

Analysis of Influencing Factors of Tourist Attractions Accessibility based on Machine Learning Algorithm

Na Liu^{1*}, Hai Zhang¹

1 School of Traffic and Transportation, Xi'an Traffic Engineering Institute, Xi'an, Shaanxi, 710300, China

*Corresponding author e-mail:1002626540@qq.com

Hai Zhang: 1371102993@qq.com

Abstract

A place of interest that draws travellers is referred to as a tourist attraction. These locations typically have historical significance, natural or man-made beauty, and options for leisure and amusement. Precisely estimating the demand for tourism is a significant and demanding undertaking, but research on this area is limited. In this work, we proposed a novel remora-optimized adaptive XGBoost (RO-AXGBoost) to forecast various factors associated with tourist attractions. We gathered data from Kaggle. The proposed method is implemented using Python software. The performance of these methods was evaluated using metrics such as MAPE (7.24), MAE (7.321), RMSE (10.241), and R^2 (85.7). The finding shows that the proposed method has achieved better performance and the key factors were identified as significantly influencing accessibility in tourist attractions.

Keywords: Tourist attractions, factors, tourism, remora optimized adaptive XGBoost (RO-AXGBoost)

1. Introduction

A concept of intention to return is widely recognized as a crucial component of tourism research, particularly in city tourism. In a attachment business framework, a typical promotional program requires not only to recruit new customers but also to enhance the attitudes of present consumers by developing devotion among consumers [1]. The way tourists' access and use information has changed considerably due to the impact of social media (SM). The primary reason for SM's growth is because it became a tool for shaping opinions, emotions, and experiences, making it an indispensable source of knowledge in the vacation process of making choices. The popularity of social networks on the internet has led to an increase in the number of social media influencers. SMI is characterized as an innovative kind of autonomous externally advocate who changes public perceptions through tweeting, blog posts, and other types of digital media use [2]. Tourism can be one of the more susceptible businesses to pandemics and disasters. Economic activities or public events are suspended according to World Health Organization (WHO) recommendations [3].

Sustainable travelling has become an important issue because of the environmental effects, and it is becoming increasingly difficult for modern-day consumers to disregard some of the adverse effects of their actions. The tourism industry has consistently understood that "sustainable tourists" are the most attractive visitors since they have the lowest environmental consequences and a higher purchasing power than ordinary tourists [4]. The tourism business relies heavily on information technology. The impact of technological advances on tourism has attracted attention with the introduction of smart tourism, and this phenomenon of digital transformation has spread to all industries. As a result, the notion of smart tourism technologies has become the most important aspect of the tourist business. It blends tourist resources and sophisticated information technology, allowing it to give relevant, timely data and facilitate interconnection among tourism stakeholders [5]. One of the greatest challenges that are hard to overlook is the stability issue, particularly the aspect of security. It's always important to look at the security and safety that a certain place has in relation to the travel and vacation plans. Indeed, the concern over safety and security is not an innovation in the modern world; nonetheless, it has evolved into a more significant concern over the recent few decades. In many nations, we observe civil conflict, increase in the terrorism, natural

disasters, diseases and even disasters as bases for increased concern with regard to safety and security [6]. To address this challenge, our work seeks to enhance the accuracy of forecasting in the tourist attraction variables using the RO-AXGBoost technique, newly developed and suitable for adaptive XGBoost models. We also aim to archive with better and accurate predictions of some key factors that affect tourist destinations so that better decisions can be made regarding tourism management as well as its planning.

2. Related work

Study [7] examined seasonal emotional shifts among travellers in natural forest environments using advanced methods such as Hrnet, SHAP, OSANet and, XGBoost. Positive feelings were most common throughout three years, with varying distributions among seasons. The study contributed to the development of sustainable forest tourism by revealing subtle emotional reactions to landscape aspects.

Research [8] used an IoT enabled, deep learning-based DNN with a multi-class classification algorithm recommendation system to improve visitor experiences in smart cities. Personalized suggestions were informed by real-time data and feedback from travellers. Compared to previous models, our multi-label classifier performed better, attaining high accuracy and precision. Potential biases in the source data and the requirement for a substantial IoT infrastructure were examples of limitations.

Paper [9] employed an improved STC-LSTM to forecast short-term vacation demand for travel. As compared to traditional approaches, the results show enhanced forecast accuracy. The limitations of the approach encompass dependence on certain data categories and possible problems with the model's regional generalization.

