ISSN: 2278-4632
Vol-13, Issue-05, May 2023

A deep learning-based cryptocurrency price prediction model that uses on-chain data

1. Mr. Saddala Praveen Kumar, Assistant professor, Department of ECE, JNTUH University College of Engineering Jagtial, Telangana, India, praveenkumar_00019@yahoo.com

2. G. Ragini, Department of ECE,JNTUH University College of Engineering Jagtial, Telangana, India, raginigolla28@gmail.com

3. Kancharla Pavan, Department of ECE, JNTUH University College of Engineering Jagtial, Telangana, India, kancharlapavan001@gmail.com

4. Kethireddy Nithin Reddy, Department of ECE, JNTUH University College Of Engineering Jagtial, Telangana, India, nithinkethireddy123@gmail.com

ABSTRACT: The underlying decentralisation and transparency of cryptocurrencies has recently sparked a lot of attention from investors. Given the volatility and distinctive qualities of cryptocurrencies, precise price prediction is crucial for creating effective trading strategies. In order to do this, the authors of this paper suggest a cutting-edge framework that forecasts the price of Bitcoin (BTC), a well-known cryptocurrency. The change point detection approach is used for steady prediction performance in unobserved price range. Time-series data are segmented in particular so that normalisation may be carried out individually based on segmentation. On-chain data is also gathered and used as an input variable to forecast prices. On-chain data refers to the distinct records recorded on the blockchain that are intrinsic in cryptocurrencies. Additionally, this paper suggests using SAM-LSTM, which combines multiple LSTM modules for on-chain variable groups and the attention mechanism, as the prediction model. SAM-LSTM stands for self-attention-based multiple long short-term memory. The usefulness of the suggested framework in predicting BTC prices has been demonstrated in experiments using real-world BTC price data and several technique settings. The greatest MAE, RMSE, MSE, and MAPE values were 0.3462, 0.5035, 0.2536, and 1.3251, respectively, and the findings are encouraging.

Keywords – *Blockchain, cryptocurrency, Bitcoin, deep learning, prediction methods, change detection algorithms.*

1. INTRODUCTION

With the introduction of blockchain technology, both the form of currency and transactions have undergone tremendous change. Since its inception, the primary function of money has been as a method of payment and a vehicle for the distribution of value. Trust in the currency, which is guaranteed and stabilised by a central organisation, is necessary for this function (e.g., government, bank). The potential for depravity that might jeopardise transaction

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

dependability is a serious weakness for central authority. The open, tamper-proof, anti-counterfeiting blockchain has given rise to a money known as bitcoin. Based on blockchain technology, bitcoin deviates from the conventional relationship by enabling confidence without the assurance of a central authority. A monetary system that eliminates fraud risks and protects privacy is possible with cryptocurrency that ensures decentralisation and transparency [2]. Regarding how it differs from currently used traditional currencies, the most popular cryptocurrency, Bitcoin (BTC), is a model cryptocurrency. Due to the 21 million cap on BTC issuances, there is almost no inflation brought on by a central authority creating money [3]. By enabling cryptocurrencies to serve as a means of value storage as well as a way of exchange, this deepens the notion of decentralisation. In fact, investing in cryptocurrencies is presently seen to be one of the most efficient methods to raise asset value, in addition to conventional investment vehicles.

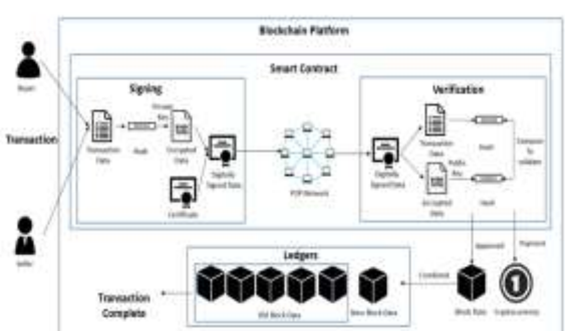

Fig.1: Example figure

Unlike conventional assets (such as gold, stocks, fiat money, etc.), cryptocurrencies lack an entity of ownership, are intangible, and are volatile (i.e., corporation). The existence of on chain data, which includes data obtained from the blockchain, is another distinguishing feature of cryptocurrencies [6]. On-chain data includes important details about the blockchain network, such as transactions, block size, and mining complexity. As a result, it is impossible to immediately apply current traditional asset classification criteria and indicators to cryptocurrencies. Given the aforementioned considerations, a creative strategy that emphasises the distinctive qualities of cryptocurrencies is essential for effective implementations.

