

# RARE: right algorithm for the right errand; a multi-model machine learning-based approach for tourism routes and spots recommendation

Ling Luo

Henan Vocational College of Information and Statistics, Zhengzhou, Henan, China

## ABSTRACT

With the globalization of the economy, tourism has emerged as a significant sector of entertainment and economic growth. Optimizing tourist attractions and routes has become crucial in modern travel planning, driven by the increasing demand for personalized recommendations. However, traditional static route-based algorithms struggle to adapt to the rapid expansion of the tourism industry, necessitating the development of dynamic, machine-learning-driven solutions. This study introduces a novel tourism recommendation system integrating multiple machine learning algorithms to provide personalized tourist spot and route recommendations. The proposed approach models the tourist map as a 2D grid of interconnected nodes, allowing for dynamic and adaptive recommendations. The framework employs long short-term memory (LSTM) for spot relevance prediction, support vector machine (SVM) for spot name classification, and depth first search (DFS) for optimal route generation. A k-means clustering approach is also utilized to designate a cluster leader (CL) responsible for managing node information within a specific zone. By inputting a simple textual query, tourists receive optimized travel routes tailored to their preferences, incorporating relevant attractions. The model is implemented in a Python-based environment and evaluated using an augmented Travel Recommendation dataset from Kaggle. Experimental results demonstrate the model's effectiveness in enhancing tourism planning and user experience, showcasing its potential for advancing intelligent tourism solutions.

Submitted 2 January 2025

Accepted 10 March 2025

Published 7 April 2025

Corresponding author

Ling Luo, Ling\_\_Luo@163.com

Academic editor

Bilal Alatas

Additional Information and  
Declarations can be found on  
page 17

DOI 10.7717/peerj-cs.2791

© Copyright  
2025 Luo

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Adaptive and Self-Organizing Systems, Agents and Multi-Agent Systems, Algorithms and Analysis of Algorithms, Autonomous Systems, Spatial and Geographic Information Systems

**Keywords** Travel routes, Tourism recommendation, Machine learning, Multi-model tourism system

## INTRODUCTION

Tourism is not a mere hobby; it is a form of entertainment that enhances the quality of people's lives (Lyu & Han, 2022). Tourists prefer to visit scenic and historical spots for serenity, solace, non-formal education, revolutionary feelings and adventurous experiences. Whether it is honeymoon tours, island tours, or red tourism, people prefer to enjoy their leisure fully, full of natural calmness and free from worldly worries. However, to endure a calm and congenial trip, route selection based on personal priorities is of utmost importance (Flisberg et al., 2012). Traditionally, tourists relied on guidebooks,

magazines, and static websites for travel recommendations. These methods lag far behind in addressing personalized priorities and dynamic demands of modern-day tourism. In the current day and age, machine learning (ML) technology may be used to revolutionize tourism, mainly to select optimal travel routes.

With the advent of internet technology, easy availability of information has become the norm of the day. This unprecedented advancement has made a remarkable surge in the tourism industry. The progress in tourism is further fueled by the emerging AI methods, which make tourism-related data ubiquitous and ensure personalized route planning (García-Madurga & Grilló-Méndez, 2023). Consequently, nowadays, tourists are privileged with many options in terms of route selection, availability of facilities and other prudent options. However, besides various challenges, man's search for an effortless model to exploit personal preference, budgetary consideration, and time constraints has yet to be fulfilled (García-Madurga & Grilló-Méndez, 2023).

Building on the success of machine learning (ML) in education, healthcare, and sports, advanced AI models can revolutionize the tourism industry by providing personalized travel recommendations. This research extends the role of AI in tourism by integrating long short-term memory (LSTM), support vector machine (SVM), and k-means clustering to generate travel routes based on tourists' preferences for attractions. The tourist map is modeled as a 2D grid of interconnected nodes, utilizing mental map (MM) and distance vector (DV) structures of Dijkstra's presentation. Since the dataset lacks distance information, additional distance and direction columns are incorporated before training. Dijkstra's algorithm is employed to calculate distances between nodes. A k-means clustering approach selects a cluster leader (CL) responsible for managing attraction data within a given zone, with CLs dynamically sharing updates.

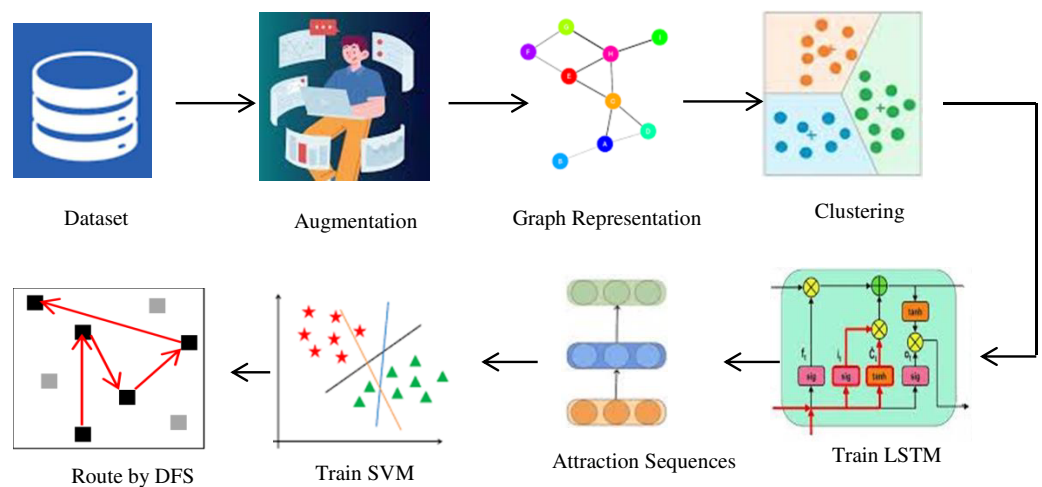
At the CL level, an LSTM network is trained on dominant attraction data, while an SVM classifier predicts spot names from spot features. Given a simple spot attraction query, LSTM generates attraction-related sequences, which are then processed by SVM to predict the optimal target spot name. The depth first search (DFS) algorithm then traces possible routes from the source to the target destination.

The model is implemented in Python using an augmented Travel Recommend Kaggle's dataset from Kaggle and evaluated on 52 randomly selected tourism queries, achieving a promising accuracy of 86%. These results validate the effectiveness and applicability of the proposed approach in enhancing personalized travel planning. The schematic representation of the RARE method is shown in Fig. 1.

Rest of the article is structured into five sections. A literature review is covered in 'Literature Review'. 'Methodology' is about the detailed methodology, whereas implementation details are presented in 'Implementation Details'. Experimentation and evaluation are reported in 'Experimentation and Result Analysis'. The last section, the 'Conclusion', concludes the article with future research directions.

## LITERATURE REVIEW

With the easy availability of transportation facilities, tourism has grown to the level of a beneficial industry. The industry plays a significant role in the economic growth of the



**Figure 1** Schematics of the RARE method.

Full-size DOI: 10.7717/peerj-cs.2791/fig-1

offering country (Malik & Kim, 2019). The effortless accessibility of tourist attraction information and luxuries enhances tourism (Basiri et al., 2018). The location attainment technologies and social and cellular networks have contributed to the uplifting of the tourism industry (Zhu et al., 2019; Cao et al., 2021). However, as mentioned in Basiri et al. (2018), challenges in tourism are yet to be resolved. The key challenges include how to tour, what to tour, and where to tour, specifically whence the destination is unfamiliar (Du et al., 2019). The prominent issue amthemu'. A number of research works have been proposed to cater to 'how to tou'. Besides the traditional approaches, data mining techniques have been utilized to help tourists 'how to tou' (Yuan et al., 2016).

Due to individual differences, tourists' preferences for various tourist attractions are different (Bao et al., 2015; Hang et al., 2018). Keeping in view the tourist preference, route recommendations can be categorized into destination recommendations and route recommendations. Where the former focuses on a single attraction satisfying the tourists' interest, the latter is about suggesting multiple routes that suit a tourist's interest. In either case, suggesting suitable routes for multiple attractions is challenging (Pu, Chen & Hu, 2012). Most of the conventional path-planning recommendation approaches are based on the shortest path (Pu, Chen & Hu, 2012). A number of methods have been suggested in this scenario, including the A\* approaches (Singh et al., 2018; Song, Liu & Bucknall, 2019), particle swarm optimization (PSO) (Pongchairerks & Kachitvichyanukul, 2016), and Genetic Algorithm (GA) (Ghosn, Drouby & Harmanani, 2016), ACO algorithm (Mirjalili, Dong & Lewis, 2020), simulated annealing (SA) (Demiral & Işik, 2020), and graph-based algorithm (Skinderowicz, 2022). The vehicle routing planning problem (Bell & McMullen, 2004) and traveling salesman problems (Zou, Yang & Zhao, 2024) are also pertinent to mention in this regard.

