

Presidential Sentiment Analysis
MSDS692 – Data Science Practicum 1

Progress Report for Week 7

Jeremy Beard

Project Details

This project is centered around a dataset which contains speeches from the State of the Union of all presidents. The project will utilize a variety of natural language processing techniques in order to answer questions that have been created surrounding the dataset. Sentiment analysis will be utilized, general word commonality will be explored and word frequency will be analyzed. The final output will be a visualization comparing all the presidents to each other within the lens of the State of the Union.

Project Timeline:

- Week 1 – Project definition and submit proposal (DONE)*
- Week 2 – Datasets selected and former related work collected (DONE)*
- Week 3 – Initial data loading and data cleaning performed (DONE)*
- Week 4 – Initial output and analysis for single speeches completed (DONE)*
- Week 5 – Output and analysis for all speeches / all presidents completed (DONE)*
- Week 6 – Result congregated, summary visualizations created, presentation began (DONE)*
- Week 7 – Presentation completed, dry runs completed (IN-PROGRESS)*
- Week 8 – Final Project Presentation*

Planned Work for the Week:

From last week, my planned work for this week was to start the report and presentation! Now that the analyses are a bit more finalized, I can begin the report on the State of the Union speeches and what I found when analyzing them with my guiding map of questions. This list of questions is a living list but the questions have been mostly finalized as follows:

- *What is the quantitative positive/negative sentiment between all the speeches?*
- *Which presidents use the widest variety of words?*
- *What is the quantitative positive/negative sentiment between all the presidents?*
- *What are the themes or buzzwords among the different speeches / presidents?*
- *What were the most common words used in each speech?*
- *Which presidents gave the longest speeches?*
- *Which presidents gave the shortest speeches?*
- *Which presidents used the most unique words?*
- *Which presidents used the least unique words?*

This week I planned to start the presentation and finalize what I was going to present. This also includes the report. This week begins the final steps of the project and what will be shown to the rest of the class.

Progress for the Week:

This week I made a lot of progress, following the previous week where I was heavily distracted purchasing a car. This week I returned to the scripts and finished everything which I wanted to output. I then started the initial presentation steps. The finishing steps on the presentation still need to be implemented but the foundation is in place.

Roadblocks/Issues:

No roadblocks or major issues to report! Following the car purchase last week, progress has continued on this analysis. All charts have been finalized and the report has been started. The reports will be finalized Monday, Tuesday, and Wednesday and will be submitted when the deadline hits.

Plan for next Week:

Next week I plan to finish the presentation side of materials Monday, Tuesday, and Wednesday. I will submit these presentation materials on Wednesday and will review other classmates' materials the rest of the week. Thank you.

Resources for the Week:

State of the Union Corpus (1790 - 2018). (2018, October 19). Kaggle.

<https://www.kaggle.com/datasets/rtatman/state-of-the-union-corpus-1989-2017>