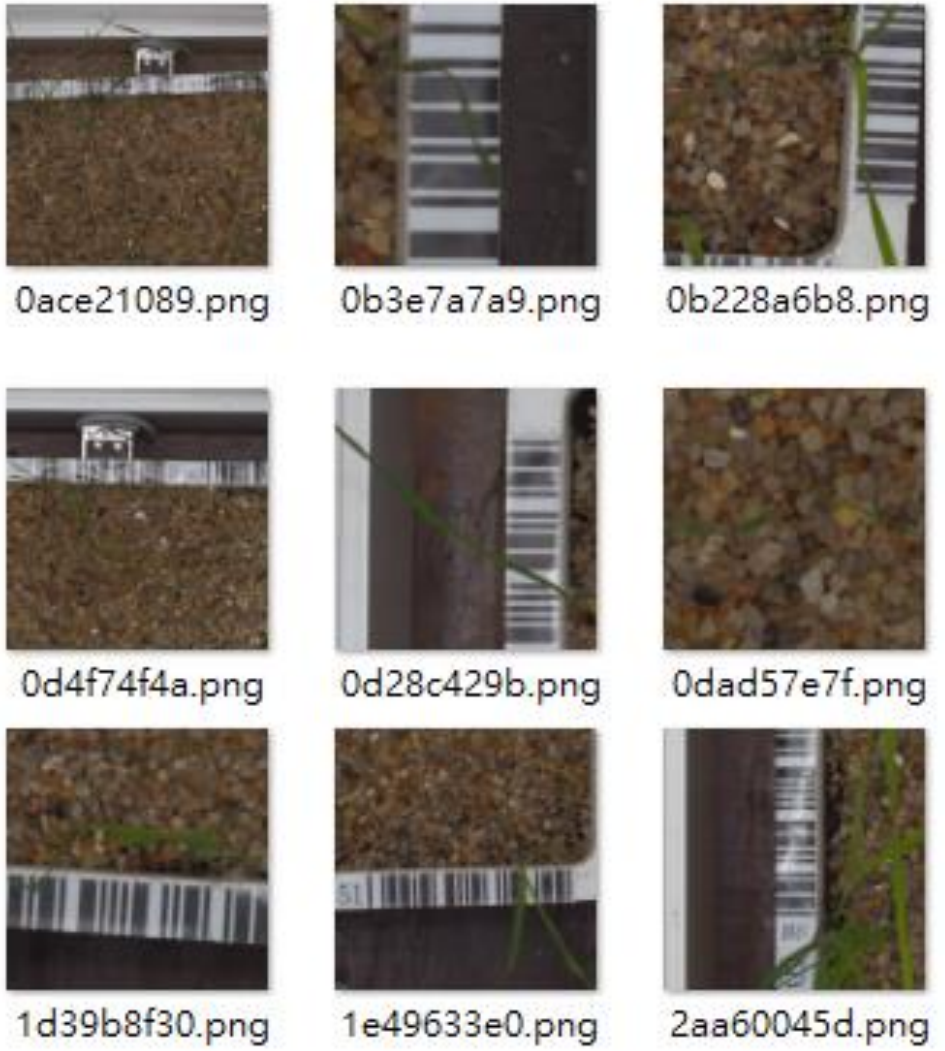
Assignment 2 : Feature Extraction

- Applying the following feature extraction methods to plant dataset:
  - Raw image
  - Color histogram
  - Local Binary Pattern
  - Co-occurrence Matrix
  - Gabor Filters
  - Histogram of Oriented Gradient (HoG)
  - Bag-of-Features

- Choose FOUR features from the list.

Kaggle Plant Classification Dataset

- Black-grass
- Charlock
- Cleavers
- Common Chickweed
- Common wheat
- Fat Hen
- Loose Silky-bent
- Maize
- Scentless Mayweed
- Shepherds Purse
- Small-flowered Cranesbill
- Sugar beet

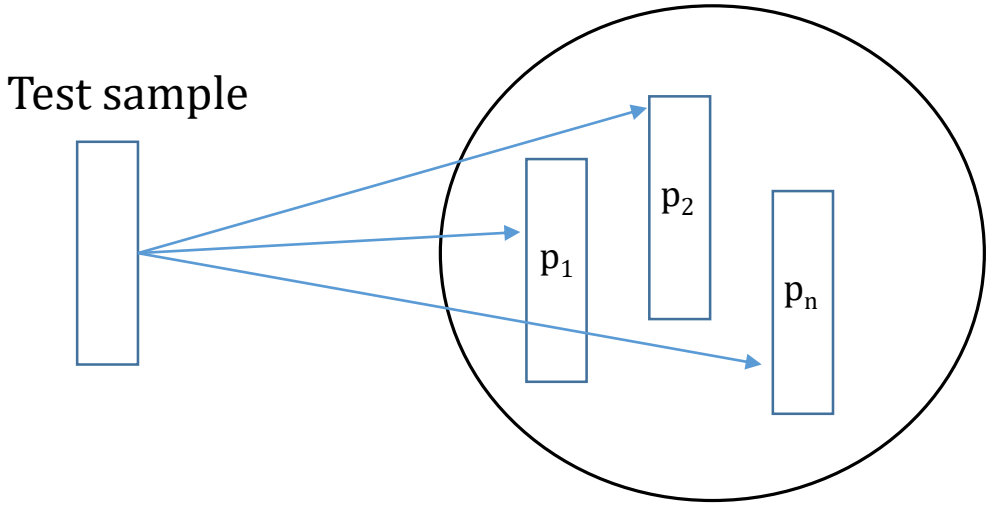


Dataset Details

- Goal
  - to create a classifier capable of determining a plant's species from a photo.
- Task
  - image classification (12 classes)
- Training set
  - 4750 images
- Test set
  - 794 images
- Ref: https://www.kaggle.com/c/plant-seedlings-classification

