Final Project (CS 2731 Fall 2023)

A major component of this course is a hands-on final project guided by students' own interests. In this project, students will demonstrate an ability to summarize current approaches and challenges in a subfield of NLP and implement some sort of contribution (however small) to this NLP area of research or practice.

Groups

Projects will be done in groups of 2-4 students. Groups will be assigned by the instructor and TA based on interests and skills, and group preferences from students, in a survey.

Deliverables

1. **Project survey.** *Due 08-31.* This survey asks about NLP research areas of interest, skills, project ideas, and any group preferences. The survey will be available as a Google Form. It will be used by the instructor and TA to match groups with similar project interests and complimentary abilities.

2. **Project area and type of contribution.** *Due 09-21.* You will provide the type of contribution you are interested in pursuing within a specific research area, as well as any inkling of idea you have, in a form. The instructor and TA will provide some feedback and guidance on this direction in meetings (see Canvas for available time slots).

3. **Project proposal and literature review.** *Due 10-12.* Please submit one per group on Canvas. This proposal will be a report with answers to the following questions:

   1. What type of contribution are you making?
   2. What is the problem or task you are focusing on?
   3. What is the nature of your contribution? That is, what is the expected output of your project? This could be a new approach and its evaluation, a new dataset, or new analysis.
   4. How does your contribution build on or extend prior work? This literature review will be of at least 4 papers relevant to your project area. It will group and summarize relevant papers into types of tasks, datasets, and/or approaches. Good places to look for NLP papers include the ACL Anthology, Semantic Scholar, and Google Scholar.
   5. What data are you using (or contributing)?
   6. What algorithm or approach are you taking to address the task?
   7. How are you evaluating your contribution? What performance metrics are you going to use?
   8. What kinds of ethical issues may be raised by your model or data?
   9. What are the proposed steps needed for completion of the project?
   10. What are roles and tasks of each person in the group? Though group members will contribute in various capacities, it is best if each person is responsible for at least one aspect of the project.

   It is recommended to start on the literature review early on, since existing work will inform your specific direction (somebody may have already tried your idea!). The instructor and TA will provide feedback on this proposal and a meeting in office hours if needed. There is no required length or format for this report, but you could use the ACL format that the final report will be formatted in. Proposals will be submitted through Canvas (one per group is fine).

4. **Project proposal presentation.** *In class 10-18.* Groups will make a brief presentation to the class outlining their proposed project, with Q&A and opportunities for feedback from other students. Please plan for 4-minute presentations with 2 minutes for questions for each group. Cover at least these key points:

   1. Project motivation (what is the value of this work?)
   2. Briefly, what 1-2 other related papers have done
   3. What data you are planning to use
   4. What approach/methods will you be taking
   5. Evaluation of your approach (or dataset, if it's a dataset contribution)

   Have each group member speak in the presentation. Please add your slides to a Google Slide presentation which you can find in a Canvas announcement. These presentations are not graded.

5. **Basic working system.** *Due 11-16.* A brief (1-2 page) progress report of a basic working system. Not everything needs to be done or fully functional, but there needs to be some sort of basic functionality. Also list any questions you have or resources you will need to successfully complete the project by the 12-14 deadline. This report should be in the ACL format that the final report will be in.

6. **Final presentation.** *In class on 12-13.* Groups will present their finished work to the group, with Q&A and feedback opportunities from students. Please prepare a maximum 5-minute presentation in which you can divide up speaking responsibilities however you see fit (not all members need to speak, and it is okay for one group member to present the whole presentation). Cover at least these key points:

   1. Project motivation (briefly)
   2. Data
   3. Methods, or annotation/collection approach for dataset projects
   4. Results

   Please add your slides to a Google Slide presentation which you can find in a Canvas announcement. These presentations are not graded.

7. **Final report and code.** *Due 12-14.* At the end of the course, groups will provide a written report of their project. This report will be in the ACL format found here (Overleaf template here). Include a section detailing the high-level tasks that each group member did. The report should be a maximum of 8 pages, not including references or the group member task breakdown. Additional figures or explanation in an appendix is allowed, but they will not necessarily be considered in grading. Outstanding reports would be of a quality and structure that could be submitted to an NLP workshop or conference, but other types of projects can also achieve an A. Here is the rubric that will be used in grading:

Rubric category	Points
Clear motivation for the work is provided	5
Research questions and/or task definition is clear	10
Sufficient grounding in relevant related literature	15
Applicable dataset/s are chosen	5
Methods are relevant. For new approach contributions, multiple methods are compared. For dataset contributions, annotation methodology is explained	15
Results are provided. For new approach contributions, results from multiple methods (at least one baseline) are presented. For dataset contributions, this may be a single set of results from a simple classifier, or other results if discussed with the instructor	20
Discussion is provided of the results and/or the potential uses or contributions of any new datasets contributed	10
Limitations of your approach or dataset are sufficiently discussed	5
Ethical issues that may be raised by your system or dataset are sufficiently discussed	5
Project content total	**80**
Meets all formatting requirements. Is maximum 8 pages, not including references or group member task breakdown	15
Writing is clear	15
Writing total	**30**
Group member had a sufficient amount of workload in the project	15
Task and roles assigned to this group member were completed sufficiently	15
Individual contribution total	**30**
Grand total	**150**

Types of contributions

Your goal is to make a contribution, even a small one, to NLP research or practice. You can select from the following types of contributions, combine multiple of them, or define a different type of contribution with instructor approval. Example project ideas and projects are provided (with a significant bias toward computational social science and hate speech, the instructor's research area). Groups are also encouraged to come up with their own ideas! Projects can be related to students' research, but should not be projects for other classes.

