Chiranjeev
P21CS007
chiranjeev.1@iitj.ac.in

Assignment - 4
Adaboost

04-10-2021

I. Dataset Description:

a. Name of Dataset : Randomly Generated {+1,-1} labelled 100 points

b. Dataset features and Target variable :

- x1
- x2
- y (Target Variable)

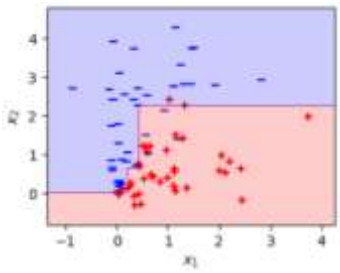
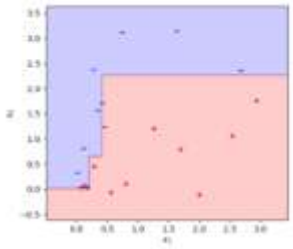
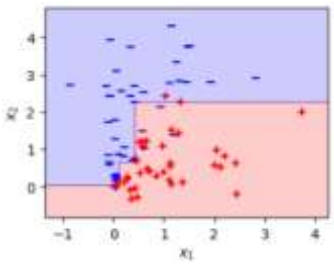
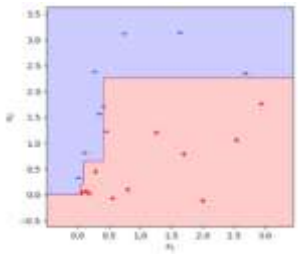
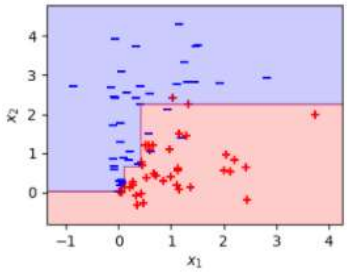
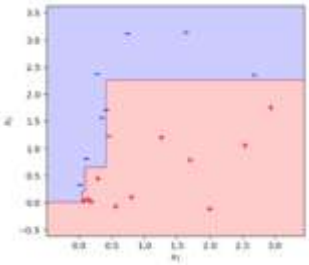
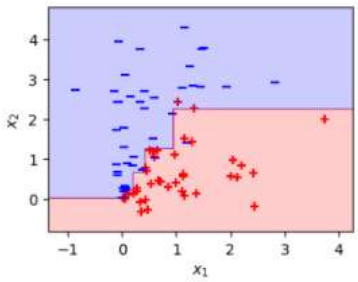
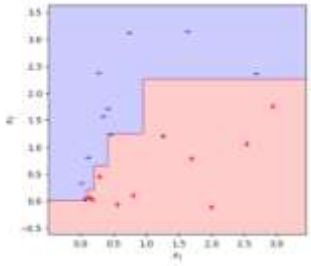
c. Target Value to be Predicted :
Class: -1(- in graph representation)
Class: +1 (+ in graph representation)

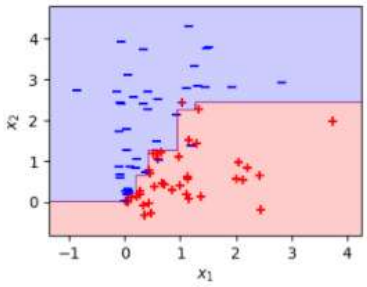
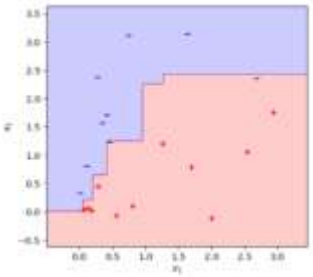
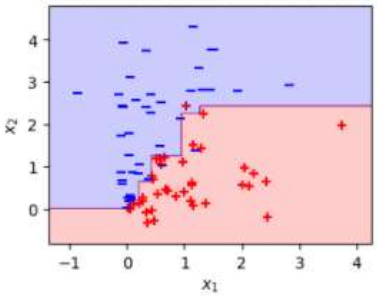
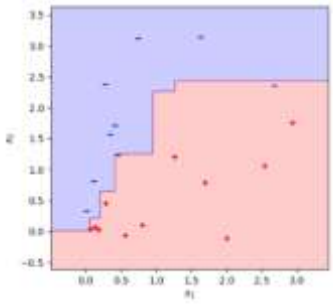
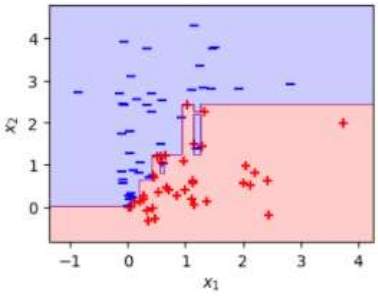
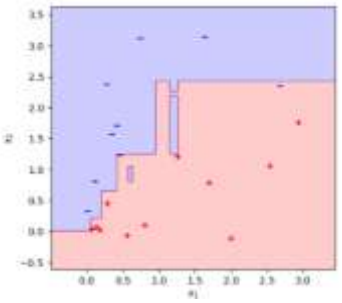
II. Splitting the Dataset:

Used train_test_split of sklearn to split the dataset into train and test.

Split the Dataset into: 80% - train set, 20% test set

III. Results

Number Of Stages (T)	Train Accuracy	Test Accuracy	Confusion Matrix on Test Set	Decision Boundary with Output Hypothesis of Adaboost on Train set	Decision Boundary with Output Hypothesis of Adaboost on Test set
5	**91.25 %**	**75.00 %**	confusion_matrix: [[7 2] [3 8]]		
10	**93.75 %**	**90.00 %**	confusion_matrix: [[7 2] [0 11]]		
15	**93.75 %**	**90.00%**	confusion_matrix: [[7 2] [0 11]]		
20	**95.00 %**	**90.00 %**	confusion_matrix: [[8 1] [1 10]]		

25	**96.25 %**	**90.00 %**	confusion_matrix: [[7 2] [0 11]]		
30	**96.25 %**	**90.00 %**	confusion_matrix: [[7 2] [0 11]]		
80	**100.00%**	**90.00 %**	confusion_matrix: [[7 2] [0 11]]		

Conclusion:

1. Experiment was conducted for Adaboost with weak classifiers by using number ofstages (T) and experiment is done by taking different number of stages.

2. Decision Tree was used to decide the stumps.

3. On visualizing and analyzing above recorded observations as the number of stages increases the Training Accuracy increases and follows ERM rule.