

CMPT 884 Spring 2018
Martin Ester

Course Projects

Here is a list of some types of research projects:

- experimental comparison of several state-of-the-art methods on a new dataset to gain insight into their strengths and weaknesses,
- modification / extension of an existing method to address some of its weaknesses / limitations,
- development of a new method for a known problem,
- definition of a new problem and development of a simple algorithm.

Course projects can be done individually or in teams of 2 (preferred) or 3 students.

If you want to do a course project using genotype and/or phenotype data, here is an annotated list of data repositories that two of my PhD students recommended:

- 1- TCGA which is usually preprocessed: <https://portal.gdc.cancer.gov/>
- 2- ICGC which is usually not preprocessed: <https://dcc.icgc.org/>
- 3- METABRIC which is a rather large gene expression dataset for Breast Cancer:
<http://discovery.dartmouth.edu/~cgreene/da-psb2015/>
- 4- GTEx for normal tissues: <https://www.gtexportal.org/home/documentationPage>
- 5- Panther, STRING, Reactom, and KEGG for network data
- 6- ENCODE project for ChIP-Seq: <https://www.encodeproject.org/search/?type=Experiment>
- 7- CbioPortal is also good for mutation and copy number studies. It's not recommended for expression data especially if you are planning to apply multiple datasets:
http://www.cbioportal.org/data_sets.jsp.

Formatting instructions

- Reports are to be written in the style of a research paper.
- Use the 2017 ACM Master Article Template, see
<https://www.acm.org/publications/proceedings-template>.
- The page limit is 5 pages for individual projects (1 student), 9 pages for teams of 2 students, and 12 pages for projects with 3 students.
- There is no minimum number of pages.
- **For team projects, please state which team member has authored which part of the report.**

Evaluation criteria

- Originality
How original is the problem addressed? How novel is the solution? Do the experiments provide interesting, new insights?
- Significance
Is the research expected to have significant applications? Are the improvements compared to existing methods significant?

- Technical quality
How hard is the problem? Is the solution principled? Is there any theoretical analysis?
How efficient is the algorithm? How convincing is the experimental evaluation?
- Presentation
How clear are the motivation and the problem definition? Is related work discussed appropriately? How understandable is the method? Are the experimental results reproducible?

Submission

- By email to ester@cs.sfu.ca, mention CMPT884 project report in the subject line.
- **Deadline is Sunday, April 22, 2018 (anytime on that day).**