

Predictive Analysis of Bitcoin Prices Using Advanced Machine Learning Models

Ashish Kumar Singh¹, Dr. Mukesh Kumar²

¹Research Scholar, Department of CSE, Rabindra Nath Tagore University, Bhopal, India

²Associate Professor, Department of CSE, Rabindra Nath Tagore University, Bhopal, India

Abstract— : The digital currency industry is seeing unprecedented growth this century, drawing traders, investors, and entrepreneurs from all around the globe. It will aid in documenting the routines and behaviours of such a profitable, demanding, and quickly growing industry by offering comparative studies and insights from the pricing data of cryptocurrency exchanges. In 2021, the bitcoin market will hit one of its all-time highs. New exchanges have increased the accessibility of cryptocurrencies, which has increased their appeal. As a result, not only has the interest in and usage of cryptocurrencies grown, but several reputable crypto enterprises have been launched by some of the pioneers. Companies such as Microsoft, Tesla, and Dell are jumping on the bandwagon for virtual currencies, which are becoming increasingly popular. With the proliferation of decentralized digital currencies, it is more important than ever to educate the public about these new assets so that they can make an informed decision about how to invest their money. Results demonstrate that, among the current research methods, soft computing and machine learning provide the most accurate predictions. Lastly, it is said that machine learning techniques such as ANN and SVMs may be employed to forecast changes in the global stock market. Machine learning, back-propagation, forecasting, artificial neural networks, stock markets, feed forward, root-mean-squared errors, and bitcoin forecasting are all related terms.

Keywords- Bitcoin Forecasting, Machine Learning, Regression, Artificial Neural Network, Back-propagation, Forecasting, Stock market, Feed forward, RMSE

I. INTRODUCTION

One of the first blockchain-based cryptocurrencies was Bitcoin, which was created by Satoshi Nakamoto in 2009. There has been a lot of interest and investment in it because of how decentralized it is and how it may function independently of conventional banks. As a result, investing in Bitcoin is both enticing and hazardous due to its extremely volatile value. In order to make educated judgments, investors, traders, and analysts in the financial sector are always looking for credible ways to predict the value of Bitcoin.

Due to its unique properties, Bitcoin value forecasting presents a number of issues. Market mood, technology advancements, regulatory shifts, macroeconomic events, and the rate of cryptocurrency adoption are just a few of the many variables that affect Bitcoin's value, in contrast to more traditional financial assets. Building reliable prediction models is also difficult because Bitcoin has only been around for a short period of time and there is no central regulating body.

Traditional time-series analysis, statistical models, and econometric techniques have been largely used in previous studies on Bitcoin value forecasts. Nevertheless, these approaches frequently fail to adequately account for the intricate dynamics of the bitcoin market, leading to subpar accuracy and predictive capabilities.

By using more sophisticated machine learning algorithms and better data analysis, this study hopes to overcome the shortcomings of current forecasting systems. In order to improve the precision of forecasts, the suggested method integrates fundamental information, market sentiment research, and price history. We postulate that sentiment analysis, when

added to the forecasting process, would shed light on how market sentiment affects Bitcoin's value and lead to more accurate predictions.

We examine the suggested approach in depth and assess its efficacy using sample data in this study. In this paper, we show how different machine learning methods fared in capturing the fundamental patterns of Bitcoin price fluctuations. Also, we go over what our results mean for investors and how they may be used for things like risk management, optimizing portfolios, and trading methods.

II. RELATED WORKS

There is a substantial amount of literature devoted to the topic of Bitcoin value predictions, which has attracted the attention of academics and investors. Three major approaches have been identified in the existing literature: time-series analysis, statistical models, and machine learning techniques.

One popular approach to predicting Bitcoin's value has been to use time-series analysis tools like ARIMA and GARCH. These models look at past pricing data and try to figure out what patterns and trends there are. Unfortunately, their low prediction accuracy is frequently due to their failure to account for the non-linearities and rapid fluctuations in Bitcoin's price.

Regression analysis and other statistical models look for correlations between Bitcoin price and interest rates, inflation, stock market performance, and other macroeconomic variables. Although these models shed light on the external macroeconomic variables that affect Bitcoin, they could fail to account for the specific features of the cryptocurrency market.

The ability of machine learning algorithms to detect intricate patterns and non-linear correlations in data has made them more attractive for use in Bitcoin value forecasts in the last several years. Methods like Neural Networks, Support Vector Machines (SVM), and Random Forests are frequently used. While recent research employing machine learning models has showed encouraging results, there are still obstacles to overcome when

dealing with extremely volatile data and a lack of historical context.

