Implementation of Chatbot using NLP

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning
with
TechSaksham – A joint CSR initiative of Microsoft & SAP

by

Name of Student:- AYUSH SINGH

Email id:- kayushsingh1@gmail.com

Under the Guidance of

Aditya Prashant Ardak

Master Trainer,

Edunet Foundation

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to everyone who contributed to the successful completion of this project.

First and foremost, I would like to thank **Aditya Prashant Ardak Sir** who was my Project Supervisor for his invaluable guidance, encouragement, and support throughout the project. His expertise and constructive feedback have been instrumental in shaping the direction and outcome of my work.

I am deeply grateful to **AICTE , MICROSOFT , SAP** and **TechSaksham** for their sincere collaboration, dedication, and teamwork, which made tackling the challenges of this project both manageable and enjoyable.

I extend my heartfelt thanks to **Edunet Foundation** for providing the necessary resources, infrastructure, and a conducive environment to carry out this thesis work on my project.

Finally, I am grateful to my family and friends for their unwavering support, understanding, and motivation throughout this journey.

.....

ABSTRACT

Traveling across a diverse country like India requires careful planning, especially when searching for accommodations that meet specific preferences. **"WanderStay"** addresses this challenge by developing an AI-powered chatbot to help users find hotels state-wise and union territory-wise. The primary objective of the project is to simplify the hotel search process, enhance user convenience, and promote tailored travel experiences.

The project employs **Natural Language Processing (NLP)** techniques for intent recognition and machine learning to classify user queries. A dataset of labeled intents, stored in a **JSON file**, forms the basis for the chatbot's training. **Logistic regression** with **TF-IDF vectorization** is used to classify user inputs, while a random response generator adds conversational variety. The chatbot is deployed using a **Streamlit interface**, which allows users to interact easily, save chat history, and access additional information about the system.

Key results demonstrate the chatbot's ability to handle diverse queries, providing curated lists of hotels based on state, city, or type of accommodation. The system also supports storing and reviewing past interactions, ensuring transparency and usability.

The project concludes that **AI-based chatbots** can significantly improve the travel planning experience by offering personalized, efficient, and real-time solutions. Future work includes expanding the dataset, **integrating real-time hotel APIs**, enabling multilingual support, and incorporating advanced machine learning models for greater accuracy and context understanding. By providing a scalable, user-friendly, and inclusive platform, **WanderStay** serves as a promising tool to revolutionize how travelers in India search for accommodations, promoting convenience, accessibility, and sustainable tourism practices.

TABLE OF CONTENT

LIST OF FIGURES

CHAPTER 1

Introduction

1.1Problem Statement:

In this vast growing technological driven world, the advancement of Artificial intelligence(AI) has played a prominent role in the field of Computer Science by solving many of the human challenges within seconds, so as is one of its key product popularly known as (CHATBOTS). Therefore, the major problem of businesses to interact with their customers in an efficient manner without any human representative cost has been a matter of concern, therefore this project majorly focuses on the implementation of an AI chat bot using NLP methodology where customer's queries can be resolved without compromising response, quality and speed.

1.2Motivation:

I have chosen this project where a chat bot will be designed using natural language processing (NLP) because NLP powered chat bots allows us to communicate with computers in a natural human like way, mimicking person to person conversation. It also allows businesses to connects with their customers and categorize their opinions and feedbacks for future references. Lastly this project will mainly help me to learn key concepts of AI and its NLP tools.

1.3Objective:

- To design a chat bot using AI that can seamlessly communicate with the people by understanding their natural language
- A particular chat bot will be developed related to specific field
- Building conversation that resembles to human conversations

1.4Scope of the Project:

Under this project we will design a Retrieval-Augmented Generation (RAG) based AI chat bot using natural language processing which can seamlessly process human natural language, allowing them to interact with digital devices as if they are communicating with a real person. Therefore, this project focuses mainly on to design a chat bot that can understand user intent and provides responses based on predefined appropriate patterns.

CHAPTER 2

Literature Survey

2.1.Review relevant literature or previous work in this domain.

The article, ***"The Impact of Artificial Intelligence on Chatbot Technology: A Study on the Current Advancements and Leading Innovations"*** by ***Farhan Aslam***, offers an insightful exploration of how AI has revolutionized chatbot technology. By delving into advancements in natural language processing (NLP), machine learning, and deep learning, the study highlights how AI-powered chatbots enhance user experiences through personalized interactions and improved contextual understanding.

