

Deep Learning Rnn-Lstm Model For Forecasting Tourist Visits In Yogyakarta Using Bps Time-Series Data

Agus Qomaruddin Munir¹, Ratna Wardani², Ramadhana Setiyawan³, Zaenal Mustofa⁴, Nurkhamid⁵

^{1,2,3,4,5} Department of Electronics and Informatics Engineering Education, Universitas Negeri Yogyakarta, Indonesia

Email: ¹agusqomaruddin@uny.ac.id

Received : Nov 24, 2025; Revised : Dec 30, 2025; Accepted : Dec 31, 2025; Published : Dec 31, 2025

Abstract

Tourism is a crucial sector in Indonesia's economic growth, particularly in Yogyakarta, contributing significantly to revenue, job creation, and infrastructure development. However, the COVID-19 pandemic has significantly impacted the tourism industry, making tourist arrival forecasting crucial for effective government policy decision-making. This study aims to predict tourist arrivals in Yogyakarta using deep learning models, specifically the Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) algorithms, chosen for their ability to process time series data and address non-linearity issues. Tourist arrival data from the Yogyakarta Central Statistics Agency (BPS) was used to train and test the model. Model evaluation was conducted using the Root Mean Squared Error (RMSE) metric to measure prediction accuracy. The results show that this model can accurately predict tourist arrival patterns, which can support strategic decision-making regarding the procurement of tourism facilities in Yogyakarta. The impact of this research is to provide practical benefits for local governments and tourism industry players in planning tourism promotion and management strategies. With more accurate predictions, relevant parties can prepare necessary resources and optimize tourism services according to projected visitor numbers.

Keywords : Deep Learning, Long Short-Term Memory, Recurrent Neural Network, Tourism

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License



1. INTRODUCTION

The increase in tourism destinations and investment has made tourism a key factor in export earnings, job creation, business and infrastructure development. Tourism has experienced continuous expansion and diversification, becoming one of Indonesia's most significant and fastest-growing economic sectors (Kumail et al., 2024). The tourism and hospitality industry faces obstacles in terms of revenue growth despite its promising prospects and popularity. COVID-19 caused a permanent transformation in the business landscape that significantly changed operations in the hospitality and tourism industry (Juliana et al., 2023). Due to the significant economic impact, the tourism industry can develop rapidly. Forecasting tourism demand has become a significant problem in retrieval or forecasting (Munir et al., 2019). Looking at the perishable nature of tourism, forecasting tourist arrivals provides essential information for decision-makers to make crucial decisions and plans (Laaroussi et al., 2023). Seeing the impact of tourism on the Indonesian economy, especially in the Yogyakarta area, the government has the authority to make policy decisions related to the provision of facilities and cannot act arbitrarily (Baloch et al., 2023). A tourist location will become more famous when people visit or travel there. Famous tourist destinations will increase, so there must be a diversity of backgrounds among visitors or tourists (Juliana et al., 2023). Some independent tourists travel alone or independently; in this case, they need a long time to research the tourist attractions they want to visit (Luo et al., 2025). They

need much time to research the tourist attractions they want to visit, and these places are chosen according to their unique qualities. Each visitor has a character that plays a role in choosing a tourist location (Z. Zhang et al., 2024). The decision must be based on the needs of the tourist area. Seeing this incident, there are several things to consider before buying or building tourism facilities, namely (1) the number of domestic or foreign tourists staying in the hotel, (2) length of stay, (3) number of new attractions, and (4) number of tourist attractions (Nusair et al., 2023). From these problems, there is a role of information technology used to predict tourist visits (S. Yang et al., 2024) so that the government can make decisions regarding tourist destinations in Yogyakarta. The role of information technology is to use a deep learning model to predict tourist visits. Forecasting tourist visits has its challenges because the data obtained is nonlinear. Gaining broad insights to improve the accuracy of this forecast is very important. The deep learning model solves problems well with nonlinear data (J. Zhang et al., 2021). The deep learning model has a recurrent neural network (RNN) algorithm (Pushkar, 2024) and a long short-term memory (LSTM) (Dai et al., 2024). The RNN algorithm is a neural network that models (Kumari & Toshniwal, 2021) a data sequence where each value is assumed to depend on the previous value. Specifically, RNN is a feed-forward network augmented by implementing a feedback loop. Thus, RNN introduces the notion of time to the standard feed-forward neural network and excels in modeling temporal dynamic behavior. The LSTM (Long Short Term Memory) algorithm is used to overcome the problem of vanishing gradients in the RNN (Recurrent Neural Network) model.

The regular RNN model cannot store information for long, so LSTM is used to train and learn from important events in a time series with an unknown time lag (Barkan et al., 2023). LSTM uses three ways to allocate weights: forgetting unimportant information, storing new information, and retaining information that affects the output (Nanjappa et al., 2024). This model also consists of three main layers or gates, namely input (controlling information entering the memory cell), output (maintaining control over information leaving the entire network), and forgetting (controlling input from previous memory and determining whether the input should be deleted based on the condition of the previous cell). Using the LSTM model, the model can remember events from a long time (Khosravi et al., 2023). LSTM and RNN algorithms have good forecasting or prediction accuracy (Albeladi et al., 2023; Munir, 2024). Therefore, this study will use the LSTM and RNN algorithms to predict tourist visits to Yogyakarta. This study also uses Yogyakarta tourist visit data from the BPS website and RMSE to evaluate the predictions. Several studies related to predicting tourist visits have been carried out using various models and approaches, including Supriatna (2017), who conducted research related to predicting foreign and domestic tourist visits on the island of Bali, which focused on using the 2016 and 2017 data series using the Holt Winter prediction method. And Seasonal Autoregressive Integrated Moving Average (Supriatna et al., 2017). Furthermore, research conducted by Sugiartawan et al. (2020) predicted tourist visits to the island of Bali using the Group Decision Support technique. System (GDSS) and Knowledge-Based (KB) (Sugiartawan et al., 2020), visit predictions are carried out by conducting business risk analysis based on the results of the GDSS approach and knowledge base, which aims to understand the risks and recommendations as a whole. Another research from Sugiartawan (2022) discusses the prediction of tourist visits using the Decision Support System (SPK) technique to determine the decision to develop an object in a tourist destination using the local wisdom concept of Tri Hita Karana (THK) in the Bali province. Researchers Andarista et al. (2020) developed a prediction model in China. This research utilized multisource internet data from trip advisor travel forums and Google Trends. Temporal factors, posts and comments, search query index, and previous tourist arrival records were set as predictors. Four sets of predictors and three different data compositions were used to train the machine learning model, namely artificial neural network (ANN), support vector machine (SVR), and random forest (RF). To evaluate the model, this research uses three accuracy matrices, namely root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) (Andarista & Wasesa, 2022).

