Course Description Form

1. Course Name:

Large Scale Machine Learning

2. Course Code:

3. Semester / Year:

First / MSc

4. Description Preparation Date:

20/10/2024

5. Available Attendance Forms:

In the class

6. Number of Credit Hours (Total) / Number of Units (Total)

Name: Prof. Dr. Ban N. Dhannoon
Email: ban.n.dhannoon@nahrainuniv.edu.iq

7. Course administrator's name (mention all, if more than one name)

8. Course Objectives

Course Objectives	• This course provides a quick and indirect introduction to the basic concepts and algorithms related to learning. • It includes the limitations and tools needed for software that can learn from data. • We'll cover a few of the benefits, the most commonly used ones (such as linear regression

9. Teaching and Learning Strategies

Strategy	• Present lectures in a PowerPoint slide show • Implement some algorithms manually to increase understanding • Request a program to implement a specific algorithm

10. Course Structure

Week	Hours	Required Learning Outcomes	Unit or subject name	Learning method	Evaluation method
1	3	learning, ML life cycle, Main Challenges of ML (Data, Model, Ethical Challenges), Types of data, Bias and Variance in ML, overfitting and underfitting	General Introduction	PowerPoint	
2	3	Neural Networks, architecture, learning MLP	ML Techniques	PowerPoint	

		Dropout Activation Functions NN example with BP			
3	3	Model Validation, Linear Regression model with example	Learning Methods	PowerPoint	
4	3	Types of GD, • Statistics of Regression Analysis (r and r-squared), with example • Normal Equation with example	Gradient Descent in Linear Regression	PowerPoint and on–board	
5	3	• Data Sets, Objects, and Attribute Types; • Measuring the Central Tendency (mean, median, mode), • Symmetric vs Skewed data • Measuring the Dispersion of Data • Graphic Displays (Boxplot & Scatter plot)	Know your data	PowerPoint and on–board	
6	3	Mid1			
7	3	Assignment 1		PowerPoint	
8	3	• Data Visualization • Pixel-Oriented Techniques • Geometric -Oriented Techniques • Measuring similarity and dissimilarity • Proximity Measures for Nominal Attributes • Dissimilarity (Nominal) Example • Proximity Measures for Binary Attributes • Dissimilarity of Numeric Data	Data Visualization	PowerPoint and on–board	
		• Vector Objects • Cosine Similarity • Tanimoto coefficient (Tanimoto similarity) • Confusion matrix (binary vs multi classes) • statistical Measures			
9	3	Classification Algorithms • Logistic Regression • Decision Tree • Random Forest • Support Vector Machine	Supervised classification	PowerPoint and on–board	
10	3	• Naïve Bayes Classifier • K- Nearest Neighbors (classification and Regression)	Classification and regression	PowerPoint and onboard	
11	3	Mid2			
12	3	• Hard and soft clustering • Clustering Algorithms: • Centroid-based Clustering • Density-based Clustering	Clustering	PowerPoint	

		• Connectivity-based Clustering			
13	3	Hierarchical clustering	Unsupervised Learning	PowerPoint	
14	3	Feature Extraction and Selection	Dimensionality reduction	PowerPoint	
15	3	Association Rule mining	Frequent items		

1. Course Evaluation

Final Exam 70
Mid1 Exam 7.5
Mid2 Exam 7.5
Two monthly projects 10
Quizes 5

2. Learning and Teaching Resources

Required textbooks (curricular books, if any	Hands-On Machine Learning with Scikit-Learn and TensorFlow: Conce Tools, and Techniques to Build Intelligent Systems, Aurelien Geron, Ori 2023.
Main references (sources)	Machine Learning for Business Analytics, 2023
Recommended books and references (scientific journals, reports…)	Machine Learning Bookcamp, 2021
Electronic References, Websites	Stanford University Machine Learning CS229 Lecture notes by Andrew Ng. Youtube