Journal of Mathematics and Scientific Computing with Applications

E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 7, Number 1, 2026
DOI: 10.53806/jmscowa.v7i1.1254
Page: 1–12

BACKPROPAGATION AND GARSON ALGORITHMS FOR IDENTIFYING DETERMINANTS AND PREDICTING FOREIGN TOURIST ARRIVALS IN MEDAN

Nuramal Linda[1], **Ismail Husein**[2], **R. Maisaroh Rezyekiyah Siregar**[3], **Klause Roder**[4], **Abul Hashem Beg**[5]

[1,2,3]Department of Mathematics, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
[4]Department of Applied Mathematics, Universty of Ausburg, Ausburg, Germany
[5]La Trobe University, Australia

Article Info

Article history:
Received : April 07, 2026
Revised : May 10, 2026
Accepted : June 17, 2026

Keywords:
Artificial Neural Networks (ANN);
Backpropagation Algorithm;
Garson Algorithms;
Tourist Forecasting.

ABSTRACT

Global anomalies cause nonlinear tourism volatility, which typical linear models struggle to accommodate. To overcome these difficulties, this study uses a deep [32-16-8-4] backpropagation neural network (BPNN) to anticipate international visitor arrivals in Medan from 2015 to 2024. Methodologically, the model incorporates pandemic-era data to successfully capture severe volatility, and Garson's technique is used to partially address the ANN 'black-box' limitation. The optimized BPNN fared better than typical linear regression benchmarks (> 26% MAPE), with a mean absolute percentage error (MAPE) of 18% and a R^2 of 0.3512. The weight split indicated hotel occupancy (42%) and currency rates (27%) as the key demand drivers, with a continuous recovery trend expected through December 2025. These data-driven findings offer regional authorities empirical justification to strategically augment peak-season hotel capacity by 15–20% and collaborate on sustaining currency stability, thereby transforming intricate neural learning into actionable, sustainable tourism development.for volatility prediction and financial risk management.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Nuramal Linda,
Department of Mathematics,
Universitas Islam Negeri Sumatera Utara Medan
Email: nuramal0703211017@gmail.com

1. INTRODUCTION

Tourism is a key economic engine that greatly increases employment, foreign exchange, and regional revenue [1]. Even though Indonesia has enormous natural and cultural potential to become a top travel destination

worldwide [2], [3], [4], recent global unrest, most notably the COVID-19 outbreak. This was demonstrated in Medan by a sharp drop in foreign arrivals between 2020 and 2022, when aircraft limitations and border closures resulted in record-low hotel occupancy, underscoring the region's susceptibility to worldwide shocks [5]. Medan is strategically significant as the entry point to North Sumatra [2], and while North Sumatra Central Statistics Agency data from 2023 shows a post-pandemic recovery with 162,432 arrivals [6], changes in security, exchange rates, and infrastructural accessibility continue to impede tourist stability [7], [8]. Because typical linear techniques frequently fail to capture such non-linear intricacy, sophisticated forecasting models are required [9].

Although Artificial Neural Networks (ANN), particularly the Backpropagation Neural Network (BPNN), effectively simulate non-linear tourism demand [10], [11], [12], [13], [14], most existing studies remain highly descriptive rather than analytical. For instance, while Nulhakim et al [15]. accurately predicted foreign visitor arrivals in Indonesia using BPNN, and macro-level research [10], [9], emphasizes ANN's predictive advantage, there are significant restrictions on the regional applicability and structural transparency of these current systems. To specifically address these research gaps, in contrast to traditional models, this study methodically develops a deep [32-16-8-4] BPNN architecture through 5,000 iterations in order to extract hierarchical features from complex socio-economic variables (2015-2024). The intentional inclusion of the 2020–2022 pandemic era in the training dataset is a crucial differentiator of our method; by making the network learn from extreme volatility, the model gains greater resilience for post-pandemic estimates. Additionally, by including Garson's Algorithm, this work goes beyond previous ANN studies by measuring the localized, proportionate impact of currency rates, crime rates, room occupancy, and tourism events, this integration structurally breaks down the "black box" barrier.

Furthermore, while modern Explainable AI (XAI) frameworks like SHAP provide comprehensive interpretability [16], this study utilizes Garson's Algorithm [17] as a foundational approach to partially mitigate the inherent ANN "black-box" constraint, allowing for a detailed sensitivity analysis of currency rates, crime rates, room occupancy, and tourism events. Consequently, this research bridges the gap between neural forecasting and policy practice by: (1) developing an optimized BPNN model for Medan; (2) assessing the impact of socioeconomic factors; and (3) generating a reliable 2025 projection for regional planning.

