Course Specification

(Bachelor)

Course Title:	Introduction to Data Science
Course Code:	CS 470
Program:	Computer Science
Department:	Computer Science
College:	Computer Science/ Information Technology
Institution:	Dr. Zamil S. Alzamil
Version:	v1.0
Last Revision Date:	15 October 2023

Table of Contents

A. General information about the course:

1. Course Identification

1. Credit hours: (3 (3,1,0))					
2. Course type					
A.	☐University	☐College	☐Department	☒ Track	☐Others
B.	☐ Required		☒ Elective		
3. Level/year at which this course is offered: (Track)					
4. Course general Description:					
Data Science is the study of the generalizable extraction of knowledge from data. Being a data scientist requires an integrated skill set spanning mathematics, statistics, machine learning, and databases and other branches of computer science along with a good understanding of the craft of problem formulation to engineer effective solutions. Students will learn concepts, techniques and tools they need to deal with various facets of data science practice, including data collection and integration, exploratory data analysis, predictive modeling, descriptive modeling, data product creation, evaluation, and effective communication.					
5. Pre-requirements for this course (if any)**:**					
STAT 102					
6. Pre-requirements for this course (if any)**:**					
7. Course Main Objective(s):					
1. **To cover the basics of data science.** 2. **To give an overview of statistical parameters used in data science.** 3. **To demonstrate how to implement various machine learning algorithms for data analysis.** 4. **To implement several machine learning algorithms in R**					

2. Teaching mode (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	60	100%

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

3. Contact Hours (based on the academic semester)

No	Activity	Contact Hours
1.	**Lectures**	40
2.	**Laboratory/Studio**	20
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	CLO1: To understand the basics of data science	**K1**	Classroom Teaching	Quiz, Mid Exam, Final Exam
1.2	CLO2: To understand the data sampling, statistical analysis, visual data exploration	**K1**	Classroom Teaching	Quiz, Mid Exam, Final Exam
...				
2.0	**Skills**			
2.1	CLO3: To apply predictive analytics techniques for real time problems	**S2**	Classroom Teaching and Lab Exercises	Lab Based Assignments, Mid Exam, Final Exam

Code	Course Learning Outcomes	Code of CLOs aligned with program	Teaching Strategies	Assessment Methods
2.2	CLO4: Analyze a complex computing problem and to apply principles of computing and other relevant disciplines to identify solutions	**S2**	Classroom Teaching and Lab Exercises	Lab Based Assignments, Mid Exam, Final Exam
2.3	CLO5: Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.	**S2**	Classroom Teaching and Lab Exercises	Lab Based Assignments
3.0	**Values, autonomy, and responsibility**			
3.1				
3.2				
...				

C. Course Content

No	List of Topics	Contact Hours
1.	INTRODUCTION: WHAT IS DATA SCIENCE? - BIG DATA AND DATA SCIENCE HYPE	4
2.	STATISTICAL INFERENCE	6
3.	EXPLORATORY DATA ANALYSIS AND THE DATA SCIENCE PROCESS	6
4.	BASIC MACHINE LEARNING ALGORITHMS	8
5.	LINEAR REGRESSION	8
6.	K-NEAREST NEIGHBOURS (K-NN)	6
7.	K-MEANS,	5
8.	MOTIVATING APPLICATION: FILTERING SPAM	5
9.	WHY LINEAR REGRESSION AND K-NN ARE POOR CHOICES FOR FILTERING SPAM	4
	NAIVE BAYES AND WHY IT WORKS FOR FILTERING SPAM	**2**
	FEATURE GENERATION AND FEATURE SELECTION DECISION TREES; RANDOM FORESTS	**2**
	RECOMMENDATION ENGINE	**2**

	DATA VISUALIZATION	
	DATA SCIENCE AND ETHICAL ISSUES	**2**
Total		**60**

D. Students Assessment Activities

No	Assessment Activities *	Assessment timing (in week no)	Percentage of Total Assessment Score
1.	Quiz 1	Week 3	5%
2.	Assignment 1	Week 4	5%
3.	Lab Exercise	Week 5	5%
4.	Lab Exercise	Week 6	5%
5.	Midterm Exam	Week 7	20%
6.	Assignment 2	Week 8	5%
7.	Quiz 2	Week 9	5%
8.	Assignment 3	Week 10	5%
9.	Lab Exam	Week 15	5%
10.	Final Exam	Week 16	40%

*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	Analytics in a Big Data World, Wiley 2014, Bart Baesens
Supportive References	
Electronic Materials	
Other Learning Materials	

2. Required Facilities and equipment

Items	Resources
facilities (Classrooms, laboratories, exhibition rooms, simulation rooms, etc.)	Classroom and laboratory

Items	Resources
Technology equipment (projector, smart board, software)	Data show and Smart Board
Other equipment (depending on the nature of the specialty)	Python and R

F. Assessment of Course Quality

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

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

Assessment Methods (Direct, Indirect)

G. Specification Approval

COUNCIL /COMMITTEE	
REFERENCE NO.	
DATE	