Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

MACHINE LEARNING-BASED TRAVEL AND TOURISM FORECASTING SYSTEM

[1]Dr. Ummadi Thirupalu, Professor**, [2]Mr. G. Ratnaiah,** Associate Professor
[3]Mr. Ch. Ramesh, Assistant Professor**, [4]Mrs. Blessi Stella Mocharla,** Assistant Professor

[1234]Audisankara College of Engineering & Technology, NH-16, By-Pass Road, Gudur, Tirupati Dist, Andhra Pradesh, India-524101, India.

ABSTRACT – Tourism and travel industries are heavily reliant on predicting future trends and demand patterns to optimize services and resources. In thisproject, we propose a Tourism and Travels Prediction System (TTPS) using Machine Learning (ML) techniques to forecast travel trends and demand. TTPS leverages historical travel data, customer preferences, seasonal variations, and external factors suchas weather, events, and economic indicators to build predictive models. The system employs various ML algorithms such as regression, classification, and clustering toanalyze past travel data and extract meaningful insights. It utilizes featureengineering techniques to preprocess data and extract relevant features for model training.

Index Terms – Machine Learning, SVM, KNN, Tourist, Prediction, Logistic Regression.

I. INTRODUCTION

The tourism industry plays a pivotal role in the global economy, contributing significantly to the GDP of many countries. As travel patterns and preferences evolve,the ability to predict tourism trends becomes increasingly crucial for stakeholders, including government bodies, travel agencies, hospitality businesses, and local communities. Accurate predictions can enhance strategic planning, optimize resource allocation, and improve customerexperiences.

The advent of machine learning (ML) hasrevolutionized various industries byproviding powerful tools to analyze and interpret complex datasets.

In the context of tourism, ML algorithms cananalyze historical travel data, seasonal trends, social media influences, economic indicators, and other relevant factors toforecast future travel patterns. This projectaims to develop a robust Tourism and Travel Prediction System utilizing machine learningtechniques to deliver precise and actionable insights for the tourism sector.

The proposed system leverages a variety of machine learning models, including time series analysis, regression models, and clustering algorithms, to analyze vast amounts of data. By integrating data

from multiple sources such as travel booking platforms, social media, weather forecasts, and economic reports, the system can predict tourism trends with high accuracy. Additionally, the system will be designed to provide user-friendly visualizations and predictive analytics to help stakeholders make informed decisions.

II. LITERATURE SURVEY

Oscar Claveria et al (2015) in their research have used three machine learning techniques is support vector regression, Gaussian process regression, and neural network to improve forecast accuracy. Also, they have built one set of forecasts by estimating models on the aggregate series, another set by using the same models to forecast the individual series prior to aggregation and then compare the accuracy of both approaches [3]. S. Cankurt et al (2015) In their analysis, they analyzed the tourism demand series by considering some potential trends for incorporating complex time and tourism demand series relationships and enhancing forecast accuracy by extracting additional information available from the related data series in the tourism demand forecast. Also, they have employed multilayer perceptron (MLP) regression, and support vector regression (SVR) machine learning models [4]. Tasfiqul Ghani et al (2018) in their research, they have used a smartphone application to introduce machine learning in its program like Google Maps API and finds out the location of all points of interest added to the database by the Admin. The user is able to log in to this app using their basic information or Facebook details. They can deselect attractions which they do not prefer. The machine learning algorithms used in this application will find the best routes to reach destinations one by one. Users also are able to write notes as a mean of diary entry which will be saved on their phones offline database. Users may allow public access to those entries or restrict them from sharing too [5]. Indri Hapsari et al (2018) in their research developed the prediction model for visiting time. Multiple linear regression becomes the popular model because there are six factors that influenced the visiting time, for example, access, government, rating, number of reviews, number of pictures, and other information. Those factors become the independent variables to predict dependent variables or visiting time. also, they used the Ordinal Logistic Regression (OLR) by transforming the interval data-independent variable into ordinal data using Expectation Maximization.

