Sherpa: Implementing a Hybrid Recommendation System for Next-Gen Tourist Experience

Inam Ullah Khan, Abdur Rahman, Muhammad Awais Khan, Dr. Madiha Sher, Dr. Yasir Saleem

Department of Computer Systems Engineering, University of Engineering and Technology Peshawar Pakistan

***Correspondence**: 20pwcse1862@uetpeshawar.edu.pk 20pwcse1878@uetpeshawar.edu.pk 20pwcse1871@uetpeshawar.edu.pk madiha@uetpeshawar.edu.pk yasirsaleem@uetpeshawar.edu.pk

Citation | Khan. I. U, Rahman. A, Khan. M. A, Sher. M, Saleem. Y, "Sherpa: Implementing a Hybrid Recommendation System for Next-Gen Tourist Experience", IJIST, Special Issue. 2 pp 37-42, May 2024

Received | May 02, 2024, **Revised |** May 10, 2024, **Accepted |** May 15, 2024, **Published |** May 20, 2024.

In the digital era, Sherpa revolutionizes personalized tourism with an AI-driven recommendation system, fostering meaningful connections between travelers and local guides. This study explores Sherpa's integration of collaborative and content-based filtering—specifically, singular value decomposition (SVD) and cosine similarity—to tailor travel experiences uniquely. Our methodology includes a detailed examination of Sherpa's algorithm and its implementation within a cross-platform, MERN Stack-powered backend. We assess the system's efficacy in aligning recommendations with individual user preferences, based on quantitative user feedback and engagement metrics. Initial results demonstrate a significant improvement in personalized experience satisfaction. The paper concludes that Sherpa's innovative approach not only enhances the quality of travel recommendations but also sets a new standard for interactive and adaptive tourism platforms. Through continuous algorithmic refinement, Sherpa is positioned to lead a transformative shift in how travelers explore new destinations, offering not just journeys, but transformative experiences.

Keywords: Artificial Intelligence; Digital Tourism; Collaborative Filtering; Content-Based Filtering; Hybrid Recommender System; Singular Value Decomposition; Cosine Similarity Matrix.

Introduction:

Travel isn't just seeing sights anymore. People want deeper experiences when they go places. They don't just want planned schedules and surface interactions. Now, travelers aim to connect with locals. They hope to uncover hidden gems and make lasting memories. The old tourism ways don't satisfy these changing desires. Connectivity and curiosity drive travel today. Travelers yearn for meaningful encounters, not just observing from outside. They aspire to immerse themselves in the cultures they visit. Sherpa shows a change in travel. It pictures tourists acting as members of host towns, not just visitors. Sherpa aims to make guided tourism different. It connects tourists with local guides, creating a good partnership where exploration is a shared project.

Traditional tourism has limits - guidebooks and online sites offer insights but lack the personal touch and real-time flexibility that only locals provide. Sherpa recognizes this limitation. It aims to give visitors and guides more influence by fostering cross-cultural dialogue and offering customized memorable experiences. Sherpa's genesis lies in addressing tourism's shortcomings. It provides a venue where cultures connect, making each trip unique through personalization. Sherpa uses an AI matching system. This system matches tourists with guides. It looks at what tourists like. It also looks at guides' skills. The AI finds hidden patterns. It uses info from tourists and guides. The goal is great tourist trips. The system matches based on tourist wants.

The goal of offering suggestions suited to each person's particular tastes and interests remains a key issue for the travel industry, despite technological advancements. Sherpa tackles this challenge head-on with an innovative hybrid recommendation system. This approach blends content-based filtering using cosine similarity matrices with a seamless integration of collaborative filtering via singular value decomposition (SVD) models. By combining these methods, Sherpa provides travelers with immersive, personalized experiences that transcend the limitations of conventional tourism models. Simply, Sherpa isn't only a mobile app. It shows a new way to do guided tours. Sherpa means adventure, connecting, and real experiences. It will make travel better for visitors and locals. As Sherpa grows, it will keep changing how we travel. One personal tip at a time.

Related Work:

The impact of technology on guided tourism is significant. Early examples like Trip Mate, TripAdvisor, and Airbnb have developed services centered on personalized travel experiences. These platforms offer personalized suggestions, local experiences, and intelligent lodging options through AI and big data. However, they sometimes compromise traveler interests, as they cater to only a limited range of customer preferences.