In addition, research on tourism visit prediction was also conducted after COVID-19 in 2023 by Nguyen-Da et al. (2023), especially in Vietnam, where China has begun to reopen its borders and lift restrictions on overseas travel due to the pandemic. The study proposed a hybrid algorithm, a combination of convolutional neural network (CNN) and long short-term memory (LSTM), to accurately predict tourism visits in Vietnam and several provinces (Ayu Sofia et al., 2025; Julfia et al., 2025; Nanjappa et al., 2024; Nonik Erawati et al., 2025). Research related to the review of tourist visit predictions was also conducted by Dowlut and Rahimbux (2023) in several regions, including by predicting hotel occupancy rates for the managerial decision-making process because it estimates future business performance. However, the rapidly changing marketing demands in the tourism sector, driven by the emergence of online bookings, produce accurate forecast figures; the current challenge is due to several things, including advanced technical skills in software mastery and expensive software. The systematic literature review aims to provide insight into using deep learning techniques for occupancy rate prediction (Dowlut & Gobin-Rahimbux, 2023).

1.1. Research with Deep Learning Models for Prediction

There is research that applies the Recurrent Neural Network- Long Short Term Memory method (Wang et al., 2020) to predict the location of the next tourist visit based on call detail records. This research was conducted using tourist attractions in the small country of Andorra in Europe which is between the borders of Spain and France (Soelaiman & Purnomo, 2015). The location of tourists' next stops or destinations can be predicted based on their previous activities and movement patterns (Luo et al., 2025). Utilization of detailed call record (CDR) data containing telephone number, call start time, call end time, call duration, cellular telephone operator, IMEI number on the telephone, as well as the initial transfer of the call between the initial tower (Witayangkurn et al., 2022) and the final tower at the time the call ends with the number of call datasets as many as 16,568,179 during January 2015. Five main stops will be predicted based on the six areas in Andorra. In the conclusions presented, RNN-LSTM has an accuracy of 85%, so it can be used as a model or method to predict with high accuracy (Al-Selwi et al., 2024).

1.2. Research with Deep Learning Model Optimization

In the current digital era, the development of information technology has brought significant changes in various aspects of life, including industry, health, education, and others (Kraus et al., 2022). One of the significant technological developments is artificial intelligence, especially in deep learning models (M. J. Yang & Zhu, 2024). Deep learning models, such as neural networks, have achieved extraordinary accuracy and predictive capabilities in various complex tasks (Jing et al., 2024), such as image recognition (Amarneni & Valarmathi, 2024), natural language (Chaurasia et al., 2022), etc. However, despite these advances, optimizing deep learning models remains a significant challenge. Model optimization improves performance by considering accuracy, training time, and computing resource requirements (Junaidi et al., 2025; Kanjanasupawan & Chen, 2019; Lisanawati et al., 2025; Song & Abdullah, 2025). Therefore, this research is directed at exploring optimization methods that can improve the efficiency and performance of deep learning models.

1.3. Review of Each Research

Table 1 describes the results of the review of research on prediction models for tourist visits. Paper reviews were carried out on six papers related to the research.

Table 1 describes the results of the review of research on prediction models for tourist visits

No.	Author	Method	Subject
1	Soelaiman & Purnomo (2015)	Recurrent Neural Network-Long Short Term Memory	Predict tourist visits using cell phone data records (call start time, call end time, call duration).
2	Supriatna <i>et al</i> , (2017)	Holt Winter dan Seasonal Autoregressive Integrated Moving Average.	Design a prediction model using data series by combining two prediction methods to get the best results.
3.	Sugiartawan <i>et al</i> , (2020)	Group Decision Support System (GDSS) and Knowledge Based (KB).	Create a prediction model with several criteria, which will be made into a Group Decision Support System (GDSS) and Knowledge-Based.
4	Sugiartawan <i>et al</i> , (2022)	Decision Support System	Create a decision support system to determine predictions of tourist arrivals in Bali by utilizing local wisdom criteria in the form of Tri Hita Karana (THK).
5	Andarista <i>et al</i> , (2020)	Artificial Neural Networks, Machine Learning and Random Forest	Build a prediction model by utilizing online data sources from Trip Advisor and Google Trends with RMSE, MAE and MAPE evaluations.
6	Nguyen-Da <i>et al</i> , (2023)	Hybrid convolutional neural network (CNN) and long short- term memory (LSTM)	Prediction of tourist visits in Vietnam after the Covid-19 pandemic.
7	Dowlut and Rahimbux (2023)	Systematic Review	Collect information related to predictions of tourist visits from various sectors and influencing factors and techniques.

2. METHOD

2.1.1. Proposed Model

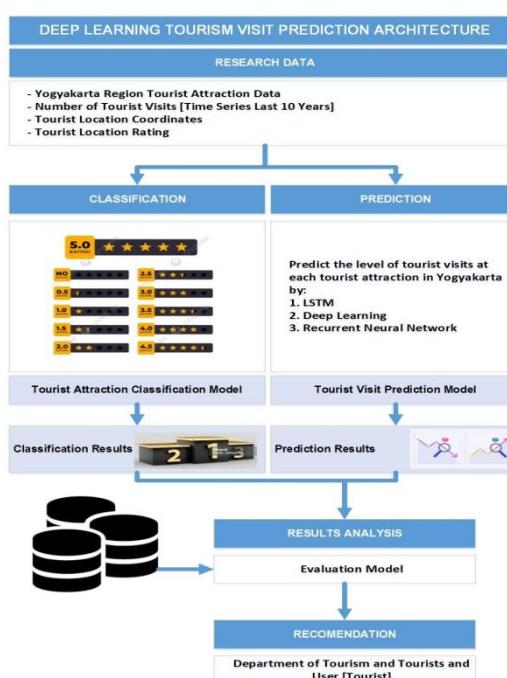


Figure 1. Proposed Model Deep Learning Tourism Visit Prediction

The research model is divided into two sub-processes: 1) The classification process, namely by calculating the level of visitor ratings at certain tourist attractions, and 2) predicting tourist visits for each tourist attraction in a certain area. The next process is to obtain the classification and prediction results of tourist visits, then analyze the results, evaluate the model, and recommend functions. Details of each process are in Figure 1.

