JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 10 Number 1 | June 2025: 77-92
DOI: 10.15575/join.v10i1.1409

The Impact of Online Reviews to Predict The Number of International Tourists

Zhasa Vashellya[1], Erna Nurmawati[2], Teguh Sugiyarto[3]
[1,2] Computational Statistics Study Program, STIS Polytechnic of Statistics, Indonesia
[3]Directorate of Finance, Information Technology, Tourism Statistics, BPS-Statistics, Indonesia

Article Info

Article history:

Received June 22, 2024
Revised November 11, 2024
Accepted December 17, 2024
Published April 01, 2025

Keywords:

International tourists
Online Review
Sentiment Analysis
Prediction
LSTM

ABSTRACT

The tourism sector is a potential resource for advancing the Indonesian economy. The development of the tourism industry is represented by the number of international tourist arrivals. Therefore, this indicator becomes an objective in development programs. To accomplish this goal and assess the demand aspect of the tourism sector, it is a must to have a precise forecast of the number of international visitors. This research attempts to develop precise methods and models for estimating the number of international tourists based on this premise. This study chooses Bali Province as its focus since nearly half, or 47%, of the tourists who visit Indonesia arrive through the entry point in Bali Province. This research uses the LSTM method and big data online reviews in building prediction models. The results of this study show that sentiment analysis of tourist attractions in Bali using the BERT model has an accuracy of 75%. The results also depict that reviews by visitors about tourist attractions in Bali Province during the period 2012-2023 contain more positive sentiments. Furthermore, the best model to predict the number of international tourists, with the smallest RMSE and MAPE values (39,470.64 and 11.25%, respectively), includes inflation, rupiah exchange rates, TPK, monthly sentiment scores, and the number of reviews as dependent variables. The prediction model also show that the review variables (sentiment score and number of reviews) can improve prediction accuracy.

Corresponding Author:

Erna Nurmawati
Computational Statistics Study Program, STIS Polytechnic of Statistics
Jl. Otto Iskandardinata No. 63 C, Jakarta, Indonesia
Email: erna.nurmawati@stis.ac.id

1. INTRODUCTION

Indonesia boasts an extensive range of natural and cultural assets. Its attractiveness as a tourist destination is prominent among international travellers. In 2019, data from the United Nations World Tourism Organization (UNWTO) ranked Indonesia 28th out of 195 countries for the highest number of international tourist arrivals [1].

The tourism sector is one of the potential areas for Indonesia in advancing its economy. Data from the Organization for Economic Co-operation and Development (OECD) in the Tourism Trends and Policies 2022 report that in 2019, the tourism sector contributed 5% to Indonesia's Gross Domestic Product (GDP) [2]. This figure is higher than in previous years [3]. In addition, the tourism sector is positioned as a leading industry which is a superior sector in advancing the economy of the Indonesian people [4]. The growth of the tourism industry ultimately leads to the growth of the regional and national economy. These facts indicate the importance of tourism industry for Indonesia’s economy.

Tourists are individuals who travel outside their usual environment, usually for less than a year, with a purpose other than being employed by the local population of the visited country [5]. Tourists are divided into three, namely domestic tourists, national tourists and international tourists. Domestic tourists (wisnus) are the number of tourists visiting other regions within their own country. While national tourists (wisnas) are the number of tourists visiting other countries. And finally, international tourists are the number of tourists from other countries visiting a country[5]. According to data from the Central Statistics Agency (BPS), the number of Indonesian international tourist visits until October 2023 reached 978,499 visits. And almost half of these visits, namely as much as 47%, were international tourist visits through the arrival gate of Bali Province [6]. Based on these statistics, this study makes Bali Province the locus of research.

The importance of the number of international tourist arrivals is reflected in the President's attention to this indicator. In a meeting on economic activity and tourism after the revocation of PPKM at the State Palace in Jakarta, the President asked that the target number of international tourists must be achieved. This is supported by increasing the number of flights and the availability of aircraft seat capacity, organizing quality events in Indonesia and facilitating regulations including visas [7]. In addition, Minister of Tourism and Creative Economy (Menparekraf), Sandiaga Uno at The Weekly Brief on Monday, October 9, 2023 said that the target of international tourists in 2024 is 14 million tourists [8].

In achieving this target, Sandiaga Uno has carried out various strategic policies, one of the strategy is digital tourism. Digital tourism is an effective method in promoting various things related to tourism in Indonesia through various platforms on the internet [9]. To see the effectiveness of the policy of the Menparekraf, research is needed on the effect of digital tourism on achieving the target number of international tourists.

As a target of intervention, data on the number of international tourists is needed especially to identify the effectiveness of interventions or strategies implemented. Thus, the government and the tourism industry are in dire need of accurate modelled predictions of the number of international tourists for effective policy planning [10]. Accurate modelled predictions is useful for the development of medium to long-term international marketing and tourism strategies, pricing policies, investment plans and strategies, and the allocation of limited resources [11].

Given the importance of provision accurate prediction models, an approach that is better than models that have been developed previously is needed. Predictions that are usually made only use historical data while the number of international tourists is influenced by many factors. In addition, the model commonly used is a conventional statistical model that provides poor results [12]. For this reason, this research will use the deep learning method by utilizing other data sources, namely big data, in building a prediction model for the number of international tourists.

Big data is one of the data sources that is widely used in the digital era, especially in the field of tourism [13]. Big data is a very large set of data or records collected from various data sources with diverse formats such as text, images, sound, and raster [14]. Utilization of big data can be done to improve the accuracy of prediction [15]. This research utilizes one of the big data, namely online review text. Reviews or opinions of tourists who have visited a place greatly influence a person's desire to visit a place. When someone plans to go on vacation they often look for information related to their vacation destination including other people's assessment of the vacation spot [13]. So online review data can affect the number of international tourists.

In building a prediction model for the number of international tourists, there are various methods that can be utilized. One of them is the deep learning method. This method produces better accuracy compared to other statistical methods [12]. One of the best deep learning methods that can be used in predicting the number of international tourists is Long Short-Term Memory(LSTM) [16]. LSTM is a development of Recurrent Neural Network (RNN) model used to reduce vanishing gradient problem [17]. Vanishing Gradient is a problem when the gradient (weight value in the neural network) shrinks as the backpropagation process goes through time. LSTM is used in building a prediction model for the number of international tourists in Bali Province.