On the other hand, Sherpa sets itself apart by using a hybrid recommender system that merges the advantages of collaborative filtering and content-based filtering. This combination enables Sherpa to offer personalized recommendations that cater to the distinct preferences of individual travelers. Lots of research has been done to improve the performance of basic Matrix Factorization used in Content-based Filtering. In [1], Zhang and colleagues introduced Weighted Non-negative Matrix Factorization (WNMF) as a method to enhance NMF (Non-negative Matrix Factorization). They utilized weights as an indicator matrix to represent the visibility of entries in the matrix R. In [2] Lee et al. gave the idea of Non-negative Matrix Factorization (NMF) to enforce non-negativity in U and V, which was proved to be useful in computer vision fields. In [3], Salakhutdinov et al. showed Probabilistic Matrix Factorization (PMF), which used Gaussian distribution to initialize U and V and applied a logistic function to limit the range of predicted to [0,1]. Koren et al. summarized this work in [4] and gave a generic framework for Matrix Factorization. Researchers also managed to incorporate information from other data sources. Zhang et al. used review sentiment analysis to construct virtual ratings for users who have not explicitly on the item [5]. Gu et al. proposed the Graph Weighted Nonnegative Matrix

Factorization (GWNMF) [6] to use user/item neighborhood graphs to preserve neighborhood information in user/item latent vectors [5][2]. utilized social network information under the assumption that friends share similar tastes and interests. In the realm of recommender systems, collaborative filtering approaches like matrix factorization and nearest-neighbor methods have received a lot of attention and have proven to be effective at recognizing similar people or products and capturing user preferences. Similar to this, item attributes and user profiles have been analyzed by content-based filtering techniques like cosine similarity and natural language processing (NLP) to produce personalized suggestions.

Methodology:

Sherpa's development process is methodical and iterative, starting with the design and implementation of its fundamental components and system architecture. The MERN (MongoDB, Express.js, React.js, Node.js) stack powers the program's backend, while a cross-platform mobile application frontend created with React Native powers the application's architecture. This architecture makes sure that various devices and operating systems work together seamlessly, giving users a consistent and easy experience. The cornerstone of Sherpa's personalized experience is its hybrid recommender system, which combines collaborative filtering with SVD and content-based filtering using cosine similarity matrices. Content-Based Filtering (Cosine Similarity Matrix):

Matrix factorization [4] is one of the most used approaches in recommender systems. Despite its efficiency, MF still suffers from sparsity problems, i.e., users who rate only a small portion of items could not get proper recommendations, and items with few ratings may not be recommended well. To cure the Sparsity problem, we have utilized the Novel approach of Cos MF from [7]. The formula for Matrix Factorization is given in eq. (1):

$$R_{i,j} = U_i V_j^t \qquad (1)$$

Where n is users and m is items. Each element of $R_{i,j}$ represents the rating of user i to item j (trip). Ui denotes the latent preference row vector of user i and Vj denotes the latent feature row vector of item j, and both have k dimensions (latent factors).

Cosine Similarity is calculated using eq. (2):

$$R_{i,j} = \frac{U_i V_j^t}{||U_i||.||V_j||} = \cos(U_i, V_j) \qquad (2)$$

Collaborative Filtering analyzes a user preferences database to predict additional products or services in which a user might be interested [8]. Collaborative filtering techniques analyze user-item interactions and user preferences to identify similar users or items, while content-based filtering techniques analyze item attributes and user profiles to generate recommendations that match users' preferences. Singular Value Decomposition is carried out using eq. (3):

$$A = U\,\mathcal{E}\,V^t \quad (3)$$

Where A is the input matrix, U is the left singular matrix, Sigma is diagonal/eigenvalues and V is the right singular matrix. By leveraging the strengths of both approaches, Sherpa can deliver highly personalized recommendations that resonate with individual users.

Sherpa's hybrid recommender system is trained on a curated dataset sourced from TripAdvisor, one of the world's largest travel platforms [9]. This dataset contains a diverse range of user interactions, reviews, and item attributes, providing Sherpa with rich and comprehensive data to inform its recommendation process. The dataset is preprocessed and transformed into a suitable format for training the recommender system, ensuring that it captures the underlying patterns and relationships in the data.