2. LITERATURE REVIEW

Stochastic neural networks for cryptocurrency price prediction:

With the development of blockchain technology over the past several years, there has been a significant rise in the use of cryptocurrencies. However, because of the market's unpredictable behaviour and excessive price volatility, cryptocurrency is not viewed as a good investment prospect. Due to their deterministic character, the majority of the techniques presented in the literature for cryptocurrency price forecasting may not be appropriate for real-time price prediction. We provide a stochastic neural network model for predicting cryptocurrency prices in response to the aforementioned problems. The idea of random walks, which is frequently employed in financial markets for stock price modelling, forms the foundation of the suggested method. In order to

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

replicate market volatility, the proposed approach introduces layer-wise randomization into the observed feature activations of neural networks. The prediction model also incorporates a method for learning the market's reaction patterns. For Bitcoin, Ethereum, and Litecoin, we trained the Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) models. The outcomes demonstrate that the suggested model is superior to the deterministic models. Cryptocurrency, Multilayer Perceptrons, Long Short-Term Memory, Random Walks, and Stochasticity are INDEX TERMS.

Privacy and cryptocurrencies—A systematic literature review

In the existing centralised banking system, our transaction history has the potential to expose a lot of personal information about each spender, both to the banking system itself and to those that surround it (e.g., governments, industry etc). The quantities spent, the items purchased with those amounts, the places where we spend our money, and the people with whom we trade currency are a few examples of information that is leaked. Those who possess this information can utilise it in a variety of ways that aren't necessarily to their advantage. Cryptocurrencies, like the well-known Bitcoin, were offered as a way to circumvent the drawbacks of centralised banking systems and to provide users with transactional data privacy. We do a thorough literature study on the subject of privacy for electronic currencies in this article. We outline the evolution of digital currency from electronic cash to cryptocurrencies and put the spotlight on the methods used to protect user privacy. We also highlight issues with the current cryptocurrency systems that compromise the privacy of users. In our last section, we outline three research areas that will help cryptocurrency users have more privacy: transaction propagation mechanisms, short ZK proof systems without a trusted setup, and specialised trustless zero-knowledge proofs.

Virtual currency bitcoin in the scope of money definition and store of value

frequently covered by the news. Users of untraceable money enjoy it for being decentralised and operating without the ability of governments to control the money supply. The benefits of bitcoin, such as its ability to send money quickly throughout the world, its ability to prevent inflation brought on by governments seeking to address their own issues, and its high level of transaction privacy, are frequently emphasised. The technical details of bitcoin and how this system functions are not the article's primary focus because they have already been extensively covered in other publications. The technical features of bitcoin are only discussed when absolutely essential, with a focus on the economic implications. The article is divided into two parts to accomplish the goal. The first section is devoted to explaining what bitcoin is. It looks at whether bitcoin satisfies the legal, theoretical, and empirical definitions of money. In general, Czech, German, and EU legislation define money compliance; nevertheless, views of the US and Chinese governments are also discussed. The results show that bitcoin cannot simply be regarded as money.

The function of the money storage is the subject of the second section. A key benefit of bitcoin should be that it serves as a better store of value than fiat money. Based on volatility calculations for bitcoin and other currencies and assets, this function analysis. Comparing the findings reveals that bitcoin's volatility (and consequent risk) is far higher than that of other currencies and assets.