Then they used the classification algorithm in machine learning by using five top algorithms which are Linear Regression, k- Nearest Neighbors, Decision Tree, Support Vector Machines, and Multi-Layer Perceptron [6]. Mohamed Elyes Ben Haj Kbaier et al (2017) in their paper, have

focused on building a personalized recommender system in the tourism field. they proposed a hybrid recommender system that combines the threemost known recommender methods whichare the collaborative filtering (CF), thecontent-based filtering (CB) and the demographic filtering (DF) in order implement these recommender methods, they have used different machine learning algorithms which are the K-nearestneighbors (K-NN) for both CB and CF and the decision tree for the DF. To enhance the recommendation accuracy, we used twohybridization techniques: switching and weighted. For the weighted approach, a novel linear programming model is appliedto obtain the optimal weights' values.

Da-Jie Lin et al (2019) in their research, proposed three approaches to predict the travel time, these approaches its Gradient Boosting Regression Tree (GBRT), K- Nearest Neighbors (KNN) and Linear regression (LR). The results showed to all machine learning methods used in this paper are well to predict the travel time [8]. MehrbakhshNilash et al (2017) In the paper, proposed a new recommendation technique based on multi-criteria CF to enhance the predictive accuracy of recommender systems in the tourism domain using clustering, DimensionalityReduction (DR), and prediction techniques. Also, they used Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Support Vector Regression (SVR) as prediction techniques, Principal Component Analysis (PCA) as a Dimensionality Reduction (DR) technique and Self-Organizing Map (SOM) and Expectation-Maximization (EM) as two well-known clustering techniques. Toimprove the recommendation accuracy of proposed multi-criteria CF, a cluster ensembles approach, Hypergraph Partitioning Algorithm (HGPA), is applied toSOM and EM clustering results [9]. Lin Shi-ting et al (2014) In their thesis, they have introduced the basic principle andprocess of ε-SVR (one algorithm of Support Vector Machine for Regression, SVR) is presented. This method is used to model tourist traffic prediction and predict one series of data.

Two different kernel functions are employed,and the former's performance is evidentlybetter than the latter. ε-SVR's performance isalso compared with that of the traditionaltime series analysis method, and the former outperforms the latter [10]. Kyoungho Son et al (2018) in their research, have implemented the LSTM machine learning method to predict visitors in a certain tourism place using machine learning.LSTM is one of the variations of RNN networks. they got some positive results showing its applicability in a real environment.

Xinquan WANG et al (2015) in their research, In order to more accurate prediction of annual tourism, using thesynthetic index method to calculate thetourism market boomindex, after timing

phase space reconstruction, merge theoriginal travel date and the tourism market boom index to get the sample. using extreme learning machine algorithm to train sample data. By comparing the support vector regression algorithm show that: the modelbased on extreme learning machine algorithm make higher precision, better fitting degree can more accurately estimate and forecast the tourism market, theapplication of this the model can provide guidance for the tourism market to achieve a reasonable allocation of resources and healthy development.

Siwen Zheng et al (2016) in their research, used kind of recommendations as a prediction problem based on the tourist's historical visiting sequences and supervised machine-learning algorithms, namely Random Forests (RF)and Lambda MART. Then, a comparison of five different machine learning algorithms, namely Random Forests, Lambda MART, RankingSVM, List Net, and Rank Boost have been taken on this feature set, results showed that the Random Forest outperformed the other algorithms with regard to prediction accuracy.

Amir Khatibiet al (2018) In their work, have suggested that access data from online social networks and travel websites, in addition to climate data, can be used to support the inference of visitation count for many touristic attractions. Social media andclimactic data each contribute to prediction accuracy. Further, the two are complementary and are well synthesized using Support Vector Regression.

NesreenKamel et al (2018) in their research, have considered different machine learning models that can be applied in the tourismprediction problem and to show the performance of seven well-known machine learning methods.