Based on the background description above, this research was conducted to build a prediction model for the number of international tourists in Bali Province by utilizing big data, namely online tourism reviews. The prediction is generated with deep learning techniques, namely Long Short-Term

Memory (LSTM). It is expected that utilizing online reviews and the LSTM method can improve the accuracy of the prediction model.

To sum up, the progress of the tourism industry is strongly influenced by the number of tourists. The presence of tourists stimulates the demand for goods and services, including the provision of accommodation, entertainment services, the culinary industry, and small souvenir-producing industries. In this sense, the sustainability of tourism industry is determined by the number of tourists, including international tourist. Data on international tourist arrivals must be available to arrange a better development planning and evaluate the existing programme. Since the development planning have to comprehend what will happen in the future condition, data about the future stance must also available. Therefore, this study focusses on the provision of an accurate prediction of international tourist arrival. This study is aimed to support the availability of valid international tourist data to arrange an effective development planning.

2. METHOD

The research methodology in general can be observed in the research flow Figure 1. The research stages begin with the process of collecting all the data used in this study. Furthermore, it is necessary to pre-process the data, divide the testing and training data. The next stage is the construction of the LSTM model, which is divided into two, namely with and without review variables (sentiment score and number of reviews). The division of these variables is done to see the effect of review variables on the accuracy of the LSTM model. After that, the model that has been formed is evaluated based on the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) values.

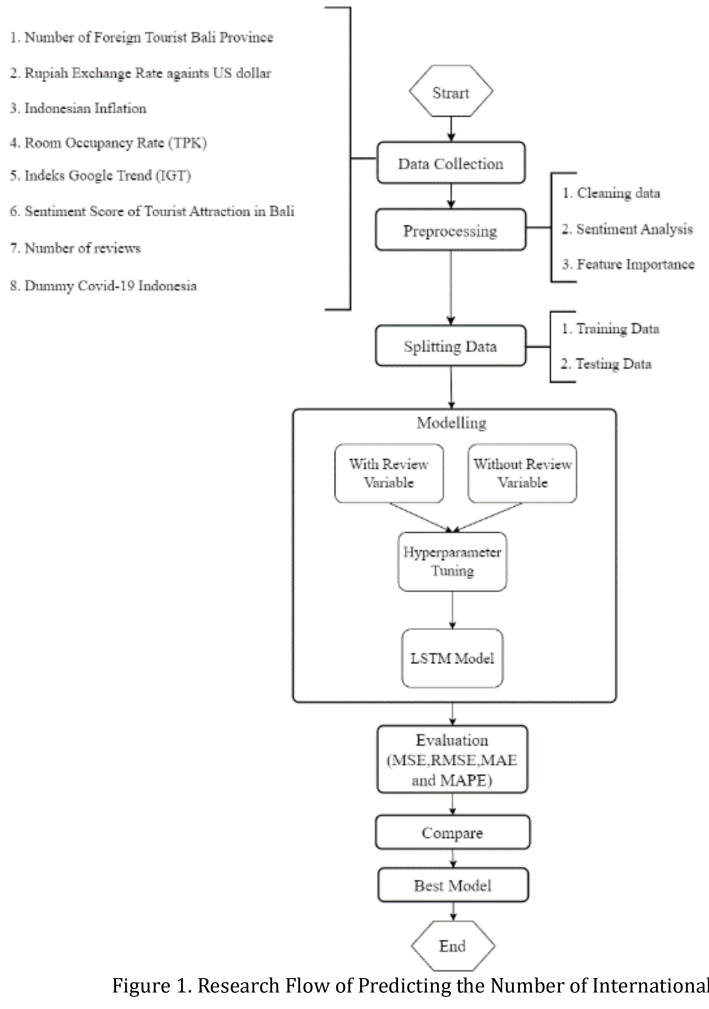

Figure 1. Research Flow of Predicting the Number of International Tourists

3.1. Data Collection

The data used in this study consists of the monthly number of foreign tourists in Bali Province. This data is obtained from the BPS Bali Province web page, https://bali.bps.go.id. Additionally, the monthly Room Occupancy Rate (TPK) data for Bali Province is also obtained from the BPS Bali Province web page. Both sets of data are obtained in .xlsx file format.

Other data used in this study include monthly Indonesian inflation data, which is obtained from the BPS RI web page, https://www.bps.go.id. Furthermore, data on the exchange rate of the US Dollar against the rupiah is used. This data is available daily and is obtained from the web page https://www.bi.go.id/. Both sets of data are also obtained in .xlsx file format.

Furthermore, Covid-19 dummy data. The data contains the number 0 in the month the Covid-19 event has been completed or has not occurred in Indonesia and contains 1 in the month Covid-19 occurred in Indonesia. The initial period of the pandemic in Indonesia began on March 31, 2020 based on Presidential Decree Number 11 of 2020 concerning the determination of the Corona Virus Disease 2019 Public Health Emergency in Indonesia [18]. Meanwhile, the pandemic was declared over based on the Presidential Decree of the Republic of Indonesia (Keppres) Number 17 of 2023 concerning the Determination of the End of the Status of the Corona Virus Disease 2019 (Covid-19) Pandemic in Indonesia on June 21, 2023 [19]. So covid data is defined as follows:

$$covid = \begin{cases} 1 & March\ 2020\ \leq t \leq June\ 2023 \\ 0 & t < March\ 2020, t > June\ 2023 \end{cases}$$

In this paper, we use Big Data as external variable to predict the International Tourist in Bali, Indonesia. One of the big data that we use is Google Trends Index (IGT). This data is taken monthly with the index value taken based on 6 keywords that match the tourism aspect [20]. The keywords are "bali tour", "bali hotel", "bali flight", "bali food", "bali tourism", "bali shop". This data was obtained on the web page https://trends.google.com/. Each of these six data is presented in the form of an .xlsx file.

The next Big data is the review data. This data was collected using web scraping techniques on the web page https://www.tripadvisor.com/.In this research, the web scraping technique is carried out by utilizing extension tools on Google Chrome called Tripadvisor® Review Scrapper. The data obtained will be reviews related to 54 tourist attractions that attract tourists in Bali Province every day. In this web scraping process, data is collected in the form of reviewer name, reviewer location, review time made, visiting time, rating, review title, and review text. In this research, review text and visiting time are used to generate variables, namely monthly sentiment scores and the number of reviews after sentiment analysis. These two variables are hereafter referred to as review variables in this study.

