Pattern Recognition - Coursework 3

Sadia Rana, King's ID: 1773869

April 16, 2018

1 Question 1

Class 1: y=-1	Class 2: y=+1
10,2	-7.5,-2.5
-8.5,3	-6,-1.5
-6,6.5;	-6,2.5
-4.5,6;	-5,-4.5
-1.5,4.5	-4.5,1
-1,8	-3.5,2.5
1,2	-3,-3
1.5,3.5	-2,1.5
1.5,6	-1,-3.5
2.5,4.5	0,0.5
3.5,8.5	2.5,-3
3,2	3,-5.5
4.5,5.5	4.5,-1
6,2.5	4.5,0.5
6.5,5.5	6,-4
7,8.5	8.5,-2
9,3	9.5,1

2 Question 2

Plotting the data from the above dataset the following graph is obtained.

The data is linearly non-seperable, since a clear line cannot be drawn between the two data sets to effectively and clearly separate all the datapoints for the two classes.

3 Question 3

Using the data point (6,2.5), we will in detail analyse how the different results are determined in this algorithm.

For the first step a random number of data points with a random selection from the dataset are chosen. As an example, for the first weak classifier 23 data points were randomly chosen. Using these datapoint a hyperplane was designed of the form:

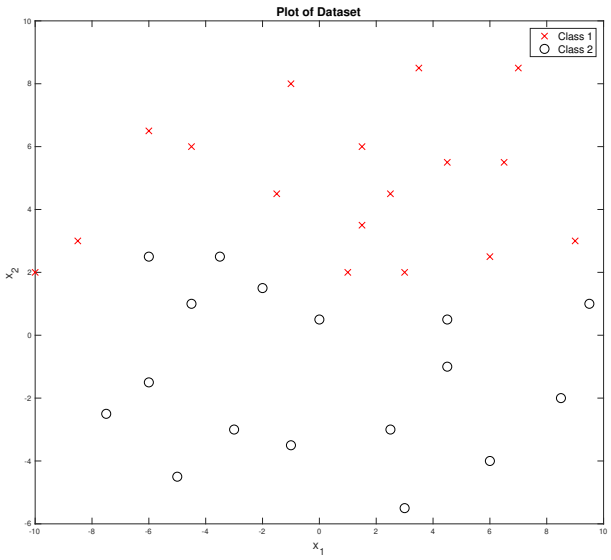

Figure 1: Data for Class 1 and 2.

$$f = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + w_0$$

to maximise the margin between the two classes. Using the calculated hyperplane, all values in the whole dataset were assigned a class as detrmined by the hyperplane. The values obtained from the above equation when applied to all datapoints were retained for use in the last step.

Assigning a class for each weak classfier was done according to result of the hyperplane, such that:

$$f \begin{cases} =0 & \text{on hyperplane} \\ <0 & \text{Class 1} \\ >0 & \text{Class 2} \end{cases} .$$

In terms of this assignment method, the following is how the Classes were chosen for the point (6,2.5) using 5 weak classifers:

	Results	Assigned value	Assigned Class
Weak Classifier 1	-0.2973	-1	Class 1
Weak Classifier 2	1.0000	+1	Class 2
Weak Classifier 3	-1.1111	-1	Class 1
Weak Classifier 4	-1.8913	-1	Class 1
Weak Classifier 5	-0.6296	-1	Class 1

The values for the overall clssification were chosen by using the values from the hyperplane results and the formula:

$$f(x) = sgn(\sum_{i=1}^{M} \frac{1}{M}(f_i(x)))$$

Using the above example one obtains the value

$$\frac{-0.2973 + 1.0000 - 1.1111 - 1.8913 - 0.6296}{5} = -0.4599$$

This gives a negative result and hence $f = -1$ and hence assigns the point (6,2.5) to class 1.

In terms of classification performance, single weak classifiers have a wide range of accuracies, as can be seen above when comparing the accuracies for all the weak classifiers. They range from 65% at the lowest value to 91%. Avaraging over the results of the weak classifiers can help to smooth out these inaccuracies and give a relatively high final accuracy, despite having a much lower one within any of the classifiers. The more weak classifiers are used, the less likely it is that a very low accuracy in one of them has an overall effect on the final result. Overall, the final accuracy tends to be higher then most, if not all individual accuracies. As can be seen that, with exception of when 4 weak clasiffiers are used, all final accuracies either equal to the highest individual accuracy or higher than all accuracies.

