

Homework-2 (ML)

1. The 1-NN algorithm finds the single nearest neighbor to the test point in the training data. In this case, the test point $x=4.2$ is closest to the training point $x=4$. So, the classification of a test sample is B.
2. The 2-NN algorithm considers the two nearest neighbors. They are $x=4$ & $x=5$. When there's a tie between the labels, we use a majority voting system. When counting the occurrences of each label, B comes first in the array. So, the answer is B.
3. The 3-NN algorithm considers three nearest neighbors, they are $x=4$ (B), $x=5$ (A), $x=3$ (A), here also we go by voting majority. So, they are 2 A's and one B. Now, the classification of test sample is A.

4. Total Errors

Misclassified points = $x=4, x=5, x=9, x=13$

An total we have 3 forced errors and 5 natural errors = $3+5=8$.

We have 18 total samples. So the answer is $8/18$.

5. We can identify 8 misclassifications. They are $x=3, x=4, x=5, x=8, x=9, x=12, x=13, x=14$. So they are 8 misclassifications out of 18 samples.

So, the answer is = $8/18$.

6. I identified 4 misclassifications. They are $x=4, x=8, x=9, x=13$.

So, the answer is = $4/18$.

7. We have to find the 4-NN for each point & make predictions. After going through each point, we have 4 misclassifications. They are $x=4, x=8, x=9$ and $x=13$.

So, the answer is $4/18$.

8. We have 18 total samples and we are using 17-NN, this means we will be considering all other points as neighbors for each test point. Comparing the predictions to the true labels, we can see that every single prediction is incorrect. So, the answer is $18/18$.