

COMPSCI 753: Algorithms for Massive Data

Semester 2, 2020

Instructors: Kaiqi Zhao, Ninh Pham

Research Project

1 Topic assignment

In order to prepare for your research projects, we provide the research topics together with its research paper under 4 main categories: locality-sensitive hashing, streaming algorithms, graph mining, and recommendation system.

You will send an email to **both of us** (ninh.pham@auckland.ac.nz, kaiqi.zhao@auckland.ac.nz) for registering your research topic and the research paper. **Registration deadline is 5pm, Friday 9 Aug 2020.** Since you will work in group for the research project, please CC your teammates (if you have) so that we will give a correct allocation.

The topic assignment follows the first-come-first-serve policy. We will update the topic assignments on the google sheet which you can view but cannot edit. Please take a look on the google sheet before sending your registration email.

In the registration email, please indicate **3 research topics** with its **priority**. We will base on your priority and our previous assignments to assign you the research topics.

2 Assessment components

2.1 Presentation (15% pts)

Each group of two or three presenters will have 20 minutes presentation (10% pts) and 5 minutes Q&A (3% pts). We will also publish the presentation schedule on the google sheet. The presentation schedule is ordered randomly. Please let us know as soon as possible any clash on your presentation schedule since it is always tough to change the schedule when we are close by. Note that each group will need to prepare at least 2 questions for Q&A sections of other groups (2%).

We expect the presentation focuses on the seed paper by describing the applications, problems, technical solutions, examples to illustrate the algorithms, and future issues.

2.2 Report (15%)

When you get the research topic together with the seed paper, you will have to branch out on your own from this paper. Please get as far as the seed paper since you will have to find **at least 4 more papers** based on the seed paper to complete the report.

The reported will be submitted in two stages so that we can keep track your progress and provide feedback if necessary. The mark will be only based on the final report.

- Stage 1: Report abstract and structure. **Due by 11:59pm, Sunday 13 September.** Note that you can change them while writing your report.
- Stage 2: Final report. **Due by 11:59pm, Sunday 1 November.**

We expect the report presents a short survey of the research topic with detailed discussion about advantages and disadvantages of surveyed algorithms used in the research topic (13%). Based on the discussion, we expect to see your “new” ideas and justification to improve previous solutions (2%).