IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Predicting Bitcoin Prices Using Machine Learning Techniques With Historical Data

Himanshu Sinha
Kelley School of Business
Indiana University, Bloomington
Naperville IL Unites States
0000-0003-4333-7028

***Abstract*—**Cryptocurrency is becoming more popular. The first cryptocurrency to be created was called Bitcoin. One of a most popular cryptocurrencies is bitcoin, and more study is being done on price prediction. Establishing a more plentiful index system and prediction model with an improved prediction effect is required to more accurately and efficiently forecast variations in the price of Bitcoin. Because of this, there is a significant risk for investors investing in Bitcoin because its price fluctuates every second. Interest in trading that is aided by artificial intelligence and machine learning has grown over the last several years. Hence, to forecast Bitcoin's price, it is recommended to use ML technologies. This paper aims to examine machine learning algorithms that have the highest predictive power for Bitcoin prices. This study addresses the critical need for accurate Bitcoin price prediction by integrating machine learning techniques with historical data. Through extensive experimentation and comparative analysis, the Extreme Gradient Boosting Classifier (XGBoost) emerges as the most effective model, achieving an impressive accuracy of 99.08%. This study helps traders in the unpredictable cryptocurrency market make better decisions by using powerful algorithms and preprocessing approaches.

***Keywords*—**Historical bitcoin price dataset, bitcoin, Machine Learning.

I. INTRODUCTION

The Bitcoin network eliminates the requirement for a governing body or intermediary in digital currency transactions by using a decentralised digital currency known as Bitcoin (BTC). Someone or more people using the identity Satoshi Nakamoto established it in 2008. Creating a decentralised, trustless transaction system that could do away with intermediaries like banks was the primary goal of Bitcoin's creators. A public ledger called the blockchain records transactions and uses complex algorithms to keep the system secure. It uses a proof-of-work consensus mechanism to safeguard the network and verify the legitimacy of each transaction before adding it to the blockchain[1].

Worldwide, virtual currencies are gaining popularity and are being utilised for business-company transactions. Reasons for its rising star power include its novel features, which include openness, simplicity, and global acceptability. Here, at the moment, the most well-known and successful virtual money is Bitcoin. Market valuations for virtual currencies fluctuate, but as of April 19, 2019, according to https://bitcoin.org, they are close to 90 billion USD. There is no central authority or bank involved in the Bitcoin network, making it a decentralised, peer-to-peer cryptocurrency. It is really difficult for a third party to mediate disputes between customers. The market price is always rising and falling. The market value of bitcoin is growing with time. More than seventy-one billion dollars are now trading hands on the market. The world's leading virtual currency is Ethereum because it is open-source, transparent, straightforward, and saves time[2][3][4].

Because of Bitcoin's status as a financial asset and its trading on multiple cryptocurrency exchanges (much like a stock market), numerous researchers have used a wide range of analytical and experimental methods to study the factors that influence Bitcoin's price and the patterns underlying its fluctuations (for examples, see the cited works of the authors of). The proliferation of DL-based

Bitcoin price prediction models is a direct result of the current explosion in ML [5][6].

The field of ML is concerned with creating statistical models and algorithms that let computers learn from data and become more efficient at a given activity [7]. Classification, regression, anomaly detection, and NLP are a few of the activities that fall under ML. The use of ML for price prediction may help investors make more educated choices by providing valuable insights into market dynamics. ML algorithms have the ability to identify patterns, trends, and correlations that human analysts might overlook via the use of historical data analysis[8][9]. ML models may use a range of methods, such as time series analysis, to estimate future price movements and collect trend and momentum information [10][11]. The goal of this study is to estimate how accurately ML expertise can be used to anticipate the rate of cryptocurrencies like Bitcoin.

A. Contribution of the Study

By presenting dimension engineering on Bitcoin price granularity—a unique strategy not before investigated—this work significantly advances the area of Bitcoin price prediction. The research outperforms conventional approaches like Logistic Regression, SVM, and Random Forest by using machine learning techniques like Extreme Gradient Boosting Classifier (XGBoost) to obtain outstanding F1-score, recall, accuracy, and precision. By giving traders a trustworthy binary classification algorithm, the suggested technique enables traders to make well-informed decisions based on sign changes in Bitcoin prices. All things considered, this study advances our knowledge of and ability to use machine learning to forecast cryptocurrency values, providing insightful information to those involved in the financial markets.

B. Structure of the study

The following is the structure of this research for the sections that follow: The literature that is pertinent to our investigation is examined in Section II. the methodological analysis of the study; the research approach utilised in this study is presented in Section III. Discuss a findings and evaluations of the study project in Section IV. The conclusion of this research and future goals are provided in Section V.

II. LITERATURE REVIEW

In this section, provide some of the related studies based on ML and DL methods to predict Bitcoin prices, also discussed this related work in Table 1.

