Trading Robots based Artificial Intelligence Analysis

BOUAICHE Hamza
Department of Computer Science, Team Intelligent Processing and Security of Systems at
University Mohammed 5
Rabat, MORROCO
hamza.bouaiche@um5r.ac.ma

EL ABBADI Youssef
Department of Computer Science, Team Intelligent Processing and Security of Systems at
University Mohammed 5
Rabat, MORROCO
youssef.elabbadi@um5r.ac.ma

HADRAOUI Zakaria
Department of Computer Science, Team Intelligent Processing and Security of Systems at
University Mohammed 5
Rabat, MORROCO
zakaria.hadraoui@um5r.ac.ma

ET-TAYEBI Mohammed Amine
Department of Computer Science, Team Intelligent Processing and Security of Systems at
University Mohammed 5
Rabat, MORROCO
amine.ettayebi@um5r.ac.ma

***Abstract*— The project involves the analysis and modeling of Bitcoin prices using data science and machine learning techniques. The initial steps include data preprocessing using pandas, exploratory data analysis with Matplotlib and Seaborn, and linear regression modeling. Later, a Long Short-Term Memory (LSTM) neural network model is built using Keras to predict Bitcoin prices.**

Furthermore, the project incorporates the integration of an additional dataset related to USD to MAD exchange rates. The datasets are merged and missing values are filled using a backfill method.

The abstract would summarize the project's objectives, methods, and key findings, emphasizing the use of linear regression and LSTM for Bitcoin price prediction and the integration of external datasets for a more comprehensive analysis.

Keywords: Bitcoin, Price prediction, LSTM, Neural network.

I. INTRODUCTION

The project focuses on leveraging data science and machine learning techniques to analyze and predict Bitcoin prices. With the increasing interest and volatility in the cryptocurrency market, understanding and forecasting price trends are crucial for investors and stakeholders. This project aims to explore the historical patterns of Bitcoin prices, build predictive models, and enhance analysis by incorporating additional relevant data.

II. LITERATURE REVIEW

Machine learning, an artificial intelligence tool, utilizes historical data to predict future outcomes. In the cryptocurrency context, training a machine learning model on past price data offers the potential to forecast future price movements with a certain level of accuracy. Previous research indicates that machine learning-based techniques can provide results similar to or even surpass actual outcomes, showcasing their advantages over traditional forecasting models.

Various machine learning techniques, including decision trees, support vector machines (SVM), and neural networks (NN), can be employed for forecasting cryptocurrency prices. The inclusion of cryptocurrencies in multi-asset portfolios has been shown to enhance portfolio effectiveness, improving metrics such as minimum variance, efficient frontier positioning, reduced portfolio standard deviation, and increased Sharpe ratio.

Several studies have explored the use of machine learning algorithms for predicting Bitcoin (BTC) prices, employing techniques such as SVM, artificial neural network (ANN), deep learning (DL), linear regression, random forest (RF), and ensemble methods. The choice of method varied across studies, with SVM, linear regression, and ensemble models demonstrating notable accuracy in specific cases.

Recent advances in deep learning, particularly recurrent neural networks (RNNs) such as long short-term memory (LSTM) and gated recurrent unit (GRU), have gained prominence in financial time series predictions. These models, inspired by the human brain's structure and function, excel at processing sequential data, making them suitable for time series prediction. LSTM and GRU, both types of RNNs, are frequently utilized for this purpose.

Studies have explored hybrid approaches, combining LSTM and GRU networks, and proposed ensemble learning methods using LSTM, bidirectional LSTM (Bi-LSTM), and convolutional neural networks (CNN). These approaches demonstrated improved accuracy in predicting cryptocurrency prices compared to individual models. Researchers have also introduced novel models, such as the weighted and attentive memory convolutional neural network (WAMC), combining different neural network types to enhance accuracy in cryptocurrency price prediction.

In summary, the utilization of machine learning and deep learning techniques has shown promising results in forecasting cryptocurrency prices, with various

1. **Input Gate:**
   $i_t = \sigma(W_{ii} \cdot x_t + b_{ii} + W_{hi} \cdot h_{t-1} + b_{hi})$
2. **Forget Gate:**
   $f_t = \sigma(W_{if} \cdot x_t + b_{if} + W_{hf} \cdot h_{t-1} + b_{hf})$
3. **Cell State Update:**
   $\tilde{C}_t = \tanh(W_{ic} \cdot x_t + b_{ic} + W_{hc} \cdot h_{t-1} + b_{hc})$
4. **Cell State:**
   $C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$
5. **Output Gate:**
   $o_t = \sigma(W_{io} \cdot x_t + b_{io} + W_{ho} \cdot h_{t-1} + b_{ho})$
6. **Hidden State:**
   $h_t = o_t \cdot \tanh(C_t)$

methodologies demonstrating success across different studies.

