International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Forecasting and Trading of the Stable Cryptocurrencies with Machine Learning for Market Conditions

Ms. P.Sindhuri [a], M.S.Kishore [b], R.Durga [c], R.Almas [d], R.Logeshwari [e]

[a] *Assistant Professor, Department Of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, India.*
[b,c,d,e] *Student, Department Of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, India.*

ABSTRACT:

Virtual currencies have been declared as one of the financial assets that are widely recognized as exchange currencies. The cryptocurrency trades caught the attention of investors as cryptocurrencies can be considered as highly profitable investments. To predict Cryptocurrency price at different frequencies using machine learning techniques, we first classify Cryptocurrency price by daily price and high-frequency price. A set of high-dimension features including property and network, trading and market, attention and gold spot price are used for Cryptocurrency daily price prediction, while the basic trading features acquired from a cryptocurrency exchange are used for 5-minute interval price prediction In today's world we can see the trend of cryptocurrency is constantly increasing every day. In the financial sector, cryptocurrency has become a huge topic and the right prediction has become more important to gain profits. For determining the right prediction with good accuracy, we performed deep analysis on dataset to understand the market behaviour by using LSTM machine learning algorithms like to predict the daily price behaviour of top 4 cryptocurrencies like Bitcoin, XRP, Ethereum, and Stellar using these machine learning algorithms. Our experimental result reaches to 95–97 percent validation accuracy

Keywords: Machine Learning, Price Prediction, Deep Analysis Dataset, LGBM Algorithm.

1. INTRODUCTION

Bitcoin is a digital currency created in 2009 by a pseudonymous developer named Satoshi Nakamoto. It works on peer-to-peer technology and does not rely on any central authority or banks. Bitcoin is open-source, and anyone can participate in its network. Transactions are verified by users called miners and stored in a public ledger called the blockchain. Bitcoin uses cryptography to ensure secure transactions and control the creation of new coins. Its value is highly volatile and differs from traditional stock markets, as it is not influenced by business events or government actions. Predicting Bitcoin prices is challenging due to these unique factors. To address this, machine learning techniques are used for price forecasting. Machine learning is a part of Artificial Intelligence that allows systems to learn from data automatically without human intervention. It improves decision-making by identifying patterns and trends in the data. There are two main types of machine learning: supervised and unsupervised. Supervised learning uses labeled data to predict outcomes, while unsupervised learning finds hidden patterns in unlabeled data. These techniques help in understanding and forecasting Bitcoin's price more effectively.

2. LITERATURE SURVEY

1.FORECASTING AND TRADING OF THE STABLE CRYPTOCURRENCIES WITH MACHINE LEARNING AND DEEP LEARNING ALGORITHMS FOR MARKET CONDITIONS

AUTHOR NAME: Hasib Shamshad 1 , Fasee Ullah 2 , Asad Ullah1 , Victor R. Kebande 3 , Sibghat Ullah1 , And Arafat Al-Dhaqm2

The digital market trend is rapidly expanding due to key characteristics like decentralization, accessibility, and market diversity enabled by blockchain technology. This study proposes a Predictive Analytics System to provide simplified reporting for the three most popular cryptocurrencies with varying digits, namely ADA Cardano, Ethereum, and Binance coin, for ten days to contribute to this emerging technology. Thus, this proposed system employs a data science-based framework and six highly advanced data-driven Machine learning and Deep learning algorithms: Support Vector Regressor, Auto-Regressive Integrated Moving Average (ARIMA), Facebook Prophet, Unidirectional LSTM, Bidirectional LSTM, Stacked LSTM. Moreover, the research experiments are repeated several times to achieve the best results by employing hyperparameter tuning of each algorithm. This involves selecting an appropriate kernel and suitable data normalization technique for SVR, determining ARIMA's

2.NOVEL DYNAMIC MODEL FOR RANKING CRYPTOCURRENCIES IN DIFFERENT TIME HORIZONS BASED ON DEEP LEARNING AND SENTIMENT ANALYSIS

