ASSIGNMENT 5

**Name:** Parth Arora

**Roll No:** 2018IMT-063

**Course:** Machine Learning Lab

**Course Code:** ITIT-4107

**Aim: Given iris dataset with 3 classes and 4 features such as sepals/petals, Length, width etc. for each flower in the dataset. There are 50 instances per class in the dataset. Use Bayes Classifier as your base classifier model. Use 60% samples for training and 40% samples for testing.**

**1. Perform feature selection on this dataset using forward search.**

**2. As you select features, until 2 features, plot your right and incorrect classification instances for all classes.**

**3. For all the set of features selected, plot the accuracies to show the best subset of selected features**

Procedure:

1. We create an empty list to which we append the relevant features.

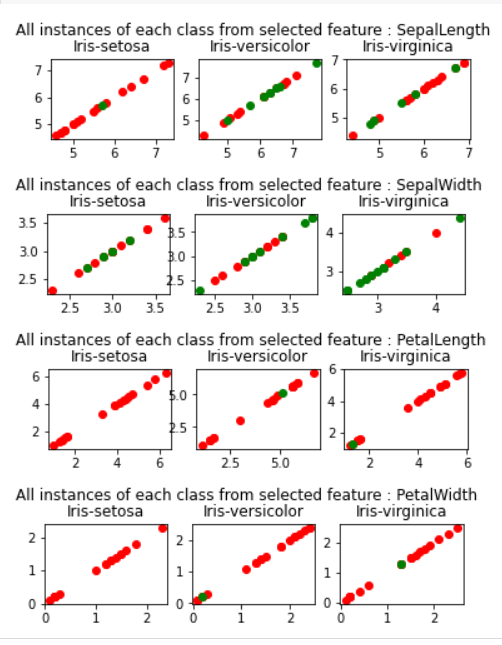
2. We start by selecting one feature and calculating the metric value for each feature on a cross-validation dataset. The feature offering the best metric value is selected and appended to the list of features.

3. The process is reiterated, this time with two features, one selected from the previous iteration and the other one selected from the set of all features not present in the set of already chosen features.

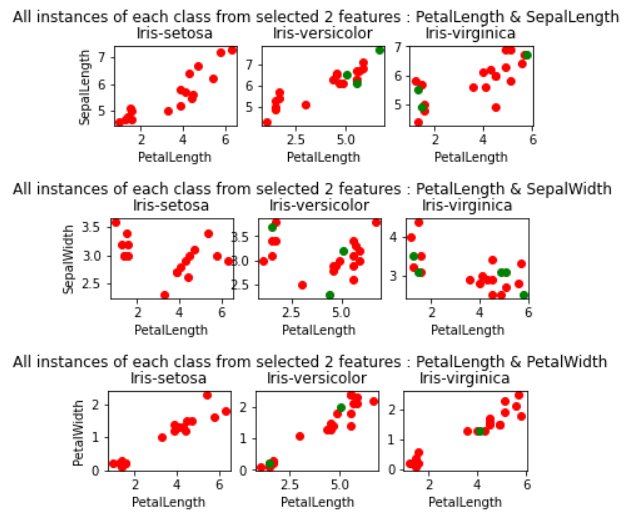
4. This process is repeated until we have the desired number of features.

Output:

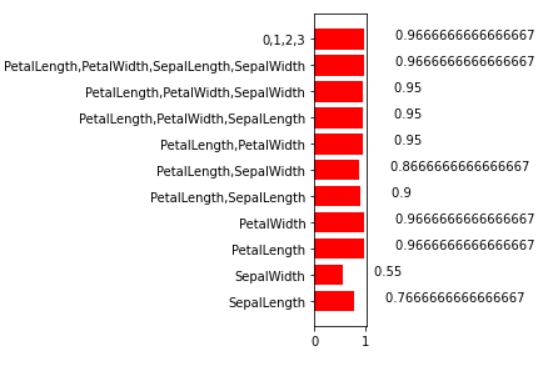
The correct and incorrect classifications for only 1 selected feature are shown below:



The correct and incorrect classifications for 2 selected features are shown below:



The accuracy plot for each subset of selected features is as follows:



Inference:

From the accuracy plot for all feature subsets, we can conclude that:

1. {Petal Length, Petal Width, Sepal Width} gives best accuracy among subsets with 3 features.

2. The dimensionality of the data set can be reduced by removing either Sepal Width or Sepal Length since the accuracy of the model remains the same even after their omission.

Code: <https://github.com/parth-arora/ITIT-4103-2021/blob/main/ML_Assignment_5%5B2018IMT_063%5D.ipynb>