ISSN: 0974-5823 Vol. 6 No. 2 September, 2021

International Journal of Mechanical Engineering

PREDICTING THE PRICE OF BITCOIN WITH MACHINE LEARNING MODELS

[1]**Dr.Shaik Saidhbi,** Associate Professor, Department of computer science, Samara University, sfajju.syed@su.edu.et, Ehiopia

[2]**Adem ali Kabo,** Department of computer science, Samara University, ademali@su.edu.et, Ehiopia

[3]**Ahmed Adem ,** Computer Science, Samara University, aniadem.adem0@gmail.com,Ethiopia

Abstract

The value of Bitcoin, one of the most well recognised cryptocurrencies, spiked above $17,000 in November 2020, a level not seen in three years. The value of the cryptocurrency has seen massive swings since its introduction in 2008. To hedge against the uncertainty in the financial markets brought on by the epidemic, investors flocked to cryptocurrencies. In this study, we compare several machine learning models' accuracy in forecasting Bitcoin's value against the dollar. Some examples of accessible models are linear regression, random forest, support vector machine, ARMA, LSTM, and RNN. The results will improve our understanding of the peculiar behaviour of the cutting-edge cryptocurrency.

1. INTRODUCTION

Bitcoin

In 2008, Satoshi Nakamoto created Bitcoin, the first decentralised cryptocurrency. The primary goal was to facilitate participant-to-participant (P2P) negotiations [1]. All Bitcoin transactions are validated by nodes currently online in the Bitcoin network. Transactions that have passed validation are recorded on the Blockchain, a public distributed ledger. Bitcoin mining is the process through which Blockchain is kept operational. Using the processing capacity of the central processing unit (CPU) to tackle difficult mathematical problems is one way for anybody to earn Bitcoins.

