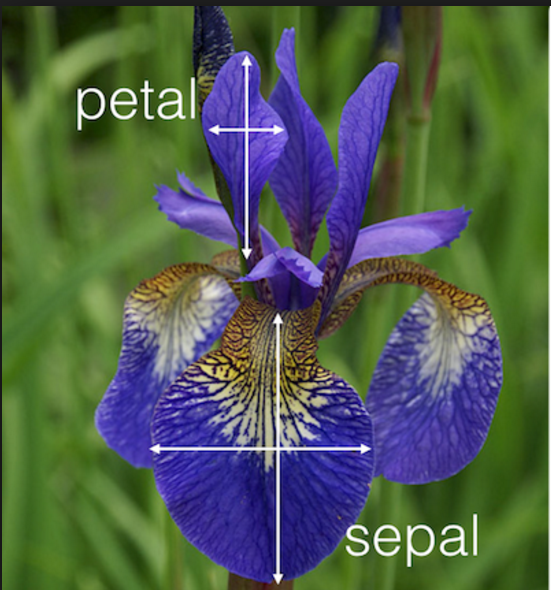
**Problem-1 [70 points]**

Design L1 and L2 distance functions to assess the similarity of bank customers. Each customer is characterized by the following attributes:

* Age (customer’s age, which is a real number with the maximum age is 90 years and minimum age 15 years)
* *Cr* (“*credit rating*”) which is ordinal attribute with values ‘very good’, ‘good, ‘medium’, ‘poor’, and ‘very poor’.
* *Av\_bal* (average account balance, which is a real number with mean 7000, standard deviation is 4000)

1. Using the L1 distance function computes the distance between the following 2 customers: c1 = (55, good, 7000) and c2 = (25, poor, 1000). **[30 points]**
2. Using the L2 distance function computes the distance between the above mentioned 2 customers. **[30 points]**

**Problem-2 [30 points]**

The iris training and test datasets are available in the homework folder. Your task is to apply kNN algorithm on the training set to detect Iris species of the test dataset. Specifically, develop an R script to

1. Read-in the training set and test set
2. Apply kNN.
3. Predict the class labels for the test set.
4. Store the item IDs and class labels in a csv file. The file format is available in the homework folder.