Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Machine Learning-Based Cryptocurrency Prediction: Enhancing Market Forecasting with Advanced Predictive Models

Md Shahidul Islam[1], Monjira Bashir[2], Siddikur Rahman[3], Md Abdullah Al Montaser[4], Joy Chakra Bortty[5], Araf Nishan[6], Muhammad Rafiuddin Haque[7]

Abstract

The cryptocurrency market, with its record volatility and breakneck speed, is a revolutionary phenomenon that is reshaping the entire world's landscape. Unlike regular markets, cryptocurrencies undergo unprecedented volatility caused by a complex interaction of factors ranging from speculative trading to updates in regulations, technological innovations, and macroeconomic trends. The central objective of this research was to develop and evaluate machine learning-driven models of cryptocurrency price trend forecasting. The focus of this research project revolved around prominent cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other prominent altcoins, within the United States. The dataset employed in this analysis comprises vast historical price data, trading volumes, and key market indicators of major cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other major altcoins. Historical price data is presented in terms of daily, hourly, and minute-level opening, closing, high, and low prices, providing detailed insights into temporal price behavior. Trading volumes, which reflect the intensity of trading action, are also provided to represent liquidity and investor participation behavior. The dataset also includes various market indicators, i.e., moving averages, relative strength index (RSI), Bollinger Bands, and other technical indicators, which play a pivotal role in establishing market patterns and momentum. Three models are chosen in this study: Logistic Regression, Random Forest Classifier, and XG Boost Classifier. For classification models, accuracy, precision, recall, and F1-score metrics are employed to evaluate the performance of the models in terms of predicting the directions of the markets (e.g., upward or downward directions). With the highest accuracy, Logistic Regression was the best-performing of the models tested, showing its relative superiority. The integration of AI forecasts into cryptocurrency trading has the potential to revolutionize the United States financial markets by providing traders and institutional investors with advanced tools to make decisions. The use of AI tools in cryptocurrency trading also has significant implications for United States regulation compliance. The integration of machine learning tools within cryptocurrency trading platforms is a significant step towards unleashing the true potential of AI in the financial markets. The field of AI-based cryptocurrency forecasting offers numerous areas of future research with the potential to break through present limitations and unlock new paths of market analysis. One of those areas is the use of deep learning models, i.e., Long Short-Term Memory (LSTM) networks, for time-series cryptocurrency forecasting.

Keywords: *Cryptocurrency, Machine Learning, Market Forecasting, Predictive Analytics, Financial Modeling.*

Introduction

According to Islam et al. (2024b), the cryptocurrency space has grown exponentially since Bitcoin's release in 2009, developing from a niche technological sandbox to a multitrillion-dollar asset class. Cryptocurrencies have upended traditional financial systems by offering decentralized, borderless, and transparent alternatives to fiat currencies. This invention has attracted a wide range of participants, from retail traders to institutional investment houses and governments, all seeking to unlock the potential of blockchain technology and digital assets. Khan et al. (2024) argued that while the growth of the market has been unprecedented, it has been accompanied by extreme volatility, with the values of cryptocurrencies experiencing double-digit percentage swings over short time frames.

In retrospect, Das et al. (2025), Bitcoin, the largest traded cryptocurrency, has seen its value fluctuate between a few cents in its early days to an all-time high of nearly $70,000 in 2021, followed by sharp

[1] MBA- Business analytics, International American University, Email: mdshahidulislam@my.unt.edu, (Corresponding Author)
[2] Master of Business Administration, International American University.
[3] MBA in Management Information Systems, International American University. Los Angeles, California, USA.
[4] MS in Business Analytics, University of North Texas.
[5] Department of Computer Science, Westcliff University, Irvine, California, USA
[6] MBA in Business Analytics, International American University. Los Angeles, California, USA
[7] MS in Business Analytics, Mercy University, New York, USA

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

corrections. Such volatility is brought on by a combination of factors ranging from sentiment to announcements of regulation, technological updates, to macroeconomic events (Byun & Shahbazi, 2021). Unlike traditional assets such as stocks or bonds, cryptocurrencies have no intrinsic value and are highly susceptible to speculative trading, thus rendering their movements unpredictable. Such unpredictability necessitates robust forecasting tools that can help crypto economy participants navigate the dynamics of the crypto economy (Sizan et al., 2023).

Problem Statement

As per Chouksey et al. (2023), traditional financial models, such as autoregressive integrated moving averages (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH), have been applied to predict asset prices in conventional markets over the years. They rely on linear assumptions and historical data to forecast future behavior. Their effectiveness in cryptocurrency markets is, however, undermined by the nature of digital assets. Cryptocurrencies exchange in a low-regulated, decentralized environment, and therefore, the markets' dynamics are influenced by a wide range of exogenous and non-linear factors (Ali et al., 2024). For example, social media sentiment, whale movements (major trades by large-scale traders), and technological updates such as hard forks or network upgrades have strong influences on prices. Such influences cannot be modeled by conventional models, and hence, forecasts and investment decisions suffer. Moreover, the high volumes and velocity of data generated by cryptocurrency markets further hinder conventional analytical tools. Under this situation, complex predictive models must be constructed that represent the dynamism and nature of crypto markets so that decision-making is done in a timely fashion (Ganesh, 2024).

Research Objective

The central objective of this research is to develop and evaluate machine learning-driven models of cryptocurrency price trend forecasting. With the power of ML algorithms, we aim to overcome the limitations of traditional financial models and provide more accurate forecasts. Specifically, we focus on supervised learning techniques, such as regression models, decision trees, and neural networks, that perform exceptionally well with large data volumes and detect non-linear patterns. We employ a method of training the models on historical price data, technical indicators, and external factors such as social media sentiment and macroeconomic indicators. We also explore the use of ensemble methods, which combine multiple models to provide stronger predictive power. Beyond building predictive models, we also test their performance using measures of mean absolute error (MAE), root mean squared error (RMSE), and R-squared values. We also provide insights into how the models could be applied in practice by investors, traders, and financial analysts, and how they could be used to better manage risk and refine investment strategies.

Scope and Relevance

The focus of this research project revolves around prominent cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other prominent altcoins, within the United States. The choice of cryptocurrencies is motivated by market capitalization, liquidity, and influence on the general crypto environment. With a focus on the United States, we aim to derive insights that are of high relevance to American institutions and individuals, who exert significant influence on global crypto trends. The use of predictive modeling here has far-reaching implications for different stakeholders. For the retail investor, accurate predictions of prices can aid in risk reduction and return maximization in an extremely volatile setting. For institutional investors, i.e., hedge funds and asset managers, predictive models can be employed to make allocation and hedging decisions. Financial analysts can employ the models to make evidence-based recommendations to clients, and the models can be employed by regulators to monitor market stability and potential risks. In general, this study is part of the growing literature on cryptocurrency market dynamics and demonstrates the revolutionary promise of machine learning in forecasting.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Literature Review

Cryptocurrency Market Trends and Challenges

Gupta & Sethi (2024) reported that the cryptocurrency space is a dynamic and rapidly evolving environment, defined by its volatility and sensitivity to a wide range of influences. One of the major drivers of the volatility of this marketplace is investor sentiment, which is shaped by news events, social media, and public perception. For instance, positive news, i.e., the adoption of Bitcoin by large businesses or favorable regulatory announcements, creates sharp increases in price, and unfavorable news, i.e., security attacks or regulatory measures, creates sharp declines. Beyond sentiment, macroeconomic influences also have a significant impact on cryptocurrency prices.

According to Islam et al. (2025c), Macro influences, i.e., inflation rates, movements in interest rates, and political events influence investor behavior, sending capital into or out of digital assets. For instance, when the economy is uncertain, cryptocurrencies like Bitcoin come to be regarded as "digital gold" and appeal to those seeking an inflation hedge (Jui et al., 2023). However, this is not always the case since cryptocurrencies also tend to correlate with traditional markets when markets are in dislocation. Another significant driver of crypto volatility is the impact of institutional investment and speculative trading. The entrance of institutional players, i.e., hedge funds, asset managers, and public companies, has brought both solidity and sophistication to the marketplace (Jagannath et al., 2021).

Poudel et al. (2023) found that while institutional investment has the potential to increase liquidity and prevent price manipulation, it also has the potential to increase volatility through large trades and algorithmic trading. On the other hand, speculative trading by retail traders, motivated by fear of missing out (FOMO) or herd behavior, increases price swings and creates challenges to reliable forecasting. These multidimensional effects indicate the complexity of cryptocurrency markets and the need for advanced analytical software to manage their volatility.

Traditional Financial Models vs. Machine Learning for Forecasting

Traditional models, i.e., autoregressive integrated moving average (ARIMA) and generalized autoregressive conditional heteroskedasticity (GARCH), have been traditionally used to predict prices in conventional financial markets (Islam et al. 2025c). The models rely on linear assumptions and historical data to predict future behavior and thus are suitable for assets with similarly stable dynamics. Their application to cryptocurrency markets is, however, hindered by the nature of digital assets. Cryptocurrencies are traded in a low-regulated environment, and thus, the dynamics of the assets rely on a multitude of exogenous and non-linear factors (Rana et al., 2025). For example, social media sentiment, whale movements (major trades by large-scale traders), and technological events such as hard forks or network updates significantly impact prices. Such factors cannot be modeled by conventional models, and thus, the models do not provide correct forecasts and optimal investment strategies (Ray et al., 2025).

