Name :- Kaustubh Shrikant Kabra

Class:- TE Computer

ERP:-38

Subject :-LP2(AI) (Chatbot)

Code:-

```
import io
import random
import string
import warnings
import numpy as np
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import cosine_similarity
import warnings
warnings.filterwarnings('ignore')
import nltk
from nltk.stem import WordNetLemmatizer
# nltk.download('popular', quiet=True)
# nltk.download('punkt')
# nltk.download('wordnet')
with open('chatbot.txt','r', encoding='utf8', errors ='ignore') as fin:
  raw = fin.read().lower()
#Tokenisation
sent_tokens = nltk.sent_tokenize(raw)
word_tokens = nltk.word_tokenize(raw)
# Preprocessing
lemmer = WordNetLemmatizer()
def LemTokens(tokens):
  return [lemmer.lemmatize(token) for token in tokens]
remove_punct_dict = dict((ord(punct), None) for punct in string.punctuation)
def LemNormalize(text):
  return LemTokens(nltk.word_tokenize(text.lower().translate(remove_punct_dict)))
# Keyword Matching
GREETING INPUTS = ("hello", "hi", "greetings", "sup", "what's up", "hey", "Helo")
GREETING_RESPONSES = ["hi", "hey", "hi there", "hello", "I am glad! You are talking to me"]
def greeting(sentence):
  for word in sentence.split():
    if word.lower() in GREETING_INPUTS:
       return random.choice(GREETING_RESPONSES)
def response(user_response):
  robo response="
```

```
sent_tokens.append(user_response)
  TfidfVec = TfidfVectorizer(tokenizer=LemNormalize, stop_words='english')
  tfidf = TfidfVec.fit_transform(sent_tokens)
  vals = cosine_similarity(tfidf[-1], tfidf)
  idx=vals.argsort()[0][-2]
  flat = vals.flatten()
  flat.sort()
  req_tfidf = flat[-2]
  if(req_tfidf==0):
     robo_response=robo_response+"I am sorry! I don't understand you"
     return robo_response
  else:
     robo response = robo response+sent tokens[idx]
     return robo_response
flag=True
print("ROBO: My name is Robo. I will answer your queries about Investments. If you want to exit, type Bye!")
while(flag==True):
  user_response = input()
  user_response=user_response.lower()
  if(user_response!='bye'):
     if(user_response=='thanks' or user_response=='thank you'):
       flag=False
       print("ROBO: You are welcome..")
     else:
       if(greeting(user response)!=None):
          print("ROBO: "+greeting(user_response))
       else:
          print("ROBO: ",end="")
          res = response(user_response)
          nlines = res.count('\n')
          if nlines > 0:
            res = res.split("\n",1)[1]
          print(res)
          sent_tokens.remove(user_response)
  else:
     flag=False
     print("ROBO: Bye! take care..")
```

Output:-

