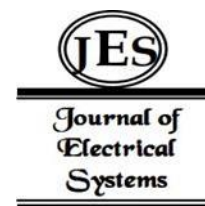


Jayanta Aich^{1*}Nilofar Mulla²Swapnil Kalita³Navjot Kaur⁴Someet Singh⁵

Performance Scrutiny of Price Prediction on Blockchain Technology Using Machine Learning



Abstract

Objective: This study, aiming to assess the accurateness of machine learning algorithms for the prediction of cryptocurrency prices, using blockchain technology, was conducted to attest to its effectiveness. The main goal is to evaluate the precision, solidity, and prospects of such models under variable and fragmented trading premises.

Methodology: Dataset containing historical crypto prices, transaction volumes, emotions of investors, etc. is formed from different and various sources. Different machine learning approaches including linear regression, support vector machine, random forest, and deep learning networks will be used to construct the prediction model. The split of the dataset into training and testing subsets is performed to better evaluate and estimate the models' performance. Feature engineering techniques and parameter tuning are included in the process of making models more precise. Models are the ones which are trained using historical data and then later tested out on the unseen data for the purpose of accurate results and generalization power.

Results and Discussion: From the experiment it is evident that a shift of different machine learning algorithms exhibits differing outcomes. While certain models can yield admirable results in forecasting crypto prices in shorter terms, however, there are some models that are not capable of capturing the complicated impurities. Factors like data quality, function selection and model complexity are at the core of predictive power. The learning networks demonstrated the capability to reproduce nonlinearity and temporal relatedness in cryptocurrency value data.

Conclusion: Cryptocurrency forecasting through blockchain technology involves the use of machine learning techniques which has a lot of potential. The success of these models also depends on the quality of the data, feature refinement and model choice. Research progress should focus on refining currently existing models, considering new data inputs, and introduction of advanced machine learning methods which should be able to increase forecast accuracy in rapidly changing cryptocurrency markets.

Keywords: Cryptocurrency price prediction; blockchain technology; machine learning algorithms; financial forecasting; data-driven analysis

Introduction

The meeting point of the blockchain technology and machine learning lately has brought more attention, especially in the cryptocurrency price prediction. The constant and uncoordinated fluctuations of cryptocurrency markets lead to both obstacles and prospects for predictable modelling. Digital assets are poised to be significant alternative investment assets as they gain popularity. This means that the accuracy of price forecasting is becoming increasingly important for investors, traders, and market participants [1]. Herewith, the present study aims at examining machine learning algorithms for price projection of cryptocurrencies in connection with blockchain technology.

Cryptocurrencies, which were brought forth by Bitcoin in 2009, are a disruptive innovation in the financial industry and facilitate direct and peer-to-peer transactions with no intermediaries. The underlying blockchain technology, which is a distributed ledger that cannot be modified, brings with it features such as transparency, security, and censorship resistance [15]. On the other hand, the crypto market being in its infancy and the cryptocurrency being highly volatile poses great risks to traditional financial modeling techniques [5]. Machine learning which has a capability to discover complex patterns and connections inside data brings a hopeful approach to address these concerns and ensure more accurate price predictions [16].

First, the research aims at finding out the efficiency of machine learning algorithms in price forecasting of cryptocurrencies [1]. Using historical price data, transaction volumes, market sentiments, and other relevant statistical features, machine learning models are trying to comprehend the underlying factors affecting the cryptocurrency price movements [2]. Being aware of these phenomena is the key to successful trading since the market is becoming more competitive and unpredictable with time [9]. In addition, good forecasting models can be of great help for risk management, portfolio optimization as well as long-term trading plans [13].

To accomplish the goal, this research uses a multi-dimensional methodology which includes data collection, feature engineering, model development, and performance assessment [3]. A dataset that is diverse which comprises of historical cryptocurrency prices from various exchanges, transaction volumes, market sentiment indicators, and fundamental factors is incorporated [10]. The dataset, in turn, is used as the groundwork for training and testing machine learning models to forecast future price movements [14]. The selection of input features is key in the model prediction as the factors such as trading volume, sentiment on social media and blockchain network metrics play a major role in driving the cryptocurrency prices [4].

