AGH University of Science and Technology

Ioannis Manousaridis

Project - 01

In the following image, it is presented the workflow implemented in the Orange Environment[1]
.

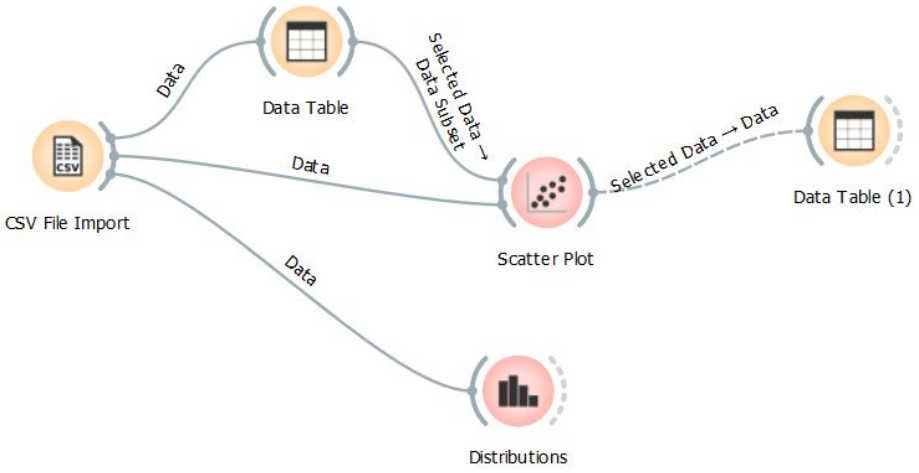
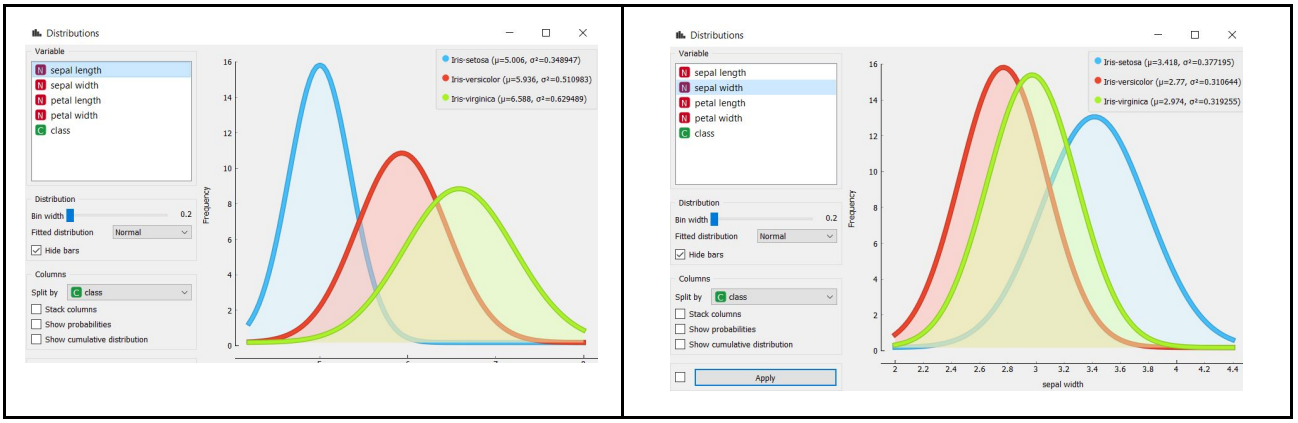


Figure 1. Workflow

For this work the Iris dataset was used. The Iris flower data set or Fisher's Iris data set is a multivariate data set. The data set consists of 50 samples from each of three species of Iris (Iris setosa, Iris virginica and Iris versicolor). Four features were measured from each sample: the length and the width of the sepals and petals, in centimeters. Based on the combination of these four features, Fisher developed a linear discriminant model to distinguish the species from each other.[2]

In the pictures in Figure 2, the distributions of the four features are presented. It is obvious that the best results can be taken from the petal width and petal length. This result can be confirmed also by figure 3.



