International Journal of Engineering, Science and InformationTechnology
Volume 5, No. 1 (2025) pp. 101-108
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i1.667
Research Paper, Short Communication, Review, Technical Paper

Performance Analysis of SVM and Linear Regression for Predicting Tourist Visits in North Sumatera

Andriyan Ginting[1], Nurdin[2*], Cut Agusniar[1]

[1]*Departement of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia*
[2]*Departement of Information Technology, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia*

**Corresponding author Email: nurdin@unimal.ac.id*

The manuscript was received on 20 June 2024, revised on 25 August 2024, and accepted on 2 December 2024, date of publication 6 January 2025

Abstract

Indonesia, an archipelago rich in cultural diversity, historical heritage, and stunning natural scenery, offers an extraordinary travel experience to visitors who make this country their vacation destination. Tourism in Indonesia plays an essential role in the domestic economy, contributing to Gross Domestic Product. With its abundant natural and cultural resources, North Sumatra has long been recognized as an attractive destination for foreign tourists. However, the tourism sector faces significant challenges related to fluctuations in the number of visits, mainly due to the impact of the COVID-19 pandemic, which has disrupted global travel patterns and caused considerable uncertainty in tourism forecasting. Therefore, predicting the number of tourist visits becomes crucial for effectively planning and managing tourist destinations. This research aims to compare the performance of two forecasting algorithms, SVM and linear regression, in predicting foreign tourist visits in North Sumatra using historical data from 2019 to 2023. The dataset was subjected to a preprocessing phase to ensure data cleanliness and consistency, focusing on key variables such as seasonal trends, external factors, and market dynamics. Both models were evaluated based on two commonly used accuracy metrics, MAPE and RMSE, to assess how well the models could predict actual tourist arrivals. The results of the study indicate that Linear Regression outperforms SVM in terms of prediction accuracy, with a MAPE of 42.40% and an RMSE of 6735.6, compared to SVM with a MAPE of 46.65% and an RMSE of 8020.42. These findings provide valuable insights for local government authorities and tourism industry stakeholders to enhance destination planning, resource allocation, and strategies to attract more foreign tourists in the post-pandemic era.

Keywords: *Linear Regression, Prediction, Support Vector Machine, Tourism.*

1. Introduction

Indonesia, an archipelago rich in cultural diversity, historical heritage, and stunning natural scenery, offers an extraordinary travel experience to visitors who make this country their vacation destination. With more than 17,000 islands spread along the equator, Indonesia's tourism is ranked 32nd globally [1]. Tourism in Indonesia is essential to the domestic economy, contributing to a Gross Domestic Product (GDP) of 4.2% in 2021, an increase from the previous year of only 4.0%. This sector also contributes significantly to the country's foreign exchange, which is recorded to generate USD 7.03 billion in 2021 [2].

North Sumatra, as one of the central provinces of Indonesia, has a variety of tourist destinations that attract domestic and international tourists. In 2019, the number of foreign tourists visiting was recorded at 258,822, although there was a significant decline in 2020 due to global travel restrictions. In 2022, foreign tourists increased again to 148,997 people [3]. This tourism sector has a positive impact on the regional economy, creating jobs and improving the income of local communities.

However, the tourism sector in North Sumatra also faces challenges, especially related to fluctuations in the number of foreign tourist visits, which can impact the policies implemented and the readiness of tourist services. This problem is increasingly complex due to the COVID-19 pandemic, which causes uncertainty in predicting the number of tourists who will come. This is a challenge for the government and tourism industry players in planning the right strategy to manage tourist visits [4].

Technology and data mining advances have brought about a significant transformation in the tourism sector. Using visitor data collected from various travel platforms, analysis through data mining techniques can provide more accurate information about trends in tourist visits. The process of using statistical, artificial intelligence, and machine learning methods to uncover valuable insights from large data sets called data mining [5]. One of the main applications of data mining in tourism is forecasting, which is useful for predicting future trends in tourist arrivals so that tourism sector managers can make more effective decisions.

