Short term price prediction of cryptocurrencies using technical indicators and neural networks architectures.

1st Leonardo Magallanes
Engineering Faculty
Universidad Panamericana
Aguascalientes, Aguascalientes
0236887@up.edu.mx

***Abstract*—This paper explores the efficacy of various neural network architectures in predicting short-term cryptocurrency prices, specifically Bitcoin. As the cryptocurrency market continues to expand, the need for accurate predictive models becomes increasingly critical. We employed a dataset enriched with technical indicators like Moving Averages, RSI, and MACD, alongside historical price data, to train and evaluate five different neural networks: Shallow Neural Network, GRUs, LSTMs, Transformer and a patch-mixing architecture. Our findings indicate that while these models can predict next-day closing prices with relatively low error, predictions of price direction (close-open) yield higher errors but are potentially more informative for trading strategies. This study not only underscores the capabilities and limitations of current deep learning approaches in financial predictions but also highlights the importance of selecting appropriate target variables to improve the practical utility of predictive models in the volatile cryptocurrency market.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

The Cryptocurrencies [1] market has witnessed exponential growth over time, with its total market capitalization now surpassing an impressive 2.5 trillion dollars. Bitcoin [2] is the biggest and leading cryptocurrency with a market capacity surpassing 1 trillion dollars. Predicting the closing price of Bitcoin is a challenging task given the inherent volatility of the cryptocurrencies market. Cryptocurrencies are widely known for the controversies and opinions of people trying to predict where the price will go next, making it challenging to trust the information sources. Unfortunately, the internet is awash with conflicting information, making it difficult to discern reliable sources from mere noise. Technical analysis is a tool used by traders and analysts to identify patterns and predict short-term trends using different technical indicators. These indicators include moving average [3], relative strength index (RSI), and more. Deep learning algorithms and architectures have proved astonishing results in detecting intricate patterns with vast datasets. Using deep learning algorithms we can focus on the raw data from the historical price movements of Bitcoin in order to find patterns and make predictions. The aim of this paper is to create a forecasting system to predict the closing price of Bitcoin in the next day.

Identify applicable funding agency here. If none, delete this.

The main goal is to create an effective forecasting system capable of predicting the closing price of Bitcoin for the next day. Achieving this requires a multi-step approach:

1) Optimal Dataset Selection: Create a dataset that combines technical indicators with the historical price chart of Bitcoin.
2) Neural Networks Architectures: Explore and compare various neural networks architectures, including feedforward networks, Recurrent Neural Networks(RNNs) [4], Long Short-Term Memory Networks (LSTMs) [5], Transformers and a state of the art model called Patchmixer [6].
3) Model Evaluation: Testing and validation is crucial. The model performance will be evaluated using metrics like Mean Absolute Error(MAE) [7] and Root Mean Squared Error(RMSE).

This paper is arranged as follows: beginning with this introduction, background and related works are explained in the immediate next section. The Materials and Methods are explained in Section 3 and Section 4 describes the experimental results of the proposed system. The conclusion described in the last section.

II. Background and Related Work

A. Stock and Cryptocurrency Prediction

Short-term trading and prediction within financial markets such as stock exchanges and cryptocurrency platforms present significant challenges due to inherent market volatilities influenced by fluctuating demands and supplies. Previous studies, such as those by Nagaraj Naik and Biju R Mohan, have highlighted the predictive potential of technical indicators in these environments [8]. These studies have traditionally focused on a limited set of technical indicators, but recent advancements have expanded this to include a broader range. For instance, Naik and Mohan utilized 33 different combinations of technical indicators, applying the Boruta feature selection technique to refine the selection process and enhance prediction accuracy .

