Course Specification
(Bachelor)

Course Title:	Data Mining
Course Code:	DSC 312
Program:	Applied Statistics & Data Management
Department:	Mathematics
College:	College of Science
Institution:	Majmaah University, Saudi Arabia
Version:	2023
Last Revision Date:	20/09/2023

Table of Contents

A. General information about the course: 3

B. Course Learning Outcomes (CLOs), Teaching Strategies and Assessment Methods 4

C. Course Content 5

D. Students Assessment Activities 7

E. Learning Resources and Facilities 7

F. Assessment of Course Quality 8

G. Specification Approval 8

A. General information about the course:

1. Course Identification

1. Credit hours: (3 (2+2))					
2. Course type					
A.	☐University	☐College	☒Department	☐Track	Others
B.	☒Required		☐Elective		
3. Level/year at which this course is offered: (7th level /4th year)					
4. Course general Description:					
This course introduces an introduction to the fundamentals of data mining principles and a basic understanding of data analysis techniques required to tackle real-world, data-rich problems in business and academia, including: Data set acquisition, cleaning, ,aggregation , dimension reduction , Feature Selection and extraction based on PCA. Information visualization methods are applied to data from many different application domains, including attributes or variables for the units of information. Data classification methods: Decision tree, Rule-based, Bayes' theorem, K-Nearest Neighbor and Support Vector Machine (AVM). Data Clustering methods: K-means.					
5. Pre-requirements for this course (if any)**:**					
Data Analysis DSC233					
6. Co-requisites for this course (if any)**:**					
7. Course Main Objective(s):					
1- **Students will develop relevant programming abilities.** **2- Students will demonstrate proficiency with statistical analysis of data.** **3- Students will demonstrate proficiency with data visualization methods.** **4- Students will develop the ability to build and assess data-based model classifications.** **5- Students will execute data analysis with professional software Python.**					

2. Teaching mode (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	30	50%
2	E-learning		
3	Hybrid • Traditional classroom • E-learning	30	50%
4	Distance learning		

3. Contact Hours (based on the academic semester)

No	Activity	Contact Hours
1.	**Lectures**	30
2.	**Laboratory/Studio**	30
3.	**Field**	
4.	**Tutorial**	
5.	**Others (specify)**	
Total		60

B. Course Learning Outcomes (CLOs), Teaching Strategies and Assessment Methods

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
1.0	**Knowledge and understanding**			
1.1	Use computer mathematical software in solving mathematical problems such as data analysis techniques, data-rich problems and Data set acquisition	K2	Direct teaching: Lectures Aimed teaching: Discovery and oral questions Indirect teaching: Cooperative Learning	. Homework · Quiz · Midterms · Final Exams
1.2			Direct teaching: Lectures	. Homework · Quiz · Midterms · Final Exams

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
			Aimed teaching: Discovery and oral questions Indirect teaching: Cooperative Learning.	
2.0	**Skills**			
2.1	Demonstrate the work independently and within a team	S2	Direct teaching: Lectures Aimed teaching: Discovery and oral questions Indirect teaching: Cooperative Learning.	. Homework · Quiz · Midterms · Final Exams
3.0	**Values, autonomy, and responsibility**			
3.1	Identify, formulate and solve mathematical problems	V2	Direct teaching: Lectures Aimed teaching: Discovery and oral questions Indirect teaching: Cooperative Learning	. Homework · Quiz · Midterms · Final Exams

C. Course Content

No	List of Topics	Contact Hours
1	**Data Mining Introduction**	12

	• What is data mining • Data mining functionalities • Data mining applications Data mining Challenges	
2	**Data Preprocessing** • Types of Data attributes • Descriptive data summarization • Data cleaning o Data integration and transformation • Data reduction: Dimension reduction & Feature extraction sing PCA • Data discretization and concept hierarchy generation Lab: Data analysis applications with Python\R\Weka	6
3.	**Data Visualization and Exploration** • Introduction to Concepts of data visualization and data exploration in science and their techniques. • Visualize spatial patterns, trends, outliers, and anomalies • Scatter graphs, scatter diagrams and Line plot • Histograms, Line histograms, and bar charts. • Heatmaps, Word Cloud and EndNote • Summary statistical exploration • Multidimensional data model Data cube and OLAP • Multidimensional Data Analysis in Cube Space Lab: Data Visualization applications with Python\R\Weka	18
4.	**Frequent Pattern Mining o Basic concepts** o Efficient and scalable frequent item set mining methods o Correlation analysis o Sequential pattern mining o Graph and tree mining Lab: Data exploration applications with Weka software or Python or R	6
5.	**Classification and Prediction**	6

	o Decision tree induction o Bayesian classification o Support vector machines o K-Nearest neighbor methods o Performance evaluation Lab: Data classification applications with Python\R\Weka	
6.	**Clustering Analysis** o Partition methods o Hierarchical methods o Density-based methods o Outlier analysis Lab: Data clustering applications with Python\R\Weka	12
7.		
8.		
9.		
Total		**60**

D. Students Assessment Activities

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
1.	Quizzes	Week 4 and Week 8	10%
2.	Assignments/Exercises	Every Week	10%
3.	Mid Term Exam 1	Week 6	15%
4.	Mid Term Exam 2	Week 12	15%
5.	Final Exam	After Week 15	40%
6.	E-exam-Lab	Week 13	10%

*Assessment Activities (i.e., Written test, oral test, oral presentation, group project, essay, etc.).

E. Learning Resources and Facilities

1. References and Learning Resources

Essential References	Michael Steinbach (Author), Vipin Kumar, Introduction to Data Mining 2nd, Amazon Warehouse, 2018

Supportive References	Jiawei Han, Jian Pei, Hanghang Tong, Data Mining Concepts and Techniques, 4th Edition, MK Elsevier Wordmark, 2022
Electronic Materials	MIT Open Courses Khanacadmy
Other Learning Materials	

2. Required Facilities and equipment

Items	Resources
facilities (Classrooms, laboratories, exhibition rooms, simulation rooms, etc.)	• The size of the room should be proportional to the number of students • Provide enough seats for students. • The number of students do not exceed on 30 in the classroom
Technology equipment (projector, smart board, software)	• Mathematics Lab is equipped with a computer. • Provide overhead projectors and related items i.e smart Board, Wi-Fi, AV. • Updated Math Software i. e Mathematica, Matlab, Maple. etc
Other equipment (depending on the nature of the specialty)	NA

F. Assessment of Course Quality

Assessment Areas/Issues	Assessor	Assessment Methods
Effectiveness of teaching	Students	Indirect
Effectiveness of Students assessment	Peer Reviewer	Direct
Quality of learning resources	Faculty	Direct
The extent to which CLOs have been achieved	Peer Reviewer	Direct
Other		

Assessors (Students, Faculty, Program Leaders, Peer Reviewer, Others (specify)

Assessment Methods (Direct, Indirect)

G. Specification Approval

COUNCIL /COMMITTEE	
REFERENCE NO.	

DATE	