Project Summary:

With access to the three datasets provided in the next section,

- Download, analyze and choose one of them as a team
- Start working through the steps outlined in the requirements section

Make sure to refer to the descriptions provided on Kaggle for each dataset before diving into the requirements.

Dataset(s):

- Gender Classification Dataset (kaggle.com)
- Garbage Classification Dataset
- Bone Fracture Multi-Region X-Ray Dataset

Requirements:

- **Data exploration and preparation:**
  - Reshape the RGB images, so that the dimension of each image is (64,64,3)
  - Convert the RGB images to greyscale.
  - Normalize each image.

- **Experiments and results:**
  - Split the data into training and testing datasets (if there is no testing dataset)
  - *First experiment:*
    - Train an SVM model on the grayscale images.
    - Test the model and provide the confusion matrix and the average f-1 scores for the suitable testing dataset.
  - Split the training dataset into training and validation datasets. (if there is no validation dataset)
  - *Second experiment:*
    - Build 2 different Neural Networks (different number of hidden layers, neurons, activations, etc.)
    - Train each one of these models on the grayscale images and plot the error and accuracy curves for the training data and validation data.
    - Save the best model in a separated file, then reload it.
    - Test the best model and provide the confusion matrix and the average f-1 scores for the testing dataset.
  - *Third experiment:*
    - Train a Convolutional Neural Network on the grayscale images and plot the error and accuracy curves for the training data and validation data.
    - Train another Convolutional Neural Network on the RGB images and plot the error and accuracy curves for the training data & validation data.
    - Save the best model in a separated file, then reload it.
    - Test the best model and provide the confusion matrix and the average f-1 score fie the suitable testing dataset.
  - Compare the results of the models and suggest the best model.

Deliverables:

You are required to submit ONE zip file containing the following:

- Your code (.py) file. If you have a (.ipynb) file, you have to save/download it as (.py) before submitting.
- A report (.pdf) containing the team members' names and IDs, the dataset you chose and the code with screenshots of the output of each part. If you have a (.ipynb) file, you can just convert it to pdf.

The zip file must follow this naming convention: ID1_ID2_ID3_ID4_ID5_ ID6_Group

Instructions:

1. The minimum number of students in a team is 3 and the maximum is 6.
2. No late submission is allowed.
3. Cheating students will take **ZERO** and no excuses will be accepted.
4. You can use any Python libraries.

Grading Criteria:

Data Preparation	2
SVM	2
Neural Networks	3
Convolutional Neural Networks	3
Total = 10 marks	