Cryptocurrencies Analytics with Machine Learning and Human-centered Explainable AI: Enhancing Decision- Making in Dynamic Market

Article *in* International Journal of Computer Applications · January 2025

DOI: 10.5120/ijca2025924418

Ubaida Fatima
NED University of Engineering and Technology

Cryptocurrencies Analytics with Machine Learning and Human-centered Explainable AI: Enhancing Decision-Making in Dynamic Market

Kinza Muneer
Department of Applied Mathematics,
NED University of Engineering & Technology,
Karachi, Pakistan

Ubaida Fatima, PhD
Department of Applied Mathematics
NED University of Engineering & Technology
Karachi, Pakistan

ABSTRACT
High volatility, common among often referred-to cryptocurrencies, is determined by numerous direct and indirect factors, and thus one of the main issues of such prediction. Such inherent instability most times causes investment uncertainty throughout the digital currency market. Looking at recent years, the forecasting of cryptocurrencies' prices has received a high level of significance owing to their high levels of volatility. This study focuses on predicting the prices of three major cryptocurrencies: These cryptocurrencies include Bitcoin, Ethereum, and Litecoin. In this regard, three ML algorithms: LSTM, SVM, and RF as well as LSTM-RF ensemble were employed to enhance the predictive precision of the predicted cryptocurrencies. Comparing all the models investigated, the hybrid LSTM-RF model exhibited the highest accuracy in predicting the relevant performance indices and outperformed the other traditional models and other single machine learning methods. In addition, this study also adopts Explainable Artificial Intelligence (XAI) approaches to create AI-generated interpretable human-centric visualizations. This approach opens the insights to simple users as it helps investors make sound decisions depending on the results given by the model.

Keywords
Bitcoin dataset, Machine learning prediction, Explainable artificial intelligence, Hybrid LSTM - RF, Cryptocurrency, SHAP.

1. INTRODUCTION
In today's globalized globe, there is no isolated country, market, or economy. Interconnectivity has become even more important to economic systems, including digital marketplaces, conventional financial markets, and macroeconomic patterns. Although cryptocurrencies constitute an emerging digital commodity, cryptocurrencies are currently playing a significant role in the global financial system [1].
The world of cryptocurrencies started with the release of Bitcoin (BTC) 1 by an individual or a collective using the pseudonym Satoshi Nakamoto (2008). It was supposed to be an alternative and decentralized payment system – digital cash. At the moment, the use of blockchain is growing, with more and more new innovative applications, and cryptocurrencies represent an attempt to decentralize various financial instruments [2].The cryptocurrency market is amongst the fastest growing of all the financial markets in the world. Unlike traditional markets, such as equities, foreign exchange and commodities, cryptocurrency market is considered to have larger volatility and illiquidity [3]. Bitcoin and other cryptocurrencies have emerged as game-changing technology in the online financial system, addressing the problem of verification risk in transaction settlements.

1.1 Blockchain Technology

1.2 Introduction to Cryptocurrency Analysis:
Cryptocurrencies are a form of digital currency that is supported by cryptography to instill security and avoid fraudulence. It is important for researchers and investors to predict the possibility of fluctuations in cryptocurrencies with great reliability. However, since the cryptocurrency market exhibits irregularity, it becomes cumbersome to analyze the little peculiarities of time-series data, thereby making it impossible to produce accurate predictions of the prices. However, as the cryptocurrency market shows irregularity, it can be difficult to evaluate the distinct features of time-series data, making it impossible to generate precise price predictions. In many uses of studies focusing on cryptocurrency price prediction, several Deep Learning (DL) founded algorithms have been applied [5].
Cryptocurrency price prediction storm is one of the biggest problems that faces the financial industry. It is important to predict cryptocurrency prices as much as possible because the prices fluctuate so quickly depending on aspects such as technology, policies and other events in the global market. One of the prominent areas of research, based on survey results both among researchers and finance professionals, is development of accurate models of price prediction, models employing elements of machine learning and deep learning, in particular. [6]
However, there is one big difference about cryptocurrencies that differentiates them from centralized currencies: cryptos do not have a centralized authority that issues them – in other words, they need not rely on traditional banks. Cryptocurrencies are said to share all of the features of the blockchain as the blockchain is fundamental to them. With Bitcoin, for example, users may easily identify the sender and receiver of a transaction while conducting safe digital transactions in a pseudo-anonymous manner. Governments all around the globe are becoming more interested in blockchain technology, which has prompted calls for regulation of the cryptocurrency industry. The reasons for this governmental involvement are rooted in worries about opportunity, sovereignty, and criminality. [8]
Cryptocurrency offers benefits such as reduced transaction costs, faster settlement times, and increased transparency in transactions. [9]. The cryptocurrency market has undergone rapid development and is considered one of the fastest-growing financial markets globally [1].

Cryptocurrencies sound more volatile than traditional assets, with periods of extreme volatility and speculative activity. These features are caused by several factors. Unlike traditional currencies, cryptocurrencies are not issued by central banks and, hence, this is also not considered under any single country budget. Their worth is not attached to any actual assets. In order to manage volatility, need a proper system that can accurately foresee it [10].

1.2.1 Bitcoin*:*

Bitcoin, the pioneering and widely recognized cryptocurrency, was created in 2009 by the intelligible figure known as Satoshi Nakamoto. Develop as a decentralized form of digital currency that enables direct transaction between peers eliminating the need for intermediaries like bank or financial institutions. Blockchain technology is used to operate bitcoin to ensure security, transparency, and verifiable transactions. Its main functions include acting as a store of value, a medium of exchange, and a safeguard against the traditional financial system. The invention of bitcoin marked a significant milestone in the evolution of digital currencies, showcasing the potential of blockchain technology in revolutionizing financial transactions [2].

1.2.2 Ethereum

Ethereum, launched in 2015 by Vitalik Buterin and a team of developers, represents a significant evolution in the realm of blockchain technology and cryptocurrencies. While Bitcoin is mainly a digital currency, Ethereum is a technology that offers a platform for software developers to create smart contracts and decentralized applications (dApps). Its native digital currency is Ether (ETH) which is employed to execute business and computational solutions on the platform. According to the data of early 2022, market capitalization of Ethereum reached over $450 billion and Ethereum is among the top five most popular cryptocurrencies in the world [9].

1.2.3 Litecoin USD (LTC-USD)

Litecoin is a decentralized virtual currency known as a peer-to-peer internet currency which started in October, 2011 by Charlie Lee. The purpose was to facilitate fast, safe, and cheap payments, due to its innovative hashing system called Script, which consumes more memory than SHA-256, prevalent in Bitcoin mining. This difference enables Litecoin to create blocks far much more quickly, that is, every 2.5 minutes compared to bitcoin’s 10 minutes making let’s better for transactions. In addition, the Litecoin is more deflationary because it has a maximum supply of eighty-four million as compared to twenty-one million for Bitcoin [13].

In this research, we utilize three cryptocurrencies to identify the most effective machine learning model for price prediction.

1.3 Machine learning vision in Cryptocurrency

The scientific study of statistical models and techniques used by computer systems to carry out specified tasks without explicit programming is known as machine learning, or ML. Machine learning, or ML, is the process of teaching machines how to process data more effectively. Occasionally, even after examining the data, researchers are unable to understand the information that has been extracted. Then, utilize machine learning. The need for machine learning is growing due to the number of datasets that are available. Machine learning is used by many sectors to retrieve pertinent data. Learning from the data is the aim of machine. Numerous research projects have been conducted on the subject of teaching computers to learn on their own without explicit programming. Numerous mathematics and programmers use a variety of techniques to solve problems, which involves large data set [11].

Machine learning is a subset of artificial intelligence technique that predicts the future by analyzing past data. From this perspective, developing a machine learning model on previous bitcoin price data may allow for the prediction of future price movements with some accuracy. Prior research has revealed that machine learning-based approaches provide a variety of benefits over conventional fore. [9].

Machine learning plays a crucial role in predicting cryptocurrency prices. It is particularly useful for managing portfolios and anticipating price changes, making it an essential tool for financial decision-making [12].Machine learning techniques work more accurately to forecast volatility of cryptocurrency as compared conventional statistical models because ML models don’t require previous assumptions about data sets and can capture nonlinear patterns in time series and perform well accordingly [10]. The primary distinction between machine learning (ML) and classical modeling is that ML algorithms understand the data on their own; initial data decomposition is not required. These algorithms "build" logic modeling using the given data, depending on the analysis's goal. By doing this, the difficult and drawn-out pre-model phase of statistical testing of several hypotheses is omitted [13].

