Journal of
Theoretical and Applied Electronic Commerce Research

MDPI

Article

Emotion and Context-Aware Artificial Intelligence Recommendation for Urban Tourism

Mashael Aldayel [1], Abeer Al-Nafjan [2,*], Reman Alwadiee [1], Sarah Altammami [1], Abeer Alnafaei [1] and Leena Alzahrani [1]

[1] Information Technology Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[2] Computer Science Department, College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia
* Correspondence: annafjan@imamu.edu.sa

Abstract

The rapid growth of digital tourism platforms has intensified information overload and decision complexity for both locals and travelers, while operators struggle to differentiate their offerings and sustain profitable, data-driven e-commerce models. This paper presents Doroob, a big data and artificial intelligence (AI)-driven, context-aware recommendation system that integrates traditional recommender techniques with real-time facial emotion recognition (FER) to enable intelligent tourism commerce. Doroob combines three AI-based recommendation strategies: smart adaptive recommendation (SAR) collaborative filtering, a Vowpal Wabbit-based context-aware model, and a LightFM hybrid model. It trained on datasets built from the Google Places API and enriched with ratings adapted from MovieLens. FER, implemented with DeepFace and OpenCV, analyzes short video segments as users browse destination details, converts emotion scores into 1–5 satisfaction ratings, and stores this implicit feedback alongside explicit ratings to support adaptive, emotion-aware personalization. Experimental results show that the context-aware model achieves the strongest top-K ranking performance, the hybrid LightFM model yields the highest AUC of 0.95, and the SAR model provides the most accurate rating predictions, demonstrating that combining contextual modeling and FER-based implicit feedback can enhance personalization, mitigate cold-start, and support data-driven promotion of local tourist services in intelligent e-commerce ecosystems.

Keywords: facial emotion recognition; hybrid recommendation system; collaborative filtering; context-aware

Check for updates

Academic Editors: Chienliang Lin, Wuke Zhang, Sung-Wen Yu and Hanqing Xu

Received: 28 January 2026
Revised: 11 March 2026
Accepted: 18 March 2026
Published: 23 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The tourism and entertainment sectors have become major drivers of urban economic growth, contributing to job creation, infrastructure development, and increased regional and global competitiveness. Cities increasingly rely on tourism and entertainment to stimulate local spending and attract residents and international visitors, as destinations diversify their offerings through malls, cultural venues, events, theme parks, and leisure activities.

Simultaneously, the explosion of user-generated content (reviews, ratings, photos, and social media posts) has created information overload: tourists and locals struggle to find options that truly match their preferences, while marketers find it difficult to make their destinations stand out in crowded digital platforms. Saudi Arabia is heavily investing in

tourism and entertainment as strategic pillars for economic diversification and quality of life, especially in cities such as Riyadh. This rapid expansion underscores the need for intelligent, personalized recommendation tools that support national tourism objectives while helping users navigate growing choices.

Recommendation systems can assist users in effortlessly discovering places that align with their interests, filtering out irrelevant options, and marketing and popularizing the increasing tourist attractions in the country [1]. Moreover, the dependence on manual reviews on travel-related websites or social media posts can swamp tourists with conflicting reviews, positive and/or negative, about the same place. In this scenario, the dependence on manual reviews and an overwhelming amount of user-generated content becomes a serious obstacle to decision-making. Visitors may be swayed by subjective opinions in their attempt to obtain the best experiences, which may not necessarily align with their personal preferences or interests.

The time-consuming nature of this process not only diminishes the joy of planning a trip but can also lead to suboptimal choices, where the tourist ends up visiting places that do not fully match their expectations. Such a problem can exist with conventional recommendation approaches driven by user-generated reviews. In tourism, traditional recommender systems usually leverage explicit user ratings to recommend destinations. Hence, they can suffer from limitations like "cold-start" issues (lack of initial user data) as users usually neglect to provide explicit feedback, resulting in cold-start problems and scarce user behavior data.

Facial emotion recognition (FER) is an effective computing technique for generating implicit feedback [2,3], bypassing the need for explicit user ratings or reviews. Therefore, in this study, we proposed the integration of recommendation system techniques with FER. This will reduce dependence on manual reviews to a large extent, thus giving personalized recommendations that are compatible with the person's preferences without wasting hours scrolling through countless reviews.

Tourism, e-commerce, and consumer behavior theories emphasize that tourism experiences and digital journeys are shaped by user engagement, perceived value, and affective responses during service encounters, rather than solely by static preferences. Within these discussions, we position our system as a form of affective personalization in tourism e-commerce. In this framework, in situ facial expressions are treated as proximal indicators of momentary affect, which complement, rather than replace, explicit ratings and behavioral histories [2].

A key research gap identified in the literature is that existing tourism recommender systems either focus primarily on contextual cues such as location and time without adequately modeling affect, or they utilize affective information while failing to deeply integrate it into multi-modal recommendation pipelines. This oversight inhibits the potential for richer, more nuanced recommendation algorithms that can enhance user experiences by adapting to emotional states in real-time [2].

We proposed "Doroob", a personalized recommendation system for tourist destinations based on user emotions combined with contextual data. The system addresses the challenge of information overload and tourist decision-making by proposing a recommender system for tourism in Saudi Arabia. Doroob delivers user-specific suggestions tailored to individual preferences by using a hybrid, collaborative filtering (CF), and context-aware recommendation system that incorporates FER. The system is designed to optimize the tourism experience for residents and tourists by integrating explicit and implicit user feedback, facilitating the discovery of destinations that align with personal preferences. Doroob also plays a key role in marketing and promoting new and classic tourist destina-

tions and services and aims to boost the economy of Saudi Arabia by attracting more local or international tourists.

This study is organized as follows: Section 2 presents the necessary background, and Section 3 reviews the related literature. The materials and methods of the proposed system are described in Section 4, and its implementation is detailed in Section 5. Section 6 reports and discusses the result, and Section 7 concludes the study.

2. Background

This section provides essential background information for understanding the topic. The main recommendation system approaches and their types are first introduced, followed by an overview of FER technology and its benefits to the proposed system.

2.1. Recommendation System Approaches

Recommender systems are widely used to help users navigate large choice sets and can be broadly grouped into four main approaches: collaborative filtering (CF), content-based (CB), context-aware, and hybrid methods.

CF assumes that users with similar past preferences will also share similar future interests and generates recommendations based on multi-user behavior patterns [4]. It can be model-based, which learns a predictive model of user preferences, or memory-based, which relies directly on user–item interactions [5]. Memory-based methods are typically divided into user-based, which finds similar users and recommends what they liked [6], and item-based, which recommends items similar to those previously preferred by a user. However, when user behavior data are limited, CF suffers from the cold-start problem for new users/items and data sparsity [2,7].

CB assumes that users who liked items with certain attributes in the past will in the future prefer items with similar attributes [4]. It typically uses machine learning to compute similarity between item features that a user has interacted with and other items, before recommending those with the closest profiles [2]. However, relying solely on this approach can reduce diversity and limit serendipitous discovery because predefined item attributes and past behavior patterns constrict recommendations. Context-aware recommendation extends traditional methods by incorporating contextual information, such as location, time, and/or weather into the prediction process to provide recommendations that are better aligned with the user's current situation [2]. This can be performed through contextual pre-filtering (filtering data by context before applying the model), contextual modeling (embedding context directly into the model), or contextual post-filtering (adjusting recommendations after generation based on context) [7].

A hybrid recommender system combines two or more recommendation approaches to overcome the limitations of individual methods and improve accuracy [5]. Hybridization can be achieved in several ways, such as weighted combinations of models, switching between methods based on predefined conditions, mixing outputs from different recommenders, or using the model or features from one approach as input to another (feature augmentation, cascade, meta-level). Hybrid systems provide more flexible, accurate, and personalized recommendations than single-approach systems by integrating the strengths of multiple techniques and knowledge sources [8].

2.2. Facial Emotion Recognition

FER uses observable facial expressions to infer a user's emotional state and is considered one of the most practical emotion-sensing modalities because it is informative and non-intrusive compared with EEG or eye-tracking, which require wearable sensors [9]. FER

systems typically detect the face, extract facial landmarks, and classify expressions using deep learning models trained on labeled datasets of faces and corresponding emotions [10].

Emotion representation models are commonly categorical (e.g., Ekman's six basic emotions: anger, disgust, fear, happiness, sadness, and surprise) or dimensional, mapping emotions in a continuous valence–arousal space that captures positivity/negativity and activation level [11,12]. FER can operate on static images (frame-based) or dynamic sequences (video-based), with the latter leveraging spatiotemporal features to improve recognition accuracy [13].

