Syllabus for MMG1345H: Foundational Computational Biology II (FCB II)

Course Instructor: Fritz Roth
Course Time and Dates: May 19 - Jun 23 (Thursdays 1-3 pm).

Course Location: Donnelly Centre Black Room or Zoom (https://utoronto.zoom.us/j/6808789917) if necessary

Total Student Contact Hours (0.25 FCE): 12 hours of lectures plus 12 hours of (optional) TA office hours.

Prerequisite: Foundational Computational Biology I or approval from the instructor. This course is targeted to 1st year graduate students with substantial computational biology experience. For example, students with computational biology experience which is too advanced for MMG1004 (A Practical Course in Programming for Biologists) will be encouraged to first take Foundational Computational Biology I prior to taking Foundational Computational Biology II. Priority will be given to first-year Molecular Genetics Department admitted graduate students who have already taken Foundational Computational Biology I. On a case-by-case basis, the instructors will also admit Molecular Genetics Department students in year 2 or later as well as students from other UofT Departments.

Enrollment is subject to Instructor approval, and will require:
1) evidence of comfort with computer programming and
2) excellence in two or more quantitative subjects, which may include: calculus, linear algebra, probability/statistics or other mathematics courses.

Course Objectives and Learning Outcomes: The Foundational Computational Biology (FCB) courses are two 6-meeting topic courses, offered through the Molecular Genetics Graduate program, covering foundational concepts and current applications for computational biology and bioinformatics.

Course Description and Format: The course will consist of lectures as well as hands on assignments. Assignments will be both pen-and-paper and practical assignments requiring programming (e.g., Python) or statistical environments (e.g., R).

Delivery of the course:

Lecture 1: Unsupervised Learning
Lecturer: Kieran Campbell
Continuous latent variable models: principal component analysis, factor analysis, non-negative matrix factorization, autoencoders
Discrete latent variable models: Gaussian and non gaussian mixture models, expectation maximization
Model complexity for unsupervised models

ASSIGNMENT #2.1 MADE AVAILABLE

**Lecture 2: Supervised Learning**
Lecturer: Fritz Roth
Feature selection, no free lunch
Linear classification, Advanced non-linear classification (random forests, decision trees, kNN, SVM, ensemble learning)
Boosting, bagging
Overfitting vs. performance inflation

**Lecture 3: Sequence Analysis 2 & Machine Learning Evaluation**
Lecturer: Alan Moses
Lecture: Sequence motif analysis: consensus seq, regular expressions, PWMs, profile HMMs
Lecturer: Michael Hoffman
Lecture: Evaluating machine learning
Paper discussion

ASSIGNMENT #2.2 MADE AVAILABLE

**Lecture 4: Bayesian Inference & Deep Learning**
Lecturer: Kieran Campbell
Bayesian inference, MCMC: Metropolis Hastings, Gibbs sampling. Variational methods.
Model diagnostics, Credible intervals.
Neural networks, MLPs, CNNs, RNNs, Auto-encoders, Variational Auto-Encoders, Backpropagation. Optimization

ASSIGNMENT #2.1 DUE

**Lecture 5: Applications of Machine Learning**
Lecturer: Bo Wang
Lecture: Applications of machine learning for single-cell analysis
Lecturer: Anne Martel
Lecture: Machine Learning for Digital Pathology

**Lecture 6: Sequence Analysis 3: Next-Generation Sequencing**
Lecturer: Jared Simpson
Lecture: Next-generation sequencing, read alignment, genome assembly and variant calling
Paper discussion

ASSIGNMENT #2.2 DUE
**Method of Evaluation:** The marking scheme will be 75% assignments, 25% participation. The lowest two participation marks will be dropped. Bring a doctor’s note for anything else! The maximum mark on late assignments will be reduced by 5% per day until the maximum penalty of 25% is reached.

**Additional Information:**
Before or during the course, here are some opportunities to brush up on essential background:

**Linear Algebra:**
Strang MIT course: [https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring2010/index.htm](https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring2010/index.htm) 3blue1brown: [https://www.youtube.com/watch?v=kjBOesZCoqc&list=PLZHQQbOWTQDPD3MizzM2xFVFitgF8hE_ab](https://www.youtube.com/watch?v=kjBOesZCoqc&list=PLZHQQbOWTQDPD3MizzM2xFVFitgF8hE_ab)  
KhanAcademy: [https://www.khanacademy.org/math/linear-algebra](https://www.khanacademy.org/math/linear-algebra)  
KNOW: eigenvectors, singular value decomposition, Moore-Penrose pseudo-inverse

**Gradients:**