Contrarily, Rahman et al. (2024) held that machine learning (ML) offers a flexible and robust approach to forecasting cryptocurrency prices. ML models, i.e., decision trees, support vector machines (SVM), and neural networks possess the capability to identify complex patterns and non-linear relationships in large data. Such models can incorporate multiple data sources, i.e., historical data of the asset's price, technical indicators, and exogenous factors such as news sentiment and macroeconomic indicators, to provide improved forecasts. Moreover, Sumsuzoha et al. (2024) added that ML models learn to adapt to varying patterns in the market, thus being suitable for the dynamic and volatile nature of cryptocurrency markets. With the aid of ML, researchers and practitioners are capable of overcoming the limitations of conventional models and constructing stronger predictive tools to predict crypto prices.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Applications of Machine Learning in Financial Prediction

According to Ali et al. (2024), Machine learning has been a game-changer when it comes to financial forecasting, offering new channels for exploring and forecasting asset prices. For cryptocurrency markets, supervised learning techniques have been widely used to predict prices. Such techniques involve training an ML model on labeled data, where input features (i.e., historical prices, trading volumes, and technical indicators) are used to predict the target variable (i.e., future prices). Popular supervised learning techniques used to predict crypto prices include linear regression, random forests, and long short-term memory (LSTM) networks. LSTM networks, a type of recurrent neural network (RNN), work well with time-series data as they can learn temporal dependencies and long-term patterns of movements in prices (Alnami et al., 2025).

Aside from supervised learning, techniques of unsupervised learning, i.e., clustering and dimensionality reduction, have been used to identify patterns in markets and segment cryptocurrencies based on the behavior of their prices. Sentiment analysis is another important application of ML in financial forecasting that involves exploring text data on news articles, social media, and other sources to gauge the mood of the markets. Sentiment analysis offers valuable information on investor sentiment and allows forecasting of movements in prices based on public sentiment (Amberrkhani et al., 2025). Technical indicators, i.e., moving averages, relative strength index (RSI), and Bollinger Bands, are also commonly used as input features in ML models. Such indicators, computed based on past prices and volumes, provide information on the behavior of markets, momentum, and volatility, thereby enabling improved forecasts. With the incorporation of these diverse data sources and techniques, ML models can generate holistic and actionable insights to be used by cryptocurrency market participants (Jagannath et al., 2021).

Research Gaps

Seabe et al. (2023) argued that despite the impressive advancement in machine learning-based cryptocurrency forecasting, several research gaps remain to be addressed. One of the most critical challenges is the lack of high-frequency forecasting models that can operate in real-time or near-real-time. Cryptocurrency markets operate 24/7, with prices moving rapidly within seconds or minutes. Current models, trained on daily or hourly data, may not be able to include these high-frequency dynamics, which lowers the utility of the models for short-term trading strategies. The development of models with the ability to process and examine data at a finer granular level, i.e., minute-by-minute or second-by-second, is critical to improve forecasting accuracy and enable decision-making in real-time (Tanwar et al, 2021).

Tripathy et al. (2024) added that another critical research gap is the need for explainable AI (XAI) in cryptocurrency forecasts. While ML models, particularly deep learning models, have demonstrated improved predictive performance, their "black-box" nature renders their decision-making processes difficult to explain. This lack of explainability is a significant barrier to adoption, particularly by institutional investors and regulators who require transparent and interpretable insights on which to make decisions. Shamshad et al. (2023) contended that Explainable AI techniques, such as Shapley additive explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), can be used to address this challenge by offering insights into the variables driving model forecasts. By increasing the interpretability of ML models, researchers can build trust and confidence in their deployments, opening the door to widespread adoption in the financial markets. Closing these research gaps will be critical to driving the field of cryptocurrency forecasting and unlocking the full potential of machine learning in financial markets.

Data Collection and Exploration

Dataset Overview

The dataset employed in this analysis comprises vast historical price data, trading volumes, and key market indicators of major cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other major altcoins. Historical price data is presented in terms of daily, hourly, and minute-level opening, closing, high, and low prices, providing detailed insights into temporal price behavior. Trading volumes, which reflect the intensity

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

of trading action, are also provided to represent liquidity and investor participation behavior. The dataset also includes various market indicators, i.e., moving averages, relative strength index (RSI), Bollinger Bands, and other technical indicators, which play a pivotal role in establishing market patterns and momentum. The data is sourced from well-known cryptocurrency exchanges, i.e., Coinbase and Binance, and financial APIs, i.e., Coin Gecko and Crypto Compare, to ensure accuracy and reliability. The sources employed provide a robust platform to examine cryptocurrency market behavior and develop predictive models. The dataset spans several years, providing scope to examine long-term patterns as well as short-term volatility, and is structured to provide scope both to exploratory data analysis and to complex machine learning applications.

Key Features Selection

S/No	Key Features	Description
01.	**Opening Price**	The opening price of a cryptocurrency at the beginning of a specific period (for instance, daily, hourly, or minute-level).
02.	**Closing Price**	The last price of a cryptocurrency at the end of a specified period.
03.	**Highest Price**	The highest price that a cryptocurrency has reached during a period.
04.	**Low Price**	The lowest level is reached by a cryptocurrency over a specific period.
05.	**Trading Volume**	The number of units of cryptocurrency traded within a specific time frame.
06.	**Moving Averages**	Average of a cryptocurrency's price over some time (e.g., 7-day, 30-day), with two forms being the Simple Moving Average (SMA) and Exponential Moving Average (EMA).
07.	**Relative Strength Index (RSI)**	It is a momentum indicator that measures the rate of change and speed of price movements, typically on a scale of 0 to 100.
08.	**Market Sentiment Score**	Quantitative indicator of investor sentiment based on text analysis of articles, social media posts, and other sources.
09.	**Return on Investment (ROI)**	The percentage change in the price of a cryptocurrency over a specific period.

Data Preprocessing

A set of common data preprocessing steps was applied in machine learning, particularly when working with time-series or tabular data. It begins with importing libraries required from pandas and scikit-learn to be applied in data manipulation, splitting, scaling, and encoding. The code then checks and converts a 'date' column to 'date-time' format if it exists, keeping in mind that it might be named otherwise. Missing values are addressed by checking for their presence and offering two options: deleting rows with any missing values or filling in the blanks with forward fill or a constant, e.g., zero. Categorical features, illustrated by a 'symbol' column, are encoded into numerical values through Label Encoding. An example 'price movement' target feature is generated by finding the 'close' price change direction. Finally, suitable features are selected, data is divided into training and test sets, and numerical features are normalized using Standard Scaler so that they have zero mean and unit variance, which is critical in many machine learning models. The resulting training and test dataset shapes are then printed to ensure.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is a crucial initial step in the research process that involves systematically investigating and summarizing a dataset to discover patterns, trends, outliers, and relationships among variables. Through data visualization techniques (histograms, scatter plots, and heatmaps), statistical summary measures (mean, median, standard deviation), and correlation analysis, EDA helps researchers learn more about the structure and nature of data. In the context of cryptocurrency market research, EDA plays a vital role in uncovering key insights, i.e., volatility patterns in prices, trading volumes, and impacts of external drivers, i.e., sentiment or macroeconomic events. It also helps in identifying data quality problems, i.e., missing values or outliers, that must be resolved before the application of advanced analytical methods. By providing a platform for hypothesis generation and informing the selection of appropriate modeling approaches, EDA ensures subsequent analysis is founded on an understood dataset and, therefore, predictive models and research findings are enhanced in terms of accuracy and reliability.

Cryptocurrency Closing Price Trend

The implemented code snippet was intended to prepare time-series cryptocurrency price data for analysis and visualization. It begins by transforming a 'date' column into a date-time format using pandas, required for time-based calculations. It then designates the 'date' column as the Data Frame index, which is common practice with time-series data to facilitate time-based indexing and efficient analysis. It then orders the data by the date-time index explicitly to be in time order, a prerequisite for any meaningful time-series analysis or modeling. The code then generates a line plot visualization of the 'close' price over time with a title, axis labels, legend, and grid for legibility, providing a good summary of the cryptocurrency price trend.

Output:

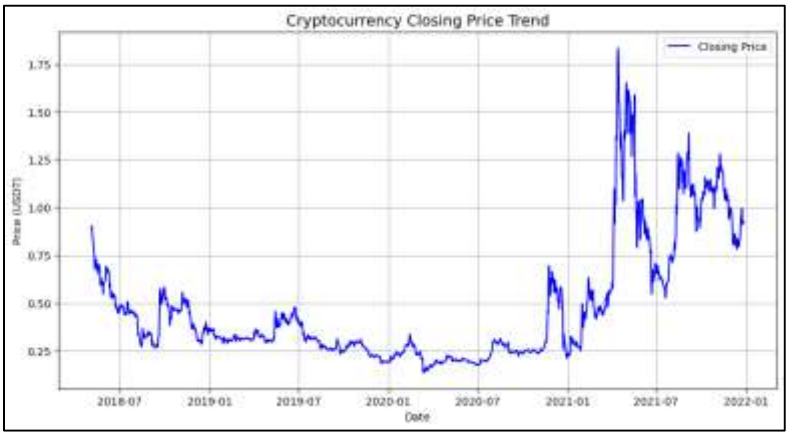


Figure 1. Cryptocurrency Closing Price Trend

The graph illustrates the closing price action of a cryptocurrency between July 2018 and January 2022, with significant volatility over time. The price initially remained quite stable, fluctuating between \$0.25 and \$0.75 until around mid-2020. It registered a sharp spike in early 2021, with the price going over \$1.75, which reflected the growing interest and investment in the cryptocurrency space then. Following the high point, the price exhibited volatility with sharp declines and comebacks, prominently seen in mid-2021. In late 2021, the price stabilized somewhat at the \$1.00 level, indicating a new level of equilibrium. Overall, the data indicates the high volatility of the cryptocurrency and how market forces influence its worth.

