

# Bitcoin Price Prediction and Forecasting

K.Ranjith Reddy<sup>1</sup>, Kiran Lenka<sup>2</sup>, N. Devi<sup>3</sup>, D.V. Rithika<sup>4</sup>

<sup>1</sup>Assistant Professor, Computer Science and Engineering, JNTUH/CMRTC, Hyderabad, Telangana, India

<sup>2,3,4</sup>B.Tech Student, Computer Science and Engineering, JNTUH/CMRTC, Hyderabad, Telangana, India

\*\*\*

**Abstract** - Bitcoin is one of the oldest and biggest cryptocurrencies being traded as of now, in terms of the volume being traded. It is a digital asset over which central banks or any financial Institutions have no control or regulations. Bitcoin has a market share of more than 55% as compared to other cryptocurrencies. It is very sporadic and this is one of the most important reasons which attracted us to analyze and predict its price. Here, we intend to study the prediction of Bitcoin prices using Machine Learning Techniques and prepare a strategy to maximize gains for investors.

**Key Words:** Bitcoin prices, LSTM, Tweets, Reddit posts.

## 1. INTRODUCTION

The stock market is one of the most volatile data available in terms of Machine learning datasets. Researchers have been long trying to predict the stock market and any breakthrough in this field would result in, literally, the people being able to mint money. Cryptocurrencies, to be specific, have gained a lot of traction in recent years from investors across the globe. There are several reasons why Bitcoin's price history has been so volatile. It is important to learn the factors that affect its market price so you can decide whether to invest, trade, or keep an eye on its development. As with most commodities, investments, assets, or other products, Bitcoin's price is heavily influenced by supply and demand. Since Bitcoin is a rapidly adopted asset by investors and traders, speculation about price movements plays a crucial role in its value at any given point in time.

Given the vastness of investments in Bitcoin, Cryptocurrencies are relatively unpredictable compared to traditional financial predictions like stock market predictions. Due to their high volatility, they offer a great opportunity for high profit if intelligent investing strategies are utilized. Here the Crypto-situation is analysed based on sentiment analysis from various platforms along with multiple learning approaches as a decent model will be able to provide the crypto investor with accurate data.

## 2. LITERATURE SURVEY

According to a research from 2016, more than 600 papers have been published on this topic. Our review of the literature focuses on work on bitcoin (BTC) price prediction utilising various methodologies, as well as the

requirement for an evaluation of recurrent neural networks (RNN) and their system design. Dennys et al.[2] applied machine learning methods such as artificial neural networks (ANN), support vector machines (SVM), and recurrent neural networks (RNN), as well as K-means clustering, to get the most important features and used different attributes selection mechanisms to get the most important features. One limitation of this study is that it is solely focused on investors. Because bitcoin has the potential to transform the dynamics of the global economy, policymakers should be regarded as major partners in the system. Sean McNally et al. [3] used a Bayesian optimized recurrent neural network and LSTM to predict the direction of Bitcoin price in USD. They also compared the deep learning approaches using the ARIMA model.

Bitcoin purchases and sales can provide enormous amounts of benefits when done correctly. It has proven to be a fortune for many people in the past and is still making them a lot of money today. The downside of this is that it doesn't come without a price. Likewise, if not planned and calculated properly, you can lose a lot of money. You should have a great understanding of how and precisely why bitcoin price changes, which means you need to know how people make their bitcoin predictions. As well as these things (supply and demand, regulations, news, etc. ), one must consider the technology of bitcoin and its progress. This aside, we now have to deal with the technical part using various algorithms and technologies which can predict precise bitcoin price movements. In terms of machine learning and deep neural networks, a time series is normally a sequence of numbers that are associated with a specific point in time. Since this is a time-series data set, it should be split into two parts: inputs and outputs. Further, LSTM can easily handle multiple input problems, compared to the classic statistics linear models. The idea of this project is to use Deep learning techniques such as the LSTM model combined with Sentiment analysis and prediction using Machine Learning techniques. Investors will be able to foresee changes in Bitcoin values with the support of this prediction model, which will also help them limit the risk of losing money. In order to analyze and predict sentiment, we will use a Machine Learning tool, which will analyze texts based on their polarity, ranging from positive to negative. With these examples of emotions in text, machines automatically learn how to detect sentiment without human input. Sentiment analysis on Reddit and Twitter are popular social media platforms where

discussion on CryptoCurrency is higher when compared to other sites. Analyzing data with visual elements such as charts, graphs, and maps and seeing patterns, outliers, and trends. Training and testing of dataset extracted from CryptoCompare tool to predict the price of Bitcoin using the LSTM neural network along with Sentiment analysis to analyze the model performance.