Key innovations, such as virtual assistants and voice-enabled chatbots such as OpenAI's GPT models like ChatGPT 4, have transformed industries like healthcare, customer service, and education. However, the article emphasizes the challenges these technologies face, particularly regarding ethical concerns and privacy. It advocates for robust frameworks to address data security and transparency issues, particularly in sensitive sectors like healthcare.

Overall, the research provides a balanced view, celebrating technological progress while calling for ethical safeguards to maximize the benefits of AI-powered chatbots responsibly

.This article also mentions several existing models, techniques, and methodologies used to address the challenges and enhance the capabilities of AI-powered chatbots:

1. **Natural Language Processing (NLP) Algorithms**:
    - NLP techniques like Natural Language Understanding (NLU) help chatbots interpret user input by decoding semantic meaning and identifying intents and entities. This enables more accurate and context-aware responses.
    - Tools like **Dialogflow** are used to implement NLP in chatbots for understanding queries and delivering human-like interactions.
2. **Machine Learning Models**:
    - Supervised and unsupervised learning algorithms are applied to train chatbots for improved response generation.
    - These models learn from user interactions to continuously refine chatbot behavior and enhance personalization.
3. **Deep Learning Techniques**:
    - Advanced deep learning architectures like recurrent neural networks (RNNs) and transformers (e.g., OpenAI's GPT models) improve conversation flow and comprehension by analyzing context and generating meaningful responses.
4. **Voice-Enabled and Virtual Assistant Technologies**:
    - Virtual assistants such as **Alexa**, **Google Assistant**, and **Cortana** use a combination of NLP, speech recognition, and machine learning to provide conversational interfaces for various industries.

5. **Sentiment Analysis**:
    - Sentiment analysis is integrated into chatbots to assess user emotions and tailor responses empathetically, enhancing customer satisfaction.
6. **Hybrid AI Systems**:
    - Combining rule-based systems with machine learning models allows chatbots to handle structured and unstructured queries, ensuring robust and adaptive behavior.

These methodologies enable businesses to deploy efficient, scalable, and user-friendly chatbots that bridge the gap between automation and personalized customer interaction.

The above review has been sited from the link below mentioned:-
https://ajpojournals.org/journals/index.php/EJT/article/view/1561?srsltid=AfmBOoqnhtNjzghlLL4T-hOh5500T_bFDX3GsHW-qhcvuxKZDFAhxiP-

2.2. Mention any existing models, techniques, or methodologies related to the problem.

Chat bots have revolutionized the field of AI with extravagant models used in different types of industries .Some of those models are advanced deep learning architectures like recurrent neural networks (RNNs) and transformers (e.g., OpenAI's GPT models) which improve conversation flow and comprehension by analyzing context and generating meaningful responses , virtual assistants such as Alexa, Google Assistant, and Cortana which use a combination of NLP, speech recognition, and machine learning to provide conversational interfaces for various industries. There are also some future-oriented techniques or methodologies that has profoundly impacted the sector of chat bots , such as **Reinforcement Learning** using the technique of Q-learning or policy gradient methods to evaluate response efficiency and quality, secondly **Sentiment Analysis Tools** for example:- VADER, TextBlob and lastly **Pre-Trained Conversational Models** like Google Dialogflow and Microsoft Bot Framework to create customizable chatbots for businesses.

2.3. Highlight the gaps or limitations in existing solutions and how your project will address them.

Existing chatbot solutions often struggle with understanding context, leading to irrelevant or repetitive responses. They may lack emotional intelligence, making interactions feel robotic. Many chatbots are limited by predefined scripts, inhibiting their ability to handle unexpected queries. Additionally, they can struggle with multi-turn conversations, failing to maintain context over longer interactions. Finally, privacy concerns regarding user data handling remain a significant limitation in building trust with users. This project aims at building a conversational coherence between a human and a Chatbot. Through its pre-defined set of patterns, it continuously compares user intents and previous responses from the memory and responds with the most optimal solution to the user's query.