Recommendation systems based on user textual data have been suggested to facilitate tourists (Du et al., 2019). Similarly, the system of Zhang et al. (2012) works on travel notes to help tourists, whereas that of Chen et al. (2025) uses tourist records. The research work

of [Ye et al. \(2011\)](#) is about improving classification preference by using textual and geographical information. The information cumulative approach of [Qiao et al. \(2024a\)](#) extracts data from tourism-related blogs to easily identify attraction of interest. The work of [Yuan et al. \(2016\)](#) utilizes the big data mining approach to trace the most popular spot in a region. The Photo2Trip method proposed in [Lu et al. \(2010\)](#) exploits photos that are geotagged adequately for recommending travel routes. The pattern-aware system of [Wei, Peng & Lee \(2013\)](#) searches out the top K paths to scenic spots and recommends the optimal one. The framework of [Hu et al. \(2015\)](#) recommends urban attractions using the DBSCAN clustering algorithm. The K+V-DBSCAN algorithm in [Pla-Sacristán et al. \(2019\)](#) is suggested to accurately identify a particular zone's most significant tourist attractions. The context-aware system of [Majid et al. \(2013\)](#) exploits geotagged graphical data to recommend places based on the interest of tourists. Similarly, the *point-of-interest*-based system of [Zuo et al. \(2024\)](#) utilizes sentimental attributes for attraction recommendation. Tourists prefer to use personalized recommendation systems ([Santos et al., 2019](#)). Therefore, collaborative techniques have also been utilized for effective attraction recommendations ([Bin et al., 2019](#)). Such systems give tourists various friendly options to improve recommendations ([Park, Park & Hu, 2021](#)). A fuzzy ontology method is proposed in [Abbasi-Moud et al. \(2022\)](#). The method analyzes the tourists' reviews for future consideration. The model of [Hong & Jung \(2021\)](#) analyzed the reviews and rated significant attractions. The stability and performance of the model are assessed correctly as well. The ant colony optimization (ACO) method of [Sun et al. \(2022\)](#) recommends optimal routes by using travel pattern sequences.

Since the last decade, the techniques of machine learning (ML) are been widely utilized in the realm of tourism. For systematic analysis of tourism data, the support vector regression (SVR) model is used in [Yang & Ren \(2024\)](#). Similarly, the models based on multilayer perceptron (MLP) are used in [Batista e Silva et al. \(2018\)](#), [Li & Law \(2020\)](#) for analysis and minimizing the occurrences of errors in tourism-related information. The cutting-edge neural network (NN) was successfully implemented by [Zhang & Tang \(2022\)](#) to predict travel routes. The convolutional neural networks (CNN) based approach of [Logesh et al. \(2020\)](#) recommends attractions after analyzing the sentiment of the tourists. The model needs the necessary dataset to enhance the effectiveness of the recommendation system. [Cheng \(2021\)](#) proposed a commendable travel route commendation system using the distance matching technique ([Raees & Ullah, 2019](#)). The system considers preference or the user interests to suggest an appropriate destination. The hierarchical multi-clue fusion (HMCF) based system of [Yang et al. \(2018\)](#) exploits user-generated content (UGC) for precise prediction. The model improves the accuracy rate in tourism and enhances travel route recommendations. Similarly, a collaborative mining and filtering process (CMFP) based method is proposed in [Nan & Wang \(2022\)](#) to reduce the mining cost in travel recommendations. The Bidirectional Encoder Representations from the Transformers (BERT) model of [Wen, Liang & Zhu \(2023\)](#) are based on sentiment analysis, whereas the LSTM networks ([Law et al., 2019](#)) make use of time series forecasting in tourism. The model effectively traces long-term dependencies for demand forecasting in the tourism domain. Though most of the recommendation systems in the literature

focused on route recommendation, there is a pressing need for an effective tourism spot recommendation system based on user's interest, hence this research.

## METHODOLOGY

To assure tourists' satisfaction, the proposed method recommends tourist spots and routes based on a single-token textual input. The method intends to enhance the tourism experience by presenting an optimized route map based on personal preference. The technique utilizes machine learning classifiers to perform each subsequent process effectually. The system's algorithm is shown in Fig. 2, whereas details about each process are presented in the following sub-section.

Preceded by graph representation, a CL is selected while nodes in the region share their information with the CL. The CL keeps a record in a CSV file, which contains each attraction's information.  $A_i$  for  $i = \{1, 2, \dots, n\}$  of a particular node with spot name  $SN_j$ ;  $A_i \in SN_j$ . At the CL level, LSTM is trained by the resources/facilities available in the region. Spot name is predicted by feeding the generated word-level sequences of LSTM to the trained SVM. The benefit of using LSTM at CL is that it provides information on related attractions at the nearest tourist spots. Suppose a tourist searches for a facility  $f_x$ , the LSTM will provide a list of facilities;  $[f_x, f_{x+1}, \dots, f_{x+n}]$ . The set of attractions or facilities is then forwarded to SVM to predict the spot name for onward operation.

### Grid representation of tourist spots

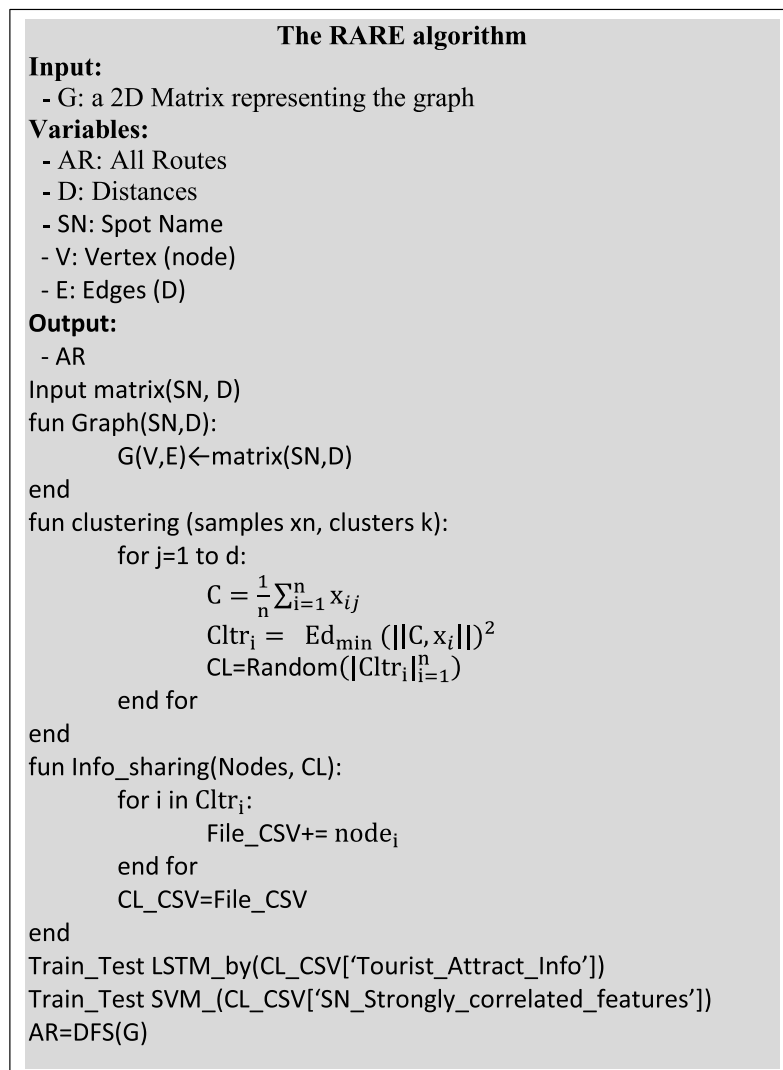
A tourism map is treated as a graph  $G$ , where  $G = \{V, E, W\}$  where the nodes  $V$  represents the visiting spots,  $E$  the edges and  $W$  is the weights (distance in Km). A path  $p(v1, v2) = [v1, v2, \dots, vn]$  represents the possible permutation between the nodes  $v1 \in V$  and  $v2 \in V$ . By utilizing the Dijkstra's Algorithm, the distance between node  $v1$ , and  $v2$  is calculated as,

$$d(v1, v2) = \min \left( \sum_{(x,y) \in T} w(e1, e2) \right) \forall \text{ paths } T \text{ from } v1 \text{ to } v2 \quad (1)$$

where  $w(e1, e2)$  is the weight between the edge  $e1$  and  $e2$ . A simple graph illustration of the following list of connected tourism spots is presented in Fig. 3.

