Forecasting Cryptocurrency Market Trends with Machine Learning and Deep Learning

Heba M. Fadhil 1*, and Noor Q. Makhool 2
1 Department of Information and Communication, Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad, Iraq.
2 Department of Information and Communication, Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad, Iraq.

Abstract. Cryptocurrency became an important participant on the financial market as it attracts large investments and interests. With this vibrant setting, the proposed cryptocurrency price prediction tool stands as a pivotal element providing direction to both enthusiasts and investors in a market that presents itself grounded on numerous complexities of digital currency. Employing feature selection enchantment and dynamic trio of ARIMA, LSTM, Linear Regression techniques the tool creates a mosaic for users to analyze data using artificial intelligence towards forecasts in real-time crypto universe. While users navigate the algorithmic labyrinth, they are offered a vast and glittering selection of high-quality cryptocurrencies to select. The ability of the tool in analyzing past data on historical prices combined with machine learning, orchestrate an appealing scene of predictions equipped with choices and information, users turn into the main characters in a financial discovery story conducted by the cryptocurrency system. The numerical results also support the effectiveness of the tool as highlighted by standout corresponding numbers such as lower RMSE value 150.96 for ETH and minimized normalized RMSE scaled down to under , which is . The quantitative successes underline the usefulness of this tool to give precise predictions and improve user interaction in an entertaining world of cryptocurrency investments.

1 Introduction

Cryptocurrencies have taken hold in the financial world as one of the most innovative assets that attract huge capital outflows. With the growth of use of digital monies, investors have tried to exploit the opportunities they offer for great profits. The appeal of cryptocurrencies as alternate asset lies in their uniqueness because of the fact that they are free from central control, internationality and higher transparency. On the other hand, this benefits carry an inbuilt vulnerability which can be quite problematic to the traders, as well as individual investors looking forward to reaping from cryptocurrencies markets. Therefore, prediction of cryptocurrencies correctly has transformed from a desired aim to vital necessity in order to make reasonable investments decisions[1],[2].

Volatility is both an advantage and disadvantage of cryptocurrency market. However, it presents chances of high profits within a short period since prices can shoot upwards. However, it has a flip side in that the volatility might entail sharp declines. Sound methods of predicting price motions are important when making wise investment decisions and reducing risks among investors and traders. On the other hand, this is what has made machine learning techniques highly useful resources in their own right. One aspect of artificial intelligence is machine learning that enables researchers, as well as traders, to draw substantial inferences
from multifarious sets of information. Machine learning models are able to pick out trends, study previous samples and forecast prices in many areas including cryptocurrencies. This study explores where finance, technology, and data science intersect, utilizing past cryptocurrency cost datasets along with sophisticated prediction models that estimate future cryptocurrency prices for pre-selected periods[3], [4]. Cryptocurrency operates in a continuous mode with sharp price changes within small periods of time. Continuous exchange, influenced by various aspects such as market sentiments, laws, inventions, globally evolving economy among others, make it for costs to fluctuate abruptly and erratically. These markets have dynamics that are not captured by conventional financial analysis models[5]. There are plenty of historical data and ready-made analytical models available for investors on traditional assets such as stock or bond. Some of these markets are relatively stable as compared to highly volatile cryptocurrency markets. Unlike their counterparts in the traditional financial markets, the market for cryptocurrencies is yet too immature and does not possess any significant depth such that it becomes quite difficult to make price predictions[6], [7]. Investors deal with uncertainties while trading cryptocurrencies and need reliable predictions for when to buy, sell or hold them. A long-term investor seeking to make profits from possible future growth of the cryptocurrency and a short-term trader targeting on intraday fluctuations of prices in the market needs to have prediction about prices. The learning machine models especially the deep learning and neural network based can analyze big datasets, discover complex trends that may have escaped the statistics. The mentioned models learn how past situations relate, depend, or trend to predict probabilities of future price changes by being trained on such old price data[8], [9].

