

Forecasting cryptocurrency prices is challenging due to its extreme volatility. The unpredictability of crypto markets differs substantially from traditional stock markets for which analysts often utilize earning reports, macroeconomic trends, and regulatory changes. In contrast, cryptocurrency prices frequently correlate more with social media sentiment and influencer opinions than traditional markets.

This project aims to tackle these challenges by examining a hybrid approach that combines traditional statistical models like ARIMA and XGBoost with deep learning techniques such as LSTM to enhance predictive accuracy.

PROJECT AIM: Construct and test the predictive accuracy of a hybrid model that integrates historical patterns, sequential dependencies, and sentiment analysis.

Hypothesis: A hybrid model that integrates historical patterns, sequential dependencies, and sentiment analysis will have a higher predictive accuracy than traditional standalone models for predicting cryptocurrency volatility.

Approach: We will combine and test the predictive accuracy of a hybrid model that incorporates ARIMA and XGBoost's ability to capture historical cryptocurrency patterns with LSTM to identify sequential dependencies (and complex patterns) and the CryptTop12 dataset to integrate sentiment analysis from social media. Utilizing open-high-low-closing (OHLC) datasets from Kaggle and Binance, we will conduct extensive data exploration, cleansing and feature engineering, and incorporate indicators such as moving averages, relative strength index (RSI), and daily return percentage.

By assessing the effectiveness of hybrid models in predicting cryptocurrency prices and specifically examining the role of sentiment analysis in enhancing forecasting, we aim to create a mechanism to account for both historical data and sentiment driven market trends. Accordingly, the results of this research will have immediate applicability in this rapidly evolving area of finance.