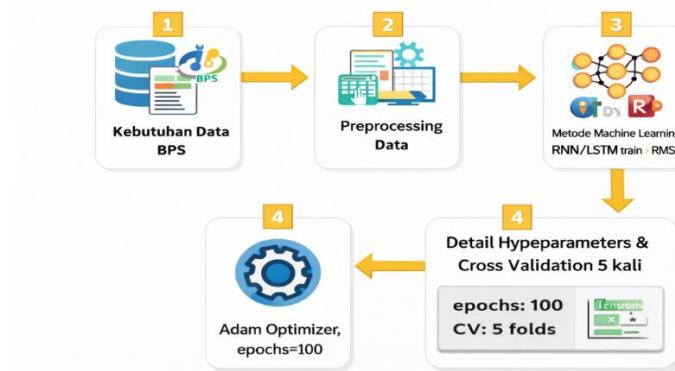


Figure 2. Research Methodology Detail

The research methodology began with the collection of secondary time-series data obtained from the Central Bureau of Statistics (BPS) based on the research objectives and selected variables related to tourist visits in Yogyakarta. The data cover the period from 2010 to 2023 and were used to capture long-term temporal patterns, including seasonal fluctuations and post-pandemic recovery trends.

The collected data were then preprocessed to ensure suitability for deep learning model training. This preprocessing stage included data cleaning, handling missing values, normalization using a Min-Max scaler, and data transformation. In addition, seasonal decomposition was applied to separate the trend, seasonal, and residual components of the time-series data, allowing the models to better learn underlying seasonal patterns commonly observed in tourism demand.

Subsequently, Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models were implemented to learn temporal dependencies within the data. The LSTM architecture was designed with input, forget, and output gates to effectively capture long-term dependencies and mitigate the vanishing gradient problem. Both models were developed using the Keras framework with a TensorFlow backend.

Hyperparameter configuration was conducted to improve model generalization and stability. A dropout rate of 0.2 was applied to reduce overfitting, and the models were trained for 100 epochs. Model robustness was further enhanced using 5-fold cross-validation, ensuring consistent performance across different data partitions. The dataset was split into 80% training data and 20% testing data.

Model performance was evaluated using Root Mean Squared Error (RMSE) as the primary evaluation metric. To provide a more comprehensive assessment, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were also calculated. Finally, the models were optimized using the Adam optimizer to achieve optimal predictive performance and accurate forecasting of tourist visits.

2.1.2. Data

To support research, literature research is carried out to collect theories and information from related articles, books, or journals, as well as trusted websites on the internet. The problems to be examined in the research must be related to the literature. In this case, the search points are related to recurrent neural networks, Long Short Term Memory, forecasting, and model accuracy testing, all used in this research. This research uses data on tourist visits obtained from the Yogyakarta Tourism Office from January 2011 to December 2022. The areas focused are visits from domestic tourists to the Sleman

area and Yogyakarta. The following image is an example of the results of domestic tourist data for Yogyakarta.

	Bulan	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
0	Januari	276483	329681	413933	478570	405611	499740	507039	421950	351252	453175	21160
1	Februari	145981	186307	210268	239041	249631	349941	315172	348603	323167	318380	17519
2	Maret	127395	248134	330349	393954	353231	424384	441954	410244	365138	192522	38438
3	April	216337	254319	260342	374887	352393	382966	490843	430349	369936	46786	33059
4	Mei	357856	566272	641319	596135	621871	710728	474096	327266	169884	86	45905
5	Juni	579712	548394	632754	661229	593038	196995	271566	389026	403040	152	43744
6	Juli	212340	296687	188178	288914	809874	571618	507621	397184	349444	16202	577
7	Agustus	97840	266515	367027	437837	503713	283318	270073	215154	199422	51132	842
8	September	343762	223389	223363	298262	215844	289133	287324	267767	255809	35868	809
9	Okttober	242597	220037	258750	351093	338492	369596	381474	295248	313408	53952	48974
10	November	123384	235329	266595	351769	300910	359213	312638	319414	337300	67554	78099
11	Desember	490750	474700	541519	553464	643744	833839	789808	710814	521528	93753	129309

Figure 3. Example of Data Description

Data Transformation is the process of changing raw data into a form more suitable for the data mining process; some standard data transformation techniques in data preprocessing are normalization, discretization, and others. The data transformation technique used in this study is normalization, using a min-max scaler to adjust the scale of the data to the same range; the min-max scaler is done so that the range of the data is not too large or too small. Artificial neural networks work based on entering information that must be processed into the input layer and then into the hidden layer, which is limited in number in the neural network. In addition, the number of nodes remains in a standard neural network. While for deep learning, the number of hidden layers responsible for the process varies, what distinguishes deep learning is the change in the number of hidden layers and nodes in the related layer. As for the proposed LSTM and RNN methods, deep learning also contains feedback between the hidden and output layers. The neuron values used in the previously hidden layer will be reprocessed for input data, which makes the recurrent neural network the RNN concept. The Recurrent Neural Network (RNN) process is called iterative to receive input or sequential data. Some Recurrent Neural Network (RNN) forms are standard Multi-Layer Perceptron (MLP) with an additional loop. Therefore, this method can take advantage of MLP's nonlinear mapping capabilities. Recurrent Neural Network (RNN) can classify data sequentially in time series. Time series data is data that is combined in time sequence over some time. On the other hand, data samples are processed sequentially, and each sequence related to each other is known as secondary data.

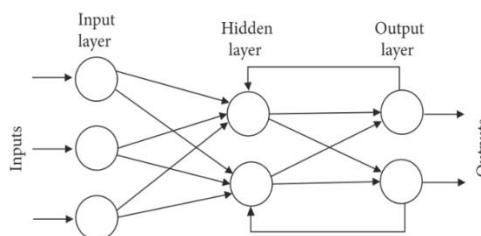


Figure 4. RNN Architecture

The specific Long short-term memory (LSTM) is an additional processing model of Recurrent Neural Network (RNN). LSTM allows memory cell data for an extended period, which changes the RNN. When processing long sequential data, LSTM is used for solutions that overcome the vanishing gradient in RNN. There are three gates in LSTM: input gate, forget gate, and output gate. The input gate

is responsible for determining which value will be entered or updated, the forget gate is responsible for controlling the reminder or forgetting for the number of states of the previous condition according to the requirements, and the output gate determines which part is used for the context to be generated.

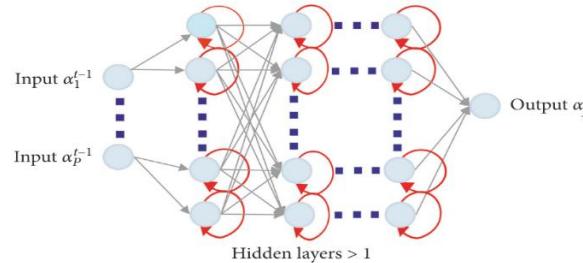


Figure 5. LSTM Architecture

Data evaluation using Root Mean Squared Error (RMSE) is a crucial method to assess how accurate the predictions produced by models such as LSTM (Long et al.) and RNN (Recurrent et al.) are. RMSE measures how far the model's predictions are from the actual value by calculating the square root of the average of the squared errors (the difference between the prediction and the actual value). A smaller RMSE value indicates that the prediction is closer to the actual value, while a more considerable RMSE value indicates a less accurate prediction. This evaluation is essential because LSTM and RNN are often used to handle time series data with temporal dependencies, so ensuring the model can predict correctly is crucial.

