AI BASED CHATBOT USING NLP

A PROJECT REPORT

Submitted by,

S BHARATH CHANDRA	**-**	**20201CAI0083**
P PAVAN KUMAR REDDY	**-**	**20201CAI0084**
MICHANAGATLA VAMSI	**-**	**20201CAI0120**
TALARI DILEEP KUMAR	**-**	**20201CAI0145**
TAPZUL SAI CHARAN	**-**	**20201CAI0149**

Under the guidance of,

Mr. JOHN BENNET J

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING
(ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING)

At

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the Project report **"AI BASED CHATBOT USING NLP"** being submitted by "Bharath Chandra, Pavan Kumar Reddy, Michanagatla Vamsi, T Sai-Charan, Talari Dileep Kumar" bearing roll number(s) "20201CAI0083, 20201CAI0084, 20201CAI0120, 20201CAI0145, 20201CAI0149" in partial fulfillment of requirement for the award of degree of Bachelor of Technology in Computer Science and Engineering (AI & ML) is a bonafide work carried out under my supervision.

Mr. John Bennet J	**Dr. Zafar Ali Khan**
Assistant Professor	Associate Professor & HOD
School of CSE	School of CSE
Presidency University	Presidency University

Dr. C. KALAIARASAN	**Dr. SHAKKEERA L**	**Dr. MD. SAMEERUDDIN KHAN**
Associate Dean	Associate Dean	Dean
School of CSE&IS	School of CSE&IS	School of CSE&IS
Presidency University	Presidency University	Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled "**AI BASED CHATBOT USING NLP"** in partial fulfilment for the award of Degree of **Bachelor of Technology** in **Computer Science & Engineering(AI & ML)**, is a record of our own investigations carried under the guidance of **Mr. John Bennet J, Assistant Professor, School of Computer Science & Engineering, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

NAME	ROLL NUMBERS	SIGNATURE
S BHARATH	20201CAI0083	
PAVAN KUMAR	20201CAI0084	
M VAMSI	20201CAI0120	
DILEEP KUMAR	20201CAI0145	
T SAI CHARAN	20201CAI0149	

ABSTRACT

A chatbot is an innovative tool powered by artificial intelligence, designed to simulate human interaction. By interpreting information and providing responses in either written or spoken form, chatbots have become indispensable in the realm of digital communication. Our revolutionary concept delves into the untapped potential of chatbot communication, paving the way for further research in this field. While most chatbots efficiently complete tasks, one recurring issue remains: the conversations can be repetitive. To tackle this, our project introduces a multi-agent system where a chatbot acts as a mediator between the user and the external world. What sets our chatbot apart is its unique ability to comprehend the user's needs and tailor responses accordingly, making technology even more efficient in minimizing human effort.

In today's world, advanced technology primarily operates through the collaboration of artificial intelligence, NLP processing, and machine learning algorithms. An essential aspect of this complex network is the use of artificial intelligence to simulate human decision-making and offer various services. As such, this paper presents a survey of chatbots and their role in artificial intelligence research. It examines the different platforms utilized to construct chatbots and their ability to serve a diverse range of users. Additionally, the design techniques employed to create chatbots vary depending on the specific services intended for the users.

The chatbot utilizes a knowledge base to compare the user's input sentence with existing patterns. When users ask questions, the chatbot uses these patterns to understand the question and provide an appropriate response. This is achieved by transforming the English sentence into a format that can be interpreted by the machine, then searching through relevant data to retrieve the necessary information and presenting it in a conversational manner. The chatbot is also equipped with the ability to identify commonly asked questions that it cannot answer, and it will notify an administrator for their input. Once the administrator provides a response, the chatbot will learn and improve, actively searching for new questions and corresponding answers to provide to users.

The chatbot's cleverness lies in its ability to combine rule-based matching and a pre-trained GPT-2 language model to generate responses. These rules and their corresponding answers are stored in a convenient JSON file. Each time a user sends a message, the chatbot first scans these rules for a match. If it finds one, it instantly reply back with the corresponding response from the JSON file. If

no match is found, the GPT-2 model seamlessly takes over to generate a response. To accurately identify key words from the user's input, the chatbot utilizes the powerful Spacy library. These words are then used to search for matching rules. Additionally, the chatbot scores potential responses based on how closely they align with the user's message, using the TF-IDF vectorization and cosine similarity method. This ensures that the chatbot's responses are both relevant and engaging

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected Dean **Dr. MD. Sameeruddin Khan**, Dean, School of Computer Science & Engineering, Presidency University for getting us permission to undergo the project.

We record our heartfelt gratitude to our beloved Associate Deans **Dr. C. Kalaiarasan and Dr. Shakkeera L,** School of Computer Science & Engineering, Presidency University for rendering timely help for the successful completion of this project.

We would like to convey our gratitude and heartfelt thanks to the University Project-II Coordinators **Dr. Sanjeev P Kaulgud, Dr. Mrutyunjaya MS** and also the department Project Coordinators.

We are greatly indebted to our guide **Mr. John Bennet J, Assistant Professor**, School of Computer Science & Engineering, Presidency University for her inspirational guidance, valuable suggestions and providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

S BHARATH

PAVAN KUMAR

M VAMSI

DILEEP KUMAR

T SAI CHARAN

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER-1
INTRODUCTION

A chatbot, also known as a chat robot, is an ingenious software that harnesses the power of artificial intelligence (AI) to engage in natural language conversations with users. Its purpose is to respond to queries, share pertinent information, and even carry out tasks, all while conversing in a relatable manner. From rudimentary rule-based systems to advanced AI-driven models, chatbots have undergone a remarkable evolution, enabling them to grasp context, learn from interactions, and deliver personalized responses.

1.1 Significance of Chatbots:

Chatbots are available 24/7, ensuring that users can access information or receive assistance at any time. Additionally, they are cost-effective by automating tasks that are typically managed by human agents. Chatbots also offer prompt responses, improving user satisfaction and efficiency. As user demand increases, chatbots are able to handle a higher volume of interactions without requiring additional resources. They also deliver consistent responses, ensuring that users receive uniform information and support. Furthermore, chatbots collect valuable user data, providing insights that can inform business strategies and decisions.

1.2. Challenges in Chatbots:

1.2.1. Natural Language Understanding (NLU):

Achieving accurate comprehension of user intent and context remains a complex task, especially with diverse language nuances and colloquialisms.

1.2.2. Training Data Limitations:

AI chatbots rely heavily on training data. Inadequate or biased datasets can lead to incorrect responses or reinforce existing biases.

1.2.3. Scalability:

While chatbots can handle increased volumes of interactions, ensuring consistent performance and responsiveness at scale remains a challenge.

1.2.4. Ethical Considerations:

Addressing privacy concerns, data security, and ensuring transparency in bot interactions are essential but complex ethical challenges.

1.2.5. Maintenance and Updates:

Continuous monitoring, updating, and refining chatbot models to adapt to changing user needs and technological advancements require significant resources.

1.3. Role of Algorithms in Chatbots:

Natural Language Processing (NLP) is a powerful tool that allows chatbots to understand, interpret, and generate human-like responses from written inputs. These algorithms, including popular techniques like word embeddings (e.g., Word2Vec, Glove), sequence-to-sequence models (e.g., LSTM, Transformer), and attention mechanisms, form the backbone of advanced language understanding and generation in NLP. On the other hand, Machine Learning (ML) plays a vital role in equipping chatbots with the ability to learn from data and improve their accuracy and responsiveness over time. By utilizing techniques such as supervised learning, where chatbots are trained on labeled data, reinforcement learning for making dynamic decisions, and unsupervised learning for exploring and categorizing data, ML enables chatbots to continuously evolve and adapt to the needs of their users

Classification and prediction algorithms are key components in guiding chatbot responses and actions. By analyzing historical data and current interactions, these algorithms can classify user intents and predict user behavior. This not only improves the chatbot's effectiveness, but also increases user engagement. In addition, recommendation systems employ algorithms to analyze user preferences and behaviors, enhancing personalization and further engaging the user. Search algorithms are also utilized by chatbots, allowing them to retrieve relevant information from databases or knowledge bases in response to user queries. And to optimize chatbot performance, optimization algorithms are employed to balance resources, response time, and user satisfaction, especially in complex environments or those with multiple agent systems.