Various methods can be used to forecast the number of foreign tourist arrivals, including linear regression, Support Vector Machine (SVM), and other methods. Linear regression models linear relationships between variables, while SVM is more effective in handling non-linear data and classification problems [6]. Based on previous research, such as that conducted by Karim et al., which compares SVM, Random Forest, and K-Nearest Neighbor methods, SVM shows better prediction accuracy.

This research is essential to provide more precise and accurate information in tourism policy planning, which will assist the government and industry players in planning development following the predicted number of tourists. With more accurate predictions, resources can be allocated efficiently, marketing campaigns can be customized, and tourism products can be developed according to market needs. To help shape more specific policies and boost local economies, this research will examine how well different algorithms forecast the amount of international tourists visiting North Sumatra.

2. Literature Review

2.1. Previous Research

Research by Manurung et al. (2022) used data on tourist visits to Toba Regency, North Sumatra, from 2013–2022, and a linear regression technique to forecast future visitor numbers. The predictor variable used is the number of events organized, while the response variable is the number of tourists. The prediction results showed that with 19 events in 2023, there would be an estimated 5,270,330 tourists, with an accuracy rate of 94.79% and a MAPE of 5.21%, categorized as good [7].

Fahirah and Sutresna Wati's (2020) study examined various methods for predicting the influx of international tourists to Indonesia. These methods included SVR, Backpropagation Neural Network, Exponential Smoothing, Weighted Moving Average, and Weighted Exponential Smoothing. According to the results, SVR offers the highest level of accuracy while having the lowest MAPE value, which is 2.5614%. SVR is the most successful than traditional and neural network approaches, with an estimated 164,723 international visitors visiting Indonesia in May 2020 using the optimum settings [8].

This is evidenced by the study titled "Application of simple linear regression method to predict the arrival of the number of foreign tourists after the COVID-19 pandemic" conducted by Binti Karomah. Based on data collected from the Central Statistics Agency (BPS) between 2015 and 2021, this research uses a simple linear regression approach to forecast the amount of international visitor arrivals after the COVID-19 pandemic. The forecast indicates that visitors will skyrocket from 23,167,450 in 2022 to 29,584,805 in 2026. Mean Absolute Percentage Error (MAPE) measures forecasting accuracy; a value of 0% indicates excellent accuracy [9].

Study Titled "Prediction of Foreign Travelers during the COVID-19 Pandemic" by Dewi Rahmawati et al. (2022). This study uses the basic linear regression approach to forecast the number of international visitors to Indonesia during the current COVID-19 epidemic. In this research, the number of COVID-19 cases in Indonesia is used as a dependent variable, and the number of travelers as an independent variable is used to predict the number of foreign passengers to Indonesia using simple linear regression. The data is used from 2020 to 2021, specifically from January to April. According to the analysis findings, there is a coefficient of determination of 26.1% and a regression value of 471,296.2073 and 3.010223266. The regression equation Y = 471,296.2073 + 3.010223266X may forecast the number of international tourist visits based on the number of COVID-19 instances, albeit it only explains 26.1% of the dependent variable's diversity [10].

This was found in the study "Support Vector Regression to Predict the Number of Foreign Tourist Visits on the Island of Bali" by Vivi Nur Wijayaningrum and Novi Nur Putriwijaya. In this work, we utilize time-series data on tourist arrivals at Ngurah Rai Airport from 2008 to 2017 to forecast the number of foreign visitor arrivals in Bali Island using the Support Vector Regression (SVR) method. Six SVR parameters were tested after data normalization to find the best ones. The ideal values for iteration, lambda, epsilon, cLR, C, and sigma were 200, 3, 0.0002, 0.00007, and 0.25, respectively, yielding a MAPE of 14.33% and an RMSE of 0.19010. The algorithm's effectiveness is heavily influenced by the quantity of training data utilized, making the selection of the correct k-fold value crucial [11].

In the study "Comparison Of Support Vector Machine And Backpropagation Models In Forecasting The Number Of Foreign Tourists In Bali Province," Imelda Alvionita Tarigan et al. (2021). With monthly visit data from 2007–2019, this research intends to assess the performance of two models—Support Vector Machine (SVM) and Backpropagation—in predicting the number of international visitors in Bali Province. Before being split into training data (2007–2016) and test data (2017–2019), the data was first adjusted using the min-max approach on a scale of 0–1 [12].

