

Chatbot Project

CS4395.001 Human Language Technologies

Brenden Healey
BMH180001

Quy Giang
QTG190000

System Description

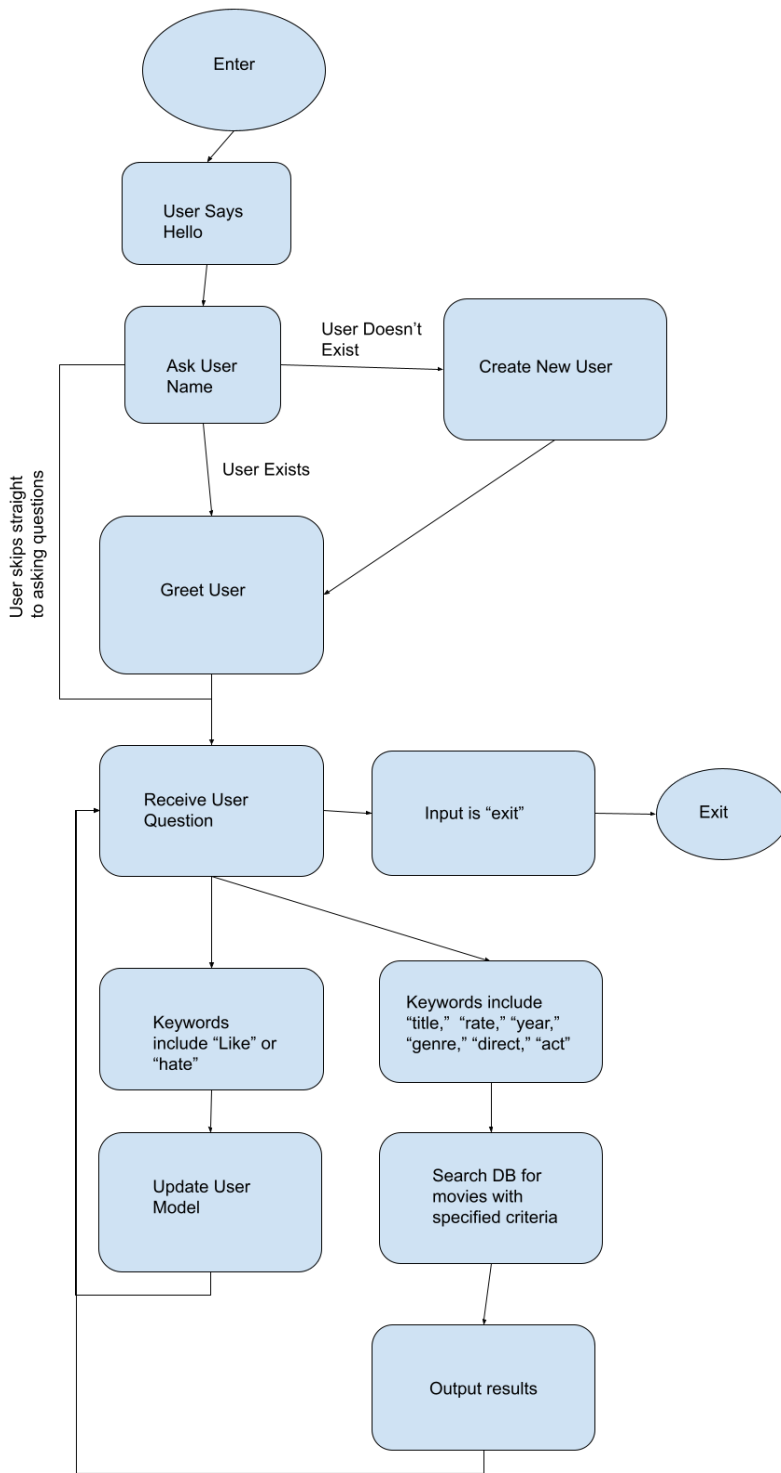
Our system, *Curator*, is a simple chatbot implemented with custom logic instead of a pre-existing framework. We chose to start our chatbot from scratch because we could not find a framework that played nicely with all aspects of our implementation, and it helped reign in our expectations of what we could accomplish in the given time. Starting with an ambitious idea of a bot that could find/generate memes for a user based on their input, we eventually settled on a movie recommendation bot.

