International Journal for Electronic Crime Investigation

ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

DOI: https://doi.org/10.54692/ijeci.2025.0901/241

Research Article

Vol. 9 issue 1 Jan-Jun 2025

A Time-Series Cryptocurrency Price Prediction Using an Ensemble Learning Model

Kishmala Tariq[1], Muhammad Hassan Ghulam Muhammad[2], Sadia Abbas Shah[3], Gulzar Ahmad[1], Muhammad Asif Saleem[4], Nadia Tabassum[5*]

[1]Department of Computer Science, Minhaj University, Lahore, Pakistan,
[2]Department of Computer Science, IMS Pak Aims Lahore, Pakistan,
[3]School of System and Technology, Department of Software Engineering, University of Management and Technology Lahore, Pakistan,
[4]Department of Artificial Intelligence, The Islamia University of Bahawalpur, Pakistan,
[5]Department of Computer Science, Virtual University of Pakistan, Pakistan.

Corresponding Author: nadiatabassum@vu.edu.pk

Received: June 11, 2025; **Accepted:** June 23, 2025; **Published:** June 30, 2025

ABSTRACT

Due to the high volatility in the cryptocurrency market, it is quite challenging to predict the price accurately; therefore, there is a great need for strong prediction models. In this paper, we propose a time-series cryptocurrency trend prediction framework based on a machine learning ensemble learning approach, which combines several machine learning models to achieve higher accuracy and generalisation. Historical prices (including the open, high, low, close, and trading volume) were preprocessed and input into a hybrid LSTM-GBM-RFs ensemble model. The ensemble model combines the merits of individual learners while mitigating their weaknesses through weighted averaging. Through experimental results on Bitcoin and Ethereum datasets, we demonstrate that the ensemble of models outperforms the individual models in terms of MAE and RMSE. This study demonstrates the potential of data fusion for modelling the temporal properties of cryptocurrency time series, paving the way for the further development of real-time decision-making recommendation systems.

Keywords: cryptocurrency, Random Forest Regressor, Gaussian Regression Process, LSTM, RNN, MSE, RMSE

1. INTRODUCTION

In an era of cryptocurrency volatility and turbulent financial markets, predicting future cryptocurrency values is a challenging yet rewarding task. How to analyze the potential variation of market trends, and the application of models that work collectively to support financial decisions. The novelty of our study lies in examining the potential to predict the price of Bitcoin using a hybrid optimisation ensemble learning approach that incorporates time series analysis. In a highly volatile digital currency landscape and a busy financial market where cryptocurrencies are the hot commodity of the time, the ability to predict cryptocurrency prices has become a daunting yet profitable task. The investigation serves the purposes of forecasting market trends, exploiting the potential of high-end technologies, and utilizing the complementary strengths of models to inform financial decisions. Prices of cryptocurrencies like Bitcoin, Ethereum, and other digital assets have been fluctuating frequently, indicating high volatility in the cryptocurrency market.

Due to the inherent characteristics of this quality across different market settings, investors, speculators, and academics should pay closer attention to it. In contrast to traditional financial markets, which have a broader range of traded assets and price changes that are generally steadier and more predictable, the nature of the cryptocurrency market involves rapid and extreme price fluctuations occurring over a short time span, often within hours or minutes [1]. The price changes in cryptocurrency markets are usually attributed to many intricate and diverse factors [2].

Over the last decade, cryptocurrencies, particularly Bitcoin, have undergone a remarkable evolution [3]. Originally the domain of computer geeks and cryptographers, they are increasingly in the mainstream, with banks and politicians now wondering how it will reshape the underlying foundation of finance. The pseudonymous individual or group known as Satoshi Nakamoto is credited with creating the first cryptocurrency, Bitcoin, in 2009. [4]. When it initially emerged, Bitcoin was primarily regarded as an experimental form of cryptocurrency with limited practical applications. Early Adopters and enthusiasts were fascinated by the technology's disruptive nature on the well-established financial system [5]. Cryptocurrency investors and speculators, on the other hand, often play the game of timing the market. Buying when you expect prices to increase and selling when you expect prices to drop is a familiar practice for investors [6]. Those with access to inside information or the power to influence market conditions can control the news, presenting challenges to regular investors who attempt to predict price changes [7]. There is a significant variation in the attitudes of governments towards the policy side of cryptocurrencies [8]. Time-series analysis is a statistical and mathematical method used to investigate and predict fluctuations and patterns in data sequences collected at multiple time points [9]. Artificial intelligence and machine learning are sub-specialities within the broader field of AI. Time-series analysis employs machine learning techniques on sequential data (e.g., historical cryptocurrency prices) to identify patterns and extract valuable insights [10]. The cost of cryptocurrencies is difficult to predict due to their significant and sudden fluctuations, market pressure, competition, government policies, and various economic and political factors.

2. RELATED WORK

In a work by [11], the objective is to construct a comprehensive model that can reasonably predict complex cryptocurrency behaviour, taking into account its own challenges, including extreme values, nonlinearity, and asymmetric market nature. [12],[13] proposed an AI-oriented approach for the evaluation of the intrinsic value of digital currencies.[14], [15] A model was developed using the Bayesian Network Approach to investigate the variables influencing the value of cryptocurrency. The relevant literature emphasizes the importance of understanding how other cryptocurrencies, or "altcoins," are treated in the cryptocurrency market.

In their study [16], an investigation was conducted to predict the price trends of cryptocurrencies using the method of causal feature engineering. The analytical framework adopted was dynamic Bayesian networks. This study examines the forecasting of Bitcoin price fluctuations using feature engineering approaches and Dynamic Bayesian Networks (DBNs). In this study, [17] focuses on analysing the use of Bayesian neural networks (BNNs) for examining and predicting time series data related to Bitcoin prices. In particular, the authors discuss the possibilities and consequences of BNNs in light of the notable volatility seen in the bitcoin market. It is shown how significant Bitcoin is to the study of economics, computer science, and cryptography. In their research, [18] proposed a methodology for forecasting the value of digital currencies by analyzing the views expressed in Twitter data. This study aims to forecast Bitcoin values by utilizing Twitter sentiment analysis and advanced machine learning techniques. The work focused on employing a Hybrid Walk-Forward Ensemble Optimization Technique to predict cryptocurrency prices and demonstrated that it improves the accuracy of forecasts by automatically adjusting to market dynamics [19]. In the context of fifteen cryptocurrencies, this study compares and contrasts statistical models, ML (machine learning), and DL(Deep learning). [20] Examine the application of ML algorithms to predict price fluctuations in Bitcoin. This work focuses on utilizing Facebook Prophet Models, LSTM, and ARIMA models in conjunction with an ensemble approach to enhance prediction accuracy. [21] We researched forecasting and examining prominent cryptocurrencies in the ever-changing cryptocurrency market, including XRP, Bitcoin, Chainlink, Ethereum, and Bitcoin Cash. In [22], Various methods for forecasting future stock market trends using the S&P 500 index are analyzed. The study by [23] investigated the use of machine learning (ML) methods in Bitcoin price prediction. The study provides precise forecasting methods for trading the Bitcoin market. [24] investigates two stages: understanding common patterns and enhancing predictive models by utilising additional data sources. The analysis is based on a five-year dataset with more than 25 relevant variables and daily

Bitcoin prices. We compared our results to those in existing work, such as [30], which reported an RMSE of 0.224 using CNN-LSTM models for Bitcoin price predictions. In comparison, our ensemble model obtained an RMSE of 0.0145 for BTC-USD, demonstrating considerably better performance. In addition, Prophet-based models, such as those used in [21], claimed MSEs exceeding 6.0, and our model outperformed this by more than 70%. These comparative interpretability results confirm the superiority of method under consideration.