FER can be integrated with recommendation systems to gain a deeper understanding of user preferences beyond explicit input. It recognizes and analyzes users' emotional states through physical measures to generate implicit feedback while they interact with the system content, thus inferring preferences from the first interaction [2]. This can provide a higher prediction rating and recommendation accuracy.

Camera-based emotion recognition in e-commerce raises significant privacy and ethical concerns that warrant careful consideration. Regulatory frameworks classify facial data as sensitive personal information requiring explicit consent and secure handling. These systems often capture data from numerous consumers simultaneously, potentially without their awareness, while struggling with data minimization requirements that create tension between technical needs and legal compliance [14]. Research examining user attitudes reveals that people view their emotions as "intimate, vulnerable, and complex", with many finding emotion recognition technologies "invasive and scary" due to perceived threats to autonomy and control [15].

Additional concerns include algorithmic bias, as these systems perform inconsistently across demographic groups with higher error rates for individuals with darker skin tones and older adults. The scientific validity of inferring specific emotions from facial expressions remains contested, with leading researchers challenging assumptions about universal expression–emotion relationships. User acceptance studies reveal that consumers are uncomfortable with camera access for non-essential tasks and expect transparency about data collection. Ethical implementation requires privacy-by-design approaches, on-device processing, and layered consent mechanisms that ensure users who decline camera access receive equivalent service quality [14].

2.3. Emotional Correlates of User Behavior and Decision-Making

Emotions play a central role in human behavior and decision-making. Psychological and consumer behavior theories indicate that individuals often rely on affect as a key source of information when evaluating options, forming judgments, and taking action. Transient emotional states influence what individuals focus on, how they perceive risk and value, and whether they choose to explore, defer, or finalize a decision. For instance, positive emotions such as interest and enjoyment are associated with greater approach behaviors and a higher willingness to choose and share experiences, while negative emotions can lead to avoidance, choice deferral, or reliance on simplified heuristics [16]. Since many of these affective reactions are expressed non-verbally through facial expressions, posture, and other observable cues, emotion-sensing technologies such as EEG-based emotion recognition and FER provide a means to capture implicit signals that may not be accurately reported through subjective self-reports and traditional explicit ratings [17].

The relationship between detected emotional responses and user preferences can be framed within the context of emotional intelligence, which posits that emotions significantly influence decision-making processes. As users engage with digital platforms, their emotional states shape their preferences, often leading them to favor options that resonate with their feelings. For instance, positive emotions may enhance receptivity to

novel experiences, while negative emotions can drive a preference for familiar options. Additionally, emotional responses serve as indicators of user satisfaction and engagement, revealing preferences that may not be explicitly articulated. By leveraging emotional data from technologies like camera-based emotion recognition, platforms can better tailor recommendations to align with users' affective states, ultimately enhancing user satisfaction and decision-making effectiveness [14,16].

In tourism and e-commerce, the role of emotion is particularly significant, as many decisions revolve around experiences. Studies on tourism experience and affective consumption demonstrate that emotions during the pre-consumption (search and planning), on-site, and post-consumption phases shape perceived value, satisfaction, memorability, and loyalty intentions [18]. On digital travel platforms, users encounter large volumes of user-generated content, images, and reviews, and their emotional reactions to this content influence their selection of destinations, the duration of their exploration of alternatives, and the options they ultimately choose to book. In e-commerce, research [16] highlights that interface design, visual stimuli, and media richness can invoke specific emotional responses that impact engagement, perceived trust, and conversion rates.

3. Literature Review

Personalized recommendations have become significant in almost every business, and the tourism industry is no exception. In recent years, this area has attracted substantial academic interest, particularly in tourism recommender systems that incorporate emotional and contextual information. Santamaria-Granados et al. [2] conducted a scientometric review of 1449 publications on emotion-aware tourist recommender systems published between 2001 and 2020; they mapped the field into four main axes (recommender systems, emotion recognition, wearable technology, and machine learning) and 16 finer categories such as sentiment analysis, CF, context, affective computing, and physiological signals, and they identified a strong post-2016 convergence around using affective signals from text and biosensors, together with advanced machine learning, to personalize tourism experiences and address data sparsity and cold-start issues. At the same time, their review highlights that most contributions remain technologically driven and calls for future work that more systematically integrates affective context, wearable/IoT data, and user-centric design guidelines into tourist recommender systems.

Oguntuase et al. [19] illustrate how context-aware tourism recommenders have moved beyond static profiles by explicitly modeling user mood, budget, companion, and distance as key determinants of place choice. Their personalized mobile POI recommender uses a hybrid approach in which Bayesian inference over contextual weights generates initial suggestions for new users, and user-based CF based on cosine similarity refines recommendations as rating data accumulate. The reported nDCG improvements (up to 83% versus 80% for traditional CF and 76% for CB methods) and high user satisfaction indicate that incorporating situational factors such as affective state and social setting can meaningfully enhance the perceived relevance and usability of recommendations.

Santamaria-Granados et al. [3] propose a tourist experience recommender system based on emotion recognition that utilizes physiological data from low-cost wearables to personalize recommendations according to the user's current affective state. The architecture integrates heart rate sensing and a mobile app for self-labeled emotions and activities, along with a pipeline to build an HR–emotion time-series dataset. Emotion-aware recommendations are generated by combining a CB filter with CF methods that utilize tourist experience categories, location, and textual descriptions. This approach achieves superior accuracy compared to traditional models, with evaluation metrics indicating improved performance.

Ghosh et al. [20] propose a mood-based personalized tourism recommendation system that uses sentiment analysis on user inputs and large-scale review data to infer travelers' current emotional state and then match them to suitable destination categories. Their model combines VADER-based sentiment classification, mood–destination mapping, and multi-criteria decision-making (including distance, budget, and accommodation) to generate itineraries optimized over multiple days, illustrating how affective signals and practical constraints jointly shape travel planning and satisfaction. This work underscores the growing recognition that tourism decisions are affect-laden and context-dependent, but it operationalizes mood through text sentiment and questionnaires rather than direct, in situ affective sensing.

Chanrueang et al. [19] develop a personalized hybrid tourist destination recommendation system that integrates emotional reactions and sentiment information from social media into a weighted combination of CF and CB. By extracting both sentiment polarity from comments and discrete emotions from reaction types (e.g., love, like, wow, sad, angry), and feeding these into SVD-based CF and SVM-based content models, their weighted hybrid framework achieves superior precision, recall, F1, and error metrics compared with single-method baselines, suggesting that emotion and sentiment are valuable behavioral signals for refining tourism recommendations.

Yang et al. [21] design a tourism recommendation system based on emotional analysis that uses a CNN to recognize users' facial expressions (positive vs. negative) from uploaded travel-related videos and then recommend destinations aligned with their prevailing mood. Their CNN-based FER model attains an emotion recognition accuracy above 80% and is used to select attractions from multiple provinces with the aim of alleviating negative affect and supporting psychological well-being in the post-pandemic context, highlighting the potential of real-time facial affect as a proximal indicator of tourism needs and preferences.

Table 1 below compares these studies in terms of their affective data sources, emotion representations, recommendation approaches, decision targets, and main findings.

Table 1. Comparative Analysis of Selected Studies on Emotion-Aware Tourism Recommender Systems.

Study	Data Source for Affect	Emotion Representation	Recommendation Approach	Target Decision	Key Findings and Limitations
Oguntuase et al. [19]	Explicit mood, budget, companion, distance	Self-reported mood; contextual variables	Hybrid: Bayesian context model + user-based CF	POI selection on mobile	Context (incl. mood) improves nDCG and satisfaction; mood not sensed in real time
Santamaria-Granados et al. [3]	Wearable HR, self-labeled emotions, activity logs	Valence–arousal quadrants from CNN–LSTM on HR data	Hybrid: CB + CNN-based CF	Tourist experiences ranking	Physiological affect improves accuracy. ER remains moderate under in-the-wild noise
Ghosh et al. [20]	Text reviews, ratings, mood questionnaire responses	Sentiment scores (VADER) and mood categories	Sentiment- and mood-driven multi-criteria model	Destination choice and daily itinerary planning	Mood and sentiment guide plans; relies on text and Q&A, not continuous sensing
Chanrueang et al. [19]	Social media comments and reaction emojis	Sentiment polarity; discrete emotions (love, sad…)	Weighted hybrid: SVD-based CF + SVM content model	Tourist attractions recommendation	Emotion/sentiment features boost precision and F1; emotions treated as static features
Yang et al. [21]	User-uploaded travel videos (facial images)	Binary facial emotion (positive/negative) via CNN	Emotion-triggered attraction recommendation	Province/attraction selection	FER-based mood (>80% accuracy) guides travel; only two emotion classes, short-term

These studies demonstrate a clear convergence around the notion that emotions and contextual factors (e.g., mood, social setting, popularity, and environmental conditions) significantly influence tourists' choices. However, they operationalize emotion in various ways, ranging from self-reported mood and sentiment scores to physiological data obtained from wearables and facial expressions. Furthermore, the table highlights unresolved issues, such as the reliance on static or self-reported affective features, limited emotional granularity (often restricted to binary or coarse categories), and a relatively weak integration with broader theories of tourism experience and digital consumer behavior. These gaps motivate our focus on real-time FER-based affective signals within a multi-modal recommendation architecture.

