Article

OurSCARA: Awareness-Based Recommendation Services for Sustainable Tourism

Luong Vuong Nguyen

Department of Artificial Intelligence, FPT University, Danang 550000, Vietnam; vuongnl31@fe.edu.vn or vuongnguyen@cau.ac.kr

Abstract: Sustainable tourism has emerged as a critical concern in contemporary society due to its potential to mitigate negative environmental and socio-cultural impacts associated with traditional tourism practices. In this context, recommendation systems (RS) are crucial in shaping travelers’ choices toward sustainable options. This research article proposes an innovative approach to RS tailored for sustainable tourism, termed Sustainability and Cultural Awareness-based Recommendation Algorithm (OurSCARA). OurSCARA integrates awareness of environmental and socio-cultural factors (sustainability attributes) into the recommendation process to facilitate informed decision-making by travelers. The system leverages data analytics techniques, including sentiment analysis, user profiling, and collaborative filtering (CF), to personalize recommendations based on users’ preferences, sustainability preferences, and contextual factors. Furthermore, OurSCARA incorporates real-time data sources such as weather conditions, local events, and community initiatives to enhance the relevance and timeliness of recommendations. A prototype implementation of OurSCARA is presented, along with a comprehensive evaluation framework to assess its effectiveness in promoting sustainable tourism behaviors. Through empirical evaluation using datasets collected from TripAdvisor, we demonstrate the potential of OurSCARA to influence traveler behavior towards more sustainable choices while enhancing their overall tourism experience. The findings underscore the significance of integrating sustainability considerations into RS and pave the way for future research and development in this emerging area at the intersection of computer science and sustainable tourism.

Keywords: recommendation system; sustainable tourism; collaborative filtering; OurSCARA

1. Introduction

Sustainable tourism has become increasingly important due to growing awareness of environmental conservation and socio-cultural preservation. Unlike conventional tourism, which often leads to over-exploitation of natural resources, degradation of ecosystems, and disruption of local communities, sustainable tourism aims to minimize negative impacts while maximizing benefits for both the environment and local communities. It emphasizes responsible travel practices that contribute to the well-being of destinations and promote long-term sustainability. Traditional tourism practices prioritize economic gains over environmental and socio-cultural considerations, leading to challenges and negative consequences. These include over-tourism, habitat destruction, cultural commodification, and loss of authenticity in tourist destinations. Moreover, the carbon footprint associated with travel contributes to climate change, further exacerbating environmental concerns.

In recent years, integrating recommendation system (RS) [1,2] into various domains has transformed how users discover and engage with products, services, and information. In tourism, RS have emerged as powerful tools to enhance user experiences, facilitate decision-making, and promote sustainable practices. By leveraging advanced algorithms and data analytics techniques, RS can offer personalized and contextually relevant suggestions tailored to travelers’ unique preferences and needs [3–5]. This study will explore how RS improve user experiences in tourism, focusing on attraction recommendations...
and ultimately enhancing sustainable tourism practices. RS are pivotal in improving user experiences by providing personalized recommendations for attractions, accommodation improvements, activities, and experiences [6]. By analyzing user preferences, historical behavior, and contextual factors such as location and time, RS can offer tailored suggestions that align with individual interests and preferences. This personalization saves users time and effort in planning their trips and enhances their overall satisfaction and enjoyment of the travel experience.

Consider a scenario where a traveler plans a trip to a new destination and seeks recommendations for attractions. Traditional methods of gathering information, such as guidebooks or online reviews, may be time-consuming and overwhelming due to the abundance of options available. However, with a recommendation system, travelers can receive personalized suggestions based on their interests, demographics, and past travel experiences. For instance, if a traveler is interested in outdoor activities and cultural heritage, the recommendation system may prioritize attractions such as national parks, historical landmarks, and local cultural festivals. Furthermore, the system can consider sustainability criteria, recommending eco-friendly attractions and community-based initiatives that promote responsible tourism practices. Travelers can discover hidden gems and off-the-beaten-path attractions that align with their interests and values by leveraging RS. This enriches their travel experiences and fosters a deeper connection with the destination and local communities. Figure 1 illustrates an example of RS using a combination of sustainability attributes and recommendations.

Figure 1. The illustration of tourism recommendation systems leverages context awareness of sustainability attributes.