The highest accuracy obtained is when using only 3 weak classifiers. Whether that is just a singular occurance or has a deeper meaning could be determined with further, repeated application of the algorithm and with a higher range of weak classifiers. Though logically, and with the knowledge from the lectures, it stands to reason that a higher number of weak classifiers would, on avarage, be more accurate than a lower one.

4 Question 4

The weak classifications for the above method were reached in the same way as those in Question 3, yet the method for choosing datapoints for consecutive weak classification and the overall classification differs.

The first weak classification requires more than 50% accuracy. The second weak classification process requires that roughly at least half of the datapoints chosen should be correctly classified by the first weak classfier and the rest should have been wrongly classified. In the last step, those points are chosen that are wrongly classified by both the first and the second classifier.

To study the overall classification,once again the point (6,2.5) is chosen. we can see that this point is not well classified if only determined by the first two classifiers. The first one determined it to be in class 1 and the second assigns it to class 2. Hence the third classifier has to be used to determine the overall classification. Since this agrees with the first result, i.e. that it is to be assigned to class 1, it is overall classification is class 1.

The argument for classification accuracy as discussed above applies.

Data	Weak Classifier 1	Weak Classifier 2	Weak Classifier 3	Overall classifier
-10,2	+1	-1	-1	-1
-8.5,3	+1	-1	-1	-1
-6,6.5;	-1	-1	-1	-1
-4.5,6;	-1	-1	-1	-1
-1.5,4.5	-1	-1	-1	-1
-1,8	-1	-1	-1	-1
1,2	+1	+1	-1	-1
1.5,3.5	-1	-1	-1	-1
1.5,6	-1	-1	-1	-1
2.5,4.5	-1	-1	-1	-1
3.5,8.5	-1	-1	-1	-1
3,2	0	+1	-1	-1
4.5,5.5	-1	-1	-1	-1
6,2.5	-1	+1	-1	-1
6.5,5.5	-1	-1	-1	-1
7,8.5	-1	-1	-1	-1
9,3	-1	-1	-1	-1
-7.5,-2.5	+1	+1	+1	+1
-6,-1.5	+1	+1	+1	+1
-6,2.5	+1	-1	-1	-1
-5,-4.5	+1	+1	+1	+1
-4.5,1	+1	+1	-1	+1
-3.5,2.5	+1	-1	-1	+1
-3,-3	+1	+1	+1	-1
-2,1.5	+1	+1	-1	+1
-1,-3.5	+1	+1	+1	+1
0,0.5	+1	+1	+1	+1
2.5,-3	+1	+1	+1	+1
3,-5.5	+1	+1	+1	+1
4.5,-1	+1	+1	+1	+1
4.5,0.5	+1	+1	+1	+1
6,-4	+1	+1	+1	+1
8.5,-2	+1	+1	+1	+1
9.5,1	-1	+1	+1	+1
Accuracy %	73%	73%	88%	94%

Table 2: Bagging method with 3 weak classifiers

Data	Weak Classifier 1	Weak Classifier 2	Weak Classifier 3	Weak Classifier 4	Overall classifier
-8.5,3	+1	+1	+1	+1	+1
-8.5,3	+1	+1	+1	+1	-1
-6,6.5;	+1	-1	-1	-1	-1
-4.5,6;	-1	-1	-1	-1	-1
-1.5,4.5	-1	-1	-1	-1	-1
-1,8	-1	-1	-1	-1	-1
-8.5,3	-1	-1	+1	-1	-1
1.5,3.5	-1	-1	-1	-1	-1
1.5,6	-1	-1	-1	-1	-1
2.5,4.5	-1	-1	-1	-1	-1
3.5,8.5	-1	-1	-1	-1	-1
3,2	-1	-1	0	-1	-1
4.5,5.5	-1	-1	-1	-1	-1
6,2.5	-1	-1	-1	-1	-1
6.5,5.5	-1	-1	-1	-1	-1
7,8.5	-1	-1	-1	-1	-1
9,3	-1	-1	-1	-1	-1
-7.5,-2.5	+1	+1	+1	+1	+1
-6,-1.5	+1	+1	+1	+1	+1
-6,2.5	+1	+1	+1	+1	+1
-5,-4.5	+1	+1	+1	+1	+1
-4.5,1	+1	+1	+1	+1	+1
-3.5,2.5	+1	+1	+1	+1	+1
-3,-3	+1	+1	+1	+1	+1
-2,1.5	-1	+1	+1	+1	+1
-1,-3.5	+1	+1	+1	+1	+1
-8.5,3	-1	+1	+1	+1	+1
2.5,-3	-1	+1	+1	+1	+1
3,-5.5	-1	+1	+1	+1	+1
4.5,-1	-1	+1	+1	+1	+1
4.5,0.5	-1	+1	+1	-1	-1
6,-4	-1	+1	+1	+1	+1
-8.5,3	-1	+1	+1	+1	+1
9.5,1	-1	-1	-1	-1	-1
Accuracy %	65%	91%	85%	88%	88%

