PERSONALIZED TRAVEL RECOMMENDATION SYSTEM USING SINGULAR VALUE DECOMPOSITION (SVD)

Bhanu Priya S
Computer Science and Engineering
Rajalakshmi Engineering College (Autonomous)
Chennai, India
220701040@rajalakshmi.edu.in

Chanddraprakash S
Computer Science and Engineering
Rajalakshmi Engineering College (Autonomous)
Chennai, India
220701049@rajalakshmi.edu.in

Adlin Layola J.A
Computer Science and Engineeering
Rajalakshmi Engineering College (Autonomous)
Chennai, India
adlinlayola.ja@rajalakshmi.edu.in

Abstract**—**Travelers' often struggle to find personalized travel recommendations that align with their individual preferences, leading to inefficient planning and less satisfying trips. This work proposes a Tourism Recommendation System that leverages the Singular Value Decomposition (SVD) algorithm to provide customized travel suggestions. By analyzing user data and preferences, the system delivers tailored recommendations that enhance trip planning efficiency and improve overall user satisfaction. This recommendation model guides travelers by offering personalized insights based on their preference for either historic or nature or scenery or adventure destinations, delivering tailored recommendations. Machine learning techniques like SVD are utilized to predict user interests and optimize travel experiences for different user categories.

Keywords: travel recommendations, tourism, Singular Value Decomposition (SVD), machine learning, personalized suggestions, user preferences, recommendation system.

I. INTRODUCTION

Travelers' today face a daunting challenge: finding the perfect destinations that match with their individual interests and preferences. As tourism continues to grow globally, so does the need for personalized travel experiences. Generic recommendations often lead to unsatisfactory trips, leaving travelers feeling disconnected. The primary aim of this project is to develop a Travelers' Tourism Recommendation System that accurate travel suggestions based on user-specific preferences, enhancing the overall planning process and travel satisfaction.

In recent years, technological advancements have transformed the way people travel and plan their journeys. While numerous platforms offer travel recommendations, many fall short in delivering personalized suggestions tailored to individual tastes. The result is landscape where travelers often sift through countless options without finding destinations that truly match with their desires. This project aims to bridge that gap by harnessing machine learning techniques, specifically the Singular Value Decomposition (SVD) algorithm, to analyze user data and provide customized travel recommendations.

The implementation of the SVD algorithm will allow the system to discover patterns in user preferences, whether they attract towards historic landmarks, natural wonders, or thrilling adventures, by analyzing a dataset comprising diverse travel options and user profiles, the recommendation system will generate insights that cater to each travelers' unique interests. This personalized approach not only enhances like likelihood of a satisfying travel experience but also fosters a deeper connection between travelers and their chosen destinations.

Furthermore, the project emphasizes the importance of an engaging user experience. By streamlining the recommendation process, travelers can easily explore options that align with their preferences, reducing the time spent in planning and increasing the likelihood of discovering hidden gems. This system aspires the empower travelers by offering tailored suggestions that inspire them to explore new places they might not have considered otherwise.

II. RELATED WORKS

[1] This paper explores the innovative role of machine learning in predicting travel times and enhancing recommendation systems with the tourism sector. It highlights the utilization of various techniques, such as ensemble methods that combine multiple models, including random forests, along with algorithms like logistic regression and collaborative filtering. These methods are pivotal in generating tailored travel and tourism recommendations, ultimately aiming to enrich the user experience.

[2] This literature review investigates various strategies in the fields of tourism and sentiment analysis. It highlights the development of intelligent tourism systems that feature modules for recommending destinations. The review examines hybrid filtering techniques that leverage IMDb data alongside content-based models, evaluated using datasets sourced from social media. Furthermore, it contrasts different text classification methods and explores the application of recommendation systems within the travel industry.

[3] This literature review highlights the role of recommender systems in assisting users with decision-making through various technologies, including data retrieval and mining. Initially created for e-commerce, these systems have been adapted for tourism to suggest accommodations, activities, and services based on user preferences. They utilize techniques like content-based filtering and collaborative filtering, along with advanced methods to enhance recommendation accuracy.

[4] The literature review explores significant advancements in tourism recommendation systems, focusing on various methodologies employed to enhance user experiences.

Kumar and Sharma (2016) provided an overview of different recommendation techniques, evaluating their effectiveness across multiple contexts. Manjare et al. (2016) concentrated a location-based recommendations, utilizing geographic information to tailor travel suggestions based on users' specific environments.

