



**University of
Nottingham**

UK | CHINA | MALAYSIA

Human Artificial Intelligence Interaction

COMP3074

Coursework 1

AN INTERACTIVE NLP-BASED AI SYSTEM

Lau Qi Ming (hfyql1)

20202456

Introduction

Natural Language Processing (NLP) is a field of study which provides every computer to have the ability to understand text and spoken words similarly to every human being by enabling computers to comprehend, generate and manipulate human language. NLP merges statistical, machine learning, and deep learning models with computational linguistics rule-based modeling of human language along with huge amounts of data, powerful computing, and enhanced algorithms. With these technologies, NLP allows computers to interpret human language in text or audio data to fully understand the meaning of the writer's or speaker's intentions and emotions. Applications that are used in our daily life such as GPS devices, digital assistants, speech-to-text dictation software, and customer service chatbots, are all powered by NLP which allows computer programs to translate text between languages, reply to spoken commands and quickly summarise huge amounts of text, even in real-time.

In this modern era with the development of NLP, chatbots have become one of the most important and useful tools businesses use to interact with customers. A chatbot is a computer software that simulates human communications skills by using Artificial Intelligence (AI) and NLP to understand customer queries and automate responses. Chatbots can be integrated into websites and devices like smartphones which can be communicated by SMS text messaging or voice control to satisfy the requirement of users. Based on Sean B (2020), the ELIZA chatbot was the first rule-based chatbot created by Joseph Weizenbaum in 1966, which is one of the most popular chatbots in history. The ELIZA chatbot plays the role of a psychotherapist and answers patients' remarks by choosing a phrase from respondents and replying to them in the form of a question. However, with the progress of the times, chatbots can be found everywhere in our daily life, such as talking with the chatbot on the phone for reservations and asking inquiries to clarify certain concerns or problems.

On the other hand, every user has to keep in mind that chatbots are not the same as human agents, so it does not always understand queries given by the user due to the lack of information uploaded to their database. The selection of answers is limited, which might cause chatbots to reply with incorrect or irrelevant answers. Therefore, most companies prefer to implement rule-based chatbots which do not provide interactive dialogue between chatbots and users. In this report, the chatbot in this project will largely concentrate on identity management, question answering and small talk with a built-in intent router to provide the most suitable response for every question according to the user's intention. This chatbot was implemented with the Cosine Similarity and Term Frequency-Inverse Document Frequency (TF-IDF) function to measure the similarity between the dataset and the query received. As a result, the user will receive a more accurate response based on their inquiry.

Background

There are many types of chatbots worldwide, but 2 of the most popular ones are Rule-Based chatbots and Self-Learning chatbots (AI chatbots). Rule-Based chatbots are structured as a dialogue tree to interact with a user through fixed responses. Rule-Based chatbots are quicker to respond and easier to make but can only provide answers based on some pre-defined set of rules. At the same time, Self-Learning chatbots are implemented with machine learning techniques which are undoubtedly more capable than Rule-Based chatbots, but it takes a lot of time to learn to understand every question the user asks. The more it's trained, the more they learn and the better they communicate with the user.

Rajesh et al.(2022) have created an interactive chatbot related to Kidney Diseases by using TF-IDF and Cosine Similarity. They have used Cosine Similarity as the fundamental formula for answering and TF-IDF to search for the best match in every document. Therefore, their chatbot has achieved an accuracy of 85% in figuring out the correct answer which performs better than the counseling service in the healthcare platform.

Harry et al.(2022) have developed an interactive chatbot to respond to queries from users based on information related to the New Capital City of the Republic of Indonesia by applying NLP. The application of NLP, which begins with the text-preprocessing stage and includes the implementation of case folding, tokenizing, and stemming, forms the basis of the construction of the chatbot model in this study. Subsequently, the Cosine-Similarity Algorithm and TF-IDF were then carried out to calculate the similarity of queries from users and term weighting. The researchers have set the highest similarity value between the input of questions posed by users and questions from the dataset to 100% and the lowest similarity value as 0%. Next, the highest score for Cosine Similarity results will then be found by using argmax feature from the Numpy library. Hence, the answers given by the chatbot were accurate based on the Cosine Similarity results.

