Course Specifications

Course Title:	**Data Science 2**
Course Code:	**AI 422**
Program:	**Information and Computer Science**
Department:	**Computer Science and Information**
College:	**Science at Al-Zulfi**
Institution:	**Majmaah**

Table of Contents

A. Course Identification

1. Credit hours: 3
2. Course type **a.** University ☐ College ☐ Department ■ Others ☐ **b.** Required ☐ Elective ☐
3. Level/year at which this course is offered: Level 9
4. Pre-requisites for this course (if any)**:** Data Science (1)- AI 324
5. Co-requisites for this course (if any)**: Nil**

6. Mode of Instruction (mark all that apply)

No	Mode of Instruction	Contact Hours	Percentage
1	**Traditional classroom**	44	80 %
2	**Blended**	3	5 %
3	**E-learning**	3	5 %
4	**Correspondence**	3	5 %
5	**Other**	3	5 %

7. Actual Learning Hours (based on academic semester)

No	Activity	Learning Hours
Contact Hours		
1	**Lecture**	30
2	**Laboratory/Studio**	20
3	**Tutorial**	5
4	**Others** (specify)	5
	Total	**60**
Other Learning Hours*		
1	**Study**	10
2	**Assignments**	20
3	**Library**	10
4	**Projects/Research Essays/Theses**	20
5	**Others** (specify)	0
	Total	**60**

* The length of time that a learner takes to complete learning activities that lead to achievement of course learning outcomes, such as study time, homework assignments, projects, preparing presentations, library times

B. Course Objectives and Learning Outcomes

1. Course Description

- **Mathematical Models** : Linear models , Non-Linear models , Descriptive Models , Stochastic and Deterministic Models.
- **Linear Algebra and Regression**: Visualizing Matrix Operation, Factoring Matrices, Eigenvalues and Eigenvectors and Singular Value Decomposition, Principal Components Analysis, Stochastic Gradient Descent, Classification and Logistic Regression, Issues in Logistic Classification.
- **Classification and Clustering:** Decision Tree Classifiers, Nearest Neighbor Classification, *k*-Nearest Neighbors and Induced Networks, k-means Clustering, Agglomerative Clustering.
- **Machine Learning**: Naive Bayes and Support Vector Machines, Supervised and Unsupervised Learning, Deep Learning.
- **Big Data**: Big Data as Bad Data , Algorithmic for Big Data , Filtering and Sampling , Data Parallelism

2. Course Main Objective

Student will:

1. have the ability to demonstrate advanced independent critical enquiry, analysis and reflection;
2. have a strong sense of intellectual integrity and the ethics of scholarship;
3. have in-depth knowledge of modern statistical methodology and computing
4. reach a high level of achievement in writing, research or project activities, problem-solving and communication;
5. be critical and creative thinkers, with an aptitude for continued self-directed learning;
6. be able to examine critically, synthesis and evaluate knowledge across a broad range of disciplines;
7. have a set of flexible and transferable skills for different types of employment; and
8. be able to initiate and implement constructive change in their communities, including professions and workplaces

3. Course Learning Outcomes

Upon successful completion, students will have the knowledge and skills to:

	CLOs	Aligned-PLOs
1	**Knowledge:**	
1.1	Demonstrate expertise in machine learning methods and strategies for advanced data mining, expertise in database systems, and expertise in computational statistics.	K3-AI
1.2	Integrate and apply this expertise to produce solutions for real-world problems using public and private data sources.	
1.3	Have the ability to demonstrate advanced independent critical enquiry, analysis and reflection;	
1.4	Have a sound knowledge of modern statistical methodology and computing that will equip them for a career in data science and enable their careers to develop as data science evolves.	

CLOs		Aligned-PLOs
2	**Skills :**	
2.1	Distinguish between different kinds of data models and identify challenges related to big data.	S3- AI
2.2	The ability to apply skills to primary research in data science and disciplines relevant to data science applications using Python programming.	
2.3	Apply skills in the evaluation and synthesis of information from big data	
3	**Competence: الكفاءات**	
3.1	Execute statistical analysis with professional software(Python).	C3-AI
3.2	Applying proficiency kinds of data models and identify challenges related to big data using Python.	
3.3	Applying supervised and unsupervised learning methods on dataset using Python.	