### 3. METHODOLOGY

To reach the goals of this study, we used historical cryptocurrency prices to train three separate models for three different types of cryptocurrency price prediction. Then, to assess the effectiveness of the proposed schemes, we compare the accuracy of our proposed model to that of existing models in five stages: (1) gathering historical AI 2021, 2 481 cryptocurrency data; (2) data exploration and visualisation; (3) training three different types of models; (4) testing the models; and (5) extracting and comparing the findings. In this section, we introduce and analyse three types of algorithms in this section: long short-term memory (LSTM), gated recurrent unit (GRU), and bidirectional LSTM (bi-LSTM)—to forecast the price of three types of cryptocurrency based on historical data: Bitcoin (BTC), Litecoin (LTC), and Ethereum (ETH) [1]. Sentiment analysis and market sentiment prediction on popular social media platforms such as Twitter and Reddit using tweets and posts related to Bitcoin will be performed. Bitcoin price data collected using the CryptoCompare API (provides highly reliable historical data for the public) and Data visualization will help investors to know the volatility in the prices of bitcoins.

Some of the previous studies' drawbacks are that they take into account a variety of features that aren't as useful in stock price prediction. Some studies miss many aspects that provide insight into the data. Existing efforts are best suited to less volatile stock market prices and cannot be used for extremely volatile cryptocurrencies like Bitcoin. Here we aim to train and test the dataset to predict the price of Bitcoin using the LSTM neural network. Visualize the prediction results in the form of a graphical representation that contains the information and the data. The proposed methodology has certain advantages because of which we will be able to predict the sentiment of tweets, news headlines and Reddit posts related to bitcoin, With the help of this system, one can predict the stock price of Bitcoin and make their investments carefully.

The risk of loss can be avoided by predicting the price of Bitcoin beforehand. This system with little modification can be used to predict the price of any cryptocurrency.

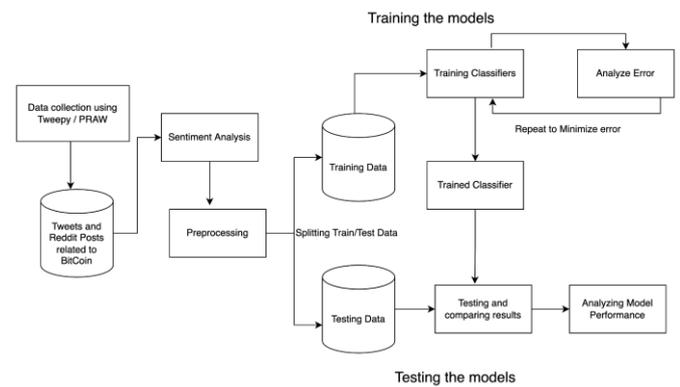


Figure 1: The model architecture of sentiment analysis

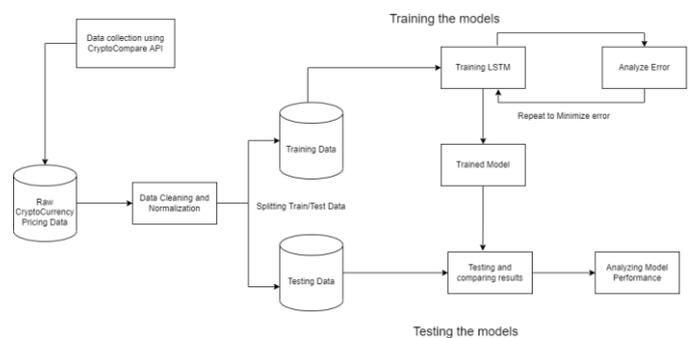


Figure 2: The model architecture of BTC price prediction

#### 3.1 DESCRIPTION

**Input Data:** There are two forms of input data in this system. The first type of input data is the tweets of current trends and prices related to Bitcoin. It is a CSV file with approximately 8000 tweets. The second type of input data is the Reddit posts related to Bitcoin prices where content is socially curated and promoted by the social website members through upvoting. This is used to create, train, and assess the models' performance.