In, Kazeminia, Sajedi and Arjmand, (2023) CNN, LSTM, and GRU () are deep learning algorithms that were utilised to automate the collection of Bitcoin datasets for the model's training. The R2 error and MAPE were 0.98166 and 0.034, respectively, using the suggested model. This model is also capable of performing real-time forecasting[12].

In, Freeda, Selvan and Hemanandhini, (2021) suggested method use DL to improve the accuracy of the data. Training and testing a RNN model using the existing dataset to achieve a long-term prediction is the innovative aspect of this study. The value of bitcoin in 2021 is forecasted in this paper. The presented study demonstrates an enhanced accuracy of 76.99% utilising the RNN model in comparison to other ML methods such as RF, GNB, SVM, and KNN techniques[13].

In, Jakubowicz and Abdelfattah, (2021) was to as opposed to particular time series price projections, ascertain if more accuracy may be attained by concentrating on a more comprehensive view of numerical ranges. Reporting the anticipated direction of the market for the next hour was the main focus of the forecasts. Five distinct ML models were tested and trained using one-hour interval trade data to create discrete classes of degrees of hourly changes. Cross-validation accuracy was attained between 96 and 100%, with the exception of one model[14].

In, Dinshaw, Jain and Hussain, (2022) three ML models—LSTM, ARIMA, and SARIMA—have been used. These models are statistically examined based on performance metrics, such as RMSE, MAPE, and coefficient of determination (R2). The results indicate that the LSTM model has computed RMSE, MAPE, and R2 values of 1447.648, 3.059%, and 0.9702, respectively; the ARIMA model has calculated values of 1288.5, 3.479%, and 0.9566, and the SARIMA model has computed values of 1802.31, 4.665%, and 0.9505[15].

In, Vestly et al., (2023) was going to inculcate the neural network models like sequence-to-sequence, ARIMA, and LSTM. This process is certainly made to analyse the bitcoin price over the past couple of years. Trained our model accordingly and found out 96.21% accuracy for the sequence-to-sequence and 94.2% for the ARIMA as well. Eventually this paper and its work insinuate regarding the highly accurate performing algorithm and its analysis as well[16].

In, Saadatmand and Zare Chahoki, (2023) intended to provide a novel method for predicting Bitcoin trends using DL algorithms. by monitoring the prior price and doing a sentiment analysis on data that was taken from Twitter. For this study, information was gathered from January 2012 to December 2020. This article examines the trends in Bitcoin price fluctuations using the LSTM, Bi-LSTM, GRU, and Bi-GRU algorithms. The Bi-GRU algorithm outperformed with a record 72% accuracy in forecasting the trend of fluctuations in the price of Bitcoin and a 20% increase in learning speed[17].

TABLE I. COMPARATIVE ANALYSIS OF BITCOIN PRICE PREDICTION USING VARIOUS METHODS AND TECHNIQUES

Ref	Methodology	Data set	Performance	Limitations & future work
Kazeminia, Sajedi and Arjmand	Hybrid model using automated web scraping, compared with CNN, LSTM, GRU	Gathered through automated web scraping technique	R2: 0.98166, MAPE: 0.034, Outperformed CNN, LSTM, GRU	Real-time forecasting capability and scalability to larger datasets.
Freeda, Selvan and Hemanandhini	RNN model trained on time series data, compared with Random Forest, Gaussian Naïve Bayes, SVM, KNN	Time series data for Bitcoin prices	Accuracy: 76.99% using RNN model	Long-term prediction validation, comparison with more recent models
Jakubowicz and Abdelfattah	Five ML models predicting market direction for the next hour	One-hour interval trading data, discrete classes of hourly changes	Cross-validation accuracy: 96-100%	Assessment of model robustness, extension to other time intervals
Dinshaw, Jain and Hussain	Comparison of LSTM, ARIMA, SARIMA models based on RMSE, MAPE, R2	Time series data from January 2012 to December 2020	LSTM: RMSE - 1447.648, MAPE - 3.059%, R2 - 0.9702	Investigation into model generalisation, inclusion of more features
Vastly	Employing sequence-to-sequence, ARIMA, LSTM models for Bitcoin price analysis	Bitcoin price data from past years	Sequence-to-sequence: 96.21%, ARIMA: 94.2%	Exploration of ensemble methods, analysis of model robustness
Saadatmand and Zare Chahoki	Sentiment analysis of Twitter data combined with previous price tracking, comparing LSTM, Bi-LSTM, GRU, Bi-GRU	Twitter data and Bitcoin price data from January 2012 to December 2020	Bi-GRU: 72% accuracy, 20% improvement in learning speed	Refinement of sentiment analysis, exploration of other social media platforms

