

Adaboost Boosting

Weak Learners

Q What are Weak Learners?

Ans Weak Learners are those type of Algorithms whose Accuracy rate is very low for example \Rightarrow The Algorithms whose Accuracy rate is just over 50% only.

Q Why to use the Weak Learners in Adaboosting?

Ans Weak Learners are important in Adaboosting because these Weak Learner Algorithms are combined and the resultant Algorithm

Spiral

Date.....

or the Model is too much powerful. Cont.

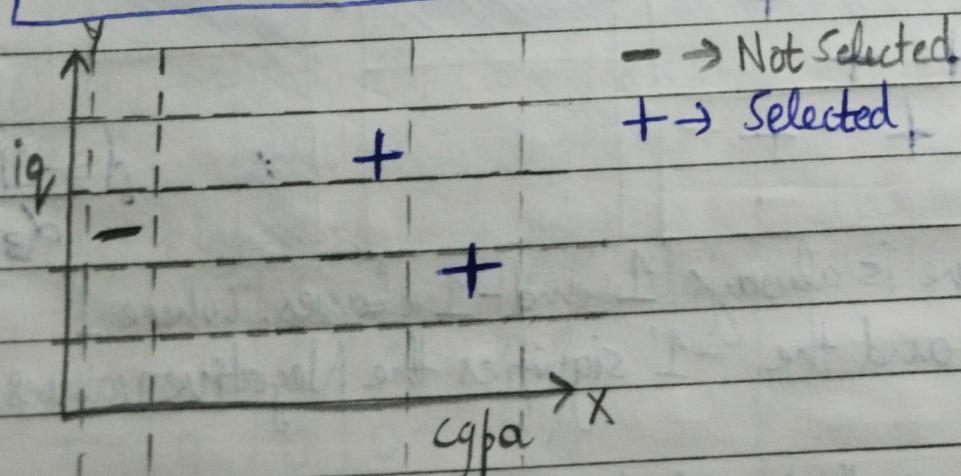
Decision Stumps

Q What is Decision Stumps?

Ans Decision Stumps is one of the type of Weak Learner in which this Algorithm only splits the Complete data for only 1 time.

This Shows that, this Algorithm is also a Weak Learner.....

Mostly, the AdaBoost uses the Decision Stumps type of Weak Learner...



For example..

