ABSTRAK

Prediksi tren harga saham merupakan tantangan penting dalam pengambilan keputusan investasi. Penelitian ini bertujuan untuk mengklasifikasikan tren harga saham menggunakan pendekatan *Multivariate Time Series Classification* berbasis arsitektur CNN-LSTM. Dataset yang digunakan adalah data harga saham Google (Alphabet Inc.). Label tren diklasifikasikan dengan dua pendekatan ambang batas , yaitu metode threshold berbasis persentil dan metode threshold absolut. Pengujian dilakukan dengan berbagai konfigurasi *window size* dan *window_out* guna menganalisis pengaruh panjang data historis serta horizon prediksi terhadap akurasi klasifikasi. Selain itu, penelitian ini membandingkan performa model LSTM dan CNN-LSTM untuk mengukur kontribusi tambahan CNN dalam menangkap pola spasial. Proses seleksi fitur menggunakan metode ANOVA juga diterapkan untuk mengevaluasi dampaknya terhadap efisiensi dan akurasi model. Hasil penelitian menunjukkan bahwa *window size* 35 dan *window_out* 1 memberikan akurasi tertinggi sebesar 0.8787, serta menunjukkan stabilitas yang baik dalam prediksi multi-hari. Window di atas 35 tidak diuji lebih lanjut karena peningkatan akurasi yang relatif kecil dan beban komputasi yang meningkat. Model LSTM murni mencapai akurasi tertinggi sebesar 0.9168, sementara CNN-LSTM mencatat 0.9034, mengindikasikan bahwa penambahan CNN tidak selalu memberikan peningkatan signifikan. Penerapan feature selection berhasil meningkatkan efisiensi tanpa mengorbankan akurasi. Temuan ini diharapkan dapat menjadi dasar dalam pengembangan sistem prediksi tren saham yang lebih akurat dan efisien.

Kata kunci : CNN-LSTM, LSTM, Saham

ABSTRACT

Stock price trend prediction is a significant challenge in investment decision-making. This study aims to identify stock price trends using a Multivariate Time Series Classification approach based on the CNN-LSTM architecture. The dataset used consists of Google (Alphabet Inc.) stock price data. Trend labels were classified using two thresholding approaches: a percentile-based threshold and an absolute threshold method. Experiments were conducted with various configurations of window size and window_out to analyze the impact of historical data length and prediction horizon on classification accuracy. Additionally, this study compares the performance of LSTM and CNN-LSTM models to evaluate the added contribution of CNN in capturing spatial patterns. Feature selection using the ANOVA method was also applied to assess its impact on model efficiency and accuracy. The results show that a window size of 35 and window_out of 1 achieved the highest accuracy of 0.8787 and demonstrated good stability in multi-day predictions. Window sizes above 35 were not further tested due to relatively small accuracy gains and increased computational costs. The pure LSTM model reached the highest accuracy of 0.9168, while the CNN-LSTM model recorded 0.9034, indicating that the addition of CNN does not always lead to significant improvement. The application of feature selection successfully enhanced efficiency without sacrificing accuracy. These findings are expected to serve as a foundation for developing more accurate and efficient stock trend prediction systems.

Keyword : LSTM, CNN-LSTM, Stock