Text As Treatment and Outcome: Applications of Word Embeddings in Ash et al. (2021) and Ash et al. (2022)

Kazuki Motohashi

April 21, 2022

Tufts University

Roadmap

Introduction

Text As Treatment: Ash et al. (2021)

Text As Outcome: Ash et al. (2022)

Text As Treatment and Outcome

- Ash et al. (2021): Text as Treatment (X)
  - Gender attitudes measured based on GloVe word embeddings
  - Examine the impacts of gender attitudes (X) on interactions with female judge
- Ash et al. (2022): Text as Outcome (Y)
  - Economic style in judicial language measured based on Word2Vec word embeddings
  - Examine the impacts of law-and-economics training on the economic language in judges' opinions (Y)

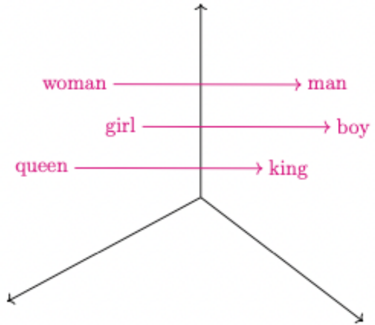
Word Embeddings

Words are distributed in a vector space based on their co-occurrence in a corpus.

- Low dimension dense representations of words ($\leftrightarrow$ one-hot-encoded vectors)
- Positions of word vectors in the space encode relations between words
- Non-contextualized representations ($\leftrightarrow$ Transformer)

Word Embeddings

Figure 1: Identifying a Gender Dimension in a Vector Space



Source: Ash et al. (2021)

Models of Word Embeddings

1. Word2Vec (Mikolov et al., 2013)
    - Continuous Bag-Of-Words: Predict the center word given the surrounding context words
    - Ski-Gram: Predict the context word based on the center word
2. GloVe (Pennington et al., 2014)
    - Learn word embeddings such that the word vectors' dot product equals the log of counts in the co-occurrence

Common Task: Computation of Similarity

1. Document Similarity
   - Judges' opinions v.s. Law and Economics dictionary (Ash et al., 2022)
   - Faculty members' publications vs University leaders' publications (Acemoglu et al., 2021)

2. Slant measures based on similarity of word embeddings within same documents (Ash et al., 2021)

Steps to Compute Document Similarity

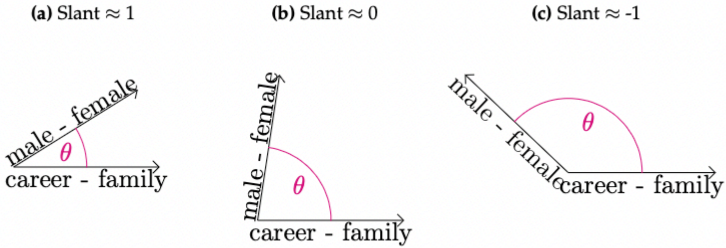
1. Text Vectorization $\rightarrow$ dimensionality reduction
   - TF - IDF (Term Frequency - Inverse Document Frequency) (Acemoglu et al., 2021)
   - LDA (Latent Dirichlet Allocation)
   - Word Embeddings (e.g. Word2Vec, GloVe) (Ash et al., 2022)
   - Document Embeddings (e.g. Doc2Vec)
   - Transformer (e.g. BERT), etc.

2. Compute similarity between generated vectors
   - Euclidean distance $\rightarrow$ Sensitive to the length of text
   - Cosine similarity (Acemoglu et al., 2021; Ash et al., 2022)

Cosine Similarity

$$\text{sim}(\vec{x}, \vec{y}) = \cos(\theta) = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \tag{1}$$

Figure 2: Measuring Gender Attitudes using Cosine Similarity



Source: Ash et al. (2021)

Roadmap

Introduction

Text As Treatment: Ash et al. (2021)

Text As Outcome: Ash et al. (2022)

Summary of Ash et al. (2021)

1. Research questions
    - Do gender attitudes influence interactions with female judges in U.S. Circuit court?
2. Methodology
    - Simple OLS based on two features
        - quasi-random assignment of judges to cases
        - control for judges' characteristics (gender, ideology, etc.)
3. Results
    - More slanted judges are more likely to vote to reverse lower-court (district court) decisions authored by female district judges.
    - Assigning judges with a higher gender slant are less likely to assign opinion authorship to a female judge.
    - More slanted judges are also less likely to cite the opinions of female judges.

