

Library Bot Report & Evaluation

Overview

I made the chatbot from scratch while trying to use as little external API's and libraries as possible while incorporating different NLP techniques we've learned. Because of this, it's not as robust as a chatbot that would utilize something like the chat-GPT API, but it was a good way to cement what we've been learning in class.

A lot of the functionality of the bot comes from regex, named entity recognition (NER), part of speech tagging (POS), and machine learning for book recommendations.

NLP & ML Techniques

Regex

- Regex was used to break down and identify certain inputs from the user.
 - An example of this is when a user wants to rate a book a regex is used to find the rating score.

Named Entity Recognition (NER)

- NER was used to identify and extract objects from sentences.
 - An example of this is the `extract_title` function which finds and extracts the title of book from a given text.

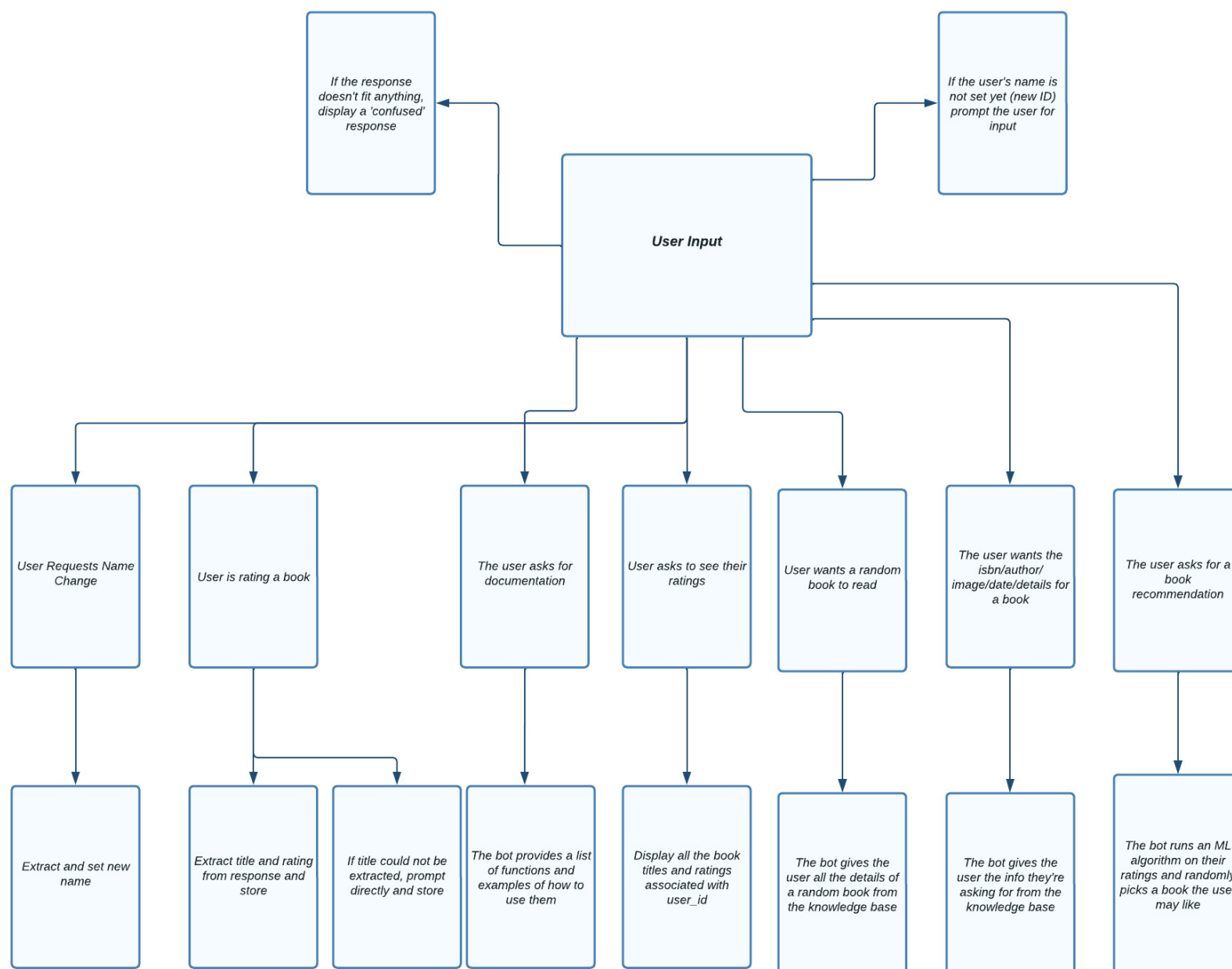
Part of Speech Tagging (POS Tagging)

- POS tagging was used similarly to NER to identify and extract objects from sentences. NER was better for certain things, and POS tagging was better for others.
 - An example of this is the chatbot's response to a name change, it finds and extracts the name so even if the user inputs "I want to change my name to John", it would extract the name as John.

