Hotel Recommendation System Using Natural Language Processing

[1]Anusha P M [2] Pradeep Kanthappa Hirebidre
[1]*Assistant Professor, Department of MCA, BIET, Davanagere*
[2] *Student,4th Semester MCA, Department of MCA, BIET, Davanagere*

Abstract

In the era of digital transformation, travelers increasingly depend on online platforms to identify accommodations that align with their unique travel intentions and preferences. The sheer volume and diversity of hotel options available online often make this selection process challenging and time-consuming. This paper presents a hotel recommendation system developed using Flask, a lightweight Python web framework, and enhanced by natural language processing (NLP) techniques to address this challenge. The proposed system enables users to input a brief, natural-language description of their travel purpose—such as business trips, honeymoons, or family vacations—alongside their preferred destination. Utilizing basic NLP methods, including tokenization, stopword removal, and lemmatization, the system processes user input and matches it against hotel tags and review metadata through a similarity-matching algorithm. Recommendations are then ranked by relevance and review scores, ensuring that users receive personalized, high-quality hotel suggestions tailored to their needs. The application features distinct roles for administrators and standard users, supporting robust management and a user-friendly experience. Administrators can manage users, FAQs, and tourist information, while travelers benefit from streamlined hotel recommendations and access to relevant travel resources. Experimental evaluation demonstrates the system's effectiveness in improving user satisfaction and efficiency in hotel selection, highlighting its practical value for both travelers and the hospitality industry.

Keywords: Hotel Recommendation System, Natural Language Processing, Flask, Text Processing, Similarity Matching, Personalized Recommendations, Travel Planning.

I.INTRODUCTION

The digital revolution has dramatically reshaped the travel and hospitality industry, empowering travelers with instant access to a vast array of accommodation options through online platforms. With the proliferation of hotel booking websites, online travel agencies, and review aggregators, users can now browse thousands of hotels, compare amenities, read guest reviews, and make reservations—all from the comfort of their homes. This abundance of information, while beneficial, has introduced a new challenge: information overload. Travelers are often confronted with an overwhelming number of choices, making it increasingly difficult to identify the hotel that best matches their unique needs and preferences. The process of sifting through countless listings, applying multiple filters, and reading extensive reviews can be both time-consuming and confusing, often resulting in suboptimal booking decisions and reduced user satisfaction.

Traditional hotel recommendation systems have attempted to address these challenges by employing collaborative filtering, content-based filtering, or hybrid approaches. While these methods can suggest hotels based on user ratings, preferences, or past behavior, they often fall short in understanding the nuanced context and intent behind a traveler's

search. For instance, a user planning a business trip may prioritize proximity to conference centers and business amenities, while a couple on a honeymoon may seek romantic settings and luxury services. Generic recommendation engines frequently overlook such contextual details, leading to recommendations that may be popular or highly rated but are not necessarily relevant to the user's specific travel purpose. This gap underscores the need for intelligent, context-aware systems that can interpret natural language input and deliver personalized hotel suggestions.

To address these limitations, this paper presents a hotel

recommendation system that leverages natural language processing (NLP) techniques and a lightweight web framework to provide tailored accommodation suggestions. The core idea is to allow users to describe their travel intentions in their own words—such as "I am going on a business trip" or "We are looking for a honeymoon suite"—along with their preferred destination. The system processes this natural language input using basic NLP methods, including tokenization, stopword removal, and lemmatization, to extract meaningful information and standardize the query. It then compares the processed input against hotel tags and review-related metadata using a similarity-matching algorithm. By ranking the results based on both relevance and average review scores, the system ensures that users receive high-quality recommendations that are closely aligned with their stated needs.

The application is built using Flask, a lightweight Python web framework known for its simplicity and flexibility, making it well-suited for rapid prototyping and deployment of web-based solutions. The backend harnesses Python's NLP libraries to analyze user queries and hotel data, while the frontend provides an intuitive interface for users to interact with the system. A key feature of the application is its support for two distinct user roles: administrators and standard users. Administrators have comprehensive control over the platform, including the ability to manage user accounts, add or remove frequently asked questions (FAQs), and update information about local tourist attractions. Standard users, on the other hand, can register, log in, and access the core recommendation features. They can receive personalized hotel suggestions, browse FAQs for guidance, and explore information about tourist places, making the application both informative and user-friendly.

