

Project Plan

Jomar Alcantara

2018-07-13

1 Background, Aims and Objectives

According to a recent report commissioned by the Alzheimer's Society in 2015, they estimate the prevalence of Dementia in the UK at approximately 815,000 people. This represents 1 in 14 of the population aged 65 or over and 1 in 79 in the general population. This report also estimates an annual healthcare spend on 4.3 billion of which approximately 85 million is spend solely on diagnosis. They also estimate that the overall impact of dementia (excluding the costs associated with early onset dementia) is 26.3 billion annually. Globally, this picture is a lot bleaker. A recent report suggests that in 2015 there were 46 million people with a diagnosis of dementia and that number is expected to hit 131.5 million by 2050. The report also states that the worldwide cost of dementia in 2018 is estimated to be in the region of one trillion US dollars.

The aim of the research I am conducting is to investigate the use of technology, particularly the use of Machine Learning and Natural Language Processing techniques to detect mild cognitive impairment (MCI) and / or Alzheimer's Disease (AD). The objectives are: To design a framework and develop. To develop a proof of concept design. To investigate the application of my framework to be used in a more naturalistic setting i.e. in a person's home, rather than in a healthcare setting.

2 Tasks

2.1 Literature Review

The Literature Review is divided into three sections as my research is interdisciplinary. I will be looking at the psychological perspective of language testing and analysis for MCI and/or AD as this will provide the theoretical foundation for any work I do with Machine Learning. I will then will look at the different possible ways to measure for language difficulties taking into account approaches that use frequencies, approaches that use measures of semantic and syntactic complexity and other more novel approaches. Finally, I will look at current developments in Machine Learning as it pertains to this area of research including theoretical approaches as well as previous research applying research in this specific area.

2.2 1st Year PhD Requirements

There are a number of requirements for the first year of my Ph.D Programme. The first part of this is the IS4001 module which consists of three assessments. These are: A presentation, A project plan and written report and a 3000 word literature review. In addition, I must complete a qualifying report which is a report in which I talk about my progress and how I have developed as a researcher in the first year. This also looks at any papers I have submitted / published and a literature review in the area I'm studying. Finally, I must defend my qualifying report at a Viva.

2.3 Skills Training and Knowledge Gathering

Under this section, I have identified a number of knowledge gaps where I need to work on. The main knowledge gap is in learning a new programming language and the extensions I will need to learn and use in my research. Given this is a fundamental part of my research, I have allowed six months to get up to speed. I also plan on attending two key conferences in my field of researched.

2.4 Research Planning

My research planning is going to be based on identifying gaps in the literature and planning some experiments to test proof of concept. Based on my literature review, I will be designing an experiment which means the objectives I have listed above. As part of the research planning stage, I intend to submit my research proposal for NHS ethics approval in order to join the clinical research network Dementias and Neurodegeneration. This process can take up to 60 days for the committee to initially look at the research proposal and I have allowed time for the research proposal to be rejected and resubmitted.

2.5 Sample Dataset Collection and Cleaning Model Conceptualisation and Production

I have identified a number of pre-existing datasets that I can use to develop my models and to test the data pipeline. However these datasets are owned by other research groups. There is an application process in order to obtain some of the datasets and they require approval of the research group before the data can be released. Using this data I plan to develop a model and test methodologies on these datasets. This is mainly a replication of previous research but allows me to validate my model and also allows me to smooth out any difficulties in the processing of large sets of data before I start to collect my own data.

3 Obstacles

There are a number of obstacles that may delay my progress. The most obvious example is my application for NHS Ethics Approval. The average time taken for

this to be accepted is 60 days. I have allowed 90 days in my project plan as my research proposal may require amending and resubmitting. Another obstacle is in the acquisition of datasets. My experience with acquiring other datasets is that applications can take up to 2 weeks to be approved. I have therefore allowed up to a month for each dataset including the application process.

In terms of managing my own time. I have assumed working eight hour days from Monday to Friday and for each task, I have estimated the time I would need to complete the task and then added on an additional 33% to 50% as contingency. This should allow me the ability to 'catch up' if I fall behind at any stage.