TourOptiGuide: A Hybrid and Personalized Tourism Recommendation System

Hilali Intissar
intissarhilali2017@gmail.com

Research Team in Intelligent Machines (RTIM)

Arfaoui Nouha
Research Team in Intelligent Machines (RTIM)

Ejbal Ridha
Research Team in Intelligent Machines (RTIM)

Research Article

**Keywords:** Tourism Recommender System, Fuzzy logic, Deep Learning, Trajectory Data Warehouse, Age estimation, POI detection.

**Posted Date:** May 8th, 2024

**DOI:** https://doi.org/10.21203/rs.3.rs-4356022/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:** No competing interests reported.
TourOptiGuide: A Hybrid and Personalized Tourism Recommendation System

Hilali Intissar, Arfaoui Nouha† and Ejbali Ridha†

*Research Team in Intelligent Machines (RTIM), National Engineering School of Gabes, University of Gabes, Gabes, 6029, Gabes, Tunisia.

*Corresponding author(s). E-mail(s): intissarhilali2017@gmail.com; Contributing authors: arfaoui.nouha@yahoo.fr; ridha_ejball@ieee.org;
†These authors contributed equally to this work.

Abstract
When visitors explore a city briefly, they must prioritize the key attractions that align with their interests. These significant points of interest (POIs) can be chosen based on specific criteria tailored to their needs. Additionally, travelers venturing into unfamiliar regions often seek help to plan their itinerary. To address this issue, we developed and presented a novel hybrid and personalized recommendation system aimed at helping tourists choose their next POI. The system tailors its suggestions based on four key factors: the tourist’s current location, single preferences, age, and historical experiences. Deep learning models play a crucial role in identifying the tourist’s current location from images and predicting age from selfies. In addition, our system leverages a trajectory data warehouse containing extensive historical data of past tourist’s experiences to provide suggestions. The core of our recommendation strategy is a fuzzy logic decision support system. This system effectively synthesizes diverse inputs to produce the top next POI to visit. By integrating various recommendation methods, our hybrid system significantly improves the precision and pertinence of its recommendations, offering a more customized and effective travel experience. Preliminary results demonstrate significant improvements in tourist satisfaction and in the efficiency of itinerary planning.

Keywords: Tourism Recommender System, Fuzzy logic, Deep Learning, Trajectory Data Warehouse, Age estimation, POI detection.
1 Introduction

Nowadays, with everything being digital, recommendation systems have become extremely useful tools in different sectors, including e-commerce [1], media and entertainment [2], social media [3], news and information [4], travel and tourism [5], and education [6]. These domains, based on data evaluation and algorithms, help users to find relevant products, services, or information in a vast universe of choices, offer personalized suggestions, and enhance their overall experience based on the user's unique preferences and needs. In the tourism industry, recommendation systems play a critical role by assisting travellers in making well-informed and personalized decisions about destinations, attractions, accommodations, travel-related services, individualized suggestions for points of interest (POIs), and even entire itineraries. The ability to offer personalized recommendations to tourists can have a significant impact on their travel decisions, ensuring memorable, comfortable, and satisfying experiences. Recommendation systems can completely change how travellers plan and enjoy their travels using data evaluation, deep learning and machine learning techniques.

Within the field of artificial intelligence, machine learning focuses on creating models and algorithms that allow computers to learn and make decisions without explicit programming. In the context of recommendation systems, machine learning algorithms can examine browsing histories, past purchases, user preferences, and other data to provide personalized suggestions that suit the interests and preferences of the user [7–9].

A branch of machine learning called "deep learning" focuses only on training deep neural networks, which are multi-layered artificial neural networks. By gradually removing more complex characteristics from unprocessed data, these deep neural networks can learn hierarchical representations of the data. This suggests that deep learning techniques can be applied to recommendation systems to extract complex features from user data, comprehend user preferences, and produce customized recommendations [10, 11].

Although there have been advancements in the tourism industry’s recommendation systems, many of them have offered helpful information and suggestions, but they continue to face challenges in offering truly personalized recommendations. Many systems ignore important factors such as user preferences, travel context, and real-time information because they only use collaborative filtering or content-based techniques. In addition, the systems did not provide tourist recommendations to a person who was already in the current location and would like to know the next place to visit [12].

To overcome these limitations, more advanced and hybrid recommendation systems that can adjust to the requirements and preferences of single tourists are desperately needed. Hence, the principal objective of this research is to suggest a hybrid, customized recommendation system specifically for travellers. The proposed system will provide a comprehensive and customized travel experience by considering multiple factors, including the traveler’s preferences, travel history, current location, and other contextual information such as age. Through this type of information, the next POI recommendation is offered to tourists. Therefore, when suggesting the top travel destinations to users, these different criteria should be considered in the discussion. To combine these criteria, we applied a fuzzy logic system, a mathematical approach that
enables the handling of imprecise or uncertain data, thus enabling the combination of many criteria for decision-making. In doing so, it enhances tourists’ travel experiences by offering highly customized, accurate, and contextually relevant suggestions.

The paper is formatted as follows: The appropriate research studies on tourism recommendation systems are reviewed in Section 2. In Section 3, the proposed hybrid recommendation system is introduced along with an overview of its special features and how it combines several recommendation methodologies. Section 4 discusses implementations and a comprehensive analysis. Section 5 offers a discussion of our system. Key conclusions and recommendations for future research are presented in Section 6.

2 Background and previous research

Recently, due to the tourism industry’s explosive growth as well as the spread of social media and online platforms, tourists have been overloaded with information in different formats (photos, text, ratings, videos, etc.). Such situations make it difficult for tourists to decide and get personalized recommendations for travel-related services. Tourism recommendation systems are becoming increasingly necessary in the travel industry to help travellers make informed decisions about their trips. In general, recommendation systems can be divided into three main categories: collaborative filtering, content-based filtering, and hybrid recommendations as presented in Figure 1.

