Course Syllabus
BIOL6385=BMEN6389 – Computational Biology
http://www.utdallas.edu/~prr105020/biol6385/

Course Information
Course number: BIOL6385=BMEN6389
Title: Computational Biology
Room: SLC 2.203
Time: Tuesday & Thursday 2:30pm-3:45pm, Semester: Spring 2014 (Jan 15-May 2, 30-Units)

Professor Information
Michael Q. Zhang
Tel: 972-883-2523
Email (preferred mode of communication): michael.zhang@utdallas.edu
Office hours: After Class and By Appointment
Office location: NSERL 4.742

Associate instructor: Dr. Pradipta Ray [prr105020@utdallas.edu]
Office hours: By Appointment
Office location: NSERL 4.74 Cubicle Complex

Teaching Assistant Information
Yuxuan Liu <yxl107120@utdallas.edu> (Official)
Office Hours: Monday 5-6 pm
Office Location: NSERL 4.74 Cubicle Complex

Course Pre-requisites, Co-requisites, and/or Other Restrictions
Elementary statistics/probability or introductory bioinformatics (such as BIOL5376=BIOL5376, or BIOL5460) or equivalent would be very helpful, but not necessary if having a good college math background. Recommendation: use Perl or Python manipulate sequence data and use R (or Matlab) do statistical computing/graphing. Serious computational biologists should use LINUX environment. There are many tutorials on-line (e.g. http://heather.cs.ucdavis.edu/~matloff/r.html for R Tutorial; http://docs.python.org/2/tutorial/ for Python Tutorial; http://perl-tutorial.org/ for Perl Tutorial)

Course Description
Computational biology (3 semester hours): Machine learning and probabilistic graphical models have become essential tools for analyzing and understanding complex systems biology data in biomedical research. This course introduces fundamental principles and methods behind the most important high throughput data analysis tools. Applications will cover molecular evolutionary models, DNA/protein motif discovery, gene prediction, high-throughput sequencing and microarray data analysis, computational modeling gene expression regulation, and biological pathway and network analysis.

Textbooks and other Material
(Required):
 Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids by Richard Durbin, S. Eddy, A. Krogh, G. Mitchison; Cambridge University Press, 1998. ISBN: 0521629713 (2006 reprinting). Click here for copies in libraries near you.

(Optional):

 Statistical Methods in Bioinformatics : An Introduction (Statistics for Biology and Health) by Warren J. Ewens, Gregory R. Grant; Springer, 2005. ISBN: 0387400826. Click here for copies in libraries near you.

Computational Molecular Evolution (Oxford Series in Ecology and Evolution) by Ziheng Yang, Oxford University Press, USA (2006). ISBN-13: 978-0198567028. Click here for copies in libraries near you.

Machine Learning and Probabilistic Graphical Model References:
Learning From Data (http://work.caltech.edu/lectures.html) by Yaser S. Abu-Mostafa.
Machine Learning: A Probabilistic Perspective by Kevin P. Murphy
The Elements of Statistical Learning: Data Mining, Inference, and Prediction by Hastie, Tibshirani and Friedman.

Grading Policy
Grades will be based on the homeworks (50%), midterm (25%) and final exam (25%).

Course & Instructor Policies
Active participation in class room discussion is expected. Attendance is mandatory.

Comp. Biol. Class schedule Spring, 2013:

Week	Date	Topic	Date	Topic	Biological Applications
1	**1/14**	Introduction	1/16	Background (cont'd): Probability Theory	
2	**1/21**	Background (cont'd): Statistical Inference	1/23	Bayes Nets I. Modeling and Estimation	Gene Regulation Network
3	1/28	Bayes Nets II. Bayes net/inference	1/30	HW1, Alignment I. Scoring Models	Decision Trees Phylogenetic trees
	1/29	***Last day to drop without a "W"***			
4	2/4	Alignment II. Dynamic Progr. & Global Align.	2/6	Alignment III. Local Align. & heuristic	Sequence homology detection

