International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 2 ǁ 2022 ǁ PP. 42-47

Survey on Cryptocurrency Price Prediction Using Machine Learning

Lloyd Kasal[1], Mihir Shetty[2], Tanmay Nayak[3], Ramanath Pai[4], Shilpa B[5]
[1234] Students, Dept. of Information Science and Engineering, CEC, Karnataka, India
[5]Asst. Professor, Dept. of Information Science and Engineering, CEC, Karnataka, India

Abstract
Cryptocurrency, in simple terms, is a digital currency in which financial transactions are verified, which is called mining, and records are maintained by a decentralized system using cryptography, rather than by a centralized authority. One of the aims of it is to decentralize the influence of institutions/organizations and to give greater control to the concerned users/investors. In India, various decisions regarding its implementation have been discussed, due to its increased importance, relevance and influence in people's lives recently, due to increased access, which requires systems consisting of powerful processors for mining, etc. The popularity and importance of cryptocurrency, along with its relevance, has been growing in the recent years. Many organizations throughout the world have acknowledged the importance of adapting to this technology to avail the various numerical benefits virtually, with rapid advances seen in this direction. Some of the fields/domains, where it is made use of include: Stock Price Prediction, Cryptocurrency Price Prediction, Real Estate Price Value Prediction, etc. Price prediction happens to be one of the biggest challenges in the present-day fiscal exchange situation, as the dilemma to sell/buy cryptocurrency is a fascinating objection witnessed by merchants. Since some past few years, it seems to have arrived at unparalleled heights resulting in theories justifying the pattern of its surge. The understanding of if the cryptocurrencies' movements and momentary fluctuations capable of being forecasted, has managed to keep investors, researchers and other economists a lot occupied as of late. To be able to enforce these judgements in the cryptocurrency market, it is mandatory to have optimal and reliable means of price prediction, asset and liability analysis, at minimum for the short-term time range. We can achieve this by using machine learning models such Long Short-Term Memory (LSTM), Multi-Layer Perceptron), Recurrent Neural Network (RNN), etc., which can learn long term dependencies. Some of the essential elements made use of are close price, available price, low price, high price, market cap and volume, along with association among a few cryptocurrencies, therefore centred at evaluating crucial characteristics that affect the transaction's unpredictability by using the system to increase the usability of this practice. Our group has been able to review 10 articles/research papers in relation to cryptocurrency, price prediction of any kind/commodity that made use of various techniques, such as RF (Random Forest), LSTM (Long-Short Term Memory), ANN (Artificial Neural Networks), MAPE (Mean Absolute Percentage Error), Mean Bias Error (MBE) and some others that are aforementioned, to name a few, in order to make a prediction of the prices of various financial commodities, such as Stocks and Cryptocurrencies. To summarize, these studies have shown an ideally considerable level of estimation accuracy. The various techniques and algorithms that have been made use of, in order to be able to make a proper estimation of the prices, varied to a considerable extent between the papers. A small amount of people took ahead and worked on findings of previous methodologies by adding newer additional layers, while the rest looked at the case through a new vision. The various techniques and methodologies, that our group touched upon, seemed to have their own advantages and drawbacks, while being suited for specific scenarios.

Keywords: *Cryptocurrency, Mean Bias Error, Long Short-Term Memory, Artificial Neural Networks.*

