



# AI based Chatbot for educational institution using Natural Language processing and Deep Learning

Dr. M. Narendra<sup>1</sup>, D.V.S.P.L. Kameswari<sup>2</sup>, M. Chetana Ramya<sup>3</sup>, K. Tarun<sup>4</sup>, G. Charan<sup>5</sup>

<sup>1</sup>Professor, Department of Computer Science and Engineering, Qis college of Engineering and Technology, Ongole, Andhra Pradesh, India

<sup>2,3,4,5</sup>Student, Department of Computer Science and Engineering, Qis college of Engineering and Technology, Ongole, Andhra Pradesh, India

**Abstract**— A chatbot is a software application which simulates humans' conversation. A Chatbot uses training data to understand the type of queries people ask, analysis correct answers to those queries and use ML and NLP to learn and get better at answering the queries. Creation and implementation of chatbot is still a developing area. Chatbots are two types: 1) Rule Based Chatbot 2) AI enabled chatbot

Rule Based Chatbots respond to queries with the help of predefined keywords. AI chatbots analysis the person's query and try to deliver the best answer based on existing data. They collect more and more data and continue and learn. This project focuses on development of chatbot for educational institutions using NLP and Deep Learning. Natural Language Processing is used by machines to understand analyze manipulate and interpret human's language. Deep Learning is a type of Machine Learning based on Artificial Neural Networks. This chatbot provides replies for the queries related to placements, counselling data, examinations and all the activities taking place in the college. This chatbot application helps the parents and students to know about the admission process and receive reply fast. Students can ask queries related to examinations, placements and events occurring in the college.

**Keywords:**- Deep learning, chat bot, knowledge base, lemmatization, stemming, natural language processing.

## I. INTRODUCTION

These days, chatbots are popular. In any particular area where it operates, chat bots are the provider of responses to users' queries. Today's most well-known illustration is Alexa from Amazon. One can see chatbots nearly everywhere; they can be seen on every second website users visit. A bot is handy for responding to questions about information that might not be easily accessible at that website. Most websites provide chatbots for visitors to use as they explore the services they offer. They are gradually serving as our virtual assistants in daily life.

### A. Chatbot basics

A chatbot is a machine with artificial intelligence that can communicate with people. This could be a verbal conversation or a text message. In essence, chatbots are employed for information gathering. Although it may be used with mobile devices and desktop computers, most people access it online. It is a conversational bot that connects with people in a particular domain or on a specific subject by taking input in the form of natural language sentences. A chatbot primarily functions when a user asks a question or proposes a new topic of

conversation. Software agents that mimic human beings are known as chat bots. These are the embedded AI bots that can respond to user questions utilising natural language processing. A predefined knowledge base aids in the development of the query response.

### ***B. Chatbot for College***

College websites are slow, an outsider wouldn't know where to look for a specific piece of information, and it is challenging for someone outside the college's domain to extract information are just a few of the reasons why a college inquiry system is necessary. The educational institution chatbot will respond to the query by summarising the user's request and then outputting responses. It also gives the user with only the information they want. All responses pertaining to areas like admissions, the examination cell, the notice board, the placement activities, college events and other such areas will be provided by the chatbot.

The important features of this educational institution chatbot are:

- Queries related to admissions could be answered through it.
- College students can obtain details regarding upcoming exams
- Details related to placement events can also be obtained

The chatbot will serve as a quick, dependable, and educational widget to improve user experience on college websites and provide users with accurate information. The bot will evaluate user inquiries, comprehend user messages, and then respond appropriately. Natural Language Processing and Deep Learning are used by the chatbot. Users' time and effort will be saved in this manner, and they will have access to effective solutions.

This application's goals are to:

- Analyze user queries and comprehend user messages.
- To give the user's question a really effective response.
- To save the user's time by eliminating the need for a personal visit to the college for inquiries.
- This mechanism will assist the student in staying informed of college activities

## **II. LITERATURE SURVEY**

Making the interaction between the machine and the human seem natural and human-like is a major difficulty in chatbot design, as this paper[1] explains. Virtual bots and other models with CUI (conversational user interface) replicate human response times by sending late responses or answers. When prompt responses are anticipated, as they are when dealing with customers, a late reply might, however, negatively affect user satisfaction. The paper[2] asserts that the requirement for for a chatbot arises for a number of reasons, such as the slowness of the college webpage, the difficulty for someone outside the college's domain to retrieve information, and the fact that a search for a specific piece of information would be unfamiliar to them. The system is necessary because of the clever answer to all the problems. The chatbot will respond by summarizing the user's enquiry and then outputting replies. It also gives the user with only the knowledge they want. An outline of the commercial platforms and tools that may be used to build and use chatbots is given in this study [3]. We specifically go through the libraries and online services that can be used to develop and deploy bots.

