BITCOIN PRICE PREDICTION USING MACHINE LEARNING

Aishwarya T,[1] Kushal M G,[2] Likitha D, [3], Prathiksha KR,[4] Madhushree YM
Dept. of Information Science and Engineering
Vidyavardhaka College of Engineering
Mysore
aishwarya.t@vvce.ac.in
kushalmg11233@gmail.com
likithapatel1741999@gmail.com
prathikshakr2002@gmail.com
madhushreemadhu179@gmail.com

Abstract — **By enabling on-demand access to computer resources and services through the internet, Bitcoin has become more widely adopted, there has been a growing interest in predicting its price movements. Machine learning techniques have emerged as a popular approach to Bitcoin price prediction, as they can leverage the large amounts of historical data available on Bitcoin transactions and market activity. In this survey, we review the state of the art in using machine learning for Bitcoin price prediction. We discuss the different types of machine learning models that have been applied to this task, including regression models, time series models, and deep learning models. We also examine the various features and data sources that have been used as inputs to these models, such as transaction volume, sentiment analysis of social media posts, and technical indicators.**

Keywords — Blockchain, Decentralized, Smart Contracts, Data Privacy, Security, Scalability, Permissioned Blockchain

1. Introduction

By enabling online, on-demand access to computer resources and services, cloud computing has completely changed the IT sector. It is a desirable alternative for both enterprises and people because of a number of advantages it provides, including cost savings, scalability, and flexibility. However, the centralised architecture of cloud systems poses significant issues with regard to data security and privacy. The complex security concerns inthe cloud environment cannot be fully addressed by conventional security techniques like firewalls and encryption.

Bitcoin is a digital or virtual currency that was created in 2009 by an unknown person or group using the name Satoshi Nakamoto. It operates on a decentralized system, which means that it is not controlled by any governmentor financial institution. One of the main features of Bitcoin is its blockchain technology, which allows for secure and transparent transactions. Each transaction is verified by a network of nodes on the blockchain, andonce verified, it is added to a permanent and unalterable record of all Bitcoin transactions.

Bitcoin can be acquired through a process called mining, where individuals or companies use powerful computers to solve complex mathematical problems and receive new bitcoins as a reward. Alternatively, people can purchase bitcoins on cryptocurrency exchanges using traditional fiat currency or other cryptocurrencies.

Bitcoin has gained popularity as a form of investment and payment due to its limited supply, its ability to bypass traditional financial institutions, and its potential for high returns. However, it is also a highly volatile asset, with prices fluctuating rapidly and dramatically.

A distributed ledger made possible by blockchain technology can securely record transactions and shield them from manipulation. It provides a decentralised, open, and unchangeable structure that can improve the security, privacy, and effectiveness of cloud computing services. Distributed consensus, smart contracts, and cryptographic algorithms are some of the salient characteristics of blockchain technology. Distributed consensus makes ensuring that all users of the network concur that a transaction is legitimate. Smart contracts streamline transaction negotiation and execution by acting as self-executing contracts. Data security and privacy are ensured by cryptographic techniques through data encryption and safe authentication.

This survey article overviews current work on a Bitcoin-based integrated architecture for preserving security in cloud computing. We examine the many elements of the integrated architecture, such as blockchain, cloud computing, and security protocols. We examine the various methods employed to incorporate permissioned and permissionless blockchains, smart contracts, and consensus mechanisms into cloud computing.

Also covered are the benefits and difficulties of incorporating Bitcoin technology into cloud computing, including interoperability, scalability, and data privacy. Based on different factors including security, performance, and cost-effectiveness, we give a comparative review of the current methods for integrating blockchain price prediction technology.

2. Background and related work

S Velankar et al. proposed a system that suggests a taxonomy for blockchains and systems built on them. The taxonomy may be used to compare blockchains and to help with the creation and assessment of blockchain-based software systems. Our taxonomy covers the key blockchain architectural traits as well as the effects of various choices. This taxonomy is meant to aid in crucial architectural deliberations on the performance and quality aspects (such as availability, security, and performance) of blockchain-based systems. Patterns are another method for categorizing and organizing the current solutions in addition to taxonomy. A new technology called blockchain enables the decentralized, transactional exchange of data across vast networks of untrusted users. This enables new distributed software architectures that enable shared state agreement without relying on a central integration point. Decentralization transfers power and authority from a central location. Bitcoin is a digital currency and its underlying technology is called the blockchain. Many banks are participating in blockchain technology experiments. The first use case explored is related to financial transactions, but there are others. Database collecting is covered in the paper's first section. We obtained the figures for Bitcoin from CoinmarketCap and Quandl, two distinct sources. It is necessary to normalise and smooth out the time-series data once it has been collected and has been recorded every day for five years at various times. We have used a variety of normalisation algorithms for this, including log transformation, z-score normalisation, box cox normalisation, and others. Data is then smoothed across the whole time period [1].

