School of Business and Management Sciences / Management Information Systems (English)

2023 - 2024 Academic Year

DATA MINING and BUSINESS INTELLIGENCE

Syllabus

Course Description					
Name	**Code**	**Semester**	**T+A Hour**	**Credit**	**ECTS**
DATA MINING and BUSINESS INTELLIGENCE	MIS4112149	Fall Semester	3+0	3	5

Prerequisites Courses	
Recommended Elective Courses	
Language of Instruction	English
Course Level	First Cycle (Bachelor's Degree)
Course Type	Required
Course Coordinator	Prof.Dr. Gökhan SİLAHTAROĞLU
Name of Lecturer(s)	Lect. Nada A. M. MISK, Prof.Dr. Gökhan SİLAHTAROĞLU
Assistant(s)	
Aim	To provide students with the ability to create a data warehouse from databases.To gain research skills on these data warehouses by using OLAP and data mining models. To get the level of knowledge to apply data mining algorithms.
Course Content	This course contains; Introduction to Data Mining ,Data warehouse –OLAP – ETL process – Database vs. Datawarehouse,Tools Used in Data Mining: Python, R, KNIME, pivoting.(KNIME Installation, Interface, Reading and viewing data),Data Preprocessing (Outliers, Missing data, Normalization, Data Transformation, Binning, Histogram),Association Rules (Shopping Cart Analysis),Supervised Learning, Classification And Decision Trees, Gini / Entropy Brief introduction. (Application: BEARS),Decision Trees (Application: MEDICINE and Wine / KNIME: filtering, train-test, validation, accuracy, Color Manager),Random Forest, Booststrapping, Ensemble Learning, Perfume Application + Telco Application,Regression and Logistic regression,Neural Network Model (principle, Hyperparameters).,Unsupervised Learning / Clustering (Application: Wholesale Customer Data, Fuzzy C-means clustering, Quality measures),Data Reduction, Synthetic Data generation, PCA, clustering with Scatter Plot (DBSCAN),Big Data concepts – HADOOP, spark, MongoDB, NOSQL,Project Discussion & Evaluation.

Course Learning Outcomes	**Teaching Methods**	**Assessment Methods**
1. Will be able to produce data warehouse from database.	10, 16, 9	A
1.1. Explains datamining.	16, 9	
1.2. Defines Data Warehouse.	16, 9	
1.3. Designs the Star Data Warehouse Model.	10, 9	
1.4. Main tables association designs the data warehouse model.	16, 9	
2. Will be able to relate Data Mining Models to each other.	16, 9	A, F
2.1. Explains data mining models.	16, 9	
2.2. Defines the concept of classification.	16, 9	
2.3. Defines the concept of clustering.	16, 9	
2.4. Defines the concept of connection analysis.	16, 9	
3. will be able to apply the classification model.	16, 9	A, F
3.1. Defines supervised learning.	16, 9	
3.2. Defines the concept of class.	9	A
3.3. sorts statistical algorithms.	16, 9	
3.4. Applies decision trees.	16, 9	F
3.5. Defines decision tree algorithms.	16, 6, 9	
3.6. Defines pruning and purity values.	9	
4. Will be able to employ the clustering model.	14, 6, 9	F
4.1. Defines unsupervised learning.	6, 9	
4.2. Explains the concept of clustering.	16, 9	
4.3. Sorts clustering algorithms.	6, 9	F
4.4. Applies K-means algorithm.	14, 16, 9	
4.5. Defines genetic algorithms.	10, 8, 9	F
5. Will be able employ the connection analysis model.	14, 16, 6, 9	A
5.1. Interprets connection analysis rules.	14, 6, 9	
5.2. Explains the concept of leverage.	16, 9	A
5.3. Applies the relationship analysis method.	14, 9	
5.4. Combines clustering with link analysis.	2	F
6. Will be able to employ Data Mining Algorithms.	2	F
6.1. Employs classification algorithms on data.	2	F
6.2. Employs clustering algorithms on data.	2	F
6.3. Employs link analysis algorithms on data.	2	F
6.4. Interprets data mining application outputs.	2	F

Teaching Methods	10: Discussion Method, 14: Self Study Method, 16: Question - Answer Technique, 2: Project Based Learning Model, 6: Experiential Learning, 8: Flipped Classroom Learning, 9: Lecture Method
Assessment Methods	A: Traditional Written Exam, F: Project Task

Lecture Schedule

Sequence	**Topics**	**Preliminary Preparation**
1	Introduction to Data Mining	
2	Data warehouse –OLAP – ETL process – Database vs. Datawarehouse	
3	Tools Used in Data Mining: Python, R, KNIME, pivoting.(KNIME Installation, Interface, Reading and viewing data)	
4	Data Preprocessing (Outliers, Missing data, Normalization, Data Transformation, Binning, Histogram)	

School of Business and Management Sciences / Management Information Systems (English)

2023 - 2024 Academic Year

DATA MINING and BUSINESS INTELLIGENCE

Syllabus

Lecture Schedule		
Sequence	**Topics**	**Preliminary Preparation**
5	Association Rules (Shopping Cart Analysis)	
6	Supervised Learning, Classification And Decision Trees, Gini / Entropy Brief introduction. (Application: BEARS)	
7	Decision Trees (Application: MEDICINE and Wine / KNIME: filtering, train-test, validation, accuracy, Color Manager)	
8	Random Forest, Booststrapping, Ensemble Learning, Perfume Application + Telco Application	
9	Regression and Logistic regression	
10	Neural Network Model (principle, Hyperparameters).	
11	Unsupervised Learning / Clustering (Application: Wholesale Customer Data, Fuzzy C-means clustering, Quality measures)	
12	Data Reduction, Synthetic Data generation, PCA, clustering with Scatter Plot (DBSCAN)	
13	Big Data concepts – HADOOP, spark, MongoDB, NOSQL	
14	Project Discussion & Evaluation	

Evaluation Methods	**Weight(%)**
Midterm Exam	40
(General Exam) Project as the General Exam	60
General Exam	60

Resources

1. Data Mining Introductory and Advanced Topics, Margaret H. Dunham, Prentice Hall.

Knime Application: https://docs.knime.com/ 1. Data Mining Concepts and Techniques , J. Han & M. Kamber, Morgan Kaufman.
2. Big Data in Practice: How 45 Successful Companies Used Big Data Analytics to Deliver Extraordinary Results , Bernard Marr, Wiley, 2016
3. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy, Cathy O'Neil ,2017
4. Naked Statistics: Stripping the Dread from the Data, Charles Wheelan, 2013
5. Kavram ve Algoritmalarıyla Temel Veri Madenciliği, Gökhan Silahtaroğlu, Papatya Yayıncılık.