Applied Machine Learning - Welcome

Contents

Introduction

This course provides an overview of key algorithms and concepts in machine learning, with a focus on applications. Introduces supervised and unsupervised learning, including logistic regression, support vector machines, neural networks, Gaussian mixture models, as well as other methods for classification, regression, clustering, and dimensionality reduction. Covers foundational concepts such as overfitting, regularization, maximum likelihood estimation, generative models, latent variables, and non-parametric methods. Applications include data analysis on images, text, time series, and other types of data using modern software tools such as numpy, scikit-learn, and pytorch.

What's Inside

Machine Learning Algorithms

A broad overview of algorithms across ML: generative models, SVMs, tree-based algorithms, neural networks, gradient boosting, etc.

Mathematical Foundations

Rigorous definitions of key concepts including: overfitting, regularization, maximum likelihood estimation, latent variable models.

Algorithm Implementations

Most algorithms are implemented from scratch in Python using standard libraries such as numpy, scipy, or sklearn.

Prerequisites

This masters-level course requires a background in mathematics and programming at the level of introductory college courses. Experience in Python is recommended, but not required. A certain degree of ease with mathematics will be helpful.

- Programming experience (ideally Python; Cornell CS 1110 or equivalent)
- Linear algebra. (Cornell MATH 2210, MATH 4310 or equivalent)
- Statistics and probability. (Cornell STSCI 2100 or equivalent)

Instructors

These lecture notes accompany CS5785 Applied Machine Learning at Cornell University and Cornell Tech, as well as the open online version of that course. They are based on materials developed at Cornell by:

- **Volodymyr Kuleshov**, Assistant Professor, Computer Science, Cornell Tech
- **Nathan Kallus**, Associate Professor, Operations Research, Cornell Tech
- **Serge Belongie**, Professor, Computer Science, University of Copenhagen

The open version of CS5785 and the accompanying online lectures have been produced by **Hongjun Wu**. We are also grateful to over a dozen teaching assistants that have helped with drafts of these lecture notes.

Table of Contents