S McNally et.al suggested a methodology that provides an empirical analysis of a spam-based "stress test" DoS attack against Bitcoin. Our cluster-based methodology shows that 385,256 (23.41%) of a total of 1,645,667 Bitcoin transactions were spam during his 10 days at the peak of the spam campaign. This attack increased the average cost by 51% (from 45 satoshis/byte to 68 satoshis/byte) and increased the processing latency of non-spam transactions by a factor of7 (from 0.33 hours to 2.67 hours). It warns that changes toBitcoin bonds may reduce some of the spam patterns seen so far. Our results point to the need for additional research into Bitcoin transaction spam filtering methods and other Bitcoin DoS mitigation solutions. This shows that a malicious party can perform a Bitcoin DoS if they are willing to distribute a small amount of Bitcoin (minimum $49,000). His subsequent DoS attacks on Bitcoin used other techniques such as “money drops” and transaction malleability, resulting in poor performance. Unlike other studies looking for patterns in Bitcoin, our clustering method removes identifying data and groups transaction attributes to identify patterns rather than aggregate transactions and anonymize individuals. To do. This studyaims to determine how accurately we can predict the direction of Bitcoin price in US dollars. The Bitcoin Price Index is a source of price information. Bayesian optimization RNN and LSTM network implementations allow different levels of achievement. LSTM performs best with 52% classification accuracy and 8% RMSE. The well-known ARIMA model is used as opposed to the deep learning model. As expected, the nonlinear deep learning method outperforms the poorer ARIMA prediction. Finally, the training times of both deep learning models are evaluated on GPU and CPU [2].

S Yogeshwaran et al. developed a system that provides a homomorphic encryption scheme based on elliptic curve cryptography (HES-ECC) is proposed in this work for secure data transmission and storage. This plan encrypts your data before storing it in the cloud. When you perform operations such as addition and multiplication on encrypted data stored in the cloud, the result is returned to the original data without being decrypted beforehand. This allows the cloud server to access only encrypted data, perform the necessary calculations, and perform the actions required by the user. This ensures the security of data storage and transmission. A modified Weil pairing is used in the proposed HES-ECC public key construction to provide additional homomorphic properties and encryption. Bilinear pairing is also used in HES-ECC for multiplicative homomorphic properties. The HES-ECC that has been presented is a system that merely makes use of the algebraic structure ofelliptic curves and pairings. Apart from those, no further computations, such as xor operations, hash functions, secure key distributors, trustworthy third parties, etc., are required. There is no way to see open messages in transit or in the public cloud. Safe communication is offered since plaintext is not utilised at any point throughout the process. The complexity of the ECDLP and the WDHP is the basis for the security of the encryption approach that they suggest to WDHP. As ensuring the security of cloud storage is our main objective, our encryption strategy makes use of homomorphic encryption techniques. The study also uses modified Weil pairing and bilinear pairing for the homomorphic property. Hence, ECDPL, WDHP, and the BDHP are the foundations for the security of homomorphic property BDHP. This creative undergraduate project aims to demonstrate how a trained machine model can forecast a cryptocurrency's price given enough data and processing capacity. It shows a graph of the anticipated values. The technology that might aid humankind in predicting future occurrences is the most well-liked. We are finally on the verge of a time when forecasts may be made with reasonable accuracy and be based on verifiable facts thanks to the massive amounts ofdata that are created and collected every day. More peoplehave also resorted to the digital market for investments as a result of the emergence of the crypto digital era. This is a chance for us to develop a model that can forecast cryptocurrencies [3].

PV Rane et al. proposed a system that emphasises the OP RETURN special programming language instruction allows the Bitcoin protocol to save any data on the blockchain. A rising number of protocols take advantage of this functionality to expand the scope of uses for the Bitcoin network beyond money transfers. This essay is an empirical examination of OP RETURN's historical use. Based on OP RETURN, we distinguish a number of protocols, which we then categorise according to their application domain. We track the changes in utilisation over time, the distribution of OP RETURN transactions across application domains, and their space usage. Finally, they calculate the amount of OP RETURN information and the ratio of OP RETURN transaction size to the total blockchain transaction size. To the best of our knowledge, it is our duty to most thoroughly test OP RETURN usage. A tool that we have created serves as the basis for all of our analysis [4].

Felizardo et al. developed a technique that emphasises symmetric algorithms with the goal of determining which one should be used for cloud-based applications and services that need link data and encryption. The paper provides a brief comparative analysis and overview of cryptographic algorithms, focusing on the symmetric approach that should be used for cloud-based applications and services that require link data and encryption. The capacity to protects data from attacks, speed and efficiency with which it does so are the two key characteristics that set one encryption method apart from another. Equally effective at safeguarding the transferred data across any communication medium are symmetric and asymmetric key methods. They weighed the advantages and disadvantages of the suggested and standard algorithms in relation to symmetric and asymmetric key cryptography. They have also examined the importance of these two cryptography methods. It has been determined that Blowfish, AES, RSA, and DES are the finest security algorithms to use in cloud computing so that data is safe and not vulnerable to hackers. The issue of data disclosure is reduced by a method that was proposed, in which encryption is used to offer security while data is being transmitted. This system employed the idea of the RSA algorithm, Hash function, and only encoded data was communicated across the channel [5].