This article [4] provides an intriguing challenge in the realm of natural language processing: the building of an Intelligent Conversational Agent combining AI and machine learning techniques (NLP). In order to enhance the quality of services offered, conversational agents, sometimes referred to as chatbots, are being employed as virtual assistants in a variety of industries, including business, government and healthcare organisations. This research [5] claims that a dataset representing a collection of user intentions, or intents, has been generated. Each intent includes illustrative user queries and the chatbot's responses to those questions. The chatbot's questions and responses were compiled based on regulatory documents for the WRC's implementation, and the dataset's implementation comprises 18 intents. There are up to 20–30 questions for each intent.

The study to present the different ways in which a chatbot may be built, as well as their categories, is represented in this work [6]. This study also examines how a chatbot may effectively respond to inquiries about colleges while also offering other details related to placements, events. As a result, the user doesn't have to go through a series of web pages to obtain the data they need. This document [7] illustrates the College Enquiry Chatbot, which aims to provide students with quick answers to their questions. Lack of

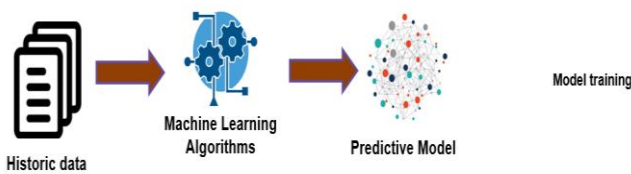
personality and conversational flow is the main drawback we discovered when using the regular rule based chatbots.

### III. PROPOSED SYSTEM

To overcome the disadvantage that is present in the existing system we are introducing the AI based chatbot.

The chatbot that we are proposing is fully trained which can answer the queries of the user more accurately.

To train the chatbot we are using the pattern matching algorithm in the NLP and Deep Learning



The steps in the proposed system are:

#### A. Data Collection

For the dataset, the college related data i.e. placements related data, admission related data, college events data, courses offered etc are collected. The collected data is stored in the form of Json file format. This Json file consists of tags, intents, patterns and responses.

#### B. Importing required libraries:

The required libraries for this chatbot are

- Numpy
- pickle
- tensorflow
- json
- random

All the required libraries are installed and imported into the program before the execution

#### C. Pre processing the data

The data preprocessing consists of following

#### Class diagram

steps

- Text case
- Stop word removal
- Tokenization
- Stemming
- Lemmatization
- Bag of words

#### D. Model building:

Utilizing weighted connections (also known as "synapses") that are computed via multiple iterations of training data, the result can be derived from an input (a classification). The weights are changed after each run through the training set so that the neural network can provide output with more "accuracy" (lower error rate).

#### E. Output response:

The chatbot takes the query input preprocesses it and give the ouput response. If there are no more queries, the user can simply type "quit". After that we can quit the chatbot application.