CHAPTER-2

LITERATURE SURVEY

Table 2.1 Summary of Existing Methods

AUTHOR	TOPIC	SUMMARY	YEAR	ADVANTAGES	DISADVANTAGES
1. Elikkiya, M Varun, P.S Mohammed Rizwan [1]	Self-Learning Conversational AI Chatbot Using Natural Language Processing	A self-learning conversational AI chatbot using Natural Language Processing (NLP) is a sophisticated system designed to engage in human-like conversations, continuously improving its performance through machine learning. NLP enables the chatbot to understand and generate human language, making interactions more natural and meaningful. The self-learning aspect implies that the chatbot can adapt and enhance its capabilities over time without manual intervention.	2022	NLP-powered self-learning chatbots provide more natural and efficient interactions.	Collection and storage of conversational data may raise privacy issues if not handled securely.
2. Parth Thosani, Jaydeep Vaghasiya, Manas Sinkar, Radha[2]	A self Learning Chat-Bot from User Interactions and Preferences	A self-learning chatbot that evolves based on user interactions and preferences utilizes machine learning to adapt and improve its responses over time, providing a personalized and	2020	The chatbot tailor's responses based on user interactions and preferences, enhancing the overall user experience.	The effectiveness of the chatbot depends on the diversity and quality of user interactions, and limited or biased data may lead to

		efficient conversational experience.			suboptimal learning outcomes.
3. Mohammad Nurazzaman, Omar Khadeer Hussain [3]	A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks	This survey explores the implementation of chatbots in the customer service industry using deep neural networks. It investigates the use of advanced technology to enhance customer interactions, focusing on the benefits and drawbacks associated with employing deep learning techniques for chatbot development in customer service.	2018	Enhanced customer service efficiency through deep neural network-powered chatbots.	Despite advancements, Chatbot may struggle with nuanced language and complex context, leading to occasional misinterpretatio-ns.
4. Aishwarya Gupta, Divya Hathwar, Anupama Vijaykumar [4]	Introducton to AI Chabot	AI ChatBots signifies the application of artificial intelligence in developing conversational agents, capable of understanding and responding to user queries, with the aim of enhancing user interactions and automating various tasks.	2021	AI ChatBots streamline processes and tasks, automating interactions and providing quick and efficient responses to user inquiries.	Despite advancements, ChatBots may struggle with nuanced language and complex context, leading to occasional Misinterpretatio-ns.

5. Prasnurzaki Anki, Alhadi Bustamam, Herley Shaori, Devvi Sarwinda [5]	High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model	The proposed conversational AI chatbot aims to achieve high accuracy and contextual understanding by leveraging deep learning techniques, specifically the BiLSTM model. The utilization of BiLSTM allows the chatbot to consider both past and future context, enhancing its ability to comprehend and respond effectively in conversational scenarios.	2020	The utilization of a BiLSTM-based deep recurrent neural network in the conversational AI chatbot promises high accuracy, improved contextual understanding, and enhanced fluency in responses	Greedy Decoding Algorithm not always guarantee to find the optimal Solutions
6. M.Ganesan, Deepika C, Harievashini B, Krithikha A S, Lokharatchana B [6]	A Survey On Chatbots Using Artificial Intelligence	This survey systematically explores the landscape of AI-driven chatbots, offering valuable insights into their development, applications, and challenges, thereby contributing to a holistic understanding of the evolving field.	2020	Provides a comprehensive overview of the current state-of-the-art, facilitating insights into trends, challenges, and advancements in the field.	Complexity, Over-fitting, takes large time to train a model.

7. Manik Rakhra, Gurram Gopinadh, Narasimha Sai Addepalli [7]	E-Commerce Assistance with a Smart Chatbot using Artificial Intelligence	The implementation of a smart chatbot in e-commerce leverages artificial intelligence to provide personalized and efficient customer support, contributing to an enhanced user experience by addressing queries, resolving issues, and ensuring accessibility to services.	2017	Provides streamlines customer interactions, enhancing user experience through personalized assistance, efficient problem resolution, and improved service accessibility.	Difficulty in interpretation and debugging
8. Lucas Kohnke[9]	A Pedagogical Chatbot: A Supplemental Language Learning Tool	As a supplemental language learning tool, the pedagogical chatbot presents advantages through personalized and interactive support, promoting engagement and offering immediate feedback. However, challenges include the difficulty in replicating human nuance and potential limitations in contextual understanding,	2022	Offers personalized and interactive support to learners, fostering engagement, and providing immediate feedback for effective language acquisition.	Chatbots may struggle to replicate the nuanced and dynamic nature of human communicatio n, potentially hindering the development of certain language skills.

		emphasizing the need for a balanced approach in language learning.			

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

1. Automatic Speech Recognition (ASR):

ASR technology is used to convert spoken words into written text. This involves analyzing the audio input and transcribing it into a format that the chatbot can understand.

2. Natural Language Understanding (NLU):

Chatbots utilize NLU to understand user input and extract relevant information. For task automation, NLU is crucial for identifying user intent and key entities related to the task. NLU systems often use techniques like intent recognition and entity extraction.

3. API Integration:

Chatbots can integrate with external APIs (Application Programming Interfaces) that provide real-time data. These APIs could be related to weather, news, financial data, stock prices, or any other information that is subject to frequent updates. By connecting to these APIs, the chatbot can fetch the latest information in real time.

4.Entity Validation:

Check the validity of entities and prompt the user for corrections if needed. For example, if a date is extracted, verify the format and ask for clarification if it is unclear.

5.Contextual Understanding:

The development of chatbots that can engage in extended, coherent conversations and understand complex user intents

6. Hybrid Approaches:

Integrates different approaches to take advantage of their strengths by integrating rule-based programming, machine learning, and NLP.

7. Generative Models:

Uses deep learning models such as recurrent neural networks (RNNs) or transformers (e.g., GPT) to derive appropriate responses as context.

8. Memory Networks:

Incorporates memory-enhancing neural networks to store and retrieve information, thereby improving contextual understanding.

9. Rule-Based Systems:

Uses predefined rules and decision trees to guide the chatbot's responses based on keywords or specific patterns.

10. Transfer Learning:

Uses a pre-trained language framework that can be fine-tuned for specific chatbot tasks.

CHAPTER-4

PROPOSED METHODOLOGY

The choice of methods for an AI-based chatbot using Natural Language Processing (NLP) depends on various factors, including the project's goals, data availability, and technical constraints. Here are some proposed methods commonly used in developing such chatbots:

1. Named Entity Recognition (NER):

Named Entity Recognition (NER) is a critical component in herbal language processing that specializes in figuring out and classifying entities inside text. Entities can consist of names of humans, organizations, places, dates, numerical values, and other precise types of facts. In the context of developing an AI-based chatbot the usage of NLP, extending NER involves incorporating advanced techniques and considerations to decorate the chatbot's understanding and reaction talents.

2. Accessibility:

Ensuring accessibility is a crucial aspect of developing an AI-based chatbot, especially in compliance with principles of inclusivity and usability for users with diverse abilities.

3. Pre-Trained Models:

Pre-trained language models play a crucial role in the development of AI-based chatbots using Natural Language Processing (NLP). These models, often trained on large corpora of text data, have learned contextual representations of language and can be fine-tuned for specific tasks.

4. API Integration:

API (Application Programming Interface) integration within the context of an AI-based chatbot using NLP is a vital element that extends the functionality and effectiveness of the chatbot. Integrating diverse APIs permits the chatbot to get entry to external services and statistics, perform moves, and provide dynamic, real-time responses.

CHAPTER-5

OBJECTIVES

1. Conversational Empathy:

Design the chatbot to recognize and respond empathetically to user emotions, creating a more human-like and understanding interaction.

2. Empathetic Interaction:

Enable the chatbot to recognize and respond empathetically to customers' emotion for an extra compassionate interplay.

3. User-Friendly Conversations:

Prioritize person-friendliness to make interactions with the chatbot feel natural, intuitive, and enjoyable.

4. Adaptable Tone of Voice:

Develop a dynamic tone of voice that adapts to users' moods and possibilities, fostering a customized and relatable communique.

5. Cultural Sensitivity:

Ensure the chatbot is culturally sensitive, respecting various backgrounds and communique nuances.

6. Proactive Assistance:

Empower the chatbot to offer proactive assistance, looking forward to user wishes and imparting applicable facts or help.

7. Interactive Learning:

Incorporate gamification and interactive elements to make learning thru the chatbot engaging and exciting.

8.Humor Integration:

Infuse the chatbot with humor and wit to create a fantastic and light-hearted conversational revel in.

9. Visual Storytelling:

Utilize visual storytelling elements to carry statistics in a extra enticing and relatable manner.

