I.J. Intelligent Systems and Applications, **2024, 4, 55-71**
Published Online on August 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2024.04.04

TourMate: A Personalized Multi-factor Based Tourist Place Recommendation System Using Machine Learning

Azmain Abid Khan
Department of computer science and engineering, Chittagong University of Engineering and Technology, Raouzan, 4349, Chittagong, Bangladesh
E-mail: u1704130@student.cuet.ac.bd
ORCID iD: https://orcid.org/0009-0002-3135-2861

Mahfuzulhoq Chowdhury*
Department of computer science and engineering, Chittagong University of Engineering and Technology, Raouzan, 4349, Chittagong, Bangladesh
E-mail: mahfuz@cuet.ac.bd
ORCID iD: https://orcid.org/0000-0002-3006-4596
*Corresponding author

Received: 23 February 2024; Revised: 15 March 2024; Accepted: 14 May 2024; Published: 08 August 2024

Abstract: Building a personalized travel recommendation system is important to enhance the satisfaction and experience of travelers. Due to the lack of an efficient online-based tourist assistance system, tourists have faced several challenges in Bangladesh, such as difficulties in planning their trips and making informed decisions. To overcome the existing challenges, in this paper, a prediction model has been developed to predict the suitability of a travel destination based on the user's preferences and some other relevant factors. Then the system offers personalized recommendations for the best local places to visit, hotels to stay in, transportation services, and travel agencies with the necessary details. This paper utilizes various machine learning classification algorithms to predict the best-suited travel destinations and local tourist spot recommendations for users based on their budget and preferences. The examined results verified that the random forest algorithm provides the best accuracy of 98 percent and is used for tourist place eligibility prediction. The user rating analysis visualized that the proposed mobile application received satisfactory remarks from more than 60 percent of reviewers regarding its effectiveness.

Index Terms: Automated Recommendation System, Personalized Travel Recommendations, User Preferences, Machine Learning, Mobile Application, Tourist Place Recommendation.

1. Introduction

The Tourism is the activity of visiting places for relaxation, happiness, or work purposes. It involves staying in hotels, enjoying attractions, exploring different lifestyles, and doing activities like shopping, hiking, sightseeing, and swimming [1-6]. Tourism is a crucial contributor to the economic growth and development of many countries, and Bangladesh is no exception. The country is endowed with rich natural beauty and historical sites, such as archaeological sites, mountains, beaches, rivers, hills, forests, waterfalls, tea gardens, and religious places, which attract a huge number of visitors annually. The tourism industry in Bangladesh has a huge impact on its GDP (gross domestic product), with a contribution of 4.4 percent in 2018–2019 and 4.2 percent in 2016–2017, according to the Bangladesh Tourism Board [1-3]. This industry has created many jobs opportunities, increased economic revenue, and improved infrastructure in the country. Bangladesh is a developing country. Interestingly, tourism can bring many environmental, economic and social benefits. It can earn a large amount of money from international tourists.

There are many opportunities to boost the tourism industry in Bangladesh [2]. It can be noted that people are interested in traveling and exploring new places. But they have no clear concept of the travel destination, time, duration, etc. Sometimes they have no idea that there are places nearby. So, they can't get the scope of travel to these places. Budget is also a factor in traveling. People may like budget friendly or luxury travel [3]. So, budget-friendly travelers hesitate to travel to long-distance places, thinking it may cross their budget. Besides, all destinations are not for family or friends. Some are best for camping and sightseeing. Some are familiar with historical values and natural beauty. So, a person's personal choice becomes a factor in traveling to a place [4]. Various trip management apps and websites, such as Trip Adviser [5] and Gozayaan [6] have many facilities in Bangladesh. People can use this facility to travel to various places. But here, suggesting the best destination according to budget, time, and interests is absent. So, travelers can't get any suggestions about a destination according to their budget and time. They also don't know what activities they can do there or what foods and restaurants they will try. The existing works did not offer any analysis for the best tourist place suitability recommendation system in Bangladesh using machine learning algorithms [7-12]. Moreover, a mobile application for determining the best local places in Bangladesh based on personal preferences are also missing. Currently, several works on tourist place recommendations exist in the literature [13-22]. To show the shortest route to the recommended tourist places, the work in [18] utilized the Dijkstra algorithm. The authors in [19] presented a deep transfer learning-based tourist spot recommendation system for Dubai by evaluating the age group and location from the image. The work in [20] developed a web application for suitable place recommendations for travelers based on their background data. The authors in [21] presented a fuzzy logic-based tourist route recommendation scheme based on the tourist spot rating by reviewers and the user's economic condition. To identify the tourist hotspot in Bangladesh along with their demographic information, the authors in [23] created a dashboard visualization technique by using a map and bar chart for future tourist assistance. The work in [24] presented a religious tourism recommender system by taking into account different religious destinations in Nepal, tourist products, security concerns, and services. However, the existing works literature [21-27] did not present any suitable local place recommendation system using machine learning for Bangladeshi travel places by taking destination, budget, and user preference into account.

Tourism is significantly related to any country's economic development [8]. Nowadays, mobile applications have become an important tool in tourism with technological advancements. Travelers can experience a unique experience and enjoy their trip using a personalized and effective travel recommendation system. But building a suitable travel recommendation system is not easy because it takes into account customer preferences and local tourist spots. It has many challenges, like collecting the appropriate dataset and selecting the best model. To address these challenges, this paper proposes a personalized travel recommendation system by using a mobile application for the assistance of Bangladeshi tourists. It will help travelers introduce the best tourist place recommendation system in Bangladesh. This application will provide the opportunity to reduce the wastage of time in selecting a destination within budget and time. The proposed system's primary goal is to help travelers by providing them with personalized recommendations based on their interests, preferences, and budgets. In this work, we have also used several machine learning classification algorithms to predict whether a place will be suitable or not for travelers in Bangladesh. This system can help travelers plan their trips before traveling based on their budgets and preferences. It would be helpful for travelers to make the correct decisions about their tour plans by providing personalized recommendations for attractions, destinations, and accommodations.

The notable contributions of this paper are noted as follows:

- In this paper, real tourism-related data from different sources was collected to create a dataset, and several algorithms were applied to the dataset for destination and activity recommendation systems.
- This paper selects the best machine learning models for tourist place eligibility prediction.
- This paper also develops a mobile application based on a machine learning model that can predict whether a tourist place is suitable or not for travelers based on user preferences. This paper also recommends the best local places based on personal preferences.
- This paper also presents survey results regarding the proposed application's performance by collecting user review comments.

In the next section, this paper highlights notable literary works and their major contributions. Section 3 gives the proposed machine-learning-based tourist place recommendation system. Sections 4 and 5 deliver the proposed mobile application and user rating analysis, respectively. Section 6 delivers the summary and future challenges.

2. Related Works

This section will deliver the related works concerning tourist place recommendation system. In the work of [9], a smart tourist guidance system has been developed in Bangladesh using machine learning. The Google Maps API is used to find out the location of any tourist place. Then this model finds the best routes (i.e., to reach the destinations). But this model depends on the user's GPS location, and the accuracy of this model is not shown. So, we can't understand how good the model is. In the work of [10], a machine learning and deep learning based tourist place recommendation and recognition system has been developed. In their work, the user's interest is taken as input by the mobile app and then recommends attractions, restaurants, and hotels using K-mode clustering. It also helps to recognize new places using a convolution neural network. But the accuracy of this model is below 50 percent. In the work of [11], a machine learning based personalized hybrid tourism recommender system has been developed. But they recommend to the user the best attractions in a particular place. In the work of [12], a tourist place recommendation system has been developed using machine learning. In their work, the new places are recommended to the user based on the user's public instagram photos. So, this model will not be useful when tourists want to know about the budget. In the work of [13], a recommendation system has been developed based on tourist or user attractions. They first collect the visiting history of the user's neighbors. Then the recommendations for attractions are constructed using the cosine method. But the model has not shown any accuracy, and there is no Android app to use it. The work in [22] provides a customized tourist route and place recommendation system based on users' personalized choice options and time constraint values using the Flutter application and tensor flow software.

In the work of [14], a machine learning model has been used to predict tourist's responses about a tourist place. It records responses from tourists to understand positive or negative judgments about a certain attraction. The main limitation of their paper says they didn't show the accuracy of this model. So, we can't understand how good the model was. In the work of [15], a cruisebased travel recommendation system is developed for users. It uses data mining, and the authors point out the limitations of the single GPS. They optimize the designed model based on the recommended patterns. In the work of [16], they designed a travel recommender system that provides tourist spots names based on the user's ratings and attractions. In their work, they used data from only one travel agency. The authors in [25] developed a tourist destination recommendation system for the citizens of Sri Lanka based on their past travel experiences and personal factors (e.g., travel mode, income level, hobbies, education status, gender, and age). The work in [26] provided a sentiment analysis scheme regarding travel destinations by collecting people's feedback. The article in [27] presented a collaborative filtering system for personalized recommendations based on user's previous ratings regarding a travel place. In the work of [17], a travel recommendation system-based system has been developed based on social media activity, specifically analyzing users' Twitter data. The work in [28] utilized a simulated annealing algorithm to recommend tourist routes. The article in [29] developed a virtual reality app for tourists in Indonesia for tourist assistance. Differing from recent and existing literature, this paper develops a personalized multi-factor-based tourist place recommendation system using machine learning.