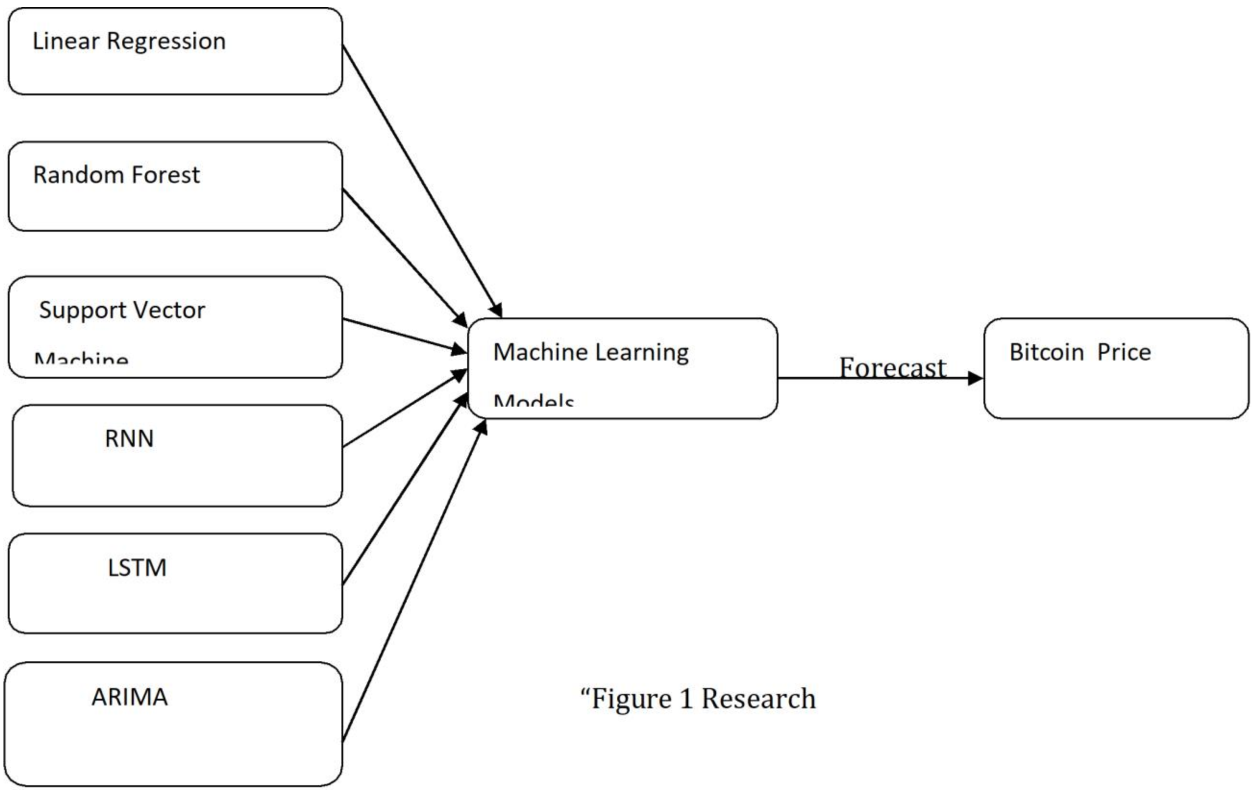
COVID-19 and Bitcoin

Capital markets, especially markets for virtual currencies, have been affected by the COVID-19 epidemic. [2] The article explains how the pandemic would affect the economy and society. The importance of the pandemic in the financial sector and the harm it brought to global financial markets are highlighted in the report. Cases of and deaths from COVID-19 have been linked to the use of cryptocurrency (such as Bitcoin, Ethereum, and XRP). The correlation between Bitcoin and reported fatalities or cases has swung around to become positive [3]. The key question is whether or not Bitcoin is a safe haven or a high-risk venture. When Bitcoin's value plummets as a result of the COVID-19 outbreak, what exactly does that make it? Conlon and McGee [4] answer both of these issues, showing that Bitcoin does not operate as a safe haven during times of economic uncertainty.

Price Prediction

The practise of predicting future prices is not new. Predicting stock prices is a hot area of study. [5] Used machine learning algorithms to provide more accurate stock price predictions. Prediction was achieved with the use of regression and LSTM models. Improved accuracy has been shown by both models, however LSTM has produced more efficient outcomes. [6] determined future stock values by analysing the correlation between market news and stock performance.

In this research, we use a variety of machine learning algorithms to forecast the daily price of Bitcoin. Here, we employ models such as the Support Vector Machine (SVM), Linear Regression (LR), Recurrent Neural Network (RNN), Random Forest (RF), Auto-Regressive Integrated Moving Average (ARIMA), and Long Short Term Memory (LSTM) (LSTM). As an overview of the literature, see Figure 1.



“Figure 1 Research

2. RELATED WORK

Bitcoin's youth and volatility mean that there is a dearth of research into predicting its value. Bitcoin's widespread adoption can be attributed, in large part, to its decentralised peer-to-peer network and ease of use. Author found link between Bitcoin price and Twitter and Google search volume [7]. The promising accuracy of 77% when analysing tweet volume and 66.66% when using Google Trends to anticipate Bitcoin's price using polynomial regression. The "latent source pattern," first proposed by Chen [8] and used to predict Bitcoin's price by [9], is a popular method. In just 50 days, investors saw an 89% ROI at a 4.10 Sharpe ratio. Finding out how different types of media coverage affect Bitcoin and other digital currencies is another goal of this investigation. [10] We compared the SVM, ANN, and Ensemble algorithms and analysed their results (k- means clustering and recurrent neural networks). In order to accurately estimate prices, the support vector machine method has shown to be the most effective. Machine learning methods including support vector machine, Random Forest, Long Short Term Memory, Quadratic Discriminant Analysis, and XGBoost were used to make daily Bitcoin price forecasts [11]. With an accuracy of 65.5%, support vector machines were the most effective. Researchers [12] found that using just one hidden layer in their feed-forward neural networks (SLFNs), they were able to forecast Bitcoin prices with an accuracy of 60.05% by analysing the edges that are used most often (approximately).

[13] The future value of Bitcoin was estimated using a recurrent neural network and a linear regression model. In contrast to the regression model, the RNN model's ability to spot long-term dependencies resulted in a lower value for the mean square error. Linear regression, RNN, and Random forest are compared and contrasted in [14]. The RNN with LSTM has enhanced the model's ability to predict Bitcoin prices, which is particularly useful given the data's tendency to fluctuate often. The model's MAE was 0.0043, which was lower than that of both linear regression and random forest. McNally evaluated the accuracy of the Automated Regressive Integrated Moving Average (ARIMA) model, the Recurrent Neural Network (RNN), and the Long Short Term Memory network and found that the latter achieved the highest accuracy (52 percent). The author, Ladislav[16], looked at how interest in Bitcoins was correlated with use of the websites Wikipedia and Google. The data also highlighted the discrepancy between the rising popularity of cryptocurrencies and the prevailing downward or upward value trends. The ARIMA model was used in [17] to provide timely predictions. ARIMA is more effective in predicting Bitcoin prices (4,1,4). The average absolute error for Day 1 estimations was 0.87, whereas the error for Day 7 forecasts was 5.98. This is why our short-term

performance has increased since we started using ARIMA.

