



BITCOIN PRICE PREDICTION USING HYBRID MACHINE LEARNING ALGORITHMS

1.D.SWETHA, 2. Dr.M.JANAKI

1.M.Phil Scholar,PG Department of Computer Science,Dr. Umayal Ramanathan College for Women,Karaikudi,Tamilnadu,India

2.Associate Professor,PG Department of Computer Science,Dr. Umayal Ramanathan College for Women, Karaikudi,Tamilnadu,India

Abstract

One of today's most explosive markets is crypto currency also called it as Bitcoin, which has drawn a lot of shareholders interest around the world. Due to the novelty of crypto currency the transaction, there is a great misunderstanding among shareholders. Implementations for predicting the value of bitcoin, Linear Regression (LR) and Random Forest (RF), Hybrid ensemble of machine learning mechanisms have been used. This proposed work aims to forecast Bitcoin values using machine learning techniques for the following day and provide a plan to optimize profits for investors. When compare, the accuracy rate of the proposed hybrid method is reasonably far above the ground; it proved 85 % percent accurate. As a result, it is demonstrated that the hybrid model is superior to the machine learning mechanisms. Accuracy of bitcoin price predictions increased by hybrid ensemble learning.

Keywords

Bitcoin,Machine Learning,Prediction,Linear Regression,Random Forest

Introduction

A crypto currency is digital price projected to take part in for trade uses, and cryptography system added to make safe its trade or business, confirm the secure transmit of belongings, have power over the formation of further crypto currencies [19]. Crypto currencies can be segregated as digital or alternate money, incongruous from habitual currencies in that they have been initiated on the theory of decentralized manage, while match up to central banking systems that common money rely on [20]. The foundation of crypto currencies starts by 2008, Satoshi Nakamoto put out a paper titled Bitcoin: A Peer-to-Peer Electronic Cash System [21] while an unknown thing beneath the pseudonym. Nakamoto unconfined bitcoin software in to the public on 2009 January, as open source code on *SourceForge* [22]. Nakamoto's enthused a prosper of public prying, also stimulates others to craft alternative crypto currencies that based on the identical basic knowledge but to be specific in intention [23].

[25] Stock market is unhinged data offered in provisions of machine learning datasets. Researchers have been doing greatest for a time-consuming period of task to estimate the stock. Crypto currencies have been grown more of traction in latest years since depositors all over the nation. Crypto currency is a innovative carry out for business deal scheme have unswerving to more mystification along with depositors and news or gossip on social media has been state to extensively influence the value of crypto currencies. Bitcoin is the head crypto currencies traded as of at present, in terminology of the quantity being trade. Hence, with the introduction of thousands of latest crypto currencies, bitcoin has a marketplace contribute to more than 55% when compared to additional crypto currencies, followed by Ethereum at 8.57%. This says a lot about the reason of Bitcoin is indispensable stock to forecast. And, Bitcoin value oscillates greatly. It is incredibly occasional and because of these reason bitcoin attained researchers to scan and conjecture the price. This study

stated a variety of tries to prognosticate its price and gives try to cause of outer surface features resembling reports influence its value.

[26] Crypto currencies are decentralized currencies, premeditated to perform as a medium of swap over that uses encryption to shield dealings and limit the formation of new currencies. It has the prospective to permit billions of people to partake in the world of internet retail without having to rely on habitual banking and credit card systems. Although decentralized currencies are hypothetical to prevent operation, some outstanding players are alleged to be by means of social media platforms to boost crypto currency fame and success. Twitter is the most common social media platform for crypto currency whales to persuade and have power over the crypto currency market. The way people use Twitter to articulate their view and mind-set makes the platform an admirable source of sentimental data. Thus, there have been several studies using Twitter for data mining. Furthermore studies have stated that there is a relationship between the behavior of Dogecoin whales and the number of tweets. However, there is no historic affiliation between Litecoin whales and tweets. This study shows that the correspondence between bitcoin price and sentiment is low. It's also to be noted that even though the correlation is low it's not completely random. It improves when a lag is introduced. Hence, Twitter does provide a slight indication to the bitcoin prices.