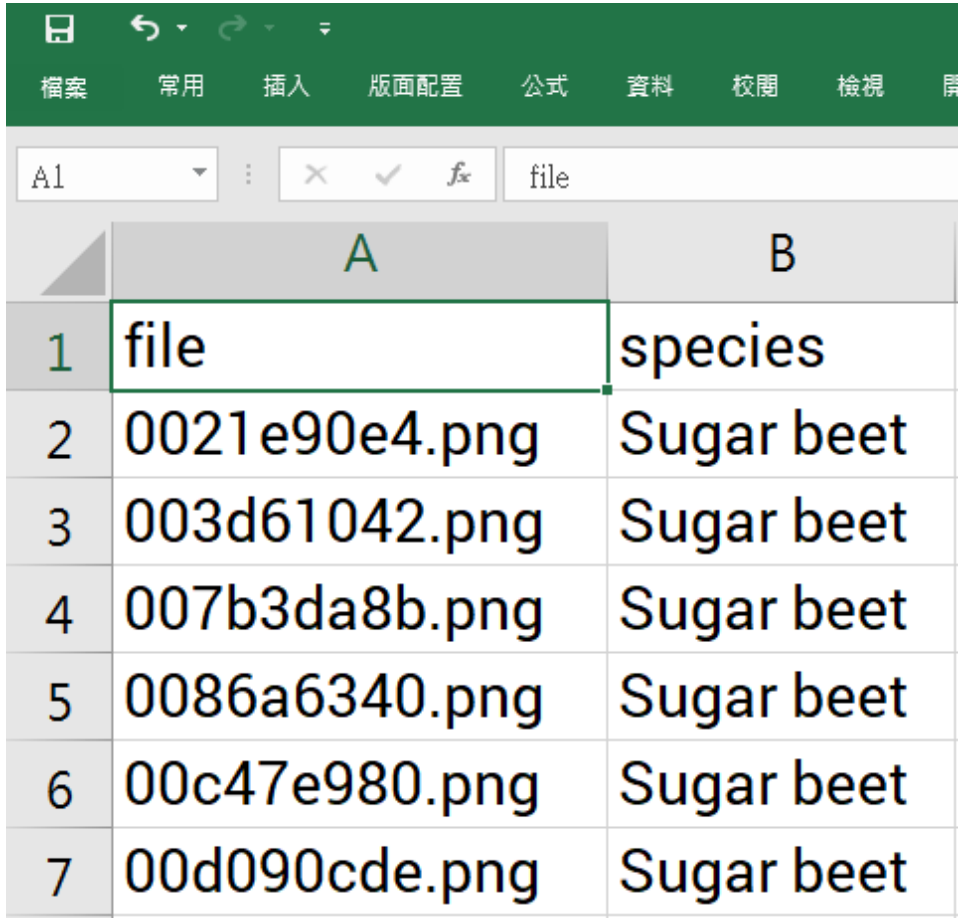
Nearest Neighbor Classification

- For the test sample, find the nearest sample in the training set.
- The nearest neighbor can be found by using
  - SAD – sum of absolute distance
  - SSD – sum of square distance
- Assign the label of the nearest neighbor to the test sample.



Submission Format

- File format: `.csv`
- Columns: `file, species`



	A	B
1	file	species
2	0021e90e4.png	Sugar beet
3	003d61042.png	Sugar beet
4	007b3da8b.png	Sugar beet
5	0086a6340.png	Sugar beet
6	00c47e980.png	Sugar beet
7	00d090cde.png	Sugar beet

Sample submission.

Results Evaluation

- Use Feature + NN to determine the plant class for each image.
- Testing results from Kaggle
  - Use the whole training dataset and test set to obtain the testing results.

Assignment #2 – Plant Classification

- You need to hand in your source code and report
- The report should cover:
  - Method description – what are your reference codes? How to run your test?
  - Experimental results
    - The comparison of classification accuracy from validation
    - The comparison of ranking from Kaggle for each feature
  - Discussion
  - Problem and difficulties
- Upload assignment #2 before 11/21 (TUE), 23:59pm
- File format – zip all your files into a single file: studentID_hw2_version, ex: 602410143_hw2_v1

Assignment Rules

- Late policy
  - You will get 20% deduction of your scores per day.
  - It means if the assignment is delayed one day for 80%, two days for 60%,…, five days for 0% .
- No-copy policy
  - Copying is strictly forbidden in our class.
  - Once the assignment is confirmed by TA as COPY, the score will be 0%.