Abstract

Bitcoin, as the most dominant cryptocurrency asset by market capitalization, is characterized by high price volatility and complex movement patterns posing significant challenges to the accuracy of conventional predictive models. This study aims to develop and evaluate hybrid ARIMA–LSTM and ARIMAX–LSTM models to improve the prediction accuracy of Bitcoin prices. The hybrid model combines the strength of ARIMA or ARIMAX in capturing short-term linear trends with the ability of LSTM to model nonlinear patterns and long-term temporal dependencies. The input variables used include daily High, Low, and Volume data, which represent market volatility. In addition, the study analyzes Bitcoin's volatility characteristics using quantitative approaches such as standard deviation, historical volatility, and rolling volatility, while also considering the impact of major events such as halving cycles and ETF launches. Daily price data from 2015 to 2024 are used for model training and testing, with two data-splitting strategies: two-stage split and single split. The models are evaluated using three primary metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). The results show that the Hybrid ARIMA–LSTM model with the two-stage split strategy performs best overall, achieving very low prediction errors: MAE of 2,216.81, RMSE of 2,984.18, and MAPE of 3.17%. Although the ARIMAX–LSTM model yields comparable MAE and RMSE values, its significantly higher MAPE indicates disproportionate errors, particularly at higher price levels. This research contributes both theoretically to the development of predictive models for complex financial data, and practically to enhancing investment decision-making in the dynamic and uncertain cryptocurrency market.

Keywords: *ARIMA, ARIMAX, Bitcoin, LSTM, Hybrid Model, Volatility, Price Prediction.*