Computer Engineering Department

A.P. Shah Institute of Technology

G.B.Road,Kasarvadavali, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2020-2021

A Project Report on

AI And Web-Based Interactive College Enquiry Chatbot

Submitted in partial fulfillment of the degree of

Bachelor of Engineering(Sem-8)

in

Computer Engineering

By

Rohan Parkar (17102022)

Yash Payare (17102064)

Keyur Mithari (17102014)

Jitesh Nambiar (17102017)

Under the Guidance of
Prof. Jaya D. Gupta

1.Project Conception and Initiation

1.1 Abstract

Students have to visit colleges to collect various information like college fees, term schedule, college activities etc. during their admission process or as per their daily needs. This process is very tedious and time consuming, hence a chatbot is the best tool which provides quick way to interact with the users. The project deals with user's request in form of question-based message and processes it to deliver a desired response in form of message. The chatbot uses Artificial Neural Network(AI) and Natural Language Processing(NLP). User can ask College-related questions, then the query is applied as an input to algorithm, which processes the message and displays the corresponding response to the user. A chatbot can work 24x7 without getting tired. It is subjected to minimal errors thus increasing the productivity.

1.2 Objectives

1. Available to everyone at any time:

Students or parents can ask their queries to the chatbot at anytime of the day and get the desired response. A chatbot will run 24 hours a day, 7 days a week, answering everyone's questions.

2. Reduce time consumption:

It can take a long time to get to college and gather all of the necessary details. Without having to go to college, a chatbot can provide a wealth of knowledge.

3. Cost effective:

There is no more work to be done once a chatbot has been deployed. In order to answer the students' questions, less human power is needed, making it more cost efficient.

4. Simple and easy use:

Parents and students who are unfamiliar with the college website can quickly access details using the chatbot.

5. Keep students updated:

In just a few minutes, students will receive the most recent news and updates about grades, tests, activities, and so on with the help of chatbot.

1.3 Literature Review

[1] The UNIBOT project is a web-based user-friendly chatbot which has minimal response time and database hits developed using HTML, CSS, jQuery and Ajax.

[2] This Chatbot project decreases customer service costs and handles multiple users at once with a FAQ supported dataset using AI Markup Language (AIML) and Latent Semantic Analysis (LSA).

[3] This chatbot is coordinated into Moodle's foundation either as a custom menu thing or as a gadget inside an HTML block utilizing Artificial Knowledge (AI) procedures like Machine Learning (ML) and Natural Language Processing (NLP).

[4] This paper presents a Natural Language Processing technology that uses subject modeling and word embeddings to build a chatbot that can answer patients' question which benefits both patients and health-care providers.

1.4 Problem Definition

- The complex nature of college website makes it difficult for an outsider to search for a particular piece of information.

- A human cannot handle several user requests at a single instance.

- Staff may not be present every time to answer user queries.

- Answering queries is a repetitive and tedious job which requires patience.

1.5 Scope

- Traditionally, a human representative addresses all of the students' doubts and questions, which is tedious and repetitive.

- Chatbots use machine learning to succeed in AI, allowing them to understand the user's question and react appropriately.

- Students may ask the chatbot questions about tests, academics, fee structures, and other topics.

- The chatbot's response can be displayed in the form of images, links, or text.

1.6 Technology stack

- Python3
- NumPy
- PyTorch
- NLTK
- Flask
- JavaScript

1.7 Benefits for environment & Society

- Chatbots will continue to operate every day throughout the year without requiring to take a break thus requiring less manpower.

- Chatbots on the other hand can simultaneously have conversations with thousands of people at any time of the day.

- Chatbots are bound by some rules, resulting them to treat a customer in the most polite and perfect way.

