femington

Int. J. Adapt. Manag. Bus. Intell. (2026) 1(1):1–10
Article ID: IJAMBI-2026-270

Data-Driven Algorithmic Trading with Market Sentiment Insights

Onyeka Alimele
onyeka.alimele@flatbookllc.com
Independent Researcher

Abstract

Cryptocurrency markets are experiencing exponential growth, creating the need for a good predictive model. In particular, this model must be able to capture the volatility and highly non-linear behaviour of crypto prices. The present study suggests a framework that is fusion-based and is hybrid in nature, using both the indicators of technical analysis and the various classical supervised ML algorithms, which help in the prediction of the price of cryptocurrency. The method detailed in this paper uses historical OHLC data for five cryptocurrencies: Ripple (XRP), Ethereum (ETH), Litecoin (LTC), Cardano (ADA), and Polkadot (DOT). The study makes use of accepted technical indicators that comprise the relative strength index (RSI), moving average convergence divergence (MACD), and Bollinger bands to perform feature engineering with these indicators, later used as the input for SVM, decision tree (DT), and random forest (RF) models. The models were trained featuring a validation procedure accounting for time series effects. The mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). The existing experimental evidence suggests that Random Forest has less prediction error and better generalization ability than SVM and Decision Tree. A combination of indicators using ensemble learning can improve the forecasting accuracy of the price of cryptocurrency.

Keywords

• Cryptocurrency Prediction • Technical Indicators • Machine Learning • Random Forest • Time Series Forecasting • Algorithmic Trading

1. Introduction

The global rise of cryptocurrency markets has introduced a dynamic and highly volatile financial environment, one that significantly diverges from traditional asset classes in terms of behavior and structure. Digital assets such as Bitcoin (BTC), Ethereum (ETH), and a wide array of altcoins are subject to frequent price swings, often influenced by speculative sentiment, social media discourse, regulatory developments, and macroeconomic news. These rapid and sometimes irrational fluctuations pose a considerable challenge to the development of accurate and robust trading strategies.

In traditional financial sectors, traders have long used technical analysis to determine trends, reversals and momentum shifts in asset prices. The Relative Strength Index (RSI) [1], Moving Average Convergence Divergence (MACD), and Bollinger Band are indicators based on historical price movement and volatility. Even though these tools remain in vogue, market participants primarily use them to assess linearly stationary behavior instead of representing the actual market condition. Moreover, they are usually rule-based and reactive, not predictive.

To enhance predictive capability and decision support, recent studies have begun to explore the integration of machine learning (ML) techniques with traditional technical indicators. Machine learning offers the advantage of data-driven modeling, enabling systems to learn complex relationships from historical data and generalize to unseen market conditions. Algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests can identify intricate patterns and dependencies that may not be visible through classical analytical techniques. When paired with engineered features from technical analysis, these models have the potential to deliver improved forecasting accuracy and trading performance.

The present study aims to investigate a hybrid approach that synthesizes technical indicators and machine learning algorithms to predict cryptocurrency prices. By combining well-established trading indicators (RSI, MACD, Bollinger Bands) with supervised ML classifiers (SVM, Decision Tree, Random Forest), the study seeks to evaluate how such integration impacts model effectiveness. Five widely traded cryptocurrencies Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Cardano (ADA), and Polkadot (DOT) are used as the basis for model training and testing [2, 3].

The overarching objective is to identify which algorithmic strategy most accurately captures price dynamics in this highly volatile domain. Accordingly, the study checks out all the models using multiple performance metrics such as MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), and unadjusted and adjusted R-squared. These metrics offer a multi-faceted view of both error magnitude and model fit.

The present study investigates a hybrid approach that combines technical indicators with classical supervised machine learning models for cryptocurrency price prediction. Rather than introducing a fundamentally novel architecture, the contribution lies in providing a consistent empirical comparison of widely used algorithms under a unified feature engineering pipeline. This enables clearer insights into model behavior across multiple assets and evaluation metrics[4, 5].

2. Related Work

Given the maturity of machine learning methods such as SVM, Decision Trees, and Random Forests, this section focuses only on their application within financial forecasting contexts rather than reiterating their standard theoretical formulations.

The intersection of machine learning and financial forecasting has gained considerable momentum, particularly with the rise of digital assets such as cryptocurrencies. Unlike traditional equity markets, cryptocurrencies are highly volatile, influenced by sentiment, liquidity shifts, and macroeconomic announcements, requiring more adaptive and nonlinear predictive systems. A variety of studies have attempted to address these challenges using both technical indicators and advanced learning algorithms.