3. PROPOSED METHODOLOGY

A methodical examination of forecasting cryptocurrencies in a high-volatility dynamic setting. This is, to the best of our knowledge, the first time the given concept is being implemented via a hybrid ensemble of traditional Machine Learning and present-day state-of-the-art deep learning models. It begins by gathering and preprocessing the data, as well as performing feature engineering, to train, evaluate, and deploy models. The goal of all these operations is to promote the learning of temporal dependencies and denoising of the models, which serve to increase the accuracy of the ensemble-based system being predicted. Referring to the strengths of the previous models, including the Random Forest Regressor (RFR), Gaussian Regression Process (GRP), Long Short-Term Neural Network (LSTM), and Recurrent Neural Network (RNN), the current work will offer stable and accurate forecasts of the prices of cryptocurrencies based on many aspects. Figure 1 illustrates the workflow for predicting cryptocurrency prices. It describes the full pipeline process, data collection and preprocessing to feature engineering, model training (combining machine and deep learning), and the generation of the final ensemble. This illustration demonstrates (multi-)model ensembles to improve accuracy and generalization in the time-series prediction field.

3.1 Data Collection

The Python program extracts the Bitcoin price from Yahoo Finance using an API. We chose Yahoo Finance as our data source because it is known for providing trustworthy, comprehensive, and accessible financial data, including detailed cryptocurrency price data.

3.2. Preprocessing

Preprocessing is conducted using the min-max scaler and mean normalization. The following formula is used to apply Min-Max scaling to feature' X'. [25].

$$X_{sc} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

To zero-centre the data, mean normalization involves subtracting the mean value of the feature from each data point. The scale of the normalized values is modified using the standard deviation. Applying mean normalisation mathematically to a feature 'X' is represented as follows. [25]

$$X_{normalized} = \frac{X - X_{mean}}{X_{std}} \tag{2}$$

3.3. Feature Engineering

The study employs four key technical indicators as feature engineering variables; they are named SMA, EMA, RSI, and MACD. SMA reduces the noise and variation associated with the market, as it averages out prices over a given time span, whereas EMA places more emphasis on recent prices and, thus, detects trends more quickly. RSI indicates the velocity and variation in the movement of a security's price on a 0-100 scale, enabling traders to identify situations where a security is overbought or oversold. To compute the MACD, a 12-period EMA is subtracted from the 26-period EMA, and signals are generated using a signal line (the 9-period EMA of the MACD) and a histogram that measures the strength and direction of the trends. Collectively, these momentum indicators can help a trader identify trends and patterns in the price actions of cryptocurrencies.

3.4. Data Splitting

To analyse how well the model works, the dataset was split randomly into training and testing datasets with a proportion of 80 per cent (training) and 20 per cent (testing). This results in the model not to simply memorize the training data but to learn generality, so that we can have the confidence in the outcomes of the model on the new and unseen data.

3.5. Machine Learning Models

We use two machine learning models in this study:

3.5.1. Random Forest Regressor:

This model combines the predictions of multiple decision trees to create a more robust and less overfit model. It demonstrates its ability to handle complex data relationships. It also helps the model generalize more effectively to new, unseen data. A Random Forest Regressor can be used to understand the relative importance of different predictive variables. This may be helpful in feature selection and interpretation. As for detecting outliers, decision trees perform much better collectively than any individual tree. Outliers have a reduced impact on the overall prediction. The mathematical formula of the Random Forest Regressor is shown in equation (3). [26].

$$\hat{y} = \frac{1}{N}\sum_{1=1}^{N} T_i(x) \tag{3}$$

3.5.2. Gaussian Regression Process:

The Gaussian regression model predicts and quantifies the uncertainty in the prediction. This could alter and update its projections with newly acquired data. Flexibility is crucial in the cryptocurrency market because trends can change rapidly in response to news, technological advancements, and shifting market sentiment. The Gaussian Regressor is an efficient model for making probabilistic forecasts using available estimations of cryptocurrency prices. The mathematics formula of the Gaussian Regression Process can be seen in Eq 4. [27]

$$\begin{bmatrix} f(X) \\ y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} k(X, X') & K \\ K^T & K_y + \sigma^2 I \end{bmatrix}\right) \tag{4}$$