The significance of this system extends beyond convenience for travelers. For the hospitality industry, the ability to deliver personalized recommendations based on user intent can enhance customer satisfaction, increase booking rates, and improve brand loyalty. Hotels can better target their offerings to the right audience, whether it be business travelers, families, or couples, thereby optimizing occupancy rates and revenue. From a technological perspective, the integration of NLP and web development demonstrates the practical application of artificial intelligence in solving real-world challenges. The system's modular design also allows for scalability and adaptation to other domains, such as restaurant or activity recommendations, further highlighting its versatility.

In summary, the proposed hotel recommendation system aims to bridge the gap between user expectations and available technology by providing a smart, efficient, and user-centric solution for hotel selection. By understanding the context and intent behind each user's query, the system not only streamlines the booking process but also enhances the overall travel experience. The remainder of this paper is organized as follows: Section II reviews related work in the field of hotel recommendation systems and NLP-based personalization; Section III details the system architecture and design; Section IV describes the implementation of NLP techniques and the recommendation algorithm; Section V presents experimental results and user evaluations; Section VI discusses the implications, limitations, and potential future enhancements; and Section VII concludes the paper with final remarks on the system's impact and contributions.

II. RELATED WORK

1. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors: Lavrentia Aravani.

This research introduces an advanced hotel recommendation system leveraging Natural Language Processing techniques, particularly BERT (Bidirectional Encoder Representations from Transformers). By analyzing user-generated hotel reviews, the framework classifies feedback into categories such as "Bad," "Good," or "Excellent," enabling highly personalized hotel suggestions. The approach enhances user satisfaction by aligning recommendations with individual preferences and emotional cues derived from past booking experiences.

2. Hotel Recommendation System, Author- Aditi MavalankarAditi Mavalankar,

When planning a trip, booking a good place to stay is crucial but can be overwhelming due to the vast number of hotel options available. To address this, we explored the Expedia hotel recommendation dataset, which provides rich user and hotel features. The main goal of this task is to predict and

recommend the top five hotel clusters (out of 100 possible clusters) that a user is most likely to book, based on their search and preference data. By analyzing user behavior and contextual information, the recommendation system aims to simplify hotel selection and provide personalized suggestions to enhance the booking experience.

3. An Intelligent Data Analysis for Hotel Recommendation Systems using Machine Learning,

Authors:Bushra Ramzan, Imran Sarwar Bajwa, Noreen Jamil, Farhaan Mirza, This paper introduces an intelligent hotel recommendation system that uses machine learning and collaborative filtering, enhanced by sentiment analysis, to process large and diverse data. By analyzing guest reviews for sentiment and profiling guest types (such as solo, family, or couple), the system personalizes hotel recommendations based on both hotel features and guest preferences. Built on a Hadoop big data platform and using fuzzy rules, the approach handles complex data efficiently. Experiments on real-world datasets show that the proposed method achieves higher accuracy, better precision, recall, and faster response times compared to traditional recommendation techniques.

4. Hotel Recommendation System Using Machine Learning, Authors: Meduri V N S S R K Sai Somayajulu, E. Mahendra, B. Seshagiri, S.Raghu Kumar, K.Bhanuji Rao,

This paper highlights the vital role of recommendation systems in the hotel industry, where they help travelers make better accommodation choices by providing personalized suggestions. Using advanced algorithms and user data, these systems have transformed how people select hotels online. The paper also discusses the key challenges and recent advancements in developing effective recommendation systems for online tourism resources.

5. Hotel Recommendation System, Authors-Aditi A. Mavalankar,Ajitesh Gupta, Chetan Gandotra, Rishabh Misra,

Choosing a hotel online can be daunting due to the vast number of options available. To address this, we analyzed Expedia's hotel recommendation dataset to better understand user preferences. The goal is to predict and recommend the top five hotel clusters, out of one hundred, that a user is most likely to book, making the hotel selection process easier and more personalized.