3.2. Data Preprocessing

Data preprocessing is carried out on online review data from scraping on the Tripadvisor platform. The results of scraping this data are still in the form of all-language text that needs to be filtered by removing Indonesian reviews using the langdetect python library. Furthermore, data cleaning is carried out, converting non-standard forms into standard, and text preprocessing, namely case folding, tokenization, stop word elimination, and stemming before sentiment analysis is carried out. The python library used is Natural Language Toolkit (NLTK).

Further pre-processing is done by adding some lag to the sentiment score data. Lag data is given to the variable because of there is a lag between the time for visiting and the time for writing reviews. The purpose of the lag variable to make the model work better. Lag data is given by shifting the sentiment score to the next few months. This is done because the n-th month sentiment can affect the number of n-1, n-2, etc. international touristss.

The last is the US Dollar exchange rate data against the rupiah. This data is initially presented in daily form. Because the prediction process of the number of international tourists is presented in monthly form, the exchange rate data is sought for its monthly average.

Next is data normalization. Before the model is built, the data is first normalized using Minmax scalar with the help of the sklearn python library. Data normalization is done because the variables used have different units and data ranges.

3.3. Sentiment Analysis

Sentiment analysis is done using a machine learning-based approach. This research uses the Bidirectional Encoder Representations from Transformers (BERT) model. The python library used in the construction of this BERT model is transformer with the help of hugging face, a machine learning library that can help the sentiment modelling process.

This stage begins by randomly selecting a sample of 10% per year for manual data labelling into 3 types of classifications, namely positive negative and neutral [21]. Taking 10% aims for time efficiency due to the large amount of data and to avoid inconsistencies if all data is labelled manually.

Samples that have been manually labelled are separated into 3 types of data, namely training data, validation data and testing data. Training data is used to train and build models, validation data to validate model performance and prevent overfitting and testing data to test the models that have been built. This data division is carried out by stratified random sampling with a ratio of 8:1:1 so that the classification division remains balanced [22].

The BERT model is developed using hyperparameter tuning to get the best model by combining several parameters, namely learning rate and number of batches and number of epochs [23]. This model is later used to label unlabelled data. After all the data is labelled, the monthly sentiment scoring is formed with the following formula:

$$\text{Monthly Sentiment Score } \frac{\text{positive}}{\text{positive+negative}} \quad (1)$$

3.4. Feature Importance

Feature importance refers to a technique that calculates scores for all input features for a given model. These scores represent the "importance" of each feature. A higher score means that a particular feature will have a greater effect on the model used to predict a specific variable. In this study, to demonstrate which variables have a significant effect on predicting the number of international tourists, feature importance using the random forest technique is employed, as this method is superior compared to other methods [24]. Feature importance is determined using the help of the Python library, sklearn.

3.5. Data Splitting

All the data collected during the data collection process will be split into two parts. These two parts are the training data and the testing data. In this study, the data splitting technique used is an 85:15 [25]. The data is split according to the monthly time sequence. In the data splitting process, the Python library used is sklearn.

3.6. Model Training

In the data training stage, the data is divided into 4 combinations of variables based on their feature importance values. These four combinations contain review variables (sentiment score or number of reviews). These four variable combinations are referred to as models with review variables. The model training is carried out and the best model is obtained. Next, the model with the best variable combination is removed from the review variable and seen which model is the best. This model is called the model without review variables.

a. Hyperparameter Tuning

Deep learning models have a set of hyperparameters. To find the optimal combination of hyperparameters, it is necessary to train the model using different sets of hyperparameters, which involve the number of epochs, the number of layer units, the learning rate, and the dropout rate [26]. The common approach to finding the best combination of hyperparameters is by using the grid search method [27], Grid search is an exhaustive search that trains the model in a manually specified hyperparameter space for every combination of hyperparameters.

Table 1. Hyperparameters Combined

No.	Hyperparameter	Size
1	Epoch	50,100,150
2	Unit Layer	150, 200, 300
3	Learning Rate	0,001; 0,01
4	Dropout Rate	0,1; 0,3

b. LSTM Model

Long Short-Term Memory (LSTM) is one of the modified algorithms of Recurrent Neural Network (RNN). In this study, the LSTM structure used to build the prediction model for the number of international tourists in Bali Province can be seen in Figure 2.

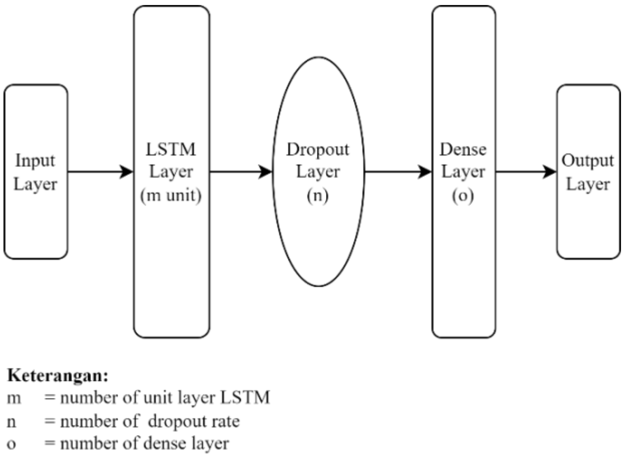

Figure 2. LSTM Model Architecture

Figure 2 shows the LSTM architecture used to build the prediction model for the number of international tourists in Bali Province. The values of m, n, and o correspond to the best parameters for each combination of variables used in the model during the hyperparameter tuning process.

3.7. Model Evaluation

To test the effectiveness of the proposed framework, four commonly used metrics—Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Squared Error (MSE)—are used to compare the performance of the model [26]. The smaller the MAE, MAPE, RMSE, and MSE, the more accurate the predicted results. MAE, MAPE, RMSE, and MSE are defined as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (2)$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (3)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (4)$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (5)$$

Formula description:

n = number of samples in the test set

y_i = the actual number of international tourists in the sample-i

$\hat{y}_i$ = number of predicted international tourists in the sample -i

3. RESULT AND DISCUSSION

3.1. Data Collection

Sentiment data collection on the Tripadvisor platform was carried out using the help of a Google Chrome extension, namely Tripadvisor Review Scraper. Scraping is done by collecting reviews from 54 tourist attractions that attract tourists spread across all districts/cities in Bali Province. These attractions are used to represent all tourist attractions in Bali. The number of reviews of each tourist attraction amounted to a maximum of 3,000 review data. Furthermore, the reviews are filtered and data other than Indonesian language data is taken. Data filtering is done to represent international tourists who use languages other than Indonesian in reading tourist attraction reviews. The results of review data that was successfully taken from tourist attractions based on districts / cities in Bali Province is shown in Table 2.

