

David Nguyen

dxn180015

Dr. Karen Mazidi

CS 4395.001

HW - Chatbot

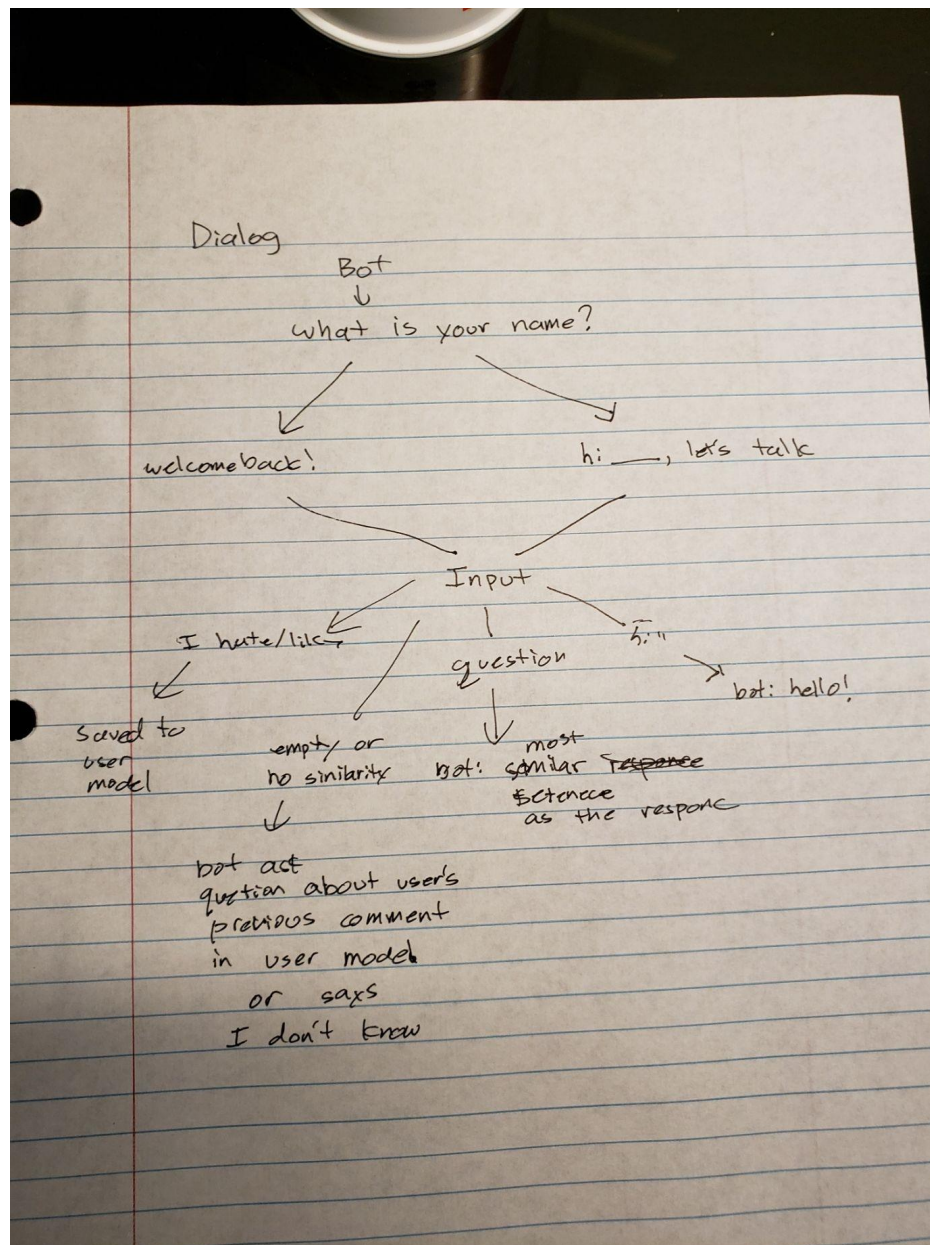
System Description:

The chatbot made for this homework was made in python from scratch and used the things we learned in class about NLTK . The program's knowledge base utilizes my web crawler's output where it puts all the sentences related to the top 10 terms for the topic: The Nintendo switch, into a text file. A dictionary knowledge base is created after reading this in where each key is one of the top 10 terms for my topic (`top_10_terms = ['nintendo','switch','game','console','system','controller','hardware','gaming','developer','mario']`), and each value is a list that will contain the related sentences.

The user can ask the bot anything about the Nintendo Switch. But first the bot asks the user his name and then adds them to a `user_models` list where it contains a list of users' likes and dislikes sentences to remember. The user can write 'bye' or 'exit' to leave the chat and then the program will write the user models to a pickle file. The bot also checks if the user asks a question and responds to it. When a user makes an opinionated statement like 'I hate/like' something, that user sentence is written to the user model. The bot also recognizes when the user says hi and says hi back.