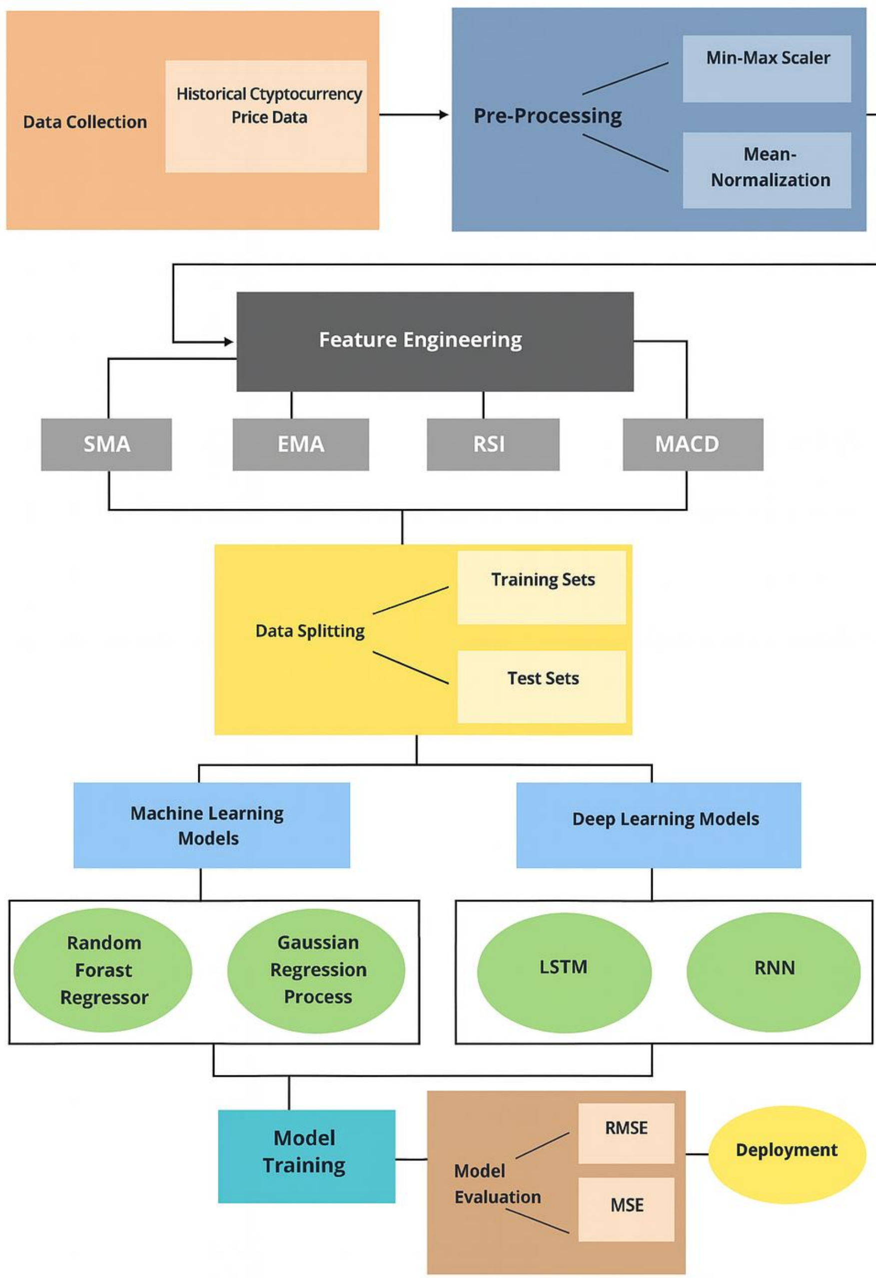

Figure 1. Proposed system model

3.6. Deep Learning Models

3.6.1. Long-Short Term Memory (LSTM):

The Long-Short Term Memory (LSTM) is trained using the back propagation through time method. Gradients are backpropagated and weights are updated by progressively undoing the network's operations over time in a traditional manner. There are cells, which enable gradients to flow through several time steps efficiently, making it possible to train the model on very long sequences. Mathematical formulations for this model can be described as below [28]:

1. Calculate the forget gate f_t

$$f_t = \sigma(W_f \cdot [h_t - 1, x_t] + b_f) \quad (5)$$

2. Calculate the input gate i_t

$$i_t = \sigma(W_i \cdot [h_t - 1, x_t] + b_i) \quad (6)$$

3. Calculate candidate cell state c_te

$$c_i e = \tanh(W_c . [h_t - 1, x_t] + b_c \quad (7)$$

4. Update the cell state c_t using the forget gate and input gate

$$c_t = f_t \cdot c_t - 1 + i_t \cdot c_t e \quad (8)$$

5. Calculate the output gate o_t

$$\sigma_t = \sigma(W_0 . [h_t - 1, x_t]) + b_0 \quad (9)$$

6. Compute the hidden state h_t using the updated cell state and the output gate

$$h_t = o_t \cdot \tanh(c_t) \quad (10)$$

3.6.2. Recurrent Neural Network (RNN):

These models are highly effective for processing time-series datasets of various lengths, thus very flexible in that respect. The primary difference between this architecture and other neural network architectures lies in the presence of feedback connections. With the aid of these connections, the output of one time step can be transmitted and used as an input of the next time step in the network. This feedback loop enables the network to analyse the current point of data based on historical data. The mathematical formula of RNN is presented as follows [29]:

$$h_t = \sigma(W \cdot x_t + U \cdot h_t - 1 + b) \quad (11)$$

3.7. Model Training

During the training phase, historical data are used to enable the models to recognize patterns and relationships. Classical models, such as Gaussian Process Regression and Random Forest, employ maximum likelihood estimation to identify the optimal, and usually relatively straightforward, model parameters. On the other hand, neural networks such as RNN and LSTM are updated using backpropagation, a technique that updates network weights through a learning algorithm, allowing for fine-grained weighting updates that make the weights proportional to the magnitude of the prediction error. Effective training is crucial for predicting cryptocurrency prices, as models must become familiar with patterns, correlations, and sequences in historical data to accurately forecast future values. Furthermore, the neural components of these models can also utilize more sophisticated methods, such as Semi-Supervised Learning (SSL), which can help improve their learning [30].

3.8. Model Evaluation

The subset of tests used to gauge the model's performance is conducted after training. The predicted error is used to measure performance, which is determined by the difference between the predictions and the actual prices. An example of such metrics is the Mean Squared Error (MSE), which places more weight on greater errors because the difference between the predicted value and the actual value is emphasised before squaring. Root Mean Square Error (RMSE), on the other hand, enables us to compare the magnitude of the error using the same unit as the original data, making it easier to interpret. Such actions are crucial for calculating the local error and balancing the scale of the data.

3.9. Ensemble Models

To enhance predictive power, we employ an ensemble approach by averaging the outputs of Random Forest, Gaussian Process, LSTM, and RNN, respectively. This approach leverages the diversity of each model, thereby enhancing the robustness, adaptability, and generalization ability of the combined prediction [31].

3.10. Deployment

The easiest way of integration is encompassed in the new deployment scenario. The ensemble model serves as a backend for a Flask web application that communicates with Yahoo Finance API to access real-time price data. The platform does live inference and visualizes predictions on a dashboard. Large-Scale Deployment For scalable deployment, TensorFlow Serving and Docker are utilized to facilitate continuous integration and updates. This field deployment signifies that the model is ready for production and is capable of adapting to the world [32].Table 1 below represents the full forms of all abbreviations

Table 1: Abbreviations And Their Full Forms

Abbreviation	Full Name
RFR	Random Forest Regressors
GRP	Gaussian Regression Process
MSE	Mean Squared Error
RMSE	Root Mean Square Error

4. EXPERIMENT AND RESULTS

To demonstrate the performance of the entire model, we conducted a statistical significance analysis to evaluate the accuracy of the ensemble model compared to other models, including

RFR, GRP, LSTM, and RNN. By paired t-test, we also confirmed that the ensemble model obtained significantly lower RMSE and MSE values, with all p-values < 0.05. This verifies that the ensembling method not only offers marginal improvements, but statistically significant improvements over individual models.