In the research of Widya Rizka Ulul Fadilah et al. (2020) with Prediction Analysis of Share Prices of PT Telekomunikasi Indonesia Using the Support Vector Machine Method. This research compares the Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) methods to predict the stock price of PT Telekomunikasi Indonesia using daily data from April 29, 2015 to April 28, 2020. After preprocessing, the data is divided into 80% training and 20% testing data. As a result, the SVM method with RBF kernel produces an accuracy of 0.9641 and RMSE of 0.0932, while KNN has an accuracy of 0.945 and RMSE of 0.1162. Thus, SVM is better than KNN in predicting stock prices [13].

Supported by the results of Abdul Karim's research entitled "Comparison of Poverty Prediction in Indonesia Using Support Vector Machine (SVM) with Linear Regression" (2020). This study uses BPS 2018 data to compare the effectiveness of linear regression and the Support Vector Machine (SVM) approach in forecasting poverty in Indonesia. The proportion of poor people is the dependent variable, while HDI by province is the independent variable. The linear regression results show that HDI affects poverty negatively; as HDI increases, the poverty rate decreases. The SVM approach outperforms linear regression in accuracy by reducing the root-mean-squared error (RMSE). The best support vector machine (SVM) model generates predictions that closely resemble the original data when the epsilon and optimum cost parameters are used. Finally, compared to linear regression, SVM is superior in predicting poverty in Indonesia [14].

2.2. Tourism

Tourism usually includes tourist activities that include travel, lodging, and recreational activities in a particular location. Tourism boosts the economy, creates new jobs, and connects cultures and information between countries. In addition, tourism has a direct effect on a country's infrastructure development. Tourism helps connect information and culture between countries. Travelers and locals can help understand other cultures tolerate, and exchange knowledge [15].

2.3. Data Mining

To assess and automate information, data mining employs several machine learning approaches. Data mining can identify hidden patterns and relationships in large data sets, which can lead to new knowledge or insights previously unseen in the database [16]. The process involves in-depth data analysis to discover trends, associations, or other valuable information that can be used to improve decision-making, strategic planning, and data-driven policy development.

2.4. Forecasting

Prediction is the process of estimating possible future events using historical data and current information to minimize errors [17]. Forecasting aims not to provide an absolute prediction but to find the most accurate estimate of what is likely to happen [18]. By combining historical data, statistical models, and related analysis techniques, forecasting attempts to predict trends or patterns that may occur in the future. While forecasts are not guaranteed to be one hundred percent accurate, they can provide valuable insights for organizations to plan appropriate actions and be better prepared for potential upcoming changes or situations.

2.5. Support Vector Machine

Support Vector Machine (SVM) is a relatively new classification and regression prediction method. SVM belongs to the supervised learning group, which requires a step-by-step training phase before proceeding with the testing phase to evaluate the trained model [19]. Linear kernels, polynomial kernels, radial basis function (RBF) kernels, and sigmoid kernels are the four main kinds of SVMs that may be used for calculations [20]. Here is the hyperplane equation systematically:

$$f(x) = w^t \emptyset(x) + b \quad \ldots\ldots(1)$$

Description:
w = weight vector
x = input vector
$\emptyset(x)$ = mapping function to a higher feature space, often used in the context of SVM kernels to handle non-linearly separable data
b = bias

2.6. Linear Regression

Statistically employ linear regression to describe the connection between one dependent variable (response variable) and one independent variable (predictor variable). Linear regression has its roots in the late 19th century, when Sir Francis Galton observed "regression to the mean" and recorded his findings. Two English mathematical statisticians, Karl Pearson and Ronald Fisher, later improved upon this idea. Statistically employ linear regression to describe the connection between one dependent variable (response variable) and one independent variable (predictor variable). Linear regression has its roots in the late 19th century, when Sir Francis Galton observed "regression to the mean" and recorded his findings. Two English mathematical statisticians, Karl Pearson and Ronald Fisher, later improved upon this idea. The linear regression method is often used when trying to ascertain the strength of a cause-and-effect connection between two variables [21]. The basic equation of linear regression is as follows:

$$y = a + bx \quad \ldots\ldots(2)$$

Description:
y = dependent variable
x = independent variable
a = intercept
b = slope

3. Research Method

There are a lot of steps to the system design that compare Linear Regression with Support Vector Machine for forecasting the amount of international tourists visiting North Sumatra in Figure 1.