Research [10] proposed a unique strategy centered on enhancing visitor experiences and reducing congestion to overcome the shortcomings of conventional IT-RSs. Used deep reinforcement learning as opposed to just single areas of interest. Its effectiveness was evaluated against three baselines using Verona tourism data from 2014 to 2023. The heavy computing load and data requirements were among the drawbacks.

Research [11] proposed a three-stage methodology to anticipate tourism demand across numerous sites. It utilized multi-dimensional scaling to identify related attractions, mixed autoregressive models with LSTM networks to reflect spatial dependency and the scale of tourism, and suggested a method for integrating predictors to increase prediction stability. The program performed exceptionally well in estimating the number of visitors to Beijing's 77 attractions, according to the results. Data accessibility and generalizability issues were two possible limitations. The objective of the research [12] was to create a DNN model for predicting how well policies would function over time to lessen the effects of crises on the travel and tourism sector in developing nations. The model indicated that the most efficient and long-lasting strategy was to concentrate on both domestic tourist marketing and disaster preparedness, drawing on both past experiences and present difficulties. The intricacy of modelling dynamic systems and the availability of data were possible constraints.

Study [13], which focused on incoming arrivals from Hong Kong, attempted to improve the accuracy of tourist volume predictions using deep learning techniques. The suggested SAE-Bi-GRU strategy beat benchmark models like PCA-Bi-GRU with Baidu index and Google trends data. It was archived by using SAE for data dimension reduction and a Bi-GRU for model forecasting. The research's dependence on a single case study, however, presented a limitation.

Study [14] improved tourist analysis by combining deep-learning text categorization and spatial clustering approaches. Using Flickr data, it created nine tourist categories using topic modelling and LSTM classification. Spatial clustering determined the ROA for each category, indicating various elements impacting attraction. The technique provided extensive insights into visitor tastes and could be used outside tourism, although its dependence on Flickr data has limits.

The purpose of the study [15] was to better understand travellers' binary emotional experiences at Dali tourist spots. SVM and LDA models were used, as well as geographic analysis. The results reflect a general trend of good attitude, with negative evaluations for select Dali city attractions. Service experience and pricing were revealed as important elements affecting travellers' attitudes. Limitations may include a dependence on internet evaluations and the study's location.

3. Methodology

In this section, the paper proposed a novel remora-optimized adaptive XGBoost (RO-AXGBoost) to forecast various factors associated with tourist attractions. We gathered data from Kaggle.

3.1. Data collection

This paper gathers a dataset from Kaggle [16]; this dataset covers tourism attraction site data from 352 Chinese cities. Each city CSV file has 100 locations. The data comprises the location name, URL, address, site introduction, opening hours, image URL, rating, proposed visit duration, suggested season, ticket information, and tips.

3.2. Remora optimized adaptive XGBoost

RO-AXGBoost is a novel technique for anticipating several aspects associated with tourist attractions. It leverages Remora Optimization technique to adjust XGboost parameters and operates at the best model optimization. By coordinating Remora's flexibility with XGBoost's robustness, RO-AXGBoost is a potent tool for tourist destination management and planning as it enhances the machine's accuracy and efficiency in predicting crucial measurable factors into the tourism sector industry. The integration of Remora's suppleness with the robustness of XGBoost is called RO-AXGBoost, which presents viable aid to key decision-makers within the tourism sector that can forecast the significant indicators necessary for strategic planning and management of destination tourism more effectively.

3.2.1. ROA

This section aims at introducing the new meta-heuristic that has been developed, known as the ROA. ROA is based on a symbiotic float trait that remoras use to cling to bigger living organisms like swordfish and whales to be in a better position to access food. Again, like all of the MAs, the ROA optimization technique's concept is derived from biological considerations. This one has incorporated some of the features from the WOA and the SCA, which is known for its efficiency in both the local and the global search. By selecting an integer variable called H (0 or 1), one can decide whether this strategy has to use a similar exploring manner like that in SCA or WOA. This flexible manner increases using search space in the case of ROA, but if compared to other approaches to refinement, this flexible manner decreases the level of accuracy in exploration.