Table 4: Bagging method with 4 weak classifiers

Data	Weak Classifier 1	Weak Classifier 2	Weak Classifier 3	Weak Classifier 4	Weak Classifier 5	Overall classifier
-8.5,3	+1	-1	+1	+1	+1	+1
-8.5,3	+1	-1	+1	+1	+1	+1
-6,6.5;	-1	-1	-1	-1	-1	-1
-4.5,6;	-1	-1	-1	-1	-1	-1
-1.5,4.5	-1	-1	-1	-1	-1	-1
-1,8	-1	-1	-1	-1	-1	-1
-8.5,3	+1	+1	+1	-1	+1	+1
1.5,3.5	-1	-1	-1	-1	-1	-1
1.5,6	-1	-1	-1	-1	-1	-1
2.5,4.5	-1	-1	-1	-1	-1	-1
3.5,8.5	-1	-1	-1	-1	-1	-1
3,2	+1	+1	0	-1	+1	+1
4.5,5.5	-1	-1	-1	-1	-1	-1
6,2.5	-1	+1	-1	-1	-1	-1
6.5,5.5	-1	-1	-1	-1	-1	-1
7,8.5	-1	-1	-1	-1	-1	-1
9,3	-1	+1	-1	-1	-1	-1
-7.5,-2.5	+1	+1	+1	+1	+1	+1
-6,-1.5	+1	+1	+1	+1	+1	+1
-6,2.5	+1	-1	+1	+1	+1	+1
-5,-4.5	+1	+1	+1	+1	+1	+1
-4.5,1	+1	+1	+1	+1	+1	+1
-3.5,2.5	+1	+1	+1	+1	+1	+1
-3,-3	+1	+1	+1	+1	+1	+1
-2,1.5	+1	+1	+1	+1	+1	+1
-1,-3.5	+1	+1	+1	+1	+1	+1
-8.5,3	+1	+1	+1	+1	+1	+1
2.5,-3	+1	+1	+1	+1	+1	+1
3,-5.5	+1	+1	+1	+1	+1	+1
4.5,-1	+1	+1	+1	+1	+1	+1
4.5,0.5	+1	+1	+1	-1	+1	+1
6,-4	+1	+1	+1	+1	+1	+1
-8.5,3	+1	+1	+1	+1	+1	+1
9.5,1	+1	+1	-1	-1	-1	+1
Accuracy %	88%	85%	85%	88%	85%	88%

Table 6: Bagging method with 5 weak classifiers

Data	Weak Classifier 1	Weak Classifier 2	Weak Classifier 3	Overall classifier
-10,2	+1	-1	+1	+1
-8.5,3	+1	-1	-1	-1
-6,6.5;	-1	-1	-1	-1
-4.5,6;	-1	-1	-1	-1
-1.5,4.5	-1	-1	-1	-1
-1,8	-1	-1	-1	-1
1,2	+1	+1	+1	+1
1.5,3.5	-1	+1	-1	-1
1.5,6	-1	-1	-1	-1
2.5,4.5	-1	+1	-1	-1
3.5,8.5	-1	-1	-1	-1
3,2	-1	+1	-1	-1
4.5,5.5	-1	+1	-1	-1
6,2.5	-1	+1	-1	-1
6.5,5.5	-1	+1	-1	-1
7,8.5	-1	-1	-1	-1
9,3	-1	+1	-1	-1
-7.5,-2.5	+1	+1	+1	+1
-6,-1.5	+1	+1	+1	+1
-6,2.5	+1	-1	+1	+1
-5,-4.5	+1	+1	+1	+1
-4.5,1	+1	+1	+1	+1
-3.5,2.5	+1	+1	-1	+1
-3,-3	+1	+1	+1	+1
-2,1.5	+1	+1	+1	+1
-1,-3.5	+1	+1	+1	+1
0,0.5	+1	+1	+1	+1
2.5,-3	+1	+1	+1	+1
3,-5.5	+1	+1	+1	+1
4.5,-1	+1	+1	+1	+1
4.5,0.5	+1	+1	+1	+1
6,-4	+1	+1	+1	+1
8.5,-2	+1	+1	+1	+1
9.5,1	-1	+1	+1	+1
Accuracy %	88%	73%	91%	94%

Table 8: Boosting method with 3 weak classifiers