**ChatBot**

A chatbot is an intelligent piece of software that is capable of communicating and performing actions similar to a human.

2 types: Rule based

Self based:

1. Generative(No pre defined responses) & 2.Retrieval(Pre defined input patterns and responses)

We have built a **Retrieval bot for A specific website.**

1. Input data from Chicken pox vaccination website.
2. Packages used:

* Nltk : Natural Language Toolkit,for the bot to understand the English language by tokenization.
* Punkt :predefined in NLTK for sentence tokenization
* newspaper3k : To gather info from the website.
* Scikitlearn : ML library

1. Libraries used :

* Random
* Article
* Numpy
* Warnings
* CountVectorizer : Convert a collection of text documents to a matrix of token counts, for feature selection.
* Cosine similarity

**Steps**:

1. importing libraries
2. Getting the article
3. Tokenization : breaking the whole text into small tokens,here,**Sentences** for the bot to understand.
4. Generating Greetings (pretty direct)
5. Creating Bot responses :

* Since machines don’t understand any language, the Strings have to converted to Numeric form using Vectorization.
* Count matrix : sentence tokens contained here in the form of 0’s and 1’s.
* Similarity Score : for comparison of user’s input to count matrix,gives us similarity score to user’s input.

//For observing difference between user input and document,learing and mapping user’s input to document. (Using cosine similarity)

* Similarity score list : for reducing the dimensionality.
* Index : will contain highest value in the similarity score list after similarity score list has been sorted.

(If the input does not have similarity to any of the sentences in text,it will be stored in the last with value=1,and compared to others =0)

* Response flag (intially =0)

If similarity score > 0 ,i.e., the bot has found something similar to user’s input,bot will return some response that it found(with highest similarity score). And response flag will become =1.

If response flag ==0 , no similar sentences found by the bot, hence ,it returns “I don’t understand”.

1. Starting Chat with the bot (pretty direct,a simple function).