Zhengyang et al. suggests a modular residue-based verification technique to validate homomorphic cryptographic computations. over the whole finite body. Depending on the underlying cryptosystems used, the performance of the proposed method has varied. However, according to the cryptosystems evaluated, this technique has a storage cost of 1.5% and a configurable computation cost of 1%. Such a cost is a reasonable compromise for cloud verification, which is very important for cloud computing. The modular residuals of the extrinsic computation are evaluated in this research in order to propose an effective DIV method for HE on Zp. The number of modules that may be employed is only constrained by the client's word count in the flexible and extendable architecture of the proposed approach. It is technically conceivable to have 264 modules on a 64-bit computer. The storage need on the client's computer is thus, based on 64-bit computers, less than 1.5% of thedata size saved on the CSP. If a check is applied to the homomorphic math, the worst computation cost incurred by the client on any of the tested cryptosystems is less than 3% of the actual homomorphic computation incurred by the CSP. Predictions are still difficult since cryptocurrencies are so complicated. In this study, we establish two widely used machine learning models, the fully-connected Artificial Neural Network (ANN) and the Long-Short-Term Memory (LSTM), to accurately predict the prices of a number of well-known cryptocurrencies, including Bitcoin (BTC), Ethereum (ETH), Ripple (XRP),Stellar Lumens (XLM), Litecoin (LTC), and Monero (XMR). To better comprehend the behaviours of our models, we assess model performance and carry out sensitivity analysis. Although LSTM appears to be better suitable for the purpose of time sequence prediction, we discover that ANN generally outperforms LSTM in our testing. The prediction of BTC might be greatly aided by the use of price data from other cryptocurrencies for combined training and prediction. Finally, the time scale of interest has a significant impact on the model's prediction inaccuracy [6].

T Phaladisailoed et al. proposed a study, the status of the UTXO sets is far from perfect right now, according to a review of the number of unspent coins in three of the most well-known cryptocurrencies in the world. A subset of transaction outputs that have not yet been spent is known as an unspent transaction output (UTXO) set. According to a fee-per-byte rate f, the Bitcoin Core client treats a certain UTXO out as dust. According to their definition, an unprofitable output is one that, after just accounting for the amount of the input that would be required, has less value than the charge that must be paid. They provide two indicators for determining if output is worthwhile investing in. They provide two indicators to determine if a manufacturing operation is worth the investment. Following these measurements, Pérez-Solà and his colleagues analyze three sets of UTXOs of the three currencies listed above. More than 50% of the UTXOs in the collection can be considered dust for low charge-per-byte values of 116 satoshi/byte or more. The researchers argue that the discovery represents a first step in solving the problem of UTXO not being worth it. The researchers plan to create both tactics to encourage dust accumulation and discourage UTXO production [7].

S Karasu et al. proposed a secure design to address the privacy issue of photos stored on his cloud servers. Create a private/public cloud infrastructure using an open-source project called OpenStack. They propose a hybrid architecture consisting of two clouds. The first is a private cloud used only for encryption and decryption and the second is a public cloud used for storing encrypted images. Mobile cloud computing has emerged as a new technology that enhances the capabilities of mobile computing. MCC provides wireless customers with enhanced access to storage and reliability. In mobile cloudcomputing settings, an architecture is suggested in this research for safeguarding offloaded pictures. This architecture is built on the idea of encryption as a service, where encryption and decryption are handled exclusively by a private cloud. To demonstrate the scheme's functioning, they constructed the water treatment technique (DWT) and the Pailier cryptosystem, an additively homomorphic encryption scheme. As the encryption and decryption processes were quick, they discovered that they might be applied as a data security/privacy solution in such a setting [8].

M Rizwan et al. developed a system where a function may be performed by a long-term memory (LSTM) network and a Bayesian recurrent hierarchical (RNH) neural network. The LSTM achieves an overall identification accuracy of 52% and an RMSE of 8%. The popular ARIMA approach for the prediction of time series is in contrast to the thorough training methods. This model is not nearly as effective as deep learning models. It was anticipated that deep learning techniques will perform better than the subpar ARIMA forecast. To anticipate the price of Bitcoin, we have employed the Gated Recurrent Network model (GRU). All deep learning models eventually have a GPU and CPU that outperform the GPU by 94.70 per cent in terms of training time. Our research's goal is to create a deep-learning model that can forecast Bitcoin's price. Since deep learning is employed in the selection of the parameter to obtain consecutive results in model development. We observed that the total parameter and dataset can affect the outcome in the latter when we applied the three recommended models, RNN, LSTM, and GRU. The previous model, which was created using RNN and LSTM, had an approximately 52% lower predicted accuracy. However, in our comparison investigation, the GRU model performed better than the LSTM model. The accuracy of the GRU results using the best model is 94.70%. The suggested model exhibits an improvement in the accuracy of roughly 42.3%. All of our GRU tests produce extremely thorough results that require time [9].

C Eom et al. proposed a system that objectively examines the statistical properties and predictions of Bitcoin returns and volatility. Bitcoin gains and volatility are distributed with a large right tail and a high core percentage. Unlike stylized data in financial time series, Bitcoin does not exhibit dynamic attributes of volatility persistence. Additionally, autoregressive models that use historical volatility perform poorly when used to predict changes in Bitcoin volatility over time. Variations in Bitcoin volatility in upcoming timeframes can be explained by changes in investor attitudes towards cryptocurrencies. These results suggest that Bitcoin appears to be an asset, rather than a monetary asset, that is subject to considerable volatility and investor opinion. We examined the distribution and dynamic characteristics of Bitcoin returns and volatility from October 2017 to May 2017. A basic autoregressive model was used to create an empirical design in which the outcome does not depend on the dominance of the predictive model. In particular, we investigated whether past lag data can be used to explain future data changes. We also examine the information value of investor attitudes towards Bitcoin in terms of data on Bitcoin as a fixed asset and new variables to improve Bitcoin predictions. We used historical data for Bitcoin and Euro exchange rates against the United States [10].