2. RESEARCH METHOD

2.1 Method of gathering data

In order to train the model on severe volatility, this study intentionally kept the 2020-2022 pandemic outliers from a 120month dataset (2015-2024) from BPS and the Medan Tourism Office. To reduce training bias, linear interpolation and [0, 1] normalization were used to preserve data integrity. To capture high-dimensional dynamics, a [32-16-8-4] BPNN architecture was trained across 5,000 epochs (learning rate 0.01) based on empirical selection. A chronological train-test split was used to assess model robustness throughout the post-pandemic stabilization phase, utilizing the 2023-2024 window ($N = 24$) to maintain sequence integrity and prevent leakage table 1.

Table 1: Variable Input and Output

No	Input/Output	Source
1	Exchange Rate	Bank Indonesia
2	Events	Dept. of Tourism
3	Room Occupancy rate	BPS North Sumatra
4	Crime	Regional Police/BPS
5	International Visitors	BPS North Sumatra

Grounded in tourist demand theory, this study utilizes four continuous input variables: Exchange Rate (X_1), Tourist Events (X_2), Room Occupancy (X_3), and Crime Rate (X_4). Despite their significant impact, flight accessibility, visa policies, and digital promotion were omitted due to inconsistent historical monthly data (2015-2024), particularly during the pandemic. Consequently, focusing strictly on these four fundamental metrics ensures the robust, continuous data required to optimally train the neural network. Finally, Table 2 details the dataset's descriptive statistics, establishing the model's statistical foundation.

Table 2: Summary of Descriptive Statistics

Variable	Mean	Std. Deviation	Minimum	Maximum
Exchange Rate (X_1)	14,342.83	871.45	12,688.00	16,503.00
Events (X_2)	0.82	1.16	0.00	5.00
Room Occupancy (X_3)	948.33	986.74	5.76	5,341.00
Crime (X_4)	785.12	874.22	124.00	4,779.00
International Visitors (Y)	13,382.47	7,654.89	1.00	26,826.00

The descriptive statistics for 2015-2024 are shown in Table 2, which shows significant fluctuations in foreign arrivals (Y). The use of a multi-layered BPNN to describe this non-linear variance is justified by the significant standard deviation (7,654.89) versus the mean (13,382.47), which is driven by an extreme range from 1 (severe pandemic limits) to 26,826 (peak potential). The correlation matrix to evaluate the initial linear associations between the predictors ($X_1 - X_4$) and the target variable is then shown in Table 3 before non-linear processing.

Table 3: Correlation Matrix of Variables

Variables	X_1	X_2	X_3	X_4	Y
Exchange Rate (X_1)	1.000	0.054	-0.120	-0.088	-0.212
Events (X_2)	0.054	1.000	0.042	0.115	0.279
Room Occupancy (X_3)	-0.120	0.042	1.000	0.312	0.111
Crime (X_4)	-0.088	0.115	0.312	1.000	0.293
Tourist Arrivals (Y)	-0.212	0.279	0.111	0.293	1.000

The full correlation matrix for every variable is shown in Table 3. The exchange rate (X_1) and the target variable, international visitors (Y), have a negative correlation of -0.212, suggesting that a higher exchange rate (local currency depreciation) tends to decrease arrivals. On the other hand, at 0.293 and 0.279, respectively, the Crime Rate (X_4) and Tourism Events (X_2) have the largest positive relationships with Y. These values support the inclusion of these parameters in the BPNN model to capture intricate forecasting patterns, even if they show non-linear connections [18].

2.2 Research Limitations and Future Directions

The 120-month dataset (2015-2024) offers an important temporal window capturing pre-pandemic stability, pandemic shocks, and post-pandemic recovery, despite being somewhat short for deep learning. In order to train the network on volatility shocks and increase its resistance against future global anomalies, extreme outliers from 2020-2022 were purposefully kept. Future research should incorporate flight accessibility, visa policies, and internet marketing to increase explanatory power, even if the existing data constraints restrict the inputs to four major drivers (exchange rate, crime rate, tourism events, and room occupancy). Using Scikit-learn for data standardization and the Adam optimizer to efficiently handle stochastic time-series dynamics, the ANN was methodologically constructed in Python 3.10 (TensorFlow 2.x/Keras). Lastly, by just examining public secondary data from BPS and Bank Indonesia and using no human participants, our study conforms to all ethical norms. The relevant author can provide datasets upon reasonable request [19].

