IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN AUTOMATED CHATBOT FOR AN EDUCATIONAL INSTITUTION USING NATURAL LANGUAGE PROCESSING

Dr.M.Senthil Kumar
Associate Professor,
Department of Computer Science and Engineering, SRM Valliammai Engineering College, Chennai.

Ms. J. Dharshani
Department of Computer Science and Engineering,
SRM Valliammai Engineering College, Chennai.

Ms. K. Divyabharathi
Department of Computer Science and Engineering,
SRM Valliammai Engineering College, Chennai.

Ms. S. Sneha
Department of Computer Science and Engineering,
SRM Valliammai Engineering College
Chennai.

Abstract - In recent years, the development of chatbots has become trendier and so far, several conversational chatbots were designed which replaces the traditional chatbots. A chatbot is a computer program that is used to interact with human beings and fulfill their needs. Chatbot gives the response for the user query and they are also capable of executing tasks. Early development of chatbots was so difficult whereas recent chatbots are much easier to develop because of the wide availability of development platforms and source code. A chatbot can be developed using either Natural Language Processing (NLP) or Deep Learning. When compared to traditional chatbots, bots designed using Deep Learning requires a huge amount of data to train. This Educational chatbot is used to provide the details asked by the user. The chatbot can efficiently answer any queries given by the user regarding the college information, details of the students and CGPA-related information. The CGPA of the students can be acquired by providing the user details and password to the chatbot. Thus, Using the NLP, the chatbot can answer all our queries which is user-friendly and also time-saving.

Keywords – Natural Language Processing, AIML

I. INTRODUCTION

A chatbot is a service that interacts with users through a chat interface using rules and artificial intelligence. A chatbot can be used on mobile apps, web apps, messaging apps, and personal assistants, among other platforms. It's a job to provide the user with human-like interaction and help. The finest chatbot in the world would be one in which the user couldn't tell the difference between it and a real person. Natural Language Processing (NLP) is a field of study and application that looks into how computers can interpret and modify natural language text or speech in order to accomplish meaningful tasks. NLP is a multidisciplinary field that is linked to linguistics. Its linguistics, computer science, information engineering subfield, and artificial intelligence are both concerned with how computers interact with human (natural) languages, specifically how to train computers to handle and evaluate huge amounts of natural language data. Computer and information sciences, psychology, and electrical and electronic engineering are all linked to it. It is, of course, linked to Artificial Intelligence. Machine translation, speech recognition, text processing, and summarization are just a few examples of NLP applications. Users can communicate with a chatbot in a variety of ways, including text, speech, and actions. Every day, tech titans like Microsoft, Amazon,

Google, and Facebook compete to release better chatbots.

This recent survey provides a complete systematic evaluation of the use of chatbots in education. Following the theoretical concept of Technology-Mediated Learning (TML), the authors establish three viewpoints for examining current research: structure (input), learning process (process), and learning outcome (output). Several dimensions have been recognized in terms of input: student profile, educational environment, and chatbot technology.

Individual student characteristics such as personality traits, technical skills, and educational and social background all influence learning outcomes.

Background Of The Project

There are numerous chatbot services available in a variety of domains. In their respective sectors, these chatbots deliver effective services. Education is a vital component in which chatbot services play a crucial part. Chatbots are required at educational institutions to know about the institution's details, student info, and so on. The institution's existing websites.

A. Limitations Of Automatic Learning

Many studies attempted to automate the creation of knowledge bases. Neural networks were utilized to create a machine-machine conversational information sharing system, in which knowledge bases are constructed on top of the knowledge bases of other chatbots. While this is great for general-purpose agents, it isn't appropriate for domain-specific cases like user support for a corporation with its own services and use cases.

A reinforcement learning approach for building knowledge bases with open-world data from conversations was established in and a self-feeding chatbot that employs sentiment analysis chooses what is fresh knowledge and should be added to the knowledge base.