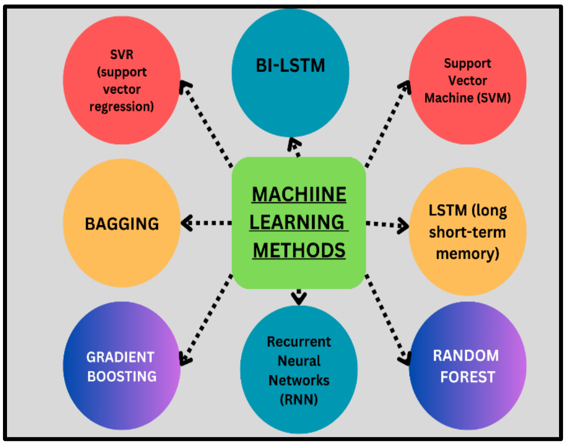

Figure 1:Popular Machine learning methods for Real-life datasets

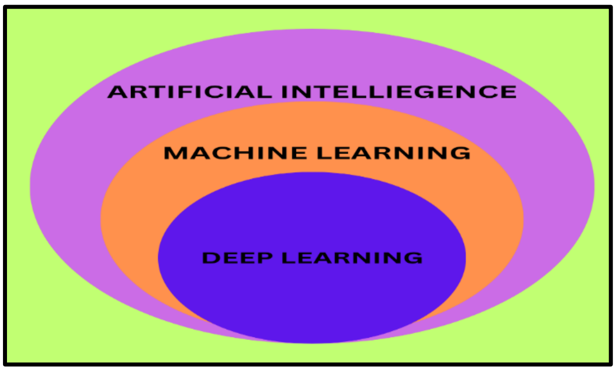

Figure 2:Machine learning and deep learning are subsets of AI

1.4 Artificial Intelligence:

AI and algorithmic decision-making are having a significant influence on daily lives [14].AI technology has a wide range of applications, including astronomy, healthcare, gaming, data security, social media, travel, transportation, and the automobile sector [15], and also having a revolutionary influence on a variety of professions and sectors After 70 years of study and development, complex AI systems are now facing external challenges such as security and ethics. Embedding AI in socio institutional systems is difficult on several levels, as explored by Sileno, Boer, and van Engers. Adversarial assaults on AI systems arise when malevolent actors (for example, software or humans) attempt to alter data, resulting in incorrect choices (such as misclassification or poor clusters). Furthermore, existing AI implementations are black-box solutions, thus the requirement for credible explanations of possible biases and adversarial assaults is fast rising [16].

1.4.1 Trustworthy AI

Prorating SHAP the rapid development of AI technology has resulted in the deployment of a variety of systems, which, although creative, frequently display flaws such as attack vulnerability, biases against disadvantaged groups, and insufficient privacy protection. These concerns can degrade user experience and diminish confidence in AI systems, making the development of trustworthy AI critical.

Key Aspects of Trustworthiness

1. **Robustness**: AI systems must be resilient against adversarial attacks and capable of maintaining performance under diverse conditions.

2. **Generalization**: Effective AI should perform well not just on training data but also on unseen data, ensuring its applicability in real-world scenarios.

3. **Explainability**: Users should be able to understand how AI systems make decisions. This transparency is vital for trust and accountability.

4. **Transparency**: Clear communication about how AI systems operate and make decisions is essential for fostering trust among users.

5. **Reproducibility**: AI models should produce consistent results when tested under the same conditions, which is crucial for validation and reliability.

6. **Fairness**: AI systems must be designed to avoid biases and ensure equitable treatment across different demographic groups.

7. **Privacy Preservation**: Protecting user data is paramount; AI systems should incorporate measures to safeguard personal information.

8. **Accountability**: There should be clear lines of responsibility for the outcomes produced by AI systems, ensuring that stakeholders can be held accountable for their actions.

Building trustworthy AI is not only a technical challenge but also a societal imperative. By focusing on robustness, fairness, transparency, and accountability throughout the lifecycle of AI development, practitioners can foster greater trust in these powerful technologies, ultimately benefiting society as a whole [14].

1.4.2 Explainable AI(XAI)

XAI is actually techniques that enlighten people about the results of machine learning so as to enhance their understanding about these black-box AI systems.By incorporating this concept, human can utilize these AI tools to support their work since the concept of AI is 'human-in-the-loop'[29]. The increasing maturity of AI technologies, having achieved high levels of model accuracy and performance, has produced demands for greater clarity on how these methods function. A great number of AI models, especially those founded on deep learning, are black boxes into which their developers often do not venture explaining the logic behind certain decisions or predictions. A recent reason for the need to develop XAI stems from the problems encountered when one implements AI in different applications, and the matters of security, transparency and accountability pertinent. To further AI adoption, organizations need information about AI algorithms to trust them and use them responsibly. This is an especially important approach in fields like health, finance and crime, where it helps to explain the reason behind certain decision-making practices and their effect on the affected individuals and societies [30].

1.4.3 Human-Centered Explainable AI:

Humanist SHAP (SHapley Additive explanations) is devoted to translating a machine learning model interpretation to human-oriented one. This forms the basis of what is being developed as Human-Centered Artificial Intelligence or HCAI; the intent of which is to develop AI systems that are sensitive to the human requirement and appreciative of human values and human knowledge. Human in the loop approach, or human-active learning emphasizes that expert knowledge from a given domain should be incorporated during the construction of machine learning [31]. Schneiderman notes that his study of the human factors supportive of AI for decision making insists on the invention of technologies that bring special efficiency and enhance people's capacity, confidence, and creativity [32]. When users apply the HCML Frame to SHAP, the working process of machine learning models can be fully understood and explained from another perspective [33]. Introducing SHAP into HCML frameworks can be beneficial to explain technical complexity and make the usage of machine learning models less complicated to the users. Auletta et al. discuss the use of SHAP to explain human behavior choosing decision making and stressing the need to model human interpretable AI [34].

Human-centered SHAP is an improvement over previous forms of explainable AI in that it is both interpretable and efficient and operate on humanistic principles. When human insights are applied to the creation and application of SHAP, machine learning and the use of AI generally is more accepted and practical and thus yield better results across various domains.

1.5 Research Objectives:

- Promises an effective algorithm for forecasting digital assets' prices based on previous experience.
- Before implementing these predictions, it is recommended to make them Human-centered Explainable AI (XAI) such as SHAP.
- Create and compare new types of hybrid forecasts to enhance the accuracy and reliability of the cryptocurrency price predictions, with reference to applying typical machine learning.

2. RELATED REVIEW

2.1 Review of Papers on Machine Learning for Price Prediction:

Koo et al. 2024 proposed a new approach towards the forecasting of prices of cryptocurrencies. To overcome the difficulties and the nonlinear tendencies that include the topic of the analysis, which is the Bitcoin value dynamics, the work uses a central decomposition technique together with LSTM networks. These price swings depend on several components inclusive of: market sentiment, trading volume, as well as other economic indicators. The authors use Empirical Mode Decomposition (EMD) to decompose Bitcoin price data into intrinsic mode functions (IMFs). This would allow LSTM models to treat different data components differently which helps to produce better predictions. They have proved that the overall performance of this method over traditional LSTM models and other machine learning algorithms in metrics such as the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE). These results are in concordance with previous researches which indicates that LSTM possess the property of identifying long term dependencies in time series data as pointed out in [18]. The conclusions of the study are significant, pointing that decomposition methods improve the accuracy of prediction significantly in the unpredictable markets such as cryptocurrencies [36].

Zubair et al. (2024), this study develops an improved approach that examines cryptocurrency prices through a futuristic hybrid model that incorporates LSTM and GRU networks with sentiment analysis driven by BERT and VADER. To their credit, the authors make an attempt to explain the inherent difficulties arising from the extremely fluctuating nature of the cryptocurrency markets in effecting drastic swings and altering investors' behavior. The RTP called for in this paper employs the historical price information of digital currencies to train the LSTM-GRU hybrid model so that it can identify sophisticated time dependency patterns in the data. Finally, in the sentiment analysis part that uses both BERT and VADER models, one can determine the reliability of certain price predictions based on the sentiment expressed in social media and news articles.. This combination does not only enhance the accuracy of the forecast of the price but also helps investors increase decision-making probability and manage risks in the unstable cryptocurrencies' market [38].

Dudek et.al (2024) This paper aims at examining various approaches towards predicting changes in prices of the major cryptocurrencies such as bitcoin, Ethereum, Litecoin and Monero. HAR, ARFIMA, GARCH, LASSO, ridge regression, SVR, MLP, FNM, RF, and LSTM models are used to see which models are the most effective for daily and weekly forecasts. As with other trading models, these types of models are evaluated according to their skills in forecasting daily price changes. One of the greatest conclusions is that there no one dominant model; all models are superior depending on the cryptocurrency, the type of prediction, and the type of error measurement used. FNM is efficient when making daily forecasts while ridge regression is efficient when used to make weekly forecast. The study also points out that more complex methods like LSTM and RF can be as effective as relatively simple models like HAR and ridge regression. In conclusion, the study demonstrates that even though sophisticated models may be employed, fundamental models are still able to give substantial outcomes, which create developments in forecasting the variability in bitcoin prices. Since most machine learning models act as a "black box" in the future studies, researchers might focus on model interpretability. Knowing why some of these forecasts are made could help traders and investors develop the needed faith and confidence in their projections [8].

Hamed et.al,2024 investigates the effectiveness of four advanced algorithms—Light Gradient Boosting Machine (LightGBM), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), and Extreme Gradient Boosting (XGBoost)—in predicting Bitcoin prices. Given the significant volatility in the cryptocurrency market, accurate price forecasting is essential for informed trading and investment decisions. The performance of these algorithms is assessed in terms of Mean Absolute Error & Root Mean Squared Error for the purpose of prediction in the study. LSTM and BiLSTM, used to learn long-term dependencies in time series data, are selected for comparison with LightGBM and XGBoost, which show better results in dealing with large amounts of data. In the proposed experimental design, these models will be trained and tested on historical Bitcoin prices enabling assessment of their predictive power. This research extends theoretical and practical understanding by establishing that state-of-the-art machine learning dramatically improves forecasting accuracy over conventional methods, suggesting that it can help deepen insights into cryptocurrency trading decision-making processes. The findings of this comparative study reveal the efficacy of each algorithm and identify future research directions that will be useful for developing similar models for financial forecasting [40].

