FINAL PROJECT

David Kauchak
CS159 Spring 2023

1

Admin

Assignment 7 due next Monday

Will update schedule for the rest of the semester by Wed

Grading

2

Final project

1. Your project should relate to something involving NLP

2. Your project must include a solid experimental evaluation

3. Your project must have an ethical component

4. Your project should be in a group of 2-4. If you'd like to do it solo, please come talk to me.
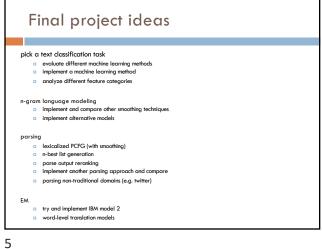
3

Final project

date	time	description
4/17	in-class	Project proposal presentation
4/19	11:59pm	Project proposal write-up
4/26	11:59pm	Status report
5/1	11:59pm	Paper draft
5/3	in-class	Presentation
5/3	11:59pm	Final paper and code

Read the final project handout ASAP!

Start forming groups and thinking about what you want to do

4

Final project ideas

pick a text classification task
- evaluate different machine learning methods
- implement a machine learning method
- analyze different feature categories

n-gram language modeling
- implement and compare other smoothing techniques
- implement alternative models

parsing
- lexicalized PCFG (with smoothing)
- n-best list generation
- parse output reranking
- implement another parsing approach and compare
- parsing non-traditional domains (e.g. twitter)

EM
- try and implement IBM model 2
- word-level translation models

5

Final project application areas

- spelling correction
- part of speech tagger
- text chunker
- dialogue generation
- pronoun resolution
- compare word similarity measures (more than the ones we looked at)
- word sense disambiguation
- machine translation
- information retrieval/search
- information extraction
- question answering
- summarization
- speech recognition

6

Final project application areas

- Ethical analysis of: Word2Vec, GPT-X, ChatGPT
- Utilize LLM (large language model), e.g., GPT
  - poety
  - lyrics
  - code generation
  - …
- Limited resource NLP (e.g., limited memory or computation)
- BART/BERT

7