CSE 842: NATURAL LANGUAGE PROCESSING
GROUP PROJECT

CSE 842 will end with a semester-long group project (instead of a final exam). Groups should consist of three to four students. The project consists of four components: proposal, short paper, long paper, group presentation.

PROJECT PROPOSAL
DUE: SEPT. 25

Groups will propose the research task they aim to solve. The written proposal should include:

- An outline of what your group is planning to do for this project. Include descriptions of the dataset(s) and model(s) you are planning to use, as well as how you tentatively plan to evaluate your results.

- Motivate why this is an interesting NLP problem worth exploring. Will it contribute a better model? A better modeling pipeline? Help interpret a socially important phenomenon? Predict some interesting real-world events?

- Include a literature review to set the project in context. What has been done before? How will your project be different? How will it improve on the current work?

- Bios of the researchers! Include your name, NetID, Master's or PhD, department, and year in the program. Also list what each team member will be responsible for or how work will be divided evenly.

- This initial proposal should be roughly similar to the Abstract, Introduction, Related Works, and Methodology sections of a typical conference paper submission.

- Length: 2 pages, ACL format

SHORT PAPER (MIDTERM REPORT)
DUE: OCT. 23

Groups will submit initial experimental results and describe their planned experimental analysis. This report should add any new related works, modify the abstract and introduction as needed, and expand on the methodology section by adding more details and initial experimental results.
Length: 4 pages; similar to a conference short paper.

LONG PAPER (FINAL REPORT) & PRESENTATION SLIDES
DUE: NOV. 27

Groups will submit their completed project. The final version of your paper should include a description of everything done for this project: literature comparison, data collection, annotation, processing, model and/or pipeline development and technical details, thorough experimental analysis, future work, and conclusion.
Length: 8 pages; similar to a conference long paper.

GROUP PRESENTATIONS
DEC. 2 & DEC. 4

Each group will present their findings through a Powerpoint presentation. The length of time for each presentation will be announced after I have a count of the number of groups.

LOOKING FOR IDEAS?

Check out past work at the top NLP conferences: ACL NAACL COLING EMNLP
NLP researchers also publish here: AAAI ICWSM IJCAI
SemEval releases trending tasks each year.

Some example categories that conferences look for research in:
- Cognitive Modeling and Psycholinguistics
- Computational Social Science and Social Media
- Dialogue and Interactive Systems
- Discourse and Pragmatics
- Ethics and NLP
- Generation and Summarization
- Information Extraction, Information Retrieval, Text Mining
- Language Grounding to Vision, Robotics and Beyond
- Machine Learning for NLP
- Machine Translation
- Phonology, Morphology and Word Segmentation
- Semantics: Lexical, Sentence Level, Textual Inference
- Sentiment Analysis, Stylistic Analysis, and Argument Mining
- Speech and Multimodality
- Syntax: Tagging, Chunking and Parsing
- Question Answering