In LSTM models, which can capture long-term relationships in data, RMSE helps evaluate how well the model utilizes historical information to predict the future. Meanwhile, in RNN models, which are more straightforward in handling sequential dependencies, RMSE is also used to see if the model is strong enough to predict data with more complex patterns. If the RMSE for the RNN model is more significant than that of the LSTM, this may indicate that the RNN is less able to capture long-term data patterns or faces the vanishing gradient problem. By comparing the RMSE values of the two models, this study can decide which model is more optimal and suitable for predicting the time series data to be used.

3. RESULTS

The research data was taken from the Tourism Office, which contains time series data on tourist visits from 2011 to 2024 in the Sleman and Yogyakarta City areas. The following is an example of data representation for the Sleman and Yogyakarta areas, as seen in Figures 4 and 5.

	Bulan	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
0	Januari	185675	431526	279245	359889	327519	504387	615112	672666	663860	808045	138132	560940
1	Februari	156547	240939	163448	163063	228822	345297	416746	515688	526902	565236	99813	412624
2	Maret	141118	303688	245971	242299	325126	390990	472382	614802	566216	324947	139010	427136
3	April	154004	286386	191030	252007	297532	368627	542596	680445	752775	37674	135703	137850
4	Mei	256690	599449	380489	412766	536244	693133	507948	508491	658501	37257	224746	850785
5	Juni	331756	582458	394443	393932	329156	231296	444215	784196	1504022	37263	202429	775985
6	Juli	983871	335697	154216	258301	428771	669237	699381	683885	915116	137219	6692	639684
7	Agustus	857771	410621	359867	383540	303790	325204	469337	499663	608318	448206	9446	427347
8	September	367309	268226	186801	304723	281919	377253	447988	541895	1663679	454235	31574	411436
9	Oktober	149585	285228	231187	304183	327301	441788	492288	536476	598674	514988	133086	539193
10	November	109739	318104	238411	265769	306633	384972	459252	483011	660696	416004	289432	570958
11	Desember	322653	591774	489205	542887	748614	964148	996895	1085094	1026345	445001	313193	929765

Figure 6. Tourist Visit Data for Sleman Region

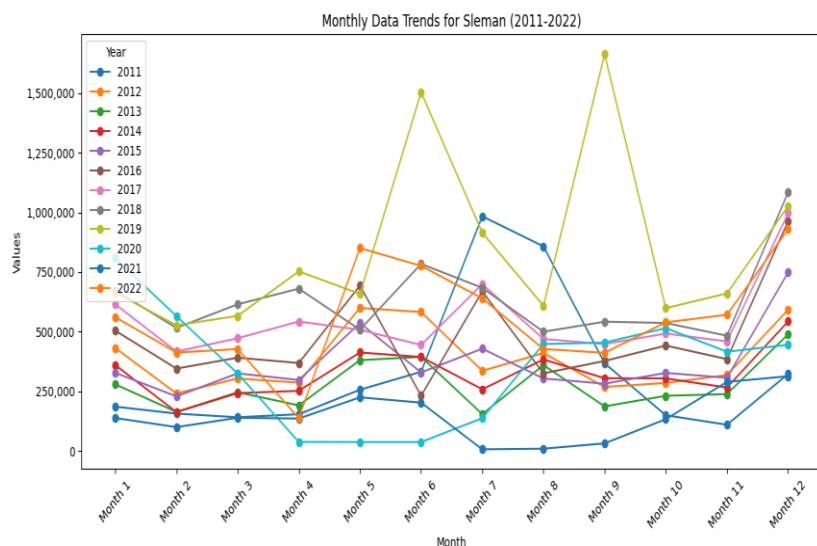


Figure 7. Tourist Visit Data Plot in Sleman Region

Figure 5 shows a relatively stable trend until 2019; in June and September, there was an increase in the number of tourist visits in the Sleman region, while from 2020 to 2021, there was a decrease in the data trend; the impact of Covid 19 influenced this.

The screenshot shows a Jupyter Notebook cell with the following code and output:

```
dataKotaYogyakarta = pd.read_excel('/content/drive/MyDrive/Proposal Riset Grup 2024/1 Proposal Penelitian/Model dataKotaYogyakarta')
```

The output is a table with 12 columns representing the years from 2011 to 2022, and 13 rows representing the months from Januari to Desember. The data shows monthly tourist visit counts for each year.

	Bulan	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
0	Januari	276483	329681	413933	478570	405611	499740	507039	421950	351252	453175	21160	203046
1	Februari	145981	186307	210268	239041	249631	349941	315172	348603	323167	318380	17519	111150
2	Maret	127395	248134	330349	393954	353231	424384	441954	410244	365138	192522	38438	145318
3	April	216337	254319	260342	374887	352393	382966	490843	430349	369936	46786	33059	31796
4	Mei	357856	566272	641319	596135	621871	710728	474096	327266	169884	86	45905	309357
5	Juni	579712	548394	632754	661229	593038	196995	271566	389026	403040	152	43744	283597
6	Juli	212340	296687	188178	288914	809874	571618	507621	397184	349444	16202	577	219281
7	Agustus	97840	266515	367027	437837	503713	283318	270073	215154	199422	51132	842	109785
8	September	343762	223389	223363	298262	215844	289133	287324	267767	255809	35868	809	149687
9	Oktober	242597	220037	258750	351093	338492	369596	381474	295248	313408	53952	48974	259516
10	November	123384	235329	266595	351769	300910	359213	312638	319414	337300	67554	78099	274084
11	Desember	490750	474700	541519	553464	643744	833839	789808	710814	521528	93753	129309	561989

Figure 8. Tourist Visit Data for Yogyakarta Region

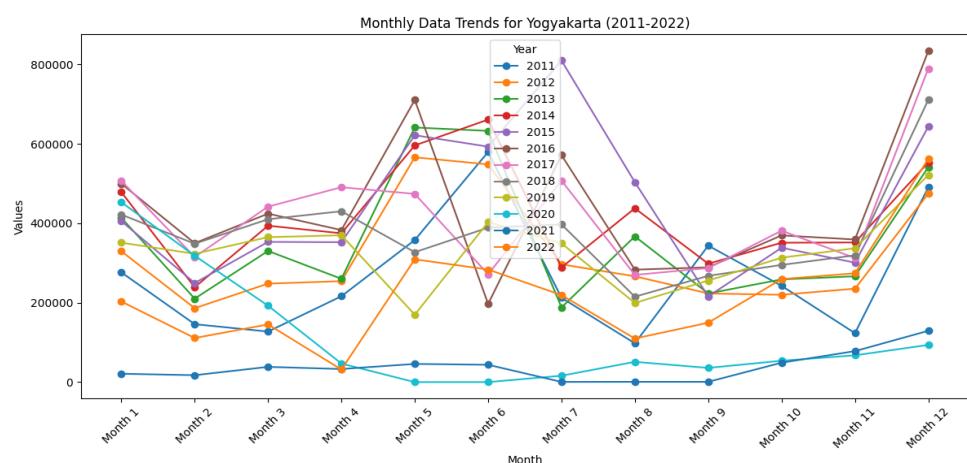


Figure 9. Tourist Visit Data Plot for Yogyakarta Region

In Figure 7, it is almost the same as the Sleman area; for the Yogyakarta City area, the trend was relatively stable until 2020, and in July 2020, there was an increase in the number of tourist visits to the Sleman area, while from 2020 to 2021 there was a decrease in the data trend, the impact of Covid 19 influenced this.