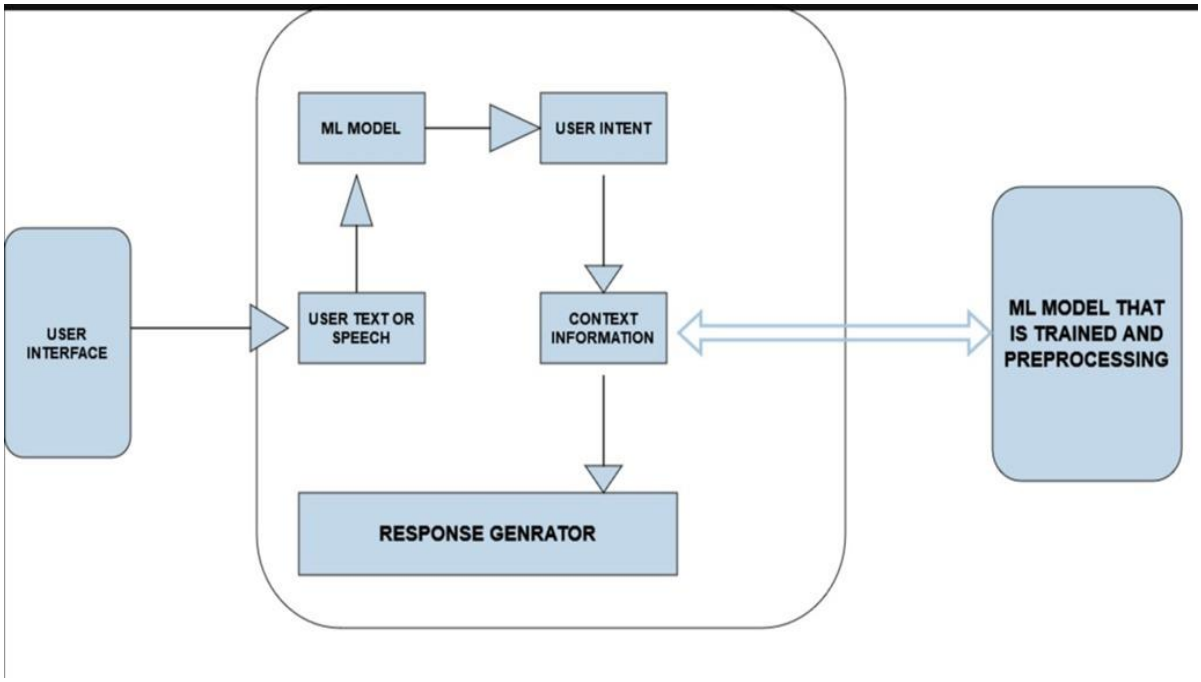
10.Educational Support:

Provide users with interactive educational content thru the chatbot, fostering a studying enjoy inside the verbal exchange.

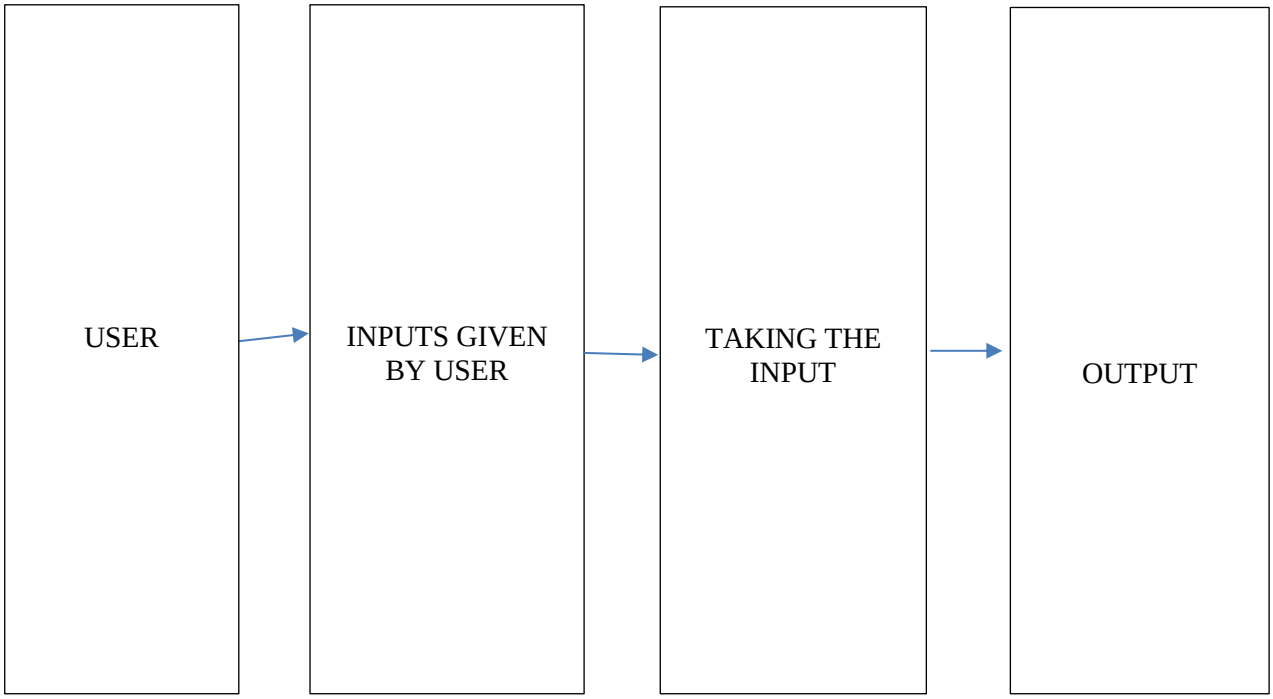
CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

6.1 ARCHITECTURAL DESIGN:



6.2 INTERFACE DESIGN:



6.3 ALGORITHMS:

1. Cosine similarity:

The Cosine similarity algorithm stands out as a widely recognized method in the field of Machine Learning. It assesses the likeness between two vectors in an inner product space, by gauging the cosine of the angle formed between them. This technique is frequently employed in text analysis to measure the resemblance between documents. By considering numerous features, such as the frequency of certain words or phrases, a document is transformed into a term-frequency vector, depicting its unique characteristics. This vector serves as a representation of the document.

How Cosine similarity Work:

- **Text Representation:**

Each piece of textual content, whether it's a consumer query or a predefined response, desires to be represented as a vector. This is generally executed by using converting the text into numerical values that seize the significance of every word inside the context of the whole dataset. Techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or phrase embeddings (along with Word2Vec, Glove, or BERT embeddings) are typically used for this purpose.

- **Vectorization:**

The selected vectorization method assigns numerical values to words or phrases within the textual content. Each size of the resulting vector represents a particular function or word within the textual content. The process includes accounting for the frequency, significance, or context of every phrase.

- **User Input Processing:**

When a user enters a query, the chatbot applies the equal vectorization method used for the predefined responses to convert the user enter right into a numerical vector. This vector represents the semantic content of the person's question in the identical excessive-dimensional space.

- **Cosine Similarity Calculation:**

The cosine similarity among the person enter vector and every predefined response vector is calculated the usage of the cosine similarity method:

$$\text{cosine\_ similarity (user\_ input, response)} = \frac{\text{user\_ input. response}}{||\text{user\_ input}|| \cdot ||\text{response}||}$$

- **Response Selection:**

The predefined reaction with the highest cosine similarity to the user input vector is chosen as the bot's reaction. The cosine similarity value tiers from -1 (diverse) to 1 (equal). The reaction with the best similarity suggests that it aligns most closely with the semantic content of the person's query.

Advantages of Cosine Similarity:

- **Scale Invariance:**

Cosine similarity isn't affected by the magnitude of the vectors however as a substitute focuses on the orientation. This belongings makes it appropriate for comparing documents or text irrespective of their length or length.

- **Efficiency:**

Computing cosine similarity is computationally efficient, particularly for high-dimensional facts. It includes simple arithmetic operations, making it suitable for real-time programs.

- **Dimensionality Reduction:**

Cosine similarity facilitates dimensionality reduction via focusing at the course of vectors in preference to their magnitude. This property is useful in lowering the computational complexity and noise in high-dimensional records.

- **Angle Measurement:**

Cosine similarity measures the cosine of the perspective between vectors, supplying a geometric interpretation of similarity. It quantifies the similarity in terms of the orientation of vectors, taking pictures semantic or thematic similarities between items.

- **Sparse Data Handling:**

Cosine similarity is effective in handling sparse facts representations, together with text records, where maximum elements are 0 (e.g., term-frequency vectors). It makes a speciality of non-zero dimensions, emphasizing relevant functions or phrases.

- **Versatility:**

Cosine similarity is flexible and applicable across numerous domain names, such as textual content mining, photo retrieval, collaborative filtering, and clustering. Its flexibility and interpretability make it a widely adopted similarity metric.

- **Robustness to Outliers:**

Cosine similarity is robust to outliers or noise in information since it emphasizes the connection between vectors rather than man or woman records points. This asset enhances the stableness and reliability of similarity measurements

Limitations:

- **Lack of Term Semantics:**

Cosine similarity treats every time or function independently and does no longer capture the semantic relationships among terms. It can also conflict to recognize context and meaning, specifically in natural language processing duties.

- **Sensitivity to Document Length:**

Cosine similarity can be touchy to the duration of files. Longer documents tend to have higher similarity rankings, as they incorporate extra terms. Normalization strategies or opportunity similarity metrics can be had to cope with this problem.

- **Difficulty with Negative Context:**

Cosine similarity on my own may not effectively handle terrible context. For instance, it may not distinguish between high-quality and terrible sentiments in text. Additional strategies, together with sentiment evaluation fashions, may be essential to deal with this predicament.

- **Impact of Stop Words:**

Common phrases (stop words) that seem frequently however bring little semantic meaning can disproportionately have an effect on cosine similarity. While prevent phrases may be eliminated in the course of preprocessing, their effect might also nevertheless persist in certain eventualities.

- **Vocabulary Mismatch:**

If the vectors are built the use of special vocabularies or have versions in terminology, cosine similarity may not efficaciously measure similarity. Techniques like stemming or lemmatization can help deal with this assignment.

- **Limited Context Understanding:**

Cosine similarity does not inherently capture complicated relationships or context dependencies in records. In tasks that require a deep expertise of context, extra superior models or embedding strategies can be greater appropriate.