B. Methodological Innovations in Financial Prediction

The emergence of advanced machine learning techniques, particularly those involving Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks, has been pivotal. These methods have proven to be highly effective in deciphering complex patterns in vast datasets, which is quintessential for predicting price movements in highly volatile markets like those of cryptocurrencies [**?**]. Notably, the implementation of ANN Regression prediction models has shown to significantly decrease error rates in stock price predictions [9]. Furthermore, the integration of LSTM with novel approaches such as the Random Connectivity LSTM (RCLSTM) model proposes potential reductions in computational costs while maintaining competitive prediction accuracies .

C. Applications in Cryptocurrency Markets

While the literature predominantly focuses on stock markets, similar methodologies are increasingly applicable to the cryptocurrency sector. Cryptocurrencies, such as Bitcoin, exhibit unique market dynamics and trading volumes that make them suitable candidates for these advanced predictive techniques. The adaptation of models tested in traditional financial systems to cryptocurrency prediction not only tests the robustness of these models but also extends their applicability to new, digital financial domains. Ji, Kim, and Im (2021) conducted a comprehensive study comparing various state-of-the-art deep learning methods for Bitcoin price prediction. Their work evaluated models including deep neural networks (DNN), long short-term memory (LSTM), convolutional neural networks (CNN), and deep residual networks (ResNet). They found that while LSTM-based models slightly outperformed others for price regression, DNN-based models were superior for predicting price direction. Their findings underscore the effectiveness of classification models over regression models for algorithmic trading, highlighting the diverse applications of deep learning in cryptocurrency markets. [10]

D. Challenges and Future Directions

Despite the progress in predictive modeling, significant challenges remain, particularly in terms of model accuracy and computational efficiency. The future of financial prediction likely lies in further refining these models to handle the specificities of different financial instruments, including emerging digital currencies. Moreover, enhancing the efficiency of these models while maintaining or improving their accuracy will be crucial for their adoption in real-time trading scenarios, where speed and reliability are paramount.

III. Proposed Methodology

A. Dataset

The dataset was extracted using the python library yfinance, which is a library that connects to yahoo finance through an API, and it allows downloading data easily. The historical bitcoin price data was extracted using a daily timeframe, spanning from April 05, 2015, to June 05, 2024. This gave us an initial dataset containing the following features:
Date.
Open $\rightarrow$ Initial bitcoin price from that day.
High $\rightarrow$ Highest bitcoin price from that day.
Low $\rightarrow$ Lowest bitcoin price from that day.
Close $\rightarrow$ Closing bitcoin price from that day.
Volume $\rightarrow$ Volume of transactions made that day.

B. Technical Indicators

Technical Indicators are a tool used by investors to help them make accurate evaluations of the market in order to have successful trades. There are a lot of different technical indicators that have been invented and used for time series analysis and day trading. We selected some of these indicators that we think were the most important.

- Relative Strength Index(RSI).
  Used to identify when a stock or cryptocurrency is being overbought or oversold. It assigns a value between 0 and 100. If the value is under 0.20 is usually considered oversold and could be a buying opportunity to traders and if there is equal or above 0.80 is being overbought and could be a sold signal for traders. The relative strength index was created by J. Welles Wilder Jr. in the late 1970s.
- Simple Moving Average(SMA).
  Simple moving averages are indicators that help us flatten the price line by calculating the average of a selected range of prices. These indicators are usually used with different time frames to predict changes in the price trend. We used 7 different simple moving averages with the following daily time frames: 5 days, 12 days, 25 days, 50 days, 100 days, 150 days and 200 days.
- Exponential Moving Average(EMA).
  Similar to simple moving average but the difference is that it gives different importance to the range data. EMA [11] gives higher importance weighting to recent prices. EMA are more reactive to the latest price changes. We used 7 different exponential moving averages with the following daily time frames: 5 days, 12 days, 25 days, 50 days, 100 days, 150 days and 200 days.
- Moving Average Convergence Divergence(MACD).
  MACD uses two different EMAs, a short-term EMA which is called fast line and a long-term EMA which is the slow line. The MACD line is the difference between the short-term and long-terms EMAs. Addittionally there is the signal line which is another EMA that has a time period even shorter than the fast line, this signal line is used to identify changes in price trends.

