APS360 Fundamentals of AI

Lisa Zhang

Lecture 10; Feb 10, 2019

Agenda

- Fun with GloVe vectors
- Sentiment analysis

GloVe

Structure (word analogies)

https://lamyiowce.github.io/word2viz/

Distance Measures

- Euclidean distance
- Cosine distance

Sentiment Analysis

Goal

Given a piece of text, identify the sentiment of the author.

Can be applied to:

- movie reviews
- feedback
- emails
- tweets
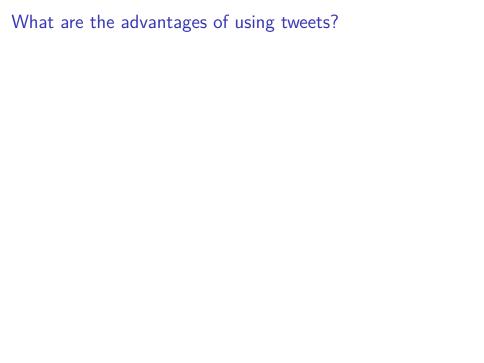
Challenges

- Difficult problem in general
- Hard to collect very clean data

Sentiment 140

- 160,000 tweets
- Sentiment determined by emoticon

Collected by students doing a course project

What are the advantages of using tweets?

What are the advantages of using tweets?

- Short
- Publicly available

What are the challenges with using tweets?

What are the challenges with using tweets?

- Spelling
- Sarcasm
- Punctuations

Code!

Let's write some code!