Applications:

- **Information Retrieval:**

In search engines like google, cosine similarity is used to determine the relevance of a file to a question. Documents with vectors closest in course (i.e., highest cosine similarity) to the query vector are considered most applicable.

- **Text Mining and Natural Language Processing (NLP):**

Cosine similarity is employed in duties which includes report clustering, text categorization, and information retrieval within big textual datasets. It allows pick out semantically comparable files or sentences.

- **Recommendation Systems:**

In collaborative filtering and content material-based totally advice structures, cosine similarity is applied to compute similarities among person profiles or objects, facilitating personalized pointers.

- **Image Retrieval:**

In image processing and computer vision, cosine similarity may be used to evaluate picture feature vectors, such as those derived from deep learning fashions, to pick out visually similar photos.

- **Clustering and Classification:**

Cosine similarity serves as a distance metric in clustering algorithms like k-means, spectral clustering, and hierarchical clustering. It helps group comparable statistics points or objects into clusters primarily based on their feature vectors.

2. TF-IDF:

TF-IDF (Term Frequency-Inverse Document Frequency) is a statistical degree used to assess the significance of a phrase in a report, that is a part of a corpus (series of files). When implemented to the context of a chatbot, TF-IDF can be an effective device for information and processing consumer queries.

Let us break down how it works, especially in the context of a chatbot:

- **Term Frequency (TF):**

This measures how regularly a term seems in a file. In a chatbot scenario, a report may be a user's input or a pre-defined reaction the bot has. The idea is that the more frequently a word appears in a document, the extra essential it is miles. However, this is normalized to save you bias in the direction of longer texts.

- **Inverse Document Frequency (IDF):**

IDF measures the importance of the term throughout a fixed of files. It assesses whether a term is commonplace or uncommon across all files. The good judgment is if a phrase seems in lots of documents, it is now not a completely unique identifier. IDF decreases the load of terms that arise very regularly inside the corpus and will increase the burden of phrases that arise hardly ever.

How TF-IDF Work:

- **Text Preprocessing:**

Initially, the text records, which incorporates each the user's input and the chatbot's ability responses, undergoes preprocessing. This involves cleaning the text (eliminating punctuation, changing to lowercase), and often putting off common but less informative phrases, called forestall phrases (like "the", "is", "at").

- **Term Frequency Calculation (TF):**

The chatbot calculates the Term Frequency for every word in every report. A file, in this case, might be a person query or a reaction stored inside the chatbot’s database. Term Frequency is the ratio of the quantity of times a selected word seems in a document to the total quantity of words in that file. It gives a measure of the importance of the word in the individual report.

- **Inverse Document Frequency Calculation (IDF):**

Next, the Inverse Document Frequency is calculated for each word throughout all files. IDF measures how commonplace or rare a phrase is throughout all documents (in this example, all consumer queries and chatbot responses).IDF is calculated by dividing the entire variety of documents by means of the number of documents containing the word, after which taking the logarithm of this quotient.

- **TF-IDF Calculation:**

The TF-IDF fee is received by using multiplying the TF and IDF values. This gives a weight to every word, indicating its significance in a report relative to the whole corpus.

- **Vectorization:**

Each document is then converted into a vector. Imagine a multi-dimensional space in which every measurement corresponds to a word from the corpus. The value in every dimension is the TF-IDF score of that word for the given file.

As a result, each the consumer's question and the chatbot's responses are transformed into vectors in this multi-dimensional space.

- **Determining Relevance:**

When a person inputs a query, the chatbot converts this query into its TF-IDF vector shape. It then compares this vector with the vectors of its capability responses. The contrast is frequently performed the usage of cosine similarity, in which the chatbot measures the cosine of the attitude between the query vector and the response vectors. The smaller the angle, the higher the similarity.

- **Response Selection:**

The chatbot ranks the responses primarily based on their similarity scores and selects the top-ranking reaction(s) to answer to the user.

Advantages of TF-IDF:

- **Enhanced Relevance Detection:**

TF-IDF excels in figuring out the maximum applicable words or terms in a consumer's query. By weighing the frequency of terms in opposition to their presence inside the complete corpus, it may successfully pinpoint which terms are maximum tremendous in a specific context, thereby aiding the chatbot in knowledge the person's precise needs or questions.

- **Effective Handling of a Wide Range of Topics:**

Since TF-IDF is based at the actual utilization of phrases within the corpus, it naturally adapts to the subjects covered within the facts. This way a chatbot the usage of TF-IDF can efficaciously take care of queries on a huge range of subjects, provided its information base is satisfactorily diverse.

- **Improved Accuracy in Response Selection:**

By changing textual content into numerical values, TF-IDF allows for extra goal and precise comparisons among the consumer's enter and capability responses. This quantitative technique can improve the accuracy of the chatbot in choosing the most suitable reaction.

- **Scalability and Flexibility:**

TF-IDF can be applied to text corpora of honestly any length, making it a scalable solution for chatbots that may want to process huge amounts of records. It’s additionally bendy in that it could be used together with different NLP strategies to enhance overall performance.

- **Reduced Bias Towards Common Words:**

The IDF aspect of the algorithm reduces the significance of phrases that arise too frequently across the corpus, which might be much less informative. This facilitates in preventing common words from overshadowing extra significant, context-particular terms in a user's question.

- **Simplicity and Efficiency:**

The set of rules is exceedingly simple and computationally efficient, making it a practical preference for real-time applications like chatbots wherein brief processing of person inputs is essential.

- **Ease of Implementation:**

TF-IDF has been a staple in text processing for a long term and is properly-documented, with several sources and implementations to be had. This ease of get right of entry to makes it a handy preference for builders.

Limitations:

- **Lack of Semantic Understanding:**

TF-IDF does not seize the semantic relationships among words. It treats every phrase as an unbiased entity, because of this that words with similar meanings might not be as it should be accounted for in similarity calculations.

- **Sensitivity to Term Variations:**

TF-IDF won't deal with versions in phrases well. For instance, plurals or exclusive verb bureaucracy may be treated as distinct phrases, lowering the set of rules potential to understand their semantic equivalence.

- **Inability to Handle Synonyms:**

Synonyms are dealt with as separate phrases in TF-IDF, that could result in a loss of flexibility in spotting similar or interchangeable words. This drawback impacts the set of rules potential to apprehend variations in person queries.

- **Doesn't Capture Word Order:**

TF-IDF does now not remember the order of phrases in a report. In herbal language, word order is crucial for understanding meaning, however TF-IDF treats documents as baggage of words, dismissing their sequential association.

- **Fixed Knowledge Base:**

TF-IDF relies closely on the corpus it is skilled on. If new terms or standards are added that were now not part of the schooling information, TF-IDF might not efficiently handle them, limiting its adaptability to evolving content material.

- **Difficulty with Short Texts:**

TF-IDF may also warfare while coping with brief texts or queries, as the frequency of terms becomes less dependable with restricted text. This can lead to much less accurate tests of term importance.

- **Bias Towards Content-Length:**

Longer files may have better TF-IDF ratings clearly because of the extra variety of words. This can lead to biased consequences favoring longer files over shorter ones.

- **Scalability Issues:**

In very massive corpora, calculating TF-IDF values for each term throughout all files can be computationally pricey and might come to be impractical for real-time programs or large-scale systems.

- **Not Effective for Sentiment Analysis:**

When it involves sentiment evaluation or understanding the emotional tone of a textual content, TF-IDF alone may not be sufficient. It doesn't capture the nuances of sentiment or emotional context**.**

- **Difficulty with Homonyms:**

Homonyms (phrases with the equal spelling but distinctive meanings) are dealt with as identical by way of TF-IDF, doubtlessly main to confusion in expertise person queries.

Applications:

- **Information Retrieval:**

TF-IDF is fundamental to steps. It allows perceive the maximum applicable documents primarily based at the consumer's question by way of giving higher weights to terms which are common within the record but rare inside the complete corpus.

- **Document Classification:**

TF-IDF is hired in file classification duties, consisting of unsolicited mail filtering, sentiment analysis, and topic categorization. It allows identify key phrases that make a contribution to distinguishing between one-of-a-kind classes.

- **Keyword Extraction:**

TF-IDF aids in extracting keywords and key phrases from documents, that's beneficial for growing concise summaries and knowledge the main topics blanketed in a textual content.

- **Information Extraction:**

TF-IDF may be used as a feature for figuring out named entities (e.g., persons, groups, places) in a text.

- **Content-Based Recommender Systems:**

In e-commerce, TF-IDF can be used to research product descriptions and endorse comparable gadgets primarily based at the text.

- **Text Clustering:**

TF-IDF is frequently employed to organization comparable files collectively in unsupervised learning situations, facilitating record clustering.

6.4 Frameworks:

6.4.1Tkinter:

Tkinter is a standard GUI (Graphical User Interface) toolkit for Python. It is used to create the graphical user interface for the chatbot application.

The Uses of Tkinter are:

- **Creating the main window:**

Tkinter is used to create the main application window that acts as the core for all GUI elements. This window provides users with a visual interaction with the chatbot.

- **Featured Information:**

The text widget in Tkinter is used to display the conversation between the user and the chatbot. This widget provides messages and replies to notify the user of the ongoing connection.