Traditional technical analysis methods such as RSI, MACD, and Bollinger Bands have long been used by traders to interpret market momentum, volatility, and potential reversals. Chakrabarty and Majumdar [1] demonstrated the predictive power of RSI in short-term stock forecasting. Studied MACD's effectiveness in identifying crossover signals in emerging markets, while examining its combination with Bollinger Bands to enhance precision.

On the machine learning front, Patel et al. [6] used artificial neural networks and SVMs for Indian stock market prediction, reporting significant gains over linear regression models. Kumar and Ravi [7] presented a comprehensive survey of machine learning techniques in financial prediction, highlighting decision trees and ensemble models as particularly effective.

Random Forests have become a strong baseline due to their robustness to noise and overfitting. Dutta et al. [8] applied Random Forests for Bitcoin price prediction, outperforming traditional autoregressive models. Similarly, Jang and Lee [9] combined RNNs with market indicators for cryptocurrency trend forecasting, illustrating the benefits of hybrid inputs.

Deep learning models such as LSTM and GRU have been widely studied for time series prediction. McNally et al. [10] employed LSTMs for Bitcoin price forecasting with encouraging results. Enhanced this approach by integrating sentiment from Twitter, reinforcing the importance of exogenous signals.

Reinforcement learning has also been explored. Deng et al. [11] developed a deep reinforcement trading system using CNNs and technical features. Yang et al. [12] extended this to a portfolio optimization context, using DDPG for continuous allocation in crypto markets.

Ensemble approaches have gained popularity for their generalization strength. Proposed an ensemble of Random Forests and Gradient Boosting for stock prediction. In cryptocurrency contexts, XGBoost is combined with MACD and RSI for altcoin trading signals.

Other studies have explored hybrid pipelines. Evaluated over 100 machine learning models using technical indicators, identifying Random Forest as a consistently strong performer. Lahmiri and Bekiros [13] integrated wavelet transforms with ML models for cryptocurrency volatility forecasting.

A growing area of research also emphasizes explainability. Ribeiro et al. [14] introduced LIME, which is now applied in fintech models to interpret predictions. Lundberg and Lee [15] proposed SHAP, used to evaluate feature importance in financial time series models.

Lastly, the significance of real-time trading systems has been emphasized in studies like those by Fischer and Krauss [16], who applied deep LSTM networks to intraday S&P 500 data, and Sezer et al. [17] who benchmarked ML techniques for algorithmic trading using historical and high-frequency features.

In summary, the literature strongly supports the integration of technical analysis with machine learning, particularly ensemble and deep learning techniques. However, there remains a gap in hybrid frameworks tailored for cryptocurrency markets that combine engineered indicators with interpretable ML models an area this paper seeks to address.

3. Materials and Methods

To develop a robust and adaptable cryptocurrency price prediction system, this study adopts a hybrid methodology that integrates conventional technical indicators with supervised machine learning algorithms. The rationale behind this integration lies in leveraging the domain-proven strength of technical indicators for signal extraction, alongside the pattern recognition capabilities of machine learning models.

The dataset comprises historical price data from five widely traded cryptocurrencies: Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Cardano (ADA), and Polkadot (DOT). Each asset's historical price record includes open, high, low, close (OHLC) values, and trading volume. This raw data is preprocessed to extract engineered features derived from selected technical indicators. These features are then used as inputs to three different machine learning models for training and evaluation. The prediction task is formulated as a regression problem targeting next-day closing price. While directional classification (price up/down) is often used in trading systems, regression provides a more granular evaluation of prediction error and enables the use of multiple continuous performance metrics. However, this choice may limit direct interpretability in trading decision contexts[18].

3.1 Technical Indicators

The selected indicators represent momentum, trend, and volatility dimensions three essential components of technical analysis.

- **Relative Strength Index (RSI):** The RSI is a momentum oscillator that measures the rate of change in prices. It involves comparing the baseline period with the preceding 14 periods after getting the value of both gain and loss. The RSI reading moves in the range of 0 and 100 with the 70 and 30 levels generally indicating overbought and oversold zones respectively. [1].
- **Moving Average Convergence MACD** Trend-following framework that uncovers changes in an asset's price strength, momentum and direction. The calculation of the EMA difference takes place between a short-term (usually 12-period) and a long-term (usually 26-period). The EMA signal line is the 9 EMAL MACD line. Signals result from crossovers.
- **Bollinger Bands:** Bollinger Bands are made of a moving average that is continuously updated, with a standard deviation line above and below. This indicator measures volatility and identifies reversal zones. When price moves outside the bounds, it indicates excessive behavior which leads to mean reversion.