### IV. DESIGN

#### Use Case Diagram:

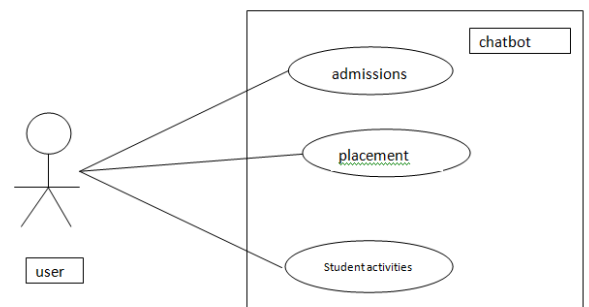


Fig 1: Use Case Diagram of chatbot

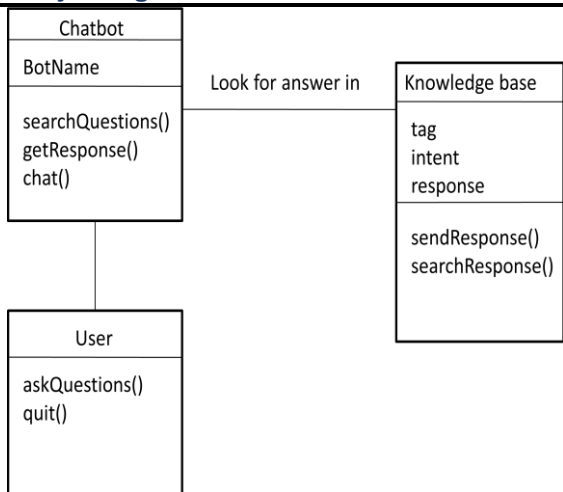


Fig 2: Class diagram

**Activity diagram**

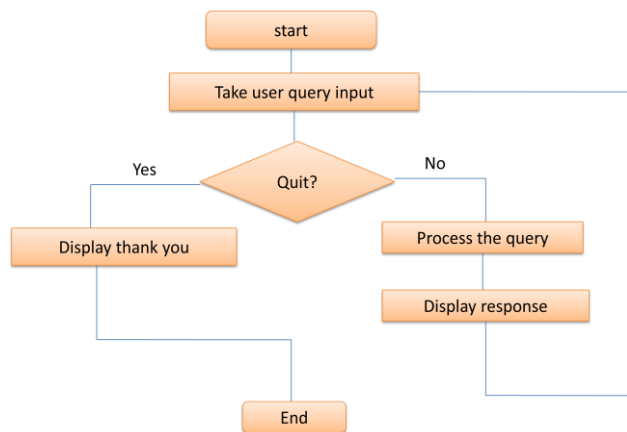


Fig 4: Activity diagram

**Sequence diagram**

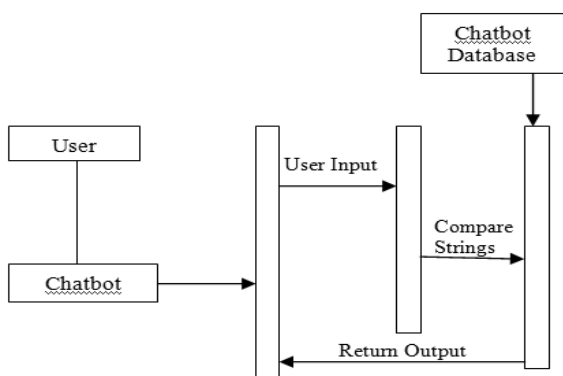


Fig 3: Sequence Diagram of chatbot

**V. IMPLEMENTATION**

**A. JSON file**

To create our data set for conversation with user, we have used JSON files to store the tag, intents and patterns. When user converses with our chat bot, the input is matched to patterns listed in JSON files and corresponding answer is returned as response.

The sample JSON file structure is as:

```

{
  "tag": "greeting",
  "patterns":["hi","hello","greetings!"],
  "responses":["hi, how can I help you?",
"hello, greetings"]
},
    
```

```

Command Prompt
C:\Users\LEO\OneDrive\codes\chatbot\NLP_PROJECT>python trainingtest.py
2022-12-18 08:47:56.546337: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'cudart64_110.dll'; dlerror: cudart64_110.dll not found
2022-12-18 08:47:56.546846: I tensorflow/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.
C:\Users\LEO\OneDrive\codes\chatbot\NLP_PROJECT>trainingtest.py:68: VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you want to do this, you must specify 'dtype=object' when creating the ndarray.
training_data = np.array(training_data)
2022-12-18 08:48:25.222822: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'nvcuda.dll'; dlerror: nvcuda.dll not found
2022-12-18 08:48:25.222884: W tensorflow/stream_executor/cuda/cuda_driver.cc:263] failed call to cuInit: UNKNOWN ERROR (383)
2022-12-18 08:48:25.235889: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:159] retrieving CUDA diagnostic information for host: DESKTOP-V508658
2022-12-18 08:48:25.236999: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-V508658
2022-12-18 08:48:25.236982: I tensorflow/core/platform/cpu_feature_guard.cc:133] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical operations: AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
Model: "sequential"

Layer (type) Output Shape Param #
-----
dense (Dense) (None, 128) 23184
dropout (Dropout) (None, 128) 0
dense_1 (Dense) (None, 64) 8256
dropout_1 (Dropout) (None, 64) 0
dense_2 (Dense) (None, 23) 1495

Total params: 31,939
Trainable params: 31,939
Non-trainable params: 0

```

## B. Data preprocessing:

### Text case:

It is mandatory to convert the text to lower case because of case sensitivity.

