UNIVERSITI
MALAYA

#NLP Applications

Detecting Hate speech refers to detecting a kind of speech that denigrates a person or multiple persons based on their membership to a group, usually defined by race, ethnicity, sexual orientation, gender identity, disability, religion, political affiliation, or views.

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6864473/
- https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0221152#sec003

- Regex: a fast string processor, use in chunking
- NLTK：Tokenization(sentence tokenization and word tokenization)
- split(): Segmentation
- wordnet: stemming&lemmatization
- spaCy
- Stanza