

ABSTRACT

Accurate stock price prediction remains a major challenge in investment decision-making, especially when models rely solely on historical data without considering market sentiment influenced by financial news. On the other hand, sentiment analysis models such as Financial RoBERTa have demonstrated superior performance in understanding financial language context, yet are insufficient on their own without incorporating technical factors like price trends. To address this gap, this study proposes a hybrid modeling approach by integrating Financial RoBERTa for sentiment analysis of financial news and Gated Recurrent Unit (GRU) for forecasting based on historical data. The research focuses on three major banking stocks in Indonesia namely BBRI, BMRI and BBCA, using sentiment data sourced from CNBC Indonesia and Bisnis.com. Four input scenarios are evaluated: historical data only, historical + CNBC sentiment, historical + Bisnis.com sentiment, and a combination of all inputs. Results show that all sentiment-integrated scenarios reduced prediction error compared to using historical data alone. The Bisnis.com sentiment scenario consistently delivered the best performance, achieving RMSE, MAE, and MAPE of 14.07, 19.72, and 0.41 respectively on BBRI, and RMSE 33.50 with MAPE 0.39 on BMRI. These findings emphasize that integrating quantitative and qualitative data enhances stock price prediction accuracy and offers a promising foundation for developing adaptive predictive systems for Indonesia's financial markets.

Keywords: *deep learning, Financial RoBERTa, GRU, sentiment analysis, stock price forecasting*

ABSTRAK

Akurasi prediksi harga saham masih menjadi tantangan dalam pengambilan keputusan investasi, terutama ketika model hanya mengandalkan data historis tanpa mempertimbangkan sentimen pasar yang dipengaruhi oleh berita keuangan. Di sisi lain, model analisis sentimen seperti Financial RoBERTa telah terbukti unggul dalam memahami konteks bahasa finansial, namun tidak cukup bila digunakan tanpa mempertimbangkan faktor teknikal seperti tren harga. Untuk mengatasi keterbatasan ini, penelitian ini mengusulkan pendekatan hybrid dengan mengintegrasikan Financial RoBERTa untuk analisis sentimen dari berita keuangan dan Gated Recurrent Unit (GRU) untuk forecasting berbasis data historis. Data dikumpulkan dari saham sektor perbankan Indonesia yaitu BBRI, BMRI, dan BBCA dengan sumber sentimen dari CNBC Indonesia dan Bisnis.com. Empat skenario input diuji: hanya data historis, historis + CNBC, historis + Bisnis.com, dan kombinasi semua. Hasil menunjukkan bahwa seluruh skenario yang melibatkan sentimen mampu menurunkan error dibandingkan penggunaan data historis saja. Skenario dengan sentimen Bisnis.com memberikan hasil terbaik, dengan nilai RMSE, MAE, dan MAPE masing-masing sebesar 14.07, 19.72, dan 0.41 pada saham BBRI serta RMSE 33.50 dan MAPE 0.39 pada saham BMRI. Temuan ini menegaskan bahwa integrasi data kuantitatif dan kualitatif mampu meningkatkan akurasi prediksi saham dan dapat menjadi fondasi sistem prediktif yang adaptif terhadap dinamika pasar keuangan Indonesia.

Kata kunci: Analisis Sentimen, *Deep learning*, *Financial RoBERTa*, GRU, Peramalan Harga Saham