3.2 Machine Learning Algorithms

Following the computation of technical indicator values for each asset and time point, the resulting feature vectors are fed into three widely used supervised machine learning algorithms, each offering distinct modeling characteristics.

- **Support Vector Machine (SVM):** SVMs are effective in classifying as well as in regression problems. SVMs do this by finding the maximum dividing hyperplane that separates the data points of different classes with the largest hat width possible. The SVM is capable of employing kernel functions like an RBF (Radial Basis Function) or polynomial kernels suitable for nonlinear decision boundaries [6].
- **Decision Tree (DT):** Decision Trees are rule-oriented models that help make decisions based on splitting the data set into sub-groups based on the values of the input features. A split is basically the condition based on an attribute. That attribute separates the target best. Because of their simplicity and low computational cost, linear regression models make for a strong baseline. However, they are vulnerable to overfitting in noisy datasets like cryptocurrencies, which are often random walks[7].
- **Random Forest (RF):** The random forest is an ensemble of decision trees that average to fight overfitting. Every tree is train on a bootstrapped copy of the data. Every tree operates independently, and the final output is generated through either majority voting for classification or averaging for regression. The robustness of Random Forest makes it very efficient for financial time series with noise and non-linearity[8].

Each model is trained on the same input feature set, allowing for fair and consistent comparison across prediction algorithms. The target variable is the next-day closing price, and the models are optimized using standard performance metrics detailed in the subsequent section.

4. Implementation

The implementation of the proposed hybrid cryptocurrency prediction system was carried out in Python 3.10, using data science libraries such as Pandas, NumPy, scikit-learn, and TA-Lib for technical indicators. The development pipeline comprises four main stages: data acquisition, feature engineering, model training, and evaluation.

4.1 Data Collection and Preprocessing

For five cryptocurrencies, the Yahoo Finance API was utilized to extract historical price and volume data. The five cryptocurrencies include Ethereum, Ripple, Litecoin, Cardano and Polkadot. The data covered the period from January 1st, 2019, to December 31st, 2023, with daily sampling. While forward-fill imputation ensures continuity in the time series, it may artificially smooth price movements and suppress short-term volatility, potentially impacting indicators sensitive to variance such as Bollinger Bands[10, 17].

4.2 Feature Engineering via Technical Indicators

For each asset, three widely used technical indicators were computed:

- **RSI (14-day)**: For overbought/oversold signals
- **MACD (12, 26, 9)**: For trend convergence/divergence
- **Bollinger Bands (20, 2)**: For volatility boundaries

These indicators were appended to the feature matrix. All features were then normalized using Min-Max scaling between 0 and 1.

4.3 Model Architecture and Training

The scikit-learn package was used to implement the Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) models. With regard to the technical features of the current day, each model was trained to predict the closing price of the next day.

Table 1: Hyperparameter Settings for ML Models

Model	**Key Parameters**	**Tuned Values**
SVM	Kernel, C	RBF, 1.0
Decision Tree	Max Depth	10
Random Forest	Estimators, Max Depth	100, 15

Hyperparameters were optimized using a time-series aware validation strategy based on expanding window splits rather than standard k-fold cross-validation, thereby preserving temporal ordering and preventing data leakage [6, 8].

4.4 Evaluation Metrics

To assess the predictive performance, the following metrics were computed:

- **Average Squared Errors**

- **Root Mean Square Error (RMSE)**
- **Average of Absolute Errors**
- **Average of Absolute Percentage Error**
- **R^2 and Adjusted R^2**

4.5 System Pipeline Diagram

Figure illustrates the end-to-end workflow of the hybrid implementation.

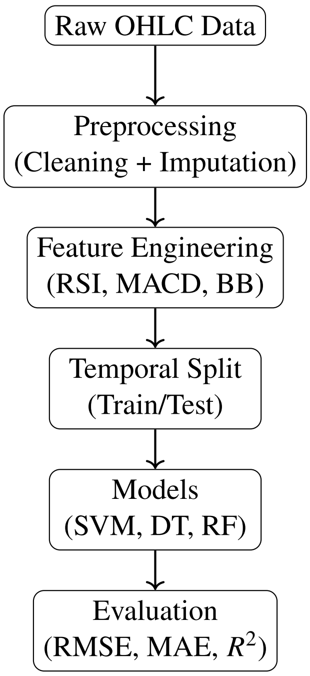

Figure 1: Enhanced pipeline showing temporal validation and evaluation flow

5. Results

This section focus on the empirical evaluation of the hybrid models for Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Cardano (ADA), and Polkadot (DOT). This project aimed to evaluate the prediction performance of SVM, DT, and RF algorithms using the same set of features obtained from technical indicators RSI, MACD, and Bollinger Bands.