**Data Cleaning and Preprocessing:** In this step, the input data is cleaned and preprocessed. This step involves the removal of emoticons, hashtags, mentions, numerical values, punctuations and stop words, stemming and lemmatization.

**Sentiment Analysis and Classification:** Here the machine visualizes how many positive, neutral and negative tweets are present using bar plots, training and testing 6 different models. It then plots the performance between these models.

**Data Visualization:** In this stage, you'll see historical price movements, compare them to the Nifty-50 and the BSE-Sensex, see volatility, collect data with CryptoCompare, train and test the LSTM model, and see how it performs.

**Building Models:** Multiple machine learning algorithms are built based on the features extracted using various natural language processing techniques. Such as Logistic Regression, Random Forest, SGD, Support-Vector Machine, LightGBM, Gradient Boosting Classifier.

**Evaluating Models:** The models built are carefully evaluated on the test dataset using various performance metrics such as Precision, Accuracy, F1 Score etc. and are compared. The whole system can be divided into two major phases such as training and testing.

**Training the models:** Here multiple models are trained using the labelled dataset and this dataset is used to train Machine Learning Algorithms and is fed with sufficient training data to learn from.

**Testing the models:** The trained models are tested and their results are considered to fine-tune the parameters of the model. It is the process where the performance of a fully trained model is evaluated on a testing set.

### 3.2 LSTM IMPLEMENTATION

In [4], the researchers evaluated a key feature of feed networks is that they do not save memory. As a result, each input is processed independently, with no preserved state in the middle. We should maintain track of future events because we are dealing with a series of situations when information from the prior Bitcoin price is needed. Long short-term memory (LSTM) networks are a sort of recurrent neural network that can learn how to predict sequences based on their order. This is a necessary operation in complex problem areas such as machine translation and speech recognition.

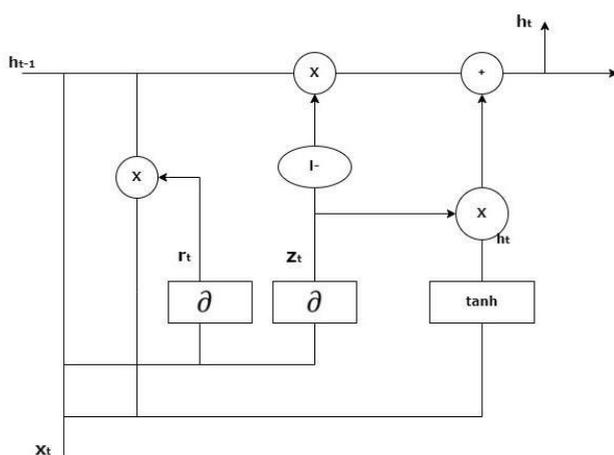


Figure 3: Long Short Term Memory Network

As LSTMs has grown in popularity, various changes to the traditional LSTM architecture have been attempted to simplify the internal design of the cell, make the cell work more efficiently, and reduce computational complexity.

Gers and Schmidhuber have introduced a peephole connection where the gate layer can always know the state of the cell. Some LSTMs also used combined input and oblivion gates instead of two separate gates to help make both decisions at the same time. Another variation was the use of a gated regression unit (GRU). This has improved design complexity by reducing the number of gates. The combination of cell state and hidden state, and forgotten update gates and input gates are merged and used.

