Pei Lun Hung
CS 410
Professor ChengXiang Zhai
November 29th, 2020

Progress Report

1. Completed tasks

- Filter *The New York Times* corpus for articles from May 2000 to October 2000 that mention "Bush" or "Gore."
- Apply latent Dirichlet allocation (LDA) to the cleaned *New York Times* corpus to identify possible topics.
- Use Granger testing to test causality and to derive the set of candidate causal topics with time lags.

2. Pending tasks

- Apply Granger testing to find the most significant causal words among top words for each candidate topic, and track these words' impact values.
- Define a prior on the parameters of the topic model using these significant terms and their corresponding impact values, and apply LDA to the *New York Times* corpus using this prior.
- Iterate, and replicate one set of experimental results by using prices from the Iowa Electronic Markets 2000 Presidential Winner-Takes-All Market.

3. Challenges

- Fully understanding how to set the strength of the prior in each iteration by using μ.