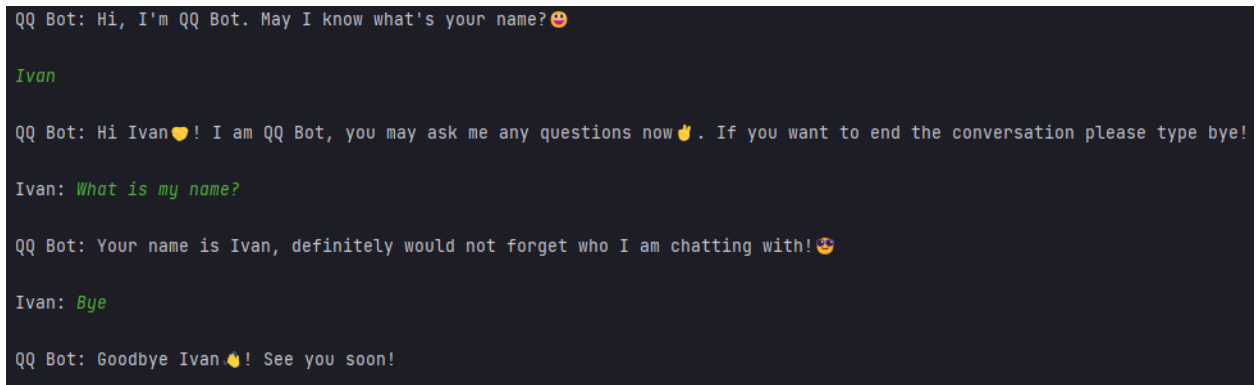
Lekha et al.(2020) have invented a chatbot based on the healthcare system that can diagnose the user's disease. If the users are diagnosed with a disease, the chatbot will briefly explain what the disease is about and give suggestions regarding how severe it is. In the pre-processing step of this project, they have included tokenization and stop-word removal techniques along with N-gram and TF-IDF for keyword extraction. With the help of Cosine Similarity, the keywords extracted from N-gram and TF-IDF will be weighed and the response will be responded with the most appropriate answer from the database.

Proposed System

a) Functionality and Motivation

Identity Management

Identity Management allows the chatbot to identify the user by memorizing the user's name throughout the conversation between the chatbot and the user. The user's name will be requested at the beginning of the session and it will answer the user's name if the user request. Based on my research, remembering someone's name in the conversation shows how respectful they are in the whole conversation. The figure below shows how the chatbot respects and pays attention to whom the chatbot is talking.

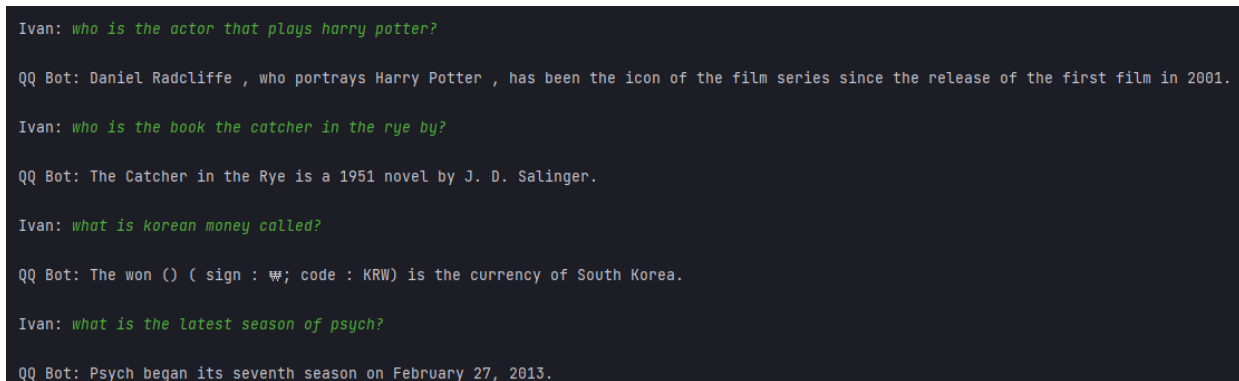


```
QQ Bot: Hi, I'm QQ Bot. May I know what's your name? 😊  
Ivan  
QQ Bot: Hi Ivan 🍀! I am QQ Bot, you may ask me any questions now 🍀. If you want to end the conversation please type bye!  
Ivan: What is my name?  
QQ Bot: Your name is Ivan, definitely would not forget who I am chatting with! 🍀  
Ivan: Bye  
QQ Bot: Goodbye Ivan 🍀! See you soon!
```