- **Contact buttons:**

Tkinter's Button widget is used to create buttons that trigger specific actions. For example, there are buttons for sending messages (self. send_ button) and for activating the microphone (self. mic_ button). These buttons provide users with a way to initiate actions.

- **Maintenance Management:**

Tkinter layout management options, such as pack and place, are used to place and organize widgets in the main window. This determines the look and feel of the chatbot interface.

6.4.2 Speech Recognition:

The Speech Recognition library is used to handle speech input from a microphone. It allows the chatbot to recognize and process spoken words. The Uses of Speech Recognition are:

- **Voice input for messages:**

Users can speak their message instead of writing it. The chatbot uses speech recognition to convert spoken words into text, enabling natural and hands-free communication.

- **Advanced Manufacturing:**

Speech recognition provides better access to chatbots, especially for users who have navigation-related issues or prefer voice input over text

6.4.3 Sklearn (scikit-learn):

Scikit-learn is a machine learning library in Python. It includes TF-IDF Vectorizer for converting a collection of raw documents into a matrix of TF-IDF features and cosine similarity for calculating the cosine similarity between vectors. The Uses of Sklearn are:

- **TF-IDF Vectorization:**

Scikit-learn's TF-IDF Vectorizer is used to transform the collected data (questions and candidate responses) into a matrix of TF-IDF (Term Frequency-Inverse Document Frequency) features and then uses this matrix to calculate the cosine similarity between the vectors.

- **Cosine similarity calculation:**

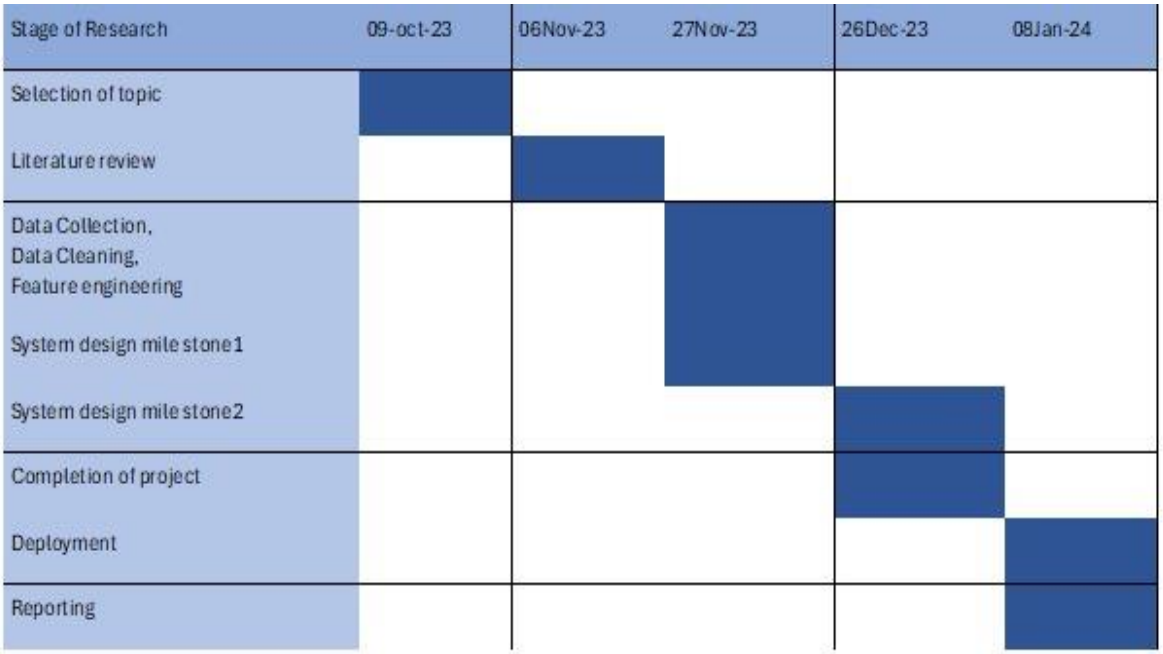
The cosine_ similarity function from scikit-learn is used to calculate the cosine similarity between vectors. In this case, it is used to measure the similarity between the user's input candidate and the answer list.

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT

(GANTT CHART)

Fig 7.1 Gantt chart:



CHAPTER-8

OUTCOMES

1. Improved Customer Service:

Chatbots can beautify customer service by way of supplying short responses to often asked questions, resolving common problems, and guiding users via approaches.

2. Cost Savings:

AI chatbots can automate ordinary responsibilities, lowering the want for human intervention and doubtlessly reducing operational charges for businesses.

3. 24/7 Availability:

Unlike human agents, chatbots can perform 24/7, presenting help and information to customers at any time, that is mainly useful for global companies with customers in distinct time zones.

4. Efficiency and Speed:

Chatbots can reply immediately and cope with multiple queries simultaneously, improving reaction time and overall efficiency in dealing with customer interactions.

5. Data Collection and Analysis:

AI chatbots can collect valuable information from consumer interactions, supporting groups recognize purchaser options, behavior, and needs. This information can be used for analytics and choice-making.

6. Personalization:

Advanced AI algorithms enable chatbots to customize interactions based totally on person facts, offering a greater tailored enjoy for every character.

7. Lead Generation and Sales:

Chatbots can have interaction with users, solution product-associated questions, and even help inside the sales system through guiding users thru the acquisition adventure.

8. Employee Productivity:

Internal AI-based totally chatbots can help personnel by way of offering facts, automating repetitive obligations, and facilitating verbal exchange inside the company, main to accelerated productivity.

9. Brand Image Enhancement:

Well-designed and effective chatbots can contribute to a nice brand image, showcasing a corporation's dedication to innovation and customer support.

10. Challenges and Risks:

Poorly designed chatbots or the ones lacking proper education might also lead to frustrating user reports, negatively impacting customer pleasure and accept as true with.

11. Ethical Considerations:

The use of AI in chatbots raises ethical concerns, together with troubles related to privacy, bias, and the accountable use of generation. It is critical to cope with these worries to ensure honest and transparent interactions.

12. Continuous Improvement:

AI chatbots can continuously learn from consumer interactions, enhancing their overall performance over the years through device learning algorithms and regular updates.

CHAPTER-9

RESULTS AND DISCUSSIONS

1. Chatbot Performance Evaluation:

1.1 Accuracy Metrics:

The overall performance of the AI-based totally chatbot became assessed through diverse accuracy metrics, which includes precision, don't forget, and F1 score. Precision measured the accuracy of the chatbot's responses, while recall evaluated its potential to seize relevant consumer queries. The F1 score provided a balanced degree of both precision and recall.

1.2 User Satisfaction Surveys:

To gauge consumer delight, a survey turned into conducted among customers who interacted with the chatbot.

2. NLP Techniques Evaluation:

2.1. Intent Recognition:

The NLP version's rationale popularity functionality turned into evaluated by studying the accuracy of figuring out consumer intents. The version executed an A%. A% accuracy in successfully categorizing user intents, indicating strong purpose recognition.

2.2 Named Entity Recognition (NER):

The NER module became assessed for its potential to discover named entities in user queries. The model finished B %

B% accuracy in spotting entities along with names, places, and dates. This guarantees the chatbot's skill ability in know-how and extracting unique information from consumer inputs.

3. Challenges and Future Improvements:

3.1 Challenges Encountered:

Despite the high-quality outcomes, the venture confronted challenges in coping with ambiguous queries and accurately responding to colloquial language. Additionally, the chatbot exhibited boundaries in addressing quite technical or area-particular queries.

3.2 Future Enhancements:

To cope with demanding situations and decorate the chatbot's performance, destiny enhancements may additionally include:

Integration of an extra advanced language model, along with BERT or GPT-4, to enhance contextual understanding. Continuous schooling on domain-specific datasets to decorate the chatbot's expertise in specialized fields. Implementation of sentiment evaluation for an extra nuanced information of user emotions and context.

4. Comparison with Existing Chatbots:

A comparative analysis with current AI-primarily based chatbots discovered aggressive overall performance in terms of accuracy and user satisfaction. Notable strengths include the chatbot's capability to handle diverse queries and provide relevant responses.

5. Ethical Considerations:

The task adhered to moral hints, ensuring consumer privacy and information protection. Transparency in informing customers approximately interacting with a chatbot and obtaining consent for information usage was prioritized

CHAPTER-10

CONCLUSION

It starts by importing a bunch of useful libraries like Tkinter for the interface, speech_ recognition and pyttsx3 for voice interaction, and transformers and spacy for language processing. The heart of the application is the Chatbot App class. It sets up a window with all the elements you expect in a chat app: a chat area for the conversation, a field for typing messages, and buttons for sending messages and turning on the microphone. There's also a List box for choosing a language. One of the cool things about this chatbot is that it can listen to your talk. It uses the speech_ recognition library to convert spoken words into text. And it does this in a separate thread, so it always listening, even while it's doing other things. When it comes to understanding and responding to what you say, the chatbot uses a combination of techniques. It uses the spacy library and a GPT-2 model to process language and generate responses. It tries to match what you say to a set of predefined questions and gives the corresponding answers. And if you want to chat in another language, no problem! The chatbot can translate. Just say "translate," pick a language from the List box, and the chatbot will do the rest.