3. Proposed Approach

The overall process is shown in detail in Fig. 1. This figure represents the system model of the proposed best tourist place suitability check and local place recommendation system. The primary objective of the system model is to provide travelers with accurate recommendations on whether a particular place is suitable for them or not based on users' answers in the proposed mobile app (Fig. 2). Fig. 3 depicts the best local place recommendation in the mobile app for the users based on their preferences (e.g., budgets, district, and specific nature of the user).

3.1. Overview of System Framework

The suggested system for best tourist place suitability recommendation and determining the best local places based on personal preferences will work mainly three steps: (i) Build an ML model on the tour-related dataset, (ii) Create an Android app, and (iii) Deploy an ML model on Android. For recommending the suitability of tourist places, we have built an ML model using several classification algorithms (KNN, random forest, decision tree). Random forest algorithm generates a classification system with multiple decision trees, each with higher predictive accuracy than the others. Random Forest is an ensemble learning technique. The decision tree algorithm falls into the category of supervised learning. The algorithm utilizes leaf nodes for final decision internal node for features, and branches for decision rules. The KNN algorithm selects k neighbors, calculates their Euclidean distance, sorts them in ascending order, and then counts the number of data points in each category. Logistic Regression is another remarkable ML technique that can model data with a binary dependent variable. Fig.4 represents the ML model framework. First, we collected data from travelers and travel agencies and then created a dataset. The features of this dataset were: sea lover, mountain lover, history lover, entertainment lover, need hotel, hotel type, need transport, days, place, budget, travel guide, prefer attractions, travel partner, prefer safety, foodie, tourist-friendly place, starting point. The data was split into training and testing sets, with 70% of the data used for training and 30% for testing. We have used machine learning algorithms like the Logistic Regression Algorithm, the KNearest Neighbor (KNN) Algorithm, Random Forest Algorithm, Decision Tree Classifier Algorithm, and Naive Bayes to build our models. After that, we have training data to train, and we also have testing data to test the model. We have chosen the best model to recommend the suitability of tourist places based on accuracy after these models have been trained with training data and tested with testing data. In our work, the model using the Random Forest algorithm was the best. Finally, we can recommend the suitability of tourist places using this model. For android mobile application, we initially required an Android studio to work with Android programming. We are aware that the UI (User Interface) is constantly produced in an XML file. After designing our UI, we wrote backend logic to accept data from the front end in a Java file. Our Android app name is TourMate. Firebase serves as a backend system that connects mobile apps and web applications to cloud storage and APIs. Fig. 5 represents the process flow diagram of our app. The user should provide valid information, such as an email and password, for login. Then, they can see the home page with some tourist place recommendation features. For deploying the ML model on Android, first we saved our machine-learning model as an a.pkl file. Then we have the flask API using Python in PyCharm. Fig. 6 represents the deployment of the ML model on Android. The form will be filled out by the user, who will then submit it. The form will then receive a POST request. Additionally, when a user submits a post request, the Flask API will accept their data and send it to the machine learning model so that it can predict the output class. We will send the predicted class as JSON to the Android app. We have used Postman to verify the API of our TourMate App. An automated and interactive tool called

Postman is used to check the project's APIs. One library called Volley is required to hit API. The logic we have created is that we will use the Android app's inputs to call the API, and then the API's answer will be displayed back in the Android app. So, there was one issue that the Android app was unable to detect because of the API we developed was running locally on our system. Therefore, we had to launch our API online, and the Railway cloud server was used to do so.

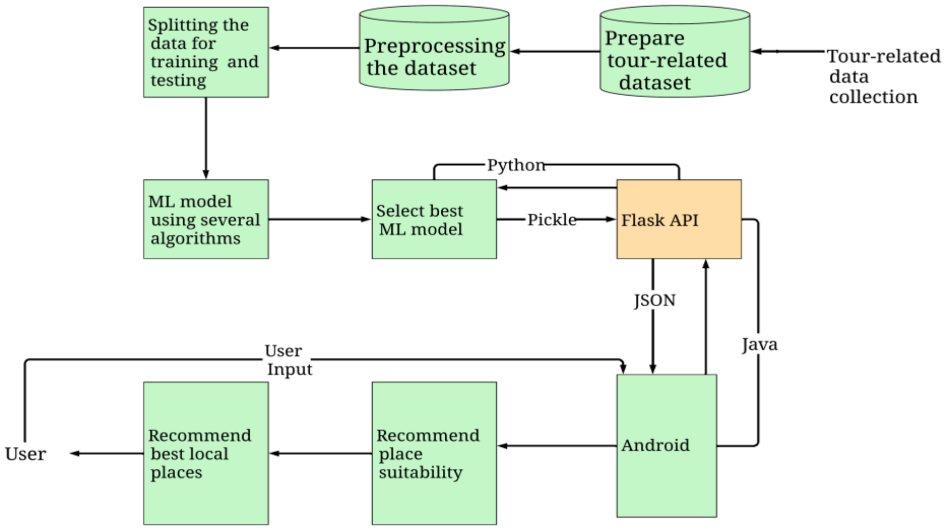


Fig.1. Overview of the tourist place recommendation system in mobile application

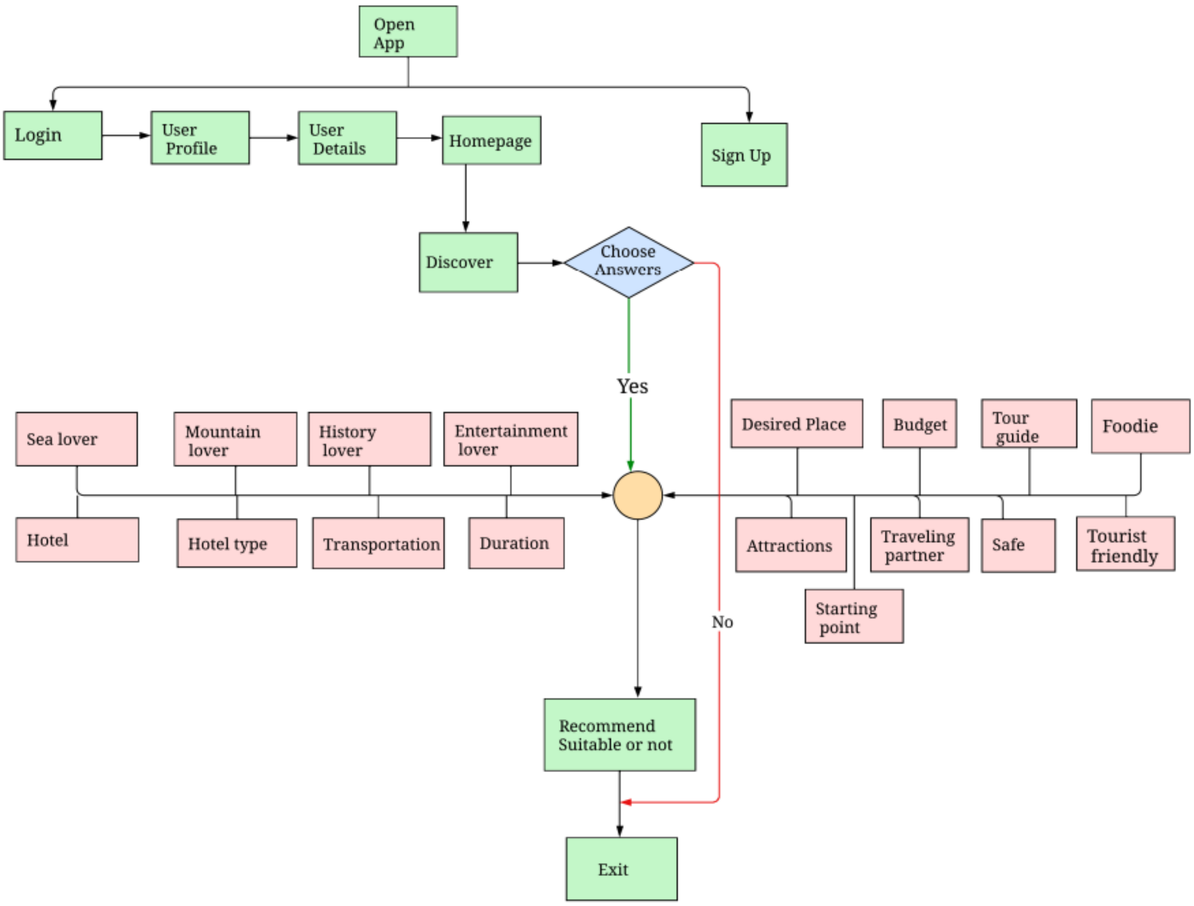


Fig.2. System Model of our multi-factor based specific tourist place suitability check system for tourists

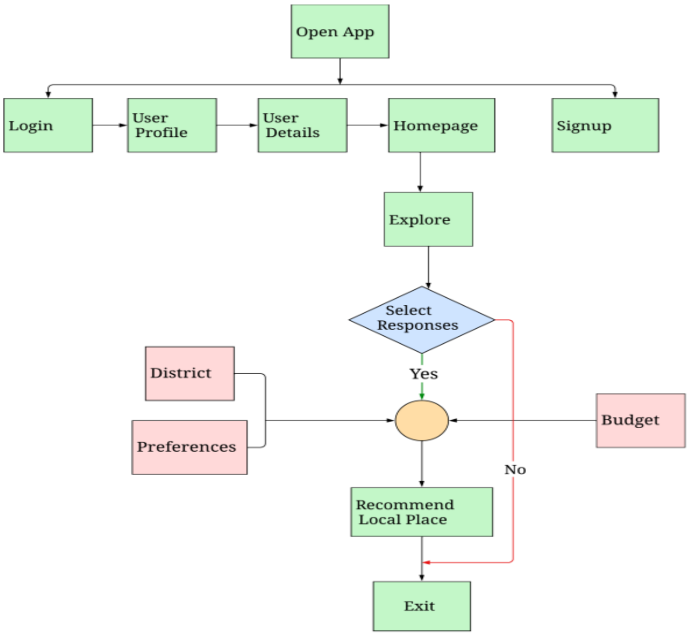


Fig.3. System Model of our multi-factor based personalized tourist place recommendation system (local place recommendation)