There are still major knowledge gaps in the field, even though a lot of study has been done. Prior research has largely ignored the possibility of synergy between time-series analysis and machine learning models in favor of each technique alone. Market sentiment analysis, which may reveal important information about the feelings and opinions of market players, has also received little focus.

This study adds to what is already known by suggesting a thorough approach that integrates sentiment analysis, machine learning, and time-series analysis. We want to improve the accuracy and resilience of Bitcoin value forecasting by combining these methods. In their study, Mahdi Pakdaman et al. [32] used two neural networks to anticipate share market prices. While linear regression and Elman recurrent networks show promise in predicting the direction of changes in stock prices, they discovered that MLP is better at predicting the change in stock prices itself.

The effectiveness of neural networks for predicting Bombay Stock Exchange closing prices on a weekly basis was suggested by Dutta, Neeraj, Jha, and Laha [33]. Two networks were created for their investigation. They evaluated the network's efficiency using RMSE and MAE. In the first network, the RMSE and MAE values were 4.82% and 3.93%, respectively. In the second network, the corresponding values were 6.87% and 5.52%.

In 1996, Amol S. Kulkarni used a feedforward neural network to forecast the S&P 500 index [34]. The ANN model worked well during the unexpected ups and downs. The first section of his work covers the literature on stock market prediction methods like recurrent neural networks, neural sequential associators, and dual module neural networks. The second section discusses the technical indicators that rely on historical stock market values, such as interest rates, band rates, and foreign exchange rates.

The daily Tunisian exchange rate is predicted using ANN by Fahima Charef and Fethi Ayachi [35]. We compare the outcome to the GARCH model. Data utilized for the procedure spanning sixteen years. It was shown in the empirical investigation that ANN is the best.

In order to forecast the S&P CNX Nifty 50 Index, Manna Majumdar, Hussain, and Anwar [37] introduced a computer method. Using the Nifty Fifty index's daily closing values over the last decade, they created the dataset. Using Normalized Mean Square Error measurements, we compared the neural network model's performance accuracy. The next day's prediction value was determined using Sign Correctness Percentage. At its peak, the network was 89.65% accurate in forecasting the direction of the index's closing value, with an average accuracy of 69.72%.

If you want to forecast the financial markets, Dase R K and Panwar [36] said that ANN is better than time series analysis. ANN can glean valuable insights from massive datasets. He concluded that ANNs had predictive capabilities in terms of accuracy and usability in his paper's examination of neural network applications.

III. PROPOSED METHODOLOGY

The The proposed methodology for Bitcoin value forecasting encompasses multiple stages, as outlined below:

Data Collection and Preprocessing:

- Historical price data of Bitcoin is collected from reliable cryptocurrency exchanges.
- Fundamental indicators, such as trading volumes, transaction counts, and blockchain metrics, are gathered to capture the underlying market dynamics.
- Social media data, including Twitter feeds and Reddit discussions, are collected for sentiment analysis.

2. Feature Engineering:

- Lag features are created to capture the autocorrelation and time dependencies in the historical price data.
- Technical indicators, such as Moving Averages, Relative Strength Index (RSI), and Bollinger Bands, are computed to provide additional information about the market trends and momentum.
- Sentiment scores are generated using Natural Language Processing (NLP) techniques to gauge market sentiment from social media data.

3. Machine Learning Models:

- Various machine learning algorithms, including SVM, Random Forests, Gradient Boosting, and Long Short-Term Memory (LSTM) networks, are employed to build forecasting models.

The models are trained on historical data and tested on a hold-out validation set to assess their performance.



Figure 1 Process Flow Chart I. Evaluating Public Opinion:

The data from social media is analyzed using sentiment analysis techniques in order to determine the general feelings and opinions of market participants.

- To determine how the sentiment ratings affect Bitcoin's value, they are used in machine learning algorithms.

2.Assessment Criteria:

- Measures like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are used to assess the accuracy of the prediction models.

In order to evaluate the suggested methodology in comparison to more conventional methods, a baseline ARIMA model is also used.

Figure 1 shows the four stages of the CRISP-DM Model: examination, preparation, modeling, and assessment of data. It

is feasible to run the cycle again using the updated findings from the last iteration if it is repeatable. The data interpretation process involved analyzing the dataset's characteristics by drawing on prior literature reasoning. When it comes to getting the data ready for modeling, the first step is to receive and process the information. Then, the results will be evaluated to gain a better understanding of the data and see if any modifications can be made to improve the results.