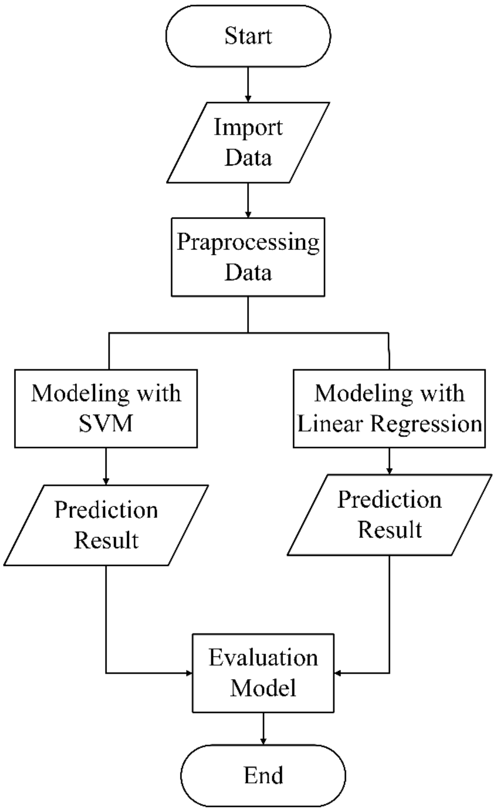

Fig 1. System Schematic

Description:

1. Import data that will be used for modeling and designing later.
2. After the data is cleaned up to remove missing values, it is split into training and testing data. The goal of using training data is to teach the model how to make correct predictions, whereas the goal of testing data is to ensure the model can perform well on data it has never seen before.
3. A prediction model is constructed from normalized and partitioned data using two support vector machine (SVM) methods and linear regression.
4. Predictions are generated by the model using the testing data. The final product looks like a graph.
5. Determine the model's efficacy by comparing the two algorithms' ability to foretell the arrival of international visitors to North Sumatra. This exercise will use RMSE and MAPE (Mean Absolute Percentage Error).

3.1. Data Import

Before applying the linear regression and Support Vector Machine (SVM) algorithms to predict the number of foreign tourist arrivals in North Sumatra, the first step is to analyze historical data on foreign tourist arrivals from January 2019 to December 2023. This analysis aims to understand the period's patterns, trends, and fluctuations.

3.2. Preprocessing Data

Data cleansing, addressing missing values or outliers, and format conversion are all part of the preprocessing phase of data analysis, which is crucial for preparing the data for the prediction model. After eliminating redundant data, missing value handling leaves specific values at 0, representing out-of-the-ordinary occurrences like COVID-19. Data transformation facilitates the model's processing of non-numerical data kinds. The next step is to split the data into training and test sets, typically 80:20, to evaluate how well the model does on new data, make it more generalizable, and avoid overfitting.

3.3. Import Library

Several libraries must be imported before the model development process to facilitate the work. The libraries used in this model are as follows:

1. Developers may utilize the open-source Streamlit library to create interactive web apps for data visualization, machine learning, and data science projects. Without worrying about complicated server setups, Streamlit makes it easy to create user interfaces and data visualizations in a flash.
2. Pandas is one of the most well-known libraries for working with and analyzing data. Among its data structures are DataFrames and Series, which facilitate activities like reading, cleaning, and processing data from various formats (e.g., CSV, Excel, SQL, etc.) and working with tabular data (rows and columns).
3. You may rely on the NumPy library for fast and accurate numerical calculations. It has several functions in mathematics, statistics, and linear algebra, and it also includes multidimensional array objects, which are great for working with numerical data.
4. Data visualizations, including line charts, bar graphs, pie charts, scatter plots, and histograms, may be done with the help of the Matplotlib package.

5. One of the most well-known machine learning libraries is Scikit-learn. Machine learning techniques for data processing, clustering, regression, and classification are all part of it. Additionally, Scikit-learn provides data preparation, model selection, validation, and assessment resources.