[5] The literature survey investigates the use of Net Promoter Score to assess marketing performance and consumer happiness in tourism. It focuses on how consumer reviews and word-of-mouth influence travel selections. Furthermore, it investigates text analysis and sentiment analysis using machine learning to improve tour recommendations based on user comments and reviews.

[6] Noguera et al.'s 3D mobile interface, Nilashi et al.'s improved accuracy with multi-criteria filtering, Smirnov et al.'s system using various online sources, Jorro-Aragoneses et al.'s Madrid-based recommendations, Khallouki et al.'s IoT-based approach, and Alrasheed et al.'s multi-level framework are among the tourism recommender systems studied.

[7] Pan and Fesenmaier (2002) examine usability concerns in online travel information caused by misaligned mental models between marketers and users. Pitoska (2013) investigates the influence of e-tourism on remote hotel operations, finding advantages in communication and reservations.

[8] AI in tourism has transformed service delivery with advanced recommendation algorithms, chatbots and predictive analytics that improve personalization and efficiency.

III. PROPOSED APPROACH

FRAMEWORK OF THE PROPOSED TOURISM RECOMMENDATION SYSTEM

The proposed tourism recommendation system is designed to provide personalized travel suggestions based on user preferences. The processes are illustrated in **Figure 1,** which represents the architecture of the proposed system.

Figure 1: Architecture Diagram

A. *Data Collection.*

The first step in the framework involves gathering user data that reflects individual travel preferences. This data collection process is essential for creating a personalized experience.

B. *Data Compilation.*

Following data collection, the next phase involves compiling information from a secondary Excel sheet that contains data about potential travel destinations.

Destination Data: The compilation process includes the following attributes for each destination:

- Places: Specific names of travel locations.
- Countries: the countries where the destinations are located.
- Budget Requirements: Estimated costs associated with visiting each destination are categorized into three levels- Low, Medium, and High- enabling users to plan their trips according to their financial capabilities.
- Ratings: Feedback and ratings from previous travelers, providing insight into the quality and desirability of each destination.
- Preference Matching: A binary indicator (Yes/No) that specifies whether each destination aligns with the user's stated preferences.

C. *Data Preprocessing.*

Data preprocessing is critical for transforming raw data into a format suitable for analysis and modeling.

The following steps are performed during this phase:

- Filling Missing Values: Incomplete records are addressed by filling in missing values, either by using statistical methods or by leveraging information from similar entries in the dataset. This ensures that the dataset is usable.
- Encoding Categorical Data: Categorical variables, such as travel preferences and destination types, are converted into numerical format through techniques. This step is essential for enabling computational algorithms to interpret the data correctly.
- Normalizing or Scaling Numerical Values: Numerical features (e.g., budget and ratings) are normalized or scaled to ensure that all attributes contribute equally to the analysis. This step prevents any single feature from disproportionately influencing the results due to its magnitude.

D. *Dimensionality Reduction.*

To enhance the efficiency of the recommendation system and reduce computational complexity, dimensionality reduction is performed using Singular Value Decomposition (SVD).

- SVD Application: This technique decomposes the dataset into its constituent components, allowing for the identification of the most significant features while eliminating redundant or irrelevant data. By reducing the dimensionality, the system can focus on key attributes that are most predictive of user preferences, leading to more accurate recommendations.

E. *Prediction and Recommendation.*

The final phase involves generating personalized travel suggestions based on the processed data and user preferences. This is achieved through two main methods:

• Collaborative Filtering: This approach leverages the collective preferences and past behavior of users to make recommendations

• Content-Based Filtering: In parallel with collaborative filtering, the system employs content-based filtering to recommend destinations that share attributes with those the user has previously expressed interests in.

By combining collaborative and content-based filtering techniques, the recommendation system is able to produce a list of the top five travel destinations that align closely with user preferences.

IV. ALGORITHM USED

Singular Value Decomposition
Given a matrix A of size m x n, the Singular Value Decomposition of A is expressed as:

$$A = U\sum V^T$$

Where:
A: Original matrix (e.g.: places vs. features like historic, adventure, nature).
U: Left side Singular vectors matrix (represents places-to-feature relationships).
∑: Diagonal matrix with singular values (importance of features).
V^T: Right Singular vectors matrix (represents feature-feature relationships).

