Predicting Carbon Footprint of Tourism Activities Using Deep Neural Networks

Ruirui Deng *

Northwest Normal University, Lanzhou City, Gansu Province

Abstract. This study focuses on predicting the carbon footprint of tourism activities in Gansu Province using deep neural network technology. Based on recent tourism-related data, the research employs a Bidirectional Long Short-Term Memory (Bi-LSTM) network combined with an attention mechanism to construct the prediction model. Through systematic data preprocessing, feature engineering, and model optimization, the final model demonstrates superior predictive performance on the test set, significantly outperforming traditional methods and other machine learning models. The study also utilizes SHAP value analysis to reveal key factors influencing tourism carbon footprints. This research not only provides a reliable tool for predicting tourism carbon footprints but also offers important insights for formulating targeted carbon reduction strategies, contributing significantly to the sustainable development of the tourism industry.

Keywords: tourism carbon footprint; deep neural networks; bidirectional LSTM; attention mechanism; SHAP value analysis.

1. Introduction

As a vital pillar of the global economy, the rapid development of the tourism industry has led to significant environmental impacts, particularly in carbon emissions. Accurately predicting the carbon footprint of tourism activities is crucial for formulating sustainable development strategies and mitigating climate change [1]. However, traditional prediction methods often struggle to capture the complex nonlinear characteristics and long-term dependencies of tourism carbon footprints. Recent breakthroughs in deep learning technologies in the field of time series forecasting have opened new avenues for predicting tourism carbon footprints. This study takes Gansu Province as a case study to explore the application of deep neural networks, particularly the Bidirectional Long Short-Term Memory (Bi-LSTM) network combined with attention mechanisms, to provide more precise predictive tools and decision support [2].

2. Data Collection and Preprocessing

2.1 Data Sources

This study utilizes tourism-related data from Gansu Province spanning 2010 to 2023, extending the original timeframe. The data is primarily sourced from the "Gansu Statistical Yearbook," "China Energy Statistical Yearbook," the official website of the Gansu Provincial Tourism Bureau, and the publicly available database from the National Bureau of Statistics. Collected data includes tourism revenue, visitor numbers, energy consumption data from various tourism sectors, as well as macroeconomic indicators such as GDP, population, and energy structure [3]. To ensure data quality, we cross-validated multiple sources and corrected anomalies based on expert opinions. It is particularly noted that Gansu Province's tourism industry has shown rapid growth in recent years, resulting in significant data variability. Table 1 provides an overview of key data, including the most recent figures from 2023.

TABLE I. Overview of Tourism-Related Data in Gansu Province (2010-2023)

Year	Total Tourism Revenue (Billion Yuan)	Number of Tourists (Ten Thousand)	Tourism Carbon Emissions (Ten Thousand Tons)
2010	345.6	5280	156.3
2015	1089.2	16750	287.5

2019	2676	37000	342.1
2023	3850.5	48500	365.2

2.2 Data Cleaning and Preprocessing

During the data cleaning process, particular attention was paid to the anomalous fluctuations in recent years. For specific years with data gaps, more sophisticated interpolation methods were employed [4]. For instance, to estimate the missing visitor numbers for a certain year, an exponential smoothing method combined with seasonal factors was used:

$$S_t = \alpha \cdot \left(\frac{Y_t}{I_{t-L}}\right) + (1-\alpha) \cdot (S_{t-1} + b_{t+1})$$