The remainder of the study is structured as shown below. A literature review is presented in Section 2. The experiments, the outcomes and assessments of the suggested approach on the bitcoin price dataset are reported in Section 4. Section 5 presents a discussion on conclusions.

Literature Review

[1] Here to envisage the bitcoin currency, the system SDAE-B is developed. To study the bond between the varied factors and bitcoin currency, mechanism named Bp, SDAE-B and LSSVM advocated predicting the bitcoin currency. After the profound scrutiny of concealed and precise factors of data, this revise ended conclusion that curb of the recent delve into work in bitcoin value is past data only considered. Predicting the price of bitcoin is a prudent problem. At the all-encompassing level, it reveals the economy course of action of a inhabitants but also reveals the shareholder's choice to attainment and put on the market deal instruments. Matkovskyy and Jalan et al. [2] stated that the exact forecast of bitcoin price not only formulate obtainable assessment sustain system for shareholder but also make offered implication for the administration of government to preparation dogmatic guidelines. Lately, diverse well-liked machine learning mechanisms such as NARX [4], SVM, LSTM, RNN, GA, and ARIMA [3] have been incorporated to forecast the price of bitcoin.

[5] Here, 24 features enclosed huge dataset composed that make known interest rates, macroeconomic variables, exchange rates, furthermore four sustaining variables, 13 crypto currencies, and across the period from 2014 December 2 to 2019 July 8. 239 comments of 5-days frequencies enclosed in this dataset, which is alienated into two different sub dataset, and named as out sample, and out-in sample. Conventional regression system is logistic regression and couple of diverse machine learning mechanisms are Support Vector Machine (SVM) and the Random Forest (RF), have been used, which reveals the prospect of forecast about the getting higher or down currency moves. [6] Due to minimum correlations of financial property, authors have been compared gold and bitcoin currency. Similarly, bitcoin currency also used to deposit in gold for enclose against rise or cost-effective ambiguity, by means of agreement of futures (Bakkt) and exchange of un course of action of crypto currency, such as Huobi, OKex and BitMex [7,8]. The intension of this Bitcoin can have a monotonous deliver, so will not experience from the reduction crisis of paper amount that take place throughout quantitative moderation. Assorted categories of methods applied in an effort to prognosticate he bitcoin value. Amongst for the most part stupendous systems are deep learning chaotic neural networks [13], random forest [9], bayesian neural networks [12], and artificial neural networks [10,11]. In the field of pragmatic research on the Bitcoin market [14–16], unforeseen features such as opinion/emotion have been assisted additionally.

Jaquart et al. [17] too incorporated machine learning systems to prognosticate high frequency of bitcoin prices over the phase March 4 2019 to December 10 2019. Also find out that all veteran systems erect statistically viable forecast, foretelling the twin market evolution by means of varying accuracy from 50.9% to 56.0%. Chen et al. [18] related some systems of machine erudition to prognosticate the high frequency of bitcoin prices. The authors' unruffled data on a daily basis between July 17 2017 and January 17 2018. Wished-for on a daily basis foretelling, they detected that the machine learning achieve 66% truthfulness and statistical mechanism achieves 65% truthfulness, correspondingly. [24] Authors suggested from the outcome of investigated Twitter opinion and volume of trade on bitcoin. Tremendous Gradient Boosting Regression Tree Mechanism provided a viable way of foretelling bitcoin fluctuations contained by market is ZClassic crypto currency.

From the above mentioned literatures can understand that current or accessible examine works on Bitcoin prophecy carried out by machine learning, sentiment analysis and deep learning mechanisms. While many existing research work focuses on sentiment analysis or deep learning methods for prediction of bitcoin, this proposed work tried to fabricate the best hybrid of machine learning algorithms for prognosticate the bitcoin price. This proposed work can smash the current research on value of bitcoin value.

