

CHATBOT IMPLEMENTATION USING NLTK AND KERAS

VIBHABHOJ, UMARANI. C (AssociateProfessor)

Dept of Master of Computer Applications

JAINUNIVERSITY

ABSTRACT

Chatbots, or conversational interfaces as they are additionally known, current a new way for folks to have interaction with pc systems. Traditionally, to get a query answered by using a software program application worried about the use of a search engine, or filling out a form. A chatbot lets in a consumer to actually ask questions in the identical manner that they would tackle a human. The most nicely regarded chatbots presently are voice chatbots: Alexa and Siri. However, chatbots are presently being adopted at an excessive price on laptop chatplatforms.

The technological power at the core of the upward jostle of the chatbot's natural-language process ("NLP").

Recent advances in machine learning have considerably accelerated the accuracy and effectiveness of natural-language process, creating chatbots a practicable selection for several organizations.

Keywords:-chatbot,natural language processing,keras,nltk,python

INTRODUCTION

A very simple chatbot can be created by means of loading an FAQ (frequently requested questions) into chatbot software. The performance of the chatbot can be accelerated by integrating it into the

company's software, permitting greater non-public questions to be answered, like "What is my balance?", or "What is the popularity of my order?".

APPLICATIONS

A chatbot is used anywhere a person is interacting with a system.

These are the areas where the quickest adoption is occurring:

Customer Service — A chatbot is used as an "assistant" to a live agent, increasing the agent's potency. Once trained, they will conjointly offer service once the decision centre is closed, or eventually even act as an freelance agent, if desired.

Sales/Marketing/Branding — Chatbots is used for sales qualification, ecommerce, promotional campaigns, or as a stigmatisation vehicle.

Human Resources — An hour chatbot will facilitate with commonly asked queries ("how several vacation days do I actually have left?")

PROBLEM STATEMENT

Changing of necessities to construct a project, we ought to observe the requirement from the user.

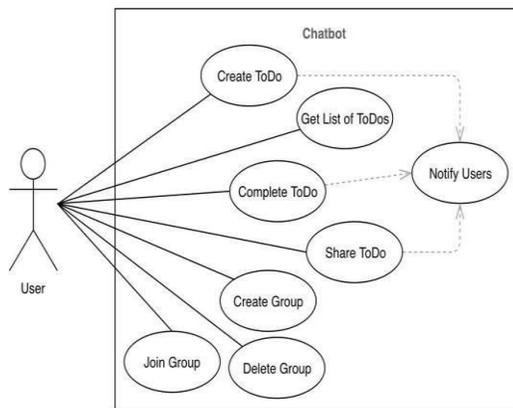
- Lacking of coaching data->The extent and satisfactory of education information is vital

to the overall performance to a desktop getting to know model.

- Develop a technique of integrating a frequent experience database with a chatbot such that it would be capable to shape particularly coherent messages primarily based on phrases related withthe user'sentry.

DESIGN

i. Use CaseDiagram



IMPLEMENTATION

The project requires you to possess good knowledge of Python, Keras, and Natural language processing (NLTK). Along side them, we'll use some helping modules which the user can download using the python-pip command.

-> pip install tensorflow, keras, pickle, nltk

Here are the 5 steps to create a chatbot in Python from scratch:

1. Import and load the datafile
2. Preprocessdata
3. Create coaching and trying outdata
4. Build themodel
5. Predict theresponse

1. Import and load the datafile

First, make a file to train the data. We import the necessary packages for our chatbot and initialize the variables we will use in our Python project. For example:

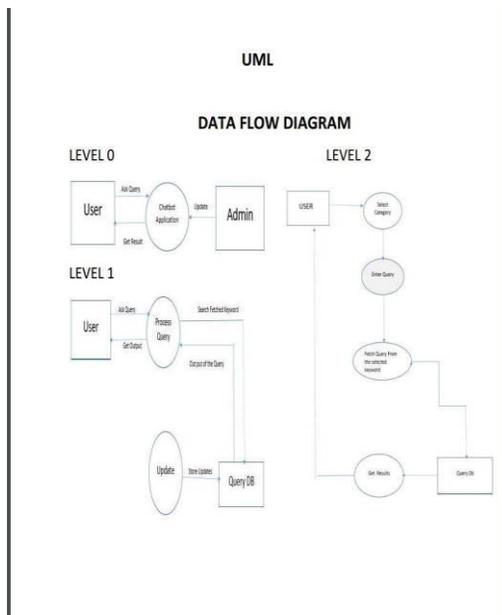
```
import nltk
```

1. from nltk.stem import WordNetLemmatizer
2. lemmatizer =WordNetLemmatizer()

2. Preprocessdata

When working with textual content data, we would like to function quite a couple of preprocessing on the statistics before we make a desktop mastering or a deep mastering model. Tokenizing is the most fundamental and first issue a user can do on textual content data. Tokenizing is the manner of breaking the total textual content into small components like words.

ii.



