

# **Flower Image Classification**

---

## **Database Management System Project**

### **Submitted by :**

---

- **Aayush Agarwal (17BCE1204)**
  - **Himanshu Tiwari (17BCE1331)**
  - **Pankaj Mangla (17BCE1265)**
  - **Ashray Saluja (17BCE1319)**
- 

### **Image Classification using Python and Machine Learning**

---

This repo contains the code to perform a simple image classification task using Python and Machine Learning. We will apply global feature descriptors such as Color Histograms, Haralick Textures and Hu Moments to extract features from FLOWER17 dataset and use machine learning models to learn and predict.

### **Summary of the project**

---

- Global Feature Descriptors such as Color Histograms, Haralick Textures and Hu Moments are used on FLOWER17 dataset.
- Classifiers used are Logistic Regression, Linear Discriminant Analysis, K-Nearest Neighbors, Decision Trees, Random Forests, Gaussian Naive Bayes and Support Vector Machine.

## Software Requirements

- Anaconda : Jupyter Notebook
- HDF5 Reader

## Hardware Requirements

- Intel P4 1.5GHz or above.
- 512MB RAM.
- 80GB HDD Minimum.

## Other Requirement

- Dataset : FLOWER17

## Execution Procedure of the Code.

- Using Anaconda Prompt , Open Jupyter Notebook .
- In Jupyter Notebook , search and open the given 'Image Classification' Folder .
- Firstly, open and run 'globalfeature.py' file in Python3 shell.
- Then after the completion of execution of the above file, open and run 'train\_test\_new.py' file.
- The output of the above file results in the FLOWER CLASSIFICATION output.

## Feature Extraction

Features are the information or list of numbers that are extracted from an image. These are real-valued numbers (integers, float or binary). There are a wider range of feature extraction algorithms in Computer Vision.

When deciding about the features that could quantify plants and flowers, we could possibly think of **Color, Texture and Shape** as the primary ones. This is an obvious choice to globally quantify and represent the plant or flower image.

But this approach is less likely to produce good results, if we choose only one feature vector, as these species have many attributes in common like **sunflower** will be similar to **daffodil** in terms of color and so on. So, we need to quantify the image by combining different feature descriptors so that it describes the image more **effectively**.