TRF-CHATBOT

**Developed By :**

Atharva Purohit Vithika Pungliya Omkar Jahagirdar

atharva.purohit20@vit.edu vithika.pungliya20@vit.edu omkar.jahagirdar20@vit.edu

Siddhant Pawar Varad Ingale Chetan Chaudhari

siddhant.pawar20@vit.edu varad.ingale20@vit.edu [chetan.chaudhari20@vit.edu](mailto:chetan.chaudhari20@vit.edu)

Ridhya Ghosh

ridhya.ghosh19@vit.edu

**Mentored By :**

Ridhya Ghosh Rohit Singh

ridhya.ghosh19@vit.edu rohit.singh19@vit.edu

**Introduction :**

A chatbot is an intelligent piece of software that is capable of communicating and performing actions similar to a human.As chatbots are used a lot in customer interaction, marketing on social network sites and instantly messaging the client.There are two basic types of chatbot models i.e Retrieval based and Generative based models. Here we have used a Retrieval based model as it uses predefined input patterns and responses.

The chatbot is built using deep learning techniques and is made in order to have a quick response for the information related to ‘The Robotics Forum’. The project aims to deliver solutions for the questions raised by the people to have better understanding about TRF.

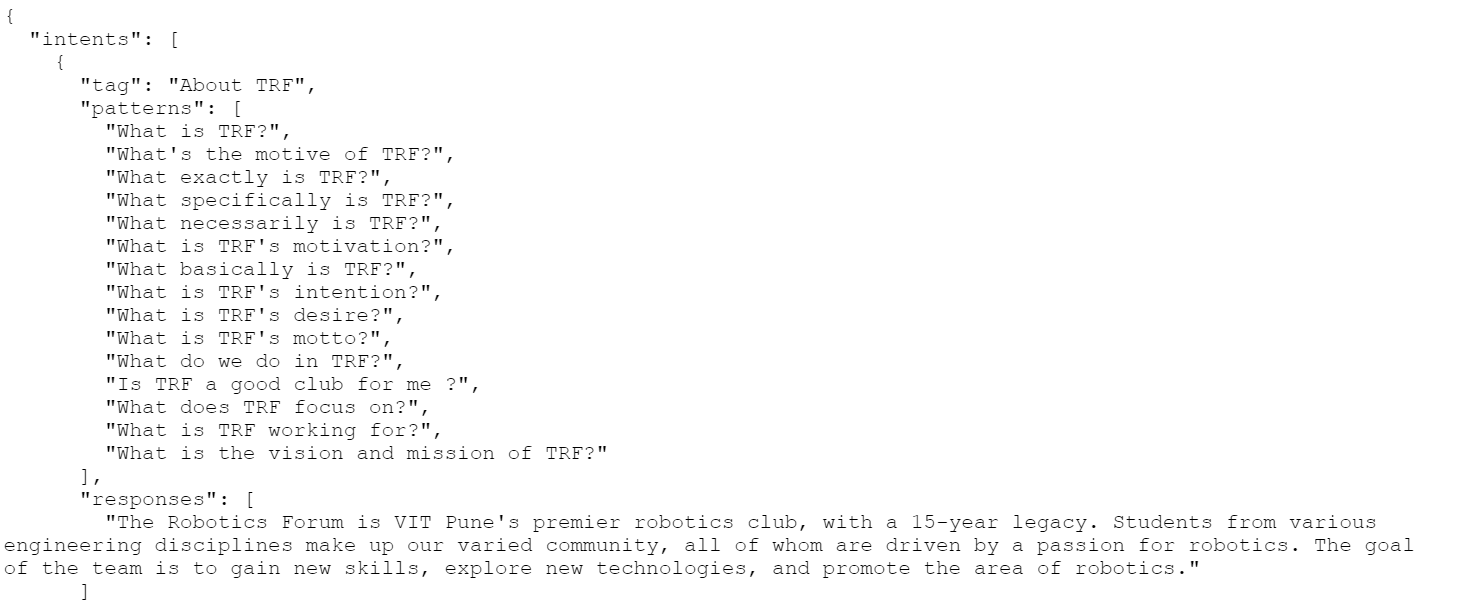
In this project we are going to build a chatbot using deep learning techniques. The chatbot will be trained on the dataset which contains categories (intents), patterns and responses. We use a special recurrent neural network (LSTM) to classify which category the user’s message belongs to and then we will give a random response from the list of responses. The project concerns more about ‘The Robotics Forum’.

