Initial Plan

Title page, originality statement

Abstract

Acknowledgments

Table of Contents

List of Figures and Tables, Formulas and Acronyms:

* Would this include all the bar graphs and WER results? Can it be included in main body?
* Definitions for WER, MER, WIL, ASR, all the terms that will be used

Literature Review

* Updated from the initial submission, up to 5 pages

Research Methods

* Comparing different word error metrics, maybe have a table of relevant info
* Outline datasets used, (Mansfield, JL, Mozilla)
* Outline commercial ASR systems being used and why they were chosen
* Outline different Open-Source ASR systems and why speech brain was chosen
* (Maybe) compare the different pre-trained models and why our one was chosen (although this may be hard since we chose our model as it was the only one we can get to work properly)

Experiments and Results

* Outline Commercial ASR for google and microsoft (Louis)
* Outline Commercial ASR for Amazon (Henry)
* Getting embeddings for the datasets (Louis)
* Finding most common words in dataset (Henry)
* Splitting data to test, train, valid (Henry)
* Outline finetuning pretrained model with datasets and testing(Louis)
* Developing graphs from finetuning results(Henry)

Discussion

* Comparing finetuned models to see the best performing one
* Compare finetuned models vs commercial ASR
* Compare commercial ASRs on NZ vs US models
* Draw conclusions on comparisons
* Outline any limitations (i.e, small dataset)
* Research Objectives and Questions met

Conclusions

* Summary of contributions
* Definitive conclusions based off discussions outlined
* Further research