Moving Averages and Bollinger Bands

The executed code calculated and plotted two technical analysis indicators, viz., the Simple Moving Average (SMA) and the Bollinger Bands. It calculates the 50-day and 20-day SMAs of the 'close' price first and then calculates the upper and lower Bollinger Bands as the 20-day SMA plus and minus two standard deviations

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

of the 'close' over the previous 20 days. The code then plots the original 'close' price along with the calculated 50-day SMA, 20-day SMA, and the Bollinger Bands, with the area between the upper and the lower bands filled in to enhance visual clarity. This plot helps in identifying the trend of the price, the potential levels of support and resistance, and the volatility of the price relative to its moving average.

Output:

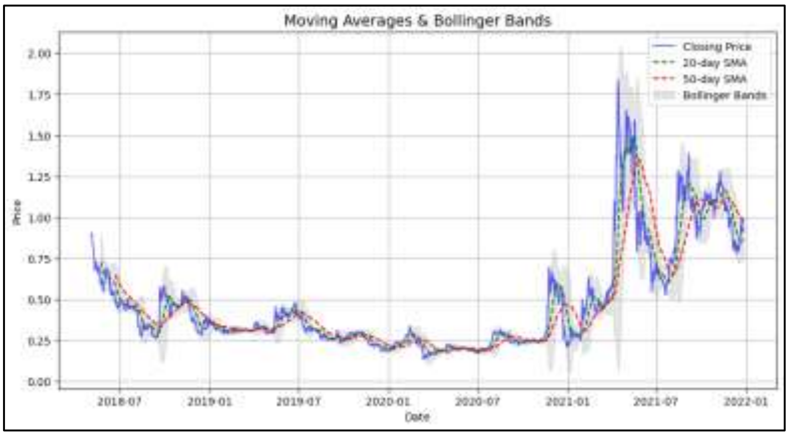


Figure 2. Moving Averages and Bollinger Bands

The chart displays the 50-day and 20-day simple moving averages (SMA) and the closing price of a cryptocurrency and its Bollinger Bands from July 2018 to January 2022. The blue line is the closing price, which is quite volatile, particularly in the first surge of 2021 when it exceeded $1.75. The 50-day SMA (dashed green line) is a smoother long-term view, and the 20-day SMA (dashed red line) closely follows the trend of the price, providing a short-term perspective. The shaded area represents the Bollinger Bands, which depict the volatility of the price; the bands increase when volatility is high, as in the surge in the price in the first few months of 2021, and decrease when volatility is low, indicating low volatility in the price. This analysis shows how moving averages and the Bollinger Bands assist in establishing the trend of the market and the potential movements of the price, depicting the volatility of the cryptocurrency and the role of the indicators in the eyes of a trader.

Correlation Heatmap of Features

The provided code created a heatmap to visualize the correlation matrix of the selected numerical features in the Data Frame, 'open', 'high', 'low', 'close', 'Volume XRP', and 'Volume USDT'. It uses the seaborn library's heatmap function to plot the pair-wise correlations in a color-coded manner, where annot=True displays the correlation values on the heatmap. The cmap='coolwarm' parameter is used to specify the color scheme to be used, where different colors represent the correlation's orientation and strength (positive or negative). The visualization is useful in establishing the linear relations between different price and volume indicators, which is important in feature selection and understanding the dynamics of the cryptocurrency market.

Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Output:

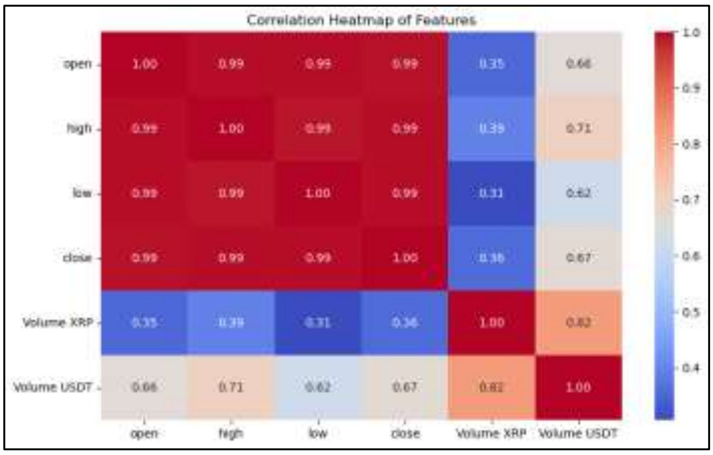


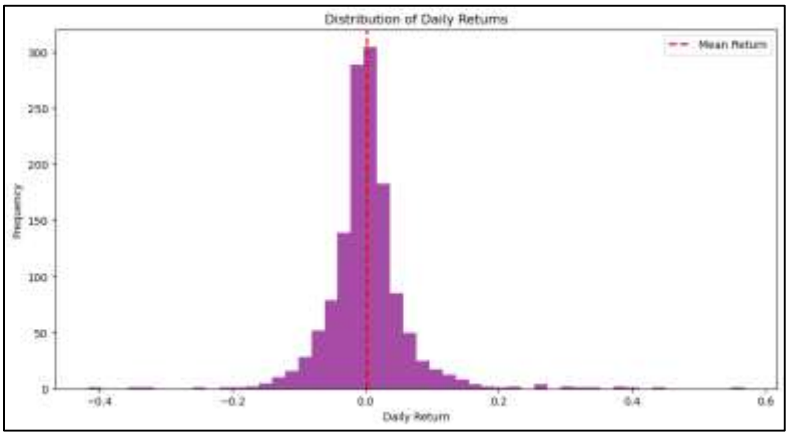
Figure 3. Correlation Heatmap of Features

The correlation heatmap displays the correlation between various features of a cryptocurrency, ranging between -1 and 1, where values signify the strength and direction of the correlation. The highly positive correlation is evident in the "open," "high," "low," and "close" features, with "close" and "open" having a correlation coefficient of 0.99, meaning that the latter two tend to increase or decrease simultaneously. Also, the "high" and "low" values of the price are highly correlated, with a correlation of 0.99, meaning uniform movements in the price throughout the trading period. However, trading volume features, i.e., "Volume XRP" and "Volume USDT," correlate less with the price features, with a low correlation of 0.35 with "open." "Volume XRP" and "Volume USDT" correlate higher at 0.82, meaning a higher correlation between the volumes compared to the prices. Overall, the heatmap shows the interdependence of the price features and relative independence of the volumes' metrics in this cryptocurrency analysis.

Distribution of Daily Returns

The code in the Python program performed a volatility analysis by calculating the daily rate of change in the 'close' price and storing it in a new 'Daily Return' column. It generated a histogram to plot the distribution of the daily returns, with 50 bins and a transparent purple color. It further added a vertical dashed red line on the histogram to mark the mean daily return and a legend to name it. The plot is then enriched with a title and axis labels, which provide a good indication of the frequency of different daily return values, which helps in determining the volatility and common movements of the asset's price.

Output:



Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Figure 4. Distribution of Daily Returns

The distribution of the daily returns of a cryptocurrency is illustrated in the histogram with a sharp spike at zero, indicating that the majority of the daily returns cluster in the area around this point. The x-axis is divided into daily returns between -0.4 and 0.6, and the y-axis is divided by frequency. The distribution is nearly symmetric with a slight tail on the left, indicating rare instances where the returns have been negative. The red dashed line represents the mean return, which is nearly zero, further indicating that the daily performance of the cryptocurrency has been stable over the period. The fact that the frequency is quite high in the narrow area around zero also suggests low volatility, but the fact that extreme values on both ends exist suggests that large swings do occur, just not as frequently. Overall, the distribution is revealing of the risk and performance profile of the cryptocurrency.

Closing Price Distribution by Weekday

The computed code made a box plot to graph the distribution of the 'close' price across various days of the week. It first makes a new 'Weekday' column in the Data Frame by extracting the name of the day from the date-time index. It then uses Sea Born's boxplot function to graph it, with 'Weekday' on the x-axis and 'close' price on the y-axis. The palette is "cool warm" for styling with colors, and the order parameter is used to get the weekdays in the correct time order. The plot is supplied with a title, axis labels, and rotated x-axis labels to facilitate easier reading, and it provides a visual comparison of the central tendency, dispersion, and potential outliers of the closing price on each weekday.