**Methodology :**

**Modules Used**

* Numpy
* Pickle
* Json

**i) Custom intent creation:** Intents were customized according to the requirements of ‘The Robotics Forum’. As the project majorly focuses on the information related to TRF. As an intent is to perform an action on the screen. It is mostly used to start an activity, send a broadcast receiver,start services and send messages between two activities. There are two intents available in android as Implicit Intents and Explicit Intents. We have worked on Explicit Intents.



**ii) Data preprocessing**

**Tokenization**: It is the process of breaking text into smaller pieces called tokens. These smaller pieces can be sentences, words, or sub-words.

We use tokenizer as it breaks the raw text into words, sentences called tokens. These tokens help in understanding the context or developing the model for the NLP. The tokenization helps in interpreting the meaning of the text by analyzing the sequence of the words.

In this model we have used ‘**OOV**’ tokens as Oov tokens are out of vocabulary tokens used to replace unknown words.

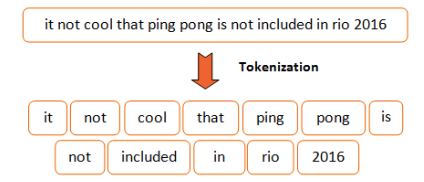


Figure: Explanation of the Tokenization

**Other parameters:**

* **Max len** is used because short messages of the bot prompt the user to act immediately where the page requires more reading.

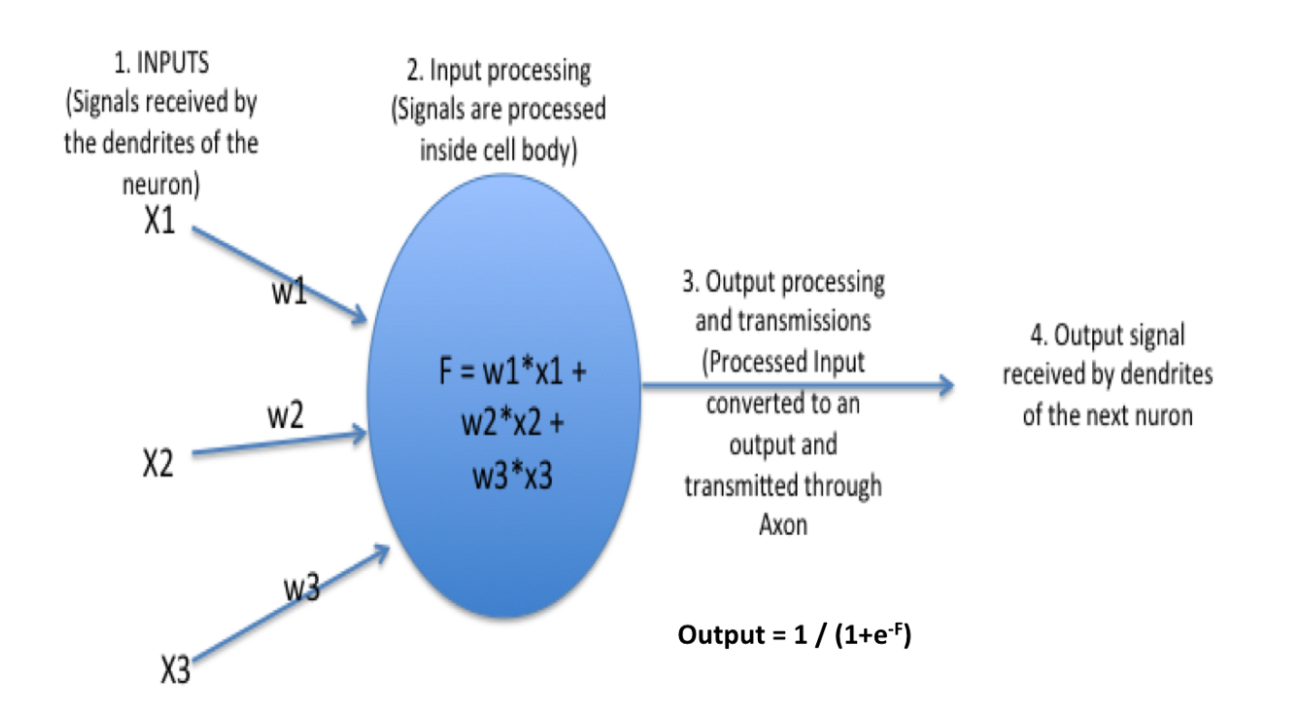
By the same token, having a long initial message would completely nullify the user engagement that a bot offers because a user has to spend time reading and scrolling through the long message.