C. Course Content

No	List of Topics	Contact Hours
1	- **Mathematical Models** : Linear vs. Non-Linear Models , Blackbox vs. Descriptive Models , First-Principle vs. Data-Driven Models , Stochastic vs. Deterministic Models , Baseline Models , Evaluating Models	12
2	- **Linear Algebra**: Visualizing Matrix Operation , Factoring Matrices , Eigenvalues and Eigenvectors , Eigenvalue Decomposition , Singular Value Decomposition (SVD) , Principal Components Analysis(PCA) .	8
3	- **Linear and Logistic Regression** : Linear Regression , Better Regression Models , Regression as Parameter Fitting , Gradient Descent Search , Stochastic Gradient Descent , Classification and Logistic Regression , Issues in Logistic Classification	8
4	- **Distance and Network Methods**: Measuring Distances Metrics , Nearest Neighbor Classification , *k*-Nearest Neighbors , Locality Sensitive Hashing , G , Weighted Graphs and Induced Networks	8
5	- **Clustering**: *k*-means Clustering , Agglomerative Clustering , Comparing Clustering's , Similarity Graphs and Cut-Based Clustering and Cluster Bombing	8
6	- **Machine Learning** : Naive Bayes , Decision Tree Classifiers , Boosting and Ensemble Learning and Support Vector Machines , Supervised and Unsupervised Learning , Deep Learning	8
7	- **Big Data**: Big Data as Bad Data , Algorithmic for Big Data , Filtering and Sampling , Data Parallelism , Cloud Computing Services , Map-Reduce Programming , Societal and Ethical Implications	8
Total		60

D. Teaching and Assessment

1. Alignment of Course Learning Outcomes with Teaching Strategies and Assessment Methods Live Learning: Lecture, PowerPoint slides and discussion

Code	Course Learning Outcomes	Teaching Strategies	Assessment Methods
1.0	**Knowledge**		
1.1	Demonstrate expertise in machine learning methods and strategies for advanced data mining, expertise in database systems, and expertise in computational statistics.	▪ **Direct Teaching:** Lectures, PowerPoint slides and discussion. ▪ **Aimed Teaching** Discovery and Oral Questions.	- Homework tasks - Quiz - Midterms - Final Exam
1.2	Integrate and apply this expertise to produce solutions for real-world problems using public and private data sources.		
1.3	Have the ability to demonstrate advanced independent critical enquiry, analysis and reflection;		- E-learning - Internet search - Oral Exam
1.4	Have a sound knowledge of modern statistical methodology and computing that will equip them for a career in data science and enable their careers to develop as data science evolves.		
2.0	**Skills**		
2.1	Distinguish between different kinds of data models and identify challenges related to big data.	**Indirect Teaching:** Brainstorming - Free Discovery – Inquiry	- Lab Exercises - Lab Exam - Oral Exam - Presentations
2.2	The ability to apply skills to primary research in data science and disciplines relevant to data science applications using Python programming.		
2.3	Apply skills in the evaluation and synthesis of information from big data		
3.0	**Competence**		
3.1	Execute statistical analysis with professional software(Python).	**Course Project: (Work group)** critical thinking and ability to seek solutions.	Introduce group project and case study approaches to enable students to have an experience in problem solving situations.
3.2	Applying proficiency kinds of data models and identify challenges related to big data using Python.		
3.3	Applying supervised and unsupervised learning methods on dataset using Python.		

2. Assessment Tasks for Students

#	Assessment task*	Week Due	Percentage of Total Assessment Score
1	Homework 1	2	2%
2	QUIZ 1	3	5%
3	Homework 2	4	2%
4	QUIZ 2	5	5%
5	Midterm 1	6	10%
6	Homework 3	7	2%
7	QUIZ 3	8	5%

#	Assessment task*	Week Due	Percentage of Total Assessment Score
8	Homework 4	9	2%
9	QUIZ 4	10	5%
10	Midterm 2	11	10%
11	Lab Exam/ Project Evaluation	14	12%
12	Final Exam	16	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:

- Determine meeting appointments for the weak' students to solve their problems and give them academic advices.
- One office hour daily
- Dealing a workshops.
- Motivate students

F. Learning Resources and Facilities

1.Learning Resources

Required Textbooks	The Data Science Design Manual , Steven S. Skiena , Springer , 2017 , ISSN 1868-0941 , ISBN 978-3-319-55443-3
Essential References Materials	The Data Science Handbook, Cady Field , Wiley, 2017, ISBN-13: 978-1119092940
Electronic Materials	https://www.kaggle.com/learn/overview
Other Learning Materials	Matlab toolboxes: Data mining/ Data Science / machine learning

2. Facilities Required

Item	Resources
Accommodation (Classrooms, laboratories, demonstration rooms/labs, etc.)	Classroom - Laboratory
Technology Resources (AV, data show, Smart Board, software, etc.)	Data show – Smart Board
Other Resources (Specify, e.g. if specific laboratory equipment is required, list requirements or attach a list)	Matlab software – Weka – Python Programming

G. Course Quality Evaluation

Evaluation Areas/Issues	Evaluators	Evaluation Methods
1. Questionnaires (course evaluation) filled by the students and acquired electronically by the University	Students	Indirect Assessment
2. Students-faculty management meetings		
3. Midterms and Final Exam	Course Coordinator Staff	Direct Assessment
4. Project Evaluation		
5. Departmental internal review of the course.	Reviewer Committee	Final Exam Evaluation
6. Course Portfolio	External Reviewer	Course Evaluation

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	