

COS557 syllabus

Course Description: The goal of this course is to introduce students to computational issues involved in analysis and display of large-scale biological data sets. Techniques covered will include clustering and machine learning techniques for gene expression microarrays and proteomics data analysis, biological networks and pathways modeling, data integration in genomics, and visualization issues for large-scale data sets.

A short introduction to the field of bioinformatics and the nature of biological data will be provided; no prior knowledge of biology is required (though some basic understanding would be helpful). In depth knowledge of computer science is not required, but students must have some understanding of computation (no need to know programming). In addition to introductory lectures on all topics, students learn through peer presentations and discussion, so willingness to participate is a definite prerequisite. Through this process, this course will not only teach students about bioinformatics topics discussed, but will also provide a foundation in critical analysis of current bioinformatics literature and in scientific presentation skills.

The course is open to graduate and advanced undergraduate students from all departments.

Format: The course will be taught in a mixed lectures and seminar format, with student presentations of current papers in the field.

Grading:

40% presentations

15% quizzes

15% participation (including attendance and participation in class discussions)

30% final project (10% project proposal, 20% final project report)

Office Hours:

11am-12pm on Thursday

Presentations: Each presentation should be 30mins, with 15mins for discussion afterwards. Presentations should be in power point (or another slides format), and you must e-mail me the power point after your presentation before I can grade it. Please put "COS557" in the message title to help me keep track of your messages.

A good presentation should include:

- a brief overview of the paper

- outline of major methods and findings

- analysis of what the paper did well

- analysis of problems/issues with the approach

- what is the future (don't just retype the "future work" section, we're looking for your analysis here)

Paper Quizzes: At the *start* of some classes (chosen randomly), we'll have a very short quiz with one question based on one of the papers you were to read for that class. These will not be very detailed questions, so if you read the papers, you should have no trouble (on the other hand, you would have a hard time with them if you haven't read the paper). If you don't remember the answer to that specific question, but have read the paper, you can try to write something to convince me you've read the paper (if you do, you get full credit). This also means that if you are late to class and miss a quiz, you get no credit!

Final Projects: Final projects can either be based on independent investigation of a bioinformatics problem related to topics we discussed in class (project papers) or can be based on literature review (review papers).

Example project papers from previous years include devising a new visualization method for microarray data, implementing and evaluating a new method for prediction of pathways, and comparison of multiple missing value estimation techniques. You will likely have some project ideas based on the papers we discuss, and I will suggest more ideas as time gets close to selecting your final project. Project papers should be up to 5 single spaced 12pt pages, including figure legends, but excluding figures.

Review papers need to provide a detailed review of current literature in a specific area of bioinformatics we discussed in class. Some examples include review of classification methods for microarray data, review of biclustering methods in bioinformatics, review of pathway prediction methods, current literature on data integration. Any of these topics would work, and if there is another topic you are interested in, discuss it with me. For examples on how such papers are written, you can refer to the published review papers we are reading in the course.

Final project work cannot be the same as work you are doing for your JP, your thesis, or your graduate research. However, it is perfectly fine for it to be a clearly defined independent project in your area of study (thesis, JP, graduate research area). For example, a student working on function prediction based on microarray data (for JP/thesis/grad research) could write a review of microarray-based classification methods or could do a project based on bi-clustering methods in microarrays. However, a project directly on function prediction based on microarray data would not be appropriate. If in doubt, ask me.

Final project write ups ***are due by 5pm on May 11th***, with no exceptions. Start working on your project on time – don't leave it all for the reading period. To turn in the writeup, put a printed copy of your report with an honor pledge and time you submitted it in the box outside my office labeled COS 557.

Auditors: Auditors are welcome, but every auditor must present one presentation and participate in all discussions (but does not need to do the final project). If you come, you should read the papers – otherwise there's no point to the course. If you want to audit, you must also come to all or most (>85%) classes – “drop in” auditing is not appropriate for this course, and you would likely not get anything out of it anyway.