AUTHOR NAME: Aida Mohagheghzadeh, Babak Amiri , And Ahmad Makui

This paper addresses the imperative task of assessing and ranking cryptocurrencies, particularly pertinent in the context of the burgeoning popularity of public blockchains. The proliferation of available options necessitates a rigorous evaluation, prompting the formulation of a novel model grounded in both objective and subjective criteria. To contend with the challenge posed by the expanding landscape of public blockchains, ten discerning criteria are delineated, encompassing facets such as Technology, TPS, Market capitalization, GitHub fork, GitHub stars, Twitter followers, Twitter hashtags, trading volume, sentiment score, and the price range differential. Leveraging expert opinions, the pairwise impact of these criteria is ascertained, and the DEMATEL method is judiciously employed to derive their respective weights. Subsequently, the PROMETHEE method is harnessed to effectuate the ranking of 20 cryptocurrencies predicated on the identified criteria. Furthermore, the integration of LSTM enables the prediction of values for four predictable criteria, seamlessly incorporated into the PROMETHEE model to furnish rankings across diverse temporal intervals. The proposed model, thus, presents a holistic and pragmatic approach to inform investment decision-making within the dynamic cryptocurrency market. By embracing a comprehensive set of criteria and integrating predictive analytics, this model stands as a valuable contribution to the field, offering nuanced insights to stakeholders navigating the complexities of cryptocurrency investment.

3.SYSTEM STUDY

EXISTING SYSTEM

After the boom and bust of crypto currencies' prices in recent years, CRYPTOCURRENCY has been increasingly regarded as an investment asset. Because of its highly volatile nature, there is a need for good predictions on which to base investment decisions.Although existing studies have leveraged machine learning for more accurate CRYPTOCURRENCY price prediction, few have focused on the feasibility of applying different modeling techniques to samples with different data structures and dimensional features.A set of high-dimension features including property and network, trading and market, attention and gold spot price are used for CRYPTOCURRENCY daily price prediction, while the basic trading features acquired from a cryptocurrency exchange are used for 5-minute interval price prediction.

PROPOSED SYSTEM

The proposed system is a predictive analytics application designed to forecast the price movements of stable cryptocurrencies using advanced machine learning algorithms, specifically Light Gradient Boosting Machine (LGBM). The system integrates data from multiple sources such as historical cryptocurrency prices, Twitter feeds, and financial news headlines. The collected data undergoes preprocessing steps including normalization, sentiment analysis, and labeling to ensure quality and relevance. Using time-series features, sentiment scores, and technical indicators, the LGBM model is trained to classify future market movements into categories like increase, decrease, or stable. Based on these predictions, the system generates buy, sell, or hold recommendations through a trading strategy engine. The final results, including accuracy metrics and visual trend analyses, are displayed on a user-friendly interface. This proposed system aims to empower investors with timely and intelligent decision-making tools in the volatile cryptocurrency market by leveraging the speed and accuracy of machine learning.

4.METHODOLOGY

4.1 SUPERVISED MACHINE LEARNING

Involves labeled training data.

Used for regression (price value prediction) and classification (price up/down/neutral).

Examples: Logistic Regression, XGBoost.

4.2 DEEP LEARNING TECHNIQUES

Neural networks for time-series prediction.

Examples: LSTM and GRU models are used to capture long-term dependencies in price movements.

4.3 STATISTICAL TIME SERIES ANALYSIS

Models the historical data based on its own past values.

Example: ARIMA is used for univariate time series forecasting.

4.4 SENTIMENT ANALYSIS

Natural Language Processing (NLP) technique to determine sentiment polarity.

Enhances predictions by incorporating social media influence.