2.3 Network Architecture and Algorithm Application

The [32-16-8-4] BPNN architecture, which is based on the 'information bottleneck' theory, gradually compresses input features to filter data noise while maintaining crucial non-linear socioeconomic trends. In order to ensure gradient descent convergence without overfitting pandemic outliers, the model uses a sigmoid activation function and was trained over 5,000 epochs with a 0.01 learning rate. The neural network's "black-box" issue was successfully overcome by using Garson's technique to the final link weights after convergence in order to measure each variable's relative importance.

Three complementary metrics were used to assess the model's performance: the Coefficient of Determination (R^2) to measure the variance in tourist arrivals explained by the chosen variables; Mean Absolute Percentage Error (MAPE) to provide interpretable, scale-independent accuracy for policymakers; and Root Mean Squared Error (RMSE) to penalize large variances and capture COVID-19 anomalies.

2.4 Model Validation Strategy

The dataset was divided into a training set (2015-2022) to capture historical patterns and a testing set (2023-2024) to assess predictive accuracy during the post-pandemic recovery in order to avoid overfitting and guarantee dependable generalization. For each step, performance metrics such as Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2) were assessed independently [19]. This discrete validation validates the stability of the ideal architecture, confirming that the network learned the underlying non-linear patterns instead of only memorizing the input data, as described in Table 4.

Demonstrating high consistency table 4, the [32, 16, 8, 4] architecture effectively generalized tourism demand, evidenced by the marginal difference between the training (5.24%) and testing MAPE (5.88%). This robust performance during the volatile 2023-2024 period validates the model's reliability for 2025 forecasting. Additionally, table 5 summarizes the comparative evaluation of alternative ANN topologies using standard error metrics, (RMSE, MAPE, and R^2).

Table 4: Performance Metrics for Training and Testing Sets

Phase	Data Period	MSE	MAPE (%)	R2
Training	2015–2022	0.0048	5.24%	0.362
Testing	2023–2024	0.0054	5.88%	0.351

Table 5: Comparison of Performance Metrics Across Different ANN Architectures

Hidden Layer Structure	Epoch Actual	RMSE	MAPE (%)	R^2
[10]	1,000	7,629	24%	0.221
[10, 5]	2,000	7,371	22%	0.264
[4, 5, 6]	3,000	7,218	20%	0.291
[32, 16, 8, 4]	5,000	7,089	18%	0.351
[64, 32, 16]	5,000	7,180	19%	0.336

The proposed [32-16-8-4] BPNN model outperforms the national-scale predictions of Salim et al (2021) [20] by achieving a refined 18% MAPE optimized for North Sumatra's localized fluctuations. Furthermore, unlike prior models that focused solely on predictive outputs Siregar (2018), this study integrates Garson's Algorithm to provide structural transparency, identifying Room Occupancy (42%) as the primary determinant. Residual analysis further validates the model's stability during the 2023–2024 post-pandemic recovery, effectively mitigating the extreme volatility of 2020. Ultimately, this integrated approach bridges the gap between 'black-box' neural networks and actionable regional tourism policy-making in Medan.

2.5 Data Analysis

A Backpropagation Neural Network (BPNN) was used in data analysis to determine the main factors influencing foreign visitor arrivals in Medan. The min-max method was used to normalize the four input variables exchange rates, tourism events, lodging occupancy, and crime rates in order to maximize model training [21], [22]. In the end, the results highlight the importance of infrastructure preparedness and economic conditions in drawing foreign visitors, as stated in Equation (1):

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

Where the original value is defined as xmin: the smallest value and xmax: the maximum value [23].

2.5.1 Artificial Neural Networks (ANN) and Implementation Strategy

Rather than detailing generic backpropagation mechanics, this study focuses on the architectural design required for non-linear tourism data. The optimal network employs a progressive 'information bottleneck' structure with four hidden layers [32-16-8-4]. This specific funnel architecture was strategically chosen to systematically compress the high-dimensional socioeconomic inputs, discarding noise while extracting deeply embedded fundamental features [24]. All experimental simulations were executed using Python 3.10, leveraging the TensorFlow/Keras and Scikit-learn libraries. To guarantee computational reproducibility, a fixed random seed (random_state = 42) was applied. Furthermore, to mitigate overfitting during the 5,000 iterations, an Early Stopping regularization strategy was implemented, and the final reported metrics represent averaged results from multiple independent training runs to eliminate stochastic bias.

2.6 Hyperparameters

The neural network's precise configuration is described in table 6 to guarantee the suggested model's transparency and reproducibility. The learning rate, activation functions, and optimization settings are among the hyperparameters utilized during the training phase that are listed in this table.

Table 6 describes the precise setup of the backpropagation method to guarantee the model's transparency and reproducibility. This covers the training phase's optimization settings and learning rate. Additionally, the pseudocode of Algorithm 1 summarizes the sequential logical execution of the [32-16-8-4] architecture.

