Fisher Iris Dataset is loaded and manipulated using Pandas, the data analysis python library.

We used matplotlib to draw the univariate and multivariate distributions.

Box and whisker plot helps us visualize the data by plotting the Minimum, first quartile, median, second quartile and maximum.

sklearn.model\_selection\_Kfold first shuffles the dataset using the random number generator whose sees is provided by us. Then splits it into n\_splits components and turn by turn trains on 9 of them and tests on the 1 component left.

So the split is:

Break the 150 observations into : 120(train+validate)+30(test) (80%,20%)

120 is further split as : 108(train)+12(validate) (90%,10%)

We train every model using the 108 observations and then validate each of them on the 12 observations of the validation set .

The algorithm giving maximum accuracy is selected and used for test data.