5. MODULES IMPLEMENTATION

5.1 LIST OF MODULES

- Data Collection
- Preprocessing
- Model Traning
- Predication
- Visualization and Output

5.2 MODULE DECRIPTION

5.2.1 DATA COLLECTION

The data collection process forms the backbone of the cryptocurrency price prediction system by providing essential information for training and testing machine learning models. In this project, two primary types of data are collected: historical cryptocurrency price data and social media sentiment data. The price data, including attributes such as open, close, high, low, volume, and weighted average prices, is gathered from trusted sources like CoinDesk through publicly available APIs and stored in CSV format for ease of analysis. Alongside this, Twitter data is collected using the Twitter REST API via the Tweepy Python library. Tweets containing keywords such as "Bitcoin" and "cryptocurrency" are extracted in real time, with each tweet capturing details like timestamp, tweet text, and user metadata. These tweets are then analyzed for sentiment using the TextBlob library to classify them as positive, negative, or neutral. The combined dataset—market prices and sentiment scores—provides a rich source of input features that enable the prediction model to consider both quantitative financial trends and qualitative public opinion. This comprehensive data collection approach significantly enhances the robustness and reliability of the prediction system.

5.2.2 PREPROCESSING

Data preprocessing is a crucial step in preparing raw data for machine learning algorithms, ensuring that the input is clean, consistent, and meaningful. In this project, the collected cryptocurrency price data and tweet sentiment data undergo several preprocessing operations. Initially, missing values and irrelevant entries are removed to ensure data integrity. The price data is then normalized using the Min-Max scaling technique, which transforms the values into a range between 0 and 1, allowing the model to learn efficiently without being affected by scale differences. For textual data, tweets are cleaned by removing special characters, links, and stop words. Sentiment scores are computed using the TextBlob library, assigning polarity values to classify tweets as positive, negative, or neutral. The processed tweet sentiment is then aligned with the corresponding price data based on timestamps, preserving the time-series nature of the dataset. This comprehensive preprocessing pipeline enhances the quality and relevance of the input features, enabling the model to detect patterns and relationships more accurately during training and prediction.

5.2.3 MODEL TRANING

The model training phase is the core component of the cryptocurrency price prediction system, where machine learning algorithms are trained on historical data to learn patterns and make accurate forecasts. In this project, the preprocessed dataset—comprising normalized price data and sentiment scores—is split into training and testing sets to evaluate model performance effectively. Various machine learning models are implemented, including Logistic Regression, LSTM, GRU, XGBoost, and LightGBM, with LightGBM chosen as the primary algorithm due to its high speed and performance with large datasets. During training, the models learn to map input features such as previous prices, technical indicators, and sentiment polarity to a target label representing future price movement (increase, decrease, or no change). Hyperparameters are tuned to optimize model accuracy and prevent overfitting. The trained models are then validated using unseen test data, and their performance is measured using metrics such as Accuracy, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). This phase ensures that the model generalizes well and is ready to make reliable predictions on new cryptocurrency data.