3.4. Building the Support Vector Machine (SVM) Model

Support Vector Machine (SVM) model to predict the number of foreign tourist arrivals by analyzing patterns in historical data. SVMs work by separating data using a line or hyperplane that maximizes the margin between different categories, allowing detection of trends and dynamic changes in the data. Implementation of SVM models can be done using libraries such as Scikit-learn, which facilitates model building, training, and evaluation. The selection and optimization of parameters such as kernel, epsilon, and gamma are essential to improve accuracy in capturing seasonal patterns and short-term trends. Once the model is built, the prediction results are evaluated to ensure the accuracy and effectiveness of the model.

3.5. Building the Linear Regression Model

This model uses linear regression to recognize patterns in historical data and predict future values by finding the best linear relationship between variables. The model can adjust to changes in the data, resulting in accurate predictions. Linear regression can be built using libraries such as Scikit-learn, which eases the process of building, training, and evaluating the model. Once the model is complete, the predicted results are assessed, and adjustments are made to improve accuracy if necessary.

3.6. Model Evaluation

Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are two standard assessment metrics used in model evaluation to quantify model performance. You may measure the model's prediction accuracy using these measures. A comparison between the actual data and the model's expected outcomes is integral to this assessment process. Visualizing the prediction results compared to the test data also helps evaluate the model's pattern recognition capabilities and pinpoint regions of imprecision. Graphs comparing predictions with actual data are also used to illustrate the model's efficacy further.

2. Result and Discussion

4.1. Model Support Vector Machine

This model has several parameters that must be determined, but the parameters used to show the test data prediction results are the best. The following are the prediction results using the Support Vector Machine (SVM) model. The prediction results of foreign tourist visits in North Sumatra are displayed in a table containing months, actual data, prediction results, and MAPE in Table 1.

Table 1. Prediction Result SVM

No	Month	Actual Data	Prediction Result	MAPE
1	January	14146	9629	31.93%
2	February	15072	9057	39.91%
3	March	16106	7694	52.23%
4	April	15497	6420	58.57%
5	May	16262	6227	61.71%
6	June	16527	7148	56.75%
7	July	17478	8134	53.46%
8	August	20843	8183	60.74%
9	September	15783	7666	51.43%
10	October	14161	8319	41.25%
11	November	16363	11538	29.49%
12	December	18237	16612	8.91%

Using the best parameter values, such as the value of c is 10, epsilon has a value of 0.01, using gamma with an auto value. The prediction results of foreign tourist arrivals in North Sumatra are obtained in the table above; there are differences in the model's prediction accuracy. The forecast shows a fairly good accuracy in certain months, but there is a significant difference between the actual data and the prediction results in other months. The best performance occurred in November, with a MAPE value of only 29.49%, indicating that the model predictions are pretty accurate. In addition, December's performance is also quite good, with the lowest MAPE of 8.91%, indicating that the model can capture changes in visits during that period well. However, the worst performance occurred in May and August, where the actual data differed significantly from the predicted data. In May, the model produced the highest MAPE of 61.71%, indicating that the model had difficulty in predicting tourist visits in that month. The same situation occurred in August, with MAPE reaching 60.74%; the table above results are visualized using a line graph in Figure 2.