5.1 Experimental Setup

The statistical models were tested on daily price data over a long time. A chronological train-test split of 80:20 was incorporated to maintain temporal ordering. All features were scaled using Min-Max normalization. All of the model performance was validated using the same input matrix. The next-day closing price of the asset is the target variable [17].

5.2 Evaluation Metrics

The following standard regression metrics were used:

- **Mean Squared Error (MSE) penalises large prediction errors.**
- **RMSE measures the erroR in the unit scale of price, so it is interpretable.**
- **The mean absolute error is the average absolute deviation.**
- **Analysis of MAPE with respect to the ongoing Indian Economy Essay.**
- **R^2 and Adjusted R^2 Indicate the proportion of variance explained.**

5.3 Ethereum Case Study

Table 2 summarizes the results for Ethereum (ETH), which showed consistent outperformance by the Random Forest model across all evaluation metrics.

Table 2: ETH Model Performance Summary

Model	MSE	RMSE	MAE	MAPE	R^2	Adj.
DT	6537	80.85	59.37	2.21%	0.9882	0.9881
RF	**2275**	**47.71**	**34.45**	**1.25%**	**0.9959**	**0.9958**
SVM	1.59M	1261.75	1048.73	35.75%	-1.86	-1.91

The Random Forest model achieved an RMSE of 47.71 and a high R^2 of 0.9959, indicating both accuracy and strong generalization. In contrast, the SVM model underperformed severely, likely due to its sensitivity to non-linearities and scale in this domain.

5.4 Performance Across Other Assets

This underperformance may be attributed to the choice of kernel and sensitivity to feature scaling, as well as the difficulty of SVMs in capturing highly nonlinear and noisy patterns typical of cryptocurrency markets without extensive hyperparameter tuning.

Similar trends were observed across other cryptocurrencies. In general:

- **Random Forest** consistently achieved the lowest prediction errors and highest R^2 values.
- **Decision Trees** offered decent performance with lower complexity.
- **SVM** struggled, especially for assets with higher volatility like DOT and XRP.

5.5 Overall Observations

The results confirm that ensemble-based methods like Random Forest are more robust in handling the non-linear, noisy, and high-variance nature of cryptocurrency markets. The superior performance of Random Forest can be attributed to its ability to average over multiple decision paths, which helps in reducing overfitting and enhancing generalization [8].

Additional result tables for XRP, LTC, ADA, and DOT can be provided in the appendix or supplementary figures, depending on space constraints.

Future evaluations should include repeated experiments across multiple rolling windows and report variability measures such as standard deviation to better assess model stability across different market conditions.

6. Conclusion

This research demonstrates a hybrid predictive modelling framework that integrates traditional guidelines on technical analysis with state-of-the-art machine learning algorithms on price prediction in the cryptocurrency market. Given the high volatility and non-linearity of digital assets, the inclusion of domain problems like RSI, MACD and Bollinger Bands with learning-based techniques represents a step towards algorithm-based trading[1, 6].

Evaluations on five representative cryptocurrencies Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Cardano (ADA) and Polkadot (DOT), state that it works. Among the tested machine learning models, the Random Forest performed better than other models i.e. Decision Tree and SVM. It had better generalization and accuracy. The findings imply that ensemble-based techniques are particularly effective in capturing the randomness and high volatility inherent in cryptocurrency price fluctuations.

The results further validated the utility of technical indicators as meaningful features when appropriately engineered and coupled with robust predictive models. The success of the hybrid strategy highlights the synergy between human-devised heuristics (technical indicators) and automated pattern discovery (machine learning), enabling more informed and adaptable forecasting systems.

While the Random Forest model proved highly accurate in backtesting scenarios, it is important to acknowledge that real-world deployment would require considerations such as latency, market impact, slippage, and real-time data streaming capabilities. Nonetheless, the strong offline performance offers a reliable foundation for future integration into live trading environments.

In summary, the proposed hybrid approach not only improves forecast accuracy but also provides a scalable and extensible architecture for financial prediction. Its modular nature allows for easy adaptation to other asset classes or additional feature sets, making it a valuable contribution to the growing field of AI-driven financial analytics [8, 13].

7. Future Work

Although the proposed hybrid modeling framework has demonstrated strong predictive capabilities in backtesting scenarios, several directions remain open for further research and enhancement.