Our chatbot utilizes many of the Natural Language Processing techniques that we have learned over the past semester. Stemming, lemmatizing, and tokenizing allow us to break up user input for processing. These tokens are then POS tagged. This allows us to identify proper nouns needed for managing user models and determining if the user wants recommendations involving certain individuals or titles. We also use keyword extraction to help the bot decide on an appropriate response to give the user.

Curator is a quirky bot that will draw from a large catalog of Netflix movies and TV shows containing over eight thousand entries in order to find the user something they might like. *Curator* is able to create a profile for a user that it can use to hone in their recommendations. It will remember a user's likes and dislikes of many different categories, including actors, genres, ratings, and more. *Curator* is able to recommend media from specific years, ratings, or topics, and can even refine its recommendations based on actors or directors.

Logic

The logic for *Curator*, modeled in the figure below, utilizes a simple control flow that processes the user's input to extract keywords, then crafts a response using a strict rules based model. *Curator* will always start off with an introduction, then ask the user to identify themselves. If the user's file is found, *Curator* will proceed with what it knows about them. If the user does not exist, a new user file will be created for them. The bot will check for any keywords in the user's input, and generate movie recommendations for each of the specified criteria. If the user mentions that they like, love, or hate something, *Curator* will add that information to the user's file and take that info into consideration when generating future recommendations.



Sample Interactions

Interactions always begin with a small blurb about what *Curator* is, then the user can begin talking to the bot. The user can choose to identify themselves, or go straight to asking questions. Should the user introduce themselves, the bot will search existing user files to determine if they are a returning user or a new one. A new user model will be created if one does not exist. User models will change over the course of the conversation as the user mentions what they like and what they don't like. *Curator* will search the user's input to determine what the user is asking for, and choose a response from a set of options. A series of example interactions are shown below.

Introduction:

```
Curator is an informant and recommender of Netflix Movies and TV Shows Up to Mid-2021
You: Hello
Curator: Hello, can I get your name?
You: My name is Sam
Curator: What can I do for you Sam ?
You: 
```

Sharing Likes and Dislikes:

```
You: I like Johnny Depp
Curator: I do like Johnny Depp
You: I hate Amber Heard
Curator: What a shame.
```

```
You: Hello
Curator: Hello, can I get your name?
You: My name is Sam
Curator: Nice to see you again, Sam who likes Johnny Depp and hates Amber Heard. How can I help you?
You: 
```

Surprise Me:

```

You: Surprise me
Curator: I will bless you with this movie/show:
          title    type      director \
6075  Aata Pita  Movie  Uttung Shelar

          cast  release_year rating \
6075  Sanjay Narvekar, Bharat Jadhav, Satish Phuleka...  2010  TV-14

          genre \
6075  Comedies, Dramas, International Movies

          description
6075  A municipal clerk with literary ambitions stal...

```

Asking About Type of Movie:

```

You: Tell me about horror movies
Curator: Here are the matches with your description:
          title    type      director \
1648  Ghosts of War  Movie  Eric Bress
6567  Darna Mana Hai  Movie  Prawal Raman
1590  Big Mouth  TV Show  NaN

          cast  release_year rating \
1648  Brenton Thwaites, Theo Rossi, Skylar Astin, Ky...  2020  R
6567  Aftab Shivdasani, Antara Mali, Boman Irani, Is...  2003  TV-MA
1590  Nick Kroll, John Mulaney, Jessi Klein, Jason M...  2020  TV-MA

          genre \
1648  Horror Movies
6567  Horror Movies, International Movies, Thrillers
1590  TV Comedies

          description
1648  Five, battle-worn Allied soldiers guarding a c...
6567  Stranded in a jungle when their car breaks dow...
1590  Teenage friends find their lives upended by th...
You: 

```

Big Mouth shows up in the results here because the description mentions the “horrors of puberty.”

Asking About Director or Actor:

You: Movies directed by Quentin Tarantino

Curator: Here are the matches directed by your requested director:

	title	type	director	\
7113	Jackie Brown	Movie	Quentin Tarantino	
392	Django Unchained	Movie	Quentin Tarantino	
3893	The Hateful Eight: Extended Version	TV Show	Quentin Tarantino	

	cast	release_year	rating	\
7113	Pam Grier, Samuel L. Jackson, Robert Forster, ...	1997	R	
392	Jamie Foxx, Christoph Waltz, Leonardo DiCaprio...	2012	R	
3893	Samuel L. Jackson, Kurt Russell, Jennifer Jaso...	2015	R	

	genre	\
7113	Dramas, Thrillers	
392	Action & Adventure, Dramas	
3893	TV Shows	

	description
7113	When an aging flight attendant's caught smuggl...
392	Accompanied by a German bounty hunter, a freed...
3893	Trapped at a stagecoach stop as a storm rages ...