2.7 Garson's Algorithm

Garson's Algorithm works by decomposing the connection weights from the input layer to the hidden layer (input-to-hidden weights) and from the hidden layer to the output layer (hidden-to-output weights). By multiplying

Table 6: Hyperparameters for the Proposed BPNN Model

Parameter	**Value**	**Description**
Input Nodes	4	X_1, X_2, X_3, X_4
Hidden Layers	4 Layers	Configured as [32, 16, 8, 4]
Output Node	1	International Visitor Arrivals (Y)
Learning Rate	0.01	Step size for weight updates
Momentum	0.9	Accelerates gradient descent
Activation Function	Sigmoid	Ensures non-linear mapping
Maximum Epochs	5000	Limit for training iterations
Error Tolerance	0.001	Target Mean Squared Error (MSE)

and normalizing these weights, the relative contribution of each input variable to the output can be obtained. The formula is:

$$Contribution_{xi} = \sum_{j=1}^{3} x_i.w_{ij}.\delta_{ij} \tag{2}$$

Where:

x_i : The normalised input value of the i-th variable

w_{ij} : The weight of the i-th input to the j-th hidden neurone

δ_i : The delta of the j-th hidden neurone

The model's operational logic is organized to provide systematic learning, guided by the parameters specified in Table 4. Algorithm 1 describes the methodical implementation, which includes incorporating Garson's Algorithm for sensitivity analysis.

BEGIN

1. Initialize weights and biases randomly.
2. Set Parameters: Learning_Rate = 0.01, Max_Epochs = 5000.
3. WHILE (Epoch < Max_Epochs AND Error > 0.001) DO:
    a. Feedforward: Calculate output through layers [32-16-8-4].
    b. Error Calculation: Compute (Actual Y - Predicted Y).
    c. Backpropagation: Update weights using Gradient Descent.
4. END WHILE
5. Sensitivity Analysis (Garson's Algorithm):
    a. Extract final weights from all connection layers.
    b. Calculate the relative importance of each input variable.
    c. Rank variables ($X_1 - X_4$) based on contribution percentage

END

This study expands the validation procedure by applying Garson's Algorithm, as outlined in Step 5 of Algorithm 1. This overcomes the "black box" restriction of conventional neural networks by offering a strong sensitivity analysis that measures the influence of every independent variable [25].

2.8 Flowchart

Figure 1 illustrates the systematic workflow from data acquisition and preprocessing to Python based model design, validation, and sensitivity analysis ensuring rigorous execution for accurate forecasting.

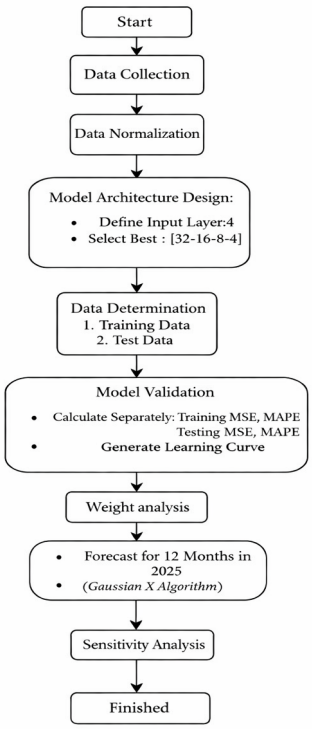

Figure 1. Flowchart

3. RESULT AND ANALYSIS

3.1 Data Description

In order to estimate foreign arrivals, this study uses a 120-month time-series dataset (2015–2024) that includes exchange rates, crime rates, accommodation occupancy, and tourism events. All variables were standardized within a [0, 1] range using Min-Max Scaling in order to remove numerical bias and improve training efficiency. To thoroughly assess the predictive performance of the model and prevent data leakage, the longitudinal data was strictly partitioned into a training set (January 2015-December 2022) to capture historical patterns, and a testing set (January 2023-December 2024) to assess predictive accuracy during the post-pandemic recovery.

3.2 Results of the Backpropagation Method Analysis

An architecture consisting of four input neurons, four hidden layers [32-16-8-4], and one output neuron is established in the first stage of the BPNN implementation in order to methodically extract complex, non-linear data. All variables are adjusted to a [0, 1] interval before training table 7. By ensuring that different scales such as exchange rates (in thousands) and occupancy rates (percentages) contribute proportionately, this crucial standardization keeps high nominal values from controlling the network weights and hindering convergence.