Date of Submission: 02-02-2022 Date of acceptance: 16-02-2022

I. INTRODUCTION

Prediction of financial markets such as the cryptocurrency market has been researched at length. Emerging cryptocurrencies such as Bitcoin, Ethereum and Litecoin present an interesting parallel to this, as it is a prediction problem of time series in a market which is still in its impermanent stage. As a result, there's high volatility inside the market and this presents an opportunity in phrases of prediction. Conventional time collection prediction methods consisting of holt-winters exponential smoothing models depend upon linear assumptions and require facts that can be divided into seasonal, trend and noise to be powerful. This sort of method is more appropriate for a mission consisting of forecasting income wherein seasonal consequences are present. Because of the dearth of seasonality inside the cryptocurrencies such as Bitcoin marketplace and its high volatility, those strategies aren't very powerful for this project. Given the complexity of the project, machine learning makes for an interesting technological answer based on its overall performance in comparable areas. Duties which include natural language processing which can be also sequential in nature and has a feature shown promising effects. This type of challenge uses records of a sequential nature and as a result is much like prediction assignment for price. The recurrent neural network (RNN) and the Long short term memory (LSTM) flavour of artificial neural networks are favoured over the traditional multilayer perceptron (MLP) because of the temporal nature of the more superior algorithms. The purpose of this research is to envision with what accuracy can the rate of cryptocurrencies like bitcoin be predicted using knowledge of machine learning. Machine learning literature for cryptocurrencies such as Bitcoin is constrained. As a result, literature relating to different economic time series prediction the usage of deep studying is also assessed as these obligations may be taken into consideration analogous. Due to the ever-growing popularity, it is important to be up to date on it's market price by keeping track of it. This project helps the user using it to receive real time trend changes of various cryptocurrency prices along with detailed statistical summary.

1.1. Literature Survey

[1] In this paper, 3 kinds of algorithms for machine learning are built and used for predicting the costs of 3 kinds of cryptocurrency—BTC, ETH, and LTC. Performance measures have been carried out to check the accuracy of various models. Then, we as compared the real and predicted prices. The effects display that GRU outperformed the other algorithms with a MAPE of 0.8267%, 0.2454%, and 0.2116% for ETH, BTC, and LTC, respectively. The RMSE for the GRU version became discovered to be 26.59, 174.129, and 0.825 for ETH, BTC, and LTC, respectively. Based on those effects, the GRU model for the focused cryptocurrencies may be considered reliable and efficient. This method is considered the satisfactory model. But, bi-LSTM represents much less accuracy than GRU and LSTM with large variations among the real and the anticipated expenses for both BTC and ETH. The experimental outcomes show that the Artificial Intelligence algorithm is proper and reliable for cryptocurrency prediction. Prediction of cryptocurrency prices of GRU is better than bi-LSTM and LSTM however all the algorithms represent predictive results which are excellent.

[2] The foremost goal is to acquire a line that nice suits the records. A line is stated to be nice suit whilst it possesses a smaller overall prediction error. Distance among factors in regression line is known as an error. To make the version examine we want to cognizance on LSTM & RNN to permit identity of smaller series patterned records and the price of the following day is predicted. The training of the records wishes a big quantity of filtered records sets. When records reach a stabilization factor it doesn't increase its outcome performance so with the aid of using the feature, we will enhance it with the aid of using loss feature, activation feature, optimizer. Results relay on training records which may be later pass checked with the aid of using the prevailing price of statistical importance test.

[3] This paper noted that although current works of research have made use of Machine Learning for better accuracy of the Bitcoin price prediction, only some seem to have concentrated on the usefulness of making use of a variety of methods to inspect with various data structures, as well as elemental characteristics. The authors of the paper, in order to estimate Bitcoin prices on various recurrences making use of ML approaches, initially classified the Bitcoin value by various parameters, such as daily price, high frequency price. A cluster of multi-dimensional characteristics, that include network and property, market and trading, gold spot price and attention have been made use of for Bitcoin everyday value forecast, whereas, the basic trading characteristics obtained from a cryptocurrency exchange have been used for the value estimation in between a 5-minute intermission value. It was also observed that, while comparing with the standard outcomes for prediction of the daily price, they accomplished an enhanced efficiency, with the maximum accuracy rates of 66% and 65.3% for the statistical methods as well ML algorithms, accordingly. One of the main observations made was that some Machine Learning techniques that include Random Forest (RF), XGBoost (XGB), Quadratic Discriminant Analysis (QDA), Support Vector Machine (SVM) and Long Short-term Memory (LSTM) for a 5-minute estimation of Bitcoin surpassed the statistical methods, with accuracies heading towards a rate of 67.2%. It was also noted that their analysis of Bitcoin rate forecast could be assumed to be a mini-survey of significance in ML approaches. Comprehensively, it was noted that the Logistic Regression and Least Discriminant Analysis approaches surpassed the other ML systems compared to the dataset of daily price, demonstrating that correctly chosen multi-dimensional characteristic groups are able to make up for the discreetness of the systems for the everyday value forecast of Bitcoin. The paper seemed to have some drawbacks, of which, one was that only two types of data for prediction had been considered, while it is actually a necessity to collect data of the price with various characteristics having more aspects. The other drawback being that the weightage of all of the Machine Learning algorithms in the studies had not been done. The final remark made was that, in order to further this analysis, they aim to study more methods, like the statistical