Bitcoin is not the new gold—A comparison of volatility, correlation, and portfolio performance

Cryptocurrencies like Bitcoin are becoming well-known as investment vehicles and are frequently referred to as the New Gold. But as this study demonstrates, the two assets couldn't be more unlike from one another. First, we examine and contrast the conditional variance features of Bitcoin, Gold, and other assets to identify structural differences. In order to estimate time-varying conditional correlations, we develop a BEKK-GARCH model. In times of market difficulty, gold is a key component of the flight to quality in the financial markets. Our findings demonstrate that Bitcoin acts in a completely different manner and that it positively correlates with market declines. Finally, we examine the characteristics of Bitcoin as a component of a portfolio and discover no proof of hedging ability. We come to the conclusion that the basic asset characteristics and connections to stock markets of Bitcoin and Gold are very different. The comprehensive cryptocurrency index CRIX upholds our findings. Asymmetric reaction in variance is the only characteristic of Gold that Bitcoin currently reflects.

Macroeconomic variables affecting the volatility of gold price

The greatest gold consumer in the world is examined in this study's analysis of the macroeconomic variables impacting gold prices (India, China, United States, Turkey and Saudi Arabia). The future correlations between gold prices and inflation, real interest rates, currency rates, crude oil prices, and gross domestic product were examined using the Statistical Package for Social Sciences (SPSS). From 1996 through 2015, there were 20 years of annual data use. The results revealed that, in addition to the negative correlations between the inflation rate, GDP, real interest rate, exchange rate, and gold price, there were also positive correlations between crude oil prices and gold prices. The regression's findings demonstrated that factors other than the exchange rate had a considerable influence on the price of gold.

An on-chain analysis-based approach to predict ethereum prices:

A large amount of data is produced by the Ethereum blockchain because of its inherent transparency and decentralised structure. It is also known as on-chain data, and it is freely available to everyone. Additionally, an open ledger is timestamped, contains the on-chain data, and validates it. We can evaluate the functionality and health of the network thanks to this critical blockchain feature. It acts as a sizable data repository for sophisticated prediction algorithms that are capable of accurately identifying systemic trends and projecting future behaviour. By

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

creating an LSTM-RNN (Long Short-Term Memory Recurrent Neural Network) using the metrics most closely related to the price as inputs, we use a quantitative approach employing a subset of these measures to establish the network's genuine monetary worth. Because multiple hyperparameters control how an RNN learns, they are quite sensitive to their settings. Therefore, choosing the best hyperparameters is crucial for efficient and speedy training. The process of choosing an RNN model's ideal parameters is time-consuming and difficult. Because of this, earlier research has created a number of self-adaptive methods to efficiently discover the ideal values for various parameters. However, none of the earlier research examined the use of on-chain data and self-adaptive algorithms in deep learning models to forecast cryptocurrency values. In this research, we provide three self-adaptive methods, each of which converges on a set of optimal parameters for the precise prediction of the price of Ethereum. We contrast our findings with a conventional LSTM model. Our method has an accuracy rate of 86.94% and a low error rate.

3. METHODOLOGY

Machine learning techniques have been extensively employed in recent years to anticipate prices for financial instruments due to their capacity to simulate non-stationarity in time-series data (in contrast to conventional approaches). However, our research has identified two problems with the literature. The first problem arises as a result of the recent rise and fall in bitcoin values. Constructed machine learning-based models are unable to effectively anticipate future prices because the price swings in an unexpected range that has never before been seen. This issue might influence nearly every prediction model built using price data with a reasonable range, not just some prediction algorithms. As a result, the change point detection (CPD) method is suggested as a unique approach in this study to solve the aforementioned issue. In particular, input data are segregated with CPD during training such that each segmented data set has unique statistical properties. Data are normalised individually based on segmentations to accurately portray extreme swings. The tests in this paper have demonstrated that this is a workable solution to the first problem. The second problem that this work tackles for the enhancement of the literature on bitcoin price prediction is that many of the works that have already been published only use cliched factors like historical prices and social media information.

Disadvantages:

1. As a result of the recent rise and fall in bitcoin values.
2. A lot of existing efforts rely solely on clichéd elements like pricing history and social media information.

