

Project report

Marco Kuhlmann

Introduction

This document contains information about the written project report in the course 732A81 / 732A92 / TDDE16 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 consist of 4–8 pages of content (text, tables, figures), plus unlimited pages for references. You must use the template made available on the course website and submit your report as a PDF document (via Lisam).

2 Abstract Together with your report, you must submit an abstract (via an online form). This abstract must be one paragraph, and may not exceed 200 words. The abstract should provide a concise summary of your project's purpose, method, and results. Note that the abstracts of all projects will be published on the course website.

3 Title page The first page of your report must contain the title of your project as well as your full name, LiU-ID, and course code (732A92 or TDDE16). Do not use a generic title such as 'Text Mining Project Report'; instead, find a title that expresses what you did in your project, such as 'Predicting Personality Types from Reddit Posts'.

4 References When relating your work to the work of others and whenever using ideas, code, or text from others, you must appropriately cite your sources. This also applies to materials obtained from the internet.

5 Proofreading Before submitting your report, make sure to carefully proofread your text and check it for errors in spelling and grammar. Consider using a writing assistant service such as *Grammarly* or *Ginger*. 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 (via an online form). This could be a public repository on GitHub or a private repository on LiU's GitLab server. In case of the latter, you must grant the teachers read access to the repository.

7 Failure to comply with these requirements If you fail to comply with any of the formal requirements listed here, we will 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.)

8 Special needs If you have a documented disability, you should speak with the course examiner (Marco Kuhlmann) as soon as possible regarding accommodations.

Assessment

When grading your report, we will assess it with respect to the following aspects: project at large, methods, and analysis. For each aspect, we will assign a component grade from the scale A, C, E (all passing grades) and F (fail). Descriptors for A, E and F are provided below. For a passing grade, your report must meet all of the criteria in the descriptor for E. Your component grade is determined by the number of criteria you meet in the descriptor for A: 0 criteria → component grade E; 1 criterion → C; at least 2 criteria → A. Your final grade is computed as follows:

Final grade 732A81 / 732A92 Your final grade is the average (on the full ECTS scale) of your two lowest component grades. (Example: If your component grades are E, C, E, your final grade is E; if they are C, E, C, your final grade is D.)

Final grade TDDE16 As for 732A81 / 732A92, but the final grade is converted into a numerical grade on the scale 3, 4, 5. In this context, A–B correspond to 5; C corresponds to 4, and D–E correspond to 3. (Example: If your component grades are C, E, C, your final grade is 3; if they are A, C, A, your final grade is 5.)

Aspect 1: Project at large

Is it clear what was done in the project, and why? Does the project go beyond what has been covered in the course? Does the project have enough substance, or would there have been room for more work?

- F The report does not contain a clear and well-motivated problem statement. The project is essentially a repetition of one of the lab assignments. For a project with this timeframe¹, I (the assessor) would have expected significantly more work.
- E The report contains a clear problem statement, and the stated problem is well-motivated. The project goes significantly beyond the lab assignments. The project represents an appropriate amount of work.
- A The problem is placed in a scientific context, including references. The project builds on technical content outside of the course, taken from scientific articles. The project represents significantly more work than expected.

¹The timeframe for the project is ca. 80 working hours.

Aspect 2: Method

Is the data used in the project suitable for the stated problem? Are technical concepts, models and algorithms applied correctly? Are the experimental results validated with appropriate evaluation methods?

- F The data does not fit the stated problem. There are flaws in how the experimental results are produced or evaluated, or there is too little information in the report to assess whether the method is appropriate.
- E The data used in the project is suitable for the stated problem. Technical concepts, models and algorithms are applied correctly. The experimental results are validated with appropriate evaluation methods.
- A The data is created specifically for the project. The project involves non-trivial modifications or combinations of models and algorithms. The experimental results are validated using several complementary evaluation methods.

Aspect 3: Analysis

How thorough and correct is the analysis of the experimental results? Does the report show awareness of the limitations of the work? Is the report clear about what conclusions can be drawn from the project, and are these conclusions well-supported?

- F There is no analysis of the experimental results, or the results are misinterpreted. The report does not contain a discussion of the limitations of the work. There are no clear conclusions, or the conclusions are not supported by analysis.
- E The analysis of the experimental results is sufficiently thorough and correct. The report shows awareness of the limitations of the work. The conclusions about the stated problem are clearly stated and supported by the analysis.
- A The analysis of the experimental results is thorough and detailed. The report contains an insightful discussion of the limitations of the work. The conclusions are placed in a scientific context, including references.