3. METHODOLOGY

Data Collection

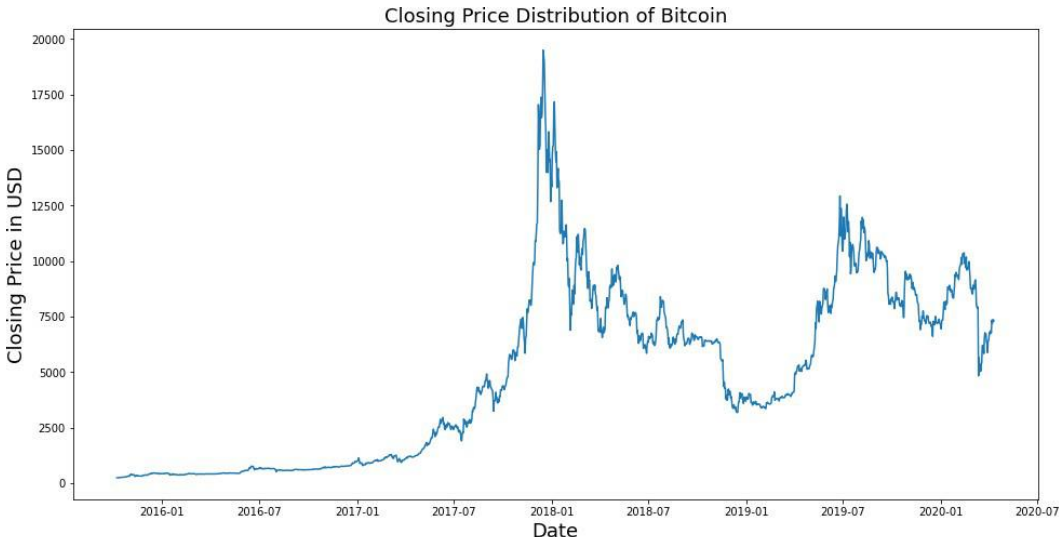
Kaggle, an online data science community, is where we get our Bitcoin data. From Oct 8th, 2015, through Apr 10th, 2020, this data collection is available. Every cell in this table contains the following data about Bitcoin: Close, Date, High, Volume in Bitcoins, Low, Volume in Bitcoins, and Opening price in Bitcoins.

Features	Explanation
Date	Bitcoin's price in specific Date
Open	Opening Price
Close	Value at the day's close
High	High Price
Low	Low Price
Volume	Quantity from Major Markets

"Table 1 Features of dataset"

Feature Engineering and Evaluation

To improve the accuracy of predictions made by machine learning algorithms, the practise of Feature Engineering is used. It's crucial to the data mining process if you want to succeed in your endeavours. Figure1 displays the frequency of respective closing prices.



"Figure 2: Bitcoin Closing Price Distribution"

Modelling

When you mine for data, you're looking for insights to glean from the numbers. Machine learning provides the computational underpinnings for information retrieval. Machine Learning (ML) is a branch of AI in which computers are taught through observation and experience rather than being given specific instructions. It's possible for instances in a data set to share many features. Supervised learning and unsupervised learning are the two primary types of machine learning. Supervised learning is a kind of machine learning that makes use of labelled data sets as a model. A case may be modelled using a set of independent variables (represented by x) and a set of dependent features (represented by y). A dependent variable might be either a discrete or continuous measure. If the target feature is continuous, a regression model should be employed, whereas a classification model should be used if the attribute is discrete. Support vector machines and neural networks, on the other hand, are instances of supervised learning. In unsupervised learning, the model takes in data and uses it to make discoveries about patterns and trends. When modelling a data set in which a certain attribute is missing, this technique is employed. The study's focus is on developing an approach for forecasting Bitcoin prices. Supervised machine learning may be used since the target is known in advance. Some of the methods that have been implemented include: Random Forest, Linear Regression, Support Vector Machine, Autoregressive Integrated Moving Average, Long Short-Term Memory, and Recurrent Neural Network.

Linear Regression

Linear regression establishes a relationship between a dependent and an independent set of variables. This equation is used to find a straight line across a collection of data points:

"Y=a+bX" (1)

In this equation, X is the independent variable, Y is the dependent variable, b is the slope, and an is the intercept.

Random Forest

Commonly used for both regression and classification, Random Forest is becoming more widespread. It optimises results by combining numerous decision trees. Several problems with categorization may be solved using decision trees. The decision tree is akin to a tree structure in that it recursively divides the feature space. If dividing does not improve the prediction, or if each node already contains samples from a single class, the recursion stops. The dataset may be associated with the tree's terminal nodes in the case of decision trees.

Support Vector Machine (SVM)

The SVM model excels at solving issues of binary classification [11]. The objective is to locate a hyperplane that maximally separates the two data sets. For the sake of argument, let's say there's a dichotomy between the two groups. If we consider a space to be n dimensions deep, the true number of points is m bigger.

"R^n is represented by matrix A (xi is i^{th} row of A). xi class can be designated by yi €"

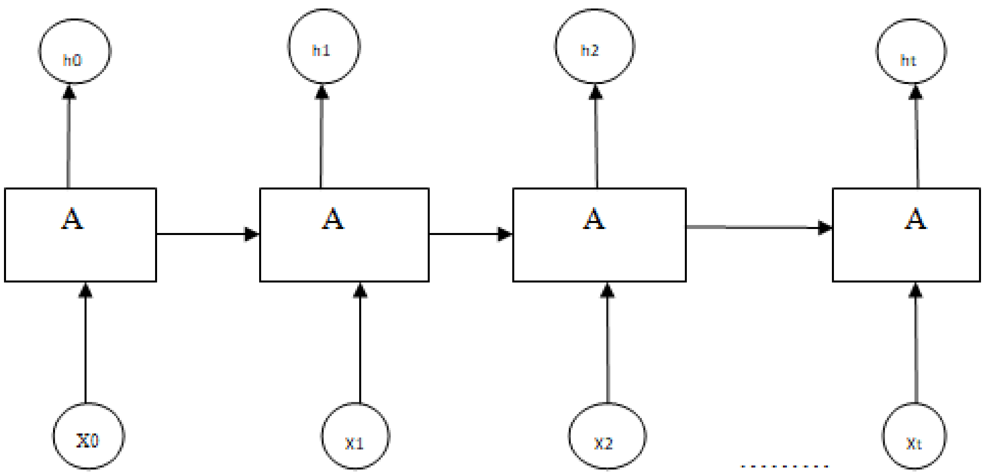
{1,-1}. Thus, we may outline the optimum decision hyperplane search space for a one-dimensional support vector machine:

"x.y + a = 0" (2)

"Where 'y' belongs to R^n, 'a' belongs to R,'y' is the weight vector that is normal to the hyperplane, 'a' is bias. The best hyperplane is achieved with minimization of ||w||."

3.3.4 An Instance-Based Recurrent Neural Network (RNN)

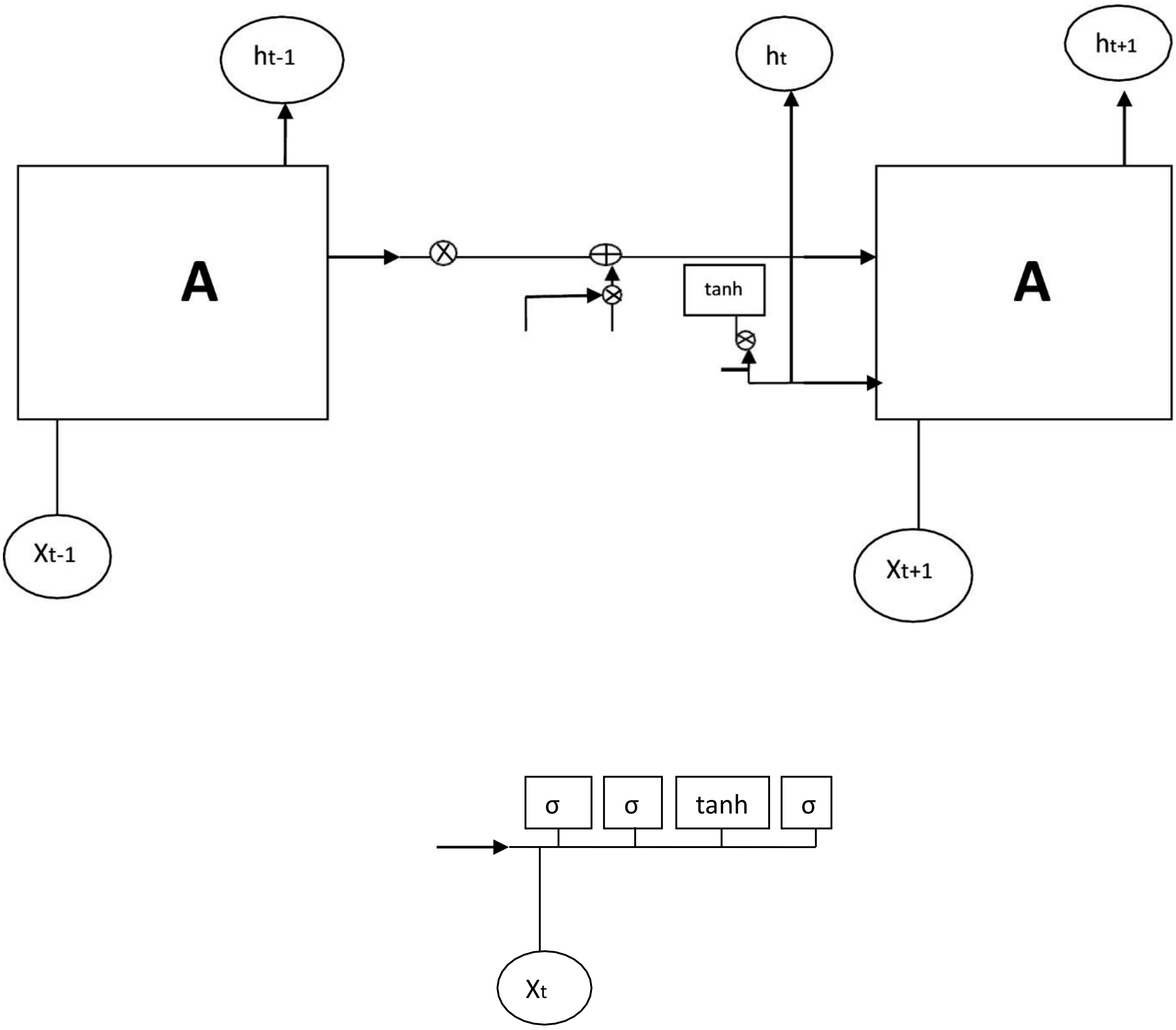
Recurrent Neural Networks were created by J.L.Elman (RNNs). In these neural networks, signals may go both in a circular fashion forwards and backwards. The context layer is what connects all these nodes together. The output of one layer is utilised as input for the next layer up by being sent via a "context" layer. At each instant in time, a new state is formed.