4. Materials and Methods

In the Doroob system, FER is integrated into the workflow during the consumption stage, capturing users' affective responses as they interact with the content [22]. Additionally, users may trigger emotional analysis by spending a certain amount of time viewing a specific destination [23]. By analyzing the user's implicit emotional responses, the system overcomes the limitations of traditional recommendation systems, such as the cold-start problem [5], which refers to the difficulty faced by traditional recommendation systems when they encounter a new user or item with little to no prior data.

In these situations, the system lacks sufficient information to make accurate recommendations, resulting in a suboptimal user experience. However, using implicit emotional responses would allow the system to learn users' preferences from their first interaction. If a user responds positively, for instance, smiling immediately after seeing some content, the system could use the emotional feedback obtained from the response to infer that other similar content could also be of interest. This enables the system to generate recommendations without an extended history of explicit user interaction. Thus, the accuracy and personalization of recommendations can be enhanced while reducing the time and effort required from users. In this study, we used CF, context-aware, and hybrid recommendation systems to overcome their limitations and improve system performance and quality.

The key innovation of this project lies in the integration of recommendation system techniques with real-time FER. The Doroob system requires the completion of the registration requirements at the beginning. Once the user is registered, they will be given the option to enable the device's camera and use the current location. The user is then directed to the homepage, where personalized destination recommendations are shown only if they have rated enough places.

Additionally, users can explore the places page, which contains a list of destinations retrieved from the Google Places API [24]. The best nearby destinations are identified using a context-aware approach if the system can detect the user's location. Concurrently, the user's device camera is activated to analyze the user's emotions and behavior while looking at the destinations. If a user is looking at a particular destination for a certain duration, FER will measure the satisfaction level on the user's face. Thus, the system generates a list of recommended destinations that meet the user's preferences (emotional state).

This integration leads to learning the user's interests from the first interaction and overcomes the cold-start problem. It also relies on the user's current emotional state along with historical data to provide a variety of suggestions. Moreover, the user's current emotional state will be used as an implicit rating to learn how users respond to different destinations positively or negatively without the need for explicit ratings. However, if the user does not allow the system to use the camera, the system can offer explicit ratings of destinations on a scale from 1 to 5.

4.1. Architectural Diagram

The Doroob system aims to provide personalized recommendations by relying on three main components: user feedback, a comprehensive database, and a recommendation system (Figure 1).

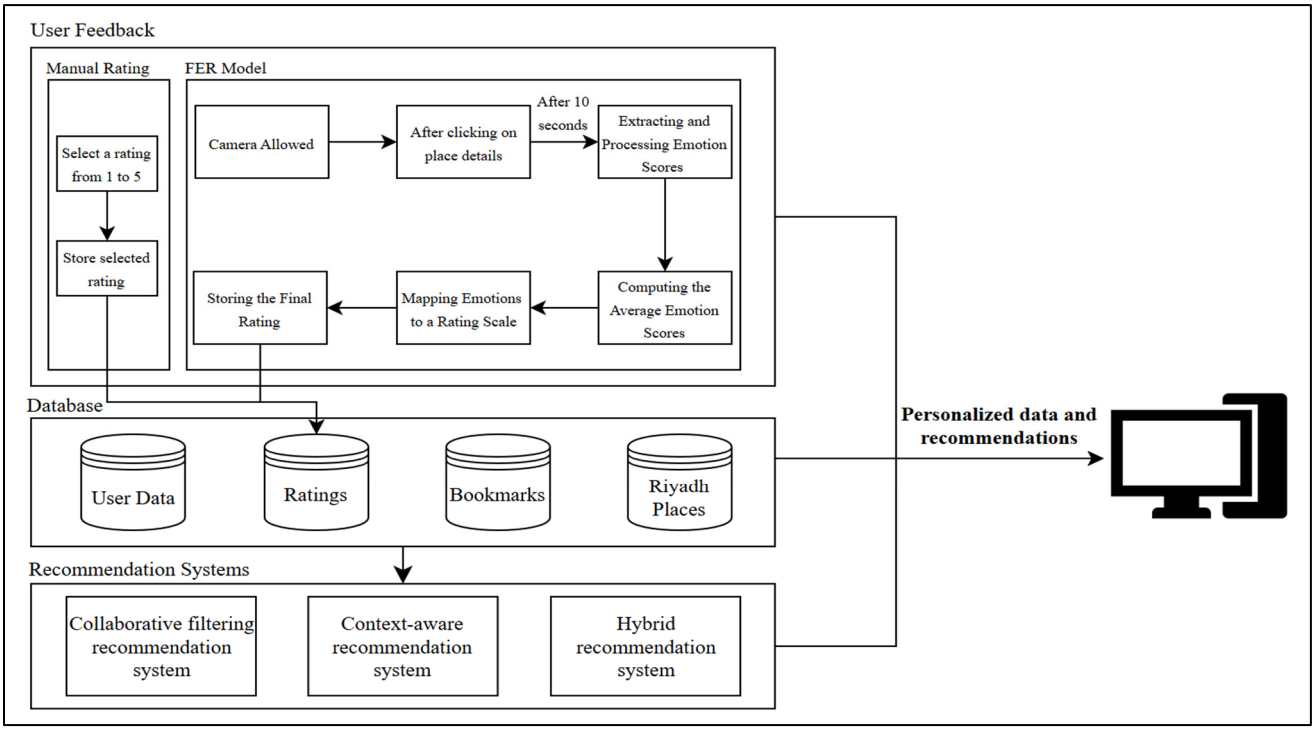


Figure 1. Doroob's System Architecture.

User feedback is collected using two primary methods: manual rating and the FER model. The manual rating method provides a straightforward approach, allowing users to assign a rating between 1 and 5 to a specific location, which is then stored in the database. In addition, the FER model offers an automated method for generating ratings by analyzing users' facial emotions. If camera access is granted, the FER process is activated after the user selects a place and spends 10 s viewing its details.

The 10 s duration for emotional data capture was selected based on prior affective computing research [25–27]. Spontaneous emotional expressions typically unfold within 0.5 to 4 s, and capturing data over a slightly longer window (10 s) ensures adequate coverage of onset, apex, and offset phases of emotional responses [25,26].

Once this period elapses, the system captures a live video feed of the user's facial emotions for an additional 10 s. Then, the model processes the extracted emotion frames, classifies them, and calculates an average emotion score. This score is mapped to a numerical scale ranging from 1 to 5. The final rating is stored in the database and used by the recommendation system to enhance personalized recommendations.

The Doroob database consists of key tables, including user data, ratings, bookmarks, and Riyadh Places extracted from the Google Places API. The recommendation system then uses these data to generate personalized recommendations. The system integrates three recommendation approaches: CF, which analyzes similar users' preferences; context-aware recommendations, which consider real-time contextual factors; and a hybrid recommendation model, which combines multiple techniques to improve accuracy. This integrated approach ensures that recommendations are tailored to user preferences and engagement with the system. Finally, the system delivers personalized data and recommendations back to the user, enhancing the overall user experience.

4.2. Recommendation System Algorithms

Doroob provides recommendations using different algorithms. Figure 2 illustrates the recommendation system flowchart.