This study focuses on two distinct machine learning models: combining linear regression, ARIMA and a HLSTM (Hybrid LSTM (Long, Short Term, Memory)) neural network. Linear regression acts as a baseline, presenting an easy-to-understand approach to price forecasting. However, it can be said that the hybrid LSTM combines a standard feedforward network with recurrent neural networks to depict time-dependent sequences from prices. To this purpose, researchers want to conduct tests with models and evaluate their efficacy in forecasting cryptocurrency values to determine if machine learning can be of assistance in this area. The research will be useful for both seasoned investors and novice traders. Additionally, it will contribute to the growing body of work on predictive modeling for digital assets.

2 Literature Survey

It presents an exhaustive review of available scholarly work on forecasts about virtual currency value and major outcomes attained from different research projects. This survey contributes to an improved perception of the approaches, models, and hurdles involved in forecasting crypto markets. In this research, the LSTM type of RNN is applied in the development and forecasting Bitcoin prices. In summary, the authors present bitcoin economy giving us an update for future prediction of bitcoin price using stocks market index, sentiments, block-chain, and coin market cap. It shows application of the LSTM networks to predict bitcoin prices for the next one month and the subsequent two months[10]. Sentiment analysis in conjunction with supervised machine learning used to forecast the price of Bitcoin in this paper. The authors examine the connection between Bitcoin prices fluctuations and the tone of tweets on Twitter and the posts of Reddit. LSTM, ARIMA and RNN are among the models they use in their analysis of price changes[11]. This research aims to accomplish this by applying LSTM and ARIMA models to the problem of Bitcoin price forecasting. The authors' results using LSTM are mixed, yielding an accuracy of 52% and an RMSE of 8%. Both the deep learning models and the GPU version are better than ARIMA model and CPU versions [12]. The aim of this study is to guess Bitcoin prices by looking at the many things that change its worth. The study watches Bitcoin's cost every day, and uses learning methods like regression and decision trees to guess what it's worth will be later on[13].

The writers suggest common tech tools used by traders and features made with a Denoising autoencoder can predict changes in the Bitcoin price. For prediction, they use both an Attentive LSTM network and an Embedding Network. ALEN works better [14]. This study compares and looks at different machine learning ways to guess Bitcoin prices. The study [4] says that the NARX model is best at guessing Bitcoin prices. This study uses things like stock market size, how much is traded and where it's sold to introduce a new way of predicting costs of online money. To predict how much digital money will be worth in the future, writers
use a type of network called LSTM and long-lasting shapes [15]. In this study, four different machine learning models are tested and compared to find out which one works best at guessing Bitcoin's future rates of return and trends in prices. They use Twitter and COVID-19 information in their models. They find that using Twitter data can improve how well the model works [16].

This study suggests a deep learning method to make better predictions about future Bitcoin prices. It helps improve the standard machine model's accuracy [17]. The authors suggest a mix model to predict the worth of Litecoin and Zcash. This uses both Gated Recurrent Units (GRU) as well Long Short-Term Memory (LSTM). This model seems more precise than others [18]. This study looks at the connections between Bitcoin and other types of crypto money. It also studies public view, trade value, older prices data and how they impact each other. This research hopes to make price prediction more exact. We will use methods like sentiment analysis and machine learning, said in simple words [19]. This paper talks about a method called LASSO for guessing Bitcoin price and compares it to other methods. The goal of the study is to figure out how Bitcoin's daily market patterns work and accurately guess where its price will change each day [20].

These books set the stage for more research in a fast-growing topic about predicting prices of cryptocurrencies. They give ideas on ways, where they get data from and models that are currently being used to do this prediction work.

3 Methodology

The proposed is generated to guess the price of coins is made with really strong methods like Linear Regression and ARIMA models. It also gets data from CryptoCompare API, prepares it well for machine learning work. The machine learning model learns from past price info in a planned way, letting it predict future cryptocurrency costs correctly. This changing process is the base of our advanced tool's guessing abilities; it relies on strong math rules and smart code designs.
3.1 Data Source

In order to get accurate predictions, so the plan is to get reliable and complete information on old prices. The proposed used the CryptoCompare API which many people use and trust to study crypto markets. The CryptoCompare API is the go-to place for accurate data on the current state of the market and future predictions for cryptocurrency values. It has an incredible amount of information about cryptocurrencies. Some of the data available through this API comprises daily price levels, trading volumes, market cap, and a list of technical indicators. The researcher and traders have a lot at their discretion because these figures cover a majority of cryptocurrencies.