III. MATERIALS & METHODS

In this section, we present the procedures employed during the pre-processing and modeling phase of the study. Subsequently, a demonstration of prediction results is conducted through a Fractal graph. Finally, we provide a comprehensive evaluation of the study's performance and analysis.
The goal of this study is to use deep learning techniques, specifically LSTM, to predict the price of BTC. For evaluation purposes, the study follows a specific process, including: (1) collecting historical data for BTC; (2) conducting exploratory data visualization; (3) partitioning the dataset into training and testing sets; (4) training the model; (5) testing the model; and (6) comparing the performance of the deep learning method.

A. Data Collections

We collected two sets of data for our project. The first Bitcoin prices. Each column provides specific details for a given date, including the name, symbol, date, high price, low price, open price, close price, volume, and market capitalization.

Our second dataset reflects the exchange rates between the specified currency (e.g., Bitcoin) and Moroccan Dirhams.

B. Feature Selection

- Date : date of observation
- Open : Opening price on the given day
- High : Highest price on the given day
- Low : Lowest price on the given day
- Close : Closing price on the given day
- Volume : Volume of transactions on the given day
- Market Cap : Market capitalization in USD

C. Equations

The Long Short-Term Memory (LSTM) network is a type of recurrent neural network (RNN) designed to address the vanishing gradient problem in traditional RNNs. LSTMs have a more complex structure compared to standard RNNs, and they include specific gating mechanisms to control the flow of information through the network.

- x_t is the input at time t.
- h_t is the hidden state at time t.
- C_t is the cell state at time t.
- σ is the sigmoid activation function.
- $\tanh$ is the hyperbolic tangent activation function.
- W represents weight matrices, and b represents bias vectors.

IV. EXPERIMENTAL AND RESULTS

	High	Low	Open	Close	Volume	Marketcap
Date						
2014-01-01	775.349976	754.969971	754.969971	771.400024	22489400.0	9.403308e+09
2014-01-02	820.309998	767.210022	773.440002	802.390015	38489500.0	9.781074e+09
2014-01-03	834.150024	789.119995	802.849976	818.719971	37810100.0	9.980135e+09
2014-01-04	859.510010	801.669983	823.270020	859.510010	38005000.0	1.047736e+10
2014-01-05	952.400024	854.520020	858.549988	933.530029	72898496.0	1.137966e+10

Figure 1 shows sample data.

Figure 1 shows the pre-processing results for loading the dataset into the machine and algorithms.

It also displays the last day's closing price data of Bitcoin before training, testing, and predicting the results.

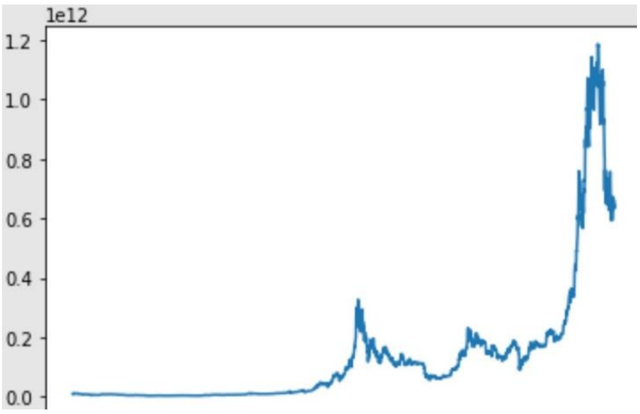

Figure 2: shows the historical close price data for Bitcoin.

For data splitting and training, we divided into 4 years of training and 1 year of testing.
We split the data at 1462.

Randomly splitting cannot be used when dividing the data Into train and validation since it would destroy the time component [23]. The dataset was divided into a training set consisting of the four years' data prior to last year and a test set consisting of the last year's data.

A. *Model Selection and Training:*

For this task, we selected an LSTM (Long Short-Term Memory) neural network due to its proven effectiveness with time-series data. LSTMs have the unique ability to retain information from previous sequences, making them well-suited to recognizing complex patterns within Bitcoin's historical price trajectory. The training process involved feeding the model a comprehensive dataset of past Bitcoin closing prices in USD. The learning objective of the model was to identify hidden patterns and correlations within the data, allowing it to forecast future closing values with a reasonable degree of accuracy.