method such as ARIMA (Autoregressive Integrated Moving Average), and the ML model RNN (Recursive Neural Network), besides, some alternative characteristics were to be assumed, also, their additional analyses would concentrate on a greater appropriate aspect of the opinion making use of textual mining as well as studies of social clusters.

[4] In this paper, some techniques such as ANN(Artificial Neural Network) and RF(Random Forest) seem to have been used, to estimate the closing price of the following day for five companies that belong to various domains of service. Here, the data related to finance relevant to stocks have been made use of to construct newer variables that are used as intake for the system. The systems have been classified making use of benchmark strategic indices, like MAPE and RMSE. The lower rates for the 2 aforementioned indices demonstrate the systems' capability of estimating the stock closing price. The paper noted that, for stock market analysis, in particular, for dealing with such type of data, an optimal system is required which can recognize some underlying trends and complicated relationships for a dataset of such large quantity. Also, ML methods in this field seemed to have demonstrated to enhance the efficiency rates by about 60%-86% in comparison to the previous approaches. To further proceed with the analysis, previous data from 5 companies were collected through Yahoo Finance, also, this dataset consists of information of 10 years of various companies, it consists of knowledge of stocks. It is to be noted that the daily closing price of the stock alone seems to have been retrieved. The correlative study based on MBE, RMSE and MAPE values seemed to, without any doubt, indicate that ANN demonstrated a more accurate estimate of stock prices, in comparison with RF. Outcomes displayed some finer rates retrieved from the ANN (Artificial Neural Network) model gave a Root Mean Square Error value of 0.42, MAPE of 0.77 and MBE of 0.013. Some of the drawbacks of this paper were that less number of parameters had been taken into consideration, exempting some factors., which possibly could have resulted in yielding better results.

[5] In the paper "Deep Neural Networks for Cryptocurrencies Price Prediction", A hybrid model of RNN, MLP and LSTM was used for the prediction. LSTMs have been found to have the highest accuracy in predicting the directional movements of cryptocurrencies. However, the limitation that trend prediction neural networks do not understand the economic cost of misclassification.

[6] In the paper "Forecasting of Cryptocurrency Prices using Machine Learning", Models such as BART, MLP, RF were used. We have confirmed that these models do not impose strict constraints on the statistical characteristics of the cryptocurrency time series investigated. The mean absolute percentage (MAPE) of BART and MLP was about 3.5% on average, while the mean absolute error percentage of RF was about 5%. We can conclude that the limit was that the smallest dataset was used using only the lag value and the prediction accuracy was minimal. This can be extended by using a dataset with more fields and values.

