

## ChatBot:-

ChatBot has implemented using python flask framework using a combination of natural language processing (NLP) techniques and machine learning. Here's an overview of the development process:

### 1. Dataset Preparation:

- a. The chatbot development starts with preparing a dataset. In this case, the dataset is stored in a JSON file called "intents.json." It contains a collection of intents, each consisting of patterns (user inputs) and corresponding responses.
- b. The dataset is loaded using the `json.load()` function, and empty lists are initialized to store words, labels, training data, and output data.

### 2. Text Preprocessing:

- a. The text preprocessing step involves tokenizing the patterns and performing stemming on the words. Tokenization breaks down the patterns into individual words or tokens, while stemming reduces words to their base or root form.
- b. The `nlk.word_tokenize()` function is used to tokenize the patterns, and the `LancasterStemmer` from `nlk.stem.lancaster` is used for stemming.
- c. The stemmed words are added to the words list, and the tokenized patterns and their corresponding intents (labels) are added to the `docs_x` and `docs_y` lists, respectively.

### 3. Building Vocabulary:

- a. The unique stemmed words are extracted from the words list, sorted, and stored back in the words variable. This creates the vocabulary of the chatbot.
- b. The unique labels are also extracted, sorted, and stored in the labels variable.

### 4. Creating Training Data:

- a. The training data is created by converting the tokenized patterns into a bag-of-words representation. Each pattern is transformed into a numerical vector where each element represents the presence or absence of a word from the vocabulary.
- b. For each pattern, a bag-of-words vector is created by iterating through the vocabulary and setting the corresponding element to 1 if the word is present, or 0 if it is absent.
- c. The training data is stored in the training list.

### 5. Creating Output Data:

- a. The output data is created by converting the intents (labels) into one-hot encoded vectors. Each intent is transformed into a binary vector where only the element corresponding to the intent is set to 1, and all other elements are set to 0.

- b. The output data is stored in the output list.

6. Model Building and Training:

- a. The tflearn library is used to build and train a neural network model for the chatbot.
- b. The model architecture consists of an input layer with a shape corresponding to the length of the training data, followed by two fully connected layers with 8 units each, and a final output layer with units equal to the number of unique intents.
- c. The model is trained using the fit() method, specifying the training data, output labels, number of epochs, batch size, and other training parameters.
- d. During training, the model learns to classify user inputs into the appropriate intents based on the provided patterns.

7. Saving the Model:

- a. After training, the model is saved in two different formats: as a pickle file ("model.pickle") and as a tflearn model file ("model.tflearn").
- b. The pickle file stores the necessary data (words, labels, training, and output) to use the trained model for making predictions.
- c. The tflearn model file stores the model architecture and weights, allowing for easy loading and further training if needed.
- d. Overall, the chatbot is developed by preprocessing the dataset, building a neural network model using tflearn, and training the model using the prepared training data. The resulting model can then be saved and used to classify user inputs and generate appropriate responses in a chatbot application.