A. Research gaps

The integration of numerous data sources and approaches to improve forecast accuracy and robustness is a significant research gap in the existing literature on Bitcoin price prediction. This gap is remarkable since it is a significant research gap. Although the current research makes use of a variety of ML and DL techniques, like LSTM, GRU, and sentiment analysis, it is common for these studies to concentrate on a single data source or method, so ignoring the potential synergies that may be achieved by integrating several approaches. Furthermore, there has been a limited amount of investigation into the possibility of combining fundamental indicators or macroeconomic aspects with technical research. This might result in a more thorough knowledge of the dynamics of the Bitcoin price. Improving prediction performance and enabling more informed investment choices in the volatile cryptocurrency market involves multidisciplinary cooperation and the creation of hybrid models that use the capabilities of diverse data sources and analytic approaches. This is necessary in order to bridge the research gap that exists.

III. RESEARCH METHODOLOGY

The methods used to develop the suggested Bitcoin price prediction system are covered in this part, along with how several ML and DL approaches were implemented.

A. Methodology

Using databases of past Bitcoin values, this research focuses on making price predictions for the cryptocurrency. Both model training and model assessment are built on top of the chosen dataset.

Starting with the preprocessing of the data, the technique removes punctuation, handles null values, and normalises the data to guarantee consistency. Subsequently, the dataset is split into training and testing sets with a 75:25 ratio to facilitate robust model evaluation. Deep learning approaches such as XGBoost, LR, SVM, and RF are employed for model implementation due to their effectiveness in handling complex financial data. Model performance evaluation includes metrics like accuracy, precision, recall, and F1-score, ensuring a comprehensive assessment of the chosen classification methods for Bitcoin price prediction. Comparative analysis is conducted to validate the efficacy of the selected model, providing insights into its suitability for accurate and reliable Bitcoin price forecasting in the financial domain. Figure 1 shows the suggested architecture. Here are in-depth explanations of each stage:

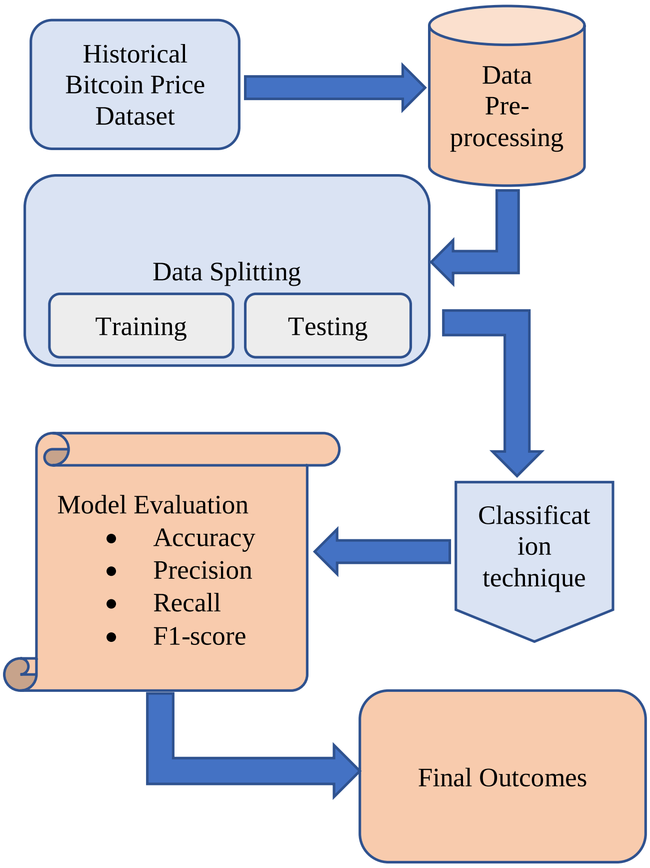

Fig. 1. Proposed flowchart for bitcoin prediction

1) Data collection

The goal of the Kaggle dataset is to find the best machine learning-based method for Bitcoin price prediction.

2) Data preprocessing

Data categorisation requires a critical step called preprocessing. For any ML to be testable and clean, preprocessing of the data is an essential step. To eliminate null values, punctuation, abbreviations, and other similar artefacts, preprocessing processes were applied to the selected dataset. We also changed the data's timestamps into widely used time formats that are easier to comprehend. Furthermore, a Min-Max scalar was used for normalisation in order to scale the values between 0 and 1. (The following are some of the stages in data preprocessing:

- **Data Cleaning:** Data cleaning involves computing the value in a filled-in format while ignoring the missing value.
- **Null values:** In data analysis and programming, handling null values correctly is essential since improper handling may result in mistakes or inaccuracy in computations or predictions.
- **Punctuation:** One important task in text normalisation is the removal of unnecessary and special characters including punctuations. The main reason for doing so is because punctuations tend to throw the model off instead of helping them predict the correct class.
- **Normalization:** To scale numerical characteristics to a specified range, usually between 0 and 1, normalisation employing Min-Max scaling is a method that is often used in machine learning and data preparation. It's especially helpful when characteristics need to be brought to a common scale in order for algorithms to function properly, since they have distinct sizes.

