CS425 – Algorithms for Web-Scale Data
Spring 2019

Instructor: M. Mustafa Ozdal (EA-420)
TA: Selcuk Gulcan (selcuk.gulcan __at__ bilkent.)

Course Hours: Tue 13:40-15:30, Thu 15:40-17:30
Classroom: EE-214

Office Hour: Tue: 15:40 - 16:30 (EA-420)

Course webpage: www.cs.bilkent.edu.tr/~mustafa.ozdal/cs425

Textbook: A. Rajaraman and J. D. Ullman, Mining of Massive Datasets, Cambridge University Press, 2011. Online free version available at: http://www.mmds.org
Credits: 3, **ECTS Credits**: 6
Prerequisites: CS202

Grading:
- Midterm: 35%
- Final: 35%
- Project: 25%
- Attendance: 5%

FZ Policy:
- Students should get a grade of 25 out of 65 to qualify for the final exam.

Fail Policy:
- Students with midterm and final average less than 20 out of 100 will automatically fail.

Course Contents

Week 1-2: PageRank	Flow & Matrix Models Google's Formulation
Week 2-3: PageRank	Efficient Algorithms for Web Graphs Topic Sensitive PageRank Web Spam
Week 4: Similarity Modeling	Document Shingling Minhashing
Week 5: Similarity Modeling	Locality Sensitive Hashing Fingerprint Matching Application
Week 6-7: Map-Reduce	Distributed Computation Model Modeling Applications with Map-Reduce
Week 7-8: Map-Reduce	Computation and Communication Complexity Matrix Multiplication using Map-Reduce
Week 9: Recommender Systems	Content-Based Recommendation Collaborative Filtering
Week 10: Recommender Systems	Latent-Factor Recommender Systems The Netflix Challenge
Week 11: Web Advertising	Online vs. Offline Algorithms Online Bipartite Matching Adwords Problem
Week 12: Project Presentations	
Week 13: Project Presentations	
Week 14: Project Presentations	