This study aims to propose awareness-based recommendation services (called OurSCARA) to enhance the user experiences in attraction recommendation tasks. To do so, we proposed a framework incorporating sustainability and culture into attraction recommendations. By integrating sustainability criteria and cultural awareness into the recommendation process, our method enhances user experiences while promoting responsible travel practices. Unlike existing RS that primarily focus on user preferences and popularity metrics, our approach considers the environmental impact, cultural significance, and community engagement associated with tourist attractions. We deploy experiments on datasets collected from Tripadvisor to demonstrate the efficiency of the proposed method regarding personalized recommendations tasks. The contributions of this study are described as follows.

- **OurSCARA** prioritizes attractions that match users’ preferences and align with sustainability principles and cultural authenticity. This can enrich users’ travel experiences by offering personalized and contextually relevant suggestions while fostering a deeper connection with the destination and its local communities.
- Through the integration of sustainability criteria, OurSCARA encourages travelers to choose attractions that minimize environmental impact, support local conservation efforts, and contribute to the socio-economic development of host communities.
OurSCARA extends traditional recommendation algorithms to incorporate sustainability and cultural awareness metrics to help OurSCARA evolve beyond personalized preferences to address broader societal and environmental concerns.

The remainder of this manuscript is organized as follows. In the next section, Section 2, we overview the context-aware recommendation system in tourism and relevant studies using these systems to ensure sustainable tourism. Section 3 describes the proposed method, OurSCARA, which infuses sustainability and cultural awareness into recommendations. The experimental results and evaluation are provided in Section 4. Then, we discussed and gave a conclusion in Sections 5 and 6, respectively.

2. Related Work

The realm of tourism RS has been a focal point for researchers seeking to enhance the travel experience while simultaneously fostering sustainable tourism practices [7,8]. Many studies have investigated various approaches to developing intelligent RS tailored explicitly for the tourism domain, encompassing diverse methodologies and technologies. This section delves into the intricacies of existing research, providing insights into the evolution of RS in tourism and highlighting their contributions to sustainable tourism efforts [9]. Hamid et al. conducted an exhaustive systematic review of innovative tourism RS, which focused primarily on applying sophisticated data management strategies to elevate the intelligence of e-tourism platforms [10]. Their review underscores the critical role of advanced data analytics and machine learning techniques in furnishing personalized recommendations for travelers navigating the digital tourism landscape. By synthesizing existing knowledge in the field, Hamid et al. elucidated key avenues for enhancing the efficiency and efficacy of innovative tourism RS catering to modern-day travelers’ evolving needs and preferences.

Borràs et al. contributed significantly to understanding intelligent tourism RS through their comprehensive survey, encompassing various approaches and technologies employed within the field [11]. Their survey shed light on the multifaceted challenges and opportunities inherent in designing effective RS tailored to assist travelers in discovering relevant destinations, accommodations, and activities. Furthermore, Gavalas et al. delved into mobile RS, emphasizing the pivotal role of context awareness and user preferences in delivering tailored recommendations [12]. Their work highlighted the importance of leveraging contextual information, such as location and user behavior, to enhance recommendation accuracy and relevance in the mobile tourism domain.

Additionally, Kbaier et al. proposed a personalized hybrid recommendation system, which amalgamated various recommendation techniques to enhance recommendation accuracy and relevance [13]. Their approach underscored the importance of leveraging diverse data sources and algorithms to cater to travelers’ heterogeneous preferences and needs. By synthesizing disparate recommendation methodologies, Kbaier et al. offered a holistic framework for developing robust and adaptive RS capable of addressing the multifaceted requirements of the tourism domain. Nguyen et al. (2020) presented OurPlaces, a cross-cultural crowdsourcing platform designed to facilitate location recommendation services [2]. Their platform harnesses the collective intelligence of diverse user communities to offer location recommendations that transcend cultural boundaries. By leveraging the crowd’s wisdom, OurPlaces provides travelers with enriched and culturally diverse recommendations, enhancing their overall tourism experience. Yochum et al. (2020) surveyed the utilization of linked open data in location-based RS within the tourism domain [14]. Their study underscores the significance of leveraging linked open data repositories to enrich RS with comprehensive and up-to-date information about tourist destinations, attractions, and events. By integrating linked open data, RS can offer users more accurate and contextually relevant suggestions, enhancing user satisfaction and engagement. Nguyen et al. (2021) proposed a novel tourism recommender system based on cognitive similarity between cross-cultural users [4]. Their system enables more personalized and culturally sensitive recommendations by considering cognitive aspects such as perception, preferences, and
decision-making processes. By recognizing and accommodating cognitive differences across diverse user groups, the proposed system enhances the effectiveness and inclusivity of tourism recommendations, catering to each user demographic’s unique needs and preferences.

