BM1 Advanced Natural Language Processing
University of Potsdam, Winter 2018/19

Final Project Guidelines

Tatjana Scheffler, Ph.D.

October 22, 2018

1 Overview

Task. The grade of the course is determined based on a **final project**. Your task is to define and carry out a small programming project using and extending the concepts covered in class. Any topic related to Natural Language Processing in a broad sense is in principle eligible. You can draw inspiration from the classes, readings, or examples of previous projects (see below on how to find a project topic).

The final project will be a group effort. This will enable you to attempt more ambitious projects, and to make use of diverse backgrounds and abilities. To ensure a fair assessment, two individual papers are also part of the grade (a pre-project proposal paper and a post-project report).

Scope. In terms of work load, the project should correspond to approximately one of the assignments, maybe a bit more. Please plan enough time. Working in groups will not decrease the overall time needed – also remember that group coordination costs time and effort, as well.

Grading. The grade takes the following criteria into account: planning paper (10%), project presentation (20%), project implementation (30%), and project report (40%).

Milestones and Requirements.

Dec. 14, 2018	Group contract due (not graded)
Jan. 11, 2019	Planning paper due
Feb. 12, 2019	Project presentations
Mar. 1, 2019	Final project due
Mar. 8, 2019	Project report due

2 Group work

Please work in groups of 2-3 students. Make some effort to create groups that include students from diverse academic backgrounds (i.e., don't just work with your friend from undergrad!). You will need much more than just technical abilities to successfully implement the project: think also about sustaining communication and cooperation, understanding the data, time management, presentation, identifying and understanding problems, etc.

Working in groups allows you to pick more challenging tasks and create more interesting projects. Remember that the benefit of group projects comes from actually working *together*, and learning from each other. Really successful group projects draw on each team member's strengths and skills and aren't just made up of several smaller "parallel" projects.

After formation, your group is required to make a group contract that will govern your collaboration. The contract will specify expectations for the group members as well as procedures for dealing with any issues that may arise. Your contract should answer (at least) the following questions:

- How will you communicate with each other?
- How often and where will you meet?
- How will you make sure that your meetings are productive?
- What will you do if a group member breaks any rules in this contract (e.g., doesn't show up for a meeting)?

Please really think about this because you all depend on each other for the success of the project.

3 Planning paper (individual)

The planning paper serves two purposes. On the one hand, it is a short project proposal which specifies what you are going to do, how, and why. On the other hand, it will require each team member to reflect on what you are hoping to learn from the project and what you will contribute to its success. The planning paper is a short (max. 2 pages/1000 words) document, which each student will write and hand in independently.

The following questions should be answered in your planning paper:

- What is your project about and what approach are you going to take?
- What tools and/or data will you use?
- What will you learn from this project? How does the successful implementation of the project help you achieve the learning objectives of the class?
- What will be your personal role in the team, which skills and what knowledge can you contribute?

The planning paper serves to show that you have thought about what you're going to do and how to go about it. If you are unsure whether your project idea is suitable for the class, contact me well ahead of the deadline to make an appointment to discuss it. I can also help identify ressources you may use, find directions/approaches, etc.

Note: Please note that I consider this planning paper (as well as all other assignments/papers you hand in) an individual exam. You must therefore write every word of it yourself. It is unacceptable to copy anything from a book, paper, or the web, or your project partners' papers! This is considered plagiarism and will be (and has been) reported. Please describe your project and any methods you use in your own words (demonstrating your own understanding of it).

4 Project presentation (group)

In the last week of class, each group will briefly present their project. These presentations serve a double purpose: First, to get you started well before the final deadline and help you evaluate (while there's still time) whether your idea and approach is going to work. Second, to foster collaboration and let you get to know the others' projects. This way, you will learn not only from your own project but also get to see (preliminary versions of) the cool things the others are doing. Each presentation should be no longer than 10 minutes, and you may designate one presenter or split it up among you. 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.

5 Project implementation (group)

The final project should be handed in as documented, running Python code (if you want/need to substitute another programming language, contact me in advance). For larger projects, a download link suffices. Please also include instructions on running the code (and thus verifying the results). Proper documentation is an important criterion for your submission.

6 Project report (individual)

The project report is a short paper summarizing your project and reflecting on your individual learning outcomes. It is a coherent essay whose length should not exceed 5 pages (2500 words). In your paper, please answer the following questions:

- What was the topic of your group's project, your approach and main results?
- How did your project support you in obtaining your learning objectives (see planning paper)?

- What did you personally contribute to the project?
- What are the main take-away messages from your project? What did you learn about your chosen topic?

The project report should stand on its own (i.e., it should not assume familiarity with your project or with your planning paper), but you may reuse material from your previous paper if you want. It is both a report on what you did, as well as a reflection of what you learned in your project.

7 Topic Ideas

Any project related to the class topics is possible. Feel free to extend any of the lecture topics, exercises and questions in the readings, or other topics that involve Natural Language Processing.

Each project should include one NLP method or algorithm applied to at least one dataset. There are different types of project: developing an application, evaluating a range of algorithms and methods for one task, applying a method to different datasets, extending an existing algorithm, etc.

If you really can't come up with any ideas, try an internet search. For example, final projects from a related (though undergraduate!) Stanford class are listed here:
`http://nlp.stanford.edu/courses/cs224n/`

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. The most common mistake by students with little university-level writing experience is to faithfully describe what they did (and how they changed their approach in the course of the project) in largely temporal order. Instead, focus on methods and results (it matters much less how you got there, especially when you don't have much space). I 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. 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. I recommend Overleaf (`https://www.overleaf.com/`), a web-based LaTeX editor, no installation required.

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

Email me with any questions: `tatjana.scheffler@uni-potsdam.de`

Have fun!