5.2.4 PREDICATION

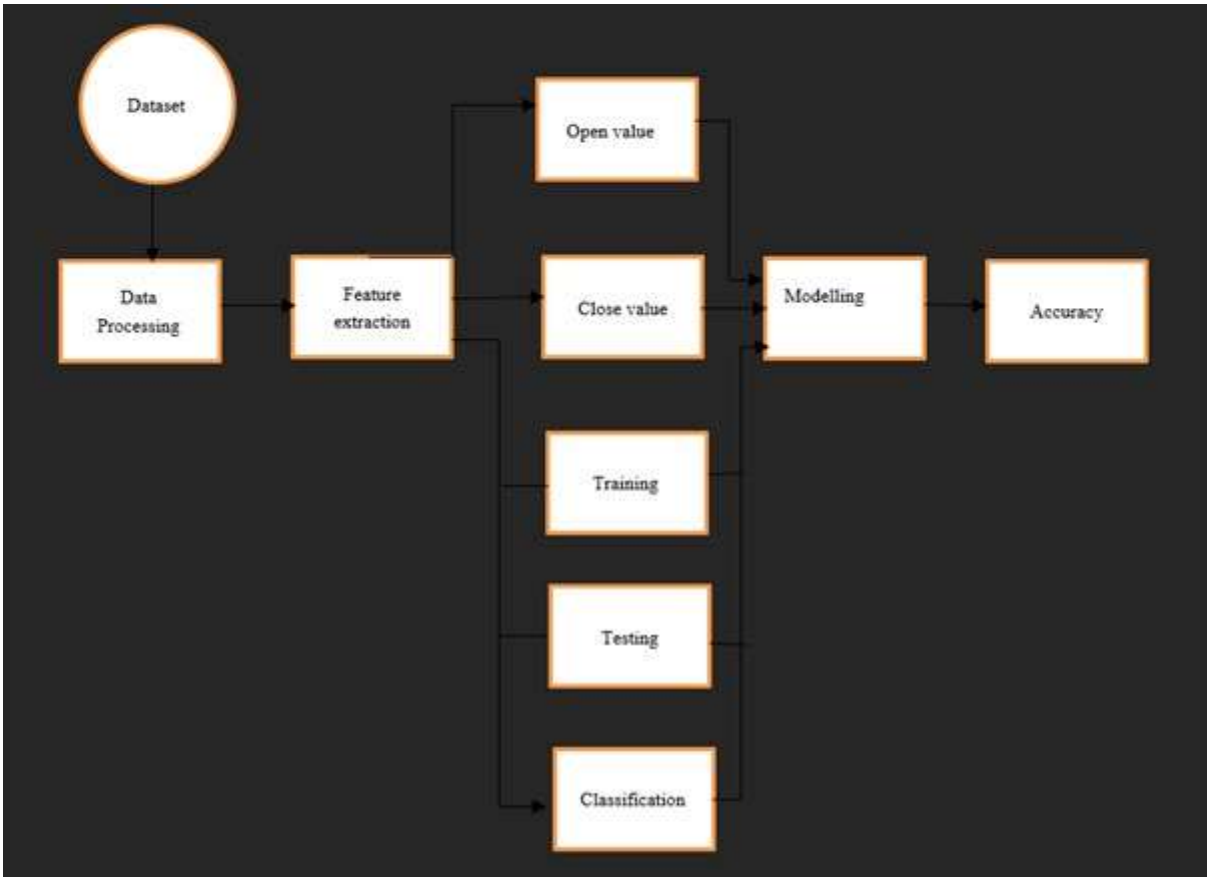
The prediction phase utilizes the trained machine learning model to forecast future cryptocurrency price movements based on new or recent data. Once the model has been validated and optimized, it is used to classify upcoming trends—whether the price of the cryptocurrency is likely to increase, decrease, or remain stable. The prediction is made using input features such as recent price history, sentiment scores derived from social media data, and technical indicators. In this system, the LightGBM model is primarily used due to its efficiency and high accuracy. The predicted output is then decoded and, if necessary, inverse-transformed to match the original scale of the price data. The final predictions are displayed in a user-friendly format, often accompanied by visualizations that compare actual and predicted values. These forecasts help users, analysts, and investors make informed decisions in the volatile cryptocurrency market by offering insights into potential short-term price behavior.

5.2.5 VISUALIZATION AND OUTPUT

The Visualization and Output module is designed to present the results of the cryptocurrency prediction system in an informative and user-friendly format. After predictions are made by the trained model, the output is visualized using graphical tools to help users better understand the model's performance and trends. Line graphs are used to compare actual vs. predicted cryptocurrency prices over time, while bar charts display classification results such as predicted price movements (increase, decrease, or stable). Sentiment trends derived from tweets are also plotted to show how public opinion correlates with price changes. Evaluation metrics such as accuracy, MAE, and RMSE are displayed alongside the graphs to give a quantitative overview of model performance. Visualization tools like Matplotlib, Seaborn, and Pandas are used for generating these plots. This module ensures that the insights gained from the prediction model are clearly communicated to users for better analysis and decision-making

