Project 2 Milestone 2

Megan Holford

DSC 680

Summer 2022

Building a Chat Bot for Diabetes Information

**Business Problem**

When searching for information online regarding healthcare topics, the amount of information can be overwhelming to dig through. One way that it could be possible to help streamline the experience of finding information for users could be by using a chat bot. Automated chat bots can be helpful for both businesses and their users. For a business, a chat bot can help with customer engagement, labor costs and more. On the user side, chat bots can be helpful if they just need a quick answer, need help navigating, or to help route them to where they need to be for assistance. For businesses, there are many pros to using a chat bot that help the online experience (Feng, 2022). For this project, I wanted to build NurseBot, a user prompted chat bot that can help answer questions about diabetes.

**Background/History**

For my second project, I wanted to see if I can use natural language processing tools to create a chat bot. This project is more of a proof of concept as I wanted to get experience building a chat bot and finding out what challenges there would be along the way. I have little experience with text analysis, so this was an interesting challenge.

**Data and Methods**

For the information used to base the chat bot on, I used information from the Mayo Clinic (Mayo Foundation, 2020), as I wanted to use a reputable medical source to collect the information to put through the model for the chat bot. For the actual text model, I used NLTK and SKLearn to use the information from the webpage to create a chat bot that users can interact with. Using NLTK, we can use the raw text from the webpage to then break down the text to their sentence parts and tokenize the sentences. Tokenization is important for this project as tokenization is basically taking a sentence, paragraph, or an entire text document and breaking it down into smaller units, such as individual words. Once the text is broken down into the smaller units, those are the tokens that can be use d for the model.

Diagram

Description automatically generated

This flowchart from NLTK.org is helpful in showing how text information is taken through the process for extraction (Bird et al., 2009). Seeing the process in a visual format helps to understand why it is important to use the natural language processing tools to help build those relationships and pieces that the model can use.

Once we have been able to tokenize the text, we can then use text vectorization to put into the model to build the chat bot. Using vectorization, we can convert the text to numerical values to then find similar responses in the original text when it is user prompted. At this point, I also built in some error handling in the code for if the bot cannot generate a response due to not recognizing the input. If the user inputs a response that is not intelligible, the bot will just give a prompt that it cannot understand. Having error handling in a chat bot is important since humans are able to mistype. With it being a chat bot, there is also not a lot of room for nuance in the responses, so trying to help streamline the experience is important.

**Conclusion**

In the end, I was able to make a basic chat bot to help answer questions about diabetes.

Text, letter

Description automatically generated

This is an image of a successful chat attempt and overall, the bot was able to answer the questions it was prompted. While the bot was able to share information to the user, there is definitely room for improvement with the bot.

**Challenges**

One of the biggest challenges I had going into this project was my lack of experience. I had little experience using NLTK prior to this and so I was not sure what successes and limitations would arise. One thing that did provide a challenge was that I do feel that more information and text could benefit the chat bot. The information from the Mayo Clinic, while useful and correct, was lacking in the amount of text that could be used to help build those bridges in the text and the tokens created from the raw information. I do think that having more text to run through the model could help and improve the chat bot experience.

**Future Uses/Additional Applications**

While I do think that there is much to be tested and improved on with the chat bot, there is definitely a good start here. With more time and information that can be used for the text analysis, I do think that a more natural chat bot could be achieved. This can be helpful in this use case of a patient trying to get more information about a medical condition, but the actual uses for a chat bot are much more far reaching than this and we do see them in practice all the time.

**Ethical Assessment**

When it comes to using a chatbot, one of the biggest ethical considerations is that your bot is able to provide accurate information in an understandable way to the user. If your chatbot is answering in gibberish or not actually answering the questions at all, then that is not a worthwhile use of time for the business or the user interacting with the bot. For this reason, I do not think it would be ethical to use this chat bot as is right now. While it can answer basic questions for users, there are a lot more nuances in human communication, so I think more work is needed to have a truly functioning chat bot.

It is also important to be transparent when you are using a chat bot, so the consumer knows they are interacting with the chat interface and not a real human. Trying to pass off a chatbot to users as a person is not ethical and a lot of times, while users can tell if the system is automated, making sure you are transparent helps to build that trust. For this reason, I made sure to make it apparent that this is a bot and not a real person.

**Resources**

Bird, S., Klein, E., & Loper, E. (2009). 7. Extracting Information from Text. In *Natural language processing with python*. essay, O'reilly.

Feng, D. (2022, July 11). *Chatbot 101: Why conversational bots are the future of Ecommerce*. The BigCommerce Blog. Retrieved July 17, 2022, from https://www.bigcommerce.com/blog/chatbots/#chatbots-now-and-in-our-future

Mayo Foundation for Medical Education and Research. (2020, October 30). *Diabetes*. Mayo Clinic. Retrieved July 17, 2022, from https://www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444

**10 Questions**

1. What were some of the biggest challenges for this project?

One of my biggest challenges for this project was my lack of experience with this type of coding and tools. I have not had much experience with creating a chat bot, so trying to use some of these packages for that was a challenge.

1. How can this be helpful to users?

This type of medical chat bot can be helpful to users who may be looking into their own health related issues. While we would always still want to refer them to their primary doctor for an actual diagnosis, having information they can take to their doctor can be helpful to begin that conversation.

1. How did the user experience influence this project?

I wanted to make the user experience as smooth as possible and took that into account with the coding. One thing was adding error handling if the bot did not understand what a user typed, it would just say it could not understand instead of giving an error.

1. Would this be able to go live now?

While the bot can respond, I definitely think it needs more before it could go live with actual users.

1. Can this type of chat bot be used for other information?

Chat bots are used in many fields already, so it is definitely applicable to other fields and businesses.

1. What are important things to consider in regard to ethics with a chat bot?

Making sure that your bot provides accurate information is very important as you do not want users interacting to get wrong information.

1. What could help improve the performance of the chat bot?

While I do not have much experience, I think having more data from other sources could help to train the model for the bot and the responses it generates.

1. What are some other applications for this chat bot?

While this project focused on using a chat bot in a medical setting, it can be used in others. One setting could be for helping users to navigate a site and where they need to be. Another could be a pre-chat bot that can get information from users before forwarding them to the correct team to assist and helping to streamline the customer service process.

1. This chat bot was just for diabetes information, but could a chat bot be made to cover more medical conditions to help answer user questions?

This just used information from a webpage on diabetes so if more information was including from other pages on other medical diagnoses, then it could be used for other conditions.