### Tokenization:

Tokenization is a technique used in NLP to break down phrases and paragraphs into simpler language-assignable elements. Tokenization divides the original text into tokens, which are words and sentences. These tokens aid in context comprehension or model development for NLP. By examining the word order in the text, tokenization aids in comprehending the text's meaning.

### stop words removal:

Deleting irrelevant words is stop words removal. Using nltk stop words like is,the,a etc from the sentences as they don't carry any information. Stop words are those words in natural language that have a very little meaning, such as "are", "an", "the", etc

### lemmatization:

Lemmatization is the combination of many spellings of the same word.

Root word in lemmatization is referred to as lemma. The canonical form, dictionary form, or citation form of a group of words is known as a lemma (plural lemmas or lemmata). Run, for instance, is the lemma of all the terms runs, running, and ran because they are all variations of the word run.

### Stemming:

Stemming is a method for reducing word

inflection to its base forms in natural language processing. Stemming is a method for eliminating affixes from words to reveal their basic form. It is analogous to trimming a tree's branches back to the trunk. For instance, the word eat is the root of the verbs eating, eats, and eaten.

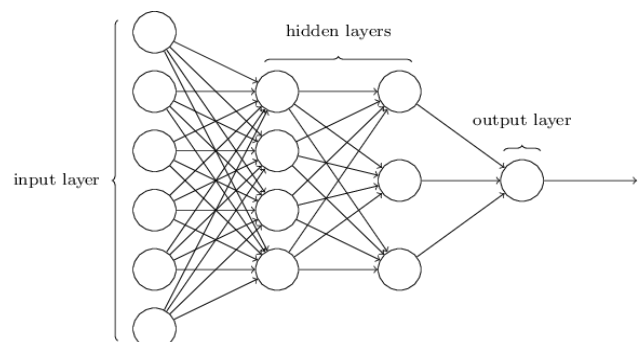
### Bag of words:

"Bag of words" is a text modelling NLP technique. It is a technique for feature extraction from text data, to put it technically. This method is a straightforward and adaptable way to extract features from documents.

A textual illustration of word recurrence in a document is called a "bag of words." We don't pay attention to grammatical conventions or word order; we only keep track of word counts. It is referred to as a "bag" of words because any details regarding the arrangement or structure of the words within the document are ignored. The model doesn't care where in the document recognised terms appear; it is only interested in whether they do.

### Building model:

There are a certain amount of example sentences provided for each class. Once more, each word in each sentence is separated into its stem and used as an input by the neural network. The synaptic weights are then determined by doing millions of iterations through the training data, with each iteration resulting in a minor adjustment of the weights for improved accuracy. The weights of every synapses are calibrated by recalculating back over numerous layers ("back-propagation") while the outcomes are compared to the output from the training data. These weights are comparable to a "strength" metric; in a neuron, synaptic weight determines whether or not something will be recalled. Each time, the "weight" grows a little bit.



### C. Output screenshots:

```

Command Prompt
C:\Users\LEKHO\Desktop\codes\chatbot\MNIT_PROJECT>python chatbottest.py
2022-12-18 08:58:44.162243: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'cudart64_118.dll'; dlerror: cudart64_118.dll not found
2022-12-18 08:58:44.162743: I tensorflow/stream_executor/cuda/cudart_stub.cc:29] Ignore above cudart dlerror if you do not have a GPU set up on your machine.
2022-12-18 08:58:53.441830: W tensorflow/stream_executor/platform/default/dso_loader.cc:64] Could not load dynamic library 'nvcuda.dll'; dlerror: nvcuda.dll not found
2022-12-18 08:58:53.441963: W tensorflow/stream_executor/cuda/cuda_driver.cc:293] Failed call to cuInit: UNKNOWN ERROR (303)
2022-12-18 08:58:53.450994: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:169] retrieving CUDA diagnostic information for host: DESKTOP-V5DR858
2022-12-18 08:58:53.461786: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:176] hostname: DESKTOP-V5DR858
2022-12-18 08:58:53.465118: I tensorflow/stream_executor/cuda/cuda_diagnostics.cc:193] this TensorFlow binary is optimized with oneAPI Deep Neural Network Library (oneDNN) to use the following GPU instructions in performance-critical operations: AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
Start chat with ChatBot of QIS
You: Hello
1/1 [=====] - 8s 388ms/step
BOT: good day!!

You: Who are you
1/1 [=====] - 8s 69ms/step
BOT: my name is QISbot

You: what is full name of qis
1/1 [=====] - 8s 84ms/step
BOT: QIS college of engineering and technology

You: what is exact code of qis
1/1 [=====] - 8s 54ms/step
BOT: exact code is QISE

You: department names in qis
1/1 [=====] - 8s 73ms/step
BOT: Computer Science and Engineering(CSE),Electronic and Communication Engineering(EC),Mechanical Engineering(MECH),Civil Engineering(CIVIL),Electronics and Electrica
Engineering(EEE),Information Technology(IT),Master of Business Administration(MBA),Master of Computer Application(MCA)