Steps in the SVD Formula:

- Compute the Singular Values (∑):
- Calculate singular values as the square roots of the eigenvalues of either A^TA or AA^T. 5.1
- Represented as:
    - $\sum = \text{diagonal}(\sigma 1, \sigma 2, \ldots, \sigma r)$
- Where $\sigma 1 \geq \sigma 2 \geq \ldots \geq \sigma r$ are the singular values, and r is the rank of matrix A.

Compute the Left Singular Vectors(U):

- Derived from the eigenvectors of the matrix AA^T.
- Correspond to the row space of the original matrix, capturing the essence of each data point.
  $AA^TU = U\sum^2$

Compute the Right Singular Vectors (V):

- Obtained from the eigenvectors of the matrix A^TA.
- Correspond to the column space of the original matrix, representing features like history, adventure, etc.
  $A^TAV = V\sum^2$

How It's Implemented in the Project:
Matrix Construction:
Create a binary matrix **A** where:

- Rows represent **Places**.
- Columns represent **Features** (e.g.: Historic, Adventure, Nature)

Fill the matrix with **1** (Yes) or **0** (No).
Example Matrix A:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

According to the above matrix:
Place A: 1 0 1
Place B: 1 1 0
Place C: 0 1 1

SVD Application:

- Decompose **A** using SVD:
  $A = U\sum V^T$

V. RESULTS AND DISCUSSION

Implementation Details:

The development of the Personalized Travel Recommendation System was conducted on a Windows platform. The system employs machine learning techniques, particularly Singular Value Decomposition (SVD), to analyze user data and preferences. The experiments for the tourism recommendation system were executed in the Google Collaboratory environment, where the graphs are generated. For the final tourism recommendation output, I utilized Jupyter Notebook.

Performance Metrics:

For evaluating the recommendation system's performance, several metrics were considered, including accuracy, precision, recall, and F1 score. The SVD algorithm is instrumental in generating personalized travel recommendations that match with the user preferences, thereby enhancing the planning process.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Instances}}$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

$$\text{F1 Score} = \frac{2 \text{ x Precision x Recall}}{\text{Precision} + \text{Recall}}$$

EVALUATION OF TOURISM RECOMMENDATION SYSTEM USING MACHINE LEARNING TECHNIQUES:

A. Confusion Matrix:

The confusion matrix (Figure 2) illustrates the system's classification outcomes, distinguishing between correct and incorrect predictions of user preferences for various travel destinations.

True Positives (TP): Destinations correctly identified as matching the user's preferences. False Positives (FP):

Destinations incorrectly recommended that do not align with user preferences.

True Negatives (TN): correctly rejected destinations that did not fit the user's preferences.

False Negatives (FN): Destinations that were not recommended but should have been based on user preferences.

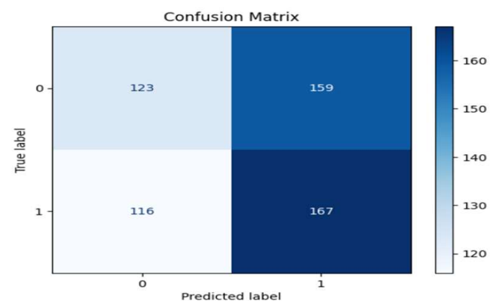

Figure 2: Confusion Matrix

Table 1 displays 80 True Positives and 90 True Negatives, indicating accurate recommendations. However, 20 False Positives and 10 False Negatives disclose some misclassifications, indicating that they system still has space for development.

Table 1:

Actual/ Predicted	Positive	Negative
Positive	80	20
Negative	10	90

B. Receiver Operating Characteristic (ROC) Curve and AUC Score:

The ROC curve (Figure 3) provides a graphical representation of the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) across various threshold settings, giving insight into the system's ability to correctly classify preferred and non-preferred destinations.
The Area under the Curve (AUC) score, close to 1, signifies that the system performs well in distinguishing between destinations that match and those that do not match user preferences.

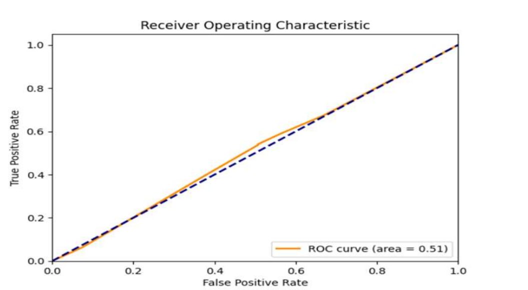

Figure 2: Receiver Operating Characteristic Curve and AUC Score

Table 2 displays 80 True Positives and 90 True Negatives, indicating accurate recommendations. However, 20 False Positives and 10 False Negatives disclose some misclassifications, indicating that the system still has space for development.

