

**Project Report**

**On**

**AI Based Chat-Bot**



**Submitted to: Submitted by:**

#### Dr.Vishan Kumar Gupta Shubham Chaudhary

(CC-CSE-B-V-Sem) Roll. No.: 2017049

**Guided by:** CSE-Section:-B:V-Sem

#### Dr.Vishan Kumar Gupta Session: 2022-2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**GRAPHIC ERA UNVERSITY, DEHRADUN**

**CERTIFICATE**

Certified that Shubham Chaudhary (University No.: 2017049) has developed mini project on “AI Based Chat-Bot” for the V Semester Mini Project in Graphic Era University, Dehradun. The project carried out by Students is their own work as best of my knowledge.

# ACKNOWLEDGMENT

I wish to thank my parents for their continuing support and encouragement. We also wish to thank them for providing us with the opportunity to reach this far in our studies.

I would like to thank our HOD sir Dr.Devesh Pratap Singh for his kind support and encouragement.

I would like to thank particularly our project Co-Ordinator Dr.Vishan Kumar Gupta for his patience, support and encouragement throughout the completion of this project and having faith in us.

We also acknowledge to teachers like Mr. Ankit Tomar who helped me in developing the project.

At last, but not the least We greatly indebted to all other persons who directly or indirectly helped us during this work.

### Shubham Chaudhary

### Roll No.- 2017049

### CSE-B-V-Sem

**Session: 2022-2023 GEU, Dehradun**

### TABLE OF CONTENTS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S NO.**  **1.** | **TITLE**  **INTRODUCTION** | | **PAGE NO.** | 1 |
| 1.1 | | About Project | 1 | |
| 1.2 | | Methodology |  | |
| 1.3 | | Libraries Used | 3 | |
| 1.4 | | Functionalities Used | 4 | |

1. [SNAPSHOT OF PROJECT](#_TOC_250000) 5-7
2. CONCLUSION 8

SUMMARY 8

REFERENCE 9

**CHAPTER 1 INTRODUCTION**

##### ABOUT PROJECT:

These AI-powered chatbots use a branch of AI called natural language processing (NLP) to provide a better user experience. Often referred to as virtual agents or [intelligent virtual assistants](https://www.ultimate.ai/blog/ai-automation/what-is-an-intelligent-virtual-assistant-iva), these NLP chatbots help human agent by taking over repetitive and time-consuming communications.

[NLP](https://www.ultimate.ai/blog/ai-automation/what-is-natural-language-processing-a-beginners-guide-nlp)**, or Natural Language Processing**, is a branch of AI that helps computers read and understand natural human language. Its main goal is to improve human-machine communication.

**METHODOLOGY:**

Steps involved:

1. We will take input string from user

2.Then after that we will do tokenization in this input string. We will do Tokenisation. In Python tokenization basically refers to splitting up a larger body of text into smaller lines, words or even creating words for a non-English language. The various tokenization functions in-built into the nltk module itself and can be used in programs as shown below.

3. After that we will perform Stemming to stem every word after tokenization. In Python tokenization basically refers to splitting up a larger body of text into smaller lines, words or even creating words for a non-English language. The various tokenization functions in-built into the nltk module itself and can be used in programs as shown below.

4. After that we will search for the words in out intents.json file and put all the words and tags inside a array called as bag\_of\_words.

5.Then after that we will mark the found words with ‘1’ and ‘0’ for words not found.

6. Then we will train our application and it will Generate output on the basis of the probability of words found in a sentence.