Output:

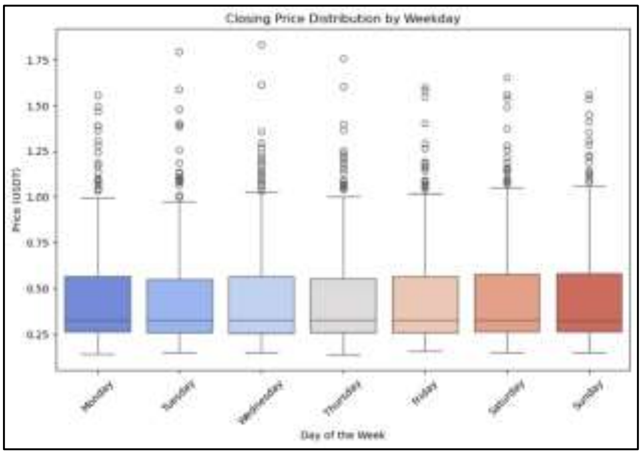


Figure 5. Closing Price Distribution by Weekday

The box plot illustrates the distribution of the cryptocurrency's closing price across different days of the week, with different price movements between Monday and Sunday. The box is the interquartile range (IQR) of the closing prices on each day, and the line inside it is the median price. Interestingly, Tuesdays and Wednesdays have higher prices, with median closing prices of between $1.00 and $1.25, as compared to other days, particularly on Saturdays and Sundays, where the median is closer to $0.75. The outliers, denoted by single points outside the whiskers, represent the occasional large price spikes, particularly on weekdays. This pattern could reflect higher trading volumes or sentiment in the middle of the week, and weekends could be less favorable to higher closing prices, possibly suggesting low trading volumes or participation in the markets. Overall, the data indicate weekday effects as being pivotal in cryptocurrency trading and pricing.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Rolling Volatility of Cryptocurrency

The implemented code calculated and printed the rolling volatility of the cryptocurrency's closing price. It computes the 10-day rolling standard deviation of the 'close' price and writes it to a new 'Rolling Volatility' column. It then makes a line plot showing how the rolling volatility changes over time, with the date on the x-axis and the volatility on the y-axis. The plot is titled, axis-labeled, and legended, so it is clear that the red line is the 10-day rolling volatility. Finally, a grid is added to the plot so it is easier to read, with a visual representation of how the volatility of the price has changed over the period being examined.

Output:

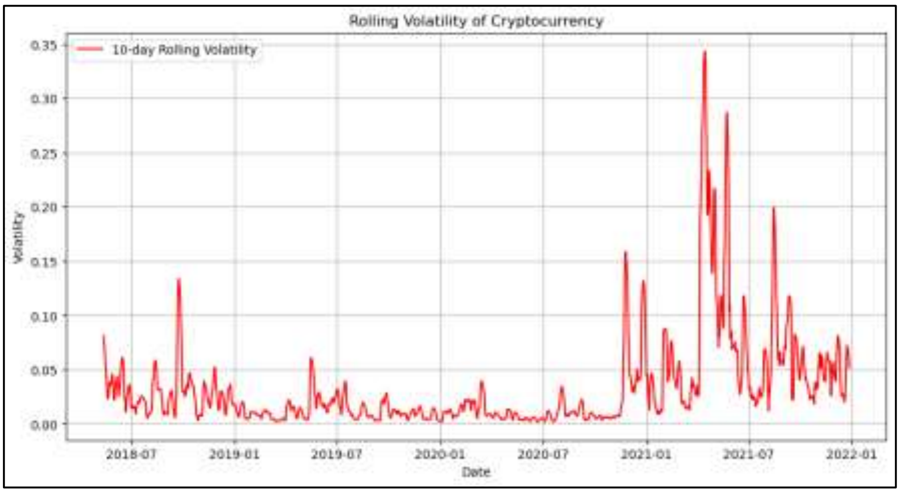


Figure 6. Rolling Volatility of Cryptocurrency

The graph shows the 10-day rolling volatility of a cryptocurrency over the period from July 2018 to January 2022, depicting significant volatility swings in the market over this period. The volatility is low initially, usually less than 0.05, indicating a stable market. Yet, there was a spectacular spike in volatility in early 2021, where it went over 0.30, which was simultaneous with heightened trading volumes and market attention toward cryptocurrencies. This volatility spike suggests very high price volatility and investor uncertainty, which is likely driven by market speculation and exogenous shocks. Subsequently, the volatility comes down but is higher than in the first episode, indicating a volatile market environment even after the initial spike. Overall, the graph shows the cryptocurrency's susceptibility to sudden changes in market conditions, emphasizing the importance of monitoring volatility to facilitate risk management and trading.

MACD Indicator for Trend Momentum

The computed code calculated and plotted the Moving Average Convergence Divergence (MACD) indicator, a popular indicator of trend momentum. It begins by calculating the 12-period and 26-period Exponential Moving Averages (EMAs) of the 'close' price. It then calculates the MACD line by subtracting the 26-period EMA from the 12-period EMA. It then calculates a 9-period EMA of the MACD line as the 'Signal Line'. The code then plots the MACD line (blue) and the Signal Line (red and dashed) over time, along with a horizontal dotted gray line at zero. This plot helps traders to identify potential buy and sell signals based on the MACD and Signal line crossovers and the position of the MACD relative to the zero line.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Output:

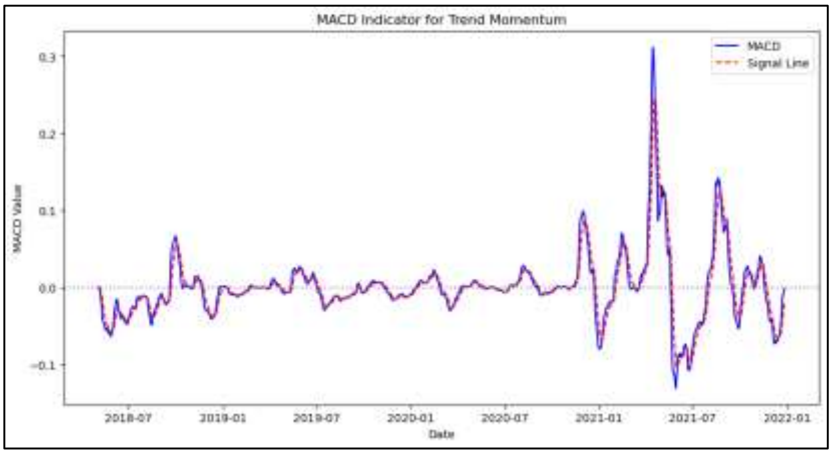


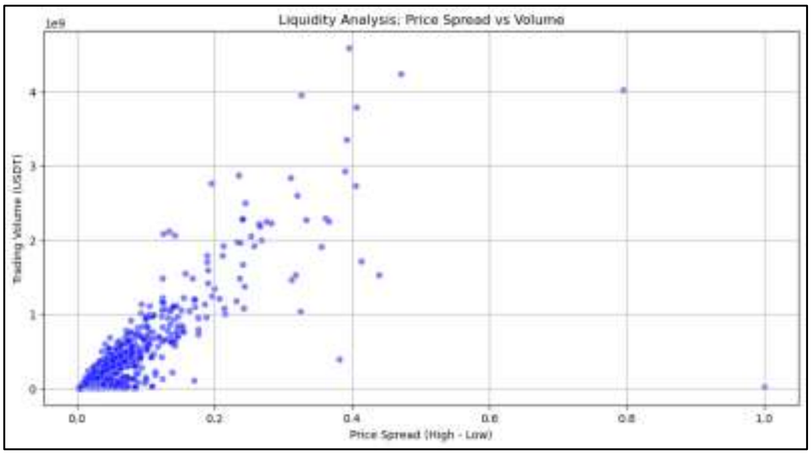
Figure 7. MACD Indicator for Trend Momentum

The graph presents the MACD (Moving Average Convergence Divergence) indicator and signal line from July 2018 through January 2022, which are critical tools for identifying trend momentum within the cryptocurrency market. The MACD line, plotted in blue, fluctuates around the zero line, highlighting zones of bullish and bearish momentum. Notably, a large spike occurs in early 2021, with the MACD exceeding 0.2, reflecting strong bullish momentum during this time, perhaps a reflection of the broader market hype surrounding cryptocurrencies. The red dashed signal line provides a smoother alternative, with buy and sell signals based on crossovers with the MACD line. Over the period under observation, several crossovers occur, particularly in late 2020 and early 2021, reflecting potential trading opportunities. Overall volatility of the MACD points to a dynamic market environment, underscoring the importance of the MACD as a trend-following momentum indicator to traders looking to profit from the price movements.

Liquidity Analysis Price Spread vs Volume

The code fragment performed a basic liquidity analysis by comparing the price spread with the trading volume. It calculated the daily spread in terms of 'high' and 'low' prices and put this in a new 'Spread' column. It then makes a scatter plot using seaborn where 'Spread' is on the x-axis and 'Volume USDT' on the y-axis. The plot includes blue markers with some transparency so that data points are visible. The title and axis labels indicate that the plot is meant to explore how the trading volume is related to the price spread, and this could help ascertain the liquidity of the market at different ranges of prices. It includes a grid to enable easier reading of the scatter plot.

Output:



Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Figure 8. Liquidity Analysis Price Spread vs Volume

The scatter plot analyzes the correlation between trading volume and price spread (as the difference between high and low) of a cryptocurrency, with both being plotted on the x and y axes. The x-axis indicates the price spread, ranging between 0 and slightly over 1 USD, and the y-axis indicates trading volume, up to 4 billion USD. The bulk of data points cluster in the bottom left quadrant, indicating that low price spreads have higher trading volumes, indicating higher liquidity in those situations. With a growing price spread, the number of data points decreases, indicating higher spreads being supported by decreasing trading volumes. This trend indicates the inverse correlation between liquidity and price spread, where tighter spreads tend to enable higher trading volumes. Overall, the chart highlights the importance of price spread as a key driver of liquidity in the cryptocurrency market, which is an important consideration when assessing markets.