Brini et.al,2024 the authors go further to explain some of the challenges that make the pricing of cryptocurrency options challenging through discussing the challenges arising from the high volatility of digital assets markets. It shows the limitations of main methods of option valuation, such as Black-Scholes and similar models, which presuppose constant volatility and normal distribution of returns, and points to the need for constructing more advanced approaches. The work suggests the extension of machine learning regression analysis to improve the functionality of cryptocurrencies when pricing options by using price history and other market characteristics. The outcomes prove the superiority of these machine learning models to the conventional pricing approaches since the results reveal the correctness of data analysis in modeling the nonlinear relationships that influence cryptocurrency prices. In addition, feature selection is also highlighted in the study because it concerns the selection of input variables for the machine learning models. The increased accuracy derived from this approach that has been underlined in this work has significant impacts upon risk management tactics when trading in cryptocurrency options; helping traders devise their trades based on accurate calculated future price probabilities. However, several research gaps have been left behind, even as different approaches continue to be developed. Future work should improve the comparison with the other advanced methods like Deep learning or ensemble approach to establish the suggested solution. Furthermore this work could be advanced by employing these models in real-time trading scenarios to understand their stability in live trading context. Other oversights imputable for low predictive accuracy included failure to incorporate sentiment analysis from social media and news articles could equally improve the results since market sentiment profoundly impacts cryptocurrency prices. Finally, it would be useful to examine how changes in the regulation affect the price of options and to study a possible approach to the use of machine learning models to comply with these regulations. Another type of research potentially useful in providing further insights into the operational stability and performance of machine learning models in different market conditions, would be longitudinal research comparing the results of various machine learning models across different market cycles.

Tripathy et al. (2023) The study evaluates three forecasting algorithms: ARIMA, LSTM, and FB-Prophet. To evaluate the accuracy of each model, the authors perform a model validation which involves a complete performance analysis that involves, MAE and RMSE among others. It can be seen from Fig 7 that LSTM model performs better than both ARIMA and FB-Prophet models based on the metrics of Mean Absolute Error and Root Mean Square Error, which substantiate the use of deep learning frameworks to model temporal sequences. To some extent, this study was epitomized by examining primarily Bitcoin. Possible future research could extend the study by adding assessments for other digital currencies to determine if the outcome is consistent [43].

Chen, (2023) The study shows that the machine-learning method, namely Random Forest is more effective than retrospective econometric models in the prediction of the Bitcoin's value. On this, the author notes that since Random Forest is able to capture non-linear correlations and interactions between the variables, the algorithm is efficient in explaining variation in Bitcoin price. The paper describes how data mining methods especially Random Forest algorithm performs better than conventional econometric models in forecasting values of Bitcoin. The author also insists that the utilization of Random Forest to reveal intertwined nonlinear relationships between variables explains the effectiveness of the methodology in capturing Bitcoin randomness. To continue, the paper analyses the avoidances of standard econometric techniques, which assume linearity in the model and latent non-linearity and volatility in the data of cryptocurrency markets. This has a strong implication for the use of higher-order machine learning techniques that in turn explain and predict prices [17].

Rathee et al. (2023) introduces a detailed framework for predicting cryptocurrency prices using advanced ensemble-based neural network methods. Acknowledging the challenges of high volatility and complexity in cryptocurrency markets, the study combines different neural architectures, including Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRU), within an ensemble learning approach. These predictions are not standard predicted prices but dynamic estimates that incorporate the movement of prices in the market by use of real time feeds. The ensemble method makes the technique more dependable through aggregation of outputs, reducing overemphases, and increasing effectiveness in cases with unseen data. The study also establishes the effectiveness of the proposed framework by using performance measures such Mean Absolute Error (MAE) & Mean Squared Error (MSE) outperforming conventional forecast methods. Also, the authors use outside data, mood from social networks and news, which allows them to consider supervisory and psychological influences of shifts in the market price—a substantial factor for the unpredictable cryptocurrency market. The results on the presented framework reveal its efficiency at capturing the intricate processes of the market, which can be helpful for investors and traders that operate in highly fluctuating markets. The paper also identifies areas for considering the hybrid models and the link between more data to be incorporated to enhance predictive accuracy and makes the way of the future research and development in cryptocurrency price prediction [45].

Akila.et.al. (2023) proposes a novel technique for the prediction of cryptocurrency prices, especially in the case of Bitcoin. To achieve enhanced prediction accuracy, the authors build upon cutting-edge deep learning approaches – the LSTM networks. Aware of the fact that these markets are very dynamic and volatile, and thus, tend to be hard to predict with a high degree of accuracy the present study includes the PELT, which stands for Pruned Exact Linear Time, change point detection algorithm in the training of the LSTM model. This flexible design enables the model to identify fundamental changes in the prices of products to suit the changing markets by incorporating shorter-term periods.These experimental findings show that the combined method is superior to LSTM models, showing higher performance Rates such as MAE, MSE, and RMSE. This evidence shows that combining deep learning with change point detection greatly improves prediction performance. This approach has great applicability to the real world: it gives more helpful instruments to investors, traders, and financial analysts, helping them to assess the situations in the highly unstable and risky cryptocurrency market [46].

Bangroo et al. (2022) studied the use of machine learning to predict the prices of crypto currencies, with specific focus on Random Forest, Support Vector Machines or Decision Trees among other. It is further important to note that price prediction is a delicate issue of which the authors discuss given the current volatile and unpredictable environment of the cryptocurrency market. To combat these problems the study uses a comprehensive dataset that comprises of price data and market conditions, which the models can analyze from the behavior which has been exhibited in the past. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are used to assess accuracy at which each algorithm performs. The results tell us that the best algorithm is the Random Forest one, which confirms that this algorithm can handle all the challenges related to fluctuations of cryptocurrencies' prices. The present study clarifies how machine learning can improve the accuracy of price estimations, giving traders and investors significant instruments for functioning in an increasingly unpredictable cryptocurrency market [50].

Derbentsev et al. (2021) emphasize the increasing importance of cryptocurrencies in the financial sphere and the challenges that relate to their price fluctations. To the authors' concern, outstanding basic techniques in forecasting may not satisfactorily capture the complexity of price movement in cryptocurrencies, this triggers them tow research on machine learning as a potential solution. Analyzing historical price of multiple cryptocurrencies, the researchers employ a number of ensemble methods like Random Forest, Gradient Boosting and others. The authors fully assess these models utilizing performance metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), as well as R-squared values. Analyzing the measurement data, I found out that the ensemble methods, especially Gradient Boosting, have better prediction ability than traditional models. Finally, as a result of the study, the critical role of feature selection and data preparation for enhancing the performance of the developed models is highlighted. In the future, extending the study would help to investigate the question of whether the tendency is constant when a wider range of digital assets is examined [53].

Hamayel et.al (2021) offers a comprehensive analysis of the forecasting of bitcoin prices with the use of new approaches to the use of machine learning techniques. Three selected digital currencies for the writers are; Ethereum, Litecoin, and Bitcoin.

To predict the future values of various cryptocurrencies, the study utilizes three different types of recurrent neural network (RNN) architecture: bidirectional Neural network model, Long Short-Term Memory model, and GRU model. Each of the models is trained using associated historical price data. During evaluation of each model, the Mean Absolute Percentage Error also known as MAPE is used. The experimental results demonstrate that the proposed GRU model is more compatible with the cryptocurrencies' data than both LSTM and bi-LSTM models for each of the three chosen cryptocurrencies and highlight its efficiency to capture the temporal patterns underlying the changes in price. The authors note that because of the high volatility of crypto-markets the more efficient learning process is a possibility by the help of GRU architecture. For even greater precision of the resulting predictions the authors note in their conclusion the opportunities for more endeavor in this area particularly concerning the combination of multiple machine learning techniques [55]

The study from **Trans et.al** gives the first hint and comes with a combined approach based on Random Forest (RF) and Long Short-Term Memory (LSTM) networks. This model is of considerable importance given that it aims at predicting optimal solutions of local weather conditions within a one-hour time horizon using localized weather sensors for data gathering. Appropriate to that, the authors underscore the need for optimizing the computational complexity for which the model yields higher accuracy here note that IoT applications may be resource-constrained17 [18].

2.2 Research Gap:

Altogether, the reviewed papers underline the progress in using machine learning methods for price prediction of cryptocurrencies, and the papers are mainly devoted to different algorithms, including LSTM, GRU, Random Forest, and ensemble techniques. While these studies demonstrate significant improvements in predictive accuracy compared to traditional statistical methods, A notable research gap exists to applied of hybrid models, so in this paper uses hybridization of Long short-term memory (LSTM) and Random Forest RF to increase the accuracy of model to predict the price of cryptocurrency. This approach would not only improve model performance but also provide a more comprehensive understanding of the factors driving price dynamics in the volatile cryptocurrency market. RNN models like LSTM and ensemble methods like Random Forest (RF) have not been hybridized in prior research so in this research create hybrid model combining LSTM and RF.

2.3 Papers implementing Explainable AI:

Taha et.al,2023 aims at investigating the application of XAI with ml models to improve the interpretative power of financial time series prediction models. Introducing XAI drastically enhances the model interpretability specifically in the context of predicting financial time series. It does this by enabling stakeholders like financial analysts and decision makers to get the rationale behind predictions therefore enabling them to trust the new automated systems. This research further shows that if EMLs are used alongside explainability frameworks, there is an improvement or at worst, no degradation in the predictive results. The researchers were able to understand key features influencing the model's predictions through model-agnostic techniques such as SHAP (Shapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations). This knowledge proves that including the outlined machine learning models with explainability frameworks retains or improves predictive performance. Using methods like SHAP (Shapley Additive explanations) and LIME (Local Interpretable Model-agnostic Explanations) the researchers highlighted important features to the prediction aiding in refining the model. The research outcomes show that not only does XAI help illuminate model behavior, but it also improves user interaction with the prediction process. When users could understand how different factors impacted the level of the model, they stated that they gained more confidence in trading decisions. The paper ends with specific recommendations to practitioners to use XAI in conjunction with conventional machine learning in financial forecasting applications. This approach can result in more dependable as well as credible solutions needed for regulation and ethics [49].