i. Travel with free

Using an elite approach, the SFO strategy is employed by ROA to perform a global search in the swordfish algorithm. The equation (1) for updating positions can be stated as follows:

$$u_i(p+1) = X_{best}(p) - (rand \times \left(\frac{x_{best}(p) + x_{rand}(p)}{2} \right) - x_{rand}(p)) \quad (1)$$

In this instance $u_i(p+1)$ represents a candidate position of the i -th remora. At the present moment, the optimal position is $X_{best}(t)$. An arbitrary remora location is denoted by $x_{rand}(p)$. The letter P stands for the number of iterations that are currently in progress. And between 0 and 1, $rand$ is a random number. Additionally, depending on its experiences, remora may switch hosts. This situation allows for the creation of a new candidate position through:

$$u'_i(p+1) = u_i(p+1) + randn \times (u_i(t+1) - x_i(t)) \quad (2)$$

And the i -th remora's candidate position is represented by $u'_i(p+1)$. The i th remora's prior location is denoted by $x_i(p)$. Additionally, a properly distributed random number is generated using $randn$.

ii. Through fully eat

Remora may also attach themselves to humpback whales to feed. Therefore, remora will move similarly to humpback whales. The WOA method is used in ROA to do local searches. To be more exact, the bubble-net attacking approach employed in WOA is utilized. The changed position updating formulae are provided below:

$$u_i(p+1) = E \times f^s \times \cos(2\pi\alpha) + x_{best}(p) \quad (3)$$

$$E = |X_{best}(p) - x_i(p)| \quad (4)$$

Here E is the separation between the food and the remora equation (4). Additionally, by employing the encircling prey mechanism in WOA, which is described as follows, the remora can generate a little step to further enhance the quality of the solution.

$$\mathbf{E}' = \mathbf{u}_i(\mathbf{p} + 1) - \mathbf{S} \times \mathbf{x}_{best}(\mathbf{p}) \quad (5)$$

$$\mathbf{x}_i(\mathbf{p} + 1) = \mathbf{v}_i(\mathbf{p} + 1) + \mathbf{B} \times \mathbf{E}' \quad (6)$$

Here the newly created location of the i -th remora is denoted by $u_i(p + 1)$. In ROA, the remora factor, represented by the letter B , is set to 0.1. When using the aforementioned techniques, ROA outperforms well-known meta-heuristic algorithms like WOA, SFO, and HHO equations (5 and 6).

3.2.2. Adaptive extreme Gradient Boosting (AXGBoost)

In this study, we employ AXGBoost as the core method to forecast gains and handle high-dimensional assembly faults. Then, to get predictions that are even more accurate, we suggest adaptiveXGBoost.

$$\hat{\mathbf{o}}_i = \mathbf{h}_2(\mathbf{E}_i) = \sum_{k_2=1}^{k_2} \mathbf{h}_{k_2}(\mathbf{E}_i), \mathbf{h}_{k_2} \in \phi(7)$$

Here $\phi = \{h(E) = \omega_{s(D)}\}$ is a grouping of decision trees. Each *tree* $h(E)$ corresponds to a structural parameter s and leaf weights ω , equation (7). The i -th analysis of tourist attractions is denoted by $\hat{\mathbf{o}}_i$. The model is developed by minimizing the total loss function:

$$\mathbf{M} = \sum_{i=1}^{N_t} \mathbf{m}(\hat{\mathbf{o}}_i, \mathbf{o}_i) + \sum_{k_2=1}^{k_2} \Omega(\mathbf{h}_{k_2}), \quad (8)$$

Here,

$$\Omega(\mathbf{h}_{k_2}) = \gamma d + \varepsilon \|\omega\|^2 \quad (9)$$

$\mathbf{m}(\hat{\mathbf{o}}_i, \mathbf{o}_i)$ Refers to the loss function used to calculate the difference between the realistic and expected tourism attraction. d Represents the number of decision tree leaves. The penalty period is represented by the symbol Ω , equations (8 and 9). The tuning parameters γ and ε regulate the complexity of decision trees. The square loss function was employed in this investigation.

$$\mathbf{m}(\hat{\mathbf{o}}_i, \mathbf{o}_i) = (\hat{\mathbf{o}}_i - \mathbf{o}_i)^2 \quad (10)$$

Iterative training is used for the loss function. For the s -th repetition of the i -th sample, we might write equation (8) as:

$$\mathbf{M}^d = \sum_{i=1}^{N_t} \mathbf{m}(\hat{\mathbf{o}}_i^{(d-1)} + \mathbf{h}_d(\mathbf{E}_i), \mathbf{o}_i) + \Omega(\mathbf{h}_s) \quad (11)$$

