

Project report

Marcel Bollmann

Introduction

This document contains information about the written project report in the course *Text Mining*, including the suggested structure, formal requirements, and grading criteria.

Suggested structure

We anticipate that most project reports can be structured as follows:

1. **Introduction.** What problem did you address in the project? Why is this problem interesting? What can we learn by solving the problem?
2. **Theory.** Present relevant theoretical background, with a focus on those concepts and methods that were not covered in the course.
3. **Data.** What data did you use in your project? How was this data created? What preprocessing did you do (if any), and why?
4. **Method.** Explain how you approached the stated problem. Aim to be detailed enough for others to reproduce your results. If necessary, refer to your code.
5. **Results.** Present your results in an objective way. Use tables and charts, but do not forget to also include a summary in text form.
6. **Discussion.** Analyse your results. Discuss the limitations of your work. Compare your study to related work, such as internet materials or scientific articles.
7. **Conclusion.** Summarise your analysis. To what extent did you solve your stated problem? What else do you take away from your project?

It is fine to deviate from this structure if you have good reasons for doing so.

Formal requirements

Your report must meet the following formal requirements:

1. *Format* Your report must use the ***ACL paper style template** and be submitted as a single PDF document (via Lisam).

2. *Page Limit* Your report must consist of **3–5 pages of content** (text, tables, figures), plus unlimited pages for references. If you want to provide additional data (e.g. more results tables) in an appendix after the references, you may do so, but your report must be complete even without the appendix.

3. *Title and Abstract* Find a **proper title** for your report that expresses what you did in your project; do *not* just use a generic title such as “Text Mining Project Report.” Your report must also contain an **abstract**, no longer than 200 words, that provides a concise summary of your project’s purpose, method, and results. The abstracts of all projects may be published on the course website.

4. *References* Whenever you use ideas, code, or text from others, you must appropriately cite your sources. Use a **consistent referencing style**, such as the APA style. Don’t forget that this also applies to materials from the internet! For datasets or software libraries without a dedicated publication, it is also fine to provide a URL in a footnote.

5. *Proofreading* Before submitting your report, make sure to **proofread your text** and check it for errors in spelling and grammar. Consider using a writing assistant service such as *Grammarly* or *Wordtune*. If you need help with your writing, consider seeking guidance from Academic English Support (*Språkverkstaden*).

6. *Code* Together with your report, you must submit a link to a **repository containing the code** that you wrote for the project. This repository can be either on GitHub or on LiU’s GitLab server. In case of a private repository, you must grant the teachers read access to it.

7. *Failure to comply with these requirements* If you fail to comply with any of the formal requirements listed here, **we may return your report to you and ask you to make the necessary changes** before submitting it for one of the additional examination opportunities. (The dates of these can be found on the course website.)

Assessment

When grading your report, we will assess it with respect to the assessment criteria laid out below. The criteria are divided into three different aspects: (1) problem & presentation, (2) methodology, and (3) evaluation & analysis. When deciding your final grade, we will weight each of these aspects equally.

Requirements for Grade E 732A81 / Grade 3 TDDE16

To receive a passing grade on this module, your report needs to fulfill *all* of the following criteria:

Problem & Presentation

- The project as a whole clearly goes beyond the lab assignments and represents an appropriate amount of work for the given timeframe (ca. 80 working hours).
- The report is clearly written and logically structured.
- The report contains a clear, relevant, and well-motivated problem statement.

Methodology

- Technical concepts, models, and algorithms are suitable for the problem.
- The data used is suitable for the problem.
- The report contains enough detail to assess that technical concepts, models, and algorithms have been applied correctly.

Evaluation & Analysis

- Results are evaluated with appropriate evaluation methods.
- Results are interpreted and analysed correctly.
- Conclusions are clearly stated, refer back to the problem statement, and are supported by the analysis.

Requirements for Grade D-B 732A81 / Grade 4 TDDE16

You will receive one of the “intermediary” grades of your grading scale if your report fulfills all the criteria given above for Grade E/3, and partially fulfills the criteria given below for Grade A/5.

Requirements for Grade A 732A81 / Grade 5 TDDE16

To receive the highest grade on this module, your report needs to fulfill all the criteria given above for Grade E/3, and additionally meet the following criteria:

Problem & Presentation

- The problem is placed in a scientific context by discussing relevant related work, taken from scientific articles.
- The report consistently uses an academic writing style and demonstrates a good command of relevant scientific terminology.
- The report makes good use of plots, illustrations, and/or data examples to communicate important parts of the work.

Methodology

- The technical concepts, models, algorithms, and/or data used for the project go substantially beyond the labs and required a significant amount of effort, for example by:
 - using concepts, models, or algorithms that were not covered in the course, taken from scientific articles;
 - making non-trivial modifications to existing models or algorithms;
 - creating data specifically for the project (e.g., manually annotating data, scraping data from the web);or a combination of the above.

Evaluation & Analysis

- The evaluation and analysis of results is thorough and provides insights that go beyond purely quantitative evaluation measures, for example by:
 - complementing a quantitative evaluation (e.g. computing F1-scores) with a qualitative analysis (e.g., performing a manual error analysis, evaluating on different partitions of the data, performing an ablation study, ...);
 - placing results and conclusions in a scientific context by relating them to relevant scientific articles;
 - providing an insightful discussion of limitations of the work and possible steps to address them;or a combination of the above.