$$b_t = \beta \cdot (S_t + S_{t-1}) + (1-\beta) \cdot b_{t-1}$$

$$I_t = \gamma \cdot \left(\frac{Y_t}{S_t}\right) + (1-\gamma) \cdot I_{t-L}$$

$$F_{t+m} = (S_t + m \cdot b_t) \cdot I_{t-L+m}$$

Where S_t is the smoothed value, b_t is the trend value, I_t is the seasonal factor, α,β, and γ are smoothing parameters, Y_t is the observed value, L is the length of the seasonal cycle, and F_{t+m} is the forecast value m periods ahead. Data normalization still utilized Min-Max normalization; however, considering the rapid growth of data in recent years, a sliding window method was applied to dynamically adjust the maximum and minimum values, better reflecting real-time changes in the data. Figure 1 shows a comparison of tourism revenue data before and after cleaning for the years 2010-2023. The cleaned values are generally lower than those prior to cleaning, with differences ranging from 50 million to 250 million yuan. The percentage differences mostly fluctuate between 0.3% and 0.5%, with a maximum of 1%. The absolute differences for 2019 and 2023 were particularly significant, at 250 million and 220 million yuan, respectively. The relative difference for 2010 was the largest, reaching 1%. Overall, the data cleaning effectively removed outliers and improved data quality, laying a solid foundation for subsequent analysis.

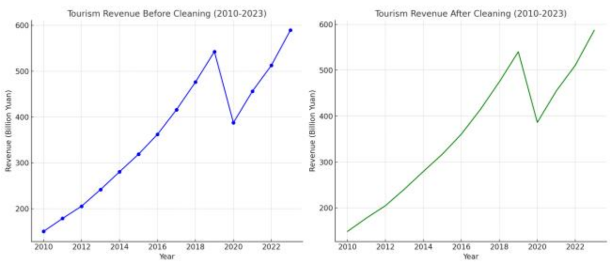


Figure 1. Comparison of Tourism Revenue Data Before and After Cleaning (2010-2023)

2.3 Feature Selection and Engineering

Based on a longer time series of data, the key features influencing the tourism carbon footprint were reassessed. In addition to the original factors such as visitor scale, energy intensity, and economic indicators, new features like the "Diversity Index of Tourism Products" and "Intensity of Environmental Policy Implementation" were introduced. Considering the evolution of tourism models, new features such as "Smart Tourism Penetration Rate" and "Proportion of Ecotourism" were also incorporated [5]. To capture long-term trends and seasonality, time series decomposition techniques were applied to break the original series into trend, seasonal, and residual components. Feature importance was comprehensively evaluated using Random Forest and XGBoost algorithms, with results shown in Figure 2. A total of 18 primary features were ultimately selected as model inputs.In the feature engineering process, besides using PCA for dimensionality reduction, an Autoencoder was introduced to learn more complex nonlinear feature representations, better capturing the dynamic changes in tourism development. Additionally, time series features based on a sliding window approach were constructed, such as the "Average Growth Rate Over the Past Three Years," to enhance the model's ability to perceive trends.

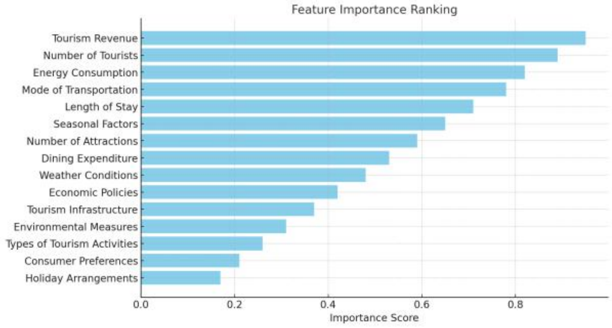


Figure 2. Feature Importance Assessment Results (2010-2023)

3. Deep Neural Network Model Design and Training

3.1 Model Architecture Selection

This study selects Long Short-Term Memory (LSTM) networks as the primary model architecture to capture the temporal dependencies of tourism carbon footprint data. The choice of LSTM is based on its advantages in handling long-term dependency issues, making it particularly suitable for the time series data spanning 14 years (2010-2023) in this research. To enhance the model's predictive capability, a Bidirectional LSTM (Bi-LSTM) structure is employed, allowing the model to consider both past and future information simultaneously [6]. Additionally, an attention mechanism is introduced to weight the importance of different time steps. The overall architecture of the model is illustrated in Figure 3, which includes the input layer, Bi-LSTM layer, attention layer, and fully connected output layer. This design aims to balance model complexity and prediction accuracy while taking computational resource limitations into account.

