

COMP316 Course Project

The projects will entail original investigation into any area of application of Natural Language Processing. Possible areas of application include *information extraction*, *information retrieval*, *question answering*, and *machine translation* which we will cover in class. You are free to do your project on any area of application, including those that will not be covered in class.

Scope

Exercise a lot of care on choosing projects, otherwise you may choose a project that will take a lot of your time. It is in your best interest to choose a project that can be implemented using some existing NLP library.

Important Dates

Abstracts Due	:	March 04, 2022
Proposals Due	:	March 18, 2022
Proposal Presentations	:	Suspended
Progress reports	:	April 25, 2022
Oral presentations	:	May 23, 2022
Final reports	:	June 03, 2022

Abstracts

Soon after choosing your project you are expected to submit an abstract on your idea. The abstract is just a short paragraph describing the project you've chosen, sketching the general approach you intend to take and the kinds of data you're going to need for testing your implementation. The abstract mainly serves to give me a chance to help you get resources you may need, and to make sure that the project you have chosen is feasible.

Proposals

The proposal is a one page description of what exactly you plan to do, designed to convince me that you've got a plan for the project and that you've started working on the project. The plan should have a set of activities that you will undertake up to completion of the project and their

timeframes. When you submit your proposals, you should have your topics completely firmed up.

Progress report

The progress report is meant to give me picture on whether you will be able to finish your project or not. You are expected to report against the activities you set out in your proposal.

Oral Presentation

The presentations are expected to be at most 15 minutes long. The presentation will be followed by 5 minutes of questions & comments from the audience and another 5 minutes for software demonstration.

The presentations will be graded by your peers using a defined assessment criterion.

Final write-up

The final write-up should not be more than 8 pages. It should describe your approach, results, data analysis, and so on. Submission of the milestones is required, but you will only receive a grade at the end, based on your write-ups and presentation, so don't stress about the formatting or polish of the intermediate submissions.

Learning Outcomes and Assessment Criteria

The learning outcomes from which the Assessment criteria is going to be drawn are:

1. Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
2. Ability to sample machine learning techniques and assessing which ones are suitable for a given problem.
3. Understand and carry out proper experimental methodology for training and empirically evaluating NLP systems.
4. Ability to implement an NLP solution using some NLP library.