6 SYSTEM ARCHITECTURE

The system architecture for the proposed cryptocurrency price prediction platform is designed as a modular pipeline to handle data acquisition, processing, modeling, and prediction. It begins with the **data collection layer**, which gathers historical and real-time cryptocurrency prices along with social media sentiment data, primarily using APIs such as Twitter REST API and Coindesk. The collected data is passed to the **pre-processing layer**, where it is cleaned, normalized using techniques like Min-Max scaling, and labeled using sentiment analysis tools such as TextBlob. In the next phase, the **feature engineering layer** extracts meaningful attributes like moving averages, sentiment scores, and lagged price features to enhance model inputs. The processed data is then fed into the **modeling layer**, which employs machine learning and deep learning algorithms such as Logistic Regression, LSTM, GRU, ARIMA, and XGBoost to train predictive models. These models are evaluated using metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and accuracy scores. The **prediction layer** uses the trained model to forecast whether the price of a cryptocurrency will rise, fall, or remain stable. Finally, the **visualization and output layer** presents results through graphical representations and reports, aiding users in making informed investment decisions. This layered architecture ensures scalability, flexibility, and accuracy in cryptocurrency price forecasting.



7.EXPIREMENTAL RESULTS

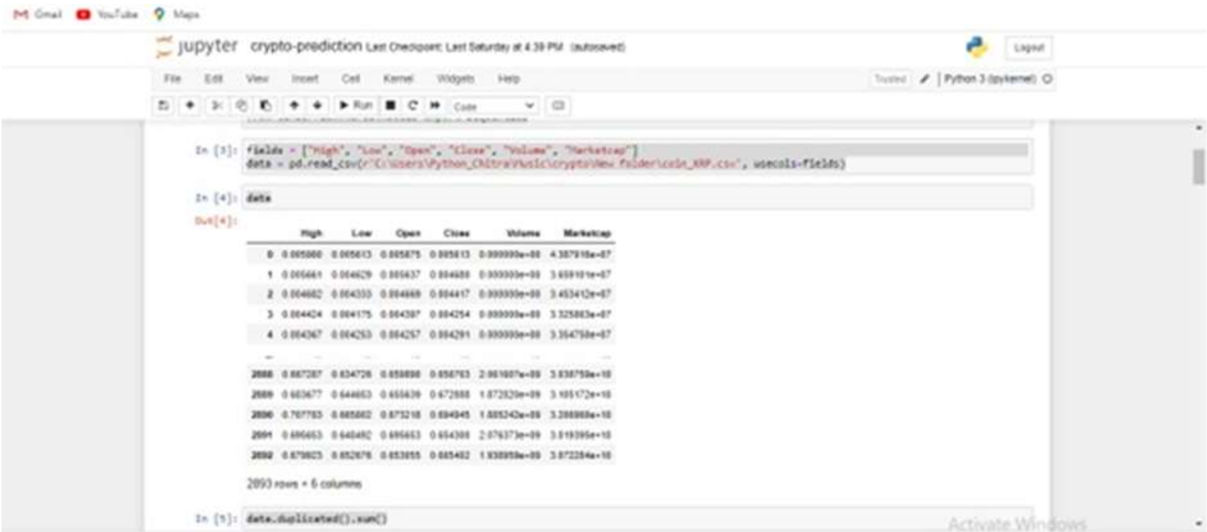


Figure 7.1: Crypto-Predication Last Checkpoint

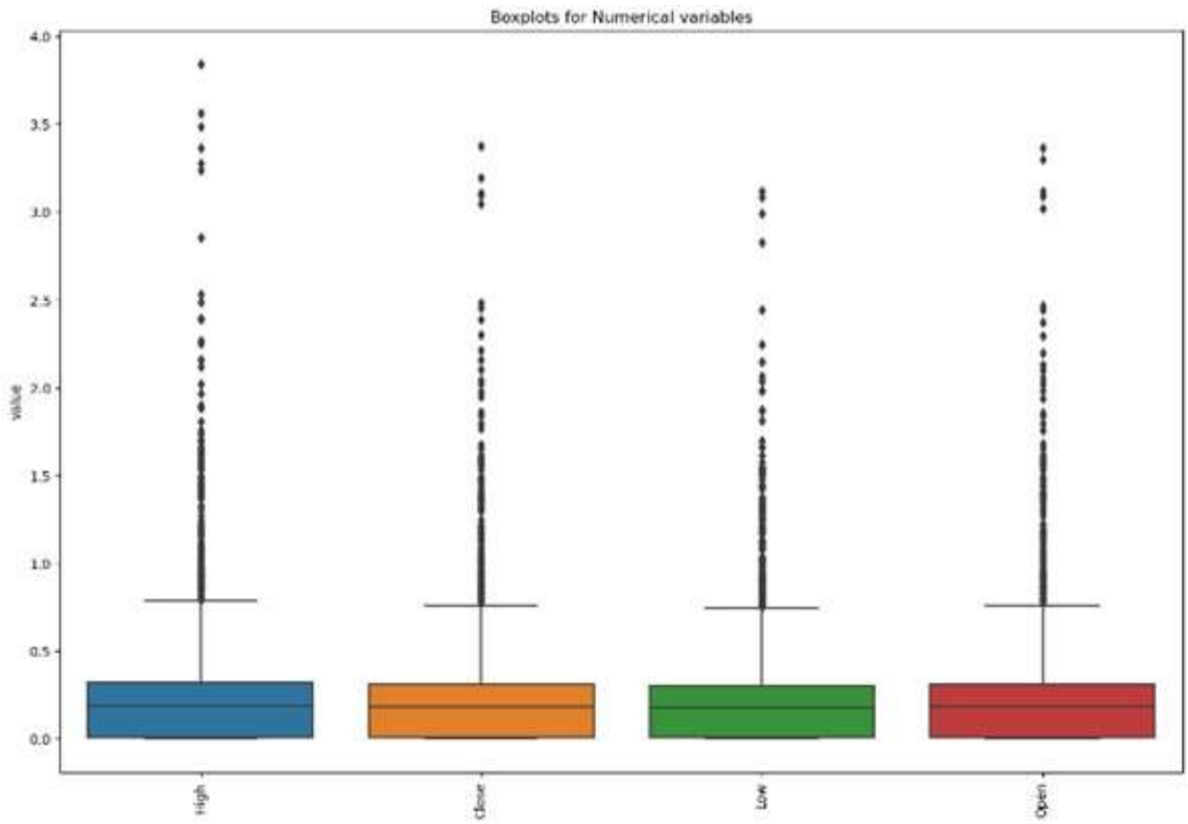
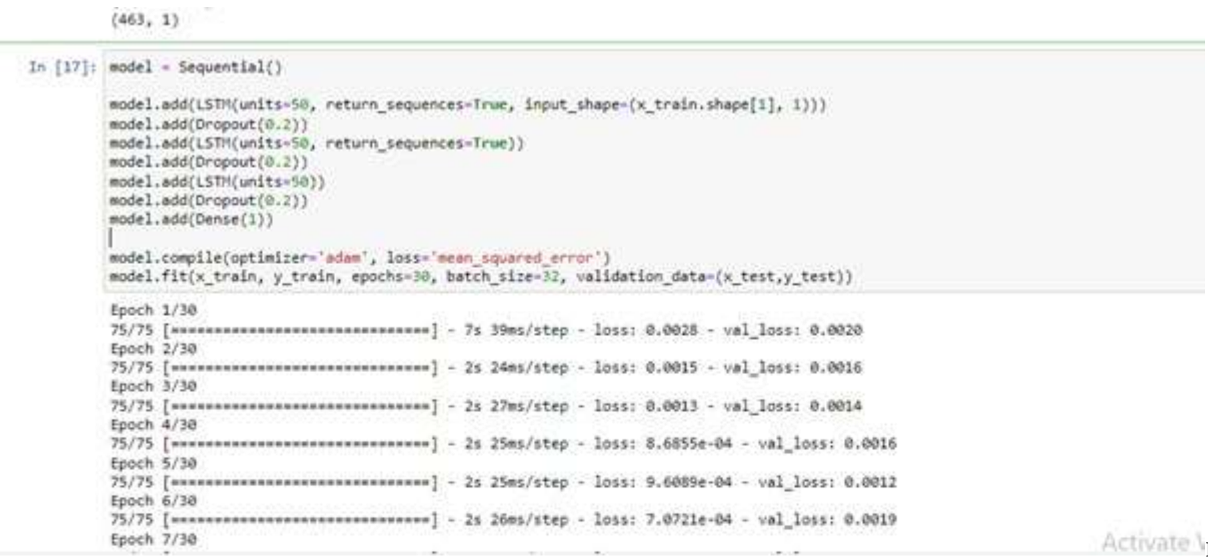