Context-aware RS have ushered in a new era in tourism, wherein personalized and timely recommendations are delivered based on contextual factors such as location, time, and user preferences. Abowd et al. laid the groundwork for context-aware computing, introducing fundamental concepts and principles that underpin context-aware RS [15–19]. Their seminal work paved the way for subsequent research endeavors to leverage contextual information to enhance the relevance and effectiveness of tourism recommendations. Subsequent studies by Barranco et al., Meehan et al., and Ashley-Dejo et al. delved deeper into the realm of context-aware mobile RS, elucidating the role of location and trajectory data in enhancing recommendation accuracy and relevance [5,20–22]. By leveraging spatial and temporal context, these studies demonstrated the potential of context-aware RS to deliver personalized and contextually relevant recommendations to travelers on the go.

Furthermore, Braunhofer et al. and Colomo-Palacios et al. explored proactive and social/context-aware recommendation approaches, emphasizing the integration of social context and user interactions into recommendation algorithms [23,24]. Their research shed light on the importance of considering social relationships and interactions in the recommendation process, enriching the user experience, and fostering community engagement within the tourism domain. Additionally, Bahramian et al. proposed a context-aware tourism recommender system based on a spreading activation method, which enhanced recommendation relevance by activating related concepts within the recommendation network [25]. Their approach leveraged semantic relationships between tourism entities to generate contextually relevant recommendations, thereby improving the overall quality and utility of the recommendation system. Esmaeili et al. introduced a pioneering tourism recommender system that integrates social commerce elements to enhance the recommendation process [26]. Their system offers travelers more personalized and trustworthy recommendations by leveraging social commerce mechanisms such as user reviews, ratings, and social network data. By incorporating social interactions and user-generated content, the recommender system enriches the travel planning experience, fostering community engagement and collaboration among travelers. Renjith et al. conducted an extensive study on the evolution of context-aware personalized travel RS, providing valuable insights into the advancements and challenges in the field [27]. Their research highlighted the dynamic nature of context-aware RS, emphasizing the need for continuous innovation and adaptation to meet travelers’ evolving needs and preferences. Moreover, Kulkarni and Rodd presented a comprehensive review of state-of-the-art context-aware recommendation techniques, offering critical insights into the latest developments and future directions in the field [28]. Their review synthesized existing knowledge and identified essential gaps in research and opportunities for further exploration, guiding future research endeavors to advance state-of-the-art context-aware RS. Furthermore, Yoon and Choi developed a real-time context-aware recommendation system for tourism, focusing on the timely delivery of personalized recommendations based on dynamic contextual information [29]. Their research demonstrated the potential of real-time RS to enhance the user experience and facilitate more informed decision-making by leveraging up-to-date contextual information.

3. **Ourscara: Awareness-Based Recommendation Services**

To infuse sustainability awareness into recommendations, we define a novel objective function that incorporates multiple factors related to environmental impact, socio-cultural preservation, and user preferences. Let \( R(u, i) \) represent the recommendation score for user \( u \) and item \( i \). The objective function can be formulated as follows.

\[
R(u, i) = \alpha \cdot P(u, i) + \beta \cdot E(i) + \gamma \cdot C(i) \tag{1}
\]
where $P(u,i)$ denotes the personalized preference score indicating the relevance of item $i$ to user $u$. $E(i)$ represents the environmental sustainability score associated with item $i$. $C(i)$ represents the cultural relevance score associated with item $i$. $\alpha$, $\beta$, and $\gamma$ are weighting factors controlling the importance of each component. The goal is to optimize the recommendation score $R(u,i)$ to reflect user preferences and sustainability considerations.

To integrate sustainability awareness into recommendations, we propose a hybrid recommendation approach called OurSCARA that combines CF with content-based filtering. Specifically, the proposed algorithm for OurSCARA is outlined in Algorithm 1.

