Fusion in Cryptocurrency Price Prediction: A Decade Survey on Recent Advancements, Architecture, and Potential Future Directions

NISARG P. PATEL1, RAJ PAREKH1, NIHAR THAKKAR1, RAJESH GUPTA1, Student Member, IEEE, SUDEEP TANWAR1, Senior Member, IEEE, GULSHAN SHARMA2, INNOCENT E. DAVIDSON3, Senior Member, IEEE, RAVI SHARMA4

1Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India (e-mails: 18bce136@nirmauni.ac.in, 18bce143@nirmauni.ac.in, 18bce133@nirmauni.ac.in, 18ftvphde31@nirmauni.ac.in, sudeep.tanwar@nirmauni.ac.in)
2Electrical Engineering Technology, University of Johannesburg, Johannesburg, South Africa (email: gulshans@uj.ac.za)
3Department of Electrical Power Engineering, Durban University of Technology, Steve Biko Campus, Durban 4001, South Africa. (email: innocentD@dut.ac.za)
4Centre for Inter-Disciplinary Research and Innovation, University of Petroleum and Energy Studies, Dehradun, 248001, India. (email: ravisharmacidri@gmail.com.)

Corresponding author: Sudeep Tanwar (e-mail: sudeep.tanwar@nirmauni.ac.in) and Innocent E. Davidson (e-mail: innocentD@dut.ac.za).

ABSTRACT Cryptographic forms of money are distributed peer-to-peer (P2P) computerized exchange mediums, where the exchanges or records are secured through a protected hash set of secure hash algorithm-256 (SHA-256) and message digest 5 (MD5) calculations. Since their initiation, the prices seem highly volatile and came to their amazing cutoff points during the COVID-19 pandemic. This factor makes them a popular choice for investors with an aim to get higher returns over a short span of time. The colossal high points and low points in digital forms of money costs have drawn in analysts from the scholarly community as well as ventures to foresee their costs. A few machines and deep learning algorithms like gated recurrent unit (GRU), long short-term memory (LSTM), autoregressive integrated moving average with explanatory variable (ARIMAX), and a lot more have been utilized to exactly predict and investigate the elements influencing cryptocurrency prices. The current literature is totally centered around the forecast of digital money costs disregarding its reliance on other cryptographic forms of money. However, Dash coin is an individual cryptocurrency, but it is derived from Bitcoin and Litecoin. The change in Bitcoin and Litecoin prices affects the Dash coin price. Motivated from these, we present a cryptocurrency price prediction framework in this paper. It acknowledges different cryptographic forms of money (which are subject to one another) as information and yields higher accuracy. To illustrate this concept, we have considered a price prediction of Dash coin through the past days’ prices of Dash, Litecoin, and Bitcoin as they have hierarchical dependency among them at the protocol level. We can portray the outcomes that the proposed scheme predicts the prices with low misfortune and high precision. The model can be applied to different digital money cost expectations.

INDEX TERMS Cryptocurrency, Price Analysis, Volume Analysis, Deep Learning, Machine Learning, Survey, Ensemble Model, Fusion in metrics

I. INTRODUCTION

Cryptocurrencies have been introduced to the market in 2008 when Satoshi Nakamoto rendered the first cryptocurrency named Bitcoin, through his white paper Bitcoin: peer-to-peer Electronic Cash System [1] intending to replace the old and widely used economic system. The traditional cash exchange system is entirely dependent on the centralized third-party financial institutions such as banks in order to validate and execute the monetary transactions. It is restricted to perform only a finite amount of money and possesses various issues such as trust, security, transparency, and flexibility. To overcome the aforementioned issues, a peer-to-peer (P2P)
network-based currency namely cryptocurrency is a perfect solution [2]. It eliminates the involvement of centralized third-party banking systems. In this, each transaction has to be acknowledged (validated) by each peer using any suitable consensus algorithm such as proof of work (PoW), proof of stack (PoS), and many more. These algorithms internally uses the transaction hash and timestamps for transaction confirmation. It requires peers in the network to solve a complex mathematical problem to add a new block in the chain. Mining is the process of solving the problem, and miners get rewarded for their work through cryptocurrency. A specific block is of no use until it is not added to a chain. Here, the question arises whether the system of cryptocurrency can be carried out in a practical scenario?

The practicality of Bitcoin can be seen through its exponential growth in terms of transaction and volume. The first transaction was done in 2010, where 10000 Bitcoins were used to buy couple of pizzas [3] and as of April 2021, nearly 2 million transactions were executed. As the number of the transaction is increasing exponentially, we can say that the cryptocurrency is accepted by the public. Specialized hardware is required to make the mining process efficient. To carry out the mining, GPU plays an important role, which utilizes its CUDA cores for quick calculation of PoW problems by recalculating the valid hash using algorithms like SHA-256 and X11 as well mentioned by the authors in [4]. The GPUs utilized in this process are expensive and they require high power supply. Then, the organization started to use a system called pool, where many users come together to mine a coin by helping each other and get proportional reward based on their contributions [5] [6]. So, is there any way to generate profit from it?

The mining activities described above are profitable, which leads to the emergence of startups and draws the attention of big companies for mining cryptocurrency. However, the complexity of making a chain of hashes increasing day-by-day due to the limited number of coins available on the web. Apart from these, other factors like the market sentiments of also affect the coin prices. Such additional factors make the cryptocurrency prices highly volatile, and the mammoth institutes are getting high returns [7] [8]. The same observation can be sum up from the price timeline, i.e., the Bitcoin’s price jumped from $0.08 in 2010 to above $64000 as of April 2021 [9], and the same trend is followed by Ethereum, Ripple, Litecoin and others as well. So, what will be the growth of the cryptocurrency market cap, and at present how many cryptocurrencies exist in the market?

Since the inception of Bitcoin, several cryptocurrencies have been introduced to the market having their advantages and disadvantages. As of April 2021, according to [9] nearly 4200 crypto coins are circulating in the market with a net worth of over $2.23 Trillion, in which Bitcoin has a share of 78% followed by Ethereum with a total market share of 12%. High market esteem brings the consideration of organizations and people to put resources into them. The cryptocurrency market will boom with a minimum expected CAGR value of 11.9% that had been derived by [10] with a detailed analysis based on different market sizes, peoples sentiments, and many other factors as well. The given growth rate is functional till 2025. Considering that the market follows the same trends then the given growth rate can be effective for cryptocurrencies. Through FIGURE 1 we can see the expected future price changes of the crypto market till 2050.

High volatility in cryptocurrency prices invokes the interest of researchers across the academia as well as industries to predict their prices using the power of machine learning and deep learning algorithms [11]. The volatility of cryptocurrencies depends on factors like volume, mining difficulty, popularity, price of alternate coins, transaction cost, sentiment on the social platform, and the laws of a specific country as well stated in [12] [13]. Due to the involvement of numerous factors, price projection becomes a challenging task for the entire research community. To solve this challenging task, we put forward a fusion of cryptocurrencies, a deep learning-based scheme using a time-series model. The function of this model is to predict the price of a specific coin by utilizing the past prices of other cryptocurrencies. For example, the price forecasting of Ethereum-Classic can be done through prices of Ethereum and Bitcoin.

A. NECESSITY OF FUSION IN CRYPTOCURRENCY PRICE PREDICTION

The sudden surge in the market value of cryptocurrencies brings huge attention to the investors. The main activities involved in cryptocurrency market are selling, buying, and submitting PoW, PoS through mining. For analyzing its efficiency in the market, the hypothesis theorems like EMH and AMH were introduced. Now, we will get insights to this theorem in context of cryptocurrency market [14] [15]. To understand the EMH, consider a situation where prices of
Bitcoin are reflecting by sentiment on the social platforms. However, the approval evidence of EMH has been arguable over a while. Oppositely, there is no approval evidence for EMH, the author of [16] has studied the cryptocurrency market through EMH and found good results. EMH being arguable for its utilization in the real scenario for determining dependency among factors. To overcome it, AMH comes with the inclusion of Behavioral finance. So now, what are AMH and behavioral finance hypotheses, and how it is important for our research?

Behavioral finance is the study of influence of psychology on the behavior of investors or financial analysts. The information by them creates a situation called herding, where the people’s actions depend on market agents [17] [7]. The bias discussed above had an important role in determining cryptocurrency prices. During this kind of scenario, EMH does not work, because it assumes the market to be frictionless without any uncertainty and biases. Therefore, an alternative theory named AMH was introduced. The study in [18] checks the proficiency in the development of Bitcoin prices, and concluded that the proof of dynamic effectiveness sticks to the recommendation of the AMH. The authors of [18] concluded that the different factors mentioned in [12] played an important role in determining the cryptocurrencies prices, so there must be a careful selection of attributes to be considered together for improving the ability of the forecasting model. Integrating these features helps to build a robust model, where the shortcomings of a particular factor can be covered by other factors. The process of integrating and utilizing these features is term fusion. So now the question arises, how can the fusion algorithm be incorporated to build a prediction model?

The fusion can be carried out in two ways. One way of fusion is to integrate several forecasting models to make a hybrid or ensemble model [19]. Through this, the results of all models can be generalized to overcome the weaknesses of a single model. Another way of fusion is our novel idea, where we consider the inter-dependency among cryptocurrencies. Here, the interdependency may be technical or non-technical. Considering these ways significantly improves the performance of the forecasting model.

B. SURVEY METHODOLOGY
This section describes the methodology considered to carry out the survey. For this, we have followed the guidelines given by Kitchenham and Charters [20]–[22].

1) Review Plan
The exhaustive study needs some survey planning for this paper, a comprehensive study has been performed to provide insights into the concrete research done in the field of cryptocurrency price prediction by utilizing it’s different depending factors. We begin with the review planning, identification of numerous research questions helping to narrow down our search criteria, reading of sorted articles for inclusion–exclusion criteria, and quality evaluation. We identified various relevant studies, articles, and publications to carry out this systematic survey. The identified material is first referred for quality before utilizing its data in the proposed survey.

2) Research Questions
The aim of this article is to systematically and comprehensively survey cryptocurrency price prediction using machine learning and deep learning algorithms and then proposed a novel method to increase the performance of the prediction model. For impelling it, various research questions were framed for this comprehensive survey as written down below:

- What are possible factors that are yet to be explored for cryptocurrency price prediction?
- What is the importance of explored factors for cryptocurrency price prediction?
- How to improvise new findings with existing work to make a powerful forecasting model?
- What are other possible fusion-based algorithms to make a robust forecasting model?
- What are the research challenges of the identified solution?

3) Data Sources
Our study had been carried out through high-quality peer-reviewed research works from the reputed online databases such as Springer, Elsevier (Science Direct), ACM digital library, Wiley, IEEE Xplore, Google Scholar, and the company’s official technical blogs.

4) Search Keywords
The vital keywords used for searching the related contents were: "cryptocurrency", "cryptocurrency price prediction", "price prediction using machine learning", "factors affecting prices of cryptocurrency", "blockchain factors affecting cryptocurrency prices", "time series based machine learning models", "regression based price prediction", "Bitcoin and Ethereum volume analysis", "Bitcoin and Ethereum price analysis", "Impact of price change on derived cryptocurrency", "Stakeholders in cryptocurrency market", "Companies and research organization involved in blockchain development", "market and sentiment trends on cryptocurrency market", "Correlation values between derived factors", "types of cryptocurrencies", "Details of Zcash, Dash, Litecoin, Ripple, Monero, Bitcoin Cash Bitcoin SV, Ravencoin, Ethereum Classic, Qum, Wrapped Bitcoin, Ripple and Binance Coin", "Normalization techniques for regression", "Working of RNN, GRU and LSTM", "Ensemble Models", "Performance evaluation metrics for price prediction models", "Transformers, Federated Learning and Reinforcement learning in field of price prediction", "Legal aspects and government policies on cryptocurrency", and "challenges in the field of cryptocurrency price prediction".
5) Criteria Of Inclusion And Exclusion
For making this survey impressive and effective, the recent and relevant papers of 2020 are included along with the early access articles. The other survey articles, tutorial papers, books, technical reports, blogs, and other resources are also included for wider coverage. The filtration process is divided into multiple stages based on the title, abstract, full paper, and investigations. Finally, we identified some relevant papers and focused on those having good citations. We have analyzed the works of curated papers and highlight the significance of papers in this field of research, incorporating various noteworthy studies in our article with our new fusion method for the prediction model.