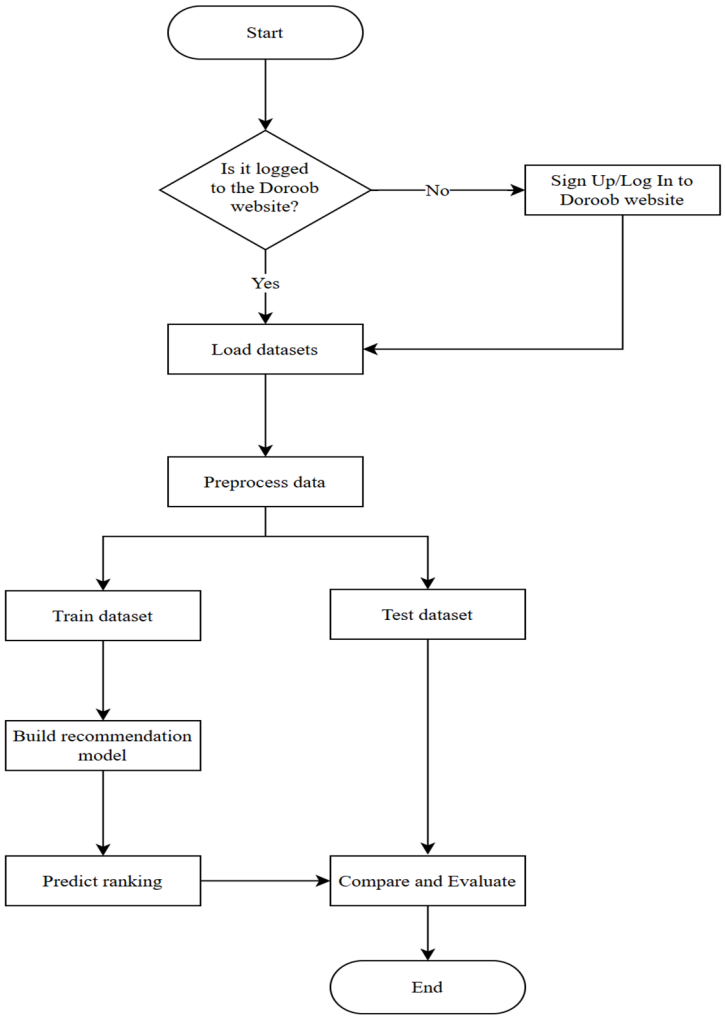


Figure 2. Recommendation System Flowchart.

The process begins by ensuring that the user is registered in the system, allowing the generation of recommendations. The model starts by uploading the data and preprocessing it. Next, the data are divided into training and testing sets. The sizes of these sets vary according to the algorithm and evaluation results. The recommendation model is then built and trained using the data. Once trained, predictions are made for the test data and compared with the actual values. Based on this comparison, the proposed model is evaluated.

4.3. Place Rating Using Facial Emotion Recognition

Figure 3 shows that the FER rating process begins when the user signs up or logs in to the Doroob system. After logging in, the system requests access to the camera. If permission is granted, the user can proceed to select a place to view its details. The system then implements a 10 s delay before capturing emotion-based feedback.

After this period, the system processes a live video feed, extracts emotion frames, and classifies them based on facial emotions. These classified frames are then used to compute the average emotion score. Subsequently, the system maps the computed emotion score to a 1–5 rating scale, aligning with the manual rating structure.

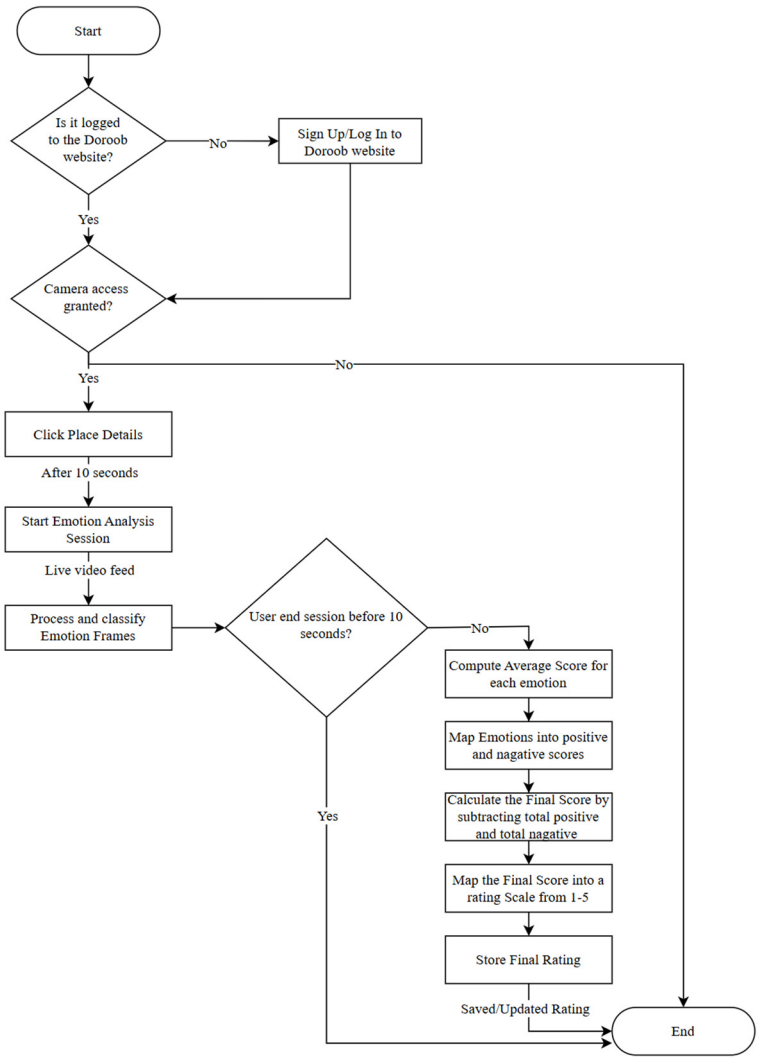


Figure 3. Facial Emotion Recognition Flowchart.

Since neutral scores do not indicate a positive or negative response, they were excluded from the final score calculations [28,29]. This decision is grounded in dimensional models of affect (Russell, 1980) and is further supported by research demonstrating that neutral expressions often introduce ambiguity in preference prediction tasks [30]. Neutral expressions are perceptually unstable and subject to systematic bias; for instance, cross-cultural studies have shown that supposedly 'neutral' faces are frequently perceived as negative [30]. Accordingly, focusing on positive and negative valence categories rather than including an ambiguous neutral category provides a more reliable foundation for assessing emotional responses to e-commerce stimuli.

The final rating value is stored in the database, which updates the user's preferences and contributes to the recommendation system. If camera access is denied, the user will be unable to rate using the FER technique and must rely on manual rating. This automated process enhances implicit feedback accuracy while reducing user effort.

4.4. Data Collection and Preparation

Data collection was essential for delivering personalized and relevant suggestions to users in developing our recommendation system. We used two datasets: one containing information about places in Riyadh and the other based on user ratings. We obtained the user ratings dataset from GroupLens Research, which was collected from the popular Movie-

Lens website. MovieLens [31] provides many datasets with different time periods and data sizes, and it is a widely used benchmark dataset in recommendation system research.

The dataset was collected during October 2024, and therefore, all records represent the attributes of the places as they were available during that collection period. The data collection process for places in Riyadh started with the identification and retrieval of POIs using the nearby search feature of the Google Places API. We developed a JavaScript-based script embedded within an HTML page to implement this, leveraging the Google Maps JavaScript API (version 3.62) and Places Service.

The script defined a comprehensive list of latitude and longitude coordinates representing diverse districts across Riyadh to ensure a geographically balanced dataset. Using these coordinates, the script conducted targeted searches programmatically for locations near each specified point. Nearby search requests were executed for each search query to extract essential details, including place_id, name, category (e.g., park, shopping mall, restaurant), rating, latitude, and longitude.

The script uses the PlacesService.nearbySearch function to query the API for locations within a defined geographic radius. Multiple geographic coordinates were selected across different districts of Riyadh to ensure broader spatial coverage of the city. For each coordinate, the script performed searches within a 5000 m radius to identify nearby places. A place was included in the dataset if it satisfied the following conditions: valid geographic coordinates, located within the specified search radius, within the geographic region of Riyadh, contained a valid place_id, which uniquely identifies the place in Google Places, and has categories relevant to urban activities and public locations such as restaurants, cafes, parks, or shopping malls. The dataset retrieved directly from the API was in JSON format (Figure 4). The results were processed and stored in a structured dataset containing the following attributes: place_id, name, category, rating, latitude and longitude.


```
[
  {
    "id": 1,
    "place_name": "Shawarma Alaz",
    "average_rating": 3,
    "granular_category": "restaurants",
    "lat": 24.6451181,
    "lng": 46.7176402,
    "place_id": "ChIJAUrpSq0FLz4RVPdqQsL3jH4"
  },
```


Figure 4. Sample of Collected Data in the JSON Format.