6. Food Recommendation System Using Machine Learning, Authors- Mr.Anil S.Manmothe, Miss.Jayashree Mahale, Miss.Bhagyashri Patil, Mr.Abhijeet patil, Prof.Kanchan Mahajan,

This paper proposes a food recommendation system for hotels that considers multiple rating criteria rather than just a single factor. By using Natural Language Processing on food reviews, the system creates a detailed user-item-feature database. It leverages both user preferences and previous user ratings to suggest the best hotels in a city. When users rate hotel food similarly, their future recommendations are shared, improving the relevance and personalization of hotel suggestions.

7. Hotel Content-Based Recommendation System, Authors- Authors:Kristian Wahyudi, Johanes Latupapua, RitchiMarch,

This recommendation system focuses on suggesting hotels in U.S. cities by calculating ratings for different hotel categories within each city. By combining features such as hotel categories and their city-specific ratings, the system recommends the hotel with the highest rating to the user, ensuring more relevant and location-based suggestions.

8. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors:Lavrentia Aravani, Emmanuel G Pintelas, Christos Pierrakeas, P.

E. Pintelas,

This study presents a hotel recommendation system that uses advanced Artificial Intelligence, specifically Deep Learning and Natural Language Processing with BERT, to analyze customer text reviews. By categorizing reviews as "Bad," "Good," or "Excellent," the framework provides personalized hotel suggestions based on user preferences and booking history. The results show that this approach significantly improves the user experience by delivering tailored recommendations for hotel bookings.

9. Hotel Recommendation System Based on Customer's Reviews Content Based Filtering Approach,Authors:H. Shah,

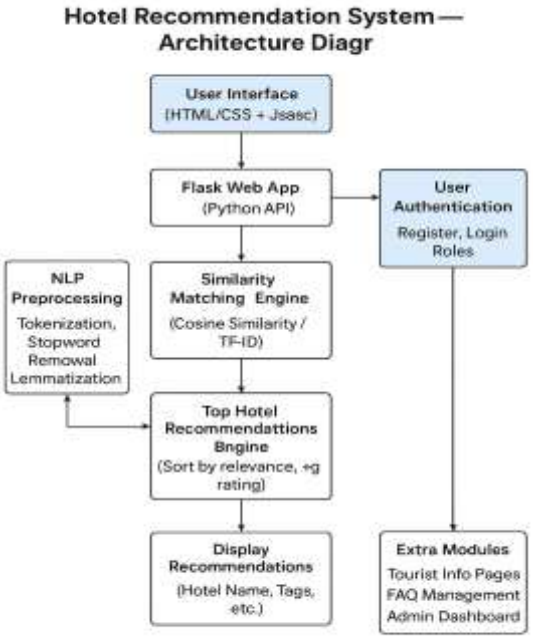
L. Jacob,

This project addresses the challenge of overwhelming online hotel choices by using content-based recommendation systems that leverage historical user preference data. By applying natural language processing techniques like word embedding, word2vec, and TF-IDF to client reviews from a Kaggle dataset, the system extracts features to recommend the top 10 hotels based on the user's previous hotel location knowledge. This approach helps users efficiently find hotels that best match their preferences.

10. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors- Lavrentia Aravani, Emmanuel Pintelas, Christos Pierrakeas, Panagiotis Pintelas This study introduces a hotel recommendation system that uses advanced AI, specifically BERT-based Natural Language Processing, to analyze customer text reviews. By categorizing reviews as "Bad," "Good," or "Excellent," the system provides personalized hotel suggestions tailored to individual preferences and booking history, significantly improving the hotel booking experience for users.

III. METHODOLOGY

The methodology for the proposed hotel recommendation system is structured to combine natural language processing (NLP) techniques with a user-friendly web interface, enabling the extraction of user intent from free-form text and matching it to relevant hotel features. The following steps outline the end-to-end process, from data collection to recommendation delivery.



1. Data Collection and Preparation

The system utilizes a dataset comprising hotel information, including tags (such as amenities, location, and room types), as well as user-generated reviews. These data sources are aggregated from online travel platforms and stored in a structured format suitable for processing. The reviews and tags provide the semantic foundation for matching user queries to hotel features.

