

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department
----------------	--

Study plan No.	2020/2021	University Specialization	Artificial Intelligence
Course No.	0142334	Course name	Data Mining
Credit Hours	3	Prerequisite Co-requisite	Introduction into Data Science
Course type	<input type="checkbox"/> MANDATORY UNIVERSITY REQUIREMENT <input type="checkbox"/> UNIVERSITY ELECTIVE REQUIREMENTS	<input type="checkbox"/> FACULTY MANDATORY REQUIREMENT <input type="checkbox"/> Support course family requirements	<input type="checkbox"/> Mandatory requirements <input checked="" type="checkbox"/> Elective requirements
Teaching style	<input type="checkbox"/> Full online learning	<input type="checkbox"/> Blended learning	Traditional learning
Teaching model	<input type="checkbox"/> 2Synchronous: 1asynchronous	<input type="checkbox"/> 2 face to face : 1synchronous	3 Traditional

### Faculty member and study divisions information (to be filled in each semester by the subject instructor)

Name	Academic rank	Office No.	Phone No.	E-mail	
Bilal Hawashin	Associate professor			b.hawashin@zuj.edu.jo	
Division number	Time	Place	Number of students	Teaching style	Approved model
1					

### Brief description

This course provides the following topics:  
Introduction to Data Mining, Classification, Clustering, Association Rule Discovery, Anomaly Detection, Web Mining, Collaborative Filtering, and various data mining topics.

### Learning resources

Course book information (Title, author, date of issue, publisher ... etc)	Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, 2016.			
Supportive learning resources (Books, databases, periodicals, software, applications, others)	1- Charu Aggarwal. Data Mining, the text book. Springer. 2015 2- Ian Witten, Eibe Frank, Mark Hall, and Christopher Pal. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann. 2016.			
Supporting websites				
The physical environment for teaching	Class room	<input type="checkbox"/> labs	<input type="checkbox"/> Virtual educational platform	<input type="checkbox"/> Others

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department
----------------	--

Necessary equipment and software	
Supporting people with special needs	
For technical support	

### Course learning outcomes (S= Skills, C= Competences K= Knowledge,)

No.	Course learning outcomes	The associated program learning output code
<b>Knowledge</b>		
<b>K1</b>	To have a good comprehension of the Classification, Clustering, and association rule discovery.	<b>MK3</b>
<b>K2</b>	To have a good understanding of various advanced data mining topics.	<b>MK3</b>
<b>Skills</b>		
<b>S1</b>	To be able to categorize each scenario to DM or Non DM task.	<b>MS3</b>
<b>S2</b>	To be able to use WEKA to solve data mining problems.	<b>MS3</b>
<b>S3</b>	To be able to compute cosine similarity, Euclidian distance between vectors	<b>MS3</b>
<b>S4</b>	To be able to select features based on TF.IDF method	<b>MS3</b>
<b>S5</b>	To be able to compute manually the results of certain data mining algorithms such as decision tree, ANN, KNN, K-Means, and Apriori.	<b>MS3</b>
<b>Competences</b>		
<b>C1</b>	To be able to use data mining methods to solve real-life problems.	<b>MC1</b>

### Mechanisms for direct evaluation of learning outcomes

Type of assessment / learning style	Fully electronic learning	Blended learning	Traditional Learning (Theory Learning)	Traditional Learning (Practical Learning)
First exam	0	0	%20	0
Second / midterm exam	%30	%30	%20	30%
Participation / practical applications	0	0	10	30%
Asynchronous interactive activities	%30	%30	0	0
final exam	%40	%40	%50	40%

**Note:** Asynchronous interactive activities are activities, tasks, projects, assignments, research, studies, projects, work within student groups ... etc, which the student carries out on his own, through the virtual platform without a direct encounter with the subject teacher.

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department
----------------	--

### Schedule of simultaneous / face-to-face encounters and their topics

Week	Subject	learning style*	Reference **
1	<b>Data Mining</b> Definition. Data Mining Categories.	Lectures	1-16 (Ref1)
2	<b>Data Preprocessing.</b> Types of Data. Enhancing Data Quality. Similarity Measurements.	Lectures	22-44 (Ref1)
3	<b>Sampling and Feature Selection.</b> Types of Sampling. Feature Selection Methods. Mean TF.IDF. CHI Square.	Lectures	47-55 (Ref1)
4	<b>Classification</b> Introduction. Decision Support Trees, K-Nearest Neighborhood, Support Vector Machines, Artificial Neural Networks.	Lectures	145-276 (Ref1)
5	<b>Classification</b> Rule Based Classification. Using WEKA to apply classification.	Lectures	145-276 (Ref1)
6	<b>Clustering</b> Introduction. KMeans Algorithm. Hierarchical Clustering. Using WEKA to apply clustering.	Lectures	487-526 (Ref1)
7	<b>Case Study</b>	Learning Through Problem Solving	
8	<b>Mid Exam</b>		
9	<b>Association Rules.</b> Apriori Algorithm. <b>Dimensionality Reduction Methods.</b> Introduction. Latent Semantic Indexing. Information Retrieval.	Lectures	327-353 (Ref1)
10	<b>Various Data Mining Topics.</b> Collaborative Filtering. Web Mining. Data ware Housing.	Lectures	Ref2
11	<b>Various Data Mining Topics.</b> Graph Mining. Social Network Analysis. Data Stream Mining.	Lectures	Ref2
12	<b>Introduction to Deep Learning.</b>	Lectures	12-27 (TB)
13	<b>Advanced Deep Learning and Data Mining Topics.</b>	Lectures	(TB)
14	<b>Project Presentations.</b>	Participatory Learning	
15	<b>Project Presentations.</b>	Participatory Learning	
16	<b>Final Exam</b>		

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department
----------------	--

- \* Learning styles: Lecture, flipped learning, learning through projects, learning through problem solving, participatory learning ... etc.  
\*\* Reference: Pages in a book, database, recorded lecture, content on the e-learning platform, video, website ... etc.

### Schedule of asynchronous interactive activities (in the case of e-learning and blended learning)

Week	Task / activity	Reference	Expected results
1	Data Mining Types		To be able to categorize real life scenarios according to their type (classification, clustering, anomaly detection, association rule mining, not data mining)
2	Types of Data Recognition		To be able to recognize data type in given scenarios
3	Feature Selection Using Mean TF.IDF		To be able to find Mean TF.IDF for one or more features and select among them
4	Data Preprocessing		To be able to perform some data preprocessing tasks
5	Data Scaling and Normalization		To apply data scaling and normalization on the data
6	Data Transformation		To apply certain data transformations on the data
7	Finding Similarity and Distance Between Items		To find the cosine similarity and Euclidean distance between two vectors
8	Applying KNN Classifier		To use KNN to classify testing records
9	Evaluating KNN		To compute some evaluation measurements for the classifier
10	Applying Decision Tree		To build manually a decision tree given a training data and use it in the classification
11	Applying ANN		To classify testing data

QF01/0408-4.0E	Course Plan for Bachelor program - Study Plan Development and Updating Procedures/ Artificial Intelligence Department
----------------	--

			using a given ANN classifier
12	Applying Clustering using K-Means		To do the first two iterations in K-Means clustering given a dataset
13	Applying Apriori		To apply Apriori to find frequent itemsets
14	Applying WEKA		To use weka to solve a real life problem
15	Applying Recommender Systems		To use item-item and user-user similarities to give recommendations
16			