The collected data were then organized into a structured format suitable for subsequent analysis and integration into the recommendation system. This approach combined the dataset with real-time and contextually relevant information, ensuring that POIs aligned with user interests were broadly covered. The dataset was further filtered to retain only the relevant categories, including restaurants, cafes, parks, hotels, shopping malls, and art galleries. Some locations did not contain ratings, extracted from the Google Places API, which resulted in missing values. In such cases, the rating value was recorded as N/A to indicate that no rating information was available.

During the data preprocessing phase, comprehensive cleaning procedures were performed to ensure the dataset's integrity and reliability. Because the script performs searches from multiple geographic points with overlapping radius, the same location may appear multiple times in the API results. To address this issue, duplicate records were identified using the place_id attribute, and only one instance of each unique place was retained in the dataset. The critical fields were validated to confirm the completeness of the essential

attributes required for the recommendation system. Missing values, particularly in ratings, were addressed by assigning default values to maintain a structured and analyzable dataset (Table 2).

Table 2. Data After Preprocessing.

Num	Name	Rating	Category	Latitude	Longitude	ID
1	Enchanted forest in Riyadh	4.5	200	24.6772285	46.7364923	Chl65WYjewFLz4RAGwPO7bhMl
2	Uzzal Zoo	N/A	200	24.6732502	46.7099585	Chl544SSQAFLz4RTHkPEGV HKo
3	BLVD Forest	4.2	200	24.7785879	46.5996631	ChisYixtoPjLj4RPUN-4IBC_Q

The data processing tasks include filtering records and merging datasets into a unified format for analysis. These preprocessing steps ensured the development of a high-quality dataset, providing a robust foundation for the subsequent implementation of the tourism recommendation system.

We modified the rating data by remapping the movie IDs to match the IDs used in the Riyadh Places dataset to customize the MovieLens dataset for our project. Specifically, since the original MovieLens dataset contained movie IDs unrelated to our locations, we reassigned each movie ID to a place ID within the valid range of our dataset (1–8836) using a modulo operation. This approach ensured that every rating in the user dataset corresponded to an existing place in Riyadh, effectively aligning user interactions with specific locations. We were able to build recommendation systems based on users' simulated preferences for attractions and places in Riyadh, providing realistic and contextually relevant recommendations.

This study utilized a large-scale interaction dataset comprising 1,048,575 interactions from 6781 users on 8836 items. With 59,916,916 possible user–item pairs, the interaction matrix exhibits a global density of 1.75%, corresponding to a sparsity level of 98.25%—a characteristic consistent with real-world e-commerce platforms where users interact with only a small fraction of available items.

User interactions range from 20 to 3893 (mean = 154.6, median = 72.0), with 36% of users contributing fewer than 50 interactions. Item interactions range from 2 to 3576 (mean = 118.7, median = 26.0), with 44% of items in the "long tail" (fewer than 20 interactions). These statistics reveal a pronounced long-tail distribution: nearly half of all items reside in the "long tail" with minimal interaction history, while a small minority of popular items account for a disproportionate share of interactions. This sparse and skewed structure provides an ecologically valid testbed for evaluating emotion-enhanced recommendation algorithms under conditions representative of real-world tourism platforms, where personalized discovery of niche destinations and accommodations represents a key challenge and opportunity. Figure 5 presents the distribution of ratings across the 0.5–5.0 scale.

4.5. Model Evaluation

In this section, the performance of the different recommendation approaches implemented in the system is evaluated. The effectiveness of each algorithm was evaluated using several ranking and rating metrics. Performance is compared based on ranking and rating metrics. These metrics allow for a comprehensive assessment of the strengths and weaknesses of CF, context-aware, and hybrid recommendation systems [32]. Ranking metrics, such as precision and recall, were used to measure how well a recommendation system orders items based on user relevance.

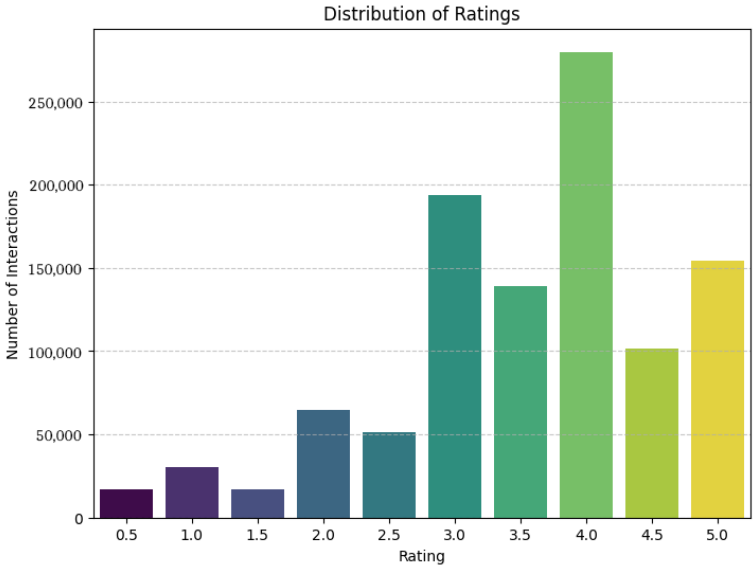


Figure 5. Ratings distribution of Collected Data.

Precision @ K: The proportion of the recommended items that are relevant within the top-K predictions is measured. Here, the Number of Relevant items in Top-k refers to the items in the top-K recommendations that are relevant to the user. Furthermore, k refers to the number of recommendations considered.

$$\text{Precision @ K} = \frac{\text{Number of Releveant items in Top}_\text{K}}{\text{k}} \tag{1}$$

Recall @ K: The proportion of relevant items successfully recommended in the top-K predictions. Where Number of Relevant Items in Top-k refers to relevant items in the top-K recommendations. The total number of relevant items refers to the total number of items relevant to the user in the dataset.

$$\text{Recall @ K} = \frac{\text{Number of Releveant items in Top}_\text{K}}{\text{Total Number of Relavent Items}} \tag{2}$$

AUC score (Area Under the Curve): Evaluates the item ranking, focusing on the distinction between relevant and irrelevant items. The number of correct pairwise comparisons refers to the number of instances where the positive class is ranked higher than the negative class by the model. Furthermore, the total number of pairwise comparisons refers to all possible pairs of instances, one positive and one negative, that the model evaluates.

$$\text{AUC} = \frac{\text{Number of Correct Pairwise Comaparisons}}{\text{Total Number of Pairwise Comaparisons}} \tag{3}$$

Normalized Discounted Cumulative Gain (NDCG): A widely used evaluation metric in information retrieval and recommendation systems to measure the quality of ranked lists. The relevance of the results in a ranked list is compared to the ideal order of relevance, with higher scores indicating better rankings. Here, rel_i refers to the item's relevance score at position i. Furthermore, i refers to the item's position in the ranking.

$$\text{DCG@K} = \sum_{\text{i=1}}^{\text{K}} \frac{rel_i}{\log_2(i+1)} \tag{4}$$

Rating metrics: The accuracy of predicted numerical ratings is measured against actual user ratings. Common rating metrics include the following:

The mean absolute error (MAE): The MAE calculates the summation of the difference between the user's and predicted ratings and divides it by the number of items considered. Here, r_i indicates the actual rating for item i, $\hat{r}_i$ indicates the predicted rating for item i. Furthermore, N indicates the total number of items.

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |r_i - \hat{r}_i| \quad (5)$$

Root Mean Square Error (RMSE): A metric used to measure the difference between the predicted and observed values in the regression and predictive models, where r_i indicates the actual rating for item i, $\hat{r}_i$ is the predicted rating for item i. Furthermore, N is the total number of items.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (r_i - \hat{r}_i)} \quad (6)$$

5. Implementation

The Doroob system was developed using multiple languages, including HTML, CSS, JavaScript, and PHP, for front- and back-end development and integrated with phpMyAdmin for real-time database management. Additionally, the Google Places API was used to connect to the Doroob interface, providing images of the displayed places. We also implemented three algorithms to recommend suitable places to users. To assess the effectiveness of the predictions, recommendation models were evaluated on both ranking and rating metrics.

5.1. Recommendation System Algorithms

Python (version 3.14) was used to program the recommendations and algorithms. The algorithms were adapted from the Recommenders repository [33]. Three algorithms were used to implement the recommendation system: Smart Adaptive Recommendation (SAR) CF algorithm, Vowpal Wabbit (VW) Context-Aware recommendation algorithm, and LightFM hybrid recommendation algorithm. These three recommendation algorithms were chosen not merely for their functional utility but because they represent distinct theoretical approaches to personalization, enabling us to investigate how emotional data might enhance different recommendation paradigms.