Proposed Methodology

In this proposed work to implement the bitcoin analysis, bitcoin price dataset have been retrieved from bitcoin website for the year 2023. This dataset contains 415 records of features such as Unix, Date, Symbol, Open, High, Low, Close, Volume AAVE, Volume BTC. From which for the purpose of this work, Date, Open, High, Low, Close these features have been taken. Using these five features bitcoin price have been analyzed and threshold value generated, based on the threshold value Logistic Regression and Random Forest algorithms has been utilized to predict the value of bitcoin. Then using these Logistic Regression and Random Forest Hybrid classification model developed for predicts the value of bitcoin. Finally, compared the performance of Logistic Regression, Random Forest and Hybrid model.

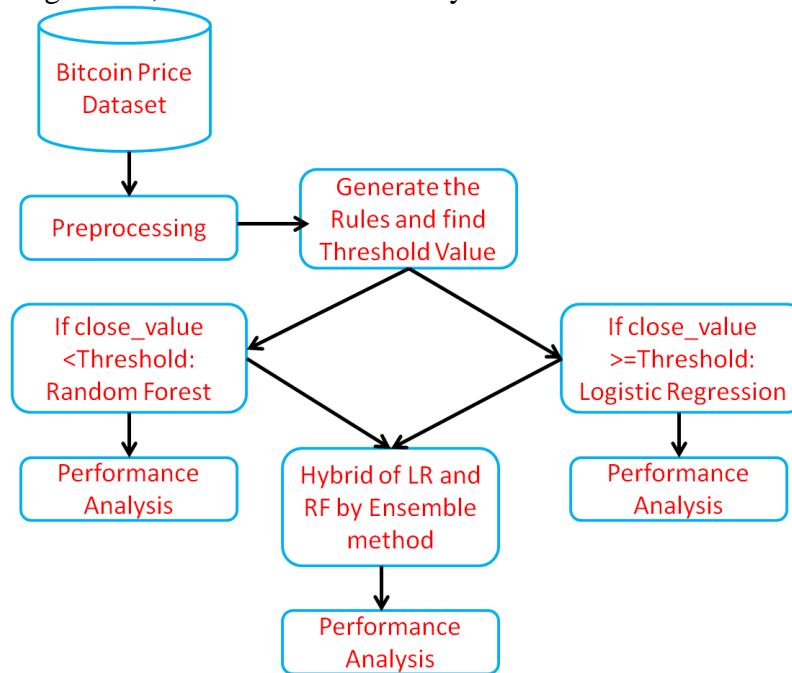


Figure.1: Architecture of the Proposed Work

From the collected bitcoin price only date, open, close, low, and high are considered for the analysis purpose. After, that in the dataset, open, close, low, and high have been grouped based on year. Variance between the open and close, low and high have been calculated. Variance of open and close then low and high are utilized as features for this proposed work. These features have been transformed as vectors by using standard scalar library in python. For the purpose of this work implementation python and their libraries utilized. In this work threshold value is calculated by taking average for variance of open and close. To train and test the utilized machine learning models such as Logistic Regression and Random Forest dataset splitted as training dataset and testing dataset. For the purpose of training 90% of dataset and for the purpose of testing 1% dataset utilized in this study. Logistic Regression and Random Forest are hybridized based on the calculated threshold value 2 different rules have been formed as follows:

- If variance of open and close value is greater than or equal to threshold value then the data is classified by the Logistic Regression (LR)
- If variance of open and close value is less than threshold value then the data is classified by the Random Forest (RF)

For these two kinds of classification performance evaluation has been performed by the metrics such as ROC-AUC value, F1-Score, Recall, Precision and Accuracy. Then by the ensemble voting classifier LR and RF hybridized. For this ensemble classifier performance evaluation performed.

