

Chatbot Project Report

System Description:

Out of the two kinds of chatbots we could make for this project (i.e., a task-oriented dialogue agent vs. a more conversational chatbot), I decided to go with the former. More specifically, my chatbot is designed specifically to answer questions about the movie, *The Nightmare Before Christmas* (1993). The knowledge base, which is called `knowledge_base.txt`, contains possible answers the chatbot could give the user (see Appendix A). All of the statements in the knowledge base are ones that I manually put in.

My chatbot works best when the user asks questions to the chatbot, such as “Who are the main characters in the movie?”, “How much money did the movie make?”, “When was the film released?”, etc. Examples of dialogue interactions can be seen later on in the report.

In regards to the NLP techniques I used for this project, I used some standard NLP processing (e.g., tokenization, lemmatization, removing stopwords and punctuation, etc.). These standard NLP processing techniques were used on the user input (i.e., whatever the user typed in) in order to (1.) get the key/important information about asking and (2.) matching a response from the knowledge base that best fits what the user is asking. The portion of this code that uses these techniques can be seen in lines 49-73.

Aside from those NLP processing techniques, I also made use of the Tf-idf vectorizer and the cosine similarity metric. Utilizing these features was really the “meat” of creating my chatbot’s response. The Tf-idf vectorizer was used to get the vocabulary and the idf of the corpus. In this case the corpus is `sentence_tokens`, which contains the input from the user. The idf, or inverse document frequency, gives us a measurement of importance or relevance of a term in a given corpus. The cosine similarity measures how close two documents are. In this case, it’d be the user’s input compared to the knowledge base. Finally, I flattened and sorted the cosine values and got the highest one. If this value, which I called `highest_tfidf_value` was *not* 0, then that means there was an appropriate response found within the knowledge base. The chatbot would then output that response. However, if this value was 0, then that means there was no response the chatbot thought was appropriate in the knowledge base. If that were the

case, the chatbot would just respond with a “Sorry, try again” message. All of this can be found in the `chatbot_response()` function, which starts on line 88.

And while not an NLP technique, I did add another feature to my chatbot that responds with fun facts if the user requests one. If the user types in “fun fact” or requests one (e.g., “Can you give me a fun fact?”, “Please give me a fun fact”, “I’d like to know a fun fact”, etc.), then the chatbot will respond with a fact about the film. These facts are also present in the knowledge base—at the very end.

The conception/implementation of the user model will be discussed in a later part of the report (see Appendix B).

Dialogue Tree Diagram/Logic:



Sample Dialogue Interactions:

```
Hi, my name is Zero and I'm a chatbot that has knowledge about the movie 'The Nightmare Before Christmas'!  
Ask me a question about the movie and I can answer it for you! If you want a fun fact about the movie, just type in 'fun fact'!  
To end conversation with me, just type 'bye' or 'goodbye'.  
hi  
Zero: *nods*  
what are the main characters of the movie?  
Zero: the main characters of the movie are jack and sally.  
tell me who the villain is  
Zero: oogie boogie is the villain.  
what's jack's dog's name?  
Zero: jack skellington has a ghostly dog named zero (that's my name too!)  
what's jack's skellington's occupation?  
Zero: jack's occupation is to work as the town's leader of their annual halloween celebration.  
how much money did the movie make?  
Zero: the amount of money the movie made was $91.5 million.  
what studio released the film?  
Zero: the film was released on october 13, 1993; the film was initially released through touchstone pictures because walt disney studios thought the film would be "too scary" for kids.  
i forgot sally's father's name... what is it?  
Zero: sally's "father" is dr. finkelstein.  
who voices the mayor?  
Zero: glenn shadix voices the mayor.  
bye  
Thanks for talking to me. Hope to talk to you again soon!  
  
Process finished with exit code 0  
|
```



```
Hi, my name is Zero and I'm a chatbot that has knowledge about the movie 'The Nightmare Before Christmas'!  
Ask me a question about the movie and I can answer it for you! If you want a fun fact about the movie, just type in 'fun fact'!  
To end conversation with me, just type 'bye' or 'goodbye'.  
hi  
Zero: hola (that's the only thing I can say in Spanish by the way)  
what's the runtime of the film?  
Zero: the runtime of the film is 76 minutes.  
what does jack call santa claus in the movie?  
Zero: jack does call santa claus "sandy claws".  
i know there are three kids in the film, but i forgot their names  
Zero: the names of the kids are lock, shock, and barrel.  
which one is lock  
Zero: lock is the one with the devil costume.  
which one is shock  
Zero: shock is the one with the witch costume.  
which one is barrel  
Zero: barrel is the one with the skeleton costume.  
that's cool, can you give me a fun fact?  
Zero: Fun fact: 12 stop-motion frames were used to create 1 second of the film.  
i'd like to know another fun fact  
Zero: Fun fact: Jack Skellington makes a cameo in 'James in the Giant Peach' (1996) as a pirate skeleton.  
That's great, thanks for answering me  
You're welcome!  
  
Process finished with exit code 0  
|
```

```
Hi, my name is Zero and I'm a chatbot that has knowledge about the movie 'The Nightmare Before Christmas'!  
Ask me a question about the movie and I can answer it for you! If you want a fun fact about the movie, just type in 'fun fact'!  
To end conversation with me, just type 'bye' or 'goodbye'.  
hola  
Zero: hey there  
what's the name of jack's enemy?  
Zero: oogie boogie's enemy is jack.  
where does the mayor live?  
Zero: the mayor does live in halloween town.  
what's the mayor's job?  
Zero: the mayor's job is the mayor of halloween town, duh.  
who voices dr. finkelstein?  
Zero: william hickey voices dr. finkelstein.  
what's dr finkelstein's job?  
Zero: dr finkelstein's job is a mad scientist.  
who's the leader of halloween town?  
Zero: jack is the leader of halloween town.  
who wrote the songs in the film?  
Zero: danny elfman wrote songs for the film.  
who voices jack?  
Zero: chris sarandon voices jack skellington's regular voice, but danny elfman voices his singing voice.  
who directed the movie?  
Zero: even though many people think tim burton directed the movie, henry selick actually directed it!  
that's cool. can you give me a fun fact?  
Zero: Fun fact: Gregg Olsson, the film's set designer, built a quarter scale mock-up that served as a model for the set.  
thanks for that  
Zero: You're welcome!  
ok, bye!!  
Thanks for talking to me. Hope to talk to you again soon!
```