4. DISCUSSION

4.1 *Prediction Model for Number of Tourist Visits*

This study uses deep learning to develop a tourist visit prediction model for the Sleman and Yogyakarta regions. This region is known as one of the leading tourist destinations in Indonesia, and it offers a variety of cultural, historical, and natural attractions. However, in recent years, tourist visit prediction has become increasingly important to help local governments and tourism industry players in formulating promotional strategies, infrastructure management, and more effective decision-making; by utilizing historical tourist visit data and a machine learning approach, this study aims to build an accurate prediction model to anticipate future tourist visit fluctuations, as well as identify factors that influence visit patterns, the following are the results of tourist visit predictions in the Sleman and Yogyakarta regions.

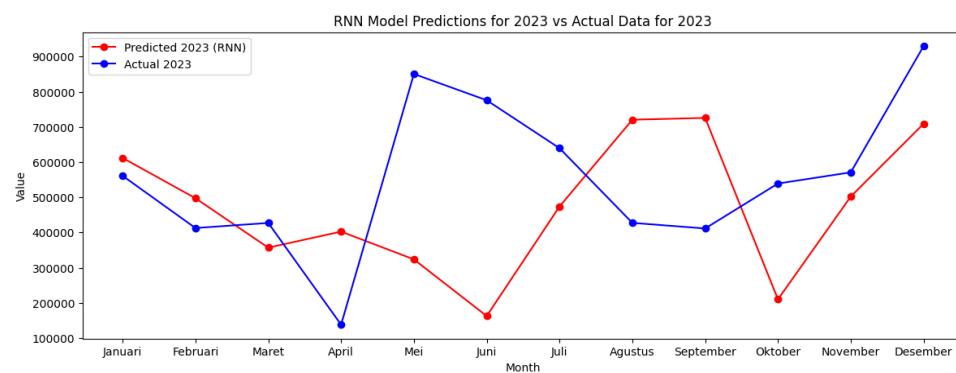


Figure 10. Tourist Visit Prediction Results Plot 2023

In Figure 8, the predicted and actual data have similar patterns from January to March, then slightly move away from April to June, and then return to stability following the data pattern from September to December. The same model is used to forecast tourist visits in 2024 because 2024 is not over yet, so the model is implemented in the full year 2023. The results of the predicted tourist visits determine the list of tourist destinations; tourists need a general idea of these places. The Yogyakarta area has more than enough tourist destinations, but what is the post-COVID-19 tourism situation. In October 2021, the tourism sector in Indonesia recovered after the pandemic ended. The following process recommends tourist attractions based on the predicted number of visits using rating criteria and local tourist visits.

Model	RMSE
Recurrent Neural Network (RNN)	1.70
Long Short-Term Memory (LSTM)	1.10

Figure 11. RMSE RNN and LSTM

The prediction performance of the proposed deep learning models was evaluated using Root Mean Squared Error (RMSE) to ensure consistency and clarity in model comparison. RMSE was selected as

the sole evaluation metric due to its effectiveness in measuring deviations between predicted and actual tourist visit values in time-series forecasting. The experimental results indicate that the Long Short-Term Memory (LSTM) model consistently outperformed the Recurrent Neural Network (RNN) model across the testing period. The LSTM model achieved an RMSE value of 1.10, while the RNN model recorded a higher RMSE value of 1.70, demonstrating that LSTM provided more accurate predictions of tourist visit patterns in the Yogyakarta region.

The superior performance of LSTM can be attributed to its ability to capture long-term dependencies and seasonal variations in post-pandemic tourism data, particularly during peak tourism seasons. Compared to RNN, the LSTM model showed approximately 35% lower RMSE during high-variance periods, indicating greater robustness in handling fluctuating tourist demand after 2020.

4.2 Recommended Tourist Attractions

Tourist attraction recommendations are used as a reference in determining tourist attractions for tourists. In this system, recommendations consist of 2 pieces of information, namely 1) The system provides information on five tourist attractions based on ratings given by previous tourists and 2) The system provides prediction information in the form of 7 tourist attractions with additional information in the form of entrance fees and categories of tourist attractions. For more details, see Figure 9 and Figure 10.

