ECE-GY 6143 / Intro to Machine Learning, Spring '20

Chinmay Hegde

Pre-requisites

Graduate status with undergraduate-level probability theory; familiarity with linear algebra and optimization; familiarity with coding/programming.

Description

This course offers a beginner-level introduction to machine learning (ML) from both a theoretical and practical perspective.

Objectives

Upon successful completion of this course, you will be able to:

- grasp the basics of modern ML,
- mathematically formulate problems arising in practical applications,
- propose algorithmic solutions to these problems using a variety of ML models,
- implement software prototypes of several algorithms of interest,
- apply these prototypes on real-world datasets, and
- evaluate their performance using statistical techniques.

Reading material

We will follow our course lecture notes for primary reading material. Lecture notes will be posted either before or right after class.

As supplementary reading, there are plenty of excellent resources for machine learning available in book form or on the web. Classical textbooks include "Pattern Recognition and Machine Learning" by Bishop, "Pattern Classification" by Duda-Hart-Stork, "Foundations of Data Science" by Hopcroft-Kannan-Blum, "Elements of Statistical Learning" by Hastie-Tibshirani, and Stanford's CS229 course notes by Andrew Ng.

Grading policy

- 40% - Homework assignments

- 30% - Midterm exam
- 30% - Final exam

There will be 6 homework assignments and 2 exams. Each homework assignment will involve both theory and programming problems. We *strongly* encourage you to use LaTeX to prepare your answers to the theory part. Code submissions should be made in the form of Jupyter Notebooks with self-contained explanations and results.

Exams will be closed book, and will be in similar spirit and difficulty as the homework problems.

Requirements

No prior machine learning knowledge will be assumed. If you have significant ML experience, there is no need to take this class.

A comfortable knowledge of probability and basic linear algebra will be assumed, but we will provide a brief review in order to be self-contained. Familiarity with optimization is a plus.

There will be a significant programming component to this course, and class/homework exercises will be in Python. You do not need to know Python a priori, but should be willing to pick it up.

Tentative list of topics

- Mathematical basics of ML
- Regression (linear, multivariate, logistic)
- Classification (kNN, perceptrons, SVMs, kernel machines)
- Neural networks, including MLPs, convnets, and autoencoders
- Data summarization (PCA, k-means, topic models)
- Applications in computer vision, NLP, time series forecasting.
- Recent ML advances.

Contact info

Chinmay Hegde (chinmay.h@nyu.edu), 1157, 370 Jay St. (Office hours: Tuesday 3:30-5pm.)