

## ABSTRACT

### DAILY STOCK CLOSING PRICE PREDICTION USING A COMBINATION OF DEEP LEARNING MODELS AND NEWS SENTIMENT FEATURES BASED ON FINBERT

Abid Nujaiba

21/478120/PA/20734

Stock price prediction is a complex challenge influenced by various factors, both technical and non-technical, such as market sentiment and external news. This study aims to analyze the impact of incorporating sentiment features extracted using *FinBERT* on the performance of daily stock price prediction models. The proposed architecture, called the *Enhanced Hybrid* model, integrates three main components: *Bidirectional Long Short-Term Memory* (BiLSTM), *Modified Transformer* (MTRAN), and *Temporal Convolutional Network* (TCN). This approach adopts an *early fusion* mechanism, combining technical features derived from historical stock price data with daily sentiment scores extracted from English-language financial news and tweets.

Experiments were conducted on four major stocks, namely Apple Inc. (AAPL), Alphabet Inc. (GOOG), Microsoft Corp. (MSFT), and Taiwan Semiconductor Manufacturing Company (TSM). Technical data were obtained from Yahoo Finance, while sentiment scores were generated using *FinBERT* to capture positive, negative, and neutral tones. The models were evaluated using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination ( $R^2$ ) as performance metrics.

The results indicate that the *Enhanced Hybrid* model consistently outperforms the baseline *BiLSTM* model. The integration of sentiment features significantly improves prediction accuracy, though the degree of improvement varies among stocks—highest for TSM (32.4%) and lowest for MSFT (9.3%). This variation is influenced by the completeness of sentiment data and the characteristics of each stock. Stocks that are more sensitive to external factors, such as TSM and GOOG, exhibit greater improvement, while more stable stocks like MSFT show limited influence from sentiment information.

In conclusion, integrating technical and sentiment features through a hybrid deep learning architecture produces more accurate and stable stock price predictions compared to single-model approaches. However, the effectiveness of sentiment features is contextual, being more relevant for stocks with high volatility and extensive media exposure. This approach can serve as a foundation for developing financial *decision support systems* that combine quantitative and textual analyses to generate more comprehensive market insights.

**Keywords:** stock price prediction, deep learning, BiLSTM, MTRAN, TCN, sentiment analysis, FinBERT, tweet sentiment, MAE, RMSE,  $R^2$ .

## INTISARI

### PREDIKSI HARGA PENUTUPAN SAHAM HARIAN MENGGUNAKAN KOMBINASI MODEL DEEP LEARNING DAN FITUR SENTIMEN BERITA BERBASIS FINBERT

Abid Nujaiba

21/478120/PA/20734

Prediksi harga saham merupakan tantangan yang kompleks karena dipengaruhi oleh banyak faktor, baik teknikal maupun non-teknikal seperti sentimen pasar dan berita eksternal. Penelitian ini bertujuan untuk menganalisis pengaruh penambahan fitur sentimen berbasis *FinBERT* terhadap performa model prediksi harga saham harian. Arsitektur yang digunakan adalah model *Enhanced Hybrid* yang menggabungkan tiga komponen utama, yaitu *Bidirectional Long Short-Term Memory* (BiLSTM), *Modified Transformer* (MTRAN), dan *Temporal Convolutional Network* (TCN). Pendekatan ini diimplementasikan menggunakan metode *early fusion*, di mana fitur teknikal dari data historis harga saham digabungkan dengan skor sentimen harian yang diekstraksi dari tweet berbahasa Inggris.

Eksperimen dilakukan terhadap empat saham utama, yaitu Apple Inc. (AAPL), Alphabet Inc. (GOOG), Microsoft Corp. (MSFT), dan Taiwan Semiconductor Manufacturing Company (TSM). Data teknikal diperoleh dari Yahoo Finance, sedangkan data sentimen diekstraksi menggunakan *FinBERT* untuk menghasilkan skor positif, negatif, dan netral yang dinormalisasi. Model kemudian dilatih dan diuji menggunakan metrik evaluasi Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan koefisien determinasi ( $R^2$ ).

Hasil penelitian menunjukkan bahwa model *Enhanced Hybrid* secara konsisten menghasilkan performa terbaik dibandingkan model *BiLSTM* sederhana. Integrasi fitur sentimen terbukti meningkatkan akurasi prediksi dengan variasi peningkatan antar saham, paling besar pada TSM (32,4%) dan paling kecil pada MSFT (9,3%). Perbedaan kontribusi ini dipengaruhi oleh tingkat kelengkapan data sentimen dan karakteristik masing-masing saham. Saham yang lebih sensitif terhadap faktor eksternal, seperti TSM dan GOOG, memperoleh peningkatan performa yang lebih signifikan, sementara saham yang lebih stabil seperti MSFT menunjukkan pengaruh sentimen yang terbatas.

Kesimpulannya, integrasi data teknikal dan sentimen melalui arsitektur hybrid memberikan hasil prediksi yang lebih akurat dan stabil dibandingkan model tunggal. Namun, efektivitas fitur sentimen bersifat kontekstual dan lebih relevan digunakan pada saham dengan volatilitas tinggi serta eksposur media yang luas. Pendekatan ini diharapkan dapat menjadi dasar bagi pengembangan sistem pendukung keputusan (*decision support system*) di bidang keuangan yang menggabungkan analisis kuantitatif dan tekstual untuk menghasilkan wawasan yang lebih komprehensif.

**Kata Kunci:** prediksi harga saham, *deep learning*, *BiLSTM*, MTRAN, TCN, analisis sentimen, *FinBERT*, sentimen *tweet*, MAE, RMSE,  $R^2$ .