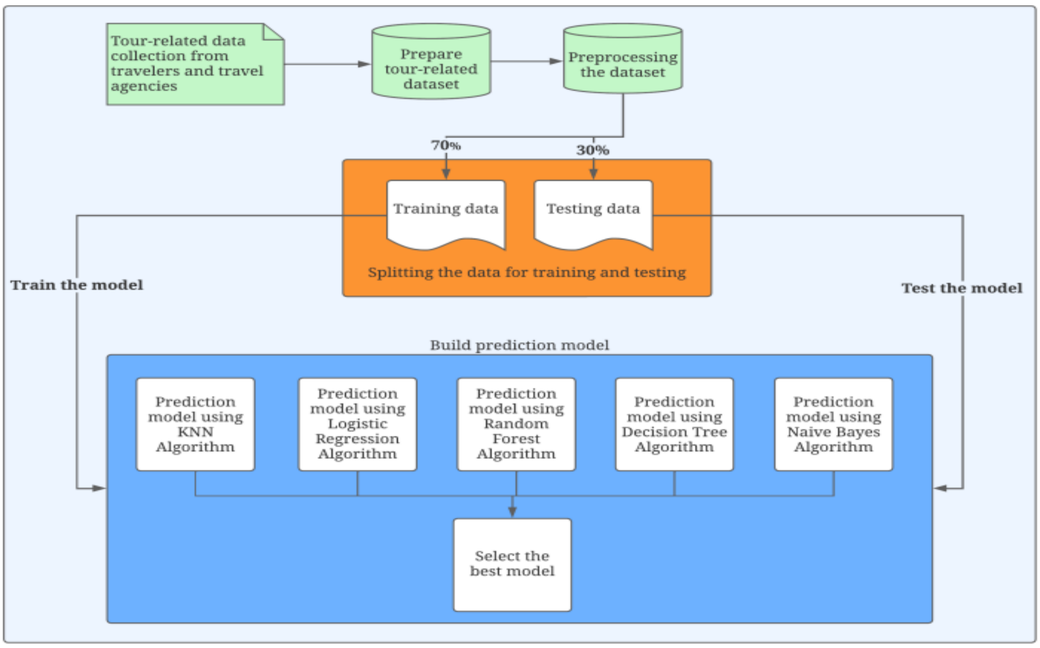


Fig.4. Proposed ML model framework

3.2. *Steps for the best Tourist Place Suitability Prediction Model*

This subsection will briefly cover the method of developing the best tourist place Suitability prediction model. Fig.4 represents the overall steps of the prediction model. Firstly, we have collected information on tour-related data from several travelers and travel agencies. I have tried to gather the dataset since the very beginning. After collecting tour-related data from travelers and travel agencies, we created a tour-related dataset. Fig.7 represents the tail data of the tourrelated dataset. This information is a very confidential issue for travel agencies, which is why we cannot collect a large number of data points. 868 instances and 17 features are present in this dataset. The figure shows that in this dataset on travelers, there was some categorical information. The categorical data cannot be understood by a machine. To create numerical data, we convert categorical data using label encoding. Then we determine whether a value is missing or not. Fortunately, our dataset did not contain any missing values. Therefore, handling the missing value wasn't necessary in our dataset. Next, feature selection is a technique that involves reducing the input variables for a model by selecting only relevant data while eliminating irrelevant data, and numerous features are relevant to predicting the suitability of tourist destinations. To determine the crucial features in our dataset, we used the feature-importance technique. We have imported Extra Trees Classifier, an ensemble method to apply this function. Next, in the model training phase, the dataset is split into 70% for the training phase and 30% dataset for the testing phase, and various ML classification algorithms such as logistic regression, KNN, or knearest neighbor, random forest, and dt or decision tree classifier, are used to train the model, and their accuracy is evaluated to choose the most effective classifier for the dataset of travelers. After that, we have selected the best parameters for these algorithms: Random Forest, KNN, and Logistic Regression Algorithm. Next step is the choosing a model. The most effective model has been chosen based on the evaluation matrices. After training the classification algorithms, we used the grid search method to fine-tune their hyper parameters and optimize their performance for the dataset of travelers. In our model, the random forest algorithm gives us the best result compared with the other algorithms.

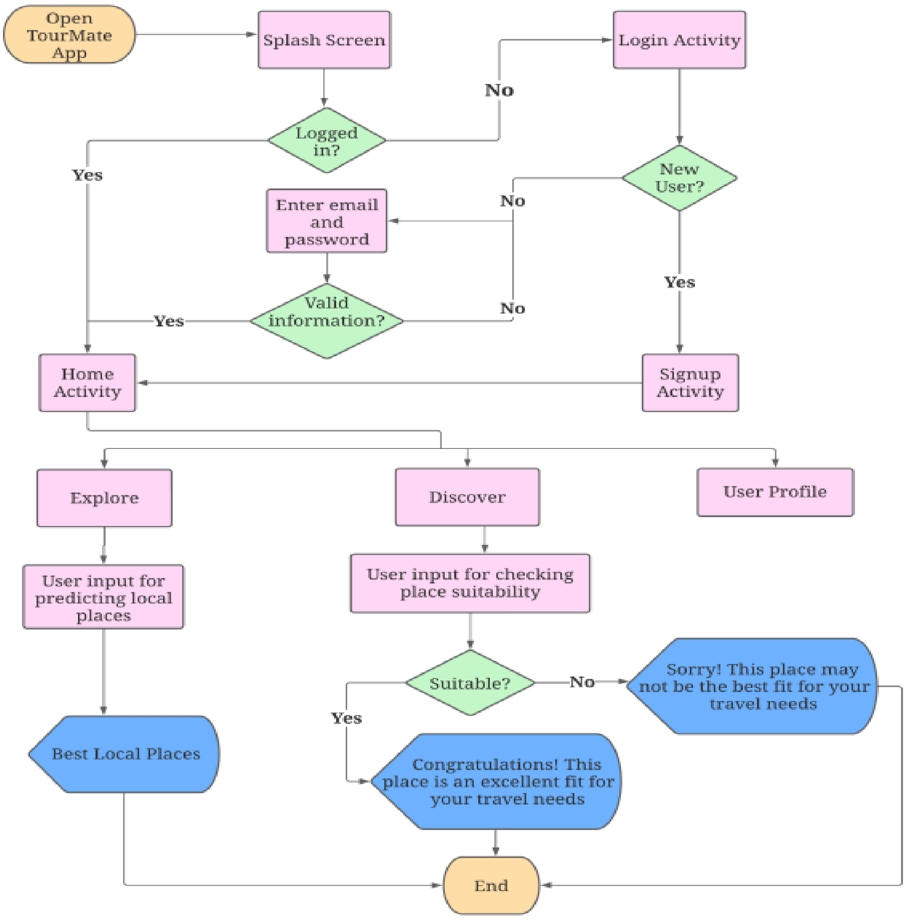


Fig.5. Process flow of android App

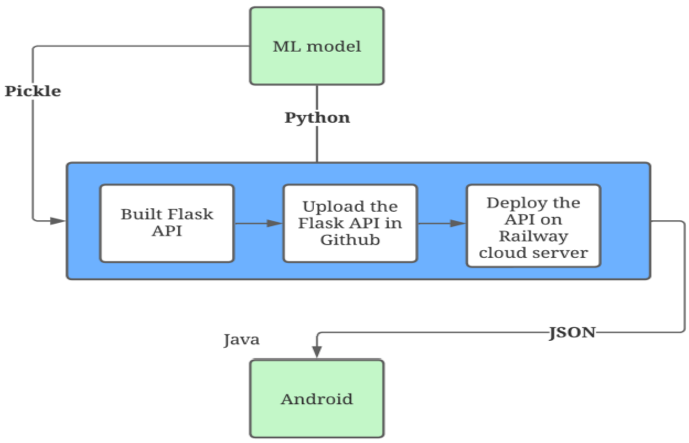


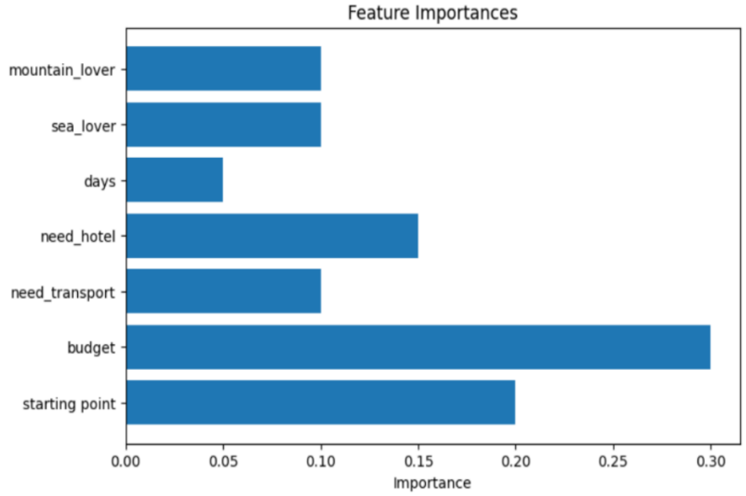
Fig.6. Deploy ML model on android

	D	E	F	G	H	I
1	Is that place right for a m	Is that place right for a hi	Available Entertainment F	Hotel needed	Transport needed	Days
16	1	0	1	0	1	1
17	1	0	1	1	1	1
18	1	0	1	1	1	2
19	1	0	1	1	1	2
20	1	0	1	1	1	2
21	1	0	1	1	1	2
22	1	0	1	1	1	2
23	1	0	1	1	1	1
24	1	0	1	1	1	2
25	1	0	1	1	1	1
26	1	0	1	1	1	2
27	1	0	1	1	1	2
28	1	0	1	1	1	2