[7] In the paper "Time Series Prediction with LSTM Networks and Its Application to Equity Investment", LSTM, LOG, RF and GBT models were used. We have found that using LOG as a baseline can bring out the benefits of complex and computationally intensive LSTM networks. Expected to be robust as it has the fewest hyperparameters. The predicted probabilities and theoretical values are in good agreement. RF allows you to combine different types of weak learners and combine them to perform diverse learning. Creates strong estimators that are high performance and less likely to overfit. GBT uses a boosting method to transform a weak learner, a decision tree, into a highly accurate learner. This works by applying weak learners to the repeatedly reweighted version of the sample. The SVM algorithm includes a powerful abstract background, perceptual geometric analysis, and other features that combine core region expansion with convex expansion. It also outputs the optimal separation hyperplane for classifying new samples. The performance of the LSTM was about the same as the LOG and SVM. Since LSTMs are LOG-like discriminative models, the predicted probabilities are called predictability. In the end, unlike LOG, I came to the following conclusions. RF and GBT showed inadequate results due to a discrepancy between the predicted probabilities and the theoretical values. This could be due to his use of decision trees because he is a poor learner. Estimating predictability using predicted probabilities and performing probability calibration processes that may be required is not easy. The SVM results show balanced behaviour, including LOG. However, LOG is a discriminative model and SVM is a discriminative feature. Therefore, it is not possible to calculate the prediction probability using SVM.

[8] In this paper, Many Algorithms are used like LSTM, GRU, RNN, and many more. In this, the authors analyzed the certainty of the bitcoin market in several prediction time-horizons ranging from one to sixty minutes. Various machine learning models were used, and it found out that it was outperforming random classifier and is well suited for Gradient boosting Classifier as it is well for prediction tasks. Many feature sets like scientific, blockchain-based, and credit-based features were used, but scientific features were the most relevant. The prediction rate will increase for extended prediction scope. In this, the ML models are able to anticipate short-run movements of the bitcoin market and are precisely able to see the exactitude lightly above 50% and show it is limited. This can be because of a couple of reasons like a right away market attitude to features or a massive quantity of facts that is past the features that impact the bitcoin marketplace. Prediction Accuracy will increase for longer prediction scope, and it makes manner for similar studies opportunities. These

models distinguish between the weight of features and rely on a few feature sets, the most influential and exact source of the predictive power is still ambiguous. Many theoretical equilibrium models say it's based on transaction cost and benefits and may partially affect the prediction. LSTM has the best accuracy to predict the directional movements of cryptocurrencies (1-min and 60 min prediction horizon). GBC weigh the intake of many decision trees and also train individual weights. GRU has more accurate predictions on 15-min horizon. The systems are capable of expecting the motion between the precision varying from 51% to 56.0%, the forecasting efficiency increases for more time scopes and GBC and RNN are appropriate and scientific features continue to be crucial and for more forecast scopes the features like transaction per second are ideal and technical features become less relevant.

[9] The authors of the paper explain predicting bitcoin using rectilinear regression and the LSTM model. In this model machine learning algorithms and artificial neural models are used. Several data sets are used to train ML and AI models. It was represented that the rectilinear regression model's accuracy is extremely high compared to other ML systems, it had been found to be 99.96% accurate. LSTM on the opposite hand shows a lesser error rate of 0.08 percent. This relation determines that the LSTM model is better than the rectilinear regression model. The largest distress is the portability and it is very much imperfect by the figure of trade. The agility of trade is very tough and the cryptocurrency market cannot fight with global leaders like Mastercard and Visa. The cryptocurrency marketplace is very erratic and there has never been a prediction that was hundred percent accurate. The dominant subjects are trend, momentum, and pattern in general. The Mean Squared Equation is used in rectilinear regression to determine the correctness of the linear graph with regard to the time frame data set. The LSTM model has the potential to discover correctness in terms of Mean Absolute Error, indicating a 0.08 percent error rate. The LSTM model has a higher prediction rate than the linear regression model, but only by a little margin.