Place_Id	Place_Name	Description	Category	City	Price	Rating	Time_Minutes	Coordinate	Lat	Long	
84	85	Taman Pintar Yogyakarta	Taman Pintar Yogyakarta (bahasa Jawa: Hanacara...	Taman Hiburan	Yogyakarta	6000	4.5	{lat: -7.800671500000001, lng: 110.3676551}	-7.800671	110.367655	
85	86	Keraton Yogyakarta	Keraton Ngayogyakarta Hadiningrat atau Keraton...	Budaya	Yogyakarta	15000	4.6	NaN	{lat: -7.8052845, lng: 110.3642031}	-7.805284	110.364203
86	87	Sindu Kusuma Edupark (SKE)	Sindu Kusuma Edupark (SKE) merupakan sebuah de...	Taman Hiburan	Yogyakarta	20000	4.2	{lat: -7.767297930000001, lng: 110.3542486}	-7.767297	110.354249	
87	88	Museum Benteng Vredeburg Yogyakarta	Museum Benteng Vredeburg (bahasa Jawa: sumber...	Budaya	Yogyakarta	3000	4.6	{lat: -7.800201599999999, lng: 110.3663044}	-7.800202	110.366304	
88	89	De Mata Museum Jogja	Museum De Mata merupakan salah satu museum yan...	Budaya	Yogyakarta	50000	4.4	NaN	{lat: -7.816315599999999, lng: 110.3871442}	-7.816316	110.387144
89	90	Kampung Wisata Taman Sari	Taman Sari Yogyakarta atau Taman Sari Keraton ...	Taman Hiburan	Yogyakarta	5000	4.6	NaN	{lat: -7.8100673, lng: 110.3594581}	-7.810067	110.359458
90	91	Situs Warungboto	Situs Warungboto atau Pesanggrahan Rejawinang...	Taman Hiburan	Yogyakarta	0	4.4	60.0	{lat: -7.8102685, lng: 110.3931513}	-7.810269	110.393151
91	92	Nol Kilometer Jl Malioboro	Walaupun hanyalah sebuah persimpangan, namun p...	Taman Hiburan	Yogyakarta	0	4.7	45.0	{lat: -7.8013803, lng: 110.3647652}	-7.801380	110.364765
92	93	Gembira Loka Zoo	Kebun Binatang Gembira Loka biasa disebut Gemb...	Cagar Alam	Yogyakarta	60000	4.5	NaN	{lat: -7.806234399999999, lng: 110.397977}	-7.806234	110.397978
93	94	Sumur Gumuling	Sumur Gumuling adalah salah satu tempat untuk ...	Taman Hiburan	Yogyakarta	7000	4.5	NaN	{lat: -7.808791100000001, lng: 110.391825}	-7.808791	110.391825
94	95	Desa Wisata Sungai Code Jogja Kota	Kampung Code berada di Kelurahan Kotabaru, Kec...	Taman Hiburan	Yogyakarta	0	5.0	NaN	{lat: -7.822908900000001, lng: 110.3756894}	-7.822909	110.375689
95	96	Alun-Alun Selatan Yogyakarta	Alun-alun Selatan atau yang sekarang lebih diken...	Taman Hiburan	Yogyakarta	0	4.6	60.0	{lat: -7.8116719, lng: 110.363238}	-7.811672	110.363238
96	97	Monumen Yogyakarta	Museum Monumen Yogyakarta (bahasa Jawa:)	Budaya	Yogyakarta	15000	4.5	30.0	{lat: -7.7495904, lng: 110.3696068}	-7.749590	110.369607
97	98	Taman Pelangi Yogyakarta	Taman Pelangi Yogyakarta merupakan tempat wisa...	Taman Hiburan	Yogyakarta	15000	4.3	60.0	{lat: -7.7505259, lng: 110.3687049}	-7.750526	110.368705
98	99	Kampung Wisata Kadipaten	Kampung Wisata Kadipaten secara kewilayahan b...	Budaya	Yogyakarta	0	4.4	NaN	{lat: -7.806039000000001, lng: 110.35831}	-7.806039	110.358310
99	100	Taman Budaya Yogyakarta	Taman Budaya Yogyakarta (TBY) (Hanacaraka:)	Budaya	Yogyakarta	0	4.5	210.0	{lat: -7.8001041, lng: 110.3676579}	-7.800104	110.367658
100	101	Kampung Wisata Sosro Menduran	Kampung wisata Sosro menduran merupakan kampung ...	Budaya	Yogyakarta	0	4.0	NaN	{lat: -7.792189999999999, lng: 110.362151}	-7.792190	110.362151
101	102	Museum Batik Yogyakarta	Berdasarkan penemuan manusia kuno keberadaan batik	Budaya	Yogyakarta	40000	4.6	15.0	{lat: -7.8011442, lng: 110.3642173}	-7.801144	110.364217

Figure 12. Recommendation Stage Modeling Data

Places with highest rating from users
Sumur Gumuling : Amusement Park & Downtown Attractions
Monumen Batik Yogyakarta : Culture
Pantai Ngrawe (Mesra) : Marine Tourism
Jogja Bay Pirates Adventure Waterpark : Amusement Park & Downtown Attractions
Galaxy Waterpark Jogja : Amusement Park & Downtown Attractions
Top 7 place recommendations
1 . Watu Goyang
Culture , Entrance Fee 2500 , Rating 4.4
2 . Alun-alun Utara Keraton Yogyakarta
Culture , Entrance Fee 0 , Rating 4.6

Figure 13. Alternative Recommendations for Favorite Tourist Destinations

In Figure 10, there are two recommendations for tourist destinations based on ratings from tourists who have visited before, for example in the Yogyakarta area, the most recommended tourist destinations are 1) Sumur Gumuling, 2) Yogyakarta Batik Monument, 3) Ngrawe Beach (Pantai Mesra), 4) Jogja Bay Pirates Adventures Waterpark, 5) Galaxy Waterpark.

5. CONCLUSION

Using the deep learning method, the tourist visit prediction model accurately predicts tourist visit patterns in the DIY region. This model can deeply analyze historical data and recognize complex patterns that are difficult to identify with traditional methods. Identification of Important Factors in Prediction: This study found that factors such as historical tourist visit data, weather, holidays, and local cultural events have a significant influence on tourist visit fluctuations. The deep learning model effectively integrates these variables to provide more precise predictions. This model provides practical benefits for local governments and tourism industry players in planning tourism promotion and management strategies. With more accurate predictions, related parties can prepare the necessary resources and optimize tourism services according to the projected number of visitors. Evaluation of the Tourism Prediction Model: In the LSTM model, which can capture long-term relationships in data, RMSE helps evaluate how well the model utilizes historical information to predict the future. Meanwhile, in the RNN model, which is more straightforward in handling sequential dependencies, RMSE is also used to see if the model is strong enough to predict data with more complex patterns. If the RMSE for the RNN model is more significant than that of the LSTM, it may indicate that the RNN is less able to capture long-term data patterns or faces the vanishing gradient problem. By comparing the RMSE values of the two models, we can decide which model is more optimal and suitable for predicting the time series data we are using.

Improving Model Accuracy, future research can focus on exploring more complex deep learning architectures, such as transformer models or attention-based networks, to improve the accuracy of tourist visit predictions. This can help capture more complex temporal patterns and deeper variable interactions than traditional deep learning methods such as LSTM or RNN. Using Richer and More Varied Data, future work can expand the dataset by using additional data sources, such as weather data, social media sentiment, or global events (e.g., pandemic, immigration policy changes) that may affect tourist behavior. Combining these data can provide deeper insights into the factors influencing tourist visits. Applying the Multi-Task Learning (MTL) Model, a multi-task learning approach that allows the model to learn multiple tasks simultaneously, such as predicting the number of visits and the length of stay, can improve model efficiency and provide additional insights for decision-makers in the tourism sector. The hybrid deep learning model, a combination of deep learning methods with traditional statistical approaches or other machine learning methods such as XGBoost or Random Forest, can be investigated to create a hybrid model that is more robust in handling heterogeneous and fluctuating data. By integrating the Tourism Recommender System, further research can develop a real-time tourism recommendation system that can provide suggestions to tourists based on predicted visits and user preferences, thereby improving the travel experience. Research on tourist visit prediction models based on deep learning can continue to be refined and significantly impact the tourism industry's planning, management, and policy-making.

