

## ABSTRACT

The tourist industry is a significant sector that has the potential to contribute to a country's gross domestic product (GDP). Notwithstanding its enormous significance, the tourism sector is highly vulnerable, especially in the wake of the COVID-19 pandemic. To facilitate the government's formulation of an effective and sustainable tourism policy, it is imperative to obtain precise estimations. According to the data provided by the Badan Pusat Statistik (BPS), there has been a significant decrease in the number of international tourist arrivals in Indonesia. The presence of this data outlier poses a significant barrier to accurately projecting tourist arrivals. The proposed method for forecasting, known as CNN-LSTM deep learning, aims to address the challenges associated with limited data sets and quickly decreasing data. The CNN-LSTM model suggested in this study comprises two primary components: The initial component comprises convolutional and pooling layers, which execute intricate mathematical processes to extract the characteristics of the input data. The subsequent component leverages the features provided by the LSTM and dense layers. Furthermore, to enhance accuracy, the inclusion of exogenous data was implemented through the utilization of the Google Trend indexes. The outcome, the CNN-LSTM approach, in conjunction with the exogenous variable of the Google Trend Indexes, has demonstrated superior performance compared to alternative models like SARIMA and Prophet. The evaluation findings indicate a Mean Absolute Percentage Error (MAPE) of 3.049, Root Mean Square Error (RMSE) of 45801.625, and Mean Absolute Error (MAE) of 36052.105.

**Keywords :** Forecasting, Tourism, CNN-LSTM, Google Trends, SARIMA, Prophet.

## INTISARI

Industri pariwisata merupakan sektor penting yang berpotensi memberikan kontribusi terhadap produk domestik bruto (PDB) suatu negara. Meskipun mempunyai arti penting, sektor pariwisata sangat rentan, terutama setelah pandemi COVID-19. Untuk memudahkan pemerintah dalam merumuskan kebijakan pariwisata yang efektif dan berkelanjutan, diperlukan estimasi yang tepat. Berdasarkan data Badan Pusat Statistik (BPS), terjadi penurunan signifikan jumlah kunjungan wisatawan mancanegara ke Indonesia. Kehadiran data *outlier* ini menimbulkan hambatan besar dalam memproyeksikan kedatangan wisatawan secara akurat. Metode peramalan yang diusulkan, yang dikenal sebagai pembelajaran mendalam CNN-LSTM, bertujuan untuk mengatasi tantangan yang terkait dengan kumpulan data yang terbatas dan penurunan data yang cepat. Model CNN-LSTM yang disarankan dalam penelitian ini terdiri dari dua komponen utama: Komponen awal terdiri dari lapisan konvolusional dan penggabungan, yang menjalankan proses matematika rumit untuk mengekstraksi karakteristik data masukan. Komponen selanjutnya memanfaatkan fitur yang disediakan oleh LSTM dan lapisan padat. Selanjutnya, untuk meningkatkan akurasi, penyertaan data eksogen dilakukan melalui pemanfaatan indeks *Google Trends*. Hasilnya, pendekatan CNN-LSTM, bersama dengan variabel eksogen dari indeks Google Tren, telah menunjukkan kinerja yang unggul dibandingkan model alternatif seperti SARIMA dan Prophet. Hasil evaluasi menunjukkan *Mean Absolute Percentage Error* (MAPE) sebesar 3,049, *Root Mean Square Error* (RMSE) sebesar 45.801,625, dan *Mean Absolute Error* (MAE) sebesar 36.052,105.

**Kata kunci** – Peramalan, Pariwisata, CNN-LSTM, Google Trends, SARIMA, Prophet.