The current feature set is limited to three well-known technical indicators: RSI, MACD, and Bollinger Bands. Future work could explore the inclusion of additional indicators such as Moving Averages, Fibonacci Retracement Levels, or On-Balance Volume (OBV) to further diversify the feature space. Moreover, combining these with fundamental indicators (e.g., blockchain activity, trading volume anomalies, or token issuance events) could provide deeper insights into market behavior.

This research makes use of classical machine learning models. As deep learning progresses, many architectures, including Long Short-Term Memory (LSTM), Temporal Convolution Network (TCN) and transformer-based time series ones, could be experimented with for their capacity to capture long-range temporal dependencies and complex price dynamics. Combining ensemble learning with deep learning can contribute to great improvement in accuracy and robustness. Using hybrid models such as stacking or boosting over the LSTM layers [16, 17].

The current pipeline operates in an offline mode, using historical data to make predictions on static test sets. Real-time prediction systems, capable of ingesting live price feeds and updating models incrementally, would be highly valuable for deployment in algorithmic trading platforms. Implementing such a real-time system would require the integration of streaming data APIs, latency-aware execution engines, and live risk management modules.

Furthermore, sentiment analysis derived from news articles, tweets, Reddit discussions, and other unstructured sources could be incorporated using Natural Language Processing (NLP) techniques. These sentiment features may serve as complementary inputs, especially during macroeconomic events or social media-driven volatility spikes [10].

Using explainable AI techniques can improve model interpretation and understanding. The use of SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-Agnostic Explanations) allows a trader to understand the reason behind the prediction, increasing trust in automated systems.

Finally, while this study focused on five cryptocurrencies, future work may involve extending the model to a broader universe of digital assets, including small-cap altcoins and tokenized derivatives. Incorporating portfolio-level objectives such as risk-adjusted returns, drawdown minimization, or Sharpe ratio maximization within the model's reward function would further align predictions with practical trading goals[14, 15].

In conclusion, the hybrid approach provides a strong foundation, and future work focused on data diversity, model complexity, real-time deployment, and interpretability can further unlock its full potential for financial forecasting in dynamic crypto markets [11, 12].

References

[1] A. Chakrabarty and A. Majumdar. The impact of directional indicators to predict the stock movement. 2024.

[2] Erdinc Akyildirim, Ahmet Goncu, and Ahmet Sensoy. Prediction of cryptocurrency returns using machine learning. *Annals of Operations Research*, 297(1):3–36, 2021.

[3] Ayman M Khedr, Iqra Arif, Mohammed El-Bannany, Saadat M Alhashmi, and Madhavan Sreedharan. Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. *Intelligent Systems in Accounting, Finance and Management*, 28(1):3–34, 2021.

[4] HA Salman, Ali Kalakech, and Ali Steiti. Random forest algorithm overview. *Babylonian Journal of Machine Learning*, 2024:69–79, 2024.

[5] Jun Hu and Silke Szymczak. A review on longitudinal data analysis with random forest. *Briefings in bioinformatics*, 24(2):bbad002, 2023.

[6] J. et al. Patel. Predicting stock and stock price index movement using machine learning techniques. *Expert Systems with Applications*, 42(1):2162–2172, 2015.

[7] M. Kumar and V. Ravi. A survey of the applications of text mining in financial domain. *Knowledge-Based Systems*, 114:128–147, 2016.

[8] A. et al. Dutta. Bitcoin price prediction using machine learning: An approach to sample variance. *Procedia Computer Science*, 167:2260–2268, 2020.

[9] H. Jang and J. Lee. An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information. *IEEE Access*, 6:5427–5437, 2018.

[10] S. et al. McNally. Predicting the price of bitcoin using machine learning. 2018.

[11] Y. et al. Deng. Deep learning for financial time series: Deep learning for algorithmic trading. In *IEEE Computational Intelligence for Financial Engineering*, 2017.

[12] H. et al. Yang. Deep reinforcement learning for portfolio management. In *Proceedings of IJCAI*, 2020.

[13] S. Lahmiri and S. Bekiros. Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons Fractals*, 118:35–40, 2019.

[14] M. T. et al. Ribeiro. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD*, pages 1135–1144, 2016.

[15] S. Lundberg and S. Lee. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 2017.

[16] T. Fischer and C. Krauss. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2):654–669, 2018.

[17] O. B. et al. Sezer. Financial time series forecasting with deep learning: A systematic literature review. *Applied Soft Computing*, 90:106181, 2020.

[18] M Poongodi, Tan N Nguyen, Mounir Hamdi, and Korhan Cengiz. Global cryptocurrency trend prediction using social media. *Information Processing & Management*, 58(6):102708, 2021.