Some specific NLP techniques I used were TfidfVectorizer and Cosine Similarity which was used to find the most similar sentence when compared to the user's input. And the most similar sentence was used as the chatbot's response to the user. I also used tf-idf to calculate the top 10 terms of my topic and used techniques like NLTK's tokenizer and text processing to tokenize text into sentences or words, and remove bad data / clean data like punctuation, stopwords. The knowledge base text was fitted and transformed with TfidfVectorizer, and the tokenizer went through the sentences to lemmatize them and clean / remove punctuation. The cosine similarity of all the sentences are put into a list, flattened, and sorted to give the most similar sentence as the response. If the cosine similarity was 0, then an empty string was given which lets the bot say it is not familiar with the topic or asks the user about previous opinionated sentences the user said based on the user model which contains sentences where the user hates or likes something.

Diagram of your dialog tree or logic:



User Model:

```
david : ["I hate Nintendo", "I like mario"]
jess  : ["I like the switch"]
```

An snippet of my printed knowledge base is here below (entire screenshot too big):

```
plastic
['Even the plastic dock bundled in the package seems very tiny in comparison to normal hardware.', 'This kickstand is made of plastic, feels flimsy, and in use you tend to fear it may break.', 'Another issue is the flimsiness of the cover for the game cards; it's literally thin tabs of plastic securing this to your console.', 'The dock also needs a mention here: it's a slab of plastic with two external USB slots, the lone out and ports complementing internal USB, USB-C, and HDMI connections.']
```

```
trailer
['In the trailer, a gamer slides what looks to be a 3DS-style cartridge into the tablet, meaning games are likely to be distributed both digitally and physically.', 'The trailer shows off lots of multiplayer gaming, either with multiple controllers connected to one system, or many Switch consoles connecting together wirelessly.', 'In the trailer you can see third-party games like The Elder Scrolls V: Skyrim and NBA 2K alongside what look to be a new Mario and Mario Kart games and Splatoon .', 'View All Comments 'Switch' is Nintendo's next game console av gaming mobile nintendo nintendoswitch nx switch tablet trailer video o zelda About Engadget Masthead About our Ads Advertise Licensing FAQ RSS Feed Sections Reviews Gear Gaming Entertainment Tomorrow The Buyers Guide Video Podcasts Deals Contribute Comment Guidelines Support International 繁體中文 Subscribe to our two newsletters: Engadget Deals - A weekly roundup of our favorite tech deals The Morning After - A daily dose of the news you need Please enter a valid email address Please select a newsletter Subscribe Follow Us © 2023 Yahoo.', 'Archive eShop Updates Amiibo Feature Interview News Previews Site Updates Trailer Analysis Contact Search for: #site-navigation #masthead Nintendo Switch Review March 16, 2017 March 16, 2017 Paul Murphy .entry-meta .entry-header It's time to switch.', 'Click To Unmute Street Fighter 6 Zangief vs Marisa Developer Gameplay Resident Evil 4 Remake Everything To Know Diablo 4 Ultimate Edition Breakdown Trailer Unlock Mad Chainsaw Mode & Secrets Weapons In Resident Evil 4 Demo Dead Island 2 Isn't What We Were Expecting... PS5 Pro Reportedly In Development | GameSpot News Sea of Thieves Season Nine Deep Dive Hearthstone: Festival of Legends Cinematic Trailer The Callisto Protocol - Contagion Bundle Trailer Call of Duty: Modern Warfare II & Warzone 2.0 Season 02 Reloaded Trailer Total War: WARHAMMER III Forge of the Chaos Dwarf's Trailer Hyper Light Breaker Gameplay Reveal Trailer Share Link Embed Size: 640 x 360 480 x 270 Start at: End at: Autoplay Loop Want us to remember this setting for all your devices?"]
```

Sample Dialog Interactions:

```
Chatbot started....
Type 'exit' or 'bye' to leave the conversation
Bot: Hello there! I can talk about the Nintendo Switch!
Bot: What is your name?
david
Welcome back david ! I'm ready to chat with you again!
hi
Bot: bonjour
```

```
Bot: why did you say ' i like mario ' again?
is there wii sports
Bot: "When asked specifically about 1, 2, Switch and how it seems like the ideal evolution of the wii sports-degree, arguing 1, 2, Switch is a more comprehensive game.
What is the cost
Bot: "You make the comment about wii Sports, and certainly, the analogy that it's a great utilization of the m that standpoint, wii Play was a standalone piece of software both with a controller at the time and without,
```

Evaluation:

The chatbot is not very smart and gives very rigid answers that may or may not be related to the question. This is a weakness of the bot since the knowledge base needs to be bigger and more comprehensive, and cleaner. If the user gives a response that is in the knowledge base, then the bot can give a good answer due to tfidf. And another positive is that the chat bot can respond to other situations like no input or say hi back.