**Algorithm 1** Sustainability and Cultural Awareness-based Recommendation Algorithm

1: procedure GENERATERECOMMENDATIONS(User $u$)
2: $U \leftarrow$ Retrieve user preferences and sustainability/cultural profile for $u$
3: $S \leftarrow$ Retrieve attraction attributes (e.g., environmental impact, cultural significance, season, temperature)
4: $R \leftarrow$ Initialize empty recommendation list
5: for each item $i$ do
6:     $P(u,i) \leftarrow$ Calculate personalized preference score based on user-item interactions
7:     $E(i) \leftarrow$ Calculate environmental sustainability score based on item attributes
8:     $C(i) \leftarrow$ Calculate cultural relevance score based on item attributes
9:     $R(u,i) \leftarrow \alpha \cdot P(u,i) + \beta \cdot E(i) + \gamma \cdot C(i)$
10:     Add $(i, R(u,i))$ to $R$
11: end for
12: Sort $R$ based on recommendation score in descending order
13: return Top five recommendations from $R$
14: end procedure

The proposed algorithms begin by initializing some variables. It retrieves the user preferences and sustainability profile for the given user $u$ and the attributes of attractions (represented as set $S$). These user preferences are represented as a vector $U$ retrieved from the system. Each vector element represents a user’s preference for a particular attribute or category relevant to the recommendations. For example, if the system recommends tourist attractions, elements of the vector might represent preferences for historical significance, natural beauty, and cultural experiences. For each item $i$ in the set of attractions $S$, the algorithm calculates three scores:

- **Personalized Preference Score ($P(u,i)$):** This score represents how much the user $u$ is likely to enjoy the attraction $i$ based on their past interactions. Here, we use collaborative filtering algorithms to predict the preference score, which is formulated as follows.

$$P(u,i) = \frac{\sum_{v \in N(u)} \text{sim}(u,v) \times r_{v,i}}{\sum_{v \in N(u)} |\text{sim}(u,v)|} \quad (2)$$

where $N(u)$ is the set of similar users to $u$, $\text{sim}(u,v)$ is the similarity between users $u$ and $v$, and $r_{v,i}$ is the rating of user $v$ for attraction $i$.

- **Environmental Sustainability Score ($E(i)$):** This score represents the environmental impact or sustainability of the attraction $i$. We consider carbon footprint, energy usage, and waste production associated with the attraction. Weights for each attribute of $E(i)$ score are determined through statistical analysis and tourist feedback. Specifically, we use Principal Component Analysis (PCA) to identify the most significant attributes and their relative importance based on variance explained in the dataset. Then, a regression analysis will be performed to understand the impact of each attribute on overall sustainability ratings provided by users or experts. Finally, we analyze the correlation between tourists’ sustainability ratings and specific attributes mentioned in reviews. Hence, the $E(i)$ can be formulated as follows.
\[ E(i) = \omega_c \times \Theta_c + \omega_e \times \Theta_e + \omega_w \times \Theta_w \]  

where, \( \omega_c, \omega_e, \omega_w \) are the weights assigned to each sustainability attribute. \( \Theta_c, \Theta_e, \) and \( \Theta_w \) denote the attribute carbon footprint, energy efficiency, and waste production, respectively. These attribute values are normalized scores representing the performance of the attraction on that attribute.

- **Cultural Relevance Score \( (C(i)) \):** This score represents the cultural significance or relevance of the attraction \( i \). This includes factors such as historical importance, cultural heritage value, popularity among locals or tourists, or recognition by relevant authorities or organizations. Similar to \( E(i) \), we calculated the \( C(i) \) by using PCA and regression analysis combine with tourist feedback. The formulation of \( C(i) \) is as follows

\[ C(i) = \delta \times \beta_h + \epsilon \times \beta_{ch} + \zeta \times \beta_{ci} \]  

where, \( \delta, \epsilon, \) and \( \zeta \) are the weights assigned to each attribute. \( \beta_h, \beta_{ch} \) and \( \beta_{ci} \) denote the attribute historical importance, cultural heritage value, and popularity, respectively. These attribute values are normalized scores representing the performance of the attraction on that attribute.

These scores are then combined using weights \( \alpha, \beta, \) and \( \gamma \) through Equation (1) to generate a recommendation score \( R(u, i) \) for each item \( i \) for the user \( u \). The top five recommendations with the highest scores are returned to the user.