- Chatbots can do repetitive tasks without errors unlike humans which in turn helps people save time and increase productivity.
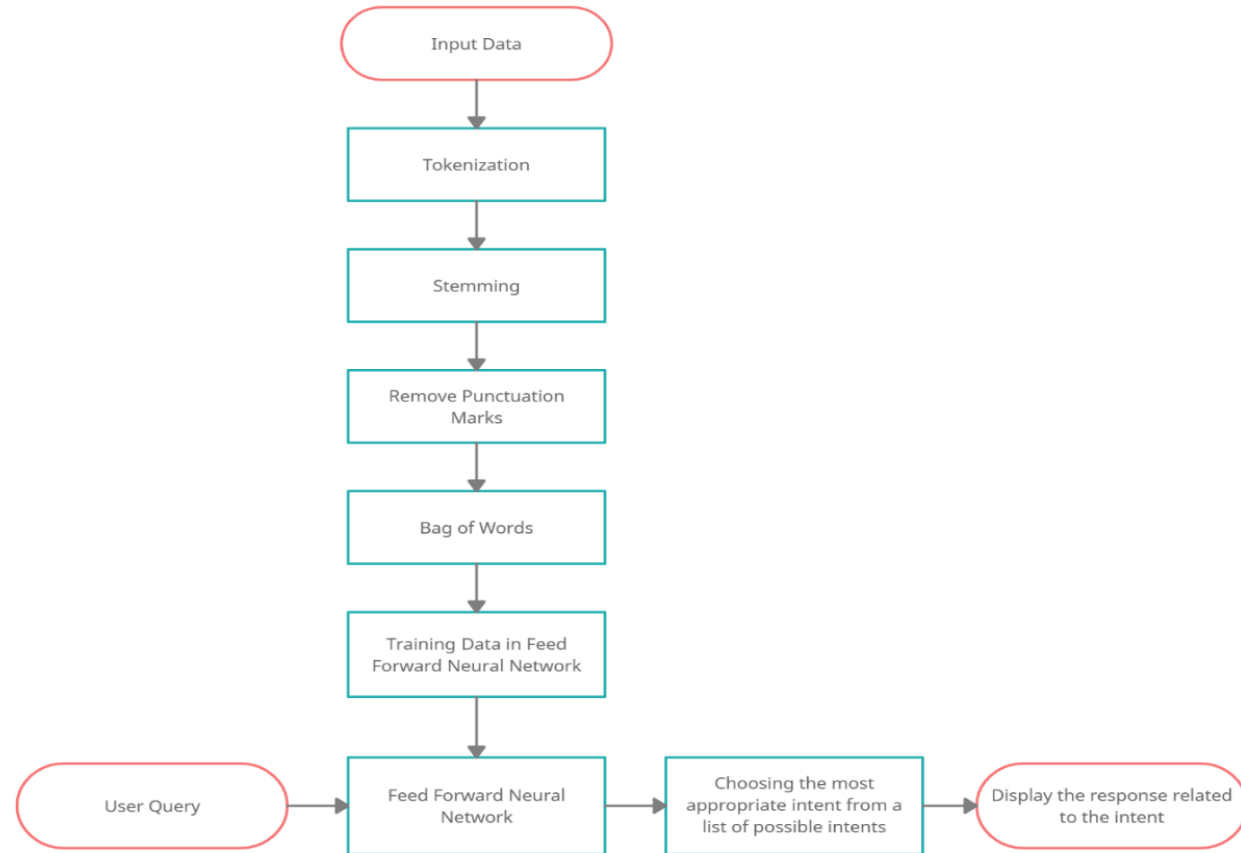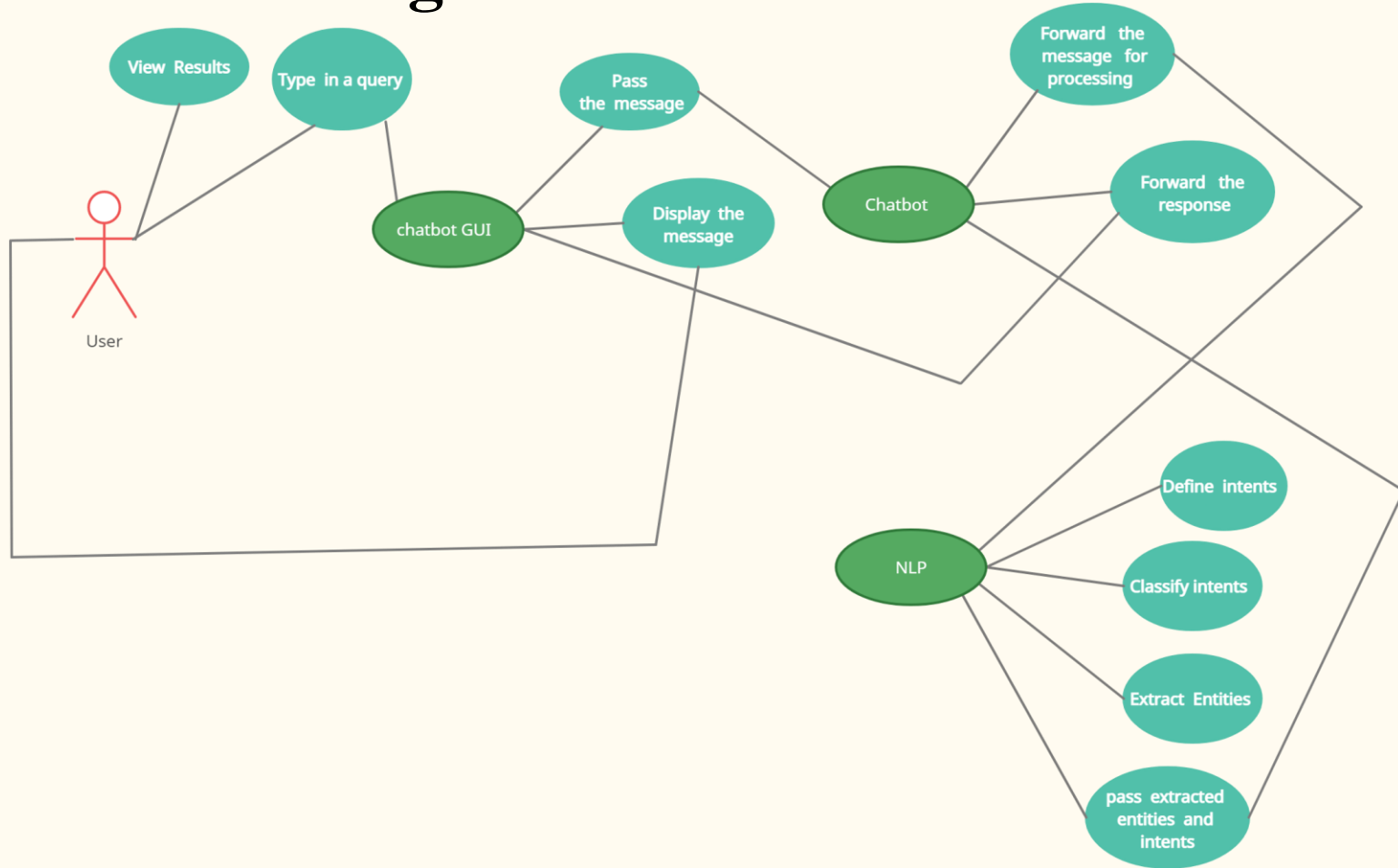
2. Project Design