Machine Learning (Random Forest Regression)

- Random forest regression was used to find recommendations of new books for the user based on their current ratings.

Dialog Tree Diagram



Dialog Interactions

- Dialog is set where you don't have to say things in a specific format, the bot looks for a range of keywords to derive what the user is wanting to do.

Terminal Name:

- The name displayed as the user when chatting is the user_id, I've set it to "John Y". This also determines the name of the user data file.

Set Name:

- If the user doesn't have a name set, the librarian bot will ask and set one.

John Y: hello

Librarian Bot: Hello!, I've noticed I don't know your name in my knowledge base. Could you provide your first name so I know how to address you?

John Y: John

Librarian Bot: Alright, i've set your name as John

John Y:

Name Change:

- Changes the user's name when prompted, will use POS tagging to determine the name in the sentence.

John Y: I'd like to change my name

Librarian Bot: Alright, let me know what you want to change your name to

John Y: Could you set my name as Matthew?

Librarian Bot: Alright, i've set your name as Matthew

John Y: Please change my username

Librarian Bot: Alright, let me know what you want to change your name to

John Y: John

Librarian Bot: Alright, i've set your name as John

Rating a Book:

- Checks if the text contains a book and stores the rating for that user.

John Y: I rate To Kill a Mockingbird by Harper Lee a 5/10.

Librarian Bot: Alright John, I've added to kill a mockingbird to your book ratings with a rating of 5

- If the user doesn't include a rating or a proper book title the chatbot will ask them to give the information to clarify.

John Y: rate book

Librarian Bot: I'm having trouble determining the title or rating, could you give me the title of the book and rating split by a comma?

John Y: the giver, 9

Librarian Bot: Alright John, I've added the book: the giver to your ratings with a score of 9

Get Help:

- If the user asks, give an overview of some of the core functions and how to use them.
- It includes some examples, but wording can differ while maintaining functionality.

John Y: I'm confused, please help.

Librarian Bot: Here's a list of some of my core functions with some examples:

Change what I call you | I'd like to change my name

Tell me how you liked or disliked books you've read | I rate To Kill a Mockingbird by Harper Lee a 5/10.

Have me tell you your ratings | I would like to see my ratings.

Ask me to recommend you a random book to read | Give me a random book to read

Ask me about the details of a book | tell me about the ISBN/author/image/date of old yeller

Ask me to recommend you a book based on your ratings

John Y: I'd like a tutorial on what you can do

Librarian Bot: Here's a list of some of my core functions with some examples:

Change what I call you | I'd like to change my name

Tell me how you liked or disliked books you've read | I rate To Kill a Mockingbird by Harper Lee a 5/10.

Have me tell you your ratings | I would like to see my ratings.

Ask me to recommend you a random book to read | Give me a random book to read

Ask me about the details of a book | tell me about the ISBN/author/image/date of old yeller

Ask me to recommend you a book based on your ratings

Get Random Book:

- When the user asks the librarian bot will randomly choose a book from its knowledge base and print all the details related to it.

John Y: give me a random book to read

Librarian Bot: Here's a random book I've chosen from my knowledge base:

ISBN: 3442248876

Book-Title: Die Krieger der Drachenlanze 06. Die Ritter des Schwerts.

Book-Author: Roland Green

Year-Of-Publication: 1999

Publisher: Goldmann

Image-URL-L: <http://images.amazon.com/images/P/3442248876.01.LZZZZZZZ.jpg>

John Y: Thanks, could you give me another random thing to read?