### 4. Experiments

To evaluate the performance of the proposed recommendation system for attractions in the context of sustainable tourism, we will conduct experiments using data collected from TripAdvisor. The experiments will assess the system’s effectiveness in providing personalized recommendations that prioritize sustainability and cultural awareness while enhancing user experiences. The following outlines the experiment design:

#### 4.1. Data Collection

We crawl TripAdvisor to scrape rankings and reviews of Vietnamese tourist attractions. This popular online platform provides user-generated reviews, ratings, and recommendations for tourist attractions worldwide. The collected dataset includes information about tourist attractions, such as location, category (e.g., historical sites, natural parks), user ratings, reviews, and sustainability attributes (extracted from user feedback). Specifically, the dataset is collected from TripAdvisor by filtering “Things to Do” content in Vietnam with details such as “Points of Interest and landmarks”. Even though data is relevant to Vietnam, we decided to use the English version because the model to analyze user feedback works well with the English. To assess the weight score for sustainability attributes, we have three labels \([1,0,0.5]\) mean affected, neutral, not affected. Specifically, the method to extract features from user feedback (plain text) and assign scores weight to sustainability and cultural relevance involves two steps. First, we identify keywords or phrases indicating sustainability (e.g., “eco-friendly”, “green energy”, “low carbon”) or cultural (e.g., “heritage site”, “traditional festival”, “local culture”) cultural relevance. To do so, we use a predefined list of keywords or by training a model to recognize these terms. In addition, we consider combinations of words, known as bigrams and trigrams, which are pairs and triplets of words that might indicate context-specific relevance. In the second step, the model assigns scores to the extracted features based on their relevance to sustainability and cultural relevance.

- **1 (affected):** Features strongly indicate sustainability or cultural relevance.
- **0.5 (neutral):** Features that moderately indicate sustainability or cultural relevance.
- **0 (not affected):** Features that do not indicate sustainability or cultural relevance.
Consider an example, Hoi An Ancient Town, with the feedback, “Hoi An was a pleasant surprise for us. Beautiful city. Full of excitement. Until 9 pm, the old quarter is completely pedestrianized, with no cars or motorbikes. The weather is good and fresh air”. By analyzing this feedback, the sustainability attributes such as carbon footprint, energy, and temperature were assigned 1, 0, and 0.5, respectively. These attributes are extracted to be stored in our database and embedded in DataFrame during the process of training data later.

The statistics of the collected data are described in Table 1. In recap, in preparing for the training step, the collected data will undergo preprocessing to clean and filter irrelevant information. We then normalize ratings and reviews to ensure consistency across different attractions. Specifically, the attraction with review under 10 was removed. Sustainability attributes are extracted and encoded for each attraction.

<table>
<thead>
<tr>
<th>Title</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Number tourist attraction</td>
<td>318</td>
</tr>
<tr>
<td>2 Number of reviews</td>
<td>1218</td>
</tr>
<tr>
<td>3 Number of city</td>
<td>24</td>
</tr>
<tr>
<td>4 Number of feedback</td>
<td>439</td>
</tr>
<tr>
<td>5 Type of sustainability attributes</td>
<td>3</td>
</tr>
<tr>
<td>6 Median number of reviews per tourist attraction</td>
<td>30</td>
</tr>
</tbody>
</table>

The ground truth for evaluating OurSCARA was obtained by collecting user feedback and TripAdvisor ratings. Users were asked to rate various attractions based on their personal experiences, considering factors such as environmental sustainability and cultural significance. These ratings served as the benchmark for evaluating the accuracy of the recommendations generated by OurSCARA. Specifically, user preferences, sustainability scores, and cultural relevance scores from their feedback were used to form a labeled dataset. This dataset was then used to validate the performance of our recommendation algorithm by comparing its predictions against these absolute user-provided ratings.

4.2. Setup

In this study, we verified the performance of OurSCARA using the collected dataset by using Python 3.9 using an Intel Core i7 processor and an NVIDIA RTX3060Ti. Our experiments start by splitting the collected dataset into training (80%) and testing sets (20%). Therefore, 254 cases were used as the training set in the collected tourism dataset, and 64 corresponding cases were used as the testing set. The number of generated recommendations (Top-N) is fixed as ten attractions to all methods to compare the accuracy during the prediction task across the proposed method and the baselines.