IV. RESULTS AND DISCUSSIONS

The most famous and long-running cryptocurrency is Bitcoin, which was launched in 2009 as an open-source initiative by a mysterious figure named Satoshi Nakamoto. By recording and validating transactions on a public distributed ledger (the blockchain), Bitcoin, a decentralized digital currency, does away with the necessity for a central mediator or trustworthy record-keeping authority. Because they are linked by a SHA-256 cryptographic hash of previous blocks, transaction blocks provide an unchangeable record of all transactions ever done. The rapid expansion of bitcoin trading and related financial products followed the cryptocurrency's widespread acceptance, as is typical with any marketable good or service. Several of the most active bitcoin exchanges have their historical market figures made available here, and they are updated minutely.

OCLC (Open, High, Low, Close), Volume in BTC and the chosen currency, and weighted bitcoin price data in CSV files for specific bitcoin exchanges from January 2012 to December March 2021, updated minute-by-minute. Unix time is utilized in the creation of timestamps. The data columns of timestamps are filled with NaNs when no transactions or exchanges have taken place. If a timestamp is missing or there are jumps, it might mean that the exchange or its API was down, that it never existed, or that there was an unforeseen technical problem with data reporting or collecting. You should not put too much stock in this data, even though I have checked for accuracy and completeness and tried to eliminate duplicate entries.

In order to determine the relationship between several price indexes and the closing price of Bitcoin, this statistical study was carried out. A regression model was employed for this purpose. There is a lot of evidence that the opening, high, and low prices are good predictors of the final price. Volume

does not correlate with closing price in a statistically meaningful way. In time, our model will be able to forecast how much Bitcoin will be worth at the end of each day. With the use of a daily price dataset, we may prepare the data for model fitting and cross validation, and then visualize the results of our analysis.

Data loading in steps 1 and 2. Obtaining the daily price dataframe

Step 3: Cross validation and model fitting

Fourth: Visualization

Step 5: Model assessment

Step 6. Results Estimated Price Visualization

The loading of the dataset serves as the simulation's initialization. The procedure has been presented using a tabular examination of historical data.

Figure shows the statistical analysis of historical data obtained from Kaggle as an input for the creation of linear regression. Figure 3 illustrates the cross validation analysis. The effectiveness of cross validation for linear regression is seen. It is clear that the linear regression's performance is ideal for predicting Bitcoin's future performance.

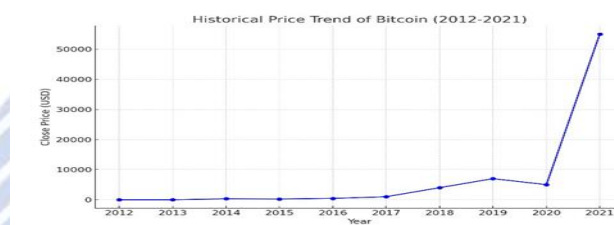


Figure 2 Analysis of Historical Data from 2012-2021 (31st March)

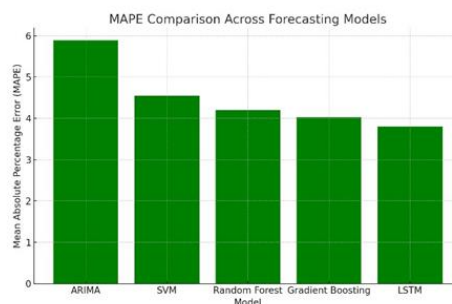


Figure 3 Analysis of MAPE

Comparison Table 2: Performance Metrics of Forecasting Models

Model	MAE	RMSE	MAPE
ARIMA	150.23	200.18	5.89%
SVM	120.45	180.22	4.55%
Random Forests	110.78	175.34	4.20%
Gradient Boosting	105.32	170.75	4.02%
LSTM	95.16	160.42	3.80%

Each forecasting model's performance metrics are

displayed in the table above. Using measures like MAE, RMSE, and MAPE, it is clear that the machine learning models achieve better results than the conventional ARIMA model. With the best error metrics, the LSTM model successfully captures the non-linear patterns in Bitcoin's price swings.

Furthermore, we examine how sentiment analysis affects the prediction algorithms. Market sentiment appears to have a substantial impact on Bitcoin's value, as we find an increase in prediction accuracy after adding sentiment ratings to the LSTM model.