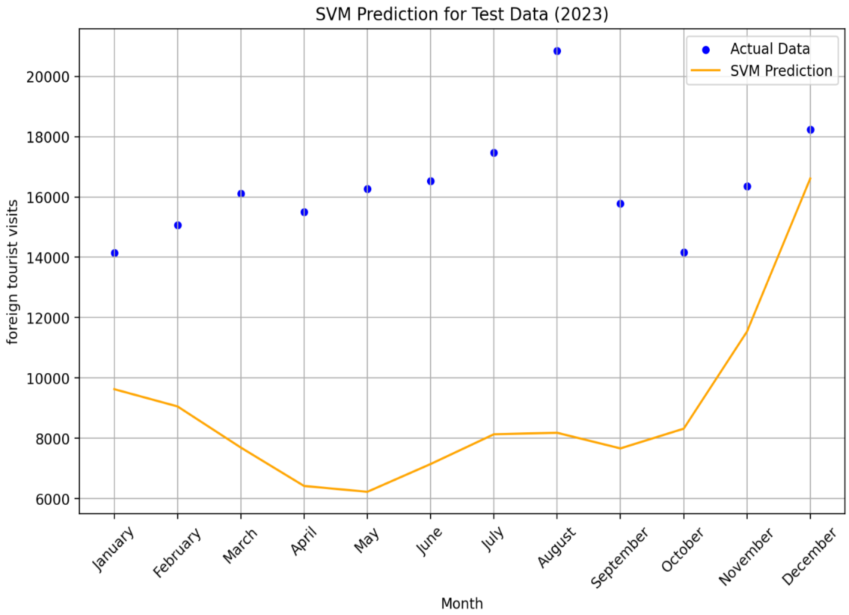

Fig 2. Result SVM

4.2. Model Linear Regression

These assumptions are still necessary for the model to function correctly, but there are fewer parameters to optimize than non-linear approaches. Here, the outcomes are predicted using the test data and linear regression, with the model trained to provide the best forecast. The outcomes of the linear regression model's predictions are as follows. The prediction results of foreign tourist arrivals in North Sumatra are displayed in a table containing months, actual data, prediction results, and MAPE in Table 2.

Table 2. Prediction Result Linear Regression

No	Month	Actual Data	Prediction Result	MAPE
1	January	14146	5719	59.57%
2	February	15072	6393	57.58%
3	March	16106	7067	56.12%
4	April	15497	7741	50.05%
5	May	16262	8415	48.25%
6	June	16527	9089	45.01%
7	July	17478	9763	44.14%
8	August	20843	10438	49.92%
9	September	15783	11112	29.60%
10	October	14161	11786	16.77%
11	November	16363	12460	23.85%
12	December	18237	13134	27.98%

Table 2 shows that the model predictions show variations in accuracy, with the highest MAPE value in January at 59.57% and the lowest MAPE in October at 16.77%, indicating relatively more accurate predictions for October. Although the model produces fairly good predictions in some months, such as October and September, the forecasts for other months still show significant errors. Based on the table above results, if visualized using a line graph in Figure 3.

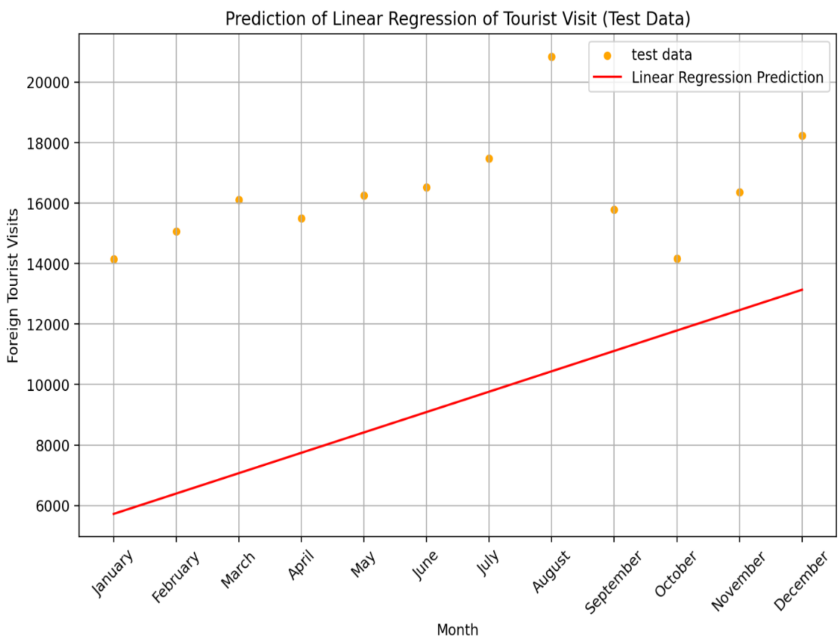

Fig 3. Result Linear Regression

4.3. Comparison of Prediction Results

It is crucial to conduct a thorough analysis using more objective evaluation metrics, like Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE), even though comparing the difference between actual data and predicted results indicates that linear regression is marginally more accurate than SVM. These two metrics will provide a clearer picture of the prediction error rate of each model and help ascertain whether linear regression remains superior when viewed from a relative and absolute error perspective. As the results shown in the previous section, table 3 below will be obtained.