- Williams Percent Range(WILL).
  Is a momentum indicator indicator that moves between 0 and -100, similarly to RSI, it indicates overbought and oversold levels. It was created by

- Momentum (MOM). The Momentum indicator compares where the current price is in relation to where the price was in the past.

- Accumulation/Distribution It uses volume and price to asses if a stock or cryptocurrency is being accumulated or distributed.

Technical Indicator	**Formula**	**Time Periods(days)**
RSI	$100 - \frac{100-}{\frac{(1+Avg(gain))}{Avg(loss)}}$	6,15,28
SMA	$\frac{C1+C2+...+Cn}{n}$	5, 12, 25, 50, 100, 200
EMA	Price $\times\, k + \text{SMA}_y \times (1-k)$	5, 12, 25, 50, 100, 200
MACD	$SMA(n) - SMA(n)$	(12_26_9), (40_80_30)
WILL	$\frac{HighestHigh-Close}{HighestHigh-LowestLow}$	14, 28, 50, 100
MOM	MOM	5, 10, 100, 200
A/D	$\frac{(Close-Low)-(High-Close)}{High-Low}$	Whole dataset

TABLE I
TABLE SHOWING WHICH TECHNICAL INDICATORS WERE USED, THE FORMULAS AND THE TIMPE PERIODS USED. Y = YESTERDAY, C_1 = CLOSING PRICE. IN A/D THE INDICATOR IS CALCULATED USING ALL OF THE DATA, WHOLE DATASET MEANS THAT ALL OF THE DATA WAS USED.

These Indicators were added to the dataset in order to give more valuable information to our Artificial Neural Network(ANN) model. After adding these features we ended up with a data set of 38 features and 3552 rows, each row representing a data entry from one day.

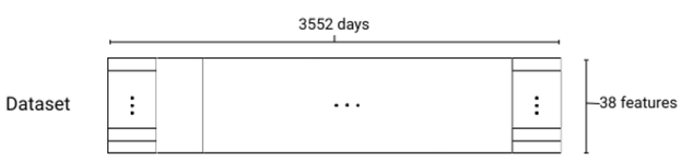


Fig. 1. The dataset. It contains 3552 days of data and each day has 38 features.

C. Data Preprocessing and preparation for the Neural Network Model

We eliminated null rows, which were generated in certain indicators with a time period of, for example, 200 days. This time period implies that data from 200 days prior is required to compute the indicator for the current day. Due to this, the initial rows in the dataset were unable to calculate some of these indicators. To address this issue, we removed the first 199 days of data to avoid null values. Prior to this removal, our dataset contained 3500 rows, but after the removal, it consisted of 3299 rows.

A minimum-maximum scaler was applied to normalize the data between 0 and 1. This scaling technique was employed to enhance the processing capabilities of the Long-Short Term Neural(LSTM) [**?**] Network model. Applying the Min-Max Scaler to the data ensures that the input features are standarized and have a consistent scale. Min Max Scaling preserves the data relationships between the original data points.

$$X_{\text{scaled}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

In this equation:
X represents the original feature value.
X_{min} is the minimum value of the feature in the dataset, representing the lower bound of the scaling range.
X_{max} is the maximum value of the feature in the dataset, representing the upper bound of the scaling range.

D. Artificial Neural Networks in forecasting

Artificial Neural Networks (ANNs) and Long Short-Term Memory Networks (LSTMs) have gained significant traction in the field of financial forecasting, particularly in predicting the volatile and complex nature of cryptocurrency markets such as Bitcoin. This section delves into the utilization of these advanced neural network architectures for the purpose of predicting Bitcoin's closing price.

1) Neural Networks in Financial Forecasting: Neural networks [12], inspired by the structure of the human brain, are adept at learning intricate patterns and relationships within data. In the realm of financial forecasting, these networks have shown promise due to their ability to handle nonlinear data and capture complex dependencies. Neural networks works really well in pattern recognition within a lot of data and make predictions [13] [14].

