CMPT 727 - G100 Statistical Machine Learning

Instructor(s): Maxwell Libbrecht

Calendar Objective/Description:
Statistical foundation for machine learning algorithms, emphasizing bias-variance tradeoff. Students will learn principles for choosing effective methods and tailoring them to fit a given learning problem. Potential topics include probabilistic graphical models, maximum likelihood estimation, latent variables and the EM algorithm, convex optimization, and variational and sampling-based methods.

Instructor’s Objectives:

1. Why we’re offering this course:
   Machine learning now plays a central role in hundreds of fields. All learning methods have common underpinnings based on probability and statistics, but these are not widely understood. It is common for machine learning practitioners to view ML algorithms as a black boxes, without understanding their inner workings. This view makes it very difficult to improve a machine learning method beyond resorting to uninformed trial and error. This course aims to give students a statistical foundation for machine learning.

   After you complete this course, you will be able to:
   (2) Understand concepts such as: overfitting/underfitting, likelihood, bias-variance tradeoff.
   (3) Choose between machine learning methods and foresee which will perform best.
   (4) Understand what aspects of a given task influence machine learning performance.

Who should take this course?
This course is intended for graduate students with an interest in machine learning or big data. You should take this course if:
(1) You intend to use or develop machine learning in your research or work.
(2) You are interested in fields where machine learning and big data is important including: vision, natural language processing/understanding, medical imaging, robotics, smart cities.
(3) You want to understand machine learning at a deep level.

Prerequisites:
None

Topics:
- Probabilistic graphical models
- Discrete and continuous distributions
- Maximum likelihood estimation
- EM algorithm inference
- Sampling based inference methods and MCMC
- Variational inference

Grading:
Grading will be based on assignments (written and coding), midterm and final exams.
**Required Books:**

Machine learning, a probabilistic perspective, Kevin Murphy, 9780262018029

**Academic Honesty Statement:**

Academic honesty plays a key role in our efforts to maintain a high standard of academic excellence and integrity. Students are advised that ALL acts of intellectual dishonesty will be handled in accordance with the SFU Academic Honesty and Student Conduct Policies (http://www.sfu.ca/policies/gazette/student.html).