3) Data Splitting

The dataset was split into two halves, with 1,646 samples used for training and 548 samples used for testing, in accordance with a 75:25 ratio.

4) Classification techniques

For the purpose of model implementation, choose machine learning techniques. In this study, choose extreme gradient boosting classifier for the bitcoin price prediction. The model is described in below:

a) Extreme gradient boosting Classifier:

The open-source gradient boosted trees approach known as XGBoost is both popular and strong. One supervised learning method that aims to increase prediction accuracy by integrating the results of many less complex models is gradient-boosting [18].

XGBoost is a library and framework that allows gradient-boosted trees to develop in a forest in parallel [19]. Gradient boosting decision trees (GBDTs) are useful since they strive to minimise the time necessary to build trees and speed up the optimisation process. In order to provide a robust prediction, a GBDT pooles the output of several weak classifiers. Due to the extensive usage of regression functions to estimate each tree, it is an enhanced decision tree model.

The classification error may be reduced step-by-step by attempting to improve the preceding tree's residual classification. The ideal split is determined by calculating the structure's scores and gain after each tree has been optimally approximated. The final outcome of the model's predictions is the total of all decision trees.

b) Logistic Regression Model

The classic multiple-variable regression technique known as logistic regression (LR) has applications in binary classifications [19][20]. The values of the features xi, where i = 1,..., K, predict the class label, which is represented by the binary response variable yi ∈ {0, 1}. One way to represent the LR model is as log it equ.1.

$$(P\,(y\,=\,1)) = log(\,P\,(y\,=\,1)1\,-\,P\,(y\,=\,1)) = \beta 0\,+\,\beta 1x1\,+\,\beta 2x2\,+\cdots+\beta kxK \quad (1)$$

where βi stands for the regression coefficient of xi and P(y = 1) denotes the likelihood of the sample being in class 1.

c) SVM Linear kernel

Data analysis, pattern identification, classification, and regression analysis are some of the many applications of SVMs, a similar supervised learning technology. Applying the statistical theory of learning, it is a kind of learning approach [21][22]. The goal is to maximise the breadth of the separation among the two groups by dividing the data points along a hyperplane, which in this instance is a line. In order to do this, we use training tuples and margins, also known as support vectors, in an effort to lower the classification error. In formal terms, this is equal to 2:

$$P(x) = w.x + b = 0 \qquad (2)$$

Here, P(x) is the set of points, w and x are the support vectors used to generate hyperplane between two groups and 'b' is the coefficient.

d) Random Forest

The RF classifier and other supervised learning algorithms blend the output of many decision trees using a bagging approach [23][24]. A random forest classifier is used with a uniform distribution of training data across classes. To solve the classification problem, we have calculated the cost of splitting the dataset using the Gini index. We used a grid search to test out different values for a number of parameters, or the amount of estimators, to train our RF model to provide reliable predictions. Numerous methods exist for partitioning decision trees, each optimised for a particular classification or regression task. After tallying up the squared probabilities for each group, take one off the total to get the Gini index. The Gini index (Gind) may be calculated using the formula 2 that follows[25][26]:

$$G_{ind} = 1 - \sum_{i=1}^{c} P_i)^2 \qquad (3)$$

To train DTs, first take a bootstrapped sample of the training data, and then for each tree, randomly choose a subset of features to use in its training. Each node of the trees is used to recursively split the data into smaller subsets according to whatever characteristic yields the best split. Every DT in the forest determines its own prediction for the label of a new input during the prediction phase. The final forecast is then determined by summing together the predictions from each tree and casting a majority vote. This method reduces the overfitting issue that DTs are known to have while improving the model's generalisability. Furthermore, it is an effective technique for classification jobs because it employs several DTs, which together might identify crucial properties that a single tree could have overlooked.

5) Model Building

Machine learning methods should be used for model building. Provided a feature importance approach that could be used to build a detection system that would be helpful in its functioning in the environment. In addition to an algorithm-based feature selection method, the proposed machine learning-based model incorporates different ratios of the sampling procedure for use in the actual world. Use methods for more accurate categorisation as well as machine learning-based techniques that make use of certain characteristics. To confirm that the proposed diabetes prediction model is successful, the validation phase is carried out in this research for every step.

6) Model Evaluation

A wide range of performance metrics are used to assess the effectiveness of Bitcoin price prediction systems. Use f1-score, recall, accuracy, and precision in your model assessment.