Table 2. Number of Tourist Attraction Reviews in Bali Province by District/City

District/City	Number of Attractions	Number of Reviews
Badung	12	19.440
Buleleng	5	3.033
Bangli	6	4.756
Denpasar	2	3.254
Gianyar	6	12.962
Jembrana	5	176
Kerangasem	9	9.145
Total	54	63.778

This total of 63,778 data was removed duplication of data and cut according to the required range, namely 2012 - 2023. So that the total data that will be included in the sentiment analysis model is 58,849 reviews. Manual labelling of 10% of data per year per type of language (English and languages other than English) is 5,868 data.

Beside review data, this paper collects another data such as monthly number of foreign tourists in Bali Province, Room Occupancy Rate (TPK) in Bali Province, Exchange Rate, Indonesian inflation data, Index Google Trends (IGT) from variety website. And the last data is dummy Covid-19 using a formula where the value is 1 for the month when Covid occurred and the value 0 for the months before and after Covid. All the data is collected monthly except for Exchange rate. The exchange rate data are converted into monthly form by finding the average value of the daily data to get monthly data. So all variables have 144 rows because the data series used is 12 years long.

3.2. Sentiment Analysis

After going through the preprocessing stage, the reviews are divided into 3 types of data, namely training data, testing data, and validation data. Data division is done by stratified random sampling based on sentiment classification due to the uneven number of positive, negative and neutral labelling in manual labelling data.

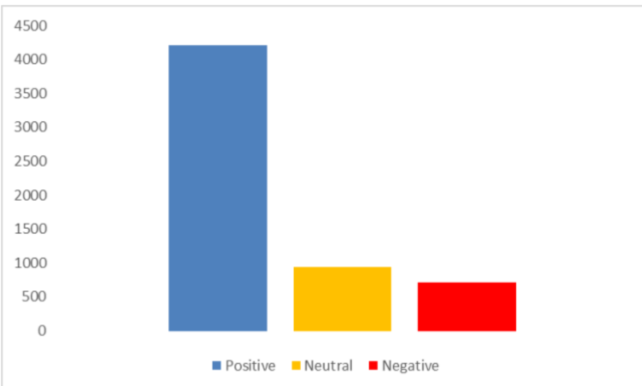

Figure 3. Number of Sentiment Classifications on Manual Labelling

It can be seen in Figure 3 that positive sentiments are much more than neutral and negative sentiments in manual labelling. This means that tourist attractions in Bali have the potential to attract foreign tourists to visit Indonesia, especially Bali. The division of training data, testing data, and validation data is done with a ratio of 8:1:1 as shown in Table 3.

Table 3. Sample Data Splitting

Number of Training Data (80%)	Number of Validation Data (10%)	Number of Testing Data (10%)
4.694	587	587

The sentiment model used is BERT which is developed by performing hyperparameter tuning. The recommended hyperparameter to get optimal performance is to use a batch size of 16 or 32 and a learning rate of 5e-5, 3e-5, 2e-5 with adam optimizer and epoch 2, 3, and 4 [26]. In this study, the batch size used is 16 because the batch size is the number of data samples used in one iteration, this refers to the ability of the computer so that it does not really affect the selection of performance. The selection of only one batch size aims at time efficiency due to the long computation time. This research also adds 5 epochs for tuning to get a broader picture.

Table 4. Hyperparameter Scenario

Scenario	Learning Rate	Epoch
1	2e-05	2
2	2e-05	3
3	2e-05	4
4	2e-05	5
5	3e-05	2
6	3e-05	3
7	3e-05	4
8	3e-05	5
9	5e-05	2
10	5e-05	3
11	5e-05	4
12	5e-05	5

The data division was used to perform several scenarios to select the optimal combination of learning rate and epoch to build the BERT model. From each scenario, a confusion matrix is created that is useful in calculating the performance of accuracy, precision, recall and f1-score values. The best scenario is the scenario with high values of these four performances. The results of the BERT model hyperparameter performance can be seen in Table 5.

Table 5. Scenario Performance

Learning rate	Epoch	Accuracy	Precision	Recall	F1-score
2e-05	2	0,74	0,71	0,74	0,69
2e-05	3	0,75	0,72	0,75	0,73
2e-05	4	0,67	0,74	0,67	0,69
2e-05	5	0,74	0,74	0,74	0,74
3e-05	2	0,75	0,70	0,75	0,71
3e-05	3	0,73	0,71	0,73	0,72
3e-05	4	0,73	0,75	0,73	0,72
3e-05	5	0,73	0,72	0,73	0,72
5e-05	2	0,75	0,72	0,75	0,73
5e-05	3	0,68	0,76	0,68	0,66
5e-05	4	0,75	0,73	0,75	0,73
5e-05	5	0,68	0,73	0,68	0,7

Based on Table 5, the scenario that has the highest accuracy, precision, recall and f1-score values is the scenario with a learning rate of 5e-5 and epoch 4. The selection of this scenario to be the best scenario can also be seen in the learning curve.

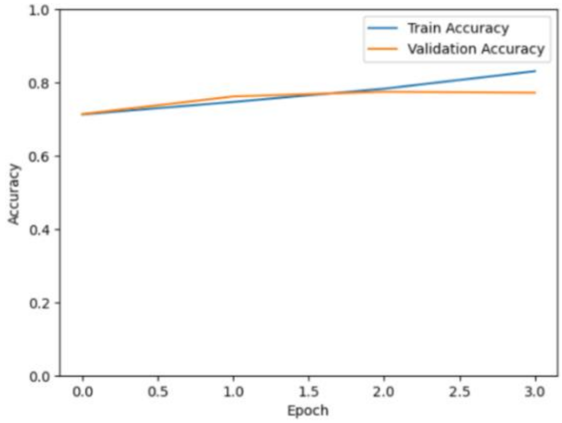

Figure 4. Learning Curve Sentiment Analysis

From Figure 4, the model has a good fit. This is because the training and validation data has a high enough accuracy. And the two lines formed are close and tend to increase. So the hyperparameters used in this research are based on Table 6.

