



Dr. Hashim Yasin

Department of Artificial Intelligence & Data Science.

## AI3002 –Machine Learning Assignment No. 2

### Assignment Submission Guidelines:

1. Submit your assignment in **soft form (Code + Report)** within the due date and time. Soft form does not mean submitting photos of the hardcopy. Late submissions will result in a deduction of marks.
2. The **report** must include a discussion, comments, and a conclusion about your solution. Submitting without a report will result in a loss of full marks.
3. Name the zip or other folder/file that you submit using the following format: **ML\_A2\_RollNo\_FirstName**.
4. Ensure that you solve each task of the assignment on your own.
5. You are allowed to do your assignment in **groups of a maximum of two members**.
6. There is no restriction on the programming language used for the tasks.
7. For programming tasks, you are NOT allowed to use any built-in functions or libraries for specific tasks.

---

### Decision Tree, Random Forest and Naïve Bayes Classifiers

Download training dataset (Iris dataset) from the following link

<https://www.kaggle.com/datasets/uciml/iris>

The **Iris dataset** consists of three different species of iris flowers: *Iris setosa*, *Iris versicolor*, and *Iris virginica*, with 50 samples per species, totaling 150 samples. The dataset includes four key features for each flower, sepal length, sepal width, petal length, and petal width.

Your task is to,

### **1. Build (implement) Decision Tree classifiers (ID3 and CART)**

- Visualize the decision tree structures for both algorithms.
- Evaluate and compare the performance of both ID3 and CART decision trees, using evaluation metrics.

### **2. Build (implement) Random Forest**

- Train a Random Forest on the same dataset (Iris dataset) with an optimal number of trees.
- Discuss and validate the optimal number of trees.
- Visualize your results.

### **3. Implement Naïve Bayes classifier**

- Train a Naïve Bayes classifier on the same dataset.
- Discuss and visualize the results.

### **4. Compare Classifiers**

- Compare the performance of Decision Trees, Random Forest, and Naïve Bayes using evaluation metrics.
- Visualize the comparison results.