2.1 Proposed System

• User discussions begin with greetings or general questions.

• NLP is used to process the question.

• Tokenization is used to divide the question into words.

• Using porter stemmer, these words are clipped to their source words.

• The punctuation marks are removed.

• A bag of words is formed.

• The bag of words is then fed into a trained neural network model.

• Then it predicts the most appropriate intent.

• The user is then given an answer based on the intent.

2.2 Design(Flow Of Modules)
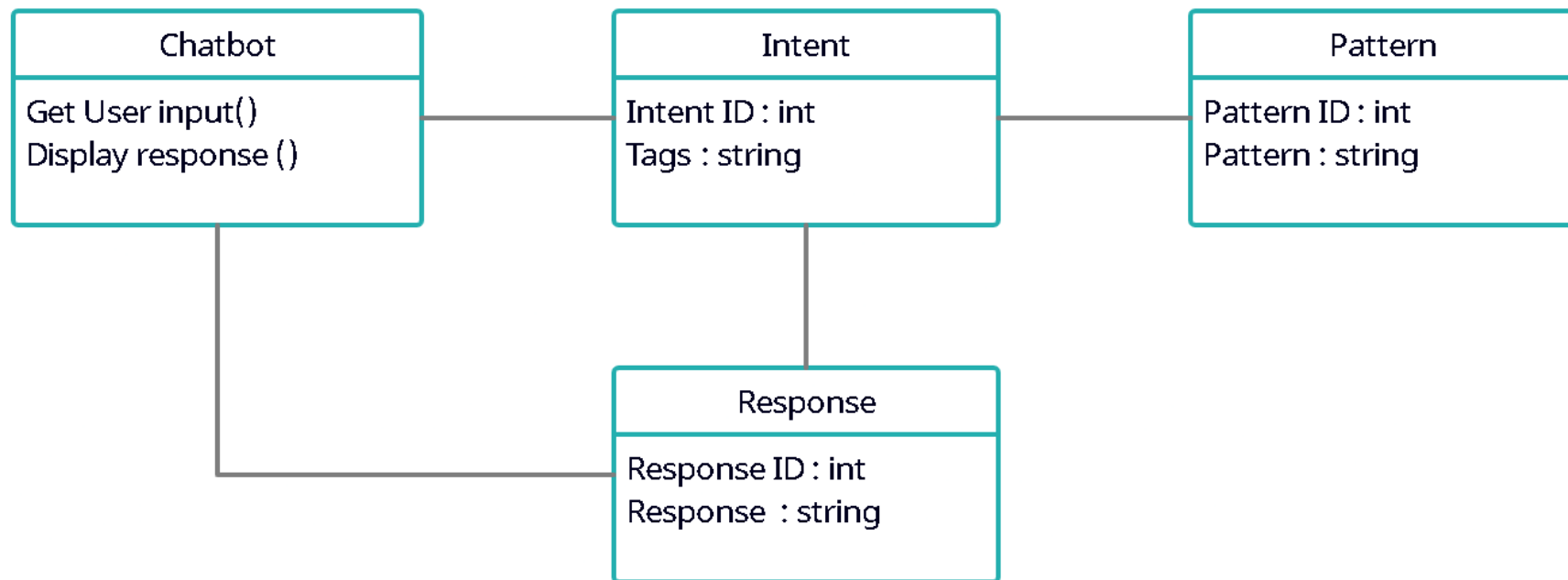
2.3 Use Case Diagram

2.4 Class Diagram

Chatbot

Get User input()
Display response ()

Intent

Intent ID : int
Tags : string

Pattern

Pattern ID : int
Pattern : string

Response
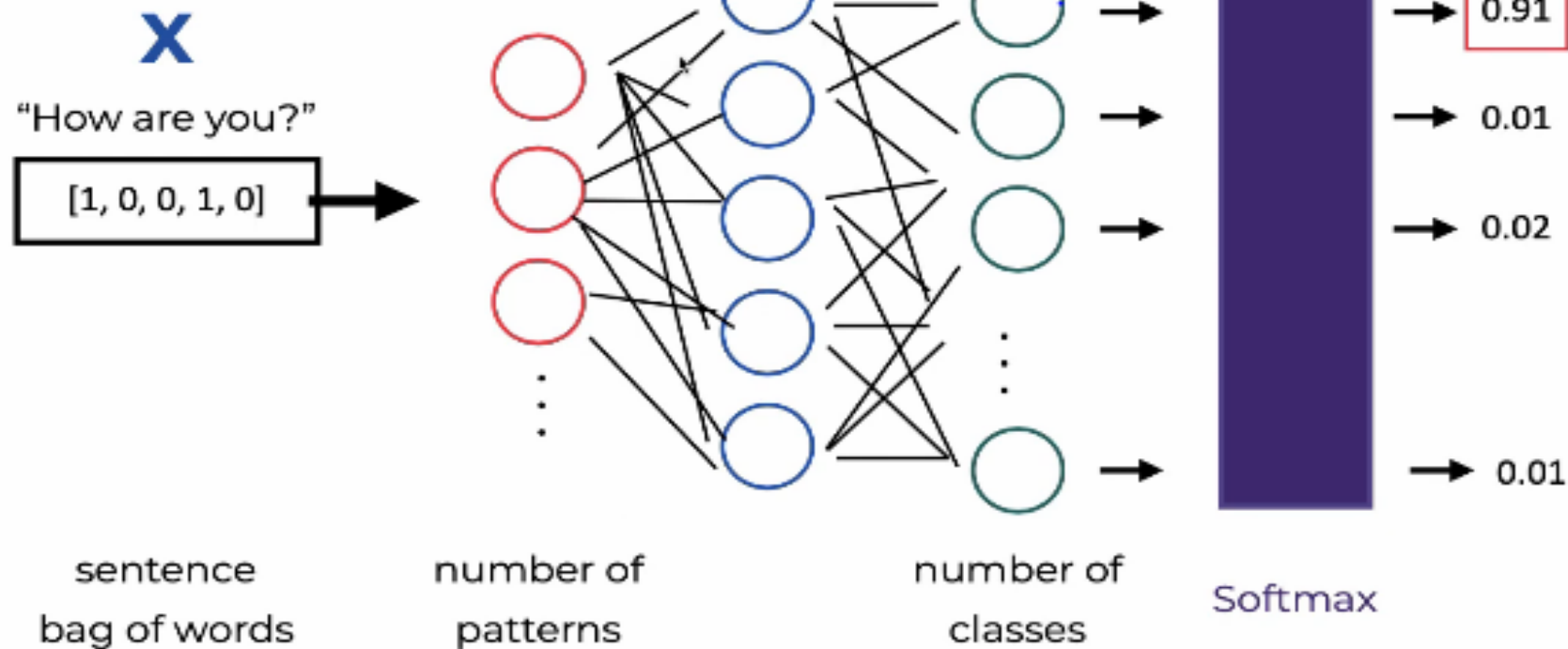
Response ID : int
Response : string

