CSE 291 - PERSONAL GENOMICS FOR BIOINFORMATICIANS - PROJECT DESCRIPTION

Overview

The final project is aimed at providing you an opportunity to perform original research related to personal genomics. The project is open ended and up to you to come up with fun creative ideas! This could take one of the following forms:

- Prototype of a novel bioinformatics tool to analyze human genomes, or an in-depth comparison and evaluation of existing tools.
- An application making personal genomics analyses accessible to non-experts.
- A novel analysis of public human genome datasets to answer an interesting biological question.
- Exploring a theoretical aspect of human medical or population genetics, and evaluating results on real genomes.

If you have a different idea for a final project not discussed here, feel free to discuss. We will also discuss potential project ideas and available datasets in class.

All projects will be performed in groups of 2-3 students. If possible, find team members with complementary expertise. e.g. if you are a strong programmer, perhaps pair with a strong biologist. Finally, you are strongly encouraged to set up a meeting with the instructor during office hours before preparing your proposal to determine if your project idea is on track and within the scope of this course.

Deliverables

The final project is worth a total of 35% of the final grade and consists of the following components:

- Proposal (5 points) (due February 14, 2017)
- Paper (20 points) (due March 17, 2017)
- Presentation (10 points) (March 16, 2017)

These are described in more detail below.

Proposal - 5 points

The proposal should be 1 page maximum and contain:

- Project title
- A list of team members and their expertise areas
- An abstract (maximum 200 words) briefly describing the project and motivation
- Resources and datasets that will be used
- Deliverables

The proposal should be sent to mgymrek@ucsd.edu with subject [CSE291 PROPOSAL:LASTNAME1_LASTNAME2] by the beginning of class on **Tuesday, February 14**.

Paper - 20 points

The final paper should be 5 pages maximum and contain:

- **Project title, author list, abstract** (200 words max).
- **Introduction** (approximately 1 page). Describe the motivation behind your project and how it fits with current literature in the field.
- **Methods** Describe the methods used in your project, including software, statistical tests, and datasets.
- **Results** Describe the major results of your study. When helpful, include figures and tables to display results.
- **Discussion** Briefly summarize the major findings. What are the limitations of your project? How could it be expanded? What future directions does it point to?
- **Author contributions** (0.5 pages) Describe the contributions of each team member to the project.
- **Reflection** (0.5 pages) Overall, how was your experience working on this project? How much time did you spend? How were the team dynamics? Did you receive adequate guidance? What were the major challenges?

The paper should be sent to mgymrek@ucsd.edu with subject [CSE291 PAPER:LASTNAME1_LASTNAME2] by midnight on **Friday, March 17**.

Presentation - 10 points

On the final day of class (Thursday March 16, 2017), each group will give a presentation about their project. Presentations will be 15 minutes total (10 minutes presentation, 5 minutes Q+A). Each team member is required to deliver part of the presentation.

Format

The format of all written portions of the final project should be maximum 0.5 inch margins, 11 point Arial font. Submit documents as PDFs.

Resources

We have an allocation on comet (account csd524) that can be used for computation for final projects. Each team can use approximately 10,000 SUs for their project. If you anticipate that your project requires additional compute resources, we can likely accommodate that but please discuss with the instructor beforehand. For more info on using comet/XSEDE, see the page on using comet for this course.

The course website will be updated with suggested public datasets that may be used for the project.