[1] The workflow can be found here in ows format: https://drive.google.com/open?id=1ULXmFplZASgNlF-YDUlZu_A8ilNqOdg6
[2] The dataset can be found here in csv format: https://drive.google.com/open?id=1FVVPBHsinj84XsBdNTsm3hJZSY67N8cD

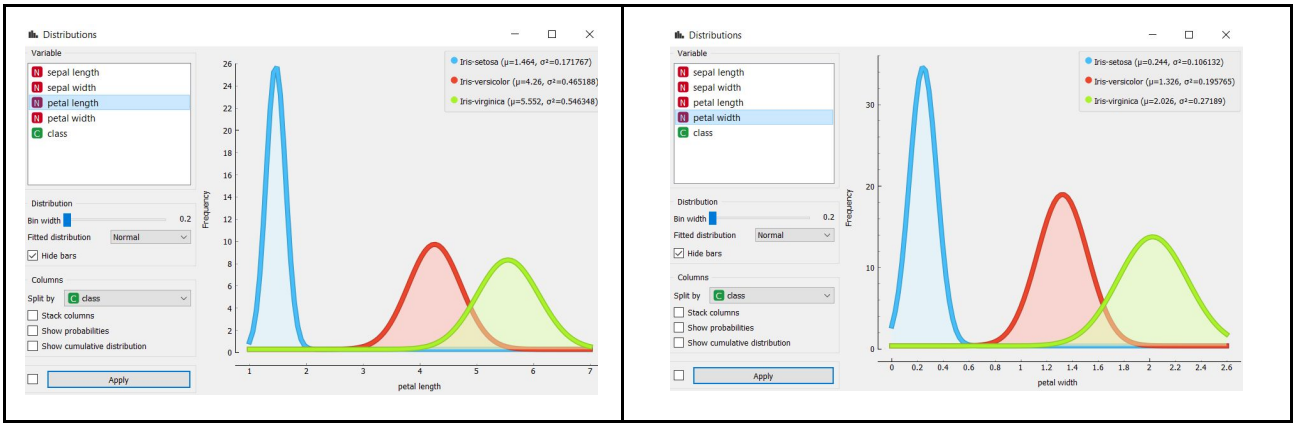


Figure 2. Distributions of the 4 features.

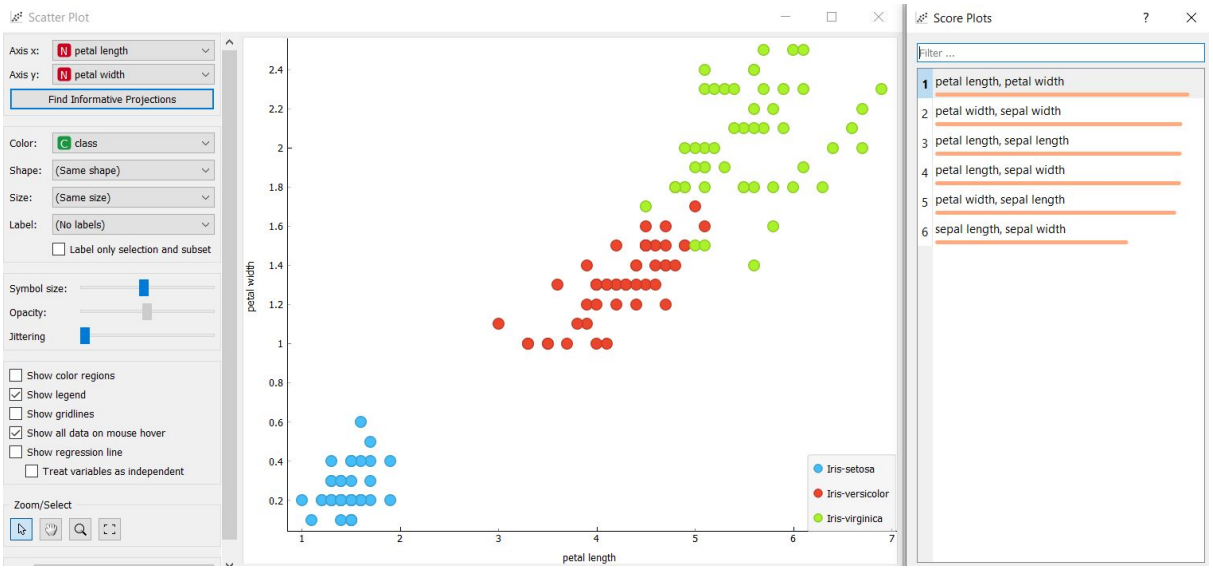


Figure 3. The result using petal width and length.