Gupta et.al (2023) also reveal that using supervised learning; regression and classification models there was prediction of more than the controllable accuracy levels of the supervision models. Other techniques of learning without supervision such as clustering and anomaly detection also provide information on the behaviors and patterns of the market that can complement the direct approach of supervised learning. The author discusses various XAI methods such as SHAP and LIME which can be useful to interpret the decision making of a range of supervised and unsupervised models. This integration is particularly crucial in building consumer confidence and understanding of AI in trading systems, especially on issues of decision making in matters related to finances. Although the present study compares supervised and unsupervised learning techniques, it does not address the question of whether it is possible to combine the two kinds. There could also be blended techniques that incorporate some utilities from both the supervised and unsupervised classes of techniques to enhance the forecast acumen as well as offer more perception into the market shifts. Future investigations might explore whether integrated application of these approaches lead to enhanced anticipatory schemes [19].

Mandeep et.al (2022) the author discusses how can XAI be further integrated with the machine learning algorithms in order to improve the prediction of the stock markets. The authors address the issues of non-linearity of stock prices and problems of either poor forecast quality or interpretational accessibility of the existing linear models. It divides methods of prediction into fundamental and technical analysis providing the highest interpretability of results but often failing to analyze and predict market conditions accurately enough due to their complexity and AI methods including random forests and neural networks providing the highest accuracy but being almost opaque to the analyst. As for this interpretability problem, the authors suggest using the XAI tools, which are LIME and SHAP in this work that can explain how a machine learning model reaches a decision. The results of experiments indicate that related XAI techniques do not only improve aspects of the interpretability of machines but also increase the participants' trust and confidence on them. As a result of exposition, XAI helps establish strong relations between input and output in stock prediction, which will serve a key purpose of effective investment. The study discusses the application of the machine learning with human-AI explainability to improve, the financial forecasting tools in decision making and risk management of the financial context [51].

John W.Goodwell et.al (2023) looks at the problem of explaining forecasts of cryptocurrency prices and the most popular one, Bitcoin. The authors present a new XAI framework which combines a new feature selector with SHAP, an XAI technique based on game theory, to improve both accuracy and interpretability of price prediction. The authors present a new XAI framework that integrates an innovative feature selection technique with SHapley Additive exPlanations (SHAP), which improved the accuracy and interpretability of machine learning models arising from the proposed methodology. It was also shown that this approach can accurately forecast the Bitcoin prices during the bearish condition like the one caused by the Russian-Ukraine war thereby asserting the model's ability to remain unscathed from potential macroeconomic shocks. In addition to that, modeling shows factors of daily financial and macroeconomic significance relevant to shift between state of low and high Bitcoin prices, useful for investors. The framework also demonstrated enhanced versatility and can be used for other cryptography and financial related systems.However, despite these advancements, the study reveals a gap in existing literature regarding the application of XAI in real-time trading environments, indicating a need for further research to explore how these explainable models can be effectively utilized in dynamic market conditions for timely decision-making [20].

2.4 Research Gap:

While Explainable AI (XAI) is becoming increasingly popular across various fields, human-centered approaches remain largely unexplored around cryptocurrency price prediction. Tools like SHAP (SHapley Additive exPlanations) are effective for explaining models, but they primarily focus on enhancing algorithmic transparency without fully addressing the human perspective or the decision-making needs of users.

This opens an opportunity to integrate human-centered AI methods alongside tools like SHAP. Doing so could provide more meaningful and user-friendly explanations tailored to the thought processes and requirements of traders, investors, and analysts. Bridging this gap would not only make XAI applications more intuitive but also improve trust and support better decision-making in cryptocurrency price forecasting.

3. METHODOLOGICAL FRAMEWORK

3.1 Dataset Description:

The dataset comes from Yahoo Finance, containing historical price data for various cryptocurrencies, including Bitcoin (BTC), Ethereum (ETH) and Dogecoin. It provides daily price information over a specified time period, enabling the analysis of price patterns, volatility, and market behavior. Features contributing to this research are:

Open price: the first price at which a cryptocurrency is traded after the opening time of the session.

Close price: price at which has been ultimately traded before the closing time deadline.

High price: highest price at which the cryptocurrency has been traded during the season.

Low price: lowest price at which the cryptocurrency has been traded during the season.

Adjusted close: in stock market, it amends the close price considering other factors like dividends, stock splits, or offerings of new stocks. In the Cryptocurrency market it does not reflect any changes with respect to the close price.

3.1.1 Bitcoin Dataset:

Conserved Bitcoin historical data from www.finance.yahoo.com. From date 01-12-2017 to 30-07-2024 for observing daily values and predicting price having 2433 observations. The historical Bitcoin price data obtained from the Bitcoin database included not only the closing price but also the high, low, open, and adjusted close, though the closing price of each day was chosen to be used in the prediction models due to its relevance in calculating the metrics under study.

3.1.2 Ethereum Dataset:

Ethereum data was retrieved from www.finance.yahoo.com and downloaded using the open-source library yfinance in Python. The dataset covers the period from December 1, 2017, to July 30, 2024, and includes a total of 2433 observations.

3.1.3 Litecoin Dataset:

Litecoin data was retrieved from www.finance.yahoo.com and downloaded using the open-source Python library yfinance. The dataset spans the period from December 1, 2017, to July 30, 2024, comprising a total of 2433 observations.

3.2 Min-Max Normalization:

Standardizing the values of cryptocurrency which involves transforming the data to a scale of 0 to 1 is regarded as an important data preprocessing step in the data analysis as well as in machine learning since the data involves hiked numbers of volatility and varying scales of cryptocurrency prices. This method ensures that all features contribute similarly to the analysis which is crucial for those algorithms based on distance computations or gradient descents [58]. One of which is Min-Max normalization which involves scaling of each value in the dataset through the formula:

$$Xnorm=\frac{X-Xmin}{Xmax-Xmin} \qquad (6)$$

Standardization is done in a way to bring all data to the same range optimal for the algorithms or ideally in the range 0-1. This step not only improves the goodness of the prediction but also the speed and reliability of training which makes it as one of the crucial steps in the data pre-processing stage of this research.

3.2.1 Machine Learning Methods:

In this study for predicting the price of various cryptocurrencies three ML models and two Hybrid models were employed to compare the value of various cryptocurrencies using Mean absolute error (MAE), Root mean square error (RMSE), Mean squared error (MSE) and R-square. According to the available research and the attributes of several intelligent approaches, selected some Machine Learning (ML) methods for different cryptocurrencies' prediction. Machine learning is a subset of artificial intelligence techniques and algorithms that use data to better themselves automatically through training. Because of the variety of techniques and concepts it employs, this field of artificial intelligence is exceptionally adaptable and has a large capacity for computing.

3.2.2 LSTM for real-life dataset evaluation:

The financial time series' one-step-ahead prediction needs both recent and historical data. The recurrent neural network (RNN) model has advantages in terms of long-term dependability due to self-feedback in the hidden layer. However there are practical implementation challenges that exist [43]. The latest research works discussed the adoption of LSTM networks for forecasting cryptocurrencies with the disclosure that these networks are perfectly capable of handling the dynamics and the inherent volatility of the markets. The combination of historical and real-time data for identifying Candlestick patterns for Bitcoin Candlestick trading as well as the analysis of candlestick layouts also demonstrates that LSTM network excels in capturing temporal characters, which are not apparent to standard models [60]. In this research LSTM is applied on distinct cryptocurrencies for carrying out the comparative analysis of LSTM with several other ML techniques. In this model no. of epochs use=100,

$$i_t = \sigma Wi\left[h_{t-1,}xi\right] + bi \quad (1)$$

$$f_t = \sigma Wf\left[h_{t-1,}xi\right] + bt \quad (2)$$

$$c_t = f_t * ct - 1 + i_t * tanh\ (Wc\left[h_{t-1,}xi\right] + bc \quad (3)$$

$$O_t = \sigma(w_0[h_{t-1}x_i] + b_0) \quad (4)$$

$$h_t = Ot * tanh(Ct) \quad (5)$$

where xt is the input at time step *t*, ht is the hidden state at time step *t*, ct is the cell state at time step *t*, and it, ft, and ot are the input gate, forget gate, and output gate, respectively, at time step *t*. *W* and *b* are the weight matrices and bias vectors, respectively. The sigmoid function and the hyperbolic tangent function (tanh) are used to bound the output between 0 and 1, and between –1 and 1, respectively.

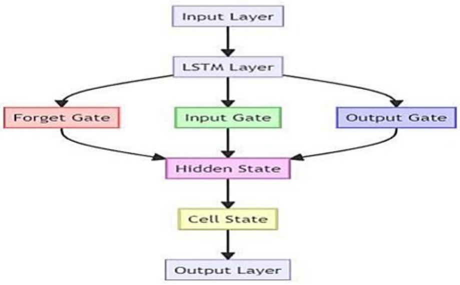

Figure 3:Structure Of LSTM

3.2.3 *SVM utilization on Real-world dataset:*

SVMs are also found to be a very effective model for classification of data, particularly in the fiendishly complex domain of cryptocurrencies where data is often complex and high-dimensional. Their capacity to model and solve multifaceted, nonsequential connections and always identify patterns has made them rather useful carrying out simulations and predictions of trends within this volatile market environment. For example, when used in concert with other machine learning approaches, the SVM has recently returned an average classification accuracy above 50% for cryptocurrencies' return predictions. This makes it possible to discover trends of predictability within the highly unstable context of Cryptocurrencies. [61]. The standard form of the Support Vector Machine (SVM) optimization problem for classification can be expressed as:

$$\underset{wb\varepsilon}{\boldsymbol{min}} \frac{1}{2}\|w\|^2 + C\ \sum_{i=0}^{n} \varepsilon i \quad (7)$$

yi (w . xi ─b)≥1─€ for i=1,...n
$\varepsilon i \geq 0$ >0 fori=1,...,n
Where:

w is the weight vector,
b is the bias,
ξi is the slack variables (for soft margin SVM),
The constant is C the regularization parameter.
Yi is value mark of i-th sample, while
Xi is the i-th data point
Finally, the SVM models simplified for optimization are used for forecasting or to update depending on signals from the monitoring systems aimed at improving the understanding of changes in price and trends in the prices of the digital assets.