"Figure 3: Recurrent Neural Network"

Long Short Term Memory (LSTM)

With the LSTM model, the RNN's vanishing gradient issue may be addressed. An LSTM consists of an input gate, a forget gate, and an output gate [11]. When dealing with sequence data, LSTM excels at both predicting the future and portraying the past.



"Figure 4 Structure of LSTM"

Here is the notation for LSTM:

"X=[xt]" (3)

"$ht - 1$"

"ft = £ (Wf. X + bf)" (4)

"it = £(Wi . X + bi)" (5)

"ot = £ (Wo. X + bo)" (6)

"Ĉt = tanh (Wc.[ht-1,xt] + bc)" (7)

"Ct = ft ʘ C t-1 +it ʘ Ĉt" (8)

"ht = ot ʘ tanh (Ct)" (9)

An LSTM offset is denoted by bc, bi, bf, and bo, whereas xt denotes input at time t, ht denotes a hidden state at time t, Wo, Wc, Wf, Wi denote a weight matrix, and so on. Activation functions are denoted by £, while the dot matrix multiplier operator is denoted by.

An ARIMA Model for Time Series Data (ARIMA)

An ARIMA model is used to analyse the data and make projections about the future.

"AR(p)+MA(q)=ARIMA(p, d, q)"

There are three orders of ARIMA, denoted by p, d, and q. Auto regression follows the letter p, whereas differencing follows the letter d, and moving average follows the letter q.

Find the link between the two time periods using Auto Regression (AR). What is the ARIMA (p,0,0)?

"Xt =µ+Ø1Xt-1+ Ø2Xt-2+……+ ØpXt-p+ et" (10)

It is the primary purpose of integration (I) to stabilise the time series.

The MA is a statistical measure of the average deviation from prior error values. Hence, we can write the ARIMA (0,0,q) as:

"Xt= µ+ et - Ø1 et-1- Ø2et-2+……+ Øq et-q" (11)

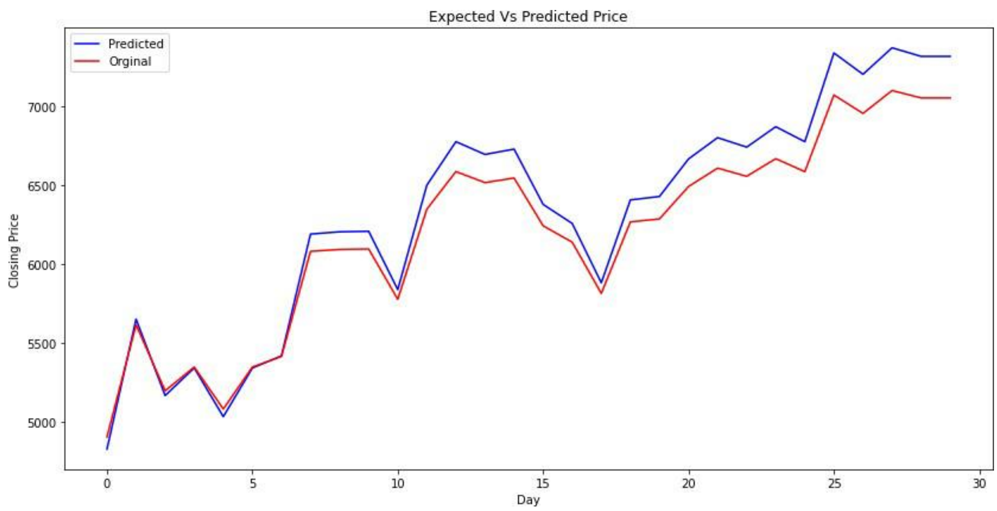
See below for the ARIMA (p,d,q):

"Xt = µ + Xt-1 + Xt-d + ...+ Ø1 Xt-1 + Ø2 Xt-2+ +"

"Øp Xt-p + et - Ø1 et-1- Ø2et-2+…+ Øq et-q" (12)