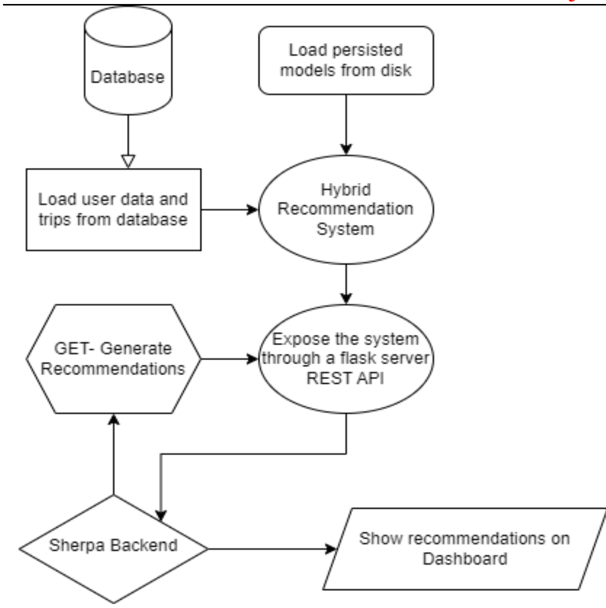

Figure 1: Training the Hybrid Recommender system.

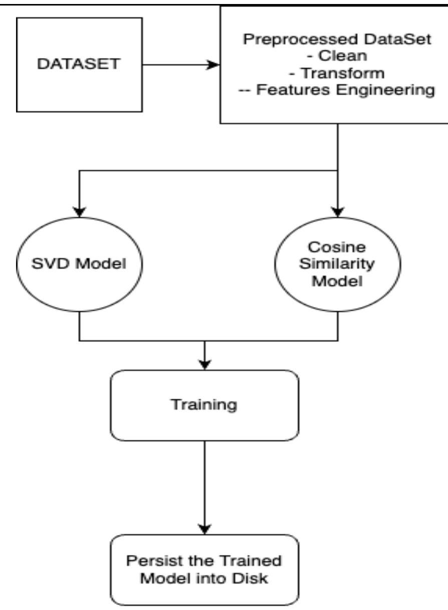

Figure 2: Detailed Process: Integration with Sherpa

Figure 1 outlines the methodology adopted to train the hybrid recommendation system using the procured Trip Advisor Dataset. First, we've carried out necessary cleaning and transformations as well as feature extraction on the dataset. This dataset is then divided into training and testing sets by a 70/30 margin. The training set is used to train the SVD and cosine similarity Models of course each model utilizes different features of the dataset, it is subsequently tested and performance metrics are calculated before persisting the model onto storage.

As Figure 2 demonstrates, the hybrid recommendation system is then integrated with the Sherpa Backend by exposing the get recommendation function through a Flask Server Rest API, which the backend calls/hits to generate recommendations for each user on the go.

Results:

As shown in the table the Hybrid Recommendation System managed to achieve a Root Mean Square Error (RMSE) of 0.8 and a Mean Absolute Error (MAE) of 0.64 both of which are indicative of good and relevant recommendations on the part of Collaborative Filtering. The content-based filtering has also proved to suggest relevant recommendations based on tags and descriptions relevant to the user's interactions with the app. The application back and front end both have been quality tested for user experience, robustness, and seamless performance. According to closed alpha test feedback, the hybrid recommendation system has streamlined trip finding for the user and the issue of cold starts for new users is almost non-existent, "recommendations were organic similar to the likes of Netflix or YouTube". We believe these results to be a good omen for the realization of our hybrid recommendation system and app "Sherpa".

