BLM5121 Web Mining

Instructor: Dr. Mehmet Aktaş
Assistant Professor, CE Department, YTU
e-mail: aktas@yildiz.edu.tr

Office Hours: By appointment (please e-mail me for times)

Text Book: Web Data Mining, Exploring Hyperlinks, Contents, and Usage Data, Bing Liu ISBN-10: 3-540-37881-2, Springer, 2007.

Course Description: This is an introductory course on Web mining. Machine learning techniques to mine the Web and other unstructured/semistructured, hypertextual, distributed information repositories. Crawling, indexing, ranking and filtering algorithms using text and link analysis. Applications to search, classification, tracking, monitoring, and Web intelligence. Group project on one of the topics covered in class.
Lecture hours: Tuesdays 13:00 – 15:30
Course Outline (tentative):

1. Course Introduction
2. Chapter 1. Introduction to Web Mining
3. Chapter 8. Web Crawling
4. Chapter 3. Supervised Learning
5. Chapter 4. Unsupervised Learning
6. Chapter 6. Information Retrieval and Web Search
7. Chapter 7. Link Analysis
8. Social Networks
9. Information Diffusion

Grading (tentative):

- Attendance: 5%
- Reading Assignments: 5%
- Homeworks: 10% (Programming Assignments)
- Midterm Exam: 20%
- Term Project: 20%
- Final : 40%

Each student will be responsible to present some papers and lead class discussion on those papers. Grading is based on student's capability to synthesize critical contributions and foster a lively discussion capable of bringing key insights and connections into light. Each student is expected to read all papers ahead of class discussion and participate actively in the discussion. Participation grade is based on demonstrated familiarity with assigned readings and capability to make critical observations and to contribute constructively to the discussion.

Academic Integrity: Cheating and disruptive talking in class will not be tolerated. Any kind of cheating and plagiarism will be severely penalized. On the other hand, asking and responding to questions in class is strongly encouraged.