Heidelberg University
Database Systems Research Group
Prof. Dr. Michael Gertz
Satya Almasian, Dennis Aumiller

November 02, 2020 Text Analytics Winter Semester 2020/21

Project Report Guidelines

The project reort is due on Monday, March 15, 2021, 18:00, via Moodle

General Notes:

- The project report should give a detailed overview over your work. You should summarize all important
 methods you utilized during your project and give insights into your results. The report counts for 35%
 of your final project grade.
- Reports should be between 15 and 20 pages.
- We require you to use the LATEX template found on the Moodle page.
- For further questions, please reach out via the Moodle forums, or contact an instructor.

1 Content of Your Report

The final project report for this class should outline what you have achieved and show the results of your work. Your report should be between 15 and 20 pages excluding references. It should be written in a way that a fellow classmate can understand your report, so you cannot assume any knowledge about your specific application, but the reader should be familiar with the basic concepts of text analytics you learned during class. The following sections can already give you a rough structure for your report, but is not set in stone, and can be altered if it serves your needs. However, your write-up should in general follow a structure similar to other scientific articles or seminar reports.

1.1 Key Information

Your report needs to give the following information:

- **Title:** The final title of your project. This does not necessarily need to be the same as in your project proposal.
- **Team Members**: The name of all team members. The participating students must match a group that is handing in assignments together. Please add your immatriculation number and course of study as well.
- Email Addresses: The preferred email of each team member.
- Member Contribution: In this report, we also expect you to outline what each member did for the
 project. Each team member needs to contribute some cover some text in which they describe what their
 contribution was, and which challenges they faced.
- **Mentor:** Give the name of your mentor.
- **Anti-plagiarism confirmation:** A short affirmation that you did the project's work at your own as already indicated in the template.

1.2 Abstract

The abstract is supposed to summarize your work in less than 300 words. It motivates your work, states the problem, and highlights your project's main contributions. The abstract serves as a short summary for the reader.

1.3 Introduction

Content-wise the introduction can be similar to the one you already gave in your project's proposal. It should explain why your problem is interesting and relevant, and motivate the rest of your report. While this may sound similar to the content of the abstract, you have more space to get your reader's attention in this section. Furthermore, the introduction outlines the key ideas of your approach without going into detail about its implementation. It is also possible to give an outlook of the results the reader can expect to see in the later sections.

1.4 Related Work

This section of your report explains what prior work was done related to your project, and put your work into the context of current research. This includes papers that you used as a basis for your approach or that solved the same problem using different methods. It can also include papers that used the same techniques as you did but applied them to a different problem. In the related work section, you should emphasize how your work differs from previous work, e.g., by outlining the limitations of previous work, or why your application domain is different from those other researchers have already investigated. However, you should not go into detail about any of the work you cite, instead you should outline how the major points of previous work relate to your own project.

1.5 Methods/Approach

The methods or approach section gives details about your applied system, and is intended for readers that want to fully understand your approach. This includes pipelines you used, algorithms you developed, and all other key methods you designed during the project. In the following, some advice for this section is given.

- Be specific about your methods. While it is important to outline the key ideas in the previous sections, in this section you can show your detailed understanding of the matter. Include equations where necessary and show figures, such that the reader can better follow and understand your approach.
- If some part of your work is original or especially creative, make it clear, such that the reader (e.g., your mentor) can understand the novelty of your approach. You do not need to show any false modesty in this part of your work. If you use any existing methods, however, you need to cite them and give references.
- Be consistent through your report and agree to a fixed vocabulary before writing. This includes mathematical terms and notations, method names, names for data sets, and the like. Few things are more irritating than changing naming during a report.
- If you have any baseline approaches, they should also be described. This can be detailed if you created the baseline yourself or just briefly if there is an external source from which you took the baseline.

1.6 Experimental Setup and Results

This section needs to contain the following information:

- **Data:** Describe the data you used for the project. If you utilized existing data, give references. If you collected your own data set, outline where, when and how you have gathered it. Providing a clear understanding of the data can be very helpful for readers to follow your following analysis. If you have any interesting or insightful metrics of your data, this is the place to show them.
- Evaluation Method: Be clear how you are going to evaluate your project. Do you use an existing quantitative metric like an F1 score? Or did you define your own evaluation metric? If you use your own method, be sure to motivate what you try to achieve with it, and what it does reflect. In any case give a definition of your evaluation metric, be it a qualitative or quantitative one, automatic or human-defined.
- Experimental Details: If your methods rely on configurable parameters, specify them, such that your results are replicable.
- **Results:** The results section of your report highly depends on your project. If you have a baseline to which you can compare your approach, you obviously should do this. In such cases tables and plots are suitable means to present your achievements. In any case, use the evaluation metrics defined before. Moreover, comment your results: are they in line with your expectations? Or are they better or worse? What are potential reasons for this?

1.7 Analysis

The analysis part of your report can be a separate section, but does not necessarily need to be. Depending on the content and the scope, it can be merged with the results section or the conclusion of your report. Nonetheless, your report should include some qualitative analysis. Did your system work as expected? Are there cases in which your approach consistently succeeds in the task, or surprisingly fails? If you have a baseline, you can also compare if they succeed and fail in the same contexts or if your approach may be suitable for other applications than the baseline. Underpin your points using examples or metrics instead of stating unproven claims.

1.8 Conclusion

Recap the main contributions of your project. Highlight your achievements, but also reflect on the limitations of your work. Potentially, you can also outline how your model can be extended or improved in the future, or you can briefly give insights into what you have learned during the project.

1.9 References

List your references at the end of your report. Since it is good practice, and avoids careless mistakes, we strongly recommend you to use BibTeX to manage and produce your references.

2 Grading

Your project report will be graded by taking into account multiple criteria. This is the originality or creativity of your approach, the performance of your methods, your scientific writing, how detailed you presented your results, in-depth understanding of the matter, etc. Note that each student gets an individual grade, i.e., in cases where team members contributed unequally to the project, grades can differ between members.