CHAPTER 3

Proposed Methodology

3.1 System Design

The diagram of the Proposed Solution is given below along with its explaination:-

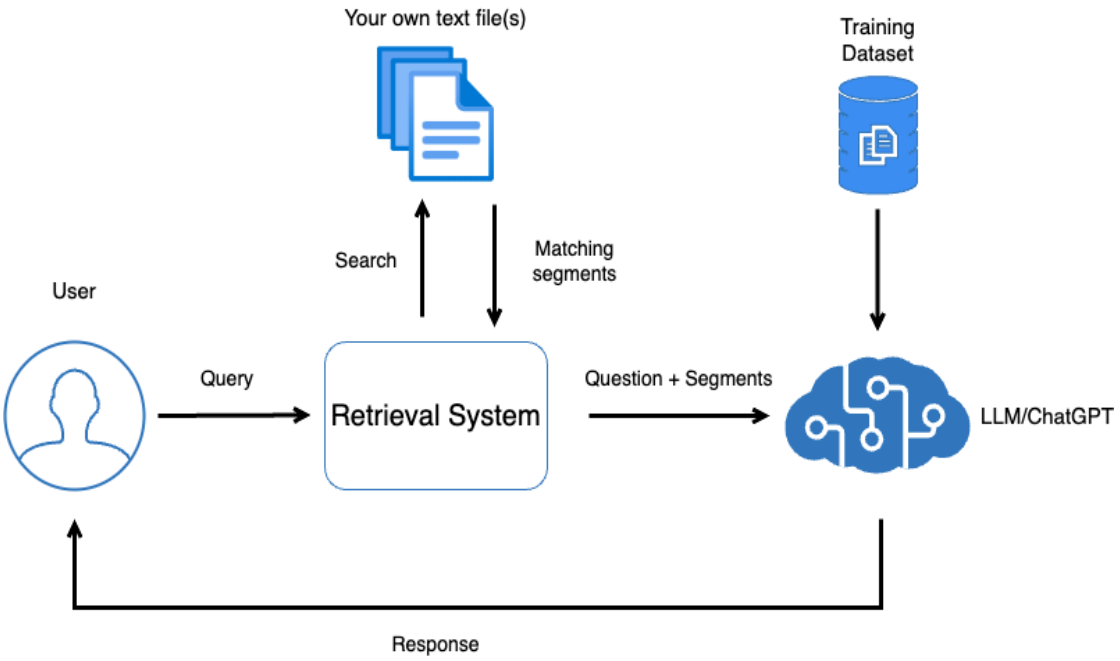

Figure:-3.1.1

DESCRIPTION:- The above ***Fig 3.1.1*** demonstrates a Retrieval-Augmented Generation (RAG) system .In a Retrieval-Augmented Generation (RAG) system, providing relevant information from the retrieval component to the generative language model involves a process of encoding, indexing, and then querying to retrieve the most pertinent data.

- **Encoding Documents into Vectors**: our text file is encoded into a vector representation. This encoding is done using a model that can convert text into a high-dimensional vector space.
- **Indexing Vectors**: Once the text file is converted into vector, it is indexed in a database. This indexing is crucial for efficient retrieval.
- **Query Processing**: When a query or prompt comes from the user, the RAG system first processes this query through the same encoding model that was used for the text file.
- **Vector Similarity Search**: The system then performs a similarity search between the query vector and the document vectors in the indexed database.
- **Retrieving Relevant Information**: Based on the similarity scores, the system retrieves the most relevant information snippets.
- **Providing Data to the Generative Model**: The retrieved snippets or their summaries (also in vector form or as raw text) are then provided to the generative language model.
- **Response Generation**: Finally, the generative model synthesizes a response, taking into account both the original query and the information from the retrieved snippets

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

3.2.1 Hardware Requirements: Laptop/Desktop with minimum i5 core to run the technologies

3.2.2 Software Requirements: They are mentioned as follows:-

- Python to build the chatbot
- NLTK library used to build the Chatbot
- Scikit-learn NLP library package to train the Chatbot
- Streamlit to create and run the application of Chatbot
- JSON for queries data

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:

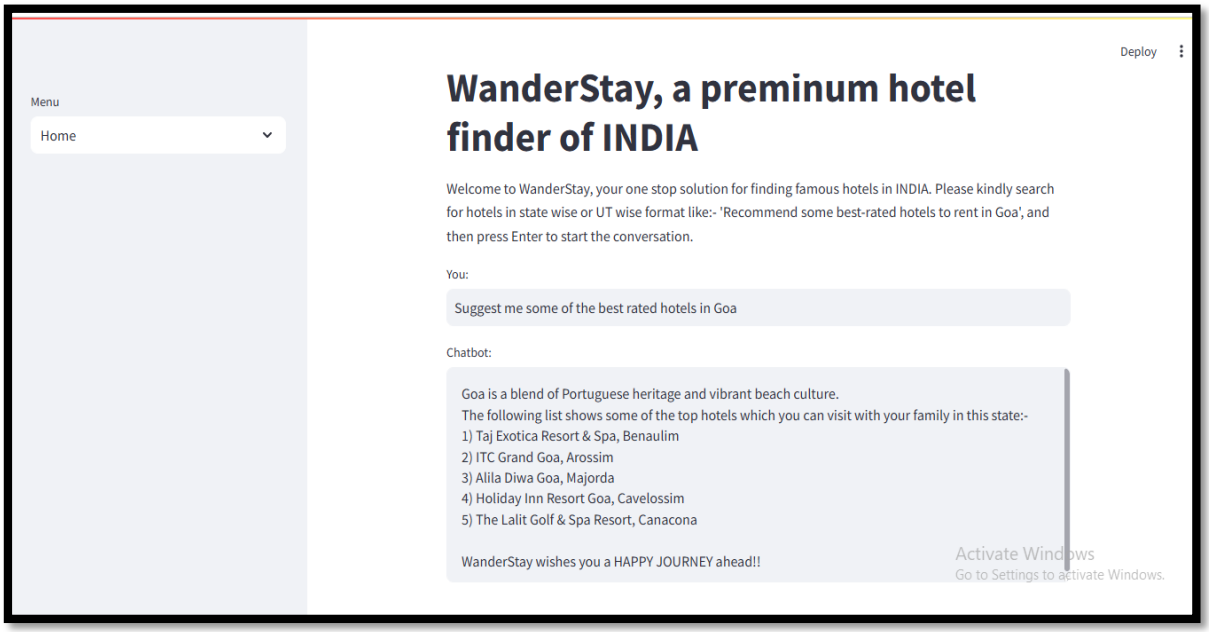

Figure 4.1.1

Description :- The above ***Fig 4.1.1*** represents the snap-shot of the establishment of a Artificial Intelligence(AI) based Chatbot named as "WanderStay", using Natural Language Proccessing(NLP).The above snap-shot shows how this chatbot is seamlessly producing predefined data/patterns as its output on the query of the user input for searching famous hotels in each State or union territories(UT) of our country INDIA.

Given Query:- "Suggest me some of the best rated hotels in Goa"

Chatbot's output:-" Goa is a blend of Portuguese heritage and vibrant beach culture.

The following list shows some of the top hotels which you can visit with your family in this state:-

1) Taj Exotica Resort & Spa, Benaulim

2) ITC Grand Goa, Arossim

3) Alila Diwa Goa, Majorda

4) Holiday Inn Resort Goa, Cavelossim

5) The Lalit Golf & Spa Resort, Canacona

WanderStay wishes you a HAPPY JOURNEY ahead!!”

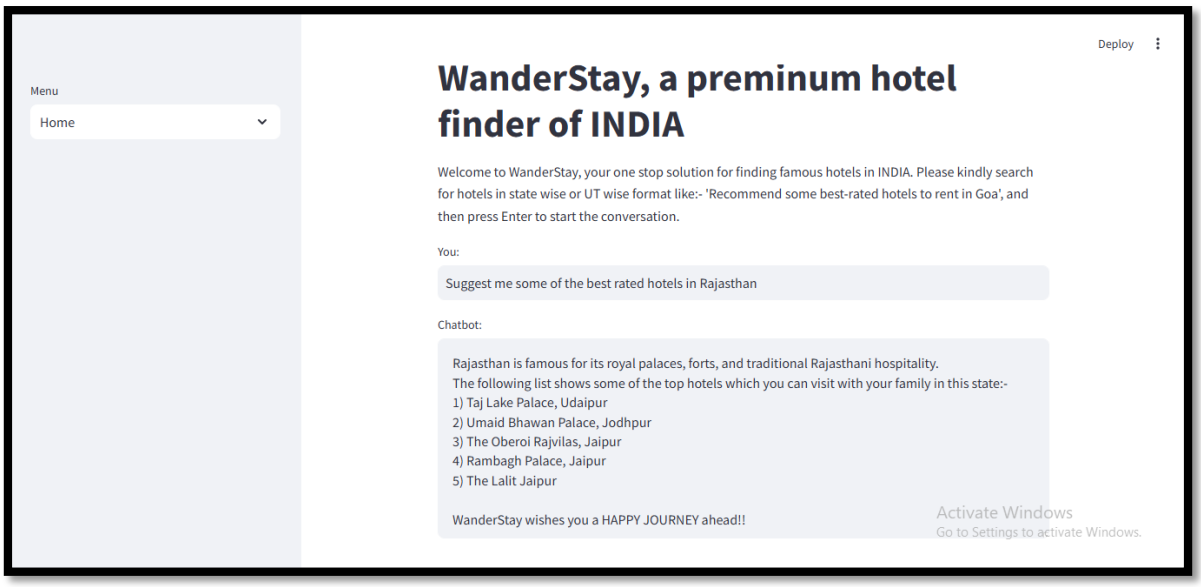

Figure 4.1.2 ***Description:-*** The above ***Fig4.1.2*** shows the following interaction