![Fig. 1: Category of recommender systems](image)

Recommendation systems based on collaborative filtering collect data users’ behaviour and preferences [13], such as hotels that they liked, activities that they enjoyed, and places that are interested, etc. These data are then used to generate recommendations for other users with similar tastes [14, 15]. Memory-based and model-based approaches constitute collaborative filtering (CF). While model-based recommendation systems use machine learning techniques and transaction data to build a model capable of delivering personalized suggestions, memory-based recommendation systems establish recommendations by directly accessing the database [16]. As an example, in this recommendation process [17], the proposed approach uses the fuzzy C-means algorithm for both user- and item-based models. In both models, the similarity measures Pearson correlation and cosine are used to assess how similar users
and items are. The suggested method’s accuracy improvement was measured using the mean absolute error (MAE) as an evaluation metric. The use of machine learning [18] is recommended for its capability to learn and advance over time. For example, in [19], the authors employ machine learning to produce customized recommendations for locations, POIs, lodging, and activities, considering the preferences and actions of users. Jomsri [16] developed a system that guides boat travel destinations along Thailand’s Om Non-Canal using machine learning and analytic order process methodologies. Based on the user’s locations, the model compiles their past travel information. The system creates models using k-means clustering to organize visitors based on the brief travel histories submitted by each user. Wang et al. [20] examined the variables influencing user’s interest preferences regarding their local and global rating data and developed a collaborative filtering recommendation algorithm model for tourist destinations based on user preferences in the presence of sparse data. The algorithm combines classic similarity algorithms with a similarity algorithm based on the Jeffries-Matusita distance. Various systems employ intelligent recommendation algorithms that operate in real-time. The authors in [21] proposed a real-time intelligent recommendation algorithm based on multi-time scale constraints for cross-regional city-level tourist routes with epidemic normalization. The objective is to enhance the capacity for intelligent real-time recommendations of traveller information on tourism routes. In a different study, Yoon et al. [13] presented a real-time travel recommendation system intended for tourism (R2Tour). The latter uses a machine learning model to recommend the top five most visited tourist destinations based on situational factors such as temperature and visitor profiles. Another recommendation system model based on content-based filtering is proposed by [22] organizes data on item characteristics, such as hotel features or activity types. This data is then used to generate recommendations for users based on their preferences. In [23], they present a context-aware recommender system for travel that classifies objects using an ontology. As the system matures, it incorporates different recommenders according to different levels of maturity, such as collaborative, popularity filtering, content filtering, and demographic filtering. To address cold-start issues, accurate item classification based on tourism ontology is achieved using natural language processing (NLP). The framework combines content-based, collaborative, and ontology-based filtering techniques for personalized recommendations, making it flexible enough to be applied to any domain with a particular ontology.

Furthermore, artificial intelligence (AI) is being used to create more intelligent recommender systems such as machine learning, NLP, and deep learning. In addition, the application of big data is used to create a recommendation system that is more comprehensive and precise [24]. Recommendation systems can handle massive datasets with user interactions, item properties, and contextual data by utilizing big data technologies like Hadoop, Spark, and distributed computing frameworks. In [25], the study suggests a framework for analyzing social media reviews and making travel destination recommendations based on five metrics using big data analytics and deep learning models. The framework was implemented to analyze unprocessed text reviews and produce a final destination recommendation score using Apache Spark and Bidirectional Encoder Representation Transformers (BERT).
A recommendation system based on a hybrid system is a combination of two or more methods used in concert to overcome the shortcomings of separate recommender systems. Due to the challenge of handling the volume of necessary data and the cold-start issue brought on by CF [26], numerous studies have integrated various techniques to mitigate the drawbacks of each approach. In [27], the study introduces a hybrid recommender system for the travel industry that combines a content-based recommendation system with a Bayesian preference elicitation component. For the first time, the system integrates a new non-hierarchy similarity metric, the weighted extended Jaccard similarity (WEJS), into a recommender system. The hybrid recommendation system presented in [28] uses deep learning algorithms to combine content-based approaches and collaborative filtering. It analyzes images of tourist attractions using convolutional neural networks (CNN) and user-generated content using recurrent neural networks (RNN).

The system may enhance the user experience and engagement, resulting in higher profits for travel-related enterprises. This research [29] proposes and integrates a fuzzy set approach travel recommender system into a fuzzy decision support system (FDSS) to provide tourists with recommendations for islands, lodging, activities, and other related factors. A fuzzy logic-based recommender system is also proposed in this research [30] that uses a fuzzy inference system (FIS) to recommend tourist spots to travellers and the top locations that satisfy most of the user’s defined preferences.

