Deep Learning based Campus Assistive Chatbot

1st Bhuvana Shivashankar
Electronics and Telecommunication Engineering
R.V. College of Engineering
Bangalore, India
sbhuvana.te19@rvce.edu.in

2nd A M Anjana Sundari
Electronics and Telecommunication Engineering
R.V. College of Engineering
Bangalore, India
amanjanas.te19@rvce.edu.in

3rd Hithaishi Surendra
Electronics and Telecommunication Engineering
R.V. College of Engineering
Bangalore, India
hithaishis.te19@rvce.edu.in

4th Atul Sai S S
Electronics and Telecommunication Engineering
R.V. College of Engineering
Bangalore, India
atulsaiss.te19@rvce.edu.in

5th Minal Moharir
Computer Science and Engineering
R.V. College of Engineering
Bangalore, India
minalmoharir@rvce.edu.in

***Abstract*—With the rapidly increasing dependence and expectations on online platforms to provide faster responses during all hours of the day, a web-based chatbot could be considered as the ideal solution to do so in the educational domain. A chatbot is a conversational agent capable of emulating human interactions. The purpose of this project is to implement a chatbot for a college website in order to make it more accessible, convenient and to enhance user experience. Campus navigation through Google Maps is an additional feature which is integrated into the chatbot to better assist the user. A menu containing frequently used links is provided which redirects the webpage according to the user's choice. Using Artificial Intelligence and Deep Learning technologies, a chatbot EVA was created. Natural Language Processing is incorporated to allow the user to communicate with the chatbot using natural language input. A Sequential Model is developed using Neural Networks and the concept of Bag of Words is utilized. Tkinter is the tool used to set up the Graphical User Interface. A basic level performance test was conducted which revealed an efficiency rate of 89%. Finally, the chatbot was integrated with a college website which resulted in the generation of a functional college chatbot. EVA not only provides concise answers to the user's queries, but also necessary links to information regarding course content, extracurricular activities as well as directions to various locations on campus as per the user's requirement.**

Keywords—Deep Learning, Natural Language Processing, Sequential Model, Neural Networks, Bag of Words

I. Introduction

Chatbots aspire to construct a machine that can reliably mimic human interaction and intelligence. With messaging apps on the rise, chatbots are growing in popularity among businesses and consumers alike. Leveraging chatbots in the educational domain can help colleges to provide faster response times, 24/7 support, reduce costs by avoiding the need to hire help desk professionals and make the college website more user friendly. This can be done using Artificial Intelligence and Deep Learning technologies which involves the use of Artificial Neural Networks (ANN) [1]. Neural Networks are a type of computational learning system that employs a network of functions to comprehend and translate data supplied in one form into another [2][3]. Deep Learning is relatively more flexible as opposed to Machine Learning as it learns to represent the world as a nested hierarchy of concepts.

The intent of this project is to use the same concept to create a Campus Assistive Chatbot which allows the user to navigate through a college website easily. Any typical college website has several categories and additional subcategories, each containing detailed information. Hence, it becomes difficult for someone who is not familiar with the website to locate what they are looking for with ease. To resolve this issue, this project introduces a web based college chatbot that could be deployed on the official college website for students and other users to utilize. The Graphical User Interface (GUI) is quite similar to existing messaging applications such as Facebook Messenger and WhatsApp. The question posed by the user is preprocessed and the most suitable answer from the database is displayed. The answers are either provided directly or a link to the required page is provided. In a way, the chatbot behaves like a search engine specifically designed to suit the needs of a college website.

In addition to all this, a feature is added to provide directions from a particular location to any Department or Office on campus using Google Maps. This allows newcomers and those who are unfamiliar with the vast campus to locate their destination without any difficulty. Also, a pop-up menu is displayed containing links to regularly visited categories of the college website.

II. Literature Review

Artificial intelligence is a branch of science that has been utilized to create a variety of applications. The college chatbot system is one example of such an application. According to the most recent technology information, one of the parameters in the development of the EVA chatbot is Natural Language Processing [4][5][6][7], which is supported by Natural Language Tool Kit (NLTK) [7]. Tokenization, lemmatization, stemming, and Bag of Words (BoW) are used for preprocessing [4][7]. The NLTK has a large range of tokenizers, including those that conform to the following standards: letters, path, words, keywords, class, pattern, and so on. The chatbot application responds to the questions posed by users using computerized reasoning.

