

### Project Statement for Milestone 3

(Optional) Your team name

Team member names

In this report you should focus on the formal description of the sequential algorithms you developed for your course project.

1. Problem statement and Algorithm description.

- a. Give a formal problem statement. It consists of (1) problem input; (2) problem output, and (3) object function/goal for your solution.
- b. Give a pseudo-code and corresponding description of your sequential algorithms. If your project targets at multiple problem, give the solution for each of the problem. Your description should correspond to the line number in your pseudo-code.

2. Explain how to extend your sequential algorithm to a distributed algorithm on distributed platforms (with detailed pseudo-code and description) and discuss possible optimization opportunities.

3. Initial experimental plan. At this point you have < 1 month to finalize code and test. You should have an initial experimental plan as a part of your report. The experimental plan has several components you should report:

- a. Dataset description (use what you have in M2)
- b. Algorithm description (specify what are baseline algorithms, and what are your algorithms you will test; give references for each of the baseline algorithm if applicable)
- c. Setting: CPU and memory information of your machine. Give description of the the distributed platform that you are using, e.g., Hadoop, Spark
- d. Plan: in sections, you specify (1) the algorithms you want to compare, (2) the performance metric (runtime, data shipment, accuracy, etc), (3) the factor that you are testing that may have impact of the result (size of datasets, size of query, etc). You should give a statement on what would be the expected result in each subsection you may expect and why.

For the presentation, it's also the time you think about what's the most important result/figure you want to show/present in the experimental study. For example, if you are testing a distributed algorithm, scalability/data shipment are two critical factor you want to show case.