NLP Trivia Bot Write Up

This bot was reasonably difficult to produce as the installation process for SpaCy was difficult at best. In the end it was possible, but through researching how to pull data from Wikipedia, I came across the python plugin “wikipedia” that could do this for me. I used it, and as this returns the data I send for. However I was posed with these questions that the bot could not solve:

*“When did World War 2 start?”*

*“When did Stalin die?”*

Or any question not asking for the date as each different Wikipedia page has too much variety to attempt to incorporate for all of the different possibilities through my testing. My program can only encompass simple queries as, due to the nature of the human language, we can rarely test – given the situation we have – all of the different possibilities that are available. In addition my program is only designed to deal with time related queries, as for now the English Language is difficult to compute based on its complexities. Due to this I output the entire summary of the Wikipedia page for the user to make their roles easier.

Improvements that could be made in the obvious regard would be to incorporate more accuracy with scanning through the summary looking for dates, as currently it is a luck based idea, with the lowest date being the one that is assumed for birth/creation date, and the largest being the death/destruction date, as one may be able to incorporate this with a deep machine learning algorithm to scan through the different Wikipedia summaries in order to pick the key details. Also an improvement would be to incorporate other queries with answers, and more “try-catch” cases in order to skip pass errors, which could work as the SpaCy library has many features in which we could find a place or another detail of the like. In addition one other improvement could be to use pyttsx and Speech Recognition libraries to make the bot talk to the end user, and these could work because of our testing with our Personal Assistants that we wrote.

Written by Will Mann