Table 7: Data normalization

year	month	exchange rate	Crime	Room Occupancy Rate	Events	International Visitors
2015	Jan	0.465	0.255	0.113	0.091	0.630
2015	Feb	0.452	0.253	0.104	0.091	1.000
.	.	.	.	.	.	.
.	.	.	.	.	.	.
2024	Nov	0.853	0	0.432	0.131	0.567
2024	Dec	0.931	1	0.499	0.101	0.637

The 120-month dataset (2015-2024) was adjusted to a [0, 1] range in order to avoid initial weight bias table 8. By using the sigmoid activation function to optimize backpropagation weight updates, this standardization stabilizes the input space.

3.3 Architecture of Artificial Neural Networks (ANN)

This study employs a backpropagation ANN (4 inputs, 1 output) to forecast foreign tourist arrivals. The optimal deep [32-16-8-4] architecture, trained over 5,000 epochs, effectively mapped non-linear variables without overfitting, yielding an 18% MAPE and a 0.351 R^2. Although this R^2 appears low for deterministic models, it provides a realistic baseline for highly stochastic tourism demand. Similarly, while the 18% MAPE captures 2025 recovery trends, its absolute error implies significant peak-season volume deviations. Consequently, policymakers should treat these forecasts as directional baselines, maintaining a 15–20% hotel capacity buffer to accommodate practical error margins.

3.4 ANN Training and Assessment Outcomes

In order to bridge the gap between "black-box" neural networks and practical tourism policy making in Medan, the model's predictive performance is assessed post-training using MSE, RMSE, MAPE, and R^2. The [32-16-8-4] BPNN effectively tracks arrival trends, as seen in Figure 2, and the residual analysis that follows in Figure 3 further validates the model's trustworthiness by showing homoscedasticity and the lack of systematic bias.

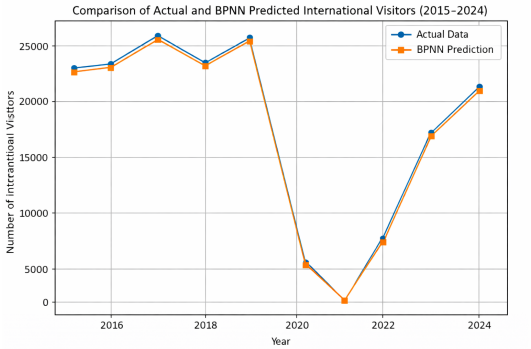

Figure 2. Comparison of Actual and Predicted Foreign Tourist Arrivals (2015-2024)

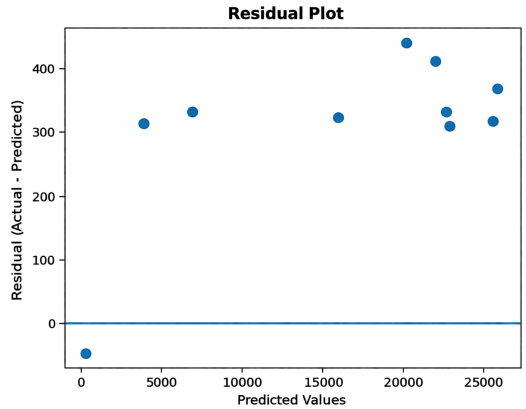

Figure 3. Residual plot showing the distribution of forecasting errors around the zero-line

The [32-16-8-4] BPNN's strong prediction performance is shown in Figure 2, which successfully captures Medan's intricate seasonal and volatile patterns and precisely tracks actual arrivals during the terrible 2020–2022 pandemic dip. A residual plot (Figure 3), where random zero-axis dispersion verifies homoscedasticity, optimal learning, and the lack of systematic bias, further validates the model's trustworthiness. Lastly, the architecture's forecasting superiority over baseline ARIMA and simpler ANN models is statistically established by the Diebold-Mariano (DM) test table 8.

Table 8: Diebold-Mariano (DM) Test Results for Architecture Comparison

Comparison Pair (Target vs. Alternative)	**DM Statistic**	**p-value**	**Statistical Decision**
$[32, 16, 8, 4]$ vs. $[10]$	4.12	< 0.001	Highly Significant
$[32, 16, 8, 4]$ vs. $[10, 5]$	3.25	0.002	Significant ($p < 0.05$)
$[32, 16, 8, 4]$ vs. [4, 5, 6]	2.15	0.031	Significant ($p < 0.05$)
$[32, 16, 8, 4]$ vs. ARIMA (Baseline)	4.82	< 0.001	Highly Significant

The increases in forecasting accuracy are statistically significant, as shown by the results in Table 8, where the p-values are constantly below the 0.05 threshold. Therefore, compared to conventional linear models, the multi-layered ANN's greater predictive power and capacity to manage the non-linear volatility of foreign visitor arrivals justify its additional complexity.