Table 3. Comparison Methods

MAPE SVM	MAPE LR	RMSE SVM	RMSE LR	Superior Methods
46.65%	42.40%	8020.42	59.57%	Linear Regression

The dataset was subjected to SVM and Linear Regression (LR) techniques, with the MAPE and RMSE metrics serving as assessment benchmarks following the data table above. On the dataset, the linear regression (LR) method obtained a MAPE value of 42.40% and RMSE 6735.67, while SVM obtained a MAPE value of 46.65% and RMSE 8020.42. These results show that the linear regression method is more accurate and superior to the SVM method.

5. Conclusion

Based on the results of research using Linear Regression and Support Vector Machine (SVM) methods to predict the number of foreign tourist visits in North Sumatra, it is concluded that the Linear Regression method is superior with a MAPE of 42.40% and RMSE of 6735.67. In comparison, SVM produces a MAPE of 46.65% and RMSE of 8020.42, showing less than optimal performance. Overall, both methods are still inadequate in capturing significant fluctuations in visit data. For future research, it is recommended to update the data, especially post-COVID-19, add tourism-related features, try other methods such as machine learning or seasonal models, and test the model in different locations to expand its application. With the implementation of these suggestions, it is hoped that future research can improve prediction accuracy and make a more significant contribution to tourism management in Indonesia.

References

[1] G. Pedro and Gomez. Pesando, "Travel & Tourism Development Index 2021 Rebuilding for a Sustainable and Resilient Future M A Y 2 0 2 2," 2022.

[2] I. M. Hasibuan, S. Mutthaqin, R. Erianto, and I. Harahap, “Kontribusi Sektor Pariwisata Terhadap Perekonomian Nasional,” *Jurnal Masharif al-Syariah: Jurnal Ekonomi dan Perbankan Syariah*, 2023, doi: 10.30651/jms.v8i2.19280.

[3] S. U. BPS, *Provinsi Sumatera Utara Dalam Angka 2022*. 2023.

[4] H. Mukhtar, R. Muhammad, T. Reny Medikawati, and Yoze Rizki, “Peramalan Kedatangan Wisatawan Mancanegara Ke Indonesia Menurut Kebangsaan Perbulannya Menggunakan Metode Multilayer Perceptron,” *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 2, no. 2, pp. 113–119, Dec. 2021, doi: 10.37859/coscitech.v2i2.3324.

[5] N. Nurdin, Rizki Mela, and Maryana, “Analisa Data Mining Dalam Memprediksi Masyarakat Kurang Mampu Menggunakan Metode K-Nearest Neighbor,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 2, Apr. 2024, doi: 10.23960/jitet.v12i2.4131.

[6] M. Maulita and N. Nurdin, “Pendekatan Data Mining Untuk Analisa Curah Hujan Menggunakan Metode Regresi Linear Berganda (Studi Kasus: Kabupaten Aceh Utara),” 2023. doi: https://doi.org/10.36080/idealis.v6i2.3034.

[7] B. A. Manurung, A. Gea, A. P. Silalahi, and N. Samosir, “Penerapan Algoritma Regresi Linear Untuk Memprediksi Jumlah Wisatawan,” 2024. [Online]. Available: https://ejurnal.methodist.ac.id/index.php/methosisfo

[8] Fahirah and S. Wati, "Comparison between Linear Regression, Support Vector Machines, and Neural Networks in Forecasting the Number of Foreign Tourist Visits to Indonesia," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 11, no. 5, pp. 262–268, Apr. 2021, doi: 10.29322/ijsrp.11.05.2021.p11330.

[9] B. Karomah, “Aplikasi Metode Regresi Linier Sederhana Untuk Memprediksi Kedatangan Jumlah Wisatawan Mancanegara Pasca Pandemi COVID 19,” *Mathematics & Applications Journal*, 2023, doi: https://doi.org/10.15548/map.v4i2.4933.