Evaluation & Analysis of Chatbot:

As stated earlier in this report, this chatbot was designed a task-oriented agent that answered questions about the movie, *The Nightmare Before Christmas*. Because of this, it is inherently less conversational than some other chatbots. I've created some responses for the chatbot that may seem conversational, but the intent of the chatbot isn't to indulge in conversation. As you can see from the sample dialogue interacts above, the chatbot works best as a question-and-answer device. You can ask for fun facts and give the chatbot your opinions about the movie, both of which should generate appropriate responses, but again, it's not necessarily made for conversation.

In regards to the responses the chatbot's able to produce when the user asks questions, I think for the most part it does a good job of answering them. As you can see from the sample dialogue interacts, the chatbot is good at answering simple and straightforward questions about the film.

Another strength of this particular chatbot is that I think it can answer a variety of questions about the film. As stated previously, the knowledge base for this chatbot was completely created by me, so I tried having as many facts about the movie as I could.

You can ask questions about how much money the movie made, when it was released, who voices certain characters, etc. You might also notice—if you look inside the knowledge base text file—that a lot of the entries in there are variations of the same fact. I did this because depending on how the user asks their question, the chatbot would be able to give it a different response. For example, if the user asks, *“how much money did the movie make?”*, the response they should get is *“the amount of money the movie made was \$91.5 million.”* However, if the user had the wording, *“how much money did the film make?”*, then the chatbot would respond using the word “film” instead of “movie”.

And as mentioned earlier in the report, the chatbot uses cosine similarity and tfidf vectorization in order to determine key words/phrases in the user’s input and responds with a statement from the knowledge base that includes some of those key words/phrases.

With that said, one of the downsides of the chatbot and its ability to produce responses is that sometimes your questions have to be phrased a certain way to get the optimal response. For example, when you type in the question, *“who are the supporting characters from the movie?”*, you’ll get the response, *“the main characters of the movie are jack and sally”*, which obviously does not answer the question. Because of this, you should probably phrase the question as, *“who is the supporting cast of the movie?”*, in which case you’ll get the response, *“the supporting cast of the movie is oogie boogie, dr. finkelstein, the mayor...”*

Appendix A

Knowledge Base

The [knowledge base](#) I've created consists of the possible responses Zero (the chatbot) can produce. Each entry in the knowledge base is one that I manually put in and answers some question about the movie, *The Nightmare Before Christmas*.

Samples:

The movie was released on October 13, 1993.

The film's box office was \$91.5 million.

The main characters are Jack and Sally.

Stan Webb was the main editor.

Danny Elfman is the composer of the movie.

Santa is the leader of Christmas Town.

Fun fact: The film required a total of 230 sets.

Fun fact: After 'The Nightmare Before Christmas', Selick directed 'James and the Giant Peach' (1996) and 'Coraline' (2009).

Depending on what the user asked, one of these responses (plus any of the other entries in the knowledge base) could be outputted by the chatbot.

Appendix B

User Models

The [user models](#) I've created are based on two things: the user's name and their likes/dislikes. Because this was designed to be a task-oriented chatbot, especially with knowledge about *The Nightmare Before Christmas*, I thought that the only thing that really seemed relevant to store were the user's name and their opinions about the movie.

The link provided in the previous paragraph shows you three user models that were created with the chatbot. Each user model is separated by the dotted line. As stated before, the only things that I thought were worth saving were the user's name and their thoughts about the movie.

Note: if you want to try this out for yourself, you have to phrase some things in a certain way. For example, you have to say "my name is" if you want the chatbot to save your name. Likewise, there are certain phrases that the chatbot will pick up (e.g., "i liked", "my favorite", "i thought", "i disliked", "i enjoyed", " i wasn't fan of", etc.) if you want your opinions about the movie to be saved. You can get a clearer idea of what phrases the chatbot will respond to starting at line 16 in the code.