In [31], the authors focused on the use of fuzzy-based multi-criteria algorithms in developing a hotel recommendation system for e-tourism platforms. The research uses the multi-criteria ratings of social media networking sites that travellers have shared to develop a novel recommender system for hotel recommendations on e-tourism platforms. This study confirms that fuzzy-based recommendation algorithms, in conjunction with machine learning for prediction and clustering, can significantly improve the accuracy of recommendations in the travel industry. A summary of the studies presented is given in Table 1.
Even with the significance of these works, we have been able to improve performance thanks to the way our new model is designed and the methods we have chosen.
<table>
<thead>
<tr>
<th>Ref</th>
<th>Main idea</th>
<th>Methods Used</th>
<th>Dataset</th>
<th>Results</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>ML approaches for developing personalized TRS</td>
<td>-Collaborative filtering, -Content-based filtering, Hybrid recommender systems</td>
<td>User data, text, images, videos, and gadgets</td>
<td>——</td>
<td>-Any metrics or results to evaluate the effectiveness of these TRS. -AB testing presents difficulties for practical testing. -A single algorithm.</td>
</tr>
<tr>
<td>27</td>
<td>Developed a novel hybrid recommender system that integrates recommendations based on content and elicits Bayesian preferences.</td>
<td>Elicitation of Bayesian preferences novel component for content-based (CB) recommendations.</td>
<td>A real-world dataset created for Agios Nikolaos, Crete, Greece.</td>
<td>—</td>
<td>No mention of metrics or results that compare the effectiveness of various algorithms.</td>
</tr>
<tr>
<td>21</td>
<td>A real-time system for recommending tourist routes that considers the normalization of epidemics and multitime scale constraints.</td>
<td>Modeling the correlation between tourists Using the particle swarm optimization technique.</td>
<td>Using small samples.</td>
<td>-Reliability and accuracy of suggested tourist routes.</td>
<td>-Absence of dataset description. -Limited Analysis.</td>
</tr>
<tr>
<td>13</td>
<td>Presented the idea of R2Tour, a real-time travel recommendation system.</td>
<td>K-NN and SVM ML models.</td>
<td>-EVGPS, the Visit Korea data lab, and Korea Meteorological Administration data. -Information on tourist attractions on Jeju Island.</td>
<td>- 77.3% accuracy rate Macro-F1 of 0.415 and micro-F1 of 0.773.</td>
<td>Does not specify the dataset size or level of representativeness.</td>
</tr>
<tr>
<td>25</td>
<td>Framework for tourism planning based on big data analytics that uses automated social media reviews to suggest locations.</td>
<td>-BERT model. -Cross entropy loss function.</td>
<td>social media platforms TripAdvisor, Twitter, and YouTube.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Page</td>
<td>Title</td>
<td>Methodology</td>
<td>Dataset</td>
<td>Benefits</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>A hybrid recommendation system for travel that mixes conventional recommendation algorithms with deep learning methods.</td>
<td>Hybrid model combining collaborative filtering and content-based approaches - CNN and RNN.</td>
<td>Dataset of tourist destinations such as images, reviews, and ratings, as these are the types of data.</td>
<td>Compared with conventional methods, the suggested system offers recommendations that are more precise and tailored. - The system increases user satisfaction and engagement in the travel and tourism sector.</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Develop an intelligent recommendation model for tourist attractions.</td>
<td>Examination of variables that influence user preferences for interests. - Using the Jeffries Matusita distance to derive a similarity algorithm for user preferences.</td>
<td>- Analysis of user rating information.</td>
<td>The enhanced algorithm performs better than alternative algorithms and conventional collaborative filtering. - On a sparser set of tourism data, the accuracy rate is high.</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>A novel recommender system for travel that uses a tourism ontology and natural language processing.</td>
<td>- Content-based, popularity-based, demographic-based, and collaborative filtering methods.</td>
<td>-</td>
<td>- Lack of common rating items among users and sparse data. - Customary similarity metrics are unable to determine user preferences.</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>A system for recommending customized travel that uses the fuzzy analytical hierarchy process.</td>
<td>Fuzzy Analytic Hierarchy Process (fuzzy-AHP) - Fuzzy Decision Support System (FDSS)</td>
<td>-</td>
<td>- Criteria for alternative performance evaluation and measurement - Evaluation of criteria and alternative rankings.</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>A multi-criteria recommender system for tourism using a fuzzy C-means algorithm.</td>
<td>- Item- and user-based fuzzy C-means algorithm - Measures of cosine similarity and Pearson correlation.</td>
<td>TripAdvisor dataset.</td>
<td>- With MAE=0.72, user-based model increases recommendation accuracy. - MAE=0.73 for item-based model</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Recommends the top tourist destinations that correspond to most of the user’s specified preferences using a fuzzy inference system (FIS).</td>
<td>Fuzzy inference system. - After the metadata has been fuzzified, send it to the fuzzy recommendation engine.</td>
<td>-</td>
<td>- Only two compound locations are available. - The user is only drawn to two different kinds of locations.</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Fuzzy-based multi-criteria algorithms improve hotel recommendations on e-tourism platforms.</td>
<td>Clustering and prediction ML techniques.</td>
<td>TripAdvisor data.</td>
<td>The quality of recommendations in the tourism industry is enhanced by fuzzy-based recommendation algorithms.</td>
<td></td>
</tr>
</tbody>
</table>
3 Research Method: Overview of the Personalized Recommendation System

In this study, we developed a personalized hybrid recommendation system designed for tourists called the OptiGuide Recommender system (OGRS). Our system is proposed to improve the relevance of the next POI recommendations by combining several sources of information: the tourist’s current location, expressed preferences, age estimated by facial analysis, and traces of other tourist’s past visits available in the trajectory data warehouse. To achieve this, we propose the use of a fuzzy logic method.

![Components of the OptiGuide Recommender System](image)

Fig. 2: Components of the OptiGuide Recommender System
to flexibly merge inputs from collaborative and content-based recommendation methods. The aim is to make the most of each method to provide the tourist with a precise, personalized recommendation of the next POI to visit.

In summary, our tourism recommendation system consists of several interlinked subsystems, including a location module, preference collection module, age estimation module, visit history database, and fuzzy logic recommendation engine. An overview of the OGRS proposed in this study is illustrated in Figure 2.

3.1 Inscription
The process begins when a tourist launches the application, the system checks whether the user is a new or existing one. They are offered the choice to log in if they are an already registered user. They are taken to a special interface where they can input their password and username. The registration process is a vital initial step for new users. They are sent to a registration interface where they are asked for basic information, including their first and last names, email, Username, Password, Confirm Password, as well as other pertinent details.

3.2 POI detection subsystem
Through image processing and scene recognition, the POI detection system uses deep learning models to identify POIs in photos taken by tourists and extract relevant information. POIs can be classified into various categories, including cultural, historical, artistic, natural, and hotel-related. The proposed POI detection mechanism can be used to help tourists navigate around a region based on POIs detected in their current environment. Many researchers have recently proposed the use of neural network architectures [32–34], transfer learning [35], and advanced reinforcement techniques to improve the performance and accuracy of POI detection. Furthermore, certain systems incorporate multimodal data, such as textual descriptions of images or contextual information, to enhance the comprehension of point-of-interest content [36]. Our research indicates that Yolov8 outperforms other object detection algorithms in terms of accuracy and performance [37]. As a result, we employ this algorithm in the first part of our suggested system to identify and locate things in travel photos.