B. Chatbot Design

Artificial Intelligence (AI) refers to a machine's ability to receive data, process it, and perform tasks depending on the information derived from the data [10]. AI computers react to incoming data on their own, and their responses are appropriate and similar to how a human would react in the same situation.

Even the most basic chatbot is an AI machine because when it receives a message, it responds on its own based on the message. The design of a chatbot can be done in a variety of ways. However, there are two primary types of chatbots: deterministic and rule-based models and probabilistic and machine learning-based models.

The rule-based technique is classified as a non-AI method by certain writers. The methodologies used in the rule-based approach, on the other hand, provide the chatbot AI capabilities. The Artificial Intelligence Markup Language (AIML), for example, is one of the tools used in the rule-based method, and it employs AI.

Chatbots that operate on established rules are known as rule-based chatbots. Because the generalization is relatively limited, the developer must list all possibilities in advance, and the chatbot will not work if the input message departs from the specified patterns. The developer can only understand the discussion context if he or she programs it directly and objectively, and it operates as a decision tree.

C. Content Management

The chatbot content management method is a continuous loop that allows for rapid iteration of material in order to adapt to, and occasionally anticipate, changing user behavior. It entails assessing talks and incorporating new information into the knowledge base.

Chatbots must be able to learn from and about their users, as well as to adapt to organizational requirements, therefore they must be constantly learning. Automatic learning is not exact, and human-supervised learning is desirable; nevertheless, it must be a well-structured process, because the chatbot's quality may begin to deteriorate over time if the human-supervised learning process is not well-defined and organized, resulting in poor content judgments.

Although the first content fed before deployment is critical, conversation-driven development for chatbot content is the key to successful customer care using machine learning-based chatbots. That is, once the chatbot is launched, the chatbot team must monitor the chats in order to better comprehend consumers and collect messages that should have been understood but weren't.

D. Evaluation Metrics

The development cycle concludes when a chatbot is ready to go live, and the assessment phase begins. The discussions must be monitored in order to assess the chatbot's effectiveness and determine if it is accomplishing its goal. The first step is to save these talks efficiently for subsequent studies. In order to identify users, channels, and message time, storing chatbot dialogues necessitates a specialized data model.

Analysts can work on collecting important information for the organization and from general performance now that the talks have been properly archived. Data from chatbots is useful for both the bot's personal

improvement and the brand it represents. User experience, information retrieval, linguistics, business, and technology are all factors to consider while evaluating chatbots.

Information retrieval evaluation examines the sufficiency of chatbot content, i.e., whether it is capable of responding to what the brand requires and whether users' queries are handled appropriately. This can be determined through quantitative measurements or by having a domain expert (business analyst) evaluate the knowledge base to ensure that the material is relevant to the company's objectives.

Because a chatbot can represent a business, poor content production can harm not only the chatbot but also the brand's reputation. It's important to pay attention to the quality of users' interactions so that you can act on areas where a lot of people are having trouble or are frustrated, and to make the dialogue flow easier.

Statement Of Problem

This educational chatbot focuses on efficiently responding to the queries of the user. So, in this telematic world, it is essential for acquiring the necessary information in a short time. This chatbot provides the user with information related to college and student details. The purpose of using a chatbot is to save time by acquiring the required information in a limited time and it is also user-friendly.

Objective of the project

The purpose of this study is to demonstrate the various ways in which a chatbot may be constructed, as well as their categories. This study also discusses how a Chatbot could efficiently answer College-related questions, with the added benefit of providing personal information such as grades and other information with proper user identification. As a result, the user can rapidly obtain the needed information without having to navigate through a succession of web pages.

II. RELATED WORK

A chatbot is divided into three categories: educational, healthcare, and business. They are built on those regions' knowledge bases to provide consistent support to the user. Educational chatbots, such as Med- Chatbot for medical students, is based on the open-source AIML and enable students to answer specialized education-related inquiries. The Unified Medical Language System (UMLS), which is freely available, is utilized to create responses and transform natural language questions into suitable SQL queries for this chatbot. The results of these SQL searches are provided to the user in a natural conversation.