2.5 Module-1

Feedforward Neural Network

- The chatbot uses Feedforward Neural Network for predicting the responses.

- The Natural Language Processing is performed on the patterns, X variable is created which has all the training dataset which is fed to the feedforward neural network model and it predicts the Y variable that is tagged.

- Feedforward neural network has one input layer and two hidden layers with each layer having an activation function between them except the last layer as it has Softmax at the end.

- After the training is complete the output of the model is saved in a file named 'data.pth' and the training loss is calculated

Feed Forward Neural Net

X

"How are you?"

[1, 0, 0, 1, 0]

y

0.91

0.01

0.02

0.01

sentence
bag of words

number of
patterns

number of
classes

Softmax

Module-2

Natural language processing (NLP):

- **Natural language processing** (NLP) facilitates human-to-machine communication without humans needing to "speak" Java or any other programming language as it allows machines to obtain and process information from written or verbal user inputs. In NLP there are various steps such as:

- **Tokenization:** Tokenization is a way of separating a piece of text into smaller units called tokens. Here, tokens can be either words, characters, or sub words. The formatting is based on spaces.

- **Stemming:** Stemming is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers. The input to the stemmer is tokenized words.

- **Bag of Words:** Bag of Words model is used to pre-process the text by converting it into a bag of words, which keeps a count of the total occurrences of most frequently used words. This model can be visualized using a table, which contains the count of words corresponding to the word itself.

Stemming

"timing","timings","times"

$\downarrow$

["time","time","time"]

Tokenization

Where is college located?

$\downarrow$

["Where","is","college","located","?"]

bag of words

all words

["Hi","How ","are","you","bye","see","later"]

"Hi" $\rightarrow$	[1, 0, 0, 0, 0, 0, 0]	
"How are you?" $\rightarrow$	[0, 1, 1, 1, 0, 0, 0]	0 (greeting)
"Bye" $\rightarrow$	[0, 0, 0, 0, 1, 0, 0]	
"see you later" $\rightarrow$	[0, 0, 0, 1, 0, 1, 1]	1 (goodbye)

X   Y

Module-3

Database

- The database of the chatbot is in **.json** format which contains the intents i.e., the goal the user has in mind when typing in a question or query to the chatbot application.

- Every intent has a tag name saved in a variable named '**tag**' which has different patterns for every tag stored in '**patterns**' variable in which a user can ask a query.

- The chatbot returns responses to the user as a message stored in '**responses**' variable if the query asked is the most matched with the pattern from the patterns variable.

- The database is created studying the college website and analyzing what queries a user can ask and what responses he or she will expect.

```json
"intents":[
    {
        "tag":"greeting",
        "patterns":["Hi","hey","hello","Good morning","Good afternoon"],
        "responses":["Hey ! how can i help you ?"]
    },
    {

        "tag":"time",
        "patterns":["Timing of college ?","When can I visit the college ?"],
        "responses":["Monday To Friday","9am to 5pm"]
    }
]
```

3.Implementation

3.1 Algorithms

- Natural language processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language.