XGBoost enhances the model by adding $\mathbf{h}_s$, equation (10). The objective is optimized using a second-order approximation.

$$\mathbf{M}^d \approx \sum_{i=1}^{N_t} \left[\mathbf{m}(\hat{\mathbf{o}}_i^{(t-1)}, \mathbf{o}_i) + \mathbf{t}_i \mathbf{h}_i(\mathbf{E}_i) + \frac{1}{2} \mathbf{p}_i \mathbf{h}_s^2(\mathbf{E}_i) \right] + \Omega(\mathbf{h}_s) \quad (12)$$

$$\mathbf{T}_i = \varphi_{\hat{\mathbf{o}}_i^{(t-1)}} \mathbf{m}(\hat{\mathbf{o}}_i^{(s-1)}, \mathbf{o}_i) = 2 \times (\hat{\mathbf{o}}_i^{(s-1)} - \mathbf{o}_i), \quad (13)$$

$$\mathbf{G}_i = \varphi_{\hat{\mathbf{o}}_i^{(s-1)}}^2 \mathbf{m}(\hat{\mathbf{o}}_i^{(s-1)}, \mathbf{o}_i) = 2 \quad (14)$$

The first and second-order gradients are denoted as $\mathbf{T}_i$ and $\mathbf{G}_i$. The equations are $\mathbf{m}(\varphi_{\hat{\mathbf{o}}_i^{(s-1)}}, \mathbf{o}_i)$ equations (12-14).

An adaptive XGBoost model serves as the foundation for transfer learning, which uses parameter-based approaches. During the process of transfer training phase, leaf weights from all trees in the adaptive XGBoost model are exchanged. It is constructed by training new trees and leaf weights. Figure 1 illustrates the adaptive XGBoost framework.

$$\hat{\mathbf{o}}_{2,i} = \mathbf{g}_3(\mathbf{E}_i) = \sum_{k_2=1}^{K_2} \mathbf{h}_{k_2}(\mathbf{E}_i) + \sum_{K_5=1}^{K_5} \mathbf{h}_{k_2}(\mathbf{E}_i), \mathbf{h}_{k_2}, \mathbf{h}_{k_5}, \quad (15)$$

The TF-model predicts tourism for the i -th sample, denoted as $\hat{\mathbf{o}}_{2,i}$. $\mathbf{h}_{k_2}$, Represents the trees in the enhanced XGBoost model, whereas $\mathbf{h}_{k_5}$, is in equation (15). To increase the accuracy of tourism attraction predictions, weights of leaves in $\mathbf{h}_{k_2}$, are shared and newer leaf strengths in $\mathbf{h}_{k_5}$, are trained.

