ABSTRACT

Bitcoin cryptocurrency is one type of cryptocurrency that has been widely used, because its transaction process is safe, easy, and fast and has the potential to gain significant profits in the future. However, bitcoin transactions are a risky activity, because cryptocurrency prices continue to fluctuate every day which is not only influenced by internal factors. External factors such as public sentiment and Google search index also affect bitcoin price fluctuations. Therefore, predictions are needed to estimate bitcoin prices by integrating sentiment variables and Google Trends Index (GTI) into a prediction model using regression techniques that are adjusted to data characteristics and classical assumption tests. This study aims to compare the XGBoost Regression and LSTM for Regression algorithms in predicting bitcoin closing prices. Both algorithms were chosen because each has advantages and disadvantages in handling the characteristics of the research data, where the data used is non-linear, not normally distributed, and contains time series elements. After the evaluation, the results showed that the LSTM for Regression Scenario 2 model with price and sentiment variables had the smallest RMSE value and the highest R2-Score of 1338.989 and 92.3%, respectively. Based on the learning curve, the model is not indicated to be overfit as evidenced by the stability of the curve at a low loss value when reaching epoch = 20. Overall, historical price variables have a significant influence in predicting bitcoin closing prices, with Open as the most dominant feature identified using the SHAP method.

Keywords: *Bitcoin Price Prediction, Sentiment, Google Trends Index, LSTM, XGBoost Regression*