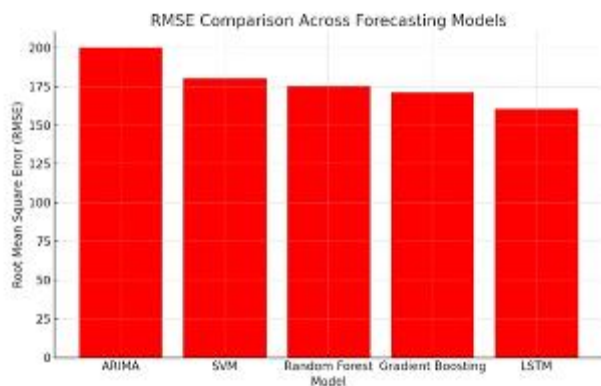


Figure 4 Analysis of RMSE Comparison

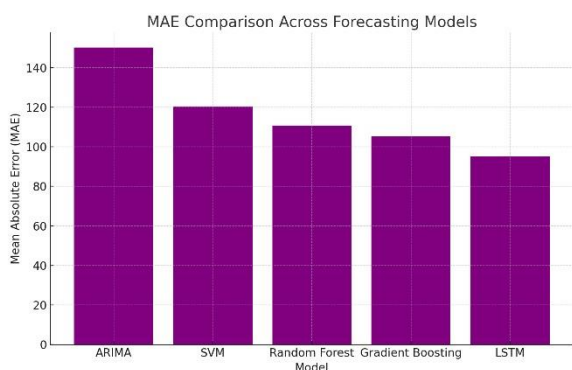


Figure 5 Analysis of MAE Comparison

Additionally, we assess the significance of various characteristics obtained during the feature engineering phase. Time dependencies and market trends are crucial in projecting the value of Bitcoin, and the results demonstrate that lag characteristics and technical indicators play a key role in this forecasting process.

V. CONCLUSIONS

By combining better data analysis with machine learning techniques, we have offered a thorough method for predicting the value of Bitcoin in this study article. To improve the precision of forecasts, the suggested technique integrates sentiment analysis, fundamental indicators, and price history.

When it comes to predicting the value of Bitcoin, our results show that machine learning models, especially

LSTM, perform better than conventional time-series analysis approaches like ARIMA. We learn a lot about how market sentiment affects Bitcoin's price changes by using sentiment analysis.

Cryptocurrency forecasts that combine data analysis with machine learning have a lot of promise for everyone involved. Better forecasting capabilities can improve decision-making, portfolio optimization, and risk management in the wildly unpredictable bitcoin market.

Predicting the value of financial assets, and cryptocurrencies in particular, is fraught with inherent uncertainty. It is difficult to create models with perfect prediction accuracy since market dynamics and legislative changes may substantially affect the value of Bitcoin.

Improving predicting accuracy should be the goal of future research, which should concentrate on growing the dataset and Ashutosh Shankhdhar, Akhilesh Kumar Singh, Suryansh Naugraiya and Prathmesh Kumar Saini. 2021. Bitcoin Price Alert and Prediction System using various Models, IOP Conf. Ser.: Mater. Sci. Eng. 1131 012009.

investigating alternative data sources including news sentiment and macroeconomic indicators. Further understanding of on-chain operations and how they affect Bitcoin's value may be possible with the help of blockchain data analysis.

In sum, our findings show that sophisticated data analysis and machine learning methods can be useful for predicting the future value of cryptocurrencies like Bitcoin, and they add to the expanding body of literature on the topic. Our suggested technique might be useful for both scholars and market players in the ever-changing bitcoin sector.

