Cryptocurrency price prediction using historical data and news sentiment analysis

Aryansh Garg
Deparment of CSE
PES University
Bangalore,India
pes2ug20cs070@pesu.pes.edu

Harsh Agrawal
Deparment of CSE
PES University
Bangalore,India
pes2ug20cs908@pesu.pes.edu

Sagarikha M
Deparment of CSE
PES University
Bangalore,India
pesug20cs292@pesu.pes.edu

Harshit Sharma
Department of CSE
PES University
Bangalore,India
pes2ug20cs909@pesu.pes.edu

Prof. Gauri S Rapate
Deparment of CSE
PES University
Bangalore,India
gauri.rapate@pesu.pes.edu

Abstract— In the fast-evolving landscape of e-commerce, finance, and blockchain, cryptocurrencies have taken center stage as a global phenomenon marked by notable price volatility. Beyond traditional predictive models relying on historical data, this paper introduces a groundbreaking approach by harnessing the power of both historical price trends and the burgeoning realm of online news. With an emphasis on the second-largest cryptocurrency by market capitalization, Ethereum, the proposed model adeptly integrates sentiment analysis to forecast price fluctuations. The research underscores the pivotal role of news data in providing investors with a strategic edge, revealing how a comprehensive understanding of market sentiment significantly enhances cryptocurrency price prediction. In a financial landscape increasingly shaped by interconnected digital ecosystems, this study demonstrates the indispensability of sentiment analysis for making informed decisions in the dynamic realm of cryptocurrency trading.

I. INTRODUCTION

The cryptocurrency market has generated interest of investors through the Internet. A cryptocurrency is a digital currency or asset created to work as a medium of exchange that exploits strong cryptography to protect financial transactions, handle the creation of additional units, and certify the transfer of assets. Bitcoin (BTC) and Ethereum (ETH) are two largest cryptocurrencies in terms of market capitalization. Both BTC and ETH have experienced significant price fluctuations on both daily and long-term valuations. Recently, cryptocurrency market movement prediction systems have emerged to help investors make an informed decision. Traditional supervised learning algorithms were used for predicting changes in cryptocurrency prices based on historical price data. However, the efficient market hypothesis (EMH) [1] states that the cryptocurrency market always follows a haphazard pattern and its prediction is always a challenging task.

As traditional financial markets, there is a relation between public moods and the cryptocurrency market. Discerning the impact of news on price movement can provide a buying and selling advantage to investors. Nowadays, more and more prevalence of related news articles and social media posts is progressively becoming available on the internet. However, it is overwhelming for people to read and figure out the most common public feelings and emotions. Therefore, automated sentiment discovery systems are needed to assist investors in making a better-informed trading decision.

The proposed model includes two stages - sentiment analysis and price prediction. In the first stage, the system takes news data as input and produce sentiment score as output. In the second stage, the sentiment score from the first stage is fed back as additional input into an artificial recurrent neural network (RNN) architecture to predict cryptocurrency prices. Long short-term memory (LSTM) network has been used which is well-suited for making predictions based on time series data. The fact that most of the researchers use LSTM network to predict cryptocurrency prices, but this system improves the accuracy and usefulness of its news sentiment data by leveraging sentimental information. It has been showed that sentiment analysis is an important perspective for cryptocurrency price prediction due to the interactive nature of financial activities..

II. LITERATURE REVIEW

Stock market always follows a haphazard pattern and its prediction always a complicated process. With the rise of technology in the world of finance, researchers have attempted a variety of analysis approaches such as technical analysis, quantitative analysis, and so on. In recent years, advances in machine learning allow many investors to do market research better, manage funds more efficiently, invest and expand more effectively. Regression and SVM are classic approaches that dominate prediction methods. However, regression models are widely applied since they allow to determine the impact of each variable included and evaluate the importance of factors by dropping them out.

Several studies have explored the application of different machine learning algorithms, such as neural networks, support vector machines, and deep learning, for cryptocurrency price prediction. These studies

demonstrate the effectiveness of machine learning ensemble techniques, random forests, stochastic gradient boosting machines, and hybrid models combining LSTM and GRU for accurate price forecasts.