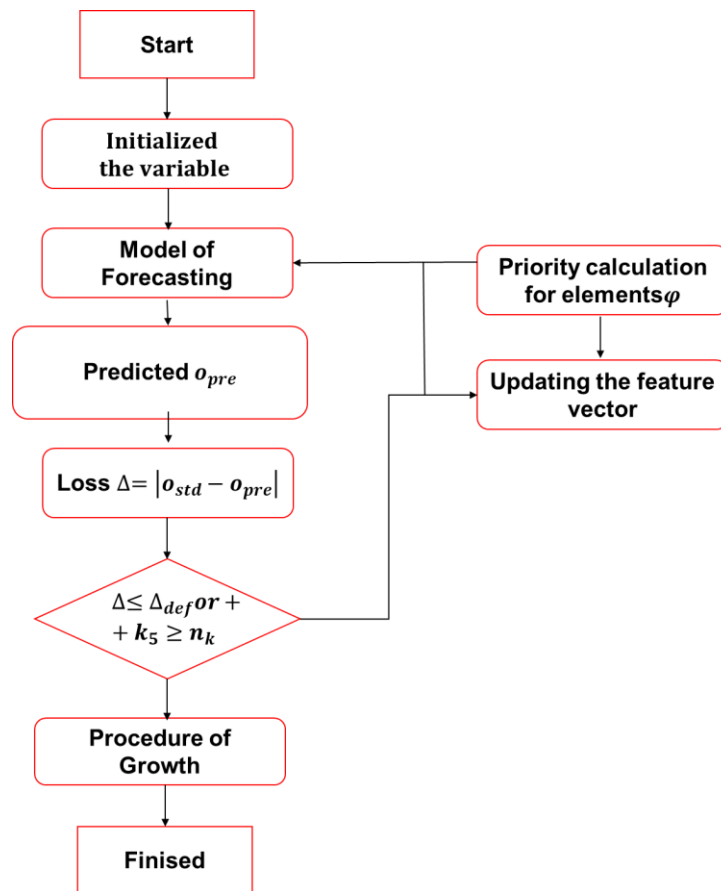


Figure 1: Adaptive XGBoost

3.2.3. Predicting RO-AXGBoost

The original technique that we used is called RO-AXGBoost, and it incorporates Remora Optimization, a newly developed meta-heuristic algorithm that refers to the behaviour of remoras about sharks in improving the capabilities of XGBoost in predicting various aspects that are related to tourist attractions. Thus, our method of utilizing the Remora Optimization procedure to enhance the adaptiveness of XGBoost hyper-parameters enhances its quality in terms of its performance and capabilities of precisely predicting labelled tourism-related variables. These two components allow for better and more consistent forecasting, helping the entities in the tourism sector to optimize their decision-making process in terms of resource management, promotion, and visitors' handling, thereby contributing to the overall improvement of the visitors' experience and the sustainability of the destinations.

4. Result and discussion

Anaconda3 for Windows is utilized in this scenario on a PC powered by an NVIDIA GeForce RTX 3090 Ti GPU and 24GB of RAM. The hard drive capacity is 1TB, and the computer's processor is the Intel Core i7-8750H running at 2.20GHz. Python 3.10 is the main programming language, using TensorFlow 1.14.0 as a development framework. In the section evaluate the performance outcome for the proposed method. The proposed method compared with existing methods and the performance is evaluated in different parameters such as MAPE, MAE, RMSE, and R^2 . The existing methods are BiLSTM [18], CNN-BiLSTM [18], MACBL [18], KNN [17], RF [17] and LR [17].

R^2 is an indicator of how accurately the variables that are independent predict the variable that is dependent. Figure 2 illustrates on R^2 value. Compared to existing methods KNN (70.6), RF (74.9), and LR (76.5), our proposed method was superior (RO-AXGB 85.7). In comparison to existing methods, the suggested method RO-AXGB showed significant improvements in tourist attraction accessibility.

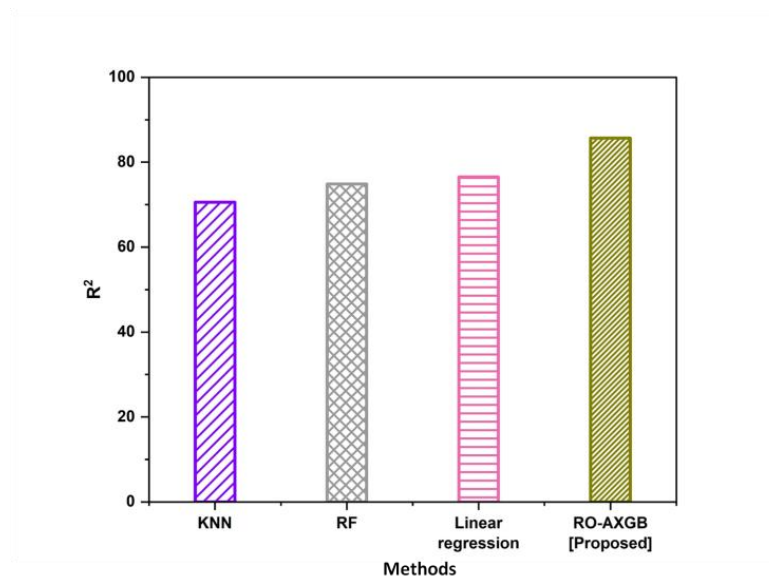


Figure 2: Outcome of R^2

The RMSE is the average of the squared discrepancies between anticipated and actual values. It indicates how much the model's predictions differ from the actual results, with lower values representing better performance. Figure 3 illustrates the RMSE result. In comparison to the existing techniques BiLSTM (23.326), CNN-BiLSTM (16.354), and MACBL (13.837), our proposed is (RO-AXGB -10.241) lower than existing methods. It shows that our suggested method, RO-AXGB is effective for tourist attractions accessibility.