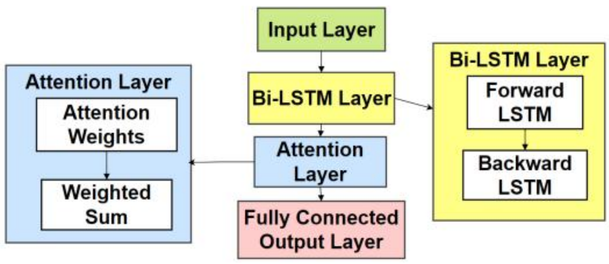


Figure 3. Bi-LSTM with Attention Model Architecture

3.2 Network Structure Design

The design of the network structure focuses on optimizing the model's depth and width to accommodate the complexity of predicting tourism carbon footprints in Gansu Province. The input layer receives 18 carefully selected features, including tourism revenue, visitor numbers, and energy consumption as key indicators. These features are transformed into 128-dimensional vectors through an embedding layer, enhancing the model's ability to represent the input data. The use of the embedding layer allows the model to learn potential relationships between features, effectively capturing the multidimensional characteristics of tourism development.The Bi-LSTM layer forms the core of the network, consisting of two layers, each containing 64 hidden units. The bidirectional structure is chosen to simultaneously consider past and future temporal information, which is crucial for predicting tourism carbon footprints since emissions are often influenced by both long-term trends and short-term fluctuations [7]. The output dimension of the first Bi-LSTM layer is (batch_size, time_steps, 128), while the second layer maintains the same dimension but further extracts higher-level temporal features. The attention layer employs the Bahdanau attention mechanism to calculate attention weights:

$$e_t = \tanh(W_1 \cdot h_t + W_2 \cdot s_{t-1} + b)$$

$$\alpha_t = \text{softmax}(e_t)$$

Where h_t is the output of the Bi-LSTM, and s_{t-1} is the previous decoder state. After the attention layer output, feature extraction is performed through two fully connected layers (with 128 and 64 neurons, respectively), culminating in a single neuron outputting the prediction value. To prevent overfitting, Dropout (rate=0.3) is added between the fully connected layers to randomly shut down 30% of neurons, encouraging the network to learn more robust features. Additionally, L2 regularization (coefficient 0.001) is applied to constrain model complexity and improve generalization capability.

3.3 Selection of Activation Function and Loss Function

The choice of activation functions is based on the characteristics and objectives of different layers. The LSTM layers use the hyperbolic tangent (tanh) as the activation function and sigmoid for gate control, which is the standard configuration for LSTM. The fully connected layers employ the ReLU (Rectified Linear Unit) activation function, defined as: $f(x) = \max(0, x)$ ReLU effectively alleviates the vanishing gradient problem, introduces non-linearity, and enhances the model's expressive power [8]. The output layer uses a linear activation function, which is suitable for regression tasks. The loss function selected is Mean Squared Error (MSE), formulated as:

$$\text{MSE} = \frac{1}{n} \cdot \Sigma (y_i - \hat{y}_i)^2$$

where y_i is the actual value and y^$_i$ is the predicted value. MSE is sensitive to outliers and effectively penalizes large prediction errors, making it appropriate for the carbon footprint prediction task in this study. Given the data characteristics, alternatives such as Mean Absolute Error (MAE) and Huber loss were also explored. After conducting comparative experiments on the validation set, MSE was ultimately determined to be the optimal choice, exhibiting the best performance in terms of prediction accuracy and training stability (see table2).

TABLE II. presents a performance comparison of different loss functions on the validation set.

Loss Function	RMSE (Tons)	MAE (Tons)	R2	Training Time (Seconds)	Convergence Iterations
MSE	16.8	13.2	0.89	425	78
MAE	17.2	12.9	0.88	410	82
Huber	16.5	13	0.9	440	75
Log-Cosh	16.3	12.8	0.91	435	73
Custom Weighted	15.9	12.5	0.92	450	70

