AGU Graduate School of Engineering and Science Bioengineering Program

COURSE RECORD

Code	**BENG546**
Name	**Data Mining**
Hour per week	3 (3 + 0)
Credit	3
ECTS	7,5
Level/Year	Graduate
Semester	Fall/Spring
Type	Elective
Location	
Prerequisites	
Special Conditions	
Coordinator(s)	Assist. Prof. Dr. Müşerref Duygu Saçar Demirci
Webpage	
Content	The course presents an introduction to popular data mining approaches. The key processes in data mining will be covered: types of attributes, common data set structures, data preprocessing, feature selection, sampling, using different statistical and machine learning techniques and visualization. Through a course project, the students will apply a data mining software on a real problem.
Objectives	- Explaining the basic concepts of Data Mining. - Using data mining software for solving practical problems. - To gain experience of analyzing real biological data. - Improving skills in independent study and research.
Learning Outcomes	Students will be, LO1 Able to describe the types, quality and influence of data. LO2 Able to describe preprocessing and feature selection methods. LO3 Able to describe classification and clustering methods and performance evaluation. LO4 Able to explain visualization techniques and anomaly detection. LO5 Able to design a data mining workflow to solve a real problem.
Requirements	
Reading List	Introduction to Data Mining: Pearson New International Edition, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson, 2014.
Ethical Rules and Course Policy	

LEARNING ACTIVITIES

Activities	**Number**	**Weight (%)**
Lecture	12	40%
Group Works	2	30%
Presentations	2	25%
Site Visits	1	5%
	Total	100

ASSESSMENT

Evaluation Criteria	**Weight** (%)
Group Project Assignments & Presentations	90%
Attendance/Participation	10%
Total	100%

For a detailed description of grading policy and scale, please refer to the website https://goo.gl/HbPM2y section 28.

COURSE LOAD

Activity	**Duration** (hour)	**Quantity**	**Work Load** (hour)
In class activities	3	14	42
Group work	8	14	112
Research (web, library)	3	14	42
Required Readings	4	14	56
Pre-work for Presentation	25	2	50
		General Sum	**302**

ECTS: 7,5(Work Load/25-30)

CONTRIBUTION TO PROGRAMME OUTCOMES*

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PO13	PO14
LO1	5	5	4	5	4	4	3	3						
LO2	5	4	4	5	4	4	3	3						
LO3	5	5	5	5	4	4	3	3						
LO4	5	5	5	5	4	4	3	3						
LO5	5	5	5	5	5	5	3	3						

* Contribution Level: 0: None, 1: Very Low, 2: Low, 3: Medium, 4: High, 5: Very High

WEEKLY SCHEDULE

W	**Topic**	**Outcomes**
1	Introduction to Data Mining Lab/Activity: data mining definition, goals, concepts	LO1
2	Data Lab/Activity: type of data, quality of data, data sources, exploring data	LO1
3	Preprocessing Activity: aggregation, sampling, dimensionality reduction, transformation	LO2
4	Feature Selection Activity: embedded, wrapper, filter approaches	LO2
5	Classification I Activity: basic classification concepts, Decision Tree	LO3
6	Classification II Activity: alternative approaches, SVM, Naïve Bayes	LO3
7	Clustering I Activity: basic issues in clustering, partitioning methods: k-means, expectation maximization (EM)	LO3
8	Clustering II Activity: hierarchical methods	LO3
9	Student Presentations Activity: students will present a research article	LO5
10	Performance Evaluation Activity: training, testing, performance evaluation, cross-validation	LO3
11	Visualization Activity: histograms, scatter plots, ROC curves	LO4
12	Anomaly Detection Activity: causes of anomalies, approaches to anomaly detection	LO4
13	Mining Real Data Activity: obtaining real data and demonstration of analysis using a software	LO5
14	Project Presentations Activity: students will present their term projects	LO5

Prepared by
Date