

Response to the Qualification Report and Viva Voce feedback

March 27, 2019

The student had a very good understanding of the literature and could summarise it well. It was a bit lacking in the machine learning area but he has the capacity for improving that and is aware of what he needs to do. His research background was good but there needed to be more focus on the research questions. These were not explored well in the QR but were much clearer in the viva. He now has some very good ideas to take forward that should lead to a Ph.D.

The main advice we would give is to be much more careful about how and what data you collect. Getting the participants to do it themselves on a daily basis by, for example, a speech diary that they record and submit, is much more practical than a battery of tests in a clinic. Plus you do not want to be asking about clinical diagnoses using, for example, PHQ9 that measures depression because this raises a lot of ethics issues that will make it harder for NHS ethics approval. Keep the data collection simple, especially as the focus of the Ph.D is about changes in language and not about the cognitive tests themselves. We do not think the latter are necessary because the focus of the Ph.D is on whether you can detect decline by differences in language use over time and not about determining the clinical underpinnings of those differences.

1 Research Aims and Questions

In my The revised objectives of my research are to explore whether the application of Natural Language Processing and Machine Learning techniques to spontaneous speech samples can improve our ability to diagnose Mild

Cognitive Impairment in patients that will directly lead to Alzheimer’s Disease and to provide some clarity as to which language features are most important to measure when looking for evidence of MCI and/or AD.

- **(Hypothesis One)** Monitoring language samples over the course of the year will be enough to detect cognitive decline in those diagnosed with Mild Cognitive Impairment and Early Alzheimer’s Disease at specific points in time.
- **(Hypothesis Two)** Given a number of samples of language over a given time frame, it is possible to detect language decline in participants such that you can differentiate between normal language decline, these with Alzheimers Disease and decline in those with MCI.
- **(Hypothesis Three)** T

2 Data Collection and Analysis Methodology

One of the main concerns I have had was about the amount of data collection that needed to be done. Given the low rate of progression from MCI to AD year on year, there is a non-zero possibility that I could have a cohort where no-one progresses to AD. Equally, working with a clinical population would invariably require NHS Ethics. Since this problem is essentially a classification problem, in theory we can look at things from a different perspective.

2.1 Why not stick to picture tasks or narrative storytelling tasks?

Semi-structured speech data in commonly available data sets such as the DementiaBank corpus or the Wisconsin Registry for Alzheimer’s Prevention (WRAP) focus on picture description tasks. These samples are usually collected once an individual is referred to a clinician for assessment.

2.2 One Class Classification

The prevalence of Dementia in the UK's population is roughly 1 in 14 or 7%. The prevalence of MCI has only briefly been explored, however some research by a team at the University of Leeds indicate that the prevalence of MCI is between 3.2 - 24%, and found that the prevalence of MCI in people who have self-reported memory problems is almost double (18%) the amount of do not self report. Whilst this means that it is entirely possible to recruit a cohort with MCI and AD, it is comparatively costly in terms of time and resources. Given that, from a statistical perspective, this is a highly imbalanced population we can look at the problem as a kind of anomaly detection.

This has some precedent in the field of . In

One-class classifiers analyze only examples of the majority class (in this case - Healthy Older Adults) to learn a classification boundary that excludes outlier examples. When dealing with heavily imbalanced datasets with text classification as an objective, the one-class classifier approach has been demonstrated to outperform conventional two-class classification.

3 Proposed Methodology

3.1 Stage 1 - Exploratory Data Analysis of DementiaBank dataset

The aim of this stage is to explore language features and acoustic features to an existing dataset. This allows the development of an automated pipeline using a more accurate speech-to-language application (IBM Watson). This also allows us the opportunity to explore the use of the one-class classifier to an existing data set as a proof in concept.

3.2 Stage 2 - Collection of Longitudinal Data

Given the aim is to collect a large sample of language from a healthy population. I now estimate the numbers of samples to be as follows. 100 Healthy

Older Adults, 20 MCI, 20 Early AD. The 20 MCI and 20 Early AD will be categorised based on a clinical diagnosis given by a mental health professional. In addition, from the Healthy Older Adults, I will ascertain a participants level of concern about their cognitive ability with a simple 2 item questionnaire. The intention, if there are enough people who express a level of concern, is to conduct an analysis of the majority class with and without those who demonstrate concern.

- Question 1. How concerned are you about your cognitive health? Rate on a scale of 0 to 10, where 0 is not concerned at all and 10 is extremely concerned.
- Question 2. If you answered 1 or more on Question 1, please explain what your concerns are.

3.3 Stage 3 - Analysis of Longitudinal Data

Once the data has been collected, we can test the automated pipeline on new data which allows the validation of this protocol. The analysis of the data will begin from the perspective of treating the healthy older adults as the Majority class, and all others as 'outliers'. The first stage of this is to see whether using a One-class classifier is sensitive enough to detect those with a diagnosis of MCI and AD without differentiating between the two.

This is important from the perspective of identifying people

4 Revised Chapters of Thesis

- Chapter 1 - Introduction
- Chapter 2 - Background and Related Work
- Chapter 3 - Feature Engineering and Classification of DementiaBank Transcripts and Audio.
- Chapter 4 - Comparison of different speech to language applications.

- Chapter 5 - Classification of language generated from Longitudinal Study.
- Chapter 6 - Techniques in feature engineering - What features matter? Comparing language generated from picture description task to language generated from our task.
- Chapter 7 - Conclusions and Future Work