Table 1: Performance Metrics of the Hybrid Recommendation System

Metric	Value
RMSE	0.8
MAE	0.64

Discussion:

The heart of Sherpa's personalized experience lies in its hybrid recommender system. The content-based component utilizes natural language processing to analyze tour descriptions and reviews, creating user and item profiles. The collaborative component leverages user-item interactions to understand and predict preferences. The integration of these methods allows

Sherpa to deliver highly relevant guide recommendations that resonate with individual user interests. The Hybrid Recommender Sherpa engine is trained on a curated dataset, ensuring diversity in user preferences and guide offerings. The iterative development of Sherpa incorporates user feedback at every stage, aligning with Agile methodologies. The repetitive cycle guarantees continuous polishing and adjusting of the application to satisfy user anticipations. The engine behind, fueled by Node.js alongside Express.js, manages data handling and API oversight, whereas the front part, shaped using React, provides an easy-to-use and swift interaction. The RESTful APIs developed in the stage of execution permit effective dialogue between the smartphone app and the server backend.

Our dataset often appeared as a sparse matrix, making it difficult to derive meaningful insights or predictions. To address this, we employed Singular Value Decomposition (SVD), which helped reduce the dimensionality of the dataset and extract underlying patterns. Additionally, we tackled the cold start problem by constructing cosine similarity matrices, enabling us to measure the similarity between users or items based on their features or preferences. By implementing these solutions, we were able to enhance the effectiveness and robustness of our recommendation system, ensuring better recommendations even in the face of sparse data and new user or item entries. The results from Sherpa's deployment indicate a successful integration of AI within a mobile tourism application.

Conclusion:

Utilizing SVD alongside the Matrix of Cosine Similarity for crafting uniquely tailored journey suggestions from data on TripAdvisor, Sherpa stands as a pioneer in the evolving era of navigated travel. Prospective endeavors could focus on advanced processing of natural language for an enriched comprehension of inclinations, integration of data instantaneously, and betterment through feedback from users. Sherpa will become the leader in personalized travel technology by including augmented reality previews, boosting sustainability, expanding suggestions to encompass more travel-related topics, and optimizing scalability. This ground-breaking method promises to revolutionize travel by providing unmatched customization and paving the way for further developments in the travel industry.

Acknowledgments:

Our profound appreciation goes out to the instructors, fellow students, and mentors who were instrumental in making the Sherpa project a reality. Their leadership, encouragement, and contagious excitement helped to transform this project into a worthwhile and prosperous endeavor. We extend a special thanks to Trip Advisor for the dataset that made the conception of our Hybrid Recommendation System possible.

References:

[1] S. Zhang, W. Wang, J. Ford, and F. Makedon, “Learning from incomplete ratings using non-negative matrix factorization,” Proceedings, vol. 2006, pp. 549–553, 2006, doi: 10.1137/1.9781611972764.58.

[2] D. D. Lee and H. S. Seung, “Algorithms for Non-negative Matrix Factorization”.

[3] R. Salakhutdinov and A. Mnih, “Probabilistic Matrix Factorization”.

[4] Y. Koren, R. Bell, and C. Volinsky, “Matrix factorization techniques for recommender systems,” Computer (Long. Beach. Calif)., vol. 42, no. 8, pp. 30–37, 2009, doi: 10.1109/MC.2009.263.

[5] W. Zhang, G. Ding, L. Chen, and C. Li, “Augmenting Chinese online video recommendations by using virtual ratings predicted by review sentiment classification,” Proc. - IEEE Int. Conf. Data Mining, ICDM, pp. 1143–1150, 2010, doi: 10.1109/ICDMW.2010.27.

[6] Q. Gu, J. Zhou, and C. Ding, “Collaborative filtering: Weighted nonnegative matrix factorization incorporating user and item graphs,” Proceedings, pp. 199–210, 2010, doi:

10.1137/1.9781611972801.18.
[7] H. Wen, G. Ding, C. Liu, and J. Wang, "Matrix Factorization Meets Cosine Similarity: Addressing Sparsity Problem in Collaborative Filtering Recommender System," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8709 LNCS, pp. 306–317, 2014, doi: 10.1007/978-3-319-11116-2_27.
[8] S. Zhang, W. Wang, J. Ford, F. Makedon, and J. Pearlman, "Using singular value decomposition approximation for collaborative filtering," Proc. - Seventh IEEE Int. Conf. E-Commerce Technol. CEC 2005, vol. 2005, pp. 257–265, 2005, doi: 10.1109/ICECT.2005.102.
[9] "TripAdvisor Dataset." Accessed: May 05, 2024. [Online]. Available: https://www.researchgate.net/publication/308968574_TripAdvisor_Dataset