**CIS 5367 Machine Learning Project Idea**

TEAM 5: Taja Castilleja, Vaishnavi Krishnan, Christian Martin, Chi Bao, and Shainan Agrawal

* **Introduce a hypothetical business scenario.**

We are a startup company that is developing an app where users take pictures of plants and the app will identify the plants using machine learning. We have premium and free services, where the premium service includes expanded plant image recognition. For this project, we will focus on the free service where plant recognition will be limited to specific types of plants.

* **Describe a way in which a machine learning system can enhance or replace certain business functions.**

We will train the machine learning algorithm with datasets of plant images to identify species of plants after taking a photo. This removes the effort and time to research plant information and type online. The app will use supervised machine learning since in the end the user knows the plant type, and the algorithm will choose the best match using statistical significance.

* **Describe the data you may use as input to your machine learning model.**

We will be using a dataset that includes images of flowers classified into many groups. In total, there are 102 flower categories where each category has 100 images or more. We may gather more photos to train the machine learning algorithm.

* **List/describe the data sources you may obtain/use for your project.**

We will use datasets of images collected by software engineers and related experts in the field of plant image recognition. The data sources are not limited to but may include:

1. [Visual Geometry Group - University of Oxford](https://www.robots.ox.ac.uk/~vgg/data/flowers/102/index.html)
2. [Flower Species Classifier. Build an image classifier to recognize… | by Silvio Mori Neto | Towards Data Science](https://towardsdatascience.com/flower-species-classifier-c8893030cd90)