

Diet Chatbot: Dieto!

Authors: Ami Patel and Jay Patel

Class: CS 4395

Instructor: Dr. Karen Mazidi

Date: December 1, 2019

Introduction:

The topic chosen for the chatbot is diets and different types of diets available. The goal of the project is to design a simple chatbot using Python and NLP techniques, and chatbot does limited conversation about diets and its types using a knowledge base scraped from the web and the knowledge it learnt from the user interaction.

The NLP techniques used for this chatbot are vector space modeling, PSG parsing, information extraction techniques including Tf-idf metrics, and conversion of raw data was to the numeric data using TfidfVectorizer.

How Dieto works:

As soon as the user responds, the bot looks for a greeting response initially. If the response is found to be one of the greeting messages, the bot starts the conversation by asking for the name of the user and if the user likes talking about healthy diet plans. The name is recorded for future conversations, and the likes or dislikes are also recorded for future references. Once the conversation is initiated, the bot waits for the user to ask about any questions related to diet.

Once the bot gets a response other than greetings, the response is preprocessed and recorded in the sentence tokens. These sentences are a list of all the sentences found in the knowledge base and the user response appended to the list. After being added in the list, the response is parsed using PSG parsing to see if the response was a simple chat or a question. A package called 'textblob' was used to extract the POS tags from the user's response. The patterns of the user response are compared to the regular patterns of any simple comment in order to classify whether the user's response is a simple comment about the bot or user.

In case of simple chats, the bot responds accordingly in a modest manner, depending on the comment made by the user. However, if the user's response was not a simple comment, we further convert the raw data to the numeric vectors of sentences using tfidfVectorizer from sklearn. After the vectorizer is done fitting and transforming the sentences to numeric vectors, we create a vector space model by finding cosine similarity among the user's response and the transformed vectors.

The results from the vector space model are further processed to find the most relevant response for the user. This is done by checking the second last sentence which was found similar in the corpus, based on the vector space model. We choose the second last sentence because the user response was the last sentence and so, the most relevant sentence should be the very next sentence after the sentences are sorted.

```

TfidfVec = TfidfVectorizer()
tfidf = TfidfVec.fit_transform(sent_tokens)
similar_vector_vals = cosine_similarity(tfidf[-1], tfidf)

# the second last sentence will have the highest cosine similarity
sentence_number=similar_vector_vals.argsort()[0][-2]

# Flatten to check if the cosine similarity is 0 or not
flat = similar_vector_vals.flatten()
flat.sort()
actual_tfidf = flat[-2]

```

Whether or not a relevant response was found for the user is checked by flattening the vectors and finding the Tf-idf value of the second last sentence. If the Tf-idf value is found to be 0, that means no information in the knowledge base matches the user's question. In this case, the bot returns with appropriate sorry response. However, if the Tf-idf value is not found to be 0, then the bot responds with the most appropriate sentence found in the knowledge base.

Sample Dialogue:

Here's a sample dialogue with Dieto. The user is an overweight person, whose name is 'Nick'. He was to know more about the diet plans that he could follow based on his BMI index. The bot dialogues are as follows:

```

DIETO: My name is DIETO. I will answer your queries about diets. If you want to exit, type Bye!
YOU: Hi
DIETO: Hey, let's start with your name. Please introduce yourself
YOU: I am Nick. I am overweight
DIETO: Nice to meet you, nick! Do you want to know more about healthy diet plans?
YOU: no
DIETO: I will try my best to make it more interesting for you! Shoot me your questions..
YOU: what is body mass index?
DIETO: the body mass index bmi is a ratio of weight in kilograms to height in meters squared.
YOU: what is normal bmi?
DIETO: someone with a bmi between 25 and 29.9 kg m2 is overweight, and someone with a bmi of 30 kg m2 or more is obese.
YOU: what diet is for person with bmi of 30 or higher?
DIETO: the very low calorie diet is prescribed by a physician on a case to case basis to promote rapid weight loss in patients with body mass index bmi of 30 or higher.
YOU: what is very low calorie diet?
DIETO: a 500 calorie diet is an extreme form of a very low calorie diet vlcd .
YOU: thank you
DIETO: You are welcome.
YOU: bye
DIETO: Bye! take care..