Figure 7.2:Boxplots for numerical variables



```
(463, 1)

In [17]: model = Sequential()

         model.add(LSTM(units=50, return_sequences=True, input_shape=(x_train.shape[1], 1)))
         model.add(Dropout(0.2))
         model.add(LSTM(units=50, return_sequences=True))
         model.add(Dropout(0.2))
         model.add(LSTM(units=50))
         model.add(Dropout(0.2))
         model.add(Dense(1))

         model.compile(optimizer='adam', loss='mean_squared_error')
         model.fit(x_train, y_train, epochs=30, batch_size=32, validation_data=(x_test,y_test))

Epoch 1/30
75/75 [==============================] - 7s 39ms/step - loss: 0.0028 - val_loss: 0.0020
Epoch 2/30
75/75 [==============================] - 2s 24ms/step - loss: 0.0015 - val_loss: 0.0016
Epoch 3/30
75/75 [==============================] - 2s 27ms/step - loss: 0.0013 - val_loss: 0.0014
Epoch 4/30
75/75 [==============================] - 2s 25ms/step - loss: 8.6855e-04 - val_loss: 0.0016
Epoch 5/30
75/75 [==============================] - 2s 25ms/step - loss: 9.6089e-04 - val_loss: 0.0012
Epoch 6/30
75/75 [==============================] - 2s 26ms/step - loss: 7.0721e-04 - val_loss: 0.0019
Epoch 7/30
```