III. Proposed System

This study proposes a accessible web-based interface using artificial intelligence concepts [8][9]. To develop and train models capable of executing the actions outlined below, multiple Deep Learning techniques are used:

- The virtual assistant is opened and it proceeds to ask the user how it can be of assistance and engages in casual conversation.

978-1-6654-0610-9/21/$31.00 ©2021 IEEE

- The queries received from the user is analyzed and preprocessed in order to extract the most appropriate response from the chatbot database.
- Suitable links and specific directions are provided in the Links menu as per the user's requirement.
- Upon completion, the chatbot cordially greets the user goodbye and its purpose is accomplished.

A. Natural Language Processing

Natural language Processing (NLP) uses algorithms to understand the user's message. NLP is supported by Natural Language Toolkit (NLTK) which is a python library having inbuilt tokenizers. Chatbots use NLP to recognize user's aim during the conversation. The amount of training data is the key to have a good conversation with the user and the data quality determines the ability of the Chatbot to recognize the intent and generate the appropriate response.

B. Libraries

Libraries like Natural Language Tool Kit (NLTK), Java Script Object Notation(JSON), Pickle, NumPy and Random were used to implement the bot. NLTK is a python library for statistical NLP which can perform many operations such as tokenizing, tagging, stemming and classification. The intents dataset is stored as a JSON file. Pickle module is used for serialization and deserialization of data structure. NumPy can perform linear algebra operations. Keras helps with the deep learning framework.

C. Intents

Creating intents and recognition is a critical factor in Chatbot's architecture (Fig.2) . This is a simple dataset with different tags. Each tag has patterns and responses. Patterns contains the keywords of questions whereas responses have the answers. During the conversation, the intents are first mapped to the tags, patterns and later generates the corresponding response. This is shown in Table I.

D. Data Preprocessing

Data preprocessing is the first and crucial step while creating a machine learning or a deep learning model (Fig.1). It involves cleaning and formatting the data to increase the accuracy and efficiency of the model. To begin, punctuation removal is performed. Since punctuations are of little use in NLP, they are ignored. In addition to this, Stop Words are also removed which are frequently occurring words. This is done for text normalization. After extracting all the words within patterns and adding to the word list, process of tokenization and lemmatization takes place.

E. Tokenization

Tokenization helps in splitting a large quantity of text into smaller segments called "Tokens" based on the requirement. These tokens helps to find patterns and is one of the base step for lemmatizations. There are two types of tokenization - word and sentence tokenization. In this model, the concept of word tokenization which breaks down a sentence into words is used. This returns a list of words as output.

TABLE I. INTENTS

Tag	***Patterns***	***Responses***
Student Clubs	-do you have any club activities -how many student clubs do you have -student clubs -technical innovative clubs -group activities -interdisciplinary clubs	-We provide a total of 14 clubs. -RVCE hosts 14 technical innovative student clubs. -Our 14 student clubs include: Ashwa Racing, Solar Car Team, Frequency Club, Entrepreneurship Development Cell, Coding Club and 9 others.

F. Lemmatization

Lemmatization is a text normalization technique for NLP. It helps to reduce the words into their roots which have actual meaning. This saves a lot of time while processing the words. WordNetLemmatizer imported from wordnet is used.

G. Predicting classes and Response Selection

Classification refers to a predictive modeling problem where a class label is predicted for a given example of input data. Once the class is predicted, a random response is extracted from that class and displayed to the user.

The flowchart of the system in Fig.1 displays how the EVA chatbot performs. Initially, the user input is obtained. The flow of control changes depending on whether the input is a query or a link. If it is a query, data preprocessing takes place using the concepts explained in section III, followed by classification and response selection. The output is then displayed to the user. On the other hand, if the user opts to explore the links provided, the web page will be redirected to link of the user's choice. Sometimes, the chatbot may urge the user to inspect the links option when a response is provided. In this case, the same flow of control continues as mentioned previously and the required task is accomplished.