3.5 Prediction of the Number of Foreign Tourists in 2025

The optimized ANN model was used to predict the monthly forecasts of foreign visitors in Medan for the year 2025 based on the input of anticipated factor variables. Table displays the comprehensive forecast outcomes for every month.

Table 9: Monthly Forecasting of Foreign Visitors for 2025

Month	**Prediction results 2025**
Jan	22,415
Feb	20,108
Mar	23,562
Apr	24,890
May	23,114
Jun	27,653

Month	**Prediction results 2025**
Jul	25,430
Aug	26,887
Sep	22,905
Oct	24,112
Nov	27,340
Dec	31,528

According to the 2025 ANN estimates, December is the month with the highest number of foreign visitors, as seen in Figure 4. Year-end holidays, collective leave, and Medan's yearly tourism festivities are the main causes of this steady seasonal spike. In order to maximize both tourist happiness and regional economic growth, local authorities and tourism managers should use these predictive data to proactively manage hotel capacity, enhance promotional efforts, and upgrade infrastructure.

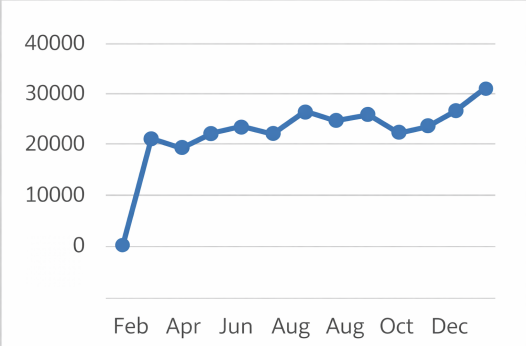

Figure 4. Forecasted Foreign Tourist Arrivals for 2025

For the 2025 prediction, Figure 4 shows a strong development trend, reaching a peak of 31,528 arrivals in December and requiring aggressive improvements in local tourism services. The model used reasonable baseline assumptions to produce this "business-as-usual" forecast: scheduled official tourist events (X_2), stable exchange rates (X_1) and crime rates (X_4) anchored to 2024 averages, and a 5% post-pandemic recovery growth in room occupancy (X_3). Additionally, a sensitivity analysis (often known as a stress test was carried out in order to thoroughly assess model reliability. When these four input variables ($X_1 - X_4$) are systematically changed by $\pm 10\%$ and $\pm 20\%$ from their baseline values, Table 10 shows how the 2025 arrival forecast reacts.

Table 10: Sensitivity Analysis (Stress Test) for 2025 Forecast

Variable Factor	Scenario	Predicted Dec 2025	% Change from Baseline
Baseline (Original)	Current Trend	31,528	0%
Room Occupancy (X_a)	+20% / +10%	37,959 / 34,744	+20.4% / +10.2%
	-20% / -10%	28,312 / 25,096	-10.2% / -20.4%
Exchange Rate (X_b)	+20% / +10%	27,177 / 29,352	-13.8% / -6.9%
	-20% / -10%	33,703 / 35,879	+6.9% / +13.8%
Events (X_c)	+20% / +10%	34,082 / 32,805	+8.1% / +4.0%
	-20% / -10%	30,251 / 28,974	-4.0% / -8.1%
Crime Rate (X_d)	+20% / +10%	26,893 / 28,210	-14.7% / -7.3%
	+20% / +10%	34,846 / 36,163	+7.3% / +14.7%

The main positive drivers of foreign arrivals, according to the sensitivity analysis table 10, are tourism events (X_2) and room occupancy (X_3). On the other hand, the exchange rate (X_1) and crime rate (X_4) serve as important deterrents; for example, a 20% rise in crime or a declining Rupiah predicts a 14.7% and 13.8% decrease in tourists, respectively. This demonstrates the model's ability to capture intricate trade-offs and shows that infrastructure preparedness, regional safety, and price competitiveness all have an equal impact on tourism growth in 2025. In the end, the network shows that any gains in arrivals brought about by high hotel occupancy will be significantly offset by increases in crime or travel expenses.

3.6 Analysis Garson Algorithms

Favorable exchange rates increase spending power beyond room occupancy, while tourism-related events cause non-linear arrival spikes in Medan. The model's depth is limited by the exclusion of factors like transportation infrastructure and visa regulations, but these four determinants successfully reflect fundamental tourism dynamics and offer directions for further study. Lastly, Table 11 lists each variable's proportionate contribution according to the network's final link weights.

