Classification Tutorial- KNN Solution

Find the class label of the following test sample based on the training dataset

Sepal Length	Sepal Width	Species
5.2	3.1	?

Sepal Length	Sepal Width	Species
5.3	2.4	Setosa
5.1	3.8	Setosa
7.2	3.0	Virginica
5.4	3.4	Setosa
5.1	3.3	Setosa
5.4	3.9	Setosa
7.4	2.8	Virginica
6.1	2.8	Verscicolor
7.3	2.9	Virginica
6.0	2.7	Verscicolor
5.8	2.8	Virginica
6.3	2.3	Verscicolor
5.1	2.5	Verscicolor
6.3	2.5	Verscicolor
5.5	2.4	Verscicolor

a) First, we should find the distance of the test data from each training sample. We use Euclidean distance to calculate the distance.

For example, distance of test data from sample1 is equal to:

$$dist(testdata, sample1) = \sqrt{(5.2 - 5.3)^2 + (3.1 - 2.4)^2} = 0.707$$

The distance of test data from sample2 is equal to:

$$dist(testdata, sample2) = \sqrt{(5.2 - 5.1)^2 + (3.1 - 3.8)^2} = 0.707$$

We do the same for other samples. In the following table we have entered the calculated distance of test data from each training sample:

Sepal Length	Sepal Width	Species	Distance
5.3	2.4	Setosa	0.707
5.1	3.8	Setosa	0.707
7.2	3.0	Virginica	2.002
5.4	3.4	Setosa	0.36
5.1	3.3	Setosa	0.22
5.4	3.9	Setosa	0.82
7.4	2.8	Virginica	2.22
6.1	2.8	Verscicolor	0.94
7.3	2.9	Virginica	2.1
6.0	2.7	Verscicolor	0.89
5.8	2.8	Virginica	0.67
6.3	2.3	Verscicolor	1.36
5.1	2.5	Verscicolor	0.608
6.3	2.5	Verscicolor	1.25
5.5	2.4	Verscicolor	0.75

b) We select the three samples that have the smallest distance from the test data

Sepal Length	Sepal Width	Species	Distance
5.3	2.4	Setosa	0.707
5.1	3.8	Setosa	0.707
7.2	3.0	Virginica	2.002
5.4	3.4	Setosa	0.36
5.1	3.3	Setosa	0.22
5.4	3.9	Setosa	0.82
7.4	2.8	Virginica	2.22
6.1	2.8	Verscicolor	0.94
7.3	2.9	Virginica	2.1
6.0	2.7	Verscicolor	0.89
5.8	2.8	Virginica	0.67
6.3	2.3	Verscicolor	1.36
5.1	2.5	Verscicolor	0.608
6.3	2.5	Verscicolor	1.25
5.5	2.4	Verscicolor	0.75

c) We use majority voting for determining the class label of test data. As we can see, among 3 nearest neighbors, 2 of them belong to Setosa, so we assign the test data to **Setosa** class as well.