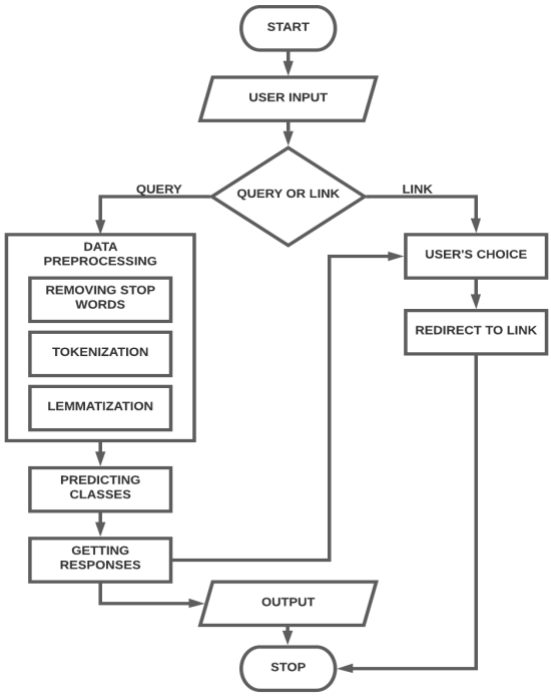

Fig. 1. Flowchart depicting the flow of control of the proposed system.

IV. Implementation

The following technologies and frameworks were used for the implementation:

A. Sequential Model

A Sequential model from Keras is one of the simplest neural networks. It allows to build a model layer by layer and each layer has weights. The developed chatbot model is made up of a three layered network: an input, a hidden and an output dense type layer. A dense type of layer is used where all nodes in the previous layer connects to the nodes of current layer. An activation function helps the model to consider nonlinear relationships. Rectified Linear Unit (ReLU) and Softmax (scales the result in the output layer) are the activation functions used here. Dropout layers in the model help us regularize the chatbot. The model is trained with Stochastic Gradient Descent (SGD) (Fig.2). After training, the model is saved as a hierarchical data format 5.
The above network is developed in order to predict which intent has to be chosen for some given data. Functions to create bag of words, predicting classes and getting responses is created to get the probability of matching to the right intent, and print the response with the highest probability by referring the JSON file.

B. Bag of Words

At a much granular level Deep learning model needs numerical data for computation and not textual one. Therefore, to represent words as numerical values we use the concept of Bag of Words (BoW). It represents the occurrence of words within a document. This model just checks if the word occurs in the document or not. It also keeps a track of the word counts and ignores the grammatical details. This technique can convert a variable-length text into a fixed-length vector i.e. it converts the word into its equivalent vector of numbers. The dataset is shuffled with patterns and intents corresponding to different variables and is divided into training and testing dataset.

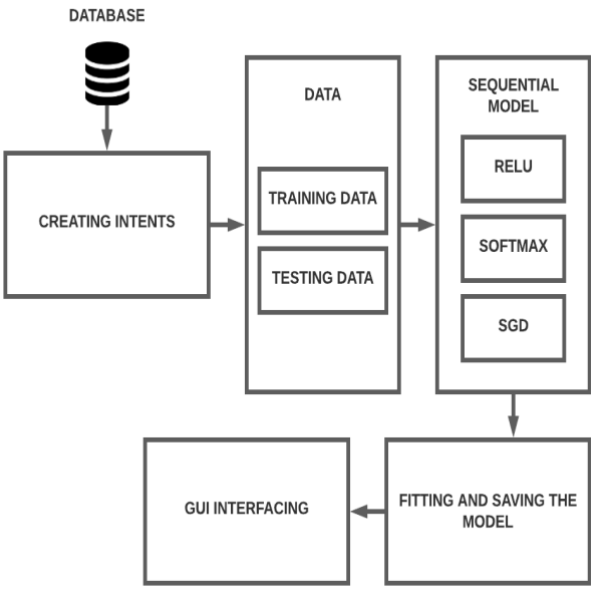

Fig. 2. System architecture.

C. Tkinter for Graphic User Interface

The most frequently used method from the multiple options Python offers for developing GUI is tkinter. Tkinter is used for Graphical user interface, development of the chat window, text box, proper positioning of buttons and to get the redirecting page for the links of the location of destinations the user wishes to reach in the campus (Fig.2).