You: █

You: Movies starring Alison Brie

Curator: Here are the matches with your requested actor(s):

	title	type	director	\
2939	Horse Girl	Movie	Jeff Baena	
2718	Community	TV Show	NaN	
2809	Search Party	Movie	Scot Armstrong	

	cast	release_year	rating	\
2939	Alison Brie, Debby Ryan, Paul Reiser, John Rey...	2020	R	
2718	Ken Jeong, Jim Rash, Donald Glover, Joel McHal...	2015	TV-14	
2809	T.J. Miller, Adam Pally, Thomas Middleditch, L...	2014	R	

	genre	\
2939	Dramas, Independent Movies	
2718	TV Comedies	
2809	Comedies	

	description
2939	A sweet misfit with a fondness for crafts, hor...
2718	When his degree is found bogus, lawyer Jeff Wi...
2809	Two guys hope to reunite their friend with his...

Ask for Movies from a Certain Year:

```

You: movies from the year 2018
Curator: Here are the movies and tv shows released in that year:
      title      type  director \
5023  Derren Brown: The Push  Movie      NaN
5086  Tom Segura: Disgraceful  Movie  Jay Karas
4311  Voltron: Legendary Defender  TV Show      NaN

      cast  release_year rating \
5023      Derren Brown      2018  TV-MA
5086      Tom Segura      2018  TV-MA
4311  Jeremy Shada, Tyler Labine, Bex Taylor-Klaus, ...      2018  TV-Y7

      genre      description
5023  Movies  Mentalist Derren Brown engineers an audacious ...
5086  Stand-Up Comedy  Tom Segura gives voice to the sordid thoughts ...
4311  Kids' TV  In an all-new series, five unlikely heroes and...
You: █

```

Ask for Movies by Title:

```

You: show me a movie called squid game
Curator: Here are the matches based on the search title:
      title      type  director \
33  Squid Game  TV Show      NaN

      cast  release_year rating \
33  Lee Jung-jae, Park Hae-soo, Wi Ha-jun, Oh Youn...      2021  TV-MA

      genre \
33  International TV Shows, TV Dramas, TV Thrillers

      description
33  Hundreds of cash-strapped players accept a str...
You: █