5.1.1. Collaborative Filtering Recommendation

Based on the CF algorithm, the Doroob system displays recommendations on the homepage and each place details page. The proposed model was implemented using the SAR algorithm [33]. SAR is based on the user's interaction history to identify similarities between items or users, proposing related items to those the user already prefers. SAR can examine whether emotion enhancement improves upon similarity-based personalization methods that dominate current e-commerce practice.

Depending on the user ID sent to the model upon logging in and navigating to the homepage, the recommendations will appear in the user interface, as shown in Figure 6.

5.1.2. Context-Aware Recommendation Approach

We used the VW model to implement a context-aware approach in Doroob. VW is a highly efficient open-source machine learning library designed for fast online learning. It supports several algorithms, including regression, classification, and contextual bandits, making it suitable for recommendation systems [34]. This algorithm enables investigation of whether emotional responses can serve as contextual signals guiding more efficient preference learning in dynamic tourism contexts where user preferences evolve based on changing travel goals or seasonal factors.

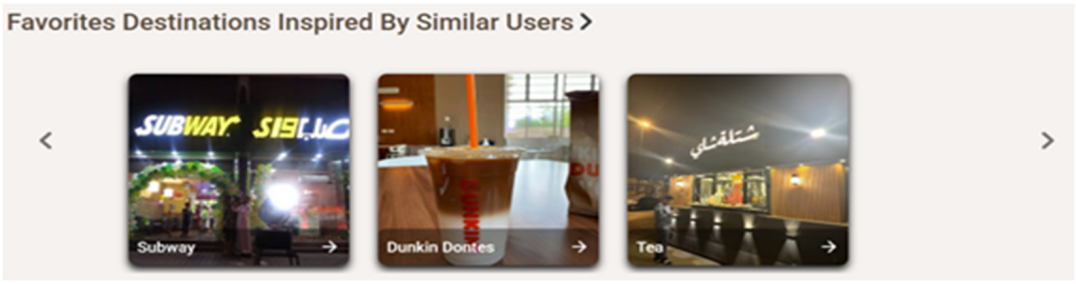


Figure 6. User View of Collaborating Filtering Recommendations.

In the Doroob system, user preferences are predicted by incorporating available contextual information into the learning process of the model through contextual modeling [7], following the equation [35]:

$$R : User \times Item \times Context \rightarrow Rating \quad (7)$$

The VW model combines user, item, and context features into a structured string of features. It learns the relationships among user preferences, item attributes, and contextual factors, enabling it to generate predicted ratings and provide accurate and personalized recommendations.

The context in VW was the distance calculated using the user's location and the coordinates of the place. This was included as a numerical feature (distance: <value>), dynamically adapting the model's predictions based on proximity. The user's location is dynamically saved using the "/api/save_location" endpoint. However, 0 is the default value of distance if the user does not allow accessing the location [34]. The recommendations appear in the user interface, as shown in Figure 7.

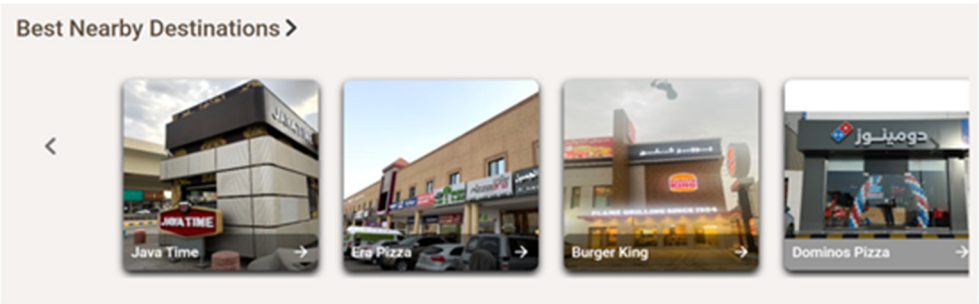


Figure 7. User View of the Context-Aware Recommendations.

5.1.3. Hybrid Recommendation Approach

A hybrid recommendation model was developed using the LightFM library, which integrates CF and CB recommendation techniques [36]. The LightFM library is a specialized tool for hybrid recommendation systems. It supports multiple loss functions, such as WARP, BPR, logistic, and warp-kos, enabling the effective integration of CF and CB approaches. The model leverages interaction data and optional metadata (item or user features). By incorporating emotional responses to visual content, LightFM enables examination of whether affective cues can address the "cold start" problem for new tourism products lacking sufficient rating history.

The recommendations appear in the user interface, as shown in Figure 8. This implementation ensures efficient handling of user requests and enables the web application to deliver personalized, context-aware suggestions to enhance user experience.

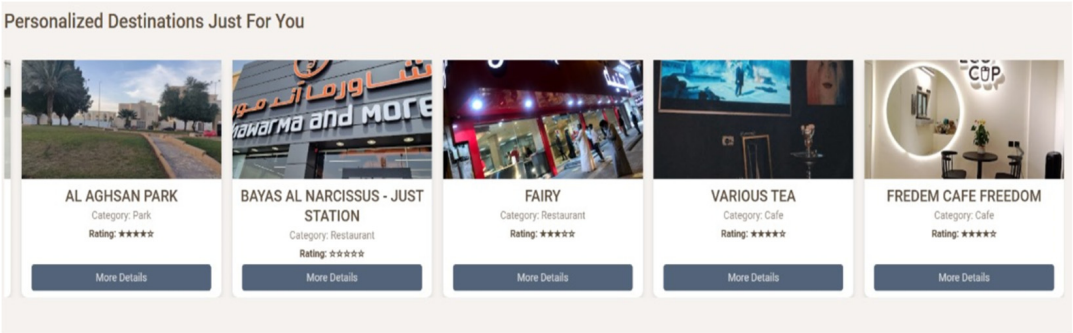


Figure 8. User View of the Hybrid Recommendations.

5.2. Facial Emotion Recognition

We implemented an emotion-based rating system using DeepFace, a lightweight Python framework designed for face recognition and facial attributes analysis, such as age, gender, emotion, and race [37]. We used it for facial recognition and emotion detection, allowing users to implicitly express their preferences. We integrated OpenCV for live video capture and DeepFace to achieve this. The process goes as follows:

1. Initiating Emotion Analysis: The emotion analysis session starts when a user interacts with a place. Then, the webcam is activated, and a live video feed begins capturing emotions. The session runs for 10 s to ensure sufficient data collection.
2. Extracting and Processing Emotion Scores: By using DeepFace.analyze, it analyzes multiple frames per second. Each frame was classified into seven emotions (Happy, Sad, Angry, Neutral, Surprise, Fear, and Disgust) with an assigned score. Figure 9 demonstrates the emotion scores during a 10 s period using MATLAB (version R2025a).

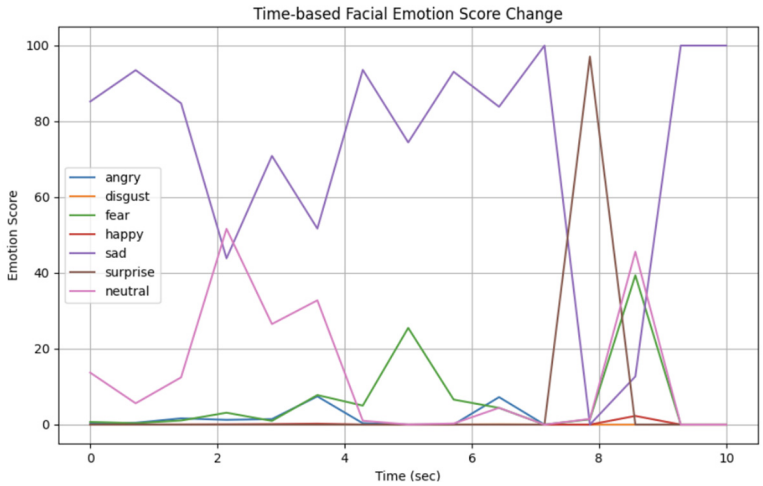


Figure 9. Emotion Score over 10 s.

Computing the average emotion scores: To ensure that the final rating reflects the overall sentiment of the user. Once all frames have been captured, the average score for each emotion is calculated using the following formula:

$$Average\ score\ = \frac{\sum Emotion\ Scores\ Across\ Frames}{Total\ Number\ of\ Frames} \quad (8)$$

3. Mapping Emotions to a Rating Scale: Emotions were classified into positive and negative scores.

$$Positive\ score = happy\ score + surprise\ score \quad (9)$$

$$Negative\ score = Angry\ score + Sad\ score + Fear\ score + Disgust\ score \quad (10)$$

The final response score is calculated by subtracting the negative emotions from the positive emotions. As the neutral score does not signify a positive or negative response, it was excluded from the final score calculations [28,29]. By using the following formula:

$$Final\ Score = Positive\ score\ - Negative\ score \quad (11)$$

The final response score is mapped to a 1–5 scale based on predefined thresholds determined through empirical analysis from previous research [28], where each rating level is assigned a weight of 40. The rating represents the satisfaction level of the user based on their emotional response (Table 3).

