CSci 5461: Functional Genomics, Systems Biology and Bioinformatics

Fall 2017 Syllabus

Prof. Rui Kuang (kuan0009@umn.edu)

Course Description: The new biotechnologies developed in the past decade are transforming molecular biology research into a quantitative science based on informatics. New platforms are now capable of collecting large amount of DNA/RNA sequences and measuring their activities in cell at a population scale. For example, a single second-generation sequencing machine can generate 6 billion basepairs of DNAs (approximately double-size of human genome), and microarray platforms can measure the intensities of hundreds of thousands of gene expressions in several days. Without doubt, the contents of the collected data offer unprecedented opportunities for a deep understanding of the molecular mechanisms in cell, which in turn will increase the hope towards a complete mapping of the linkage between individual's phenotypes and genetic/genomic patterns. Facing the challenges of making sense out of the sheer volume of data, computer algorithms and data analysis models are playing the central role. This course covers an introduction to various types of functional genomic data available and current computational/statistical methods used for analyzing the data to answer questions in functional genomics, systems biology and their applications in disease research. We will cover the analysis of gene expression data, proteomic data and protein-protein interaction data, with a special focus on how they can be used to understand and infer networks. In particular, the course will also cover the recent advances in cancer genomics by applications of the sequencing, array-based and proteomic technologies. The topics are organized as: 1) Introduction to Genomics and Genomic Techniques, 2) Statistical Analysis of Genomic Data, 3) Data Mining and Machine Learning Methods for Genomics, 4) Cancer Genomics and 5) Biological Network Analysis.

Goals: There will be 4 homework assignments, each requiring the implementation of a computational method and its application to a real functional genomic dataset. Each student will also present a research paper on cancer genomics in the class. Students will also choose a topic for the final project, and present the results in-class. All the homework assignments are required to be programmed in matlab/java/C/C++ and submitted through moodle.

Grading: Homework Assignments: 60%, Course Project: 30%, and Participation/Presentation: 10%.

Text Books:

- Biological Network Analysis, Björn H. Junker, Falk Schreiber, ISBN: 9780470253489 (free online access from library http://onlinelibrary.wiley.com/book/10.1002/9780470253489)
- Cancer Genomics, Edited by:Graham Dellaire, Jason N. Berman and Robert J. Arceci ISBN: 978-0-12-396967-5 (free online access at http://www.sciencedirect.com/science/book/9780123969675)
- Introduction to Biological Networks, Alpan Raval, Animesh Ray, ISBN 9781584884637

Prerequisites: Some programming skills are required for this course. Biology or other non-CS students are required to take Csci 3003 or an equivalent programming course as a prerequisite or get instructor approval. Prior knowledge of biology is highly recommended.

Intended Audience: This course is primarily for graduate and senior undergraduate students in computer science, math, statistics, biological sciences and biomedical sciences with interest in computational biology.

Time and Location: 11:15 P.M. - 12:30 P.M. , M,W; ME 108.