



Course Specifications

Course Title:	Introduction to Data Science
Course Code:	CS 471
Program:	Computer Science
Department:	Computer Science
College:	College of Computer and Information Sciences
Institution:	Majmaah University



Table of Contents

A. Course Identification	3
6. Mode of Instruction (mark all that apply)	3
B. Course Objectives and Learning Outcomes	3
1. Course Description.....	3
2. Course Main Objective.....	3
3. Course Learning Outcomes	4
C. Course Content	4
D. Teaching and Assessment	5
1. Alignment of Course Learning Outcomes with Teaching Strategies and Assessment Methods.....	5
2. Assessment Tasks for Students	5
E. Student Academic Counseling and Support	6
F. Learning Resources and Facilities	6
1. Learning Resources	6
2. Facilities Required.....	7
G. Course Quality Evaluation	7
H. Specification Approval Data	7



A. Course Identification

1. Credit hours:	3 (3+0+1)
2. Course type	
a.	University <input type="checkbox"/> College <input checked="" type="checkbox"/> Department <input type="checkbox"/> Others <input type="checkbox"/>
b.	Required <input checked="" type="checkbox"/> Elective <input type="checkbox"/>
3. Level/year at which this course is offered:	Level-5 / 3
4. Pre-requisites for this course (if any):	CS 120
5. Co-requisites for this course (if any):	

6. Mode of Instruction (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	Traditional classroom	44	100%
2	Blended		
3	E-learning		
4	Distance learning		
5	Other		

7. Contact Hours (based on academic semester)

No	Activity	Contact Hours
1	Lecture	22
2	Laboratory/Studio	22
3	Tutorial	
4	Others (specify)	
	Total	44

B. Course Objectives and Learning Outcomes

1. Course 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, 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.
2. Course Main Objective
1) To cover the basics of data science



2) To give overview of statistical parameters used in data science
To demonstrate how to implement various machine learning algorithms for data analysis 4) To implement several machine learning algorithms in R

3. Course Learning Outcomes

CLOs		Aligned PLOs
1	Knowledge:	
1.1	CLO1: Identify probability distributions commonly used as foundations for statistical modeling	K1
1.2	CLO4: Identify basic Feature Selection algorithms Decision Trees, Random Forests and use in applications	K1
1.3		
1...		
2	Skills :	
2.1	CLO2: Apply basic tools (plots, graphs, summary statistics) to carry out Exploratory Data Analysis.	S2
2.2	CLO3: Apply basic machine learning algorithms (Linear Regression, k-Nearest Neighbors (kNN), k-means, Naive Bayes) for predictive modeling.	S2
2.3	CLO5: Use R language to carry out basic statistical modeling and analysis	S2
2...		
3	Values:	
3.1		
3.2		
3.3		
3...		

C. Course Content

No	List of Topics	Contact Hours
1	Introduction: What is Data Science? - Big Data and Data Science hype - Current landscape of perspectives, Skill sets needed	4
2	Statistical Inference - Populations and samples - Statistical modeling, probability distributions, fitting a model	4
3	Data Design and Implementation (algorithm analysis, growth of functions, ADTs)	4
4	Exploratory Data Analysis and the Data Science Process	4
5	Basic tools (plots, graphs and summary statistics) of EDA	4
6	Three Basic Machine Learning Algorithms - Linear Regression Naive Bayes and why it works for Filtering Spam	4

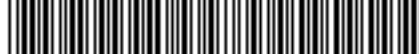


7	k-Nearest Neighbors (k-NN)& k-means	4
8	One More Machine Learning Algorithm and Usage in Applications , Filtering Spam - Why Linear Regression and k-NN are poor choices for Filtering Spam	4
9	Algorithmic ingredients of a Recommendation Engine - Dimensionality Reduction - Singular Value Decomposition - Principal Component Analysis	4
10	Data Visualization - Basic principles, ideas and tools for data visualization	4
11	Data Science and Ethical Issues - Discussions on privacy, security, ethics	4
Total		44