Table 3. Emotions Mapping.

Result of the Final Score	Satisfaction Level	Rating
Final Score ≥ 60	Highly Satisfied	5
$20 \leq$ Final Score < 60	Satisfied	4
$-20 \leq$ Final Score < 20	Moderately Satisfied	3
$-60 \leq$ Final Score < -20	Unsatisfied	2
Final Score < -60	Highly Unsatisfied	1

4. Storing the final rating: The system checks whether the user has previously rated the place. If a previous rating exists, it is updated. Otherwise, a new rating is inserted into the MySQL database Rating Table. In addition, the session ends in one of two cases: either when the system completes the 10 s duration and records the rating or when the user closes the place page before the 10 s duration is completed, causing the session to stop early without computing the emotion rating.

6. Result and Discussion

Table 4 demonstrates that the performance of the recommendation models varies across evaluation metrics.

Table 4. Results of Algorithm Evaluation.

Recommendation Approach	Top-K	Ranking Metrics				Rating Metrics	
		AUC	NDCG	Precision	Recall	RMSE	MAE
Doroob Models							
SAR (Collaborative)	3	54.94%	0.32	30.70%	5.95%	0.89	0.72
VW (Context)	5	/	1.05	85.80%	63.41%	1.02	0.81
LightFM (Hybrid)	5	95.88%	0.35	78.82%	5.40%	5.98	5.43
Baseline Comparison Models							
SVD (Collaborative)	10	73.56%	0.01	0.61%	1.87%	0.94	0.73
TF-IDF (Context)	5	/	67.3%	19.2%	0.9%	/	/

The SAR model demonstrates moderate performance across multiple evaluation metrics, underscoring its effectiveness in the Doroob system's CF. With an AUC of 54.94%,

SAR exhibits a moderate ability to rank relevant recommendations above non-relevant ones, while its NDCG of 0.32 indicates a meaningful ranking structure that prioritizes highly relevant items. The model achieves a precision of 30.70%, indicating that nearly one-third of the top-K recommendations are relevant, and a recall of 5.95%, indicating its ability to retrieve relevant items across the dataset. Additionally, SAR demonstrated the highest rating accuracy among the tested models, achieving the lowest RMSE (0.89) and MAE of 0.72, indicating its superior precision in predicting user preferences. SAR consistently outperforms the Singular Value Decomposition (SVD) model in ranking quality, recommendation relevance, and rating prediction accuracy, demonstrating its robustness as a reliable recommendation model for the Doroob system.

The VW model exhibits strong recommendation performance, achieving the highest NDCG of 1.05, indicating excellent ranking quality. Precision (85.80%) and recall (63.41%) demonstrate the strongest ability to retrieve relevant recommendations across a broad range of applications. However, the RMSE (1.02) and MAE (0.81) are slightly higher than those of SAR, indicating minor rating inaccuracies despite its superior ranking effectiveness. When compared to TF-IDF, VW outperforms TF-IDF across all ranking metrics. Both methods use TOP-K = 5, but VW achieves significantly higher NDCG (1.05 vs. 67.3%), precision (85.80% vs. 19.2%), and recall (63.41% vs. 0.9%), indicating better ranking performance. In addition, VW provides rating predictions, whereas TF-IDF lacks RMSE and MAE because it is a purely ranking-based approach without numerical rating predictions. These results suggest that VW is a more effective context-aware recommendation approach, excelling in ranking quality and rating prediction.

The LightFM model achieves the highest AUC of 95.88%, reflecting excellent ranking capability. The precision (78.82%) is notably high, but the recall (5.40%) suggests limited coverage, with fewer relevant recommendations being retrieved. The NDCG of 0.35 indicates moderate ranking effectiveness, while the RMSE (5.98) and MAE (5.43) are the highest among the models, suggesting that while LightFM excels in ranking, its rating predictions are significantly less accurate.

Each recommendation model is used based on its strengths in specific evaluation metrics to align with our objectives. SAR is leveraged for its superior rating accuracy, ensuring precise user preferences predictions. VW is employed for its high-ranking quality and broad recommendation coverage, making it highly effective in retrieving relevant content. LightFM, with the highest AUC, is integrated for strong ranking performance, ensuring that the most relevant recommendations are prioritized.

Our results clarify how emotion relates to decision-making in tourism e-commerce along several dimensions. Overall, they do not support viewing emotional responses as direct substitutes for explicit preferences; emotion-based ratings behave as additional, affective signals that are distinct from evaluative judgments and therefore complement, rather than replace, traditional rating-based inputs. This aligns with componential perspectives on emotion (e.g., Mauss and Robinson [16]) by showing that behavioral indicators such as facial expressions and experiential reports of preference have different predictive roles in recommendation models. At the same time, the contribution of emotion features is not uniform: they are most useful for items with sparse interaction histories, where standard collaborative filtering has limited evidence, which indicates that the value of affective information depends on factors such as item popularity, data sparsity, and decision context. Our findings further support the view that emotion and preference are related but distinct constructs, with emotion features showing modest yet systematic predictive power that varies across item types and user segments, implying partial rather than complete coupling. Finally, the observed patterns point to boundary conditions on the use of emotion recognition in commercial systems: emotion-based signals appear more beneficial for

exploratory, lower-stakes choices (e.g., discovering new destinations) than as the primary basis for high-stakes decisions, suggesting that affective computing is best deployed in a context-sensitive, selective manner within tourism e-commerce.

Despite these promising findings, several limitations of our evaluation design warrant consideration. The reliance on remapped MovieLens ratings to simulate user preferences in a tourism context raises questions about ecological validity, impacting the generalizability of our results. The MovieLens dataset originates from movie recommendations, where preference structures differ substantially from those in tourism decision-making. Our remapping approach assumes cross-domain transferability, which may limit its applicability to tourism contexts. The remapped ratings represent simulated rather than genuine tourism preferences, suggesting that the emotion-preference relationships observed may differ from those in real travel planning or post-stay evaluations.

7. Conclusions

Doroob is a context-aware tourism recommender that combines CF, contextual signals, and FER to provide personalized destination suggestions while mitigating cold-start and overreliance on textual reviews. The proposed system integrates SAR-based CF, a visit-weighted context model, and a hybrid LightFM model over a Google Places–based Riyadh dataset, using FER-derived implicit ratings alongside explicit ratings to learn preferences from the first interactions and adapt recommendations to the user's current emotional state. Experiments show that SAR offers the best rating accuracy (lowest RMSE and MAE) with solid ranking quality, the context-aware model achieves high precision and recall at top-K, and further tuning is needed for the hybrid model, indicating that emotion-aware implicit feedback combined with CF and context modeling meaningfully improves tourism recommendations.

Despite its contributions, the study has limitations. The reliance on remapped MovieLens ratings raises ecological validity concerns, as simulated preferences may not fully capture genuine tourism decision-making. Additionally, the observed modest convergence between behavioral measures (facial expressions) and experiential states suggests that affective signals may function primarily as a complementary layer to collaborative filtering. Contextual concerns, such as user acceptance of emotion recognition technologies, should also be considered. Future research should focus on cross-cultural evaluations, authentic tourism stimuli, privacy considerations, and longitudinal studies to develop resilient, ethical AI frameworks that can be applied across various e-commerce contexts.

From an intelligent e-commerce management perspective, these findings demonstrate how emotion-aware, context-aware recommendation can enhance user engagement, support the discovery and promotion of local tourist destinations, and enable data-driven decision-making for tourism providers operating in competitive digital marketplaces. Future work will explore richer contextual factors (e.g., temporal and social context), more granular affective modeling, and online testing with real users and real transactions, to build resilient, ethical AI frameworks that can generalize to other smart-city tourism platforms and broader e-commerce environments.

Author Contributions: R.A., S.A., A.A. and L.A. conceived, performed, analyzed, and interpreted the experiments and also drafted the manuscript. M.A. designed, analyzed, validated, reviewed, and edited the manuscript. M.A. and A.A.-N. supervised, reviewed, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This scientific paper was supported by a research grant from the Research, Development, and Innovation Authority (RDIA) of the Kingdom of Saudi Arabia (Grant number: 13461-imamu-2023-IMIU-R-3-1-HW-).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is available upon request.

