

20BCE1538 – MAYANK GUPTA
Dr. Geetha S. | Dean Research, SCOPE

Abstract:

Chatbots are more scalable and popular than earlier rule-based chatbots. The VIT-FAQAssist is a custom-built conversational agent designed to address frequently asked questions related to VIT. This chatbot is intended to be used in university education sector for answering frequently asked questions about the university and its related information. In this report, we explore the ways of communication through neural network chatbot by using two distinct models, one based on TensorFlow and the other on PyTorch, handle intent classification, efficiently categorizing user queries into predefined intent categories.

In this report, we introduce a novel question-answering chatbot that employs Convolutional Neural Networks (CNNs) for intent recognition, followed by data augmentation and a combination of Pytorch and Tensorflow models to find the right answer from the output.json provided by the university.

Keywords: Artificial Intelligence Markup Language(AIML), Pytorch, Tensorflow, Pattern Matching, Chatbot, Human Computer Interaction(HCI).

Introduction:

In spite of the number of techniques, models and datasets, Question Answering is still an exacting problem because of the issues in understanding the question and extracting the correct answer. The influx of new students and faculty members creates a surge in queries, adding complexity to the management of communication channels between authorities and the academic community.

In this project, we recruit NLP methods of pattern matching and information retrieval to create an answer candidate pool. Before scoring similarities between the question and answers, we map them into some feature space. Our approach solves this task through distributional representations of the words and sentences wherein encodings store their lexical, semantic, and syntactic aspects..

Our model learns an optimal representation for the input question and answer sentences and a matching function to relate each such pair in a supervised manner from training data.

Proposed Methodology:

Algorithms Implemented: *Text Synthesis*

The following outlines the algorithm for answer processing:

Step 1: Mapping Words Tokenization involves breaking down a sequence of text into individual words or tokens. Words are mapped to discrete units to facilitate further analysis and processing. Tokenization helps in preparing the input data for the subsequent steps in the natural language processing pipeline.

Step 2: Intent Classification In the process of building an intent classification system, the training of the PyTorch model involves several key stages. Pytorch & Tensorflow model uses a supervised learning is employed to train the model on the dataset, with the model adjusting its internal parameters to effectively map input queries to their intended classifications. The training process involves the definition of a loss function to quantify the model's performance, coupled with optimization algorithms like stochastic gradient descent to iteratively refine the model's parameters.

Step 3: Random Answer

It is used by the Chatbot to understand the intents of the user and reply with a particular action to the user. The Domain has all the information that is needed by the chatbot to understand what the user wants to convey and provide an appropriate answer to the query raised by the user

Step 4: Login to DB Record and store user interactions, including queries, intents, and relevant information in a database (DB). Maintain a user record in the database for future reference and analysis.



Result:

Chatbot is implemented to meet the academic needs of the visitors. The chatbot is based on AIML language for VIT University. This will help the student to fetch information like ranking of university, availability of services, university environment, updates regarding activities happening inside campus and many more and other academic information. A snapshot of the proposed chatbot is shown in Fig.2. The above conversation is a result of partially trained dataset where the chatbot is able to correctly identify intents of the user and provide answers from database. The more the model is trained the better will be the accuracy of the bot.

Figure 2 Response of VIT-FAQAssist API

```

Curl
curl -X 'POST' \
  http://127.0.0.1:8080/chatbot/ \
  -H 'accept: application/json' \
  -H 'Content-Type: application/json' \
  -d '{
    "user_input": "Hi What is tttt",
    "model_selection": 2
  }'

Request URL
http://127.0.0.1:8080/chatbot

Server response
Code    Details
200     Response body
{
  "user query": "Hi What is tttt",
  "predicted intent": "tense what",
  "output to user": "As the name suggests, we are a student",
  "model used": "PyTorch",
  "response time": "0.025136072918221",
  "mongoDb append status": "appended"
}

```

Conclusion:

In conclusion, the development of this chatbot will provide an easier way for students and any person to solve their queries faster and in an easier way. Not only it will help people to get their queries solved easily but it will also help reduce the work stress of the receptionist. The chatbot will be developed using RNN and LSTM, because of The The above conversation is a result of partially trained dataset where the chatbot is able to correctly identify intents of the user and provide answers from database. The more the model is trained the better will be the accuracy of the bot which the chatbot will be able to frame its answer if the answer for that particular question is not available in the database..

Contact Details: mayank.gupta2020@vitstudent.ac.in

References:

- [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), Udupi, India, 2017, pp. 1525-1530, doi: 10.1109/ICACCI.2017.8126057.
- [2] A. Huddar, C. Bysani, C. Suchak, U. D. Kolekar and K. Upadhyaya, "Dexter the College FAQ Chatbot," 2020 International Conference on Convergence to Digital World - Quo Vadis (ICCDW), Mumbai, India, 2020, pp. 1-5, doi: 10.1109/ICCDW45521.2020.9318648.