B. Proposed Algorithm

In this section, provide the proposed algorithm to be used to build the bitcoin price prediction system.

Proposed Algorithm: For Bitcoin price prediction
Step 1: Data Collection
• Collect the Bitcoin price dataset from Kaggle.
Step 2: Data Preprocessing
• Preprocess the data for data cleaning, remove null values, punctuation and normalisation.
Step 3: Data Splitting
• Split the data into training and testing.
Step 4: classification techniques
• Apply extreme gradient boosting classifier.
Step 5: Model Training
• Train the model with preprocessed data
Step 6: Model Evaluation
• Use some performance parameters like accuracy, precision, recall and f1-score.

Step 7: predict outcome
Finish!!!

IV. RESULTS & DISCUSSIONS

Provide the findings from the assessment of the dataset used for this study, together with the classifier's statistics, in this section.

A. Dataset Description

The CSV files of the chosen historical Bitcoin price dataset offer minute-by-minute updates on a range of metrics, such as OHLC (Open, Low, High, Close) prices, BTC trading volume, currency information, and the weighted Bitcoin price, which consisted of 3,161,057 rows. The datasets were chosen because they had undergone the least amount of modification.

TABLE II. FEATURES DETAILS OF THE HISTORICAL BITCOIN PRICE DATASET

Features	Definition
Open	Opening trade
Close	Latest trade
Low	Lowest trade during day
High	Highest trade during day
Weighted price	Mean Bitcoin price
Volume_(BTC)	Total trade volume of day in USD
Volume_(Currency)	Total trade volume of day in USD
Timestamp	Date recorded time

B. EDA (Exploratory Data Analysis)

Important first steps in data analysis include exploratory data analysis (EDA), which makes use of summary statistics and visual representations to find patterns, identify outliers, test hypotheses, and confirm assumptions [27].

Fig. 2. Correlation feature heatmap

Above, figure 2 displays a heatmap graph of the Bitcoin price dataset. In the graphical representation x-axis shows the close, high, low, open and price attributes and y-axis represents the same. Heatmap achieves the score of 100% diagonally.

C. Performance Measures

Evaluation metrics may be used to calculate a model's performance. One important aspect of assessment measures is their capacity to distinguish between various model outcomes. Confusion matrix, accuracy score, precision score, recall score, f1 score, and ROC metrics will all be employed in this research to assess the suggested employee attrition algorithm.

1) Confusion Matrix

Figure 3 illustrates how the confusion matrix, which is used to gauge the introduction of two class issues for the given instructional record, works. True positive (TP) and true negative (TN) parts describe instances correctly from corner to corner, but false positive (FP) and false negative (FN) parts incorrectly request instances. Matrix of Confusion Appropriately Categorise Example TP + TN Misclassify Instances.

		Actual Values	
		Positive(1)	Negative(0)
Predicted Values	Positive(1)	TP	FP
	Negative(0)	FN	TN

Fig. 3. Representation class of Confusion matrix

Four distinct columns are formed when comparing the actual values with the predictions: true negative (TN), false positive (FP), false negative (FN), and true positive (TP). For example, it is considered a false positive if an instance was expected to have diabetes but did not. Accuracy, recall, and precision are computed using values from the confusion matrix.

2) Accuracy

The degree of accuracy is the degree to which the value is near to its true value. Trueness and precision are used to calculate accuracy (the degree to which measurement findings are close to the real value). It is given as eq.4-

$$Accuracy = \frac{TP+TN}{TP+Fp+TN+FN} \quad (4)$$

3) Precision

Precision is a statistical metric for gauging the variation in results from a set of trials. As a metric, it tracks the proportion of positive predictions that lead to accurate outcomes. It is expressed as eq.5-

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

4) Recall

Recall is a metric for accurately determining which positive results were really produced. It is recommended to look at both recall and precision when evaluating a technique's efficacy. In mathematical form it is given as eq.6-

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

5) F1 score

The F-score may alternatively be expressed as the geometric mean of the Diversity and Coverage metrics, with weights applied. Integrating recall and precision yields the F1-score. Mathematically, it is given as eq.7-

$$F1 - Score = \frac{2(Precision*Recall)}{Precision+Recall} \quad (7)$$

D. Experimental results

This section's analysis focuses on the outcomes of an experiments that were carried out to anticipate a price of bitcoin. Python is a computer language used to show the inquiry outcomes. The following are the outcomes of the experiment:

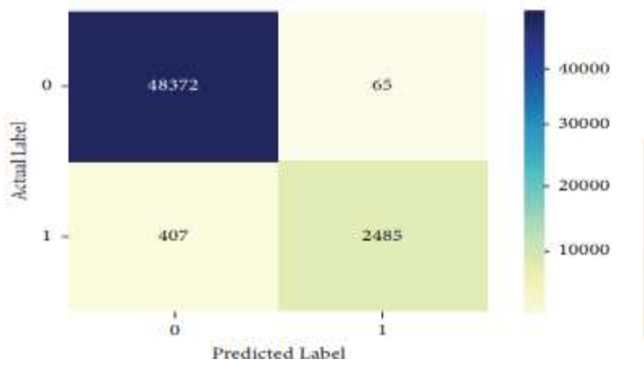

Fig. 4. XGBoost confusion matrix

The above figure 4 represent the Confusion Matrix for XGBoost where x-axis is predicted label and y-axis is Actual Label. Class 1 represent the value of false negative (407) and true positive (48372). Class 0 represent the value of true negative (2485) and false positive (65).