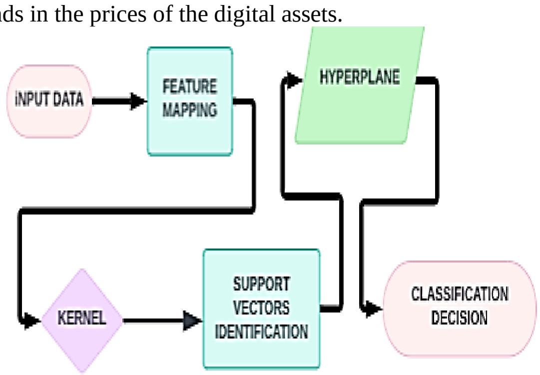

Figure 4:Structure of SVM

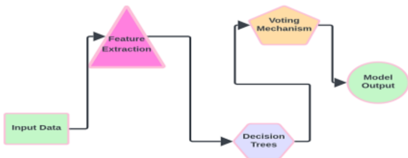

Figure 5:Structure of Random Forest

3.2.4 *Random Forest Applied to Real-World Dataset*

Random Forest (RF), which proved to be inherently effective at predicting prices and categorizing markets, has become so popular in the cryptocurrency market. For example, while Shen was working on his study, RF was used for the analysis of Bitcoin prices, and it proved to be more effective than the other machine learning algorithms. The model delivered outstanding results: Mean Absolute Error (MAE) is 138.39 and the Root Mean Squared Error (RMSE) indicating the stability and accuracy of the model in the extremely fluctuating industry [62]. Random Forest is ideal for the analysis of big-sized data containing numerous features, which is why it can be the best solution for complex and rather unstable cryptocurrency markets.

4 LSTM and RF Hybrid Model(As a novel approach in cryptocurrency)

Recently, the integration of LSTM with other approaches was implemented effectively using the RF algorithms in numerous predictive applications like time series forecasting or the financial market prediction. Using this combination of strategies, the strengths of both models are included, where LSTM works better in analyzing temporal dependencies and sequential data patterns, and Random Forest has stability and feature importance comparison.

The hybrid LSTM-RF model is distinguished by considerable improvement in reducing the prediction errors. For instance, Tran et al suggest that RF combined with LSTM can enhance short term local weather prediction accuracy and produced lower MAE, R-squared, and RMSE than individual models [56].

The hybrid LSTM-RF model can be described in two stages:

1. $t = f\,\mathrm{LSTM}(x1, x2, \ldots, xt)$ (8)
2. $\tilde{y}t = f\,\mathrm{RF}\ (xt, \hat{y}t)$ (9)

Here:

y t is the level of output that LSTM returns at time t.f LSTM stands for the value that LSTM gives as its future estimate represents the LSTM model's prediction

fRF is the Random Forest model of prediction based on the output of LSTM as well as extra variables from the input dataset.

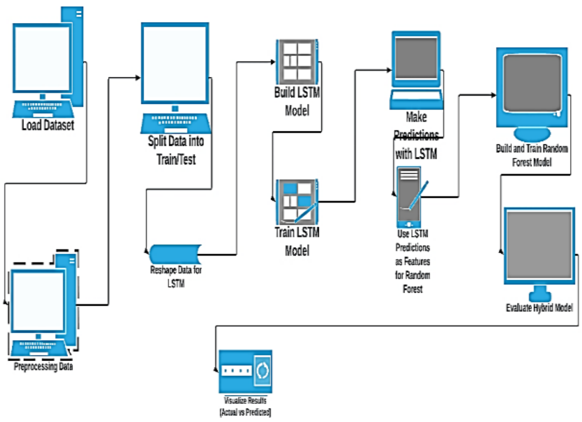

Figure 6:Workflow of Hybrid LSTM-RF

3.3 Human-Centered AI with SHAP: A Novel Approach in Cryptocurrency Analysis

Introducing synergies of human-anchored AI and the SHAP technique allows for designing an innovative approach to analyzing the cryptocurrency markets. This approach aims to improve interpretability, a critical aspect in financial decision-making since decision-makers should make decision based on model's predictions. Because of SHAP values, it is easier for researchers to see how various factors influence the outcomes of a model and build more trust in the information produced by AI. This is even more important in the highly fragile market such as that of cryptocurrencies where constant and frequent changes require detailed explanations for machine learning models.

Specifically, cryptocurrency through SHAP gives a deeper understanding of factors such as market sentiment, trading volume, and price history that impacts the predictions of the prices. Apart from addressing the imprecision in the models, this human-oriented approach allows the user to gain understanding of the data dependencies required to draw useful conclusions. Thanks to the steady advancement of the cryptocurrency market, creating and implementing human-centered AI interacted with SHAP as a potential way to improve the choice-making approach and develop trust in the automatic trading system. Classic SHAP offers an algorithmic method of splitting the contribution of each characteristic to the obtainable model result using the principles of the cooperative games theory, with particular reference to Shapley values.

$$\phi i = \sum_{S \subseteq N\{i\}} \frac{|S|!(|N|-|S|-1}{|N|} \left(f(SU\{i\}) - f(S)\right) \quad (10)$$

where:

N is the set of all features.

S⊂N and i∉S.

The vector f(S) shall refer to the result obtained with the model based solely on the features in S.

F(S∪{i}) is the value which the model provides after adding the third feature iii to the set S.

|S| represents the number or size of the subset S of Featurespace, and |N| is total number of features in Featurespace.

The factorial terms |S|! and (|N|−|S|−1)! ensure that the contributions of each feature are fairly weighted based on the number of subsets they appear in.

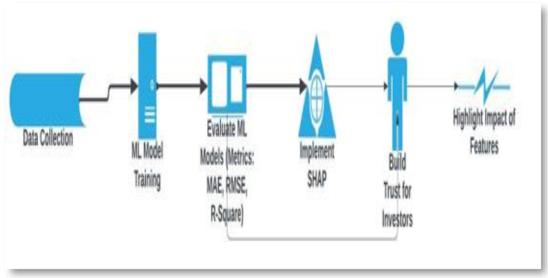

Figure 7:Explainable AI Workflow for Investor Trust through SHAP

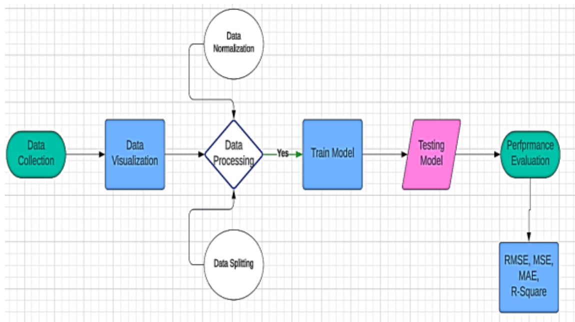

Figure 8:Methodology of Machine Learning Models

3.4 Tools for visualization

3.4.1 Python module:

They import the various libraries depending on the machine learning method used. The cryptocurrency forecast makes use of the Python modules which include among others; Pandas for data cleaning and feature creation, NumPy for computation of numbers and array manipulation, and Stats models for statistical purposes. Matplotlib and Seaborn are used for data visualization, and Scikit-learn has various armaments for model making and optimization. These libraries assist the investor to make informed decisions because he can build the forecast models based on Bitcoin prices and technicals. The Python 3.8.12 language was selected for the implementation and evaluation of these six selected machine learning methods for which four libraries were used including Keras 2.10.0 and Tensorflow 2.10.0 and MathPlot Lib and SKlearn 1.0.1.

3.5 Significance of Metrics

3.5.1 Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is one of the most popular measures for calculating the accuracy of predictive models more often regarding regression scenarios. It determines the mean value of the size of the error in the set of predictions considering whether the error is positive or negative. This metric gives one simple measure of the accuracy of the model with no ambiguity about what it measures. The formula for MAE is:

$$\text{MAE}=\frac{\sum_{i=1}^{n}|y_i-x_i|}{n} \qquad (11)$$

yi= prediction

xi =true value

n= total number of data points

MAE tends to have a lower value, meaning machine learning models have less error when made and recognized the efficiency in the model. This makes MAE an important measure for assessing the performance of models because it directly tells the user how close the forecasted value is to the real value [64]. MAE is used in assessing and comparing the predictive accuracy of a model where it is used to compare different predictive algorithms [65].

3.5.2 Root Mean Square Error (RMSE)

In regression analysis evaluation context, the usage of RMSE is common to determine the accuracy of established regression models. It calculates the difference between estimate value and actual values which gives information about the efficiency of a model to forecast the outcome variable from the input variable. RMSE is widely used because it gives a clear measure of the difference between the expected values and the observed value which would be quite useful in regression and forecasting application analysis [30][66][67].The formula for RMSE is:

$$\text{RMSE}=\sqrt{\frac{\sum_{i=1}^{N}(x_i-\hat{x}_i)^2}{N}} \qquad (12)$$

i = variable i

N = number of non-missing data point

Xi= field observation of times series

X̂i= estimated time series

RMSE is therefore one of the most important measures because it provides more information that includes the size of the errors. While Mean Absolute Error (MAE) norm does not differentiate the errors and treats all to be equal, RMSE has a more conclusive measure of impact wherever big error is squared which is very helpful in cases of larger variation of prediction values [68, 69]. With a smaller value of RMSE, higher level of accuracy in prediction is expected when compared with a large value of RMSE; which explains less accurate models [70].

