Machine Learning Principles
Computer Science 461
4 credits
Spring 2023

Instructor: Qiong Zhang
Meeting location: SEC 205
Prerequisites: M 250 (Introductory Linear Algebra), CS 112 (Data Structures), CS 206 (Introduction to Discrete Structures II) OR M 477 (Mathematical Theory of Probability) OR S 379 (Basic Probability Theory)
Meeting Times: M/H 10:20am-11:40am
Email: qiong.z@rutgers.edu

Course Description

This course is a systematic introduction to machine learning, covering theoretical as well as practical aspects of the use of statistical methods. Topics include linear models for classification and regression, support vector machines, regularization and model selection, and introduction to deep learning. Machine learning is a vast, fast-changing field. The course is designed to provide a solid starting point by focusing on technical foundations. Students will learn fundamental concepts in the field such as generalization and model selection. Students will also learn how to implement key techniques in the field from scratch including gradient descent, kernel trick, boosting, and expectation maximization. The course content will be delivered by a combination of written assignments that require rigorous proofs and self-contained Python Juypter Notebooks that require completing missing portions and empirically investigating the behavior of models on datasets. The course will reinforce learning with a series of short quizzes throughout the semester.The course content is designed to be accessible to all SAS students regardless of their majors who have an adequate background in linear algebra, calculus, probability, and programming.

Learning Objectives

The course objectives are: (1) understanding the goals, capabilities, and principles of machine learning, (2) acquiring mathematical tools to formalize machine learning problems, and (3) acquiring implementation skills to build practical machine learning systems. At the end of the course, students will have the knowledge of core machine learning techniques, so that they can use them to solve real-world prediction problems using basic Python and a small set of libraries. The knowledge they acquire in this course will directly transfer to any future courses they may take that go deeper into specific application areas of machine learning.

Text/Resources

There is no required text. Lecture slides are self-contained. To augment lectures, there will be optional reading from publicly available online textbooks such as:

Machine Learning, (Tom Mitchell, 1997)
Machine Learning: A Probabilistic Perspective (Murphy, 2012)

There will also be suggested reading from other publicly available technical notes written by various instructors/researchers.

Coursework Requirements

To assess that students have acquired basic literacy in all the concepts, tools, and techniques they are taught, they will be given 3 quizzes (each 15%) periodically through the semester.

There will be five assignments each given 2 weeks to complete (each 10%). The assignments are designed to be challenging but rewarding. Collaboration is allowed and encouraged as long as students (1) write their own solutions entirely on their own, and (2) specify names of student(s) they collaborated with in their writeups. The students will need to digest and reformulate the lecture content in order to successfully complete the assignments. Discussion about assignment problems on Canvas is encouraged.

Students are expected to attend all classes and take notes on the most basic and important concepts discussed in each class. There will be ten in-class surveys distributed randomly across the semester (each 0.5%). They consist of short true and false questions which serve as attendance and attention check for that class.

Grade Evaluation

Quizzes	45%
Assignments	50%
Attendance survey	5%

Week 1
  General Introduction (01/19)

Week 2
  Decision Tree Learning (01/23, 01/26)

Week 3
  Estimating Probabilities (01/30, 02/02)

Week 4
  Naïve Bayes (02/06, 02/09)

Week 5
  Logistic Regression (02/13, 02/16)

Week 6
  Generalization I. (02/20)
  Quiz 1 (02/23)

Week 7
  Generalization II. (02/27)
  Model Selection I. (03/02)

Week 8
  Model Selection II. (03/06)
  Clustering I. (03/09)

Spring break 03/11-03/19

Week 9

Clustering II. (03/20)
SVM and Kernel I. (03/23)

Week 10

SVM and Kernel II. (03/27)
Linear Regression I. (03/30)

Week 11

Quiz 2 (04/03)
Linear Regression II. (04/06)

Week 12

Neural Network Basics (04/10, 04/13)

Week 13

Deep Neural Networks (04/17)
Reinforcement Learning I (04/20)

Week 14

Reinforcement Learning II (04/24)
Advanced topics, Review (04/27)

Week 15

Quiz 3 (05/01)

Academic Integrity Policies

Rutgers University regards acts of dishonesty (e.g. plagiarism, cheating on examinations, obtaining unfair advantage, and falsification of records and official documents) as serious offenses against the values of intellectual honesty. These policies are detailed here:

https://nbprovost.rutgers.edu/academic-integrity-students

In addition, the Computer Science departments has established policies for academic integrity that pertain specifically to programming assignments:

https://www.cs.rutgers.edu/academics/undergraduate/academic-integrity-policy/programming-assignments