Lecture 0 - Overview of Machine Learning

Dahua Lin

The Chinese University of Hong Kong

About this Course

- This is a *graduate level* introduction to *advanced statistical learning*.

About this Course

- This is a *graduate level* introduction to *advanced statistical learning*.
- **No Exams!**
- Topic driven
- For each topic:
  - Introductory lecture
  - Paper reading and Homework
  - In-class discussion
- You will present a paper/subject at the end of the course

What is Machine Learning?

Machine learning *is a scientific discipline that explores the construction and study of algorithms that can learn from data. Such algorithms operate by building a model based on inputs and using that to make predictions and decisions, rather than following only explicitly programmed instructions. – Wikipedia*

Elements of Machine Learning

- Elements:
  - Data
  - Model
  - Learning Algorithms
  - Prediction
- Learn from old data, make predictions on new data. The common aspects of both the old and new data are captured by the model.

Please write down five machine learning algorithms that you know.

Don't write *Deep Learning*.

Overview of Machine Learning

- Tasks
  - Supervised learning
  - Unsupervised learning
  - Semi-supervised learning
  - Reinforcement learning
- Problems
  - Classification
  - Regression
  - Clustering
  - Dimension Reduction
  - Density Estimation

What this Course is About

- The course is *not* to teach you:
  - Support Vector Machine
  - Linear Regression
  - ...
  - Deep Learning
- Instead, you are going to learn *foundational theories and tools* for developing *your own* models and algorithms.

Topics

- Exponential family distributions and conjugate prior
- Generalized linear model
- Empirical risk minimization and Stochastic gradient descent
- Proximal methods for optimization
- Graphical models: Bayesian Networks and Markov random fields
- Sum-product and max-product algorithms, Belief propagation
- Variational inference methods
- Markov Chain Monte Carlo
- Gaussian Processes and Copula Processes
- Handling Big Data: Streaming process and Core sets