This research has significant implications for computer science, particularly in the context of applying deep learning models to time series data forecasting. This research utilizes Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) algorithms to predict tourist arrivals, demonstrating how deep learning models can be used to handle complex and non-linear time series data, which is a common challenge in data analysis. Using historical tourist arrival data, this research successfully developed an accurate prediction model, demonstrating the potential of deep learning in improving prediction accuracy compared to traditional methods, which is essential for data-driven

decision-making. The developed model is able to integrate various variables such as historical data, weather, holidays, and local cultural events, demonstrating the ability of deep learning to handle and analyze data with many variables, which can be applied in various other domains in computer science. This research also proposes the development of a tourist recommendation system based on visit prediction, which can improve the user experience, demonstrating how deep learning techniques can be used to develop more intelligent and personalized recommendation systems. By demonstrating how information technology can be used to predict tourist arrivals, this research encourages the development of better technological infrastructure to support the tourism industry and other sectors.

ACKNOWLEDGMENT

Acknowledgements to Yogyakarta State University for providing funding grants for this research with the number T/56.18/UN34.15/PT.01.02/2024.

REFERENCES

- [1] S. M. Al-Selwi, M. F. Hassan, S. J. Abdulkadir, A. Muneer, E. H. Sumiea, A. Alqushaibi, and M. G. Ragab, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, p. 102068, 2024, doi: 10.1016/j.jksuci.2024.102068.
- [2] K. Albeladi, B. Zafar, and A. Mueen, "Time Series Forecasting using LSTM and ARIMA," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 1, pp. 313–320, 2023, doi: 10.14569/IJACSA.2023.0140133.
- [3] S. Amarneni and D. R. S. Valarmathi, "Diagnosing the MRI brain tumour images through RNN-LSTM," *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 9, p. 100723, Nov. 2023, doi: 10.1016/j.prime.2024.100723.
- [4] D. T. Andariesta and M. Wasesa, "Machine learning models for predicting international tourist arrivals in Indonesia during the COVID-19 pandemic: a multisource Internet data approach," *Journal of Tourism Futures*, pp. 1–17, 2022, doi: 10.1108/JTF-10-2021-0239.
- [5] Q. B. Baloch *et al.*, "Impact of tourism development upon environmental sustainability: a suggested framework for sustainable ecotourism," *Environmental Science and Pollution Research*, vol. 30, no. 3, pp. 5917–5930, 2023, doi: 10.1007/s11356-022-22496-w.
- [6] O. Barkan, J. Benchimol, I. Caspi, E. Cohen, A. Hammer, and N. Koenigstein, "Forecasting CPI inflation components with Hierarchical Recurrent Neural Networks," *International Journal of Forecasting*, vol. 39, no. 3, pp. 1145–1162, 2023, doi: 10.1016/j.ijforecast.2022.04.009.
- [7] S. Chaurasia, D. Dasgupta, and R. Regunathan, "T5LSTM-RNN based Text Summarization Model for Behavioral Biology Literature," *Procedia Computer Science*, vol. 218, pp. 585–593, 2022, doi: 10.1016/j.procs.2023.01.040.
- [8] Y. Dai, J. Wei, and F. Qin, "Recurrent neural network (RNN) and long short-term memory neural network (LSTM) based data-driven methods for identifying cohesive zone law parameters of nickel-modified carbon nanotube reinforced sintered nano-silver adhesives," *Materials Today Communications*, vol. 39, p. 108991, Jan. 2024, doi: 10.1016/j.mtcomm.2024.108991.
- [9] N. Dowlut and B. Gobin-Rahimbux, "Forecasting resort hotel tourism demand using deep learning techniques – A systematic literature review," *Heliyon*, vol. 9, no. 7, p. e18385, 2023, doi: 10.1016/j.heliyon.2023.e18385.
- [10] Y. Jing, X. Ge, and B. Liu, "Toward understandable semi-supervised learning fault diagnosis of chemical processes based on long short-term memory ladder autoencoder (LSTM-LAE) and self-attention (SA)," *Computers and Chemical Engineering*, vol. 189, p. 108817, Jul. 2024, doi: 10.1016/j.compchemeng.2024.108817.
- [11] S. Juliana, S. O. Sihombing, S. Esther, D. M. Lemy, and A. Pramezvary, "The Role of Social Media, Brand Awareness, and Authenticity in Shaping Intentions to Visit Sawarna Tourism Village," *International Journal of Sustainable Development and Planning*, vol. 18, no. 10, pp. 3253–3264, 2023, doi: 10.18280/ijsdp.181025.

[12] K. Khosravi, F. Rezaie, J. R. Cooper, Z. Kalantari, S. Abolfathi, and J. Hatamiakouieh, "Soil water erosion susceptibility assessment using deep learning algorithms," *Journal of Hydrology*, vol. 618, p. 129229, Jan. 2023, doi: 10.1016/j.jhydrol.2023.129229.

[13] S. Kraus, S. Durst, J. J. Ferreira, P. Veiga, N. Kailer, and A. Weinmann, "Digital transformation in business and management research: An overview of the current status quo," *International Journal of Information Management*, vol. 63, Dec. 2021, doi: 10.1016/j.ijinfomgt.2021.102466.

[14] T. Kumail, A. Mandić, H. Li, and F. Sadiq, "Uncovering the interconnectedness of tourism growth, green technological advancements and climate change in prominent Asian tourism destinations," *Tourism Management Perspectives*, vol. 53, Mar. 2023, doi: 10.1016/j.tmp.2024.101284.

[15] P. Kumari and D. Toshniwal, "Deep learning models for solar irradiance forecasting: A comprehensive review," *Journal of Cleaner Production*, vol. 318, p. 128566, Jul. 2021, doi: 10.1016/j.jclepro.2021.128566.

[16] H. Laaroussi, F. Guerouate, and M. Sbihi, "Deep Learning Approach and Topic Modelling for Forecasting Tourist Arrivals," *International Journal of Electrical and Computer Engineering Systems*, vol. 14, no. 4, pp. 411–419, 2023.

[17] Q. Luo, S. Cai, N. Lv, and X. Fu, "Daily forecasting of tourism demand: An ST-LSTM model with social network service co-occurrence similarity," *Information and Management*, vol. 62, no. 1, p. 104056, 2025, doi: 10.1016/j.im.2024.104056.

[18] A. Q. Munir, "Spatial weighting: modification of the spatial area proximity matrix weights," *International Journal of Information Technology (Singapore)*, 2024, doi: 10.1007/s41870-024-01865-3.

[19] A. Q. Munir, S. Hartati, and A. Musdholifah, "Early identification model for dengue haemorrhagic fever (DHF) outbreak areas using rule-based stratification approach," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 2, pp. 246–260, 2019, doi: 10.22266/IJIES2019.0430.24.

[20] Y. Nanjappa, V. Kumar Nassa, G. Varshney, B. Lal, S. Pandey, and A. V. Turukmane, "Improving migration forecasting for transitory foreign tourists using an Ensemble DNN-LSTM model," *Entertainment Computing*, vol. 50, p. 100665, Jan. 2024, doi: 10.1016/j.entcom.2024.100665.

