

CS425 – Algorithms for Web-Scale Data

Fall 2015

Instructor: M. Mustafa Ozdal (EA-420)

TA: Huseyin Gokhan Akcay (EA-427)

Tutor: Orcun Gumus

Course Hours: Tuesday 15:40-16:30

Friday 13:40-15:30

Spare Hour: Tuesday 16:40-17:30

Classroom: EA-Z03

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

Students are responsible for checking the webpage regularly for important announcements

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: 30%
- Final: 30%
- Project: 30%
- Attendance: 10%

Course Contents

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