Exploring LSTM for Bitcoin Price Prediction: A basic approach

1st Mickael O. de Oliveira
dept. Computer Engineering
IFMG - Campus Bambuí
Bambuí, Brazil
mickaelosvaldo1999@gmail.com

2nd Yanky M. Fonte Boa
dept. Computer Engineering
IFMG - Campus Bambuí
Bambuí, Brazil
yankyboa@gmail.com

***Abstract*—This study focuses on investigating the relationship between Bitcoin price, trading fees, and volume to predict prices at 15-minute candle intervals. Additionally, the research proposes the use of Long Short-Term Memory (LSTM) neural networks to develop a reliable prediction model. Historical Bitcoin data was obtained through the Binance API and underwent collection and preprocessing. Subsequently, the performance of the prediction model is evaluated by comparing it with actual values.**

The findings of this study contribute to the analysis of cryptocurrency markets, enabling investors to make informed real-time decisions. Furthermore, they facilitate the development of advanced trading algorithms and risk management strategies in the cryptocurrency space.

***Index Terms*—Bitcoin, cryptocurrency, Binance, price analysis, volume analysis, neural networks, sentiment analysis, market sentiment.**

I. Introduction

The price volatility of Bitcoin presents both opportunities and challenges for investors and traders. On one hand, it allows for potential high returns and profit opportunities during price surges, on the other hand, it poses risks and challenges in terms of risk management. Furthermore, the volatility of Cryptomarket has implications for wider financial ones and regulatory frameworks. Leading to discussions and debates around market stability and regulatory measures. The graph below illustrates the price fluctuations of Bitcoin from January to March of the 2018.

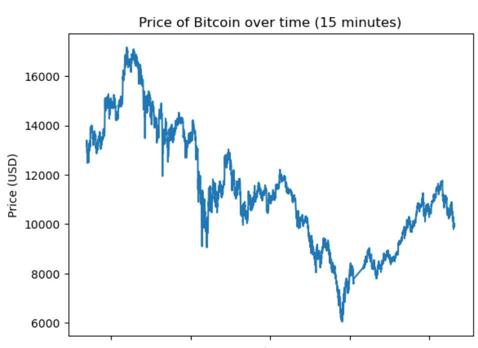

Fig. 1. Chart of Bitcoin price from january to martch of 2018...

Long Short-Term Memory Networks (LSTMs) are a specialized type of recurrent neural networks (RNNs) renowned for their ability to effectively capture long-term dependencies. First introduced by Hochreiter and Schmidhuber in 1997 [1], LSTMs have undergone continuous refinement and gained widespread adoption among researchers due to their exceptional performance across diverse problem domains.

To investigate this premise, our study employed a dataset comprising 5000 parameters, namely timestamp, safe trades, safe volume, risky trades, risky volume, and price. Each row of the dataset corresponds to a 15-minute candle, providing a temporal sequence of observations for analysis.

- **Timestamp**: An auxiliary column used to ensure the integrity of the temporal series and prevent time-related errors.
- **Safe trades**: The number of trades conducted using fiat currency (traditional government-issued currency) or stable coins such as USDT or BUSD.
- **Risky trades**: The number of trades conducted using other cryptocurrencies.
- **Safe volume**: The transaction volume involving stable coins or fiat currency.
- **Risky volume**: The transaction volume involving other cryptocurrencies.

By incorporating these parameters into the model, we aim to analyze the relationship between various trading factors and the price.