3.3 Previous history of the visit
In this part of our tourism recommendation system, we get valuable information from travel routes in the past to generate pertinent recommendations for current tourists. To obtain historical data about tourists, we use our Trajectory Data Warehouse, which stores data on past itineraries, visited locations, preferences, and other relevant details. The purpose of this data warehouse is to handle massive amounts of trajectory data in an organized and effective manner [38]. Historical data can be accessed by executing specific queries on the database. Our method for extracting relevant information involves SQL queries based on specific search criteria. This could include determining the number of tourists who have visited a specific place after visiting another. The objective of this part was to analyze visit sequences, detect frequent movement patterns, and identify significant correlations between different POIs. An example of a
query might be to find a same sequence of POIs followed by many tourists. This subsystem helps to improve and power our hybrid recommendation system by proposing locations that suit the individual preferences of each user.

### 3.4 Preferences Selection

Modelling user preferences is an important part of our personalized and hybrid recommendation system. By understanding each user’s unique preferences, this system component enables us to provide recommendations that most closely align with their choices. Tourists are asked to choose their preferences by sorting them in order of priority. For example, the user can choose culture, history, and then nature. Subsequently, a database is created to save user preferences for future use. Each user’s preference model is considered by the recommendation system when it makes suggestions. It looks for POIs that match the user’s selected preferences and filters the recommendations based on those POIs. For example, in the case where the user indicates that culture is his top preference, the system prioritizes cultural POIs when making recommendations.

### 3.5 Tourist Age Classification (TAC)

Choosing a tourist destination that is appropriate for this age is the main problem that a visitor faces when visiting a country’s tourist attractions. Therefore, in our hybrid and personalized recommendation system, visitors are asked to take selfies showing their faces when they use our system. Subsequently, these selfies are processed and fed into our deep learning model to infer the tourist’s age from the features of their selfie. As a result, we can obtain a clear idea and useful demographic data to produce recommendations that are more relevant and personalized. The table 2 provides an example of age recommendations based on the current tourist location.

<table>
<thead>
<tr>
<th>Table 2: Tourist age recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>POI detection</td>
</tr>
<tr>
<td>Chenini Oasis</td>
</tr>
<tr>
<td>Chenini Oasis</td>
</tr>
<tr>
<td>Activity Park</td>
</tr>
<tr>
<td>Farhoud Parc</td>
</tr>
<tr>
<td>Mosque of Jara</td>
</tr>
<tr>
<td>The Lodge</td>
</tr>
<tr>
<td>Star warshotel</td>
</tr>
<tr>
<td>Star warshotel</td>
</tr>
<tr>
<td>Ksar Toujane</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
3.6 Fuzzy logic recommendation system (FLRS)

In this section, we present a novel and flexible approach to make decision through the integration of fuzzy logic into our recommendation system. In this context, fuzzy logic is used to combine our inputs: tourist preference, tourist age, current POI, and historical visit, and recommend the next POI to visit, our solution enables the generation of more personalized recommendations adapted to the tastes and characteristics of every visitor. Figure 3 shows the fuzzy logic decision support system and its components.

![Fig. 3: Fuzzy Logic Decision](image)

Our Fuzzy logic recommendation system (FLRS) is based on four inputs: TP (Tourist Preference), TA (Tourist Age), CP (Current POI), and HV (Historical Visit). The application of fuzzy logic in our recommendation system can be divided into three steps: fuzzification, inference, and defuzzification.

3.6.1 Fuzzification Step

In the fuzzification step, the input variables are converted into fuzzy sets or linguistic terms through the definition of membership functions. The latter can take different shapes, such as triangular, trapezoidal, and Gaussian, depending on the most suitable form for each variable. Because there are no known rules for choosing the shape of linguistic variables, we use in each variable's triangular and each other's trapezoidal membership functions during the fuzzification step. A trapezoidal membership function is typically defined by four parameters a, b, c, and d, representing the start and end points of the base of the trapezoid and the points where the function reaches its...
maximum values. The general equation 1 of a trapezoidal membership function [30] is:

\[
\mu(x) = \begin{cases} 
0 & \text{if } x \leq a \\
\frac{x-a}{b-a} & \text{if } a \leq x < b \\
1 & \text{if } b \leq x \leq c \\
\frac{d-x}{d-c} & \text{if } c \leq x \leq d \\
0 & \text{if } x \geq d
\end{cases}
\] (1)

Where \(a\), \(b\), \(c\), and \(d\), are the parameters defining the shape of the trapezoid, and \(x\) is the value for which the membership function is evaluated. A triangular membership function is defined by three parameters \(a\), \(b\), and \(c\), representing the start, peak, and end points of the triangle. The general equation 2 of a triangular membership function is [30]:

\[
\mu(x) = \begin{cases} 
0 & \text{if } x \leq a \\
\frac{x-a}{b-a} & \text{if } a \leq x < b \\
\frac{x-c}{b-c} & \text{if } b \leq x \leq c \\
0 & \text{if } x \geq c
\end{cases}
\] (2)

Where \(a\), \(b\), and \(c\) are the parameters defining the shape of the triangle, and \(x\) is the value for which the membership function is evaluated. These equations describe how input values are converted into fuzzy membership degrees based on the trapezoidal and triangular membership functions defined in the code. Fuzzy rules then use these membership degrees to make decisions about recommending the next point of interest.

### 3.6.2 Inference

Following the fuzzification of the input data, the inference engine receives the associated fuzzy sets. The linguistic rules that have been obtained from the rule base are used by the inference engine to process the current inputs. The "if-then" structure is commonly used to represent fuzzy rules, which define the relationships between linguistic variables. There are several methods for the inference phase in a fuzzy logic system, each one has advantages and disadvantages. These approaches include MAX-MIN, MAX-PROD, SOM-PROD, ANFIS, Sugeno, and Tsukamoto. Each method uses fuzzy rules to translate fuzzy inputs into fuzzy outputs, but they differ in how they compute these outputs based on the defined rules. In this study, we apply the Mamdani method for inference. It is widely used in the literature [30] because of its simplicity and effectiveness in fusion tasks, employing fuzzy linguistic rules to represent human reasoning. Like other methods, this approach goes through three stages:

- **Rule Activation Stage:** In this stage, the relevant rules are identified based on the input values. Each rule’s firing strength is determined, indicating the degree to which it is applicable. Implication Calculation Stage: The implication stage involves combining the activated rules to derive intermediate fuzzy outputs. This step captures the impact of each rule on the system’s decision.
- **Aggregation Stage:** The final step involves aggregating the intermediate outputs to generate a crisp output, providing a concrete recommendation for the next point of interest.
POI to visit. This aggregation often involves techniques such as the centroid or weighted average. By following these three stages, the Mamdani method facilitates a comprehensive and interpretable inference process within the fuzzy logic system.