3.5.3 R-Square (R^2)

R-Squared (R^2) then is literally a measure of the goodness of fit of your regression model or regression planes to the data. It measures the extent of explained variation of the dependent variable by the independent variables in the model so as to be an index of the fit of the model**Invalid source specified.Invalid source specified.**.

$$R^2 = 1 - \frac{RSS}{TSS} \qquad (13)$$

R^2= coefficient of determination

RSS= sum of square of residuals

TSS=total sum of squares

An R squared value of .70 – .90 is desirable for predictive models because it could mean that an overwhelming proportion of the data has been captured by the model.

4. RESULT AND DISCUSSION

4.1 Bitcoin dataset analysis discussion:

From the analysis of **Table 1** the predicting technique assigned to the Hybrid LSTM-RF model yields the lowest MAE indicates the small errors across the whole process, prediction are more likely to be accurate as compare to other models and the least value of RMSE refers that model is performing well as it reflects high accuracy, accompanied by the highest R-Squared values implies that the model fits the data well, that shows the Hybrid LSTM-RF model provides the best predicting outcome for Bitcoin. These results are supported through **Figure 15** which show that the model prediction is near the actual price set by the Hybrid LSTM-RF for BTC. Actual versus predicted values for the five models for the training dataset are presented on the following figures: **Figure 9,Figure 10, Figure 11, Figure 12**.

Table 1:Performance evaluation of BTC for proposed models

Datasets	ML Models	MAE	RMSE	R^2
Bitcoin	LSTM	0.0142	0.0210	0.9918
	SVM	0.0304	0.0360	0.9762
	RF	0.0134	0.0233	0.9900
	Hybrid LSTM-RF	**0.0108**	**0.0184**	**0.9938**

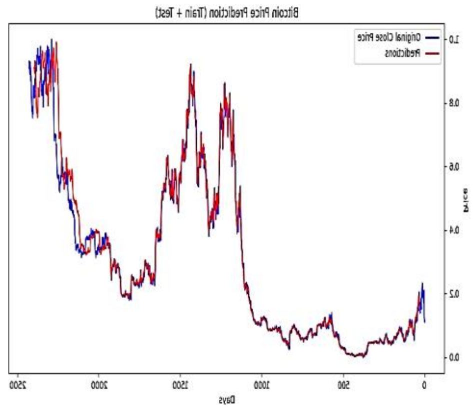

Figure 9:Illustration of Actual Vs Predicted BTC price by LSTM

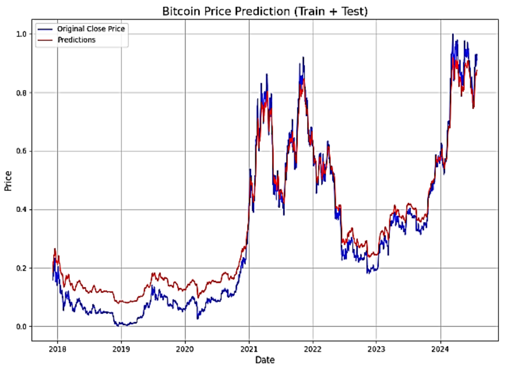

Figure 10:Illustration of Actual Vs Predicted BTC price by SVM

Figure 11:Illustration of Actual Vs Predicted BTC price by RF

Figure 12:Illustration of Actual Vs Predicted BTC price by Hybrid LSTM-RF

Figure 9 presents the relationship between the actual and predicted BTC prices using the LSTM model. The blue line represents the original closing price, while the red line represents the predicted closing price. The close alignment of the two lines indicates that the model performs well.

Figure 10 compares the actual and predicted BTC prices using the SVM model. The larger gap between the lines, as compared to other models, suggests that this model does not perform as effectively as the others.

Figure 11 depicts the relationship between the actual and predicted BTC prices using the Random Forest model. The differences between the lines are minimal, with slight deviations toward the end, indicating that this model ranks as the second-best in terms of prediction accuracy among all the proposed models

Figure 12 illustrates the comparison of actual and predicted BTC prices using the Hybrid LSTM-RF model. The lines are nearly identical throughout the graph, with only very small differences at certain points, indicating strong performance of the hybrid model.

Table 2:Forecasted Bitcoin Prices (Normalized and Original)

Date	Normalized Predicted Close	Predicted Close Original
2024-07-30	0.944064	69176.55
2024-07-31	0.666633	49798.90
2024-08-01	0.685668	51128.44
2024-08-02	0.592858	44645.96
2024-08-03	0.782385	57883.80
2024-08-04	0.668853	49953.96
2024-08-05	0.710485	52861.82
2024-08-06	0.702861	52329.31
2024-08-07	0.743125	55141.62
2024-08-08	0.718056	53390.63

In **Table 3** the forecasted Bitcoin closing prices for the next 10 days by using Hybrid LSTM-RF as this model shows the best result among other models. It includes both normalized predicted values and their corresponding original prices (in USD) after scaling back to the actual range. The original predicted values may deviate from real-world prices due to the exclusion of external factors that significantly influence cryptocurrency markets. As this study utilizes a limited set of variables, the model's predictions may not fully capture the complexity and dynamics of actual market conditions. The focus on a limited variable set was intentional to evaluate the predictive capabilities of the chosen models under constrained conditions. This approach provides a foundation for future research to build upon by incorporating additional features.

4.2 Ethereum dataset analysis discussion

As shown in **Table 3**, the hybrid LSTM-RF model achieved the best performance on the Ethereum dataset, exhibiting the lowest MAE and RMSE values, as well as the highest R-squared value, indicating an excellent fit of the data to the model. The Random Forest model also performed well, showing metric values close to those of the hybrid LSTM-RF model, making it the second-best model. In contrast, the SVM model did not perform well, even after applying Min-Max normalization. The comparison of actual and predicted values for the training dataset across the five models is illustrated in **Figures 13, 14, 15, and 16.**

Table 3:Performance evaluation of ETH for proposed models:

Datasets	ML Models	MAE	RMSE	R^2
Ethereum	LSTM	0.0116	0.0168	0.9871
	SVM	0.0420	0.0468	0.9089
	RF	0.0056	0.0086	0.9969
	Hybrid LSTM-RF	**0.0055**	**0.0080**	**0.9974**

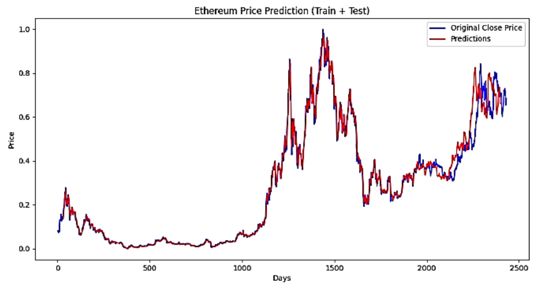

Figure 13:Illustration of Actual Vs Predicted ETH price by LSTM

Figure 14:Illustration of Actual Vs Predicted ETH price by SVM

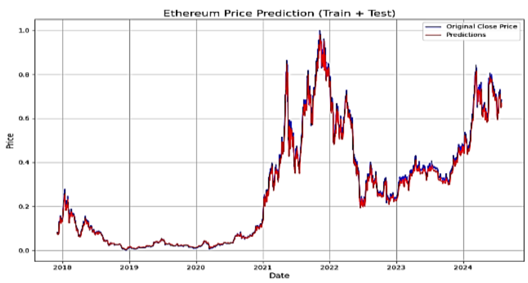

Figure 15:Illustration of Actual Vs Predicted ETH price by Hybrid LSTM-RF

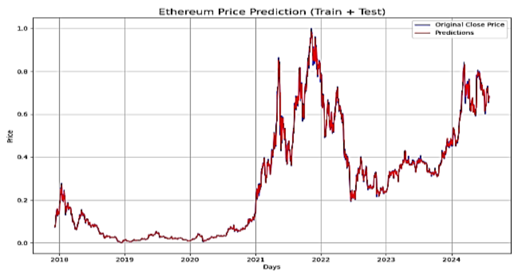

Figure 16:Illustration of Actual Vs Predicted ETH price by RF

Figure 13 illustrates the relationship between the actual and predicted Ethereum (ETH) prices using the LSTM model. The blue line represents the original closing price, while the red line represents the predicted closing price. The close proximity of the two lines indicates that the model performs well, though not as effectively as the Random Forest and Hybrid LSTM-RF models, as demonstrated in Figures 15 and 16. There are slight deviations between the two lines at certain points, reflecting minor discrepancies in the predictions.

Figure 14 displays the comparison between the actual and predicted closing prices of Ethereum using the SVM model. This means that there are a larger gap between actual and predicted values in this models as represented by the significant gap between the red and blue lines in the graph. This implies that the current SVM model is not as effective in the current classification as expected on the given dataset.

The closing prices of Ethereum predicted by using the Random Forest (RF) model are presented in **figure 15** with the plot of actual closing prices for the same period. The RF model presents the second best fit among the proposed models and the gap between the blue (actual) and the red (predicted) lines is almost impractical making us conclude that the models' predicted values have a strong resemblance with actual values.

To see how well the chosen Hybrid LSTM-RF model predicts the closing price of Ethereum, **Figure 16** shows the comparison of the closing prices predicted by the model and actual closing prices of Ethereum. As seen from the above graph the proposed models give the least errors compared to other models with a closer fit between the actual and the predicted values.

