Julian Marin

9/18/2018

Lab 9 Report

**3.1**

Sepal -length sepal-width petal-length petal-width class

0 5.1 3.5 1.4 0.2 Iris-setosa

1 4.9 3.0 1.4 0.2 Iris-setosa

2 4.7 3.2 1.3 0.2 Iris-setosa

3 4.6 3.1 1.5 0.2 Iris-setosa

4 5.0 3.6 1.4 0.2 Iris-setosa

5 5.4 3.9 1.7 0.4 Iris-setosa

6 4.6 3.4 1.4 0.3 Iris-setosa

7 5.0 3.4 1.5 0.2 Iris-setosa

8 4.4 2.9 1.4 0.2 Iris-setosa

9 4.9 3.1 1.5 0.1 Iris-setosa

10 5.4 3.7 1.5 0.2 Iris-setosa

11 4.8 3.4 1.6 0.2 Iris-setosa

12 4.8 3.0 1.4 0.1 Iris-setosa

13 4.3 3.0 1.1 0.1 Iris-setosa

14 5.8 4.0 1.2 0.2 Iris-setosa

15 5.7 4.4 1.5 0.4 Iris-setosa

16 5.4 3.9 1.3 0.4 Iris-setosa

17 5.1 3.5 1.4 0.3 Iris-setosa

18 5.7 3.8 1.7 0.3 Iris-setosa

19 5.1 3.8 1.5 0.3 Iris-setosa

**3.2** sepal-length sepal-width petal-length petal-width

count 150.000000 150.000000 150.000000 150.000000

mean 5.843333 3.054000 3.758667 1.198667

std 0.828066 0.433594 1.764420 0.763161

min 4.300000 2.000000 1.000000 0.100000

25% 5.100000 2.800000 1.600000 0.300000

50% 5.800000 3.000000 4.350000 1.300000

75% 6.400000 3.300000 5.100000 1.800000

max 7.900000 4.400000 6.900000 2.500000

**3.3**

**class**

**Iris-setosa 50**

**Iris-versicolor 50**

**Iris-virginica 50**

**4.1**

Box Plot or Whisker Plot: Shows the range of the values in the data set. The bottom of the box is the first quartile, the top is the 3rd quartile. The middle line is the median while the top and bottom of the graph are the minimum and maximum. The circles are the outliers with white circles representing suspected outliers and black circles representing outliers.

Histogram: A representation of the distribution of data. It is similar to a bar graph, however instead of using exact values it uses ranges.

**4.2**

Scatterplot: Plot that displays values for typically two variables in a set of data. Data is displayed as a collection of points with each having the value one variable on the x-axis and one variable on the y-axis.

**5.3**

KNN with an accuracy of 98.333 %.

Iris-setosa sample has an accuracy of 100%

**6.0**

The accuracy score is 0.9 which is 90%