جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Course Syllabus: Special Topics in Data Sciences - CS 390FF

Division	Computer, Electrical and Mathematical Sciences & Engineering
Course Number	CS 390FF
Course Title	Special Topics in Data Sciences
Academic Semester	Fall
Academic Year	2018/2019
Semester Start Date	08/26/2018
Semester End Date	12/11/2018
Class Schedule (Days & Time)	09:00 AM - 10:30 AM | Sun Tue

Instructor(s)

Name	Email	Phone	Office Location	Office Hours
Peter Richtarik	PETER.RICHTARIK@KAUST.EDU.SA			I will be ready to answer any questions after each lecture (10:30-11:00) in the lecture room. If you want to chat with me during some other time, please send an email to me to ask for an appointment.

Teaching Assistant(s)

Name	Email
Samuel Horvath	samuel.horvath@kaust.edu.sa

Course Information

Comprehensive Course Description	**Special Topics in Data Sciences: Big Data Optimization** The course is a mathematically rigorous introduction to the emerging field of big data optimization. It covers theory, algorithms and applications. Randomized/stochastic algorithms play a dominant role. The course is based on a novel and unified approach to recent developments in the field developed by the lecturer. Big data optimization is the study of optimization problems described by big quantities of data, where "big" is loosely defined as large enough for traditional approaches to suffer or not be applicable at all. As we live in a digital age where it is increasingly easier to collect and store data in digital form (e.g., transaction records, YouTube clicks, internet activity, Wikipedia, twitter, customer behaviour databases, government records, image collections), big data problems are becoming ubiquitous. New methods and tools are needed to analyze such vast datasets, and optimization algorithms are at the heart of such efforts, underpinning much of data science, including machine learning, operations research and statistical analysis. Alongside computer science and statistics, optimization is one of the pillars of big data analysis.
Course Description from Program Guide	

Goals and Objectives	-Detailed understanding of the role of randomization as a decomposition tool for solving optimization problems of big sizes. -Understanding the underlying mathematical theory. -Ability to apply the methodologies to selected applications in machine learning and data science. -Preparation for original theoretical and applied research in the field.
Required Knowledge	-Strong experience with at least one high level computing language (e.g.: MATLAB, Python, Julia, C, ...) -Mathematical maturity. -Ability to comprehend and generate proofs. -Linear algebra (abstract vector spaces, linear independence, basis, linear operators, quadratic forms, Euclidean spaces, inner product, norm, …) -Matrix theory (matrices, determinants, singular values, eigenvalues, matrix decompositions, …) -Multivariate calculus (gradient, Hessian, Taylor approximation, chain rule, …) -Probability theory (probability spaces, expectation, law of large numbers, tower property, …)
Reference Texts	-Detailed slides (these will be handed out before each lecture) -Relevant papers (optional reading)
Method of evaluation	**30.00**% - Final exam **20.00**% - Midterm exam **50.00**% - Homework /Assignments
Nature of the assignments	2 assignments: computational (2 x 25%) 1 midterm exam: theory (20%) 1 final exam: theory (30%)
Course Policies	-No late submissions will be accepted. Submit all you have done (even if you did not fully complete the assignment) by the deadline. -A no-show at midterm / exam results in a 0 mark. In very exceptional circumstances, **agreed with the lecturer well in advance** (e.g., serious illnes with a document from the doctor; attendance at a conference abroad, if approved by the lecturer in advance), a make-up midterm/exam might be arranged. However, this should be avoided at all costs.

Additional Information	The course is divided into 4 parts. There is a MIDTERM after the first two parts, and a FINAL EXAM at the end (covering the last two parts only). The "Tentative Course Schedule" on next page refers to the following 4 parts and topics: Part 1.Stochastic Methods for Linear Systems and Quadratic Optimization - Stochastic reformulations of linear systems - Basic method: stochastic gradient descent, stochastic Newton descent, stochastic proximal point method, stochastic projection method, stochastic fixed point method - Equivalence and exactness - Convergence analysis of the basic method - Parallel and accelerated methods - Duality and stochastic dual subspace ascent - Applications Part 2. Stochastic Methods for Convex Feasibility Problems - Stochastic projection methods - Convergence theory - Applications Part 3. Stochastic Methods for Convex Optimization - Stochastic subspace ascent - Stepsizes - Importance sampling - Parallelism - Convergence theory - Applications Part 4. Stochastic Methods for Machine Learning - Empirical risk minimization - Dual methods - Primal methods - Variance reduction - Variance reduced primal methods - Distributed methods - Nonconvex problems - Applications

Tentative Course Schedule *(Time, topic/emphasis & resources)*		
Week	**Lectures**	**Topic**
1	Sun 08/26/2018 Tue 08/28/2018	Introduction 1 Introduction 2
2	Sun 09/02/2018 Tue 09/04/2018	Introduction 3 Introduction 4
3	Sun 09/09/2018 Tue 09/11/2018	1.1 1.2
4	Sun 09/16/2018 Tue 09/18/2018	1.3 1.4
5	Sun 09/23/2018 Tue 09/25/2018	No Class (Saudi National Day) 1.5 - No Class (will be rescheduled to a different week)
6	Sun 09/30/2018 Tue 10/02/2018	1.6 1.7 - Assignment 1 released
7	Sun 10/07/2018 Tue 10/09/2018	No Class No Class
8	Sun 10/14/2018 Tue 10/16/2018	2.1 2.2 - Assignment 1 deadline
9	Sun 10/21/2018 Tue 10/23/2018	2.3 MIDTERM EXAM
10	Sun 10/28/2018 Tue 10/30/2018	3.1 - No Class (will be rescheduled to a different week) 3.2 - No Class (will be rescheduled to a different week)
11	Sun 11/04/2018 Tue 11/06/2018	3.3 3.4
12	Sun 11/11/2018 Tue 11/13/2018	3.5 3.6
13	Sun 11/18/2018 Tue 11/20/2018	4.1 4.2 - Assignment 2 released
14	Sun 11/25/2018 Tue 11/27/2018	4.3 4.4
15	Sun 12/02/2018 Tue 12/04/2018	4.5 4.6 - Assignment 2 deadline
16	Sun 12/09/2018 Tue 12/11/2018	FINAL EXAM No Class
17		
18		

Note

The instructor reserves the right to make changes to this syllabus as necessary.