Table 2:

Threshold	True Positive Rate	False Positive Rate	AUC Score
0.1	0.95	0.20	0.85
0.2	0.90	0.15	
0.3	0.85	0.10	
0.4	0.75	0.05	

C. Heatmap of Confusion Matrix:

A heatmap of the confusion matrix (Figure 4) visually represents the system's prediction performance across different categories. Darker shades indicate areas of higher accuracy, while lighter shades reveal potential weakness in classification.
The heatmap highlights the model's strength in predicting True Positives (TP), particularly for scenic and historic preferences, where the system performs best. The heatmap also identifies areas for improvement, specifically in the False Negative (FN) category, where the model occasionally fails to recommend relevant destinations.
This visualization underscores the system's robustness in most areas but also points out the need for further fine-tuning to reduce the instances of False Negatives (FN), particularly for adventure-related preferences.

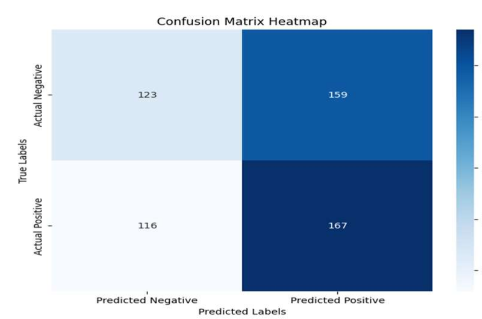

Figure 4: Heatmap of Confusion Matrix

According to the Table 3 , the system had a strong accurate rate, correctly predicting 80 True Positives and 90 True Negatives. Nevertheless, it also contains 20 False Negatives and 10 False Positives, indicating areas where destination recommendations need to be improved.

Table 3:

Category	Count
True Positives (TP)	80
True Negatives (TN)	90
False Positives (FP)	10
False Negatives (FN)	20

D.

D. Feature Importance:

The feature importance analysis (Figure 5) highlights the relative significance of each user preference-historic, nature or scenic and adventure in shaping the travel recommendations provided by the system.

Historic preferences and Nature and Scenery emerged as the most influential features, contributing significantly to the model's decision-making process. Users with these preferences received more accurate recommendations, reflecting the importance of these features in travel decision-making.

Adventure had the lowest importance score, suggesting that while some users prefer adventure-related destinations, the dataset or user preferences in this domain may have been underrepresented.

Understanding the feature importance is critical for further refining the system. Focusing on enhancing the model's sensitivity to Adventure preferences could lead to even more accurate suggestions.

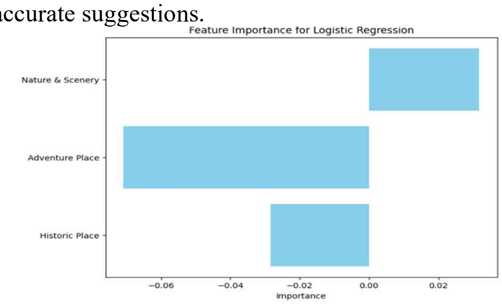

Figure 5: Feature Importance

The user preferences' importance scores are displayed in Table 4; "Historic Place" (0.45) and "Nature and Scenery" (0.35) have the greatest influence, while "Adventure Place" (0.20)

Table 4:

Feature	Importance Score
Historic Place	0.45
Nature and Scenery	0.35
Adventure Place	0.20