1.2 METHODOLOGY

Neural network is an amazing concept where we train the machine to use the past experiences (data) for a current problem. It is based on the Neural networks present in a human brain which sends millions of signals in an instance and face a situation based on past knowledge. RNN are special kind of neural networks that uses loops in them, allowing information to persist. A RNN can be considered as multiple clones of the same network, each passing the message of the predecessor to a successor. Upon unrolling an RNN, we get the following network The chain-like structure shows that RNN are closely associated to sequences and lists. They are the normal design of Neural Networks to apply for such data. In the previous gone by years, there have been major achievement using RNN to a diversity of complications: Speech recognition, Language modelling, paraphrase etc. These successes were possible due to the use of "LSTM", a very special kind of RNN which works, for various tasks, way better than any orthodox edition. Almost all the above results based on RNN are accomplished using them.

1.2.1. Drawback of RNN for Long term dependencies.

Key selling points of RNN is the notion that they are able to link past information to the current project. If we are trying to predict the last word of a sentence "The Sun rises in the…." We can directly tell, without any context that the word is going to be 'East'. During these situations, where the interval between the appropriate info and the position that it needs to be is minimal, RNN's can study the application of the previous information.

However, for some cases, minimal context is insufficient. Consider predicting the last word of the sentence "I was raised in Barcelona… I can fluently converse in Spanish." Latest information advocates that the next word is undoubtedly a language, but if we want to limit which language, we need the background idea of Barcelona. It is likely for the interval between the appropriate info and the position where it is required to become huge. Regrettably, as this limit increases, RNNs are not able to learn to link the info.

Hence, we can conclude that RNNs are not capable of handling "Long-Term Dependencies." A human can carefully pick the parameters for them to solve small problems. But, in practice, RNNs can't learn them. Thankfully, using LSTM, we can solve this issue.

1.2.2 LSTM Networks

Long Short-Term Memory Networks (LSTM) is a special type of RNN that can learn long-term dependencies. They are widely used in a variety of problems and have surprising results.

LSTMs are specially created to circumvent long-term dependency issues. Remembering information for a long time is their practical nature, not something they have to learn carefully! Every RNN has a chain of iterative neural network modules. In a standard RNN, this module has a very simple structure, such as a hyperbolic tan layer. LSTMs have a similar structure, but repeating modules have different structures. Instead of a single neural network layer, there are four layers that interact in a special way.

The key to the LSTM is the state of the cell, the horizontal line that runs through the top of the diagram. The state of the cell is like an assembly line. It goes straight through the chain with just a few small linear interactions. Information just flows without change. LSTMs have the ability to delete or insert cell state information that is carefully controlled by a structure called a gate. Gates are the best way to get information through. They comprise of a sigmoid neural network layer and pointwise multiplication. The number of outputs in the sigmoid layer is between 0 and 1, indicating the capacity of each component. Zero indicates "allow none" and 1 indicates "allow all". The LSTM has three gates to protect and control the cell gate.

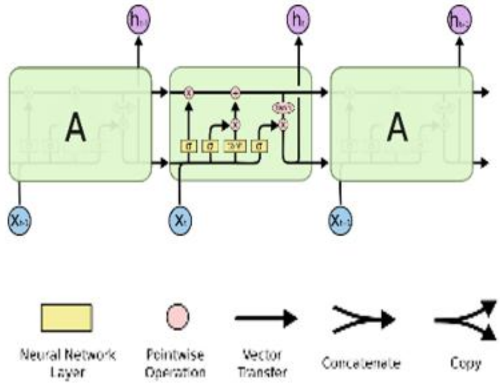

Figure 1: LSTM

1.3. DATA-FLOW DIAGRAM

First, Dataset is imported to the project. Datasets are the historical values of different cryptocurrencies, like Bitcoin, Ethereum, Tether etc. Then python libraries like keras from TensorFlow, Theano etc. are loaded which has complicated mathematical functions that are needed to predict the values of the imported data. Next, the datasets are split into training data and testing data. The machine plots the cryptocurrency prices and normalizes these values and extracts the data. This data is sent as the input data to the LSTM cells, building a model. The parameters are set up to train the model and finally the machine plots and predicts the prices