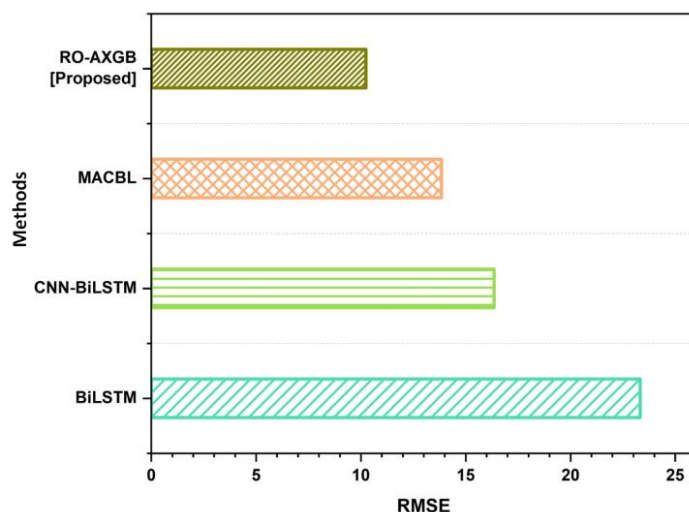


Figure 3: Outcome of RMSE

MAE estimates the average of the absolute differences between anticipated and actual values. It reveals information about the model's accuracy in forecasting, but it is less sensitive to extremes since it does not square the errors. Figure 4 illustrates the MAE values. When compared to existing approaches BiLSTM (19.146), CNN-BiLSTM (11.354), and MACBL (8.928), our suggested technique is higher (RO-AXGB -7.321) than the existing method. It demonstrates that our suggested approach, RO-AXGB, successfully detects tourist attractions accessibility.

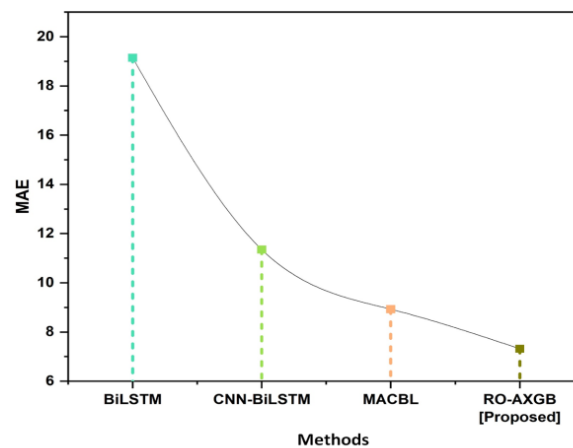


Figure 4: Outcome of MAE

MAPE is the average of the absolute percentage deviations between projected and actual values, represented as a percentage of actual values. Figure 5 depicts the MAPE result and Table 1 illustrates the comparison of existing. When compared to the existing methods, BiLSTM (21.42), CNN-BiLSTM (16.96), and MACBL (10.21), our suggested method performs better (RO-AXGB 7.24). The successful demonstration of a tourist attraction's accessibility using the RO-AXGB approach indicates its effectiveness.

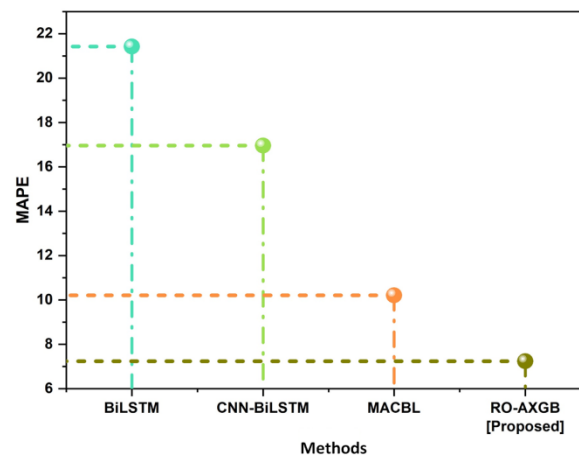


Figure 5: Outcome of MAPE

Table 1: Comparison of existing methods with proposed method

Method	RMSE	MAE	MAPE	Method	R ²
BiLSTM	23.326	19.146	21.42	KNN	70.6
CNN-BiLSTM	16.354	11.354	16.96	RF	74.9
MACBL	13.837	8.928	10.21	LR	76.5
RO-AXGB	10.241	7.321	7.24	RO-AXGB	85.7