II PROPOSED SYSTEM

The overview of our proposed system is shown in the below figure.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

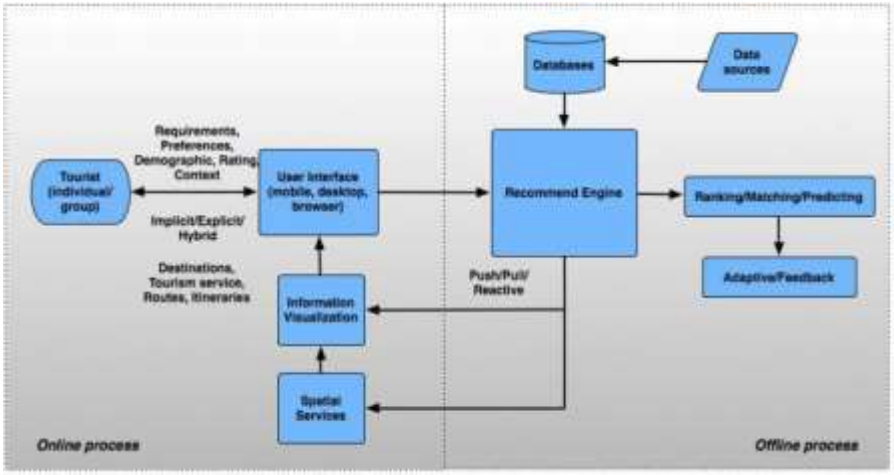


Fig. 1: System Overview

Implementation Modules

- Admin can login to system usingusername and password as 'admin and 'admin'. After login admin will add employee details and can viewavailable employee details. Admin can view all authenticated employee's login details by selecting FROM and TO date.

Employee

- Employee can login to system by using login details provided by admin. After login employee will create, modify Itinerary details, create, and modify transport and hotel details. Employee can view all bookings and feedback given by user
- **Register:** new user can sign up with the application and then view and book Itinerary, transport and hotel. User can cancel or give feedback and ratings to provided services
- **FAQ:** can view answers from frequently asked questions.

Implementation Algorithms

Support Vector Machine

- In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyse data for classification and regression analysis. An SVM training algorithm builds a

  model that assigns new examples to one category or the other, making it a non- probabilistic binary linear classifier.

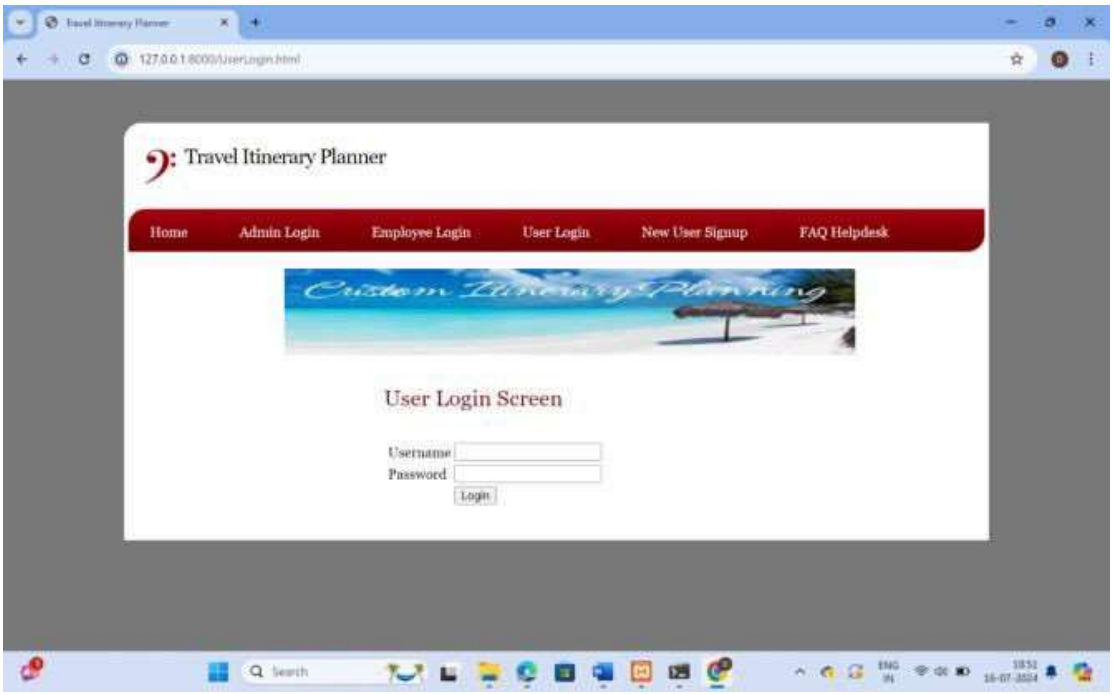
Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