C. SURVEY MOTIVATION
The properties like immutability, transparency, security, decentralization, and exponential growth make the cryptocurrency market an important part of the investment for investors and individuals. Because of its volatile nature with high returns, people started trading, which works similarly to the stock market. The study in [23] briefs about the crypto trading. FIGURE 2 concluded the same which reveals the exponential growth in the number of crypto wallets since 2011. However, the price is much more volatile compare to other investment areas because of its dependency on a vast number of parameters as discussed above. The above reason motivates us to work in the research field of cryptocurrency price prediction.

The state-of-the-art forecasting models uses machine learning and deep learning algorithms. But most of them focus on the limited and famous cryptocurrencies such as Bitcoin, Ethereum, Ripple, and Litecoin [19]. The factors they have considered are also limited and direct factors like opening and closing price, the complexity of mining problem, volume, and flow of the coins. Others have considered indirect factors like sentiments of cryptocurrencies on social platforms [24], socio-economic factors [25], epidemic modeling [26] and online social networks (OSN) models for detecting price bubbles. Apart from this, there are any significant factors that can be considered for designing a precise and robust model?

There exist several other properties that plays an important role while forging the forecasting model. One of the properties is the dependency of one coin on others. Although it is an indirect factor as most of them are derived from Bitcoin and Ethereum with additional features, so the price tendency of the derived coins is dependent indirectly on them. The inverse can be justified as the price changes of the derived coins can reflect their impact indirectly on the parent one. This factor is vital, which further motivates us to build a forecasting model utilizing the described feature. Our paper aims to incorporate this factor through fusion in state-of-the-art models to improve the performance of the price prediction algorithm.

FIGURE 2: Growth in number of cryptocurrency wallet users till 2021 [9].

D. RESEARCH CONTRIBUTIONS
Price volatility and the dynamic nature of cryptocurrency make the task of price forecasting hard and tedious. Several state-of-the-art models have been introduced by researchers across the globe that analyzed the dependency of several direct and indirect factors for the price prediction. Still, many dependencies are yet to be explored. The existing studies have explored such attribute and following are the key objectives of the paper.

- We present a comprehensive study on various existing cryptocurrency price prediction models that uses both machine and deep learning approaches.
- We propose a novel deep learning-based fusion model to predict the cryptocurrency prices by considering their interdependencies among other cryptocurrencies.
- To evaluate the performance of the proposed fusion-based model, we consider different window sizes, i.e., 1, 3, 7, and 30 for previous cryptocurrency prices.

E. ORGANIZATION
FIGURE 3 represents the paper organization, where of this paper. The rest of the paper is organized as follows. Section II discusses the existing research work in the field of cryptocurrency price predictions. Section III gives the detailed analysis of major factors mentioned in [12], affecting the price of the cryptocurrency. Section IV describes the importance of stakeholders associated with cryptocurrency price prediction. Section V analyzes the important metrics that affect the prices of cryptocurrency. Section VI consists study of dependency among the cryptocurrency based on technical and non-technical factors. Section VII discusses problem formulation and making a system model for the proposed architecture. Section VIII describes the architecture of proposed fusion-based model. Section IX contains the
The introduction section explains the working of cryptocurrency and factors affecting its price. It also focuses on the estimated growth of the cryptocurrency market between 2020-2050 with a CAGR value of 11.9% and its timeline in FIGURE 1. Following, a short explanation of the proposed forecasting model. After that, there is a brief description of factors influencing cryptocurrency price based on the EMH, AMH, and how a robust model can be built through a fusion by overcoming the demerits of one feature through other features. Consequently, we discussed a brief survey of a few research works and our motivation behind this model. After that, we described the role of each section in the paper.

II. INSIGHTS TO THE EXISTING WORKS IN CRYPTOCURRENCY PRICE PREDICTION

Due to the extreme fluctuation in cryptocurrency prices, their price prediction has become an attractive topic for the researcher since the last decade. High volatility in the prices of the cryptocurrency was noticed for Bitcoin. Researchers are successfully able to predict the price for the stock market but not able to predict the price of cryptocurrency with good accuracy compared to the stock market because of the large number of depending factors affecting the price of cryptocurrency significantly.

Average price, opening price, closing price, highest price, lowest price, and the volume traded of cryptocurrency for the past days are some of the main factors to forecast the next day’s price significantly. The factors affecting the price of cryptocurrency also include social media sentiments-based comments/posts. Some researchers have tried to implement the relation of the cryptocurrency with gold and then try to predict the cryptocurrency prices. The change in the price of Bitcoin for the near future can be explained by the investor’s sentiment significantly. This study was carried out by R. C. Philip in [47] that justify the behavioral finance property, which was also discussed in Section I. In [36], researchers have also concluded that there is a relation between Bitcoin price and search engines. The relationship between Bitcoin and search engines, was established based on the rank of documents from specific search engines. In terms of forecasting, machine learning and deep learning have made some essential contributions to researchers in recent years [48]. Deep learning algorithms recently became famous for forecasting purposes. Many researchers have implemented models using deep learning methods for forecasting cryptocurrency prices and get great results with short-term and long-term data. Substantial work has been carried out for cryptocurrency price prediction using machine learning and deep learning methods, but there is still a lot to be done.

This section gives insight into the existing work that has been done in the field of cryptocurrency price prediction.

A. CLASSIC MACHINE LEARNING-BASED PREDICTION

This subsection discussed the different machine learning algorithms [48], [49] used by various researchers to predict the cryptocurrency price. For example, Madan et al. in 2014, got an accuracy of 57.4% using decision tree for Bitcoin, with a 10-minute sample price as the main input feature [50]. Then, Hegazy and Mumford [51] in 2016 predicted the price of Bitcoin, for every 8-minute interval and fed into the decision tree with an accuracy of 57.11%. Further, Chen et al. implemented both machine and deep learning models to predict the price changes in Ethereum [27]. Study in [27] had used machine learning-based algorithms like linear regression, naive bayes, SVM, random forest, ARIMA, and deep learning-based models like RNN and NN. They got the highest accuracy using ARIMA 61.17% for prediction.

Later, Connor et al. [52] used Naive Bayes, Regression
models, and SVM for predicting Bitcoin, Ethereum, and Litecoin prices considering the sentiments of news and posts of social platforms. Using Google trends data and Twitter texts, Jethin et al. [24] tried to predict the price of Bitcoin and Ethereum where the Twitter data was considered as a main feature. Further, the data cleaning was performed using the hashtags like #btc and #eth respectively, where #btc represents the tweet data for Bitcoin and #eth the tweet data for Ethereum. After pre-processing, they have applied a regression model for price prediction. Then, the authors of [53] considered the daily price of Bitcoin for prediction purposes and applied more than three normalization techniques on the dataset. Forecasting was performed via feature selection and implemented machine learning models such as random forest and Bayesian models considering five features, which were block size, the total volume of Bitcoin, day high, day low, number of transactions, and trade volume. Further, in [37], the authors have applied decision tree and linear regression techniques for forecasting the price of Bitcoin and compared the result of both and found that the linear regression technique outperforms the other on price prediction with high accuracy.

### TABLE 1: Comparative study of existing techniques for cryptocurrency price prediction.

| Ref. | Year | Description | Fusion? | Technique used | Expected result | Merit | Demerit |
|------|------|-------------|---------|----------------|----------------|-------|---------|
| [24] | 2018 | Authors had used different regression techniques for Bitcoin price prediction. | No | Linear regression, SVM, NA, Bayes, ADMA, and BNN | RMSE accuracy = 41.17% | Algorithmic changes on existing machine learning and deep learning algorithms done for Ethereum price prediction. | Not considered time-series-based models like LSTM and GRU. |
| [25] | 2018 | Authors had used different regression techniques for Bitcoin price prediction. | No | Theil-Sen Regression, Huber Regression, LSTM, and GRU | MSE ≤ 0.00002, R2 ≥ 0.992 (GRU) | Data Collection was focused on a 1-minute interval raw. | Ignored other impacting factors like sentiment. |
| [26] | 2018 | Machine learning-based algorithms were used for price prediction. | No | ANN | Accuracy = 75% - 95% | Prediction of cryptocurrencies prices on a hourly and daily basis. | Not explored other kind of machine learning and deep learning algorithms. |
| [27] | 2019 | Authors had used RNN and LSTM to predict Bitcoin prices with 10-fold cross-validation. | No | RNN/LSTM, Random Forest, and Linear Regression | LSTM MLE = 0.0063 | They had implemented 10 cross-fold validation to make a robust prediction model. | Absence of other impacting factors like opinion mining-based factors. |
| [28] | 2019 | Authors had compared and studied various models like DNN, LSTM, CNN, DBN for price prediction. | No | DNN, LSTM, and CNN | Accuracy, precision, recall, specificity, F1 score | Anomalous or seasonal hyper parameter factors and post-processing function on model. | Absence of wall-ticker users for factor predicting price direction. |
| [29] | 2019 | Authors of the paper proposed a model by conducting research on efficiency of LSTM-MN. | No | LSTM | RMSE = 0.046, MAE = 0.003 | Implemented statistical test on LSTM-based model. | Bi-directional LSTM can be considered before optimizing. |
| [30] | 2019 | Relationship between field prices and Bitcoin price prediction had been incorporated by authors through deep learning algorithms. | No | CSN, LSTM, and GRU | LSTM RMSE = 0.98 | Established the relationship between cryptocurrencies with gold and twitter sentiments. | Absence of live dataset input streams of various parameters affecting Bitcoin prices. |
| [31] | 2019 | Price prediction for Bitcoin using Yahoo Finance data was carried out using LSTM. | No | LSTM | RMSE = 265.5966 | Implemented a complex LSTM-based model for good results. | RMSE value is quite high. |
| [32] | 2019 | Relationship between Twitter sentiments and Bitcoin price prediction had been incorporated by authors through VADER algorithm. | No | VADER algorithm | 3-day window test accuracy = 57.84% | Usage of Twitter sentiments for determining the direction of price. | Not considered data from Facebook, Bitcoin forums, and other social media platforms. |
| [33] | 2019 | Authors had incorporated CSN, LSTM based hybrid model for price prediction. | No | CSN-LSTM | Prediction loss MAE = 209.65, RMSE = 203.1, MAPE used for comparison of different window sizes for Bitcoin price prediction. | Technical indicators and macroeconomic variables were used for price prediction through hybrid model. | The loss for prediction is quite high as compared to others. |
| [34] | 2019 | For cryptocurrency price prediction authors had used decision tree and regression techniques. | No | Decision tree and regression | Decision tree accuracy = 91.8011%, regression accuracy = 97.9526 | Explored decision-tree-based algorithm for cryptocurrency price prediction. | Deep learning based algorithms were not considered. |
| [35] | 2019 | Authors had done price prediction for Ethereum, Ripple using LSTM. | No | LSTM | MSE | Statistical analysis to justify the price prediction of cryptocurrencies was worth it. | Hybrid or ensemble model can be utilized. |
| [36] | 2020 | Authors had used MLP and LSTM for cryptocurrency price prediction. | No | MLP, and LSTM | MAPE MAE RMSE MSE | Consideration of business sector information for Bitcoin price prediction. | Two-tuning of hyper parameters is required to improve performance. |
| [37] | 2020 | Price prediction was carried out with a 2-day to 60-day window through BNN and MLP. | No | MLP and BNN | MLP accuracy = 91.70%, precision = 81%, and recall = 94.70% | Authors had done a detailed comparison of different window sizes for Bitcoin price prediction. | Instead of using BNN, the author can also go for LSTM-GRU for better long-term memory. |
| [38] | 2020 | Ensemble model was used for Litecoin and Monero price prediction. | No | GRU-LSTM | RMSE, MAPE, MAE, MAE used for evaluation purpose for Monero & Litecoin for window size 1, 3 & 7 days. | Introduced ensemble models for cryptocurrency price prediction. | Few technical factors from the protocol level can also be considered. |
| [39] | 2020 | Authors have demonstrated correlation between sentiments and cryptocurrencies prices through LSTM. | No | LSTM | MSE = 0.00151082 | Sentiment-based cryptocurrencies price prediction. | BNN or 6-directional model can be used for making more robust model. |
| [40] | 2021 | A deep learning-based hybrid model of LSTM and GRU used to predict the price of cryptocurrencies using inter-dependent relations with parents coin. | No | LSTM, and GRU | MSE | Dependability of Litecoin on the major coin like Bitcoin was also considered. | The bimodal of prices can be used as an input feature for better performance, which is not considered. |
| [41] | 2021 | The authors here used a hybrid model of LSTM, GRU, and TSN (Temporal Convolutional Networks) for price prediction based on historical data of Ether. | No | LSTM, GRU, and TSN | Accuracy: 1-day 84.2%, 4-day 79.9% | A novel type of ensemble model was used, which was a combination of LSTM, GRU, and TSN for better prediction. | The proposed model was not tested for the higher window like 15 days or 30 days. |
| [42] | 2021 | The proposed model, LSTM and GRU-based model, was tested on two different datasets. | No | LSTM, GRU | RMSE, MAE | The performance of the model was tested on two different datasets. | The influence of other major coins was not considered. |
| [43] | 2021 | The authors have proposed an improved RNN-based model for price prediction of Bitcoin. | No | Improved RNN | RMSE = 0.577440 | The model has proposed an improved RNN model which shows better performance compared to traditional RNN, and logistic regression. | The influence of other major coins was not considered. |
| [44] | 2021 | The authors have proposed a price prediction scheme based on long and short term integrated learning using SVR (Support Vector Machine). | No | Integrated learning with SVR | RMSE, R2 | The performance of the model was tested on many cryptocurrencies like Bitcoin, Binance Coin, Cardano, Dogecoin, Ethereum, XRP, and XMR. | The authors can also consider other factors like influence of major coin, which was not considered. |
| [45] | 2022 | The authors have proposed a hybrid scheme to predict the price of cryptocurrency with considering the impact of major coins like Bitcoin, and Litecoin on flash crash using fusion techniques. | Yes | LSTM-GRU | MSE | Dependency of the target coin was a major coin like Bitcoin, and Litecoin considered. | - |
TABLE 2: Acronyms

