

# **RAWAH: An AI-based Tourism Recommendation System using Embedding Models**

**Ghala Aldeheem**

Imam Abdulrahman Bin Faisal  
University, Dammam, Saudi  
Arabia  
Department of Computer  
Information Systems.

**Lamees AlDaej**

Imam Abdulrahman Bin Faisal  
University, Dammam, Saudi  
Arabia  
Department of Computer  
Information Systems.

**Roqayah Bahubail**

Imam Abdulrahman Bin Faisal  
University, Dammam, Saudi  
Arabia  
Department of Computer  
Information Systems.

**Retaj Alahmari**

Imam Abdulrahman Bin Faisal University,  
Dammam, Saudi Arabia  
Department of Computer Information Systems.

**Ghala AlFawzan**

Imam Abdulrahman Bin Faisal University,  
Dammam, Saudi Arabia  
Department of Computer Information Systems

## **ABSTRACT**

This paper presents RAWAH, an AI-based mobile application which provides personalized tourism recommendations for Saudi Arabia. The system uses advanced embedding models to match individual preferences with appropriate travel destinations which helps to enhance user experience. The four embedding models tested in this study included OpenAI text-embedding-3-small and Sentence-BERT and BGE (bge-small-en-v1.5). The models were tested under the same conditions using real user preference data and place descriptions and their performance was measured using Precision@5 and Recall@5 and F1-score metrics. The results show that the SBERT (all-mpnet-base-v2) model achieved the best overall performance which showed a higher ability to capture semantic similarity and create precise recommendations. The selected model was integrated into the RAWAH application to improve its recommendation system. The proposed system creates intelligent tourism solutions through its combination of artificial intelligence and real-time data and user-centered design.

## **General Terms**

Artificial Intelligence, Recommendation System, Mobile Application, Tourism Systems

## **Keywords**

personalization, embeddings, BGE, SBERT, tourism recommendations

## **1. INTRODUCTION**

The tourism applications have become a critical tool for travelers due to the wide range of services they provide, including navigation, booking, and destination information. Popular tourism applications, such as Google Maps and TripAdvisor, offer useful services like maps, reviews, and ratings. However, existing applications provide services in a generalized manner without providing personalized recommendations based on user preferences.

Today, personalization has become an integral part of most modern technologies, especially in the tourism industry, where user preferences are highly diverse. However, most of the applications currently used do not utilize an intelligent

recommendation system or incorporate user information with location-based services in order to provide more appropriate suggestions. As a result, users often receive generic recommendations instead of personalized ones.

Therefore, to address this issue, the RAWAH mobile application is proposed, which is a smart tourism application providing personalized recommendations based on user preferences. Unlike traditional tourism applications, Rawah integrates semantic embedding models with real-time user preferences to generate context-aware personalized tourism recommendation. The application is developed using modern approaches, including the utilization of technologies such as Flutter for developing the front-end of the application and Firebase for real-time data management, along with interactive maps [1] and geolocation tools. One of the main features of the RAWAH system is an artificial intelligence model used to improve recommendation generation.

Four advanced embedding models are selected to generate personalized recommendations. The first one is OpenAI's text-embedding-3-small model. Another two models are Sentence-BERT models (all-MiniLM-L6-v2 and all-mpnet-base-v2). In addition, the fourth model is BGE (bge-small-en-v1.5). All four models are applied under the same conditions and compared based on their efficiency.

## **2. RELATED WORK**

Tourism recommendation systems have gained significant attention with the rapid growth of mobile applications and artificial intelligence. Popular tourism applications such as Google Maps and TripAdvisor provide essential services including navigation, reviews, and ratings. However, these platforms mainly offer generalized recommendations and do not fully adapt to individual user preferences.

Content-based recommender systems have been widely used to address personalization challenges in recommendation systems [2]. These systems analyze user interests and match them with item features to generate relevant suggestions. According to Lops et al. [3], content-based filtering is particularly effective in scenarios where user interaction data is limited, making it suitable for handling the cold-start problem.