Y Li et al. proposed a system that a hybrid neural network model based on a convolutional neural network (CNN) and long short-term memory (LSTM) neural network is presented to address the issue that the price of Bitcoin changes significantly and is difficult to anticipate. The inputs used include the Bitcoin transaction data as well as externaldata like macroeconomic factors and investor interest. First, feature extraction is performed using CNN. The feature vectors are then fed into the LSTM to train it and make short-term Bitcoin price predictions. The outcome demonstrates that when compared to a single-structure neural network, the CNN-LSTM hybrid neural network may significantly increase the accuracy of value prediction and direction prediction. The discovery has significant ramifications for researchers and digital currency investors. In contrast to conventional research, this article thoroughly employs a variety of elements that may affect Bitcoin values by taking into account both internal factors, such as Bitcoin's own transaction data, as well as external factors, such as macroeconomic indicators and investor interest. In terms of value prediction and direction prediction, this hybrid neural network is contrasted with the single-structure neural network. The findings indicate that the hybrid CNN-LSTM neural network works well and is better suited for Bitcoin prediction [11].

JZ Huang et al developed a system that a investigates whether or not investors can forecast Bitcoin profits. There has been significant discussion regarding the predictability of returns on financial assets, such as stock returns or commodity returns, but relatively few researchers have looked into the predictability of bitcoin returns. In the analysis, we concentrate on bitcoins, the most well-known is cryptocurrency. Our goal is to forecast the 21 intervals in the daily Bitcoin return domain the next day's return will fall. Using 124 technical indicators that are solely dependent on past Bitcoin prices, we create a return predictor/model [12].

Rathan et al. developed a technique that deals with identifying the price trend on daily variations in the price of bitcoin while providing information on price trends. The dataset up to the present contains information on the open, high, low, and close prices of Bitcoin. Machine learning is utilised to anticipateprice values by utilising the dataset. The purpose of this research is to determine and evaluate the Bitcoin forecast accuracy of several machine-learning algorithms. The decision tree and regression models' test results are contrasted. Numerous research in the economics and priceprediction fields have been conducted, and the market for Bitcoin is growing. Our suggested study analyses the Bitcoin dataset from 2011 to the present using machine learning models like Decision Trees and Linear Regression. To forecast prices for the following five days, decision trees and linear regression models are also employed. The best algorithm for the bitcoin prediction challenge should be chosen and used, according to the specified learning strategy. The results of the experimental study show that linear regression provides high accuracy price prediction more effectively than the alternative [13].

S Ji et al. proposed a system in which several cutting-edge deep learning methods for forecasting Bitcoin values are examined and compared in this study including deep neural networks (DNN), long short-term memory (LSTM) models, convolutional neural networks, deep residual networks, and their combinations. The results of the experiments showed that while LSTM-based models performed marginally better for price regression (regression), DNN-based models greatly outperformed LSTM-based models for price classification (ups and downs). A simple analysis of profitability also showed that classification models outperformed regression models for algorithmic trading. The findings of the proposed deep learning-based prediction models were comparable in general. Issues with regression and classification, where the former predicts the price of bitcoin in the future and the latter predicts whether the price will increase or decrease, were both addressed. In contrast to past studies on the prediction of Bitcoin prices, DNN somewhat outperformed the competition for classification problems while LSTM slightly exceeded it for regression problems. CNN and ResNet are known for being extremely effectivein many applications, such as analysing sequence data, however their performance in forecasting the price of Bitcoin was not particularly spectacular. Overall, there was no clear winner, and the performance of each of the deep learning models was comparable [14].

Z Chen et al. developed a system that we investigated a machine learning method to predict Bitcoin price based on sample properties and dimensions. We show that sample granularity and feature dimension should be considered, in contrast to most other efforts that simply use machine learning techniques to predict Bitcoin prices. The inclusion of higher dimensional aspects such as ownership and networks, trades and markets, attention, and the current price of gold is made possible by the daily average price of Bitcoin supplied by Coin Market Cap. Binance's exchange feature helps facilitate Bitcoin trading prices at 5-minute intervals. He used statistical methods for Bitcoin's daily price predictions and machine learning models for his 5-minute interval price predictions for Bitcoin. It is based on the Occam's Razor concept anduses a paradigm applied to real-world prediction problems using machine learning algorithms. The results show that machine learning models outperform statistical approaches on high-frequency data, while statistical methods outperform statistical methods on low-frequency data with high-dimensional features. is showing. Most of our results outperform benchmark results of comparable machine learning algorithms. We believe more domains with Bitcoin-like properties can benefit from his sample sizing method, which uses machine learning models for prediction [15].