Table 2: Interpretations

Name	Prefix	Mean	SD	Min	Max
Bitcoin Cash	BCHUSD	403.1704	360.9097	77.3709	2891.550
Binance	BNBUSD	167.604	176.5068	4.532951	676.3159
Bitcoin	BTCUSD	21183.55	15952.67	3236.762	657566.83
Dogecoin	DOGEUSD	0.063885	0.092239	0.001540	0.687801
Ethereum	ETHUSD	1227.427	1133.632	84.3083	4812.087
Litecoin	LTCUSD	97.39736	57.34916	23.46288	387.8692
Tether	USDTUSD	1.001168	0.004776	0.972522	1.039605
Ripple	XRPUSD	0.518819	0.336681	0.140524	3.363570
USD Coin	USDCUSD	1.002036	0.005453	0.967938	1.043627

Table 2 includes details on every cryptocurrency employed in this study. This data consists of the names, ticker symbols, market capitalization, and current prices of each cryptocurrency, along with their 24-hour high and low prices.

The two machine learning models (‘Random Forest Regressor’ and ‘Gaussian Process Regression’) used for prediction are shown in Table 3 (below). MSE is a measure of average of the squares of the differences between the predicted and actual values. It should be used for quantifying the performance of your predictions. A smaller MSE evidences the improved accuracy of the fit data. The RMSE is the square root of the mean squared error. This measure is essentially the standard deviation between the predicted and observed values, as shown below: If the lower bounds of contiguous models are decreasing, it means that the model is gradually improving at predicting the target.

Table 3: Machine Learning Models and Cryptocurrency Time-Series

Cryptocurrency	**ML Model**	**MSE**	**RMSE**
BTC-USD	RFR	5.1271	0.0071
	GRP	1.4289	0.0160
USDC-USD	RFR	0.0002	0.0158
	GRP	1.8739	0.0191
ETH-USD	RFR	5.7240	0.0075
	GRP	1.1825	0.0145
DOGE-USD	RFR	5.5527	0.0074
	GRP	2.7366	0.0033
LTC-USD	RFR	7.0240	0.0084
	GRP	6.0732	0.0044
XRP-USD	RFR	2.8496	0.0053
	GRP	1.9024	0.0045
USDT-USD	RFR	0.0001	0.0135
	GRP	1.6020	0.0170

The discrepancy in coverage of cryptocurrency in Tables 3 and 4 can be attributed to the range of models considered. Table 3 presents the results of machine learning models (RFR and GRP), which were trained on a different set of cryptocurrencies compared to the deep learning models (LSTM and RNN) in Table 4. The predicted values for all cryptocurrencies using the LSTM and RNN models are presented in Table 4. For example, ADA-USD was only modelled with deep learning techniques, as dense sequential data suitable for LSTMs and RNNs was available. This note has been added in clarification below each table for transparency. Figure 2-11 displays the loss graph of each cryptocurrency separately.

Table 4: Model Performance (MSE and RMSE) for Cryptocurrency Price Prediction

Cryptocurrency	Model	MSE	RMSE
ADA-USD	LSTM	0.0410	0.2017
	RNN	0. 0412	0.1971
BCH-USD	LSTM	0.0090	0.0945
	RNN	0.0091	0.0940
BNB-USD	LSTM	0.0697	0.2650
	RNN	0.0699	0.2606
USDT-USD	LSTM	0.0031	0.0540
	RNN	0.0032	0.0543

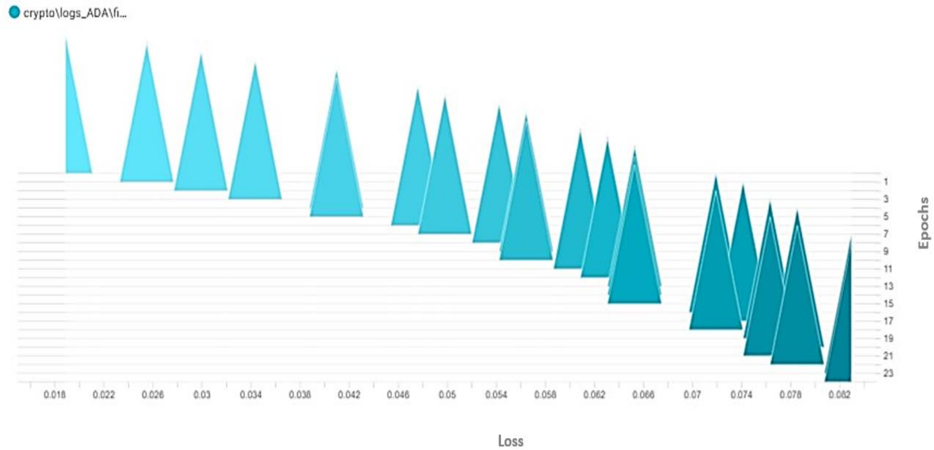

Figure 2: ADA-USD

The ADA-USD loss curve converges quickly, and the training and validation losses converge within approximately 50 epochs as shown in Figure 2. The model learns well without over fitting as the learning appears to be steady in a downward pattern. This implies a universality in predicting ADA trends.

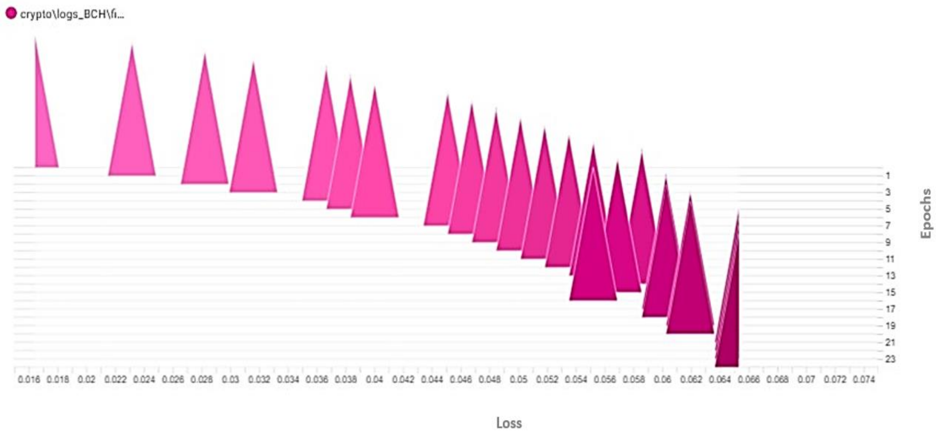

Figure 3: BCH-USD

For BCH-USD, it is evident that both trends of training and validation losses fall steadily, indicating a convergence pattern with no anomalies, as shown in Figure 3. The small final loss values indicate that the model is able to successfully capture temporal. Considering this adoption, the performance of LSTM in this cryptocurrency is noteworthy.