5	2/11	Karlin-Altschul Statistics & Score Significance	2/13	Markov Nets I. Markov Chain	Database search Dayhoff matrix
6	2/18	HW2, Markov Nets II HHM: Segmentation	2/20	Markov Nets III HMM: Viterbi, Forward/Backward	CpG Islands ChIP-seq
7	2/25	Markov Nets IV B-W algorithm, PHMM	2/27	Markov Nets V Profile HMM	Sequence functional domains
8	3/4	**Midterm Exam review**	3/6	**Midterm Exam** (covers all material up to 2/27)	
9	3/10	SPRING BREAK	3/15	SPRING BREAK	No lecture
10	3/18	Markov Nets VI HMM vs. CRF	3/20	Multiple Alignment	Gene prediction Functional site detection
11	3/25	Evolutionary models I.	3/27	Evolutionary models II.	Evolution dynamics
12	4/1	HW3,Phylogenetic trees I.	4/3	Phylogenetic trees II.	Phylogenetic relationship
	4/3	*Last day to withdraw with an automatic "W"* (Graduate students)			
13	4/8	Motif finding I (Greedy)	4/10	Motif finding II (EM, Gibbs sampling)	TF-binding sites
14	4/14	Evaluation of significance of motifs	4/17	Discriminant motif finding (DWE/DME) Functional motif finding (Regression: CART,MARS)	Cis-regulatory elements discovery
15	4/22	SVM and Kernel method	4/24	Ensemble learning, Boosting (Random Forest)	DNA methylation prediction, Promoter, enahncers
16	5/6	**Final Exam review**	5/8	**Final Exam**	

GRADING: HW (50%); two 75-min open book exams (25% each).

Field Trip Policies. Off-campus Instruction and Course Activities: Off-campus, out-of-state, and foreign instruction and activities are subject to state law and University policies and procedures regarding travel and risk-related activities. Information regarding these rules and regulations may be found at the website address: http://www.utdallas.edu/BusinessAffairs/Travel_Risk_Activities.htm.
Additional information is available from the office of the school dean. Below is a description of any travel and/or risk related activity associated with this course.

Student Conduct & Discipline: The University of Texas System and The University of Texas at Dallas have rules and regulations for the orderly and efficient conduct of their business. It is the responsibility of each student and each student organization to be knowledgeable about the rules and regulations which govern student conduct and activities. General information on student conduct and discipline is contained in the UTD publication, *A to Z Guide*, which is provided to all registered students each academic year. The University of Texas at Dallas administers student discipline within the procedures of recognized and established due process. Procedures are defined and described in the *Rules and Regulations, Board of Regents, The University of Texas System, Part 1, Chapter VI, Section 3*, and in Title V, Rules on Student Services and Activities of the university's *Handbook of Operating Procedures*. Copies of these rules and regulations are available to students in the Office of the Dean of Students, where staff members are available to assist students in interpreting the rules and regulations (SU 1.602, 972/883-6391). A student at the university neither loses the rights nor escapes the responsibilities of citizenship. He or she is expected to obey federal, state, and local laws as well as the Regents' Rules, university regulations, and administrative rules. Students are subject to discipline for violating the standards of conduct whether such conduct takes place on or off campus, or whether civil or criminal penalties are also imposed for such conduct.

Academic Integrity: The faculty expects from its students a high level of responsibility and academic honesty. Because the value of an academic degree depends upon the absolute integrity of the work done by the student for that degree, it is imperative that a student demonstrate a high standard of individual honor in his or her scholastic work. Scholastic dishonesty includes, but is not limited to, statements, acts or omissions related to applications for enrollment or the award of a degree, and/or the submission as one's own work or material that is not one's own. As a general rule, scholastic dishonesty involves one of the following acts: cheating, plagiarism, collusion and/or falsifying academic records. Students suspected of academic dishonesty are subject to disciplinary proceedings. Plagiarism, especially from the web, from portions of papers for other classes, and from any other source is unacceptable and will be dealt with under the university's policy on plagiarism (see general catalog for details). This course will use the resources of turnitin.com, which searches the web for possible plagiarism and is over 90% effective.

Email Use: The University of Texas at Dallas recognizes the value and efficiency of communication between faculty/staff and students through electronic mail. At the same time, email raises some issues concerning security and the identity of each individual in an email exchange. The university encourages all official student email correspondence be sent only to a student's U.T. Dallas email address and that faculty and staff consider email from students official only if it originates from a UTD student account. This allows the university to maintain a high degree of confidence in the identity of all individual corresponding and the security of the transmitted information. UTD furnishes each student with a free email account that is to be used in all communication with university personnel. The Department of Information Resources at U.T. Dallas provides a method for students to have their U.T. Dallas mail forwarded to other accounts.