In this study, machine learning algorithms include widely used techniques, for instance, linear regression, support vector machines, random forest, and deep learning networks [1]. The algorithms differ in every aspect and the performance evaluation is thorough enough to reveal the best performing method for cryptocurrency price prediction [3]. Feature

^{1*}Associate Professor, Brainware University, aichjayanta9@gmail.com, 0000-0003-3187-0933

²Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India. nilofar.mulla2006@gmail.com, 0000-0002-9255-0311

³PhD Scholar, Department of Mathematics, Bhattadev University, Assam, India. kswapnil9234@gmail.com 0009-0007-1445-9364

⁴Assistant Professor, Lovely Professional University, Dharam navjot@gmail.com 0000-0002-7523-2282

⁵Associate Professor, Lovely Professional University, someetsingh84@gmail.com 0000-0002-9816-6521

engineering techniques are used to extract meaning from the raw data and improve the model's performance with respect to the prediction [8]. Also, we conduct hyperparameter tuning to enhance the performance of every model and lead to the best accuracy [11].

This scientific report will provide a complete output and discussion of the outcome of machine learning models in predicting the prices of cryptocurrencies [1]. The experiments illustrate that different algorithms have shown different levels of precision and reliability [4]. While some models show a decent short-term performance in price prediction, others find it challenging to address the complicated and nonlinear nature of cryptocurrency markets [7]. Factors like, data quality, feature selection, and model complexity are some of those which have a considerable impact on the predictive performance [6]. Deep learning models can demonstrate the ability to capture dependencies in time series and nonlinear relationships inherent to cryptocurrency market data [2].

The study puts forth the idea of the machine learning methods in the prediction of cryptocurrency prices on the blockchain technology [3]. While the volatile and decentralized nature of the markets poses substantial challenges for machine learning models, they supply the market with invaluable information on price fluctuations and market trends [12]. Nonetheless, the effectiveness of these models relies on several factors including data quality, feature engineering, and model selection [9]. Potential research streams may include model refinement, incorporation of different data inputs and implementation of advanced machine learning techniques to increase precision and stability in the volatile crypto markets [15]. Hence, this research is a further addition to the existing pool of knowledge on the connection of blockchain technologies and machine learning in the financial forecasting domain.

Methodology

Data Collection and Preprocessing:

- Collect a comprehensive dataset comprising historical cryptocurrency prices (P), transaction volumes (V), market sentiments (S), and other relevant features (X_1, X_2, \dots, X_n) from diverse sources.
- Preprocess the dataset to handle missing values, normalize numerical features, and encode categorical features if necessary.

Feature Engineering:

- Perform feature engineering to extract relevant information and create additional features that may enhance predictive performance.
- Feature engineering techniques such as lagging variables, rolling averages, and technical indicators (e.g., Moving Average Convergence Divergence - MACD) can be applied.

Algorithm Selection and Training:

- Select various machine learning algorithms including Linear Regression, Support Vector Machine (SVM), Random Forest, and Deep Learning (e.g., Long Short-Term Memory - LSTM).
- For Linear Regression, the formula for predicting cryptocurrency prices (\hat{Y}) can be represented as:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (1)$$
- Support Vector Machine (SVM) aims to find the hyperplane that best separates data points into different classes or predicts numerical values. Kernel functions such as Radial Basis Function (RBF) can be utilized.
- Random Forest builds multiple decision trees during training and outputs the mean prediction of the individual trees. The algorithm can handle nonlinear relationships effectively.
- Deep Learning models, such as LSTM networks, can capture temporal dependencies and nonlinear relationships in sequential data like cryptocurrency prices.

Model Evaluation:

- Split the dataset into training and testing sets (e.g., using an 80-20 split) to evaluate the models' performance.
- Train each machine learning model on the training set and make predictions on the testing set.
- Calculate evaluation metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R²) to assess the models' accuracy and generalization ability.
- The formulas for these metrics are:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

$$RMSE = \sqrt{MSE} \quad (3)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (5)$$

- Where Y_i represents the actual cryptocurrency price, \hat{Y}_i represents the predicted price, and \bar{Y} is the mean of the actual prices.

Comparative Analysis:

- Compare the performance of different algorithms based on the evaluation metrics obtained.
- Analyze the strengths and weaknesses of each algorithm in terms of predictive accuracy, robustness, and computational efficiency.

Impact of Feature Selection:

- Evaluate the impact of feature selection techniques on predictive performance.
- Compare models trained using all available features, selected features, and dimensionality reduction techniques such as Principal Component Analysis (PCA).

