CMPS 242 Project Report Guideline, 2006

Project reports must be typeset. I would like you turn in a hardcopy of your report to me, as well as e-mailing me a .pdf file. See me if you have code or unusual data that you would like to make available to future classes. The early deadline for the projects is December 1 (the last day of classes). All project reports are due by 3:00pm[1] on Wednesday December 6. The text in your report should be your own. Quoted text must be set off by quotes (" ") and the source clearly attributed, even if it is as small as a single phrase. If you use someone else's figures or tables the appropriate attribution should appear in the caption.

In the past, project reports have been from 8 to about 18 pages of text long, with most in the 10-14 page range (not counting appendices and large tables, which can add quite a bit of length). Please do not turn in large sections of code listings or massive tables of raw data (although some information on the data is important, and a table indicating what a few typical examples look like would be OK, especially if you have an unusual data set). The report should be easy to read, if I can't tell what you are talking about it is difficult to give you a good grade. Every figure or table in the report body should be discussed in the report body. If you would like to present additional experiments that are not evaluated in the body of your report, include them as an appendix.

Your report should have an abstract as well as introduction/problem description, related work, methods used, results, and conclusion sections, as well as a bibliography. I am flexible on the exact section breakdown, add or merge sections if it makes writing/reading the report easier. Readability is important, so be sure to define your terms before using them and present things in a logical order.

Your report should start with a short 1-paragraph abstract that mentions the problem you attacked, your main methodology, and your results (perhaps 3-4 sentences total).

The introduction should contain a description of your problem at a level that any student in CMPS 242 could understand, any area-specific jargon should be explained. It can also give an overview of your results, how you obtained your data, etc. However this additional information is likely to appear elsewhere, and so should be just summarized in the introduction to avoid too much redundancy. If your particular problem is technical or difficult to describe precisely then you might give just an overview of it in the introduction and use a different section to describe it precisely. The introduction should provide an overview of what the problem is, why it is interesting/important (why did you choose it) how you attacked the problem, and an idea of the success and/or failure of your methods.

The related work section should contain a survey of relevant previous work for your problem and possibly the methods you used. This is sometimes a good place to clearly spell out what you did for the course as opposed to what was done by others or outside of the course. Feel free to cite textbooks or articles etc. for descriptions of algorithms. However, the best related work sections are not just lists of references, but evaluate and put into context the previous contributions, as well as relating them to the current work.

The methodology section should describe the details of your experiments. This includes the data source (what was measured, how, what kinds of errors are likely) and any preprocessing done. Describe the learning techniques used and what software packages (Weka, SVM light) you used. Ideally, there will be enough information here so that another student could reproduce your results.

[1] I will be in a meeting from noon until 2:30 on December 6, you can leave your project reports outside my door (but indicate in your e-mail that you have done so).

Although I am not interested in a printout of any code you wrote, you could include a link or pointer to where it could be obtained (as well as your datasets). You should also explain here (if not earlier in your report) why you picked the methods you did.

The experimental results section should describe what happened. Is it what was expected? What were the surprises/anomalies? In retrospect, why do you think the results come out the way they did? How do your results compare with others? Ideally, each experiment is a question and the results provide an answer. Tables and graphs are appropriate ways to summarize information. If you are doing many experiments or varying many parameters, a good way of structuring your presentation is to have a baseline situation and compare each of the individual experiments to the baseline.

The conclusions section should include a short self-evaluation of your project (what went right and what went wrong) together with a summary of what was learned from the experiments and what you yourself learned and a recap of what you accomplished. If there are other things you would have liked to try but didn't get around to, you can include future work in the conclusions section (or even make further work its own section).

You should acknowledge any help you have been given on the project and anything else that made the project possible (such as data or machinery/code).

The bibliography should contain relevant publications (articles, books, etc.) that you read in conjunction with your project.

Progress Report

The progress report is a less formal (and much briefer, perhaps 3–4 pages of text) version of the final report. It should have three sections: an introduction (like that for the final report, but with a complete problem statement and you probably won't have results to report), a methodology/plans section, and a progress/problems section.

The methodology/plans section should describe:

1. the data you are using, how it was obtained (you should have all your data in hand by now),
2. the pre-processing and/or feature extraction you have done (or are planning to do)
3. the learning methods and tools you plan to use (or implement)
4. what parameters of the learning algorithm you will need to tune
5. what experiments you plan, and how you will evaluate the results

The progess/problems section should list the progress you have made as well as any significant problems you have encountered (or can visualize down the road). One purpose of the progress report is to get you to think about any potential difficulties while there is still time to work around them.