| 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** | ☐ MANDATORY UNIVERSITY REQUIREMENT | ☐ UNIVERSITY ELECTIVE REQUIREMENTS | ☐ FACULTY MANDATORY REQUIREMENT | ☐ Support course family requirements | ☐ Mandatory requirements ☑ Elective requirements |
| **Teaching style** | ☐ Full online learning | | ☐ Blended learning | | Traditional learning |
| **Teaching model** | ☐ **2Synchronous: 1asynchronous** | | ☐ **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 | |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| **Division number** | **Time** | **Place** | **Number of students** | **Teaching style** | **Approved model** |
| | | | | | |
| | | | | | |
| | | | | | |

## 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**<br><br>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 | ☐ labs | ☐ Virtual educational platform | ☐ Others |
| 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 |
|---|---|---|
| | **Knowledge** | |
| **K1** | To have a good comprehension of the Classification, Clustering, and association rule discovery. | **MK3** |
| **K2** | To have a good understanding of various advanced data mining topics. | MK3 |
| | **Skills** | |
| **S1** | To be able to categorize each scenario to DM or Non DM task. | **MS3** |
| **S2** | To be able to use WEKA to solve data mining problems. | **MS3** |
| **S3** | To be able to compute cosine similarity, Euclidian distance between vectors | **MS3** |
| **S4** | To be able to select features based on TF.IDF method | **MS3** |
| **S5** | To be able to compute manually the results of certain data mining algorithms such as decision tree, ANN, KNN, K-Means, and Apriori. | **MS3** |
| | **Competences** | |
| **C1** | To be able to use data mining methods to solve real-life problems. | **MC1** |

**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.

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

| Week | Subject | learning style* | Reference ** |
|---|---|---|---|
| **1** | **Data Mining** Definition. Data Mining Categories. | Lectures | **1-16 (Ref1)** |

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

*** 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 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 | | | |