V. Results

A chatbot is successfully implemented using the technologies and algorithms discussed in sections III and IV. Fig.3 shows EVA answering the user's queries with direct and concise information.

Fig. 3. EVA chatbot answering the user's queries.

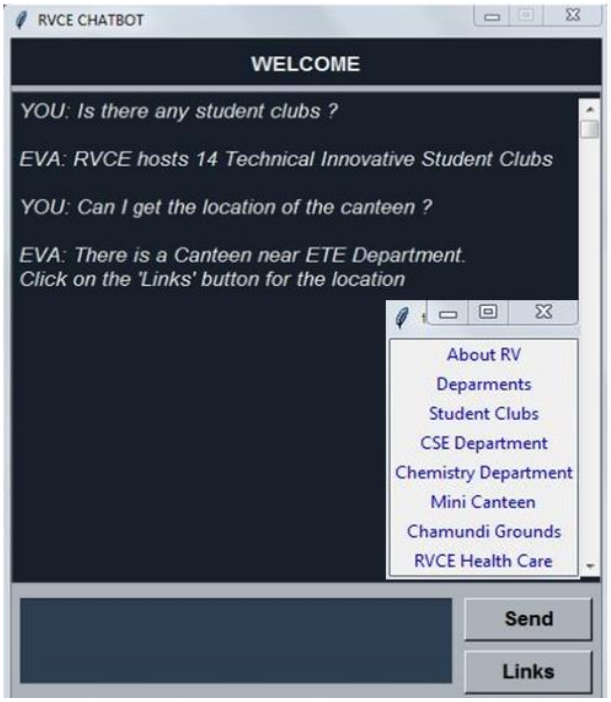

Fig. 4. EVA chatbot providing links pop up menu.

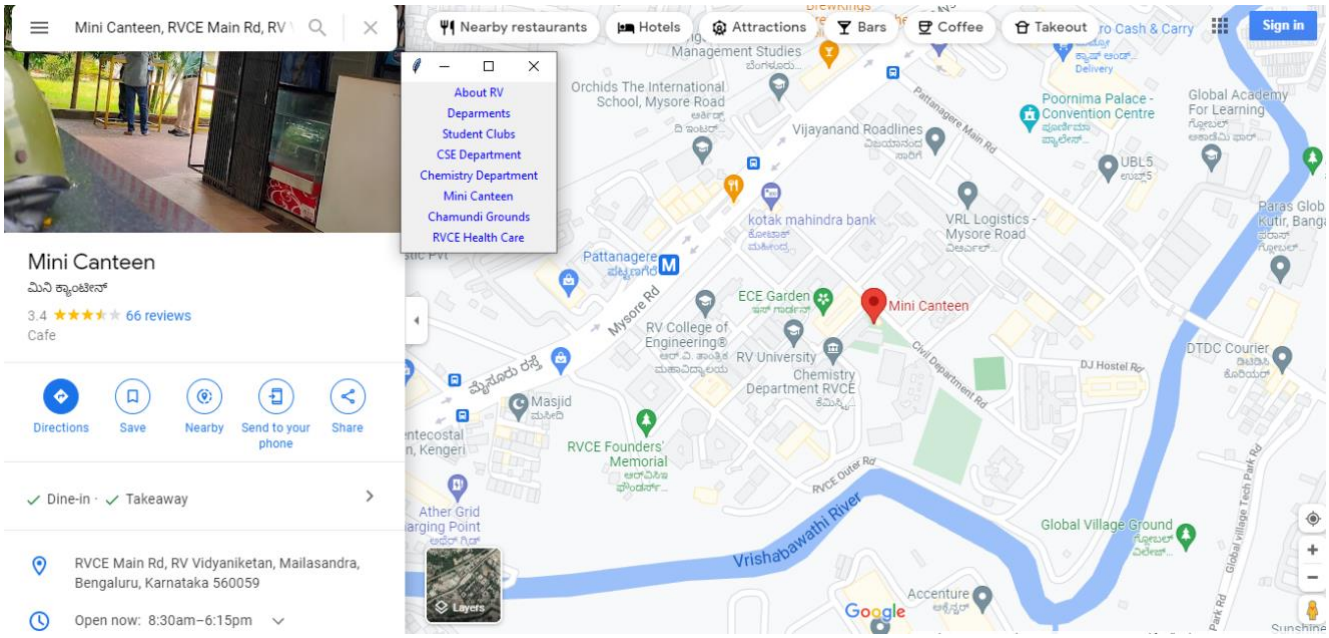

Fig. 5. Redirected webpage providing the location of Mini Canteen in Google Maps.