3.6.3 Defuzzification

The final step is defuzzification, which transforms the aggregated fuzzy output into a crisp value or decision. It allows transitioning from a fuzzy response to a precise and understandable one. Several defuzzification methods were used, including the centroid (average of maxima), center of gravity, and maximum method [39]. Each method has its own characteristics and is chosen on the basis of the specific needs of the problem. In this study, we will use the center-of-gravity method. Although it is the most time-consuming method, it remains the most widely used [39]. For a resulting membership function $\mu_A(x)$, the center of gravity (COG) can be determined using the following equation 3:

$$COG = \frac{\int_{x \in X} x \mu_A(x) dx}{\int_{x \in X} \mu_A(x) dx}$$

Where:

- $x$ stands for the discourse universe.
- The degree of membership of $x$ is represented by $\mu_A(x)$, which is the membership function of set $A$ at point $x$.
- The symbol $\int x$ indicates integration throughout the full discourse universe.

4 Running example: Use case

In this section, we give a thorough explanation of the fuzzy sets, rules, input variables, and fuzzy logic decision-making procedure in this part. Subsequently, we explore a particular tourism scenario to demonstrate the real-world implementation of our Fuzzy Logic Recommendation System (FLRS) in Gabes.

4.1 Input Variables and Fuzzy Sets

Our fuzzy logic recommendation system (FLRS) is based on four input variables, each specified by a fuzzy set:

- For the CP variable, four fuzzy sets are defined: 'Chenini Oasis,' 'Mosque of Jara,' 'Star Wars Hotel,' and 'Ksar Toujane.' The intervals established for every location according to the "poi_current" variable are interpreted with the use of visitor trajectory data kept in a data warehouse. The distance in geographical terms between every point of interest and the visitor’s current location is indicated by these intervals. For Chenini Oasis, for instance, an interval [0, 0, 1] indicates that a visit or is considered to be very close to the location when the "poi_current" value is near 0 and less close as the value approaches 1. This strategy is predicated on the idea that travellers are more likely to visit locations that are near to where they are right now. In order to
give travellers more relevant recommendations for places of interest based on their current position, the intervals are selected based on this geographic proximity. "TA" is the second input variable, and three fuzzy sets—"young," "adult," and "senior" are defined. The classification is as follows: Young users fall within the age range of 0 to 18 years, adults have ages between 18 and 55 years, and seniors encompass users over 55 years old. For the "TP" variable, four fuzzy sets are defined: 'Nature,' 'Culture,' 'Hotel,' and 'Historic'. When a tourist launches the application, he/she is directed to an interface where he/she can specify his/her preferences for different categories such as culture, history, and nature. For each category, the tourist assigns a numerical value to his/her level of interest. The numerical values assigned by the tourist serve as the basis for assessing their membership level for each category. For example, if a tourist wants to express an interest in nature, he/she selects a numerical value for that specific category. The system applies the corresponding membership function to this value, thus calculating the tourist’s degree of membership. The final variable is "HV," represented by a fuzzy set comprising the categories Low (L), Medium (M), and High (H) based on the number of historical visits. Low refers to a range of 0 to 30, Medium refers to a range of 30 to 70, and High refers to a range of 70 to 100 tourists. Figure 14 and 15 provides the linguistic (fuzzy) representations of these variables.

4.2 Fuzzy rules

As previously mentioned, our goal is to suggest the next point of interest for tourists, considering their age, preferences, current location, and past experiences. In the inference process, we integrated these four input variables to identify the next recommended destination. Here are the fuzzy rules employed for decision making.

1. If CP is "Chenini Oasis" and TA is "young" and TP is "Nature" and HV is "High", then Next_POI is "Chenini Zoo."
2. If CP is "Chenini Oasis" and TA is "Adult" and TP is "Nature" and HV is "Low", then Next_POI is "The Ksar of Chenini Charchara Coffee".
3. If CP is "Chenini Oasis" and TA is "Senior" and TP is "Nature" and HV is "High", then Next_POI is "Chenini Museum".
4. If CP is "Mosque of Jara" and TA is "young" and TP is "Culture" and HV is "High", then Next_POI is "Farhoud Parc".
5. If CP is "Mosque of Jara" and TA is "Adult" and TP is "Culture" and HV is "High", then Next_POI is "Traditional Souk Jara".
6. If CP is "Mosque of Jara" and TA is "Senior" and TP is "Culture" and HV is "High", then Next_POI is "National Office of Artisanats".
7. If CP is "Star Wars hotel" and TA is "young" and TP is "Hotel" and HV is "Low", then Next_POI is "Quad".
8. If CP is "Star Wars hotel" and TA is "Adult" and TP is "Hotel" and HV is "High", then Next_POI is "Camping".
9. If CP is "Star Wars hotel" and TA is "Senior" and TP is "Hotel" and HV is "High", then Next_POI is "Dar Khadija".
10. If CP is "Ksar Toujane" and TA is "young" and TP is "Historical" and HV is "High", then Next_POI is "Artisanat el-hamzaoui".

11. If CP is "Ksar Toujane" and TA is "Adult" and TP is "Historical" and HV is "Medium", then Next_POI is "Hamzaoui Coffee Break".

12. If CP is "Ksar Toujane" and TA is "Senior" and TP is "Historical" and HV is "High", then Next_POI is "Diar Toujane OR Toujane Guesthouse".

4.3 Defuzzification

Finally, the center-of-gravity approach is used to defuzzify the combined fuzzy output. A specific recommendation for next Point of Interest in Gabes is produced by this procedure.