Acknowledgments: The authors would like to thank the Research, Development, and Innovation Authority (RDIA) of the Kingdom of Saudi Arabia for funding this research through Grant No. (13461-imamu-2023-IMIU-R-3-1-HW-).

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Mufeed, S.A.; Gulzar, R. Tourism in Saudi Arabia. *Glob. Rev. Res. Tour. Hosp. Leis. Manag.* **2014**, *1*, 167–178.
2. Santamaria-Granados, L.; Mendoza-Moreno, J.F.; Ramirez-Gonzalez, G. Tourist Recommender Systems Based on Emotion Recognition—A Scientometric Review. *Future Internet* **2021**, *13*, 2. [CrossRef]
3. Santamaria-Granados, L.; Mendoza-Moreno, J.F.; Chantre-Astaiza, A.; Munoz-Organero, M.; Ramirez-Gonzalez, G. Tourist Experiences Recommender System Based on Emotion Recognition with Wearable Data. *Sensors* **2021**, *21*, 7854. [CrossRef]
4. Çano, E.; Morisio, M. Hybrid Recommender Systems: A Systematic Literature Review. *Intell. Data Anal.* **2017**, *21*, 1487–1524. [CrossRef]
5. Al-Bashiri, H.; Abdulgabber, M.A.; Romli, A.; Hujainah, F. Collaborative Filtering Recommender System: Overview and Challenges. *Adv. Sci. Lett.* **2017**, *23*, 9045–9049. [CrossRef]
6. AlSaeed, D. LOCUS: A Mobile Tourism Application and Recommender System for Personalized Places and Activities. *Informatica* **2023**, *47*, 201–212. [CrossRef]
7. Seyednezhad, S.M.M.; Cozart, K.N.; Bowllan, J.A.; Smith, A.O. A Review on Recommendation Systems: Context-Aware to Social-Based. *arXiv* **2018**, arXiv:1811.11866. [CrossRef]
8. Widayanti, R.; Chakim, M.H.R.; Lukita, C.; Rahardja, U.; Lutfiani, N. Improving Recommender Systems Using Hybrid Techniques of Collaborative Filtering and Content-Based Filtering. *J. Appl. Data Sci.* **2023**, *4*, 289–302. [CrossRef]
9. Ko, B.C. A Brief Review of Facial Emotion Recognition Based on Visual Information. *Sensors* **2018**, *18*, 401. [CrossRef]
10. JS, G.; Gleran Lobo, D.; Aman, M. Reading Faces, Recommending Choices: A Systematic Review of Facial Emotion Recognition and Recommendation Systems. *Int. J. Comput. Digit. Syst.* **2024**, *15*, 1–12.
11. Chutia, T.; Baruah, N. *A Review on Emotion Detection by Using Deep Learning Techniques*; Springer: Dordrecht, The Netherlands, 2024; Volume 57, ISBN 0123456789.
12. Ekman, P. Basic Emotion. In *Handbook of Cognition and Emotion*; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 1999.
13. Sailunaz, K.; Dhaliwal, M.; Rokne, J.; Alhajj, R. Emotion Detection from Text and Speech: A Survey. *Soc. Netw. Anal. Min.* **2018**, *8*, 28. [CrossRef]
14. Barrett, L.F.; Adolphs, R.; Marsella, S.; Martinez, A.M.; Pollak, S.D. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychol. Sci. Public Interest* **2019**, *20*, 1–68. [CrossRef] [PubMed]
15. Andalibi, N.; Buss, J. The Human in Emotion Recognition on Social Media: Attitudes, Outcomes, Risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2020; pp. 1–16.
16. Mauss, I.B.; Robinson, M.D. Measures of Emotion: A Review. *Cogn. Emot.* **2009**, *23*, 209–237. [CrossRef] [PubMed]
17. Scherer, K.R. What Are Emotions? And How Can They Be Measured? *Soc. Sci. Inf.* **2005**, *44*, 695–729. [CrossRef]
18. Strack, F.; Martin, L.L.; Stepper, S. Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis. *J. Pers. Soc. Psychol.* **1988**, *54*, 768–777. [CrossRef]
19. Oguntuase, R.A.; Gabriel, A.J.; Ojokoh, B.A. A Personalized Context-Aware Places of Interest Recommender System. *J. Comput. Theor. Appl.* **2025**, *2*, 481–497. [CrossRef]
20. Ghosh, P.; Kumar, A.; Sinha, P.; Neogy, S.; Das, S.; Ghosh, T.T.; Goto, T.; Sen, S. Mood-Based Personalized Tourism Recommendation System Using Sentiment Analysis. In *Proceedings of the 2025 IEEE/ACIS 29th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*; IEEE: New York, NY, USA, 2025; pp. 232–237.
21. Yang, W.; Wei, D.; Huang, W.; Huang, X. Research and Application of a Tourism Recommendation System Based on Emotional Analysis. In *Proceedings of the 2023 4th International Conference on Machine Learning and Computer Application*; ACM: New York, NY, USA, 2023; pp. 386–391.
22. Lin, S.Y.; Wu, C.M.; Chen, S.L.; Lin, T.L.; Tseng, Y.W. Continuous Facial Emotion Recognition Method Based on Deep Learning of Academic Emotions. *Sens. Mater.* **2020**, *32*, 3243–3259. [CrossRef]
23. Dhanith, J.; Venkatraman, S.; Narendra, M.; Sharma, V.; Malarvannan, S.; Gandomi, A.H. Multimodal Emotion Recognition Using Audio-Video Transformer Fusion with Cross Attention. *arXiv* **2024**, arXiv:2407.18552. [CrossRef]

24. Google. Google Places API. Available online: https://developers.google.com/maps/documentation/places/web-service/overview (accessed on 1 August 2024).
25. Hess, U.; Kleck, R.E. Differentiating Emotion Elicited and Deliberate Emotional Facial Expressions. In *What the Face RevealsBasic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*; Oxford University Press: Oxford, UK, 2005; pp. 271–287.
26. Cohn, J.F.; Kanade, T. Use of Automated Facial Image Analysis for Measurement of Emotion Expression. In *Handbook of Emotion Elicitation and Assessment*; Oxford University Press: New York, NY, USA, 2007; pp. 222–238, ISBN 019972766X, 9780199727667.
27. Tangnimitchok, S. Non-Intrusive Affective Assessment in the Circumplex Model from Pupil Diameter and Facial Expression Monitoring. Doctoral Dissertation, Florida International University, Miami, FL, USA, 2019.
28. Bandyopadhyay, S.; Thakur, S.S.; Mandal, J.K. Online Recommendation System Using Human Facial Expression Based Emotion Detection: A Proposed Method Human Facial Expression Based Emotion Detection: A Proposed Method. In *Proceedings of the International Conference on Advanced Computing Applications*; Springer: Singapore, 2022. [CrossRef]
29. Bandyopadhyay, S.; Thakur, S.S.; Mandal, J.K. Emotion Detection for Online Recommender System Using Deep Learning: A Proposed Method. *Innov. Syst. Softw. Eng.* **2022**, *20*, 719–726. [CrossRef]
30. Tae, J.; Nam, Y.; Lee, Y.; Weldon, R.B.; Sohn, M.-H. Neutral but Not in the Middle: Cross-Cultural Comparisons of Negative Bias of "Neutral" Emotional Stimuli. *Cogn. Emot.* **2020**, *34*, 1171–1182. [CrossRef]
31. Movielens. Available online: https://grouplens.org/datasets/movielens/ (accessed on 30 August 2024).
32. Vineela, A.; Devi, G.L.; Nelaturi, N.; Yadav, G.D. A Comprehensive Study and Evaluation of Recommender Systems. In *Microelectronics, Electromagnetics and Telecommunications: Proceedings of the Fifth ICMEET 2019*; Springer: Singapore, 2019; ISBN 9789811538285.
33. Smart Adaptive Recommendations (SAR)-Microsoft. Available online: https://github.com/Microsoft/Product-Recommendations/blob/master/doc/sar.md (accessed on 3 August 2024).
34. Vowpal Wabbit. Available online: https://github.com/VowpalWabbit/vowpal_wabbit (accessed on 3 August 2024).
35. Ishanka, U.A.P. A Context-Aware Travel Recommendation System Based on User Emotion and Personality. Ph.D. Thesis, Nagaoka University of Technology, Niigata, Japan, 2018.
36. Kula, M. Metadata Embeddings for User and Item Cold-Start Recommendations. *CEUR Workshop Proc.* **2015**, *1448*, 14–21.
37. DeepFace. Available online: https://pypi.org/project/deepface/ (accessed on 3 August 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.