- Natural language processing includes many different techniques for interpreting human language, ranging from statistical and machine learning methods to rules-based and algorithmic approaches.

- Basic NLP tasks include tokenization and parsing, lemmatization/stemming, part-of-speech tagging, language detection and identification of semantic relationships.

- In general terms, NLP tasks break down language into shorter, elemental pieces, try to understand relationships between the pieces and explore how the pieces work together to create meaning.

3.2 Pseudo Code

```
program start
take message from user
send message to model using Socketio
function chat(message)
            while True
                        tokenize (message)
                        stem (tokenized message)
                        create bag of words of stemmed message
                        give bag of words as X variable to model
                        model predicts a tag
                        if prediction >0.75
                                    loop through the tags in dataset
                                                if predicted tag is equal to tag in dataset
                                                            return response related to that tag from dataset
                                                end if
                                                end loop
                        else
                                    return tag not found message
                        end if
            end while
```

3.3 Platforms for Execution

1. Anaconda: Anaconda is an open-source distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment.

2. Visual Studio Code: Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

4. Results

The ApsitBot uses NLP to process the user query and then this processed query is used by the neural network model to predict the response. The feedforward neural network model was trained with 54 intents with a learning rate of 0.00017. The number of epochs was set to 1000 and after 10 successful iterations, the final loss of the trained model was found to be 0.0002. Training loss was calculated after each iteration and it is plotted using a graph.The plotted graph shows that the training loss is decreased to the point of stability. This indicated that this is an optimal fit.

Training Loss
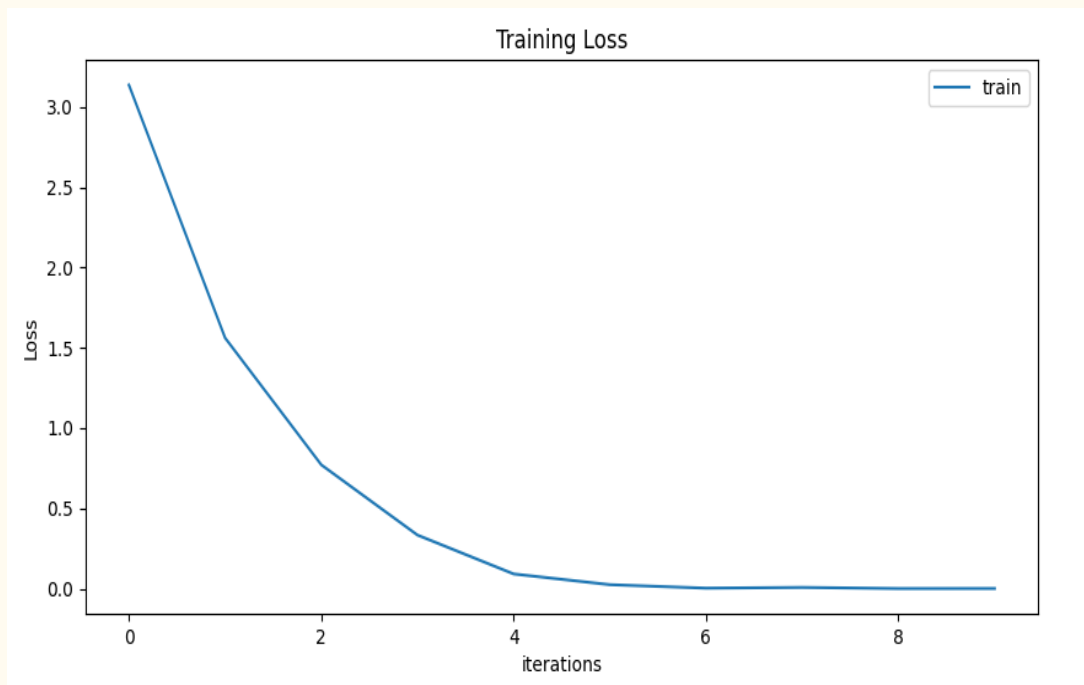
Iterations	Loss
1	3.1367
2	1.5614
3	0.7708
4	0.3333
5	0.0911
6	0.0243
7	0.0073
8	0.0026
9	0.0005
10	0.0002

5.Conclusion

ApsitBot is a stepping stone technology developed with NLP and AI. Even if not being familiar with the college website, the user can ask their queries to the bot and get desired answers. The development of this chatbot addresses the issues that may arise when collecting information. Any doubts related to college can be easily solved using this application. This application can be relied on by students with considerable ease and without difficulties.