Table 6. Selected Hyperparameters

No.	Hyperparameter	Size
1.	Batch Size	16
2.	Learning Rate	5e-5
3.	Epoch	4

Next, BERT modelling is performed using the selected hyperparameters as shown in Table 6. This model is applied to unlabeled data to get sentiment classification results. The following are the sentiment classification results of the entire data.

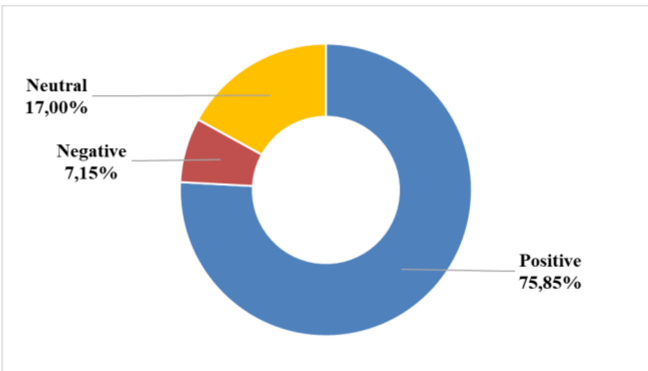

Figure 5. The Percentage of Positive, Negative, and Neutral Sentiments

Based on Figure 5, positive-sentiment reviews have a greater occurrence than negative and positive-sentiment reviews. Positive sentiment often appears because it is in line with the average rating of all reviews, which is 4.05, which leads to positive reviews.

Next calculate the monthly sentiment score using the sentiment of reviews that are in the same month. The sentiment score is calculated as the percentage of the number of positive reviews to the total number of positive and negative reviews in that month. This data is obtained from formula (1). Neutral sentiment is not included in the calculation because it is assumed that the neutral sentiment has no effect on the number of international tourists. The example of the monthly sentiment percentage results can be seen in Table 7.

Table7. Monthly Sentiment Score Percentage

Month	Percentage (%)
January 2023	87,30
February 2023	90,14
March 2023	82,35
April 2023	84,11
May 2023	86,61
June 2023	78,79
July 2023	84,38
August 2023	82,78
September 2023	82,69
October 2023	85,71
November 2023	88,41
December 2023	90,20

A monthly sentiment scores close to 1 means that positive reviews outweigh the total reviews, whereas if it is close to 0 then negative reviews outweigh the total reviews. The histogram of the monthly sentiment score distribution can be seen in Figure 6.

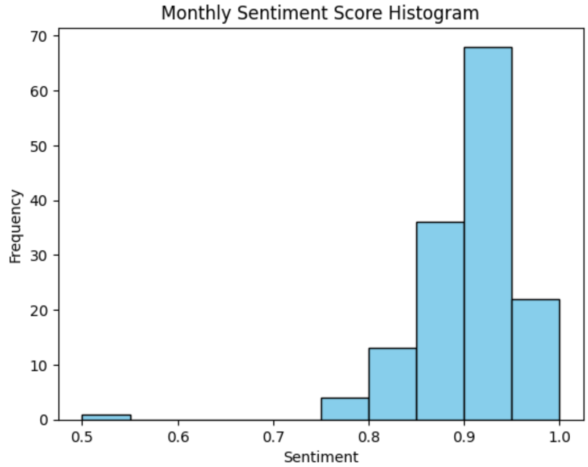

Figure 6. Histogram of Monthly Sentiment Score

From the histogram in Figure 6, most of the monthly sentiment score data tends to be greater than 0.5, which means that positive reviews outweigh negative reviews. This is in line with the results of scrapping data which shows that the most sentiment towards Bali tourist attractions is positive sentiment.

3.3. Feature Importance

The best feature selection for the LSTM modelling process is selected using feature importance. The first feature selection is the selection of which sentiment score data lag is most influential on the number of international tourists. The lag data sought are n-1, n-2, n-3, ..., n-12.

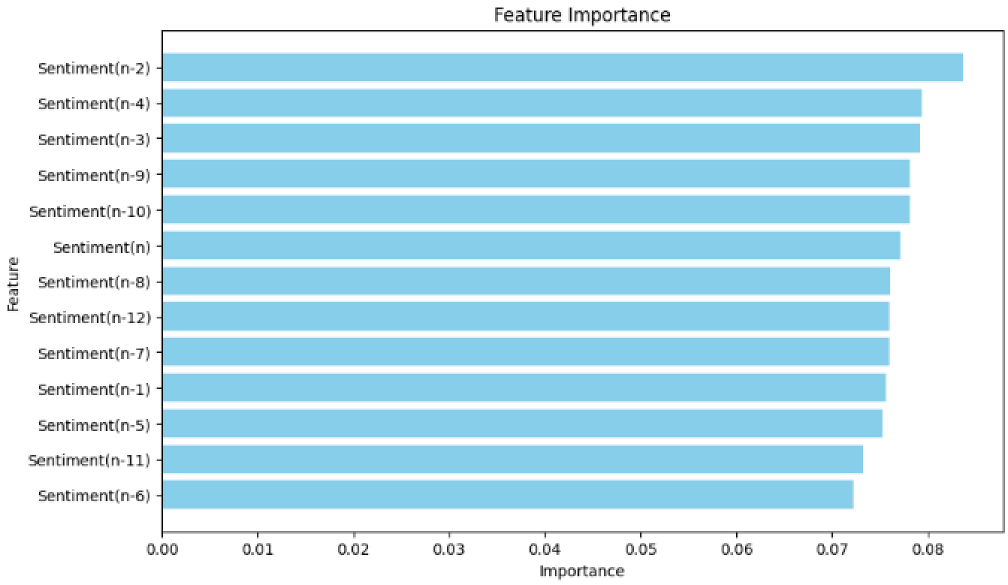

Figure 7. Monthly Sentiment Score Feature Importance

Based on Figure 7, the best feature is Sentiment(n-2), which means the best lag data is 2 months. This indicates that the sentiment of the current month will affect the number of international tourists in the following 2 months. Therefore, the sentiment score variable taken is sentiment(n-2).

Figure 7 also reveals that sentiment derived from reviews over a relatively long period still has an impact. This is due to the digital footprint of those reviews, which can still be accessed and referenced by potential tourists. For tourism industry players, especially tourist attraction managers, it is crucial to be cautious of this fact to ensure business sustainability and maintain/increase the popularity of tourist destinations.