People often consider the cryptocurrency market as another version of the stock market with no difference. But the fact is that cryptocurrency market is very much different from the ordinary stock market in tangible value, user base, price volatility, and so on. Cryptocurrency is a means of exchange whose value is based solely on popularity. Therefore, our model achieved satisfactory experimental results by relying on particular information obtained from news data.
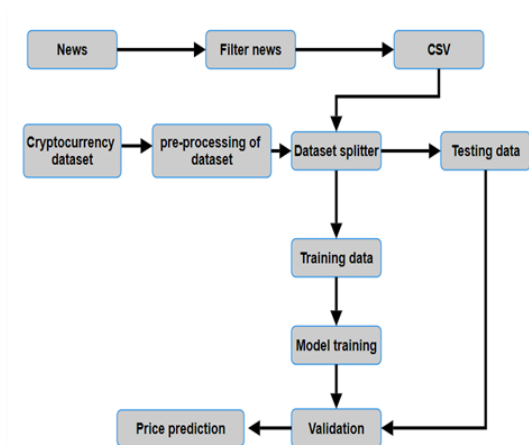
III. METHODOLOGY

Fig. 1 Overview of Prediction Model

As a first step, historical price data of Ethereum has been collected which includes columns "Open", "Close", "High", "Low" and "Adj close". The proposed model is a multivariate LSTM model which utilizes all five price indicators as input to the model. The model incorporates four layers which include two activation layer, one dense layer and one dropout layer. The model being multivariate, it can predict the price only one day in advance since it requires all five features together to make the prediction.

Date	Open	High	Low	Close	Adj Close	Volume
2018-01-01	755.757019	782.530029	742.004028	772.640991	772.640991	2595760128
2018-01-02	772.346008	914.830017	772.346008	884.443970	884.443970	5783349760
2018-01-03	886.000000	974.471008	868.450989	962.719971	962.719971	5093159936
2018-01-04	961.713013	1045.079956	946.085999	980.921997	980.921997	6502859776
2018-01-05	975.750000	1075.390015	956.325012	997.719971	997.719971	6683149824

Fig. 2 Data values of Crypto

To stay up-to-date with the evolving cryptocurrency market, discerning the impact of news on price movement can provide a buying and selling advantage to investors. The analysis of web and social media conversations brings a lot of value to better understanding the specific expectations of the public and the emerging changes in market trends. To do this kind of market psychology analysis, the ability of news data to predict price fluctuations for the largest

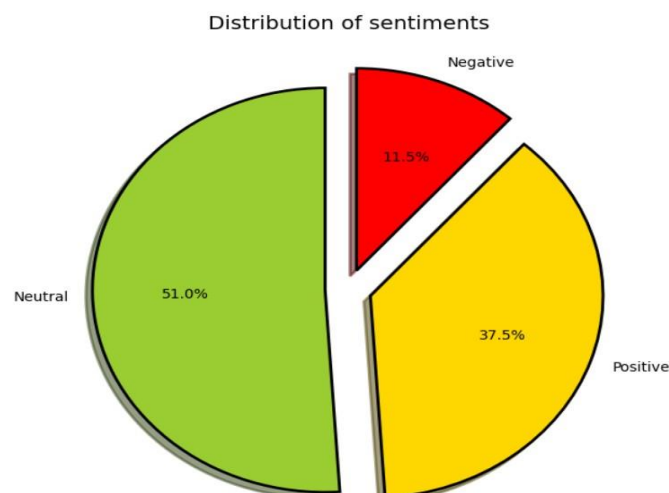cryptocurrency in terms of market capitalization: Ethereum has been analysed.

Fig. 3 Chart representing sentiment

The final version of the model was able to correctly predict cryptocurrency price using historical data and sentimental information gained from news data. It has been showed that sentiment analysis is an important perspective for cryptocurrency price prediction due to the interactive nature of financial activities.

Since it is a time series forecasting model, number of inputs of the historical data plays a crucial role. For the proposed model, the number of past data inputs taken were 15 which means the model takes past 15 days prices and predict the prices of 16 days from the first day giving one day prediction in advance.