Methodology

Feature Engineering

Feature engineering is a crucial step in the machine learning pipeline that transforms raw data into meaningful inputs that enhance the predictive power of models. In predicting cryptocurrency prices, some of the most predictive features include moving averages, Relative Strength Index (RSI), trading volume, and sentiment scores. Moving averages, i.e., Simple Moving Average (SMA) and Exponential Moving Average (EMA), smoothen price volatility and identify trends over a specific time horizon. These features help models filter out the noise in the short term from the price movement in the long term. The RSI, a momentum oscillator, provides information about overbought or oversold conditions, which are critical in predicting potential reversals in the price. Trading volume, a measure of market activity, is an indicator of liquidity and investors' interest and tends to be associated with significant price changes. Sentiment scores, based on textual analysis of news articles and social media, put a number on the emotional tone of the market, incorporating a qualitative element into the dataset.

Lagged features are produced to better represent short-term and long-term patterns to further enhance the predictive ability of the models. Lagged features involve the use of past data at previous time steps (for example, past 7, 30, or 90 days' worth of volumes or prices) as input to predict future values. For example, last night's closing price or last week's average trading volume might be helpful context to describe current movements in the price. Lagged features enable models to include temporal dependencies and cyclical patterns in the data, particularly applicable in highly volatile markets like cryptocurrencies. With the addition of these engineered features, the dataset is enriched and rendered more informative, enabling machine learning models to pick up on complex patterns and provide better forecasts.

Model Selection and Training

The use of machine learning models for cryptocurrency price prediction is based on selecting models that can capture the unique nature of the market. Three models are chosen in this study: Logistic Regression, Random Forest Classifier, and XG Boost Classifier. Logistic Regression is included as a baseline for predicting price directions (upward or downward trends) based on linear relationships between input features and the target. Simple as it may be, the model provides a baseline for measuring the performance of sophisticated models. Random Forest Classifier, as an ensemble of decision trees, is included to capture non-linear relationships and interaction among features. Random Forest aggregates the decisions of several trees to prevent overfitting and generalize well to new, unseen data. XG Boost Classifier, as a high-performing gradient boosting algorithm, is included due to its high predictive performance and efficiency in handling large data. XG Boost constructs decision trees in a sequential manner, where each tree attempts to correct the mistakes of the preceding tree, and thus it develops a highly optimized model.

The use of ensemble learning techniques, i.e., Random Forest and XG Boost, is backed by the argument that they can harness the strengths of multiple models and offset each other's weaknesses. Ensemble techniques improve predictive capacity by reducing variance, boosting robustness, and modeling diverse

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

patterns in data. For instance, a decision tree may overfit noise in data, but an ensemble of trees averages out such mistakes, making more accurate predictions. Using these models, the study aims to achieve a trade-off between explainability (Logistic Regression), flexibility (Random Forest), and performance (XG Boost), allowing comprehensive coverage of the problem space.

Model Optimization and Performance Analysis

For maximum possible accuracy and robustness of the selected models, hyperparameter tuning is done using Grid Search, an exhaustive method of evaluating different combinations of hyperparameters. Grid Search systematically tries predefined ranges of hyperparameters, e.g., the number of trees in Random Forest or the learning rate in XGBoost, to determine the optimal setup of each model. This process ensures that the models are tuned to the nuances of the cryptocurrency dataset, enhancing their predictive capability. Additionally, cross-validation is applied to test the models' ability to generalize and prevent overfitting. Cross-validation divides the dataset into several folds, trains the model on part of the data, and then validates it on the rest of the folds. This approach provides a better estimate of the performance of the models by reducing the impact of data variability and ensuring the models generalize well to different subsets of data.

The optimization also includes an evaluation of the model's performance on both the training and validation data to check for overfitting. Overfitting occurs when a model learns to identify noise in the training data rather than underlying patterns, leading to poor performance on new data. Through careful hyperparameter tuning and the use of cross-validation, the study ensures that the models find a balance between complexity and generality, leading to stable and reliable predictions.

Evaluation Metrics

The performance of the models is evaluated using a combination of regression metrics when predicting prices and classification metrics when forecasting movements in the markets. For regression tasks, Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are used to measure the accuracy of the projected prices relative to actual prices. MSE computes the average square of the difference between actual and projected values, and RMSE provides a better understandable measure by finding the square root of MSE. Such measures are particularly useful when assessing the accuracy of projected prices and finding models that minimize errors in projections.

For classification models, accuracy, precision, recall, and F1-score metrics are employed to evaluate the performance of the models in terms of predicting the directions of the markets (e.g., upward or downward directions). Accuracy measures the proportion of correct predictions, and precision and recall measure the ability of the models to predict positive cases (e.g., upward directions) and not misrepresent false positives or false negatives, respectively. The F1-score, being the harmonic mean of precision and recall, provides a balanced measure of performance, particularly when dealing with imbalanced data where one class is dominant. A comparison of performances is carried out to evaluate the strengths and weaknesses of both models. Logistic Regression is interpretable and easy to understand, but it may not be capable of capturing complex non-linear relationships in data. Random Forest, being an ensembling method, is better in terms of accuracy and reliability but may be computationally intensive. XG-Boost has advanced optimization techniques.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Results and Analysis

Cryptocurrency Market Trend Analysis

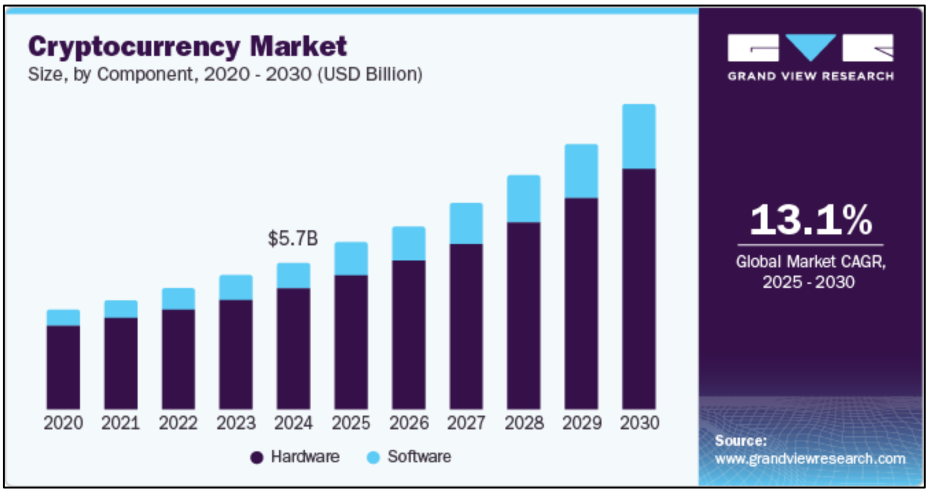


Figure 9. Cryptocurrency Market Trend

The graph shows the size of the growth of the cryptocurrency market, split between hardware and software segments, estimated between 2020 and 2030 in terms of USD billions. The stacked bar graph shows a steady growth in the size of the total market over the decade, with hardware (dark purple segment) and software (light blue segment) both contributing to the growth. The software segment is seen to have a slightly greater share in the size of the market than hardware in the majority of the years estimated. The graph also shows a global market Compound Annual Growth Rate (CAGR) of 13.1% between 2025 and 2030, with significant expected growth in the second half of the estimated period. The data is reported as being taken from Grand View Research.

Model Performance Evaluation:

Logistic Regression Modelling

The code utilizes a Logistic Regression classification algorithm, perhaps modeling the 'price movement' target set up in earlier steps. It initially imports scikit-learn modules, the Logistic Regression class, Grid Search CV for hyperparameter tuning, and metrics to be used in the evaluation. It initializes a Logistic Regression model with the 'liblinear' solver. It then creates a parameter grid to search over values of the regularization strength ('C') and the type of penalty ('penalty': 'l1' and 'l2'). It then utilizes Grid Search CV to execute a grid search over this parameter grid to discover the hyperparameter combination that achieves the best performance on the training set. After it fits the grid search, it prints the best parameters and utilizes the model with the best parameters to predict the scaled test set. It then utilizes a classification report, a confusion matrix, and the accuracy score to assess the performance of the tuned Logistic Regression model and prints it to measure the effectiveness of the model.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Output:

Table 1. Logistic Regression Results

```
precision     recall   f1-score    support

           0       0.69      0.22      0.34       129
           1       0.53      0.90      0.67       128

    accuracy                           0.56       257
   macro avg       0.61      0.56      0.50       257
weighted avg       0.61      0.56      0.50       257

Accuracy for Logistic Regression: 0.5603
```

Accuracy for Logistic Regression: 0.5603

The above table shows the performance metrics of an Optimized Logistic Regression model with specific parameters, i.e., a regularization strength (C) of 10 and penalty type 'l1'. Precision, recall, and F1-score values of class 0 are 0.69, 0.22, and 0.34, respectively, indicating that the model is very good at classifying class 0 instances but is poor in recall, suggesting a high number of false negatives. For class 1, the low precision is 0.53, recall is 0.90, and thus the F1-score is 0.67, indicating better detection of class 1 instances. The accuracy of the model is 56.03%, with a total support of 257 instances in both classes. Macro average metrics suggest well-balanced detection across classes, despite the low recall of class 0, suggesting scope for improvement. The confusion matrix indicates that the model misclassified 29 class 0 instances as class 1 and 13 class 1 instances as class 0, and thus the need to improve further in discriminating between the two classes.