The analysis of the data aligns with the current era of digitalization. It is highly likely that information spread on the internet becomes an important aspect in shaping the mindset of the population. The availability of adequate infrastructure supports high accessibility to information by the population. Additionally, the influence of the tech-savvy younger generation tends to utilize their capacity to make choices, including in determining travel destinations.

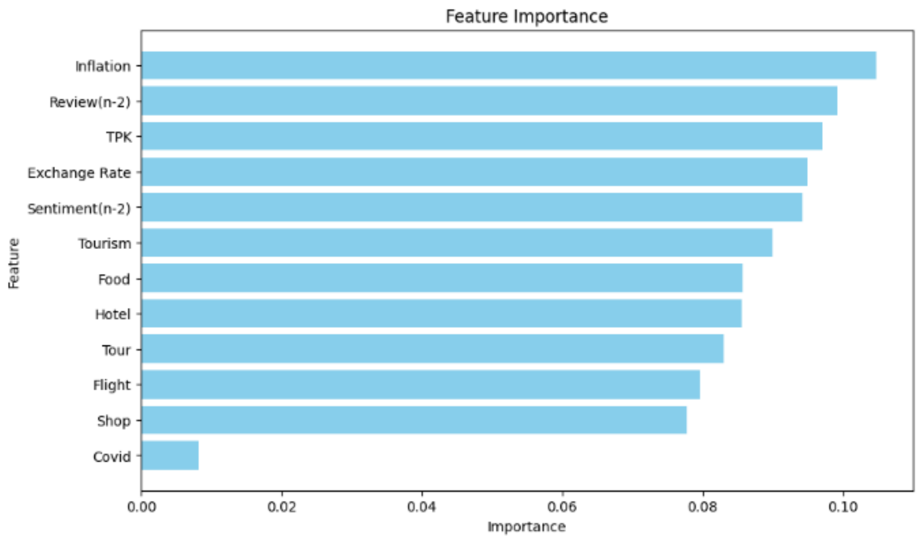

Figure 8. Feature Importance of All Variables

Furthermore, the feature importance for all research variables on the number of international tourists. Based on Figure 8, it can be seen that the best features in predicting the number of international tourists are inflation, the number of reviews on lag 2, TPK, and so on. With this important feature, 4 variable combinations are taken for modelling the number of international tourists in Bali Province as shown in Table 8. The four scenarios use review variables (sentiment score or number of reviews).

Inflation will generally affect purchasing power. The higher the inflation, the lower the purchasing power of consumers. Inflation can have an impact on the tourism industry both in the area of origin of tourists and tourist destinations. [28]. Inflation that occurs in the tourist's home region can reduce the tendency to travel. Likewise, inflation in tourist destination areas can also be taken into consideration by prospective tourists in determining travel destinations.

The variable number of reviews (n-2) and sentiment (n-2) confirms the importance of visitor experience spread through social media that can influence the image of tourist attractions and the decision of others to visit them. In fact, it can be concluded that the development of a tourist attraction does not only depend on its natural potential or attractiveness. However, word of mouth through reviews given by visitors is one of the determinants of the sustainability of the tourism industry in that destination [29].

Table 8. Combination Of Variables with Review Variables

No.	Number of Most Influential Variables	Variable Combination
1	12	Inflation, number of reviews, TPK, exchange rate, sentiment score, IGT (tourism, food, hotel, tour, flight, shop)
2	10	Inflation, number of reviews, TPK, exchange rate, sentiment score, IGT (tourism, food, hotel, tour, shop)
3	5	Inflation, number of reviews, TPK, exchange rate, sentiment score
4	3	Inflation, number of reviews, TPK

3.4. Prediction Modelling

Modelling the prediction of the number of foreign tourists in Bali Province is done with the LSTM method for each combination of variables. Before that, hyperparameters optimization for each combination of variables was done with grid search to find the best parameters with the smallest MSE value as shown in Table 9.

Table 9. Best Hyperparameter Combination of Variable with Review Variable

Variable Combination	Parameter
Inflation, number of reviews, TPK, exchange rate, sentiment score, IGT (tourism, food, hotel, tour, flight, shop)	{'dropout_rate': 0.1, 'epochs': 100, 'learning_rate': 0.01, 'num_units': 150}
Inflation, number of reviews, TPK, exchange rate, sentiment score, IGT (tourism, food, hotel, tour, shop)	{'dropout_rate': 0.1, 'epochs': 150, 'learning_rate': 0.001, 'num_units': 300}
Inflation, number of reviews, TPK, exchange rate, sentiment score	{'dropout_rate': 0.3, 'epochs': 150, 'learning_rate': 0.001, 'num_units': 300}
Inflation, number of reviews, TPK	{'dropout_rate': 0.3, 'epochs': 50, 'learning_rate': 0.01, 'num_units': 150}

After the best hyperparameter combination of each model is obtained, then the model is used to predict the number of foreign tourists in Bali Province with LSTM for all the combination variables as shown in Figure 9 until Figure 12.

