

STOCK PRICE PREDICTION USING A HYBRID CNN-LSTM MODEL WITH SENTIMENT DATA INTEGRATION FROM GOOGLE NEWS

(Case Study: Tesla Stock)

Demas Risdho Listianto

Data Science Study Program, Faculty of Science and Technology

University of Technology Yogyakarta

Jl. Ringroad Utara Jombor Sleman Yogyakarta

E-mail: demasrisdho11@gmail.com

ABSTRACT

This study aims to forecast Tesla's stock price (TSLA) using a hybrid CNN-LSTM model integrated with financial news sentiment analysis. The dataset includes historical TSLA stock prices sourced from Yahoo Finance and over 100,000 news articles collected from Google News between 2020 and 2025. Sentiment analysis was conducted using the VADER algorithm to classify articles as positive, neutral, or negative. These sentiment scores were then integrated with the stock data and engineered into time-series features, including lag variables. The Convolutional Neural Network (CNN) component was used to extract local spatial patterns from the input data, while the Long Short-Term Memory (LSTM) component captured long-term temporal dependencies. The model was trained and evaluated across eight experimental scenarios to assess the impact of combining stock price data with sentiment information on prediction accuracy. Model performance was measured using RMSE, MAE, and MAPE metrics. Results indicate that incorporating sentiment data significantly reduces prediction error compared to using stock price data alone. The best performance was achieved in the Lag-7 configuration, with MAE, RMSE, and MAPE values of 14.9379, 10.7227, and 4.05%, respectively, substantially outperforming previous benchmark studies in terms of prediction error.

Keywords: CNN, LSTM, TSLA, VADER, Hybrid Model, Sentiment.