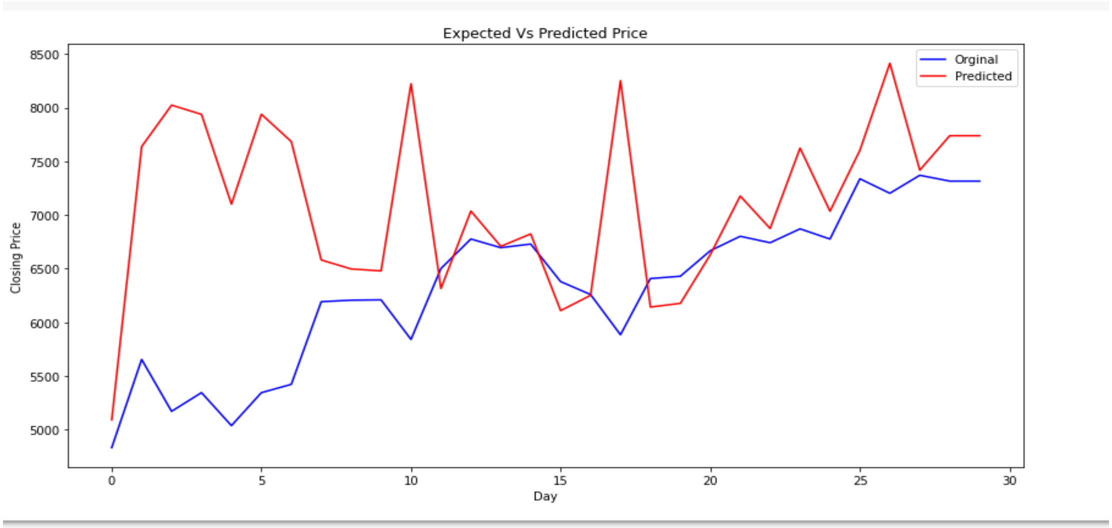
4. IMPLEMENTATION

The next 30 day Bitcoin price forecasts were generated using a dataset that was divided into a training set of 80% and a testing set of 20%. Actual and anticipated graphs are shown as follows:

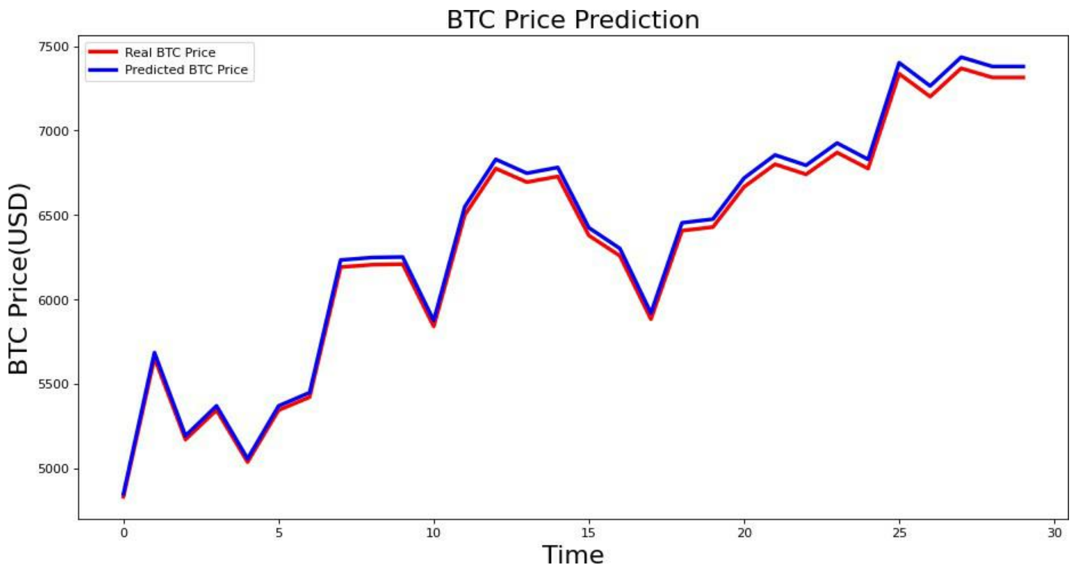


"Figure 5:Actual and Predicted Price (Linear Regression)"

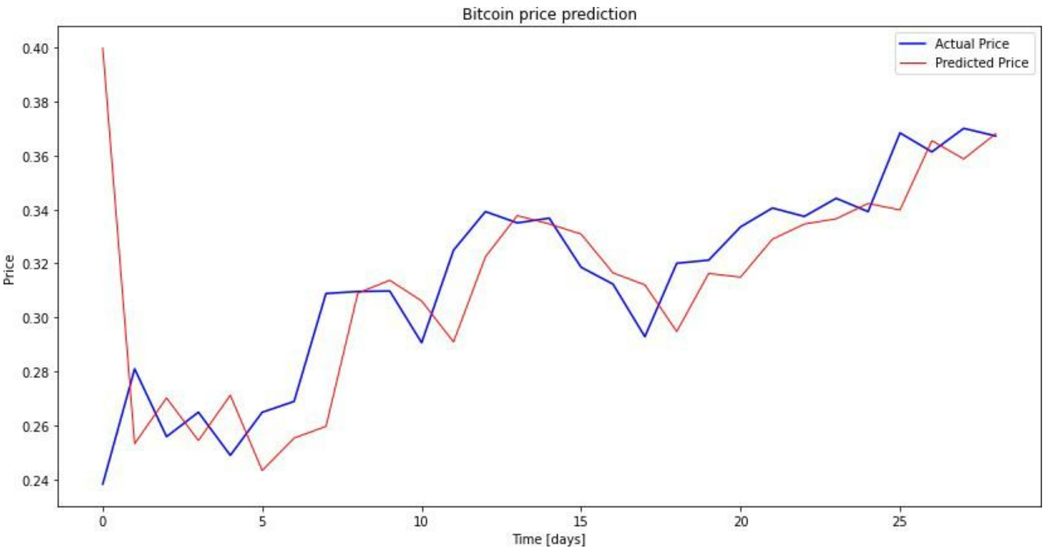
“Figure 6:Actual and Predicted Price (Random Forest)”



