**Option 2: "Problem"**

General guidelines

* Find a data set and corresponding problem that you'd like to solve.
* Find a work bench of data mining / machine learning algorithms.
* Apply at least two different algorithms to your problem.
* Write a report following the proposal guidelines below.

Proposal guidelines

The proposal can be half a page. It should contain the following items.

* A short description of the problem you want to solve and corresponding data.
* A description of or link to the software that you want to use to solve your problem.
* Pointers to the literature that you use for inspiration and comparison.

Below you can find some suggestions for data sets and software. If you're up to a real challenge, you can enter one of the data mining or machine learning competitions.

Report guidelines

The report should read like a scientific paper of about 6 to 8 pages, containing the following items.

* A descriptive title.
* An abstract (about 100 words) that summarizes the research problem, approach/method, and main results. In the body of the report, you should explain clearly, not necessarily in the following order:

1. the application domain and the research problem;
2. related previous work that attempted to address the same or a similar problem;
3. your data set and data collection, preprocessing methods;
4. your approach/method(s) and, often more importantly, the rationale behind your decisions;
5. the main results of your project;
6. an analysis of results and evaluation of your project outcome: to what extent it matches your expectation and why if there is a gap;
7. possible future directions to improve the results;
8. in the appendix, add a listing of the (main) files that you used to generate the results.

Remember that it is important not only to tell what you did, but also to motivate why you made those choices. Also clearly identify any innovative ideas that you explored in your project, by comparing related work to the best of your knowledge.

Criteria for evaluation

The report will be evaluated based on the following criteria.

* Understanding of material: does it show that you know what you are talking about?
* Clarity of exposition: is the report easy to read and understand?
* Correctness: are there any major flaws?
* Completeness: did you address all the points listed above?
* Originality: does the project contain novel, innovative ideas?