| Acronym    | Description                                      |
|------------|--------------------------------------------------|
| ARCH       | Autoregressive Conditional Heteroskedasticity    |
| AMH        | Adaptive Market Hypothesis                       |
| ANN        | Artificial Neural Network                        |
| ARIMA      | Autoregressive Integrated Moving Average         |
| ARIMAX     | Autoregressive Integrated Moving Average with Exogenous variable |
| Binance    | Binance chain Evolution Proposal                 |
| BFT        | Byzantine Fault Tolerance                        |
| CAGR       | Compound Annual Growth Rate                      |
| CNN        | Convolutional neural network                     |
| CUDA       | Compute Unied Device Architecture                |
| DAO        | Decentralized Autonomous Organization            |
| DNN        | Deep Neural Network                              |
| DRN        | Dilated Residual Network                         |
| EMH        | Efficient Market Hypothesis                      |
| ERC        | Ethereum Request of Comments                     |
| FFNN       | Feed Forward Neural Network                      |
| GASEN      | Genetic Algorithm based Selective Neural Network Ensemble |
| GPU        | Graphics Processing Unit                         |
| GRU        | Gated Recurrent Unit                             |
| HMM        | Hidden Markov Model                              |
| LSTM       | Long Short Term Memory                           |
| MLP        | Multi Layer Perceptron                           |
| MSE        | Mean squared error                               |
| NGO        | Non Governmental Organization                    |
| NN         | Neural Network                                   |
| PoS        | Proof of Stack                                   |
| PoW        | Proof of Work                                    |
| RNN        | Recurrent Neural Network                         |
| SARIMA     | Seasonal Autoregressive Integrated Moving Average |
| SHA        | Secure Hash Algorithm                            |
| SVM        | Support Vector Machine                           |
| VADER      | Valence Aware Dictionary and Sentiment Reasoner  |

B. DEEP LEARNING-BASED PREDICTION

Various existing works on deep learning-based forecasting models to predict cryptocurrency prices are discussed in this subsection. Deep learning approaches are highly accurate compared to the classic machine learning models [11]. It is because the deep learning models can able to find complex patterns in comparison to the classic machine learning models. For example, Sebastian et al. predicted the price of Bitcoin with a minute sample price as an input feature and passed to the feed-forward network to predict the directionality of Bitcoin, which gave 60% accuracy [54]. LSTM with its variant was used for forecasting the price of Bitcoin by Wu et al. in [55] and Tandon et al. in [31] with the same purpose of implementing the LSTM model along with the 10-fold cross-validation and got the MAE value of 0.0043s. Another deep learning model ANN was used by Radiyo et al. in [56], which was optimized using a genetic algorithm to avoid local minima by optimizing the initial weights of ANN. One variant of the genetic algorithm GASEN was used with ANN ensemble approach for Bitcoin price prediction. A hybrid of the Hidden Markov model and LSTM was proposed by Hashish et al. in [57] and got improved results compared to the traditional approach of ARIMA and LSTM. Bitcoin, Ethereum, and Ripple’s prices were analyzed using LSTM and ANN by Yiying et al. in [38]. Then, Mohil et al. in [19] proposed an hybrid LSTM and GRU-based scheme to predict the price of Litecoin and Monero with the RMSE of 2.2986 and 3.2715 respectively for 1-day prediction window, RMSE of 2.0327 and 5.5005 respectively for 3-day prediction window, and RMSE of 4.5521 and 20.2437 respectively for 7-day prediction window. Then, Rahmat et al. in [40] proposed an MLP and RNN-based scheme to predict the Bitcoin price for short and long time windows, where MLP showed the best accuracy, precision, and recall for both scenarios. However, Suhwan et al. in [32] implemented different combinations of deep learning models such as DNN, LSTM, CNN, and ResNet to forecast the price of Bitcoin. They also studied the effect of the window size prediction, the effect of using log values as inputs, the effect of data splitting, and the effect of normalization on classification and prediction problems. Then, Temiloluwa et al. in [33] applied different statistical tests to the LSTM model for better prediction because statistical tests are important for time series analysis and forecasting. In [25], the author considered the Bitcoin parameters and social parameters for forecasting purposes and compared the results of the deep learning models like CNN, GRU, and LSTM, and concluded that LSTM showed the best result among all with 4-layers. Then, the author of [34] used yahoo finance stock market data for predicting the price of Bitcoin using LSTM with 500 epochs and without a dropout layer achieved the RMSE of 288.59866. Later, Serafini et al. in [41] proposed a sentiment-driven scheme using statistical and deep learning approaches. ARIMAX and LSTM-based RNN were trained to predict the sentiment-based price prediction of Bitcoin with different feature combinations like (i) BTC volume, sentiment, tweets volume, (ii) BTC volume, sentiment, (iii) BTC volume, tweets volume, (iv) BTC sentiment, tweets volume, (v) BTC volume, (vi) BTC sentiment and (vii) BTC tweets volume. The author found that ARIMAX outperforms well compared to LSTM-based RNN and has an MSE of 0.00030187 on new predictions. Tanwar et al. [42] proposed a hybrid LSTM and GRU-based deep learning model, which outperformed the state-of-the-art techniques, to predict the price of cryptocurrencies. They have also considered the influence of the major parent coins on other altcoins, such as Bitcoin’s influence on Litecoin and Zcash. Politis et al. [43] has proposed an ensemble model of LSTM, GRU, and TSN (Temporal Convolutional Networks) to predict the price of Ether based on its historical price data and achieved 84.2% and 78.9% accuracy respectively for 1-day and 1-week. In the same manner, to predict the price of Bitcoin, Ethereum, and Litecoin, the author of the paper [44] proposed a system with LSTM and GRU. The price prediction was performed on two types of a data sample of Bitcoin, Ethereum, and Litecoin.

C. REINFORCEMENT AND FEDERATED LEARNING-BASED PREDICTION

Reinforcement learning is one of the types of machine learning techniques that focuses on learning based on its own experience and environment. It enables an agent that learns to achieve a goal in an uncertain, complex, interactive environment by trial and error using previous actions and...
This section discussed the work that was already being carried out to forecast the price of various cryptocurrencies. Starting off the section discussed why researchers find interest in this field of research. The relation of cryptocurrency price with gold, search engines, sentiment-based comments also tried to be established by researchers, discussed in the following section. Existing work on machine learning and deep learning techniques is also discussed in detail. Some methods that are not explored in literature and are available in suggested future research.

III. INFORMATION FUSION IN CRYPTOCURRENCY PRICE PREDICTION

This section revolves around the analysis of several factors involved in cryptocurrency price prediction. The prices are highly volatile, so a detailed analysis of important factors is necessary to build a robust forecasting model. Features like volume and past days’ prices plays a crucial role in cryptocurrency price prediction. So, the statistical analysis of volume and prices are highly required. The importance of amount concerning cryptocurrency price is demonstrated in [61]. For price prediction, [57] utilizes Hidden Markov theory for making a price prediction model. According to Hidden Markov theory, there exist hidden patterns among the data and this pattern can be learned by the forecasting model. So, to statistically conclude these results, we have to perform a cryptocurrency price analysis. Through this, we can find the hidden pattern, which is statistically significant [62]. Popular choice among investors, cryptocurrency risk analysis is a must. Even organization like Forbes [63] has surveyed and analyzed the risks related to cryptocurrency. So, risks also play an important role in cryptocurrency prices, and their analyses play an important role in price prediction. Further discussion on this topic is done in the following subsections.

A. CRYPTOCURRENCY VOLUME ANALYSIS

The total volume of the cryptocurrency here is referred to as the volume that is actively traded and not actively traded. The mass of cryptocurrency keeps increasing over the period. FIGURE 4 shows the total volume of Bitcoins from the year the currency was introduced, till 2021, similarly FIGURE 5 shows the total volume of Ethereum coins from the year the currency was introduced, till 2021.Proficient merchants and chartists use volume to extraordinary benefit, following the formula that assuming the cost falls alongside volume, it by and large denotes a reason behind weariness, flagging an inversion will happen soon.

While then again, an ascent in cost with a drop in complete volume presents a more grounded case for the bears as they drag costs for a lower bid, typically after gathering a key opposition zone. By mining, you can earn cryptocurrency without using cash. Mining rewards are paid to the digger who finds an answer for a complex hashing puzzle first. The likelihood that a member will be the one to find the arrangement is identified with the segment of the absolute mining power on the organization. The Bitcoin reward is halved every 210,000 units or roughly four years [64]. A portion of the Bitcoins available for use are accepted to be lost perpetually or unspendable, for example in case of lost passwords, a wrong destination address, or mix-ups in the output contents [64].
B. CRYPTOCURRENCY PRICE ANALYSIS

In this section, we analyze the various highs and lows of the price of different cryptocurrencies. The analysis of the price history of cryptocurrency helps us to know the nature of that particular cryptocurrency. This will help us to establish an interdependence among the cryptocurrencies, which is an aim of the paper. So the fake bubbles and crashes for the prices of the cryptocurrencies are discussed in detail as follows:

1) Bitcoin Price Analysis

Among the various assets such as gold, silver, cryptocurrencies, and other investment areas, the Bitcoin has a broad range of price variation in its trading history. The first time the insurrection in the price of Bitcoin was occurred in 2011, i.e., from $1 in April to $32 in June 2011, an increase of 3200% [65]. This lofty increase in the price was followed by a sharp downturn in cryptocurrency prices. The price of Bitcoin reached to $2 in November 2011 [65]. Following year price of Bitcoin went from $4 to $13.5 from May to August 2012 [65] and become one of the important years for the price of Bitcoin. In 2013, from the past data, we observed two price bubbles. The first bubble happened at the beginning of April 2013, when the price of Bitcoin reached the skies of $220 from $13.5 at the start of the year [65]. Due to processing delays in payment processes in April 2013 by processors BitInstant and Mt.Gox, the price dropped from $226 to $76 and then again returning to $160 within 6 hours [9]. But it was followed by a steep decline within 15 days, Bitcoin’s price reached to $70 in the mid of April 2013 [65].