Several issues shaped the decision on using the CryptoCompare API. To begin with, through the API, it becomes easy to get cryptocurrency data without physically retrieving and collating the information from different places. This simplifies the data gathering process and makes sure that the collected information is good-quality and updated. The second is that CryptoCompare is renowned for providing accurate and reliable information which is essential in performing credible and serious research. Lastly, the depth of the available data using the API allows for an in-depth analysis of multiple cryptocurrencies, offering insights into the broader market trends.
3.2 Data Preprocessing

Data retrieved from CryptoCompare APIs is important but needs cleaning up for training with machine models. This pre-processing stage entails some of the most important operations of transforming raw data into useful structures for our models.

Outlier Removal: Indeed, cryptocurrency markets are characterized by short-term price discrepancies and high volatilities. Such unusual instances could also bring noise in the data set, which may bias our model’s results. The method used to remove strange data points that were far from average or middle observations. This method checked that the information was real, which made our guessing models stronger.

Handling Missing Values: Incomplete data is something that often happens in sets of information, particularly those taken from places like CryptoCompare API. Sometimes, machine learning can be stopped by not having enough data. To make sure that any parts missing, data imputations to fill in the gaps was used.

Standardization: It is difficult to compare and analyze cryptocurrency data, which is usually displayed in different forms and units. To normalized the data by making it uniform. Scaling all dimensions to a similar range of normal distributions that often have a zero mean and unit deviation is called standardization. The process makes sure that all parts are measured in the same way, so they can be linked and compared to a machine learning model used.

Cleaning data is very important for making the set of information right and good to use with machines that learn. We tried to create a sample that includes everyone and can be changed. The proposed handled extreme cases, unknown values, and made things equal. Also used this data to improve our model and make better guesses about the value of cryptocurrencies.

4 Prediction Model Selection

Picking a good guessing model for predicting prices and cryptocurrencies is one of the most important choices. This section, outlines the model selection process and discuss the two models were employed: Both, the basic Linear Regression and ARIMA model as well as the more advanced Hybrid LSTM (Long Short-Term Memory) model. So, the researchers picked these models to study because they wanted to see how strong and useful they are in hard-to-predict crypto markets.

4.1 Linear Regression

Linear Regression helps create simple mathematical connections between two or more things. It's a key part of making statistics models work! For guessing the price of cryptocurrency, it thinks that how much money a person spends to buy or sell depends on things like market overall value and amount people trade. It also considers news around this topic. The simple math behind Linear Regression is easy to understand, so people without technical skills can get it. Once trained, Linear Regression models can usefully guess prices for more data points. Because it works fast, Linear Regression can be used in real-time high speed trading.

4.2 ARIMA

ARIMA, which stands for Autoregressive Integrated Moving Average, looks at how much cryptocurrency costs change over time. The ARIMA model carefully looks at old data to find patterns, seasonal changes and outliers. ARIMA is good at guessing changes in the market quickly. This helps traders find chances to make money fast. Bitcoin prices can be well-made using ARIMA because it's good at handling time data patterns. When there are many ups and downs, ARIMA can make its predictions better by including seasonal patterns in its models.

4.3 Hybrid LSTM

Hybrid LSTM fixes the problems of its earlier versions by mixing deep learning with regular machine learning. Figure shows how Hybrid LSTM catches both short-term and long-term connections by using Long Short-Term Memory (LSTM) networks. This lets it get detailed patterns from large amounts of information. (2) Hybrid LSTM can take on complex relationships between things, helping it to better show how
cryptocurrency prices move. Hybrid LSTM can quickly handle lots of data and find hidden patterns that easy methods might miss. Hybrid LSTM's ability to get long-term connections makes it good for guessing price changes over a lot of time.

Fig. 2 : Hybrid LSTM Layers.

5 Experiments Environment

The study focuses on the testing phase with three types, Linear Regression, ARIMA and Hybrid LSTM. This part explains how were organized in experiment, what type of feature used throughout the activity and what was taught. It also describes measurements taken along with results from studies as well.
Making new features is very important in any machine learning project because the input values greatly affect how good your model can predict things. Researchers picked out some key points that we think will make sure the prediction of cryptocurrency prices is trustworthy. These features included:

- **Historical Prices**: Old prices, one of the main parts that any forecasting price model needs to have. The model looks at past actions of the money to find any patterns and guess what will happen in future. This gets helped by these signs or clues.