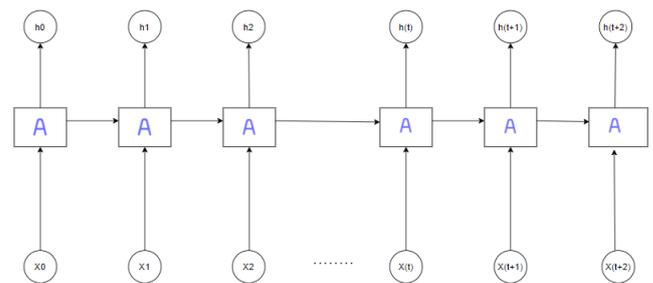
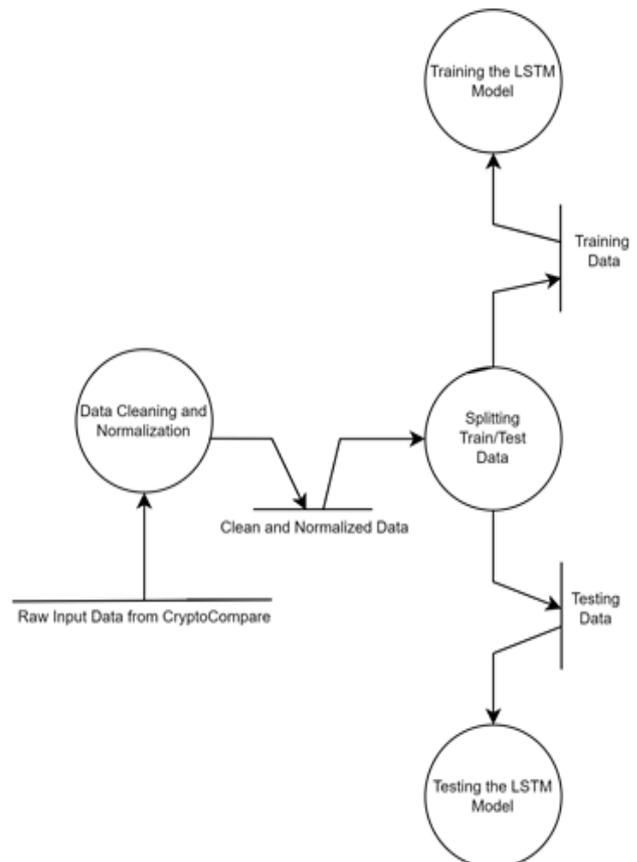


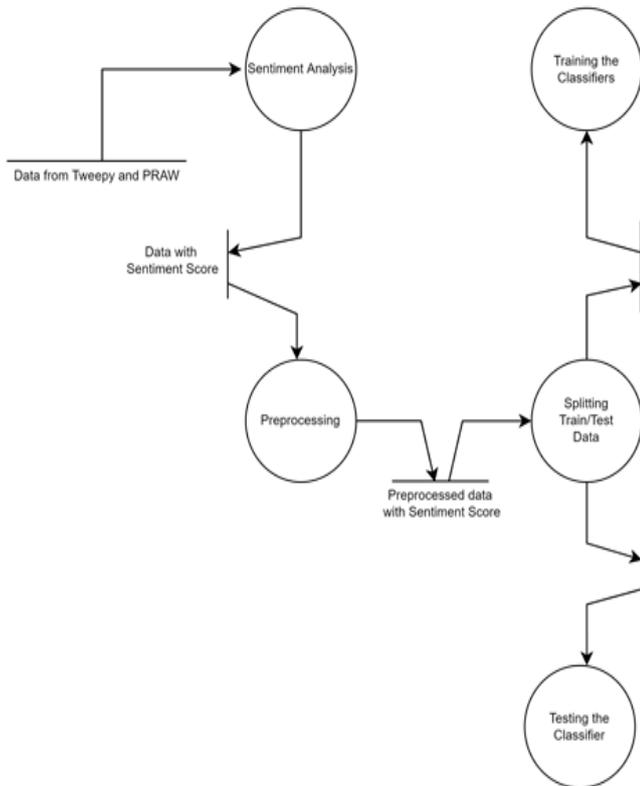
Figure 4: An unrolled recurrent neural network