RS Gadey et al. proposed the most well-known and popular machine learning library is described by the authors. In order to develop machine learning methods, this tool is quite helpful. The writers highlight the benefits of the Sci-kit Learn and describe how easy and effective the library is. Italso discusses how the library was integrated into the Python ecosystem and outlines the difficulties that developers encountered while utilising this tool. Keras and TensorFlow are the additional libraries required for the machine learning algorithm's implementation. The results contributed to the expansion of the Bitcoin sector with the aid of straightforward design and algorithms. As the market capitalisation increased in 2017, cryptocurrencies gained popularity. In the current environment, more than 1500 cryptocurrencies are being traded. You may create and utilise cryptocurrency for online transactions. Cryptocurrency technology is Bitcoin. Every second, the value of Bitcoin changes. As a result, we employ the LSTM Architecture to forecast the value ofBitcoin here. We are attempting to demonstrate with the aid of this design that the LSTM architecture offers outcomes that are more accurate than those produced by any other machine learning algorithms or architecture [16].

SM Raju et al. developed a system which utilizes a variety of deep learning models such as Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU), a comparative investigations of the many aspects impacting bitcoin price prediction is carried out based on Root Mean Square Error (RMSE). We have investigated how the price of gold affects the cost of Bitcoin. This study compares several deep learning models based on RMSE values for a variety of characteristics influencing bitcoin price prediction. The findings demonstrate how well the various deep-learning models predict the price of Bitcoin. When bitcoin price prediction factors are taken into account, LSTM yields the lowest RMSE value. When used as the parameter for predicting bitcoin prices, LSTM does not, however, show a strong correlation between the price of gold and that of bitcoin. As a result, there may be a sizable difference between the actual and predicted numbers on the graph. According to the popularity of the tweeter, which is measured by the number of followers the tweeter has on Twitter, it is possible to predict that when a tweet is passed that is positive about bitcoin, the prices are expected to rise and when a tweet is passed that is negative about bitcoin, the prices do get affected and are expected to falldown. The influence of Twitter on the price of Bitcoin may be evaluated in further research [17].

Aggarwal et al. developed a technology to accurately predict the price of Bitcoin while taking into consideration a variety of variables that affect its value in this study. In the initial phase of our investigation, we want to grasp and identify everyday market patterns for Bitcoin while learning the ideal conditions pertaining to its pricing. Numerous elements ofthe Bitcoin pricing and payment network during a five-year period that were daily documented are included in our data collection. We will generate the most precise forecast of the daily price change sign for the second phase of our investigation using the information at our disposal. The first portion of the study discusses database collection. The data for Bitcoin came from two different sources, CoinmarketCap and Quandl. It is necessary to normalise and smooth out the time-series data once it has been collected and has been recorded every day for five years at various times. We have used a variety of normalisation algorithms for this, including log transformation, z-score normalisation, box cox normalisation, and others. Data is then smoothed across the whole time period [18].

Z Chen et al. proposed a system with an accurate estimationof the Bitcoin price while taking into account a number of variables that influence the Bitcoin value.In our work, we aimed to comprehend and recognise everyday fluctuationsin the Bitcoin market while gaining knowledge of the most pertinent aspects affecting the price of Bitcoin. We'llforecast daily price changes as accurately as we can. Currently, there are more than $230 billion worth of openly traded cryptocurrencies on the market. The main purpose of Bitcoin, the most valuable cryptocurrency, which also has the best price predictability, is to act as a digital store of wealth. The basic features of Bitcoin, as they are discussed in detail in the relevant publications, are summarised in the subsection that follows. In order to forecast the future price of Bitcoin, we took into account prior Bitcoin transactions where the price and timestamp were the determining factors. For price forecasting, we employed four different techniques: logistic regression, support vector machines, RNN, and ARIMA [19].

S McNally et al. developed a system in which GRU has the best accuracy but requires more computation time than Huber regression. The findings, however, might be impacted by parameter settings and the overall number of datasets. Additionally, the chosen features of Open, Close, High, and Low may not be sufficient to forecast Bitcoin prices because a variety of factors, including social media reactions and the policies and laws that each nation announces to regulate digital currency, can influence the rise and fall of Bitcoin prices. As a result, the objective of this study is to examine several machine learning algorithms in order to identify the model that can estimate Bitcoin prices most effectively and accurately. The scikit-learn and Keras libraries were used to evaluate a range of regression models using 1-minute interval trade data from the Bitstamp Bitcoin exchange website from January 1, 2012 to January 8, 2018. The best results showed that the R-Square (R^2) was up to 99.2% and the Mean Squared Error (MSE) was down to 0.00002 [20].

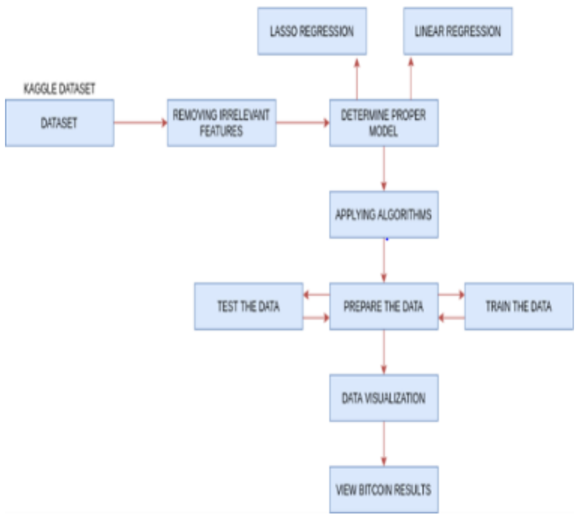

Fig 2.1 Methodology

A kind of artificial neural networks called recurrent neural networks (RNNs) is created to model sequential data. They are extensively utilised in time series analysis, natural language processing, voice recognition, and other applications where the input data is presented as a sequence or time series. The essential feature of RNNs is the recurrent link between the hidden layers, which enables them to keep track of prior inputs. Because of this, they are especially well suited for representing sequences of varying duration in which the current input is dependent on the sequence's past inputs.