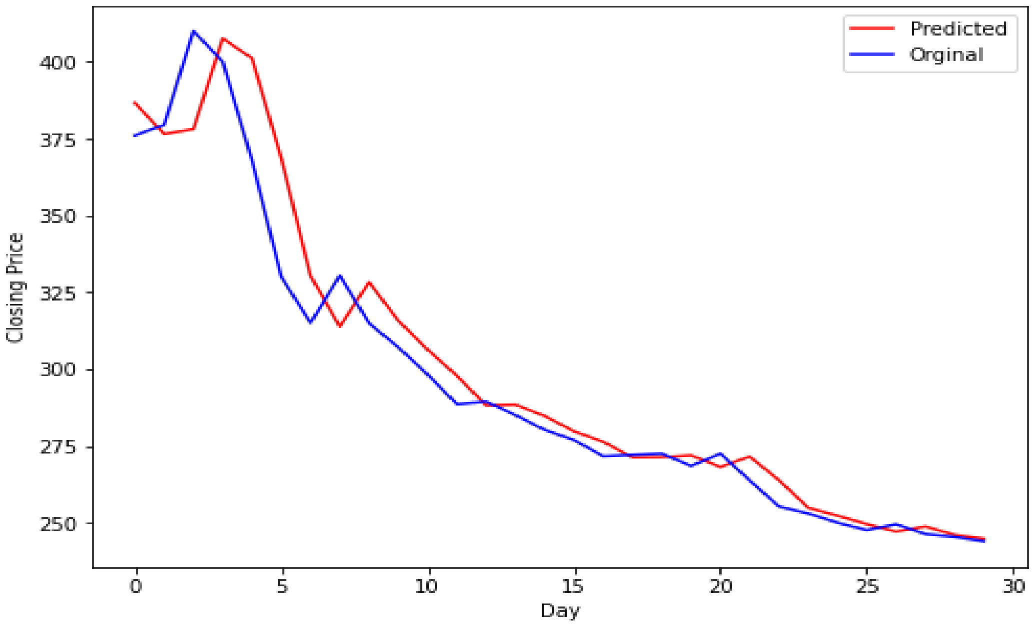
“Figure 7 Actual and Predicted Price (SVM)”



“Figure 8 Actual and Predicted Price (RNN)”



"Figure 9 Actual and Predicted Price (LSTM)"



"Figure 10 Actual and Predicted Price (ARIMA)"

5. EVALUATION

The performance of the models is evaluated using the following metrics:

- "Mean Absolute Percentage Error (MAPE)"

The Mean Absolute Prediction Error (MAPE) measures how off predictions are on average. This is one method in which the efficacy of a forecasting system may be measured.

"$MAPE = \frac{100}{n} \sum_{i=1}^{n} \frac{|actual\ i - predicted\ i|}{actual\ i}$

- “Root Mean Square Percentage Error (RMSPE)”

“RMSPE tells about the percentage of error with respect to actual values.”

“RMSPE= (np.Sqrt (np.Mean (np.Square ((actual-predicted) /actual)))) *100”

