

Assignment 2

Groups

Please form groups of exactly three students and add your group to this spreadsheet:

<https://docs.google.com/spreadsheets/d/12c4923l-nX5yRcwK2pEv9DVbJ9qxFGDhKQ5gc35SAgo/>

If you have difficulties finding a third member, add your group to the corresponding column “We are looking for a third member” in the “Blackboard” tab. Similarly, if you are looking for a group, add yourself to the corresponding column “I am looking for a group” in the “Blackboard” tab.

We are only accepting submissions from groups with exactly three students. If there are multiple two student groups by February 1st, we will randomly split and merge these groups.

Assignment Description

For this assignment, work as a group towards a project proposal for the last assignment. To this end, you will prepare a report (maximum eight pages; excluding references) that delivers (i) a **Literature Review**, critically discussing existing related work for the project that you have in mind, (ii) your research hypothesis and a description of your plan for the project (**Project Proposal**), and (iii) an analysis including descriptive statistics and qualitative examples of the data that you will be working with (**Feasibility Study**).

You can choose to either work on one of the default projects that we outline below or to propose your own project. We highly encourage you to do the latter and to propose your own project. Your project should (a) have a strong machine learning or deep learning component and (b) make extensive use of natural language data.

Literature Review

- Discuss existing work and explain how these works relate to your proposal
- Put the related work into a broader perspective, identifying shortcomings and next steps that you want to address in your project
- Discuss one of the related papers in more depth, assessing its strengths and weaknesses; provide a critical analysis of the claims of the paper

Project Proposal

- State your research hypothesis clearly
- Describe a concrete plan for testing this hypothesis
- Lay out steps for measuring progress towards the goal of your project
- What data are you going to use?
- Which existing frameworks and pipelines are you using (if any)?
- Is there an existing open-source project you are directly building upon?
- What kind of models do you have in mind?

Feasibility Study

- Look at the data!
- Use descriptive statistics to analyze the data
- Give qualitative examples demonstrating the feasibility and motivation for your research project

Suggested Report Structure

- Short abstract (500 words maximum)
- 1 page: Introduction describing the problem/research question on a high-level
- 3 pages: Literature Review
- 2 pages: Project Proposal
- 2 pages: Feasibility Study

Early Feedback

Optionally, you can request early feedback before February 7th for your project idea by directly contacting our TAs Pasquale Minervini (p.minervini@ucl.ac.uk), Patrick Lewis (ucabhle@ucl.ac.uk), Yuxiang Wu (yuxiang.wu.18@ucl.ac.uk), Qiang Zhang (qiang.zhang.16@ucl.ac.uk)

Default Projects

Develop a new model for one of the challenges below:

- CommonsenseQA challenge (A Question Answering Challenge Targeting Commonsense Knowledge)
 - Dataset: <https://www.tau-nlp.org/commonsenseqa>
 - Paper: <https://arxiv.org/abs/1811.00937>
- CoNaLa challenge
 - Dataset: <https://conala-corpus.github.io/>
 - Paper: <https://arxiv.org/abs/1805.08949>
- WikiHow Text Summarization
 - Dataset: <https://github.com/mahnazkoupaei/WikiHow-Dataset>
 - Paper <https://arxiv.org/abs/1810.09305>

Options for Computational Resources

- Colab
 - You get one free GPU per person for 12 consecutive hours
 - 3 people = 3 GPUs
 - You could train for longer than 12 hours if you are clever about checkpointing and regular saving your models
- UCL Cluster

Style Guides and Submission

You are required to use the ACL 2019 style template for your report. You can either use LaTeX (<http://www.acl2019.org/medias/340-acl2019-latex.zip>) or Word (<http://www.acl2019.org/medias/341-acl2019-word.zip>). Submit a single PDF of eight pages maximum (excluding references) to Moodle. Use only your group ID as the name (e.g. 123.pdf), but mention the authors on the title page. Do not change the template (e.g. reducing or increasing the font size, margins etc.).

Marking

Your report will be assessed based on the three criteria above (Literature Review, Project Proposal, Feasibility Study). In addition, we will ask you to anonymously rate your co-authors regarding the quality of their contribution to your joint report.