ABDULLAH GÜL UNIVERSITY
INSTITUTE OF SCIENCE
ELECTRIC and COMPUTER ENGİNEERING ANABİLİM DALI
INDIVIDUAL COURSE DESCRIPTION

Course Ttitle	Code	Semester	T+U Hours	Credit	ECTS
DATA MINING	ECE-565	FALL	3 + 0	3	10

Prerequisities and co-re-requisities	Introduction to Computer Programming, Probability and Statistics, Linear Algebra

Type	Elective
Language	English
Coordinator	Assist. Prof. Dr. Zafer Aydin Assist. Prof. Dr. Bekir Hakan Aksebzeci
Instructor	Assist. Prof. Dr. Zafer Aydin Assist. Prof. Dr. Bekir Hakan Aksebzeci
Adjunt	None
Aim	This course provides an introduction to data mining. It covers fundamental pattern discovery techniques such as frequent itemset and association rule detection, efficient data structures and algorithms that have good scaling properties and fundamental classification and clustering algorithms. Through a course project, the students will program a data mining software and apply the concepts to a real problem.
Learning Outcomes	1. Describe the techniques that are used to represent data and knowledge 2. Describe the data transformation techniques 3. Explain frequent itemset detection and association rule mining methods 4. Explain the fundamental classification and clustering methods 5. Outline the approaches used for model training, testing, comparison and performance evaluation 6. Operate a data mining software (e.g. WEKA) 7. Apply a data mining method to a real problem
Course Content	• Introduction to data mining • Data representation techniques • Knowledge representation techniques • Data preprocessing techniques • Frequent itemset and association rule mining methods • Basic classification methods • Basic clustering methods Performance evaluation methods

WEEKLY TOPICS AND PRELIMINARY STUDY

Week	Topics	Prelimanary Study
1	**Introduction to Data Mining:** Knowledge discovery, steps in a data mining system, data and pattern types, learning algorithms, data mining methods, databases and data warehouses	
2	**Data representation:** Concepts, instances, attributes, statistical descriptions, graphical descriprions, data similarity and distance measures	
3	**Knowledge representation:** linear models, trees, rules, clusters, instance based representation	
4	**Data Preprocessing:** Missing values, data cleansing, noise reduction, data integration, aggregation	
5	**Data Preprocessing:** Data normalization, transformation, dimension reduction, attribute selection	

6	**Pattern Discovery:** Market basket analysis, frequent itemsets, closed itemsets, association rules, apriori algorithm	
7	**Pattern Discovery:** FP growth algorithm, discovering closed and maximum patterns	
8	**Pattern Discovery:** Deriving association rules from frequent itemsets, pattern evaluation measures	
9	**Midterm Exam**	
10	**Basic Concepts and Methods of Classification:** Decision tree induction, Bayes classification methods	
11	**Basic Concepts and Methods of Classification:** Rule-based classification, Model evaluation and selection	
12	**Basic Concepts and Methods of Clustering:** Definition of clustering, Comparing clustering methods, Requirements for clustering, Partitioning methods	
13	**Basic Concepts and Methods of Clustering:** Hierarchical methods, Probabilistic hierarchical clustering	
14	**Performance Evaluation and Testing Methods:** Model training, testing, performance evaluation, cross-validation, bootstrap method	
15	**Performance Evaluation and Testing Methods:** Comparing data mining schemes, ROC curves, recall-precision curves, minimum description length principle	
16	**Final Exam**	

SOURCES	
Lecture Notes	Lecture slides
Other Sources	**Course Textbook:** 1. Data Mining: Concepts and Techniques", 3rd edition, Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann Publishers, 2011. **Additional Materials**: **1.** "Data Mining: Practical Machine Learning Tools and Techniques", Ian H. Witten, Eibe Frank, Mark A. Hall, Morgan Kaufmann Publishers, 2011. **2.** "Introduction to Data Mining", Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Addison-Wesley, 2005.

COURSE MATERIALS SHARING	
Documents	
Homeworks	10
Exams	2 Midterm and 1 Final Exam

EVALUATION SYSTEM		
SEMESTER STUDY	**NUMBER**	**CONTRIBUTION**
Midterm	2	30
Project Homework	10	10
Semester Project	2	40
Final Exam	1	20
Contribution of Semester Study		80
Contribution of Final Exam		20
TOTAL		100

Course Category	
Sciences and Mathemetics	%50
Engineering	%50
Social Sciences	%0

RELATIONSHIPS BETWEEN LEARNING OUTCOMES AND PROGRAM QUALIFICATIONS		
No	Program Qualitications	Contribution Level

		1	2	3	4	5
1	The skills of using mathematics, science and engineering information in advanced research				X	
2	The skills of analyzing, designing and/or implementing an original system that will be able to solve an engineering problem					X
3	The skills of using the required software, hardware and modern measurement equipments in their field of research					X
4	The skills of planning independent research and implementing in detail			X		
5	The skills of following literature, listening to and making technical presentation, writing a paper in academic level			X		
6	The skills of innovative and interrogative thinking and finding original solutions		X			

Increasing from 1 to 5

ECTS/ WORK LOAD TABLE			
Activities	Number	Duration (Hours)	Total Work Load
Course Length (includes exam weeks: 16x total course hours)	16	3	48
Out of class study time (pre-study, practice)	16	8	128
İnternet search, library work, literature search	1	5	5
Presantation			
Homework	10	5	50
Midterm Exam	2	20	40
Final Exam	1	30	30
Total Work Load			291
Total Work Load/ 30			291/30
Course ECTS Credit			10