CPSC 540: Machine Learning

Term 2: Mark Schmidt

Machine Learning

- Automatically detecting patterns in data to:
  - Make predictions.
  - Make decisions.
- One of the fastest-growing areas:
  - We are collecting data at unprecedented rates.
  - ML is a key tool in making sense of this data.
  - Startups being bought for $$$$$$$.
- Key technology behind:
  - Optical character recognition, spam filtering, Amazon product recommendation, Microsoft Kinect, speech recognition, object detection, Facebook M, and self-driving cars.
  - Your graduate research???

CPSC 540

- Course Objectives:
  - Develop your own machine learning models.
  - Understand and extend math behind models.
  - Implement large-scale methods.
- Big class:
  - 52 CS and ECE students registered.
  - More coming from other departments and 340.

CPSC 340 vs. CPSC 540

- Previous years:
  - Substantial overlap between CPSC 340 and 540.
- This year:
  - Both courses cover a few core topics (~25%).
  - 340: more data mining and ML applications.
  - 540: research-level methods and theory.
- 540 will not cover important but useful topics:
  - Random forests, clustering, collaborative filtering.
- For most students: take CPSC 340 first.

CPSC 340

1) Data preprocessing and summary statistics
2) Frequency-based supervised learning (naïve Bayes, nearest neighbours, decision trees, random forests).
3) Data clustering and association rules.
4) Linear prediction, regularization, and kernels.
5) Neural networks and deep learning.
6) Outlier detection, dimensionality reduction, and visualization.
7) Link analysis and collaborative filtering.
8) Sequences and time-series.

CPSC 540

- Linear prediction, regularization, and kernels.
- Neural networks and deep learning.
- Density estimation.
- A subset of the following:
  - Bayesian methods.
  - Latent-factor models.
  - Large-scale methods.
  - Online learning.
  - Graphical models.
  - Reinforcement learning.
  - Spectral methods.
  - Sub-modularity.

Prerequisites.

- Ideal prerequisites:
  - Linear algebra, probability, multivariate calculus, scientific computing, algorithms and complexity, statistical inference, machine learning.
  - Missing one is ok, course could be very hard if missing more than one (take CPSC 340 instead).
  - First two weeks of classes: fill out a form describing how you fulfill the prerequisites.
  - Weekly tutorials to help with the background.
  - Consider auditing if you don't have prerequisites.

Course Work

- Assignments: math and Matlab programming.
- Midterm: similar to assignment questions.
- Final project: small research project.
- Coding project: expand on matLearn project.

- I hope to see you there!