

CHATBOT PROJECT

Overview:

This chatbot scrapes the dialogue from the play *Hamlet*. Once the text has been processed to only include lines of dialogue, it is tokenized. This separates each line or sentence of dialogue into its own entry in the knowledge base. The program measures the sentiment analysis of the user's response using sentiment analysis from VADER. Using cosine similarity, the program matches an entry of the knowledge base to a particular response. This entry is then displayed to the user.

NLP Techniques:

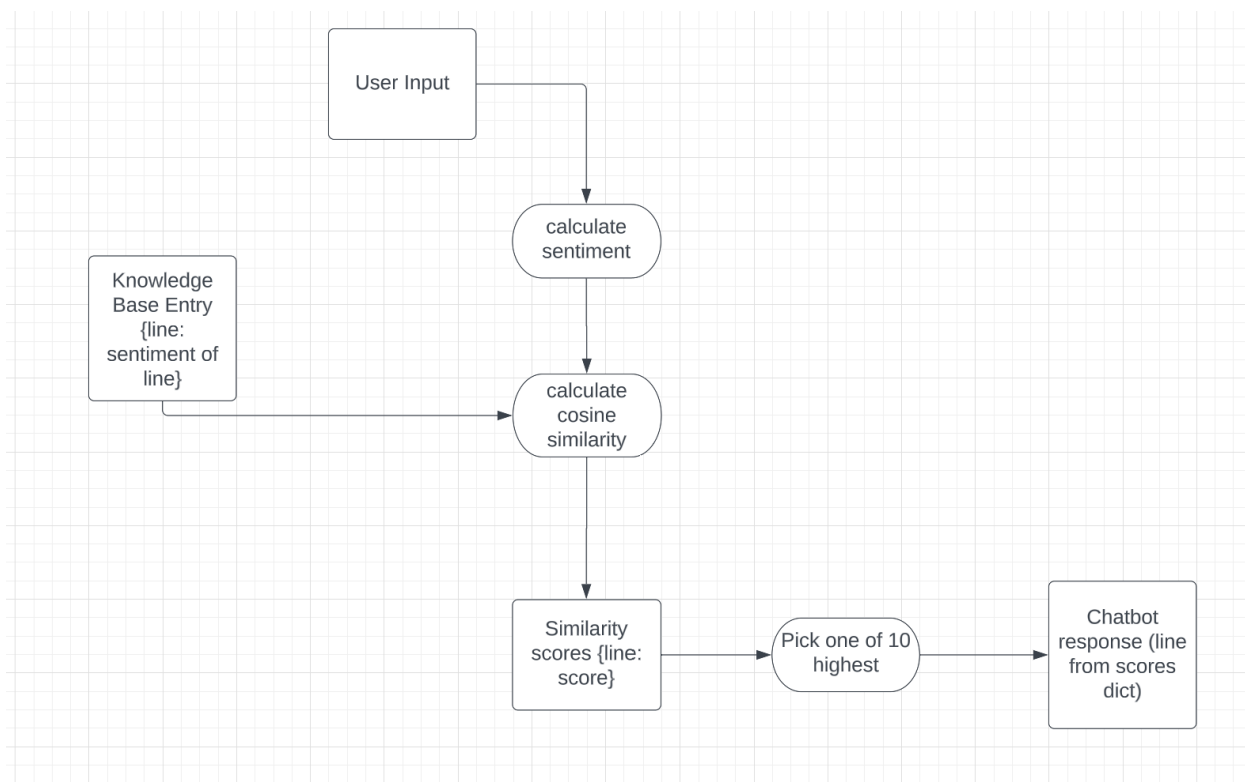
- BeautifulSoup to get the content of the website (Hamlet from Project Gutenberg)
- Sent_tokenize for lines of dialogue
- Tf-idf used to calculate cosine similarity
- Cosine similarity
- Sentiment analysis with VADER

Execution Details:

This project uses Project Gutenberg's *Hamlet* webpage as the basis of its knowledge base. First, BeautifulSoup is used to get a soup object of the contents of the website. Any text that is visible is stored to be processed. The text is now made up of dialogue, scene descriptions, act headers, and speaker indicators. We only need the dialogue for this knowledge base, so anything that isn't dialogue is removed.

The knowledge base is a dictionary that consists of the line of dialogue and its sentiment. Then the cosine similarity of the user's response and each piece of dialogue is calculated. The user is shown one of the most similar pieces of dialogue. The user's information is stored in a file. The user must respond with 'bye', 'bye!', 'Bye', 'Bye!', 'Farewell', 'Fare thee well', 'farewell', 'fare thee well', 'see you!', 'Goodbye', or 'goodbye' to exit.

