Workplan

Jomar Alcantara

2018-04-17

1 Interest Statement

I am interested in whether technology can be used to aid the early diagnosis of dementia specifically by analysing speech and language in a natural environment. As part of this, I am exploring the use of technologies such as the Amazon Echo or Google Home kit to capture speech and language. I am using Natural Language Processing to generate features from the captured speech such that we can categorise people into specific categories and/or predict test results on scores using machine learning.

2 Literature Review

2.1 Dementia - Context

What is the Problem? Why is this important? What is the potential impact?

2.2 How do Psychologists assess dementia

What tools do psychologists use to diagnose dementia? Mini mental state examination (MMSE), Addenbrooke's Cognitive Examination(ACE-III) Wechsler Adult Intelligence Scale(WAIS) The Free and Cued Selective Reminding Test (FCSRT)

2.3 How do Psychologists understand how language is affected by dementia

What language deficits are there and how are they characterised? What speech deficits are there and how are they characterised?

2.4 How is Natural Language Processing best positioned to help in this area?

2.5 Are there any other technologies that can be used to help diagnosis?

2.6 What Research has been done in the area so far?

2.7 Conclusion

3 Stages of research

4 Stage 1 - Data capture

5 Training needs analysis

5.1 Psychology

Training in delivering the WAIS or psychological tools that might be needed

5.2 Programming

- Online courses in Python Programming
- Online courses in Machine Learning

6 Data collection

6.1 DementiaBank and AphasiaBank

Holland Corpus
Consists of 2 participants a 62 year old man with mild AD and a 78 year old woman with more advanced AD.
Kempler Corpus
Consists of 6 participants with a range of 65 to 86 (mean, sd)
Pitt Corpus
Consists of 104 control participants, 208 with dementia and 85 with an unknown diagnosis.
DePaul Corpus
Consists of 1 participant, 66 year old female with PPA.
Hopkins Corpus
Consists of 36 participants, 23 males and 13 females

6.2 Three Authors, P.D. James, Iris Murdoch and Agatha Christie

Data to be downloaded

6.3 Presidents Speeches

Data to be downloaded

6.4 Alexa Skills Kit and Amazon Web Services

- **Training in Alexa Skills Kit - Booked**
- Training in Amazon Web Services

7 Stage 2 - Data Processing, Feature Construction

7.1 Natural Language Processing

Online courses in Natural Language Processing

- Jurafsky and Martin - Speech and Language Processing
- Manning and Schutze - Statistical NLP
- Manning and Raghavan - Intro to IR

8 Stage 3 - Data Analysis and Machine Learning

8.1 Application of Machine Learning Algorithms

8.2 Experimental Design

Battery of tests to differentiate between Controls, those with MCI, those with AD.

- FCSRDT - As most effective

9 Proposed Timeline, Key Dates and Deliverables

9.1 April 2018

Wednesday, 18th April - Workplan (First Draft Agreed)

9.2 May 2018

Wednesday, 9th May - Alexa Skills Training (All Day)

9.3 June 2018

Friday 15th June - First Draft of Literature Review to be handed in.
Monday 18th June - Monday 25th June - Annual Leave

9.4 July 2018

9.5 August 2018

9.6 September 2018

9.7 October 2018

9.8 November 2018

9.9 December 2018

9.10 January 2019

Friday 11th January - Rough Draft of Qualifying Due Friday 18th January - Complete first rough pass of capturing and processing speech.

9.11 February 2019

Friday, 1st February - Qualifying Report Due

9.12 March 2018

9.13 April 2018

Monday, 1st April - Viva Due by this point.