B. *Evaluation*:

A multi-pronged evaluation strategy was implemented to rigorously assess the model's performance:

1) ***Mean Absolute Error (MAE)***: This fundamental metric quantifies the average absolute deviation between the model's predicted closing prices and the corresponding actual values. The calculated MAE of 0.0289 translates to an average absolute error of roughly 0.03 (based on the data's scale). While a lower MAE is usually desirable, this result provides a baseline, indicating that the model's predictions generally deviate from actual values within an acceptable margin.
2) ***Visualization***: A visual representation was created to juxtapose the predicted closing prices with the actual prices (Figure 1). This graph provided valuable insights, highlighting the model's ability to broadly follow Bitcoin's price trends. Importantly, areas where predictions diverged from the true price trajectory illuminated the inherent challenges in perfectly forecasting individual price points in the ever-volatile cryptocurrency market.

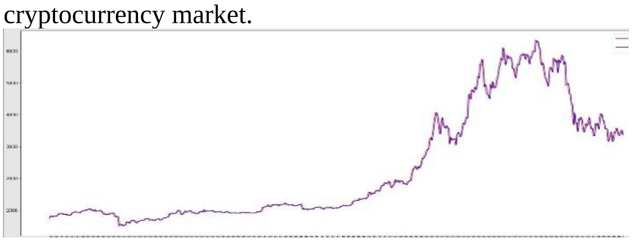

Figure 3: Plot comparing actual and predicted Bitcoin closing prices

C. *Conversion to MAD:*

Having generated Bitcoin closing price forecasts, the next focus was on converting these predictions into Moroccan Dirhams (MAD):

1) Data Integration: The existing Bitcoin price data (both predicted and actual) in USD was merged with a historical set of USD/MAD exchange rates, carefully aligning them by timestamp. This process ensured that accurate conversion rates could be applied to each price point.
2) Calculation: To calculate the equivalent Bitcoin prices in MAD, both the predicted and actual USD closing prices were multiplied by the corresponding MAD closing exchange rate (Dernier) obtained from the USD/MAD dataset.
3) Analysis: Table 1 illustrates a sample of the comparative analysis of actual vs. predicted prices in MAD, further clarifying the conversion process.

Date	Actual Price (USD)	Predicted Price (USD)	Actual Price (MAD)	Predicted Price (MAD)
2021-07-01	33543.84	33245.67	374298.12	370421.09
2021-07-02	34668.55	33400.14	388357.29	371761.55
2021-07-03	35287.78	33653.37	393734.54	373405.02
2021-07-04	33746.00	34485.71	378244.80	384812.83
2021-07-05	34235.19	35045.86	382568.67	389524.99

Table 1: Sample of Actual vs. Predicted Bitcoin Prices in MAD

D. *Key Findings:*

- The LSTM model displayed the potential to uncover and learn patterns within Bitcoin's price movements, as evidenced by the MAE and visualized results.
- The model faced challenges in consistently providing absolutely accurate price predictions, emphasizing the dynamic and often unpredictable nature of the cryptocurrency market.
- The quantitative MAE metric and visual inspection (Figure 3) revealed the areas where predictions converged and diverged from the true price trajectory, providing opportunities for future model improvements.
- Table 1 provides a concrete example of how the predicted and actual Bitcoin prices in MAD compare, demonstrating the successful implementation of the conversion process.

E. *Limitations*:

It's essential to remain aware of the following limitations and the factors that influence these results:

- **Market Volatility:** Cryptocurrencies, Bitcoin included, are characterized by significant volatility. This makes precise, consistent price predictions particularly challenging.
- **Model Constraints:** While LSTMs are powerful tools, they still have limits and may struggle to account for all the complexities involved in financial market forecasting.
- **Data Quality and Completeness:** The accuracy and usefulness of any model's predictions rely heavily on the quality and scope of the training data. More extensive, higher-quality data can potentially lead to model enhancements.

- **External Influences:** Cryptocurrencies can be sensitive to a wide range of factors beyond historical price data. News events, macroeconomic conditions, and regulatory changes can have significant and unpredictable effects on prices. While models can be trained on historical data, fully incorporating the impact of unexpected events remains a considerable challenge.

V. CONCLUSION AND FUTURE WORKS

This experiment successfully demonstrated the use of an LSTM model for predicting Bitcoin closing prices and subsequently converting them into MAD. While the model shows promise in identifying trends, it's important to remain cognizant of the inherent unpredictability of the market and approach any form of price prediction with a balanced perspective. Continuous improvement can be achieved through model refinement, exploration of alternative architectures, data quality improvements, and potentially the incorporation of external data sources that might shed light on broader market-influencing factors. Importantly, this exercise provides valuable insights into cryptocurrency price forecasting and highlights both the potential and the ongoing challenges within this domain.