The writers Rupesh Singh, Manmath Paste, and Nirmala Shinde created a chatbot for small businesses using Tensor Flow. Chatbots, they claim, are software utilized in the entertainment sector, corporations, and customer service. Chatbots are created utilizing a variety of techniques, including knowledge bases and machine learning. Chatbots that use machine learning produce more useful results. User-friendly chatbots are those that respond based on the context of the interaction. The proposed chatbot provides a method for creating a chatbot that can follow the conversation's context. This method employs TensorFlow to create the chatbot's neural network model, and NLP techniques to keep the discussion in context. These chatbots can be employed in small businesses or sectors to automate customer service by answering user questions.[1]

The Automated Thai-FAQ Chatbot was created using RNN-LSTM by the authors Panitan Muangkammuen and Narong Intiruk. They came to the conclusion that in an e-commerce model with online customer assistance, such as email or live chat, people choose live chat since it is quick and convenient. As a result, a business must hire and pay for administrative staff. However, this creates an issue because administrators must devote a significant amount of time to producing responses, and customers must wait for them. There are several chatbots accessible, however, they all require users to manually set up keywords. The goal is to propose and create a Frequently Asked Questions (FAQs) Chatbot that uses a Recurrent Neural Network (RNN) in the form of a Long Short-Term Memory (LSTM) for text classification and automatically responds to clients. The chatbot recognized 86.36 percent of the questions and responded with 93.2 percent accuracy, according to the results of the experiment. [2]

The authors Pablo Rivas, and Kerstin Holzmayer have proposed a Chatbots and talk bots are intelligent programs that can establish written and oral communication with human beings, usually with the purpose of helping them achieve a specific goal. More and more companies are now implementing bots in order to reduce operational costs. Most bots use machine learning algorithms that are deployed on a company’s websites, cloud services, or distributed mobile systems so that customers are always able to speak with `someone' to inquire about products or services. Most bots are trained using data from interactions among human beings so that they can learn speech patterns and answer questions. [3]

III. PROPOSED SYSTEM

The proposed system is an educational chatbot that is used to provide the user with information regarding the college, student details, and CGPA the students. It consists of the database of the students, a gallery of the college, and detailed information about the institution as required by the user. The CGPA can be acquired by providing the user's name and password to the system. The Chatbot will provide the correct details if the appropriate credentials are given.

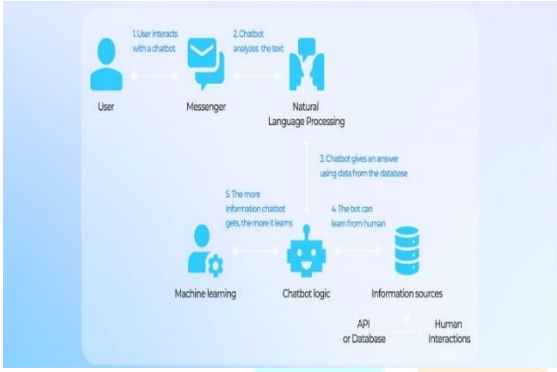

Fig.1 Architecture of the proposed system

IV. METHODOLOGY

List Of Modules

- Pre-Processing
- Feature Extraction
- Context Identification
- Personal Query Response System

Pre-Processing

Pre-processing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis.

Feature Extraction

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. The redundant data is reduced from the data set.

Context Identification

Pre-processing is applied to the input text to standardize the input as per the system's requirement. Based on the keywords used in the text, appropriate context is recognized.

Personal Query Response System

Upon receiving personal queries like CGPA, attendance, etc., the authenticity of the user is checked through user-id and password.

4.1 Natural Language Processing Algorithm

"Natural Language Processing, also known as NLP, is an area of computer science, artificial intelligence, and machine learning concerned with the interactions between computers and humans. Natural Language Processing (NLP) helps provide context and meaning to text-based user inputs so that it can come up with the best response. It undergoes a series of steps as follows,

Sentence Segmentation

Splitting sentences apart every time you see a punctuation mark is a straightforward way to code a Sentence Segmentation model.