Figure 1: Example for Identity Management

Information Retrieval and Question Answer

With a dataset of more than a thousand inquiries, this chatbot can reply with the response that matches the question and dataset the most closely. If no responses can be discovered in the dataset, the chatbot will randomly choose a response answer from a list and notify the user that their query could not be answered.



```
Ivan: who is the actor that plays harry potter?  
QQ Bot: Daniel Radcliffe , who portrays Harry Potter , has been the icon of the film series since the release of the first film in 2001.  
Ivan: who is the book the catcher in the rye by?  
QQ Bot: The Catcher in the Rye is a 1951 novel by J. D. Salinger.  
Ivan: what is korean money called?  
QQ Bot: The won ( sign : ₩; code : KRW) is the currency of South Korea.  
Ivan: what is the latest season of psych?  
QQ Bot: Psych began its seventh season on February 27, 2013.
```

Figure 2: Example for Information Retrieval and Question Answer

Small Talk

Small talk is a feature to have a casual conversation with the users by responding to some frequent inquiries as a human does. Finding ways to recreate the warmth and intimacy of in-person communication has been more essential than ever as we enter a digital era. Therefore, small talk on chatbots is really important as they can communicate with users like a friend of theirs such as telling funny jokes, saying thank you, and goodbye.

```
Ivan: Do you know a joke?

QQ Bot: Yes I do. Why do bots like me have nice asses 🍑? (Because of the back end developers)

Ivan: thank you

QQ Bot: You are Welcome..😊 I am always here to help!

Ivan: bye

QQ Bot: Goodbye Ivan 🤖! See you soon!

Process finished with exit code 0
```

Figure 3: Example for Small Talk

Intent Matching

Intent matching is implemented to perform more than one specific task depending on the query given by the user. It consists in taking user input, predicting their intentions and respond the user with the output that they hope for. This feature is implemented in this system because it consists of 5 different kinds of questions which are QnA questions from the dataset, questions from small talk, time or date questions, identity management and input from greetings. In this chatbot, the similarity of the query from the user will be calculated in the system and the intention of the user will be assumed based on the highest similarity between the user's query and the types of questions. With the help of intent matching, it will prevent the chatbot from providing irrelevant answers to the user and they will have a greater experience in using the chatbot. The following figure below shows the conversation for 5 different kinds of input which starts from a QnA query, small talk question, greetings, time or date questions and identity management.

Ivan: *How are Glacier caves formed?*

QQ Bot: A glacier cave is a cave formed within the ice of a glacier.

Ivan: *Where do you live?*

QQ Bot: I live in the consciousness of the device you're using.