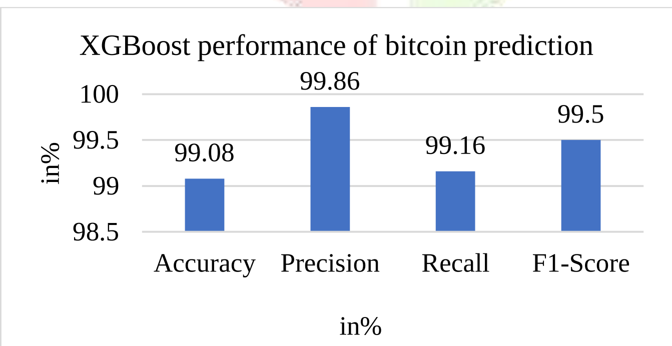

Fig. 5. Parameters performance of XGBoost for bitcoin

Figure 5 displays the performance of the XGBoost model for bitcoin prediction according to F1-score, recall, accuracy, and precision with 99% performance.

E. Comparative analysis

In this section, present the comparative result analysis of various machine performance according to F1-score, recall, accuracy, and precision. In this study, comparison between the model, shows in Table 3.

TABLE III. COMPARATIVE RESULTS OF MODEL PERFORMANCE

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression[12]	0.66667	0.538462	0.411765	0.466667
SVM[12]	0.604167	0.416667	0.294118	0.344828
Random Forest[12]	0.5625	0.409091	0.529412	0.461538
XGBoost	99.08	99.86	99.16	99.5

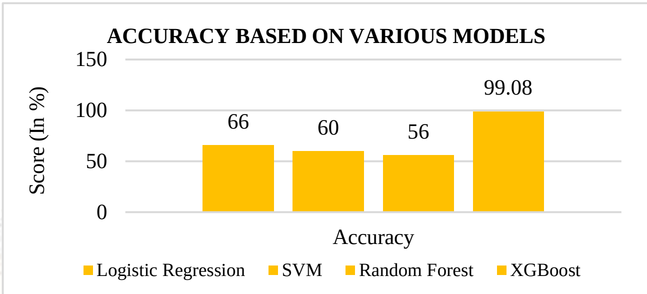

Fig. 6. Performance measures based on models Accuracy

Figure 6 displays several models' accuracy. With the greatest accuracy score of models, 99.08%, achieved by XGBoost. With XGBoost showing a little superior performance in this particular category.

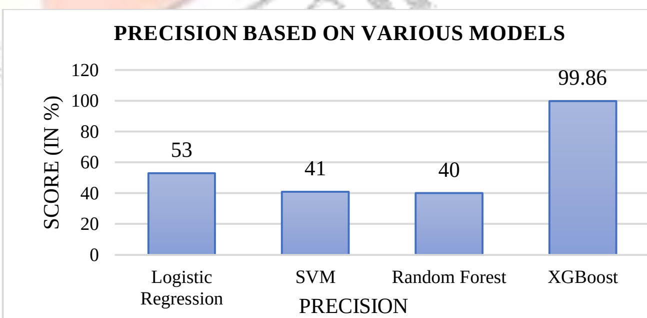

Fig. 7. Performance measures based on models' precision

Figure 7 shows the precision of models. At 99.86%, XGBoost achieved the highest precision score. Out of all the models in this specific category, the XGBoost model has the best performance.

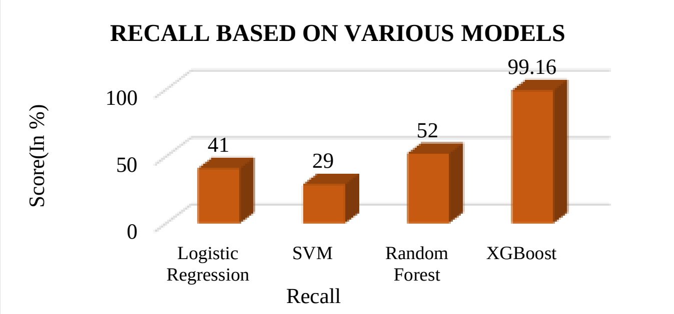

Fig. 8. Performance measures based on models recall

Figure 8 shows the several models' recall. XGBoost achieved the greatest recall score, getting in at 99.16%. The XGBoost model performs the best among all the models in this particular category.