* **Truncation** is done if a sequence length is larger than the provided maxlen value then, these values will be truncated to the maxlen. It consists of two subtypes ‘pre’ and ‘post’ wherein ‘pre’ is default. 'pre' option will truncate at the beginning whereas 'post' will truncate at the end of the sequences.

In this model we have used the ‘**post**’ method.

**iii) Model building**

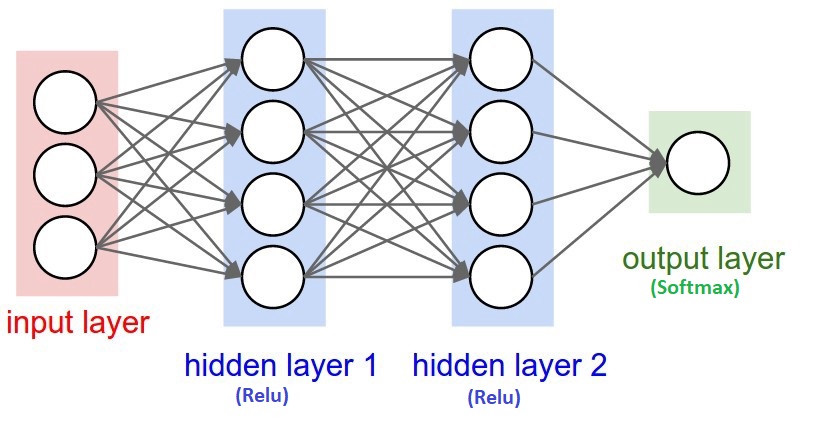
**ANN:**An Artificial Neural Network (ANN) is similar, but a computing network in science that resembles the properties of the human brain.



It uses the processing of the brain as a basis to develop algorithms that can be used to model complex patterns and prediction problems.

* Artificial Neural Networks can be best viewed as weighted directed graphs, where the nodes are formed by the artificial neurons and the connection between the neuron outputs and neuron inputs can be represented by the directed edges with weights. The Artificial Neural Network receives the input signal from the external world in the form of a pattern and image in the form of a vector. These inputs are then mathematically designated by the notations x(n) for every n number of inputs.
* Each of this input is then multiplied by its corresponding weights. In general terms, these weights typically represent the strength of the interconnection amongst neurons inside the artificial neural network. All the weighted inputs are summed up inside the computing unit.
* If the weighted sum equates to zero, a bias is added to make the output non-zero or else to scale up to the system’s response. Bias has the weight and the input to it is always equal to 1. Here the sum of weighted inputs can be in the range of 0 to positive infinity. To keep the response in the limits of the desired value, a certain threshold value is benchmarked. And then the sum of weighted inputs is passed through the activation function.
* The activation function, in general, is the set of transfer functions used to get the desired output of it. There are various flavors of the activation function, but mainly either linear or non-linear sets of functions. Some of the most commonly used sets of activation functions are the Binary, Sigmoidal (linear) and Tan hyperbolic sigmoidal (non-linear) activation functions.