Here we iterate via the patterns and tokenize the sentence the use of `nltk.word_tokenize()` feature and append every phrase in the phrases list. We additionally create a listing of lessons for our tags.

Now we will lemmatize every phrase and discard the reproduction phrases from the list. Lemmatizing is the method of changing a phrase into its lemma shape and then developing a pickle file to keep the Python objects which we will use whilst predicting. For example:

1. # lemmatize, lower each word and removeduplicates
2. words=
[lemmatizer.lemmatize(w.lower()) for w in words if w not in ignore_words]
3. words =sorted(list(set(words)))

3. Create coaching and trying out data
Now, we will create the coaching statistics in which we will grant the enter and the output. Our enter will be the sample and output will be the classification our enter sample belongs to. But the pc doesn't apprehend textual content so we will convert textual content into numbers.

4. Build the model

By using the coaching data that is ready, we will build a deep neural network that has 3 layers. We use the Keras sequential API for this. After training the model for 200 epochs, we achieved 100% accuracy on our model.

5. Predict the response (Graphical User Interface)

Now to predict the sentences and get a response from the user to a new file is created. We will load the trained model and then use a graphical user interface that will predict the response from the bot.

The model will only tell us the class it belongs to, so we will implement some functions which will identify the class and then retrieve us a random response from the list of responses.

CONCLUSION

It is typically not possible to induce all the info on one interface while not the complications of prying multiple forms and windows. The chatbot aims to take away this problem by providing a common and easy interface to resolve queries of the users. The cause of a chatbot is to simulate a human conversation

Its design integrates a language model and procedure algorithmic program to emulate info on-line communication between a human and a laptop exploitation language. The chatbot provides quick and economical rummage around for answers to the queries and gets the relevant links to their question.

A background analysis took place, including an associate summary of the voice communication procedure and tries to find out the relevant keywords associated with that question to offer the correct link. The information storage includes info regarding queries, answers, keywords, and logs.

FUTURE ENHANCEMENTS

Instead of a simple bot, different algorithms can be implemented. We can encompass voice-based queries and AIML primarily based bots. The customers will have to provide voice enter and the gadget will supply the textual content output. Also, after profitable execution of chatbot, we can enforce it in different domains like medical, forensic, sports, etc. It will be really useful in all the fields besides spending lots of time, we are having access to the applicable records and that too except any sorting.

REFERENCES

[1] Yuhua Li, David McLean, Zuhair A. Bandar, James D. O'Shea, KeeleyCrockett, "Sentence Similarity Based on Semantic Nets and Corpus Statistics", IEEE Transactions on Knowledge and Data Engineering, Volume 18 -No. 8, August 2006.

[2] Emanuela Haller, Traian Rebedea, "Designing a Chat-bot that Simulates a Historical Figure", IEEE Conference Publications, July 2013.

[3] Pratik Slave, Vishruta Patil, Vyankatesh Gaikwad, Girish Wadhwa, "College Enquiry Chat Bot", International Journal on Recent and Innovation Trends in Computing and Communication, Volume 5, Issue 3, March 2015.

[4] "AIML Based Voice Enabled Artificial Intelligent Chatterbot", International Journal of u-and e-Service, Science and Technology Volume 8 -No. 2, 2015.

[5] Amey Tiwari, Rahul Talekar, Prof. S. M. Patil, "College Information Chatbot System", International Journal of Engineering Research and General Science, Volume 2, Issue 2, April 2017.

[6] Rachit Kulkarni, Ankit Methwani, Nakul Pawar, Charmi Valecha, Pooja Shetty, "College Chat-bot", International Journal of Advanced Research in Computer Engineering & Technology, Volume 6, Issue 4, April 2017.

[7] https://www.ijrcst.org/DOC/2_irp620.pdf