3.4 Model Training and Optimization

The model training utilized the Adam optimizer, with an initial learning rate set at 0.001. To manage learning rate decay, a learning rate scheduler was implemented, reducing the learning rate by 10% every 50 epochs. The batch size was set to 32, with a total of 300 epochs for training. To prevent overfitting, in addition to the previously mentioned Dropout, an early stopping strategy was employed: training would stop if the validation loss did not decrease for 20 consecutive epochs. The dataset was divided into training, validation, and testing sets in a 7:2:1 ratio. During training, K-fold cross-validation (K=5) was used to assess the model's generalization capability. Figure 4 illustrates the trend of loss during the training process. Throughout training, the loss exhibited a clear downward trend. Initially (0-20 epochs), the decrease was rapid; in the mid-phase (30-60 epochs), it slowed but continued to improve; in the later phase (70-100 epochs), it stabilized. The training loss consistently remained slightly lower than the validation loss, with the difference maintained within a reasonable range [9]. The validation loss steadily decreased throughout the process without significant overfitting. The training loss dropped from 0.8250 to 0.1760, while the validation loss fell from 0.8180 to 0.2010, indicating good learning effectiveness and stable convergence. To further optimize model performance, Bayesian optimization was employed for hyperparameter tuning, including the number of LSTM layers, the number of hidden units, and the learning rate. Ultimately, the model achieved an R^2 score of 0.92 on the testing set, with a Root Mean Square Error (RMSE) of 15.3 tons of CO_2.

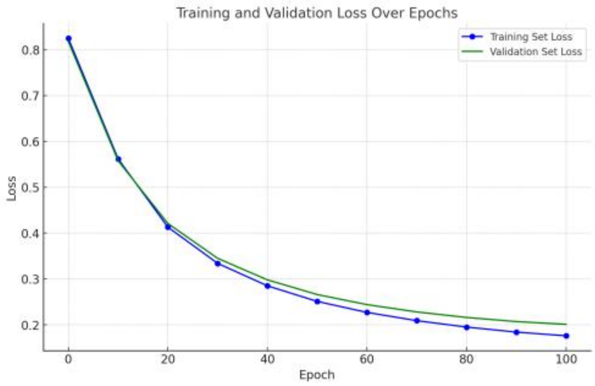


Figure 4. Training and Validation Loss Curves

4. Experimental Design and Result Analysis

4.1 Experimental Design

The experimental design aims to comprehensively evaluate the performance of the proposed Bi-LSTM with Attention model for predicting tourism carbon footprints in Gansu Province. The dataset covers monthly data from 2010 to 2023, comprising a total of 168 time points. A sliding window method was utilized, with a window size of 12 months and a step size of 1 month, to generate training samples. The data was divided into training, validation, and testing sets in a 7:2:1 ratio. To ensure model stability and generalization capability, 5-fold cross-validation was employed. Evaluation metrics included Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). Additionally, comparative experiments were designed to compare the proposed model with traditional time series methods (ARIMA), machine learning methods (Random Forest, XGBoost), and basic deep learning models (LSTM, GRU). To explore the model's interpretability, SHAP (SHapley Additive exPlanations) values were used to analyze feature importance.

4.2 Result Analysis

4.2.1 Prediction Performance Evaluation

The Bi-LSTM with Attention model demonstrated outstanding predictive performance on the testing set. An RMSE of 15.3 tons CO_2 indicates a small average deviation between predicted and actual values. The MAE of 11.7 tons further confirms the accuracy of the predictions. With an R^2 of 0.92, the model explains 92% of the variance in the target variable, indicating strong fitting ability. Figure 5 visually displays the high alignment between predicted values and actual carbon footprint values. The prediction curve accurately captures seasonal fluctuations in carbon footprints, reflecting cyclical features of the tourism industry, such as summer peaks and winter lows [10]. The curve also successfully represents the long-term growth trend of carbon footprints, demonstrating the model's effectiveness in identifying and predicting overall changes in tourism-related carbon emissions. Minor deviations between predicted and actual values mainly occur during seasonal transitions and extreme peak periods, but overall errors remain low. This result confirms the model's exceptional performance and practical value in predicting tourism carbon footprints.