2. Text Preprocessing

To ensure effective analysis, both user input and hotel metadata undergo several preprocessing steps:

- Tokenization: User queries and hotel tags are broken down into individual words or tokens.
- Stopword Removal: Common words that do not contribute significant meaning (e.g., "the", "is", "at") are removed to reduce noise.
- Lemmatization: Words are converted to their base or root forms, standardizing the vocabulary for more accurate matching.

3. Feature Extraction and Vectorization

Processed text data is transformed into numeric representations. For basic systems, this may involve simple vectorization techniques such as Bag-of-Words or TF-IDF. More advanced implementations may leverage transformer-based models like BERT, which generate contextual embeddings that capture the semantic meaning of the input text.

4. Similarity Matching Algorithm

The core of the recommendation engine is a similarity-matching algorithm that compares the processed user query against hotel tags and review metadata. The algorithm calculates the degree of similarity—using measures such as cosine similarity—between the user's described preferences and the features associated with each hotel. Hotels with the highest similarity scores are considered most relevant to the user's intent.

5. Sentiment and Review Analysis

To further refine recommendations, the system analyzes hotel reviews using sentiment analysis. This step classifies reviews

as positive, neutral, or negative, and aggregates this information to assign an overall sentiment score to each hotel. Advanced systems may use deep learning models like BERT for more nuanced sentiment classification, categorizing reviews into "Bad," "Good," or "Excellent".

6. Ranking and Recommendation Delivery

The hotels are ranked based on a combination of relevance (similarity to user intent) and average review sentiment scores. The system presents the top recommendations to the user, ensuring that the results are both contextually appropriate and highly rated by previous guests.

7. Web Application Integration

The entire recommendation pipeline is integrated into a Flask-based web application. The frontend allows users to register, log in, and enter their travel preferences in natural language. The backend processes these requests, executes the recommendation algorithm, and returns the most suitable hotel options. Administrators can manage users, FAQs, and tourist information via dedicated interfaces

IV. TECHNOLOGIES USED

The proposed hotel recommendation system integrates a range of modern technologies and frameworks to deliver personalized, context-aware hotel suggestions. Below are the primary technologies and tools employed in the development and deployment of the system:

1. Python Programming Language

Python serves as the core programming language for both backend and data processing tasks due to its simplicity, extensive libraries, and community support.

2. Flask Web Framework

Flask, a lightweight and flexible Python web framework, is used to build the web application interface, manage user authentication, handle requests, and serve hotel recommendations to users in real time.

3. Natural Language Processing (NLP) Libraries

NLTK (Natural Language Toolkit): Utilized for basic NLP tasks such as tokenization, stopword removal, and lemmatization, which are essential for cleaning and standardizing user input and hotel metadata.

spaCy: May be used for efficient and scalable advanced NLP processing, including part-of-speech tagging and entity recognition.

Transformers (e.g., BERT): For advanced implementations, transformer-based models like BERT are employed to extract semantic features from user queries and hotel reviews, enabling deep contextual understanding and sentiment analysis.

4. Data Science and Machine Learning Libraries

Pandas and NumPy: Used for data manipulation, cleaning, and numerical operations throughout the preprocessing and analysis pipeline.

scikit-learn: Provides algorithms for feature extraction (e.g., TF-IDF vectorization), similarity calculation (e.g., cosine similarity), and model evaluation.

Deep Learning Frameworks (e.g., PyTorch, TensorFlow): These may be used for training and deploying deep learning models, especially when leveraging transformer architectures for text classification and sentiment analysis.

5. Database Management

A relational or NoSQL database is used to store hotel information, user profiles, reviews, and system metadata, ensuring efficient data retrieval and scalability.

6. Frontend Technologies

The user interface is developed using HTML, CSS, and JavaScript, providing a responsive and intuitive experience for both standard users and administrators.

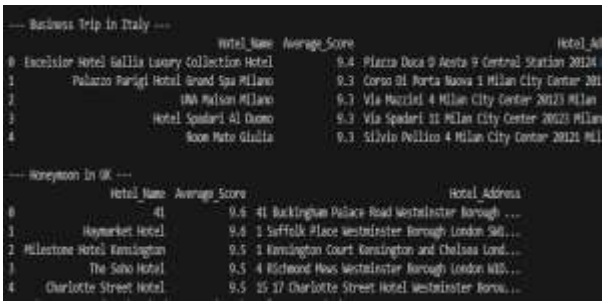
7. Jupyter Notebook

Jupyter Notebook is employed for exploratory data analysis, model prototyping, and visualization during the development phase.

