



# AIML

# MODULE PROJECT

Computer Vision

TOTAL  
SCORE

60

Part A - 30 Marks

- **DOMAIN:** Botanical Research
- **CONTEXT:** University X is currently undergoing some research involving understanding the characteristics of plant and plant seedlings at various stages of growth. They already have have invested on curating sample images. They require an automation which can create a classifier capable of determining a plant's species from a photo.
- **DATA DESCRIPTION:** The dataset comprises of images from 12 plant species.  
Source: <https://www.kaggle.com/c/plant-seedlings-classification/data>.
- **PROJECT OBJECTIVE:** To create a classifier capable of determining a plant's species from a photo.

Steps and tasks: [ Total Score: 30 Marks]

1. Import and Understand the data [12 Marks]
  - A. Extract 'plant-seedlings-classification.zip' into new folder (unzipped) using python. [2 Marks]  
*Hint: You can extract it Manually by losing 2 marks.*
  - B. Map the images from train folder with train labels to form a DataFrame. [6 Marks]  
*Hint: Create a DataFrame with 3 columns: Name of image, Species/class/type of image & actual image..*
  - C. Write a function that will select n random images and display images along with its species. [4 Marks]  
*Hint: If input for function is 5, it should print 5 random images along with its labels.*
2. Data preprocessing [8 Marks]
  - A. Create X & Y from the DataFrame. [2 Marks]
  - B. Encode labels of the images. [2 Marks]
  - C. Unify shape of all the images. [2 Marks]
  - D. Normalise all the images. [2 Marks]
3. Model training [10 Marks]  
*Checkpoint: Please make sure if shape of X is (No.of images, height, width, No. Of channels). If not, you need to correct it otherwise it will be issue during model training.*
  - A. Split the data into train and test data. [2 Marks]
  - B. Create new CNN architecture to train the model. [4 Marks]
  - C. Train the model on train data and validate on test data. [2 Marks]
  - D. Select a random image and print actual label and predicted label for the same. [2 Marks]

Part B - 30 Marks

- **DOMAIN:** Botanical Research
- **CONTEXT:** University X is currently undergoing some research involving understanding the characteristics of flowers. They already have have invested on curating sample images. They require an automation which can create a classifier capable of determining a flower’s species from a photo.
- **DATA DESCRIPTION:** The dataset comprises of images from 17 plant species.
- **PROJECT OBJECTIVE:** To experiment with various approaches to train an image classifier to predict type of flower from the image.

Steps and tasks: [ Total Score: 30 Marks]

1. Import and Understand the data [5 Marks]
  - A. Import and read oxflower17 dataset from tflearn and split into X and Y while loading. [2 Marks]

*Hint: It can be imported from tflearn.datasets. If tflearn is not installed, install it.*

*It can be loaded using: x, y = oxflower17.load\_data()*
  - B. Print Number of images and shape of the images. [1 Marks]
  - C. Print count of each class from y. [2 Marks]
2. Image Exploration & Transformation [Learning purpose - Not related to final model] [10 Marks]
  - A. Display 5 random images. [1 Marks]
  - B. Select any image from the dataset and assign it to a variable. [1 Marks]
  - C. Transform the image into grayscale format and display the same. [3 Marks]
  - D. Apply a filter to sharpen the image and display the image before and after sharpening. [2 Marks]
  - E. Apply a filter to blur the image and display the image before and after blur. [2 Marks]
  - F. Display all the 4 images from above questions besides each other to observe the difference. [1 Marks]
3. Model training and Tuning: [15 Marks]
  - A. Split the data into train and test with 80:20 proportion. [2 Marks]
  - B. Train a model using any Supervised Learning algorithm and share performance metrics on test data. [3 Marks]
  - C. Train a model using Neural Network and share performance metrics on test data. [4 Marks]
  - D. Train a model using a basic CNN and share performance metrics on test data. [4 Marks]
  - E. Predict the class/label of image ‘Prediction.jpg’ using best performing model and share predicted label. [2 Marks]