Feedforward neural networks, characterized by their layered architecture where information flows unidirectionally from input to output, are commonly used for regression tasks like price prediction. However, their limitations in capturing temporal dependencies make them less suited for time series data like historical price movements.

2) Recurrent Neural Networks (RNNs) for Time Series Analysis: RNNs address the temporal aspect by incorporating feedback loops, allowing them to retain information about past inputs. While effective in capturing short-term dependencies, traditional RNNs can suffer from the vanishing gradient problem, hindering their ability to capture long-term dependencies effectively. We use a Gated Recurrent Neural Network.

3) Long Short-Term Memory Networks (LSTMs) for Time Series Forecasting: LSTMs, a variant of RNNs, mitigate the vanishing gradient problem through gated mechanisms, enabling them to capture long-term dependencies more robustly. Their architecture, comprising memory cells and gates, facilitates learning patterns over varying time scales, making them well-suited for financial time series forecasting [15]. This system serves the purpose of instructing nodes on the extent to which they should retain or discard information, as well as determining how much new information they should acquire. This mechanism plays a crucial role in enhancing the model's ability to learn from sequential data and make accurate short-term predictions in the context of cryptocurrency price movements.

4) Transformer Models for Time Series Prediction: Transformers, initially developed for natural language processing, have recently been adapted for time series prediction due to their capability to capture long-range dependencies through self-attention mechanisms. Unlike RNNs and LSTMs, Transformers do not rely on sequential data processing, allowing them to parallelize computations and improve efficiency. However, a fundamental challenge with Transformers is the permutation-invariant self-attention mechanism, which can lead to a loss of temporal information. Despite this, Transformers have shown strong performance in time series forecasting tasks by leveraging positional encoding to retain temporal order [16].

5) PatchMixer Model: The PatchMixer model [6] addresses the challenge of preserving temporal information in Transformers by introducing a permutation-variant convolutional structure. This novel CNN-based model diverges from conventional CNNs, which often employ multiple scales or numerous branches, by relying exclusively on depthwise separable convolutions. This allows PatchMixer to extract both local features and global correlations using a single-scale architecture.

PatchMixer employs dual forecasting heads that encompass both linear and nonlinear components, enabling it to model future curve trends and details more effectively. Experimental results on seven time-series forecasting benchmarks indicate that PatchMixer outperforms state-of-the-art methods and the best-performing CNNs, yielding 3.9% and 21.2% relative improvements, respectively, while being 2-3 times faster than the most advanced methods. The model's innovative approach and superior performance make it a promising tool for predicting Bitcoin price movements.

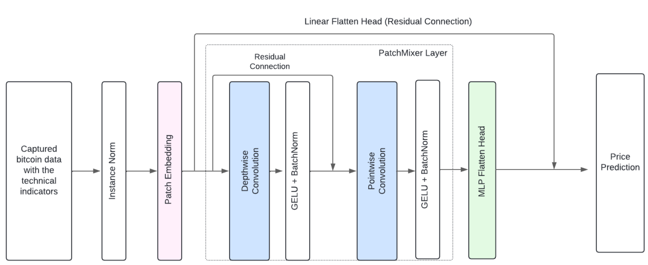


Fig. 2. An illustration of a LSTM architecture.

Approach for Bitcoin Price Prediction

In this study, we leverage both traditional feedforward networks and advanced LSTM architectures to predict Bitcoin's closing price. The integration of technical indicators with historical price data enriches the input features, enhancing the models' predictive capabilities.

