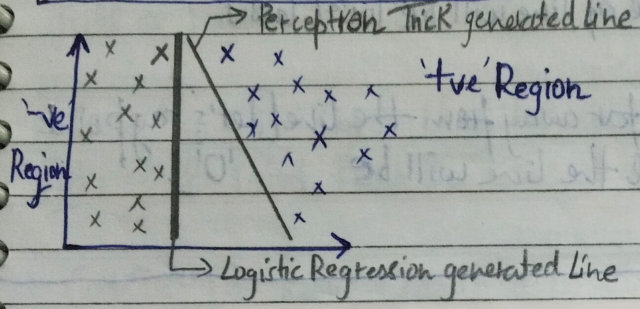
Date.....

Logistic Expression

Firstly, I will talk about that, what is the problem comming in the Preceptron Trick so that, we moved to Logistic Regression.

Results as per Perceptron Trick



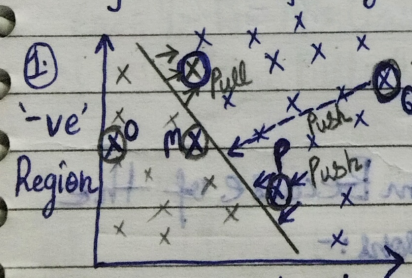
Scale:-

x ⇒ Selected Candidates
x ⇒ Not Selected Candidates

We can analyse that the Perceptron Trick had generated the line which is more closer to the points that are selected Candidates. But, when the new data will come it will move more closly to the Positive region.

So, Applying Logistic Regression:-

Algorithm of Logistic Regression



Scale:-

x ⇒ Selected Candidates
x ⇒ Not Selected Candidates

(1) How or What changes we have to make to make Perceptron line to Logistic Regression line. $\because (y - \hat{y}) + +$

Step 1:- Misclassified points ⇒ Line ⇒ Pull
Right Classified points ⇒ Line ⇒ Push

Step 2:- The Magnitude also matters in Push and Pull of points line.

Case(a) Classified [Right Classified] Points.

↳ The right classified point is too much closer to the line. [Let's say Point P]
∵ So, the Magnitude will be high of Pushing the line.

Spiral

Date.....

Case (b) The right classified point is far away from the line [Let's say point 'O']
∴ So in that case, Magnitude of pushing the line will be low.

Wrong Classified Points

Case (a) When the mis classified point is more closer to the line. [let's say point 'M'].

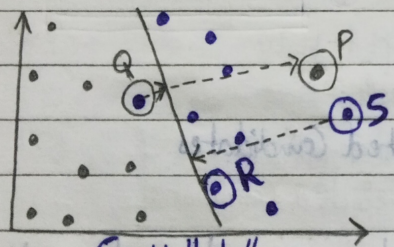
∴ So in that case the Magnitude of Pull of line will be low [less].

Case (b) When the mis classified point is far away from the line [let's say point 'O'].
∴ So, in that case the Magnitude of Pull the line will be High [More].

Sigmoid Function

Sigmoid Algorithm

Scale:-
• → Selected
• → Not Selected



Graph "abc"

In Sigmoid Function we are changing the Algorithm Because of the following reasons:-

$$W_{new} = W_{old} + \eta (Y_i - \hat{Y}_i) \cdot X_i$$

(i) In Perception Trick we only talk about the Mis classified points like point 'P' and 'Q', and move the line accordingly so that they can properly classified.

(ii) But, in this Algorithm include both Classified and Mis classified points and the concept is used as: ~~Right~~ Classified Points → Push the Line.
Mis Classified Points → Pull the Line.
[P, Q]

Spiral

Date.....

Q What strategy is used in Sigmoid Function?

Ans According to the Graph "abc", the strategy of Sigmoid Function will be:-

(i) In this Algorithm we include both the points Classified [Right] [RS] points and Misclassified Points [P,Q].

(ii) The Classified Points [R, S] ⇒ Push the line.

(iii) The Misclassified Points [P,Q] ⇒ Pull the line.

(iv) One more important things, "Magnitude" in which Magnitude they Pull or Push the line. This includes 4 cases:-

Case 1:- For Right Classified which is nearer to the line. i.e; Point 'R'.

Sol: ⇒ 'R' Point Push the line at a higher Magnitude.

Case 2:- For Right Classified which is far away to the line. i.e; Point 'S'.

sol: ⇒ 'S" Point Push the line at a lower Magnitude.

Case 3:- For Mis Classified which is nearer to the line, i.e; Point 'Q'.

sol: ⇒ 'Q' Point Pulls the line at a lower Magnitude.

Case 4:- For Mis Classified which is far away to the line. i.e; Point 'P'.

sol: ⇒ 'P' Point Pulls the line at a higher Magnitude.

Q Working to the Algorithm:-

sol: $W_{new} = W_{old} + \eta (y_i - \hat{y}_i) X_i$

According to Case 1 and Case 2:-

$W_{new} = W_{old} + \eta (y - \hat{y}) X_i$

$W_{new} = W_{old} + (0)$

⇒ No Change in the line.