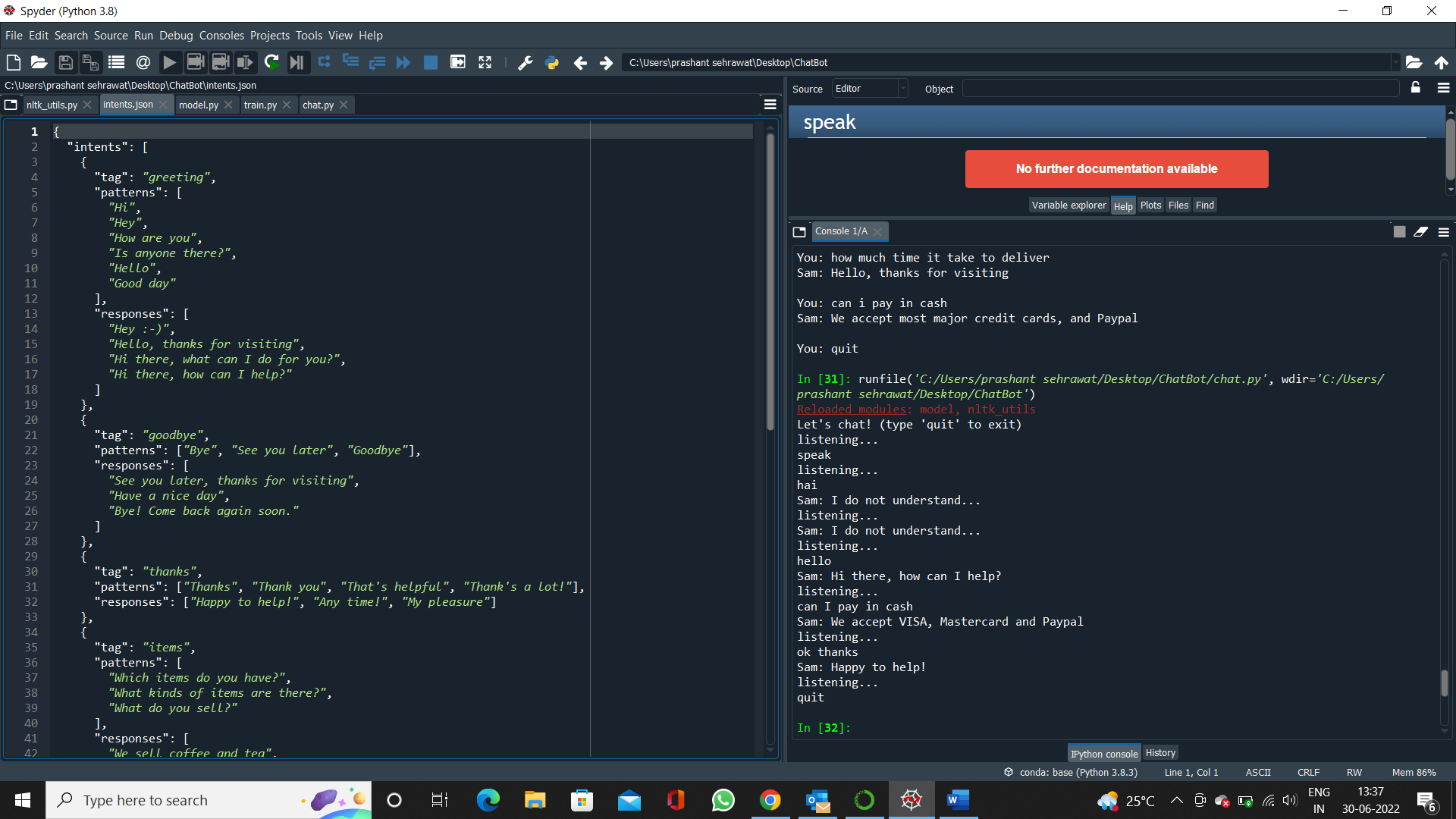
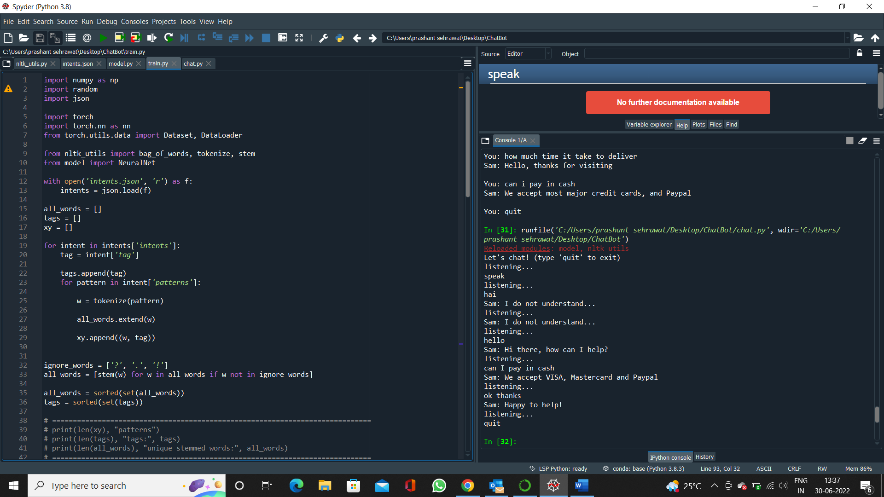
**Libraries** Used :