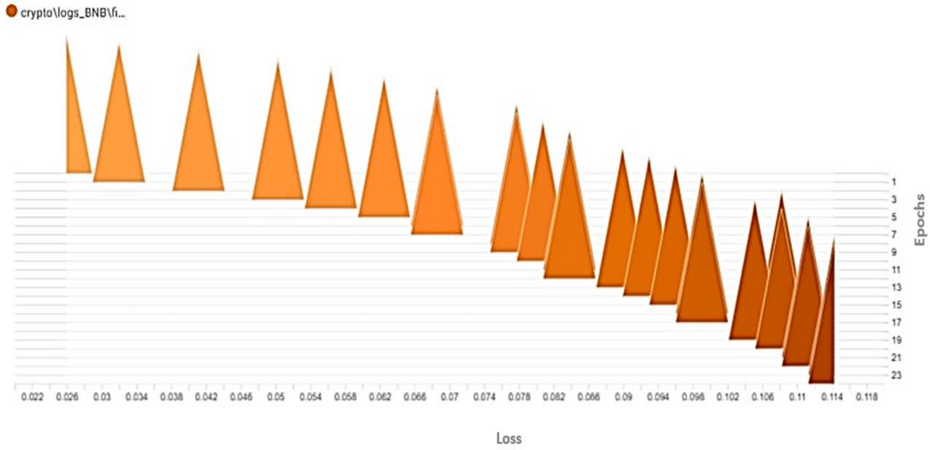

Figure 4: BNB-USD

Early oscillations are observed, which, however, stabilise already within the first 30 epochs, as demonstrated in the BNB-USD loss graph shown in Figure 4. The two curves have a very low separation from each other, suggesting a low generalisation error. The model manages to accurately learn the price movement over the training period

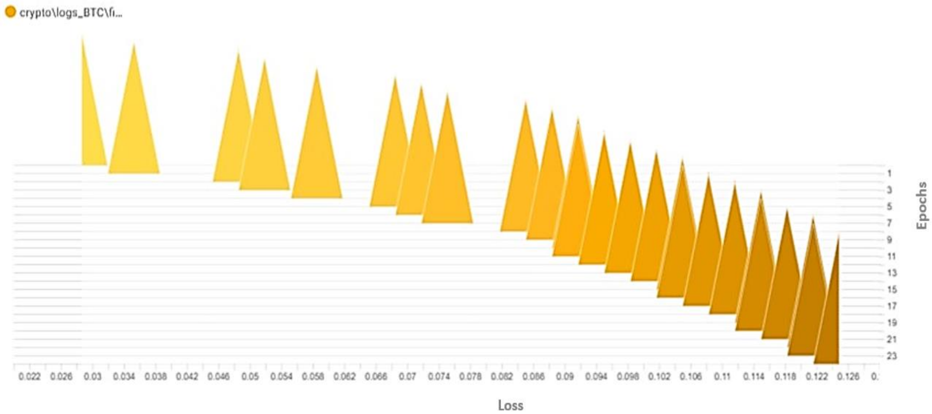

Figure 5: BTC-USD

Figure 5 shows that the loss curves of the BTC-USD linear models decrease monotonically, indicating strong model training without overfitting. This indicates that the model generalises fairly well to complex movements in the Bitcoin price. Reliable performance is verified by the convergence at the final iteration.

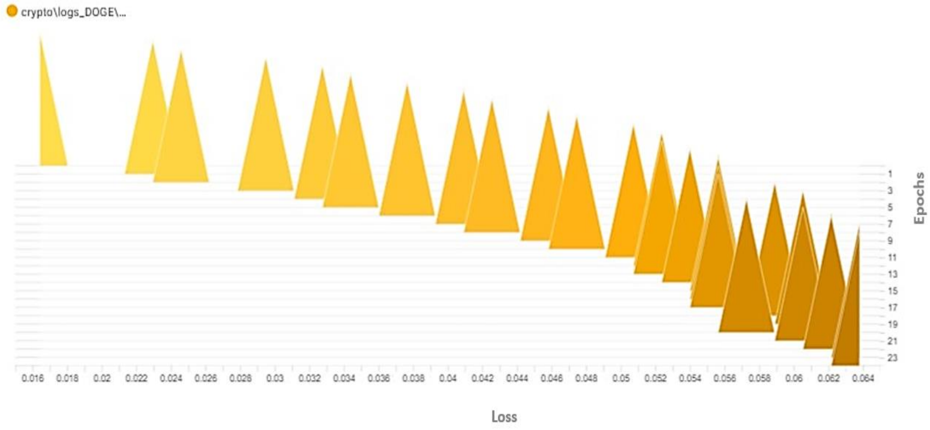

Figure 6: DOGE-USD

Figure 6 shows that the DOGE-USD variant has a stable convergence trend, but exhibits slightly more fluctuation than the collection. Losses tend to decrease, with validation loss following a delay of a few more updates. This indicates moderate volatility in DOGE data, but good model generalization.

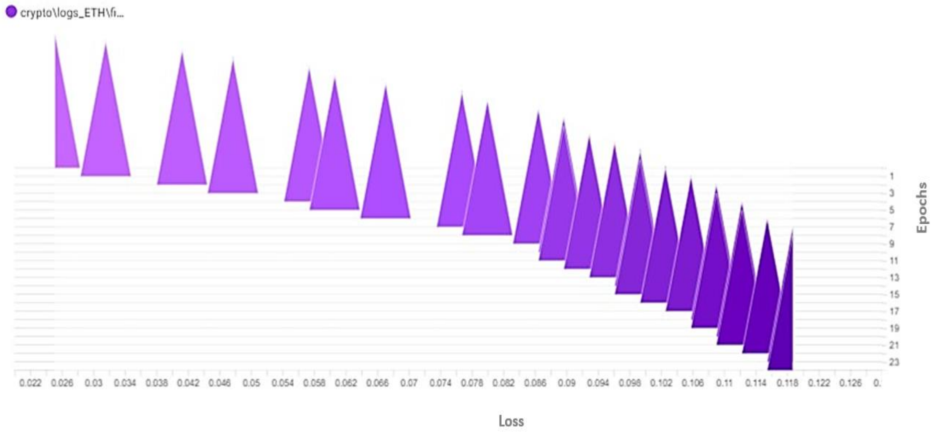

Figure 7: ETH-USD

Results for ETH-USD also suggest strong training behavior as both losses decrease and plateau by epoch 60 as shown in Figure 7. The validation loss closely follows the training loss, indicating that little overfitting is occurring. This model consistently performs well in predicting Ethereum.

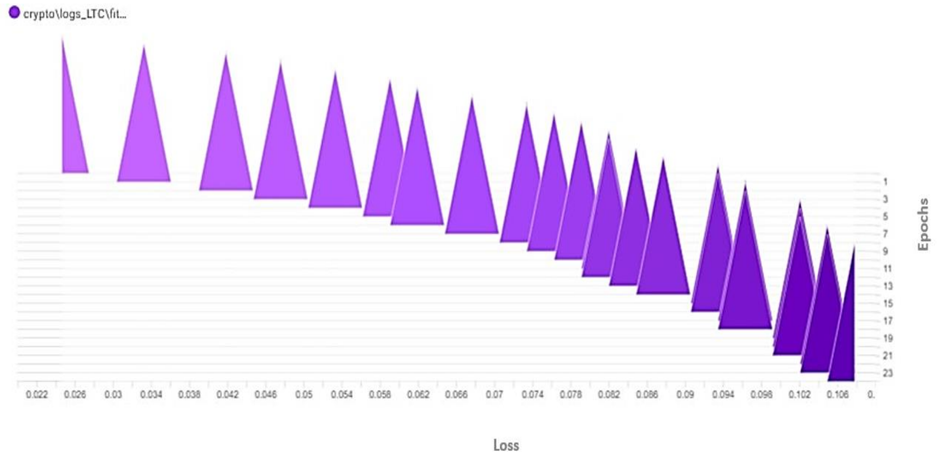

Figure 8: LTC-USD

Figure 8 shows that LTC-USD decreases sharply in the early epochs and remains stable afterwards. The training and validation curves are very closer, fitting well. This indicates the model's proficiency in learning past trends of Litecoin.