Backpropagation through time (BPTT), a variant of the backpropagation method that takes into consideration the temporal aspect of the data, is often used to train RNNs. The weights of the network are modified during training in order to reduce the error between the expected output and the actual output. The Long Short-Term Memory (LSTM) network is a popular form of RNN that was created to overcome the vanishing gradient issue that can arise in conventional RNNs. It might be challenging to identify long-term relationships in the data because of the vanishing gradient problem, which happens when the gradients used to update the weights of the network are very tiny. To solve this issue, LSTMs introduce a gating mechanism that enables the network to selectively recall or forget certain information.

RNNs are superior to conventional machine learning methods for sequential data in a number of ways, including their capacity to handle variable-length sequences and simulate complicated temporal connections. However, they may also be difficult to train computationally and be susceptible to problems like overfitting and disappearing gradients. RNNs are an effective modelling technique for sequential data and have been applied in a variety of fields, including time series analysis, speech recognition, and natural language processing. Future RNN applicationsare expected to be much more creative as long as research in this field is conducted.

A form of recurrent neural network (RNN) called Long Short-Term Memory (LSTM) was created to solve the issue of disappearing gradients in conventional RNNs. In applications where the input data is in a sequence or time-series format, such as natural language processing, audio recognition, time series analysis, etc.,

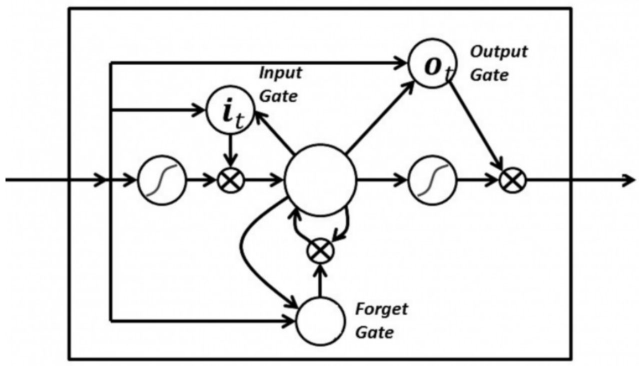

Fig 2.2 RNN

LSTMs are frequently utilised. The inclusion of memory cells, which enable the network to selectively recall or forget data fromearlier time steps, is the crucial component of LSTMs. Three gates—the input gate, the forget gate, and the output gate—control the memory cells. The memory cell's input gate selects the data from the current time step that should be stored there, and the forget gate selects the data that should be erased from the memory cell.

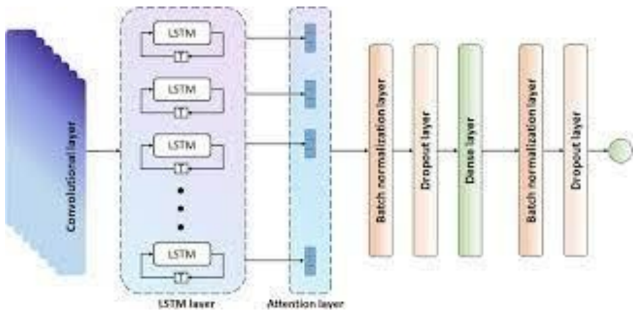

Fig 2.3 LSTM Architecture

Some common methods used to predict Bitcoin prices include:

1. Technical analysis: This involves analyzing past price patterns and market trends to identify potential future price movements. Technical analysts use charts and mathematical indicators to identify support and resistance levels, price targets, and other key patterns.
2. Fundamental analysis: This involves analyzing the underlying factors that drive Bitcoin's value, such as adoption rates, network activity, and regulatory developments. Fundamental analysts look at macroeconomic trends and events that can impact Bitcoin's price, such as changes in interest rates or geopolitical tensions.
3. Sentiment analysis: This involves analyzing the overall sentiment of market participants towards Bitcoin. Sentiment analysis can be done using various tools, such as social media monitoring, news sentiment analysis, and surveys of market participants.
4. Machine learning algorithms: This involves using complex algorithms to analyze large amounts of data and identify patterns that can be used to predict future price movements.