[('Spot1', 'Spot4'), ('Spot0', 'Spot4'), ('Spot0', 'Spot2'), ('Spot4', 'Spot3')]

Determining node-to-node weights depends on examining a basic  $3 \times 3$  tourist spot grid represented by nodes. The calculated weights for nodes represent edge connections between nodes in the graph. Thus, node  $v1$  at (0,0) maintains a weight value of five units when connected to node  $v2$  at (0,1), as well as a link weight of three units between node  $v2$  and node  $v3$  at (1,1). We begin the algorithm at node  $v1$  while setting its distance to 0, then compute the minimum routes to neighboring nodes by changing their distance values according to weight information. The system uses Dijkstra's algorithm to discover the most direct path between  $v1$  (0,0) and  $v9$  (2,2) by evaluating several nodes and choosing the shortest connection. This algorithm's shortest distance tracking capability determines an optimal  $v1$ -to- $v9$  path through successive step progression.



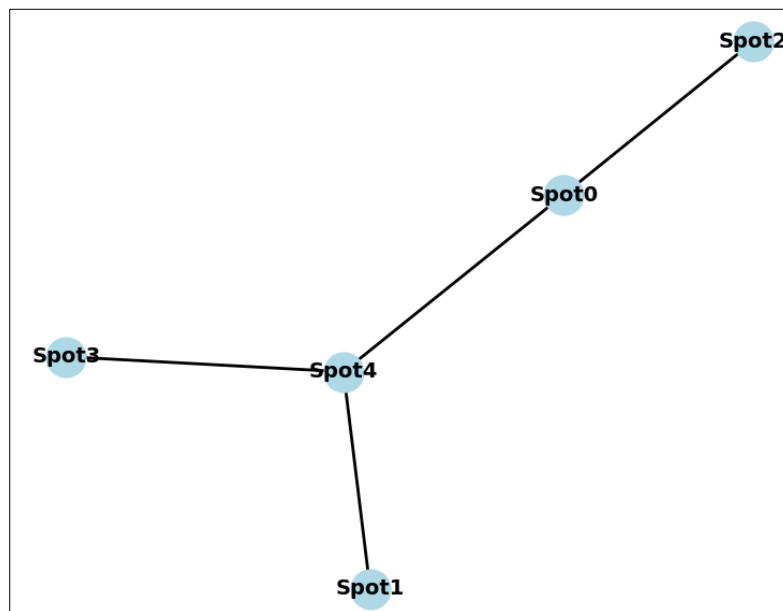
**Figure 2** The rare algorithm.

Full-size  DOI: 10.7717/peerj-cs.2791/fig-2

## Cluster leader selection

Detection of clustering patterns using k-means proves effective because it efficiently manages substantial dynamic datasets like those found in tourism data. Unsupervised learning algorithm k-means partitions data into separate clusters based on data similarity distributions. This method's simplicity and computational power suit applications that need rapid and efficient clustering solutions while processing datasets with frequently changing features, like tourist preferences and regional features. Despite the dynamic characteristics of tourism data, k-means analyzes the data consistently into distinct clusters, which help identify key interests, making this method a strong selection for our system. With the k-means clustering approach, a CL node is selected for a region containing k tourist spots. In the process of finding CL, initially k





**Figure 3** A generated graph of five connected nodes. [Full-size](#) DOI: 10.7717/peerj-cs.2791/fig-3

centroids- $\{U_1, U_2, \dots, U_k\}$  are set to a visiting spot  $S_x$  which in turn is assigned to the nearest cluster centroid ( $C_i$ ) using the given criteria,

$$C_i = \left\{ S_j : E_d ||S_j - U_i||^2 \leq E_d ||S_j - U_m||^2 \forall m = 1, \dots, k \right\} \quad (2)$$

where  $E_d$  represents Euclidean distance (Li et al., 2023). Next, the mean of the centroid is updated to accommodate the newly added spot information,

$$U_i = \frac{1}{|C_i|} \sum_{S_x \in C_i} S_x \quad (3)$$

where  $|C_i|$  represents the number of nodes in the cluster. The convergence is checked to ensure that the centroid has no change or is negligible. For this purpose, convergence is achieved as,

$$\sum_{i=1}^k ||U_i^{t+1} - U_i^t||^2 < \delta \quad (4)$$

In the given equation,  $t$  is the number of iterations and  $\delta$  being the threshold value. Our approach used empirical cross-validation testing on the tourism dataset to select the threshold value ( $\delta$ ) for the k-means clustering algorithm. Our study performed cluster analysis under various  $\delta$  settings while measuring cluster outcomes to confirm their agreement and suitability for tourist preferences. The tested threshold  $\delta$ , 0.5, served as the optimal value, generating clusters with high accuracy and efficient computation demands.

It is assumed that all the nodes of a cluster,  $n \in Cltr_x$  Have their own attraction and facilities information except the distance measured from the rest of the nodes. Though the

Source text	Training Sample
Temple Buddah Statues idles	(Temple, Buddah) (Temple, Statues) (Temple, Idles)
Hotel Family-friend hotel Pet-hotel	(Hotel, Family-friend hotel) (Hotel, Pet-hotel)
Museum Historical sites Nature park	(Museum, Historical sites) (Museum, Nature park)

**Figure 4** Attraction sequence prediction by LSTM.

Full-size  DOI: 10.7717/peerj-cs.2791/fig-4

user may feed explicit distance information, Dijkstra's algorithm may be utilized to find out the distance measures, whereas all the nodes  $n \in Cltr_x$  Shared their information with the CL. The dataset  $D$  with rows  $r$  and columns  $C$  at the CL is updated with the new feature  $f = [f_1, f_2, \dots, f_n]$  and is inserted at the second last column;  $C - 2$ . For this purpose,  $D$  is split into sub-matrices  $D_L$  with columns  $C_L$  and  $D_R$  with  $C - C_L$  columns as;

$$D_L = D[:, : C_L] \quad (5)$$

$$D_R = D[:, C_L : ] \quad (6)$$

To get the resultant dataset  $D'$ ,  $D_L$ , the new column  $C_f$  and  $D_R$  are concatenated as;

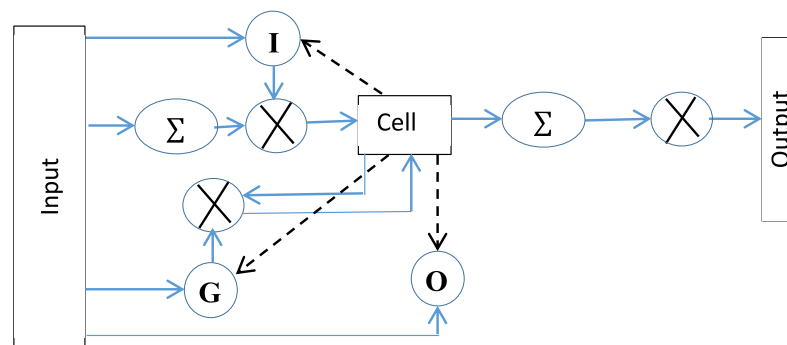
$$D' = [D_L, C_f, D_R] \quad (7)$$

### Tracing relevant facilities by LSTM

In the proposed method, the deep learning-based approach is followed to recommend attraction-related information in the textual entry of a tourist. The method predicts travel preferences after analyzing the input query by utilizing the sequence modeling and attention mechanisms. LSTM is particularly exploited to suggest optimized tourist attraction facilities based on fashion preferences. At the CL level, LSTM is trained based on the attraction information offered by region nodes. Taking the input token from the tourist, LSTM presents related attraction information of the tourist spots, as shown in Fig. 4. The output of LSTM is further fed to SVM to name the appropriate tourist spot.

LSTM is the enhanced recurrent neural network (RNN) deep learning model where the vanishing gradient issue is overcome by backpropagation (Karmiani et al., 2019), where RNN is an artificial neural network (ANN). In RNN, the nodes are connected to form a temporal sequence directed graph. At each stage, the output of RNN is to coincide to generate new output. However, the network suffers from the issue of preserving the extensive past information (Supakar, Satvaya & Chakrabarti, 2022). LSTM is the modified and advanced form of RNN (Supakar, Satvaya & Chakrabarti, 2022) with the ability to





**Figure 5** The standard LSTM structure, represents activation function.

Full-size DOI: 10.7717/peerj-cs.2791/fig-5

back-propagate to retain a large number of related information If' is the hidden node at a particular time instance t, its output Oh t is given as

$$O_h^t = a^h \left( \sum_{i=0}^{I-1} CW^{hi} x_t^i + \sum_{m=0}^{Th-1} u^{hm} b_{t-1}^m + c^h \right) \quad (8)$$

where,  $a^h$  is the activation function at node h and  $c^h$  is the bias value at a particular node". I and Th represent the total number of inputs and the number of nodes in a layer, respectively.  $x_t^i \in R$  is the ith input occurring at node h at time t,  $CW \in R$  connection weight between input 'i' and node 'h'. Similarly,  $u^{hm} \in R$  is the connection weight between the nodes h and m, which belong to the same layer, whereas  $b_{t-1}^m$  is the output of at m node in time t – 1.

