# Stylometry

Deepika(CSE)  Pujitha Jaji(ECE)  Bhavishya(IT)

June 23, 2017

# What is Stylometry?

- Study of the writing style of an author.
- Legal, literary, academic applications.
- Used to attribute authorship to anonymous or disputed documents.

# Steps invoved in our project

- Data collection.
- Analysis of authors style.
- Comparision of authors styles.
- Prediction of the authorship of given unknown text.

# Analysis of authors style

- Average sentence length
- Word endings
- Bi-grams
- Filtering of stopwords

# What we've learnt

- nltk
- collections
- LaTeX

# Statistics

- Written code for author-analysis - 81 lines
- Written code for author-prediction - 95 lines
- Functions Used
- nltk.tokenize - word-tokenize, sent-tokenize
- nltk.corpus - PlaintextCorpusReader, stopwords