Table 4:Forecasted Ethereum Prices (Normalized and Original)

Date	Normalized Predicted Close	Predicted Close Original
2024-07-30	0.944064	69176.55
2024-07-31	0.666633	49798.90
2024-08-01	0.685668	51128.44
2024-08-02	0.592858	44645.96
2024-08-03	0.782385	57883.80

2024-08-04	0.668853	49953.96
2024-08-05	0.710485	52861.82
2024-08-06	0.702861	52329.31
2024-08-07	0.743125	55141.62
2024-08-08	0.718056	53390.63

Table 4 displays the normalized predicted prices of Etherium and their corresponding values in the original scale for the next 10 days. The predicted values might differ from real-world prices because external factors that greatly impact cryptocurrency markets were not included in this study. By focusing on a limited set of variables, the research aimed to assess the performance of the selected models within a constrained framework. While this approach highlights the models' capabilities, it also opens the door for future studies to enhance predictions by incorporating a broader range of features.

4.3 Litecoin (LTC) dataset analysis discussion:

The **table 5** showcases the performance of various machine learning models in predicting Litecoin prices, evaluated using three key metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 (coefficient of determination). The Hybrid LSTM-RF model stands out as the top performer, achieving the lowest MAE (0.0023) and RMSE (0.0036), along with the highest R^2 score (0.9868), demonstrating its superior accuracy and ability to explain the data's variability. Among the individual models, LSTM also performs well, with an MAE of 0.0025, RMSE of 0.0036, and an R^2 of 0.9867, closely matching the hybrid model's performance. The Random Forest (RF) model delivers slightly higher error rates, with an MAE of 0.0025, RMSE of 0.0040, and R^2 of 0.9838, showing solid but slightly less accurate predictions. On the other hand, the SVM model demonstrates significantly higher errors (MAE = 0.0593, RMSE = 0.0601) and a noticeably lower R^2 of 0.9162, indicating it is less effective for this task. **Overall, the Hybrid LSTM-RF model emerges as the most accurate and reliable approach for predicting Litecoin prices in this analysis.**

Table 5:Performance evaluation of LTC for proposed models

Datasets	ML Models	MAE	RMSE	R^2
Litecoin	LSTM	0.0025	**0.0036**	0.9867
	SVM	0.0593	0.0601	0.9162
	RF	0.0025	0.0040	0.9838
	Hybrid LSTM-RF	**0.0023**	**0.0036**	**0.9868**

Figure 17:IIustration of Actual Vs Predicted LTC price by LSTM

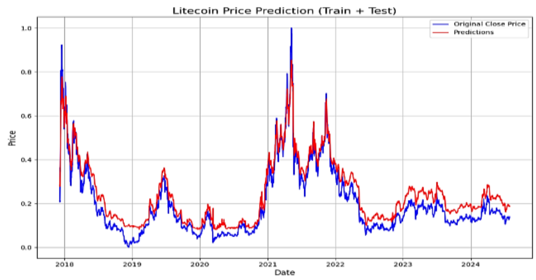

Figure 18:Illustration of Actual Vs Predicted LTC price by SVM

Figure 19:Illustration of Actual Vs Predicted LTC price by Random Forest

Figure 20:Illustration of Actual Vs Predicted LTC price by Hybrid LSTM-RF

Figure 17 shows the Litecoin Price Prediction based on the LSTM (Long Short-Term Memory) model, to let understand the contrast between actual and predicted closing rates in the training set and the overall set. Nonetheless, the model is capable of reproducing the price dynamics and its fluctuations, which are deviations from the idealized curves, are marked by the highest peaks and troughs observed in the real data. This alignment shows that LSTM is able to grasp the time-series data as the networks were trained to remember lengthy dependencies.

In figure 18 the graph of Actual vs Predicted LiteCoin (LTC) Price based on SVM is shown for Training and Testing data where line is plotted to show the difference between real world quantity values and the values of the same quantity quantities generated by the model. Both the red (predict) and blue (real) curves of the graph repeat the general plan of fluctuations in the historical data set, which is quite suitable. The SVM model captures the general trends up and downs such as high volatility in 2018 and 2021 as shown below. However, there are larger variations in the actual and forecasted prices, especially where there is a steep increase or decrease in prices. This implies that even though the basic trends were predicted well, there were problems in capturing extreme values, which are characteristic of trading cryptocurrencies.

Figure 19 The predicted and actual LiteCoin (LTC) prices Random Forest model cumulative for training and testing set Predicted and Actual LiteCoin (LTC) prices . By comparing the red (predicted) and blue (actual) lines, the proposed model well explains the overall behavior of the total number of daily codes changes and preserves the primary features, such as peaks and troughs in 2018 and 2021. The model has especially instilled great optimism over the ability to simulate patterns of past price fluctuations, and it offers minimal oscillations mostly at a time when it is volatile. From this, one can infer that Random Forest model is perfectly suited for the complexity of such data and provides sensible predictions when it comes to cryptocurrency price prediction.

The actual versus the predicted Lite Coin (LTC) because of the Hybrid LSTM-RF model is presented in **figure 20** where training and testing datasets are used. Whereas the blue line refers to the actual closing prices of Litecoin and the red line referring to the predicted values defined by the model. The result shows a high correlation of two lines which indicate that, Hybrid LSTM-RF model efficiently consider the variation of prices over time. In cases of high fluctuation, as seen in 2021, they are convincing, as well as in the later years, following the model's steps.

This alignment demonstrates the model's ability to adapt to the dynamic nature of cryptocurrency markets, providing accurate predictions that can be used for informed decision-making. The visualization reinforces the robustness and precision of the Hybrid LSTM-RF approach in forecasting cryptocurrency prices.

These results suggest that the combination of models yields better results particularly when one combines a conventional machine learning model with a deep learning model yields the best results upon normalization of Bitcoin, Ethereum, Litecoin datasets

4.4 Implementation of Human-centered SHAP

Therefore, from the above results, it is concluded that Hybrid LSTM-RF is performing best among others. So, now implementing **Human-Centered SHAP XAI** to that model to make model and prediction more understandable.

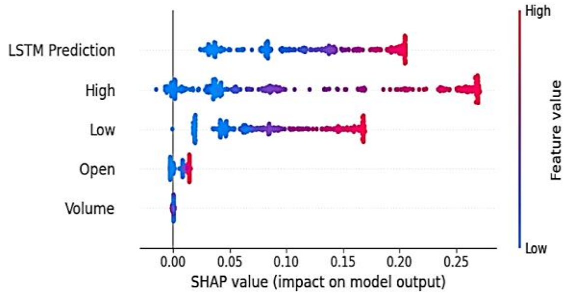

Figure 21:SHAP Analysis of Feature Impact on Hybrid LSTM-RF Predictions for Bitcoin Price

The **figure 21, SHAP summary plot** visually represents how different features (such as High, LSTM Prediction, Low, Open, and Volume) of Bitcoin dataset contribute to the model's predictions (**Hybrid LSTM-RF)**. Each dot on the graph corresponds to a single instance (observation), and the x-axis represents the **SHAP value**, which indicates how much a specific feature contributes to either increasing or decreasing the model's output (the target variable). The **figure 21** i.e Traditional XAI show The SHAP summary plot illustrates the impact of various features on the predictions of the hybrid LSTM-Random Forest model. Notably, "LSTM Prediction" emerges as the most influential feature, with SHAP values ranging from 0 to 0.25, indicating that the predictions generated by the LSTM model substantially contribute to the final output of the Random Forest. Higher LSTM predictions (denoted by red points) positively influence the output, while lower values (blue points) have a more neutral effect.

The "High" and "Low" price features also play significant roles, albeit with a slightly smaller impact compared to "LSTM Prediction." Higher values for these features (in red) correlate with positive SHAP values, which push the model's prediction upward, while lower values (in blue) have a downward or neutral influence. "Open" and "Volume" have comparatively smaller SHAP value ranges, indicating a lesser influence on the final model prediction. However, high "Open" and "Volume" values still exhibit a positive impact, albeit to a smaller degree.

Overall, the plot shows that while traditional price features ("High," "Low," "Open," "Volume") contribute to the model's performance, the "LSTM Prediction" feature dominates in shaping the outcome, showcasing the strength of the hybrid approach where LSTM-generated predictions are fed into the Random Forest model to improve forecasting accuracy.

Figure 21 provides a general explanation of how each feature influences the model's output, showing feature importance in a broad sense. It is effective for machine learning practitioners and data scientists but may be harder for non-expert users to interpret without guidance. To make it more user friendly we proposed human centered XAI graph in **figure 22.**

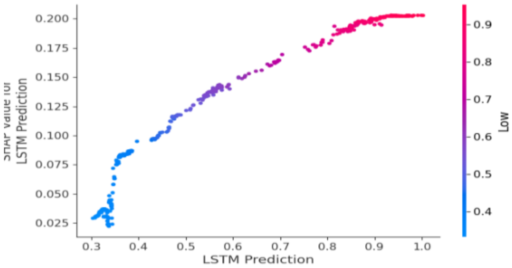

Figure 22:"Interpreting LSTM Predictions: SHAP Value Gradient Reveals Feature Influence"

Figure 22 shows the SHAP **dependence plot**, which visualizes the relationship between the "LSTM Prediction" feature and its SHAP values. In this graph, both the X-axis and Y-axis represent the LSTM predictions, while the color gradient indicates the values of another feature, Low. There is a clear positive linear relationship between the LSTM predictions (X-axis) and their corresponding SHAP values (Y-axis). As the LSTM prediction increases, the SHAP value also increases, indicating that higher LSTM predictions have a greater positive influence on the model's final output. The gradation from blue to red also means that as the "Low" value of the price rises, so does the LSTM forecast. This implies that the "Low" price has a positive contribution on the forecast of the LSTM model; hence, high "Low" values increase the predicted prices. The dependence plot of the Random Forest model helps to visualize what feature, LSTM Prediction', is influencing the model as follows; The dependence plot correlates the higher LSTM Prediction' with higher 'SHAP' hence, higher predictions. Further, the characteristics of low price and LSTM prediction: the larger the Low' number is, the larger the predictive value resulting in a direct correlation between the two measurements. This revelation further supports the significance of the 'Low' price that directs the behavior of the LSTM model impacting the general hybrid model.

