Math 728 D: Machine Learning & Data Science, Spring 2019

Course Description

Ever improving sensor technology and computational capacity have provided an unprecedented wealth of data. A drastically increasing impact on society and culture is blatantly reflected by the role of social media. In a perhaps less obvious why virtually each branch of science is affected as well and the newly forming discipline "Data Science" is expected to play a transformative role.

Data Science covers a wide spectrum of activities at interfaces between several different disciplines like Computer Science, Engineering, Economics, Statistics, and Mathematics, due to its many facets with varying emphasis on processing or analyzing large data sets. Of course, there is no clear division between the two corner stones but this course focuses primarily on basic mathematical foundations. Generally speaking, its main emphasis is on providing the main techniques needed for extracting quantifiable information from possibly very large data sets which is commonly referred to as Machine Learning. The spectrum of applications is enormous covering, for instance hyperspectral imaging in material science, genome expression, autonomous driving, face recognition, or web search. Two underlying main task categories are "regression" for unveiling "expected" functional interrelations between often noisy data, and "classification" that aims at assigning one of finitely many labels to an observation or measurement. For instance, does a medical image represent healthy or cancerous tissue? Recurring embedded tasks are "denoising" techniques to provide for instance a clearer view of a noisy image, "clustering" i.e., grouping together portions of data exhibiting strong structural similarities, for instance, in face recognition, or conversely "discriminating" data. A common challenge ln these applications is the fact that the involved mathematical objects involve functions of a large number of variables. Learning mathematical principles and quantification concepts to deal with resulting often counterintuitive effects, is therefore a key objective of the course. Dimension reduction, sparse recovery, dictionary learning, low discrepancy or importance sampling, are examples of such concepts. The actual numerical algorithms for addressing these tasks are often based on convex as well as non-convex optimization or nonlinear approximation techniques which, in particular, play a pivotal role in the "deep learning" paradigm.

Course Overview

The following tentative list of lecture topics is to address the key issues discussed above along with simple motivating examplary applications. The introductory subjects aim at a reasonable level of self - consistency. The list may be subject to later modifications depending on how we are able to progress with the material.

- Basics from functional analysis, linear algebra, and numerical analysis

- Regularization and optimization
  convex and non-convex constrained optimization, descent methods, saddle points

- Basics from probability and information theory
  Probability distributions, concentration inequalities, Bayes rule

- High dimensionality: Dimension reduction, concentration phenomena, approximation and recovery concepts

- Sparse recovery and compressed sensing

- Fundamentals of learning algorithms
  estimators, variance and bias, regression, classification, clustering, supervised- unsupervised learning

- Deep neural and convolutional networks
  basic ingredients, architecture design, network types, optimization aspects, stochastic gradient descent, regularization

- Advanced topics: models and data

Learning Outcomes[1]

On successful completion of this course, graduate students will understand the mathematical and computational prerequisites required for applying and designing learning algorithms. This includes a sound familiarity with relevant optimization concepts. A proper understanding of the main estimation tasks and of related performance measures for estimators is important. The students will be then be able to apply these principle in conjunction with major types of learning algorithms such as deep networks.

He central goal of this course is to prepare for subsequent in depth studies of topical developments in machine learning and data science such as "learning models" or synthesizing data driven and model based methodologies to be taken up in subsequent specialized courses or seminars.

Suggested Reading

Avrim Blum, John Hopcroft and Ravindran Kannan: Foundations of Data Science

Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep Learning, The MIT Press, Cambridge, Massachusetts, London, England

Course Assignments

Graduate students are required to actively contribute to discussion sessions, occasionally present homework assignments which will then be discussed. This is to help developing and continuously growing a good understanding of the conceptual foundations and related implementation skills. The weight of this assignment complex is 60%. The students are also expected to complete a small accompanying project weighing 40 % which is to consolidate understanding and trains application skills.

Grading Scale

90-100: A

87-89: B+

80-86: B

77-79: C+

[1] if this is a course that can be taken for either undergraduate or graduate credit, there **must** be separate learning outcomes for undergraduate students and graduate students. There also must be **additional** assignments for graduate students)

70-76: C

60-69: D

0-59: F

Short description:

This course addresses the basis mathematical concepts modern machine learning algorithms are based upon. This concerns central machine learning tasks like regression or classification as well as the main constituents of corresponding learning algorithms including recent developments in “deep learning”. The course targets building conceptual understanding along with basic implementation skills.