Results and Discussion

The obtained performances of the LR, RF and ensemble classifier are compared by performance analysis values.

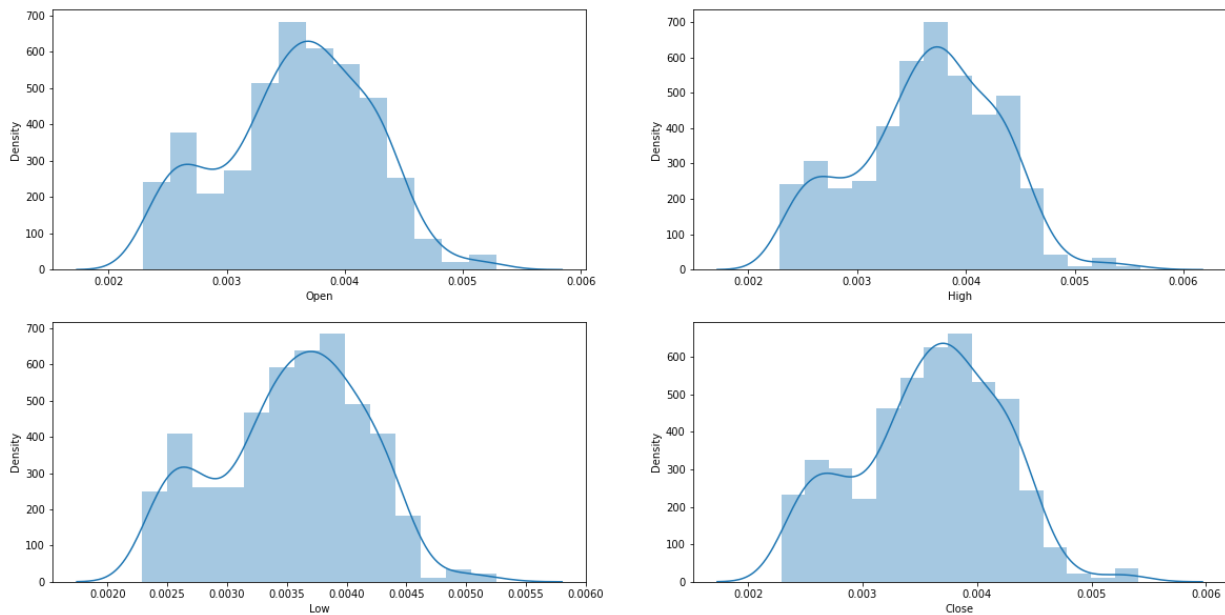


Figure.2. Bitcoin Price

The above figure 2. Demonstrate the real time bitcoin data, such as open, close, low, and high for the year 2022 and 2023.