The baselines utilized in this study encompass a diverse range of algorithms, each renowned for its unique approach to prediction tasks in RS. Firstly, the K-Nearest Neighbors (KNN) algorithm [30] operates on the principle of proximity, where instances are classified based on the labels of their nearest neighbors in the feature space. Secondly, the Support Vector Machine (SVM) algorithm, discussed in [31], seeks to find the optimal hyperplane that best separates data points of different classes within the feature space. Thirdly, the Random Forest algorithm, presented by [32], forms an ensemble of decision trees, each trained on a random subset of the data, and combines their predictions to achieve robust classification performance. Lastly, the XGBoost (R package version 0.4-2) algorithm, introduced by [33], is a gradient-boosting technique that sequentially builds a series of weak learners, with each subsequent learner aiming to correct the errors of its predecessors, thereby enhancing the overall predictive accuracy.

We evaluate algorithms on the testing sets using various evaluation metrics. In particular, accuracy metrics include Precision, Recall, and F1-score, which assess the relevance of recommended attractions compared to ground truth. The metrics are defined as follows.
Precision = \frac{TP}{TP + FP}, \quad (5)

where:
- TP (True Positives): Number of relevant attractions correctly recommended.
- FP (False Positives): Number of non-relevant attractions incorrectly recommended.

Recall = \frac{TP}{TP + FN}, \quad (6)

where:
- TP (True Positives): Number of relevant attractions correctly recommended.
- FN (False Negatives): Number of relevant attractions not recommended.

F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (7)

4.3. Experimental Results

The experimental results highlight the performance disparities between OurSCARA and traditional CF-based RS (KNN, SVM, Random Forest, XGBoost) in tourist prediction tasks and real-time contexts. OurSCARA exhibits significantly better precision, recall, and F1-score in both scenarios than its counterparts. This suggests that OurSCARA has potential advantages, such as adaptability and effectiveness, demonstrated by traditional RS.

Regarding the tourist prediction task, where models predict tourist preferences or behaviors based on historical data. As shown in Table 2, OurSCARA emerges as the top-performing model, followed by SVM and KNN. Conversely, Random Forest and XGBoost performance need to catch up, implying limitations in their ability to predict tourist preferences accurately in these models. These results showcase superior predictive capabilities, indicating OurSCARA's effectiveness in understanding complex patterns in tourist behavior.

Table 2. Comparison between OurSCARA and CF-based recommendation system (KNN, SVM, Random Forest, XGBoost) regarding tourist attraction prediction task.

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.408</td>
<td>0.424</td>
<td>0.415</td>
</tr>
<tr>
<td>Support Vector Machine (SVM)</td>
<td>0.438</td>
<td>0.398</td>
<td>0.417</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.295</td>
<td>0.228</td>
<td>0.257</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.255</td>
<td>0.239</td>
<td>0.247</td>
</tr>
<tr>
<td>OurSCARA</td>
<td>0.512</td>
<td>0.475</td>
<td>0.492</td>
</tr>
</tbody>
</table>

For the real-time context, this study considers real-time to mean monitoring user interaction and real-time feedback to improve recommendation quality continuously. The purpose is to generate recommendations in response to immediate user queries or interactions. We kept data in our data frame and added more by crawling new data from TripAdvisor (considered a new real-time interaction). Then, we re-ran the model and evaluation. In the results shown in Table 3, OurScara outperforms other models, closely followed by SVM, Random Forest, KNN, and XGBoost exhibit subpar performance, reinforcing their struggle to compete in real-time decision-making scenarios.
Table 3. Comparison between OurSCARA and CF-based recommendation system (KNN, SVM, Random Forest, XGBoost) according to real-time context.

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.277</td>
<td>0.215</td>
<td>0.242</td>
</tr>
<tr>
<td>Support Vector Machine (SVM)</td>
<td>0.733</td>
<td>0.548</td>
<td>0.627</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.339</td>
<td>0.246</td>
<td>0.285</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.258</td>
<td>0.182</td>
<td>0.213</td>
</tr>
<tr>
<td>OurSCARA</td>
<td>0.747</td>
<td>0.568</td>
<td>0.645</td>
</tr>
</tbody>
</table>

5. Discussion

Despite the performance differences, the ultimate goal of tourist attraction recommendations should extend beyond mere effectiveness to incorporate sustainability considerations. Sustainable tourism aims to minimize negative environmental, socio-cultural, and economic impacts while maximizing benefits for local communities and preserving natural and cultural heritage. Therefore, when recommending tourist attractions, it is essential to prioritize destinations and activities that align with sustainability principles.