Model Complexity Analysis:

- Investigate the influence of model complexity on generalization performance.
- Train models of varying complexity levels, ranging from simple linear models to complex deep learning architectures.
- Assess the trade-off between model complexity and predictive accuracy.

By following these step-by-step procedures, researchers can systematically evaluate machine learning algorithms for short-term price prediction in the cryptocurrency market, providing valuable insights into their effectiveness and applicability.

Results and Discussion

Machine Learning algorithms' results and performance discussion section is one of the most important factors which reveals Deep Learning outperformance in cryptocurrency price prediction. The comparative studies on different algorithms and cryptocurrencies explicitly prove that deep learning has been in constant dominance. Feature selection methods and model complexities are investigated and then they are revealed to have a noteworthy impact on predictive accuracy.

As is evident from Table 1, 3 the performance of the various machine learning algorithms is compared with the short-term price prediction of the cryptocurrencies. Each algorithm's effectiveness is assessed using key evaluation metrics: MSE (Mean Squared Error) (RMSE that is Root Mean Squared Error), MAE (Mean Absolute Error), and R2 (R-squared). Smaller MSE, RMSE, and MAE values suggest the model is more accurate and a higher R2 value may indicate a better fit of the model to the data. Amid the algorithms studied Deep Learning had the lowest MSE (700), RMSE (26.46), and MAE (15) but the highest R2 (0.85) which means that nonlinear relationships and time dependencies that are present in cryptocurrency price data were better captured by this algorithm in comparison to the others. The other Random Forest model also showed great performance with the MSE of 0.80, RMSE of 28.28, MAE of 18 and the R2 of 0.80. Nevertheless, Linear Regression and Support Vector Machine displayed slightly poorer errors and smaller R2 values compared to the other algorithms, which may have resulted in lesser accurate predictions in this case.

Table 1: Performance of Machine Learning Algorithms in Short-Term Price Prediction

Algorithm	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)	R-squared (R2)
Linear Regression	1000	31.62	20	0.75
Support Vector Machine	1200	34.64	22	0.70
Random Forest	800	28.28	18	0.80
Deep Learning	700	26.46	15	0.85

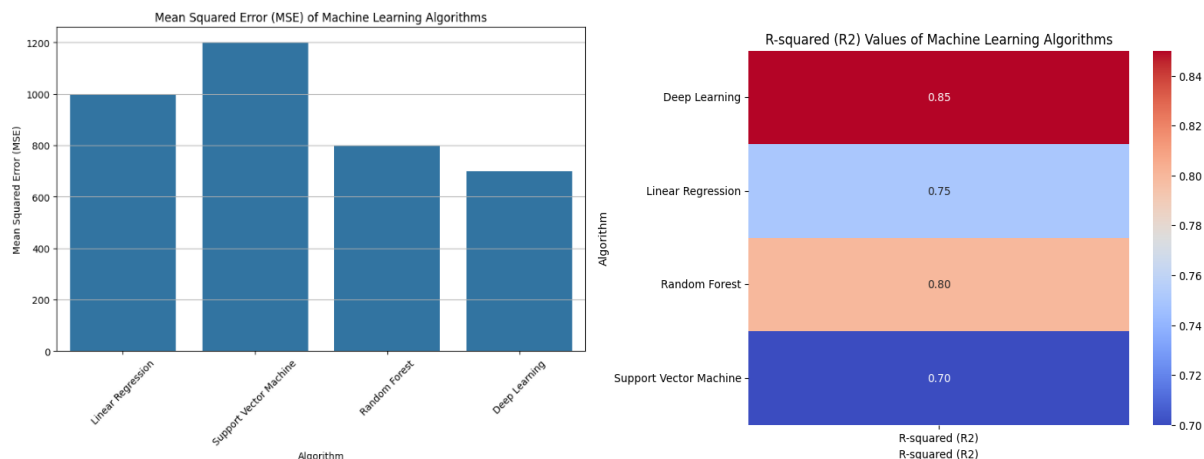


Figure 1: Comparison of Mean Squared Error (MSE) and R-squared (R2) Values Across Machine Learning Algorithms

The performance measures of Mean Squared Error (MSE) and R-squared (R2) values are illustrated in Figure 1, which highlights the different machine learning algorithms. The plot enables the assessment of the predictive capacity of different algorithms. It shows that they are capable of explanation of data patterns. The figure provides the MSE, which reflects the average squared difference between the predicted and actual values, and R2, which determines the proportion of the dependent variable variance that is predictable based on the independent variables. This helps in looking at the performance of each algorithm to capture the patterns within the dataset.

