

**BLM4800- Introduction to Data Mining Syllabus**

Instructor Name: Assoc. Prof. Serkan Ayvaz
 Tel: (212) 383 5780
 Email: sayvaz@yildiz.edu.tr
 Website: <https://avesis.yildiz.edu.tr/17212>
 Office: D-220

Course Hours Fri 14:00--16:50PM
Office Hours Fri 13:00--14:00PM or by appointment

Text Books

- Data Mining , J. Han – M. Kamber, Morgan-Kaufman, Academic Press, 2001, ISBN: 1-55860-901-6
- Data Mining – Concepts, Models, Methods and Algorithms, Mehmed Kantardzic, ISBN:0-471-22852-4

Course Description

Data mining is discovery of the knowledge from the huge amount of data. The objective of data mining is to find useful patterns in bulky data and to use discovered patterns to help explain behavior or to predict future outcomes. Data Mining Concepts, Preparing the Data, Statistical Classification Method (Naïve Bayes), Clustering Methods(K-Means,Hierarchical), Decision Trees and Decision Rules, Association Rules.

Upon completion of the course, students will:

- 1- have the basic knowledge and application ability about the data mining methods.
- 2- have the ability of exploring the knowledge from bulky data.
- 3- have the ability to analyze, clear and find the outliers from bulky data.
- 4- have knowledge about supervised classification methods.
- 5- have knowledge about unsupervised clustering methods.

Course Format

The lectures will include readings tutorials. The PowerPoint presentations/class notes will also be available on the website following each class. During class meetings, there will be some lecturing. You will tackle different data mining problems individually.

Participation: Participation will be a part of your course grade. You will be responsible for recovering any information you have missed if you do not attend a particular day's lecture.

Exams: The exams will be 90-120 minutes, closed book, closed notes exam. The use of any reference material is strictly forbidden.

Project: There will be one project for the course. It can be a group project or an individual project. For group project, the students are expected to provide more comprehensive work. For group project, all study group members will receive the same score. In general, students may freely communicate within their group, but you may not discuss group work with members of other groups. Details of this assignment will be discussed later in the semester.

Students are responsible for forming and managing their study groups. We expect that students will manage their study groups so that everyone performs a fair share of the work, and that all perspectives are heard and considered. The assignments should be turned in electronically by end of the due date.

Correspondence: If you want to get response to your e-mails, always include your name, your course name. Observe grammatical rules while composing your e-mails.

Grading

Final course grades will be based on:

| | |
|---------------------|-----|
| Project | 30% |
| Midterm Examination | 30% |
| Final Examination | 40% |

No late assignments will be accepted. No make-up will be administered for the midterm.

TENTATIVE SCHEDULE

| | |
|------------------------------|---|
| WEEK 1 | March /3 |
| Topics to be Covered: | Introduction to Data Mining |
| WEEK 2 | March /10 |
| Topics to be Covered: | Data Mining Concepts |
| WEEK 3 | March /17 |
| Topics to be Covered: | Preparing the Data |
| WEEK 4 | March /24 |
| Topics to be Covered: | Data Reduction |
| WEEK 5 | March /31 |
| Topics to be Covered: | Statistical Classification Method (Naïve Bayes) |
| WEEK 6 | April /7 |
| Topics to be Covered: | Decision Trees and Decision Rules |
| WEEK 7 | April /14 |
| Topics to be Covered: | Clustering and Similarity Measures |
| WEEK 8 | April /28 |
| Topics to be Covered: | MIDTERM EXAM |
| WEEK 9 | May /5 |
| Topics to be Covered: | Clustering Methods(Hierarchical Clustering) |
| WEEK 10 | May /12 |
| Topics to be Covered: | Evaluation of Classification Methods |
| WEEK 11 | May /26 |
| Topics to be Covered: | Association Rule Mining |

| | |
|------------------------------|--|
| Covered: | |
| WEEK 12 | June /2 |
| Topics to be Covered: | Classification with Artificial Neural Networks |
| WEEK 13 | June /3 |
| Topics to be Covered: | PRESENTATIONS |
| WEEK 14 | June /3 |
| Topics to be Covered: | PRESENTATIONS |
| WEEK 15 | |
| Topics to be Covered: | FINAL |