

150B/355B  
Introduction to Machine Learning for Social Science  
TA Section 9

Haemin Jee and Tongtong Zhang

March 9, 2018

# Road Map

- 1 Review key terms
- 2 Questions for final

# Key Concepts and Terms

## Dictionary Methods

### Distinctive words

- Difference in frequencies
- Difference in rates
- Standardized mean difference

### Distance

- Euclidian
- Cosine similarity

## K-means clustering

- Centroid
- Inputs and outputs
- Single membership model
- Interpretation

## Topic Modeling

- Inputs and outputs
- Mixed membership model
- Interpretation

# Review Questions

- 1 How do you turn raw text into a DTM? (give example question here)
- 2 How would you apply a dictionary method to a DTM?
- 3 What are different metrics of distinctive words? What are some pitfalls of the different metrics?
- 4 How do you find distance between two documents?
- 5 Describe a k-means algorithm. What is a centroid?
- 6 Why might two researchers, using the same raw text data, get different results when running k-means?
- 7 What are the inputs and outputs of a topic model?
- 8 For both k-means and topic models, how do we interpret the results that we get?

How do you turn raw text into a DTM?

- “Thank you to all the incredible women paving the way. Happy International Women’s Day. H”
- “Happy International Women’s Day to every single woman in the world”
- “happy international women’s day to my bad ass girls!!!”
- “Celebrate International Women’s Day by explaining a man’s joke to him”

# Review Questions

How would you apply a dictionary method to a DTM?

- Positive words (score of 1): incredible, happy, celebrate
- Positive words (score of 0.5): joke, thank, world
- Negative words (score of -0.5): single, paving
- Negative words (score of -1): ass, bad
- Neutral words (score of 0): international, every, explain, man, women

Give each of the tweets a score based on the sentiment dictionary above.

# Review Questions

How would you apply a dictionary method to a DTM?

- Positive words (score of 1): incredible, happy, celebrate
- Positive words (score of 0.5): joke, thank, world
- Negative words (score of -0.5): single, paving
- Negative words (score of -1): ass, bad
- Neutral words (score of 0): international, every, explain, man, women

Give each of the tweets a score based on the sentiment dictionary above.

**A:**

- Tweet 1:  $0.5 + 1 - 0.5 + 1 = 2$
- Tweet 2:  $1 - 0.5 = 0.5$
- Tweet 3:  $1 - 1 - 1 = -2$
- Tweet 4:  $1 + 0.5 = 1.5$

# Review Questions

What are different metrics of distinctive words? What are some pitfalls of the metrics?



What are different metrics of distinctive words? What are some pitfalls of the metrics?

- Difference in frequencies
  - Does not take into account difference in total number of words used
- Difference in rates (compare the average rate each author/topic uses a word)
  - Favors more frequent words
- Standardized Mean Difference

# Review Questions

How do you find distance between two documents?

# Review Questions

How do you find distance between two documents?

**A:** Euclidian distance, cosine similarity

What is one advantage of using cosine similarity?

# Review Questions

Describe a k-means algorithm. What is a centroid?

# Review Questions

Describe a k-means algorithm. What is a centroid?

**A:** Start with random means, divide documents into clusters (based on Euclidian distance). Re-calculate the means, and then again divide documents into clusters. Continue until the partitions stabilize.

# Review Questions

Describe a k-means algorithm. What is a centroid?

**A:** Start with random means, divide documents into clusters (based on Euclidian distance). Re-calculate the means, and then again divide documents into clusters. Continue until the partitions stabilize.

**A:** Centroid: The “exemplar” document for that particular cluster – a vector representing the average values of all observations in that cluster.

# Review Questions

Why might two researchers, using the same raw text data, get different results when running k-means?

Why might two researchers, using the same raw text data, get different results when running k-means?

**A:** 1) Different starting values; 2) Different ways of pre-processing the data; 3) Different number of clusters



# Review Questions

What are the inputs and outputs of a topic model?

# Review Questions

What are the inputs and outputs of a topic model?

**A:** Inputs: DTM, desired number of topics. Outputs: Topic distribution over words; Document distribution over topics

# Review Questions

For both k-means and topic models, how do we interpret the results that we get?

# Review Questions

For both k-means and topic models, how do we interpret the results that we get?

**A:** We can 1) find distinctive words of each cluster or topic; 2) Read a selected number of documents in each cluster or topic.

**Q:** How does unsupervised learning differ from supervised learning?

**Q:** How does unsupervised learning differ from supervised learning?

In **supervised** learning, we know the labels or classes that we are interested in and the goal is to assign each document into those pre-defined groups. We can also validate our results by comparing it to human coding.

**Q:** How does unsupervised learning differ from supervised learning?

In **supervised** learning, we know the labels or classes that we are interested in and the goal is to assign each document into those pre-defined groups. We can also validate our results by comparing it to human coding.

In **unsupervised** learning, we **do not** know what the interesting or meaningful classes are. We find the meaningful groups while assigning things to those groups at the same time. There is no clear way to check the quality of our results.