- **Trading Volumes**: The amount of a cryptocurrency bought or sold tells market feeling and how easy it is to trade. It shows how much interest and action there is for a specific type of virtual money. It also helps the models to understand these levels better.

- **Technical Indicators**: They also used tools like moving averages and RSI to better understand how the market was acting. The moving average helps find a trend. The RSI tells about times when things are too cheap or expensive, which could be signs of upcoming changes in price movement coming soon.

These parts were chosen very carefully to show a view of the crypto money markets. The proposed used them for both Linear Regression and a mixed LSTM model. This let us test their skills in an equal testing place.

### 6 Performance Measurements

The model is evaluated according to several certain metrics called Mean absolute error (MAE), mean squared error (MSE) root, a mean square error (RMSE).

Mean Absolute Error (MAE): MAE is the measure of average difference between two continuous variables X and Y, that measures how far each (xi, yi) data point deviate from the line y = x. Equation represents the MAE Eq. (1) [18], in which n denotes the number of data points and eᵢ stands for model errors.

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |e_i|
\]  

(1)
Mean Squared Error (MSE): MSE means the sum of squared values noted for actual observations and estimate. Squaring errors before aggregation guarantees positive values, and lower MSE rates reflect better model performance. Eq. (2) provides the MSE computation showing how $X_i$ stands for original sample data and $Y_i$ represents processed data [19].

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (X_i - Y_i)^2 \frac{1}{2}$$

(2)

Root Mean Squared Error (RMSE): RMSE, that is obtained by taking the square root of MSE can be used as a measure of typical size errors in accordance to units [20]. It is derived in Eq. (3).

$$RMSE = \sqrt{MSE}$$

(3)

These measures are essential to evaluate how well the model predicts future electricity demand. Lower values of MAE, MSE and RMSE mean the better model performance since projected value approaches to real value.

7 Results and Discussion

Guessing the worth of a crypto coin later on needs using many types of math and ways. To guess the cost of some cryptocurrency over different dates, we used a mix of Autoregressive Integrated Moving Average (ARIMA) and Linear Regression models in this study. The models were taught and checked using information gathered from old price lists. Guessing costs in the cryptocurrency market is hard, but it's important because this field can change suddenly. In this study, we use ML models to guess how the cost of one crypto will change over time. For this job, the popular ARIMA and Linear Regression methods were used.

To ensure the reliability of our results, we divided our dataset into two distinct sets: training data and test data. In this work, used 80% of data to learn and kept the rest 20% for testing. They started the study by getting old price records and setting up important libraries or tools. The models were taught with a part of the data and then checked over future times. The predictions were kept and shown to see how accurate each model was, plus if it was constant.

Fig 4. ARIMA model predication for seven days.
The produced data exhibit both projected and real prices, providing a thorough evaluation of the model's performances. The ARIMA model, for instance, accurately projected prices on individual days that were quite near to the actual values. However, Linear Regression showed stable forecasts, accurately reflecting the overall trend. The ARIMA model showed impressive precision, especially on individual dates like November 15th, 2023. The power of the algorithm comes in its ability to predict time series, as it is able to capture temporal dependencies in the data. However, it could be vulnerable to external forces and rapid shifts in the market as shown in Figure 4 to 6.

**Fig. 5.** ARIMA model prediction for 24 hours.

**Fig. 6.** ARIMA model prediction for 12 months.
Although not as time-aware as ARIMA, linear regression showed steady and reliable predictions over the whole time frame as illustrated in Figure 7 through 9. Its ease of use and resistance to short-term changes make it ideal for a wide range of applications. When looking at the two models side-by-side, it is clear that ARIMA is superior when it comes to capturing short-term variations, making it appropriate for traders concerned with immediate market moves. Linear Regression, on the other hand, can help long-term investors grasp the big picture when it comes to market movements. The results are much easier to comprehend thanks to the visualizations. The plotted graphs show how closely the forecasted prices match the observed ones. In addition, for each prediction, an annotation is added to show which algorithm was utilized.