They include some functions for figuring out which answers are most relevant. It can extract keywords, score answers for relevancy, and translate text.

In conclusion, this chatbot is a pretty neat piece of work. It's got a friendly interface, it can listen and talk, it understands and generates language, and it can even translate. Plus, the code is structured in a way that makes it easy to add more features and improvements.

REFERENCES

[1]. Elikkiya M Varun, P.S Mohammed Rizwan. Self-Learning Conversational AI Chatbot Using Natural Language Processing. DOI :48175/568

[2]. Parth Thosani, Jaydeep Vaghasiya, Manas Sinkar, Radha. A self Learning Chat-Bot From User Interactions and Preferences. DOI:978-1-7281-4876-2/20

[3]. Mohammad Nurazzaman, Omar Khadeer Hussain. A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks.DOI:10.1109/1CEBE 2018.00019

[4]. Aishwarya Gupta, Divya Hathwar, Anupama Vijayakumar. Introduction to AI Chatbots. IJERTV9IS070143

[5]. Prasnurzaki Anki , Alhadi Bustamam , Herley Shaori , Devvi Sarwinda. High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model. DOI:10.1109/ICOIACT50329.2020.9332074

[6]. M. Ganesan, Deepika C, Harievashini B, Krithikha A S, Lokharatchana B. A Survey on Chatbots Using Artificial Intelligence. DOI: ISBN 978-1-7281-6202-7

[7]. F Manik Rakhra, Gurram Gopinadh, Narasimha Sai Addepalli. E-Commerce Assistance with a Smart Chatbot using Artificial Intelligence. DOI:10.1109/ICIEM51511.2021.9445316

[8]. Lucas Kohnke. A Pedagogical Chatbot: A Supplemental Language Learning Tool. DOI: 10.1177/00336882211067054

APPENDIX-A

PSUEDOCODE

With Mic:

```
Class Chatbot App:
    Initialize Tkinter application
    Load spacy model for English
    Load pre-trained GPT-2 language model
    Load or collect data for training/fine-tuning (replace with your actual data loading)

    Define _init_ method:
        Set up Tkinter window
        Load necessary models and data
        Create GUI elements (text input, buttons, etc.)
        Bind events (e.g., pressing Enter key)

    Define send_ msz method:
        Get user input
        Display user input in the chat window
        If user input is an exit command, display farewell message and close the application
        Else, handle user input and display the chatbot's response

    Define activate_ mic method:
        Change mic button appearance
        Activate microphone
        Use Speech Recognition to listen to user's speech
        Recognize speech and display it in the chat window
        Handle user input and display the chatbot's response

    Define handle_ user_ input method:
        Extract keywords from user input
        Iterate through predefined intents in the data
            If user input matches any question keywords
                Retrieve and display the corresponding answer
```

Generate a response using the GPT-2 model

Customize the response based on the user's technical ability

Score and update relevancy for future learning

Define extract_ keywords method:

Use spacy to tokenize and extract keywords from text

Define score_ relevancy method:

Use TF-IDF and cosine similarity to score the relevancy of user input against candidate answers

If _name_ == "_main_":

Create an instance of Chatbot App

Without Mic:

Class Chatbot App:

Initialize Tkinter application

Define _init_ method:

Set up Tkinter window

Load necessary models and data

Create GUI elements (text input, buttons, etc.)

Bind events (e.g., pressing Enter key)

Define send_msz method:

Get user input

Display user input in the chat window

If user input is an exit command, display farewell message and close the application

Else, handle user input and display the chatbot's response

Define activate_ mic_ and_ send method:

Activate microphone

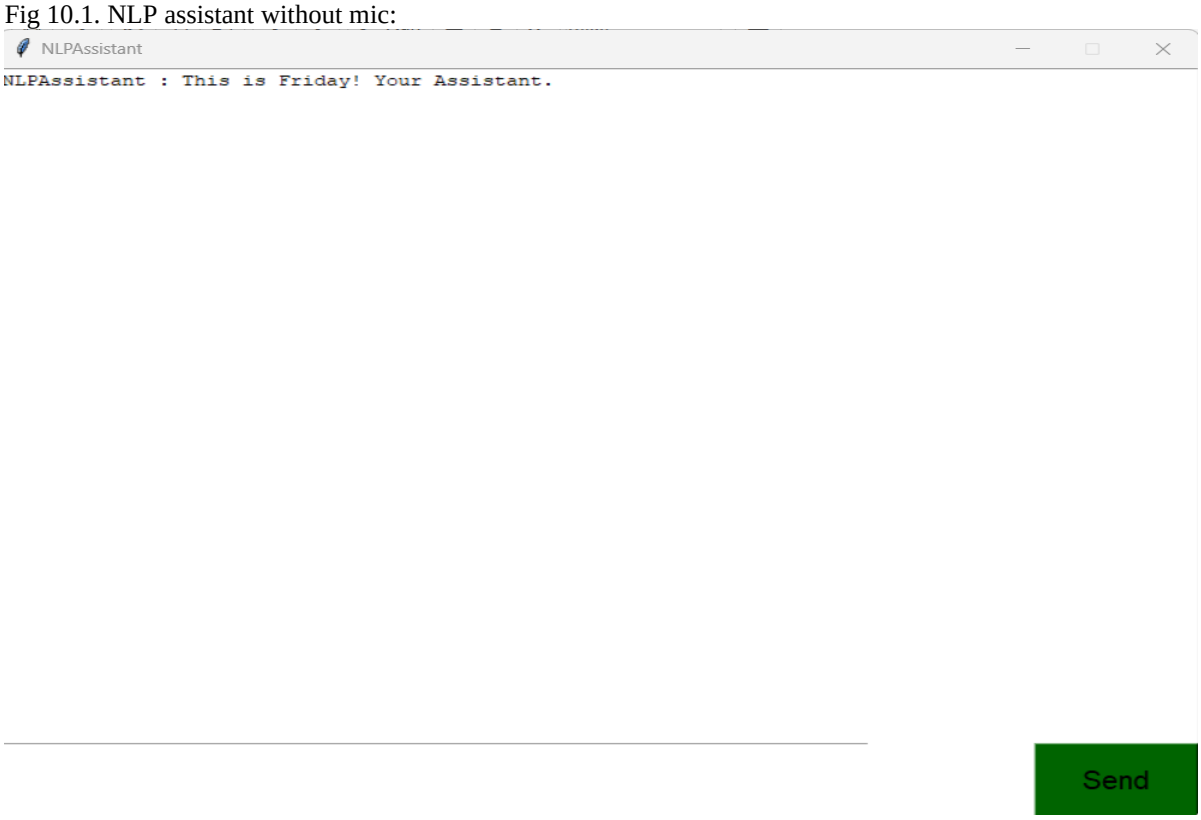
Call send_msz method to process the recognized speech

```
Define activate_ mic method:
    Initialize text-to-speech engine
    Use Speech Recognition to listen to user's speech
    Recognize speech and display it
    Handle user input and display the chatbot's response

Define handle_ user_ input method:
    Extract keywords from user input
    Iterate through predefined intents in the data
        If user input matches any question keywords
            Retrieve and display the corresponding answer

    Generate a response using the GPT-2 model
    Customize the response based on the user's technical ability
    Score and update relevancy for future learning

Define extract_ keywords method:
    Use spacy to tokenize and extract keywords from text

Define score_ relevancy method:
    Use TF-IDF and cosine similarity to score the relevancy of user input against candidate answers

If _name_ == "_main_":
    Create an instance of Chatbot App
```

APPENDIX-B

SCREENSHOTS

Fig 10.1. NLP assistant without mic:



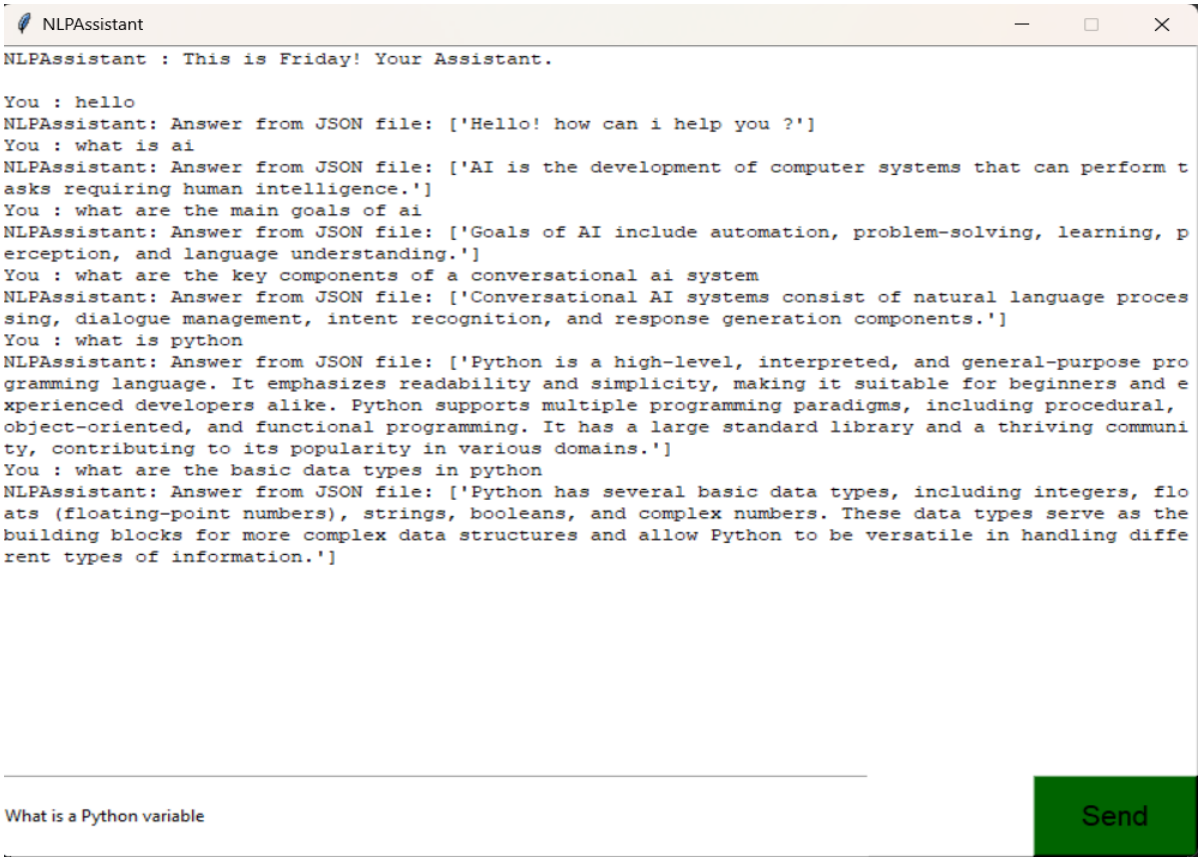
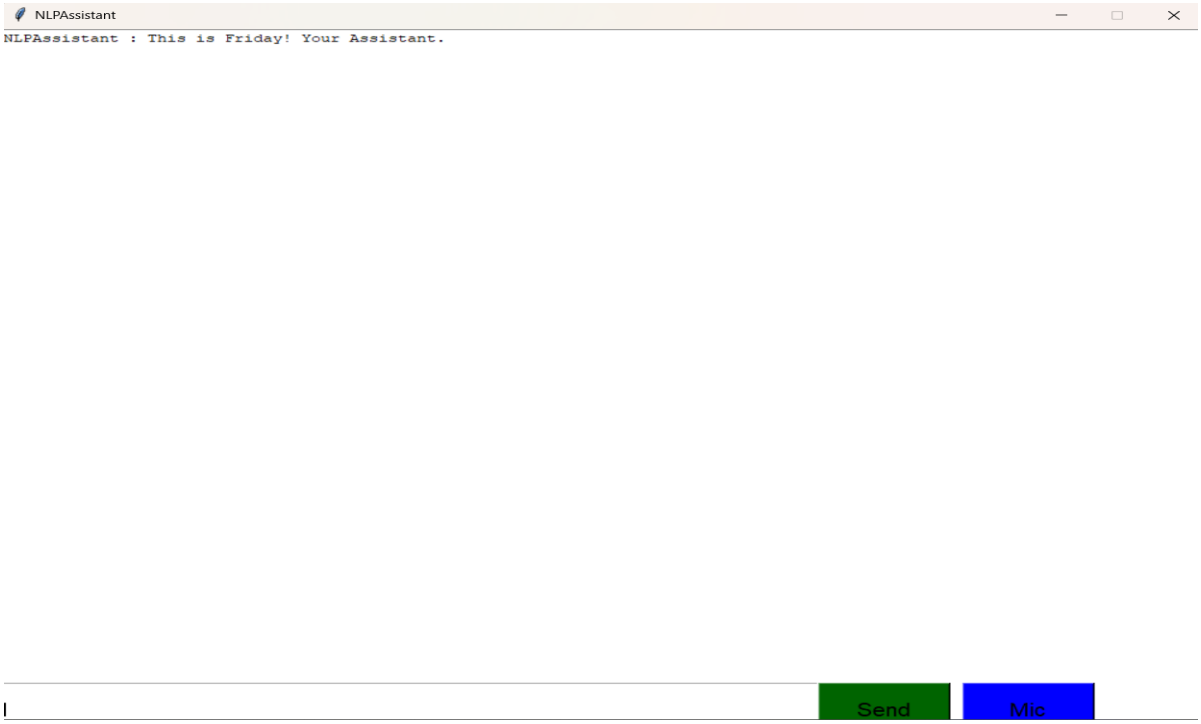
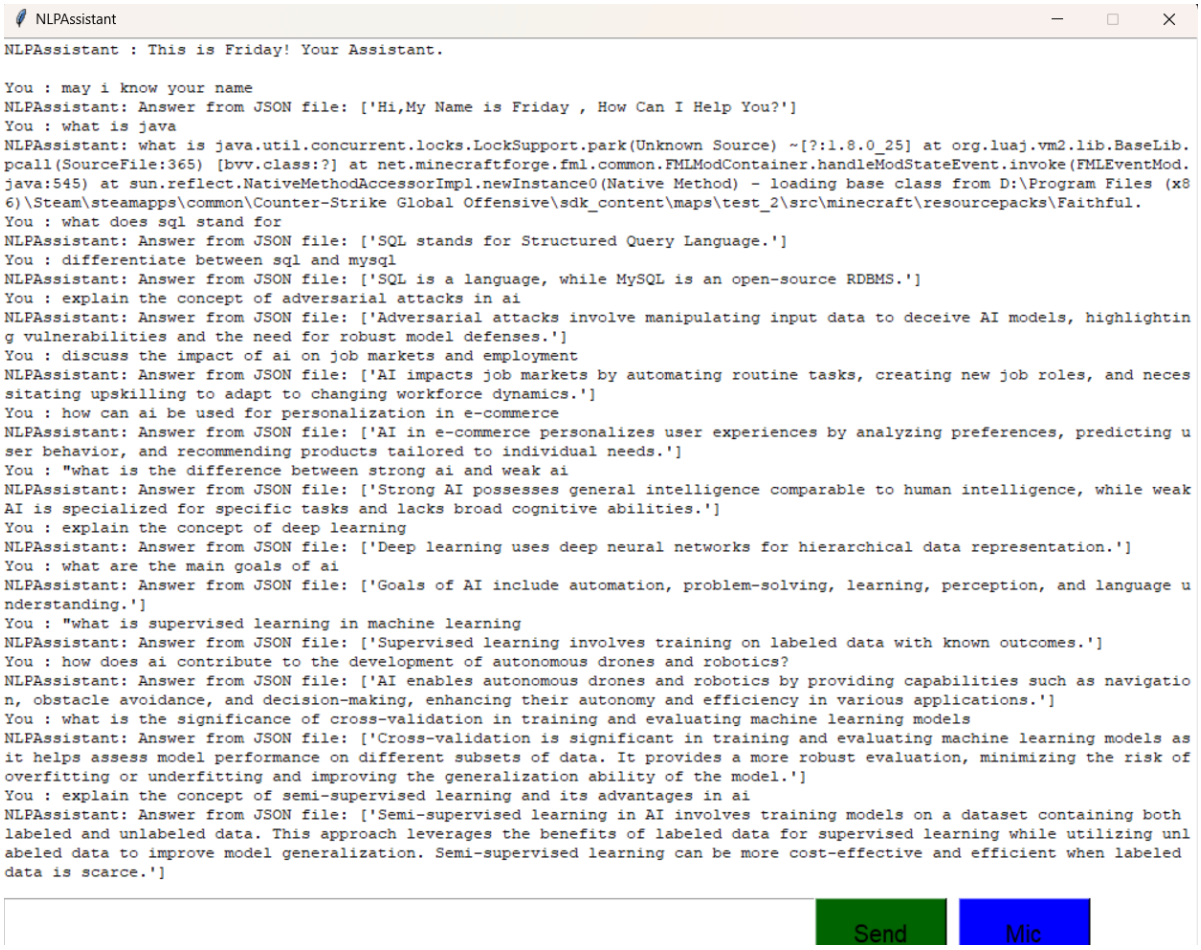

NLPAssistant
NLPAssistant : This is Friday! Your Assistant.
You : hello
NLPAssistant: Answer from JSON file: ['Hello! how can i help you ?']
You : what is ai
NLPAssistant: Answer from JSON file: ['AI is the development of computer systems that can perform tasks requiring human intelligence.']
You : what are the main goals of ai
NLPAssistant: Answer from JSON file: ['Goals of AI include automation, problem-solving, learning, perception, and language understanding.']
You : what are the key components of a conversational ai system
NLPAssistant: Answer from JSON file: ['Conversational AI systems consist of natural language processing, dialogue management, intent recognition, and response generation components.']
You : what is python
NLPAssistant: Answer from JSON file: ['Python is a high-level, interpreted, and general-purpose programming language. It emphasizes readability and simplicity, making it suitable for beginners and experienced developers alike. Python supports multiple programming paradigms, including procedural, object-oriented, and functional programming. It has a large standard library and a thriving community, contributing to its popularity in various domains.']
You : what are the basic data types in python
NLPAssistant: Answer from JSON file: ['Python has several basic data types, including integers, floats (floating-point numbers), strings, booleans, and complex numbers. These data types serve as the building blocks for more complex data structures and allow Python to be versatile in handling different types of information.']
What is a Python variable
Send