In Fig.4, on requesting the location of Mini Canteen, the chatbot provides a pop-up menu containing all the links available. As seen in the figure, several options including health care centre and grounds is provided. Once a link is selected, the webpage is redirected, either to give more information or to the location of the user's choice on Google Maps. This is depicted in Fig.5.

VI. Performance Test

For testing the accuracy of the virtual assistant, 100 interactions were recorded [10]. The observations from the conducted survey are provided in Table II. According to the results obtained, it is observed that the chatbot has an accuracy rate of 89%. 'Correct' means that the chatbot processes the query or response input by the user and provides the necessary information as opposed to displaying a default error message "Unable to understand. Could you please repeat?"

TABLE II. Results from Perfromance Test

Total Interactions: 100	
Correct	***Incorrect***
89	11

VII. Conclusion

A Campus Assistive Chatbot was developed to ease the process of accessing relevant information from a college website. In this paper, details about design, algorithm used, implementation and performance results of the chatbot are presented. Additional features such as campus navigation are successfully incorporated as well. The efficiency of the chatbot obtained can be further increased to provide more meaningful and accurate data by training it with more information regarding the college and expanding its database. Overall, the virtual assistant is convenient and easy to use. With the growing reliance on online platforms to give speedier, intelligent responses at all hours of the day and night, a web-based chatbot that can imitate human interactions could be considered the best solution in the educational arena.

References

[1] Di Franco, Giovanni & Santurro, Michele. (2021). "Machine learning, artificial neural networks and social research." Quality and Quantity. 55. 10.1007/s11135-020-01037-y.

[2] Dhyani, Manyu & Kumar, Rajiv. (2020). "An intelligent chatbot using deep learning with bidirectional RNN and attention model." Materials Today: Proceedings. 34. 10.1016/j.matpr.2020.05.450.

[3] Collobert R. and Weston J., "A unified architecture for natural language processing: deep neural networks with multitask learning.," in In Proceedings of the 25th international conference on Machine learning, 2008.

[4] 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), 2020, pp. 1-5, doi: 10.1109/ICSCAN49426.2020.9262426.

[5] J. Purohit, A. Bagwe, R. Mehta, O. Mangaonkar and E. George, "Natural language processing based Jaro-the interviewing chatbot," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 134-136, doi: 10.1109/ICCMC.2019.8819708.

[6] Mondal, M. Dey, D. Das, S. Nagpal and K. Garda, "Chatbot: an automated conversation system for the educational domain," 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), 2018, pp. 1-5, doi: 10.1109/iSAI-NLP.2018.8692927.

[7] Naeun Lee, Kirak Kim, Taeseon Yoon, "Implementation of robot journalism by programming custombot using tokenization and custom tagging" International Conference on Advanced Communications Technology (ICACT) Page no: 566-570| Feb 2017.

[8] N. P. Patel, D. R. Parikh, D. A. Patel and R. R. Patel, "AI and web-based human-like interactive university chatbot (UNIBOT)," 2019 3rd International conference on Electronics, Communi- cation and Aerospace Technology (ICECA), 2019, pp. 148-150, doi: 10.1109/ICECA.2019.8822176.

[9] Amey Tiwari, Rahul Talekar, Prof.S.M.Patil, "College information chat bot system" International Journal of Engineering Research and General Science (IJERGS) Volume: 5, Issue: 2, Page no: 131-137| March-April 2017.

[10] B. Kohli, T. Choudhury, S. Sharma and P. Kumar, "A platform for human-chatbot interaction using python," 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT), 2018, pp. 439-444, doi: 10.1109/ICGCIoT.2018.8753031.