Recent advancements in natural language processing have introduced embedding-based models that significantly enhance recommendation accuracy. Sentence-BERT, proposed by Reimers and Gurevych [4], enables semantic understanding by converting textual data into dense vector representations. This allows systems to capture contextual meaning rather than relying solely on keyword matching. Similarly, Devlin et al. [5] introduced BERT, which laid the foundation for many modern embedding techniques used in recommendation systems.

In addition, cosine similarity is commonly used to measure the relevance between embeddings and rank recommendations accordingly [6]. Other studies [4] have also explored machine learning techniques to improve personalization and recommendation accuracy in tourism recommender systems [7], demonstrating the effectiveness of neural approaches in capturing user preferences more accurately [8].

Furthermore, the concept of smart tourism has emerged, integrating intelligent technologies to enhance user experience and decision-making processes [4]. Despite these advancements, many existing tourism applications still lack the integration of real-time user preferences with intelligent recommendation engines.

This gap highlights the need for systems like RAWAH, which combine content-based filtering, embedding models, and real-time user data to deliver personalized tourism recommendations.

### **3. METHODOLOGY**

#### **3.1 Method Overview**

The AI-based recommendation feature of the RAWAH mobile application was supported by a content-based recommender system. The system uses semantic similarity methods to align user preferences with the appropriate travel destinations.

The information about user preferences and place is presented as text and converted into a form of vector embedding. Such embeddings are compared to find out the relevant recommendations for each user.

#### **3.2 Selected Models**

Four embedding models were tested in this research:

- OpenAI text-embedding-3-small
- SBERT all-MiniLM-L6-v2 [4]
- SBERT all-mpnet-base-v2 [4]
- BGE (bge-small-en-v1.5)

The choice of these models was to enable a fair comparison of the models in terms of their performance in semantic similarity tasks.

#### **3.3 Experimental Setup**

Only one experimental system was used to provide a uniform assessment to all the models. The data on user preferences were retrieved from Firebase Firestore [9] and converted into a single textual input. Moreover, the evaluation process used user feedback stored in the `liked_places` field to measure recommendation performance. Place data were collected from the `places` and `business_places` collections, where only approved records were included to ensure data quality and reliability.

The experimental dataset included 49 destination records and 56 users' preference records obtained from the Firebase Firestore dataset. To validate the proposed methods, 10 users were chosen to analyze. The performance of four embedding

models was analyzed, and recommendations for all users based on the four different embedding models were calculated, yielding 40 recommendation generation. The places were represented as structured textual data in terms of different features such as description, category, activity, budget, and city.

All the models were used to produce embeddings of user preferences and place data. Cosine similarity was used to measure the relevance between embeddings. The similarity scores were used to rank all places, and the Top-5 most relevant places were used as recommendations to each user.

### **3.4 Evaluation Metrics**

The models were assessed using standard evaluation metrics, including Precision@5, Recall@5, and F1-score [10]. These metrics were used to compare the effectiveness of the recommendation models.

## **4. IMPLEMENTATION**

The Rawah system is designed as an AI-based mobile application that is intended to offer personalized tourism recommendations in Saudi Arabia. It is composed of four main components: the mobile application, the backend service, the AI recommendation engine, and the admin web panel.

The mobile application is built using Flutter to support multiple platforms and provide a smooth user experience. Through the app the users can create accounts, choose their preferences in the AI planner page and view recommended places. In addition, businesses owners can submit their own places to be included in the system. The interfaces were designed to be simple and easy to navigate for nontechnical users without any confusion.

Firebase is used as the backend service [9]. It handles user authentication and stores data using Firestore. The system saves user preferences such as categories, budget, weather, and selected region, which are later used to generate recommendations.

The AI recommendation engine is developed using Python and FastAPI [11], and it is deployed in Hugging Face [12]. The system uses the SBERT (all-mpnet-base-v2) model [4] to convert both user preference and place information into an embedding. User preferences are first converted into a text and then into embedding vectors, while places data are also transformed into embedding based on multiple attributes such as description, category, activities, budget, and weather.