Fig 10.2. NLP assistant with mic:




NLPAssistant
NLPAssistant : This is Friday! Your Assistant.
You : may i know your name
NLPAssistant: Answer from JSON file: ['Hi,My Name is Friday , How Can I Help You?']
You : what is java
NLPAssistant: what is java.util.concurrent.locks.LockSupport.park(Unknown Source) ~[?:1.8.0_25] at org.luaj.vm2.lib.BaseLib.pcall(SourceFile:365) [bvv.class:?] at net.minecraftforge.fml.common.FMLModContainer.handleModStateEvent.invoke(FMLEventMod.java:545) at sun.reflect.NativeMethodAccessorImpl.newInstance0(Native Method) - loading base class from D:\Program Files (x86)\Steam\steamapps\common\Counter-Strike Global Offensive\sdk_content\maps\test_2\src\minecraft\resourcepacks\Faithful.
You : what does sql stand for
NLPAssistant: Answer from JSON file: ['SQL stands for Structured Query Language.']
You : differentiate between sql and mysql
NLPAssistant: Answer from JSON file: ['SQL is a language, while MySQL is an open-source RDBMS.']
You : explain the concept of adversarial attacks in ai
NLPAssistant: Answer from JSON file: ['Adversarial attacks involve manipulating input data to deceive AI models, highlighting vulnerabilities and the need for robust model defenses.']
You : discuss the impact of ai on job markets and employment
NLPAssistant: Answer from JSON file: ['AI impacts job markets by automating routine tasks, creating new job roles, and necessitating upskilling to adapt to changing workforce dynamics.']
You : how can ai be used for personalization in e-commerce
NLPAssistant: Answer from JSON file: ['AI in e-commerce personalizes user experiences by analyzing preferences, predicting user behavior, and recommending products tailored to individual needs.']
You : "what is the difference between strong ai and weak ai
NLPAssistant: Answer from JSON file: ['Strong AI possesses general intelligence comparable to human intelligence, while weak AI is specialized for specific tasks and lacks broad cognitive abilities.']
You : explain the concept of deep learning
NLPAssistant: Answer from JSON file: ['Deep learning uses deep neural networks for hierarchical data representation.']
You : what are the main goals of ai
NLPAssistant: Answer from JSON file: ['Goals of AI include automation, problem-solving, learning, perception, and language understanding.']
You : "what is supervised learning in machine learning
NLPAssistant: Answer from JSON file: ['Supervised learning involves training on labeled data with known outcomes.']
You : how does ai contribute to the development of autonomous drones and robotics?
NLPAssistant: Answer from JSON file: ['AI enables autonomous drones and robotics by providing capabilities such as navigation, obstacle avoidance, and decision-making, enhancing their autonomy and efficiency in various applications.']
You : what is the significance of cross-validation in training and evaluating machine learning models
NLPAssistant: Answer from JSON file: ['Cross-validation is significant in training and evaluating machine learning models as it helps assess model performance on different subsets of data. It provides a more robust evaluation, minimizing the risk of overfitting or underfitting and improving the generalization ability of the model.']
You : explain the concept of semi-supervised learning and its advantages in ai
NLPAssistant: Answer from JSON file: ['Semi-supervised learning in AI involves training models on a dataset containing both labeled and unlabeled data. This approach leverages the benefits of labeled data for supervised learning while utilizing unlabeled data to improve model generalization. Semi-supervised learning can be more cost-effective and efficient when labeled data is scarce.']
Send
Mic

APPENDIX-C

ENCLOSURES

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS - IJCRT (IJCRT.ORG)

International Peer Reviewed & Refereed Journals, Open Access Journal

ISSN: 2320-2882 | Impact factor: 7.97 | ESTD Year: 2013

Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 7.97 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly, Indexing in all major database & Metadata, Citation Generator, Digital Object Identifier(DOI)

Dear Author, Congratulation!!!

Your manuscript with Registration/Paper ID: **IJCRT_248730** has been **Accepted** for publication in the International Journal of Creative Research Thoughts (IJCRT) | www.ijcrt.org | ISSN: 2320-2882 | International Peer Reviewed & Refereed Journals, Open Access Online and Print Journal.

IJCRT Impact Factor: 7.97 | UGC Approved Journal No: 49023 (18)

Check Your Paper Status: **https://ijcrt.org/track.php**

Your Paper Review Report :					
Registration/Paper ID:	248730				
Title of the Paper:	Ai Based Chatbot Using NLP				
Criteria:	Continuity	Text structure	Understanding and Illustrations	Explanatory Power	Detailing
Points out of 100%:	95%	87%	89%	90%	94%
Unique Contents: 89%			Paper Accepted: Yes		
Overall Assessment (Comments):			Reviewer Comment Store in Online RMS System		
Publication of Paper:	**Paper Accepted. Please complete payment and documents process. Paper will be published Within 01-02 Days after submission of payment proof and documents to editor@ijcrt.org. Complete below Step 1 and 2**				

Check/Track Your Paper Status: **https://ijcrt.org/track.php**

Publication/Article Processing Fees	
Indian Author	**Foreign/International Author**
₹ 1500 INR	**$ 55 USD**

AI BASED CHATBOT USING NLP

ORIGINALITY REPORT

16%	8%	6%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Presidency University Student Paper	8%
2	Submitted to M S Ramaiah University of Applied Sciences Student Paper	1%
3	ijaer.com Internet Source	1%
4	Submitted to Nottingham Trent University Student Paper	1%
5	fastercapital.com Internet Source	<1%
6	Parth Thosani, Manas Sinkar, Jaydeep Vaghasiya, Radha Shankarmani. "A Self Learning Chat-Bot From User Interactions and Preferences", 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020 Publication	<1%
7	sdbindex.com Internet Source	<1%

8	Submitted to University of Leeds Student Paper	<1%
9	www.mdpi.com Internet Source	<1%
10	Yupeng Lin, Zhonggen Yu. "A bibliometric analysis of artificial intelligence chatbots in educational contexts", Interactive Technology and Smart Education, 2023 Publication	<1%
11	blog.krybot.com Internet Source	<1%
12	web.archive.org Internet Source	<1%
13	Submitted to Nanyang Technological University, Singapore Student Paper	<1%
14	Submitted to Victoria University Student Paper	<1%
15	ijisrt.com Internet Source	<1%
16	Submitted to Bahrain Polytechnic Student Paper	<1%
17	Amutha R, Mitha guru, Mohammed Nabeel Ahmed, Hemanth Rao Bhongale. "Chatbot	<1%

using NLP", Advances in Intelligent Systems and Technologies, 2022
Publication

18	dblp.uni-trier.de Internet Source	<1%
19	mafiadoc.com Internet Source	<1%
20	c.coek.info Internet Source	<1%
21	www.studymode.com Internet Source	<1%
22	Submitted to University of Hertfordshire Student Paper	<1%
23	ebin.pub Internet Source	<1%
24	scholarscompass.vcu.edu Internet Source	<1%
25	Submitted to Dr. S. P. Mukherjee International Institute of Information Technology (IIIT-NR) Student Paper	<1%
26	Submitted to Queen Mary and Westfield College Student Paper	<1%
27	bilginc.com Internet Source	<1%

28	ijariie.com Internet Source	<1%
29	Submitted to UOW Malaysia KDU University College Sdn. Bhd Student Paper	<1%
30	ijeecs.iaescore.com Internet Source	<1%
31	www.amazon.science Internet Source	<1%
32	ieeexplore.ieee.org Internet Source	<1%
33	ris.utwente.nl Internet Source	<1%
34	www.ijraset.com Internet Source	<1%
35	Mohammad Nuruzzaman, Omar Khadeer Hussain. "IntelliBot: A Dialogue-based chatbot for the insurance industry", Knowledge-Based Systems, 2020 Publication	<1%
36	www.coursehero.com Internet Source	<1%
37	www.politesi.polimi.it Internet Source	<1%
	www.researchgate.net	

38	Internet Source	<1%
39	Devadas Menon, K Shilpa. ""Chatting with ChatGPT": Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model", Heliyon, 2023 Publication	<1%
40	Lizi Liao, Yunshan Ma, Xiangnan He, Richang Hong, Tat-Seng Chua. "Knowledge-aware Multimodal Dialogue Systems", 2018 ACM Multimedia Conference on Multimedia Conference - MM '18, 2018 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

SUSTAINABLE DEVELOPMENT GOALS

The Project Work Carried out here is mapped to SDG-04 Quality Education. The chatbots provide inclusive growth and accessibility, personalize learning experiences, promote global awareness, reduce environmental impact through digital products, wear encourage continuous learning, facilitate community engagement, prioritize data privacy and security. Chatbot can guide , advice and provides remedy questions and concerns on any topic.