## 4. SYSTEM DESIGN

### 4.1 DATA FLOW DIAGRAM DURING BTC PRICE PREDICTION USING LSTM

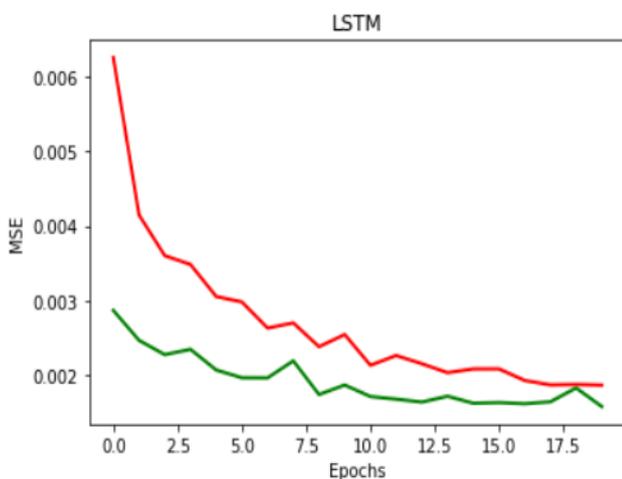


#### 4.2 DATA FLOW DIAGRAM DURING BITCOIN MARKET SENTIMENT ANALYSIS USING DATA FROM TWITTER AND REDDIT POSTS

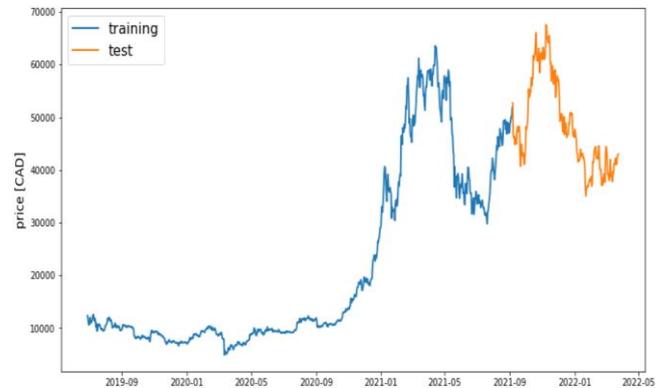


### 5. CRYPTOCURRENCY PRICES

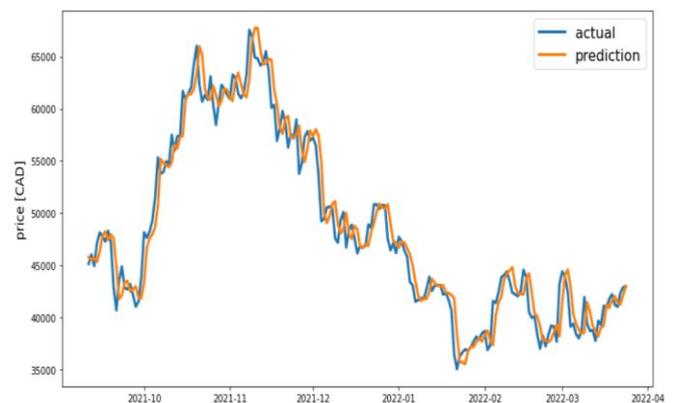
#### 5.1 GRAPH REPRESENTING MSME AND EPOCHS USING LSTM



#### 5.2 CRYPTOCURRENCY PRICES IN INDIAN RUPEES AS A FUNCTION OF TIME



#### 5.3 GRAPH PLOT OF ACTUAL AND PREDICTED PRICES



### 6. CONCLUSION AND FUTURE SCOPE

BitCoin Prediction is the actual demand for beneficial business as the use of cryptocurrency is growing every day. Predictions are always helpful to decrease risk factors in any business environment. Risk factors can be analyzed on the basis of historical data and previous business trends. The results of this project will help us predict the trends beforehand when investing in any cryptocurrency and help us reduce the chances of facing a loss.

With the help of this prediction model, investors will be able to know the changes in the prices of Bitcoin beforehand. Deep learning techniques such as the LSTM model combined with Sentiment analysis and prediction using Machine Learning, will help them reduce the chances of facing a loss. This system with little modification can be used to predict the price of any cryptocurrency in the near future.

With regards to future work, we seek to expand our dataset, both in terms of numbers and diversity, as a first step to achieving better results. More relevant labels can be added according to the status quo, and various deep learning models can be experimented on.

## REFERENCES

- [1] Hamayel, Mohammad J., and Amani Y. Owda. 2021. "A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms" *AI* 2, no. 4: 477-496. <https://doi.org/10.3390/ai2040030>
- [2] Mallqui, Dennys C.A., and Ricardo A.S. Fernandes. 2018. "Predicting The Direction, Maximum, Minimum And Closing Prices Of Daily Bitcoin Exchange Rate Using Machine Learning Techniques". *Applied Soft Computing* 75: 596-606. doi:10.1016/j.asoc.2018.11.038.
- [3] McNally, Sean, Jason Roche, and Simon Caton. 2018. "Predicting The Price Of Bitcoin Using Machine Learning". In *2018 26Th Euromicro International Conference On Parallel, Distributed And Network-Based Processing (PDP)*, 23. IEEE. <https://ieeexplore.ieee.org/document/8374483>.
- [4] M, Vaidehi, Alivia Pandit, Bhaskar Jindal, Minu Kumari, and Rupali Singh. 2021. "BITCOIN PRICE PREDICTION USING MACHINE LEARNING". *International Journal Of Engineering Technologies And Management Research* 8 (5): 20-28. doi:10.29121/ijetmr.v8.i5.2021.953.