The dataset includes 150 observations of Iris flowers. These observations have been labeled in 3 classes as Setosa, Vericolor, and Virginica (50 observations in each class). For each observation, 4 features have been measured as sepal length. sepal width, petal length, and petal width. The objective is to use these observations for training a classifier.

Since each observation has 4 features, their visualization is not possible (4D feature space). However, an intuition can be gained by pairwise illustration of features. Sepal length versus sepal width, and petal length versus and petal width of the observations are to be depicted.

SEPAL LENGTH VS SEPAL WIDTH shows that if only sepal length and sepal width are employed, Setosa is expected to be linearly separated from the other two classes. However, these two features alone are not able to linearly separate Versicolor from Virginical.

PETAL LENGTH VS PETAL WIDTH shows that application of petal length and petal width will have a better result than the previous set, This case is expected to linearly separate Setosa from the other two classes and alms, separate Versicolor and Virginical linearly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistic | Sepal Length | Sepal Width | Petal Length | Petal Width |
| Minimum | 4.3000 | 2.0000 | 1.0000 | 0.1000 |
| Maximum | 7.9000 | 4.4000 | 6.9000  \_ | 2.5000 |
| Mean | 5.8433 | 3.0573 | 3.7580 | 1.1993 |
| Variance | 0.6857 | 0,1900 | 3.1163 | 0.5810 |
| Within-class Variance | 0.2650 | 0.1154 | 0.1852 | 0.0419 |
| Between-Class Variance | 0.4214 | 0.0756 | 2.9140 | 0.5361 |

Petal length and petal width have interesting statistics; they both have relatively low within-class variance and high between-class variance. If it is intended that 2 of 4 features to be applied for classification, this characteristic makes these two features favorable for separating the classes. This is resulted from the property of the mentioned statistics; a small within-class variance of a feature means that there is a small difference between the values of this feature for different members of the same class, while a large between-class variance states that there is a meaningful difference in the feature values for members of different classes.

correlation coefficients of features and class labels. Sepal width shows an interesting pattern: it has negative correlation with the other three features and the class label. However, the absolute values of the correlation coefficients are not big enough to draw a comprehensive conclusion, Further, petal width has a high correlation with petal width and class label. It means that the flowers with bigger petal width have bigger length and higher odds of belonging to classes with bigger labels (e.g. Virginica rather than Setosa). There is a similar pattern in the correlation between petal length and class label but with lower odds (due to the smaller correlation coefficient). Due to the high correlation of petal width and petal width, it may be decided to use only one of them for the classification.