Project:
Group report

Due: Monday 23 December

Required Submission Format

Provide a write-up of your research in the form of an academic paper that could be submitted to a conference on data mining/machine learning. Your paper should be self-contained. Everyone who has read the assigned reading materials from the course should be able to read and understand your paper. That means that in your paper you can be brief about machine learning methods that are described in the assigned readings, but that you need to provide sufficient details about the problem domain, the dataset, as well as about any other machine learning methods that you used that were not covered in class. The reasons for this are: (1) a description of the problem domain and the dataset will allow to share your paper with interested parties who have not taken ift6758 but who have general knowledge of machine learning; (2) a description of machine learning methods not covered in class will allow to evaluate whether you truly understood those methods instead of treating them as a black box. Your paper can for instance be divided into sections as follows (but if another structure works better for you, don't feel restricted to the one below):

1. Introduction: a description of the problem (profiling of Facebook users), what the goals of the study are, and a very brief description of the results.

2. Methodology: a brief description of the machine learning methods used.

3. Dataset and metrics: a description of the datasets and the evaluation measures used.

4. Results: an overview of the results you obtained by applying the methods from section 2 to the dataset from section 3 using the metrics from section 3. In addition to reporting numbers, your analysis of the results should also contain your insights into the results, i.e. why did a particular method work well/did not work well?

5. Conclusion and future work: briefly summarize your results and list opportunities for future research that seem promising to you but for which you did not find the time within this quarter.

Formatting guidelines: up to 8 pages, double column, ACM Proceedings format.[1] In case you need more than 8 pages, consider splitting your material in a main paper and an appendix. Submit a pdf file on the course Gradescope (8 points).

Code and documentation

Submit an additional documentation file in PDF, called `readme.pdf`. The documentation should contain enough details to allow a student who will take the machine learning course next year to understand the general structure of your code, compile and run it (8 points). **Your software and its documentation will be collected at some time point between Tuesday 9am and Thu 9am. Make sure that on Monday December 23 before 11:59pm, the command `ift6758` invokes a stable version of your software on your account (/submissions), and that this stays available until at least Tuesday 9am. We automatically collect your software on Monday at midnight.**

---

[1] `https://www.acm.org/publications/proceedings-template`