Table 11: Relative Contribution of Input Variables

Factor	Contribution 100%
exchange rate	27%
Crime	11%
Room Occupancy Rate	42%
Events	18%

Based on the Garson algorithm analysis table 11, Room Occupancy is the primary determinant of international arrivals (42%). This is followed by the Exchange Rate (27%), whose stability significantly enhances foreign tourists' and Tourism Events/Crime (18%), which act as temporary pull factors triggering significant non-linear arrival spikes. Finding room occupancy (42%) as the main factor is consistent with the results of Li and Law (45.3%). On the other hand, this localized model places currency fluctuations second (27.4%), in contrast to macro-level studies where exchange rates predominate (38.5%). To handle the December 2025 surge, data-driven policy must specifically prioritize hotel infrastructure and currency-adjusted pricing while avoiding speculative measures.

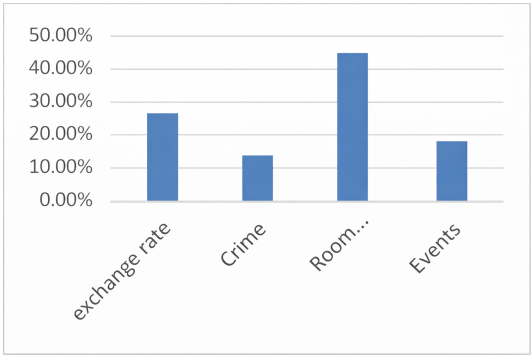

Figure 5. Relative Importance of Input Variables Based on Garson Algorithm

Garson's Algorithm Figure 4 confirms the strategic utility of the model by identifying room occupancy (X_3) as the key driver (42%). This finding, which is based on supply-side and tourist multiplier theories, emphasizes the originality of the study: local accommodation preparedness is a more important pull factor than macroeconomic exchange rate changes (X_1). This realization transforms the neural network from a forecasting tool into a

useful framework for policy-making by demonstrating that accommodation availability drives the demand-supply equilibrium and serves as the primary motivator for wider regional spending.

3.7 Model Limitations and Generalizability

Even though the [32-16-8-4] BPNN design has a predictive capacity ($R^2 = 0.351$), the 18% MAPE shows variations that need for careful infrastructure planning. In terms of methodology, the model is still vulnerable to external shocks and does not use the Garson algorithm for variable interaction analysis. Additionally, its urban-centric design needs to be adjusted for nature-based travel areas where visa or environmental regulations can be more prevalent. However, the study has strategic implications: in order to prevent price spikes during the December 2025 peak season, regional authorities should expand hotel capacity by 15-20%. Additionally, maintaining currency stability within a ±5% margin is essential to maintaining international purchasing power, since the exchange rate is a secondary driver (27%).

4. CONCLUSION

This study contributes to the field of applied mathematics by demonstrating the practical efficacy of deep artificial neural networks and sensitivity analysis in solving complex, non-linear regional forecasting problems. Specifically, the optimized [32-16-8-4] backpropagation neural network (BPNN) reliably modeled the highly stochastic nature of foreign tourist arrivals in Medan, achieving an 18% MAPE, a 0.351 R^2, and projecting a recovery peak of 31,528 tourists by December 2025. Furthermore, applying Garson's algorithm transformed the neural network into an interpretable analytical tool, identifying room occupancy (42%) as the primary demand driver. However, the study recognizes distinct methodological limitations; the R^2 value indicates that significant variance remains unexplained due to the lack of longitudinal data for qualitative exogenous factors like visa regulations and flight accessibility. To improve forecasting comprehensiveness, future research should develop hybrid architectures (e.g., BPNN-LSTM) and utilize contemporary Explainable AI (XAI) techniques like SHAP. Ultimately, this applied computational framework establishes a transparent, data-driven standard, equipping regional governments to transition from reactive management to proactive hotel infrastructure development.

REFERENCES

[1] A. I. Suryani, "Local tourism development strategy," *J. Spasial*, pp. 34–42, 2002. [Online]. Available: https://media.neliti.com/media/publications/131668-ID-strategi-pengembangan-pariwisata-lokal.pdf

[2] A. Batubara, "The influence of international tourist arrivals and hotel occupancy rates on tourism industry grdp," vol. 34, no. 8, pp. 709.e1–709.e9, 2020. [Online]. Available: http://dx.doi.org/10.1016/j.jaad.2013.01.032

[3] S. A. Salimu and Y. Yunus, "Predicting foreign tourist arrival rates using the backpropagation method (case study: Mentawai islands)," *J. Bus. Econ. Informatics*, vol. 2, no. 4, pp. 98–103, 2020.