In future chatbot can have both voice-based queries and voice-based responses. Accuracy can be increased by providing more data to the database and training the neural network model. Also, after successful execution of chatbot in college domain, we can implement it in other domains like medical, forensic, sports, etc. It will be beneficial in all the fields as without spending much time, we are accessing the relevant information and that too without any sorting.

6. References

[1] N. P. Patel, D. R. Parikh, D. A. Patel and R. R. Patel, "AI and WebBased Human-Like Interactive University Chatbot (UNIBOT)," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019.

[2] B. R. Ranoliya, N. Raghuwanshi and S. Singh, "Chatbot for University Related FAQs", 2017 International Conference on Advances in Computing Communications and Informatics (ICACCI), 2017.

[3] Souali, K., Rahmaoui, O., Ouzzif, M.,—&. El Haddioui, I. (2019). "Recommending Moodle Resources Using Chatbots", 2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS).

[4] Lal, H., & Lal, P. (2019). NLP chatbot for Discharge Summaries. 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT).

[5] Fernando A. Mikic Fonte, Mart´ın Llamas Nistal, Juan C. Burguillo Rial and Manuel Caeiro Rodr´ıguez, "NLAST: A natural language assistant for students" in , Vigo, Spain:Department of Telematics Engineering, University of Vigo.

[6] Amey Tiwari, Rahul Talekar and S.M. Patil, "College Information Chat Bot System", International Journal of Engineering Research and General Science, 2017.

[7] Balbir Singh Bani and Ajay Pratap Singh, "College Enquiry Chatbot Using A.L.I.C.E", International Journal of New Technology and Research, 2017.

[8] Mondal, A., Dey, M., Das, D., Nagpal, S., & Garda, K. (2018). Chatbot: An automated conversation system for the educational domain. 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP).

[9] H. K. K., A. K. Palakurthi, V. Putnala and A. Kumar K., "Smart College Chatbot using ML and Python," 2020 International Conference on System, Computation, Automation and Networking" (ICSCAN), Pondicherry, India, 2020.

[10] Chun Ho, C., Lee, H. L., Lo, W. K., & Lui, K. F. A. (2018). "Developing a Chatbot for College Student Programme Advisement", 2018 International Symposium on Educational Technology (ISET).

[11] Prof.K. Bala, Mukesh Kumar, Sayali Hulawale and Sahil Pandita, "ChatBot For College Management System Using A.I", International Research Journal of Engineering and Technology (IRJET), vol. 04, no. 11, pp. 2030- 2033, Nov 2017.

[12] Guruswami Hiremath, Aishwarya Hajare, Priyanka Bhosale, Rasika Nanaware and K. S. Wagh, "Chatbot for education system", International Journal of Advance Research Ideas and Innovations in Technology (IJARIIT), vol. 4, no. 3, pp. 37-43, 2018.

7.Bibliography

[1] B. R. Ranoliya, N. Raghuwanshi and S. Singh, "Chatbot for University Related FAQs", 2017 International Conference on Advances in Computing Communications and Informatics (ICACCI), 2017.

[2] Youtube Channels:

        1. PythonEngineer: https://www.youtube.com/c/PythonEngineer

        2. edureka!: https://www.youtube.com/c/edurekaIN

[3] https://www.ultimate.ai/blog/ai-automation/how-nlp-text-based-chatbots-work

[4] https://chatbotsmagazine.com/contextual-chat-bots-with-tensorflow-4391749d0077

[5] https://towardsdatascience.com/your-guide-to-natural-language-processing-nlp-48ea2511f6e1

8. Publication

Paper entitled **"AI and Web-Based Interactive College Enquiry Chatbot"** is accepted at *"The 3rd World Symposium on Artificial Intelligence (WSAI 2021)"* by Rohan Parkar, Yash Payare, Keyur Mithari, Jitesh Nambiar, and Prof. Jaya Gupta.

Thank You