A simple architecture of LSTM is represented by three gates with a memory cell. The gates control information to and from the cell. The input gate manages the input of the flow of information into the memory. The forget gate controls how long information be retained in the cell, whereas the output gate is responsible for generating the output of the LSTM. The working of LSTM, as depicted in Fig. 5 is represented mathematically as follows,

$$I_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (9)$$

$$F_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (10)$$

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (11)$$

$$C_t = A_f \circ C_{t-1} + A_i \circ \sigma_c (W_c x_t + U_c h_{t-1} - 1 + b_c). \quad (12)$$

In the given equations, input, forget, and output gates are represented by I, F and O, respectively. Weight matrices are represented by W (e.g.,  $W_i$ ,  $W_f$ ,  $W_o$  are weight matrices for input, forget and output neuron, respectively). The previous state of LSTM, at timestamp t – 1, is represented by  $h_{t-1}$  whereas the bias value for a gate is represented by

$b_x$ . The network's output activation functions (softmax) is represented by  $\sigma$ . Similarly,  $\sigma_c$  is the cell's hyperbolic tangent activation function.  $W_c$  and  $U_c$  is the matrix weights of the connections and recurrent connections respectively to the memory cell of LSTM.  $A_f$ ,  $A_i$  and  $C_t$  are the activation functions of the memory cell's forget gate, input gate and state vector.  $b$  represents bias it may be  $b_i$ ,  $b_f$ ,  $b_o$ , or  $b_c$  which shows bias at input, forget, output gate, or at the memory cell, respectively. The operation ( $\circ$ ) represents the Hadamard product.

### Tracing of tourist Spot by SVM

The SVM classifier is one of the effective supervised ML models used for regression and classification. In its simple form (binary classification), the model intends to find a hyperplane between two classes with class labels  $L_i \in \{-1, 1\}$  for the given training dataset  $\{x_i, L_i\}_{i=1}^n$  where  $x_i \in \mathbb{R}^d$  the feature vector. The hyperplane is given as,

$$w \cdot x + b = 0$$

where  $w$  is the weight and  $b$  the bias.

For an unknown input point  $x_x$ , the decision function  $f(x_x)$  is given as,

$$f(x_x) = \text{sign}(w \cdot x_x + b) \quad (13)$$

In case of high dimensional data, the decision function be as,

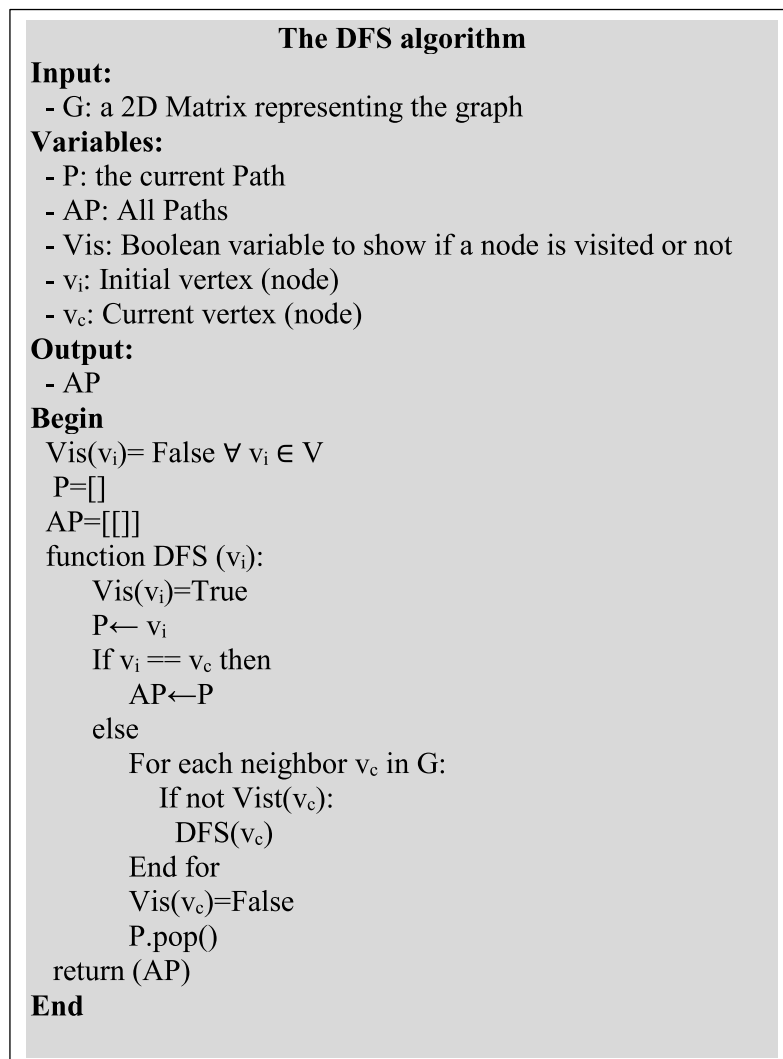
$$f(x_x) = \text{sign}\left(\sum_{i=1}^n \phi_i c_i(x_i \cdot x_x) + b\right) \quad (14)$$

where  $\phi_i$  is the Lagrange multiplier.

In the proposed method, the SVM is trained by features strongly correlated with the spot name. The attraction facilities of tourist spots-  $X = \{f_1, f_2, \dots, f_n\}$  are taken as features, whereas the names of spots as labels;  $N = \{c_1, c_2, c_3, \dots, c_n\}$ . The features, as found out by LSTM, are fed to the SVM to prompt the most appropriate tourist spot name, SN. Multiple elements determine the LSTM algorithm's decreasing accuracy during extended sequence length prediction. The algorithm's nature makes it difficult to process elaborate sequences because the vanishing gradient problem affects it even with its built-in long-range dependency capabilities. Longer sequences cause the model to lose its ability to capture important sequence data from previous parts, making predictions less precise. The escalating number of sequence configurations leads to improved model overfitting, which degrades the model's ability to adapt to new data.

### Routes recommendation by DFS

The effective DFS is exploited for route recommendation ([Hagerup, 2020](#); [Gers & Schmidhuber, 2001](#); [Gedela & Karthikeyan, 2022](#); [Ni et al., 2021](#); [Li et al., 2021](#); [Qiao et al., 2024b](#); [Zeng et al., 2015](#); [Paulson & Tzanavari, 2003](#); [Abadi et al., 2016](#); [Rafay, Suleman & Alim, 2020](#)). The algorithm offers a comprehensive exploration of all possible routes based on the priorities and constraints of a user. Once a spot name is suggested, routes from the predicted spot to the rest of the spots in a *Cltr* are searched out by DFS.



**Figure 6** The routes finding DFS algorithm.

Full-size  DOI: 10.7717/peerj-cs.2791/fig-6

The algorithm is designed to explore all possible paths recursively till the end of the destination node. Let  $G = \{V, E\}$  represents the tourist spots in a *Cltr* with nodes  $V = \{v_1, v_2, \dots, v_n\}$  and edges  $E = \{e_1, e_2, \dots, e_n\}$ . Any edge  $e_i \in E$  is a pair  $(v_1, v_2)$ , where  $v_1, v_2 \in V$ . The steps of the algorithm are presented in Fig. 6. Suppose a cluster  $Cltr_x$  Has five such nodes (A, B, C, D, E) containing the required or matching tourism facilities; as predicted by the LSTM, the possible routes, as recommended by DFS, are shown in Fig. 7.

In a nutshell, key contributions of this research work include: (1) a machine learning framework is proposed to provide tourist spots with travel routes based on the requirements or priorities of the tourists. (2) The method is implemented in a case-study project by using a real-world dataset and augmented with the required attributes. (3) The results obtained are analyzed, and they show promising improvements in contemporary systems.