Dialogue Tree:



Sample Dialogue 1:

BOT:

What is thy name?

USER:

Bob

BOT:

How fares my lord?

USER:

I'm great!

BOT:

Meghana Kambhampati

In the most high and palmy state of Rome, A little ere the mightiest Julius fell, The graves stood tenantless and the sheeted dead Did squeak and gibber in the Roman streets; As stars with trains of fire and dews of blood, Disasters in the sun; and the moist star, Upon whose influence Neptune's empire stands, Was sick almost to doomsday with eclipse.

BOB:

Are you ok

BOT:

But, woe is me, you are so sick of late, So far from cheer and from your former state, That I distrust you.

BOB:

Wow what did I do :(

BOT:

If he be now return'd, As checking at his voyage, and that he means No more to undertake it, I will work him To exploit, now ripe in my device, Under the which he shall not choose but fall; And for his death no wind shall breathe, But even his mother shall uncharge the practice And call it accident.

BOB:

So this sounds a little murder-y

BOT:

Murder!

BOB:

Riiiiiiight ok then

Meghana Kambhampati

BOT:

Up, sword, and know thou a more horrid hent: When he is drunk asleep; or in his rage,
Or in th'incestuous pleasure of his bed, At gaming, swearing; or about some act That
has no relish of salvation in't, Then trip him, that his heels may kick at heaven, And that
his soul may be as damn'd and black As hell, whereto it goes.

BOB:

Goodbye

Fare ye well, Bob

User Model 1:

[[["I'm great!", 0.6588], ["Are you ok", 0.296], ["Wow what did I do :(", 0.2263], ["So this
sounds a little murder-y", 0.0], ["Riiiiiiight ok then", 0.296]]

Sample Dialogue 2:

BOT:

What is thy name?

USER:

Meghana

BOT:

How fares my lord?

USER:

I'm ok

BOT:

And your water is a sore decayer of your whoreson dead body.

Meghana Kambhampati

MEGHANA:

Excuse me? How rude

BOT:

But thou wouldst not think how ill all's here about my heart: but it is no matter.

MEGHANA:

I do not care

BOT:

Give me that man That is not passion's slave, and I will wear him In my heart's core, ay,
in my heart of heart, As I do thee.

MEGHANA:

Flattery will get you everywhere

BOT:

Who would these fardels bear, To grunt and sweat under a weary life, But that the dread
of something after death, The undiscover'd country, from whose bourn No traveller
returns, puzzles the will, And makes us rather bear those ills we have Than fly to others
that we know not of?

MEGHANA:

I don't follow

BOT:

Follow her close; give her good watch, I pray you.

MEGHANA:

WHO?

Meghana Kambhampati

BOT:

Peace, who comes here?

MEGHANA:

Meghana

BOT:

But that I love thee best, O most best, believe it.

MEGHANA:

Are you flirting with me

BOT:

But, woe is me, you are so sick of late, So far from cheer and from your former state,
That I distrust you.

MEGHANA:

Ok that was unexpected

BOT:

You must not put another scandal on him, That he is open to incontinency; That's not
my meaning: but breathe his faults so quaintly That they may seem the taints of liberty;
The flash and outbreak of a fiery mind, A savageness in unreclaimed blood, Of general
assault.

MEGHANA:

Goodbye

Fare ye well, Meghana

User Model 2:

Meghana Kambhampati

[["I'm ok", 0.296], ["Excuse me? How rude", -0.4019], ["I do not care", -0.3875], ["Flattery will get you everywhere", 0.1027], ["I don't follow", 0.0], ["WHO?", 0.0], ["Meghana", 0.0], ["Are you flirting with me", 0.2023], ["Ok that was unexpected", 0.296]]

Knowledge Base:

The knowledge base has the following structure: {line: Sentiment of line}

{...

"Heaven secure him.": 0.6908,...

"[A dead march.]": -0.6486,

...}

Strengths:

The chatbot matches the sentiment of the user's responses fairly well. It also efficiently conveys the existential dread of *Hamlet* by focusing heavily on sentiment analysis. It is good at matching the user's tone and carrying that through a conversation. In general, this chatbot is simple.

Weaknesses:

The chatbot isn't great at carrying a conversation about a specific topic, seeing as the knowledge base is entirely made up of *Hamlet* dialogue. This chatbot is not knowledgeable about specific topics, but is a more general chat tool.