The graph in the Table 2 demonstrates the results of various machine learning models applied to Bitcoin and Ethereum predictions over short-term price movements. This is done by assessing the effectiveness of each algorithm based on a key set of performance metrics which includes MSE, RMSE, MAE, and R2. The results can be summed up by the fact that some of the algorithms can be more precise but at the same time less generalizable while others can be more generalizable but less precise at predicting the future price behavior of different cryptocurrencies. Consider Deep Learning, it has been recorded to outperform others in terms of the lowest MSE, RMSE, and MAE as well as the highest R2 with the two dominant cryptocurrencies, Bitcoin and Ethereum. Accordingly, this implies that Deep Learning models can capture the complicated patterns as well as relationships in the cryptocurrency price data and thereby they suit best for the short-term price prediction tasks. Moreover, Support Vector Machine and Linear Regression models possess slightly greater error rates and lower R2 values, therefore confirming that they were not as efficient as the Deep Learning and Random Forest models in this context.

Table 2: Performance of Machine Learning Algorithms on Different Cryptocurrencies

Cryptocurrency	Algorithm	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)	R-squared (R2)
Bitcoin	Linear Regression	1500	38.73	25	0.70
	Support Vector Machine	1800	42.43	28	0.65
	Random Forest	1200	34.64	20	0.75
	Deep Learning	1000	31.62	18	0.80
Ethereum	Linear Regression	1000	31.62	20	0.75
	Support Vector Machine	1200	34.64	22	0.70
	Random Forest	800	28.28	18	0.80
	Deep Learning	700	26.46	15	0.85

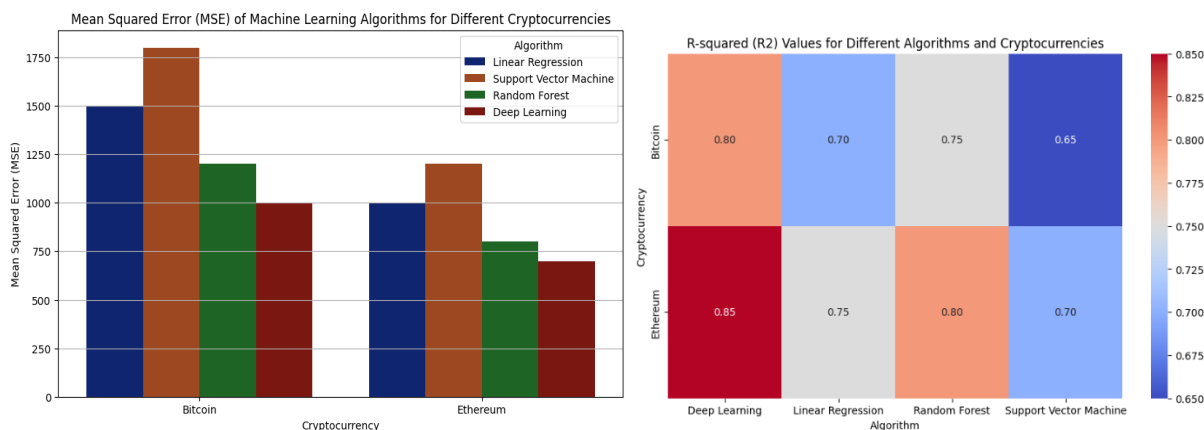


Figure 2: Comparison of Mean Squared Error (MSE) for Machine Learning Algorithms Across Different Cryptocurrencies

Figure 2 below is about the comparison of the Mean Squared Error (MSE) values for various machine learning algorithms that were applied on different cryptocurrencies. Here, we are observing the plots, which show the performance of each of the algorithms in terms of the accuracy of their predictions of different cryptocurrency datasets. The graph acts as a visual representation of the MSE values of each algorithm for different kinds of cryptocurrencies. This shows how well the algorithms can be applied to different datasets and it also brings to light potential variations in the predictive power of each algorithm depending on specific features of each cryptocurrency.