Librarian Bot: Here's a random book I've chosen from my knowledge base:

ISBN: 1580171060

Book-Title: Herbal Tea Gardens: 22 Plans for Your Enjoyment & Well-Being

Book-Author: Marietta Marshall Marcin

Year-Of-Publication: 1999

Publisher: Storey Books

Image-URL-L: <http://images.amazon.com/images/P/1580171060.01.LZZZZZZZ.jpg>

Get information about a Book:

- There's a number of information you can ask the librarian bot about books.

```
John Y: Could you tell me about old yeller?

Librarian Bot: Here's all the details for Old Yeller:
ISBN: 006080971X
Book-Title: Old Yeller
Book-Author: Fred Gipson
Year-Of-Publication: 1989
Publisher: Perennial
Image-URL-L: http://images.amazon.com/images/P/006080971X.01.LZZZZZZZ.jpg

John Y: could you tell me the author for old yeller?

Librarian Bot: The Author I found for Old Yeller is: Fred Gipson

John Y: you could tell me the ISBN for old yeller?

Librarian Bot: The ISBN I found for Old Yeller is: 006080971X

John Y: what year was old yeller published?

Librarian Bot: The Year of Publication I found for Old Yeller is: 1989

John Y: show me an image of old yeller

Librarian Bot: The Picture I found for Old Yeller is: http://images.amazon.com/images/P/006080971X.01.LZZZZZZZ.jpg
```

View Ratings

- Shows the user their ratings and score/10 when prompted.

```
John Y: Can you show me my ratings for books?

Librarian Bot: John, your ratings for books are:
- the giver: 9/10
- bird: 10/10
- enders game: 10/10
- old yeller: 4/10

John Y: Give me a list of ratings i've made

Librarian Bot: John, your ratings for books are:
- the giver: 9/10
- bird: 10/10
- enders game: 10/10
- old yeller: 4/10
```

Get Book Recommendation:

- When prompted the librarian bot will run a ML algorithm and predict the top 20 books it thinks you'll like. It then returns a random book from the 20.

```
John Y: Can you give me a book recommendation

Librarian Bot: Here is the book I recommend based on your ratings:
ISBN: 087113375X
Book-Title: Modern Manners: An Etiquette Book for Rude People
Book-Author: P.J. O'Rourke
Year-Of-Publication: 1990
Publisher: Atlantic Monthly Press
Image-URL-L: http://images.amazon.com/images/P/087113375X.01.LZZZZZZZ.jpg

John Y: Can you give me another recommendation

Librarian Bot: Here is the book I recommend based on your ratings:
ISBN: 0394895894
Book-Title: The Ruby in the Smoke (Sally Lockhart Trilogy, Book 1)
Book-Author: PHILIP PULLMAN
Year-Of-Publication: 1988
Publisher: Laurel Leaf
Image-URL-L: http://images.amazon.com/images/P/0394895894.01.LZZZZZZZ.jpg
```

Confused response:

- If the user inputs something the librarian bot doesn't understand and doesn't know what to do with it'll give a confused output.

```
John Y: askldjasjdka
```

```
Librarian Bot: I'm not sure what you're asking, John. Could you please rephrase?  
If you need documentation let me know
```

Knowledge Base

The knowledge base I'm using is a dataset of over 270,000 books with their ISBN's, titles, authors, years they were published, publisher, and image urls of the cover.

The bot uses this knowledge base for a number of things such as finding details about the books the user is rating, retrieving information the user is asking for, or machine learning functions.

User Models

The UserModel is a class that contains the user's name, their ratings, and contains flags for different chat options and choices. Each user gets their own object of the usermodel that represents their 'account' based on the ID they provide.

```
class UserModel:  
    def __init__(self, name):  
        self.name = name  
        self.book_ratings = {}  
        self.name_prompt = False  
        self.add_book_prompt = False  
        self.rating_prompt = False
```

The PersonalLibrarian class loads and saves different user models as a pickled class based on what ID is set when the chat starts.

```
def load_user_data(self):  
    try:  
        with open(self.user_data_file, 'rb') as f:  
            return pickle.load(f)  
    except (FileNotFoundError, EOFError):  
        return {}  
  
def save_user_data(self):  
    with open(self.user_data_file, 'wb') as f:  
        pickle.dump(self.user, f)  
  
def get_user_model(self, user_id):  
    if user_id not in self.user:  
        self.user[user_id] = UserModel("unknown")  
    return self.user[user_id]
```

Evaluation

The librarian bot is pretty decent at what it does such as looking up books by title or extracting book titles, and is very quick to run because it's built from scratch on limited libraries.

However, because this bot doesn't use more extensive libraries it's more limited in the syntactic understanding of what it's being told and It also relies on regex for some of the prompt flagging and understanding so it may get commands wrong or get confused from time to time.

Overall I feel it functions pretty well for what it is and what it's built on, and it was a good project to work on to cement the understanding of the material we're learning.