Fig.7. A Glimpse of user dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 868 entries, 0 to 867
Data columns (total 28 columns):
 #   Column                                                  Non-Null Count  Dtype
---  ------                                                  --------------  -----
 0   Timestamp                                               868 non-null    object
 1   Email                                                   868 non-null    object
 2   Is that place right for a sea lover?                    868 non-null    int64
 3   Is that place right for a man who loves mountains?      868 non-null    int64
 4   Is that place right for a history lover?                868 non-null    int64
 5   Available Entertainment Facilities                      868 non-null    int64
 6   Hotel needed                                            868 non-null    int64
 7   Transport needed                                        868 non-null    int64
 8   Days                                                    868 non-null    int64
 9   Suitable trip                                           868 non-null    int64
 10  Place                                                   868 non-null    int64
 11  Budget                                                  868 non-null    int64
 12  Travel guide                                            868 non-null    int64
 13  Hotel type                                              868 non-null    int64
 14  Is the place full of natural and man-made attractions?  868 non-null    int64
 15  Traveling partner                                       868 non-null    int64
 16  Is the place safe for a tourist?                        868 non-null    int64
 17  Transport type                                          868 non-null    int64
 18  Hotel cost                                              868 non-null    int64
 19  Are there enough food stalls to enjoy various types of foods?  868 non-null    int64
 20  Are the local people friendly with tourists?            868 non-null    int64
 21  Food cost                                               868 non-null    int64
 22  Transport_cost                                          868 non-null    int64
 23  famous Places                                           867 non-null    object
 24  Foods                                                   868 non-null    object
 25  Restaurants                                             868 non-null    object
 26  Your starting point                                     868 non-null    int64
 27  Ratings                                                 868 non-null    int64
dtypes: int64(23), object(5)
memory usage: 190.0+ KB
```

a

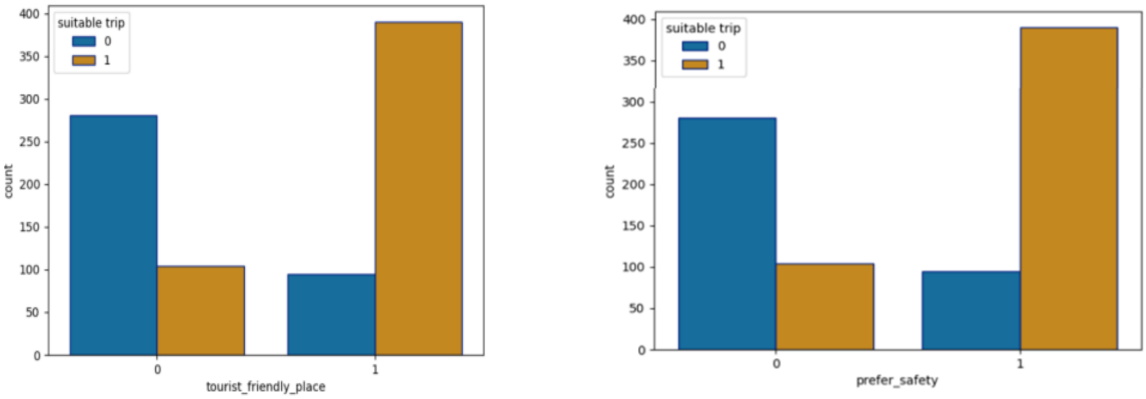


b

Fig.8. a. Features of dataset, b. Comparison of features in order of importance

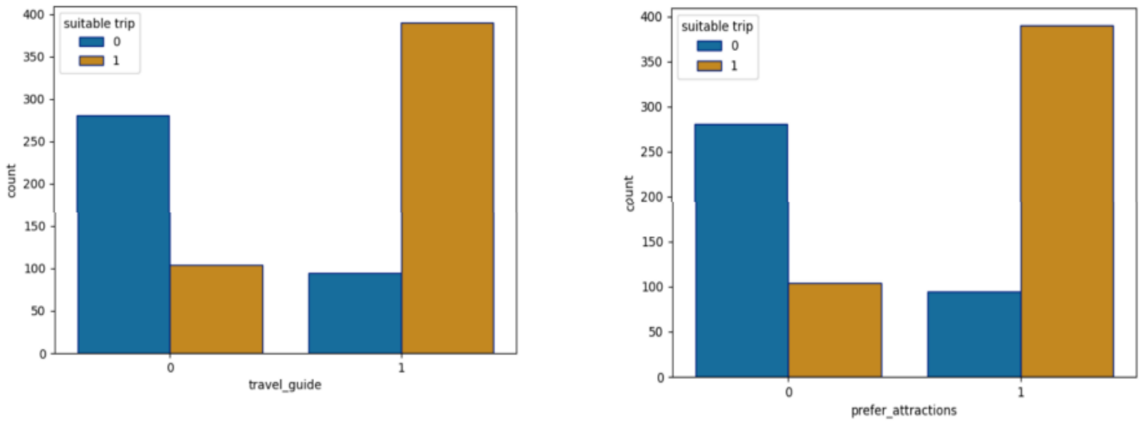
3.3. *Dataset Collection, Data Visualization, and Best Parameter Selection*

In this paper, we have created a dataset by collecting answers from Google Forms. We asked travelers and travel agencies to provide us with tour-related information. In total, we got a response of 868 data points, which is the size of our dataset. Fig.8 (a) shows the features of the dataset. Here we can see five object-type data points. Some of the sample data from the dataset is shown in Fig.7. An analysis of the importance of features in the dataset is also done, as shown in Fig.8(b). Fig.9 represents the possibility of recommending the suitability of a place against tourist-friendly places and safety. It is found from the count vs. tourist friendly place graph and from the count vs. prefer safety, which denotes the suitability status against tourist-friendly places, that if the traveler prefers a friendly place and safety, then the possibility of recommending the suitability of a place is much higher than if they prefer a non-friendly place and non-safety. Fig.10 represents the possibility of recommending the suitability of a place against a travel guide and preferring attractions. Fig.11 represents the possibility of recommending the suitability of a place against the transport needs of a hotel. For the best accuracy of the model using the K-Nearest Neighbor Algorithm, we have selected the optimal number of neighbours. In our model, we have found the best result for n neighbour=4. The optimal number of neighbors was 4. Fig.12 shows that when the number of neighbors is 4, then the accuracy result is best. For the best accuracy of the model using logistic regression, we have selected the optimal number of C. In this model, we have found the highest test roc auc is 0.936754367812 at C = 1. The optimal value of C was 1. Fig.13 shows the AUC ROC curve for C = 1. Fig.14 represents the best parameter for the Random Forest Algorithm in our model.



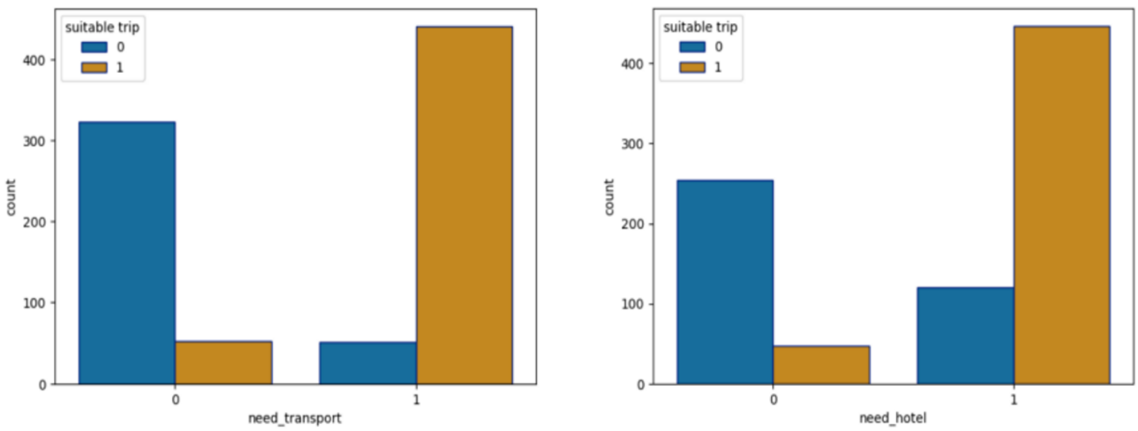
a b

Fig.9. Suitable trip data visualization based on a. tourist friendly place, b. safety criteria



a b

Fig.10. Suitable trip data visualization based on customer prefer a. travel guide, b. attractions



a b

Fig.11. Suitable trip data visualization based on a. transport need, b. hotel need preference

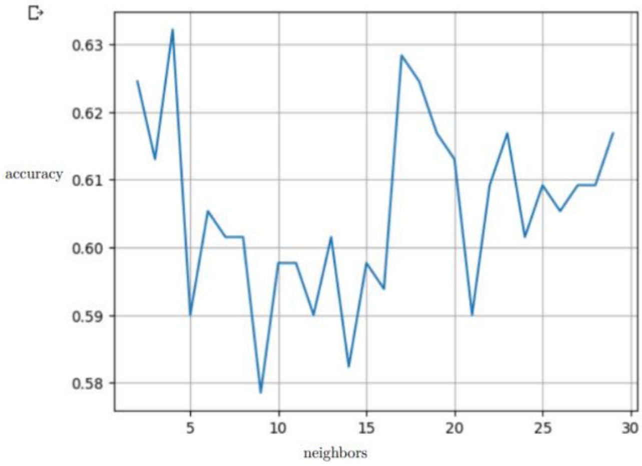


Fig.12. Model accuracy comparison against neighbors

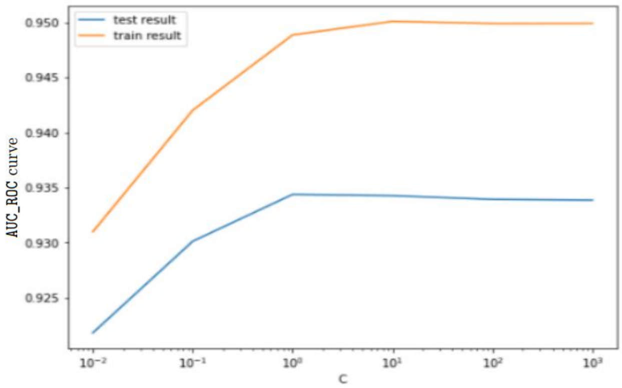


Fig.13. AUC-ROC curve for logistic regression