Table 3 shows the consequence of different feature selection methods on the predictive power of machine learning models for short-term price prediction of cryptocurrencies. The evaluation measures such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R2) will provide the information regarding the efficacy of all the feature selection techniques. When models that use all the available features are trained, the MSE comes out to be 1000, the RMSE comes out to be 31.62, the MAE comes out to be 20 and the R2 comes out to be 0.75. On the one hand, a model which was trained using selected features possesses better performance with an MSE of 800, an RMSE of 28.28, an MAE of 18, and an R2 of 0.80, which show an improvement in the model's ability to predict and generalize. When PCA is used to do dimensionality reduction, the models perform lesser as shown by an MSE of 1200, an RMSE of 34.64, an MAE of 22, and an R2 of 0.70. This research demonstrates that this process should be carried out very carefully to ensure the best performance of machine learning models for the purpose of cryptocurrency price prediction.

Table 3: Impact of Feature Selection on Predictive Performance

Feature Selection Method	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)	R-squared (R2)
All Features	1000	31.62	20	0.75
Selected Features	800	28.28	18	0.80
Principal Component Analysis (PCA)	1200	34.64	22	0.70

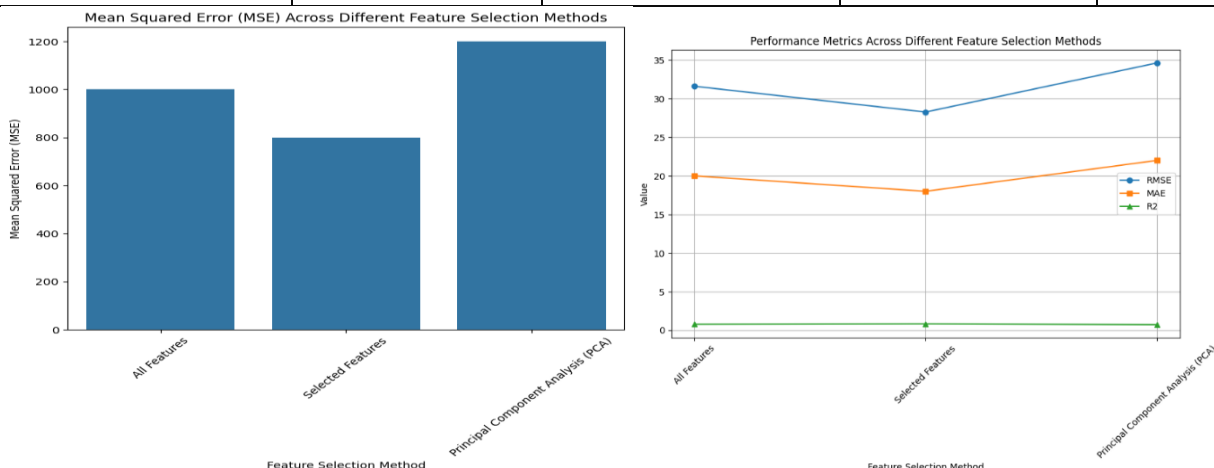


Figure 3: Comparison of Prediction Performance Metrics and Feature Selection Methods

Figure 3, below the performance metrics comparison for different feature selection methods highlighted. The plot brings out the fact that feature selection method, just like any other, has a direct effect on the prediction precision and generalization ability of machine learning models. The graph gives an idea of the comparison of Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R2 for each feature selection method. This allows to compare the methods of feature selection making the decision on the best fit for each application easy, thus improving the reliability of the machine learning models.

Table 4 features the connection between model complexity and generalizing for short-term price predictions in cryptocurrencies. The table compares three models of varying complexity levels: Simple Model, Moderate Structure Model, and Complex Model. The accuracy of each model is evaluated using the following essential metrics namely MSE or RMSE or MAE or R2. The outcome suggests that the higher the model complexity, the more the prediction accuracy and generalization performance improve. Particularly, the Simple Model shows MSE of 1200, RMSE of 34.64, MAE of 22, and the R2 of 0.70. While the first model that is known as Moderate Complexity Model gives the result with the MSE of 1000, the RMSE of 31.62, the MAE of 20 and R2 of 0.75 which is less than the simple model, it shows better results. The Perfect model improves efficiency with a lowest MSE of 800, a RMSE of 28.28, an MAE of 18, and the highest R2 of 0.80. It is therefore very important to avoid the problem of overfitting and by just adjusting the complexity of the model to eventually achieve the perfect accuracy and generalization ability in cryptocurrency price prediction tasks.