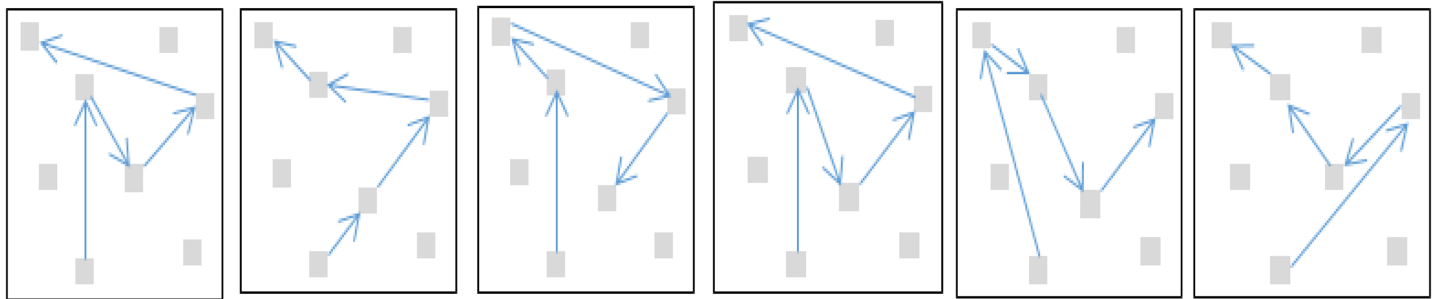


Figure 7 The recommended routes as generated by the DFS algorithm.

Full-size DOI: 10.7717/peerj-cs.2791/fig-7

## IMPLEMENTATION DETAILS

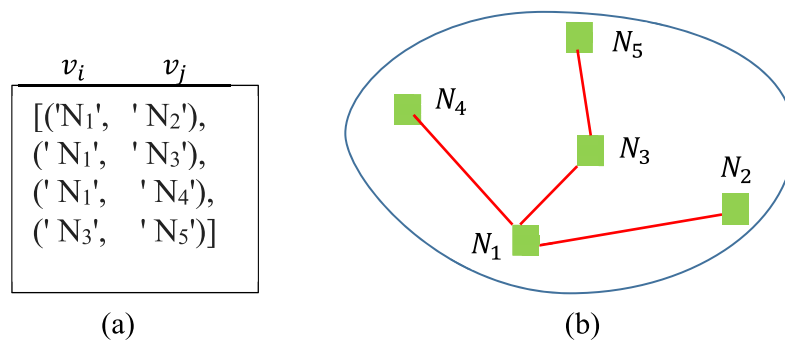
The proposed method is implemented in VSCode using various advanced machine learning libraries, including TensorFlow, Keras, and Python networks. The system contains four modules: representation, selection, information tracing and spot-name tracing. In the *representation* module, data is taken from the dataset and is represented in 2D graph form using Dijkstra's algorithm from the Networkx library (Chollet, 2015; Jabbar, Abass & Hasan, 2023). In the said module, the spot names and their distances are taken as inputs. A simple demonstration of the module is shown in Fig. 8. The computational complexity of the algorithms varies. LSTM has higher complexity due to its recurrent nature, while SVM can be slow with large datasets. DFS is efficient for small datasets but could be optimized for larger ones.

In the *selection* module, the CL is selected out of 'n' nodes in the cluster using the centroid-based approach. Each node 'n' shares its available attraction information with the CL, which is appended to the dataset and is stored in a CSV file. Some of the tourism information of a CL in tabular form is presented in Table 1.

In the *information-tracing* module, LSTM is trained by the available information  $I_i$  of the node  $N_i$  where  $(I_i, N_i) \in CL$ . LSTM is an effective model for learning context-free and context sensitive languages (Gers & Schmidhuber, 2001). LSTM retains previously learned knowledge through a memory cell at the hidden layer. In the proposed system, LSTM is used to map input sequence  $I\beta_i = (s_1, s_2 \dots s_n)$  to output sequence  $\Omega = (os_1, os_2, \dots os_n)$ . By applying unit activation iteratively, where  $I\beta$  and  $\Omega$  represent tourist information by which it is trained. A predicted sequence  $\Omega_x$  is concatenated to the formerly predicted sequence  $\Omega_{x-1}$  to generate the next sequence  $\Omega_{x+1}$ . This sequence prediction process is repeated till the last token term;  $k_1$  as shown in the following equations for the predicted sequence.

$$PS_r = \sum_{i=0}^{k1} \omega(\Omega_i). \quad (15)$$

An illustration of generated PS is shown in Fig. 9 by entering Restaurant as an input to the LSTM.

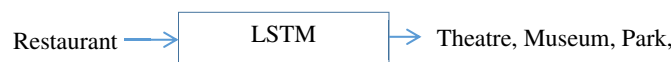


**Figure 8** (A) List of the connected nodes and (B) graph representation of the algorithm.

Full-size DOI: 10.7717/peerj-cs.2791/fig-8

**Table 1** Tourism information maintained at CL level.

Name	State	Type	Popularity	Best time to visit	Currency	Culture	Distance (Km)
Taj Mahal	Uttar Pradesh	Historical	8.6919062031	Nov–Feb	Indian (Rs)	Indian	123



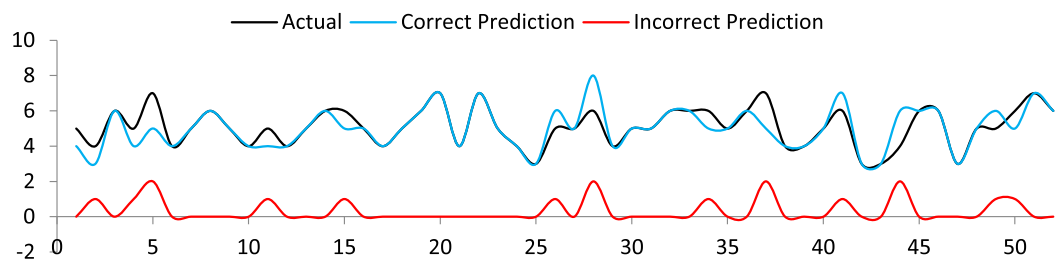
**Figure 9** Input sequence and predicated sequence by LSTM.

Full-size DOI: 10.7717/peerj-cs.2791/fig-9

The spot-name tracing module deals with the tracing of spot names based on the generated PS of LSTM. To predict the optimal Spot-Name, the features train SN, SVM  $x_i \in X$ ,  $X$  being the training dataset and  $y_i \in Y$  as class label where  $Y = X[SN]$ .

## EXPERIMENTATION AND RESULT ANALYSIS

Unlike other deep learning methods where a classifier is trained, validated and tested by different dataset splits, the proposed system is trained by all tuples of the dataset that lies at a CL. The entire dataset was used solely to ensure the most optimal results. A data augmentation approach was implemented to boost model robustness and enhance generalization because we added artificial dataset elements to the input data. Our approach expanded rows through multiple data transformation tools, including random dimensional changes, scaling operations and flipping mechanics on tourist location records. We developed synthetic features (columns) from historical travel data to analyze potential tourist behavior characteristics, including seasonality choices and travel window behaviors. Observational tests verified the usefulness of the enlarged dataset, which resulted in a 15% increase in model accuracy, thus validating the performance-enhancing effects of these data improvements. The original Kaggle dataset was augmented heuristically by 83 rows and 12 columns. The system was evaluated 52 times by entering textual tourism query— $TQ_i$  of variant dimensions;  $TQ_i = \{1, 2, \dots, 7\}$ . A  $TQ_x$  was



**Figure 10** Actual sequence, correct and incorrect sequence prediction by LSTM.

Full-size DOI: [10.7717/peerj-cs.2791/fig-10](https://doi.org/10.7717/peerj-cs.2791/fig-10)

**Table 2** Results of the sequence prediction by LSTM.

Accuracy (%)	Precision	F1 score	Specificity	Sensitivity
82	0.9	0.89	0.5	0.88

obtained by feeding a random attraction term to the information tracing module. Though each of the evaluation iterations assessed the entire system, from sequence-related tourism information to possible routes, the results of each section were analyzed individually. The 52 queries used for evaluation were carefully selected to cover a broad spectrum of tourism scenarios, ensuring that they reflect diverse tourist preferences and behaviors. These queries were designed to represent various trip types, including cultural, historical, and recreational tourism, and they included different levels of complexity, such as simple requests for specific attractions or more detailed queries involving routes between multiple spots. The selection process aimed to ensure the model was tested under conditions that closely mirror real-world tourism planning, capturing the full range of user preferences and geographic contexts.