5 A Tourist Scenario: Using FLRS

In our hypothetical tourism situation, let us use Mr. Abdessalem, who is currently visiting the "Ksar Toujane" historic site, as an example. Abdessalem, who is over 55 and categorized as a "Senior" traveler, has a special interest in historical sites. Our fuzzy logic recommendation system (FLRS) uses the following unique rule: "If CP is 'Ksar Toujane' and TA is 'Senior' and TP is 'Historical' and HV is 'High', then Next_POI is 'Diar Toujane OR Toujane Guesthouse'". Based on a significant number of historical visits, this system falls into the "High" category. Thus, our FLRS provides him with a customized recommendation for his next destination, advising him to check out the historically significant "Diar Toujane OU Toujane Guesthouse". This customized plan illustrates Abdessalem’s flexibility and effectively matches his tastes. Figure 16 illustrates this scenario’s results.

6 Experimental Evaluation

The experiments were performed on a CPU with a 2.5 GHz Core i5 processor and 16 GB RAM. The 64-bit operating system used was Windows 10 Professional. Numerous Python and Google Collab libraries were used in the implementation.

6.1 Dataset

In this study, our system used three different data sources. First, a constructed dataset of images is collected manually and described using truth labels. In this study, we have four locations. These locations are represented by ROI Gabes. The second dataset contains data about the historic Trajectory Data warehouse [38], which is already the history of the ancient tourists. The third dataset consists of images from the UTK face data of people divided into different age groups. These datasets were employed as multi-datasets to evaluate the performance of our system recommendations. Table 3 describes our tourism dataset in terms of the total number of tourist images for each class. Table 4 shows the number of images for each category of age.

https://susanqq.github.io/UTKFace/
The following is a detailed description of both datasets: The tourist POI dataset is composed of 689 images of tourist POI like Chenini Oasis (class 0), Ksar Toujane (class 1), Mosque of Jara (class 2) and Star Wars Hotel (class 3). Some of the images were taken manually using a mobile phone camera and some of them were obtained from internet sources. For training and testing, the images were divided into an 80:20 ratio. An overview of our dataset is shown in Figure 4. The age group dataset comprised 9972 images. The images are collected from the UTK face dataset. An overview of the dataset is shown in Figure 5.

The database of trajectories data warehouse contains historical information (e.g. Point of interest, distance, time, and other relevant details).

---

**Table 3:** Tourist POI Dataset items.

<table>
<thead>
<tr>
<th>Class</th>
<th>Tourist POI</th>
<th>Number of images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 0</td>
<td>Chenini Oasis</td>
<td>205</td>
</tr>
<tr>
<td>Class 1</td>
<td>Ksar Toujane</td>
<td>209</td>
</tr>
<tr>
<td>Class 2</td>
<td>Mosque of Jara</td>
<td>180</td>
</tr>
<tr>
<td>Class 3</td>
<td>Star Wars Hotel</td>
<td>95</td>
</tr>
</tbody>
</table>

---

Fig. 4: Sample images of the tourist POI dataset.
Fig. 5: Samples images of the age group dataset.

Table 4: Image number for each category of age.

<table>
<thead>
<tr>
<th>Class</th>
<th>Tourist Age</th>
<th>Number of images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 0</td>
<td>&lt;18</td>
<td>4050</td>
</tr>
<tr>
<td>Class 1</td>
<td>18-55</td>
<td>4024</td>
</tr>
<tr>
<td>Class 2</td>
<td>&gt;55</td>
<td>1898</td>
</tr>
</tbody>
</table>

6.2 POI detection subsystem

This study examines YOLOv7 and YOLOv8, two well-known object detection algorithms, in the context of the larger YOLO (You Only Look Once) model family. These models are highly praised for their improved real-time processing power and increased object detection accuracy. Another notable object detection technique is the Faster R-CNN algorithm, which is renowned for its high detection accuracy. Regarding the experimental design, we use a customized tourism dataset for this research, which suggests a purposeful adjustment to the unique difficulties and features present in tourism photos. A comprehensive set of metrics was used to evaluate the three algorithms: Yolov7, Yolov8, and Faster R-CNN. A more nuanced understanding of the algorithm’s performance can be obtained by comparing the precision and recall metrics, which measure correctly predicted positive instances out of all predicted positives and correctly predicted positives out of all actual positives, respectively. A popular metric in object detection, mean average precision (mAP), combines recall and precision across different confidence levels to provide a comprehensive evaluation. The experimental
findings are then displayed in a figure showing that YOLOv8 outperforms the other algorithms in terms of precision. Thus, the paper concludes that, among YOLOv7, YOLOv8, and Faster R-CNN, YOLOv8 is the most successful algorithm for object detection in travel-related photos. Showing the results of extensive testing, figure 8 illustrates how different classes were evaluated. A notable 94.5% accuracy rate was obtained, highlighting the model’s exceptional ability to correctly identify and classify the various classes being studied. These findings underscore the importance of aligning model selection with specific application requirements to optimize performance for diverse use cases.

![Fig. 6: Results of object detection models](image)

<table>
<thead>
<tr>
<th></th>
<th>YoloV8</th>
<th>YoloV7</th>
<th>Faster R-CNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision (%)</td>
<td>92.8</td>
<td>92.1</td>
<td>91.8</td>
</tr>
<tr>
<td>Recall (%)</td>
<td>93.0</td>
<td>92.7</td>
<td>86.3</td>
</tr>
<tr>
<td>mAP (%)</td>
<td>94.5</td>
<td>93.3</td>
<td>90.0</td>
</tr>
</tbody>
</table>
6.3 Tourist Age Classification