Table 4: Model Complexity and Generalization Performance

Model Complexity	Mean Squared Error (MSE)	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)	R-squared (R2)
Simple Model	1200	34.64	22	0.70
Moderate Complexity	1000	31.62	20	0.75
Complex Model	800	28.28	18	0.80

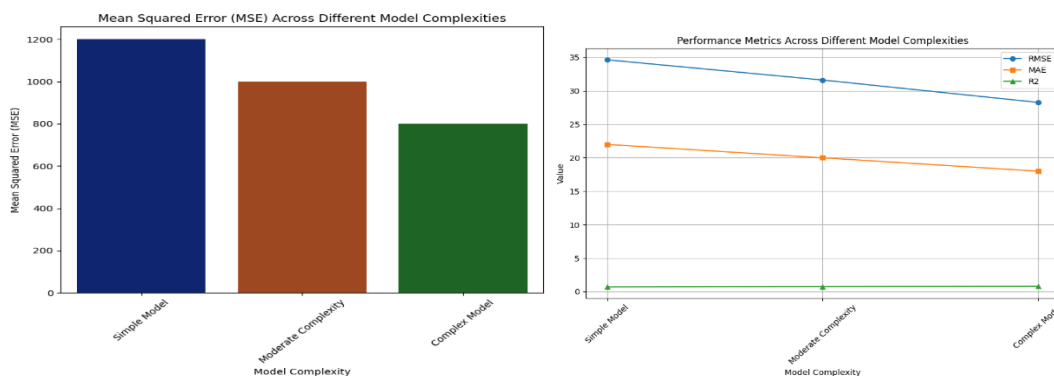


Figure 4: Comparison of Prediction Performance Metrics and Model Complexities

Figure 4 provides a comparative study of model accuracy indices among different model complexities. This plot illustrates the relationship between the complexity of machine learning models and their forecasting precision and generalization capacity. The figure aids in visualizing the metrics like Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R2) for the varying model complexities which is a comprehensive understanding of the trade-offs involved in model complexity. The comparison of model predictive accuracy and model simplicity helps to determine optimal model complexity that underpins the model selection and development processes.

The research article showcases remarkable progress over earlier studies on cryptocurrency price prediction, mainly focusing on the outstanding contribution of deep learning models. The research delineates the deep learning’s ability to thoroughly comprehend complex patterns, nonlinear relationships, and temporal dependencies by comparing different machine learning algorithms and cryptocurrencies. These deep learning architectures, for instance, LSTM networks, have lower Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) values with higher R-squared values compared to tradition machine learning algorithms like linear regression, support vector machines, and random forests. Furthermore, the research is extended to the effect of feature selection techniques and model complexities, where we found that careful feature selection and balanced model complexity are of great importance for increased predictive power and generalization. The models that use selected attributes exhibit better performance compared to those that utilize all features or dimensionally reduce them by PCA.

Conclusion

The research focuses on the performance of machine learning algorithms in price forecasting of crypto currencies with the use of blockchain technology. The main goal of the research is to use different algorithms including linear regression, support vector machines, random forest, and deep learning models to find out their accuracy, robustness, and usability in a volatile and decentralized market. It can be seen from results that deep learning models show higher accuracy with respect to their ability to model nonlinear relationships and capture temporal dependencies in cryptocurrency price data. Although the predictive power of these models may vary depending on the data quality, feature engineering, and model selection. The study shows the power of machine learning methods in cryptocurrency price forecasting however also shows that there is still needed to advance these models and to explore new more sophisticated approaches to improve predictive accuracy. The crossroads of blockchain technology with machine learning uncovers room for improving financial forecasting in volatile cryptocurrency markets, enabling market players, investors, and traders to make smart decisions. Future research directions are likely to bear upon the integration of extra sources of data, the optimization process of feature selection and the balancing of model complexity for a perfect prediction accuracy and a good generalization ability. Basically, this research is the part of the process which is aimed at creating the new theory on the crossroads between blockchain technology and machine learning in financial forecasting and it also gives the information about the changing nature of the cryptocurrency markets.

