

STA 414/2104 (Fall 2015): Statistical Methods for Machine Learning and Data Mining

No Class on Monday, Sep 21st.

Instructor :

- Russ Salakhutdinov, Office: Pratt Building, Room 290F
- Email: rsalakhu [at] cs [dot] toronto [dot] edu
- **Lectures: Mondays 2-5pm, MC 254**
- **Office hours:** Mondays 11:00am - 12:00pm in Pratt Building, Room 290F.

Marking Scheme:

- For undergraduate students
 - 3 assignments: 40%
 - 2-hour midterm 20%
 - 3-hour final exam 40%
- For graduate students
 - 3 assignments: 40%
 - 2-hour midterm 20%
 - 3-hour final exam 30%
 - 10% A 12-minute individual presentation on a conference paper that you have read.

Midterm is on TBD:

You can use a nonprogrammable calculator and an 8 by 11 inch Crib Sheet - **Single-sided** .

Final is on TBD.

You can use a nonprogrammable calculator and an 8 by 11 inch Crib Sheet - **Double-sided** .

Course Outline:

This course covers some of the theory and methodology of statistical aspects of machine learning. The preliminary set of topics to be covered include:

- Linear methods for regression, Bayesian linear regression
- Linear models for classification
- Probabilistic Generative and Discriminative models
- Regularization methods
- Model Comparison and BIC
- Neural Networks
- Radial basis function networks
- Kernel Methods, Gaussian processes, Support Vector Machines
- Mixture models and EM algorithm
- Graphical Models and Bayesian Networks

Prerequisite: Either STA302H or CSC411H

Books :

Christopher M. Bishop (2006) [Pattern Recognition and Machine Learning](#), Springer.

You can also use these books for additional reference:

- Machine Learning: A Probabilistic Perspective, by Kevin P. Murphy.

- Trevor Hastie, Robert Tibshirani, Jerome Friedman (2009) [The Elements of Statistical Learning](#)
- David MacKay (2003) [Information Theory, Inference, and Learning Algorithms](#)