- **Environmental Impact**: Recommendations should prioritize attractions with minimal ecological footprint and promote environmental stewardship. This could include nature-based activities such as hiking, wildlife viewing, or visiting protected areas like national parks and marine reserves. Additionally, eco-friendly accommodations and tours that adhere to sustainable practices, such as waste reduction, energy efficiency, and conservation education, should be highlighted. Models like OurSCARA, Random Forest, and XGBoost, which demonstrate robust predictive performance, can be leveraged to identify and recommend such environmentally responsible attractions.

- **Cultural Significance**: Cultural experiences play a vital role in tourism, offering opportunities for cross-cultural exchange and fostering mutual understanding. When recommending attractions, emphasis should be placed on sites of cultural significance, such as historical landmarks, museums, indigenous heritage sites, and traditional arts and crafts workshops. These attractions provide enriching experiences for tourists and contribute to preserving cultural heritage and identity. OurSCARA, which excels in capturing intricate patterns in data, can aid in recommending culturally immersive experiences tailored to individual preferences.

- **Community Involvement**: Sustainable tourism prioritizes the participation and empowerment of local communities, ensuring that tourism benefits are equitably distributed and contribute to community development. Recommendations should highlight community-based tourism initiatives, locally-owned businesses, and social enterprises that directly involve and benefit residents. This could include homestays, agritourism experiences, guided tours led by community members, and artisanal markets. Besides OurSCARA, models like Random Forest and KNN can also predict tourist preferences. They can be crucial in directing tourists toward community-driven attractions that promote socioeconomic inclusion and cultural authenticity.

While integrating sustainability attributes into tourist recommendations is essential for promoting responsible tourism practices, several challenges must be addressed. One challenge is the availability and quality of data on sustainability indicators, such as carbon footprint, cultural authenticity, and community engagement. Collecting and incorporating such data into recommendation algorithms requires collaboration between tourism stakeholders, researchers, and data scientists. Another challenge is the complexity of balancing multiple sustainability dimensions in recommendations. For example, a tourist attraction may have a low environmental impact but limited cultural authenticity, or vice versa. RS must consider trade-offs and prioritize attractions that offer a holistic sustainability experience. Despite these challenges, there are significant opportunities for innovation and collaboration in developing more sustainable RS. Advances in machine learning techniques, such as deep learning and reinforcement learning, can enhance the accuracy and
adaptability of recommendation algorithms. Additionally, partnerships between tourism industry players, technology companies, and sustainability organizations can facilitate the integration of sustainability metrics into recommendation platforms.

6. Conclusions

In this study, we have presented a comprehensive approach to enhancing user experiences in sustainable tourism by developing and evaluating a novel recommendation system for attractions. Leveraging data collected from TripAdvisor, we proposed the Sustainability and Cultural Awareness-based Recommendation Algorithm (OurSCARA), which integrates sustainability criteria and cultural awareness into the recommendation process. Through experiments and evaluations, we have demonstrated the effectiveness and significance of SCARA in promoting responsible travel practices and enriching user experiences. Our experiments showcased OurSCARA's superior performance compared to traditional baseline methods such as CF-based (using KNN, SVM, Random Forest, and XGBoost). OurSCARA exhibited higher accuracy in recommending attractions aligned with users’ preferences while prioritizing sustainability and cultural authenticity. Moreover, OurSCARA demonstrated superior diversity, fairness, and utility in its recommendations, catering to a wide range of user demographics and preferences. The evaluation results provided valuable insights into the strengths and limitations of OurSCARA, enabling us to identify areas for further optimization and improvement. For the future, we aim to continue refining and optimizing OurSCARA and envision its broader adoption across the tourism industry, ultimately contributing to preserving natural and cultural heritage and promoting sustainable tourism practices worldwide. For future work, we are planning to conduct experiments on other datasets to ensure the evaluate to demonstrate the improved performance of our proposed method.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This article includes the data presented in the study.

Acknowledgments: This research was supported by AIT Laboratory, FPT University, Danang Campus, Vietnam 2024.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

RS Recommendation Systems
CB Content Filtering
CF Collaborative Filtering
AI Artificial Intelligence

References


**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.