```
[49] rf_grid_model.best_params_

     {'bootstrap': True,
      'max_features': 'auto',
      'min_samples_leaf': 1,
      'min_samples_split': 5,
      'n_estimators': 30}

[50]
     rf_grid_model.best_estimator_

     RandomForestClassifier
     RandomForestClassifier(max_features='auto', min_samples_split=5,
                            n_estimators=30)
```

Fig.14. Best parameter for random forest algorithm

3.4. Performance Evaluation Using Model Accuracy and Confusion Matrix

Fig.15 (a) represents the confusion matrix of the K-Nearest Neighbor Algorithm. The prediction model using this algorithm can predict the 100 places that have been correctly identified as suitable for travelers, and the 66 places that have been appropriately categorized but are unsuitable for travelers. This model predicted 45 places that were mistakenly labeled as suitable when they were not, and also predicted 50 places that were mistakenly excluded from place suitability. Fig. 15(b) also represents the confusion matrix of the Logistic Regression Algorithm. The prediction model using this algorithm can predict 140 places which have been correctly identified as suitable for travelers, and the 98 places which have been appropriately categorized but are unsuitable for travelers. This model predicted 13 places that were mistakenly labeled as suitable when they were not and also predicted 10 places that were mistakenly excluded from suitability. Fig.15(c) represents the confusion matrix of the so-called random forest algorithm. The prediction model using this algorithm can predict 149 places which have been correctly identified as suitable for travelers, and the 108 places which have been appropriately categorized but are unsuitable for travelers. This model predicted 3 places that were mistakenly labeled as suitable when they were not and also predicted one place, which was mistakenly excluded from place suitability.

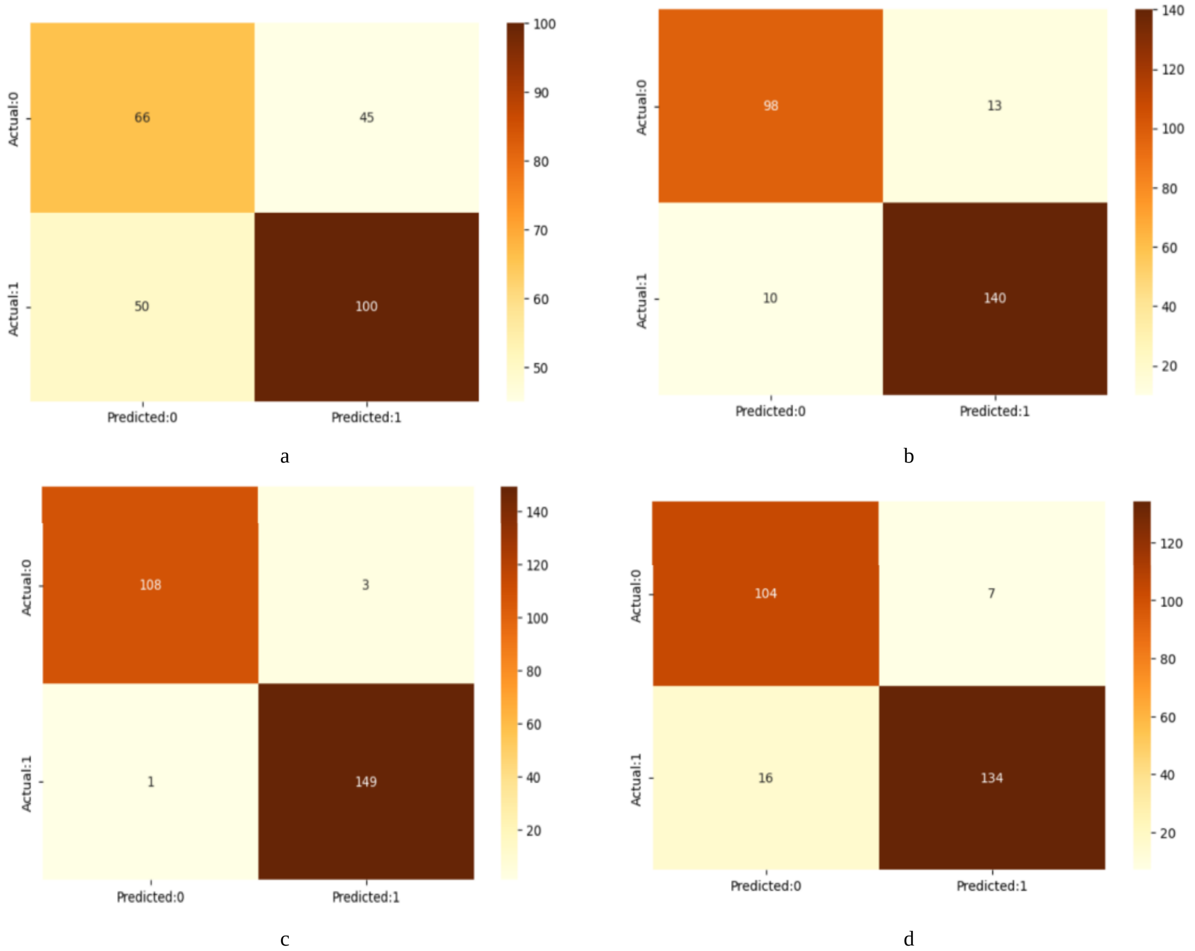


Fig.15. a. Confusion matrix of K-nearest neighbor algorithm, b. Confusion matrix of logistic regression algorithm, c. Confusion matrix of random forest algorithm, d. Confusion matrix of naive bayes algorithm

	Predicted Not Eligible	Predicted Eligible
Actual Not Eligible	True Negative	False Positive
Actual Eligible	False Negative	True Positive

Fig.16. Confusion matrix definition in four axes

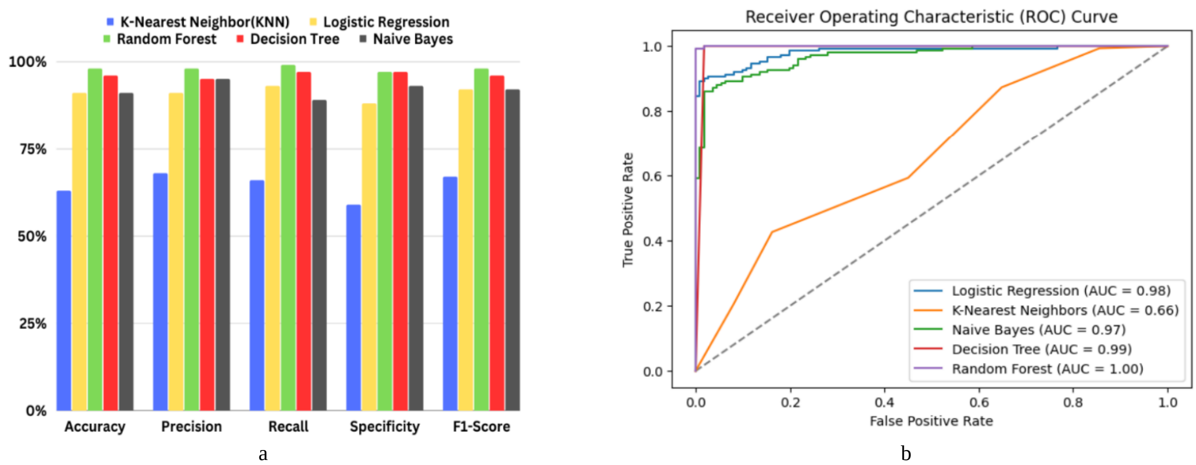


Fig.17. a. Accuracy, Recall, Precision, Specificity, and F1-score Performance Comparison, b. ROC-AUC curve

Fig.15(d) represents the confusion matrix of the so-called naive Bayes algorithm. The prediction model using this algorithm can predict the 134 places that have been correctly identified as suitable for travelers and the 104 places that have been appropriately categorized but are unsuitable for travelers. This model predicted 7 places that were mistakenly labeled as suitable when they were not, and it also predicted 15 places that were mistakenly excluded from place suitability. Fig.16 shows the four axes of the confusion matrix. Here, TruePos (true positive) indicates the number of tourists who have been correctly identified as a suitable place to visit. TrueNeg (true negative) is a measure of the number of places that have been appropriately categorized but are not suitable for tourists to visit. FalsePos is a measure of the number of places that were mistakenly labeled as suitable for tourists to visit when, in fact, they were not. FalseNeg indicates the number of places that were mistakenly excluded as suitable for tourists to visit.

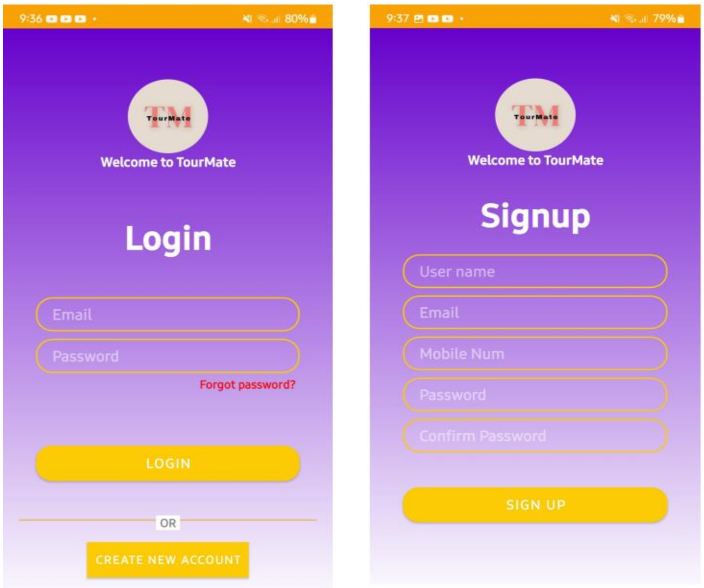
Fig.17(a) represents the performance comparison of all five algorithms using accuracy, precision, recall, specificity, and F1- score. The model accuracy for six different classification algorithms (KNN Classifier, the logistic regression, the random forest, the Decision Tree, the linear regression, and the naive bayes) are 63%, 91%, 98%, 96%, 74%, and 91%, respectively, and the Random Forest algorithm has the highest accuracy. So, finally, the random forest classifier is selected as the predictive model of the proposed system, giving the best accuracy (98%) among the six of these algorithms. The precision values for KNN, the logistic regression, the random forest, the DT, or the decision tree, and naive bayes algorithms are 68%, 91%, 98%, 95%, and 95%, respectively. Among these algorithms, the Random Forest model achieves the highest precision. Recall measures the proportion of actual positive observations that were correctly predicted, out of all positive observations. The recall values for KNN, the logistic regression, the random forest, the dt or the decision tree, and naive bayes algorithms are respectively 66%, 93%, 99%, 97%, and 89%. It can be observed that the Random Forest algorithm has the highest recall among these five algorithms. The specificity for the KNN, or K-Nearest Neighbor, the Logistic Regression, the Random Forest, the Decision Tree, and the Naive Bayes algorithms are respectively 59%, 88%, 97%, 97%, and 93%. It is seen that the Random Forest algorithm gives the best specificity among the four of these algorithms. The F1-score for the KNN, or K-Nearest Neighbour, Logistic Regression, Random Forest, Decision Tree, and Naive Bayes algorithms are respectively 67%, 92%, 98%, 96%, and 92%. It is observed that the Random Forest algorithm outperforms undeniably the other examined algorithms in terms of the F1 score. We may conclude that the model using the Random Forest method provides the best prediction result after reviewing the overall performance. To predict the suitability of the desired place for travelers, we have selected the model with the Random Forest algorithm. Accuracy, recall, precision, and F1-score calculation equation is given below:

$$Accuracy = \frac{TruePos+TrueNeg}{TruePos+TrueNeg+FalsePos+FalseNeg} \tag{1}$$

$$Recall = \frac{TruePos}{TruePos+FalseNeg} \tag{2}$$

$$Precision = \frac{TruePos}{TruePos+FalsePos} \tag{3}$$

$$F1-score = 2 * \frac{Precision*Recall}{Precision+Recall} \tag{4}$$



a b

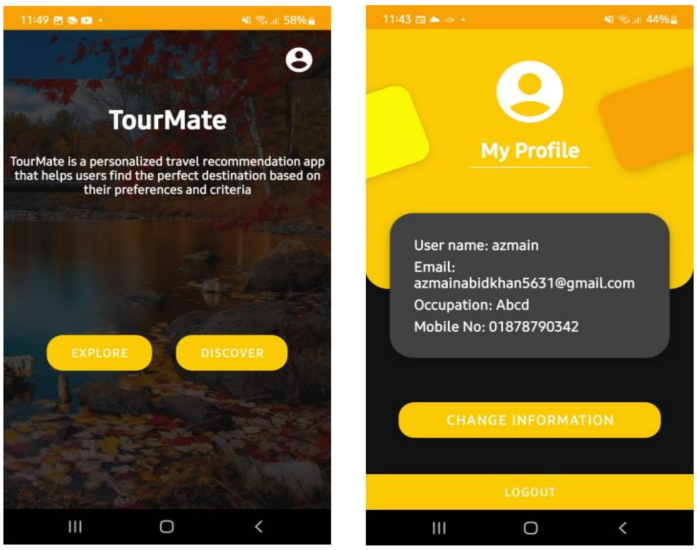