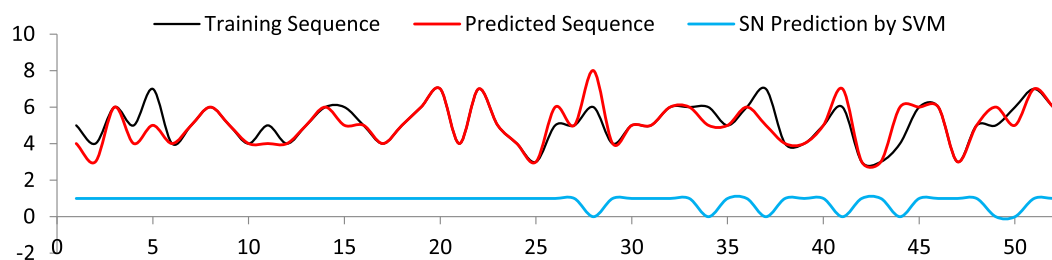
LSTM was assessed to generate the available tourist attraction information in 1 and 7 sequences. A user only needs to input a textual term and the number of sequences to be generated. The accuracy of LSTM was up to the mark for a lesser number of sequences. However, irrelevant terms were predicted in case of a greater number of sequences. This is clear from the correct and irrelevant sequence prediction graph shown in Fig. 10. As a whole, reasonable accuracy, precision, and F1 score were obtained, as shown in Table 2. The formulas for precision, sensitivity, specificity and F1 score (Gedela & Karthikeyan, 2022) are follows.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (16)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (17)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (18)$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (19)$$



**Figure 11** Spot name prediction with respect to training sequence, and predicted sequence.

Full-size DOI: [10.7717/peerj-cs.2791/fig-11](https://doi.org/10.7717/peerj-cs.2791/fig-11)

**Table 3** Results of the SVM model.

Accuracy (%)	Precision	F1 score	Specificity	Sensitivity
90	0.93	0.94	0.4	0.95

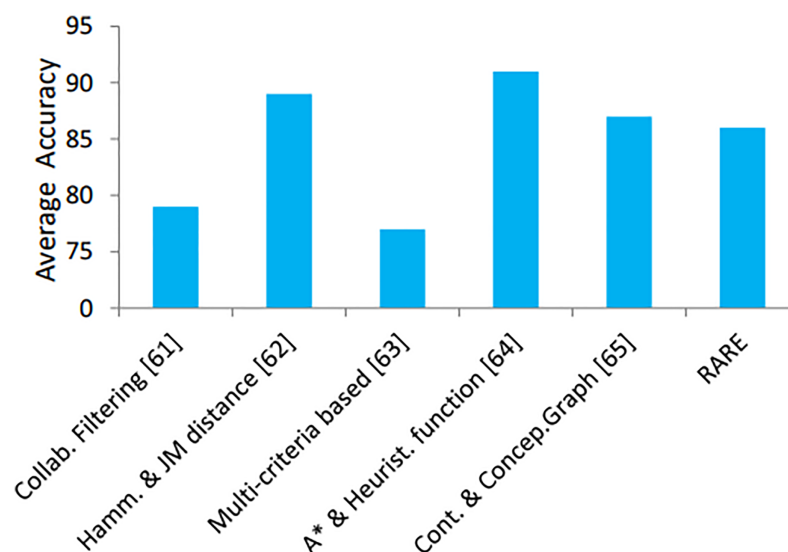
**Table 4** Details of the state-of-the-art recommendation methods.

Author(s)	Title	Method used	Remarks
Ni et al. (2021)	“Collaborative Filtering Recommendation Algorithm Based on TF-IDF and User Characteristics”	Collaborative filtering based on (TF-IDF) and fuzzy membership	Needs to find similar users and resources
Li et al. (2021)	“Intelligent recommendation model of tourist places based on collaborative filtering and user preferences”	Hamming and Jeffries-Matusita distance approach	User preferences is supported and the issue of sparse data is addressed
Qiao et al. (2024a)	“Tourist Recommender System using Hybrid Filtering”	Multi-criteria based containing content, C-filtering, and Ontology	Predicts the ratings of demography and POIs to satisfy needs of users
Zeng et al. (2015)	“Optimal Route Search with the Coverage of Users’ Preferences”	Algorithm based on A* and heuristic function	The weighted user preference is considered in route researching, proposes convergence solution for optimal route.
Paulson & Tzanavari (2003)	“Combining Collaborative and Content-Based Filtering Using Conceptual Graphs”	Conceptual graphs and dataset contents for prediction and error reduction	Support only a limited fuzzy value fields and tourism attractions
RARE	The proposed system	Multi-model based method including SVM, LSTM and DFS	Tourist preference is supported to trace spot and routes.

The generated sequences are forwarded to SVM to predict the appropriate spot names. Most of the spot name classes were accurately predicted. In the odd cases where the predicted sequences are fewer than the training sequences, incorrect class names were expected, see Fig. 11. Analysis of the results obtained for the SVM is presented in Table 3.

The DFS algorithm precisely depicts all the routes between a current node,  $n_c$  and the predicted destination route  $n_p$ . Hence, the accuracy of the algorithm was simply 100%. Our augmentation technique that builds synthetic features using historical information has the potential to strengthen existing patterns within the original dataset. Specific tourist





**Figure 12** Accuracy comparison of the state-of-the-art recommendation methods.

Full-size DOI: [10.7717/peerj-cs.2791/fig-12](https://doi.org/10.7717/peerj-cs.2791/fig-12)

demographic biases often become more pronounced when we use historical data to determine popular destinations. Instrumented modifications to original data through special transformations such as scaling and random rotations might not adequately interpret natural modifications in tourism behavior patterns. The model's generalization capability faces challenges when using the data for areas lacking adequate representation within the initial dataset. Real-world data performance assessments led to continual updates of the augmentation system that reduced the effect of biases discovered in the process.

In addition, the proposed method is compared with some up-to-date, cutting-edge approaches, see Table 4. Comparison in terms of accuracy is shown in Fig. 12. The performance of RARE surpasses contemporary methods presented in Table 4 because it utilizes SVM and LSTM together with DFS in its architectural design. RARE defeats traditional techniques such as collaborative filtering and hybrid filtering by implementing LSTM to track unique visitor preferences. The DFS algorithm delivers efficient route optimization through tourist priority interests and geographic limitations, which other methods like A\* do not handle properly. The unified framework of models enables RARE to generate travel guidance that surpasses the accuracy, adaptability, and tailored nature of standalone approaches.

The model does not currently incorporate temporal data, such as opening and closing times, which could affect the accuracy of recommendations for time-sensitive travel plans. We plan to address this limitation in future work by integrating such data to improve the model's relevance in dynamic tourism scenarios. Future work will also focus on integrating GIS data to enable the automatic detection and localization of tourist spots, reducing reliance on explicit user input and further improving the system's usability.

## CONCLUSION

Tourism is the modern-day valuable entertainment. Tourism adventure can be made more attractive if routes and spots are based on personal preference. This research proposes a multimodal approach for tourist routes and spot recommendations based on the personal priority of the visitor. The method exploits the approach of a right algorithm for the right task. To keep the attraction information of all the nodes, the k-mean clustering algorithm is used to select a cluster leader in a particular zone. The LSTM model is trained to generate relevant attraction information, whereas the SVM classifier is used to suggest the most optimal spot name based on the generated input sequences. By inputting a simple textual demand term, the matching and relevant spots are suggested, as well as presenting travel routes to the various tourist spots. The method is implemented in a Python based project and is evaluated by an augmented real-world dataset. The proposed method is systematically assessed and compared with the contemporary state-of-the-art recommendation systems. The satisfactory accuracy revealed that the process is applicable for enhancing tourism. The system has strong potential for real-world applications in large-scale tourism datasets, such as in smart tourism platforms. By leveraging its ability to process diverse tourist preferences and optimize routes, it can scale to accommodate millions of users. Future improvements, including better data integration and performance optimization, will enhance its scalability for large, dynamic datasets.

Though the research presents a novel approach for tracing tourist spot based on user's preferences, the system fails to consider the opening and closing times, often critical for tourism planning. As our future strategy, we are scheming to enhance the model. Moreover, the GIS data will be incorporated into the dataset to automatically locate tourist spots without explicit inputting. As a whole, the research contributes to the realm of tourism and paves the way for future research.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study was supported by the Annual project of Philosophy and Social Science Planning of Henan Province in 2024 (No. 2024BJJ166). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Annual Project of Philosophy and Social Science Planning of Henan Province in 2024: 2024BJJ166.

### Competing Interests

The authors declare that they have no competing interests

## Author Contributions

- Ling Luo conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

The code and raw data are available in the [Supplemental Files](#) and at GitHub and Zenodo:

- <https://github.com/halowisata/ViVe-Machine-Learning-Flask>.

- Ling, L. (2025). RARE: Right Algorithm for the Right Errand; A multi-model machine learning-based approach for tourism routes and spots recommendation [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.15038384>.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.2791#supplemental-information>.

