



Syllabus

CSCI 344

Data Mining and Decision Support

Spring Term 2016

Prof. Mark Sterling mark.sterling@nu.edu.kz Office 7-248	Ainur Rysbekova ainur.rysbekova@nu.edu.kz Office 7-510	
---	---	--

Spring Term Schedule

CSCI 344 MWF 13:00 – 13:50, 6.507, Sterling

Course Description

Introduction to data mining techniques, including data preprocessing, data mining primitives, association rules, decision trees, cluster analysis, classification and machine learning, data visualization, and data warehousing. Applications from a wide variety of domains will be analyzed.

Course Outcomes

After taking and successfully passing this course, students will:

- Establish a foundational understanding of basic data mining concepts and techniques
- Understand a variety of supervised and unsupervised learning algorithms and what types of problems each is appropriate for
- Be able to assess the efficacy of data mining techniques with quantitative metrics
- Gain practical experience with small programming projects

Class Structure

The model of instruction will emphasize lectures but we will also have presentations, demonstrations, and hands-on exercises. The class meetings will be conducted in 6.507.

Students are expected to attend class. Credit will not be given for homework assignments handed in late except under extenuating circumstances. In addition to homework, in class lab sessions will be conducted regularly and will be relevant to the week's topic.

In-class exercises are designed to help students practice the current material and may be done and submitted in pairs. Discussion with classmates, the instructor and the TAs is encouraged during these exercises. Homework assignments are intended to be individual work, unless instructed otherwise. Tests and exams are always individual work.

Course Assessment

The final grade will be calculated as follows:

- Mid-Term Exam 15%
- Final Exam 25%
- Homework 20%
- Course Project 15%
- Quizzes 15%
- Class Participation 10%

Grading Scale

A	95 and above
A-	90 - 94
B+	85 - 89
B	80 - 84
B-	75 - 79
C+	70 - 74
C	65 - 69
C-	60 - 64
D+	55 - 59
D	50 - 54
F	49 or below

Tentative Course Outline

Week	Dates	Topics	Homework
1	Jan 11,13,15	Introduction, Probability Review	
2	Jan 18,20,22	Linear Algebra Review, Regression	
3	Jan 25,27,29	Classification, Optimization	HW #1
4	Feb 1,3,5	Learning: LDA, k-Means, nearest-neighbor	
5	Feb 8,10,12	SVM, Hierarchical Clustering	Quiz #1
6	Feb 15,17,19	Features and Feature Selection	
7	Feb 22,24,26	Data Cleaning and Collection	HW #2
8	Feb 29, Mar 2,4	Evaluation, Cross Validation	
9	Mar 7,9,11	Introduction to Big Data	Mid-Term Exam
10	Mar 14,16,18	Data Visualization	
	Mar 21,23,25	Spring Break	
11	Mar 28,30 Apr 1	Cloud and Infrastructure	HW #3, Quiz #2
12	Apr 4,6,8	“ “	
13	Apr 11,13,15	Data Warehousing	HW #4
14	Apr 18,20,22	Security and Privacy	Quiz #3
15	Apr 25,27,29	Review of Materials	
Exam Period	May 2-14	Exams	Final Exam

Academic Integrity

Nazarbayev University and The School of Science and Technology have established high standards for academic integrity, using an approach in which students are trained to produce original work according to professional standards, and to properly cite and reference the work of others when it is appropriate to do so.

The specific guidelines are published in the NU Student Handbook. In particular,

- The assignments in this class are designed to introduce important concepts and techniques, and enable you to explore the material independently so as to gain insight and comprehension of the subject. Doing the work is much more important than getting the right answer.
- The course is designed such that each new week's material builds on the skills developed in the preceding week, thus, any action that interferes with this process (missing class, skipping the assignment, copying) will seriously impede your progress.
- You are welcome—and encouraged—to talk through concepts and ideas with your fellow students and to study with them, but do not give or receive direct help from your classmates on a graded assignment.
- Homework should be completed individually. If you distribute your work to others, even if you are not intending them to copy it, this is still considered academic misconduct.
- Even the appearance of cheating or inappropriate copying should be avoided.
- Students should be aware that the homework submission process incorporates an automated plagiarism detector.
- You may only get help on graded assignments from designated people—the professors or TAs for the course.

If you are struggling with an assignment, by all means, please seek help from them. In the event that academic misconduct such as plagiarism or cheating is discovered, the student will receive no credit for the work, and the event reported to the Senior Administrator for Students. Egregious cases, or a second offense, can result in failure of the course and potential suspension or expulsion from the university. When a student suspects that another student has violated the academic honesty policy, a report should be made to the appropriate faculty member.

Behavior

Students are expected to maintain respectful decorum in the classroom and laboratories, and in all interactions with fellow classmates, Teaching Assistants, Research Assistants and NU faculty and staff. Class time is short, and valuable, and thus should be used effectively; students are expected to refrain from such distractions as texting, phone calls, on-line chats, personal web browsing, the use of social networking sites, and excessive chatting or greetings during class time.

Students should come to class well-prepared, having completed the background reading and related assignments and possessing proper resources for the class meeting (books, paper, writing implements, computers, etc.), as needed.

Important Note: Please do not raise the stipend to me as an issue I should consider in changing your grade. I will consider this to be inconsistent with “respectful decorum” as stated above. Also, be aware that with excessive absences you risk impacting your final grade significantly.