```

Knowledge Base

Our knowledge base is a Kaggle dataset of Netflix movies and TV shows, titled 'Netflix Movies and TV Shows.' It has been modified for our purposes to exclude certain columns that were not relevant. The dataset contains over 8800 entries, last updated in mid-2021. It is stored as a .csv file with the following columns: title, type, director, cast, release_year, rating, genre, description. An example of the first 25 entries are shown in the image below, though they are cut short due to the length of some of the fields.

```

1 title,type,director,cast,release_year,rating,genre,description
2 Dick Johnson Is Dead,Movie,Kirsten Johnson,,2020,PG-13,Documentaries,"As her father nears the end of h
3 Blood & Water,TV Show,,Ama Qamata, Khosi Ngema, Gail Mabalane, Thabang Molaba, Dillon Windvogel, Nata
4 Ganglands,TV Show,Julien Leclercq,"Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabiha Akkari, Sofia Lesa
5 Jailbirds New Orleans,TV Show,,,2021,TV-MA,"Docuseries, Reality TV","Feuds, flirtations and toilet tal
6 Kota Factory,TV Show,,Mayur More, Jitendra Kumar, Ranjan Raj, Alam Khan, Ahsaas Channa, Revathi Pilla
7 Midnight Mass,TV Show,Mike Flanagan,"Kate Siegel, Zach Gilford, Hamish Linklater, Henry Thomas, Kristi
8 My Little Pony: A New Generation,Movie,"Robert Cullen, Jos   Luis Ucha","Vanessa Hudgens, Kimiko Glenn
9 Sankofa,Movie,Haile Gerima,"Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra Duah, Nick Medley, Mutabaruka
10 The Great British Baking Show,TV Show,Andy Devonshire,"Mel Giedroyc, Sue Perkins, Mary Berry, Paul Hol
11 The Starling,Movie,Theodore Melfi,"Melissa McCarthy, Chris O'Dowd, Kevin Kline, Timothy Olyphant, Dave
12 "Vendetta: Truth, Lies and The Mafia",TV Show,,,2021,TV-MA,"Crime TV Shows, Docuseries, International
13 Bangkok Breaking,TV Show,Kongkiat Komesiri,"Sukollawat Kanarot, Sushar Manaying, Pavarit Mongkolpisit,
14 Je Suis Karl,Movie,Christian Schwochow,"Luna Wedler, Jannis Niew  hner, Milan Peschel, Edin Hasanovi  
15 Confessions of an Invisible Girl,Movie,Bruno Garotti,"Klara Castanho, Lucca Picon, J  lia Gomes, Marcu
16 Crime Stories: India Detectives,TV Show,,,2021,TV-MA,"British TV Shows, Crime TV Shows, Docuseries",Ca
17 Dear White People,TV Show,,Logan Browning, Brandon P. Bell, DeRon Horton, Antoinette Robertson, John
18 Europe's Most Dangerous Man: Otto Skorzeny in Spain,Movie,"Pedro de Echave Garc  a, Pablo Azor  n Willia
19 Falsa identidad,TV Show,,Luis Ernesto Franco, Camila Sodi, Sergio Goyri, Samadhi Zendejas, Eduardo Y  
20 Intrusion,Movie,Adam Salky,"Freida Pinto, Logan Marshall-Green, Robert John Burke, Megan Elisabeth Kel
21 Jaguar,TV Show,,Blanca Su  rez, Iv  n Marcos,   scar Casas, Adri  n Lastra, Francesc Garrido, Stefan
22 Monsters Inside: The 24 Faces of Billy Milligan,TV Show,Olivier Megaton,,2021,TV-14,"Crime TV Shows, D
23 Resurrection: Ertugrul,TV Show,,Engin Altan D  zyatan, Serdar G  khan, H  lya Darcan, Kaan Ta  yaner,
24 Avvai Shanmughi,Movie,K.S. Ravikumar,"Kamal Hassan, Meena, Gemini Ganesan, Heera Rajgopal, Nassar, S.P
25 Go! Go! Cory Carson: Chrissy Takes the Wheel,Movie,"Alex Woo, Stanley Moore","Maisie Benson, Paul Kill
26 Jeans,Movie,S. Shankar,"Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi, Nassar",1998,TV-14,"Comedies,

```

Sample User Models

Our User Models were kept very simple, just storing the user's name, likes, and dislikes as a dictionary that can be saved as a pickle object. An example of both the code and our test User Model are shown below.

Code:

```

# User Model
user = {'name': None, 'like': None, 'hate': None}

```

Test User:

```

{'name': 'Tester1', 'like': 'Lee Jung-jae from Squid Game', 'hate': 'Bollywood Movies'}

```

Real Person Who Tested Our Bot:

```

{'name': 'Sam', 'like': 'Johnny Depp', 'hate': 'Amber Heard'}

```

Evaluation

Our chatbot was tested and evaluated by 10 individuals. Each tester was given 10 minutes to try out the bot, with some guidance on some aspects of the bot to try out, such as . Testers were then given a Likert-style survey with 5 questions. Each question was answered with a rating from 1 to 5, along with an optional field for comment. Below are the questions, along with their average rating and feedback from the testers.

1 Curator was able to identify movies I would like

Average score: 4.5

Feedback: "Some recommendations seemed odd"

2 Curator was able to remember what I liked and disliked

Average score: 4.8

Feedback:

3 Curator remembered me if I left and came back

Average score: 3.7

Feedback: "It had some trouble finding my name"

4 Curator understood my questions

Average score: 4

Feedback: "It occasionally didn't know what I was asking"

5 Curator was able to find me movies from a certain year

Average score: 5

Feedback:

6 Curator was able to find me movies with a certain actor

Average score: 5

Feedback:

7 Curator was able to find me movies from a certain director

Average score: 5

Feedback:

8 Curator was able to find me movies from a certain genre

Average score: 5

Feedback:

9 Curator was able to find me movies with a specific rating

Average score: 5

Feedback:

10 Curator was able to suggest a random movie

Average score: 5

Feedback:

References

Bansal, Shivam. Netflix Movies and TV Shows, 2021. Available online at

[\[https://www.kaggle.com/datasets/shivamb/netflix-shows\]](https://www.kaggle.com/datasets/shivamb/netflix-shows) from Kaggle. Web. 8

November 2022.