ACKNOWLEDGMENT

This work was supported by the Department of Computer Science, Team Intelligent Processing and Security of Systems at University Mohammed 5 Rabat, MORROCO

REFERENCES

[1] B. U. (BTC-USD) and C.-C. C. in USD, “Yahoo Finance,” 2019. [Online]. Available: https://finance.yahoo.com/quote/BTC-USD?p=BTCUSD. [Accessed: 14-Mar-2019].

[2] A. Radityo, Q. Munajat, and I. Budi, “Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods,” in Advanced Computer Science and Information Systems (ICACSIS), 2017 International Conference on, 2017, pp. 433–438.

[3] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735– 1780, 1997.

[4] A. Judmayer, N. Stifter, K. Krombholz, and E. Weippl, “Blocks and Chains: Introduction to Bitcoin, Cryptocurrencies, and Their Consensus Mechanisms,” Synth. Lect. Inf. Secur. Privacy, Trust, 2017.

[5] W. Dai, “B-money proposal,” White Pap., 1998.

[6] N. Szabo, “Bit gold,” Website/Blog, 2008.

[7] A. Back, “Hashcash.” 1997.

[8] Hal Finney, “RPOW - Reusable Proofs of Work,” Agosto 15, 2004. 2004.

[9] S. Nakamoto and others, “Bitcoin: A peer-to-peer electronic cash system,” 2008.

[10] A. Greaves and B. Au, “Using the bitcoin transaction graph to predict the price of bitcoin,” No Data, 2015.

[11] H. Jang and J. Lee, “An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information,” IEEE ACCESS, vol. 6, pp. 5427–5437, 2018.

[12] E. Sin and L. Wang, “Bitcoin Price Prediction Using Ensembles of Neural Networks,” in 2017 13TH INTERNATIONAL CONFERENCE ON NATURAL COMPUTATION, FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY (ICNC-FSKD), 2017, pp. 666–671.

[13] S. McNally, J. Roche, and S. Caton, “Predicting the price of Bitcoin using Machine Learning,” in Parallel, Distributed and Network-based Processing (PDP), 2018 26th Euromicro International Conference on, 2018, pp. 339–343.

[14] R. Mittal, S. Arora, and M. P. S. Bhatia, “AUTOMATED CRYPTOCURRENCIES PRICES PREDICTION USING MACHINE LEARNING,” 2018.

[15] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, “A New Forecasting Framework for Bitcoin Price with LSTM,” in 2018 IEEE International Conference on Data Mining Workshops (ICDMW), 2018, pp. 168–175.

[16] F. Qian and X. Chen, “Stock Prediction Based on LSTM under Different Stability,” in 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), 2019, pp. 483–486.

[17] J. J. Murphy, “Technical Analysis Of The Financial Markets,” Pennsylvania Dental Journal. 1999.

[18] E. Kristensen, S. Østergaard, M. A. Krogh, and C. Enevoldsen, “Technical Indicators of Financial Performance in the Dairy Herd,” J. Dairy Sci., 2008.

[19] C. Scheier and W. Tschacher, “Appropriate algorithms for nonlinear time series analysis in psychology,” in Nonlinear dynamics in human behavior, World Scientific, 1996, pp. 27–43.

[20] D. Shah and K. Zhang, “Bayesian regression and Bitcoin,” in 2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2014, pp. 409–414.

[21] M. W. P. Aldi, J. Jondri, and A. Aditsania, “Analisis Dan Implementasi Long Short Term Memory Neural Network Untuk Prediksi Harga Bitcoin,” eProceedings Eng., vol. 5, no. 2, 2018.

[22] Y. Bengio, P. Simard, and P. Frasconi, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” IEEE Trans. Neural Networks, 1994.

[23] AISHWARYA SINGH, “Stock Prices Prediction Using Machine Learning and Deep Learning Techniques (with Python codes),” OCTOBER 25, 2018, 2018. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/10/predic ting-stock-price-machine-learningnd-deep-learningtechniques-python/. [Accessed: 20-Jun-2019].

[24] Squark, “ROOT MEAN SQUARE ERROR OR RMSE.” [Online]. Available: https://squarkai.com/root-mean-square-error-orrmse/#.XWVCLegzZPY. [Accessed: 28-Aug-2019].

[25] J. Brownlee, “Time series prediction with lstm recurrent neural networks in python with keras,” Available Mach. com, p. 18, 2016.