Fig.18. Login and signup page of our app

A common tool for evaluating binary classification models is the Receiver Operating Characteristic (ROC) curve. The ROC plots the TPR (true positive rate) against the FPR (false positive rate) at various threshold values. It can distinguish between ''signal" and ''noise." The classifier's ability to accurately distinguish between classes is measured by the Area Under the Curve (AUC), which summarizes the ROC curve. A higher X-axis value on the ROC curve indicates a higher proportion of false positives compared to true negatives, while a higher Yaxis value indicates a higher proportion of true positives compared to false negatives. Fig.17 (b) represents the ROC-AUC curve of five algorithms such as KNearest Neighbor, Random Forest, Logistic Regression, Decision Tree, and Naive Bayes algorithms, which are 66%, 100%, 98%, 99%, and 97%, respectively. The AUC provides a measure of the model's ability to distinguish between positive and negative classes, with a higher AUC indicating better performance. Among the five algorithms, the Random Forest algorithm achieved the best AUC score. It can be noted that, a perfect model has an AUC score of 1.

4. Mobile Application Development

This section will discuss the proposed tourist assistance mobile application features. Fig.18 represents the user signup and login interface of our proposed tourist assistance mobile application (named as Tourmate). The traveler can enter the home section immediately if he or she has previously signed in to the app. If not, he or she must enter the login phase. This app requires an email address and password for access. The home task interface is shown if the email is verified. If the traveler is new to this Tour Mate app, he or she must first register. A username, email address, mobile number, password for your occupation, and a confirmed password are needed to register for this app. The home activity interface is shown if the password matches the confirmed password. Fig.19(a) represents the home page of the Tour Mate mobile app. The user profile is in the upper-right corner of the home page. In the middle section, there is a Discover module where travelers can check the suitability of their desired place. Travelers can access their profile by clicking on the user profile section. There is also an Explore section where users will answer some questions to see the best local places.

Fig.19 (b) shows the user profile. It can be accessed by clicking the profile logo on the home page (Fig.19(a)). After choosing a user profile, a traveler can view and edit the information on their profile. Travelers can see their personnel information in the User Profile interface, and they can change their information if required. They can also log out of the system if they want. The interface for recommending the suitability of the desired place prediction is shown in Fig. 20. It can be accessed by clicking the discover box in Figure 19(a). With the necessary inputs, travelers can determine their place's suitability. They must enter the necessary information, including the required inputs, such as as whether they are sea lovers, mountain lovers, history lovers, entertainment lovers, or not. Besides, do they need any hotels? type of hotel they need; do they need any transport service or not must be included. Finally, the duration of the trip, desired place, budget, whether they require a travel guide or not, whether they prefer attractions, whether they have a travel partner or not, whether they prefer safety, whether they are foodies or not, whether they prefer a tourist-friendly place or not, their starting point must also be given. Here, this paper uses a machine learning model that determines or predicts the suitability of the respective tourist places based on user preference. Fig.21 shows the resultant output of the place suitability prediction model (from Fig.20). If the user's desired place is suitable, then it will show a message:" Best wishes! This place is an excellent fit for your travel needs". Then the user can click on the" Try Again" option for further checking or the" Go Home" option for checking other features. Otherwise, it will show that:" Sorry! This place may not be the best fit for your travel needs." After that, the user can click on the" Try Again" option. Fig. 22(a) shows the recommendations of local tourist places based on users' selected destination (city name), budget range, and users' preferences (e.g., sea lover, nature lover).



a b

Fig.19. Home page and profile page of our app

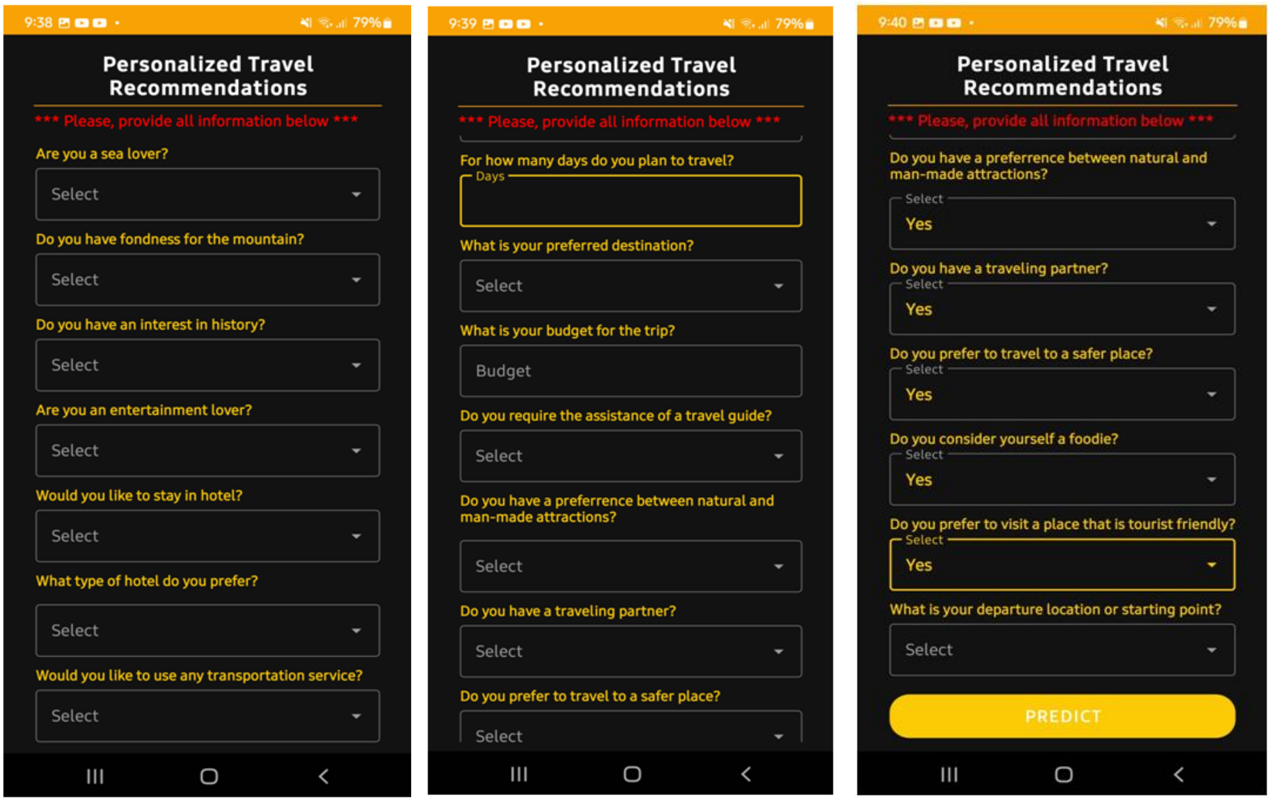
Fig.22 (a) can be accessed by clicking the explore button in Fig.19 (a). Fig. 22(b) shows some local places' recommendation results based on users' input in Fig.22 (a). Then, by selecting the respective local places' names from Fig. 22(b), the user can be forwarded to Fig. 23(a). Fig. 23(a) gives some necessary travel related information (e.g., hotel name, travel agency, and transport services information for booking) about a tourist place.

4.1. Comparison with Existing Works

The existing system in [17] used a machine learning classifier to identify travel-related tweets and incorporated time-sensitive emergency weights to provide personalized recommendations based on the user's most recent interests. Their proposed model has shown an overall accuracy of 75 percent. But in our proposed system, various algorithms have been tested, and a comparison has been shown, such as the confusion matrix and the Roc-Auc curve, to get better ideas of these algorithms. We have noticed that the random forest algorithm provides the best accuracy of 98 percent for suitable place prediction. Different from the previous works, our proposed system focused on various criteria to evaluate the suitability of travel destinations, like preferences for mountain or sea environments, desired duration, interest in entertainment or history, budget constraints, accommodation needs, safety concerns, and desire for a tour guide, etc. Our proposed model can help travelers improve their overall travel experience and save effort and time. Differ from existing works, the proposed work creates a dataset and gathers traveling information from travelers and travel agencies. The proposed application differs from existing applications (like Trip Advisor or Google Trips) in several ways. First, they did not provide multi-factor-based tourist place eligibility prediction using machine learning algorithms. They did not develop a mobile application that provides machine learning-based local place recommendation for travelers by examining cost, safety, budget, transport, and the user's specific interests (like mountains or nature), among others.