## REFERENCES

- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M, Levenberg J, Monga R, Moore S, Murray Dg, Steiner B, Tucker P, Vasudevan V, Warden P, Wicke M, Yuan Y, Zheng X. 2016. {TensorFlow}: a system for {Large-Scale} machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265–283.
- Abbasi-Moud Z, Hosseinabadi S, Kelarestaghi M, Eshghi F. 2022. CAFOB: context-aware fuzzy-ontology-based tourism recommendation system. *Expert Systems with Applications* **199**:116877 DOI [10.1016/j.eswa.2022.116877](https://doi.org/10.1016/j.eswa.2022.116877).
- Bao J, Zheng Yu, Wilkie D, Mokbel M. 2015. Recommendations in location-based social networks: a survey. *Geo-Information* **19**(3):525–565 DOI [10.1007/s10707-014-0220-8](https://doi.org/10.1007/s10707-014-0220-8).
- Basiri A, Amirian P, Winstanley A, Moore T. 2018. Making tourist guidance systems more intelligent, adaptive and personalized using crowd sourced movement data. *Journal of Ambient Intelligence and Humanized Computing* **9**(2):413–427 DOI [10.1007/s12652-017-0550-0](https://doi.org/10.1007/s12652-017-0550-0).
- Batista e Silva F, Marín Herrera MA, Rosina K, Ribeiro Barranco R, Freire S, Schiavina M. 2018. Analyzing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management* **68**(3):101–115 DOI [10.1016/j.tourman.2018.02.020](https://doi.org/10.1016/j.tourman.2018.02.020).
- Bell JE, McMullen PR. 2004. Ant colony optimization techniques for the vehicle routing problem. *Advanced Engineering Informatics* **18**(1):41–48 DOI [10.1016/j.aei.2004.07.001](https://doi.org/10.1016/j.aei.2004.07.001).
- Bin C, Gu T, Sun Y, Chang L. 2019. A personalized POI route recommendation system based on heterogeneous tourism data and sequential pattern mining. *Multimedia Tools and Applications* **78**(24):35135–35156 DOI [10.1007/s11042-019-08096-w](https://doi.org/10.1007/s11042-019-08096-w).
- Cao D, Zeng K, Wang J, Sharma PK, Ma X, Liu Y, Zhou S. 2021. BERT-based deep spatial-temporal network for taxi demand prediction. *IEEE Transactions on Intelligent Transportation Systems* **23**(7):9442–9454 DOI [10.1109/TITS.2021.3122114](https://doi.org/10.1109/TITS.2021.3122114).

- Chen H, Bei Y, Huang W, Chen S, Huang F, Huang X. 2025.** Graph cross-correlated network for recommendation. *IEEE Transactions on Knowledge and Data Engineering* **37(2)**:710–723 DOI [10.1109/TKDE.2024.3491778](https://doi.org/10.1109/TKDE.2024.3491778).
- Cheng X. 2021.** A travel route recommendation algorithm based on interest theme and distance matching. *EURASIP Journal on Advances in Signal Processing* **2021(1)**:57 DOI [10.1186/s13634-021-00759-x](https://doi.org/10.1186/s13634-021-00759-x).
- Chollet F. 2015.** Keras. GitHub. Available at <https://github.com/fchollet/keras>.
- Demiral MF, Işık AH. 2020.** Simulated annealing algorithm for a medium-sized TSP data. In: *Artificial Intelligence and Applied Mathematics in Engineering Problems: Proceedings of the International Conference on Artificial Intelligence and Applied Mathematics in Engineering (ICAIAE 2019)*. Cham: Springer International Publishing, 457–465.
- Du S, Zhang H, Xu H, Yang J, Tu O. 2019.** To make the travel healthier: a new tourism personalized route recommendation algorithm. *Journal of Ambient Intelligence and Humanized Computing* **10(9)**:3551–3562 DOI [10.1007/s12652-018-1081-z](https://doi.org/10.1007/s12652-018-1081-z).
- Flisberg P, Lidén B, Rönqvist M, Selander J. 2012.** Route selection for best distances in road databases based on drivers' and customers' preferences. *Canadian Journal of Forest Research* **42(6)**:1126–1140 DOI [10.1139/x2012-063](https://doi.org/10.1139/x2012-063).
- García-Madurga M-Á, Grilló-Méndez A-J. 2023.** Artificial intelligence in the tourism industry: an overview of reviews. *Administrative Sciences* **13(8)**:172 DOI [10.3390/admsci13080172](https://doi.org/10.3390/admsci13080172).
- Gedela B, Karthikeyan PR. 2022.** Credit card fraud detection using AdaBoost algorithm in comparison with various machine learning algorithms to measure accuracy, sensitivity, specificity, precision and F-score. In: *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. Piscataway: IEEE, 1–6.
- Gers FA, Schmidhuber J. 2001.** Long short-term memory learns context free and context sensitive languages. In: *Artificial Neural Nets and Genetic Algorithms: Proceedings of the International Conference in Prague, Czech Republic*. Vienna: Springer, 134–137.
- Ghosn SB, Drouby F, Harmanani HM. 2016.** A parallel genetic algorithm for the open-shop scheduling problem using deterministic and random moves. *International Journal of Artificial Intelligence* **14(1)**:130–144.
- Hagerup T. 2020.** Space-efficient DFS and applications to connectivity problems: simpler, leaner, faster. *Algorithmica* **82(4)**:1033–1056 DOI [10.1007/s00453-019-00629-x](https://doi.org/10.1007/s00453-019-00629-x).
- Hang L, Kang S-H, Jin W, Kim D-H. 2018.** Design and implementation of an optimal travel route recommender system on big data for tourists in Jeju. *Processes* **6(8)**:133 DOI [10.3390/pr6080133](https://doi.org/10.3390/pr6080133).
- Hong M, Jung JJ. 2021.** Multi-criteria tensor model for tourism recommender systems. *Expert Systems with Applications* **170**:114537 DOI [10.1016/j.eswa.2020.114537](https://doi.org/10.1016/j.eswa.2020.114537).
- Hu Y, Gao S, Janowicz K, Yu B, Li W, Prasad S. 2015.** Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems* **54(4)**:240–254 DOI [10.1016/j.compenvurbsys.2015.09.001](https://doi.org/10.1016/j.compenvurbsys.2015.09.001).
- Jabbar LS, Abass EI, Hasan SD. 2023.** A modification of shortest path algorithm according to adjustable weights based on Dijkstra algorithm. *Engineering and Technology Journal* **41(2)**:359–374 DOI [10.30684/etj.2022.136107.1296](https://doi.org/10.30684/etj.2022.136107.1296).
- Karmiani D, Kazi R, Nambisan A, Shah A, Kamble V. 2019.** Comparison of predictive algorithms: backpropagation, SVM, LSTM and Kalman Filter for stock market. In: *2019 Amity International Conference on Artificial Intelligence (AICAI)*. Piscataway: IEEE, 228–234.
- Law R, Li G, Fong DKC, Han X. 2019.** Tourism demand forecasting: a deep learning approach. *Annals of Tourism Research* **75**:410–423 DOI [10.1016/j.annals.2019.01.014](https://doi.org/10.1016/j.annals.2019.01.014).