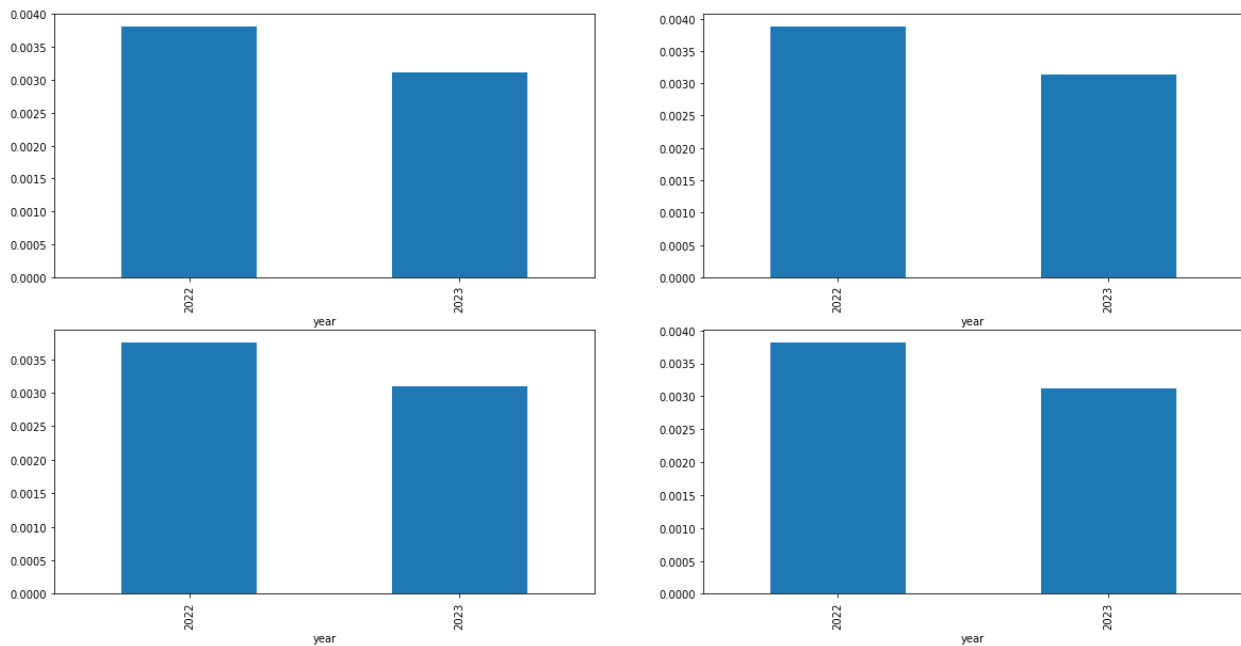


Figure.3. Bitcoin Price Report for 2022 and 2023

Figure.3 demonstrates the bitcoin price report comparison for the year 2022 and 2023.

Table 1: Performance Analysis Report

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
LR	83.20	83.27	83.25	83.20	0.8305
RF	83.72	83.77	83.76	83.69	0.8377
Hybrid	85.75	85.77	85.76	85.77	0.8523

Table.1. illustrates the performance analysis report for the involved machine learning models and hybrid models. Here , ROC-RUC value, F1-Score, Recall, Precision, and Accuracy have been calculated for LR , RF, and Ensemble Hybrid model. LR achieved 83.20 % accuracy, 83.27% precision, 83.25% recall, 83.20% f1-score, and 0.8305 ROC-AUC values. RF achieved 73.72% accuracy, 83.77 % precision, 83.76% recall, 83.69% f1-score, and 0.8377 ROC-AUC values. And Ensemble model provided 85.75% accuracy, 85.77% precision, 85.76% recall, 85.77% f1-score, and 0.8523 ROC-AUC values. From these can understand that RF achieved better results than LR for prediction of bitcoin value. But ensemble model attained highest values than LR and RF. Ensemble is outperformed than LR and RF.