Random Forest Classifier Modelling

The implemented code employed a Random Forest Classifier for prediction, again perhaps on the 'price movement' target variable. It employs the scikit-learn required classes, i.e., Random Forest Classifier, Grid Search CV, and evaluation metrics. The Random Forest Classifier is initialized with a set random state to provide reproducibility. A large hyperparameter grid is defined to tune the number of trees (n-estimators), maximum tree depth (max_depth), the number of samples required to split an internal node (min_samples_split), and the number of samples required to be at a leaf node (min_samples_leaf) 1 2. GridSearchCV is then employed to perform 5-fold cross-validation on all possible combinations of the hyperparameters to find the set that provides maximum performance on the training data. The optimal hyperparameters learned by GridSearchCV are then printed, and the trained model with the optimal hyperparameters is employed to predict the scaled test data. Finally, the accuracy of the tuned Random Forest is tested and reported using a classification report, confusion matrix, and overall accuracy score.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Output:

Table 2. Random Forest Classifier Results

```
precision    recall  f1-score   support

           0       0.46      0.46      0.46       129
           1       0.46      0.47      0.47       128

    accuracy                           0.46       257
   macro avg       0.46      0.46      0.46       257
weighted avg       0.46      0.46      0.46       257

Accuracy for Random Forest: 0.4630
```

The table above presents the performance metrics of a Random Forest model, with precision, recall, and F1-score provided for two classes, 0 and 1. For both classes, the precision is 0.46, indicating that the model is correct in 46% of positive calls. The recall is also 0.46 for class 0, indicating that the model is catching just 46% of true class 0 instances, and class 1 is slightly better at 0.47 recall, indicating it has a slightly better ability to detect true positives in that class. The F1-scores for both classes are identical at 0.46, indicating both precision and recall are well matched, yet overall performance is low. The accuracy of the model is 46.30%, with total support of 257 instances in both classes, indicating much potential for improvement. The confusion matrix indicates that the model is misclassifying 59 class 0 instances as class 1 and 68 class 1 instances as class 0, again indicating the challenge in being able to tell the two classes apart.

XG-Boost Classifier Modelling

The code script utilized an XG Boost classifier, yet another effective algorithm well-suited to classification tasks like predicting 'price movement'. It utilizes the XG Boost package and the scikit-learn modules required to select the model and evaluate it. An XG Boost classifier is initialized with a random state to provide reproducibility, disables the label encoder, and provides the evaluation function as 'log loss'. It initializes a parameter grid to search over the key hyperparameters such as number of trees (n-estimators), maximum depth of each tree (max depth), learning rate (learning rate), proportion of samples used when training each tree (subsample), and proportion of features used when constructing each tree (colsample_bytree). GridSearchCV is then used to determine the optimal set of these hyperparameters using 5-fold cross-validation on the scaled training data. The optimal set of parameters discovered by Grid Search CV is then printed, and the corresponding model is used to predict the scaled test data. Finally, the performance of the model is tested and reported using a classification report, confusion matrix, and accuracy score.

Output:

Table 3. XG-Boost Results

```
 precision    recall  f1-score   support

           0       0.46      0.49      0.47       129
           1       0.45      0.41      0.43       128

    accuracy                           0.45       257
   macro avg       0.45      0.45      0.45       257
weighted avg       0.45      0.45      0.45       257

Accuracy for XGBoost: 0.4514
```

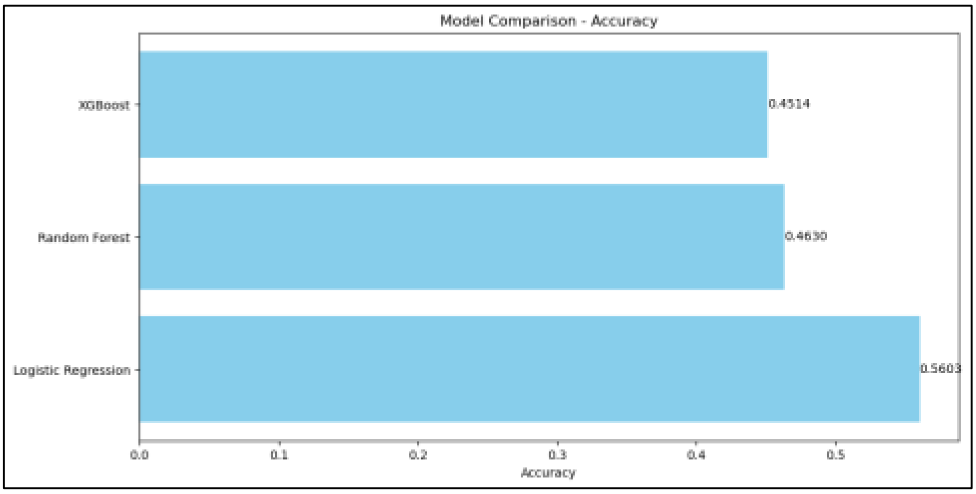
Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

The table above provides the performance metrics of an XG-Boost model, showing its recall, precision, and F1 score on two classes, 0 and 1. The precision of class 0 is 0.46, indicating that 46% of class 0 instances it classified as correct, and class 1 is slightly less at 0.45. For recall, class 0 is at 0.49, indicating the model is correct on 49% of true class 0 instances, and class 1 is at 0.41, indicating it is correct on 41% of true positives. The respective F1 scores are at 0.47 for class 0 and 0.43 for class 1, indicating a moderate level of balance between recall and precision. The accuracy of the overall model is at 45.14%, with 257 instances supported in both classes. The confusion matrix shows misclassifying 63 class 0 instances as class 1 and 75 class 1 instances as class 0, indicating huge challenges in being correct in discriminating between both classes, indicating the need to further tune the model.

Comparison of All Models

The implemented code focused on comparing the accuracy of Logistic Regression, Random Forest, and XG-Boost classification models. It first initializes a dictionary named model comparison to store the names of the models and corresponding accuracy values, which were possibly calculated in the earlier code cells. It then converts the dictionary to a pandas DataFrame named comparison_df to make it easier to handle and visualize. The Data Frame is sorted by the 'Accuracy' column in descending order so that the best-performing model is seen readily. Finally, a horizontal bar chart is generated using matplotlib to visually compare the accuracies of the three models, with the accuracy values on the bars to make it easier to compare. This provides an efficient and effective summary of the performance of the different classification models on the problem at hand.

Output:



The bar chart is a comparison of the accuracy of three machine learning models: Logistic Regression, Random Forest, and XG Boost. With the highest accuracy of 56.03%, Logistic Regression was the best-performing of the models tested, showing its relative superiority. Next, the Random Forest model achieved 46.30% accuracy, which was close but reflected that it captures market trends well enough, but not as well as Logistic Regression. XG Boost, which had the lowest ranking with 45.14% accuracy, did not live up to expectations in its predictive capability despite being an advanced algorithm. Overall, the graph is a clear indication of how each of these models has performed based on accuracy, and Logistic Regression proved to be the most reliable model for the dataset analyzed.

The results of the comparison of the models shed light on the importance of a feature importance analysis in establishing which indicators have a major impact on movements in cryptocurrency prices. Examining the features used in the models, i.e., moving averages, RSI, trading volumes, and sentiment scores, we can establish which factors have the biggest impact on movements in price. Knowledge of said contributions is critical to traders and analysts who must make decisions based on behavior within markets. The trade-offs in selecting a model, however, should also be considered, i.e., the trade-off between explainability and

Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

predictive power. While models like Logistic Regression provide greater transparency and explainability, allowing stakeholders to be aware of the reasons behind the forecasts, they may not be as predictive as intricate models like XG Boost and Random Forest. Thus, the choice relies on the balance of the strengths of transparent insights compared to the need for robust forecasts, ultimately establishing the decision-making process within a dynamic and volatile environment.

Practical Applications in the USA

Impact on Financial Markets

The integration of AI forecasts into cryptocurrency trading has the potential to revolutionize the United States financial markets by providing traders and institutional investors with advanced tools to make decisions. For crypto traders, machine learning (ML) models deliver actionable data on patterns of change in prices, enabling them to spot profitable points of entry and exit with greater accuracy. With predictive analytics, traders can anticipate movements in markets, i.e., bullish or bearish, and adjust strategies accordingly. This is particularly helpful in the highly volatile cryptocurrency market, where steep movements in prices over short durations can accrue massive profits or losses. Institutional players, i.e., asset managers and hedge funds, also stand to benefit from AI-driven forecasts by incorporating them into portfolio management and risk evaluation. For example, ML models are capable of picking up patterns of correlation between cryptocurrencies and traditional assets, enabling investors to diversify portfolios and mitigate risks. Predictive analytics also enhances risk management strategies by providing warnings of impending market downturns or liquidity crises. With machine learning insights being woven into processes, financial institutions can make data-driven decisions that drive higher profitability and reduce exposure to market uncertainties.

Besides, AI tools can be employed to create algorithmic trading strategies where trades are executed automatically based on pre-decided criteria. Such strategies could be based on predictive models and real-time data to capitalize on arbitrage opportunities and inefficiencies in the markets. For instance, ML-driven algorithmic strategies could analyze order book data, trading volumes, and price movements to execute high-frequency trades with low latency. This not only makes trading efficient, but also assists in providing liquidity and price discovery in the markets. Overall, the application of AI-driven forecasts in the United States' markets has the potential to make the trading environment better informed, efficient, and robust, both at the level of individual traders and institutional players.

The Regulatory Impact of Trading in the USA

The use of AI tools in cryptocurrency trading also has significant implications for United States regulation compliance. Cryptocurrencies come under the purview of oversight agencies such as the Securities and Exchange Commission (SEC) and the Commodity Futures Trading Commission (CFTC), which implement regulations to protect investors and maintain market integrity. AI predictive models have the potential to assist market participants in complying with regulations by providing transparent and auditable insights into trading activities. For example, ML models can be used to identify transactions that demonstrate signs of market manipulation, e.g., wash trading or spoofing, and label the activities as suspect so that further investigation is feasible. This is in line with the SEC's mission to prevent fraud and ensure fair and orderly markets.