LOGISTIC REGRESSION

- Logistic regression is one of the mostpopular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

- Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can beeither Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, **it gives the probabilistic values which lie between 0 and 1.**

IV RESULTS



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

Fig. 2: Home Page

Fig. 3: Admin Login

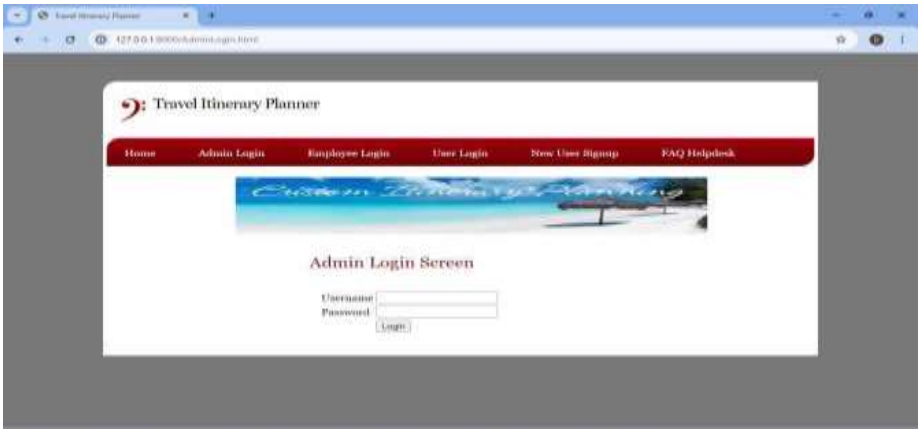
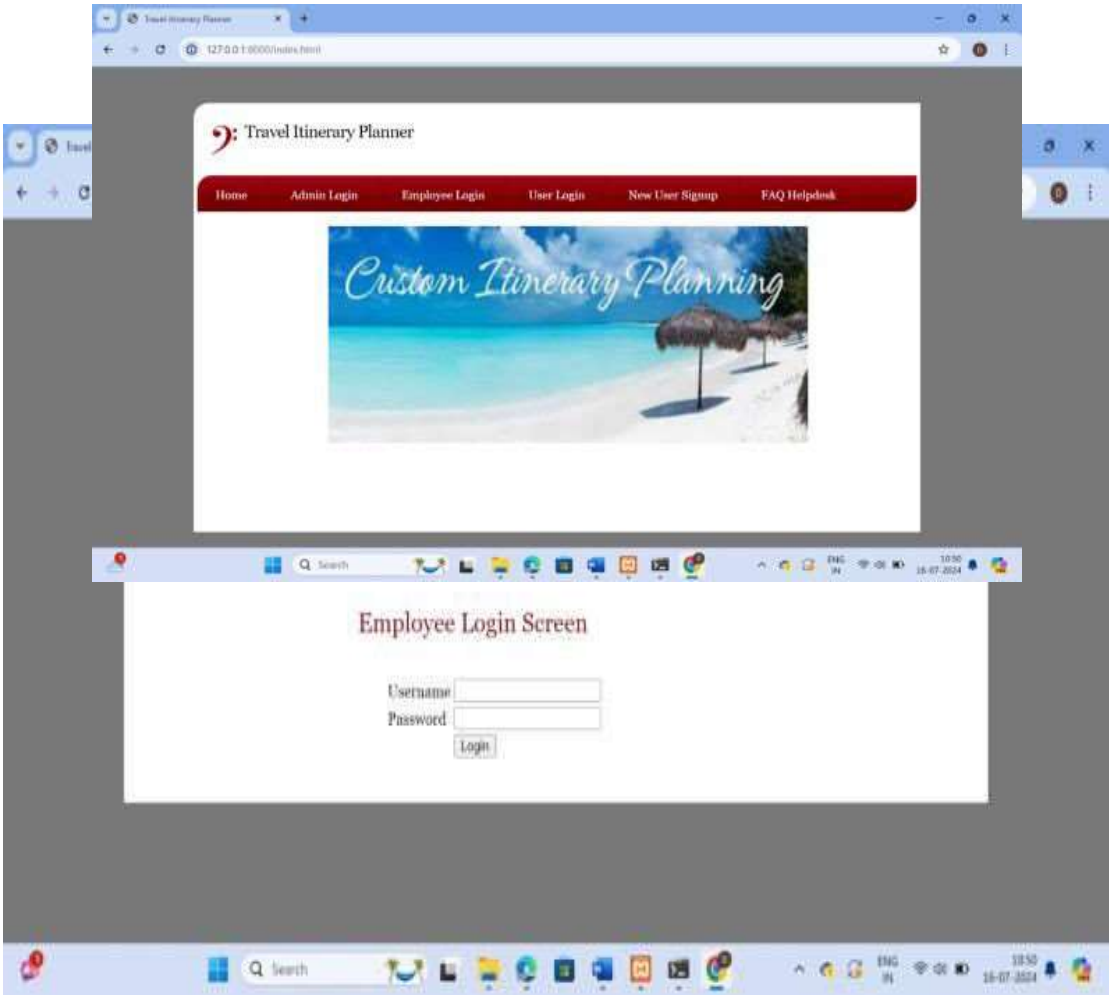


