

Presidential Sentiment Analysis
MSDS692 – Data Science Practicum 1

Progress Report for Week 6

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 (DELAYED)*
- Week 7 – Presentation completed, dry runs completed*
- 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 was actually heavily distracted. Following some cold weather we got, I have been in the car market for awhile and ended up closing on a pre-owned car last Friday. I didn't make much progress on the report this week so I need to sit down tomorrow and make some progress with it. This past week I was just heavily researching cars and the market and what models are out there, so honestly I really didn't make progress with the report yet! But all that is done and I plan to hit it more heavily this week. I did see your message about not being available for the project review so maybe we can go into that this coming week.

Roadblocks/Issues:

No roadblocks or major issues to report yet, just the distractions like I mentioned! Now that I have a car for the winter, I can get back to focusing on this project 😊 that was the main roadblock this week, unknowingly.

Plan for next Week:

Next week I plan to hold a project review of some sort to hopefully gain a sense of anything that needs to be added to the project. I also will definitely build out the report and presentation a bit more as that is really the focus as we close in on these last couple weeks. 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>