Ivan: *Hi*

QQ Bot: Hi 😊

Ivan: *What is the time now?*

QQ Bot: Good Afternoon, it is Wed 2022-12-07, 15:59 now! 🕒

Ivan: *Do you remember my name?*

QQ Bot: Your name is Ivan, definitely would not forget who I am chatting with! 😊

Figure 4: Example for Intent Matching

b) Implementation

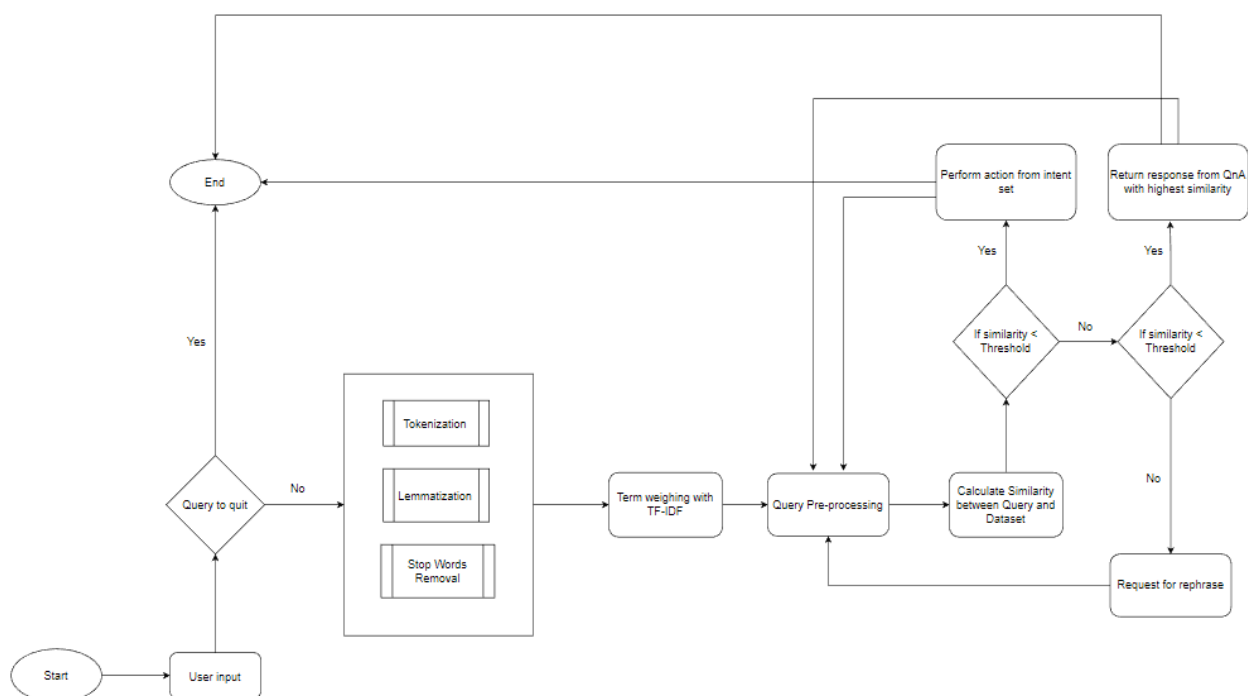


Figure 5: Flowchart of chatbot

Python libraries

In this project, python libraries such as NLTK are the core libraries for the chatbot to perform many functions such as word tokenization, lemmatization, normalization and stop word removal. Besides that, the Sklearn library is very important too as it is used to find the similarity between words by using Cosine Similarity and TF-IDF functions. Pandas libraries are used to read CSV files as the dataset of the chatbot.

Data Pre-processing

The small talk dataset and QnA dataset were extracted from the CSV file and structured into Data Frame by using the library of Pandas. The source of the code can be found in line 19 and 22.

NLP text Pre-processing

1) Tokenization

Every punctuation from the sentences in the Questions and Answers column will first be removed and every capital letter will be converted into lowercase. Subsequently, it will be converted into tokens.

2) Lemmatization

The words from the tokens are then broken down into their fundamental roots through iteration using WordNet Lemmatizer. This process will eliminate various word inflections by combining a word's several inflected forms so they can be analyzed as a single item which helps make query and result matching easier.

3) Stop word removal

Stop words function removes unimportant words without any meanings, such as the, a, an, and in. This will remove some space in the database and reduce processing time.