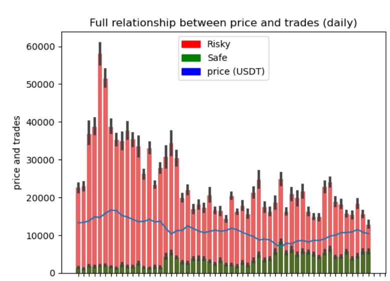

Fig. 2. Relationship between price and trades daily

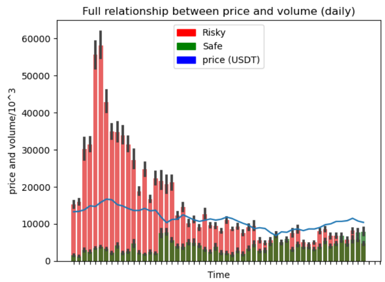

Fig. 3. Relationship between price and volume daily

II. Development

A. Problem Delimitation

Accurate price prediction models can significantly assist in maximizing profitability and managing risks in the cryptocurrency market. However, predicting Bitcoin prices accurately is a complex task due to various factors such as market sentiment, trading rates, and volume. Traditional time-series analysis techniques may struggle to capture the intricate relationships and patterns inherent in volatile graphs. Therefore, advanced machine learning methods, such as the LSTM, are employed to make more acceptable predictions. It's important to note that this study does not claim to provide an absolute values or warranty investments, the dynamic nature of the cryptocurrency market and the presence of unforeseen events make it impossible to accurately predict future price movements with complete certainty. Instead, this research aims to provide a directional indication and aid decision-making processes by identifying trends and patterns within the data.

B. Related Previous Studies

The origin of cryptocurrency, also known as cryptographic currency, can be traced back to the 1980s with the contributions of David Chaum. In his seminal paper, Chaum introduced an innovative cryptographic scheme that involved blinding the content of a message before it is signed, thereby preventing the signer from determining its actual content. These blind signatures could be publicly verified, similar to regular digital signatures. Chaum's proposal aimed to create a digital cash system that was untraceable by third parties, ensuring privacy and anonymity [2]. The development of cryptocurrency gained momentum with the introduction of B-money in 1998. Wei Dai proposed B-money as an anonymous and distributed electronic cash system [3]. This approach entailed the use of two protocols based on a network that couldn't be traced. In this system, both senders and receivers were identified solely by their digital public keys, and each message was signed by the sender to the receiver, ensuring the integrity of transactions and the security of the network.

In Ferdiansyah's paper, a machine learning approach is employed, where historical Bitcoin data is fed into an LSTM for training[3]. Created by Hochreiter and Schmidhuber in 1980 the model is designed to learn patterns in the input data and generate future predictions based on this information[1]. The data used includes features such as opening price, closing price, trading volume, and other relevant indicators.

After training the model, the researchers evaluate its effectiveness in predicting future Bitcoin prices. Performance metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE) are analyzed to assess the accuracy of the predictions compared to actual values. The results of the study demonstrate that the proposed method, based on LSTMs, is capable of providing reasonably accurate predictions of future Bitcoin prices.

Regarding the works by Nametala in 2020[5] and 2022[6], we can observe that time series forecasting can be applied not only to bitcoins but also to other domains. Nametala's work focuses on predicting volatility in bilateral electricity contracting auctions, which involve negotiations and agreements of electricity purchase and sale contracts between buyers and sellers. The volatility index is designed to capture the level of uncertainty and risk associated with electricity prices in these auctions. However, the neural network implemented in this system is quite complex, involving multiple variables and factors that can influence the prediction. In contrast, our method is simpler, utilizing only the opening and closing prices. Therefore, we can leverage both works and implement a neural network capable of predicting Bitcoin prices.

III. Theoretical Framework

The main advantage of LSTMs is their ability to handle long-term dependencies. Unlike conventional RNNs, which struggle to remember information for extended periods, LSTMs were explicitly designed to overcome this problem. In fact, retaining information for long periods is an inherent behavior of LSTMs, making them highly effective.

All recurrent neural networks have a chain-like structure, composed of repeated modules of neural networks. In the case of traditional RNNs, these repeating modules have a simple structure, such as a single tanh layer.

The following figure illustrates this simple structure:

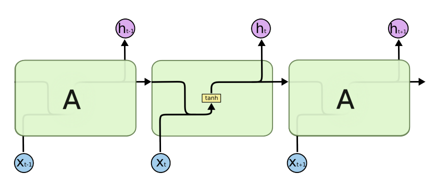

Fig. 4. Example of a simple RNN.