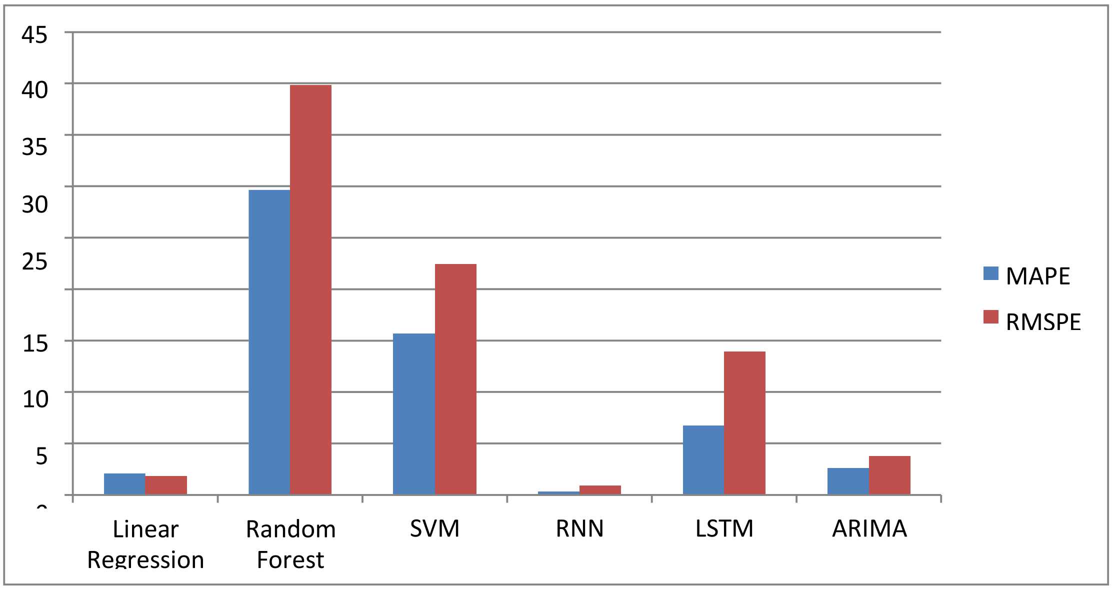
6. RESULTS

The focus of this research is to evaluate and contrast numerous machine learning models that are used to predict the future value of Bitcoin. The outcomes of several statistical models (ARIMA, LSTM, RNN, and Linear Regression) are shown in Table 2. The greatest performance was accomplished by the recurrent neural network, which showed a Mean Absolute Percentage Error of 0.3174 and a Root-Mean-Squared Performance Error of 0.8853.

Model	**MAPE**	**RMSPE**
Linear Regression	2.0795	1.8279
Random Forest	29.6536	39.84291
Recommendation Machine	15.6801	22.4580
Synaptic Recurrent Unit	0.3174	0.8853
Ability to Remember Things Quickly and Completely	6.7365	13.9277
Model	**MAPE**	**RMSPE**
Averaging across time with a regression component	2.5931	3.7452

“Table 2 Model comparison”

“Figure 11 MAPE and RMSPE”



7. CONCLUSION AND FUTURE WORK

Over the last several years, the cryptocurrency market has expanded significantly, catching the attention of both established corporations and newcomers to the space. By providing analysis and conclusions based on Bitcoin price data, it will aid in understanding the complicated and rapidly evolving sector. In conclusion, this research looks at how Bitcoin price predictions may be made using machine learning algorithms. Bitcoin dataset implementation was done using the Google Collaboratory. Additional study is required to ascertain which data sets and qualities will provide the most reliable Bitcoin price forecast.

References

[1] S. Nakamoto, Bitcoin: A peer to peer electronic cash system, 2008

[2] Goodell, J.W,"COVID-19 and finance: Agendas for future research", Finance Research Letters 35 (2020), pp. 101512

[3] Ender Demir,"The relationship between cryptocurrencies and COVID 19 pandemic", Eurasian Economic Review 10 (2020),pp. 349-360

[4] Colon & Mc Gee, R.J. (2020),"Safe haven or risky hazard? Bitcoin during the COVID-19 bear market", Finance Research Letters, 35 (2020), pp. 101607

[5] Palmer et al.,"Stock Market Prediction Using Machine Learning," First International Conference on Secure Cyber Computing and Communication (ICSCCC), 2018,pp. 574-576

[6] Z. Wang,S.Ho and Z.Lin,"Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment," IEEE International Conference on Data Mining Workshops (ICDMW),2018,pp. 1375-1380

[7] A. Mittal, V. Dhiman, A. Singh and C. Prakash,"Short –Term Bitcoin Price Fluctuation Prediction Using Social Media and Web Search, Data", Twelfth International Conference on Contemporary Computing (IC3), 2019, pp. 1-6

[8] G.H.Chen,S.Nikolov,and D. Shah,"A latent source model for non-parametric time series classification," Advances in Neural information Processing Systems, 2013, pp. 1088-1096

[9] D.Shah and K. Zhang,"Bayesian Regression and Bitcoin", Communication, Control and Computing (Allerton), 2014,52nd Annual Allerton Conference on IEEE, 2014, pp. 409-414.

[10] D.C.A. Mallqui and R.A.S. Fernandes, "Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques", Applied Soft Computing Journal (2018), https://doi.org/10.1016/j.asoc.2018.11.038

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

[12] Marcell Tamas Kurbucz,"Predicting the price of Bitcoin by the most frequent edges of its transaction network." Economics Letters 184 (2019) 108655.

[13] H. Kavitha, U.K. Sinha and S.S. Jain, "Performance Evaluation Of Machine learning Algorithms for Bitcoin Price Prediction", Fourth International Conference on Inventive Systems and Control (ICISC), 2020, pp. 110-114

[14] S. Tandon, S. Tripathi, P. Saraswat and C. Dabas, "Bitcoin Price Forecasting using LSTM and 10 Fold Cross Validation", International Conference on Signal Processing and Communication (ICSC), 2019, pp. 323-328

[15] S. McNally, J. Roche and S. Caton," Predicting the price of Bitcoin Using Machine Learning", 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), 2018, pp. 339-343

[16] L.Kristoufek,"Bitcoin meets google trends and Wikipedia: quantifying the relationship between the phenomena of the internet era", Sci.Rep. 3 (2013) pp. 3415

[17] I.M. Wirawan, T. Widiyaningtyas and M. M. Hasan, "Short Term Prediction on Bitcoin Price using ARIMA method", International Seminar on Application for Technology of information and communication, 2019, pp. 260-265