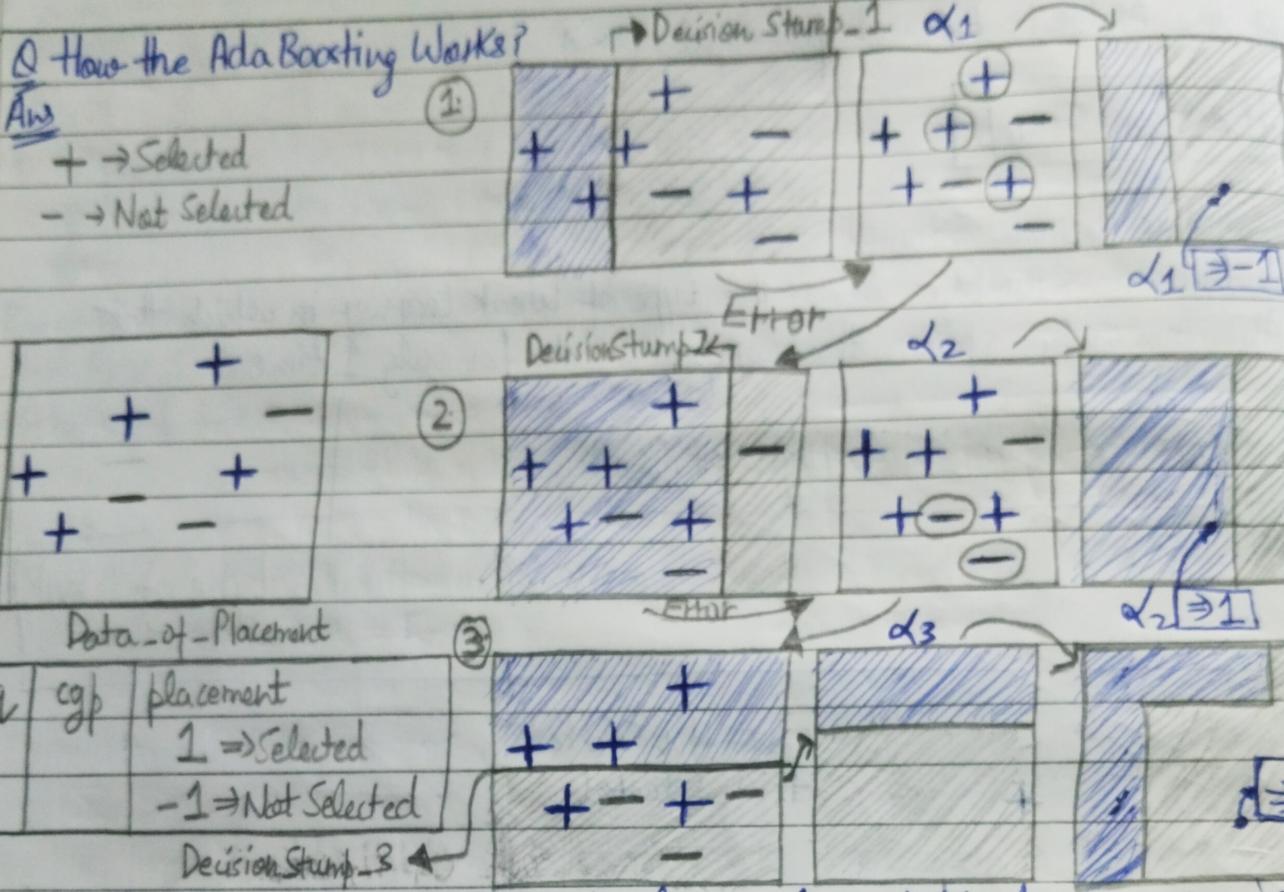
Date.....

Q How the Ada Boosting Works?

Ans

+ → Selected

- → Not Selected



Note:- ① In Ada Boosting there is always 1 and -1 classes. Where '1' signifies the Positive values and the '-1' signifies the Negative values.

② The Decision Stump is always that applied, whose Entropy is Minimum.
For example:- Decision Stump-1, Decision Stump-2, Decision Stump-3.

i) Now these α_1, α_2 and α_3 are the Weights provided to these Models as per their Accuracy.

ii) These Errors let's say in ① Case the Error is transferred to ② Case in which it tells that the 3 Marked points are misclassified. So, in another case, keep focus on these points. This focusing is done through up Sampling.

Date.....

(iii) Finally to calculate the Ada Boosting the Formula is:-

$$= \text{Sign} (\alpha_1 * h_1(x) + \alpha_2 * h_2(x) + \alpha_3 * h_3(x))$$

For example:- let's say $\alpha_1 = 2, \alpha_2 = 10, \alpha_3 = 3$

Input $\Rightarrow Y \{9.3, 8.0\}$ \Rightarrow Predict Select or Not Selected...
(gpa, iq)

As, per the formula:-

$$= (2 * -1 + 10 * 1 + 3 * -1)$$

$$\Rightarrow (-2 + 10 - 3)$$

$\Rightarrow 5 \Rightarrow$ i.e, the Input Candidate is Selected....

Q How to Actually Calculate all the things like x_1 , Up Sampling and
Finally the Ada Boost Formula...

Sol: Step by Step Explanation "Ada Boost Classification"

x_1	x_2	y	Weights	$y_{predicted}$
3	7	1	0.2	1
2	9	0	0.2	0
1	4	1	0.2	0
9	8	0	0.2	0
3	7	0	0.2	0

Step-1 \Rightarrow Initially, assign a Weight to all the rows.

Sol \Rightarrow To Assign a common Weight $\Rightarrow 1/n$ where $n = \text{number of rows}$

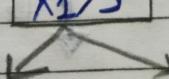
$$\Rightarrow 1/5 \Rightarrow [0.2]$$

Stage 1

1> Apply a Decision Tree having a $\text{max_depth} = 1$.