```

Knowledge base:

The knowledge base was created using the web scraping methods and a crawler that crawled through different websites to gather the relevant text about diets. From the text scrapped, manually 10 important words were chosen using Tf-idf metrics. These 10 important words were:

['calorie', 'weight', 'food', 'eating', 'loss', 'fasting', 'disease', 'body', 'protein', 'meal']

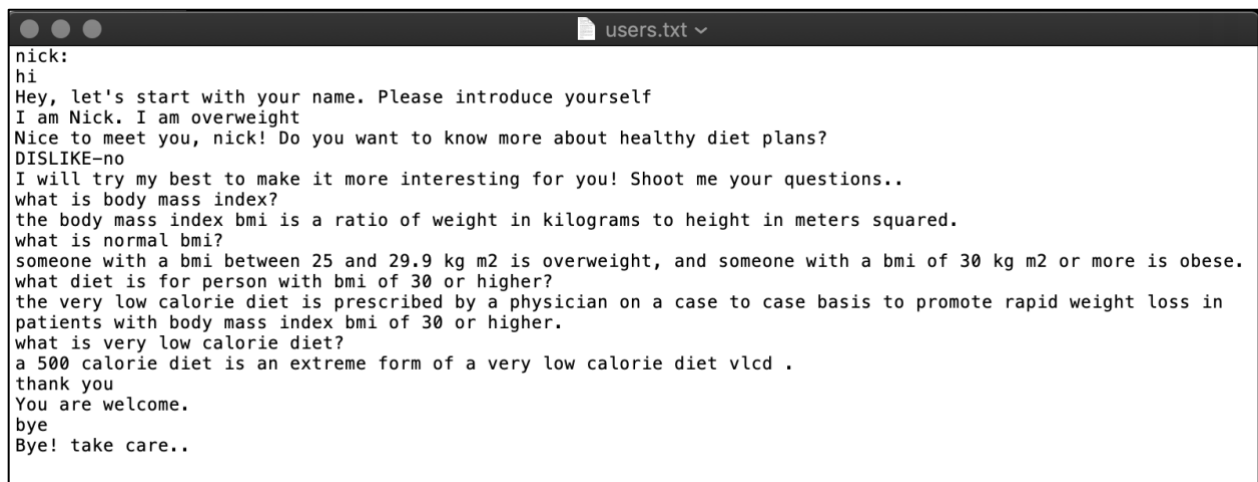
Based on these words, sentences from extracted text were added to the knowledge base, 'kb.txt'. Once this knowledge base was created by the crawler, it was thoroughly studied and processed to remove stop words or any inappropriate words. Here's a sneak peek of the final knowledge base used for Dieto:

a 500 calorie diet is an extreme form of a very low calorie diet vlcd .
it can help you eat fewer calories and optimize numerous hormones related to fat loss. read more evidence based health benefits of intermittent fasting written by kris gunnars, basically this is a detailed review of the health benefits of intermittent fasting.
the very low calorie diet is a doctor supervised diet that allows 600 800 calories per day.
mostly very low calorie diets consist of high protein shake containing vitamin and mineral supplements to meet recommended daily requirements.
the very low calorie diet is prescribed by a physician on a case to case basis to promote rapid weight loss in patients with body mass index bmi of 30 or higher.
the very low calorie diet is a very effective treatment for rapid weight loss and allows patients to lose three to five pounds per week.
the minor side effects of the very low calorie diet include fatigue, nausea, constipation or diarrhea.
the serious side effect of the very low calorie diet is gallstone formation caused by rapid weight loss.
low calorie diets allowing 800 1000 calories a day have shown to be equally effective in helping patients lose weight.
the very low calorie diet can cause large amount of weight loss and improvement in health in obese patients.
the very low calorie diet always requires medical supervision as it can be misused by the patient and lead to health risks.
you must seek medical help before trying any very low calorie diet because dangerous side effects can occur if the diet is abused.
a protein pack feeds your body through the tube, by dispensing drops of a liquid mix of nutrients but no carbs totaling about 800 calories a day.
fish, chicken and meat are excellent protein sources, and soy protein can sometimes be helpful for people who are willing to try it.
experts recommend that healthy adult women consume about 2,000 calories and man consume 2,500 calories daily.
when you cut calories that drastically, the body responds by going into a state of ketosis, in which it starts to burn stored fat rather than consuming sugar to keep normal body functions going.
a severely calorie restricted diet can start to eat away at muscle.
but the fact that women are dropping pounds by starving themselves is no surprise whenever you cut calories and deprive your body of food, you ll lose weight.
it won t take long before some unscrupulous doctors start offering the tubes with less than adequate supervision, and patients start pushing their calorie intake ever lower, stressing the body s metabolic system with too many dips into the dangerous

User models:

Whenever Dieto interacts with any user, all the interactions are recorded in a text file, 'users.txt' for all the users. This file has all the user's as well as bot's dialogues recorded. The dialogues are separated for different users using user's name. If the bot starts for the first time, this file is created by the script. Otherwise, every time a user starts a conversation, the bot checks if it has interacted with the same user before. If so, the bot tries to extract the data from the file, including whether the user is interested in talking about diet plans or not. These likes and dislikes of the users, along with the dialogues, gets updated in users.txt every time any user returns.

An example of the stored user model is shown below:



```
nick:
hi
Hey, let's start with your name. Please introduce yourself
I am Nick. I am overweight
Nice to meet you, nick! Do you want to know more about healthy diet plans?
DISLIKE-no
I will try my best to make it more interesting for you! Shoot me your questions..
what is body mass index?
the body mass index bmi is a ratio of weight in kilograms to height in meters squared.
what is normal bmi?
someone with a bmi between 25 and 29.9 kg m2 is overweight, and someone with a bmi of 30 kg m2 or more is obese.
what diet is for person with bmi of 30 or higher?
the very low calorie diet is prescribed by a physician on a case to case basis to promote rapid weight loss in
patients with body mass index bmi of 30 or higher.
what is very low calorie diet?
a 500 calorie diet is an extreme form of a very low calorie diet vlcd .
thank you
You are welcome.
bye
Bye! take care..
```

Note that the bot stores the data collected from ALL the users it interacts, using their names as the key to read, write or update in 'users.txt'.

Evaluating the Chatbot:

Overall, the chatbot seems to complete the task it is designed to do. First, it greets the user, and encourage the user to talk more about diets and answer the general questions. The chatbot gave satisfied and informative answer when user ask about some specific topic such as weight loss, various types of diets and diabetes. As shown in the screenshot below, user “Jay” and “Jason” tried to get general knowledge of healthy diets and the chatbot seems to provide appropriate answer and information. Moreover, we have shown a special case with user “Nick”, where Nick is getting advice from chatbot to lose weight. In the above cases, the vector space model used is able to find the most relevant responses from user response and the chatbot gives appropriate output. Also, the chatbot keep track of user data, so chatbot greet user in specific way if it is a returning user.

As we communicated more with the chatbot, we found out that user need to ask more informative question so that chatbot answer it appropriately. Short question like definitions create difficulties in finding cosine similarities in the sentences from knowledge base and hence, the chatbot ends up giving inappropriate answers. For example, in screenshot of bad chat, user does not get appropriate answers. Moreover, there is limited information the knowledge base so the chatbot cannot answer all the definitions asked.