1. New dataset, annotations, or analysis of existing datasets

Data is at the heart of machine learning and NLP systems; it enables further modeling and encapsulates what NLP systems "know".

Example project ideas

- Hate speech is often expressed in a "joking" manner, yet hate speech classifiers struggle to identify hate speech in irony and humor. Build a dataset of hate speech in jokes ("light talk") to aid future NLP classifiers. This dataset could be collected from sites that are known for both hateful content and a "humorous" tone, and portions could be annotated for hate speech and/or types of humor. This could be an expansion of prior work that provides a dataset of offense in humor: https://competitions.codalab.org/competitions/27446
- Hate speech is culturally specific, yet the majority of NLP work focuses on English in North American and European contexts. A quantitative analysis of different features of datasets annotated for hate speech in multiple languages and from multiple cultural contexts would illuminate global similarities and culturally specific contexts.
- Speech acts are the "actions" that speakers intend with utterances (such as asking a question or giving an explanation). Emojis may help automated systems determine speech acts. Build a dataset of utterances with and without emojis and annotate these for speech acts. Build a simple classifier with interpretable features and see if emojis are informative.
- Build a dataset of text from different personas from fiction roleplaying sites ("language cosplaying" in China). This could be useful for dialogue systems that adopt personas, or for story generation.

Example projects

- Newell et al. 2017: "Verifiability" of claims made in news. New datasets often accompany the introduction of a new NLP task, which could be a previously untackled phenomenon like verifiability in this one.

2. New approach or application

This is perhaps the most common sort of NLP research contribution, in which a new method or algorithm for approaching a task (which could be a new task) is presented. Applying an existing method in a new context or task (as might be necessary in an industry setting) would also fit within this contribution.

Example project ideas

New tasks and applications:

- Automatically identify the targeted group of hate speech in English from individual posts. This may seem "easy" by simply recognizing identity terms in the text, but there are many difficult cases where multiple identities are mentioned, or references to groups are ambiguous or indirect. Hate speech annotated for targeted identities would be provided by the instructor and could act as training data. This classifier could be applied to audit the distribution of targeted identities in commonly used hate speech datasets.
- New identity terms are commonly developed in online communities, some of them hateful. Develop methods to find in-group hate jargon and identity terms.
- Hate speech identification without the text: Identify the discourse contexts in which hate speech is likely to occur, without allowing classifiers to look at the exact text of the hate speech.
- Stancetaking, a concept from sociolinguistics, is when speakers take an evaluative position toward the concept ("No, I actually don't like Taylor Swift's music that much, but she's great"). Develop automated methods for identifying the "stance object", who the speaker is evaluating, likely from Reddit data that would be provided by the instructor.
- From a set of descriptions of characters, develop a classifier for who are the most fanfiction? This could be a lens into online community and media norms.
- Examine the framing of different entities in police Facebook posts from the Plain View Project.

Existing tasks (some ideas are pulled from Graham Neubig's Advanced NLP class):

- Any of the SemEval 2021 tasks
- X-FACTR multilingual knowledge probing in QA
- GoEmotions Fine-grained Emotion Detection Dataset
- SciREX Scientific Information Extraction
- Subjective Intent Classification in Discourse
- Very Low Resource machine translation
- Predict the points at which speakers switch languages when code-switching
- Sign language translation
- Develop best approaches for training hate speech classifiers that generalize across targeted identities. Data would be provided by the instructor
- Style transfer of offensive language into inoffensive language. See existing paper and dataset+code.
- Develop new approaches for hate speech detection by comparing with knowledge bases or pretrained models of stereotypes from a stereotype dataset
- Cross-lingual emotion detection. See existing paper.
- Active learning for NLP. See existing paper and code.
- Multilingual content moderation on Reddit. See existing paper and dataset.

3. New evaluation

Good automated evaluations of machine learning systems are hard to come by. Ideally they correlate with human judgments of quality and capture key elements of the phenomenon being modeled while also being robust to adversarial attacks. This may be a difficult contribution to make without being very familiar with a research area, but there is often a need for new evaluations for specific NLP applications.

Example projects

- Moosavi and Strube 2016. New coreference resolution metric.
- Freitag et al. 2022. Neural evaluation metrics for machine translation.

