Plan of Work Task 1

Karnkumar Dalmia

June 2020

1 Introduction

A key technology that our team is working on is on finding a better representation of speeches, or other corpora of texts, in a graphical format, and using that representation to convert back to a more appropriate textual format, notably an abstract. In particular, we focus on creating a knowledge graph representation of a speech, and then in an latter phase, converting that knowledge graph to an abstract textual representation.

2 Speech to Knowledge Graph

In this first phase of our system, we effectively want to create a knowledge graph representation of a provided audio speech. A knowledge graph is a graph-based representation of ideas with parts of speech dependencies denoted via graph links, with varying parts of speech shapes as nodes. Within this first phase, we follow a three step pipeline. First, the audio sample is ingested into an LSTM model, that outputs words stutters, or unfiltered word predictions of said audio. Next, our system prepossesses the unfiltered text, and subsequently outputs a cleaner, punctuated format of the text. This is to help our NLP system, described in the third step, better create a knowledge graph. Finally, based on the filtered text, we create an NLP system with POS tagging to effectively decide word constituent dependencies in the text, and then proceed to create a knowledge graph out of that.

3 Knowledge Graph to Abstract

In this section of the pipeline, our goal is to create a mild abstract of the knowledge graph representation of the speech. In doing so, we utilize a GraphWriter technology that learns, via machine learning, the graph dependencies. After doing so, it outputs a stream of text words, that get utilized to create a textual representation of the graph, aforementioned earlier.

4 Karn's Work

I will be mostly working on appending to the first phase of the model, that is, the "Speech to Knowledge Graph" portion. In this section, a key appendage is to also include audio features in a particular speech to be included in the knowledge graph representation. This could mean capturing the tone, or level of volume, of the speech as well. As for now, my current plan for this add-on is to utilize color technology to denote varying speeches tone and volume. A darker gradient on words in the knowledge graph would correspond to lower volume, whilst lighter higher volume. A color mechanism can be introduced to denote the tone of the speech. For example, a red hue could denote anger, or passion in the select utterances, whereas a bluer hue could denote calm. As such, both the described aspects of audio features can accurately be denoted in the knowledge graph with varying color and brightness.