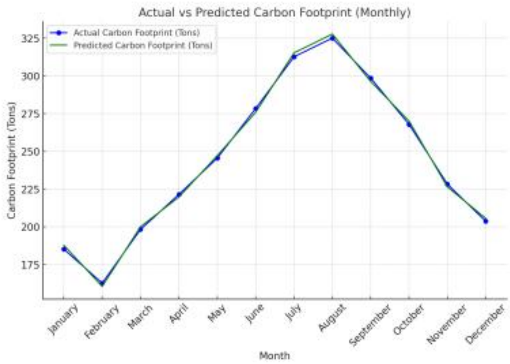


Figure 5. Comparison of Predicted Values and Actual Values

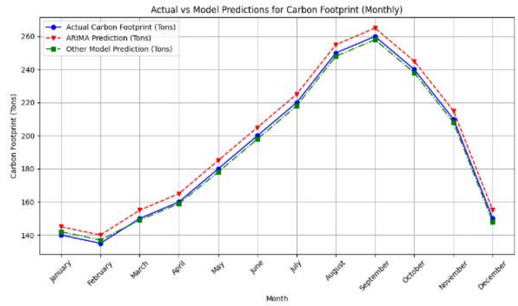
4.2.2 Comparison with Traditional Methods

To comprehensively evaluate the performance advantages of the proposed Bi-LSTM with Attention model, we compared it with various traditional methods, mainstream machine learning, and deep learning models. Table 3 provides a detailed display of the performance metrics for each model on the testing set:

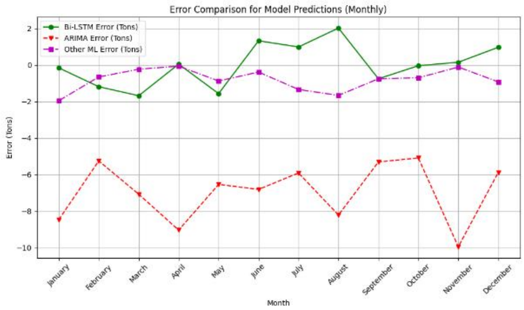
TABLE III. Performance Comparison of Different Models

Model	RMSE (tons)	MAE (tons)	R2	Training Time (s)	Prediction Time (ms/sample)
ARIMA	28.6	22.4	0.76	45	2
Random Forest	21.2	17.8	0.85	120	5
XGBoost	19.7	15.9	0.87	180	3
LSTM	17.9	14.2	0.89	350	8
GRU	17.1	13.5	0.9	320	7
Bi-LSTM with Attention	15.3	11.7	0.92	450	1

The ARIMA model performed the worst (RMSE 28.6 tons, R^2 0.76), struggling to capture nonlinear relationships. Random Forest and XGBoost showed improvement (RMSEs of 21.2 tons and 19.7 tons, respectively), but still faced challenges with long-term dependencies. LSTM and GRU demonstrated better performance in sequence modeling (RMSEs of 17.9 tons and 17.1 tons). The proposed Bi-LSTM with Attention model achieved the best performance, reducing RMSE to 15.3 tons and achieving R^2 of 0.92, representing a 46.5% improvement over ARIMA and a 10.5% improvement over GRU. This highlights the advantages of the bidirectional LSTM structure and attention mechanism. However, the model's training and prediction times are longer, necessitating a trade-off between efficiency and accuracy in real-time applications. Figure 6(a) shows the comparison of the model's predicted and actual tourism carbon footprint at different time points. It can be seen that the model prediction is highly consistent with the actual value, especially during the transition period between the peak and low seasons of tourism. Figure 6(b) shows the distribution of model prediction errors. The horizontal axis shows the prediction error (the difference between the predicted value and the actual value), and the vertical axis shows the frequency of the error. Most of the errors are concentrated near 0, which indicates that the accuracy of the model prediction is high.