We evaluated multiple classifications to determine the ages of visitors to provide useful demographic information that would allow the development of recommendations that are not only more accurate but also highly personalized for each tourist. We invite users to take a selfie of their faces when they use our application. The tourist’s age is then inferred from the attributes of their selfie by processing and feeding it into
our deep learning model. This can assist visitors in selecting attractions that are best suited to their age group, which can greatly improve their whole tourist experience. We conducted multiple tests on a dataset of various tourist selfies with age labels to evaluate the performance of our classifier. To assess the performance of the model, we separated the dataset into 10% test, 20% validation, and 70% training sets. Different distinct deep-learning models for age categorization were trained and evaluated: MobileNet, ResNet50, ResNet101, YOLOv5, DenseNet-169 and InceptionV3. Every model was trained using the training set, and different evaluation criteria were used to evaluate each model’s performance. A variety of metrics and assessment methods, such as accuracy, precision, recall, F1 score can be used to evaluate our tourist age classification model. These metrics help in the understanding of the model’s classification abilities, age prediction accuracy, and facial recognition accuracy. The table compares the results of several neural network models for a classification task. YOLOv5 achieves the highest accuracy of 93.5 percent, indicating that it can accurately predict the class of a picture with a 93.5 percent success rate. With a precision of 91%, ResNet50 comes in second place, followed by MobileNet and InceptionV3, which have precisions of 88% and 86%, respectively. The lower precisions of DenseNet-169 and ResNet101 are 80% and 67%, respectively. ResNet50 has the lowest loss (0.039), indicating that it can learn the classification task with the fewest errors. The second-lowest loss is YOLOv5, followed by DenseNet169 (0.061), MobileNet and InceptionV3 (0.051), and lastly ResNet101 (1.36). The highest F1 score, obtained by YOLOv5, is 0.97, indicating a good balance between recall and precision. ResNet50 comes in second place with a score of F1, followed by MobileNet and InceptionV3, who have scores of F1 of 0, 88 and 0, 86, respectively. DenseNet169 and ResNet101 had lower F1 ratings, with 0.81 and 0.65, respectively. The highest precision (0.98) is achieved by YOLOv5, indicating that it can accurately predict the positive classes (pictures belonging to

Fig. 9: Comparison of the different classifiers.
Fig. 10: Tested images on yolov5.

Table 5: Results for age group classification.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>Loss</th>
<th>F1 Score</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>MobileNet</td>
<td>88</td>
<td>0.51</td>
<td>0.88</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>Resnet50</td>
<td>91</td>
<td>0.39</td>
<td>0.91</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>ResNet101</td>
<td>67</td>
<td>1.36</td>
<td>0.65</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>InceptionV3</td>
<td>86</td>
<td>0.51</td>
<td>0.86</td>
<td>0.88</td>
<td>0.86</td>
</tr>
<tr>
<td>DenseNet169</td>
<td>80</td>
<td>0.61</td>
<td>0.81</td>
<td>0.83</td>
<td>0.80</td>
</tr>
<tr>
<td>YOLOv5</td>
<td>93.5</td>
<td>0.64</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>

The target class). With a precision of 0.91, ResNet50 comes in second place, followed by MobileNet and InceptionV3, with precisions of 0.88 and 0.86, respectively. The lower precision values of DenseNet169 and ResNet101 are 0.83 and 0.69, respectively. YOLOv5 receives the highest recall (0.98), indicating that it can accurately predict all images belonging to the target class. ResNet50 ranks second with a recall of 0.91, followed by MobileNet and InceptionV3, with recalls of 0.88 and 0.86, respectively. The lower regressions for DenseNet169 and ResNet101 are 0.80 and 0.67, respectively. We summarize in figure 9 the results of the different models. We choose YOLOv5 as
the most reliable and accurate model for age classification in our hybrid recommendation system based on the experimental findings. The model’s ability to accurately predict the age groups of tourists was confirmed by its consistent performance over various evaluation metrics. YOLOv5, which was applied to face detection and ROI extraction, performed exceptionally well, achieving a validation accuracy of 93.5% on the test set. This helped with the age classification procedure by precisely localizing facial features as shown in Figure 10 and Figure 11.

6.4 Preferences selection

An important part of implementing our personalized and hybrid recommendation system is comprehending and simulating user preferences. This part of the system seeks to understand each user’s distinct tastes to personalize recommendations for them. To make this easier, travelers actively participate in the preference selection process by prioritizing and ranking their choices. As seen in the illustration 12, travelers can
identify their personal interests by expressing their preferences in several categories through checkboxes. A user might, for instance, place culture above history, then nature, and so on. After that, the preferences are collected and stored for later use in a specific database. This database acts as a storehouse for user preferences, providing the framework for the system to comprehend personal preferences.

![Trip Preference](image)

**Fig. 12:** Recommendation.

### 6.5 Data manipulation

This section outlines the experimental design and presents the findings of our method for gathering pertinent data for travel recommendations from a trajectory data repository. We utilized our big trajectory data warehouse, which holds historical information on traveler routes, destinations, and other pertinent facts. Our big trajectory data warehouse is an extension of big data warehouse [38] [40] where we integrate data from various sources to analyze and store large-scale trajectory data.

To get pertinent data out of the data warehouse, we created a set of SQL queries. The purpose of this search was to sequentially identify POIs that are regularly frequented by visitors. An example of a query we used is shown below:

**Example 1**

```sql
SELECT COUNT (DISTINCT T.ID_Tourist) AS Tourist_Count
FROM Trajectory T
JOIN POI P1 ON T.ID_POI_1 = P1.ID_POI
JOIN POI P2 ON T.ID_POI_2 = P2.ID_POI
WHERE P1.Name_POI = 'Chenini Oasis'
AND P2.Name_POI = 'Chenini Museum';
```

**Example 2**

```sql
SELECT COUNT (DISTINCT T.ID_Tourist) AS Tourist_Count
FROM Trajectory T
JOIN POI P1 ON T.ID_POI_1 = P1.ID_POI
```

24
JOIN POI P2 ON T.ID\_POI\_2 = P2.ID\_POI
WHERE P1. Name\_POI = ‘Ksar Toujane’
AND P2. Name\_POI = ‘Artisanat el-hamzaoui’.

A hybrid tourist recommendation system can be enhanced and powered by the data retrieved from the big trajectory data warehouse. We can provide travellers with more precise and tailored recommendations by detecting POIs that are regularly visited in a particular order. For instance, since ‘Chenini Museum’ is a well-liked point of interest that is commonly visited after ‘Chenini Oasis, we can suggest it as a possible next destination.

(a) POI current input variable.   (b) Age input variable.

