**D604 Task 1 Image Scenario**

Scenario

Computer Vision—Image Classification—Plant-Seedling Classification (updated)

Botanists have been categorizing species for centuries, especially in an effort to separate crop seedlings from weeds. The ability to effectively separate crop seedlings from weeds can mean customized care for the crop, resulting in better crop yields and better stewardship of the environment. For several decades, researchers have worked on various approaches aimed at performing site-specific species and weed detection. However, the past decade has seen substantial progress in content-based image retrieval and analysis of images using neural networks. The goal of this task is to solve the seedling classification problem for botanists using modern-day image classification techniques.

Dataset Details

The given dataset contains 4750 RGB images of species divided across 12 categories: "Small-flowered Cranesbill," "Fat Hen," "Shepherds Purse," "Common wheat", "Common Chickweed," "Charlock," "Cleavers," "Scentless Mayweed," "Sugar beet," "Maize," "Black-grass," and "Loose Silky-bent."

A graph of different colored bars

Description automatically generated

The dataset contains two files:

* images.npy—images in a numpy array format
* labels.csv—the labels in a .csv file
* [Link to files](https://lrps.wgu.edu/provision/452249989)

There are 4750 RGB images of shape 128 x 128 X 3 (i.e., each image has three channels).

**Reading the Files**

#Reading the files—assuming the files are in the same folder as the notebook/ide.

# Otherwise use the file path

images = images = np.load("images.npy")

labels = pd.read\_csv("labels.csv")

print(images.shape)

>> (4750, 128, 128, 3)

print(labels.shape)

>> (4750, 1)

**Example Images**

A collage of photos of plants

Description automatically generated