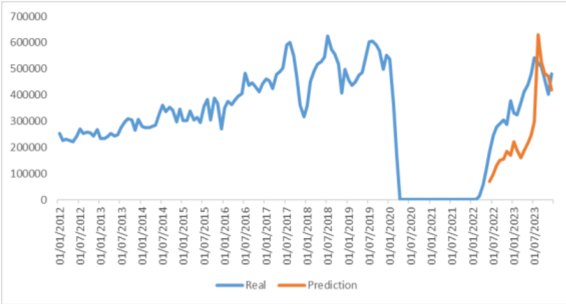

Figure 9. Graph of Prediction Results of 12 Variables with Review Variables

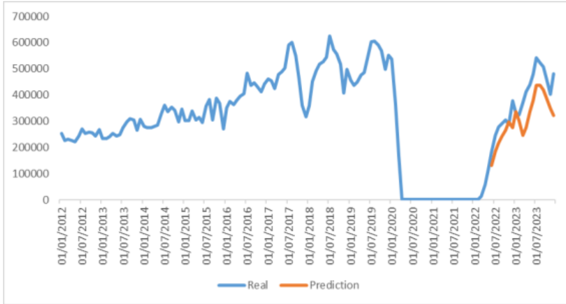

Figure 10. Graph of Prediction Results of 10 Variables with Review Variables

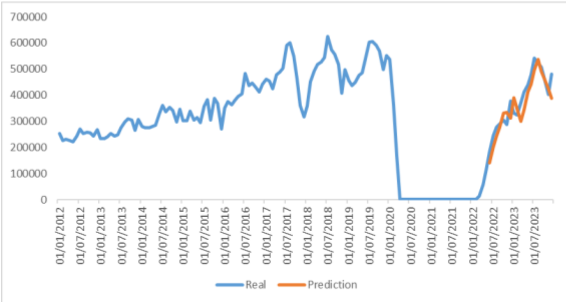

Figure 11. Graph of Prediction Results of 5 Variables with Review Variables

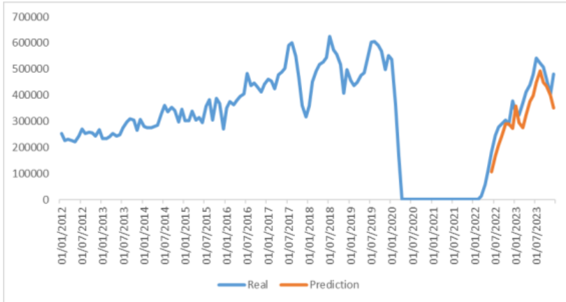

Figure 12. Graph of Prediction Results of 3 Variables with Review Variables

The best model is selected based on the smallest value of the model evaluation based on formula (2), (3), (4) and (5). The LSTM model with 5 feature importance result variables is the best LSTM model with a MAPE evaluation value of 11.25% as shown in Table 10. This shows that the best variables in predicting the number of foreign tourists in Bali Province are Indonesian inflation, rupiah exchange rate with the US dollar, TPK, number of reviews and monthly sentiment score.

Table 10. The Evaluation Value of The Variable Combination Prediction Model with The Review Variable

Variable	MAE	MAPE (%)	RMSE	MSE
12 variables	140.158.,43	39,55	155.704,62	24.243.929.321,32
10 variables	75.042,19	19,44	85.475,35	7.306.035.514,43
5 variables	39.470,64	11,25	45.568,58	2.076.495.208,33
3 variables	58.476,81	16,31	68.291,70	4.663.756.399,14

Furthermore, an LSTM model with the 5 best variables was built, but the review variables (sentiment score and number of reviews) were excluded so that only the inflation, exchange rate and TPK variables were used to see if the LSTM model with review variables could affect the prediction accuracy. The hyperparameter for this model is shown in Table 11.

Table 11. Best Hyperparameter Variable Combination Without Review Variable

Variable Combination	Parameter
Inflation, TPK, exchange rate	{'dropout_rate': 0.1, 'epochs': 100, 'learning_rate': 0.01, 'num_units': 200}

Next, the LSTM model is built with variables without sentiment.

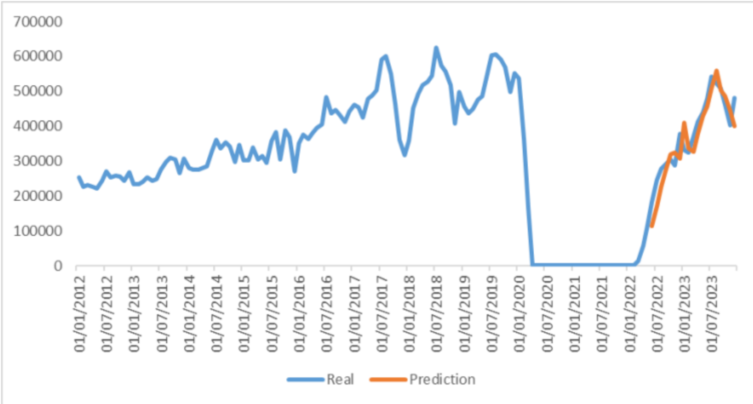

Figure 13. Graph of Prediction Results of 3 Variables without Review Variables

Table 12 shows the model evaluation comparison between the LSTM model with and without the review variable.

Table 12. Evaluation Value of Variable Combination Prediction Model without Review Variable

Variable	MAE	MAPE (%)	RMSE	MSE
With Review Variable	39.470,64	11,25	45.568,58	2.076.495.208,33
Without Review Variable	42.151,58	12,58	49.909,66	2.490.974.238,02

Based on Table 12, the model with the review variable provides a better evaluation value than the model without the review variable. So, it is evident that the presence of review variables, namely sentiment score and number of reviews, can improve the performance of the LSTM model.

4. CONCLUSION

The data analysis in this study shows the influence of online reviews by tourists in determining the number of foreign tourists. The penetration of internet usage has contributed to changes in the way people plan to travel. It is common that people rely on the available online information to be the main reference in deciding destination. The available information on social media or websites is often more representative to capture current conditions of tourist destinations. In addition, online information provide various options including figures or videos not just news narratives.

This research utilizes one of the deep learning methods, LSTM, in predicting the number of international tourists. The data analysis reveals a good result. This finding might be used as an approach for providing an accurate prediction in the future. However, improvements must still be made, especially to provide prediction for longer periods.

This study is also useful for policymakers or business owners. For planners or policy makers this method can be considered as one solution in providing future predictions of demand in the tourist industry. For service providers, this finding should be a concern. The sustainability of tourism businesses can be affected by the existence of online reviews by tourists. Positive reviews can create high demand and vice versa. Negative reviews that spread easily in cyberspace will reduce demand.

To maintain the sustainability of businesses, the services providers might do two main things. First, using online media or digital marketing as one of the marketing strategies. The quantity analysis of this study clearly shows the importance of information disclosure in cyberspace in the form of online reviews that can increase the accuracy of foreign tourist predictions. Second, tourism industry stakeholders must be responsive to tourists' online reviews. Negative reviews must be immediately answered with explanations and improvements. This step may reduce the unexpected impact of negative sentiment.

ACKNOWLEDGEMENTS

This research was supported and funded for publishing by STIS Polytechnic of Statistics, Indonesia. We thank our colleagues from STIS Polytechnic of Statistics who provided insight and expertise that greatly assisted the research.

REFERENCES

[1] UNWTO, "UNWTO | World Tourism Organization a UN Specialized Agency." 2019. [Online]. Available: https://www.unwto.org/

[2] OECD, "OECD Data," *OECD website*. 2021. [Online]. Available: https://data.oecd.org/

[3] BPS, "Proporsi Kontribusi Pariwisata Terhadap PDB," *Https://Www.Bps.Go.Id/Indicator/16/1188/1/Proporsi-Kontribusi-Pariwisata-Terhadap-Pdb.Html*. 2021.