Another price shot/bubble occurred nearly at the end of the year 2013. In October 2013, Bitcoin price was $123, but till the end of December 2013, the price shot up to $1156 and again slumped to $760 within the three days [65]. From this time, it started to show the volatility in the price of Bitcoin. Following years, the prices slumped many times, and at the beginning of 2015, the price reached $315 [65]. An increase in the price of Bitcoin was about to occur in the year 2017, which was the 5th bubble [65]. At the beginning of the year 2017, the trading price was $1000. After a decrease in the price for 2-months, in March 2017 the price was $975 [65]. But now there was a steep increase. At the end of the year, in December 2017 prices skyrocketed to $20,000 [65].

FIGURE 6 shows the same in graphical manner for better understanding. As seen from the FIGURE 6, for the next two years, the price of Bitcoin moved sideways, and there was not a big streak. In June 2019, the price surpassed $10,000, which created the hopes of another rally in Bitcoin prices. But the price declined to $7,112 in December 2019 [65]. In 2020 until the lockdown and economic shutdown occurred prices of Bitcoin came to the spotlight again. The economy shut down, and the government policies inset a fear in the investor’s mind which then fuelled the rise in the price of Bitcoin. In November 2020, Bitcoin was traded at $18,000 [65]. Following the Bitcoin prices rose to $24,000 in December 2020 and $40,000 in January 2021 but within three days price declined to $30,253 [65].

2) Ethereum Price Analysis

Ethereum is the second-biggest cryptocurrency, followed by Bitcoin based on its market cap. As of April 2021, according to [9], Ethereum has a share of 12% in cryptocurrency market cap. After its surge in 2015 with an initial supply of 72 million coins. After the surge of Ethereum in 2015, there was a bull run in the price of Ethereum till the 15th of January 2018. The same can be sum up from FIGURE 7. Ethereum prices show several ups and downs, and like other cryptocurrencies, it is also stochastic in nature. There is also a vast gap between the minimum and maximum annual returns. FIGURE 10 shows minimum annual return of -82.70% and max annual return of 9159.40%. So the analysis of Ethereum prices is required.

In this subsection, we will analyze Ethereum prices and relate its downfall with the factors listed in [12] [13]. Although the bull run of January 2018 was stochastic in nature, the prices of Ethereum increased from $0.67 to
$1332.62 [9]. The first bull run jumps the cryptocurrency prices by 1988.98%. After this long bull, the prices start to fall, and within a year its prices fall to $92.65. This fall decreases the price of Ethereum by 93%. Forbes article [67] said that this downfall was because of ICO’s projects. On the other hand, Financial Times [68] says that this fall was because of its enormous retail investors rather than the ICO projects. After this downfall, there is again a big and long bull run which increases its prices to $2346 at 18th April 2021. Although the bull run was stochastic in nature, the prices of Ethereum managed a jump of nearly 2532% from its last min price of $92.65. As of the latest trends, financial analysts [69] expect that this bull run is going on with a stochastic nature.

3) Litecoin and Ripple Price Analysis
Among 4200 cryptocurrencies, there are only a few that gives a high return on investment. Bitcoin and Ethereum are the ones that are popular among investors because of their high returns. There are other cryptocurrencies like Dash, Ripple, Litecoin, which are good options for investment. As of April 2021, yahoo finance [70] sorted a few cryptocurrencies for investment. To generalize the analysis, we have taken Litecoin and Ripple. As seen from the FIGURE 8, Litecoin shows a great amount of surge in some specific period of time. Similarly, FIGURE 9 shows the historical data of price of Ripple. FIGURE 9 shows some similarity with the FIGURE 7, which shows similar behavior of Ripple as Ethereum for certain period of time. By analyzing data from [9], there was a bull run till January 2018, and after that, there was a fall down till March 2020. Now again, there is a bull run, and analysts [69] also consider that there will be a stochastic bull run ahead. Litecoin price behavior has a lot of short up and downfalls [9]. So we can say that till now, Litecoin has not properly shown its downfalls like Ethereum or Ripple. At present, the Litecoin in its downfall, but the analysts [69] considered that there is a stochastic bull run ahead.

C. CRYPTOCURRENCY PATTERN PREDICTION
In this subsection, we will analyze the price history of cryptocurrency and attempt to predict a pattern for increase and decrease in the price of the cryptocurrency. These patterns can be useful to determine the direction of the cryptocurrency prices. The knowledge can be utilized for analyzing the bullish and bearish trend running on is real or fake. And this will helps us statistically to justify the idea of a forecasting model for cryptocurrency price prediction. FIGURE 10 shows the annual return patterns of Bitcoin and Ethereum. Accordingly, the statistical pattern will help investors to notice the trend. The following paragraph discusses various statistical patterns to determine trends in cryptocurrency prices.

To predict the pattern of uptrend or downfall, we need to analyze the historical data of cryptocurrency. Here, we concentrate on the returns given by Bitcoin and Ethereum in past to determine future patterns. The daily return study of Bitcoin and Ethereum, over the last four years, was done by the authors in [73]. The returns in [73], were derived with an assumption that $1 of each asset was purchased on the same day, and all the dividends were reinvested in the market. However, as we can notice from the graph that both of them follow the same pattern of increase and decrease.
in the return value, which means when the return value of Ethereum increases at the same time, the return value of Bitcoin also increases. As we can see that the spikes and steeps were simultaneous in both currencies. From the studies of [73] conclusion can be carried out that both Bitcoin and Ethereum must be interdependent in some manner because of the simultaneous increase or decrease in return percentage value.

Now, we will see some general patterns in cryptocurrency prices, which helps to predict the price movements from the graph. The graph patterns can be categorized into three classes: bullish, bearish, and reversal. Bullish patterns are the types of patterns, from which we can say that the price of the cryptocurrency is likely to go up and helps us to determine our profit margin. Similarly, bearish patterns are the opposite of bullish patterns. Whenever bearish patterns occur in the graph, cryptocurrency prices are likely to go down. Through this analysis, investors can book early profits. Reversal patterns are the shapes of graphs through which we can expect the trend reversal in cryptocurrency prices. Further, the detailed knowledge related to the three patterns discussed here can be found in [62].

### D. CRYPTOCURRENCY RISK ANALYSIS

The main distinguishing characteristic of cryptocurrency is, that they are not issued by any government or central authority. Hence, they have their advantages as well as disadvantages. The absence of backing by any central authority makes their value strictly volatile and leads to a loss of confidence among the investors and an abrupt price drop. The presence of a central authority guarantees the transaction comes with the facility to reverse the transaction in a condition of adversity. To maintain a unique security feature, no such facility is available in the case of cryptocurrencies. The lack of central authority means, if access to keys or an account is lost, then it can’t be restored and money will be lost forever.

Many users store their cryptocurrencies with third-party services, they have the risk of getting compromised. There is a large pool of criminal community which can infect computers with malware that steals the cryptocurrency. Similarly, bugs can also cause huge harm to the cryptocurrencies. It majorly affects when cryptocurrencies are combined with smart contracts, these are programs that automatically get triggered on the exchange of cryptocurrency. The first big catastrophic event once occurred when DAO tried to create an independent mutual fund. This system was affected by a bug that led the attacker to steal all the Ethereum in their account. The theft was in vain as the Ethereum developers soon released a new version of the system that invalidated the stolen cryptocurrencies [75]. There is no uncertainty that cryptocurrencies are going to stay as innovation progresses. Public acknowledgment and their trust will take some time yet, the dangers will continue as before. Some giving off an impression of being more material and raised for money and business [76].

![Graph](image.png)

**FIGURE 10:** Comparison of annual return rate between Bitcoin and Ethereum over the period of 2017-2021 [73] [74]

This section was, based on the analysis of important factors related to cryptocurrency price prediction. Firstly, we discussed the importance of statistical analysis of important factors. After that, we have a brief analysis of cryptocurrency volume. Following is the detailed analysis of cryptocurrency prices especially Bitcoin and Ethereum. Following is the cryptocurrency pattern analysis where we discussed bearish and bullish patterns for cryptocurrency prices. Further through the last subsection, we discussed the risk involved with an investment in cryptocurrency, and its effect on the rates of crypto coins.

### IV. STAKEHOLDERS ASSOCIATED IN CRYPTOCURRENCY PRICE PREDICTION

Stakeholders can be stated as all social actors that can be affected either positively or negatively due to decisions, objectives, and action plan of the company [77]. In spite of the fact that cryptocurrencies being transparent, secure, and decentralized systems. By connecting the dots, there is an inference that no organization can control cryptocurrency prices. But, somebody should be governing them otherwise it will cause havoc in the cryptocurrency market. These facts were studied by the author of [78] discussing the governance crisis of the cryptocurrency market. So, a collaborative organization like Bitcoin organization has been formed for its governance worldwide. Their main work is to make two users’ software i.e. wallets, compatible with each other for transactions, but they don’t have any control over its prices [79]. So now, are there any other stakeholders for cryptocurrency price prediction?

Yes, their are stakeholders are affecting the price of cryptocurrencies. Messages and tweets from stakeholders can manipulate the cryptocurrency prices. Sentiments of their messages and tweets influence the prices [25]. For the sake of understanding purpose, let’s consider an imaginary...
situation where Barry Silbert, one of the major Bitcoin investors [80], stated from his Twitter handle that he convert all of his Bitcoin to USD. Considering the behavioral finance hypothesis, within few hours we might see a significant drop in Bitcoin prices. Similar types of features were used for making a sentiments-based forecasting model in [81]. So we can conclude that there are stakeholders associated with cryptocurrency price prediction. Based on [77], categorization of stakeholders is done into two categories:

- **Primary Stakeholders:** They are the one whose decision directly affect the price of cryptocurrency. Generally, developers of Bitcoin organization, N.G.O’s like Bitcoin foundation dealing with countries government for making Bitcoin as a valid mode of transaction, but has been dissolved in 2015 [82]. Wallet companies like Coinbase, then venture capitalists or investors and miners are the ones considered as primary stakeholders. Few cryptocurrencies are also there, which is being controlled by its possessing company who had developed it like Libra is owned by Facebook [83].

- **Secondary Stakeholders:** They are the one whose work or research indirectly affects the price of a cryptocurrency. Associated with this category are companies or researchers working at the protocol layer of any cryptocurrency for improving specific factors like transaction time. Companies or researchers working on implementing a bot for cryptocurrency portfolio management like [84]. Companies or researchers working on forecasting models like the ones in Table 1 and this paper authors also fall under this category.

### A. COMPANIES/ORGANIZATIONS ASSOCIATED

Companies associated with cryptocurrency price prediction are of both types. Few of the cryptocurrencies are still under the control of the companies who developed it like Air-Asia is the proprietor of BigCoin [86]. So they might have control over its price. Startups and companies designing protocols like Ripple [87] are also important stakeholders in price prediction. Their main work is at the core level so how are they important for the price prediction might come to your mind. Let’s consider a hypothetical situation, where a Ripple makes few changes in its cryptocurrency by which it can make transactions three times faster than its previous generation. Through these, big giant companies started to make transactions through Ripple, leads to increase demand in the market. An increase in demand will make the prices of Ripple reach unbelievable limits. So yes, protocol companies and organizations play an important role in the prices of cryptocurrencies.