[10] D. Rahmawati, T. Kristanto, B. F. Setya Pratama, and D. B. Abiansa, "Prediksi Pelaku Perjalanan Luar Negeri Di Masa Pandemi COVID-19 Menggunakan Metode Regresi Linier Sederhana," *Journal of Information System Research (JOSH)*, vol. 3, no. 3, pp. 338–343, Apr. 2022, doi: 10.47065/josh.v3i3.1507.

[11] V. N. Wijayaningrum and N. N. Putriwijaya, “Support Vector Regression Untuk Memprediksi Jumlah Kunjungan Wisatawan Mancanegara Di PULAU Bali,” *Prosiding SENTIA*, 2020, [Online]. Available: www.bps.go.id

[12] I. A. Tarigan, I. P. A. Bayupati, and G. A. A. Putri, "Comparison of support vector machine and backpropagation models in forecasting the number of foreign tourists in Bali province," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 2, pp. 90–95, Apr. 2021, doi: 10.14710/jtsiskom.2021.13847.

[13] W. R. U. Fadilah, D. Agfiannisa, and Y. Azhar, “Analisis Prediksi Harga Saham PT. Telekomunikasi Indonesia Menggunakan Metode Support Vector Machine,” *Fountain of Informatics Journal*, vol. 5, no. 2, p. 45, Sep. 2020, doi: 10.21111/fij.v5i2.4449.

[14] A. Karim, “Perbandingan Prediksi Kemiskinan di Indonesia Menggunakan Support Vector Machine (SVM) dengan Regresi Linear,” *Jurnal Sains Matematika dan Statistika*, vol. 6, no. 1, p. 107, Jan. 2020, doi: 10.24014/jsms.v6i1.9259.

[15] M. Sari, D. Agustini, M. Farida, U. Islam Kalimantan Muhammad Arsyad Al Banjari Banjarmasin, and K. Selatan, “Model Prediksi Kunjungan Wisata: Mengoptimalkan Arsitektur Algoritma Backpropagation untuk Prediksi Kunjungan Wisata Mancanegara (ASIA),” *Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)*, vol. 5, no. 1, pp. 240–254, 2024, doi: https://doi.org/10.30645/kesatria.v5i1.332.g329.

[16] V. Jessfry and M. Siddik, “Penerapan Data Mining Menggunakan Algoritma Apriori Dalam Membangun Sistem Persediaan Barang,” *Journal Of Information Systems And Informatics Engineering*, vol. 8, no. 1, pp. 187–199, 2024.

[17] A. Prasetyo, N. Nurdin, and H. A. K. Aidilof, "Comparison of Triple Exponential Smoothing and ARIMA in Predicting Cryptocurrency Prices," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 63–71, Oct. 2024, doi: 10.52088/ijesty.v4i4.577.

[18] N. Nurdin, F. Fajriana, M. Maryana, and A. Zanati, "Information System for Predicting Fisheries Outcomes Using Regression Algorithm Multiple Linear," *Journal of Informatics and Telecommunication Engineering*, vol. 5, no. 2, pp. 247–258, Jan. 2022, doi: 10.31289/jite.v5i2.6023.

[19] A. Anggara, N. Nurdin, and R. Meiyanti, "Sentiment Analysis of the MK Decision Trial of the Result of the 2024 President and Vice President General Election on Social Media X Using the Support Vector Machine Method," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 125–134, Oct. 2024, doi: 10.52088/ijesty.v4i4.591.

[20] V. A. Zahrah, N. Nurdin, and R. Risawandi, "Sentiment Analysis of Google Maps User Reviews on the Play Store Using Support Vector Machine and Latent Dirichlet Allocation Topic Modeling," *International Journal of Engineering, Science and Information Technology*, vol. 4, no. 4, pp. 87–100, Oct. 2024, doi: 10.52088/ijesty.v4i4.580.

[21] W. Andriani, Gunawan, and A. E. Prayoga, “Prediksi Nilai Emas Menggunakan Algoritma Regresi Linear,” *Jurnal Ilmiah Informatika Komputer*, vol. 28, no. 1, pp. 27–35, 2023, doi: 10.35760/ik.2023.v28i1.8096.