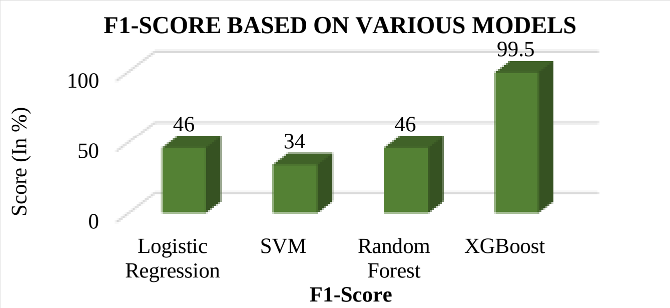

Fig. 9. Performance measures based on models f1-score

Figure 9 displays the f1-scores for the various models. With 99.5%, XGBoost obtained the highest f1-score. Of all the models in this specific category, the XGBoost model performs the best.

A performance analysis of different ML models for bitcoin prediction discussed in above, highlights the superiority of the XGBoost model across multiple metrics. The confusion matrix demonstrates a high TPR (48,372) and a low FPR (65) for XGBoost. When comparing models, XGBoost outperforms Logistic Regression, SVM, and Random Forest, achieving the highest scores in accuracy (99.08%), precision (99.86%), recall (99.16%), and F1-score (99.5%). These results underscore XGBoost's exceptional performance, making it the most effective model for Bitcoin prediction in this study due to its ability to balance precision and recall, ultimately leading to a superior F1 score.

V. CONCLUSION AND FUTURE WORK

This study shows an effectiveness of ML techniques, particularly the Extreme Gradient Boosting Classifier (XGBoost), in accurately predicting Bitcoin prices. By leveraging historical Bitcoin price datasets and employing advanced algorithms, the XGBoost model achieved remarkable F1-score, recall, precision, and accuracy, outperforming other traditional classification methods like LR, SVM, and RF with 99%. The high predictive power of the XGBoost model underscores the potential of machine learning in forecasting cryptocurrency prices, thereby assisting traders in making informed investment decisions in the volatile cryptocurrency market. Future work in this area could focus on further enhancing the predictive capabilities of machine learning models by incorporating additional data sources, such as macroeconomic indicators, social media sentiment analysis, and market news. Moreover, exploring ensemble learning techniques and hybrid models that combine multiple algorithms could potentially improve prediction accuracy and robustness. Additionally, conducting research on real-time prediction models that adapt to dynamic market conditions would be valuable for practical applications in cryptocurrency trading. By continually refining and innovating machine learning approaches, researchers can contribute to the development of more effective tools for cryptocurrency price forecasting and risk management.

REFERENCES

[1] M. Frohmann, M. Karner, S. Khudoyan, R. Wagner, and M. Schedl, “Predicting the Price of Bitcoin Using Sentiment-Enriched Time Series Forecasting,” *Big Data Cogn. Comput.*, vol. 7, no. 3, 2023, doi: 10.3390/bdcc7030137.

[2] T. Awoke, M. Rout, L. Mohanty, and S. C. Satapathy, *Bitcoin Price Prediction and Analysis Using Deep Learning Models*, vol. 134, no. October. Springer Singapore, 2021. doi: 10.1007/978-981-15-5397-4_63.

[3] M. K. Salman and A. A. Ibrahim, “Price prediction of different cryptocurrencies using technical trade indicators and machine learning,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 928, no. 3, 2020, doi: 10.1088/1757-899X/928/3/032007.

[4] J. Thomas, “Optimizing Bio-energy Supply Chain to Achieve Alternative Energy Targets,” pp. 2260–2273, 2024.

[5] S. Ji, J. Kim, and H. Im, “A comparative study of bitcoin price prediction using deep learning,” *Mathematics*, 2019, doi: 10.3390/math7100898.

[6] A. S. Shaista Naz, Humera Amin, “Maternal Mortality in Pakistan: The Potential Role of Community Midwives,” *J. Dev. Soc. Sci.*, vol. 5, no. 2, pp. 45–52, 2024.

[7] S. Mathur., “Supervised Machine Learning-Based Classification and Prediction of Breast Cancer,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 12(3), 2024.

[8] S. A. Pranav Khare, “THE IMPACT OF MACHINE LEARNING AND AI ON ENHANCING RISK-BASED IDENTITY

VERIFICATION PROCESSES," no. 05, pp. 1–6, 2019.

[9] S. Arora and S. R. Thota, "Using Artificial Intelligence with Big Data Analytics for Targeted Marketing Campaigns," no. June, 2024, doi: 10.48175/IJARSCT-18967.