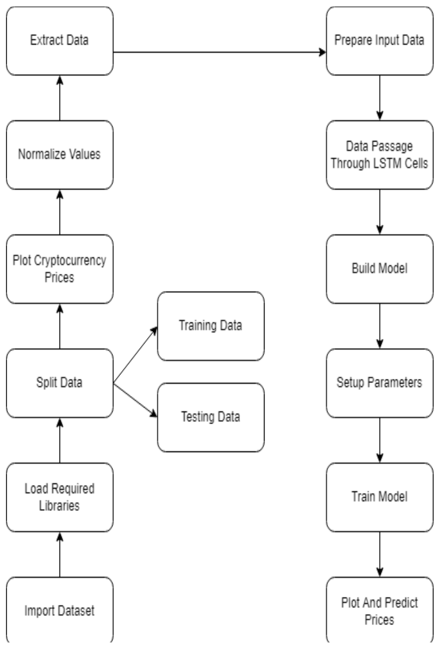

Figure 1: Data Flow Diagram

II. CONCLUSION

Researchers have had a tough time predicting cryptocurrency values because of outside factors or external social aspects that influence price forecasting. In addition to this, a variety of machine learning and deep learning techniques are utilized to make forecasts. In recent years, neural networks have shown promise in the prediction of time-series information. To analyze cryptocurrency values, numerous different kinds of neural networks may be utilized. The most successful of them all has been determined to be LSTM. This might be because of their ability to recall and derive materialistic features of information. In conclusion to the outline of this project, the proposed internet site interface goals to help users/clients/investors predict the change in prices of diverse cryptocurrencies happening in actual time, in addition to offering an in-depth statistical summary and rate assessment of diverse cryptocurrencies, which could assist them to evaluate a huge variety of records concerning the crypto market scenario of that time. In addition to this, it's highly viable to improve the overall performance of this model through utilizing higher and more efficient techniques, and additionally, through the use of extra systems equipped with powerful Graphical Processing Units (GPUs) and a quicker processor, which could doubtlessly bring about more training speed for the data, alongside its accuracy.

REFERENCES

[1]. Hamayel, M.J.; Owda, A.Y. A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. AI 2021, 2,477-496. https://doi.org/10.3390/ai2040030
[2]. Laxmi, K.R., Reddy M.A., Shivasai, CH., Reddy, P.S. (2020). Cryptocurrency Price Prediction Using Machine Learning. SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, (2020); DOI : 10.18090/samriddhi.v12iS3.3
[3]. Spilak, B. Deep Neural Networks for Cryptocurrencies Price prediction. Masterarbeit, DOI:10.18452/19249
[4]. Derbentsev, V., Matviychuk, A., Soloviev, V.N.; Forecasting of Cyptocurrency Prices Using Machine Learning; Advanced Studies of Financial Technologies and Cryptocurrency Markets. doi:10.1007/978-981-15-4498-9
[5]. Matsumoto, K., Makimoto, N.; Time Series Prediction with LSTM Networks and Its Application to Equity Investment; Advanced Studies of Financial Technologies and Cryptocurrency Markets. doi:10.1007/978-981-15-4498-9
[6]. Chen, Z., Li, C., & Sun, W. (2019). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 112395. doi:10.1016/j.cam.2019.112395
[7]. Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. Procedia Computer Science, 167,599–606. doi:10.1016/j.procs.2020.03.326
[8]. Jaquart, P., Dann, D., & Weinhardt, C. (2021). Short-term bitcoin market prediction via machine learning. The Journal of Finance and Data Science, 7, 45–66. doi:10.1016/j.jfds.2021.03.001
[9]. Ho A, Vatambeti R, Ravichandran SK (2021) Bitcoin Price Prediction Using Machine Learning and Artificial Neural Network Model. Indian Journal of Science and Technology 14(27): 2300-2308. https://doi.org/10.17485/IJST/v14i27.878