If investors are aware that high LSTM predictions accompanied by high Low prices result in higher final price estimations, they therefore can determine when they want to invest or disinvest.

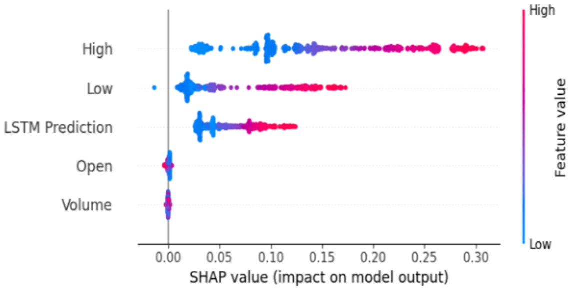

Figure 23:SHAP Analysis of Feature Impact on Hybrid LSTM-RF Predictions for Ethereum Price

Figure 23 represents the following SHAP summary plot depicted below have analyzed and presented to show the prominence and contribution of various features of the hybrid LSTM-Random Forest model on Ethereum dataset. On the Y-axis, features like "High," "Low," "LSTM Prediction," "Open," and "Volume" are listed, while the X-axis shows their corresponding SHAP values, indicating the impact each feature has on the model's output.

The "High" feature has the largest spread of SHAP values, ranging up to 0.30, indicating that it significantly influences the model's predictions. In **Figure 23** the "Low" feature also has a considerable impact on the model's predictions, with SHAP values extending up to 0.20. Similar to "High," lower values (blue) generally decrease the model's predictions, while higher values (red) increase them. However, the SHAP value range for "Low" is slightly smaller than for "High."

Figure 23 SHAP summary plot reveals that the hybrid LSTM-Random Forest model relies heavily on the "High" and "Low" price features, which have the most substantial impact on the final predictions. The "LSTM Prediction" also contributes significantly, but to a slightly lesser extent than "High" and "Low."

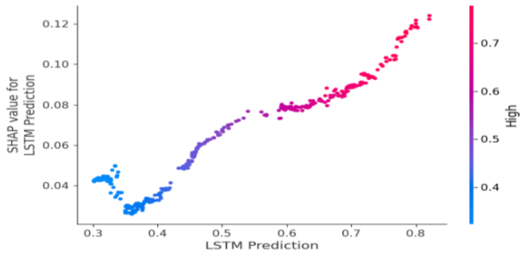

Figure 24:Interpreting LSTM Predictions: SHAP Value Gradient Reveals Feature Influence"

Figure 24, This SHAP dependence plot shows the relationship between the "LSTM Prediction" (on both the X-axis and Y-axis) and its SHAP value, illustrating the feature's contribution to the model's output. The color gradient represents the "High" price feature, where blue corresponds to lower values and red to higher values. As "LSTM Prediction" increases, its SHAP value also increases, indicating a positive correlation between the LSTM predictions and the final model output. Higher "High" prices (red) further amplify this impact, reinforcing the predictive strength of both features in the model.

If investors know that high **LSTM predictions** combined with high **High prices** lead to higher final price predictions, they can strategically decide when to enter or exit the market. For example, they might wish to get a buying signal when the LSTM Prediction and the High price are both rising because they both show a higher rate of increase.

The proposed LSTM-Random Forest model has brought a new and efficient paradigm to predict cryptocurrency prices that are highly accurate than other models.

5. CONCLUSION AND FUTURE DIRECTION

The machine learning models in this study were LSTM, RF, SVM, and an LSTM and RF combined model. The assessment was mainly aimed at the identification of the potential of each of the models in the analysis of the not very stable cryptocurrency market and their ability to outline critical patterns as well as generate relatively accurate forecasts. These results implied that the proposed hybrid LSTM-RF model had a higher accuracy and reliability than the single models. The LSTM part successfully modelled temporal features within the data, with the RF part providing additional stability, by handling non-linearities and preventing over-training. This ability enabled the hybrid model to give more accurate and reliable predictions than the single models were able to facilitate. Although both LSTM and RF models provided well for the task, they did not benefit from the synergies of the hybrid model as was with the case to the models but otherwise, SVM seemed to have a hard time handling the volume and complexity of the data. These results point to improvement from the use of such a hybrid model such as LSTM-RF, which is informative in dynamic financial forecasts, which facilitates improved predictive abilities and tools for investors and traders.

To enhance the possibility of using the model, the Explainable AI (XAI) procedures were used. These methods offered explanation and visualizations that made automobile comprehensible to the human intellect, rather than the artificial one. The implementation of XAI served two primary purposes: Interpretability: Through the feature importance graphs, the

SHAP (SHapley Additive exPlanations) plots, and decision tree maps that were made available users were able to see what factors such as previous prices, volumes, or trends played a huge role in the predictions. This makes the people have a trust in the model so that they can produce the required output. Accessibility: With these specifications, and the general approach outlined in the previous section, the work helped individuals without extensive background in machine learning to understand what the model was doing and why. This accessibility makes it easy to derive decision-grounded conclusions from the model to assist the user make decisions regarding their investments or trading plans.

Thus, the integration of XAI serves as a useful improvement to the LSTM-RF model since it tries to close the gap between comprehending how the machine learning algorithm optimal the given equations and making it understandable for users in the real world.

The promising results of the hybrid LSTM-RF model open several avenues for future exploration:

1. Application to Diverse Cryptocurrency Datasets: Future works can use the LSTM-RF model to cover more cryptocurrencies including alt-coinage cryptocurrencies. This would help to evaluate the stability and adaptability of the model to changes in market conditions and taking into account different types of assets.

2. Comprehensive Model Comparisons: Exploring a complex of other advanced machine learning algorithms (for example, XGBoost, CatBoost, Transformer based) would give a better setting of LSTM-RF model comparing to the others. They may also contribute to the refinement of implementing hybrid techniques such as the one under study.

6. REFERENCES

[1] H. P. a. Ana Todorovska, "Using ML and Explainable AI to understand the interdependency networks between classical economic indicators and crypto-markets," Elsevier, 2023.

[2] A. Stencel, "What is a meme coin? Dogecoin to the moon!," 2023.

[3] W. C. C. V. M. B. L. K. L. L. a. F. W. Fan Fang, "Ascertaining price formation in cryptocurrency markets with machine learning," THE EUROPEAN JOURNAL OF FINANCE, 2024.

[4] Y. N. K. D. Mohammad Hashemi Joo1, "Cryptocurrency, a successful application of blockchain technology," 2019.

[5] Q. A. N. L. Sudaryono Sudaryono1, "Application of Blockchain Technology for iLearning Student Assessment," 2020.

[6] R. Li1, "Research on the Investment Value of Cryptocurrency and Blockchain Technology," 2023.

[7] S. O. Olabanji, "Technological Tools in Facilitating Cryptocurrency Tax Compliance: An Exploration of Software and Platforms Supporting Individual and Business Adherence to Tax Norms," CJAST, 2023.

[8] L. Badea and M. Mungiu-Pupăzan, "The economic and environmental impact of bitcoin," IEEE Access, 2021.

[9] C. R. B. M. a. E. P. Phumudzo Lloyd Seabe, "Forecasting Cryptocurrency Prices Using LSTM, GRU, and Bi-Directional LSTM: A Deep Learning Approach," MDPI, 2024.

[10] P. F. K. W. O. Grzegorz Dudek, "Forecasting cryptocurrencies volatility using statistical and machine learning methods: A comparative study," Elsevier, 2024.

[11] B. Mahesh, "Machine Learning Algorithms - A Review," 2018.

[12] M. R. M. R. &. M. M. R. Chowdhury, "An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning",," 2020.

[13] A. M. &. V. N. S. Vasily Derbentsev, "Forecasting of Cryptocurrency Prices Using Machine Learning," springer link, 2020.

[14] S. U. J. R. D. Davinder Kaur, "Trustworthy Artificial Intelligence: A Review," ACM, 2022.

[15] T. Zhan, "TRUSTWORTHY ARTIFICIAL INTELLIGENCE FOR BLOCKCHAIN-BASED CRYPTOCURRENCY," International Journal of Computer Science and Information Technology (IJCSIT).

[16] K. S. M. H. u. R. S. Mohamed Nassar, "Blockchain for explainable and trustworthy artificial intelligence," WIREs, 2019.

[17] J. Chen, "Analysis of Bitcoin Price Prediction Using Machine Learning," MDPI, 2023.

[18] T. K. S. T. X. V. M. V. Tran, "Exploring Hybrid Models For Short-Term Local Weather Forecasting in IoT Environment," MENDEL soft computing , vol. 29, 2023.

[19] A. Gupta, D. Vyas, P. Nale, H. Jain, S. Mishra and R. V. Bidwe, "Cryptocurrency Prediction and Analysis between Supervised and Unsupervised Learning with XAI," IEEE, 2023.

[20] S. B. J. F. S. John W. Goodell, "Explainable artificial intelligence modeling to forecast bitcoin prices," Elsevier, 2023.