(a)



(b)

Figure 6. Comparison of Prediction Curves and Error Analysis

4.2.3 Model Interpretability Analysis

To enhance the interpretability of the model, SHAP values were used to analyze feature importance. Figure 7 presents the top 10 important features and their impact on predictions:

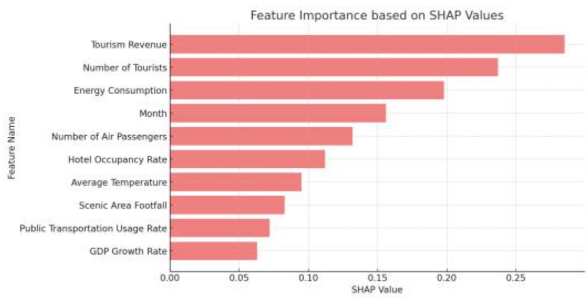


Figure 7. Top 10 Feature Importance SHAP Value Analysis

The results indicate that tourism revenue, visitor numbers, and energy consumption are the three most significant factors influencing carbon footprint predictions. Notably, tourism revenue has the highest SHAP value, highlighting the crucial role of economic factors in carbon emissions. Seasonal factors, such as the month, also show a significant impact, reflecting the cyclical nature of tourism activities. The choice of transportation modes (e.g., number of airline passengers) also plays a vital role in predicting carbon footprints, underscoring the transportation sector's importance in tourism-related emissions. This analysis not only enhances the model's interpretability but also provides a basis for developing targeted carbon reduction strategies. For instance, the findings suggest that improving energy efficiency management during peak tourism seasons and promoting low-carbon transportation options could be key measures for effectively reducing the tourism carbon footprint. Overall, the experimental results validate the effectiveness and superiority of the proposed model. The Bi-LSTM with Attention structure successfully captures the complex spatial-temporal patterns of tourism carbon footprints in Gansu Province, providing reliable predictive tools and insights for relevant decision-making.

5. Conclusion

This study successfully developed a high-accuracy tourism carbon footprint prediction model for Gansu Province using a Bidirectional Long Short-Term Memory (Bi-LSTM) network combined with an attention mechanism. The model achieved an R^2 score of 0.92 and an RMSE of 15.3 tons CO_2 on the test set, significantly outperforming traditional methods and other machine learning models. SHAP value analysis revealed that tourism revenue, visitor numbers, and energy consumption are key factors affecting carbon footprints. This model not only provides accurate predictions but also serves as an important basis for developing targeted carbon reduction strategies. The results demonstrate that this approach has a significant advantage in capturing the complex spatial-temporal patterns of tourism carbon footprints, offering reliable predictive tools and insights for related decision-making.

References

[1] Roumiani A, Shayan H, Sharifinia Z, et al. Estimation of ecological footprint based on tourism development indicators using neural networks and multivariate regression[J]. Environmental Science and Pollution Research, 2023(12):30.

[2] Xu Y, Martínez-Fernández, Silverio, Martinez M, et al. Energy Efficiency of Training Neural Network Architectures: An Empirical Study[J]. arXiv e-prints, 2023.

[3] Arévalo, Paúl, Cano A, Jurado F. Mitigation of carbon footprint with 100% renewable energy system by 2050: The case of Galapagos islands[J]. Energy, 2022, 245.

[4] Pei S, Luo J, Liang S, et al. Carbon Emissions Reduction of Neural Network by Discrete Rank Pruning[J]. CCF Transactions on High Performance Computing, 2023(3):5.

[5] Gao W, Zhao J. Prediction of geothermal temperature field by multi-attribute neural network[J]. Geothermal Energy, 2024, 12(1).

[6] Belloni E, Grasso F, Lozito G M, et al. Neural-assisted HVACs optimal scheduling for renewable energy communities[J]. Energy and buildings, 2023(Dec.):301.

[7] Zhao H, Zhou T, Long G, et al. Vote for Nearest Neighbors Meta-Pruning of Self-Supervised Networks[J]. 2022.

[8] Chen, T., Wang, M.: Deep learning-based carbon emission forecasting and peak carbon pathways in China’s logistics industry. Sustainability 16, 1826 (2024).

[9] Dong D, Jiang H, Wei X, et al. ETNAS: An energy consumption task-driven neural architecture search[J]. Sustainable computing: Informatics and systems, 2023(Dec.):40.

[10] Li B, Li W, Zhang Y. Prediction and Analysis of Agricultural Eco-Efficiency in Henan Province Based on GM-BP Neural Network[J]. The Journal of grey system, 2022.