Text As Treatment in Ash et al. (2021)

- Gender attitude (slant) based on the **GloVe word embeddings** of judge's authored opinions
- Cosine similarity between the gender dimension ($\overrightarrow{male} - \overrightarrow{\text{female}}$) and stereotypical dimension ($\overrightarrow{career} - \overrightarrow{\text{family}}$)
    - High $\rightarrow$ stereotyped language
    - Low $\rightarrow$ non-stereotyped language

GloVe

1. Compute global co-occurence matrix
   - Reports the number of times two words ave occurred within a given context window (X_{ij})

2. Minimize the following objective function to obtain word embeddings $\boldsymbol{w}$:

$$J(\boldsymbol{w}) = \sum_{i,j} f\left(X_{ij}\right)\left(w_i^T w_j - \log\left(X_{ij}\right)\right)^2 \tag{2}$$

3. Two key hyperparameters:
   - Dimension of word vectors
   - Window size for computing co-occurence matrix

Ref: Co-occurence matrix

Corpus: [I like deep learning. I like NLP. I enjoy flying.]

Window length: 1

counts	I	like	enjoy	deep	learning	NLP	flying	.
I	0	2	1	0	0	0	0	0
like	2	0	0	1	0	1	0	0
enjoy	1	0	0	0	0	0	1	0
deep	0	1	0	0	1	0	0	0
learning	0	0	0	1	0	0	0	1
NLP	0	1	0	0	0	0	0	1
flying	0	0	1	0	0	0	0	1
.	0	0	0	0	1	1	1	0

Source: Stanford CS224d Lecture 2

Calculating Vector Differences

1. Select word sets to identify the gender and career-family dimensions
   - Obtain word sets from Linguistic Inquiry and Word Count Dictionaries for the concepts of male, female, work, and family
   - Eliminate words that could be ambiguous or have specific legal meanings
   - Select 10 most frequent words in the full judicial corpus

2. Calculate vector differences based on selected word sets

$$\overrightarrow{\text{male}} - \overrightarrow{\text{female}} = \frac{\sum_n \overrightarrow{\text{maleword}}_n}{|N_{\text{male}}|} - \frac{\sum_n \overrightarrow{\text{femaleword}}_n}{|N_{\text{female}}|} \quad (3)$$

Selected Word Sets in Ash et al. (2021)

Table 1: Word Sets

Male	his, he, him, mr, himself, man, men, king, male, fellow
Female	her, she, ms, women, woman, female, herself, girl, girls, queen
Career	company, inc, work, business, service, pay, corp, employee, employment, benefits
Family	family, wife, husband, mother, father, parents, son, brother, parent, brothers

Source: Ash et al. (2021)

Measuring Gender Attitudes

1. Obtain the universe of published 380,000 opinions in thirteen circuit courts from Bloomberg Law for the years 1890-2003
2. Preprocessing data
   - Exclude punctuation, numbers
   - Transform all words to be lower cased
   - Retain only the most common 50,000 words in all corpus
   - Opinions separated into sentences and tokenized into words
3. Train GloVe embeddings on each judge's corpus
   - To address the issue of small corpora and achieve stable results, use median measures after training the 25 bootstrap samples
   - Time-varying measures cannot be constructed due to this small corpora
4. Validation of Measures
   - Human evaluation: law students evaluated empathy toward women on two randomly paired opinions
   - Count-based measures

Roadmap

Introduction

Text As Treatment: Ash et al. (2021)

Text As Outcome: Ash et al. (2022)

Summary of Ash et al. (2022)

1. Research questions
    - The Effects of the early law-and-economics movement on the U.S. judiciary
2. Methodology
    - Difference in difference specification based on staggered attendance of judges in the law-and-economics training
3. Results
    - After attending economics training, participating judges
        - use more economics language in their opinions (NLP part)
        - issue more conservative decisions in economics-related cases
        - rule against regulatory agencies more often
        - favor more lax enforcement in antitrust cases
        - impose more/longer criminal sentences