3. DISCUSSIONS

Author Name	Title Name	Advantages	Disadvantages	Methodology
Aggarwal, Apoorva, Isha Gupta, Novesh Garg, and Anurag Goel. [1]	Deep Learning Approach to Determine the Impact of Socio Economic Factors on Bitcoin Price Prediction.	This is helpful sincenot all of the variablesthat might affect Bitcoin's price are equally significant. The use of LASSO can assist to increase prediction accuracy by highlighting the key traits. This can involve using autoregressive integrated moving average (ARIMA) models, seasonal decomposition of time series (STL), or other models specifically designed for time series forecasting.	The linear nature of the LASSO model necessitates the assumption of a linear connection between the characteristics and the target variable. This could not always be the case, as there might be intricate relationships between the characteristics that a linear model is unable to account for.	Key characteristics that are most effective at predicting Bitcoin price using LASSO. This entails choosing the features with non-zero coefficients and fitting a linear model with a LASSO penalty. Root Mean Square Error (RMSE) values of all the three models namely LSTM, CNN and GRU are 201.34, 151.67 and 179.23.
Akyildirim, Erdinc, Oguzhan Cepni, Shaen Corbet, and Gazi Salah Uddin. [2]	Forecasting mid-price movement of Bitcoin futures using machine learning.	Because they can be depicted visually and display the decision-making process, they are simple to understand and intuitive. As aresult, it is simpler for traders and investors to comprehend themodel and thevariables that affect the price of Bitcoin.	Overfitting, which happens when a model is overly complicated and catches noise in the data instead of underlying patterns. In order to avoid overfitting, procedures like pruning or regularisation may be necessary.	An empirical analysis of a recent spam campaign that resulted in a denial-of-service assault on Bitcoin in this paper. Our investigation's goal is to comprehend spammers' tactics and how they affect Bitcoin users. The random forest algorithm provides the highest in-sample success rate that can reach up to 87% for the 0.3 hold-out sample
Awoke, Temesgen, Minakhi Rout, Lipika Mohanty, and Suresh Chandra Satapathy. [3]	Bitcoin Price Prediction and Analysis Using Deep Learning Models.	The kNN method is straightforward andsimple to use, making it a suitable option for people who are new tomachine learning or those looking for a quick fix forpredicting the price of Bitcoin. Uses proximity to make classifications or predictions about the grouping of an individual data point.	When there are many features and a huge dataset, the kNN approach might be computationally costly. Due to this, scaling the process to very big datasets may be challenging. To increase performance, specialised hardware or optimisation techniques may be needed.	Identify the key characteristics that most accurately forecast the price of Bitcoin. This mightentail using feature selection methods like recursive feature removal, mutual information analysis, or correlation analysis. The RMSE value of the LSTM is 0.067.
Basak, Suryoday, Saibal Kar, Snehanshu Saha, Luckyson Khaidem, and Sudeepa Roy Dey [4]	Predicting the direction of stock market prices using tree-based classifiers.	The significance and direction of the link between the independent factors and the dependent variable are represented by the coefficients of the linear regression model. This makes it simple to evaluate the findings and derive insightful conclusions.	The dependent variable and independent variables are thought to have a linear relationship according to the linear regression model.	This might entail using feature selection methods like recursive feature removal, mutual information analysis, or correlation analysis. Create training and test sets from the preprocessed data, allocating the bulk of the data to the training set and the remainder to the test set. The accuracy of the proposed system is 96%.

Baur, Dirk G., and Thomas Dimpfl. [5]	The volatility of Bitcoin and its role as a medium of exchange and a store of value.	Without the need for manually designing features, CNNs may automatically extract pertinent characteristics from the incoming data. This might help you see patterns and trends in the data on Bitcoin price changes.	The computational cost of CNNs may be high, especially when working with huge datasets. When working with high-frequency Bitcoin price data, this might be a drawback.	To automatically extract pertinent characteristics from the input data, use a CNN. Setting hyperparameters like the learning rate and creating an acceptable architecture for the CNN are possible steps in this process. 1,000 replications of a Monte Carlo simulation's 95% critical value results and SADF test statistics were also compared to identify bubbles periods.
Blake, R [6]	An Econometric Analysis of the Relationship between Bitcoin & Gold.	Compared to other, more sophisticated machine learning algorithms, Nave Bayes has a quick training period, which is advantageous when working with big datasets. Generative learning algorithms, meaning that it seeks to model the distribution of inputs of a given class or category.	Despite the fact that the scheme resolves the issue of ensuring the accuracy of the data calculations, the client may be required to supply storage space and additional labour to complete the verification step. Hence, in future work, we intend to move the verification process to distributed fog nodes that converse via consensus.	In order to enforce data confidentiality, privacy, and integrity during an outsourced computation, this study suggests a verification technique based on the modular residue to authenticate homomorphic encryption computation over integer finite field. The analysis of this paper includes gold and bitcoin data from the 8-year time period between 2014 and 2022 is presented.
Carbó, José Manuel, and Sergio Gorjón. [7]	Application of Machine Learning Models and Interpretability Techniques to Identify the Determinants of Price of Bitcoin.	Because Random Forests can handle ahuge number of input characteristics, they are appropriate for challenging issues like predicting the price of Bitcoin. Combines the output of multiple decision trees to reach a single result.	Because Random Forests are regarded as "black box" models, it might be challenging to understand how the model made its predictions. It may be difficult to determine which input properties are responsible for the model's predictions as a result.	By changing hyperparameters like the number of trees, the depth of each tree, or the bare minimum of samples needed to divide a node, the Random Forest model may be improved. The model replicates reasonably well the behaviour of the price of bitcoin over different periods of time with an accuracy of 92.8%.