Word Tokenization

Tokenization is the process of breaking down large pieces of text into smaller ones. Tokenization divides the raw text into words and sentences, which are referred to as tokens. These tokens aid in the comprehension of the context or the development of the NLP model.

Predicting Parts Of Speech For Each Token

Part-of-speech is a term that refers to a certain part of speech (POS) Tagging is a common Natural Language Processing procedure that involves categorizing words in a text (corpus) in accordance with a specific part of speech, based on the word's definition and context.

Text Lemmatization

Lemmatization is a term that relates to performing things correctly using a vocabulary and morphological analysis of words, with the goal of removing inflectional endings solely and returning the base or dictionary form of a word, also known as the lemma.

Identifying Stop Words

A stop word is a word that has the same chance of appearing in documents that aren't relevant to a query as it does in documents that are relevant to the query. In this study, we illustrate how the requirement of being highly evaluated by a similarity measure can be used to replace the concept of relevance.

Dependency Parsing

The practice of evaluating the dependencies between the phrases of a sentence in order to determine its grammatical structure is known as Dependency Parsing (DP). The majority of the time, a sentence is broken into multiple portions depending on this.

Named Entity Recognition (Ner)

Introduction NER is a subtask of information extraction that locates and classifies different entities in a sentence, such as names, organizations, people, and so on.

Coreference Resolution

The objective of coreference resolution is to locate all expressions in a text that refer to the same thing. It's a key step in a variety of higher-level NLP activities including document summarising, question answering, and information extraction that need natural language understanding.