4. New survey or position paper

Surveys are especially needed for new, emerging research areas. All projects will require a literature review, but a survey paper would be both broader and go much more in depth. It would summarize key approaches and key challenges and present lines for future work. Some sort of implementation is necessary for this type of contribution as well, such as applying multiple established methods to a new dataset or in a new context to show challenges that need to be addressed. Position papers argue for a certain viewpoint or shortcoming of existing approaches, e.g. arguing for the utility of techniques from a discipline outside NLP in NLP tasks.

Example project ideas

- Survey how NLP is used and applied in other fields. What has been our most useful contributions to writers in the social sciences, physical sciences, or humanities? This survey would assemble papers across disciplines for mentions of NLP and summarize what is most useful, what is lacking, and what approaches from NLP could be helpful to others.
- Computational social science using NLP generally relies on data from online communities. But this is missing non-online interactions and the practices of those who are not active online. Survey datasets and approaches that use quantitative and computational techniques on recordings of offline linguistic interaction.
- A growing area of research in computational social science aims to capture the framing and portrayal of entities across large text corpora (such as in news media). Survey existing approaches and challenges.

NLP research areas

Early on, groups will select an NLP research area to focus their project on (or multiple areas, if applicable). Here is a list of NLP research areas in alphabetical order, drawn from the ACL Rolling Review list. You can also pursue a project in a different area with instructor permission.

- *Computational Social Science and Cultural Analytics.* Using NLP to study societies and cultural texts, as well as developing tools that process language in social context. This is the instructor's main research area and many of the example project ideas focusing on hate speech would fit within this area.
- *Dialogue and Interactive Systems.* NLP systems that interact conversationally, as well as systems that analyze human conversation.
- *Discourse and Pragmatics.* Discourse here refers to linguistic structure beyond the a sentence level. This includes predicting the coherence of different sentences together (discourse coherence) and the relations between sentences (discourse relations). It also includes parsing rhetorical structures of texts such as essays.
- *Efficient/Low-Resource Methods for NLP.* This includes computational and energy efficiency. How do we get similar performance with models with fewer parameters and less demanding computational requirements?
- *Ethics, Bias, and Fairness.* There's a lot of focus on this area with LLMs right now.
- *Generation.* This includes language and code generation, as well as interesting applications like story and narrative generation.
- *Information Extraction.* This includes extracting and linking events and particular types of entities, timelines, and locations.
- *Information Retrieval and Text Mining.* This includes searching for particular information from large text corpora.
- *Interpretability and Analysis of Models for NLP.* Many NLP and ML models are quite uninterpretable; this area addresses this.
- *Linguistic theories, Cognitive Modeling and Psycholinguistics.* This area includes work that applies or evaluates linguistic and cognitive theories of language with computational models.
- *Machine Learning for NLP.* This area includes core ML improvements for NLP, as well as applying new ML methods to language.
- *Machine Translation.* Automatic translation between languages.
- *Multilinguality and Language Diversity.* Too much NLP work is solely in English or other "high-resource" languages. This area addresses this problem and also handles language diversity such as code-switching.
- *Multimodality and Language Grounding to Vision, Robotics and Beyond.* Language is often used by humans in tandem with vision and motion. NLP models within this area take this into account.
- *NLP Applications.* This area would include applications such as educational applications (like automated essay grading) or in areas such as human rights (see aporaphobia paper) or industry settings.
- *Phonology, Morphology and Word Segmentation.* This area includes linguistic areas that study how sounds and parts of words fit together to produce meaning. Computational approaches from NLP can extract and model this subword information at scale.
- *Question Answering.* Automatically answering questions posed by humans.
- *Semantics: Lexical.* Approaches that model word meanings and semantic relationships between words.
- *Semantics: Sentence-level Semantics, Textual Inference and Other areas.* This area includes tasks of natural language inference and coreference resolution.
- *Sentiment Analysis, Stylistic Analysis, and Argument Mining.* Extracting and analyzing sentiments (pro vs con), styles, and arguments.
- *Speech recognition, text-to-speech and Spoken Language understanding.*
- *Summarization.* Automatic approaches to summarizing texts.
- *Syntax: Tagging, Chunking and Parsing.* Computational approaches to extracting grammatical structure from text, including part-of-speech tagging, dependency parsing, and constituency parsing.

How your project will be graded

To get an A, your group's project should make progress toward an achievable, concrete contribution specified in your project proposal. The project does not necessarily need to be successful in the sense that it outperforms baselines or contributes to our knowledge of a phenomenon. Sometimes ideas don't work, and that's okay. But you need to provide evidence of progress toward that contribution. If you are building a dataset, for example, the dataset needs to be built in some form, even if it is as not as large or as useful as you may have hoped. If you are evaluating a new method for a task, you must have an implementation that tests that method against other baselines, even if it doesn't perform as well as you would have hoped (or you didn't get to evaluate it against all the baselines you wanted to. If you are doing a survey, you must distill a sufficient number of papers in that comprehensively describe a research area, even if you don't end up finding groundbreaking gaps in knowledge that must be addressed. Feel free to take on more risky ideas, but only if you know you'll have something to show for it at the end. The instructor and TA will guide you toward scoping projects that should fulfill this goal in the planning phase through the proposal.