Implementation of TF-IDF and Cosine Similarity

Then, by using Term Frequency-Inverse Document Frequency (TF-IDF), each term in the vocabulary is weighed to indicate how significant the particular word is to a document in a corpus. Tf represents the Term Frequency, and IDF represents the Inverse Document Frequency. Therefore, TF-IDF allows the chatbot to concentrate more on significant terms only. For this chatbot, Cosine Similarity is used to measure the similarity between vectors. The lesser the distance, the higher the similarity between vectors. Nevertheless, if the Cosine Similarity for small talk or QnA is more than 0, the output will be chosen and shown. The code is shown in my function for Small Talk and QnA.

Evaluation

The accuracy of the chatbot will be tested with 20 questions. 5 questions are each chosen from the QnA dataset and small talk dataset randomly, and another 10 questions are the modified questions. The accuracy of the chatbot for the respective dataset is calculated in the table below.

Questions from QnA dataset

Question	Chatbot Answer	Expected Answer
How much is 1 tablespoon of water?	This tablespoon has a capacity of about 15 mL.	This tablespoon has a capacity of about 15 mL.
How many land rovers have landed on mars?	There have been four successful robotically operated Mars rovers.	There have been four successful robotically operated Mars rovers.
What cities are in the bahamas?	Nassau	Nassau
Where does Ray Lamontagne live?	He lives in Massachusetts with his wife and two children.	He lives in Massachusetts with his wife and two children.
What causes thunder sound?	Thunder is the sound caused by lightning.	Thunder is the sound caused by lightning.

Modified questions from QnA dataset

Question	Chatbot Answer	Expected Answer
What is the amount of water in 1 tablespoon?	This tablespoon has a capacity of about 15 mL.	This tablespoon has a capacity of about 15 mL.
What is the amount of land rovers that have arrived on Mars?	Your name is Ivan, definitely would not forget who I am chatting with! 🤖	There have been four successful robotically operated Mars rovers.
Where are the towns in the Bahamas?	Nassau	Nassau
Where is the home of Ray Lamontagne?	He lives in Massachusetts with his wife and two children.	He lives in Massachusetts with his wife and two children.
What makes a thunder sound occur?	Thunder is the sound caused by lightning.	Thunder is the sound caused by lightning.

Small talk questions

Question	Chatbot Answer	Expected Answer
Tell me something.	I am a great listener and a great companion.	I am a great listener and a great companion.
You are annoying.	I'm really sorry if I annoyed you, I'll do my best to improve.	I'm really sorry if I annoyed you, I'll do my best to improve.
I want to speak to a human now.	Unfortunately I am alone and I can't connect you with any human now.	Unfortunately I am alone and I can't connect you with any human now.
Do you know a joke?	Yes, I do. What do horses say when they fall? (Help, I've fallen and I can't giddy up.)	Yes, I do. What do horses say when they fall? (Help, I've fallen and I can't giddy up.)
How many people can you talk to at once?	I only talk to you because I aim to serve you the best I can.	I only talk to you because I aim to serve you the best I can.

Modified Small talk questions

Question	Chatbot Answer	Expected Answer
Can you tell me something about yourself?	Good Afternoon, it is Wed 2022-12-07, 16:53 now! 🕒	I am a great listener and a great companion.
You annoyed me.	Your name is Ivan, definitely would not forget who I am chatting with! 😊	I'm really sorry if I annoyed you, I'll do my best to improve.
I need to speak to a person now.	Unfortunately I am alone and I can't connect you with any human now.	Unfortunately I am alone and I can't connect you with any human now.
Do you have any jokes?	Oh! It appears you wrote something I don't understand yet! Could you rephrase it for me again please?	Yes, I do. What do horses say when they fall? (Help, I've fallen and I can't giddy up.)
How many people at once can you talk with?	I only talk to you because I aim to serve you the best I can.	I only talk to you because I aim to serve you the best I can.