Withdrawal from Class: The administration of this institution has set deadlines for withdrawal of any college-level courses. These dates and times are published in that semester's course catalog. Administration procedures must be followed. It is the student's responsibility to handle

withdrawal requirements from any class. In other words, I cannot drop or withdraw any student. You must do the proper paperwork to ensure that you will not receive a final grade of "F" in a course if you choose not to attend the class once you are enrolled.

Student Grievance Procedures: Procedures for student grievances are found in Title V, Rules on Student Services and Activities, of the university's *Handbook of Operating Procedures*. In attempting to resolve any student grievance regarding grades, evaluations, or other fulfillments of academic responsibility, it is the obligation of the student first to make a serious effort to resolve the matter with the instructor, supervisor, administrator, or committee with whom the grievance originates (hereafter called "the respondent"). Individual faculty members retain primary responsibility for assigning grades and evaluations. If the matter cannot be resolved at that level, the grievance must be submitted in writing to the respondent with a copy of the respondent's School Dean. If the matter is not resolved by the written response provided by the respondent, the student may submit a written appeal to the School Dean. If the grievance is not resolved by the School Dean's decision, the student may make a written appeal to the Dean of Graduate or Undergraduate Education, and the deal will appoint and convene an Academic Appeals Panel. The decision of the Academic Appeals Panel is final. The results of the academic appeals process will be distributed to all involved parties. Copies of these rules and regulations are available to students in the Office of the Dean of Students, where staff members are available to assist students in interpreting the rules and regulations.

Incomplete Grade Policy: As per university policy, incomplete grades will be granted only for work unavoidably missed at the semester's end and only if 70% of the course work has been completed. An incomplete grade must be resolved within eight (8) weeks from the first day of the subsequent long semester. If the required work to complete the course and to remove the incomplete grade is not submitted by the specified deadline, the incomplete grade is changed automatically to a grade of **F**.

Disability Services: The goal of Disability Services is to provide students with disabilities educational opportunities equal to those of their non-disabled peers. Disability Services is located in room 1.610 in the Student Union.

The contact information for the Office of Disability Services is: The University of Texas at Dallas, SU 22. PO Box 830688. Richardson, Texas 75083-0688. (972) 883-2098 (voice or TTY)

Essentially, the law requires that colleges and universities make those reasonable adjustments necessary to eliminate discrimination on the basis of disability. For example, it may be necessary to remove classroom prohibitions against tape recorders or animals (in the case of dog guides) for students who are blind. Occasionally an assignment requirement may be substituted (for example, a research paper versus an oral presentation for a student who is hearing impaired). Classes enrolled students with mobility impairments may have to be rescheduled in accessible facilities. The college or university may need to provide special services such as registration, note-taking, or

mobility assistance. It is the student's responsibility to notify his or her professors of the need for such an accommodation. Disability Services provides students with letters to present to faculty members to verify that the student has a disability and needs accommodations. Individuals requiring special accommodation should contact the professor after class or during office hours.

Religious Holy Days: The University of Texas at Dallas will excuse a student from class or other required activities for the travel to and observance of a religious holy day for a religion whose places of worship are exempt from property tax under Section 11.20, Tax Code, Texas Code Annotated. The student is encouraged to notify the instructor or activity sponsor as soon as possible regarding the absence, preferably in advance of the assignment. The student, so excused, will be allowed to take the exam or complete the assignment within a reasonable time after the absence: a period equal to the length of the absence, up to a maximum of one week. A student who notifies the instructor and completes any missed exam or assignment may not be penalized for the absence. A student who fails to complete the exam or assignment within the

prescribed period may receive a failing grade for that exam or assignment. If a student or an instructor disagrees about the nature of the absence [i.e., for the purpose of observing a religious holy day] or if there is similar disagreement about whether the student has been given a reasonable time to complete any missed assignments or examinations, either the student or the instructor may request a ruling from the chief executive officer of the institution, or his or her designee. The chief executive officer or designee must take into account the legislative intent of TEC 51.911(b), and the student and instructor will abide by the decision of the chief executive officer or designee.

These descriptions and timelines are subject to change at the discretion of the Professor.

Some additional guidelines for the course as well as these policies will be kept updated at: http://www.utdallas.edu/~prr105020/biol6385/about.html