Text Analysis

Instructor: Byengseon Bae
Email: byengseon.bae@cgu.edu

Course Description

This course examines text analysis as a method. It covers both qualitative and quantitative approaches, including skills in interpreting political texts and cutting-edge natural language processing techniques.

Learning Objectives

- Understand the differences between qualitative and quantitative approaches.
- Learn how to interpret political texts qualitatively.
- Apply quantitative techniques to analyze textual data.
- Improve our ability to analyze texts.

Required Texts

- Readings and software tutorials will be provided

Course Requirements

Grading	Weight
Participation	10%
Assignments	20%
Presentation	20%
Midterm Exam	25%
Final Exam	25%

Attendance: I will expect you to attend all scheduled classes. If emergencies prevent you from attending, please let me know as soon as possible.

Participation (10%): Please read, take notes, and think about all assigned readings before class. I encourage participation in the form of questions and comments.

Assignments (20%): All assignments must be submitted by the specified deadlines.

Presentation (20%): You will be required to give a presentation at the end of the semester.

Midterm Exam (25%) and **Final Exam** (25%): You will be required to write your answers in response to the prompts I provide on the exam dates.

Course Schedule

Week 1: Introduction
What is text analysis, and why does it matter?

- No required readings.

Week 2: Interpreting Political Philosophy Texts
The contextual approach; The Straussian approach.

- Bevir, Mark, 2011 'The Contextual Approach', in George Klosko (ed.), *The Oxford Handbook of the History of Political Philosophy.* Oxford Academic.
- Zuckert, Catherine, 2011 'The Straussian Approach', in George Klosko (ed.), *The Oxford Handbook of the History of Political Philosophy.* Oxford Academic.

Week 3: Interpreting the Constitution
Textualism; Semantic Originalism; Original Intentions; Original Public Meaning; Moral Readings; Pragmatism; etc.

- Sunstein, Cass R. 2023. *How to Interpret the Constitution.* Princeton, New Jersey: Princeton University Press.

Week 4: Preparing for Text-as-Data
Python Basics; Google Colab.

- Watch a python tutorial here: [Link]

Week 5: Introduction to Natural Language Processing (NLP)
Token; Vocabulary; Document; Corpus; Tokenization; POS tagging; Syntactic parsing; Name-entity recognition

- Watch a python tutorial here: [Link]

Week 6: Word Representation
One-hot-encoding; Embedding; PMI; SVD; Word2vec; GloVe

- Wilkerson, John, and Andreu Casas. 2017. "Large-scale Computerized Text Analysis in Political Science: Opportunities and Challenges." *Annual Review of Political Science* 20: 529-544.

Week 7: Document Representation
Bag-of-words; Word clouds; TF-IDF.

- Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing for Unsupervised Learning: Why It Matters, When It Misleads, and What To Do about It." *Political Analysis* 26, no. 2: 168-189.

Week 8: Midterm Exam

Week 9: Topic Modeling
Unsupervised machine learning; K-means clustering; Latent dirichlet allocation (LDA).

- Pan, Jennifer, and Kaiping Chen. 2018. "Concealing Corruption: How Chinese Officials Distort Upward Reporting of Online Grievances." *American Political Science Review* 112.3: 602-620.

Week 10: Introduction to Neural Networks

MLP; Hidden layers; Activation function; Loss function; Stochastic Gradient Descent; Vanishing gradients; Adam

- Watch a video on NN here: [Link]

Week 11: Neural Networks for NLP

MLP; RNN; LSTM; Bidirectional LSTM; Encoder; Decoder; Attention; Self-attention; Transformer.

- Vaswani, Ashish et al. 2017. "Attention is All You Need." *Advances in Neural Information Processing Systems* 30.

Week 12: Measuring Sentiment

Sentiment analysis

- Osmundsen, Mathias et al. "Partisan Polarization is the Primary Psychological Motivation Behind Political Fake News Sharing on Twitter." *American Political Science Review* 115, no. 3 (2021): 999-1015.

Week 13: Measuring Populism

Measurement theory; Validity; Reliability

- Dai, Yaoyao, and Alexander Kustov. 2022. "When Do Politicians Use Populist Rhetoric? Populism as a Campaign Gamble." *Political Communication* 39.3: 383-404.

Week 14: The Authorship Dispute of *The Federalist Papers*

- Breitzman, Anthony. 2023. "An Accessible Python based Author Identification Process" *SciPy Proceedings.* DOI: https://doi.org/10.25080/gerudo-f2bc6f59-003.

Week 15: Student Presentations

Week 16: Final Exam