This study recommends leveraging a wide range of blockchain-related factors to improve price prediction techniques. The most crucial elements for predicting cryptocurrency prices are on-chain data, which is used in the proposed framework as independent variables. Given the volatility and

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

distinctive qualities of cryptocurrencies, precise price prediction is crucial for creating effective trading strategies. In order to do this, the authors of this paper suggest a cutting-edge framework that forecasts the price of Bitcoin (BTC), a well-known cryptocurrency. The change point detection approach is used for steady prediction performance in unobserved price range. Time-series data are segmented in particular so that normalisation may be carried out individually based on segmentation. On-chain data is also gathered and used as an input variable to forecast prices. On-chain data refers to the distinct records recorded on the blockchain that are intrinsic in cryptocurrencies. Additionally, this paper suggests using SAM-LSTM, which combines multiple LSTM modules for on-chain variable groups and the attention mechanism, as the prediction model. SAM-LSTM stands for self-attention-based multiple long short-term memory.

Advantages:

1. How well the suggested framework predicts BTC prices.
2. Strict experiments are used to confirm the efficacy of CPD and SAM-LSTM in BTC price prediction.

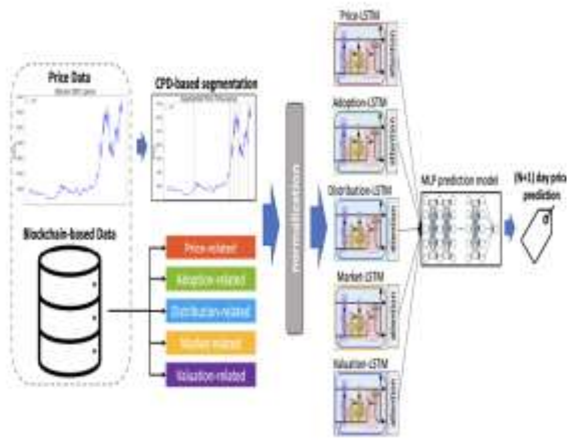

Fig.2: System architecture

MODULES:

The following modules were created to carry out the aforementioned project.

- Data exploration: We will load data into the system using this module.
- Processing: We will read data from the module and process it.
- Data splitting: Using this module, data will be split into train and test.
- Model creation includes the following: LSTM, LSTM + CPD, Attention + CPD + LSTM, Linear Regression, Lasso Regression, Ridge Regression, XGBooster Regression, Voting Regression.
- User registration and login are obtained by using this module.
- Using this module will provide input for prediction, according to user input.

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

- Final expected result showed

4. IMPLEMENTATION

ALGORITHMS:

LSTM:

Long short-term memory networks, or LSTMs, are employed in deep learning. Many recurrent neural networks (RNNs) are able to learn long-term dependencies, particularly in tasks involving sequence prediction.

Linear Regression:

A machine learning algorithm based on supervised learning is linear regression. It executes a regression operation. Regression uses independent variables to model a goal prediction value. It is mostly used to determine how variables and forecasting relate to one another.

Lasso Regression:

Less absolute shrinkage and selection operator, also known as lasso or LASSO, is a regression analysis technique used in statistics and machine learning that performs both variable selection and regularisation in order to improve the predictability and understandability of the resulting statistical model.

Ridge Regression:

Any data that exhibits multicollinearity can be analysed using the model tuning technique known as ridge regression. This technique carries out L2 regularisation. Predicted values differ much from real values when the problem of multicollinearity arises, least-squares are unbiased, and variances are significant.

XGBooster Regression:

The gradient boosted trees approach is widely used and well implemented in open-source software called XGBoost. Gradient boosting is a supervised learning process that combines the predictions of a number of weaker, simpler models to attempt to properly predict a target variable.

Voting Regression:

An ensemble meta-estimator called a voting regressor fits a number of base regressors, one after the other, to the whole dataset. The ultimate forecast is created by averaging the different guesses.

