Artificial Intelligence and Machine Learning

Term Project

Classification/Regression

**Due Date (31/5/2025)**

---

In this project you are asked to apply the concepts that you learned in this course to solve a problem of your choice (regression or Classification), you could obtain dataset from **Kaggle** or any other data resources. You have to follow the instructions below:

1- **Dataset Selection**: You have to look for a dataset that you can use in your project. Here is some general guidelines on how to select the dataset:
   a. The number of sample in the dataset should be large enough to be able to develop an accurate model. For tabular data more than 700 samples should be sufficient.
   b. The dataset should be balanced.
   c. Avoid working on dataset with images due to the limited hardware resources. If you have laptop with GPUs or you can use the Colabs then you will not face this issue.
   d. You should get the **approval** on the dataset before start working on the project. *You can get the approval by sending a message to me on MS teams with the following details*:
      1- Dataset name
      2- Number of Features
      3- Number of Instances (should be more 1000)
      4- Target Column
      5- In case of classification you have to check if the data is balanced Is the data balanced (the number of instances in each target classes is balanced.)
   e. In the Jupyter notebook you should write general description about the data used. Number of feature and in which domain this data will be used.
2- **Data Preparation:** You have to prepare the data for training using the techniques that you learned in this course including data visualization and analysis.
3- **Model Training and Testing:** You should train different models using different metrics and select the model with the best performance results. Results analysis should be thorough.

You should submit your work before the deadline. The Jupyter notebook should be comprehensive and contains explanation and reasoning about the used procedure.

**Submission Guidelines:**

- Submit your work as a PDF file (exported Jupyter notebook) + Jupyter notebook on Ms teams.
- You are allowed to work on this assignment in teams of two
- **Students names and their IDs should be clearly listed at the beginning** of the notebook
- Add a brief text before each code cell to explain the purpose of the code block
- The submitted work should be original and not copied from any online resource or other teams in the class.