Apart from fraud detection, AI tools can also assist in maintaining market stability by identifying system-wide threats and providing warnings of potential disruptions. For instance, predictive models can scan markets to identify signs of excessive speculation or a lack of liquidity and enable regulators to undertake preventive measures to stabilize the market. This is particularly important in the cryptocurrency markets, where the lack of central oversight and speculative trading pose potential volatility. With the use of AI in regulatory compliance and market monitoring, U.S. regulators will be better equipped to regulate quickly evolving cryptocurrency markets and protect investors from potential threats. Secondly, the design of explainable AI (XAI) techniques will be able to mitigate concerns over the explainability and transparency of machine learning models. Explainable AI explains why the models make certain predictions, so it is easier

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

for regulators to assess the reliability and accuracy of AI tools. This is critical in building trust in AI tools and having them used responsibly within the markets. With cooperation between tech providers, market participants, and regulators, the United States is capable of creating a regulatory environment that fosters innovation without compromising market integrity.

Integration with Trading Platforms

The integration of machine learning tools within cryptocurrency trading platforms is a significant step towards unleashing the true potential of AI in the financial markets. One of the most promising applications is the use of ML-based trading bots that automatically execute trades based on predictive models and pre-programmed strategies. The bots can scan large data volumes in real-time, ranging from trading volumes and price movements to market sentiment, to identify profitable opportunities and make trades with minimal intervention. For example, a trading bot could use an ML model to predict short-term price movements and make buy or sell orders accordingly, maximizing returns and minimizing risk. Automated trading bots find special utility in the cryptocurrency markets, where prices fluctuate rapidly and trading is done 24/7. Apart from trading robots, predictive analytics is also used in automated investment strategies, i.e., robo-advisors, which suggest customized investment recommendations based on an individual's risk profile and investment goals.

With the incorporation of ML models, robo-advisors can suggest improved and dynamic recommendations, dynamically updating portfolios based on changing market scenarios. This brings complex investment strategies within everyone's reach, enabling retail investors to utilize the same complex tools used by institutional investors. Besides, predictive analytics can be used to fortify risk management on cryptocurrency exchanges by identifying potential threats and vulnerabilities. For instance, ML models can analyze trading behavior to identify signs of market manipulation or liquidity risk so that exchanges can take proactive measures to protect users. Predictive analytics can be used to optimize margin requirements and collateral management, reducing the risk of defaults and keeping the trading platform stable. With machine learning in their business, cryptocurrency exchanges can create a safer and better trading environment, encouraging trust and confidence in users.

Discussion and Future Directions in the USA

Challenges in AI-Based Cryptocurrency Forecasting

Notwithstanding the significant advancement in AI-driven cryptocurrency forecasting, several challenges persist that hinder the development and application of predictive models. One of the major challenges is the rampant existence of manipulative activities and false volumes of trading in the cryptocurrency markets. Unlike the highly regulated traditional financial markets, the cryptocurrency platform is a decentralized network that is not uniformly regulated, making it susceptible to manipulative tactics such as wash trading, spoofing, and pump-and-dump. Such activities distort trading volumes and price action, injecting noise into data and affecting the accuracy of predictive models. For instance, false volumes of trading mislead models to recognize false trends, hence providing incorrect forecasts. Overcoming this challenge necessitates the design of sophisticated anomaly detection tools that identify and filter out manipulative activities so that models learn on clean and genuine data.

One of the biggest challenges is the volatility and extreme price movements that characterize cryptocurrency markets. While volatility offers profit opportunities, it also renders it unpredictable, as the prices change rapidly over short intervals of time due to speculative trading, announcements, or macroeconomic events. Traditional machine learning models, which learn from historical data, may not be capable of dealing with sudden and unprecedented movements in prices and thus suffer in terms of predictive accuracy. To combat this issue, researchers must explore adaptive models that adapt dynamically to reflect changing market conditions in real-time. Also, incorporating external data sources, i.e., news sentiment and macroeconomic data, can provide context to movements in price and make predictive models stronger. The challenges indicate the necessity of constant innovation and optimization of AI-driven forecasting techniques so that they remain effective in the highly dynamic cryptocurrency environment.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Limitations of the Study

While this study provides valuable information on leveraging machine learning to predict cryptocurrency prices, it is not without limitations. One limitation is data availability on less well-known or smaller cryptocurrencies, referred to as altcoins. While well-known cryptocurrencies such as Bitcoin and Ethereum have a rich history and high liquidity, smaller cryptocurrencies have low volumes and short histories. This limited data prevents the effective creation of predictive models since machine learning models require large and diverse data to learn meaningful patterns. Additionally, data on smaller cryptocurrencies could be of low quality with gaps or inconsistencies affecting the performance of models. Overcoming this limitation requires higher-quality and fuller data, as well as techniques to train models on sparse or incomplete data.

Another limitation is the potential biases in predictive models caused by unpredictable events in the markets. Cryptocurrency markets rely on a wide range of factors, from regulatory announcements to technological innovations and political events, many of which cannot be foreseen or quantifiable. For example, an unanticipated regulatory crackdown or a major security breach initiates large-scale movements in the price that diverge from historical patterns, leading to incorrect projections. Such unpredictable events introduce uncertainty into the process of forecasting, and it is hard to develop models that are always reliable. To overcome this limitation, future research could include incorporating event detection systems in real time that are capable of detecting and responding to unanticipated market events, making predictive models more dynamic.

Future Research Directions

The field of AI-based cryptocurrency forecasting offers numerous areas of future research with the potential to break through present limitations and unlock new paths of market analysis. One of those areas is the use of deep learning models, i.e., Long Short-Term Memory (LSTM) networks, to time-series cryptocurrency forecasting. LSTMs are a type of recurrent neural network (RNN) that is well-suited to modeling temporal dependencies and long-term patterns in serial data and thus ideal to apply on cryptocurrency price movements. With the assistance of LSTMs, researchers can construct models that better capture the complex and non-linear dynamics of cryptocurrency markets, with improved predictive power. One of the areas of potential is incorporating social media sentiment analysis into predictive models. Social media outlets, such as Twitter and Reddit, have a significant impact on the sentiment of the cryptocurrency markets, with opinion leaders and groups influencing the action in the markets through opinion and discussion. With sentiment analysis, predictive models will be capable of sensing the emotional sentiment of the markets and identifying patterns that may not be visible through price data. For example, an increase in positive sentiment on social media could be predictive of an imminent rally in the price, and conversely.

Advanced NLP techniques, such as transformer models (e.g., BERT), can be used on social media data to extract actionable insights to be applied in forecasting markets. Aside from that, tremendous potential lies in creating AI-driven investment advisory systems tailored to the requirements of financial institutions. Such systems can be capable of leveraging machine learning models to provide investment recommendations, make portfolio allocations, and manage risk in real-time. For instance, an AI-driven advisory system could assess a client's risk appetite, investment objective, and current market scenario to recommend a diversified basket of cryptocurrencies and traditional assets. Such systems, through automation, bring efficiency, reduce expenses, and provide improved results to investors. Additionally, the integration of explainable AI (XAI) techniques holds the potential to introduce transparency to the decision-making process, building confidence and trust in users.

Conclusion

The central objective of this research was to develop and evaluate machine learning-driven models of cryptocurrency price trend forecasting. The focus of this research project revolved around prominent cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other prominent altcoins, within the United States. The dataset employed in this analysis comprises vast historical price data, trading volumes, and key

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

market indicators of major cryptocurrencies, i.e., Bitcoin (BTC), Ethereum (ETH), and other major altcoins. Historical price data is presented in terms of daily, hourly, and minute-level opening, closing, high, and low prices, providing detailed insights into temporal price behavior. Trading volumes, which reflect the intensity of trading action, are also provided to represent liquidity and investor participation behavior. The dataset also includes various market indicators, i.e., moving averages, relative strength index (RSI), Bollinger Bands, and other technical indicators, which play a pivotal role in establishing market patterns and momentum. Three models are chosen in this study: Logistic Regression, Random Forest Classifier, and XG Boost Classifier. For classification models, accuracy, precision, recall, and F1-score metrics are employed to evaluate the performance of the models in terms of predicting the directions of the markets (e.g., upward or downward directions). With the highest accuracy, Logistic Regression was the best-performing of the models tested, showing its relative superiority. The integration of AI forecasts into cryptocurrency trading has the potential to revolutionize the United States financial markets by providing traders and institutional investors with advanced tools to make decisions. The use of AI tools in cryptocurrency trading also has significant implications for United States regulation compliance. The integration of machine learning tools within cryptocurrency trading platforms is a significant step towards unleashing the true potential of AI in the financial markets. The field of AI-based cryptocurrency forecasting offers numerous areas of future research with the potential to break through present limitations and unlock new paths of market analysis. One of those areas is the use of deep learning models, i.e., Long Short-Term Memory (LSTM) networks, to time-series cryptocurrency forecasting.

References

Ahmad, T., & Abbas, M. (2024). AI Based Cryptocurrency Price Prediction: A Comparative Analysis of Traditional & Deep Learning Models with Sentiment Integration (Master's thesis, UIS).

Akula, R. & Pinky, J. N., (2024). Enhancing Cryptocurrency Market Forecasting: Advanced Machine Learning Techniques and Industrial Engineering Contributions. arXiv preprint arXiv:2410.14475.