Investors, startups, organization, and individuals plays an important role in cryptocurrency price prediction. It is because, they are the one who creates a demand and supply chain in it [88]. As discussed above, based on the behavioral finance hypothesis, their statement can cause a shift in the market. Mining organizations like Genesis-Mining, are also prime stakeholders for price prediction because they are the ones who generate volumes of cryptocurrencies in the market. Mining-based startups are versatile, a couple of them like BitFury [89] generates revenue by making and selling mining machines and rigs in the market. On the other hand, companies like Genesis-Mining [90] is a cloud platform where users can mine cryptocurrency by paying the cost of hardware power consumption they utilized. Companies like BitMain generates revenue by mining Bitcoin only [91].

B. RESEARCHERS ASSOCIATED

Researchers associated with cryptocurrency are categorized as secondary stakeholders. The field of research in the domain of cryptocurrency and blockchain is vast. So, researchers from a different area other than price prediction are also cover under the secondary stakeholders. Article [92] gives a detailed explanation regarding the research development in the discipline of cryptocurrency and blockchain. A couple of them plays a crucial role as stakeholders. The evolution of cryptocurrency is also one of it. The discussion of this factor is essential because whenever a new cryptocurrency is initiated they are derived from other cryptocurrencies with few changes. Simultaneously, the introduction of new crypto coins increases the market capitalization and volatility of the cryptocurrency market too. For example, Bitcoin Cash has been derived from Bitcoin, based on price analysis, when Bitcoin Cash set in motion during 2017 after one year in a period of 2018 a cryptocurrency market crash has happened. In this market crash, Bitcoin was the worst affected. As Bitcoin Cash was just inaugurated, they faced the worst during the end of the year, forcing the developers to split into two parts, namely Bitcoin Cash and Bitcoin [93]. Although, there are many other factors as well [94], from present drift whenever Bitcoin price escalates, Bitcoin Cash price also rises simultaneously.

Researchers are also involved in the zone of designing forecasting systems, similar to the prediction model stated in Table 1 and portfolio management bot like [84] have a major contribution to cryptocurrency price prediction. Big investors and organizations use the statistical output of forecasting models to conclude whether they should invest or back off their money from the market. Their research works are indirectly affecting the price of cryptocurrency. The same goes for the research work done in the field of portfolio...
management bot. Through their past learning, the bot can invest clients’ money or back off their money from the market to obtain profit. Based on the demand-supply concept [95] when thousands of such bot conclude the same result from the environment videlicet from the market, it can cause a sudden surge or fall in the price of the cryptocurrency. So, indirectly the researchers working in this field of research also affect the cost of the cryptocurrency. FIGURE 11 gives the bibliometric analysis of researchers associated with the domain of cryptocurrency price prediction, the data is retrieved using Microsoft Academic API.
Many factors that affect the change in cryptocurrency prices and this section discusses those metrics in detail. Some factors affect the price of cryptocurrency directly and some which affect it indirectly. Public perception is one of the most significant factors that indirectly affect the change in the price of any cryptocurrency. In [99], D’Alfonso et al. in 2016 studied the relation of online news with the price of Bitcoin and found that the number of online talk increases, which means that positive talk about Bitcoin leads to an increase in price and negative talk about Bitcoin lead to a decrease in price. Google search is also one of the indirect factors that affect cryptocurrency price. Google trends for Ethereum show that there is a high relationship between the maximum price, average price, and the total traded volume. In sentiment-based models, one can find features as crowd-sourced data having sentiment-based comments from the online social platform like Google Trends, and Twitter. The utilization of factors from the blockchain part can be helpful for cryptocurrency price prediction. A similar approach is done by the author of [32] for predicting the price of Bitcoin. They also found the correlation between various features and Bitcoin prices in the right amounts.

Apart from the factors mentioned above, some other factors have a significant impact on the price of cryptocurrency. They are as follows:

- **PoW**: It is one of the most important metrics that affect the price of cryptocurrency significantly. For each mining process, miners should submit the PoW having genuine work. A new cryptocurrency has modified PoW and name it as a PoS having rewards for the miners based on their holdings.
- **Volume**: In cryptocurrency price prediction, the volume of any cryptocurrency plays an important role as it can be broken down in many ways. By observing the volume change, the researcher can find the probable movements and direction of the coin.
- **Closing and opening price**: Many researchers have successfully implemented the model having closing price and opening price of the cryptocurrency data day wise. Closing price and opening price pattern also helps to determine the probable ups and downs of price.

### A. Correlation between Metrics

Several correlations have been found among factors stated in [12]. Volume and price correlations are altogether famous among the traders as mentioned in [101]. The author of [101] also concluded that the relationship between volume...
and price is inversely proportional. Next comes the analysis of volume and mining algorithms, two factors that affect the volume of newly introduced cryptocurrencies. To study the affecting attributes, we have chosen Bitcoin. Firstly the mining rate of the new Bitcoin is fixed. New Bitcoins are introduced into the market when miners process blocks of transactions. The rate at which new coins are introduced by miners is slowing over time \[102\]. If we analyze the above statement, we can derive the rate at which the volume of cryptocurrency introduced to the market is inversely proportional to the complexity of mining algorithms (PoW or PoS). Since price is inversely proportional to the supply of a cryptocurrency, we can also conclude that price and mining algorithms are directly proportional. The studies in \[102\] also stated that indirect factors like market sentiments, opinions, regulations and legal matters, and other factors are also directly proportional to the price of cryptocurrency.

### VI. DEPENDENCIES BETWEEN CRYPTOCURRENCIES

The forecasting model proposed in this paper through fusion has considered dependency between the cryptocurrencies as an important metric. Top 100 cryptocurrencies based on their market cap \[9\] has been considered to study dependency among them. From the price analyzes, we found that cryptocurrencies are of two types:

- **Stable cryptocurrency**: This type of cryptocurrency price is always fixed, specifically their exchange rate is always equal to 1 USD. Cryptocurrencies like Tether, USD Coin, and Paxos are a few of the examples.
- **Unstable cryptocurrency**: This type of cryptocurrency price keeps fluctuating similar to stocks, in other words, their exchange for 1 unit keeps on changing by minutes. Cryptocurrencies like Bitcoin, Ethereum, Litecoin, and Ripple are a few of the examples.

Based on the above categorization, the forecasting model can be used only for unstable cryptocurrencies. As there is no requirement to predict the price of a stable one because their prices are always fixed. Although, the forecasting model for the exchange rate of stable coins for other countries' currencies like Rupees, Euros, and Pounds keeps fluctuating. This factor is not from cryptocurrency volatility but is from the volatility of a specific country’s currency with respect to U.S. Dollar. Models similar to \[103\] \[104\] \[105\] can be utilized to predict exchange rates for a specific country’s currency with USD. This can be incorporated as a forecasting model for stable cryptocurrencies concerning a country’s currency other than USD. So, the research field on the forecasting model revolves around unstable cryptocurrencies only. For the statistical inference, FIGURE 13 shows the correlation among Bitcoin, Bytecoin, Dash, Litecoin, Monero, Ethereum, and Zcash. In the following subsections, there is a brief analysis of main or parent cryptocurrencies.

#### A. BITCOIN

Bitcoin being the first most cryptocurrency, is considered as the parent for most of the other cryptocurrencies. The newly introduced cryptocurrencies made a few changes in their protocol to improve the performance in terms of anonymity, speed of transaction, and other crucial features. Bitcoin holds nearly 78% value of the market cap during April 2021 \[9\], and its effect on other cryptocurrencies makes it a popular choice as a crypto coin for researchers to determine the performance of the forecasting model. There are many

---

**FIGURE 13: Heat map representing the correlation among popular cryptocurrencies \[66\].**

In the beginning of the section, we have done a short analysis of existing work and discusses important metrics related to cryptocurrency price prediction. After that, we have tried to find the correlation between the important metrics like prices, volume, and complexity of a mining problem. After that we have done a thorough discussion on the requirement of metrics fusion and also have an overview for our next section.
currencies derived from Bitcoin, further we discuss the major ones in the following paragraph.

Bitcoin Cash is one of the derived cryptocurrencies from Bitcoin by improving the size of the block to increase scalability and performance of transaction [93]. Litecoin’s is also among the cryptocurrencies that are derived from Bitcoin. Litecoin developers aim to increase the speed at which new blocks are added to blockchain [106]. Bitcoin SV is derived from Bitcoin Cash [107], and Bitcoin Cash, furtherly derived from Bitcoin [93]. So, indirectly Bitcoin SV depends on Bitcoin. In addition, SV stands for “Satoshi Version” [106] and aims to restore the original version of Bitcoin introduced in [1] with the belief of cheaper fees. Bitcoin Gold is also the one forked from Bitcoin. Bitcoin Gold developers aim to change the PoW algorithm to democratize mining processes [106].

Zcash is a fork of a Bitcoin code-base. Zcash developers did not change anything at the protocol level, but they introduced a new option for users where they can make their transactions public [106]. This feature has attracted many governments because the dealings can be kept visible with the will of users. Through this government can trace the fraud easily which is not possible for other cryptocurrencies [108]. Ravencoin is a fork of a Bitcoin code-base. Ravencoin developers had made four crucial changes: they modified the issuance schedule (with a block reward of 5000 Ravencoin), the block time reduced to one minute, their coin supply capped at 21 billion (ten times more than Bitcoin), and finally, a new mining algorithm (KAWPOW, formerly X16R and X16RV2 respectively) intended to mitigate the centralization of mining through Application-specific integrated circuit (ASIC) hardware [109]. Dash is a fork of Litecoin, through which it is indirectly a fork of Bitcoin [9]. There are many others as well with slight changes in the Bitcoin code-base. Through this, we observe the importance of Bitcoin in cryptocurrency’s market cap. And their influence both technically and financially on other cryptocurrencies.

B. ETHEREUM

Ethereum is one of the famous and second-most versatile cryptocurrencies contributing nearly 12% to cryptocurrencies market cap [9]. Ethereum resembles a similar feature with Bitcoin [110]. However, it is not forked from the Bitcoin code-base. On the other hand, the motivation of its introduction is different from Bitcoin. Bitcoin was introduced as a digital currency, on the other hand, Ethereum is a ledger technology that companies are using to build new programs, and Ethereum coin is a part of this ledger. Ethereum is far more robust than Bitcoin [111]. Many other properties at the protocol level are also different, such as decrement in the amount of time for confirmation of the transaction, developers of Ethereum had used ethash instead of SHA-256 [110]. For this reason, we have considered Ethereum as a separate cluster from Bitcoin, and also there are a large number of cryptocurrencies forked from Ethereum. In the following paragraph few of them will be discussed.

Ethereum Classic is the one that has been emerged from the Ethereum blockchain. Initially, Ethereum Classic and Ethereum were the same till the incident of DAO. After that, few developers had made a hard fork of Ethereum and introduced Ethereum crypto coin. And other developers decided to use the existing code-base, with no changes as Ethereum Classic [112]. Binance Coin initial release had incorporated the ERC-20 token, the same token used in Ethereum Blockchain. However as of now, ERC-20 tokens are replaced with BEP-2. Developers replace PoW with BFT, so mining is not possible [9]. Chainlink is also based on ERC-20 tokens with additional ERC-223 for transfer and call functionality [9]. Aave utilizes ERC-20 tokens of Ethereum ledger with open source implementation [113]. Currently, they are transferring the PoW algorithm with PoS algorithm [9]. There are many others as well who made a fork from Ethereum. From this, we can observe the importance of Ethereum to cryptocurrency’s market cap. And its influence both technically and financially on other cryptocurrencies.

