[6+2 = 8 Marks]

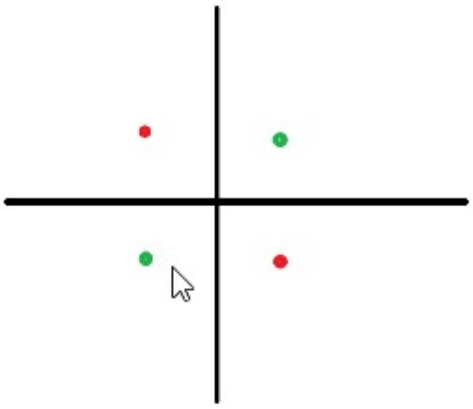
Consider the following dataset:

X1	X2	Class
1	1	Positive
-1	1	Negative
-1	-1	Positive
1	-1	Negative

a) Following classifiers are used to classify the given dataset. For each of the classifier state whether it can perfectly classify the given dataset or not? Justify the answer."

1. Logistic Regression
2. Decision Tree
3. 1 Nearest neighbour

Plotting those points on a graph we get:



Here the green samples are of the positive class and the red samples are of the negative class.

Now let's analyse which algorithm will be able to classify these points.

First let us look at logistic regression. We know that logistic regression has a linear decision boundary, but it is evident from the above picture that no straight line will be able to seperate the positive and negative samples. So logistic regression won't work.

Now we will analyse 1-Nearest Neighbor. To analyse this, let's first find the distances between the samples:

We take the example of the point at (-1, -1) which is of positive class. It's closest neighbors are the two negative samples which are both at a distance of 2. The other positive sample is at a distance of $\sqrt{5}$. So the positive sample is farther than the negative samples. Hence 1-Nearest Neighbor will give wrong prediction.

Now, Decision Tree can construct a non-linear decision boundary. If we take the black lines above as the decision boundary of the decision tree, we can see that it can easily classify the points correctly.

So **Decision Tree can perfectly classify the given dataset.**