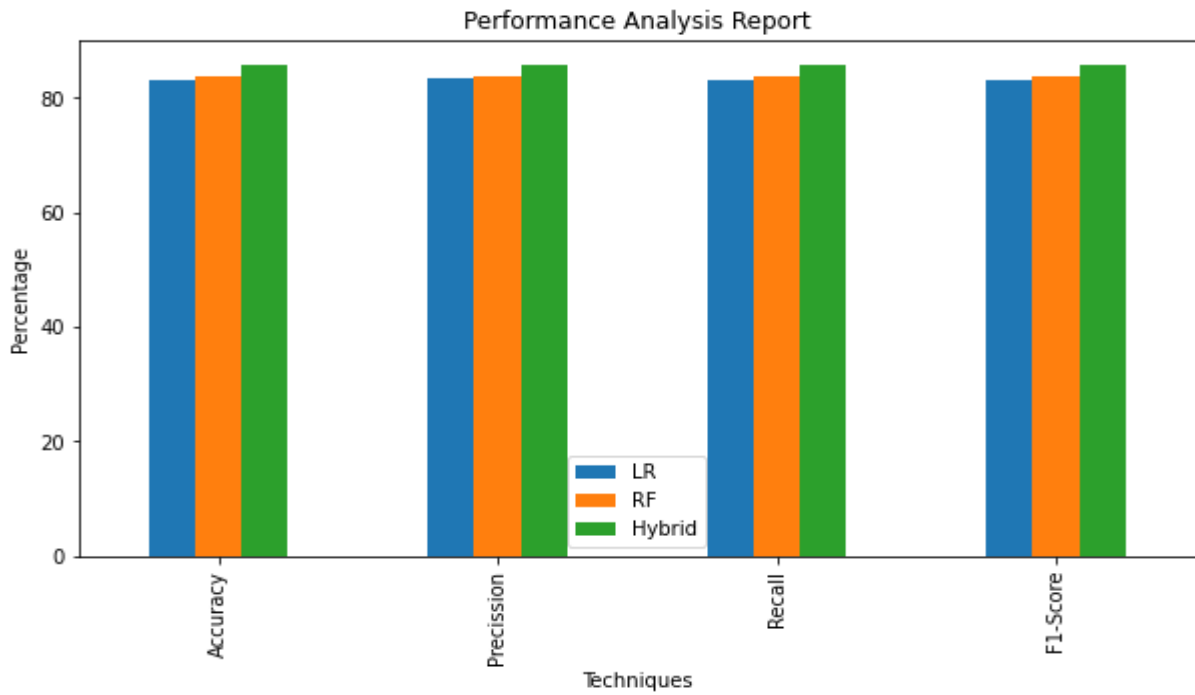


Figure.4. Comparison of Performance Analysis

Figure.4, describes the performance analysis report of LR, RF, and Ensemble models. Shows that ensemble is better than other two classifiers.

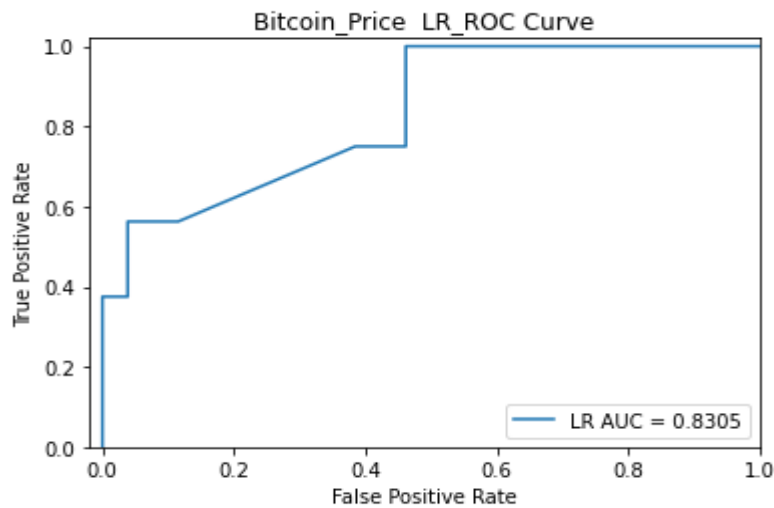


Figure.5. ROC_Curve for Logistic Regression Method

Figure.5. shows the ROC-AUC value and ROC-AUC curve for the model of Logistic Regression on prediction of bitcoin values.

