ABSTRACT

Stock market fluctuations are influenced not only by numerical indicators but also by investor sentiment and perceptions shaped by economic behavior. However, the integration of news sentiment through sentiment analysis features into stock forecasting remains limited, with the use of sentiment probabilities particularly underexplored despite their potential to enhance prediction performance. To develop a more robust and precise forecasting framework, this study integrates sentiment probabilities from Indonesian financial news titles into deep learning models, focusing on the LQ45 stock index from 1 January 2024 to 31 March 2025, covering 297 open market days. Sentiment features such as probabilities for positive, negative, and neutral classes as well as their corresponding hard labels were extracted from 158,514 Indonesian financial news titles using a fine-tuned IndoBERT model. These sentiment features were combined with the historical close index of LQ45 and configured with lag and time step settings to form the input of the LSTM forecasting model. Two sentiment integration methods were compared: direct input of sentiment probabilities from IndoBERT's softmax output and aggregated sentiment scores derived from either probability or hard labels. The results show that integrating sentiment features improves the model's predictive performance. Notably, Scenario 5, which used aggregated sentiment scores based on probabilities, achieved the best results with an MAE of 14.16, RMSE of 17.95, and MAPE of 1.88%. This also significantly outperformed the aggregated sentiment score derived from hard labels.

Keywords: LQ45 stock index prediction, sentiment analysis, LSTM, Indonesian financial news