However, LSTMs have a more complex structure that enables them to capture and store relevant information over extended periods, they include memory units called "cells" that are capable of retaining information and deciding which information to keep or forget. These cells are combined with gates that regulate the flow of information within the network.

The following figure illustrates the structure of an LSTM:

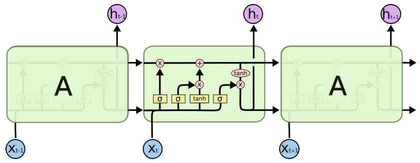

Fig. 5. Example of the LSTM.

IV. Methodology

In the initial step, the data underwent normalization using the MinMaxScaler function from the Scikit-Learn library. This normalization technique facilitates scaling the data within the range of zero to one, promoting consistency across the features. Consequently, despite potential variations in the scales of individual columns, the network benefits from enhanced interpretability and inferential capabilities based on the normalized values. To adapt this problem for LSTM, the dataset is divided into time window frames, where each frame represents a one-hour interval.

The frame dataset was divided into two subsets: training and validation, with each subset comprising 70% and 30% of the data, respectively. It is of utmost importance to preserve the sequential order and avoid shuffling the data, maintaining the original order of the windows is crucial as it ensures the temporal integrity of the data.

The LSTM model is constructed with three layers and is trained to predict the Bitcoin price at 1:15 by utilizing the information from the previous four candlesticks and forecasting the price beyond the known time frame. The input consists of a matrix with 4 rows and 5 columns, representing the candlestick data, while the output is a scalar value representing the predicted price.

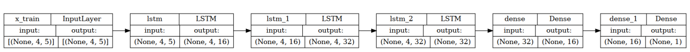

Fig. 6. LSTM Layers Structure.

The network was trained using the training set and Nadam optimizer with a learning rate of 0.0001, implemented through the Keras library with TensorFlow backend. The Nadam optimizer combines the benefits of Nesterov accelerated gradient descent (NAG) and Adam optimization algorithms. It incorporates NAG's momentum-based approach to accelerate convergence and Adam's adaptive learning rate to handle different gradient scales effectively. During the configuration, a portion of the data was designated for validation. Typically, this division is done by setting a fixed proportion of the data for validation, is this case the value is 20% The model was trained over 1000 epochs, with each epoch representing a complete pass through the training data. The metrics used to make decisions are mean square error.

V. Results and Discussion

In our results, we can observe that the LSTM neural network was able to predict the price of Bitcoin with an accuracy of 97.45% using mean squared error (MSE). The MSE is used to measure the quality of predictions made by an LSTM neural network in regression problems, then, the LSTM is trained to minimize the MSE during the training process by adjusting its weights and biases to better approximate the real values of the training data. With the MSE values, we can derive the coefficient of determination, called R^2 ("R-Squared"), which is a statistical measure indicating the proportion of variability in the output data explained by the model.

R^2 is calculated as 1 - MSE / Variance of the output data. Therefore, the higher the R^2, the better the model fits the data. In our case, we obtained an R^2 value of 0.9745546312041986, which is a very good value for the problem at hand.

We also calculated the MAPE (Mean Absolute Percentage Error), which is a metric expressed in percentage and makes it easier to interpret how far the model's predictions are from the real values. The lower the MAPE value, the more accurate the model's predictions, a MAPE of 0% would mean that the model made perfect predictions, while a MAPE of 100% would mean that the model, on average, erred by 100% compared to the real values. In our case, we obtained a MAPE of 1.6321638077870935, which is excellent.