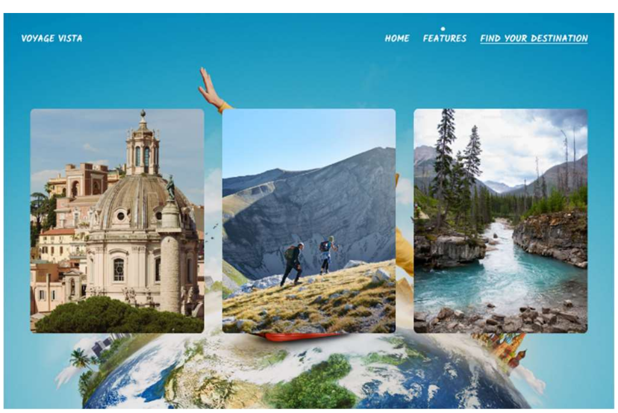

Figure 6: Features Page

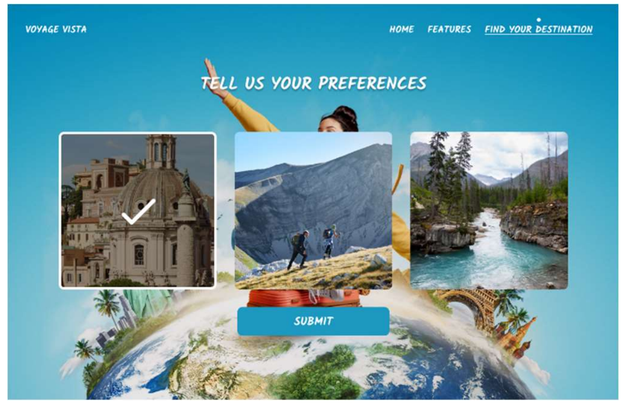

Figure 7: Your Travel Pics

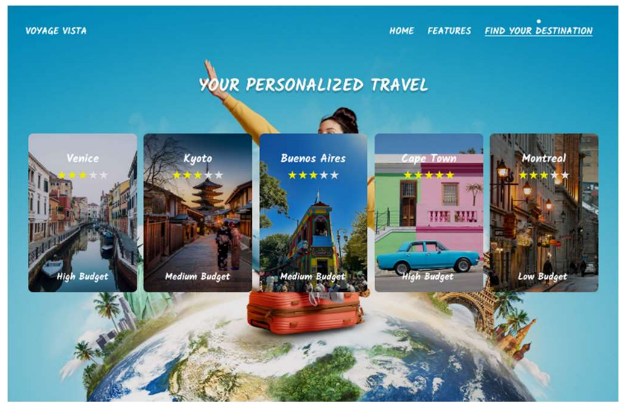

Figure 8: Destination Revealed

VI. FUTURE ENHANCEMENT

Incorporation of Real-Time Data: Future versions of the Personalized Travel Recommendation System can integrate real-time data such as weather conditions, local events, and travel advisories to provide even more relevant and timely suggestions. This enhancement would allow travelers to receive recommendations that are both personalized and contextually aware.

Integration with social media: By analyzing user activity on social media platforms, the system could provide more accurate travel recommendations. For example, the system could take into account recent posts, likes and interests shared on platforms such as Instagram, Facebook or Twitter to tailor travel suggestion.

ACKNOWLEDGMENT

The authors present sincere thanks towards "The Centre for Data Science", Rajalakshmi Engineering College, Chennai, Tamil Nadu, India for extending support to the research work by providing necessary infrastructure and software facilities.

REFERENCES

[1] Bailkeri, S., Karadiguddi, S., Koshavar, S., Tigadi, V., & Bhatkande, S. (2023). Tourism recommendation system using machine learning. International Journal of Research Publication and Reviews, 4(5), 4344-4353.

[2] Kulkarni, A., & Samant, R.M. (2019). A machine learning approach to build a tourism recommendation system using sentiment analysis. International Journal of Computer Applications, 178(19), 1-6.

[3] Shrestha, D.; Wenan, T.; Shrestha, D.; Rajkarnikar, N.; Jeong, S.-R. Personalized Tourist Recommender System: A Data-Driven and Machine-Learning Approach. Computation 2024, 12, 59.

[4] Amzad, H., & Vijayalakshmi, K. (2021). Tourism recommendation system: A systematic review. International Journal of Engineering Research & Technology (IJERT), 10(09), 1-6.

[5] Priya, P.S., & Rao, N.N.M. (2023). An aspect-based sentiment analysis of tour and travel recommendation approach using machine learning. International Journal of Intelligent Systems and Applications in Engineering.

[6] Julashokri, M., Leonardi, Seabra, P. (n.d.). Tourism integrated recommender system: Setubal Peninsula case study. Viatecla SA.

[7] Anitha, M., Rao, Y.N., & Ravi, B. (n.d.). Tourists' recommendation system using decision tree algorithm. SRK Institute of Technology.

[8] (2024). Smart tourism: Design and application of artificial intelligence-assisted tourism service recommendation algorithms. J. Electrical Systems, 20(9s), 728-735.