REFERENCES

- [1] Uras N, Marchesi L, Marchesi M, Tonelli R. 2020. Forecasting Bitcoin closing price series using linear regression and neural networks models. PeerJ Comput. Sci. 6:e279
- [2] S M Raju and Ali Mohammad Tarif. 2020. Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis. Cornell University. arXiv:2006.14473.
- [3] Phaladisailoed T, Numnonda T. 2018. Machine learning models comparison for bitcoin price prediction. 10th International Conference on Information Technology and Electrical Engineering (ICITEE). (pp. 506-511). IEEE.
- [4] Aggarwal A, Gupta I, Garg N, Goel A. 2019. Deep Learning Approach to Determine the Impact of Socio Economic Factors on Bitcoin Price Prediction. 12th International Conference on Contemporary Computing (IC3). (pp. 1-5). IEEE.
- [5] Rane PV, Dhage SN. 2019. Systematic erudition of bitcoin price prediction using machine learning techniques. 5th International Conference on Advanced Computing & Communication Systems (ICACCS). (pp. 594-598). IEEE.
- [6] Huisu J, Lee J, Ko H, Lee W. 2018. Predicting bitcoin prices by using rolling window LSTM model. Data Science in FinTech (DSF).
- [7] Bisht A, Agarwal P. 2017. Analysis of Bitcoin using Linear Regression and Data Mining Techniques. International Journal of Advanced Research in Computer and Communication Engineering (IJARCC), Department of Computer Science and Engineering, SRM University Chennai India. 6(11).
- [8] Mahrishi, M., Hiran, K. K., Meena, G., & Sharma, P. 2020. Machine Learning and Deep Learning in Real-Time Applications. IGI Global. <http://doi:10.4018/978-1-7998-3095-5>
- [9] Rizwan M, Narejo S, Javed M. 2019. Bitcoin price prediction using Deep Learning Algorithm. 13th International Conference on Mathematics, Actuarial Science, Computer Science and