Long Short-Term Memory (LSTM)

For various learning issues involving sequential data, recurrent neural networks with long short-term memory (LSTM) have emerged as an effective and scalable approach. They are useful for capturing long-term temporal dependencies since they are generic and effective. The LSTM is an RNN-style architecture with gates that govern the flow of information between cells. The input and forget gate structures can modify information traveling along the cell state, with the ultimate output being a filtered version of the cell state based on context from the inputs.
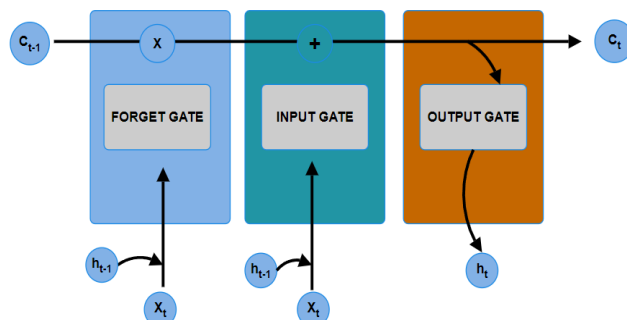

Fig. 4 Flow of LSTM Model

For sentiment analysis, polarity has been calculated and the values has been normalized to be in the range -1 to 1 and this value is called sentiment score where the lesser the

score the negative the sentiments. News data has been extracted using SentiTrade API.

In the final stage, the sentiment score from the first stage is fed back as additional input into a RNN network to predict cryptocurrency prices. The prediction is built on a LSTM network, which is well-suited for making predictions based on time series data. The LSTM takes historical cryptocurrency price datasets and obtained sentiment scores as input and produces a predicted price as output.

IV. CONCLUSION

The final predictions made by the proposed model are compared wit the actual prices of the cryptocurrency Ethereum and plotted to get an idea of how accurate the model is working.
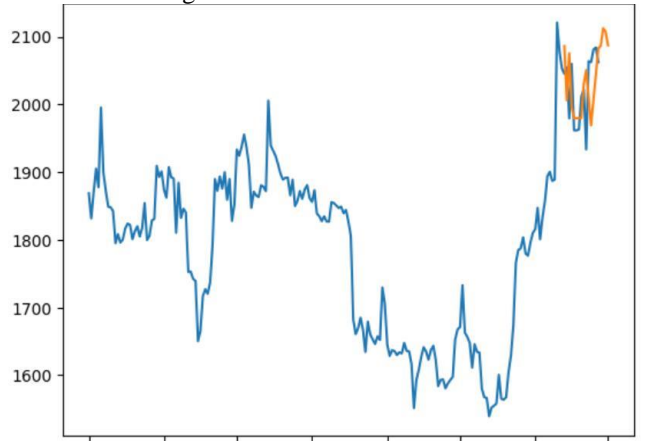

Fig. 5 Graph of Actual value vs. Predicted value

The blue line shows the actual price of Ethereum and orange line shows the price predicted by the proposed model.

The graphs shows that the proposed model is successful in predicting the prices of Ethereum despite it having volatile nature.

The final version of the model was able to correctly predict cryptocurrency price using historical data and sentimental information gained from news data. It has been showed that sentiment analysis is an important perspective for cryptocurrency price prediction due to the interactive nature of financial activities. This strategy offers a new approach to addressing cryptocurrency price prediction and can be applied to challenging tasks such as social media tracking, popularity, trending, and sentiment analysis-based event detection.

V. REFERENCES

[1] A. D Vo, Q. P. Nguyen, and C. Y. Ock, "Opinion–Aspect relations in cognizing customer feelings via reviews," IEEE Access, vol. 6, p. 6, 2018

[2] Andrei-Alexandru Ecean , Daniel Zinca : "Cryptocurrency Price Prediction using LSTM and GRU Networks" , 2022 International Symposium on Electronic and Telecommunications

[3] R. Parekh et al., "DL-GuesS: Deep Learning and Sentiment Analysis-Based Cryptocurrency Price Prediction," in IEEE Access, vol. 10, pp. 35398-35409, 2022, doi: 10.1109/ACCESS.2022.3163305.

[4] Seng hansun, Arya Wicaksana, Abdul Q.M. Khaliq : "Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches" – in Journal of Big Data, Article no. 50, April 2022

[5] Lekkala Sreekanth Reddy, Dr.P. Sriramya : "A Research On Bitcoin Price Prediction Using Machine Learning Algorithms" - INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 04,APRIL 2020