Final Project Guidelines

AM: Deep Learning for NLP, MSc Cognitive Systems, Summer 2020

Sharid Loáiciga, PhD

June 23rd, 2020

1 Overview

Your final project will determine 60% of your final grade. The project can be carried out alone or in groups of up to 3 people. If it is a group project, all members are assumed to be involved in all parts of the project and by default will receive the same grade. This default can be overwritten by the group contract (see 1.4).

The final project should implement a system related to deep learning for NLP using the PyTorch library and test it. The project is documented in an ACL-style paper that adheres to the standards of practice in computational linguistics.

1.1 Content and scope

This a 6 ECTS course (30 study hours = 1 ECTS credit point), which means that in total, it should take you about a quarter of your work load this semester (assuming full time studies). You have the option to choose from one of the topics published in the course website (`https://compling-potsdam.github.io/sose20-deep-nlp/final_project/instructions`) or to choose your own topic. If you opt for the second, your topic needs to be approved by Sharid first.

1.2 Project implementation

It is recommended that you use pre-trained word embeddings or pre-trained visual features. In contrast, existing implementations of the systems proposed should not be used. One of the goals of this course is for you to implement your own deep learning models using the PyTorch library. You may look at existing implementations as examples. However, you need to prove a **deep understanding** of the implementation of your chosen system; you need to show why it worked or didn't work and try to improve it.

The final project should be handed in as documented, running Python/PyTorch code. For larger projects, a download link (e.g., to Github, my id is `sharidloaiciga`) suffices. Please also include instructions on running the code (and thus verifying the results). Proper documentation is an important criterion for your submission.

1.3 Supervision meetings

Sharid is available for individual meetings until **July 24th**. Please send her an email to arrange a meeting if you'd like guidance or advice.

1.4 Group contract

An example of group contract is provided in the course website. Most of the items are optional except for the item **Conflict**. You should decide in advance how to handle conflicts where a

group member doesn't contribute their part to the completion of the project. In particular, you should specify how to handle the equal grading default.

1.5 Grading

As a reminder, these are the requirements to pass the course:

Weekly reaction paragraph for each pack of preparation material	mandatory not graded	weight 10%
3 programming assignments completed	mandatory not graded	weight 30%
Project work: – presentation of a project proposal – written report of completed project – code	mandatory & graded	weight 60%

The grading for the project will be based on the written report and code mainly. The grade takes into account the following criteria:

Understanding and correctness

- Rationale and goals of the project are introduced: Are there any problems in conceptualizing the problem, and mistakes in your understanding of basic issues?
- Related work is assimilated and tied to the research question: Do you adhere to academic standards by citing relevant papers and comparing to them?
- Methods are understood: Is anything you wrote in your report blatantly wrong? Are the algorithms/methods described correctly?

Quality of description

- Information is organized coherently: Is there a coherent thread throughout the paper from the motivation through approach to results?
- Ideas are communicated accurately and appropriately: Is the formatting consistent and does it make the paper easy to read?
- Graphs and tables: Do you include visual supports such as graphs and tables? Do you refer to them? Do they support your arguments / display your results?
- Formality: Is the formatting consistent and does it make the paper easy to read? No spelling errors, formal style!

Code

- Code runs
- Code is documented and well-structured, so it can be understood and run or adapted by someone who has read your paper (or yourself in one year)
- Results can be verified in inference mode (please submit all necessary materials to check your system).

These criteria are not necessarily weighted equally, but they will all influence your final grade.

1.6 Timeline

03.07.2020	Group contracts
21.07.2020	Project proposal presentation
28.08.2020	Hand in group report

If you fail to meet the deadline, you can still submit your final report with a penalty of 2% off your final grade per day. The report will not be accepted after 5 late working days.

1.7 Project report

The project report is an academic paper which documents your research questions, methods and results. Please follow the current ACL style file, and write a standard ACL style paper. The study regulations specify 20 pages for the paper, but I'm counting the weekly reaction paragraphs as part of your report. In addition, the ACL style is quite dense. I expect 9.5-10 pages (including references) in the ACL style.

If you're working in a group, you are jointly responsible for writing the report, and each person should understand and be familiar with all parts of the report.

The project report should be self-contained, it should not assume familiarity with your project.

1.8 Academic Writing

The papers you submit wrt. your final project are part of a specific genre, academic writing. Student papers are special since they serve the purpose of clearly documenting what you did (as in scientific papers), but also to demonstrate what you know/have learned. Please remember that your instructors can only grade you based on what's written in your papers – so try to make it easy on us to understand what you know and what you have achieved. Initially, it can be hard to judge which things are "obvious" and shouldn't be explained and which things need repeating (even though we both know them). It usually helps to picture an intelligent peer who hasn't taken this class yet as your target audience. I.e., you can assume most basic knowledge about the field (of the kind that undergraduates possess) as given, but should explain or at least mention anything that is beyond that level, though briefly.

In terms of style, academic writing is very results-oriented, focusing on outcomes and argumentation much more than on the process. Use sensible sectioning. Focus on methods and results (it matters much less how you got there, especially when you don't have much space). I strongly recommend stating your main result early on (in the introduction): "An academic paper is not a crime novel."

All non-obvious statements need proof, either by argumentation (why did you conclude this?) or by citation. Make sure to cite appropriate literature. Prefer official conference proceedings and journals over arXiv papers. It is clear that you won't have time for a thorough literature review, but in many cases there are straightforward previous approaches, and if you rely on them, you must cite them correctly. The same applies to code packages and data sets that you use in your project.

Use LATEX if you can. A user friendly alternative is Overleaf (https://www.overleaf.com/), a web-based LATEX editor, no installation required. Here's the ACL overleaf template: `https://www.overleaf.com/latex/templates/acl-2020-proceedings-template/zsrkcwjptpcd`

Finally, respect your instructors' time by proofreading your papers.

1.9 Project Proposal Presentations

On the last week of the course (eventually the last two weeks), each group will briefly present their project idea and approach. The goal of the presentation is to share with the class what you're doing and to have a space to give each other feedback. The idea is for you to support each other.

Each presentation should be no longer than 15 minutes. It is up to the students in each group to decide how to organize the presentation. Each student should present some part of the work. All students in the group should know the contents of the whole presentation and be prepared to answer questions.

Your presentation will likely only include preliminary results (from previous work or a baseline). You should also provide a clear motivation and problem statement, a description of your approach, and your plan on how to evaluate your project. It is typically necessary to refer to previous work, as well, in order to provide context, background, and ideas for what works and doesn't work.

While the others are presenting, think of some questions or comments to contribute! This presentation is mandatory but not graded individually.

I suggest looking at these recommendations by Annemarie Friedrich: `https://drive.google.com/file/d/1Jw7iUfRX2WDpZ7wndi7cfeXqjPujEGDE/view`.

Aknowledgements

This document is based on a similar document by Tatjana Scheffler, Questions and NLP: Final Project Guidelines. The author thanks her for sharing it.