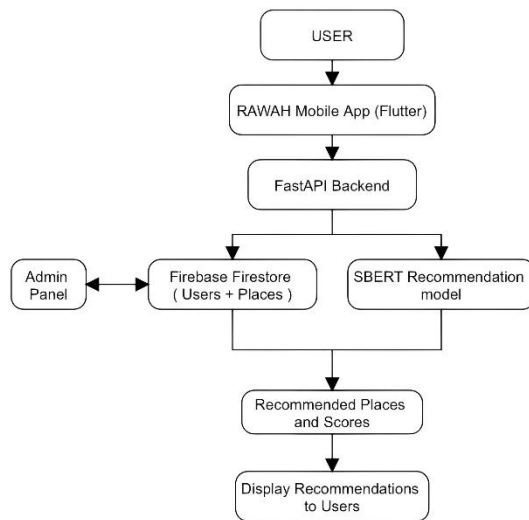
The integration between the mobile application and the AI engine is achieved through HTTP requests. The Flutter app sends user preferences to the API, which processes the data

and computes similarity scores using cosine similarity for comparing embeddings [6], and returns the top recommended places. The app then retrieves full place details from Firestore and displays them to the user.

The admin website panel was developed as a separate system connected to the application, which allows the admin to manage users and review business submissions, where the places must be approved before appearing to the users. The web panel was developed using Flutter Web and connected to the Firebase to allow real-time communication between these two components.

In such a way, the implementation of the software allows scalability, flexibility and interaction of user interfaces with the AI system in real time. This approach makes it possible to efficiently transfer requests and recommendations between the

Flutter app, Firebase server, and AI recommendation engine with minimal latency.



**Figure 1** proposed RAWAH Recommendation System Architecture

Figure 1 demonstrates the Rawah architecture and consists of the following elements: Flutter mobile application, FastAPI backend service, Firebase Firestore database, and SBERT recommendation model. The architecture demonstrates how user preferences are processed through the AI engine and integrated with the backend service to generate personalized tourism recommendations.

## 5. RESULTS

The recommendation system was evaluated by testing four different embedding models: OpenAI text-embedding-3-small, SBERT all-MiniLM-L6-v2, SBERT all-mpnet-base-v2, and BGE (bge-small-en-v1.5).

The system evaluation used user feedback through the liked\_places field, which represents the places that users liked. The system generated Top-5 recommendations for each user, which it used to evaluate their match with the users' liked places.

The system performance was measured using three evaluation metrics, which included Precision@5, Recall@5 and F1-score, which are commonly used for evaluating classification and recommendation systems [10]. The Precision@5 metric measures how many of the recommended places are relevant, while the Recall@5 measures how many relevant places were successfully retrieved. The F1-score provides a balance between precision and recall.

From the experiment findings, it became evident that the SBERT all-mpnet-base-v2 outperformed other tested models as it provided the highest Precision@5, Recall@5, and F1-Score. This shows that it is more accurate while providing higher recommendation coverage. Therefore, it can be concluded that this model has high potential for capturing semantic relations between users' preferences and the description of tourism destinations.

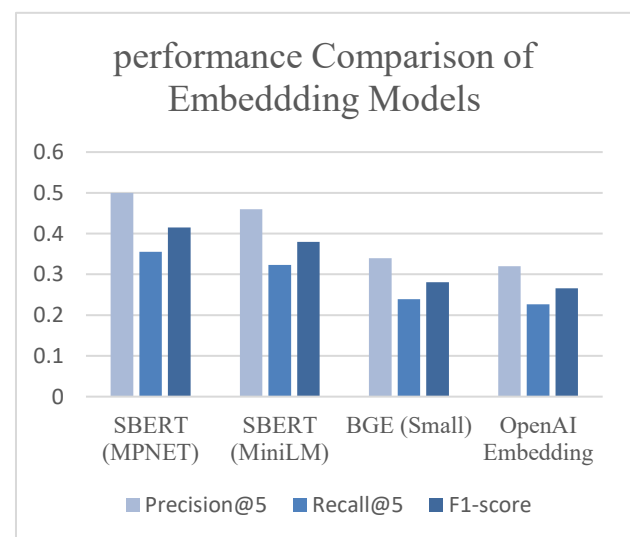
The MiniLM and BGE models were also quite effective. However, the ability to analyze context is not their strong side. At the same time, the OpenAI embedding model produced good results but was less efficient than SBERT.

