THE MATHEMATICS OF
MACHINE LEARNING

MAY 16-22, 2020

IAS | WOMEN AND MATHEMATICS

ORGANIZERS

SUN-YUNG ALICE CHANG Princeton University

MICHELLE HUGUENIN Institute for Advanced Study

DUSA MCDUFF Barnard College & Columbia University

MARGARET READDY University of Kentucky

TERNG LECTURE

Cynthia Rudin, Duke University: Introduction to Interpretable Machine Learning

Machine learning is now used throughout society and is the driving force behind the accuracy of online recommendation systems, credit-scoring mechanisms, healthcare systems and beyond. Machine learning models have the reputation of being "black boxes" meaning that their computations are so complicated that no human would be capable of understanding them. However, machine-learning models do not actually need to be black boxes. It is possible, with some mathematical sophistication, to derive algorithms for machine learning that produce models that are interpretable by humans.

This self-contained short course will introduce the basic concepts of supervised machine learning, including loss functions, overfitting and regularization, and how interpretability is a form of regularization. It will focus on two important topics in interpretable machine learning: (1) sparse decision trees and (2) interpretable neural networks. For each topic, a brief history of the field, going all the way to the state-of-the-art in current methodology, and an introduction of specific technical and mathematical challenges will be presented.

UHLENBECK LECTURE

Maria Florina Balcan, Carnegie Mellon University: Foundations for Learning in the Age of Big Data

With the variety of applications of machine learning across science, engineering, and computing in the age of Big Data, re-examining the underlying foundations of machine learning has become imperative. This lecture will introduce new models and algorithms for important emerging paradigms, specifically, interactive learning and distributed learning.

Most classic machine learning methods depend on the assumption that humans can annotate or label all the data available for training. However, many modern machine-learning applications have massive amounts of unannotated or unlabeled data. Consequently, there is new interest in machine learning and its application to design-active learning algorithms that intelligently select which data to request to be labeled, with the goal of dramatically reducing the human labeling effort. This course will discuss recent advances on designing active learning algorithms with provable guarantees that are computationally efficient, noise tolerant, and enjoy strong label efficiency guarantees.

Also to be reviewed is the problem of learning from distributed data and the analysis of fundamental algorithmic and communication complexity questions involved. A framework where massive amounts of data is distributed among several locations will be broadly considered. The goal is to learn a low-error predictor with respect to the overall distribution of data using as little communication and as few rounds of interaction as possible. Finally, the course will consider general upper and lower bounds on the amount of communication needed to learn a given class, as well as broadly applicable techniques for achieving communication-efficient learning.

PREREQUISITES: Undergraduate knowledge of linear algebra & probability, as well as an introductory computer science course.

Application Deadline: February 17, 2020
For more information visit: math.ias.edu/wam/2020

A Program of the Institute for Advanced Study
THE MATHEMATICS OF MACHINE LEARNING

IAS | WOMEN AND MATHEMATICS

PLEASE POST: THE MATHEMATICS OF MACHINE LEARNING

Program for Women and Mathematics
Institute for Advanced Study
1 Einstein Drive
Princeton, NJ 08540
(609) 734-8115
womensprogram@ias.edu

IAS | WOMEN AND MATHEMATICS