C. HYBRID

Till now, we have seen that cryptocurrencies were derived from popular or versatile cryptocurrencies like Bitcoin and Ethereum. However, many developers tried to combine two or many cryptocurrencies at the protocol level to make a hybrid one. This subsection discusses a few of them. Wrapped Bitcoin is the tokenized version of Bitcoin running on Ethereum Blockchain [9]. The main motive of its developers is to use Bitcoin in the Ethereum decentralized environment [114]. Qum is another cryptocurrency that is a hybrid of Bitcoin and Ethereum. Qum tokens were derived from Ethereum tokens, specifically ERC-20. And their transaction model is UTXO derived from Bitcoin [9] [115].

D. INDEPENDENT

Till now, we have seen how new cryptocurrencies are derived from other ones. But there are cryptocurrencies from which new ones were derived, so their study is also required. This subsection discusses a few of them. Bitcoin is the first cryptocurrency introduced in [1], which makes it an independent one. Ethereum is the first and original smart contract-based blockchain project. Ethereum is independent because of its ERC-20 token derived from the Ethereum blockchain. Binance Coin during their initial period was derived from Ethereum. But as of now, Binance Coin is independent. They introduced their own token BEP-2 [9]. Bytecoin being the first cryptocurrency based on cryptonote technology makes it an independent cryptocurrency [116]. Cryptocurrencies like Monero [9] and Dashcoin [117] are derived from Bytecoin. Ripple is another independent cryptocurrency that runs on RippleNet and Ripple ledger [9].
This section analyzes the facts and discusses interdependency among cryptocurrencies. Initially, we have categorized cryptocurrencies into two categories based on their price or exchange rates. Later, in two subsections we discuss cryptocurrencies that are directly or indirectly derived from Bitcoin and Ethereum. After that, we had a brief discussion on hybrid cryptocurrencies. This kind of cryptocurrency was formed by combining the properties of two or more cryptocurrencies. Following is a subsection that discusses the cryptocurrencies that are independent of other cryptocurrencies. Their independent nature can be analyzed, through their code-base, protocols, and other technical aspects.

VII. PROBLEM FORMULATION

The proposed model is based on the fusion of cryptocurrencies in the context of interdependency among them. From the discussion in Section VI, we have picked out that there is an interdependency among Bitcoin and Litecoin, Litecoin, and Dash. So, hierarchically there is an interdependency between Bitcoin and Dash. To justify the fusion, we consider the prediction of Dash based on the prices of Dash with the inclusion of Litecoin and Bitcoin prices as well. The forecasting problem is a supervised learning problem. And under supervised learning, it is a regression problem. As we are predicting prices, the output from the model will be the predicted price for that particular day. But what will be feed as an input to the model?

We have the price history of all cryptocurrencies. This data can be incorporated to forecast the price of the present day. Let the cryptocurrency prices at the individual timestamps be \( p_1, p_2, p_3, \ldots, p_n \), where \( p_i \) denotes price at time stamp \( i \). Let the input window length be \( w \), the input vector be \( i_p \), price of Bitcoin be \( pb \), price of Litecoin be \( pl \), price of Dash be \( pd \) and the output be \( o \), mathematically \( i_p \) and \( o \) can be denoted as:

\[
\begin{align*}
    ip_{pb} &= [pb_{i-w+1}, pb_{i-w+2}, \ldots, pb_{i-2}, pb_{i-1}] \\
    ip_{pl} &= [pl_{i-w+1}, pl_{i-w+2}, \ldots, pl_{i-2}, pl_{i-1}] \\
    ip_{pd} &= [pd_{i-w+1}, pd_{i-w+2}, \ldots, pd_{i-2}, pd_{i-1}] \\
    ip &= [ip_{pb}, ip_{pl}, ip_{pd}] \\
    o &= pd_i
\end{align*}
\]

where \( w \) is the window size, \( ip_{pb} \) is the preprocessed input of Bitcoin, \( ip_{pl} \) is the preprocessed input of Litecoin, and \( ip_{pd} \) is the preprocessed input of Dash. Our motive is to predict output \( o \) using input vector \( ip \), which contains prices of past \( w \) days value of Bitcoin, Litecoin, and Dash. The data for training and testing is arranged into multiple input-output pairs as shown in (4) and (5) respectively.

FIGURE 14 shows the working of the proposed system model. For training, the data is collected from Investing.com [66] and then preprocessed. For preprocessing we have normalized through a division operator. Let \( x \) be the current price and \( x_{\text{max}} \) be the maximum price seen till now, and \( y \) is the difference between the number of digits in 1 and \( x_{\text{max}} \), and \( x_{\text{new}} \) be the normalized price then \( x_{\text{new}} \) can be denoted as:

\[
x_{\text{new}} = \frac{x}{10^y}
\]

Through the normalization used (6), values will be in the range of 0 to 10. After preprocessing, the merging of data is done according to (4) to make an input tuple. After merging, the data is further divided into two parts training and testing data. Model weights are trained using training data and their performance is evaluated using testing data. For predicting price at \( i \)-th day, we have to feed past \( w \) days (window length) data of all dependent cryptocurrencies (here Bitcoin, Litecoin, and Dash). The forecasted price is fed into the model with past \((w - 1)\) days of prices to predict the \((i + 1)\)-th day price consequently. This process is repeated several times equal to the prediction window length. The new \( w \) days price is compared with test data to evaluate the performance. The proposed scheme uses a hybrid model combining GRU and LSTM. In the following subsection, we will see the working of LSTM and GRU neurons and why the classic RNN is not used for making system models.

1) Recurrent Neural Network

RNN is a type of feed forward neural network, whose output is derived from the input features and the outputs of all hidden neurons of that RNN layer at the previous timestamp or previous stage. This special property helps the RNN to solve the problems where the result of the present situation is based on the past scenarios. Price prediction, text classification, and language generation are the few problems that can be solved using RNN [118]. Mathematically RNN is denoted as [119]:

\[
a_t = \sigma (W_{\text{hidden}} a_{t-1} + W_{\text{input}} x_t + b_a) \\
y_t = \sigma (W_{\text{output}} a_t)
\]

where \( x_t \) is the input at timestamp \( t \), \( a_t \) is the output of hidden neuron at timestamp \( t \). \( W_{\text{hidden}}, W_{\text{input}} \) and \( W_{\text{output}} \) are the weights of hidden neuron, input neuron and output neuron respectively. Activation function in both hidden and output neuron is sigmoid function. Detailed mathematical description of RNN is given in [120], and [121] gives the overview of RNN. FIGURE 15 represents graphical structure of RNN neuron. So, is the RNN sufficient to solve our forecasting problem?

The specialty of RNN helps us to solve many problems, including the forecasting problem. But, there are few drawbacks. The RNN is, trained through back-propagation. They are cyclic in nature which makes them difficult to train and take time to converge [121]. They also suffer from a vanishing gradient problem due to which we cannot get proper output for long sequences or long window size [122] [123]. To overcome the vanishing gradient problem few steps in preprocessing [124] or mathematical operations like
gradient clipping might help. However, this also fails when the data is having long sequences or window sizes. Two new models LSTM [125] and GRU [126] are introduced. They are derived from RNN with an additional feature called memory cell. This memory cell helps them to avoid vanishing gradient problems by remembering the essential past features only. Both of them also works well while feeding long sequences or long window size of data. In the following subsection, we will take a brief look at LSTM and GRU structure.

2) Long Short Term Memory
LSTM is a variant of RNN that contains memory cells in it [125]. Each neuron of LSTM consist of four other sub neurons in which act as a memory cell. The weights of these sub neurons helps LSTM to remember long sequences. A classic LSTM cell consist of an input gate, forget gate, and output gate. (9), (10), and (14) represents mathematical form of the respective gates. Their mathematical equation can be denoted as [119]:

\[
\begin{align*}
  i_t &= \sigma(x_t U_i + a_{t-1} W_i) \\
  f_t &= \sigma(x_t U_f + a_{t-1} W_f) \\
  o_t &= \sigma(x_t U_o + a_{t-1} W_o) \\
  \tilde{C}_t &= \tanh(x_t U_g + a_{t-1} W_g) \\
  C_t &= i_t * \tilde{C}_t + f_t * C_{t-1} \\
  a_t &= \tanh(C_t) * o_t
\end{align*}
\]
where \( x_t, a_{t-1} \) and \( C_{t-1} \) will be the inputs, and \( a_t \) and \( C_t \) will be the output from the cell. Symbols \( i \) is for input gate, \( f \) is for forget gate, and \( o \) is for output gate. And symbols \( W, U \) and \( b \) are the weights of each sub neuron or memory cells. FIGURE 17 represents graphical structure of the GRU cell.

3) Gated Recurrent Unit

GRU is the latest variation of RNN introduced by [126]. The working of GRU is similar to that of LSTM. Even though, the structure is similar but [126] remove one hidden sub neuron and one gate from LSTM. GRU has three sub neurons and two gates which helps them to resolve the vanishing gradient problem. Two gates are an update gate and a reset gate. Like LSTM, the gates in GRU controls the flow of information through the network. Update or output gate controls information flow from the past that needs to be passed further. The reset gate determines the amount of information to forget. (15), and (16) represents mathematical representation of respective gates and output of GRU is represented by (18). Their mathematical equation are [119]:

\[
\begin{align*}
    u_t &= \sigma(Uu x_t + Wa_{t-1} + b_u) \quad (15) \\
    r_t &= \sigma(Ur x_t + Wr a_{t-1} + b_r) \quad (16) \\
    \tilde{a}_t &= \tanh(Ua x_t + Wo(r_t \odot a_{t-1}) + b_o) \quad (17) \\
    a_t &= (1 - u_t) \odot a_{t-1} + u_t \odot \tilde{a}_t \quad (18)
\end{align*}
\]

where \( x_t, a_t, u_t \), and \( r_t \) is the input, output, update gate output, and reset gate output respectively. Symbols denotes the Hadamard product and \( U, W, \) and \( b \) are the weights and biases matrices. FIGURE 17 represents graphical structure of the GRU cell.

VIII. PROPOSED FUSION-BASED SOLUTION MODEL

In this section, we elaborate the proposed fusion-based cryptocurrency price prediction model. To represent fusion, through dependency among cryptocurrencies, we predict the price of Dash based on the previous days’ prices of Dash, Litecoin, and Bitcoin. In section VI, we have seen that Litecoin is derived from Bitcoin and Dash is derived from Litecoin. The model for the prediction task is based on LSTM and GRU. RNNs are well known to remember sequential data, but as discussed in Section VII it suffer from vanishing gradient problem. LSTM and GRU overcome this problem with memory cells which makes them fit for this task. The authors in [28], [39], [32], [34] managed to get good results using RNN, LSTM and GRU.

First of all, data available with us are continuous historical data. The prices of Bitcoin are in the range of $60,000 whereas Litecoin and Dash are in the range of $200. This vast gap between the range of prices will makes the task complicated, and requires a lot of computational time for model training. So, we have normalize the data using (6) that make every data in the range of 0 to 10. Reasons for not using standard normalization technique like Min-Max normalization is, discussed in Section IX. After normalization, the data is arranged in a suitable format of input-output pairs as per (4) and (5). The input is a sequence of past \( w \) day’s prices of all three cryptocurrencies and output will be the present-day price of Dash. Here \( w \) is a window size which is a hyper-parameter. These multiple input-output
tuples are divided into two parts. The first part is used to calculate the optimal values of parameters for the model, and the second part was used to determine its robustness and performance.