4.2. Challenges during Implementation

One of the major challenges that we faced during implementation was dataset collection and criteria selection for tourist place recommendations. But we have overcome these challenges by contacting real tourists and tourist agencies. By contacting them, we have collected a huge amount of data, which is essential for our machine-learning-based tourist place eligibility prediction system.

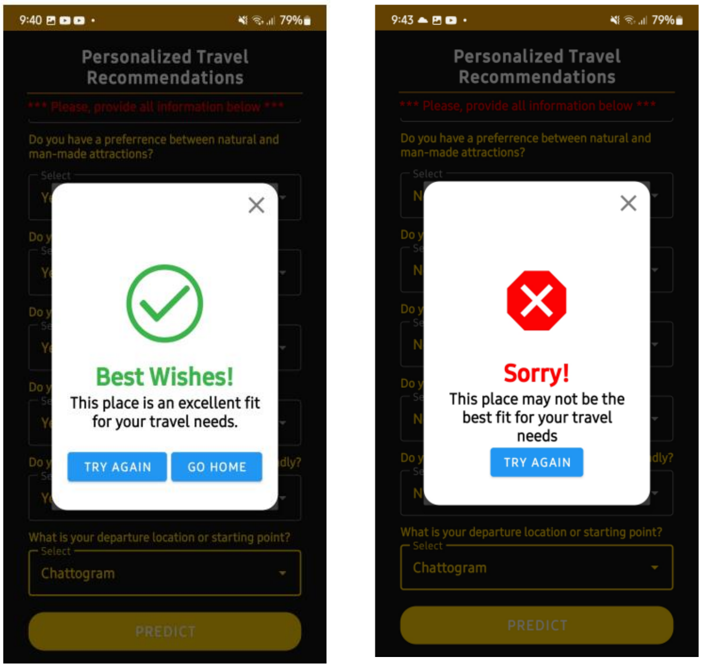


a b c

Fig.20. Gathering information from travelers for checking place suitability

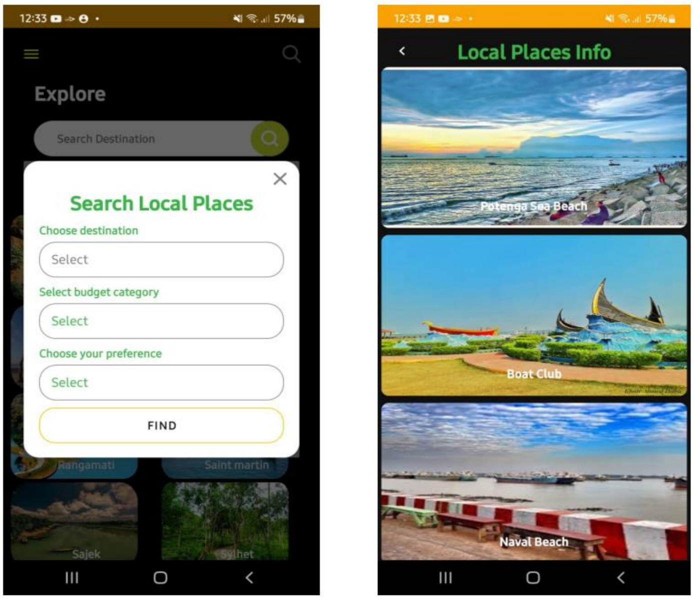
5. Evaluation Results

To evaluate our mobile application features, we have collected user ratings for this Tour Mate app from one hundred random reviewers. These ratings are shown in Fig.23 (b). From this user rating graph, it is seen that the good rating comments secured first place among all other ratings for the proposed application features (e.g., place suitability prediction using machine learning). The number of medium ratings, no comments, and bad ratings secured second, fourth, and third place, respectively. The number of bad ratings is very low compared with the good ratings. It can be seen from the figure that more than 60 percent of reviewers selected a good rating by evaluating all the app features. The number of users with medium ratings, no comments, and bad ratings for all features is less than 15 percent, 10 percent, and 5 percent, respectively. Table 1 evaluates application benefit analysis using 80 reviewer comments. Table 1 shows that more than 70% of reviewer comments are positive in terms of features, benefits, security, and market readiness.



a b

Fig.21. Place suitability result display in the mobile app



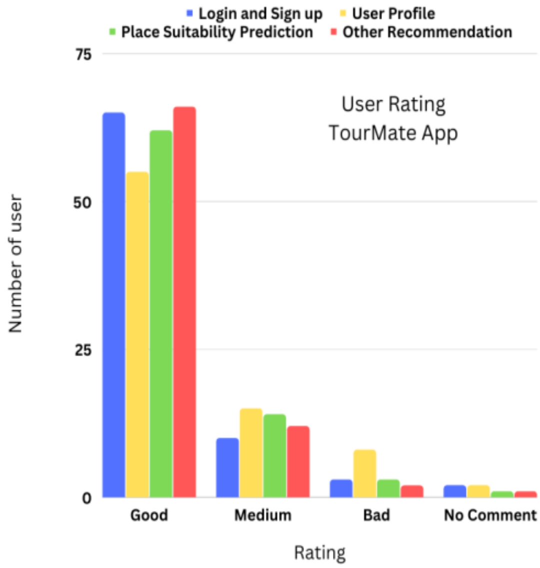
a b

Fig.22. Recommendations of local places

Table 1. Proposed application benefit analysis

question	Good	Not-good	Silent user
Useful for suitable travel place recommendation	70%	20%	10%
Proposed application can help tourist interms of hotel and transport finding	75%	20%	5%
Application considers multi-factor based place selection	70%	20%	10%
Suitable in terms of security and business	75%	15%	10%

a



b

Fig.23. a. Others travel related information page, b. User rating

6. Conclusions

Based This paper develops a prediction model for determining the suitability of travel destinations for individual users based on their preferences, destination information, budget, travel interests, and previous experiences. For this prediction model, we have collected data on various factors such as whether they (tourists) are sea lovers, mountain lovers, history lovers, or entertainment lovers, or not, hotel and transport service information, duration of the trip, desired place, budget, and the safety of tourists, among others. We then trained several machine learning algorithms, such as the decision tree or DT, the naive bayes, the KNN or k-nearest neighbor, the logistic regression, and the random forest, to predict the suitability of a place for a given user. In terms of accuracy, recall, specificity, and F1 score, we have seen that the Random Forest algorithm worked better. Besides, it had the highest AUC, indicating that it could better differentiate between true positives and true negatives than the other algorithms. That’s why this algorithm predicts the suitability of the destination. Our proposed mobile application offers several features, including travel and hotel information for users regarding a tourist place, specific place suitability prediction, and local tourist place recommendations based on the destination, budget, and user nature. The user rating analysis verified that the proposed mobile application features received more than 60 percent positive reviews from the reviewers regarding the usefulness of the proposed tourist assistance application.

Future Research Directions

The current recommendation engine focuses primarily on local places, hotels, transportation, travel agencies, and activities to do. But there is much potential to extend the scope of recommendations, including other factors like cultural experiences and dining options, to provide a more comprehensive travel planning experience for travelers. The current prediction model and recommendation engine were developed using data from a specific population and geographic region. There is a huge potential to test the model on different populations and geographic regions to determine whether it is applicable in other contexts and to identify any necessary adjustments. Any tourist can find the suitable local places for travel by using our proposed mobile application based on his or her criteria fulfillment. There is also much significant potential to improve the user experience and the user interface of the travel application. In the future, this work can incorporate personalized travel planning tools, potential bias analysis regarding machine learning algorithm selection, add social media sharing features, user safety alert, intruder detection, or generate better navigation and search facilities along with inclusion of different geographical area, different context like elder people’s choice, and location considerations, among others. For more enhanced security, in future this work can incorporate quantum cryptography and deep learning-based solution for personalized travel recommendation application.

References

[1] B. E. Moscow. Tourism in bangladesh. https://bangladeshembassy.ru/about-bangladesh/tourism/, last accessed on 20 may, 2023.