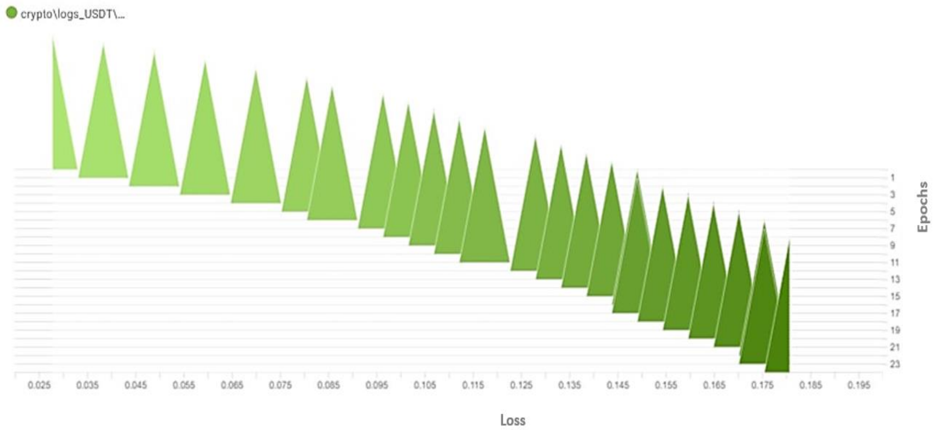

Figure 9: USDT-USD

The USDT-USD model exhibits a lower loss during training, as indicated in Figure 9. Its early convergence and parallel curves manifests its good accuracy and generalization properties. This is indicative of stability and predictability of the Tether dataset.

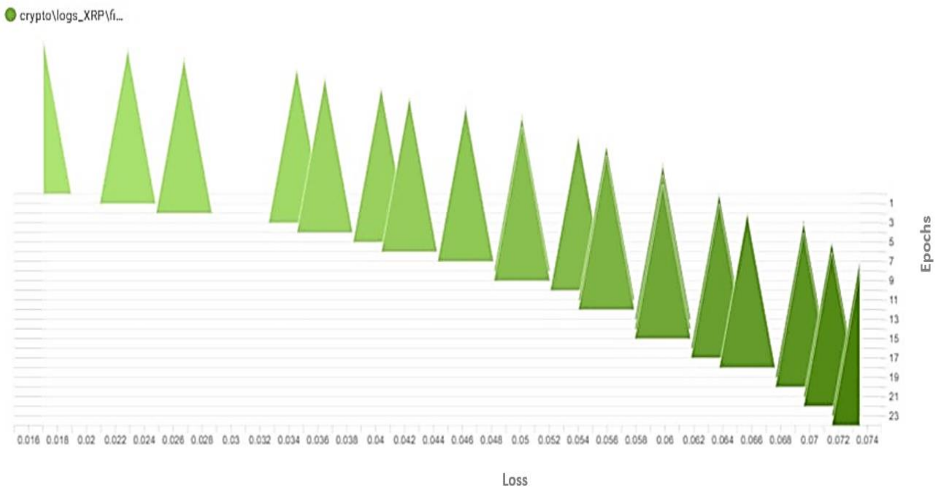

Figure 10: XRP-USD

Figure 10 depicts that XRP-USD is one of the cryptocurrencies with the most well-behaved convergence profiles, meaning there is very little gap between the training and validation losses. The model can achieve quick learning and high accuracy. This continuous loss trend exhibits significant generalisation on XRP data.

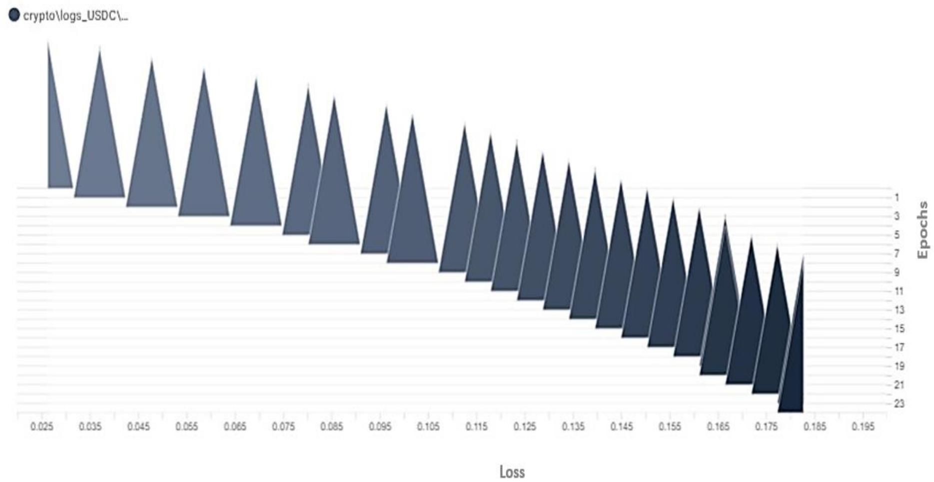

Figure 11: USDC-USD

For USDC-USD, the model trains well and consistently with evident convergence. The loss curves are nearly parallel, as shown in Figure 11. This is a sign of good learning and a properly balanced model with little overtraining.

Table 5: Baseline vs. Ensemble Model Performance

Cryptocurrency	Model	MSE	RMSE
BTC-USD	ARIMA	5.98	2.45
	Prophet	6.45	2.54
	Ensemble	**1.42**	**1.19**
ETH-USD	ARIMA	6.22	2.49
	Prophet	6.58	2.56
	Ensemble	**1.18**	**1.08**
ADA-USD	ARIMA	0.089	0.298
	Prophet	0.076	0.275
	Ensemble	**0.041**	**0.201**
BCH-USD	ARIMA	0.017	0.130
	Prophet	0.014	0.118
	Ensemble	**0.009**	**0.094**
BNB-USD	ARIMA	0.098	0.313
	Prophet	0.085	0.292
	Ensemble	**0.069**	**0.265**
DOGE-USD	ARIMA	5.23	2.29
	Prophet	4.85	2.20
	Ensemble	**2.73**	**1.65**
LTC-USD	ARIMA	8.42	2.90
	Prophet	7.56	2.75
	Ensemble	**6.07**	**2.46**
XRP-USD	ARIMA	3.34	1.83
	Prophet	3.09	1.76
	Ensemble	**1.90**	**1.38**
USDC-USD	ARIMA	2.53	1.59
	Prophet	2.41	1.55
	Ensemble	**1.43**	**1.20**
USDT-USD	ARIMA	0.006	0.077
	Prophet	0.005	0.071
	Ensemble	**0.003**	**0.054**

Table 5 presents the prediction errors of ARIMA, Prophet, and the proposed ensemble model for different cryptocurrencies. The sample that produces the minimum MSE and RMSE across all samples is the ensemble, which verifies that this sample has the best performance. It demonstrates the superiority of the hybrid learning model over classical time series forecasting techniques.