The next step is to train the model using training data and analyze its performance on the test data. The model consists of three sub-models one for each cryptocurrency. Through these three sub-models, latent features are extracted and after that, all three latent features are concatenated to make a dense model. This dense model is further given us the forecasted price. Input is fed to the 100 neurons LSTM layer. The output of this LSTM layer is passed to the 100 neurons GRU layer. After that, we have 100 neurons of the dense layer and after that another dense layer of 10 neurons. Now we have 10 latent features from each cryptocurrency. So, in total, we have 30 latent features after the concatenation layer. Following the concatenation, we have a single neuron dense layer, which is the output layer that gives us the forecasted price. All the layers except the last layer that has ReLU as an activation function, and the last layer has a linear activation function. The difference between the output from the model and actual value will be used to determine loss, and through Adam optimizer this loss will be decreased to a global minimum value. The model has been trained for 50 epochs. FIGURE 18 is the graphical representation of above described fusion-based model.

After training, the losses are compared to determine its performance. To justify our model, we train our model with various input patterns such as Dash only, then Dash and Litecoin, and after that Dash, Litecoin, and Bitcoin. For comparison purposes, we have used an MSE as our loss function. And for justifying the use of the (LSTM + GRU) model, we compared the loss with LSTM and GRU. In the below subsection, there will be a graphical comparison among LSTM, GRU, and (LSTM + GRU). The comparison was carried out through window sizes of one, three, seven, and thirty.

IX. PERFORMANCE EVALUATION

This section gives an insight to the dataset used for the proposed prediction model, feature selection for various cryptocurrencies, data preprocessing phase for each cryptocurrency, evaluation measurement for the proposed fusion-based model, and the final result is obtained by the proposed model which is described in detail analysis.

A. DATASET DESCRIPTION

The data was collected from the global portal Investing.com [66] for research purposes, which provides real-time data like the news and analysis of the financial market. The data for cryptocurrencies like Bitcoin, Bytecoin, Dash, Ethereum, Litecoin, and Monero was collected daily which adds features as follows.

- Price: Average price of each cryptocurrency day-wise
- Open: Opening price of each cryptocurrency day-wise
- High: Highest price of each cryptocurrency day-wise
- Low: Lowest price of each cryptocurrency day-wise

Data for each currency were collected for January 24, 2018 - April 01, 2021 (1164 data points). The proposed fusion-based model takes average daily prices of Bitcoin, Dash, and Litecoin as an input feature and FIGURE 13 concluded their dependency.

B. DATA PREPROCESSING

In the proposed fusion-based model, we pass the price data of Bitcoin, Litecoin, and Dash as an input for predicting the price of Dash. The range of price data of each currency is not the same and varies a lot. Due to these reasons, we cannot use the data directly for training purposes. We found some major issues with the existing normalization techniques and their applicability to this type of data for predicting prices. Techniques like min-max normalization are popular among the research community. To predict the price of Dash, if we provide input as the min-max normalized form then the model will not able to predict the price of Dash when it gets higher than the maximum value of training data and lower than the minimum value of training data. Due to this, we cannot use this kind of normalization technique and eventually can not fulfill our purpose for predicting the price accurately. (19) is mathematical presentation of min-max normalization [127].

\[ x_{\text{min-max normalized}} = \frac{x_{\text{original}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

where \( x_{\text{min-max normalized}} \) is the transformed value, \( x_{\text{original}} \) is the original input value, \( x_{\text{min}} \) is the minimum value from the input data and \( x_{\text{max}} \) is the maximum value from the input data. To avoid the above-discussed situation and for aligning price data in one range we have incorporated weighted normalization technique where is already explained in the above section, (6) that transforms the data in the range of 0 to 10. After normalization, we select price as a primary feature for the proposed fusion-based model and give input as the price of Bitcoin, Litecoin, and Dash.

C. EVALUATION METRICS

Evaluation of the proposed fusion-based model is done using loss function namely MSE, which is observed for the singular model, dual dependency model, and hierarchical dependency model. Mathematically an MSE can be denoted as [128]:

\[ MSE = \frac{1}{N} \sum_{i=0}^{N} (\hat{p}_i - p_i)^2 \]  

where \( \hat{p}_i \) represent the predicted price, \( p_i \) represents actual price, and \( N \) is the number of observations (data points).

D. RESULTS

The result of the proposed model and various simple models compared in this subsection. The fusion-based
proposed model experimented with different predicting window lengths: 1 day, 3-days, 7-days, and 30 days. We provide 1000 data points for each cryptocurrency as input for training, and testing data is different from training data on which the prediction task was performed. Training dataset length fixed for all window size that is 1000, and testing dataset length is variable that are as follow:

- 1 day: Test dataset length: 163 data points
- 3 days: Test dataset length: 161 data points
- 7 days: Test dataset length: 157 data points
- 30 days: Test dataset length: 134 data points

To predict the price of Dash, two models were implemented besides the fusion-based prediction model. Without considering any dependency, the simple model having Dash price as an input feature implemented. Considering dual dependency, the second model with an input feature price of Dash and Litecoin implemented. The final fusion-based model is implemented by considering the multiple hierarchies, and have the price of Dash, Litecoin, and Bitcoin as input features. These three models were implemented with three varieties of the model that are as follows: only LSTM, only GRU, and fusion of LSTM and GRU models.

A simple model, for predicting the price of Dash with a different window size was implemented on LSTM, GRU, and LSTM-GRU based model with 1000 data points for the training. A single dependent model takes two input features, i.e., the price of Dash and Litecoin’s 1000 data points of each for training, whereas in the hierarchical dependent model the 1000 data points of each of Bitcoin, Litecoin, and Dash are given as an input feature.

1) Scenario 1: 1-day price prediction

In a 1-day price prediction window size, 1000 data points passed as input for training and 163 data points for testing purpose to calculate loss. In the 1-day price prediction window size, the J-day price passed as an input, and the price of the next day is estimated.

A simple model with input as a Dash’s price with LSTM unit gives an MSE loss of 0.0192. A GRU unit produces an MSE loss of 0.0214, and (LSTM + GRU)-based model delivers the MSE loss of 0.0208. A single dependent model with input as Dash and Litecoin’s price with LSTM unit gives the MSE loss of 0.0176. A GRU unit produces an MSE loss of 0.0180, and LSTM+GRU based model delivers an MSE loss of 0.0192. A multiple dependent model with input as the price of Dash, Litecoin, and Bitcoin with LSTM unit gives the MSE loss of 0.0205. A GRU unit produces an MSE loss of 0.0185, and LSTM+GRU based model delivers an MSE loss of 0.0205. The loss observed by all prediction models for 1-day window size is displayed in FIGURE 19a, and FIGURE 19b shows the actual price and predicted price by our proposed fusion-based model (LSTM + GRU) in USD.

2) Scenario 2: 3-day price prediction

In a 3-day price prediction window size, 1000 data points passed as input for training and 161 data points for testing purpose to calculate loss. In the 3-day price prediction window size, the 3-day price passed as an input, and the price of next day is estimated.

A simple model with input as Dash’s price with LSTM unit gives the MSE loss of 0.0202. A GRU unit produces an MSE loss of 0.0331, and (LSTM + GRU) based model delivers an MSE loss of 0.0187. A single dependent model with input as Dash and Litecoin’s price with LSTM unit gives the MSE loss of 0.0186. A GRU unit produces an MSE loss of 0.0187, and (LSTM + GRU) based model delivers an MSE loss of 0.0277. A multiple dependent model with input as Dash, Litecoin, and Bitcoin’s price with LSTM unit gives the MSE loss of 0.0236. A GRU unit produces an MSE loss of 0.0201, and (LSTM + GRU) based model delivers an MSE loss of 0.0192. The loss observed by all prediction models for 1-day window size is displayed in FIGURE 20a, and FIGURE 20b shows the actual price and predicted price by our proposed fusion-based model (LSTM + GRU) in USD.
3) Scenario 3: 7-day price prediction

In a 7-day price prediction window size, 1000 data points passed as input for training and 157 data points for testing purpose to calculate loss. In the 7-day price prediction window size, the 7-day price passed as an input, and the price of the next day is estimated.

A simple model with input as Dash’s price with LSTM unit gives the MSE loss of 0.0188. A GRU unit produces an MSE loss of 0.0216, and (LSTM + GRU) based model delivers an MSE loss of 0.0267. A single dependent model with input as Dash and Litecoin’s price with LSTM unit gives the MSE loss of 0.0178. A GRU unit produces an MSE loss of 0.0178, and (LSTM + GRU) based model delivers an MSE loss of 0.0200. A multiple dependent model with input as Dash, Litecoin, and Bitcoin’s price with LSTM unit gives the MSE loss of 0.0196. A GRU unit produces an MSE loss of 0.0180, and (LSTM + GRU) based model delivers an MSE loss of 0.0188. The loss observed by all prediction models for 1-day window size is displayed in FIGURE 21a, and FIGURE 21b shows the actual price and predicted price by our proposed fusion-based model (LSTM + GRU) in USD.

4) Scenario 4: 30-day price prediction

In a 30-day price prediction window size, 1000 data points passed as input for training and 134 data points for testing purposes to calculate loss. In a 30-day price prediction window size, the 30-day price passed as an input, and the price of the next day is estimated.

A simple model with input as Dash’s price, with an LSTM unit, gives the MSE loss of 0.0205. A GRU unit produces an MSE loss of 0.0218, and (LSTM + GRU) based model delivers an MSE loss of 0.0271. A single dependent model with input as Dash and Litecoin’s price with LSTM unit gives the MSE loss of 0.0201. A GRU unit produces an MSE loss of 0.0220, and the (LSTM + GRU) based model delivers an MSE loss of 0.0260. A multiple dependent model with input as Dash, Litecoin, and Bitcoin’s price with LSTM unit gives the MSE loss of 0.0201. A GRU unit produces an MSE...
loss of 0.0201, and (LSTM + GRU) based model delivers an MSE loss of 0.0203. The loss observed by all prediction models for 1-day window size is displayed in FIGURE 22a, and FIGURE 22b shows the actual price and predicted price by our proposed fusion-based model (LSTM + GRU) in USD.

E. OBSERVATION OF RESULTS FROM GRAPHS

In this subsection, we are going to mark down some important observations captured from FIGURE 19,20,21, and 22. One of the main observations observed from all the graphs is that the proposed fusion-based hierarchical model and dual-dependency model performs better than the simple model for Dash price prediction. The hierarchical model shows the least variation in the loss amongst the dual-dependent and simple fusion-based algorithms for all model architecture and chosen window sizes.

For a 1-day window size, it was observed from FIGURE 19b that the dual-dependency model performs better than the other two models. For a 3-days window size, the GRU model for dual-dependency outperforms by having more than 40% less loss compared to the simple Dash prediction model where neither hierarchical model nor dual-dependent model performs better than the simple Dash prediction model for the fusion-based model of (LSTM + GRU). For a 7-days window size, the GRU model gives the lowest MSE loss for dual-dependency approach that is 0.0178. Performance increment of 30% is observed for (LSTM + GRU) by loss values, which is the highest increment amongst rest models for the hierarchical model. FIGURE 21b provides the evidence for the above statement. For the 1-days window size, the GRU model gives the lowest MSE loss for the hierarchical approach, which is 0.0201. The performance increment of 25% is observed for (LSTM + GRU) by loss values, which is the highest increment amongst rest models for the hierarchical model. FIGURE 22b provides the evidence for the above statement.

From the above study, we can conclude that the fusion-based forecasting model in the context of the interdependency of cryptocurrency gives better results compared to the simple forecasting model. We have proposed the (LSTM + GRU) architecture based on the fusion of the cryptocurrency prices. To justify this, a particular model is compared with the simple LSTM and simple GRU fusion-based model. From the figures, it’s evident that most of the time GRU and (LSTM + GRU) give better performance compared to the LSTM model. The observation leads us to confusion about the performance of models as we find the GRU and (LSTM + GRU) both perform well for the hierarchical models. To provide an evidence that the (LSTM + GRU) performs better than the GRU model, we perform the prediction based on the hierarchical model for GRU and (LSTM + GRU) for 1, 3, 5, 7, 10, 15, 20, 25, 30, 35, and 40 days window sizes. We found that (LSTM + GRU) is more promising than the GRU model for forecasting the price of cryptocurrency Dash based on previous days price of Dash, Litecoin, and Bitcoin. FIGURE 23a shows the comparison of training performance while FIGURE 23b shows the comparison of validation performance.