[21] K. Nusair, F. Okumus, O. M. Karatepe, U. F. Alfarhan, and G. L. de Larrea, "From tourist motivations to buying decisions: A multilevel engagement perspective," *Tourism Management Perspectives*, vol. 48, p. 101148, Sep. 2022, doi: 10.1016/j.tmp.2023.101148.

[22] T. A. B. S. Pushkar, "IWF-ECTIC: Improved Wiener filtering and ensemble of classification model for tourism image classification Histogram of Oriented Gradients," *Multimedia Tools and Applications*, 2024, doi: 10.1007/s11042-024-20452-z.

[23] R. Soelaiman and M. H. Purnomo, "Implementation of recurrent neural network and boosting method for time-series forecasting," in *Proc. ICICI-BME*, Dec. 2009, doi: 10.1109/ICICI-BME.2009.5417296.

[24] P. Sugiartawan, S. Hartati, and A. Musdholifah, "Modeling of a tourism group decision support system using risk analysis based knowledge base," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, pp. 354–363, 2020, doi: 10.14569/IJACSA.2020.0110747.

[25] A. Supriatna, E. Hertini, B. Subtini, D. Susanti, and S. Supian, "Prediction of Tourist Arrivals To the Island of Bali With Holt Method of Winter and Seasonal Autoregressive Integrated Moving Average (Sarima)," *J. Sains Dasar*, vol. 6, no. 2, pp. 81–90, 2017.

[26] Y. Wang, Y. Zhao, and S. Addepalli, "Remaining useful life prediction using deep learning approaches: A review," *Procedia Manufacturing*, vol. 49, pp. 81–88, 2020, doi: 10.1016/j.promfg.2020.06.015.

[27] A. Witayangkurn, A. Arai, and R. Shibusaki, "Development of Big Data-Analysis Pipeline for Mobile Phone Data with Mobipack and Spatial Enhancement," *ISPRS International Journal of Geo-Information*, vol. 11, no. 3, 2022, doi: 10.3390/ijgi11030196.

[28] M. J. Yang and N. Zhu, "Online public opinion attention, digital transformation, and green investment: A deep learning model based on artificial intelligence," *Journal of Environmental Management*, vol. 371, p. 123294, Oct. 2024, doi: 10.1016/j.jenvman.2024.123294.

[29] S. Yang, Q. Li, D. Jang, and J. Kim, "Deep learning mechanism and big data in hospitality and tourism: Developing personalized restaurant recommendation model to customer decision-making," *International Journal of Hospitality Management*, vol. 121, p. 103803, May 2024, doi: 10.1016/j.ijhm.2024.103803.

[30] J. Zhang, F. Chen, Z. Cui, Y. Guo, and Y. Zhu, "Deep Learning Architecture for Short-Term Passenger Flow Forecasting in Urban Rail Transit," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7004–7014, 2021, doi: 10.1109/TITS.2020.3000761.

[31] Z. Zhang, F. Wang, and L. Deng, "Identifying node-corridor-network of tourist flow and influencing factors using GPS big data: A case study in Gansu and Qinghai provinces, China," *International Journal of Applied Earth Observation and Geoinformation*, vol. 135, p. 104271, Oct. 2024, doi: 10.1016/j.jag.2024.104271.

[32] Ayu Sofia, Dien, Z., & Erda, G. (2025). Forecasting International Tourist Arrivals to Indonesia Using LSTM: Post-Pandemic Analysis for 2024-2025. *Journal of Mathematics, Computations and Statistics*, 8(2), 316–324. <https://doi.org/10.35580/jmathcos.v8i2.7309>

[33] Julfia, F. T., Sibuea, S., & Satryawati, E. (2025). Penerapan Algoritma Neural Network dalam Prediksi Kedatangan Wisatawan Mancanegara di DKI Jakarta Melalui Pintu Masuk Bandara Soekarno-Hatta. *Joined Journal (Journal of Informatics Education)*, 8(1), 25. <https://doi.org/10.31331/joined.v8i1.3813>

[34] Junaidi, T., Hariyadi, M. A., & Abidin, Z. (2025). OPTIMASI METODE LONG SHORT-TERM MEMORY MENGGUNAKAN SELEKSI FITUR BORUTA UNTUK MEMPREDIKSI KEDATANGAN WISATAWAN. *JURNAL APLIKASI DAN INOVASI IPTEKS "SOLIDITAS" (J-SOLID)*, 8(1), 26–33. <https://doi.org/10.31328/js.v8i1.7089>

[35] Kanjanasupawan, J., & Chen, Y. (2019). Prediction of Tourist Behaviour : Tourist Visiting Places by Adapting Convolutional Long Short-Term Deep Learning. *2019 International Conference on System Science and Engineering (ICSSE)*, 12–17.

[36] Lisnawati, A., Alam, S., & Kurniawan, I. (2025). APLIKASI PREDIKSI KEDATANGAN WISATAWAN MANCANEGARA KE INDONESIA DENGAN ALGORITMA LONG SHORT-TERM MEMORY (LSTM). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(5), 8343–8349. <https://doi.org/10.36040/jati.v9i5.15055>

[37] Luo, Q., Cai, S., Lv, N., & Fu, X. (2025). Information & Management Daily forecasting of tourism demand : An ST-LSTM model with social network service co-occurrence similarity. *Information & Management*, 62(1), 104056. <https://doi.org/10.1016/j.im.2024.104056>

[38] Nanjappa, Y., Kumar, V., Varshney, G., Lal, B., Pandey, S., & Turukmane, A. V. (2024). Improving migration forecasting for transitory foreign tourists using an Ensemble DNN-LSTM model. *Entertainment Computing*, 50(November 2023), 100665. <https://doi.org/10.1016/j.entcom.2024.100665>

[39] Nonik Erawati, K., Sugiartawan, P., Nengah, N., Ardiani, D., Dewa, I., Bayu, A., Hartama, M., Ngurah, G., Frasetya, H., Gede, I., Mahendra, O., Erawati, K. N., Sugiartawan, P., & Ardiani, N. N. D. (2025). LSTM Neural Network for Predicting Tourist Arrivals to Bali. *Jurnal Sistem Informasi Dan Komputer Terapan Indonesia (JSIKTI)*, 8(1), 81–90. <https://infoteks.org/journals/index.php/jsikti/article/view/211>

[40] Song, M., & Abdullah, M. I. (2025). Application of LSTM in the dynamic changes of tourism market in Hunan Province. *Proceedings of 2025 2nd International Conference on Digital Economy, Blockchain and Artificial Intelligence, DEBAI 2025*, 36–40. <https://doi.org/10.1145/3762249.3762256>