Table 6: Impact of Feature Removal on Ensemble Model Performance (BTC-USD)

Configuration	Included Indicators	MSE	RMSE
All Indicators (Baseline)	SMA, EMA, MACD, RSI	1.42	1.19
Without MACD	SMA, EMA, RSI	2.01	1.42
Without RSI	SMA, EMA, MACD	1.88	1.37
Without EMA	SMA, MACD, RSI	1.63	1.27
Without SMA	EMA, MACD, RSI	1.58	1.25
Only MACD and RSI	MACD, RSI	1.75	1.32
Only SMA and EMA	SMA, EMA	2.28	1.51

Table 6 presents an ablation analysis that evaluates the impact of performance using different technical indicators (SMA, EMA, MACD, RSI) individually to determine accuracy. The elimination of MACD and RSI has led to an extremely high RMSE, indicating their importance in prediction. However, these findings highlight the significance of momentum indicators in predicting cryptocurrency trends.

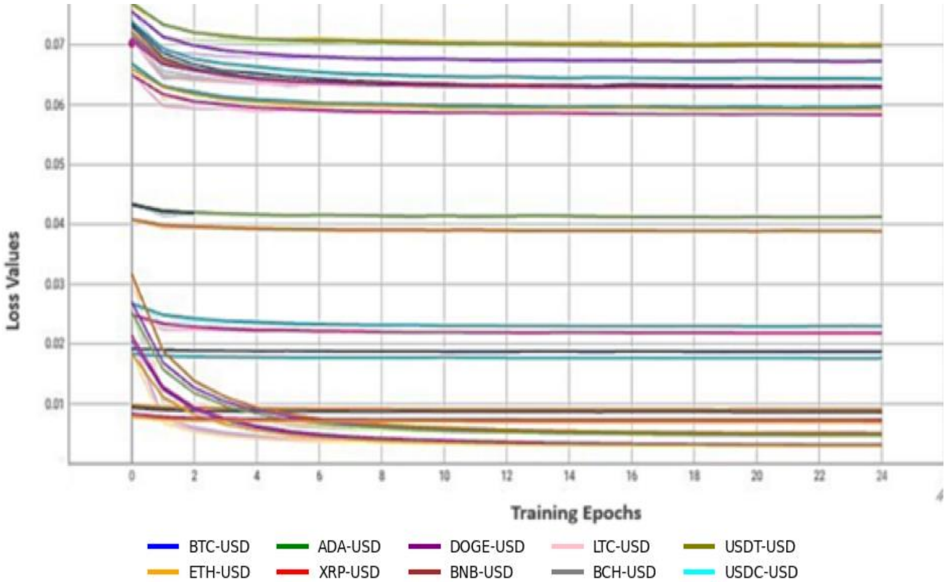

Figure 12: Loss Comparison Graph of Cryptocurrencies

The comparison graph of the loss values for each cryptocurrency employed in this study is displayed in Figure 12. The loss function of each cryptocurrency used in this investigation is illustrated in the graph above. Lighter lines represent the validation loss, while darker lines represent the training loss. The value of the loss per cryptocurrency reached its minimum as the number of epochs was getting increased. The losing value of a cryptocurrency decreases with the number of epochs. The models appear to have converged to a stable solution, as indicated by the comparatively stable loss curves.

Table 7: Computational Time

Algorithms	Mean	Standard Deviation (sec/loop)
RFRr	609.6599	15.7422
GRP	596.8471	29.3247
LSTM	1106.7652	30.5389
RNN	542.5957	27.5115

The calculation time of each machine learning and deep learning model is presented in Table 7. The standard deviation of the Random Forest Regressor used to indicate how predictable the time is has been determined as 15.7422 seconds per cycle, i.e., 609.6599 seconds to perform a repeating task. The standard deviation of variance for the Gaussian Regression

Process, expressed in seconds per loop, is 29.3247. With a somewhat higher standard deviation than the Random Forest Regressor, the projected average time of each cycle is 596.85 seconds. The results showed that the LSTM loops had an average duration of 1106.7652 seconds and a standard deviation of 30.5389 seconds for the number of iterations. The average time for each repetition in the RNN model was 542.5957 seconds. It is computed that the time taken for each loop iteration has a standard deviation of 27.5115 seconds. The model under consideration has a shorter mean time per cycle (about 542.60 seconds) compared to the LSTM. However, the standard deviation remains comparable, indicating a notable degree of variability in computing time.

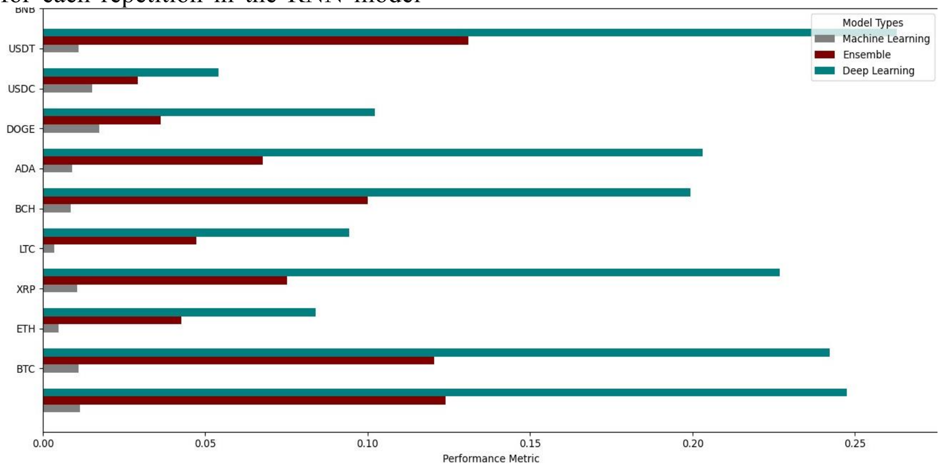

Figure 13: Comparison Graph of Machine Learning, Deep Learning and Ensemble Model

Figure 13 demonstrates that machine learning models produced the best outcomes among all the models employed in this study, as they achieved the lowest loss values. The ensemble model outperforms the LSTM and RNN deep learning models, yielding the second-best predictions for bitcoin prices.