To feed the model with data and training it we follow the following process:

- We have a variable called backcandles which is used to indicate the model how many days in the past look in order to make a prediction. We are going to see through data in the past to make a prediction.
- Build a 3 dimensional array which is going to contain in each element 30 one dimensional arrays, and each one dimensional contains the features of one day. This is the method used to give the model some numbers of days to look in the past to make a prediction.
- Split the time series data into 80% for training and 20% for validation
- Create the X_train, X_test, y_train and y_test arrays to be ready for training and validation

IV. Results

To evaluate the models we use three different error metrics and also we predicted two different values, the first value that we tried to predict was the closing price of bitcoin of the next day and then we tried to predict the movement/difference of the close price of bitcoin of the next day. Meaning the difference of the next closing price minus the next opening price. With this we try to predict how much will the price go up or go down.

$$Target = nextClosingPrice - nextOpeningPrice$$

We use 3 different metrics: Metrics:

- Mean Squared Error(MSE). The Mean Squared Error is a commonly used measure of the differences between values predicted by a model and the values actually observed from the environment that is being modeled. For the prediction of cryptocurrency prices, MSE captures the average of the squares of the errors—that is, the average squared difference between the estimated values and the actual value:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

- Mean Absolute Error(MAE). The Mean Absolute Error is a measure of errors between paired observations expressing the same phenomenon. For models predicting financial metrics such as cryptocurrency prices, the MAE provides a straightforward representation of average error:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$$

- Mean Absolute Percentage Error(MAPE). The Mean Absolute Percentage Error expresses prediction accuracy as a percentage, providing a clear depiction of error magnitude in relation to the actual values, which is particularly interpretable in financial and economic contexts:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

A. Using 'NetxClosePrice' as Target

If we use the closing price of the next day as our target prediction, these are the results that we get using the three different neural networks model architectures. The simple neural network, the recurrent neural network and the LSTM neural network.

TABLE II
PERFORMANCE METRICS OF NEURAL NETWORK MODELS, USING 'NEXTCLOSEPRICE' AS OUR PREDICTION

Model	MSE	MAE	MAPE
Shallow NN	0.002202	0.042263	0.118945
GRU	0.000790	0.016724	0.032724
LSTM	0.000528	0.019906	0.053577
Transformer	0.019896	0.097130	0.204902
Patch Mixer	0.03282202	0.14413296	0.30065334

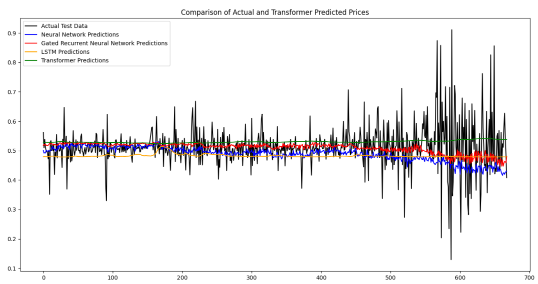


Fig. 3. Comparison of Neural Network Model Performances. The values in the y axis are scaled from 0 to 1. Using target as the price movement of the next day.

We got some really good metrics using the next closing price of bitcoin as our value target to predict. But there may be some bias in the predictions that we are seeing, the prediction might be too good to be true. We believe the models are simply learning to predict a value that is not too far from the lowest bitcoin price and also not too far from the highest value of the given day.

B. Using 'Target'= 'nextClosingPrice - nextOpeningPrice' as Target

We used this value as a target because in this way we can predict how the price will move, if it will go up or go down an also in which quantities it will go up or down. If we use the difference of the

$$nextclosingprice - nextOpeningPrice$$

day as our target prediction, these are the results that we get using the three different neural networks model architectures.

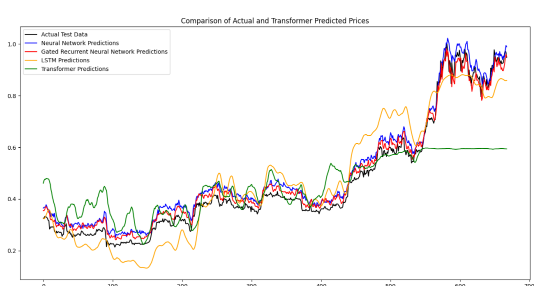


Fig. 4. Comparison of Neural Network Model Performances. The values in the y axis are scaled from 0 to 1. Using target as the closing price of the next day.