- Li X, Law R. 2020. Network analysis of big data research in tourism. *Tourism Management Perspectives* 33(5):100608 DOI 10.1016/j.tmp.2019.100608.
- Li T, Li Y, Xia T, Hui P. 2021. Finding spatiotemporal patterns of mobile application usage. *IEEE Transactions on Network Science and Engineering* 99:1 DOI 10.1109/TNSE.2021.3131194.
- Li T, Li Y, Zhang M, Tarkoma S, Hui P. 2023. You are how you use apps: user profiling based on spatiotemporal app usage behavior. *ACM Transactions on Intelligent Systems and Technology* 14(4):1–21 DOI 10.1145/3597212.
- Logesh R, Subramaniaswamy V, Malathi D, Sivaramakrishnan N, Vijayakumar V. 2020. Enhancing recommendation stability of collaborative filtering recommender system through bio-inspired clustering ensemble method. *Neural Computing and Applications* 32(7):2141–2164 DOI 10.1007/s00521-018-3891-5.
- Lu X, Wang C, Yang J-M, Pang Y, Zhang L. 2010. Photo2Trip: generating travel routes from geo-tagged photos for trip planning. In: *Proceedings of the 18th ACM international conference on Multimedia*. New York: ACM, 143–152.
- Lyu H, Han Z. 2022. Algorithm of scenic spot tourism route planning based on convolution neural network. *Wireless Communications and Mobile Computing* 2022(1):5330611 DOI 10.1155/2022/5330611.
- Majid A, Chen L, Chen G, Mirza HT, Hussain I, Woodward J. 2013. A context-aware personalized travel recommendation system based on geotagged social media data mining. *International Journal of Geographical Information Science* 27(4):662–684 DOI 10.1080/13658816.2012.696649.
- Malik S, Kim DH. 2019. Optimal travel route recommendation mechanism based on neural networks and particle swarm optimization for efficient tourism using tourist vehicular data. *Sustainability* 11(12):3357 DOI 10.3390/su11123357.
- Mirjalili S, Dong JS, Lewis A. 2020. Ant colony optimizer: theory, literature review, and application in AUV path planning. *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications* 811:7–21 DOI 10.1007/978-3-030-12127-3.
- Nan X, Wang X. 2022. Design and implementation of a personalized tourism recommendation system based on the data mining and collaborative filtering algorithm. *Computational Intelligence and Neuroscience* 2022(1):1424097 DOI 10.1155/2022/1424097.
- Ni J, Cai Yu, Tang G, Xie Y. 2021. Collaborative filtering recommendation algorithm based on TF-IDF and user characteristics. *Applied Sciences* 11(20):9554 DOI 10.3390/app11209554.
- Park E, Park J, Hu M. 2021. Tourism demand forecasting with online news data mining. *Annals of Tourism Research* 90(4):103273 DOI 10.1016/j.annals.2021.103273.
- Paulson P, Tzanavari A. 2003. Combining collaborative and content-based filtering using conceptual graphs. In: *Modelling with Words: Learning, Fusion, and Reasoning within a Formal Linguistic Representation Framework*. Cham: Springer, 168–185.
- Pla-Sacristán E, González-Díaz I, Martínez-Cortés T, Díaz-de-Maria F. 2019. Finding landmarks within settled areas using hierarchical density-based clustering and meta-data from publicly available images. *Expert Systems with Applications* 123(2011):315–327 DOI 10.1016/j.eswa.2019.01.046.
- Pongchairerks P, Kachitvichyanukul V. 2016. A two-level particle swarm optimization algorithm for open-shop scheduling problem. *International Journal of Computing Science and Mathematics* 7(6):575–585 DOI 10.1504/IJCSM.2016.081693.
- Pu P, Chen Li, Hu R. 2012. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction* 22(4–5):317–355 DOI 10.1007/s11257-011-9115-7.

- Qiao G, Chen H, Li G, Liu H, Wang X. 2024a.** The role of filial piety in filial tourism: an intergenerational analysis of decision-making. *Asia Pacific Journal of Tourism Research* **29**(8):1017–1031 DOI [10.1080/10941665.2024.2362199](https://doi.org/10.1080/10941665.2024.2362199).
- Qiao G, Hou S, Chen Q, Xiang G, Prideaux B. 2024b.** Role of body in travel: wheelchair users' experience from a multi-sensory perspective. *Journal of Travel Research* **5**(11):57 DOI [10.1177/00472875241249391](https://doi.org/10.1177/00472875241249391).
- Raees M, Ullah S. 2019.** DOP: discover objects and paths, a model for automated navigation and selection in virtual environments. *Turkish Journal of Electrical Engineering and Computer Sciences* **27**(4):2784–2797 DOI [10.3906/elk-1807-288](https://doi.org/10.3906/elk-1807-288).
- Rafay A, Suleman M, Alim A. 2020.** Robust review rating prediction model based on machine and deep learning: yelp dataset. In: *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*. Piscataway: IEEE, 8138–8143.
- Santos F, Ana A, Constantino M, Ramiro G, José M. 2019.** Using POI functionality and accessibility levels for delivering personalized tourism recommendations. *Computers, Environment and Urban Systems* **77**(sup1):101173 DOI [10.1016/j.compenvurbsys.2017.08.007](https://doi.org/10.1016/j.compenvurbsys.2017.08.007).
- Singh Y, Sharma S, Sutton R, Hatton D, Khan A. 2018.** A constrained A\* approach towards optimal path planning for an unmanned surface vehicle in a maritime environment containing dynamic obstacles and ocean currents. *Ocean Engineering* **169**(2):187–201 DOI [10.1016/j.oceaneng.2018.09.016](https://doi.org/10.1016/j.oceaneng.2018.09.016).
- Skinderowicz RŁ. 2022.** Improving ant colony optimization efficiency for solving large TSP instances. *Applied Soft Computing* **120**(9):108653 DOI [10.1016/j.asoc.2022.108653](https://doi.org/10.1016/j.asoc.2022.108653).
- Song R, Liu Y, Bucknall R. 2019.** Smoothed A\* algorithm for practical unmanned surface vehicle path planning. *Applied Ocean Research* **83**(6):9–20 DOI [10.1016/j.apor.2018.12.001](https://doi.org/10.1016/j.apor.2018.12.001).
- Sun H, Chen Y, Ma J, Wang Y, Liu X, Wang J. 2022.** Multi-objective optimal travel route recommendation for tourists by improved ant colony optimization algorithm. *Journal of Advanced Transportation* **2022**(1):6386119 DOI [10.1155/2022/6386119](https://doi.org/10.1155/2022/6386119).
- Supakar R, Satvaya P, Chakrabarti P. 2022.** A deep learning based model using RNN-LSTM for the detection of Schizophrenia from EEG data. *Computers in Biology and Medicine* **151**(6):106225 DOI [10.1016/j.compbimed.2022.106225](https://doi.org/10.1016/j.compbimed.2022.106225).
- Wei L-Y, Peng W-C, Lee W-C. 2013.** Exploring pattern-aware travel routes for trajectory search. *ACM Transactions on Intelligent Systems and Technology (TIST)* **4**(3):1–25 DOI [10.1145/2483669.2483681](https://doi.org/10.1145/2483669.2483681).
- Wen Y, Liang Y, Zhu X. 2023.** Sentiment analysis of hotel online reviews using the BERT model and ERNIE model—data from China. *PLOS ONE* **18**(3):e0275382 DOI [10.1371/journal.pone.0275382](https://doi.org/10.1371/journal.pone.0275382).
- Yang Y, Duan Y, Wang X, Huang Zi, Xie N, Shen HT. 2018.** Hierarchical multi-clue modelling for POI popularity prediction with heterogeneous tourist information. *IEEE Transactions on Knowledge and Data Engineering* **31**(4):757–768 DOI [10.1109/TKDE.2018.2842190](https://doi.org/10.1109/TKDE.2018.2842190).
- Yang H, Ren X. 2024.** Design and development of a rural tourism marketing system using deep learning. *IEEE Access* **12**:64795–64806 DOI [10.1109/ACCESS.2024.3396081](https://doi.org/10.1109/ACCESS.2024.3396081).
- Ye M, Xiao R, Lee W-C, Xie X. 2011.** On theme location discovery for travelogue services. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York: ACM, 465–474.
- Yuan H, Xu H, Qian Y, Li Y. 2016.** Make your travel smarter: summarizing urban tourism information from massive blog data. *International Journal of Information Management* **36**(6):1306–1319 DOI [10.1016/j.ijinfomgt.2016.02.009](https://doi.org/10.1016/j.ijinfomgt.2016.02.009).



- Zeng Y, Chen X, Cao X, Qin S, Cavazza M, Xiang Y. 2015.** Optimal route search with the coverage of users' preferences. In: *24th International Joint Conference on Artificial Intelligence*. New York: ACM.
- Zhang Y, Tang Z. 2022.** PSO-weighted random forest for attractive tourism spots recommendation. *Future Generation Computer Systems* **127(1)**:421–425  
DOI [10.1016/j.future.2021.09.029](https://doi.org/10.1016/j.future.2021.09.029).
- Zhang S, Zhang R, Liu X, Sun H. 2012.** A personalized trust-aware model for travelogue discovering. In: *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*. Vol. 3, Piscataway: IEEE, 112–116.
- Zhu X, Sun T, Yuan H, Hu Z, Miao J. 2019.** Exploring group movement pattern through cellular data: a case study of tourists in Hainan. *ISPRS International Journal of Geo-Information* **8(2)**:74  
DOI [10.3390/ijgi8020074](https://doi.org/10.3390/ijgi8020074).
- Zou Z, Yang S, Zhao L. 2024.** Dual-loop control and state prediction analysis of QUAV trajectory tracking based on biological swarm intelligent optimization algorithm. *Scientific Reports* **14(1)**:19091 DOI [10.1038/s41598-024-69911-5](https://doi.org/10.1038/s41598-024-69911-5).
- Zuo C, Zhang X, Yan L, Zhang Z. 2024.** GUGEN: global user graph enhanced network for next POI recommendation. *IEEE Transactions on Mobile Computing* **23(12)**:14975–14986  
DOI [10.1109/TMC.2024.3455107](https://doi.org/10.1109/TMC.2024.3455107).