Screenshot of good chats:

Chat 1:

```
DIETO: My name is DIETO. I will answer your queries about diets. If you want to exit, type Bye!
YOU: hi
DIETO: Hey, let's start with your name. Please introduce yourself
YOU: my name is Jay
DIETO: Nice to meet you, jay! Do you want to know more about healthy diet plans?
YOU: yes
DIETO: Perfect! Let's get started. You can ask me any questions you might have!
YOU: what are different types of diet
DIETO: some different types of diet are low fat diet, low glycemic index diet, low protein diet, low sodium diet, low sulfur diet, some common macrobiotic ingredients macrobiotic diet: a diet in which processed food is avoided.
YOU: okay
DIETO: Please feel free to ask me any further questions you might have! If you want to exit, type Bye!
YOU: what diets are popular
DIETO: high protein, low carbohydrate diets, such as the atkins diet, are popular and controversial.
YOU: what are sources of protein
DIETO: fish, chicken and meat are excellent protein sources, and soy protein can sometimes be helpful for people who are willing to try it.
YOU: thank you
DIETO: You are welcome.
YOU: bye
DIETO: Bye! take care..
```

Chat 2:

DIETO: My name is DIETO. I will answer your queries about diets. If you want to exit, type Bye!
YOU: Hello
DIETO: Hey, let's start with your name. Please introduce yourself
YOU: Jason
DIETO: Welcome back, jason! Glad to know you are interested in talking about healthy foods! Shoot me your questions..
YOU: I am diabetic what I should do?
DIETO: for diabetic person, avoid all artificial sweeteners found in diet drinks and some diabetic foods as these are toxic to the liver and can cause hypoglycaemia and fatigue.
YOU: what are some good eating habits?
DIETO: building simple habits for healthy weight loss lasting, healthy weight loss demands that you transform your eating and exercise habits.
YOU: how to increase body weight?
DIETO: the reason for weight increase is not what or how much you eat that causes weight gain but rather when and in what combinations the food is eaten.
YOU: Which diet help in building muscles?
DIETO: high protein diet: a diet in which high quantities of protein are consumed with the intention of building muscle.
YOU: okay
DIETO: Please feel free to ask me any further questions you might have! If you want to exit, type Bye!
YOU: bye
DIETO: Bye! take care..

Chat 3:

DIETO: My name is DIETO. I will answer your queries about diets. If you want to exit, type Bye!
YOU: Hi
DIETO: Hey, let's start with your name. Please introduce yourself
YOU: I am Nick, and I am overweight.
DIETO: Welcome back, nick! Glad to know you are interested in talking about healthy foods! Shoot me your questions..
YOU: What is body mass index?
DIETO: the body mass index bmi is a ratio of weight in kilograms to height in meters squared.
YOU: What is normal bmi?
DIETO: someone with a bmi between 25 and 29.9 kg m² is overweight, and someone with a bmi of 30 kg m² or more is obese.
YOU: What diet is for person with bmi of 30 or higher?
DIETO: the very low calorie diet is prescribed by a physician on a case to case basis to promote rapid weight loss in patients with body mass index bmi of 30 or higher.
YOU: What is very low calorie diet?
DIETO: a 500 calorie diet is an extreme form of a very low calorie diet vlcd .
YOU: okay
DIETO: Please feel free to ask me any further questions you might have! If you want to exit, type Bye!
YOU: bye
DIETO: Bye! take care..

Screenshot of a bad chat:

DIETO: My name is DIETO. I will answer your queries about diets. If you want to exit, type Bye!

YOU: hi

DIETO: Hey, let's start with your name. Please introduce yourself

YOU: Joanna

DIETO: Nice to meet you, joanna! Do you want to know more about healthy diet plans?

YOU: what diet has more carbs?

YOU: What diet has more carbs

DIETO: the zone diet: the zone diet is another breed of low carb diet that has very specific recommendations about the amount of carbs and protein that you can eat in a day.

YOU: what diet has less gluten?

DIETO: if you eat less than five pounds of grapes, you get even less protein.

YOU: how to increase weight?

DIETO: the reason for weight increase is not what or how much you eat that causes weight gain but rather when and in what combinations the food is eaten.

YOU: bye

DIETO: Bye! take care..