X. THREAT TO VALIDITY

This section discusses threats to the validity for recognizing it and counter them in a research design for a robust study.

- Scholarly and analysis articles: The articles and analyses done by financial institutes played a big role to carry out the survey done in this paper. Somehow, if the surveyed articles are false or wrong, it will pose a threat to the survey. To tackle it, we have considered scholarly articles from reputed peer-reviewed journals, and analytical articles were derived from the reputed financial institutes. Moreover, the filtration of articles was done based on the number of citations, quality of the journal, and several views on it.
- Dataset validation: For the proposed model, the dataset play an important. A slight change in the price history
will affect the performance at a drastic rate. The price history used were compared with other websites and then it’s been considered to do a performance evaluation of our proposed work.

- Architecture validation: Proposed model uses a sophisticated algorithm. If the algorithms’ implementation was wrong, the whole performance evaluation became useless. So, the standard and open-source libraries were used while setting up the whole environment to calculate the performance metrics of the proposed framework.

XI. DISCUSSION ON OPEN ISSUES AND FUTURE ASPECTS

This section discusses various future research challenges and open issues in cryptocurrency price prediction. Some of the major issues are discussed below.

A. ARCHITECTURAL ADVANCEMENT

The proposed model comprises of simple LSTM and GRU layers, which significantly improves the efficacy of cryptocurrency price prediction. But, there is a scope to further improve the prediction accuracy and robustness using ensemble learning, bidirectional layers, or the present trending transformers-based models. Research community can also go for Federated Learning as after few years the amount of data will be vast and it will be hard to train model on one single machine.

B. STOCHASTIC NATURE OF CRYPTOocurrency PRICES

The prices of cryptocurrency are stochastic. As, most of the times it follows random walk hypothesis, which is independent of time. External changes affect the price significantly that make the forecasting highly complex and tedious. This can be resolved to some extent by considering the fusion of multiple features by creating a robust forecasting model. In the proposed work we have used price of other but similar cryptocurrency for price prediction, further this can be fused with market sentiment, costs of mining hardwares, volume of cryptos in the market, current active users, worldwide governents policy and many more. The model might become complex and harder to train, but if all these features are fused with proper mapping and weightage, researchers can create a robust price prediction model.

C. LOW UTILITY OF FORECASTING MODEL

Based on the literature, the prediction models predicts the price of one cryptocurrency only. But, in fusion-based model, there can be multiple inputs, which comprises of diverse cryptocurrency prices. Thus, this can improve the models efficacy, which predicts the price of multiple cryptocurrencies at the same quantum of time.

The paper we have surveyed in Table 1 are kind of one-to-one price prediction model, i.e. the price history of Dash is used to predict current price of Dash. The model we have proposed is many-to-one, the price history of Litecoin, Bitcoin and Dash is used for price prediction of Dash. Now, we can also explore the path where many-to-many model can be implemented, to put it another way the price history of Litecoin, Bitcoin and Dash will predict the upcoming prices of all three cryptocurrencies based on each others history, through this we have completely cut off training time of 3 different models to one model and also the weights of models are efficiently used for prediction of different cryptocurrency.

D. ALGORITHMIC ADVANCEMENT

The proposed fusion-based model considered the interdependency among the cryptocurrencies and further improved the accuracy by considering various other factors (except the past prices) like market sentiments and opinion mining to improve the robustness and accuracy of the prediction model. The accuracy can further be improved by incorporating reinforcement learning that train the model in real-time by considering the environmental factors.

E. LEGAL ASPECTS

Although, the cryptocurrencies becomes a popular choice for investors nowadays, but there is no regulatory body that controls its circulation. In recent times, many countries
legally accept cryptocurrency as a medium of transaction with numerous rules and laws. Few countries also started to ban cryptocurrency because of their belief that the anonymous nature of cryptocurrency is a boon for organizations that are involved in illegal businesses and terrorist activities. These factors affect public perception, which indirectly affects the prices of cryptocurrency.

**XII. CONCLUSION**

Price prediction of any cryptocurrency is quite challenging and popular among researchers due to its volatile nature and its dependency among various internal and external factors, as discussed in Section I. For price prediction, time-series-based models are used because of their dependency on past-day prices. Researchers have explored various machine learning-based time-series models like ARIMA, SARIMA, ARCH, ARIMAX, and other regression-based models for cryptocurrency price prediction. Since the introduction of deep learning-based algorithms in the time-series prediction field, researchers have also started to incorporate their functionality in this field of research. Several variants of deep learning-based algorithms like MLP, LSTM, GRU, CNN, and hybrid or ensemble models had been used by the researchers for cryptocurrency price prediction as discussed in Section II. In consideration of price prediction, statistical analysis on the price history of cryptocurrency is also important with the risks related to their prices as discussed in Section III. The study of companies, startups, organizations, and researchers related to the cryptocurrency market is also to be considered as they impact cryptocurrency prices as discussed in Section IV.

From the literature, we found that the researchers have not considered the interdependency among the crypto coins. But, we have considered this as our main factor for the proposed model. To find the appropriate factors for the proposed model, we have analyzed various existing factors and dependencies among the cryptocurrencies. In this paper, we have proposed an LSTM and GRU-based fusion model with a detailed comparison in terms of performance with existing models. For performance evaluation, we have considered the price prediction of Dash with the inclusion of daily price data from Dash, Litecoin, and Bitcoin. Based on the performance for predicting Dash’s price, our proposed model works better compared with the model trained using only Dash price and other model trained using data from Dash and Litecoin. Hence, to increase the performance of the forecasting model, hierarchical dependency among cryptocurrencies was found useful. The concept of hierarchical dependency has termed the concept of fusion.

In the future, we will further improve the accuracy of the cryptocurrency price prediction using federated and distributed learning.

**REFERENCES**

[1] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2009.

[2] V. A. Shaik, P. Malik, R. Singh, A. Gehlot, and S. Tanwar, “Adoption of blockchain technology in various realms: Opportunities and challenges,” Security and Privacy, vol. 1, p. 1198, 2019.

[3] A. HANKIN, “Bitcoin pizza day: Celebrating the $80 million pizza order.” https://www.vestopedia.com/news/bitcoin-pizza-day-celebrating-20-million-pizza-order/. Accessed: 2019.

[4] M. Trajevski, “What is gpu mining?” https://coindesk.com/article/what-is-gpu-mining. Accessed: 2019.

[5] L. Cong, Z. He, and J. Li, “Decentralized mining in centralized pools,” SRRN Electronic Journal, 01 2018.

[6] A. Tule, “How do mining pools work and how to choose a pool to join?” https://captainaltcoin.com/what-is-pool-mining/. Accessed: 2019.

[7] G. Giudici, A. Milne, and D. Vinogradov, “Cryptocurrencies: market analysis and perspectives,” Journal of Industrial and Business Economics, vol. 47, pp. 1–18, Mar 2020.

[8] Coindesk, “Bitcoin + ether: An investor’s perspective.” https://www.coindesk.com/research/reports/bitcoin-ether-an-investors-perspective. Accessed: 2021.

[9] Coinmarketcap. https://coinmarketcap.com/. Accessed: 2021.

[10] M. Watch, “Cryptocurrency market grow with a high cagr–global industry analysis, key manufacturers, trends, size, and forecasts during 2025.” https://www.marketresearch.com/press-release/cryptocurrency-market-grow-with-a-high-cagr--global-industry-analysis-key-manufacturers-trends-size-and-forecasts-during-2025-2021-03-01?siteid=bigcharts&dist=bigcharts&tesla=y#:~:text=According%20to%20the%20research%20report,11.9%25%2020 during%202019%2020%202024.

[11] K. Sheth, K. Patel, H. Shah, S. Tanwar, R. Gupta, and N. Kumar, “A taxonomy of ai techniques for 6g communication networks,” Computer Communications, vol. 161, pp. 279–303, 2020.

[12] Y. Sovbetov, “Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litecoin, and monero,” Journal of Economics and Financial Analysis, vol. 2, pp. 1–27, 02 2018.

[13] O. Angela and Y. Sun, “Factors affecting cryptocurrency: Evidence from etherum,” in 2020 International Conference on Information Management and Technology (ICIMTech), pp. 318–323, 2020.

[14] A. G. E. T. Fu, “The efficient market hypothesis: Review of specialized literature and empirical research,” Procedia Economics and Finance, vol. 32, pp. 442–449, 2015. Emerging Markets Queries in Finance and Business 2014, EMQFB 2014, 24-25 October 2014, Bucharest, Romania.

[15] A. Thakkar and K. Chaudhari, “Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions,” Information Fusion, vol. 65, pp. 95–107, 2021.

[16] V. L. Tran and T. Lavril, “Efficiency in the markets of cryptocurrencies,” Finance Research Letters, vol. 35, p. 101382, 2020.

[17] C. Gurdgiev and D. O’Loughlin, “Herding and anchoring in cryptocurrency markets: Investor reaction to fear and uncertainty,” Journal of Behavioral and Experimental Finance, vol. 25, p. 100271, 2018.

[18] S. Linkman, “Systematic literature reviews in software engineering – a systematic literature review,” Journal of Systems and Software, vol. 80, no. 4, pp. 571 – 583, 2007. Software Performance. 20during%202019%20%2D%202024.

[19] S. Khuntia and J. Pattanayak, “Adaptive market hypothesis and evolving predictability of bitcoin,” Economics Letters, vol. 167, pp. 26–28, 2018.

[20] M. M. Patel, S. Tanwar, R. Gupta, and N. Kumar, “A deep learning-based cryptocurrency price prediction scheme for financial institutions,” Journal of Information Security and Applications, vol. 55, p. 105832, 2020.

[21] B. Kitchenham, O. P. Breerton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, “Systematic literature reviews in software engineering – a systematic literature review,” Information and Software Technology, vol. 51, no. 1, pp. 7 – 15, 2009. Special Section - Most Cited Articles in 2002 and Regular Research Papers.

[22] B. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering,” 2007.

[23] P. Breerton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, “Lessons from applying the systematic literature review process within the software engineering domain,” Journal of Systems and Software, vol. 80, no. 4, pp. 571 – 583, 2007. Software Performance.

[24] F. Fan, C. Vente, M. Basios, H. Kong, L. Kanthan, L. Li, D. Martinez-Regoband, and F. Wu, “Cryptocurrency trading: A comprehensive survey.” https://arxiv.org/pdf/2003.11352.pdf/. Accessed: 2020.

[25] J. Abraham, D. Higdon, J. Nelson, and J. Ibarra, “Cryptocurrency price prediction using tweet volumes and sentiment analysis.” 2018.

[26] A. Aggarwal, I. Gupta, N. Garg, and A. Goel, “Deep learning approach to determine the impact of soco economic factors on bitcoin price prediction,” in 2019 Twelfth International Conference on Contemporary Computing (IC3), pp. 1–5, 2019.
Fusion in Cryptocurrency Price Prediction

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3163023, IEEE Access
[125] F. A. Gers, J. Schmidhuber, and F. Cummins, “Learning to Forget: Continual Prediction with LSTM,” Neural Computation, vol. 12, pp. 2451–2471, 2000.

[126] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” 2014.

[127] J. Han, M. Kamber, and J. Pei, “3 - data preprocessing,” in Data Mining (Third Edition) (J. Han, M. Kamber, and J. Pei, eds.), The Morgan Kaufmann Series in Data Management Systems, pp. 83–124, Boston: Morgan Kaufmann, third edition ed., 2012.

[128] A. Botchkarev, “A new typology design of performance metrics to measure errors in machine learning regression algorithms,” Interdisciplinary Journal of Information, Knowledge, and Management, vol. 14, p. 045–076, 2019.

***