Table 1 Relevant studies on advantages and disadvantages of various bitcoin price prediction

4. Conclusion

Bitcoin price predictions are subject to a high level of uncertainty and volatility due to the cryptocurrency market's unpredictable nature. Some analysts predict that Bitcoin's price could reach as high as $100,000 or even $1 million in the future, while others are more cautious and predict that the price will continue to fluctuate and may even decrease in the short term. Bitcoin price prediction using machine learning has been a popular area of research in recent years. Several machine learning techniques, including artificial neural networks, support vector machines, and decision trees, have been used to predict the price of bitcoin. While some studies have reported promising results, others have shown that the prediction accuracy varies widely depending on the data and the model used.

There are several factors that may influence the price of Bitcoin, including adoption rates, regulatory developments, technological advancements, and market sentiment. It's important to note that Bitcoin's price has historically been very volatile, with significant price swings occurring within a short period of time.

References

[1] Velankar, Siddhi, Sakshi Valecha, and Shreya Maji. "Bitcoin price prediction using machine learning." In 2018 20th International Conference on Advanced Communication Technology (ICACT), pp. 144-147. IEEE, 2018.

[2] McNally, Sean, Jason Roche, and Simon Caton. "Predicting the price of bitcoin using machine learning." In 2018 26th Euromicro international conference on parallel, distributed and network-based processing (PDP), pp. 339-343. IEEE, 2018.

[3] Yogeshwaran, S., Maninder Jeet Kaur, and Piyush Maheshwari. "Project-based learning: predicting bitcoin prices using deep learning." In 2019 IEEE Global Engineering Education Conference (EDUCON), pp. 1449-1454. IEEE, 2019.

[4] Rane, Prachi Vivek, and Sudhir N. Dhage. "Systematic erudition of bitcoin price prediction using machinelearning techniques." In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 594-598. IEEE, 2019.

[5] Felizardo, Leonardo, Roberth Oliveira, Emilio Del-Moral-Hernandez, and Fabio Cozman. "Comparative study of bitcoin price prediction using wavelets, recurrent neural networks and other machine learning methods." In 2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESC), pp. 1-6. IEEE, 2019.

[6] Zhengyang, Wang, Li Xingzhou, Ruan Jinjin, and Kou Jiaqing. "Prediction of cryptocurrency price dynamics with multiple machine learning techniques." In Proceedings of the 2019 4th International Conference on Machine Learning Technologies, pp. 15-19. 2019.

[7] Phaladisailoed, Thearasak, and Thanisa Numnonda. "Machine learning models comparison for bitcoin price prediction." In 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), pp. 506-511. IEEE, 2018.

[8] Karasu, Seçkin, Aytaç Altan, Zehra Saraç, and Rifat Hacioğlu. "Prediction of Bitcoin prices with machine learning methods using time series data." In 2018 26th signal processing and communications applications conference (SIU), pp. 1-4. IEEE, 2018.

[9] Rizwan, Muhammad, Sanam Narejo, and Moazzam Javed. "Bitcoin price prediction using deep learning algorithm." In 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), pp. 1-7. IEEE, 2019.

[10] Eom, Cheoljun, Taisei Kaizoji, Sang Hoon Kang, and Lukas Pichl. "Bitcoin and investor sentiment: statistical characteristics and predictability." Physica A: Statistical Mechanics and its Applications 514 (2019): 511-521.

[11] Li, Yan, and Wei Dai. "Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model." The Journal of Engineering 2020, no. 13 (2020): 344-347.

[12] Huang, Jing-Zhi, William Huang, and Jun Ni. "Predicting bitcoin returns using high-dimensional technical indicators." The Journal of Finance and Data Science 5, no. 3 (2019): 140-155.

[13] Rathan, Karunya, Somarouthu Venkat Sai, and Tubati Sai Manikanta. "Crypto-currency price prediction using decision tree and regression techniques." In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 190-194. IEEE, 2019.

[14] Ji, Suhwan, Jongmin Kim, and Hyeonseung Im. "A comparative study of bitcoin price prediction using deep learning." Mathematics 7, no. 10 (2019): 898.

[15] Chen, Zheshi, Chunhong Li, and Wenjun Sun. "Bitcoin price prediction using machine learning: An approach to sample dimension engineering." Journal of Computational and Applied Mathematics 365 (2020): 112395.

[16] Gadey, Reshma Sundari, Nikita Thakur, Naveen Charan, and R. Reddy. "Price prediction of bitcoin using machine learning." (2020).

[17] Raju, S. M., and Ali Mohammad Tarif. "Real-time prediction of BITCOIN price using machine learning techniques and public sentiment analysis." arXiv preprint arXiv:2006.14473 (2020).

[18] Aggarwal, Apoorva, Isha Gupta, Novesh Garg, and Anurag Goel. "Deep learning approach to determine the impact of socio-economic factors on the bitcoin price prediction." In 2019 Twelfth International Conference on Contemporary Computing (IC3), pp. 1-5. IEEE, 2019.

[19] Chen, Zheshi, Chunhong Li, and Wenjun Sun. "Bitcoin price prediction using machine learning: An approach to sample dimension engineering." Journal of Computational and Applied Mathematics 365 (2020): 112395.

[20] McNally, Sean, Jason Roche, and Simon Caton. "Predicting the price of bitcoin using machine learning." In 2018 26th Euromicro international conference on parallel, distributed and network-based processing (PDP), pp. 339-343. IEEE, 2018.