CS410 - Text Analysis
Satyo Iswara (iswara2)

Technology Review - State of the Art of Medical Text Analysis

- massive data (example #)
- challenges:
    - scalability
    - accuracy
    - unstructured data
    - resources (time + man power)
- state of the art
    - named entity recognition (NER)
        - automating medical term identification
        - how?
            - detection & determination of entities
            - selection & extraction of entities
        - example
    - hypothesis generation/knowledge discovery
        - using natural language processing
        - example
            - gathering of data from clinical records + clinical data
    - text classification
        - categorization
        - example
            - TODO example
    - text summarization
        - topic identification + presentation
        - simplification of data
        - faster user consumption
    - text classification
- example of commercial use
    - Quertle
    - medbioinformatics