D. Teaching and Assessment

1. Alignment of Course Learning Outcomes with Teaching Strategies and Assessment Methods

Code	Course Learning Outcomes	Teaching Strategies	Assessment Methods
1.0	Knowledge		
1.1	CL01: Identify probability distributions commonly used as foundations for statistical modeling	Classroom and practice Teaching Laboratory	Midterm Exam, Quizzes, Final Exam
2.0	Skills		
2.1	CL02: Apply basic tools (plots, graphs, summary statistics) to carry out Exploratory Data Analysis.	Classroom and practice Teaching Laboratory	Midterm Exam, Quizzes, Final Exam Lab Exercise
2.2	CL03: Apply basic machine learning algorithms (Linear Regression, k-Nearest Neighbors (KNN), k-means, Naive Bayes) for predictive modeling.	Classroom and practice Teaching Laboratory	Lab Exercise
2.3	CL04: Identify basic Feature Selection algorithms Decision Trees, Random Forests and use in applications	Classroom and practice Teaching Laboratory	Midterm Exam, Quizzes, Final Exam Lab Exercise
2.4	CL05: Use R language to carry out basic statistical modeling and analysis.	Classroom and practice Teaching Laboratory	Midterm Exam, Quizzes, Final Exam Lab Exercise
3.0	Values		



Code	Course Learning Outcomes	Teaching Strategies	Assessment Methods
3.1			
3.2			
...			

2. Assessment Tasks for Students

#	Assessment task*	Week Due	Percentage of Total Assessment Score
1	Quiz 1	Week 3	10%
2	Lab Assignment 1	Week 3	10%
3	Mid Term	Week 6	20%
4	Lab Assignment 1	Week 7	10%
5	Quiz 2	Week 8	10%
6	Final Exam	Week 12	40%

*Assessment task (i.e., written test, oral test, oral presentation, group project, essay, etc.)

E. Student Academic Counseling and Support

Arrangements for availability of faculty and teaching staff for individual student consultations and academic advice :

- Each student is allotted to an academic advisor for guidance and counselling.
- Available for a minimum of 4 hours per week/course, as communicated to the students.
- Student also contacts through social networking websites / D2L/ Email for advice and consultations

F. Learning Resources and Facilities

1. Learning Resources

Required Textbooks	Cathy O'Neil and Rachel Schutt. <i>Doing Data Science, Straight Talk From The Frontline</i> . O'Reilly. 2014
Essential References Materials	<ul style="list-style-type: none"> •Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. <i>Mining of Massive Datasets</i>. v2.1, Cambridge University Press. 2014. •Kevin P. Murphy. <i>Machine Learning: A Probabilistic Perspective</i>. ISBN 0262018020. 2013 <i>NINA ZUMEL, JOHN MOUNT, Practical Data Science with R, Manning Publications Co.,2014</i>



Electronic Materials	
Other Learning Materials	

2. Facilities Required

Item	Resources
Accommodation (Classrooms, laboratories, demonstration rooms/labs, etc.)	Classrooms, laboratories
Technology Resources (AV, data show, Smart Board, software, etc.)	PC or Laptop with Windows/Linux, Smart Board, LCD Projector, 'R' programming Tool, Blackboard
Other Resources (Specify, e.g. if specific laboratory equipment is required, list requirements or attach a list)	Internet connection

G. Course Quality Evaluation

Evaluation Areas/Issues	Evaluators	Evaluation Methods
Final Exam Answer Scripts Verification	Peer faculty members	Direct Review
Course Feedback	Students	Indirect course Survey

Evaluation areas (e.g., Effectiveness of teaching and assessment, Extent of achievement of course learning outcomes, Quality of learning resources, etc.)

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

Assessment Methods (Direct, Indirect)

H. Specification Approval Data

Council / Committee	
Reference No.	
Date	