[2] Share Trip. Sharetrip - the largest travel portal in bangladesh. https://sharetrip.net/, last accessed on 20 may, 2023.

[3] Bangladesh Tourism Board. Official website of bangladesh tourism board. http://tourismboard.gov.bd/, last accessed on 20 may, 2023.

[4] Bangladesh Embassy Vienna. Tourism in bangladesh. https://bangladootvienna.gov.bd/tourismin-bangladesh/, last accessed on 15 may, 2022.
[5] TripAdvisor. Bangladesh tourism: Best of bangladesh. https://www.tripadvisor.com/Tourismg293935-Bangladesh-Vacations.html, last accessed on 15 may, 2023.
[6] Gozayaan. Gozayaan - travel to bangladesh and beyond. https://www.gozayaan.com/, last accessed on 15 may, 2023.
[7] Obokash. Obokash - your travel partner. https://www.obokash.com/, last accessed on 15 may, 2023.
[8] Amy Bangladesh. Amy Bangladesh - the largest online marketplace in bangladesh, https://www.amybd.com/index?v=987654484, last accessed on 10 may, 2023.
[9] T. Ghani et al., Amar bangladesh - a machine learning based smart tourist guidance system. in 2018 2nd International Conference on Electronics. Materials Engineering Nano-Technology (IEMENTech), 1–5, 2018.
[10] V. Parikh et al., A tourist place recommendation and recognition system. Second International Conference on Inventive Communication and Computational Technologies (ICICCT), India, 218-222, 2018.
[11] M. E. B. H. Kbaier et al., A personalized hybrid tourism recommender system. IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Tunisia, 244-250, 2018.
[12] C. Srisawatsakul et al., Tourism Recommender System using Machine Learning Based on User's Public Instagram Photos. International Conference on Information Technology (InCIT), Chonburi, Thailand, 276-281, 2020.
[13] P. A. Manjare et al., Recommendation System Based on Tourist Attraction. International Research Journal of Engineering and Technology (IRJET), 3(4): 877-881, 2016.
[14] E. Pantano et al., you will like it!' using open data to predict tourists' response to a tourist attraction. Tourism Management, 60(1):430–438, 2017.
[15] R. Jiang et al., Personalized cruise travel recommendation system based on data mining and glonass tools. 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), india, 1172-1175, 2022.
[16] R. Logesh et al., Exploring hybrid recommender systems for personalized travel applications. in Cognitive Informatics and Soft Computing, 535–544, 2018.
[17] P. Nitu et al., Improvising personalized travel recommendation system with recency effects. Big Data Mining and Analytics, 4(3):139-154, 2021.
[18] E. Paskahlia Gunawan et al., Development of an Application for Tourism Route Recommendations with the Dijkstra Algorithm. International Conference on Information Management and Technology (ICIMTech), Jakarta, Indonesia, 343-347, 2021.
[19] G. Rastogi et al., Tourist Spot Recommendation from Images based on Age Group and Location for Dubai using Deep Transfer Learning. 8th International Conference on Signal Processing and Communication (ICSC), Noida, India, 514-519, 2022.
[20] S. Y. Shu et al., A Web-based Application for Interesting Place Recommendation. International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 215-220, 2022.
[21] B. N. M. Casilla et al., Recommendation System for Tourist Routes Using Fuzzy Logic In The Arequipa City. IEEE Colombian Conference on Communications and Computing (COLCOM), 1-6, 2022.
[22] S. M. Hari Krishna et al., Trip Planner and Recommender using Flutter and Tensor Flow. IEEE 7th International conference for Convergence in Technology (I2CT), Mumbai, India, 2022, 1-7.
[23] M. T. Islam et al., "Designing Dashboard for Exploring Tourist Hotspots in Bangladesh, " 23rd International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2020, 1-6, 2020.
[24] T. Wenan et al., Analysis and Design of Tourism Recommender System for Religious Destinations of Nepal. IEEE International Conference on Sustainable Engineering and Creative Computing (ICSECC), Indonesia, 214-220,2020.
[25] R. Premakumara et al., Travel Recommendation System for Domestic Travelers in Sri Lanka. 3rd International Conference on Advanced Research in Computing (ICARC), Belihuloya, Sri Lanka, 2023, pp. 238-243.
[26] H. T. Kanakia et al., Trip Saathi: Travel Planning System based on Sentiment Analysis of Reviews. 4th International Conference on Electronics and Sustainable Communication Systems (ICESC) , Coimbatore, India, pp. 1216-1220, 2023.
[27] A. K. Oktavius et al., TRecommendation System Based Collaborative Filtering for Deciding Travelling Place. 2023 4th International Conference on Artificial Intelligence and Data Sciences (AiDAS), IPOH, Malaysia, 262-268, 2023.
[28] M. D. Mahardika et al., Recommender System for Tourist Routes in Yogyakarta Using Simulated Annealing Algorithm. IEEE *8th International Conference for Convergence in Technology (I2CT)*, India, 1-6, 2023.
[29] Idris et al., Assisting Smart Tourism Through Virtual Reality Apps for Tourists Destination in Indonesia. *8th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, Indonesia, 1-6, 2023.

Authors' Profiles

Azmain Abid Khan is a researcher at the computer science and engineering department, Chittagong university of engineering and technology. His major research interests include machine learning and mobile app development. He finished his bachelor of computer science degree in 2023.

Dr. Mahfuzulhoq Chowdhury obtained his PhD in 2018. He is currently a faculty member and researcher at computer science and engineering department, Chittagong University of Engineering and Technology. His major research interest includes machine learning, cloud computing, computer network, and mobile app development. He has published several journal and conference articles in reputated IEEE journals and conferences.

How to cite this paper: Azmain Abid Khan, Mahfuzulhoq Chowdhury, "TourMate: A Personalized Multi-factor Based Tourist Place Recommendation System Using Machine Learning", International Journal of Intelligent Systems and Applications(IJISA), Vol.16, No.4, pp.55-71, 2024. DOI:10.5815/ijisa.2024.04.04