The simple neural network, the recurrent neural network and the LSTM neural network.

TABLE III
PERFORMANCE METRICS OF NEURAL NETWORK MODELS, USING 'CLOSE - OPEN' AS OUR PREDICTION

Model	MSE	MAE	MAPE
Shallow NN	0.004791	0.042560	0.086552
GRU	0.004859	0.048007	0.103109
LSTM	0.004416	0.042312	0.090873
Transformer	0.0053824305	0.04602859	0.10171234
Patch Mixer	0.0064404577	0.05082225	0.10021754

Using this target as our prediction, we get a more realistic prediction, it has bigger error metrics and it can be seen in the graphics section on the last pages that the predicted line might not be as close to the true line. Comparing this result with the results obtained when trying to predict the next closing price it is worst but it might have more valuable information for day traders, because knowing in which direction the price will move it is more relevant that simply predicting the next closing price. This target prediction gives a good prediction if the price will go up or down and this is helpful for traders.

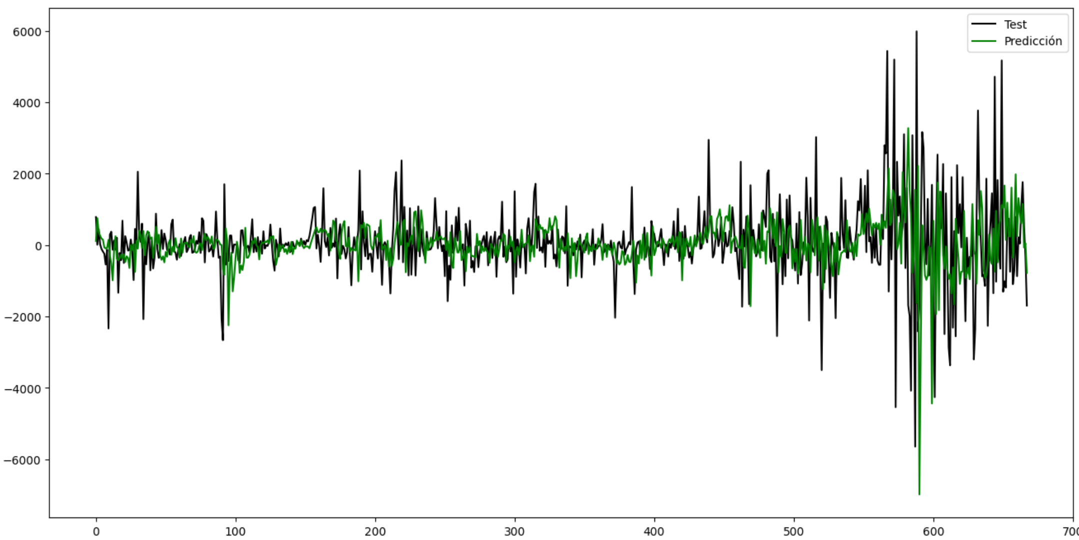
V. CONCLUSIONS

This study demonstrated the potential of neural network models, specifically Shallow Neural Networks, GRUs, and LSTMs, Transformers and the Patch Mixer model in predicting short-term cryptocurrency price movements, with a focus on Bitcoin. Our experiments utilized a combination of technical indicators and historical price data to train our models. The evaluation of these models revealed differing levels of accuracy, as quantified by MSE, MAE, and MAPE metrics, depending on the target variable and model architecture used.

Our results indicate that while using the next day's closing price as a target provided lower error rates, the predictions might be overly optimistic, potentially due to the model's learning patterns that are closely tied to recent price movements rather than predicting true future values. In contrast, using the price movement direction (close-open) as a target resulted in higher error metrics but offered potentially more useful predictions for applications such as day trading, where understanding the direction of price movement is more critical than exact future values.