- **User Query:-**“ Recommend some best rated hotels to rent in Rajasthan”
- **Chatbot’s Response**:-“ Rajasthan is famous for its royal palaces, forts, and traditional Rajasthani hospitality.The following list shows some of the top hotels which you can visit with your family in this state:-

1) Taj Lake Palace, Udaipur

2)Umaid Bhawan Palace, Jodhpur

3) The Oberoi Rajvilas, Jaipur

4) Rambagh Palace, Jaipur

5) The Lalit Jaipur

WanderStay wishes you a HAPPY JOURNEY ahead!!”

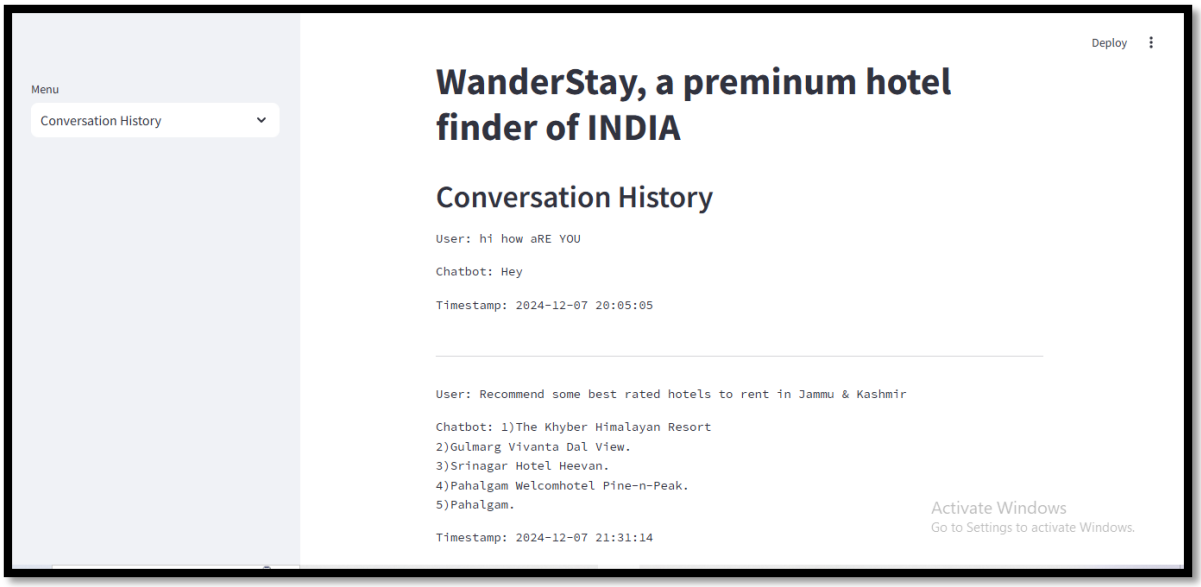

Figure 4.1.3

Description:- The above ***Fig 4.1.3*** explains the **Conversation History** between the user and the chat-bot along with its appropriate **Timestamp** so that the chatbot-‘WanderStay’ can analyze its previous conversations and data-sets using pre-defined AIML models to give relevant responses and predictions in its future conversations.