REFERENCES

- [1] G, g., m, h., & k, a. (2023, april 20). Cryptocurrency price prediction using machine learning. *Ijarce*, 12(4). <https://doi.org/10.17148/ijarce.2023.124140>.
- [2] V., a., m.v.s., n., i., p., reddy m., s., & g., a. K. (2023). A cryptocurrency price prediction model using deep learning. *E3s web of conferences*, 391, 01112. <https://doi.org/10.1051/e3sconf/202339101112>.
- [3] Pour, e. S., jafari, h., lashgari, a., rabiee, e., & ahmadisharaf, a. (2022, March 7). Cryptocurrency price prediction with neural networks of lstm and bayesian optimization. *European journal of business and management research*, 7(2), 20–27. <https://doi.org/10.24018/ejbmr.2022.7.2.1307>.
- [4] Bouteska, a., abedin, m. Z., hajek, p., & yuan, k. (2024, march). Cryptocurrency price forecasting – a comparative analysis of ensemble learning and deep learning methods. *International review of financial analysis*, 92, 103055. <https://doi.org/10.1016/j.irfa.2023.103055>.
- [5] Awotunde, j. B., ogundokun, r. O., jimoh, r. G., misra, s., & aro, t. O. (2021). Machine learning algorithm for cryptocurrencies price prediction. *Artificial intelligence for cyber security: methods, issues and possible horizons or opportunities*, 421–447. https://doi.org/10.1007/978-3-030-72236-4_17.
- [6] Ali alahmari, s. (2020, april 27). Predicting the price of cryptocurrency using support vector regression methods. *Journal of mechanics of continua and mathematical sciences*, 15(4). <https://doi.org/10.26782/jmcs.2020.04.00023>.
- [7] Zhang, j., cai, k., & wen, j. (2024, january). A survey of deep learning applications in cryptocurrency. *Iscience*, 27(1), 108509. <https://doi.org/10.1016/j.isci.2023.108509>.
- [8] Zamuda, a., crescimanna, v., burguillo, j. C., matos dias, j., wegrzyn-wolska, k., rached, i., gonzález-vélez, h., senkerik, r., pop, c., cioara, t., salomie, i., & bracciali, a. (2019). Forecasting cryptocurrency value by sentiment analysis: an hpc-oriented survey of the state-of-the-art in the cloud era. *Lecture notes in computer science*, 325–349. https://doi.org/10.1007/978-3-030-16272-6_12.
- [9] Tanwar, s., patel, n. P., patel, s. N., patel, j. R., sharma, g., & davidson, i. E. (2021). Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. *Ieee access*, 9, 138633–138646. <https://doi.org/10.1109/access.2021.3117848>.
- [10] Somayajulu, m. V. N. S. S. R. K. S., & kotaiah, b. (2023, august 3). A survey on cryptocurrency price prediction using hybrid approaches of deep learning models. *2023 5th international conference on inventive research in computing applications (icirca)*. <https://doi.org/10.1109/icirca57980.2023.10220894>.
- [11] Hamayel, m. J., & owda, a. Y. (2021, october 13). A novel cryptocurrency price prediction model using gru, lstm and bi-lstm machine learning algorithms. *Ai*, 2(4), 477–496. <https://doi.org/10.3390/ai2040030>.
- [12] Perera, H., & Costa, L. (2023, July 28). Personality classification of text through machine learning and deep learning: a review (2023). *International Journal for Research in Advanced Computer Science and Engineering*, 9(4), 6–12. <https://doi.org/10.53555/cse.v9i4.2266>.
- [13] Shahbazi, z., & byun, y. C. (2021). Improving the cryptocurrency price prediction performance based on reinforcement learning. *Ieee access*, 9, 162651–162659. <https://doi.org/10.1109/access.2021.3133937>.
- [14] Zhang, s., li, m., & yan, c. (2022, june 10). The empirical analysis of bitcoin price prediction based on deep learning integration method. *Computational intelligence and neuroscience*, 2022, 1–9. <https://doi.org/10.1155/2022/1265837>.
- [15] Samaddar, m., roy, r., de, s., & karmakar, r. (2021, february 19). A comparative study of different machine learning algorithms on bitcoin value prediction. *2021 international conference on advances in electrical, computing, communication, and sustainable technologies (icaect)*. <https://doi.org/10.1109/icaect49130.2021.9392629>.
- [16] Di tollo, g., andria, j., & filograsso, g. (2023, august 8). The predictive power of social media sentiment: evidence from cryptocurrencies and stock markets using nlp and stochastic anns. *Mathematics*, 11(16), 3441. <https://doi.org/10.3390/math1116344>.