[4] S. Wibowo, O. Rusmana, and Z. Zuhelfa, "Pengembangan Ekonomi Melalui Sektor Pariwisata Tourism," *J. Kepariwisataan Destin. Hosp. dan Perjalanan*, vol. 1, no. 2, pp. 83–89, 2017, doi: 10.34013/jk.v1i2.13.

[5] IRTS, *Role of the International Recommendations for Tourism Statistics 2008*. 2017. doi: 10.18356/05265168-en.

[6] Badan Pusat Statistik, "Kunjungan Wisatawan Mancanegara per bulan Menurut Kebangsaan (Kunjungan) 2023." pp. 335–58, 2023. [Online]. Available: https://www.bps.go.id/linkTableDinamis/view/id/960.

[7] Kemenparekraf, "Siaran Pers_ Menparekraf_ Presiden Minta Target Wisman dan Wisnus Harus Tercapai."

[8] Kemenparekraf, Indonesia. *The Weekly Brief with Sandi Uno - 9 Oktober 2023*, (2023). [Online]. Available: https://www.youtube.com/live/kikH5cMHhbI?si=XH03Ns_K814we-Rf

[9] Kemenparekraf, "Strategi Digital Tourism dalam Menggaet Wisatawan," *Kemenparekraf.Go.Id*. 2021. [Online]. Available: https://kemenparekraf.go.id/ragam-pariwisata/Strategi-Digital-Tourism-dalam-Menggaet-Wisatawan

[10] P. F. Pai, K. C. Hung, and K. P. Lin, "Tourism demand forecasting using novel hybrid system," *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3691–3702, 2014, doi: 10.1016/j.eswa.2013.12.007.

[11] T. Dergiades, E. Mavragani, and B. Pan, "Google Trends and tourists' arrivals: Emerging biases and proposed corrections," *Tour. Manag.*, vol. 66, pp. 108–120, 2018, doi: 10.1016/j.tourman.2017.10.014.

[12] J. Lemmel *et al.*, "Deep-Learning vs Regression: Prediction of Tourism Flow with Limited Data," 2022, [Online]. Available: http://arxiv.org/abs/2206.13274

[13] J. Li, L. Xu, L. Tang, S. Wang, and L. Li, "Big data in tourism research: A literature review," *Tour. Manag.*, vol. 68, pp. 301–323, 2018, doi: 10.1016/j.tourman.2018.03.009.

[14] G. Management, B. Data, and W. Model, "Research in Business and Gossip Management at Universities using Big," vol. 5, no. 1, pp. 1–14, 2016.

[15] E. Park, J. Park, and M. Hu, "Tourism demand forecasting with online news data mining," *Ann. Tour. Res.*, vol. 90, p. 103273, 2021, doi: 10.1016/j.annals.2021.103273.

[16] H. Laaroussi and F. Guerouate, "Deep learning framework for forecasting tourism demand," in *2020 IEEE international conference on technology management, operations and decisions (ICTMOD)*, IEEE, 2020, pp. 1–4.

[17] S. C. Hsieh, "Tourism demand forecasting based on an lstm network and its variants," *Algorithms*, vol. 14, no. 8, 2021, doi: 10.3390/a14080243.

[18] Presiden RI, "Keputusan Presiden RI No. 11 Tahun 2020," *Fundam. Nurs.*, no. 01, p. 18=30, 2020.

[19] P. P. Indonesia, "Keputusan Presiden Nomor 17 Tahun 2023 tentang Penetapan Berakhirnya Status Pandemi Corona Virus Disease 2019 (COVID-19) di Indonesia," *Jakarta Pemerintah Pus.*, no. 167292, pp. 1–3, 2023.

[20] X. Li, B. Pan, R. Law, and X. Huang, "Forecasting tourism demand with composite search index," Tour. Manag., vol. 59, pp. 57–66, 2017, doi: 10.1016/j.tourman.2016.07.005.

[21] T. Bey and I. K. Ari, "Implementasi BERT pada Analisis Sentimen Ulasan Destinasi Wisata Bali," *J. Elektron. Ilmu Komput. Udayana*, vol. 12, no. 2, pp. 409–420, 2023.

[22] G. Danilov, T. Ishankulov, K. Kotik, Y. Orlov, M. Shifrin, and A. Potapov, "The classification of short scientific texts using pretrained BERT model," *Public Heal. Informatics Proc. MIE 2021*, vol. 0, pp. 83–87, 2021, doi: 10.3233/SHTI210125.

[23] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, no. Mlm, pp. 4171–4186, 2019.

[24] H. Alsagri and M. Ykhlef, "Quantifying feature importance for detecting depression using random forest," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 5, pp. 628–635, 2020, doi: 10.14569/IJACSA.2020.0110577.

[25] R. I. Agustin, "Peramalan Data Intermiten Menggunakan Metode Autoregresive Integrated Moving Average dan Neural Network," *Dep. Sist. Inf. Fak. Teknol. Inf. Dan Komun. Inst. Teknol. Sepuluh Nop. Surabaya*, 2018.

[26] Z. Wang *et al.*, "Climate and environmental data contribute to the prediction of grain commodity prices using deep learning," *J. Sustain. Agric. Environ.*, vol. 2, no. 3, pp. 251–265, 2023, doi: 10.1002/sae2.12041.

[27] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," *ACM Int. Conf. Proceeding Ser.*, vol. 227, no. 2006, pp. 473–480, 2007, doi:

10.1145/1273496.1273556.

[28] U. Khalid, L. E. Okafor, and M. Shafiullah, “The Effects of Economic and Financial Crises on International Tourist Flows: A Cross-Country Analysis,” *J. Travel Res.*, vol. 59, no. 2, pp. 315–334, 2020, doi: 10.1177/0047287519834360.

[29] A. Suryani, S. Soedarso, D. Rahmawati, E. Endarko, A. Muklason, and B. M. Wibawa, “Why What they Say Matters: The Impacts of Visitors’ Experiences on Tourism Sustainability,” *Int. J. Soc. Sci. Bus.*, vol. 5, no. 1, pp. 99–110, 2021, doi: 10.23887/ijssb.v5i1.31355.