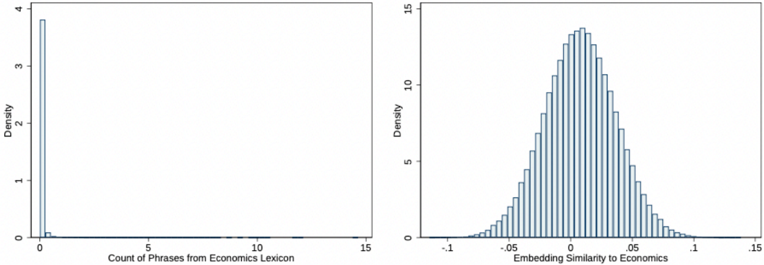
Text As Outcome in Ash et al. (2022)

- Economics style measured based on the **Word2Vec word embeddings** of judge's authored opinions
- Compute cosine similarity between the judge's opinions and law-and-economics phrased used by Ellickson (2000)
  - externality, externalities, transaction, transactions, cost, costs, efficient, efficiency, deterrence, benefit, benefits, capital, market, markets, marketplace, economic, economics

Count-based Measures vs Embedding-Based Measures

Rare occurence of law-and-economics phrases in opinions →
Cosine similarity based on word embeddings (right plot) is better

Figure A.5: Distributions of Count-Based and Embedding-Based Econ Language Measures



Notes. Histograms by case of the number of words in a case from the Ellickson lexicon (left graph), vs the embedding-based economics language similarity measure (right graph).

Source: Ash et al. (2022)

Words Correlated with Law-and-Economics Lexicon

Panel (a): Cosine similarity close to 1

Panel (b): Cosine similarity close to -1

Figure A.4: Words Correlated with Law-and-Economics Lexicon Dimension

(a) Positively Associated Words

(a) Negatively Associated Words

Notes. The left word cloud lists the set of words that have the highest cosine similarity to the average word vector for Ellickson phrases in the word embedding space. The right word cloud gives the words that have the lowest (most negative) cosine similarity to this vector.

Source: Ash et al. (2022)

How to Compute Similarity between Documents in Practice

1. Word Embeddings are computed for all tokens (words) of two documents.

2. Aggregate these word embeddings to document-level embeddings by
   - taking an element-wise average, minimum or maximum
   - fancier technique such as smooth inverse frequency (SIF)

3. Compute the cosine similarity between document-level embeddings

Ref)
https://aajanki.github.io/fi-sentence-embeddings-eval/models.html

Alternative measure: Supervised learning approach

1. Obtain corpus metadata on labels for whether it is an economics-related case (regulation or labor)

2. Use an L2-penalized logistic regression to predict this label on the text features based on Arora et al. (2017)

3. Apply the training model to the full corpus to obtain the text-predicted probability that a case is on an economics topic

4. This probability in the non-economics-related cases can be thought of as a measure of how much economics language was used
   $\rightarrow$ Similar empirical results when this probability is used as an outcome instead of the embedding-based similarity measure

Thank you!

Kazuki.Motohashi@tufts.edu

Acemoglu, Daron, David Y Yang, and Jie Zhou (2021) "Political Pressure and the Direction of Research: Evidence from Chinas Academia."

Arora, Sanjeev, Yingyu Liang, and Tengyu Ma (2017) "A simple but tough-to-beat baseline for sentence embeddings," in *International conference on learning representations*.

Ash, Elliott, Daniel L Chen, and Suresh Naidu (2022) "Ideas have consequences: The impact of law and economics on american justice," National Bureau of Economic Research (NBER) Working Paper 29788.

Ash, Elliott, Daniel L Chen, and Arianna Ornaghi (2021) "Gender attitudes in the judiciary: Evidence from us circuit courts," *Center for Law & Economics Working Paper Series*, Vol. 2019, No. 02.

Ellickson, Robert C (2000) "Trends in legal scholarship: A statistical study," *The Journal of Legal Studies*, Vol. 29, No. S1, pp. 517–543.

References ii

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean (2013) "Distributed representations of words and phrases and their compositionality," *Advances in neural information processing systems*, Vol. 26.

Pennington, Jeffrey, Richard Socher, and Christopher D Manning (2014) "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543.