Figure 7.3:Boxplots for numerical variables

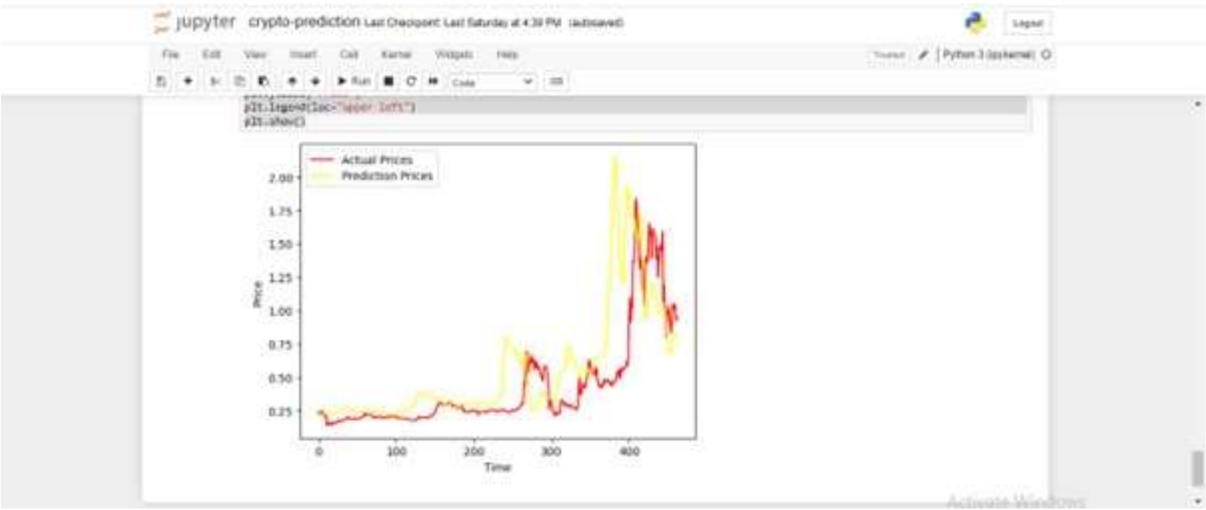


Figure 7.4: Architecture For Time Series Prediction

CONCLUSION AND FUTURE ENHANCEMENTS

CONCLUSION

Over the last several years, the cryptocurrency market has expanded significantly, catching the attention of both established corporations and newcomers to the space. By providing analysis and conclusions based on CRYPTOCURRENCY price data, it will aid in understanding the complicated and rapidly evolving sector. Although various theories and algorithms have been developed for the prediction of the price of bitcoins, most of them have been proved that they needed to be reconsidered for reducing problems of overfitting and errors resulting from high sized datasets. The value of bitcoin in the future can be predicted using the LSTM algorithm. Because of the usage of this algorithm, we can save a large amount of data and predict the most accurate results.

FUTURE ENHANCEMENTS

The work in progress, includes implementing hyperparameter tuning, in order to get a more accurate network architecture. Also, other features can be considered (although from our experiments with Bitcoin, more features have not always led to better results). Microeconomic factors might be included in the model for a better predictive result. Anyway, maybe the data we gathered for Bitcoin, even though it has been collected through the years, might have become interesting, producing historic interpretations only in the last couple of years. Furthermore, a breakthrough evolution in peer-to-peer transactions is ongoing and transforming the landscape of payment services. While it seems all doubts have not been settled, time might be perfect to act. We think it's difficult to give a mature thought on Bitcoin for the future.

REFERENCES

[1] CoinMarketCap. Global Cryptocurrency Charts. Accessed: Jan. 13, 2023. [Online]. Available: https://coinmarketcap.com/charts/

[2] CoinMarketCap. Top Cryptocurrency Spot Exchanges. Accessed: Jan. 13, 2023. [Online]. Available: https://coinmarketcap.com/rankings/ exchanges/

[3] A. M. Khedr, I. Arif, M. El-Bannany, S. M. Alhashmi, and M. Sreedharan, ''Cryptocurrency price prediction using traditional statistical and machinelearning techniques: A survey,'' Intell. Syst. Accounting, Finance Manage., vol. 28, no. 1, pp. 3–34, Jan. 2021, doi: 10.1002/isaf.1488.

[4] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel, and B. K. Lama, ''Recurrent neural network based bitcoin price prediction by Twitter sentiment analysis,'' in Proc. IEEE 3rd Int. Conf. Comput., Commun. Secur. (ICCCS), Oct. 2018, pp. 128–132, doi: 10.1109/CCCS.2018.8586824.

[5] T. Shintate and L. Pichl, ''Trend prediction classification for high frequency bitcoin time series with deep learning,'' J. Risk Financial Manage., vol. 12, no. 1, p. 17, Jan. 2019. [Online]. Available: https://www.mdpi.com/1911-8074/12/1/17

[6] A.-D. Vo, ''Sentiment analysis of news for effective cryptocurrency price prediction,'' Int. J. Knowl. Eng., vol. 5, no. 2, pp. 47–52, 2019, doi: 10.18178/ijke.2019.5.2.116.

[7] J. Chen. Technical Indicator: Definition, Analyst Uses, Types and Examples. Investopedia. Accessed: Jan. 16, 2023. [Online]. Available: https://www.investopedia.com/terms/t/technicalindicator.asp

[8] C. Zhuang. Stockstats (0.5.1). [Online]. Available: https://github.com/ jealous/stockstats [9] B. Dolan. MACD Indicator Explained, With Formula, Examples, and Limitations. Accessed: Jan. 16, 2023. [Online]. Available: https://www. investopedia.com/terms/m/macd.asp

[10] J. Fernando. Relative Strength Index (RSI) Indicator Explained With Formula. Accessed: Jan. 16, 2023. [Online]. Available: https://www. investopedia.com/terms/r/rsi.asp

[11] C. Mitchell. What is the Commodity Channel Index (CCI)? How to Calculate. Accessed: Jan. 16, 2023. [Online]. Available: https://www.investopedia.com/terms/c/commoditychannelindex.asp