From the above analysis, it can be concluded that the SBERT all-mpnet-base-v2 is the most suitable one for the RAWAH recommendation system. Such a choice was made based on semantic representation capability.

**Table 1** shows the comparison between the four models based on Precision@5, Recall@5, and F1-score.

Model	Precision@5	Recall@5	F1-score
SBERT (MPNET)	0.50	0.3554	0.4154
SBERT (MiniLM)	0.46	0.3232	0.3795
BGE (Small)	0.34	0.2393	0.2808
OpenAI Embedding	0.32	0.2268	0.2654

As shown in Table 1, the SBERT (MPNET) model achieved the highest score across all evaluation metrics, this indicates that the model was more effective in identifying relevant tourism destinations while maintaining a better balance between precision and recall. The SBERT (MiniLM) model produced competitive results but consistently scored slightly lower across all evaluation metrics. Moreover, the BGE and openAI Embedding models achieved lower performance and suggested weaker semantic matching capabilities for this recommendation task.



**Figure 2** Performance Comparison of Embedding Models Based on Precision@5, Recall@5, and F1-score

Figure 2 visually compares the performance of the evaluated models, clearly showing that the SBERT (MPNET) model outperforms the other models across all evaluation metrics.

## 6. DISCUSSION

This research's experimental results offer valuable insights into the suitability of transformer embeddings for local tourism recommendation systems. This section examines the meaning of the performance measures and how they translate to the RAWAH system.

### 6.1 Interpretation of Performance Metrics

The SBERT (all-mpnet-base-v2) model's high Precision@5, Recall@5, and F1-score indicate that dense vector embeddings are very effective at capturing the "vibe" and feel of tourist attractions. Experimentally, SBERT was statistically superior to all the evaluated models, such as MiniLM, BGE and OpenAI [4]. All-mpnet-base-v2's superior performance is due to its ability to learn subtle sentence meaning by looking at nuanced

sentence-level semantics, instead of strict, isolated token matches. The chosen model, unlike the traditional filtering method, did not have difficulty in understanding the semantics of the relationship between what the user was searching for "relaxation" and the description of the quiet beach. In the "vocabulary mismatch" problem, a traditional keyword or rule-based filter is unable to detect a recommendation because the exact terms of this recommendation are not found in the metadata of the location. Semantic embeddings are designed to address this challenge by associating the user preferences and destination profiles in a common vector space where they are semantically similar. The results indicate that by using the 768-dimensional embedding space provided by SBERT, the semantic gap was partially bridged, allowing capturing of richer relationships in the context, and increasing the recommendation accuracy and coverage in the ranking.

## 6.2 Open-Source vs. Proprietary Models

An important aspect for discussion is the comparison between the OpenAI proprietary model and the open-source SBERT model. OpenAI's text-embedding-3-small is a very powerful model, trained on large-scale, diverse data from across the world, but in this task of tourism recommendations, it failed to perform as well or capture the semantics as other models. These models have been trained on diverse tasks, including the generation of OpenAI's GPT series, and often make generalizations that highlight the overall meaning of the language, which may not capture the local details important for comprehending the attributes of the Saudi Arabian tourism sector (e.g., cultural events, cost of living, or weather conditions).

Additionally, the decision to utilize SBERT-MPNet [4] is driven by its structural need for an enclosed, protected, and sustainable design. The use of proprietary APIs, on the other hand, means that the system is always dependent on the external party, has varying financial operational expenses, and there are strict data privacy requirements for user logs. On the other hand, deploying the open-source SBERT model as a custom FastAPI backend on Hugging Face Spaces reduced the need for third-party service availability and recurring usage fees. It is also crucial for the architectural independence of building scalable, high throughput academic or startup initiatives, requiring little system latency when generating real-time recommendations.

