

Midterm Report

Massey University 159333 - TAIC report summarizer

James Thompson (21011195) - 1jamesthompson1@gmail.com

11/09/2023

Introduction

Work and progress has been steady over the first half of the semester. The project is being developed openly here. In the proposal I discussed the three deliverables. The first one has been completed and the second one is in progress. The third deliverable is set to be defined after feedback from the second deliverable is received.

Completed tasks

There have been three main tasks completed so far. Framework, basic summarizer and basic theme generation.

Framework

By framework I am referencing the whole process of downloading, parsing and then looping through all of the reports. But in general the work here can be seen in the Gather_Wrangle module. This module is the simplest of the three modules and is the most complete. Various other scripts/classes have been made to help speak with the openAI api and managing the various input/out files; Config, Themes and OutputFolderReader.

Basic summarizer

Basic theme generation

Remaining tasks

There are broadly three remaining tasks.

Convergence

Firstly is arguably the most important. Reliability of the engine and guaranteeing the output is correct is outside the scope of the project. However making sure that the output is consistent and converges to some values will be quite important. Due to its importance this is the main goal of the second deliverable.

Summarization

Theme generation

viewer app

As discussed in the here the viewer app is a deliverable that has been added to the project. This will be a Shiny app that will allow the user to view the output of the engine in a user friendly way. This will be a simple app that will allow the user to select a report and then view the summary and themes. It is currently

at the point of having a basic search function. Work will be done before the second deliverable so that the search function is fuzzy (#23).

It is being made use Shiny for python as it is the simplest to get a basic web app up and running. I am using a web app so that the end user can simply visit a website to see the app in effect.

Extraction of more information

As discussed in the initial proposal there will need to be the addition of more information from each of the reports. This will allow the viewer app to provide more information to the user. I expect this work to be completed in the third deliverable.

Scope and and Objective changes

The scope of the project remains mostly unchanged. The two changes that have been made are the addition of a Shiny app to give the output folder a user friendly GUI to demonstrate. The second change is the widening of the engine which has changed the third deliverable.

Shiny app

Engine widening

Conclusion