![Cryptocurrency Price Predictions](image)

**Fig. 7.** Linear Regression model prediction for seven days.
In overall, when it comes to predicting the value of cryptocurrencies, both the ARIMA and the Linear Regression models have their advantages. Which one is better depends on the user's needs, whether for day trading or long-term investing. This analysis yields useful information for market players, giving them the tools they need to make educated choices. It further shows that the hybrid LSTM model is highly predictive accurate compared to other models under review as seen in Figure 10 and 11. Having combined the ability of the LSTM architecture to recognize long term connections and that of a simple feedforward
Neural network for capturing short term changes, the hybrid LSTM gave a clear edge to it as far as making sense out of complex market dynamics relating to cryptocurrencies is concerned.

**Fig. 10.** Hybrid LSTM model prediction.

**Fig. 11.** Hybrid LSTM model val and loss.

The performance evaluation metrics reveal essential details about the operation of cryptocurrency prediction models as shown in Figure 12. First of all, under the measurer in terms of individual cryptocurrencies, MEA, MSE, RMSE and NRMSE are considered for Bitcoin (BTC), Ethereum (ETH), Monero (XMR) and XRP. The MSA of BTC is 1685. The MSE, 460.68 is a measure of the total variance in prediction errors. A noticeable reduction in the RMSE value of 200.67 is seen, pointing out toward the square root of mean squared residuals between predicted and actual values. The Normalized RMSE, equal to 0.071, is the normalization measure and highlights the precision of predictions regarding cryptocurrency values’ range. ETH has a MEA of 131.34, which is lower than the given value and signifies enhanced accuracy. The MSE and RMSE are 28731.66, respectively implying minimization of the errors as well as its variability. The Normalized RMSE of 0.050 indicates how excellent Ethereum is at prediction in terms of a normalized measure. Monero (XMR) has a MEA of 12.46, indicating that its predictions are relatively accurate ones. The MSE and RMSE are 290.16, correspondingly, as well as the value of 15.03 points to a lower strength of
mistakes and enhanced accuracy level due to their nature. The NRMSE of 0.059 further strengthens the model’s validity for forecasting Monero prices. The prediction accuracy of the smallest MEA —0.107 can be observed for Ripple (XRP). The MSE and RMSE are 0.0212 and 0. The Normalized RMSE, with a value of 0.107, highlights the impressive precision that is further emphasizing by considering it to be in context towards normalization about performance.

Going to the measures of dropout rates and epochs, changes in these parameters influence predictive accuracy. The 0.2 dropout rate has the best performance because it achieves a reduced RMSE of 150.96 and normalized RMSE score, a minimal number of epochs such as 25 gives best results with an RMSE value of 150.42.

Fig. 12. The performance evaluation of MEA MSE, RMSE and NRMSE.

8 Conclusion

In summary, this research examined hybrid LSTM machine learning models that can be used for the prediction of cryptocurrency prices especially Bitcoin. These models, using historical price data and technical indicators were able to predict market movements. However, it is important to note that these models have their constraints. The study used historical information from the CryptoCompare API, using machine learning to predict cryptocurrency prices. It is the hybrid LSTM model that revealed a unique capacity to uncover temporal patterns in cryptocurrency prices demonstrating impressive outcomes. This demonstrates the high predictive capability of LSTM model in unstable cryptocurrency market.

Cryptocurrency being an attractive alternative asset class provides huge profits despite the price factor complexities arising out of its volatile nature. A valuable contribution to the extant literature is made by this study, which demonstrates that a hybrid LSTM algorithm efficiently addresses such challenges. But it is essential to underline that the vulnerabilities of this study are factors such as external influences and market disruptions, in addition to sentiment analysis, real-time news and dynamic markets were omitted leading uncertainty. Based on these results, the study can be considered a tentative investigation of machine learning-driven Bitcoin price forecasting. The emergence of complex and efficient models is crucial for the development of cryptocurrency, as this market evolves. Real-time data, sentiment analysis and other
The numerical outcomes attested the fruitfulness of hybrid LSTM model with impressive results like 150.96 RMSE reduction for ETH and reduced normalized one-RMSE of only 0.05 remained until modified dropout rates and epochs number are considered. The outcomes of these studies offer numerical evidence in favor of the suggested models’ ability to predict cryptocurrency prices.

**Reference**