Therefore, with the convergence graph, we were able to prove our theory that the LSTM neural network can predict the price of Bitcoin with an accuracy of 97.45%:

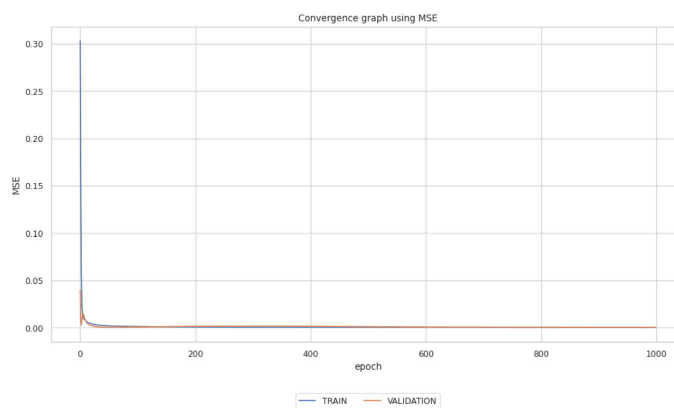

Fig. 7. Convergence Graph using MSE.

Next, we have the graph of the real Bitcoin price vs. the predicted Bitcoin price by the LSTM neural network. The prediction was done correctly, and in certain parts of the graph, it is evident that the network predicted the exact same result. However, in some parts, it did not execute the prediction correctly, by comparing the real points with the predictions, it is possible to visualize how close or far the predictions are from the real values. If the predictions are very close to the real values, the predicted lines or points will overlap or be close to the real points on the graph. On the other hand, if there is a significant difference between the predictions and the real values, there will be a visible separation between the real and predicted points on the graph.

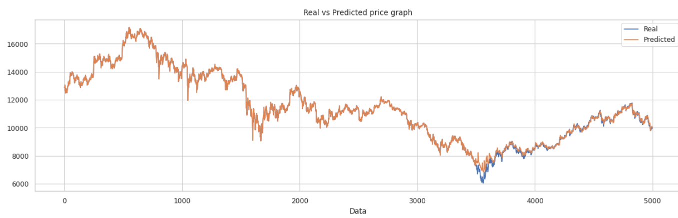

Fig. 8. Real vs Predicted Price Graph.

VI. Conclusion

Unfortunately, the current model cannot be used in this way due to the risks associated with the generated values. Although it is relatively close to reality, fully relying on these results can be dangerous. However, there are additional techniques that can be applied to improve its effectiveness, such as using Transformer-based models. The Transformer is a neural network architecture that has excelled in natural language processing tasks, such as automatic translation and text generation. Its ability to capture complex relationships in sequence data can be applied to the problem of financial prediction.

Other data can be incorporated into the model, such as information from the futures market. This data can provide additional insights into future market trends, enabling a more comprehensive analysis and more accurate predictions. A complementary approach would be to integrate neural networks for sentiment analysis, which can capture information about market sentiment and opinion regarding specific financial assets. By combining the predictions generated by the model with these sentiment analyses, a more complete view of the market can be obtained, leading to more informed investment decisions.

VII. References

[1] S. Hochreiter and J. Schmidhuber, 'Long Short-Term Memory', Neural Comput., vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[2] D. Chaum, 'Blind Signatures for Untraceable Payments', in Advances in Cryptology, 1983, pp. 199–203

[3] W. Dai, "B-money proposal," White Pap., 1998.

[4] F. Ferdiansyah, et al. "A lstm-method for bitcoin price prediction: A case study yahoo finance stock market." 2019 International Conference on Electrical Engineering and Computer Science (ICECOS). IEEE, 2019.

[5] C. Nametala, W. de Faria and B. Junior, "A Volatility Index for Bilateral Electricity Contracting Auctions," in IEEE Latin America Transactions, vol. 18, no. 05, pp. 938-946, May 2020, doi: 10.1109/TLA.2020.9082923.

[6] C. Nametala, W. Faria, G. Lage, and B. Júnior, 'Analysis of Hourly Price Granularity Implementation in the Brazilian Deregulated Electricity Contracting Environment', SSRN Electronic Journal, 01 2022.