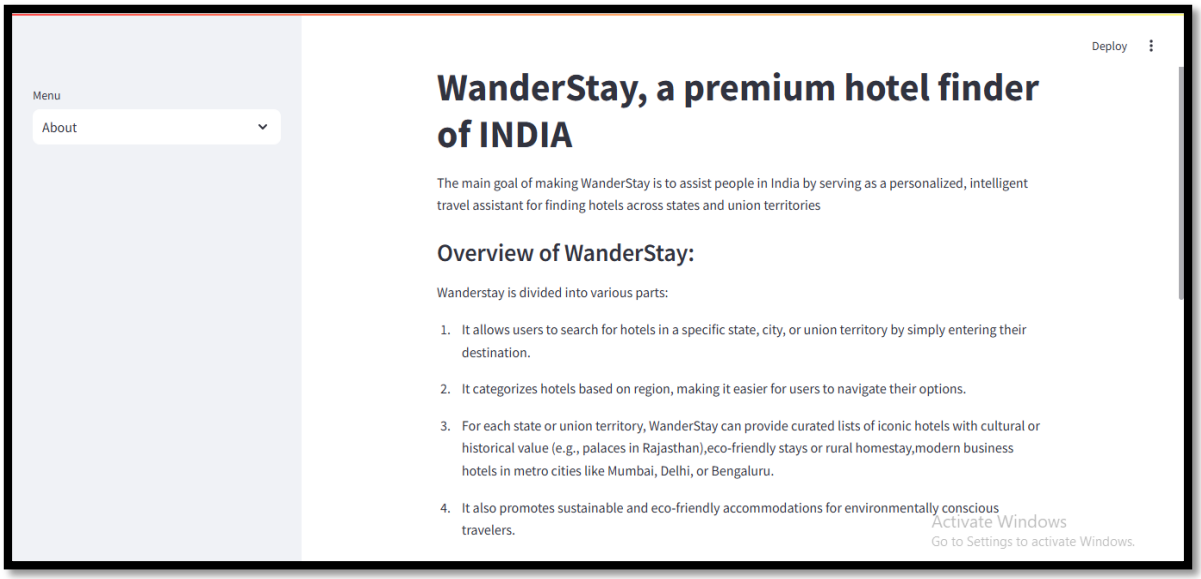

Figure 4.1.4

Description:- The above ***Fig 4.1.4*** explains the main motive of making "WanderStay" for the users of INDIA so that they can freely search for hotels in case of spending a vacation out of station with their utmost convenience.

4.2 GitHub Link for Code:

***GitHub-Link*:-** CLICK-HERE

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

When it comes to the future work of the given AI based chat bot- “WanderStay”, I will like it to collect real user interactions to expand the dataset with additional patterns, tags, and responses. Secondly we can use libraries like **spaCy** or **Hugging Face Transformers** to extract entities (e.g., "Goa" → state, "luxury" → hotel type). Lastly it will be very beneficial if we replace the logistic regression model with a pre-trained transformer model (e.g., **BERT**, **DistilBERT**, or **ChatGPT API**).

There are also some important improvement suggestions which can transform the entire AI chatbot into a robust, scalable, and user-friendly solution for travelers in India and beyond. Some of them are listed below:-

- **Interactive Map Integration**: Allow users to view hotels on a map for better visualization.
- **User Ratings and Reviews**: Enable users to rate hotels and read reviews directly from the app.
- **Itinerary Suggestions**: Suggest nearby attractions and activities along with hotels.
- **Travel Assistance**: Add train/flight booking suggestions and transportation options.

5.2 Conclusion:

The **major conclusion** of developing **WanderStay** is that technology-driven solutions can significantly enhance the travel planning experience by providing **personalized, efficient, and accessible hotel-finding services**, especially in a diverse country like India. WanderStay simplifies the hotel search process by allowing users to query hotels state-wise or UT-wise, reducing the time and effort required to find accommodations. It proves that chatbots can act as effective **travel assistants**, capable of curating information tailored to individual user preferences. India's vast diversity makes travel planning challenging but this project shows that technology can bridge this gap by accommodating regional variations, cultural nuances, and specific travel preferences (e.g., eco-tourism, heritage stays, or business hotels). It highlights the importance of **inclusivity** in digital solutions. It also reflects the growing importance of **AI-driven personalization, real-time insights, and inclusivity in travel planning tools**. By combining conversational AI, machine learning, and a user-friendly interface, this project paves the way for revolutionizing the

way travelers search for accommodations, setting the stage for future innovations in the tourism industry.The success of **"*WanderStay*"** points to a broader trend of using AI in tourism to:

- Empower travelers with **accessible, real-time solutions**.
- Encourage digital adoption in the hospitality industry.
- Demonstrate how AI can improve operational efficiency for businesses while enhancing customer satisfaction.

- **My Stand on the completion of this project:-** Lastly, by the help of this project I have learned major sustainable development of profound AI based Tools(For eg:- NLTK, streamlit etc),which has positively enhanced my skills and confidence in the sector of AIML and have also made me significantly curious to learn much more about python language and its vital libraries.

REFERENCES

[1]. Article Review Source Link:-
https://ajpojournals.org/journals/index.php/EJT/article/view/1561?srsltid=AfmBOoqnhtNjzghlLL4T-hOh5500T_bFDX3GsHW-qhcvuxKZDFAhxiP-

[2]. My GitHub Link For Code:- CLICK-HERE