The **softmax** function is used as the activation function in the output layer of neural network models that predict a multinomial probability distribution. That is, softmax is used as the activation function for multi-class classification problems where class membership is required on more than two class labels.



**Global Average Pooling** is a pooling operation designed to replace fully connected layers in classical CNN**s**. The idea is to generate one feature map for each corresponding category of the classification task in the last mlpconv layer.

**LSTM:**Long Short Term Memory (LSTM) is an advanced Recurrent Neural Network (RNN), a sequential network, that allows information to persist. It is capable of handling the vanishing gradient problem faced by RNN. A recurrent neural network is also known as RNN is used for persistent memory.

The output of an LSTM at a particular point in time is dependant on three things:

▹ The current long-term memory of the network — known as the *cell state.*

▹ The output at the previous point in time — known as the previous *hidden state.*

▹ The input data at the current time step.

LSTMs use a series of ‘gates’ which control how the information in a sequence of data comes into, is stored in and leaves the network. There are three gates in a typical LSTM; forget gate, input gate and output gate. These gates can be thought of as filters and are each their own neural network.

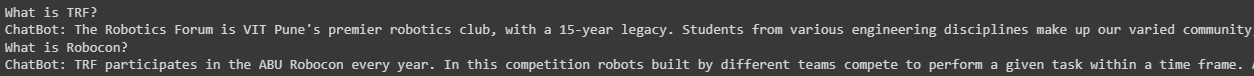
This model comprises two hidden layers and one output layer.

**iv) Testing**

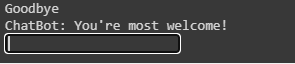
**Greeting:**The Chatbot would greet at the start.



**Questions:**Answer the following questions raised by the people.



**Closing:**Chatbot welcome’s people by answering questions.



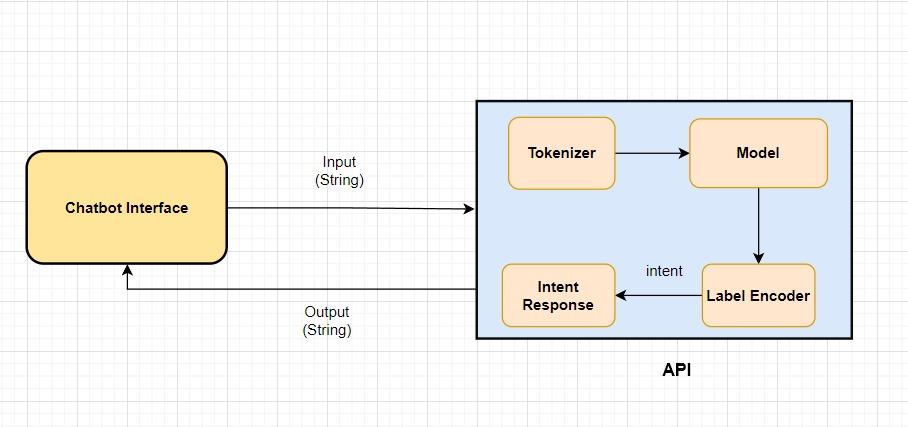
**v) Model Conversion**

The model is converted so that the chatbot would efficiently work on the website. In order to do so following commands are necessary to be run as per the mentioned syntax.

The command is particularly used for the '.h5' extension.

1. **$ tensorflowjs\_converter --input\_format=keras 'folder directory/file name.h5'**
2. **!tensorflowjs\_converter --input\_format keras folder directory/file name.h5**

**vi) Assumed data flow**



The above figure elaborates the working of the chatbot via flowchart. Wherein greeting to the fellow people at initially and also resolving their queries related to ‘The Robotics Forum’ is kept on priority and providing accurate responses to the asked questions. Moreover, chatbots run endlessly until one is satisfied with his/her questions.

**vii) Summary**

In this project, we understood about the chatbots and their implementation techniques used. Also further intents would be customized according to the need. The chatbot majorly focuses on ‘The Robotics Forum’ related information. The chatbot would help fellow people to know a lot of information about the motive of the club.