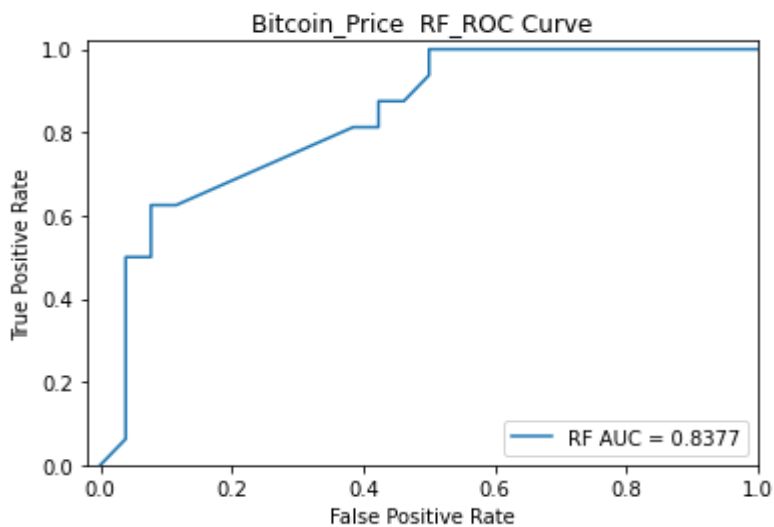


Figure.6. ROC_Curve for Random Forest Method

Figure.6. depicts the ROC-AUC value and ROC-AUC curve for the model of Random Forest on prediction of bitcoin values.

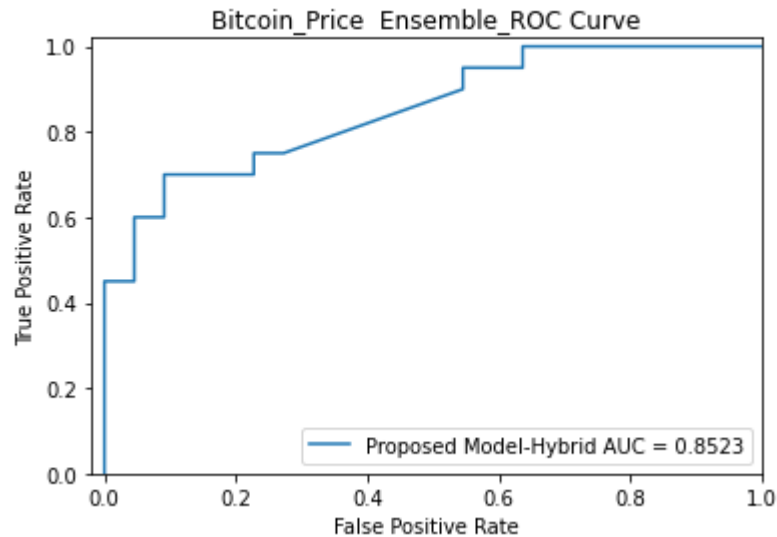


Figure.7. ROC_Curve for Proposed Hybrid Method

Figure.7. describes the ROC-AUC value and curve for the model of Ensemble model on prediction of bitcoin values.

The above shown all results proves that the proposed hybrid model achieved best results than the other two classifiers.

Conclusion

The main purpose that all those who understand bitcoin currently utilize it for is as a reliable store of value and safe haven. Bitcoin is likely to eventually turn into a national strategic reserve resource similar to gold given the ongoing advancements in blockchain technology and the growing acceptance of the currency by nations. Although market professionals place a lot of importance on the price of bitcoin, there hasn't been much theoretical investigation in this area. The existing methods of predicting the price of bitcoin, which simply took into account historical data, so have their flaws to be fixed. Getting data is the first step in empirical research, followed by data processing and model training. Using the training data, we can make an accurate prediction of the price of bitcoin using hybrid ensemble model. The hybrid ensemble model has higher accuracy, than other two models such as Logistic Regression and Random Forest. It states that this hybrid technique can be utilized as an effective tool for predicting the value of bitcoin.