## 6.3 Practical Implications for RAWAH

The practical analysis demonstrates the successful solving of the critical problem of "Cold Start", which is one of the most popular bottlenecks of traditional collaborative systems, by means of the semantic content based filtering method [3]. RAWAH moves away from historic user logs or large datasets of past behaviors to real-time inputs which are collected directly from the user's AI Planner page and used to calculate similarity vectors. For the end-user, this equates to giving new tourists or visitors extremely localized, context and person centered travel plans from the moment they interact with the app. The instant personalization is crucial for the development of the Saudi tourism market under Vision 2030 [13] and it addresses a wide spectrum of international and domestic tourists with personalized experiences. Combining these formidable embeddings with user-focused mobile design, RAWAH brings AI's power to users, offering an automated and fair platform for micro-businesses and local SMEs to enhance their online presence and conversion rates.

## 6.4 Limitations and Future Directions

Although the current system is very successful for the Eastern Province, it is acknowledged that augmenting the data set with more dialects and/or Arabic-inflected descriptions would improve accuracy. Future versions could integrate these successful embeddings with user input to dynamically refine the recommendation system.

## 7. CONCLUSION

This paper presented RAWAH, an AI-based mobile application designed to provide personalized tourism recommendations in Saudi Arabia. The system integrates a content-based recommendation approach with advanced embedding models to match user preferences with suitable destinations.

Four embedding models were evaluated, including OpenAI text-embedding-3-small, SBERT (all-MiniLM-L6-v2), SBERT (all-mpnet-base-v2), and BGE (bge-small-en-v1.5). The experimental results demonstrated that the SBERT (all-mpnet-base-v2) model achieved the best overall performance based on Precision@5, Recall@5, and F1-score metrics, indicating its strong ability to capture semantic relationships between user preferences and place descriptions.

The proposed system successfully addresses key challenges in tourism recommendation systems, including personalization and the cold-start problem, by relying on real-time user input rather than historical data. In addition, the integration of modern technologies such as Flutter, Firebase, and FastAPI ensures scalability and efficient system performance.

Overall, RAWAH contributes to the development of smart tourism solutions by combining artificial intelligence with user-centered design. Future improvements may include expanding the dataset, enhancing Arabic language support, and incorporating hybrid recommendation techniques to further improve accuracy and user satisfaction. The findings demonstrate that embedding-based recommendation systems can effectively enhance tourism personalization and support smart tourism initiatives in Saudi Arabia.

## 8. REFERENCES

- [1] O. Contributors, "OpenStreetMap," 2024. [Online]. Available: <https://www.openstreetmap.org/#map=5/24.15/45.08>. [Accessed 20 May 2025].
- [2] L. R. B. S. Francesco Ricci, Ed., *Recommender Systems Handbook*, 2 ed., Springer New York, NY, 2015.
- [3] P. Lops, M. D. Gemmis and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender systems handbook*, Springer, 2011.
- [4] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China, 2019.
- [5] J. Devlin, M. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of NAACL-HLT*, Minneapolis, MN, 2019.
- [6] H. Steck, C. Ekanadham and N. Ekallus, "Is Cosine-Similarity of Embeddings Really About Similarity?," in *Companion Proceedings of the ACM Web Conference*

- 2024 (WWW '24 Companion), Singapore, Singapore, 2024.
- [7] M. Nilashi, K. Bagherifard, M. Rahmani and V. Rafe, "A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques," *Computers & Industrial Engineering*, vol. 109, pp. 357-368, 2017.
- [8] X. Zhang, Y. Chen and Q. Li, "Deep Learning based Recommender Systems: A Survey and New Perspectives," in *ACM Computing Surveys*, 2019.
- [9] Google, "Firebase Documentation," Google, 2024. [Online]. Available: <https://firebase.google.com/docs>. [Accessed 20 January 2026].
- [10] R. Yacouby and D. Axman, "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models," in *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems (Eval4NLP)*, Online, 2020.
- [11] S. Ramírez, "FastAPI," 1 January 2024. [Online]. Available: <https://fastapi.tiangolo.com/>. [Accessed 2 April 2026].
- [12] T. e. al., "Transformers: State-of-the-Art Natural Language Processing," online, 2020.
- [13] U. Gretzel, M. Sigala, Z. Xiang and C. Koo, "Smart tourism: foundations and developments," *Electronic Markets*, vol. 25, no. 3, 2015.