8. Web Hosting and Deployment

The application can be deployed on cloud platforms or traditional web servers, ensuring accessibility and scalability for a broad user base.

V Result



Recommendation of hotels based on text.

VI. CONCLUSION

The proposed hotel recommendation system effectively addresses the challenges faced by modern travelers in selecting suitable accommodations by leveraging natural language processing and a user-friendly web interface built with Flask. By allowing users to describe their travel needs in their own words and intelligently matching these preferences with hotel tags and review data, the system delivers personalized and high-quality recommendations that enhance both user satisfaction and booking efficiency. The integration of sentiment analysis, robust backend processing, and distinct user roles ensures a comprehensive solution that benefits travelers and administrators alike. This work demonstrates the practical impact of AI-driven personalization in the hospitality industry and lays a strong foundation for future enhancements and broader applications in intelligent travel planning.

REFERENCES

[1]. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors: Lavrentia Aravani, Emmanuel
G. Pintelas (University of Peloponnese), Christos Pierrakeas (University of Patras),P. E. Pintelas (University of Patras), DOI: 10.48550/arXiv.2408.00716, Publisher: arXiv (Cornell University)

[2]. Hotel Recommendation System, Author- Aditi MavalankarAditi Mavalankar, Publisher: arXiv, DOI: 10.48550/arXiv.1908.07498, arXiv preprint in August 2019

[3]. An Intelligent Data Analysis for Hotel Recommendation Systems using Machine Learning, Authors:Bushra Ramzan, Imran Sarwar Bajwa, Noreen Jamil, Farhaan Mirza, publisher- October 2019, DOI:10.48550/arXiv.1910.06669

[4]. Hotel Recommendation System Using Machine Learning, Authors: Meduri V N S S R K Sai Somayajulu, E. Mahendra, B. Seshagiri, S.Raghu Kumar, K.Bhanuji Rao, Publisher- 2024 JETIR April 2024, Volume 11, Issue 4 www.jetir.org(ISSN-2349-5162)

[5]. Hotel Recommendation System, Authors-Aditi A. Mavalankar,Ajitesh Gupta, Chetan Gandotra, Rishabh Misra, Publisher- arXiv:1908.07498v2 [cs.LG] 21 Aug 2019

[6]. Food Recommendation System Using Machine Learning, Authors- Mr.Anil S.Manmothe, Miss.Jayashree Mahale, Miss.Bhagyashri Patil, Mr.Abhijeet patil, Prof.Kanchan Mahajan, Publisher- April 2020 | IJIRT | Volume 6 Issue 11 | ISSN: 2349-6002

[7]. Hotel Content-Based Recommendation System, Authors- Authors:Kristian Wahyudi, Johanes Latupapua, RitchiMarch, Publisher- 2020Journal of Physics Conference Series 1485(1):012017, DOI:10.1088/1742-6596/1485/1/012017, LicenseCC BY 3.0e Chandra, Ganda Girsang,

[8]. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors:Lavrentia Aravani, Emmanuel G Pintelas, Christos Pierrakeas, P. E. Pintelas, Publisher- August 2024, DOI:10.48550/arXiv.2408.00716, LicenseCC BY 4.0

[9]. Hotel Recommendation System Based on Customer's Reviews Content Based Filtering Approach,Authors:H. Shah, L. Jacob, Publisher: IEEE (proceedings of the 4th International Conference on Advances in Computing, Communication, Control and Networking – ICAC3N, 2022), DOI: 10.1109/ICAC3N56670.2022.10074228

[10]. A Natural Language Processing Framework for Hotel Recommendation Based on Users' Text Reviews, Authors- Lavrentia Aravani, Emmanuel Pintelas, Christos Pierrakeas, Panagiotis Pintelas, Publisher- 01 Aug 2024 (Cornell University),Journal Article 10.48550/arxiv.2408.00716