CS 7331 Data Mining

How to write a successful project/consulting report

Think of your project report as of a report a consultant would prepare for a client. The main goal is to provide value to the client.

Structure

- Use a title page with the project title and your name.
- Add an abstract/executive summary (1⁄4 page) that introduces the problem and **highlights the major results**. Be concrete with your major results (e.g., "This report shows that...").
- Organize the report using CRISP-DM and then by issues (e.g., the impact of gender on ...)
- Pages should be numbered. Use an automatic generated table of contents for longer documents (>10-15 pages).
- The document should end with a conclusion that summarizes the main findings and a list of recommendations.

Business Understanding

- Who are the main actors in the field? What do they want/need?
- Can you get some numbers from other reports or studies?
- What is published about the topic in the news/in academic publications? Do at least an internet search.

Tables

- Do not copy tables from your tool (copy&paste or screen shots). Create tables in Word/Latex and only add the information that is necessary. Highlight important information.
- Tables should be **numbered and referenced in the text**. Discuss what we can learn from the numbers in the table.
- Use a table to describe features in a consistent and easy to read way.

Figures

- Every figure/plot should have a **caption with a number** under it (Example caption "Figure 1: The impact of ..."). It should be referenced (by number) and discussed in the text. What do we see and how is it helpful?
- If there are several options to plot the same thing then choose the best visualization. You may highlight important aspects in the graph.
- Your graph should use proper names for both axes (use xlab and ylab in plot).

- Make sure your graphs are readable (text/numbers in graphs should be about the same size as the regular text in your document). Graphs should be pretty (add a legend, use color where appropriate, etc). However, figures should not waste space!

Code

- Upload **code as a separate file.** The code should be executable as is and contain comments.
- Short pieces of code can be included in the Appendix of the main document if they are particularly interesting.

Experiments

- Use an appropriate way to organize you experiments (e.g., different classification algorithms) and present results. Justify why you choose the experiments you did and why you choose not to pursue others.
- It is important to compare the results from different experiments. What worked best and why?

References

- Add as many references (books, papers, URLs) as possible so other people know where you took information from. State sources of code, tools or additional data used. The aim is to make your work reproducible by someone who reads the report.
- You may use footnotes or any consistent reference style (ACM, IEEE, Named, etc.).