MLP:

A fully connected kind of feedforward artificial neural network is called a multilayer perceptron (MLP) (ANN). The term "MLP" is used ambiguously; sometimes it is used broadly to refer to any feedforward ANN, and other times it is used specifically to describe networks made up of several layers of perceptrons (with threshold activation); see Terminology. When multilayer perceptrons contain just one hidden layer, they are frequently referred to as "vanilla" neural networks.

RNN:

Recurrent neural networks (RNNs) are the most advanced algorithm for sequential data and are the foundation of Google voice search and Apple's Siri. Due to its internal memory, it is the first algorithm to recall its input, making it ideal for machine learning issues involving sequential data.

CNN:

A CNN is a particular type of network design for deep learning algorithms that is utilised for tasks like image recognition and pixel data processing. Although there are different kinds of neural networks in deep learning, CNNs are the preferred network design for identifying and recognising objects.

5. EXPERIMENTAL RESULTS

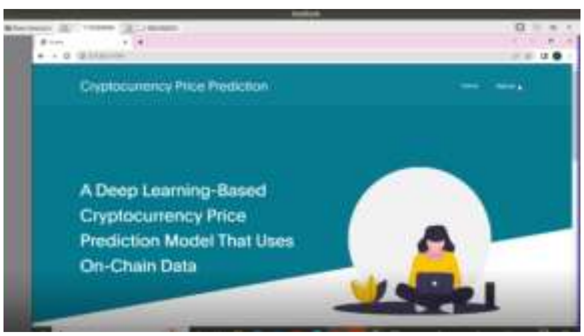

Fig.3: Home screen

Fig.4: User registration

Fig.5: user login

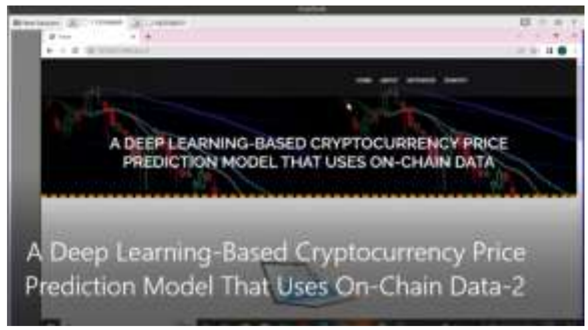

Fig.6: Main screen

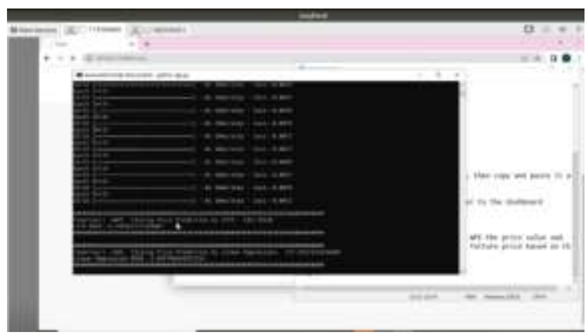

Fig.7: Model generation

Fig.8: Prediction calculations

ISSN: 2278-4632
Vol-13, Issue-05, May 2023

Fig.9: Prediction graph

Fig.10: user input

Fig.11: prediction result

6. CONCLUSION

The authors provide a unique method for forecasting cryptocurrency values using multivariate on-chain time-series data. The suggested technique includes BTC price prediction. A CPD-based normalising approach enables price prediction models to forecast unforeseen price ranges, in contrast to conventional machine learning-based models. As input variables for price prediction, a variety of on-chain variables are chosen, categorised based on their intrinsic qualities, and then utilised. The suggested price prediction model (i.e., SAMLSTM), which is made up of many LSTM modules with independent attention processes and an MLP-based aggregate module, draws out distinguishing characteristics from a set of on-chain data. Five major steps make up this work. Using on-chain data, a thorough variable collection is first carried out. Second, on the basis of CCFs, significant on-chain variables are chosen as input variables and categorised. Third, using a CPD approach known as PELT, time-series data are segmented and normalised throughout each segmentation. Fourth, SAM-LSTM, a hypothesised attention mechanism and multiple LSTM for various on-chain variable groups, is used to forecast prices. Finally, thorough experiments are used to demonstrate the efficiency of CPD and SAM-LSTM in BTC price prediction. The absence of a performance comparison with other bitcoin price prediction techniques is one of the work's limitations. Actually, there are a number of reasons why comparison tests cannot be performed. First off, every piece of literature employs unique input data, whether it