Reference

1. Zhang, S., Li, M., & Yan, C. (2022). The Empirical Analysis of Bitcoin Price Prediction Based on Deep Learning Integration Method. *Computational Intelligence and Neuroscience*, 2022.
2. Matkovskyy, R., & Jalan, A. (2019). From financial markets to Bitcoin markets: A fresh look at the contagion effect. *Finance Research Letters*, 31, 93-97.
3. Shi, S., & Xu, G. (2018). Novel performance prediction model of a biofilm system treating domestic wastewater based on stacked denoising auto-encoders deep learning network. *Chemical Engineering Journal*, 347, 280-290.
4. Han, J. B., Kim, S. H., Jang, M. H., & Ri, K. S. (2020). Using genetic algorithm and NARX neural network to forecast daily bitcoin price. *Computational Economics*, 56, 337-353.
5. Dimitriadou, A., & Gregoriou, A. (2023). Predicting Bitcoin Prices Using Machine Learning. *Entropy*, 25(5), 777.
6. Henriques, I., & Sadorsky, P. (2018). Can bitcoin replace gold in an investment portfolio?. *Journal of Risk and Financial Management*, 11(3), 48.
7. Junttila, J., Pesonen, J., & Raatikainen, J. (2018). Commodity market based hedging against stock market risk in times of financial crisis: The case of crude oil and gold. *Journal of International Financial Markets, Institutions and Money*, 56, 255-280.
8. Tronzano, M. (2021). Financial crises, macroeconomic variables, and long-run risk: An econometric analysis of stock returns correlations (2000 to 2019). *Journal of Risk and Financial Management*, 14(3), 127.

9. Basher, S. A., & Sadorsky, P. (2022). Forecasting Bitcoin price direction with random forests: How important are interest rates, inflation, and market volatility?. *Machine Learning with Applications*, 9, 100355.
10. Adcock, R., & Gradojevic, N. (2019). Non-fundamental, non-parametric Bitcoin forecasting. *Physica A: Statistical Mechanics and its Applications*, 531, 121727.
11. Nakano, M., Takahashi, A., & Takahashi, S. (2018). Bitcoin technical trading with artificial neural network. *Physica A: Statistical Mechanics and its Applications*, 510, 587-609.
12. Jang, H.; Lee, J. An Empirical Study on Modeling and Prediction of Bitcoin Prices with Bayesian Neural Networks Based on Blockchain Information. *IEEE Access* **2018**, 6, 5427–5437. [CrossRef]
13. Jang, H., & Lee, J. (2017). An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *Ieee Access*, 6, 5427-5437.
14. Jain, A., Tripathi, S., Dwivedi, H. D., & Saxena, P. (2018, August). Forecasting price of cryptocurrencies using tweets sentiment analysis. In *2018 eleventh international conference on contemporary computing (IC3)* (pp. 1-7). IEEE.
15. Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65, 101188.
16. Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy*, 21(6), 589.
17. 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.
18. Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, 365, 112395.
19. Greenberg A. Crypto currency. (2011, April). Available online at: <https://www.forbes.com/forbes/2011/0509/technology-psilocybin-bitcoins-gavinandresen-crypto-currency.html> (accessed April 30, 2018).
20. Allison, I. (2015). If banks want benefits of blockchains, they must go permissionless. *International Business Times*.
21. Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*.
22. Davis, J. The crypto-currency: bitcoin and its mysterious inventor. *The New Yorker*. (2014, October). Available online at: <https://www.newyorker.com/magazine/2011/10/10/the-crypto-currency> (accessed April 30, 2018).
23. Liljeström, O. (2019). Efficiency of cryptocurrency exchanges: Risk exposure analysis of identical assets.
24. Li, T. R., Chamrajnagar, A. S., Fong, X. R., Rizik, N. R., & Fu, F. (2019). Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model, arXiv preprint, 2018. *Applied Mathematics Journal of Hindawi* www.hindawi.com, 2018.
25. Velankar, S., Valecha, S., & Maji, S. (2018, February). Bitcoin price prediction using machine learning. In *2018 20th International Conference on Advanced Communication Technology (ICACT)* (pp. 144-147). IEEE.
26. Edgari, E., Thiojaya, J., & Qomariyah, N. N. (2022, March). The impact of Twitter sentiment analysis on bitcoin price during COVID-19 with XGBoost. In *2022 5th International Conference on Computing and Informatics (ICCI)* (pp. 337-342). IEEE.