5. CONCLUSION

In the field of data science, time series forecasting is a fundamental methodology used in the analytical processes of businesses and other organizations. Time series forecasting is a practice that uses a wide range of approaches and techniques, much like other data science methods. This paper presents a novel and ideal hybrid optimisation model that incorporates ensemble learning, specifically designed for time series forecasting. The model presented utilises mean normalisation and min-max scalar techniques to address the issue of missing data. To the best of our knowledge, previous studies in the field of real-time bitcoin prediction have not employed the hybrid optimization with ensemble learning approach described in our proposed model, despite its promising outcomes. However, the subject has recently garnered significant interest. This work aims to explore the usefulness of average ensemble models in machine learning and deep learning applications. It is noteworthy that, in this specific context, the solution we provide is more effective than any other technique previously reported. Several variables constrain the scope of our research, chief among them being the possibility that external influences such as news, legislation, or political issues could impact bitcoin pricing. This is the case even though our research has yielded some significant findings.

6. FUTURE WORK

Subsequent research will employ a broad range of machine learning models, including GRU, CNN, BERT, and Cubist, to assess the performance and resilience of bitcoin price prediction through hybrid optimization and ensemble learning. Furthermore, the model can be applied to various sectors, including cryptocurrency and stock markets, as well as any other datasets that can be characterised by time series. This allows for real-time monitoring and prediction. This increases its relevance and usefulness in equal measure.

7. REFERENCES

[1] Li, T., Shin, D., & Wang, B. (2021). Cryptocurrency pump-and-dump schemes. Available at SSRN 3267041.

[2] Chaim, P., & Laurini, M. P. (2019). Nonlinear dependence in cryptocurrency markets. The North American Journal of Economics and Finance, 48, 32-47.

[3] Watorek, M., Drozd˙ z, S., Kwapie˙ n, J., Minati, L., O´ swiecimka, P., &´ Stanuszek, M. (2021). Multiscale characteristics of the emerging global cryptocurrency market. Physics Reports, 901, 1-82.

[4] Swartz, L. (2018). What was Bitcoin, what will it be? The technoeconomic imaginaries of a new money technology. Cultural studies, 32(4), 623-650.

[5] Jardine, E. (2015). The Dark Web dilemma: Tor, anonymity and online policing. Global Commission on Internet Governance Paper Series(21).

[6] Damodaran, A. (2012). Investment valuation: Tools and techniques for determining the value of any asset (Vol. 666). John Wiley & Sons.

[7] Peterson, R. L. (2016). Trading on sentiment: The power of minds over markets. John Wiley & Sons.

[8] Albayati, H., Kim, S. K., & Rho, J. J. (2020). Accepting financial transactions using blockchain technology and cryptocurrency: A customer perspective approach. Technology in Society, 62, 101320.

[9] Yu, Y., Zhu, Y., Li, S., & Wan, D. (2014). Time series outlier detection based on sliding window prediction. Mathematical problems in Engineering, 2014.

[10] Lim, B., Arık, S. O., Loeff, N., & Pfister, T. (2021). Temporal fusion¨ transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748-1764.

[11] Catania, L., & Grassi, S. (2017). Modelling crypto-currencies financial time-series. Available at SSRN 3028486.

[12] Liu, Y., & Zhang, L. (2023). Cryptocurrency valuation: An

explainable ai approach. Science and Information Conference.

[13] Hao, M., & Lenskiy, A. (2023). Short-Term Volatility Prediction Using Deep CNNs Trained on Order Flow. arXiv preprint arXiv:2304.02472.

[14] Jiang, Z., & Liang, J. (2017). Cryptocurrency portfolio management with deep reinforcement learning. 2017 Intelligent systems conference (IntelliSys).

[15] Amirzadeh, R., Nazari, A., Thiruvady, D., & Ee, M. S. (2023a). Causal Feature Engineering of Price Directions of Cryptocurrencies using Dynamic Bayesian Networks. arXiv preprint arXiv:2306.08157.

[16] Amirzadeh, R., Nazari, A., Thiruvady, D., & Ee, M. S. (2023b). Modelling Determinants of Cryptocurrency Prices: A Bayesian Network Approach. arXiv preprint arXiv:2303.16148.

[17] 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.

[18] Haritha, G., & N.B, S. (2023). Cryptocurrency Price Prediction using Twitter Sentiment Analysis. ArXiv, abs/2303.09397.

[19] Oyewola, D. O., Dada, E. G., & Ndunagu, J. N. (2022). A novel hybrid walk-forward ensemble optimization for time series cryptocurrency prediction. Heliyon, 8(11).

[20] Dunnala, S., Bandla, A., Sunkara, K. S. A., & Jangam, E. (2022). Predicting the fluctuations of the bitcoin using machine learning. AIP Conference Proceedings.

[21] Indulkar, Y. (2021). Time series analysis of cryptocurrencies using deep learning & fbprophet. 2021 International Conference on Emerging Smart Computing and Informatics (ESCI).

[22] Sheta, A. F., Ahmed, S. E. M., & Faris, H. (2015). A comparison between regression, artificial neural networks and support vector machines for predicting stock market index. Soft Computing, 7(8), 2.

[23] Mudassir, M., Bennbaia, S., Unal, D., & Hammoudeh, M. (2020). Timeseries forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. Neural computing and applications, 1-15.

[24] Madan, I., Saluja, S., & Zhao, A. (2015). Automated bitcoin trading via machine learning algorithms. URL: http://cs229. stanford. edu/proj2014/Isaac% 20Madan, 20.

[25] Struga, K., & Qirici, O. (2018). Bitcoin Price Prediction with Neural Networks. RTA-CSIT.

[26] Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.

[27] Schulz, E., Speekenbrink, M., & Krause, A. (2018). A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. Journal of Mathematical Psychology, 85, 1-16.

[28] Hamayel, M. J., & Owda, A. Y. (2021). A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. AI, 2(4), 477-496.

[29] Yi, D., Bu, S., & Kim, I. (2019). An enhanced algorithm of RNN using trend in time-series. Symmetry, 11(7), 912.

[30] Bouteska, A., Abedin, M. Z., Hajek, P., & Yuan, K. (2024).

Cryptocurrency price forecasting–A comparative analysis of ensemble learning and deep learning methods. International Review of Financial Analysis, 92, 103055.

[31] Asuquo, M., & Umoren, I. (2024). A Hybrid Machine Learning Model for Clustering and Prediction of Closing Price of Cryptocurrency. International Journal of Network and Communication Research, 8(1), 1-22.

[32] M. I. Sarwar et al., "Data Vaults for Blockchain-Empowered Accounting Information Systems," in IEEE Access, vol. 9, pp. 117306-117324, 2021, doi: 10.1109/ACCESS.2021.3107484