be in terms of time frames, data kinds (such as social media data and Google Trends), preprocessing techniques, etc. In example, it is not assured that previous research that use price data from before a recent decline can produce comparable forecast findings. In a related line, a comparison with current research that assert to have impressive performance in price prediction, such as, will be made in the future. One potential area of research for the future is creating a comprehensive framework for predicting bitcoin prices. In particular, a thorough aggregation model should be built in addition to a unified framework that incorporates price-related factors, such as on-chain and social media data, to simulate the price dynamics of the cryptocurrency market. Additionally, it would be beneficial to create a real-time price prediction model that makes forecasts on an hourly or minutely basis using a variety of input data.

REFERENCES

[1] P. Jay, V. Kalariya, P. Parmar, S. Tanwar, N. Kumar, and M. Alazab, ‘‘Stochastic neural networks for cryptocurrency price prediction,’’ IEEE Access, vol. 8, pp. 82804–82818, 2020, doi: 10.1109/ACCESS.2020.2990659.

[2] L. Herskind, P. Katsikouli, and N. Dragoni, ‘‘Privacy and cryptocurrencies—A systematic literature review,’’ IEEE Access, vol. 8, pp. 54044–54059, 2020, doi: 10.1109/ACCESS.2020.2980950.

[3] M. KubÆt, ‘‘Virtual currency bitcoin in the scope of money definition and store of value,’’ Proc. Econ. Finance, vol. 30, pp. 409–416, Jan. 2015, doi: 10.1016/S2212-5671(15)01308-8.

[4] T. Klein, H. P. Thu, and T. Walther, ‘‘Bitcoin is not the new gold—A comparison of volatility, correlation, and portfolio performance,’’ Int. Rev. Financial Anal., vol. 59, pp. 105–116, Oct. 2018, doi: 10.1016/j.irfa.2018.07.010.

[5] S. Hashim, H. Ramlan, N. Razali, and N. Nordin, ‘‘Macroeconomic variables affecting the volatility of gold price,’’ J. Global Bus. Soc. Entrep., vol. 3, no. 5, pp. 97–106, 2017.

[6] N. Jagannath, T. Barbulescu, K. M. Sallam, I. Elgendi, B. Mcgrath, A. Jamalipour, M. Abdel-Basset, and K. Munasinghe, ‘‘An on-chain analysis-based approach to predict ethereum prices,’’ IEEE Access, vol. 9, pp. 167972–167989, 2021, doi: 10.1109/ACCESS.2021.3135620.

[7] S. Nanayakkara, A. Wanniarachchi, and D. Vidanagama, ‘‘Adaptive stock market portfolio management and stock prices prediction platform for Colombo stock exchange of sri Lanka,’’ in Proc. 5th SLAAI Int. Conf. Artif. Intell. (SLAAI-ICAI), Dec. 2021, pp. 1–6, doi: 10.1109/SLAAIICAI54477.2021.9664735.

[8] Preeti, R. Bala, and R. P. Singh, ‘‘Financial and non-stationary time series forecasting using LSTM recurrent neural network for short and long horizon,’’ in Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT), Jul. 2019, pp. 1–7, doi: 10.1109/ICCCNT45670.2019.8944624.

[9] S. Nakamoto. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. [Online]. Available: https://www.debr.io/article/21260.pdf

[10] M. Nofer, P. Gomber, O. Hinz, and D. Schiereck, ‘‘Blockchain,’’ Bus. Inf. Syst. Eng., vol. 59, no. 3, pp. 183–187, Mar. 2017, doi: 10.1007/s12599- 017-0467-3.