```

```

Command Prompt
You: about transport facilities
1/1 [=====] - 8s 67ms/step
BOT: yes,our college will provide transport facilities for 21 routes for long routes 2 buses will be provided based on the number of students

You: canteen
1/1 [=====] - 8s 61ms/step
BOT: There are totally 4 canteens in the college and a separate canteen for the girls

You: training classes
1/1 [=====] - 8s 56ms/step
BOT: The company specific training classes will start from the 1st year and continued till final year. The CRT will be provided on aptitude, reasoning and coding.

You: quit
C:\Users\LEKHO\Desktop\codes\chatbot\MNIT_PROJECT>

```

### D. APPLICATIONS

- This chatbot helps the students to be updated with events in the college
- It saves time for teaching and non-teaching staffs during admissions time.
- It is easily accessible, saving time and money
- Gives students quick access to information without requiring any physical effort.

### E. CONCLUSIONS

A chatbot system's goal is to mimic a human conversation. Faculty and college students are welcome to post their inquiries. The chatbot quickly and effectively searches for the answers to users' questions.

This chatbot provides replies for the queries related to placements, counseling data, examinations and all the activities taking place in the college. This chatbot application helps the parents and students to know about the admission process and receive reply fast. Students can ask queries related to examinations, placements and events occurring in the college.

### F. FUTURE ENHANCEMENTS

Instead of AI enabled chatbot, other algorithms can be implemented. We can provide option for voice-based queries. The users will have to give voice input and the system will give the text output or the system can also give voice output. Additionally, if the chatbot has been successfully implemented in the educational domain, it can be used in other domains such as medical, forensic, sports, etc. It will be helpful in all disciplines because we can quickly get the key information and do so without having to sort it.

### REFERENCES

- [1] Abhay Chopde , Mohit Agarwal."Chatbot using Deep Learning".Vishwakarma Institute of Technology, Pune, 411037 ,India.
- [2] Lalwani, Tarun, Shashank Bhalotia, Ashish Pal, Vasundhara Rathod, and Shreya Bisen. "Implementation of a Chatbot System using AI and NLP." International Journal of Innovative Research in Computer Science & Technology (IJRCST) Volume-6, Issue-3 (2018).
- [3] Vagelis Hristidis . "Chatbot Technologies and Challenges ". UC Riverside and SmartBot360 .
- [4] Dr. Sunanda Mulik ,Dr. Poonam Sawant ,Dr. Vaishali Bhosale . "Application of NPL : Design of Chatbot for New Research Scholars ".Turkish Online Journal of Qualitative Inquiry (TOJQI) Volume 12, Issue 8,July 2021 : 2817-2823.
- [5] Dmitry Alekseev , Polina Shagalova and Eleonora Sokolova . "Development of a Chatbot Using Machine Learning Algorithms to Automate Educational Processes ". NNSTU n.a. R.E.Alekseev ,Minina str.,24 Nizhny Novgorod,603950,Russia.
- [6] Dr.M.Senthil Kumar ,Ms.J.Dharshani ,Ms. K.Divyabharathi ,Ms. S.Sneha ."An Automated Chatbot for an Educational Institution using Natural Language Processing ".SRM Valliammai Engineering College ,Chennai .Volume 10,Issue 5 May 2022 ,ISSN 2320-2882
- [7] Ujaliben Kalpesh Bavishi ."Implementing a college enquiry chatbot ".California State University ,Sacramento ,2019.