[10] R. K. Vinita Rohilla, Sudeshna Chakraborty, "Deep learning based feature extraction and a bidirectional hybrid optimized model for location based advertising," *Multimed. Tools Appl.*, vol. 81, no. 11, pp. 16067–16095.

[11] A. El Youssefi, A. Hessane, I. Zeroual, and Y. Farhaoui, "Utilizing Machine Learning and Deep Learning for Predicting Crypto-currency Trends," *Salud, Cienc. y Tecnol. - Ser. Conf.*, 2024, doi: 10.56294/sctconf2024638.

[12] S. Kazeminia, H. Sajedi, and M. Arjmand, "Real-Time Bitcoin Price Prediction Using Hybrid 2D-CNN LSTM Model," in *2023 9th International Conference on Web Research, ICWR 2023*, 2023. doi: 10.1109/ICWR57742.2023.10139275.

[13] S. E. Freeda, T. C. E. Selvan, and I. G. Hemanandhini, "Prediction of Bitcoin Price using Deep Learning Model," in *Proceedings of the 5th International Conference on Electronics, Communication and Aerospace Technology, ICECA 2021*, 2021. doi: 10.1109/ICECA52323.2021.9676048.

[14] E. Jakubowicz and E. Abdelfattah, "The Rise and Fall of Bitcoin: Predicting Market Direction Using Machine Learning Models," in *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2021*, 2021. doi: 10.1109/IEMCON53756.2021.9623119.

[15] C. Dinshaw, R. Jain, and S. A. I. Hussain, "Statistical Scrutiny of the Prediction Capability of Different Time Series Machine Learning Models in Forecasting Bitcoin Prices," in *Proceedings of 4th International Conference on Cybernetics, Cognition and Machine Learning Applications, ICCCMLA 2022*, 2022. doi: 10.1109/ICCCMLA56841.2022.9989057.

[16] D. J. Vestly, N. Thangallapelly, G. S. Manish, S. Hariharan, V. Kukreja, and J. PR, "Bitcoin Price Analysis With Deep Learning Models," in *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, 2023, pp. 419–424. doi: 10.1109/ICAICCIT60255.2023.10465994.

[17] F. Saadatmand and M. A. Zare Chahoki, "Time Series Analysis by Bi-GRU for Forecasting Bitcoin Trends based on Sentiment Analysis," in *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2023, pp. 323–328. doi: 10.1109/ICCKE60553.2023.10326259.

[18] T. Islam, A. Vuyia, M. Hasan, and M. M. Rana, "Cardiovascular Disease Prediction Using Machine Learning Approaches," in *Proceedings of International Conference on Computational Intelligence and Sustainable Engineering Solution, CISES 2023*, 2023. doi: 10.1109/CISES58720.2023.10183490.

[19] Z. Chen, C. Li, and W. Sun, "Bitcoin price prediction using machine learning: An approach to sample dimension engineering," *J. Comput. Appl. Math.*, 2020, doi: 10.1016/j.cam.2019.112395.

[20] V. Rohilla, S. Chakraborty, and R. Kumar, "Car Auomation Simulator Using Machine Learning," *SSRN Electron. J.*, 2020, doi: 10.2139/ssrn.3566915.

[21] G. R. Kumar, G. A. Ramachandra, and K. Nagamani, "An Efficient Prediction of Breast Cancer Data using Data Mining Techniques," *InternationalJournalofInnovationsinEngineeringandTechnology(IJIET)Vol.2Issue4August2013139SSN:2319-1058*, 2013.

[22] M. H. Hashmi, M. Affan, and R. Tandon, "A Customize Battery Management Approach for Satellite," in *2023 24th International Carpathian Control Conference (ICCC)*, IEEE, Jun. 2023, pp. 173–178. doi: 10.1109/ICCC57093.2023.10178893.

[23] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," *J. Supercomput.*, 2021, doi: 10.1007/s11227-020-03481-x.

[24] P. Khare, "AI-Powered Fraud Prevention : A Comprehensive Analysis of Machine Learning Applications in Online Transactions," vol. 10, no. 9, pp. 518–525, 2023.

[25] R. K. Pearson, *EXPLORATORY DATA ANALYSIS USING R*. 2018. doi: 10.1201/9781315382111.

[26] V. Rohilla, D. S. Chakraborty, and D. R. kumar, "Random Forest with Harmony Search Optimization for Location Based Advertising," *Int. J. Innov. Technol. Explor. Eng.*, 2019, doi: 10.35940/ijitee.i7761.078919.

[27] J. D. Miranda-Calle, V. Reddy C, P. Dhawan,

and P. Churi, "Exploratory data analysis for cybersecurity," *World J. Eng.*, 2021, doi: 10.1108/WJE-11-2020-0560.