DA 514 – Machine Learning
Course Syllabus

Course Instructor: Kemal Kılıç
Office: FENS 1032
Tel: 216- 483 9596
E-mail: kkilic@sabanciuniv.edu

Course Description:
In this course, we will cover fundamental aspects of Machine Learning. We will start with fundamentals of machine learning, including different learning paradigms, regression and classification problems, evaluation methods, generalization and overfitting. We will then cover some of the fundamental machine learning techniques such as decision trees, Bayesian approaches, Naive Bayes classifier, and logistic regression, k-Nearest neighbor, and online learning algorithms. Besides understanding the basic theory behind the techniques, students are expected to apply them on platforms such as Rapid Miner.

Course Web Site: *SuCourse*
Lecture materials, including the class overheads, readings, assignments etc. will be available at the course web site prior to the lectures. Students are expected to check the web site regularly in order to attain the recently posted material.

Reference Texts:
TBA. Please check the *SuCourse*.

Marking Scheme

Assignments	30%
Project	30%
Exams	40%

Assignments
There will be five assignments. Some of the questions will require a software (i.e., Rapid Miner).

Term Project
Students should submit a term project which will be essential part of their evaluation. Term projects will be a practical application of a data mining software to a particular problem (which includes some literature review as well as discussion on the result of the analysis). Students should submit a proposal (*presentation*) and a final report and deliver a 20 minutes presentation at the end of the semester.

Final Report is due to TBA.

Objection Policy
Concerns regarding marks will be accepted until two days before the last date of grades submission

Academic Conduct
Do not plagiarize other people's work. Students should be aware that anyone who engages in actions prohibited by the University's policy on academic honesty will be subject to disciplinary action.

Course Content

1. Introduction
2. Validation
    a. Metrics for classifier evaluation
    b. Methods for classifier evaluation
    c. Costs in data mining
3. Classification
    a. 1 –R
    b. Naive Bayes
    c. Decision Trees
    d. Instance Based Learning
    e. Logistic Regression
    f. Support Vector Machines
4. Regression
    a. Simple Linear Regression
    b. Multiple Linear Regression
5. Meta Heuristics
    a. Local Search
    b. Simulated Annealing
    c. Genetic Algorithms
    d. Tabu Search
    e. Beam Search
6. Feature Subset Selection
    a. Filtering Methods
    b. Wrappers
    c. Embedded Techniques
    d. Feature Reduction, e.g. Principal Component Analysis
7. Interactive Machine Learning

Course Schedule and Disclaimer

The course schedule can be found in Spring 2017 DA 514 Schedule.xls which will be posted to the SuCourse. However, the instructor reserves the right, when necessary, to change examination dates, and modify the syllabus and course content. Modifications will be announced in class. Students are responsible for announced changes.