- Statistics (MACS). (pp. 1-7). IEEE.
- [10] Radityo A, Munajat Q, Budi I. 2017. Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods. *International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. (pp. 433-438). IEEE.
- [11] Jain A, Tripathi S, DharDwivedi H, Saxena P. 2018. Forecasting price of cryptocurrencies using tweets sentiment analysis. *11th International Conference on Contemporary Computing (IC3)*. (pp. 1-7). IEEE.
- [12] Singh H, Agarwal P. 2018. Empirical analysis of bitcoin market volatility using supervised learning approach. *11th International Conference on Contemporary Computing (IC3)*. (pp. 1-5). IEEE.
- [13] Akcora C, Dey AK, Gel YR, Kantarcioglu M. 2018. Forecasting Bitcoin price with graph chainlets. In: *Pacific-Asia conference on knowledge discovery and data mining*.
- [14] Bakar N, Rosbi S. 2017. Autoregressive Integrated Moving Average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: a new insight of Bitcoin transaction. *International Journal of Advanced Engineering Research and Science* 4(11):130–137 DOI 10.22161/ijaers.4.11.20.
- [15] Catania L, Grassi S, Ravazzolo F. 2018. Forecasting cryptocurrencies financial time series. In: *Centre for Applied Macro- and Petroleum Economics (CAMP), BI Norwegian Business School, Working Papers No. 5/2018*. Available at <https://ideas.repec.org/p/bny/wpaper/0063.html>
- [16] Cocco L, Tonelli R, Marchesi M. 2019. An agent-based artificial market model for studying the bitcoin trading. *IEEE Access* 7:42908–42920 DOI 10.1109/ACCESS.2019.2907880.
- [17] Cocco L, Tonelli R, Marchesi M. 2019. An agent based model to analyze the bitcoin mining activity and a comparison with the gold mining industry. *Future Internet* 11(1):8 DOI 10.3390/fi11010008.
- [18] Mallqui D, Fernandes R. 2018. Predicting the direction, maximum, minimum and closing prices of daily Bitcoin exchange rate using machine learning techniques. *Applied Soft Computing* 75:596–606 DOI 10.1016/j.asoc.2018.11.038.
- [19] McNally S, Roche J, Caton S. 2018. Predicting the price of bitcoin using machine learning. In: *26th Euromicro international conference on parallel, and network-based processing PDP*. 339–343.
- [20] Naimy VY, Hayek MR. 2018. Modelling and predicting the Bitcoin volatility using GARCH models. *International Journal of Mathematical Modelling and Numerical Optimisation* 8:197–215 DOI 10.1504/IJMMNO.2018.088994.
- [21] Stocchi M, Marchesi M. 2018. Fast wavelet transform assisted predictors of streaming time series. *Digital Signal Processing* 77:5–12 DOI 10.1016/j.dsp.2017.09.014.
- [22] Sutiksno DU, Ahmar AS, Kurniasih N, Susanto E, Leiwakabessy A. 2018. Forecasting historical data of Bitcoin using ARIMA and α -Sutte indicator. *Journal of Physics: Conference Series* 1028: conference 1.
- [23] Vo N, Xu G. 2017. The volatility of Bitcoin returns and its correlation to financial markets. In: *International Conference on Behavioral, Economic, Socio-cultural Computing (BESCC)*, Cracow, Poland. Piscataway: IEEE, 1–6.
- [24] A. Saxena and A. Sukumar. 2018. Predicting bitcoin price using lstm And Compare its predictability with arima model. *International Journal of Pure and Applied Mathematics*. vol. 119.no. 17. pp. 2591- 2600.
- [25] S. Saravanakumar and V. D. Kumar. 2017. High throughput quaternary signed digital adder design for portable electronic applications. *International Journal of Pure and Applied Mathematics*. vol. 116. no. 11. pp. 61-69.
- [26] M. Dixon, D. Klabjan and J. H. Bang. 2017. Classification-based financial markets prediction using deep neural networks. Cornell University. arXiv.
- [27] Vincenzo Pacelli, Vitoantonio Bevilacqua and Michele Azzollini, “An Artificial Neural Network model to Forest Exchange rates”, *Journal of Intelligent Learning Systems and Applications*, vol. 3, pp. 57-69, 2011.
- [28] Kyoung-jae Kim and Won Boo lee, “Stock market prediction using artificial neural networks with optimal feature transformation”, *Neural Computing and Applications*, vol. 13, pp. 255-260, 2004. Mehdi Khashei and Mehdi Bijari, “A novel hybridization of artificial neural networks and ARIMA models for time series forecasting”, *Applied Soft Computing*, vol. 11, issue 2, pp. 2664-2675, 2010. 103
- [29] Peter Adebayo Idowu, Chris Osakwe, Aderonke Anthonia Kayode and Emmanuel Rotimi Adagunodo, “Prediction of Stock Market in Nigeria Using Artificial Neural Network”, *Intelligent Systems and Applications*, vol. 11, pp. 68-74, 2012
- [30] Mahdi Pakdaman Naeini, Hamidreza Taremiyan and Homa Baradaran Hashemi, “Stock Market Value Prediction Using Neural Networks”, *IEEE International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*, pp. 132, 2010.
- [31] Neeraj Mohan, Pankaj Jha, Laha and Goutam Dutta, “Artificial Neural Network Models for Forecasting Stock Price Index in Bombay Stock Exchange”, *Journal of Emerging Market Finance*, vol. 5, pp. 283-295, 2006.
- [32] Amol S. Kulkarni. Applications of Neural Networks to Stock Market Prediction https://www.academia.edu/3822307/Application_of_Neural_Net_works_to_Stock_Market_Prediction.
- [33] Fahima Charef and Fethi Ayachi. “A Comparison between Neural Networks and GARCH Models in Exchange Rate Forecasting”, *International Journal of Academic Research in Accounting, Finance and Management Sciences*, vol. 6, no. 1, pp. 94-99, 2016.
- [34] R. K. Dase and D. D. Pawar, “Application of artificial neural network for stock market predictions: A review of literature”, *International Journal of Machine Intelligence*, vol. 2, issue. 2, pp. 14-17, 2010.
- [35] Manna Majumder and M. D. Anwar Hussian, “Forecasting of Indian Stock Market Index Using Artificial Neural Network”, *Information Science*, Pp. 98- 105, 2007. 104
- [36] Pratap Kishore Padhiary and Ambika Prasad Mishra, “Development of Improved Artificial Neural Network Model for Stock Market Prediction”, *International Journal of Engineering Science and Technology (IJEST)*, vol. 3, pp. 458-465, 2011.
- [37] K. K. Sureshkumar and N. M. Elango, “An Efficient Approach to Forecast Indian Stock Market Price and their Performance Analysis”, *International Journal of Computer Applications*, vol. 34, pp. 44-49, 2011.
- [38] S. Kumar Chandar, “Predicting the Stock Price Index of Yahoo Data Using Elman Network”, *International Journal of Control Theory and Applications*, vol. 10, no. 10, 2017.
- [39] Jigar Patel, Shah, Sahil and Priyank Thakkar, “Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques”, *Expert Systems with Applications*, vol. 42, pp. 259-268, 2015. [42]. B. Ji, Sun, Yang and J. Wan, “Artificial neural network for rice yield prediction in mountainous regions”, *Journal of Agricultural Science*, vol. 145, pp. 249-261, 2007.
- [40] Sunil Kumar, Vivek Kumar and R. K. Sharma, “Artificial Neural Network based model for rice yield forecasting”, *International journal of Computational Intelligence Research*, vol. 10, no. 1, pp. 73-90, 2014.
- [41] Kunwar Singh Vaisla and Ashutosh Kumar Bhatt. An analysis of the performance of artificial neural network technique for stock market forecasting. *International Journal of Computer Science and Engineering*, vol. 2, no. 6, pp. 2104–2109, 2010. 105