(a) Historical visit variable.   (b) Preferences

Fig. 14: Input Variables and Fuzzy Sets.
7 Discussion

TourOptiGuide’s development responds adeptly to the shortcomings inherent in existing recommendation systems within the tourism industry, which frequently fail to deliver truly personalized suggestions. Through a comprehensive hybrid approach that considers multiple influencing factors, we have successfully crafted an enriched and personalized travel experience for tourists. The deployment of deep learning models for location identification and age estimation significantly elevates the precision of recommendations, ensuring that tourists receive pertinent suggestions tailored to
their specific context. The incorporation of historical data evaluation enables TourOptiGuide to draw insights from past experiences, providing a foundation for valuable recommendations in future travel scenarios. The infusion of fuzzy logic decision support serves as a valuable augmentation, further empowering the system to synthesize diverse inputs and generate the next top point of interest (POI) suggestions. This nuanced approach allows for a more intricate understanding of user preferences and context, resulting in highly individualized recommendations. Moreover, the hybrid nature of TourOptiGuide combines the strengths of collaborative filtering, content-based filtering, and other techniques, effectively mitigating the limitations inherent in single recommendation approaches. When compared to previous literature, our FLRS system stands out due to several novel features [19]. First, in line with several other studies, we employ machine learning to personalize suggestions by deftly fusing collaborative and content-based methods. However, we distinguish ourselves by adding other data, such as present location, approximate age, and previous visits, which improves the precision of our suggestions. Using an advanced fuzzy logic system to combine the output of recommendation methods is a noteworthy feature that offers more accuracy and flexibility than current models [19]. Moreover, our comparative analysis with other studies demonstrates our dedication to offering more sophisticated and pertinent suggestions. Our method appears to include an additional fuzzy logic model for the fusion of results, allowing for a better evaluation of the numerous aspects, based on an analysis of multiple studies [27] [28]. Furthermore, the incorporation of further data on location, age, and previous visits bolsters the applicability of our suggestions. Lastly, our strategy places a particular focus on customizing recommendations to each visitor, accounting for their age, tastes, and past visits [21] [25]. With the help of fuzzy logic and this data, our system can handle data imprecision and uncertainty more effectively than other approaches in the literature [13], providing recommendations that are more exact, tailored to the individual, and flexible. Preliminary evaluations affirm notable advancements in both tourist satisfaction and the efficiency of itinerary planning. By delivering personalized and contextually relevant recommendations, TourOptiGuide enables tourists to make well-informed decisions concerning travel destinations, attractions, accommodations, and overall itineraries. The system has the potential to revolutionize how travelers plan and relish their journeys, offering a more enjoyable and memorable travel experience. Future research could concentrate on refining the recommendation algorithms and exploring the integration of additional data sources, such as real-time information and social media data, to further enhance the system’s recommendations. In addition, a continuous feedback loop and ongoing evaluation are imperative for optimizing TourOptiGuide to ensure its effectiveness in real-world scenarios. Overall, TourOptiGuide marks a significant stride forward in the realm of personalized tourism recommendation systems, poised to positively impact the travel industry and enhance tourist’s overall travel experiences. Although TourOptiGuide’s hybrid approach overcomes the drawbacks of single recommendation methods, data balancing’s possible effects should be considered in the Tourist POI Dataset, especially in class 3. We include expanding the dataset by collecting additional image sets pertaining to various tourist destinations. The goal of data balancing approaches is to correct imbalances in the training data that may occur from under representation
of classes or categories. The recommendation system may be biased towards the over represented categories if the training data is unbalanced, which would lead to fewer accurate recommendations for the underrepresented categories. Additional possible negatives include Absence of rating: The recommender system’s rating is not mentioned in our article. In future assessments, we want to employ evaluation criteria including F1-score, mean absolute loss (MAE), recall, and precision.

8 Conclusion

In this work, we proposed OptiGuide Recommender system that offers a comprehensive and innovative approach to personalized tourism recommendations. By integrating various sources of information and utilizing advanced techniques such as fuzzy logic, the system aims to enhance the relevance and precision of Points of Interest recommendations for tourists. With its interconnected subsystems and the ability to consider factors such as current location, expressed preferences, estimated age, and visit history, OGRS is poised to provide tourists with tailored and accurate recommendations for their travel experiences. As the tourism industry continues to evolve, the development and implementation of such personalized recommendation systems are essential for improving the overall travel experience for tourists worldwide.

Acknowledgements. The authors would like to acknowledge the financial support of this work by grants from the General Direction of Scientific Research (DGRST), Tunisia, under the ARUB.

Declarations

We affirm that all named authors have made substantial contributions to the research and manuscript preparation. Furthermore, we declare that the submitted work is original and has not been published previously.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The authors declare that they have no competing interests related to this research.

Ethics Approval and Consent to Participate

Ethical approval was obtained for this study, which involved analysis of the existing literature. No human subjects or animals were directly involved; therefore, consent to participate was not required.

Consent for Publication

Not applicable because this study did not include any individual data.
Data Availability
Data sharing does apply to this article, as new custom data were created and applied in this study; and other data used are available and cited.

Materials Availability
Not applicable, as this study does not involve the creation or development of new materials.

Code Availability
Applicable, as this study does include the development or implementation of a new code.

Author Contribution
All authors contributed to the conception and design of the study and manuscript preparation. Each author played a significant role in critically revising the content and approving the final version for submission.

Conflict of Interest Statement
"I, Intissar Hilali, hereby declare that I have no personal or financial ties to any person or organization that could improperly influence, or be perceived to influence, my work in the preparation and publication of this paper, which is titled ‘TourOptiGuide: A Hybrid and Personalized Tourism Recommendation System. There were no conflicts of interest, and the research reported in this study was conducted objectively. Still, I recognize that the topic includes using deep learning in the tourism field, so it is crucial to reveal any possible conflicts to maintain transparency."

References


[16] Aditya, P., Budi, I., Munajat, Q.: A comparative analysis of memory-based and model-based collaborative filtering on the implementation of recommender system


[27] Chalkiadakis, G., Ziogas, I.P., Koutsmanis, M., Errikos Streviniotis, Panagiotakis,