Ali, J., Alhussein, M., Hassan, S., Aurangzeb, K., & Umair, M. (2024). An improved machine learning-driven framework for cryptocurrencies price prediction with sentimental cautioning. IEEE Access.

Alnami, H., Mohzary, M., Assiri, B., & Zangoti, H. (2025). An Integrated Framework for Cryptocurrency Price Forecasting and Anomaly Detection Using Machine Learning. Applied Sciences, 15(4), 1864.

Amberkhani, A., Bolisetty, H., Narasimhaiah, R., Jilani, G., Baheri, B., Muhajab, H., ... & Shubbar, S. (2025, March). Revolutionizing Cryptocurrency Price Prediction: Advanced Insights from Machine Learning, Deep Learning and Hybrid Models. In Future of Information and Communication Conference (pp. 274-286). Cham: Springer Nature Switzerland.

Byun, Y. C. & Shahbazi, Z., (2021). Improving the cryptocurrency price prediction performance based on reinforcement learning. IEEE Access, 9, 162651-162659.

Chouksey, A., Shovon, M. S. S., Tannier, N. R., Bhowmik, P. K., Hossain, M., Rahman, M. S., ... & Hossain, M. S. (2023). Machine Learning-Based Risk Prediction Model for Loan Applications: Enhancing Decision-Making and Default Prevention. Journal of Business and Management Studies, 5(6), 160-176.

Das, B. C., Sarker, B., Saha, A., Bishnu, K. K., Sartaz, M. S., Hasanuzzaman, M., ... & Khan, M. M. (2025). Detecting Cryptocurrency Scams in the USA: A Machine Learning-Based Analysis of Scam Patterns and Behaviors. Journal of Ecohumanism, 4(2), 2091-2111.

El Mahjouby, M., Bennani, M. T., Lamrini, M., El Far, M., Bossoufi, B., & Alghamdi, T. A. (2024). Predicting market performance using machine and deep learning techniques. IEEE Access.

Ganesh, D., & Harshavardhan, V. C. (2024, May). Cryptocurrency Market Dynamics: A Machine Learning-Based Approach for Price Prediction. In 2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES) (pp. 1-6). IEEE.

Gupta, S. P., & Sethi, M. (2024, December). Predictive Modeling in Cryptocurrency: A case Study of Bitcoin using RNN. In 2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N) (pp. 635-641). IEEE.

Islam, M. R., Nasiruddin, M., Karmakar, M., Akter, R., Khan, M. T., Sayeed, A. A., & Amin, A. (2024). Leveraging Advanced Machine Learning Algorithms for Enhanced Cyberattack Detection on US Business Networks. Journal of Business and Management Studies, 6(5), 213-224.

Islam, M. Z., Islam, M. S., Al Montaser, M. A., Rasel, M. A. B., Bhowmik, P. K., & Dalim, H. M. (2024). EVALUATING THE EFFECTIVENESS OF MACHINE LEARNING ALGORITHMS IN PREDICTING CRYPTOCURRENCY PRICES UNDER MARKET VOLATILITY: A STUDY BASED ON THE USA FINANCIAL MARKET. The American Journal of Management and Economics Innovations, 6(12), 15-38.

Islam, M. Z. ., Islam, M. S., das, B. C. ., Reza, S. A. ., Bhowmik, P. K. ., Bishnu, K. K. ., Rahman, M. S. ., Chowdhury, R., & Pant, L. . (2025). Machine Learning-Based Detection and Analysis of Suspicious Activities in Bitcoin Wallet Transactions in the USA. Journal of Ecohumanism, 4(1), 3714 –. https://doi.org/10.62754/joe.v4i1.6214

Jagannath, N., Barbulescu, T., Sallam, K. M., Elgendi, I., Okon, A. A., McGrath, B., ... & Munasinghe, K. (2021). A self-adaptive deep learning-based algorithm for predictive analysis of bitcoin price. IEEE Access, 9, 34054-34066.

Journal of Ecohumanism
2025
Volume: 4, No: 2, pp. 2498 – 2519
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i2.6663

Khan, M. T., Akter, R., Dalim, H. M., Sayeed, A. A., Anonna, F. R., Mohaimin, M. R., & Karmakar, M. (2024). Predictive Modeling of US Stock Market and Commodities: Impact of Economic Indicators and Geopolitical Events Using Machine. Journal of Economics, Finance and Accounting Studies, 6(6), 17-33.

Jui, A. H., Alam, S., Nasiruddin, M., Ahmed, A., Mohaimin, M. R., Rahman, M. K., ... & Akter, R. (2023). Understanding Negative Equity Trends in US Housing Markets: A Machine Learning Approach to Predictive Analysis. Journal of Economics, Finance and Accounting Studies, 5(6), 99-120.

Mohaimin, M. R., Das, B. C., Akter, R., Anonna, F. R., Hasanuzzaman, M., Chowdhury, B. R., & Alam, S. (2025). Predictive Analytics for Telecom Customer Churn: Enhancing Retention Strategies in the US Market. Journal of Computer Science and Technology Studies, 7(1), 30-45.

Poudel, S., Paudyal, R., Cankaya, B., Sterlingsdottir, N., Murphy, M., Pandey, S., ... & Poudel, K. (2023). Cryptocurrency price and volatility predictions with machine learning. Journal of Marketing Analytics, 11(4), 642-660.

Seabe, P. L., Moutsinga, C. R. B., & Pindza, E. (2023). Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: a deep learning approach. Fractal and Fractional, 7(2), 203.

Tanwar, S., Patel, N. P., Patel, S. N., Patel, J. R., Sharma, G., & Davidson, I. E. (2021). Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations. IEEE Access, 9, 138633-138646.

Tripathy, N., Hota, S., Mishra, D., Nayak, S. K., Dalai, S. S., & Tripathy, J. (2024, November). Cryptocurrency Price Prediction: A Comprehensive Analysis of High-dimensional Features with Deep Learning Techniques. In 2024 International Conference on Intelligent Computing and Sustainable Innovations in Technology (IC-SIT) (pp. 1-6). IEEE.

Rana, M. S., Chouksey, A., Hossain, S., Sumsuzoha, M., Bhowmik, P. K., Hossain, M., ... & Zeeshan, M. A. F. (2025). AI-Driven Predictive Modeling for Banking Customer Churn: Insights for the US Financial Sector. Journal of Ecohumanism, 4(1), 3478-3497.

Ray, R. K., Sumsuzoha, M., Faisal, M. H., Chowdhury, S. S., Rahman, Z., Hossain, E., ... & Rahman, M. S. (2025). Harnessing Machine Learning and AI to Analyze the Impact of Digital Finance on Urban Economic Resilience in the USA. Journal of Ecohumanism, 4(2), 1417-1442.

Rahman, A., Debnath, P., Ahmed, A., Dalim, H. M., Karmakar, M., Sumon, M. F. I., & Khan, M. A. (2024). Machine learning and network analysis for financial crime detection: Mapping and identifying illicit transaction patterns in global black money transactions. Gulf Journal of Advance Business Research, 2(6), 250-272.

Sabeena, J., & Sagar, P. L. (2023, November). Enhancing Predictive Accuracy for Real Time Cryptocurrency Market Prices with Machine Learning Techniques. In 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA) (pp. 1288-1292). IEEE.

Shamshad, H., Ullah, F., Ullah, A., Kebande, V. R., Ullah, S., & Al-Dhaqm, A. (2023). Forecasting and Trading of the Stable Cryptocurrencies With Machine Learning and Deep Learning Algorithms for Market Conditions. IEEE Access, 11, 122205-122220.

Sizan, M. M. H., Chouksey, A., Miah, M. N. I., Pant, L., Ridoy, M. H., Sayeed, A. A., & Khan, M. T. (2025). Bankruptcy Prediction for US Businesses: Leveraging Machine Learning for Financial Stability. Journal of Business and Management Studies, 7(1), 01-14.

Sizan, M. M. H., Das, B. C., Shawon, R. E. R., Rana, M. S., Al Montaser, M. A., Chouksey, A., & Pant, L. (2023). AI-Enhanced Stock Market Prediction: Evaluating Machine Learning Models for Financial Forecasting in the USA. Journal of Business and Management Studies, 5(4), 152-166.

Sizan, M. M. H. ., Chouksey, A. ., Tannier, N. R. ., Jobaer, M. A. A., Akter, J. ., Roy, A. ., Ridoy, M. H. ., Sartaz, M. S. ., & Islam, D. A. (2025). Advanced Machine Learning Approaches for Credit Card Fraud Detection in the USA: A Comprehensive Analysis. Journal of Ecohumanism, 4(2), 883 –. https://doi.org/10.62754/joe.v4i2.6377

Sumsuzoha, M., Rana, M. S., Islam, M. S., Rahman, M. K., Karmakar, M., Hossain, M. S., & Shawon, R. E. R. (2024). LEVERAGING MACHINE LEARNING FOR RESOURCE OPTIMIZATION IN USA DATA CENTERS: A FOCUS ON INCOMPLETE DATA AND BUSINESS DEVELOPMENT. The American Journal of Engineering and Technology, 6(12), 119-140.

Sun, G. (2024). Cryptocurrency price prediction based on Xgboost, LightGBM and BNN. Applied and Computational Engineering, 49, 273-279.

Zhao, F., Zhang, M., Zhou, S., & Lou, Q. (2024). Application of Deep Reinforcement Learning for Cryptocurrency Market Trend Forecasting and Risk Management. Journal of Industrial Engineering and Applied Science, 2(5), 48-55.