

### **1. VARIABLE RANKING BY LINEAR CORRELATION COEFFICIENT**

Download the Leaf dataset from the UCI repository:

<http://archive.ics.uci.edu/ml/datasets/Leaf>

Rank the importance of the variables for classifying the species according with their linear correlation coefficient.

### **2. VARIABLE RANKING BY MUTUAL INFORMATION**

Download the Congressional Voting Records Data Set

<http://archive.ics.uci.edu/ml/datasets/Congressional+Voting+Records>

And rank the importance of the variables for classifying the party of vote according with the mutual information criteria.

### **3. PCA.**

Program a PCA and use it to analyze the dataset for radar returns from the ionosphere.

<http://archive.ics.uci.edu/ml/datasets/Ionosphere>

### **4. DISTANCES AND SIMILARITIES**

Program and compute the Mahalanobis distance for the leaf dataset using the coordinates selected in Point 1.

Program and compute the Euclidean distance in the PCA space between the entries in the ionosphere data set.

Program and compute the Hamming distance for the Voting records dataset