V. RESULTS

A. *Basic Information*

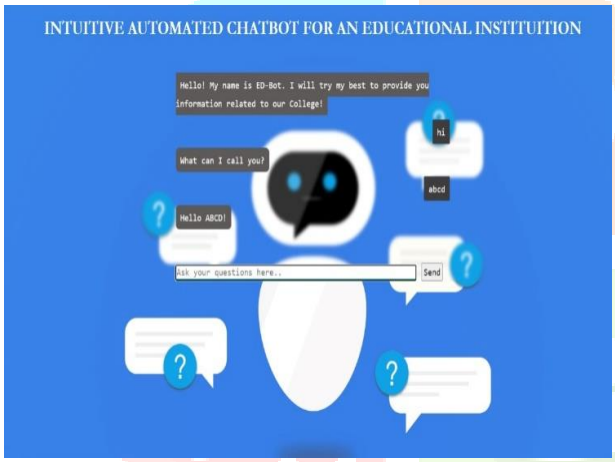

Fig.2 Basic Information

B. *Query And Response About The Institution*

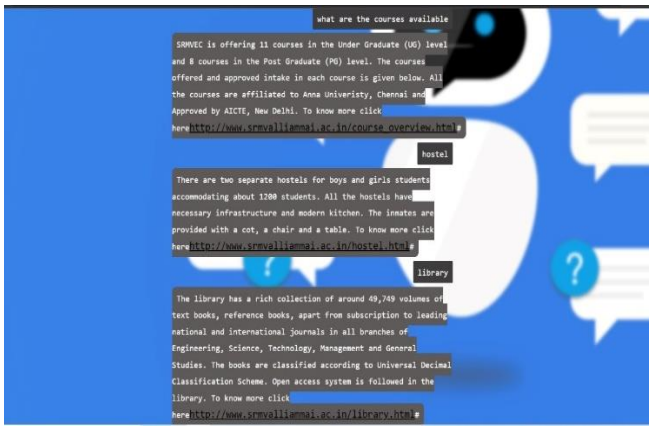

Fig.3 Query and response

C. *Students' CGPA Details*

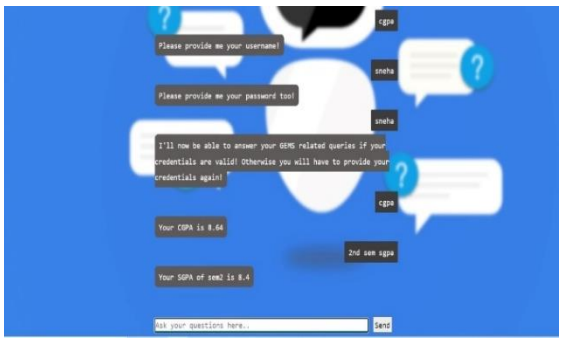

Fig. 4 Students' CGPA Details

D. *Gallery*

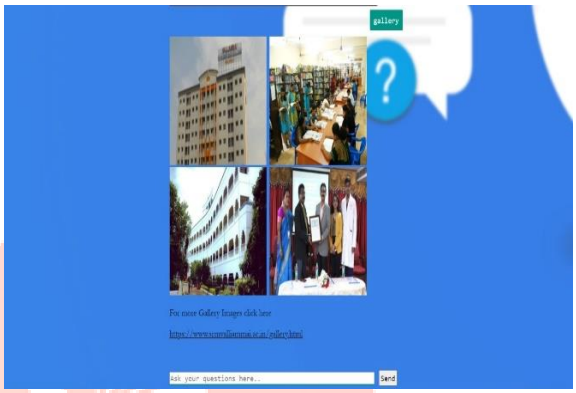

Fig. 5 Gallery

E. *Database of the Students*

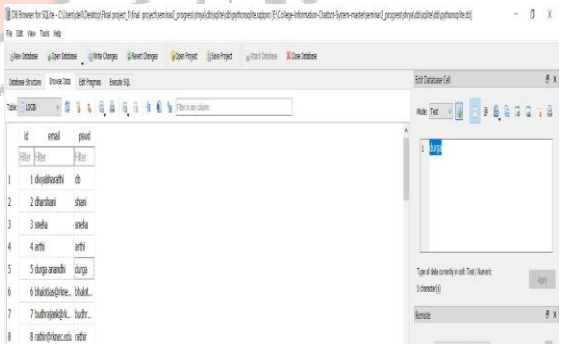

Fig.6 Database of the students

VI. FUTURE WORK

In future the Image recognition can be enhanced with much more details about the image captured through the camera. Enhancement to this system can be done by adding the features of currency recognition. We can make a chatbot that is a blend of AIML and LSA. This will enable a client to interact with the chatbot in a more natural fashion. We can enhance the features by including and changing patterns and templates for general client queries using AIML and the right response are given more often than not utilizing LSA.

VII. CONCLUSION

Chatbots are extremely valuable for businesses and the value will only increase as time goes by. While the technology to simulate conversation with a computer has been around for decades, bots are adaptable versions with powerful integration of Artificial Intelligence and Natural Language Processing. It creates an assisting guide for all the users. It increases efficiency by maintaining known standard responses. It provides improved question responsiveness and accuracy. It is used for increasing the ability to track and monitor queries and highlighting gaps in available information. The impact of NLP in machines will be greater than the impact of microprocessor technology in the last 20 years because Natural Language is fundamental to almost all business, military & social activities. Therefore, the application of NLP has no end. In the future, the proposed system will be able to interpret the textual description in a much better way.

REFERENCES

1. KBot: A Knowledge Graph-Based ChatBot for Natural Language Understanding Over Linked Data-2020 Addi Ait-Mlouk and Lili Jiang.

2. A Conversation-Driven Approach for Chatbot Management Giovanni Almeida Santos 1, (Member, Ieee), Guilherme Guy De Andrade 1, 2022.

3. AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT), Neelkumar P. Patel; Devangi R. Parikh; Darshan A. Patel; Ronak R. Patel, 2019

4. JAICOB: A Data Science Chatbot Daniel Carlander-Reuterfelt, Álvaro Carrera, 2020.

5. Excitement and Concerns about Machine Learning-Based Chatbots and Talkbots: A Survey Pablo Rivas; Kerstin Holzmayer; Cristian Hernandez; Charles Grippaldi, 2018.

.