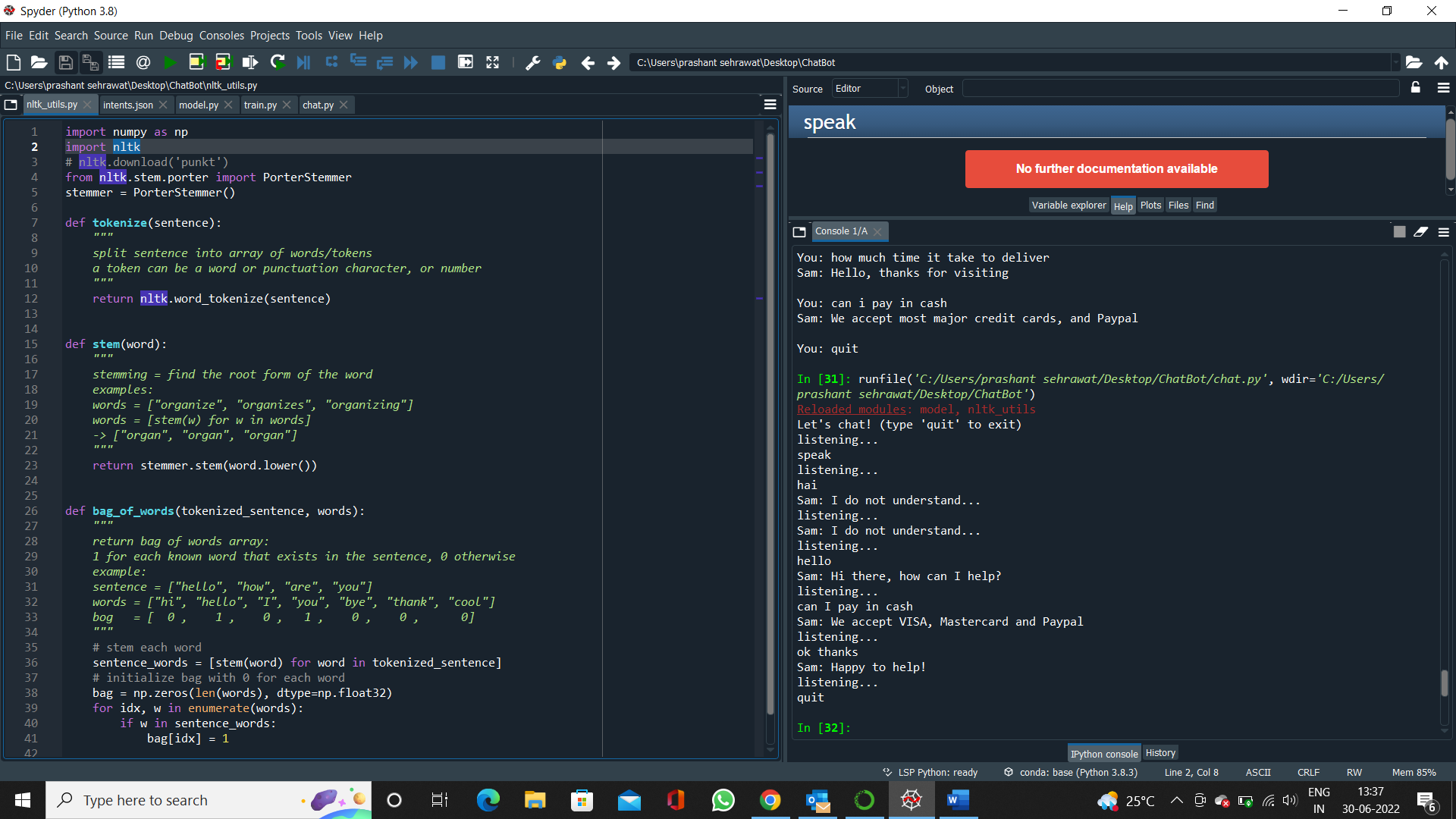
* **NUMPY:** NumPy is a Python library used for working with arrays. It supports a wide range of hardware and computing platforms.It is accessible and productive for programmers from any background or experience level.
* **JSON:** The full-form of JSON is JavaScript Object Notation. It is a popular data format used for representing structured data.It’s common to transmit and receive data between a server and web application in JSON format.
* **Speech Recognition:** Speech Recognition is an important feature in several applications used such as home automation, artificial intelligence, etc. It is the ability of a machine or program to identify words spoken and convert them into readable text.
* **Pyttsx3:** pyttsx3 is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline and is compatible with both Python 2 and 3. An application invokes the pyttsx3. init() factory function to get a reference to a pyttsx3.
* **Torch:** Torch is an open-source machine learning library, a scientific computing framework, and a script language based on the Lua programming language. It provides tensor computation and Deep neural networks build on a tape-based autograd system.
* **Nltk.Stem.Porter: Stemming**is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers. A stemming algorithm reduces the words to root words.For ex: “having” will be reduced to “have”.
* **NLTK:** The Natural Language Toolkit (NLTK) is **a platform used for building Python programs that work with human language data for applying in statistical natural language processing (NLP)**. In this interaction,understanding and response are made by a computer instead of a human.

