TOPICS IN COMPUTER SCIENCE
OPTIMIZATION FOR MACHINE LEARNING AND DATA SCIENCE

ITCS 6010/8010 - Fall 2023

Course Syllabus (Version of August 21, 2023)

Instructor: Dr. Christian Kümmerle **Email:** kuemmerle@uncc.edu	**Course Number:** ITCS 6010/8010
Course Lectures: Mondays, 2:30-3:45 PM in Woodward 140 Wednesdays, 2:30-3:45 PM in CHHS 159 (Week 1 on Zoom)	**Instructor's Office Hours:** Mondays, 4:00-5:00 PM Tuesdays, 3:00-4:00 PM in Woodward 410D & Zoom
Teaching Assistant: Kalvik Jakkala **Email:** (kjakkala@charlotte.edu)	**TA's Office Hours:** Wednesdays, 3:45-4:45 PM (Woodward 200 & Zoom) Friday, 11:00 AM-12:00 PM (Zoom)
Zoom information:	https://charlotte-edu.zoom.us/j/95526959172

Summary: Optimization formulations and algorithms play a central role in modern data science, statistics and machine learning. This graduate-level course is an introduction to the basics of continuous optimization, with an emphasis on techniques that are most relevant to data science and machine learning and are amenable to large-scale implementations.
We also discuss basic convergence properties and computational aspects of relevant optimization techniques. The relationship and applications of optimization to common machine learning models such as (generalized) principal component analysis, sparse and logistic regression, regularized empirical risk minimization and the training of deep neural networks is another main focus of this course.
In addition to theory and algorithms, common optimization modeling software and computer solvers will be introduced to students.

Learning Objectives:

- Know how to use optimization to fit, train and evaluate machine learning models
- Understand the fundamental statistical and computational challenges in models motivating optimization
- Identify and understand the differences between linear, convex, non-convex, smooth and non-smooth optimization problems to select appropriate solution algorithms
- Develop a solid understanding of the basics of continuous optimization
- Understand the trade-offs and limitations of different numerical optimization methods
- Learn how to work with optimization modeling software and solvers
- Understand research papers at the intersection of optimization and data science/ machine learning (more focus on this in ITCS 8010)
- Gain practical experience implementing state-of-the-art customized optimization algorithms (more focus on this in ITCS 8010)

Prerequisites/ Expected Background:

- Calculus including multivariate differentiation,
- Linear algebra
- One course in probability and statistics
- Some experience with Python
- Recommended: One course in machine learning

Course Topics (Tentative List):

Data Models and Applications in Machine Learning

- Unsupervised learning: Principal component analysis, k-means clustering
- Supervised learning: Logistic regression, sparse regression, matrix completion, recommender systems, computer vision
- Training of deep neural networks
- Optimal transport: Image matching, shape registration

Algorithms and Theory of Continuous Optimization

- Basics of convex analysis & optimization
- Multidimensional calculus: gradients, optimality conditions
- Gradient and descent methods
- Accelerated gradient
- Stochastic gradient descent
- Coordinate descent
- Proximal gradient
- Subgradient methods
- Second-order methods
- Variable metric techniques
- Conditional gradient
- Minimax Optimization

Homeworks:

- Weekly to biweekly problem solving questions.
- At times, they include a programming exercises in Python.

Collaboration rule: Discussing problems with colleagues is acceptable *if discussion partners are indicated in list of collaborators*. However, answers need to be written down individually. Submission format: Handwritten (scanned as PDF) or PDF created by LaTeX (or LaTeX enabled Markdown).
Please upload your solution to Gradescope.

Midterm: A central part of the assessment in the course is a late midterm (date tbd, between October 25 and November 1). Comprehensive, written, in-person exam. There is **no** final in-person exam (see below for final course project requirements).

Final Course Projects: An important part of the course is a final project that is worked on by a *group* of 2-3 students (6000 level) or individually or in a group of two (8000 level). The goal of the project to learn how to deeply understand an optimization algorithm, learn the ability to use ideas from optimization to solve real-world world problems, gain the ability to implement algorithms/ use optimization modeling software, and/or learn how to read, understand and build on existing research literature.

Possible directions of the course project:

- Apply optimization methods to practical problem (possibly from within data science / machine learning); implement, evaluate, compare methods and/or formulations
- Problem can be inspired from models discussed in class or from your interests
- Explore whether & how theory of a method discussed in class applies/ extends to different problems

For the Ph.D.-level course cohort in ITCS 8010, possible directions could also include:

- How to use optimization to solve problem in your line of research,
- Theoretical or empirical survey on different algorithms,
- Build on recent research: Understand & expand on research paper in ML&Opt (e.g., from Mathematical Programming, SIAM Journal on Optimization, ICML/ICLR/NeurIPS)
- Design new algorithm for specific problem

For ITCS 6010: 2-3 students (default: 3), for ITCS 8010 cohort: 1-2 students per project.

Timeline.

1. Until End-September: Find group partners. Think about which task you are interested in, conduct preliminary research on potential topics. Submit 1-2 paragraphs about tentative project direction.
2. Until Mid-/End-October: Do some preliminary work, test your ideas, e.g., through preliminary experiments. Submit an $\approx 0.75 - 1$ page project proposal.
3. November 30-December 7: Final Project presentations (7-10 minutes length). Present the motivation of your project, your problem, your methodology, and the results of your project to your classmates. Attendance of *all* student presentations is mandatory.
4. By December 13: Submission of final report. Length of 3-4 pages for ITCS 6010 and 4-8 pages for ITCS 8010 projects. Incorporate feedback from final presentations.

Grading: Your final grade for the class will be given as a weighted average with the weights given as follows:

- Homework: 30%
- Midterm exam: 30%
- Final Project: 40%

Academic Integrity: You are expected to do your own work for your homework missions (and beyond). Any violation of the Code of Student Academic Integrity will be taken very seriously.
Discussion homework questions on a high level with students is allowed **if this is indicated** in the list of collaborators when submitting a homework. The write up of the homework must happen individually.

Literature: The textbook that we follow most closely the following to references:

- Stephen J. Wright and Benjamin Recht (2022), *Optimization for Data Analysis*, Cambridge University Press.
- Bernhard Gärtner and Martin Jaggi, Optimization for Machine Learning [Github], Lecture Notes, EPF Lausanne, Spring 2023.

The following book touches more on the relationship between structured data models and optimization, and its material will also be heavily used in this course.

- John Wright, Yi Ma (2022), *High-Dimensional Data Analysis with Low-Dimensional Models: Principles, Computation, and Applications*, Cambridge University Press.

Further references:

- Stephen Boyd and Lieven Vandenberghe (2009), *Convex Optimization*, Cambridge University Press.
- Marc P. Deisenroth and A. Aldo Faisal and Cheng Soon Ong (2020), *Mathematics for Machine Learning*, Cambridge University Press.
- Amir Beck (2017), *First-Order Methods in Optimization, SIAM.*
- Sébastien Bubeck (2016), *Convex Optimization: Algorithms and Complexity*, Foundations and Trends in Machine Learning.
- Elad Hazan (2019), *Lecture Notes: Optimization for Machine Learning*, Princeton University, arXiv:1909:03550.