4.1. Discussion

Though the current methods of classifying influential factors of tourist attractions' accessibilities as BiLSTM, CNN-BiLSTM, MACBL, KNN, RF, and LR shed light on the circumstances of the objective, they all have their drawbacks. For that reason, both BiLSTM[18] and CNN-BiLSTM [18] models perform well in modelling sequential dependencies as well as spatial relations but can fail to adequately model long-range dependency and

global context information. To address this issue, MACBL [18] employs a multiple-attentional mechanism in the feature extraction area to improve performance, but the additional attention mechanisms might increase the number of parameters for the resulting model and risk overfitting.

KNN [17], although the highlighted algorithm is easy to understand and implement, it has its limitations regarding its scalability and effectiveness when applied to datasets containing high dimensions. RF [17] has high resistance to overfitting and nonlinearity but can have a low capacity for modelling high-level interactions among the variables. As a flexible, interpretable and easy-to-implement method, LR [17] might not have the ability to capture the interactions and complex functional forms between the variables. Thus, although all of these methods yield valuable insights to analyze tourist attractions' accessibility, their flawed aspects point out to continued need for research and possibly a combination of all the methods to provide effective solutions regarding the various angles of the problem.

The drawback tackled, in this paper RO-AXGBoost provides an innovative approach for the interpretation of perturbing variables in the case of tourist attraction availability. Effectively applying the optimization techniques and the incorporation of adaptability to the equations make RO-AXGBoost highly efficient in capturing patterns of accessibility, which can aid in sound decision-making processes in the tourism sector. Its strength can be attributed to its capability to deal with big data, particularly concerning operating and learning in fluctuating environments and delivering solutions with noteworthy speed and accuracy.

5. Conclusion

A tourist attraction is a point of interest that draws visitors. These destinations usually feature historical value, natural or man-made beauty, and opportunities for recreation and entertainment. In this paper, we present a unique remora-optimized adaptive XGBoost (RO-AXGBoost) for forecasting different tourism attractions-related parameters. We collected data from Kaggle. The proposed approach was implemented with Python software. The performance of these approaches was assessed using measures such as MAPE (7.24), MAE (7.321), RMSE (10.241), and R^2 (85.7). The suggested approach was compared against other known algorithms. The study discovered that the suggested strategy performed better, with key parameters identified as significantly impacting accessibility to tourist destinations.

5.1. Limitation and future scope

There are some key disadvantages of inadequate transportation, some of these are; accessibility, which may result from poor connectivity; there are few linkages in this area or there are very few connection points. Since road, rail, and air transport infrastructures are constructed in education and urban planning facilities, the physically challenged are awarded in accessibility. Alleviating it to such trends as increased ecological means of transport, for example, electric buses, or bike paths, will also contribute to the improvement of the availability criterion and reduce adverse effects on the environment.

Reference

- [1] Hussein, A.S., 2020. City branding and urban tourist revisit intention: The mediation role of city image and visitor satisfaction. *International Journal of Tourism Policy*, 10(3), pp.262-279.
- [2] Pop, R.A., Săplăcan, Z., Dabija, D.C. and Alt, M.A., 2022. The impact of social media influencers on travel decisions: The role of trust in consumer decision journey. *Current Issues in Tourism*, 25(5), pp.823-843.
- [3] Rather, R.A., 2021. Monitoring the impacts of tourism-based social media, risk perception, and fear on tourist's attitudes and revisiting behavior in the wake of the COVID-19 pandemic. *Current Issues in Tourism*, 24(23), pp.3275-3283.
- [4] Holmes, M.R., Dodds, R. and Frochot, I., 2021. At home or abroad, does our behavior change? Examining how everyday behavior influences sustainable travel behavior and tourist clusters. *Journal of Travel Research*, 60(1), pp.102-116.
- [5] Azis, N., Amin, M., Chan, S. and Aprilia, C., 2020. How smart tourism technologies affect tourist destination loyalty. *Journal of hospitality and tourism technology*, 11(4), pp.603-625.