Model-1

$x_1 > 5$



On this model 1, we calculate the $y_{predicted}$ by providing the x_1 and x_2 Values... Spiral

model_1

Q How to calculate this α_1 ?

$$\text{Ans} \quad \alpha = \frac{1}{2} \log \left(\frac{1 - \text{error}}{\text{error}} \right)$$

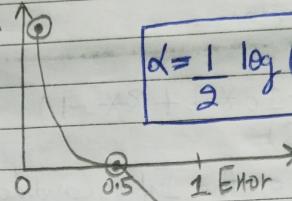
Error rate

$$[a] \rightarrow \text{Error Rate} = 0\% \approx 0$$

$$[b] \rightarrow \text{Error Rate} = 100\% \approx 1$$

$$[c] \rightarrow \text{Error Rate} = 50\% \approx 0.5$$

$$\alpha = \frac{1}{2} \log \left(\frac{1 - \text{error}}{\text{error}} \right)$$



Q How to calculate this "error"?

$$\text{Ans} \quad \sum \text{Misclassified Weights}$$

X_1	X_2	Y	Y-Predicted	Weight
3	7	1	1	0.2
2	9	0	1	0.2
1	4	1	0	0.2
9	8	0	0	0.2
3	7	0	0	0.2

model_1

$$\text{error} = \sum \text{Misclassified Weights}$$

$$\Rightarrow 0.2 + 0.2 = 0.4$$

$$\text{model-1} \Rightarrow \alpha_1 \Rightarrow \frac{1}{2} \log \left(\frac{1 - \text{error}}{\text{error}} \right) \Rightarrow \frac{1}{2} \log \left(\frac{1 - 0.4}{0.4} \right)$$

$$\alpha_1 \Rightarrow 0.20$$

Step-2:-

Stage-1 UpSampling Boosting the Weights.

For Misclassified Points

Increase the Weights

Right Classified Points

Decrease the Weights

Date.....

Q How much weight is to be Increased or Decreased?

Ans For Misclassified Points.. For example:-

$$\text{New-Weight} = \text{Current-Weight} * e^{-\alpha_1}$$

$$\text{New-Weight} = 0.2 * e^{0.20} \\ \Rightarrow [0.24]$$

For Classified Points..

$$\text{New-Weight} = \text{Current-Weight} * e^{-\alpha_1}$$

$$\text{New-Weight} = 0.2 * e^{-0.20} \\ \Rightarrow [0.16]$$

x_1	x_2	y	$y_{\text{Predicted}}$	Weights	Updated Weights	Normalized Weights
3	7	1	1	0.2	0.16	0.166
2	9	0	1	0.2	0.24	0.250
1	4	1	0	0.2	0.24	0.250
9	8	0	0	0.2	0.16	0.166
3	7	0	0	0.2	0.16	0.166
					0.96	1

x_1	x_2	y	Updated-Weights	Range	"5 Random Numbers between 0 to 1"
3	7	1	0.16	0-0.16	① 0.13
2	9	0	0.24	0.16-0.40	② 0.38
1	4	1	0.24	0.40-0.64	③ 0.49
9	8	0	0.16	0.64-0.80	④ 0.62
3	7	0	0.16	0.80-0.96	⑤ 0.98

"1, 3, 3, 5, 5" these are the Rows in which these Random Numbers are falling....

Now the Complete Data Set is Changed and this New data Set includes the Rows as "1, 3, 3, 5, 5". For example:-

Stage-2

Date.....

X_1	X_2	Y	Updated-Weights	Range	Again all the steps will be followed :-
3	7	1	0.16	0-0.16	(1) First Decision Stump will be designed.
1	4	1	0.24	0.40-0.64	(2) Then, we calculate the α_2 .
1	4	1	0.24	0.40-0.64	(3) Calculate the error.
3	7	0	0.16	0.80-0.96	(4) Then, Up Sampling will be there...
3	7	0	0.16	0.80-0.96	