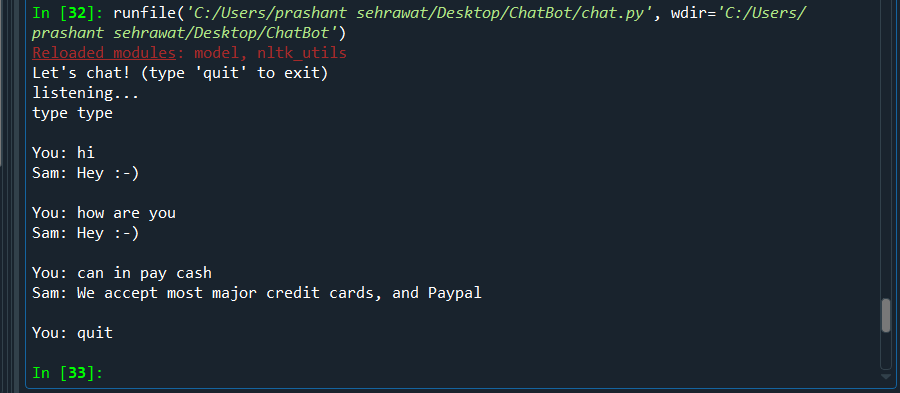
**Functionalities Used:**

1. **Stemming: It** is the process of producing morphological variants of a root/base word. Stemming programs are commonly referred to as stemming algorithms or stemmers. A stemming algorithm reduces the words “having” to the root words “have”.
2. **Tokenization**: In Python tokenization basically refers to splitting up a larger body of text into smaller lines, words or even creating words for a non-English language.It can be used for training machine learning models,Natural Language Processing text cleaning.
3. **CrossEntropyLoss:** Cross entropy can be used to define a loss function (cost function) in machine learning and optimization. It is defined on probability distributions, not single values. It is also known as the negative log likelihood.It is useful when training a classification problem.

**4.Torch.Softmax:** Takes two parameters input and dim,class torch.nn. Softmax (dim=None) [source] Applies the Softmax function to an n-dimensional input Tensor rescaling them so that the elements of the n-dimensional output Tensor lie in the range [0,1] and sum to 1.

**SnapShots:** 





**Conclusion:**

AI Based chatbot is a python-based application which uses python libraries to help in making a chatbot. We have lots of examples of chat bots today for example Flipkart, Zomato etc. These chat bots help to reduce the human work by providing customer support in these websites.

And can solve basic problems related to any task or queries. My project is Based of AI/NLP (Natural Language Processing) in this we use functionalities like tokenization, stemming and torch to make the output better and get the accuracy of our application. The result is Calculated on basis of Probability of words find in a sentence and probability. So, this is how we can get the desired output.

**References:**

1. YouTube
2. Google
3. Stack Overflow
4. Geek for Geeks