- [6] Sharma, A.K., Birendra, K.C. and Calderon, A.H., 2020. Perceived safety and security concerns among tourists in Thamel-a tourism hub in Kathmandu Valley, Nepal. *Indonesian Journal of Tourism and Leisure*, 1(2), pp.92-102.
- [7] Chen, Z., Ye, C., Yang, H., Ye, P., Xie, Y. and Ding, Z., 2024. Exploring the impact of seasonal forest landscapes on tourist emotions using Machine learning. *Ecological Indicators*, p.112115.
- [8] Cepeda-Pacheco, J.C. and Domingo, M.C., 2022. Deep learning and Internet of Things for tourist attraction recommendations in smart cities. *Neural Computing and Applications*, 34(10), pp.7691-7709.
- [9] Li, W., Guan, H., Han, Y., Zhu, H. and Wang, A., 2022. Short-term holiday travel demand prediction for urban tour transportation: a combined model based on STC-LSTM deep learning approach. *KSCE Journal of Civil Engineering*, 26(9), pp.4086-4102.
- [10] DallaVecchia, A., Migliorini, S., Quintarelli, E., Gambini, M. and Belussi, A., 2024. Promoting sustainable tourism by recommending sequences of attractions with deep reinforcement learning. *Information Technology & Tourism*, pp.1-36.
- [11] Bi, J.W., Han, T.Y. and Yao, Y., 2024. Collaborative forecasting of tourism demand for multiple tourist attractions with spatial dependence: A combined deep learning model. *Tourism Economics*, 30(2), pp.361-388.
- [12] Neshat, N., Moayedfar, S., Rezaee, K. and AmrollahiBiuki, N., 2024. Sustainable planning of developing tourism destinations after COVID-19 outbreak: A deep learning approach. *Journal of Policy Research in Tourism, Leisure and Events*, 16(1), pp.1-21.
- [13] Li, M., Zhang, C., Sun, S. and Wang, S., 2023. A novel deep learning approach for tourism volume forecasting with tourist search data. *International Journal of Tourism Research*, 25(2), pp.183-197.
- [14] Lee, H. and Kang, Y., 2021. Mining tourists' destinations and preferences through LSTM-based text classification and spatial clustering using Flickr data. *Spatial Information Research*, 29(6), pp.825-839.
- [15] Yin, X. and Jung, T., 2024. Analyzing the causes of tourists' emotional experience related to tourist attractions from a binary emotions perspective utilizing machine learning models. *Asia Pacific Journal of Tourism Research*, pp.1-20.
- [16] The data is available in online: <https://www.kaggle.com/datasets/audreyhengruizhang/china-city-attraction-details?resource=download>
- [17] Bravo, J., Alarcón, R., Valdivia, C. and Serquén, O., 2023. Application of Machine Learning Techniques to Predict Visitors to the Tourist Attractions of the Moche Route in Peru. *Sustainability*, 15(11), p.8967.
- [18] Tang, Q., Yang, L. and Pan, L., 2023. Passenger flow forecast of tourist attraction based on MACBL in LBS big data environment. *Open Geosciences*, 15(1), p.20220577.

APPENDIX-I

Description	Abbreviation
High-Resolution Net	Hrnet
Social Media	SM
Shapley Additive Explanations	SHAP
Object Semantic Attention Network	OSANet
Extreme Gradient Boosting	XGBoost
Internet of Things	IoT
Deep Neural Network	DNN
Spatial & Temporalis Correlation Long Short-Term Memory Model	STC-LSTM
Touristic Recommender Systems	T-RSs
Stacked Autoencoder-Bidirectional Gated Recurrent Unit	SAE-Bi-GRU
Principal Component Analysis	PCA
Stacked Autoencoders	SAE
Region of Attractions	ROA

Support Vector Machine	SVM
Latent Dirichlet Allocation	LDA
Remora Optimization Algorithm	ROA
Metaheuristic Algorithms	MA
Whale Optimization Algorithm	WOA
Sine Cosine Algorithm	SCA
Bi-Directional Long Short-Term Memory	BiLSTM
Convolutional Neural Network	CNN
Multi-Attentional Convolutional Bi-Directional Long Short-Term Memory	MACBL
K-Nearest Neighbors	KNN
Random Forest	RF
Linear Regression	LR
Mean Absolute Percentage Error	MAPE
Harris Hawks Optimization	HHO
Mean Absolute Error	MAE,
Root Mean Square Error	RMSE,
Coefficient of Determination	R^2
Stochastic Fractal Optimization	SFO