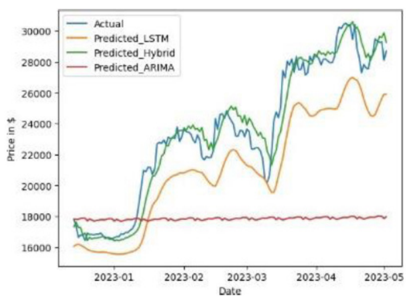


Figure 5

Bitcoin price predictions on test data

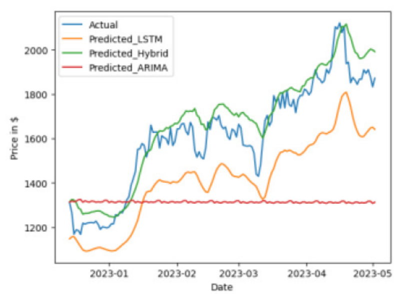


Figure 6

Ethereum price predictions on test data

We can see in "Figure 5", on bitcoin price test set and in "Figure 6" on ethereum price test set, our model most accurately predicts the bitcoin prices and ethereum prices respectively. In addition, we also plotted the predicted prices by other two models against the actual prices to visualize the performance on test data. We can see in these two figures; our model captures the trends and patterns present in the test data and provides accurate predictions for future prices.

5. Conclusion

In our study, we suggested a hybrid model for time-series-based forecasting of cryptocurrencies using CNN and LSTM. Our model was trained and evaluated on a dataset of daily closing prices of Bitcoin and Ethereum. Our model achieved better performance compared to traditional time series forecasting and deep learning models such as ARIMA and LSTM.

The model architecture integrates the advantages of CNN and LSTM, where the CNN layer's ability to extract significant features from the time-series-based data is responsible for the model's effectiveness, these features are then given to the LSTM layer, which uses them to identify long-term dependencies. The LSTM layer's usage of dropout contributes to the reduction of overfitting.

However, there are limitations to our proposed model. The model uses historical crypto price and does not incorporate other important features such as volume and volatility. In addition, the model was tested only on two cryptocurrencies and may not be generalizable to other cryptocurrencies price. The model was able to capture the overall trend of the both cryptocurrencies prices and performed well in predicting the turning points. However, there were some in- stances where the predicted prices deviated significantly from the actual prices.

In future additional features can also be incorporated into the model, like volume, social media sentiment and news articles, to improve forecasting accuracy. The model can also be tested on more other cryptocurrencies to determine its generalizability. The performance can also be investigated in different market conditions, such as bull and bear markets. There are more deep learning architectures, which can also be combined, such as attention-based models, for time-series-based analysis.

In conclusion, the proposed hybrid deep learning model shows promise in forecasting the highly volatile crypto price and it holds the potential for utilization in real-world applications such as trading, investment decision-making and portfolio management. Moreover, the hybrid approach of combining deep learning techniques can also be applied to different other time series analysis problems and domains, like in healthcare, transportation, and weather forecasting. However, further research is needed to improve its performance and generalizability.

References

[1] Zheshi Chen, Chunhong Li, and Wenjun Sun. Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, 365 : 112395 (2020).

[2] Vasily Derbentsev, Andriy Matviychuk, and Vladimir N Soloviev. Fore- casting of cryptocurrency prices using machine learning. *Advanced Studies of Financial Technologies and Cryptocurrency Markets*, pages 211–231 (2020).

[3] M Poongodi, V Vijayakumar, and Naveen Chilamkurti. Bitcoin price prediction using arima model. *International Journal of Internet Technology and Secured Transactions*, 10(4) : 396–406 (2020).

[4] Fa˘t Codrut,a Maria, Dezsi Eva, et al. Exchange-rates forecasting: Exponential smoothing techniques and arima models. *Annals of Faculty of Economics*, 1(1) : 499–508 (2011).

[5] Andreea-Cristina Petrica˘, Stelian Stancu, and Alexandru Tindeche. Limi- tation of arima models in financial and monetary economics. *Theoretical & Applied Economics*, 23(4) (2016).

[6] Yuxiu Hua, Zhifeng Zhao, Rongpeng Li, Xianfu Chen, Zhiming Liu, and Honggang Zhang. Deep learning with long short-term memory for time series prediction. *IEEE Communications Magazine*, 57(6) : 114–119 (2019).

[7] Jorg Bornschein, Francesco Visin, and Simon Osindero. Small data, big decisions: Model selection in the small-data regime. In International conference on machine learning, pages 1035–1044. PMLR, (2020).

[8] Suhwan Ji, Jongmin Kim, and Hyeonseung Im. A comparative study of bitcoin price prediction using deep learning. *Mathematics*, 7(10) : 898, (2019).

[9] Jan Koutnik, Klaus Greff, Faustino Gomez, and Juergen Schmidhuber. A clockwork rnn. In International conference on machine learning, pages 1863–1871. PMLR (2014).

[10] Leon O Chua and Tamas Roska. The cnn paradigm. IEEE Transactions on Circuits and Systems I: *Fundamental Theory and Applications*, 40(3) : 147–156 (1993).

[11] Yong Yu, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. A review of recurrent neural networks: Lstm cells and network architectures. *Neural computation*, 31(7) : 1235–1270 (2019).

[12] Mohil Maheshkumar Patel, Sudeep Tanwar, Rajesh Gupta, and Neeraj Kumar. A deep learning-based cryptocurrency price prediction scheme for financial institutions. *Journal of information security and applications*, 55:102583 (2020).

[13] Bhatnagar, Vaibhav, et al. "Descriptive analysis of COVID-19 patients in the context of India." *Journal of Interdisciplinary Mathematics* 24.3 : 489-504 (2021).

[14] Kapoor, V. & Dey, S. A genetic algorithm based decision support system for forecasting security prices in stock index, *Journal of Information and Optimization Sciences*, 43 : 8, 2153-2166 (2022).

[15] Kumar, Ankit, et al. "An enhanced quantum key distribution protocol for security authentication." *Journal of Discrete Mathematical Sciences and Cryptography* 22.4 : 499-507 (2019).

[16] Mohammad J Hamayel and Amani Yousef Owda. A novel cryptocurrency price prediction model using gru, lstm and bi-lstm machine learning algorithms. AI, 2(4):477–496 (2021).

[17] Shukla, Gargi Pant, Balwani, Nitin & Kumar, Santosh. Predicting stock prices with LSTM: A hybrid machine learning model for financial forecasting, *Journal of Information and Optimization Sciences*, 44:3, 575–584 (2023).

[18] Ebru Seyma Karakoyun and AO Cibikdiken. Comparison of arima time series model and lstm deep learning algorithm for bitcoin price forecasting. In The 13th multidisciplinary academic conference in Prague, volume 2018, pages 171–180 (2018).

[19] Agis Politis, Katerina Doka, and Nectarios Koziris. Ether price prediction using advanced deep learning models. In 2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pages 1–3. IEEE (2021).

[20] Sashank Sridhar and Sowmya Sanagavarapu. Multi-head self-attention transformer for dogecoin price prediction. In 2021 14th International Conference on Human System Interaction (HSI), pages 1–6. IEEE (2021).

[21] Yan Li and Wei Dai. Bitcoin price forecasting method based on cnn-lstm hybrid neural network model. *The Journal of Engineering*, 2020(13) : 344–347 (2020).