Fig. 4: Employee Login



Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

Fig. 5: User Login

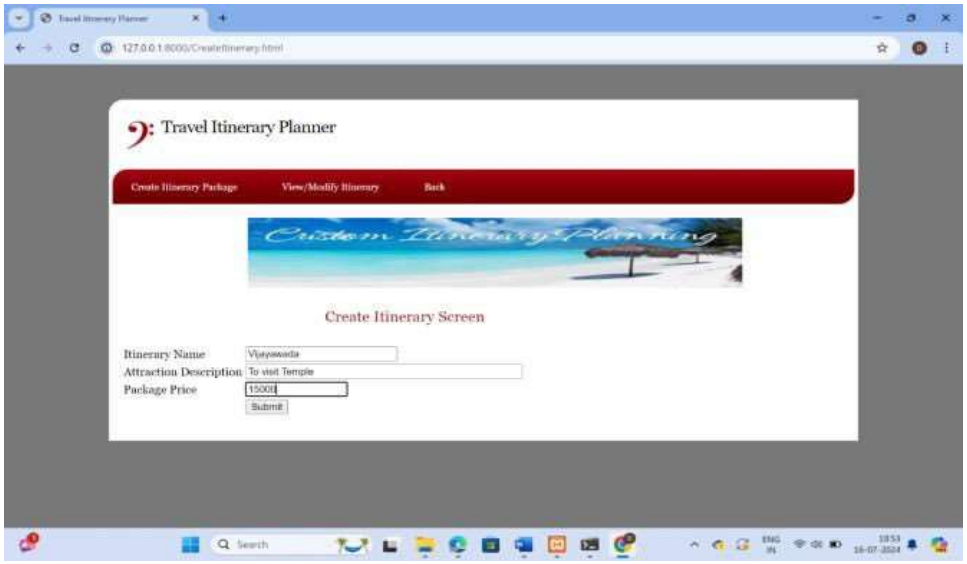


Fig. 6: Create a Itinerary

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 11, November : 2024

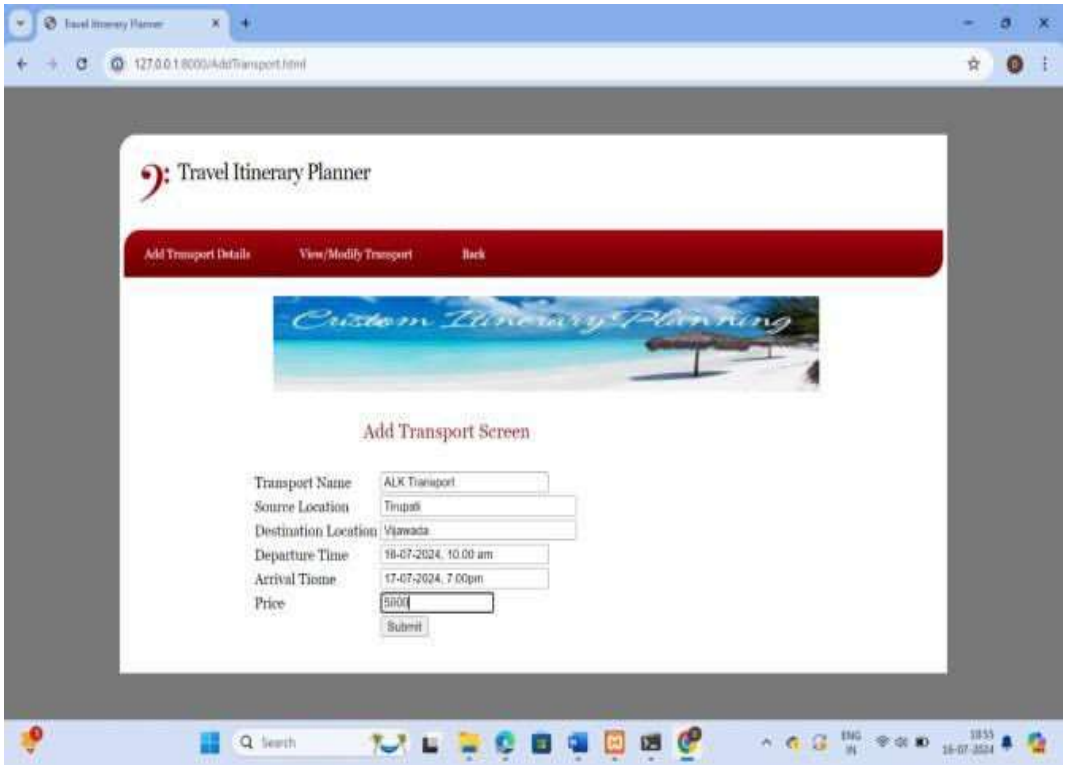


Fig. 7: Add Transport

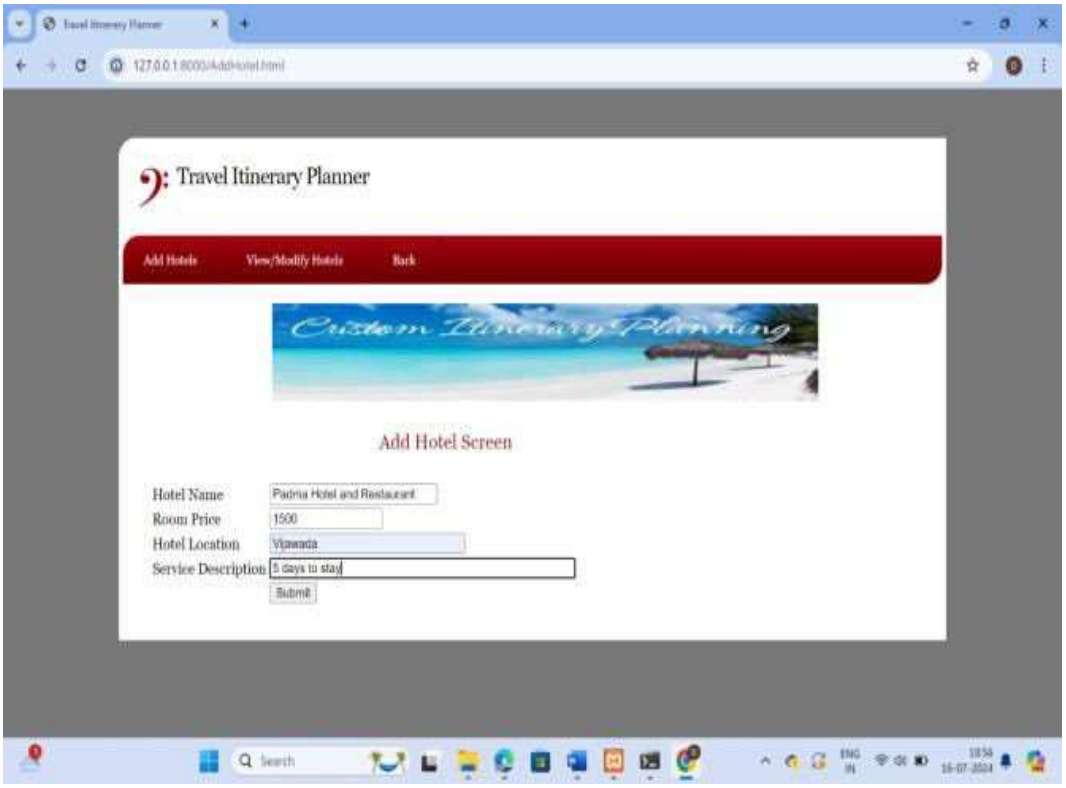


Fig. 8: Add Hotel

V. CONCLUSION

The development and implementation of a Tourism and Travel Prediction System using machine learning represent a significant advancement in the tourism industry. This system leverages historical data, patterns, and trends to predict future tourism activities, thereby enabling stakeholders tomake informed decisions. the Tourism and Travel Prediction System using machine learning is a transformative tool that offers substantial benefits to the tourism industry. By providing accurate predictions, optimizing resources, and enabling data- driven decision-making, it paves the way for a more efficient, sustainable, andpersonalized tourism experience. Future enhancements will continue to build on this foundation, driving further innovation and excellence in the sector.

