

Supervision Meeting Notes

Jomar Alcantara

2018-06-27

1 My Progress so far

In the past two months, I have attended one training course on the development of applications for the Amazon Echo. This was a beginners guide to using the Amazon Echo and it's API as a tool for development. I have also attended the research module IS4001. This is part of the introduction to completing a Ph.D and consists of a 3000 Literature Review, 10 minute presentation and Organisational Plan. Finally, I attended the Alzheimer's Society Annual Conference in London. This allowed me to get an idea of the context in which my research is placed. It also allows me to understand what research is taking place in this area. At this conference I encountered ESCAPE (Nicholas Firth, See below) and this tool allows me to capture the transcribed speech of everything captured by the Amazon Echo.

I have started a trial of the Amazon Echo tool (Echo SCraper and CAssifier of Persons) using the Amazon Echo in my own home. The goal is to test the effectiveness and accuracy of the software at tracking what I'm saying, as well as looking at the richness of the text generated by my use. I'm going to do this over the course of a month.

As such I have conceptualised a datapipe line. This is currently based on the assumption that I will be using the Amazon Echo to collect data. Firstly, I will use ESCAPE to capture and transcribe the speech and output to a data file. I will then split each interaction into sentences and run this through the stanford parser to tag for various features. This will then be converted into features that can then be used in a Machine Learning algorithm to classify AD or Not AD.

I have collected a number of datasets, most notably the DementiaBank corpus of cognitive interviews. This has a total of 240 Participants with Dementia as well as controls who complete the Boston Cookie Theft exercise. I have access to other data such as the Presidents Speeches and Conferences archive. Currently I'm in the process of detagging these interviews as these are marked up in a different format.

I have conducted literature searches using Web of Science and Scopus which have highlighted a number of papers to read in the relevant areas. I have collected a number of these papers to read through which will contribute to my literature review. Periodically I conduct new searches and add to these papers and I have searches set up to automatically update me on papers.

2 What I intend to do over the next two months - end of August

In the next two months I plan to complete and submit an initial literature review as well as an initial experimental idea formalised and finally I would like to submit a first work-in-progress draft of the Qualifying Report. I also intend to look into the NHS ethics procedure as it's likely that I will need to be able to submit this before

I would like to look into how to measure the complexity of a sentence as I think a predictor could be how complex a sentence is. This certainly fits in with the idea that generating sentences increases cognitive load. Are the sentences required by Alexa to activate commands complex enough?

I would like to explore further the idea of Deep Learning for Natural Language Processing - the idea of pre-processing (One-Hot Encoding) vs vector representations.

3 Further down the line

It's my intention to attend the 2019 Alzheimer's Society Conference and as part of this I would like to submit a paper/poster presentation detailing my research so far. In addition, I will be attending the Natural Language Processing conference in Brussels (October / November 2018).

4 Questions and thoughts

Question - Is there any training you can recommend on Natural Language Processing? Particularly the Stanford Parser.

Question - The literature points to voice patterns as an area to explore, is this necessary? How might this work? Pauses might be useful to measure?

Thought - What if the language used in interacting with the Amazon Echo is not diverse enough? Test to see with my data? Vs DementiaBank

Thought - Is there a stage before collecting data with Amazon Echo where we just look at natural language in terms of speech between controls and those with dementia. Cookie Theft Picture + Semi Structured Interview?