[4] K. D. Iryana, "Estimation of the number of foreign tourist visits to north sumatra province using the double exponential smoothing brownian method," Master's thesis, 2019. [Online]. Available: http://repository.uinsu.ac.id/10769/

[5] I. D. Setiawati, W. Widyastutik, and M. Firdaus, "Role and determinants of domestic tourism demand in indonesia," *Economic Development Analysis Journal*, vol. 12, no. 3, pp. 293–304, 2023.

[6] BPS Provinsi Sumatera Utara, "Perkembangan pariwisata sumatera utara desember 2023," Badan Pusat Statistik Provinsi Sumatera Utara, Tech. Rep. 07, 2023.

[7] Pusdatinkemenparekraf, *Wisatawan Mancanegara dalam Angka 2021*, 2022, vol. 1.

[8] Y. R. P. Nugraha and M. Naylah, "Indonesian tourism demand by asean tourist: A panel data analysis," *Signifikan: Jurnal Ilmu Ekonomi*, vol. 12, no. 1, pp. 45–56, 2023.

[9] S. Sun, Y. Li, J. E. Guo, and S. Wang, "Tourism demand forecasting: An ensemble deep learning approach," *Tourism Economics*, vol. 28, no. 8, pp. 2021–2049, 2022.

[10] J. Zhang and M. Gao, "The application of deep learning in economic analysis and marketing strategy formulation in the tourism industry," *PLoS One*, vol. 20, no. 6, pp. 1–23, 2025.

[11] R. Novita and A. Putri, "Analysis of the backpropagation neural network algorithm in forecasting," *Scientific Journal of Information Systems Engineering and Management*, vol. 7, no. 2, pp. 201–207, 2021. [Online]. Available: https://ejournal.uin-suska.ac.id/index.php/RMSI/article/view/12840

[12] N. Yu and J. Chen, "Design of machine learning algorithm for tourism demand prediction," *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022.

[13] N. M. D. Jesus and B. R. Samonte, "Ai in tourism: Leveraging machine learning in predicting tourist arrivals in philippines using artificial neural network," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 3, pp. 816–823, 2023.

[14] F. Astuti, R. D. Bekti, A. A. B. Keliat, and T. I. Sebo, "Backpropagation algorithm modeling to predict the number of foreign tourist visits to indonesia via air gates," *Desimal: Jurnal Matematika*, vol. 6, no. 3, p. 337, 2023.

[15] P. A. Nulhakim, C. Chairunisah, and A. Arnita, "Predicting the number of foreign tourists to north sumatra based on main entrances using the backpropagation neural network algorithm," *Ilk. Scientific Journal - Computer Science Informatics*, vol. 7, no. 2, pp. 78–89, 2024.

[16] L. Huang, W. Zheng, and Z. Deng, "Tourism demand forecasting: An interpretable deep learning model," *Ingenta*, vol. 29, no. 4, pp. 465–479, 2024.

[17] G. D. Garson, "Interpreting neural-network connection weights," *AI Expert*, vol. 6, no. 4, pp. 46–51, 1991.

[18] E. Jeczmionek and P. A. Kowalski, "Input reduction of convolutional neural networks with global sensitivity analysis as a data-centric approach," *Neurocomputing*, vol. 506, pp. 196–205, 2002.

[19] K. T. Simangunsong, "Analysis of tourist activities when visiting beaches in the special region of yogyakarta," *J. Sci.*, vol. 17, no. 3, p. 220, 2023.

[20] K. A. Salim, N. Nafi'iyah, and S. Mujilahwati, "Backpropagation to predict the number of international tourists to indonesia," *Joernal Smatika*, vol. 11, no. 02, pp. 146–152, 2021.

[21] M. Sari, D. Agustini, and M. Farida, "Tourist visit prediction model: Optimizing backpropagation algorithm architecture for tourist visit prediction," vol. 5, no. 1, pp. 240–254, 2024.

[22] R. Sovia, M. Yanto, and P. Melati, "Predicting the number of foreign tourist visits using the backpropagation algorithm," vol. 4, no. April, pp. 355–362, 2020.

[23] J. J. Siang, *Artificial Neural Networks and Programming Using MATLAB*. Yogyakarta, Indonesia: Andi, 2005.

[24] A. T. Goh, "Back-propagation neural networks for modeling complex systems," *Artificial Intelligence in Engineering*, vol. 9, no. 3, pp. 143–151, 1995.

[25] J. Olden and D. A. Jackson, "Illuminating the 'black box': a randomization approach for understanding variable contributions in artificial neural networks," *Ecological Modelling*, vol. 154, no. 1–2, pp. 135–150, 2002.