REFERENCES

[1] Rout, H.B., Mishra, P.K. and Pradhan, B.B., 2016. Trend and progress of tourism in India: An empirical analysis.

[2] Karakitsiou, A. and Mavrommati, A., 2017. Machine learning methods in tourism demand forecasting: Some evidence from Greece. MIBES Transactions, 11(1), pp.92-105.

[3] Claveria, O., Monte, E. and Torra, S.,2016. Combination forecasts of tourism demand with machine learning models. Applied Economics Letters, 23(6), pp.428-431.

[4] Cankurt, S. and Subasi, A., 2015. Developing tourism demand forecasting models using machine learning techniques with trend, seasonal, and cyclic components. Balkan Journal of Electrical and Computer Engineering,3(1), pp.42-49.

[5] Ghani, T., Jahan, N., Ridoy, S.H., Khan, A.T., Khan, S. and Khan, M.M., 2018, May. Amar Bangladesh-a Machine Learning Based Smart Tourist Guidance System. In 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech) (pp. 1-5). IEEE.

[6] Hapsari, I. and Surjandari, I., 2018, May. Visiting Time Prediction Using Machine Learning Regression Algorithm. In 2018 6th International Conference on Information and CommunicationTechnology (ICoICT) (pp. 495-500). IEEE.

[7] Kbaier, M.E.B.H., Masri, H. and Krichen, S., 2017, October. A personalized hybrid tourism recommender system. In 2017IEEE/ACS 14th International Conferenceon Computer Systems and Applications (AICCSA) (pp. 244-250). IEEE.

[8] Lin, D.J., Tsao, W.L., Yu, C.H., Liu, H.H. and Chang, Y.C., 2019, April. The Travel Time Prediction by Machine Learning Methods with Traffic Data in Chiayi City, Taiwan. In 2019 4th International Conference on Electromechanical Control Technology and Transportation (ICECTT) (pp. 257- 260).

[9] Nilashi, M., Bagherifard, K., Rahmani, M. and Rafe, V., 2017. A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques. Computers & industrial engineering, 109, pp.357-368.

[10] Shi-Ting, L. and Bo, X., 2014, October. The application of improved svm for dataanalysis in tourism economy. In 2014 7th International Conference on Intelligent Computation Technology andAutomation (pp. 769-772). IEEE.