The study highlights the importance of careful target selection in predictive modeling for cryptocurrencies. It also underscores the challenges of modeling highly volatile financial time series data, where the choice of model, features, and target variables can significantly influence predictive performance and practical utility.

In future work, we aim to explore using more data from the bitcoin blockchain and feature engineering techniques to enhance prediction accuracy and reliability. Additionally, incorporating external data sources such as market sentiment analysis or macroeconomic indicators could provide further insights into the complex dynamics of cryptocurrency.



[c]0.15

Fig. 5. Predictions of the Patch Mixer Model using target as the price movement of the next day.

REFERENCES

[1] L. Tredinnick, "Cryptocurrencies and the blockchain," *Business Information Review*, vol. 36, no. 1, pp. 39–44, 2019.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[3] M. J. S. de Souza, D. G. F. Ramos, M. G. Pena, V. A. Sobreiro, and H. Kimura, "Examination of the profitability of technical analysis based on moving average strategies in brics," *Financial Innovation*, vol. 4, pp. 1–18, 2018.

[4] S. Grossberg, "Recurrent neural networks," *Scholarpedia*, vol. 8, no. 2, p. 1888, 2013.

[5] A. Graves and A. Graves, "Long short-term memory," *Supervised sequence labelling with recurrent neural networks*, pp. 37–45, 2012.

[6] Z. Gong, Y. Tang, and J. Liang, "Patchmixer: A patch-mixing architecture for long-term time series forecasting," *arXiv preprint arXiv:2310.00655*, 2023.

[7] T. O. Hodson, "Root mean square error (rmse) or mean absolute error (mae): When to use them or not," *Geoscientific Model Development Discussions*, vol. 2022, pp. 1–10, 2022.

[8] N. Naik and B. R. Mohan, "Optimal feature selection of technical indicator and stock prediction using machine learning technique," in *Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics: Second International Conference, ICETCE 2019, Jaipur, India, February 1–2, 2019, Revised Selected Papers 2*. Springer, 2019, pp. 261–268.

[9] M. Hiransha, E. A. Gopalakrishnan, V. K. Menon, and K. Soman, "Nse stock market prediction using deep-learning models," *Procedia computer science*, vol. 132, pp. 1351–1362, 2018.

[10] S. Ji, J. Kim, and H. Im, "A comparative study of bitcoin price prediction using deep learning," *Mathematics*, vol. 7, no. 10, p. 898, 2019.

[11] M. Nakano, A. Takahashi, and S. Takahashi, "Generalized exponential moving average (ema) model with particle filtering and anomaly detection," *Expert Systems with Applications*, vol. 73, pp. 187–200, 2017.

[12] A. Krogh, "What are artificial neural networks?" *Nature biotechnology*, vol. 26, no. 2, pp. 195–197, 2008.

[13] E. Guresen, G. Kayakutlu, and T. U. Daim, "Using artificial neural network models in stock market index prediction," *Expert systems with Applications*, vol. 38, no. 8, pp. 10 389–10 397, 2011.

[14] X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, "An innovative neural network approach for stock market prediction," *The Journal of Supercomputing*, vol. 76, pp. 2098–2118, 2020.

[15] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Communications Magazine*, vol. 57, no. 6, pp. 114–119, 2019.

[16] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, "Informer: Beyond efficient transformer for long sequence time-series forecasting," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 12, 2021, pp. 11 106–11 115.

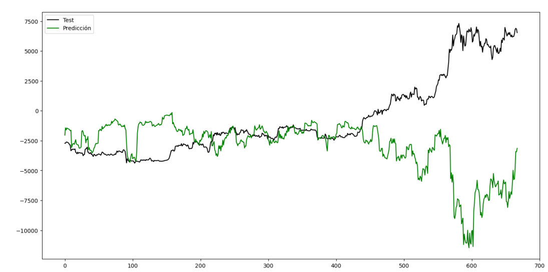


Fig. 6. Prediction of the Patch Mixer Model using target as the price movement of the next day.