

Artificial Intelligence Chatbot

Using Python.

Harry Clarkson (N0774446)

# Chosen Topic

The Topic I have chosen to use in my project is formula 1. I think this is a suitable topic as there are lots of areas to talk about and many useful questions could be asked.

# Requirements

The chatbot will have a certain list of requirements such as if someone asks about how the scoring system works it will reply with the answer using a rule-based method where it will recognise *exact* text to reply with the appropriate answer. However, if someone has another question a bit vaguer or the chatbot doesn’t have a specific response then the user will get a response which has been generated from Wikipedia. Such as if the user asks “What is a car” the response given will be what is wrote in the first few paragraphs on the Wikipedia page.

Another “method”/ requirement is that when the user asks a question if it can not directly recognise it from the xml file using the rule-based method or isn’t a vague question, then the system will refer to the csv file where it will compare the similarity between questions already recorded. When the system comes across a question that is *very* similar to a question it already knows the answer to, then it will give that answer back. However, if the system does not find any similarity it will report to the user that it does not know what the question is and that it must either be rephrased, or a different question must be asked.

# Modules Used

The first module that the system will run through will be the rule-based method where it will see if there is an exact match for the question given, If so it will respond with an answer.

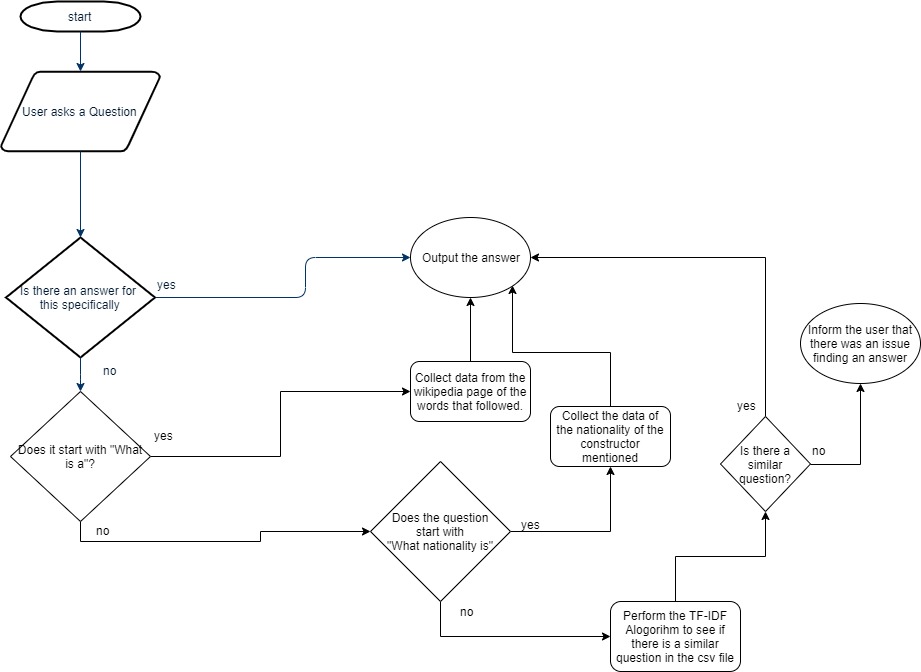
If it doesn’t find an answer straight away it will then see if the user has typed “What is a” then it will use the Wikipedia API to get the first paragraphs on the site on the specific topic given after the needed phrase.

After that, the third module is another API based on a F1 data site called ergast. It will ask what the nationality of a constructor is, such as “What is the nationality of McLaren” and it will return the data set in json file where the chatbot will single out the nationality part for McLaren.

The last module used is the similarity-based question and answers where when the user asks a question it may not directly correspond to a question set it has but will then compare the question to other questions stored in a csv format. From here it will look at the words used and compare the frequency between all the other questions to find the best match to give out a sufficient answer.

If the question cannot be found/ similar to the questions it has stored, then the chatbot will reply saying that it doesn’t understand meaning that the question will have to be rephrased or a different question should be asked.

# Flowchart



# Latest modules used

One of the latest modules used is the Sequence to Sequence model to try and give a reasonable response to the user asking/ talking to the chatbot. It is based off the other previous conversations to best predict something that makes sense and is relevant. As there are limited data sets, I have used a Movie conversation record to use rather than something related to Formula 1 as I simply couldn’t find any. Because of this rather than implementing it into my current chatbot it is separated by its own to distinguish that it is not related to F1. This was sourced from following a tutorial and example code at: <https://github.com/paulakash36/Chatbot>

Next I followed a tutorial to create a taxi game using OpenAI gym, using Q learning. It in cooperates Reinforcement learning. I followed a tutorial at <https://www.learndatasci.com/tutorials/reinforcement-q-learning-scratch-python-openai-gym/>. This was done by using a penalty and reward-based system. The AI gets +20 points for every drop off but loses -1 point for every time step it takes.