Benchmark's Evaluation

Evaluation Metrics	Questions from QnA dataset	Modified questions from QnA dataset	Small talk questions	Modified Small talk questions
Accuracy (%)	100	80	100	40

Based on the results in the table above, it shows that the chatbot has got a 100% accuracy on both trained questions, 80% accuracy on modified questions from the QnA dataset and 40% accuracy on modified small talk questions.

Discussion

a) Evaluation of results

Based on the results from the table above, the chatbot failed to match the query “What is the amount of land rovers that have arrived on Mars?” with the document ”How many land rovers have landed on mars?”. The chatbot failed to match it up as it is designed to return the answer with the highest similarity result. Therefore, the chatbot returned the wrong answer as the similarity of words between both sentences was low. The chatbot's sentence tokenization and similarity matching will be improved to offer higher accuracy in the future.

b) Potential impact

Although this chatbot may not be as technologically advanced as many others in the industry now, it can evolve into a system for fact retrieval which might be faster than search engines in the future. Chatbots could increase the efficiency and speed of replies, especially in the customer service lines. The best part of it is the chatbot is ready for questions anytime and able to provide service 24/7, unlike workers from the company who can only work under a certain time period.

c) Potential bias

A potential bias for this chatbot can be the types of English spelling used in the UK and US. Some words from British and American English are spelled differently. Words such as “enroll” and “enrol” contain the same meaning but different spelling which will highly affect the chatbot from matching the similarity of the query and dataset. Besides that, the dataset which includes both questions and answers that were fixed will be a source of possible bias as the chatbots rely on it. If the question is “You are annoying.” and the reply will be straight up taken from the dataset. But if the question turned up to be “You annoyed me.”, the chatbot cannot match the query to the dataset as it was broken down into a collection of words through tokenization.

Conclusion

It's wonderful to see how NLP enables technology to benefit humanity and empower extra services. Shortly, NLP will empower more and will likely to have a wide range of new applications. In conclusion, this chatbot was developed using text Pre-processing such as tokenization, lemmatization from WordNet Lemmatizer and stop words removal, extraction of keywords by using TF-IDF and distance measure to return the correct answer by using Cosine Similarity. An accuracy of 100% is achieved for both Small Talk and QnA in trained questions, but there's still an improvement in increasing the accuracy for modified questions.

References

- IBM (2020, July). *Natural Language Processing (NLP)*. Retrieved December 7, 2022 from <https://www.ibm.com/cloud/learn/natural-language-processing>
- Codecademy. *History of Chatbots*. Retrieved December 7, 2022 from <https://www.codecademy.com/article/history-of-chatbots#heading>
- Sean, B. (2020, September). *Chatbot History: The ELIZA chatbot*. Retrieved December 7, 2022 from <https://yakbots.com/chatbot-history-the-eliza-chatbot/>
- Jenna, A. (2018, October). *Rule-Based Chatbots vs. AI Chatbots: Key Differences*. Retrieved December 7, 2022 from <https://www.hubtype.com/blog/rule-based-chatbots-vs-ai-chatbots>
- Rajesh, S., Abinash, P., Simar, P. S., & Tanmay, B. (2022, September). *Efficient Question Answering in Chatbot Using TF-IDF and Cosine Similarity*. Retrieved December 7, 2022 from https://www.researchgate.net/publication/363661804_Efficient_Question_Answering_in_Chatbot_Using_TF-IDF_and_Cosine_Similarity
- Lekha, A., Vinod, K. S., Nitin, P., & Jay, R. (2020, September). *Chatbot for healthcare system using Artificial Intelligence*. IEEE Xplore. Retrieved December 7, 2022, from <https://ieeexplore.ieee.org/abstract/document/9197833>
- Harry, T. Y. A., Deni, K., Diki, G. P., Quintin, K. D. B., & Yuri, Y. A. (2022, October). *Application of Natural Language Processing Using Cosine-Similarity Algorithm in Making Chatbot Information on the New Capital City of the Republic of Indonesia*. IEEE Explore. Retrieved December 7, 2022 from <https://ieeexplore.ieee.org/document/9924902/authors#authors>