



UNIVERSITY OF SOUTHERN MAINE

Natural Language Processing, Spring 2023, Course Project

Instructor: Behrooz Mansouri (behrooz.mansouri@maine.edu)

Due: Proposal on Session 17 (10 minutes presentation + maximum of 2 pages writing)
Final Report on Exam Date (15 minutes presentation + minimum of 4 pages writing)

Notes for submission:

1. Submit your file(s) with the correct naming as NLP_Project_StudentName(s)
2. Two files should be uploaded, one .zip file having all the codes (the directory is zipped, and it is named codes) and one .pdf file. There would be a penalty for uploading wrong formatted files. Any other formatting will be ignored
3. Codes should be well-structured with comments to run
4. Codes should be available on your GitHub Repo. Failure to have codes publicly available results in a 20% reduction in your grade. Make sure to include the GitHub link in your PDF file.
5. Any assumptions made by students should be explicitly mentioned in the submitted document
6. The final writing should be done in Overleaf and both the source and PDF files should be submitted.

Natural language processing at the University of Southern Maine is a project-based course. Students will work, in groups of 2 or individually, on a project of their choice. In Session 17 (one week after the Spring break), students will present their project proposal.

Upon request, students can work on the topic proposed by the instructor. However, students are strongly encouraged to first do some investigating to find a topic they are interested in. Some possible options are:

- In our class, there would be invited speakers to introduce NLP Tasks, and students are welcome to work on them.
- Students can look at [SemEval](#) and [CLEF](#) tasks and actually participate in these labs.
- Stack Exchange topics were introduced in Session 2. Students can define their own tasks

on new topics. For instance, each question on Stack Exchange has a topic. Students can consider the task of topic generation for a question. Another task is to predict whether a question will receive an answer. Question classification (based on tags on Stack Exchange) is another task. Finally, students can work on comments posted on questions, viewing them as clarifying questions asked about a question.

Upon acceptance of the project proposal, students will continue working on the project and can consult with the instructor about their approach during student hours. There would be a dedicated time in the last week of classes (during class time) to review the progress of the project.

Criteria for accepting a proposal include the following:

- **Relevance:** The project should be relevant to the course objectives and align with the current state-of-the-art in NLP research.
- **Originality:** The project should be innovative and contribute to the existing NLP research in some way.
- **Feasibility:** The project should be feasible within the given time frame and resources available, and the project plan should demonstrate that the students have a clear understanding of the steps involved in completing the project.
- **Clarity of objectives:** The project goals and objectives should be clearly stated and well-defined so that they can be evaluated effectively.
- **Technical proficiency:** The students should demonstrate that they have the technical skills necessary to complete the project, or a plan to acquire these skills if necessary.
- **Real-world impact:** The project should have the potential to have a real-world impact, such as improving the accuracy of NLP models or solving a real-world NLP problem.
- **Ethical considerations:** The project should take into account ethical considerations, such as the privacy and security of data, and the potential impact of the project on society.

Note that students are not allowed to change the topic after the proposal is accepted. The final presentation will be during exam time for this course. Students that work in groups, should both present their contribution to the project.