	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$
Case 1:	1	1	0
Case 2:	0	0	0
Case 3:	1	0	1
Case 4:	0	1	-1

According to Case 3:-

$W_{new} = W_{old} + \eta X_i$ { Becoz $(y - \hat{y}) = 1$ }

According to Case 4:- $W_{new} = W_{old} - \eta X_i$ { Becoz, $(y - \hat{y}) = -1$ }

Spiral

Date.....

So can see that in Case 1 and Case 2, there is no change in the line. But we want to shift the line even the points are rightly classified.

Q Why this line is not shifting [in Case 1 and Case 2]?

Sol: Because, $(Y-\hat{Y})=0$ it becomes zero, $\Rightarrow$ $W_{new} = W_{old}$

∴ To stop this, firstly stop "$Y-\hat{Y}\neq 0$" i.e, we want that $Y-\hat{Y}\neq 0$ should not equal to Zero. (Y → Actual data, $\hat{Y}$ → Predicted Value.)

· We can make changes only in the Predicted Value [$\hat{Y}$]. Assume · how to do it?

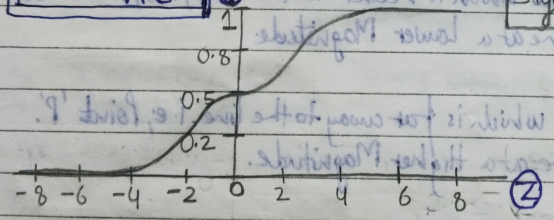
Sol: Firstly, we calculate the $\hat{Y} \Rightarrow \sum W_i X_i = [0,1]$ → Step Function

To overcome this, we use the Sigmoid Function.

Q What is Sigmoid Function?

Ans $\sigma(z) = \frac{1}{1+e^{-z}}$ (Y)

Sigmoid Graph



According to the Graph :-

'Z' lies between $\Rightarrow [-\infty, \infty]$

i.e; $Z \Rightarrow [-\infty, Z, \infty]$

It means

Y lies between [0,1]

and when Z=0

Y = 0.5

It means at a Infinite Negative Value, z's Value will be zero.

It means, at a Infinite Positive Value, z's Value becomes the positive Zero.

So, this function solves our program, that $(Y-\hat{Y}) \neq 0$ because it provides the Smallest Value also which is between 0 → 1.

Spiral

{7·5, 81}

Date.....

$\hat{Y}_i = \Sigma W_i X_i$

So, in Perception Trick, we generally calculate the $\hat{Y}_i$ as:- $W_1 X 7{\cdot}5 + W_2 X 81 + W_0 = \boxed{Z}$

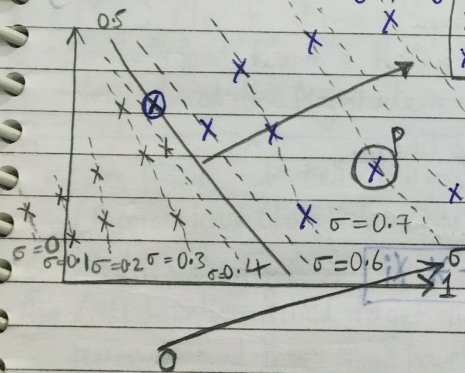
If this Value i.e; $\boxed{Z} > 0 \rightarrow$ Selected
and If this Value $\boxed{Z} < 0 \rightarrow$ Not Selected.

Now, We use the Sigmoid Function to calculate the $\hat{Y}_i$.

$\hat{Y}_i = \Sigma W_i X_i = \boxed{Z}$

- This Value which comes, i.e. 'Z' is passed to Sigmoid Function.
- If the Value after the Sigmoid Function is Positive $\Rightarrow \hat{Y} > 0{\cdot}5 \Rightarrow 1$
  the Sigmoid Function is Negative $\Rightarrow \hat{Y} < 0{\cdot}5 \Rightarrow 0$

① Additional Functionality of Sigmoid Function :-



① $\Sigma W_i X_i = 0 \rightarrow \sigma = 0{\cdot}5$

② If $\Sigma W_i X_i > 0 \rightarrow \sigma = 0{\cdot}6$

③ If $\Sigma W_i X_i > 0 \rightarrow \sigma = 0{\cdot}7$
Similarly $\Rightarrow \sigma \uparrow$

④ If $\Sigma W_i X_i < 0 \rightarrow \sigma = 0{\cdot}4$

⑤ If $\Sigma W_i X_i < 0 \rightarrow \sigma = 0{\cdot}3$
Similarly $\sigma \downarrow$

$\sigma() \Longrightarrow P()$ Probability [Positive → P, Negative → N]

We can say that the $P(P) = 0{\cdot}8$ or above.

Sigmoid Function is beneficial because of following:-

① We can say that Jetha Hamin Line Sei Dutt Jayenge $\Rightarrow P(P)$ will increase

② We can also say that, Jetha Hamin line Sei Dutt Jayenge [Negative Side]
$\Rightarrow P(N)$ will Decrease.

③ i.e; With the help of Sigmoid Function, we can calculate the Probability of both Selected and Not Selected. let's say 'P' Point Selected Probability $= 0{\cdot}8$

P(Not Selected) $= 1 - 0{\cdot}8$
$= \boxed{0{\cdot}2}$

Date

Scale:-

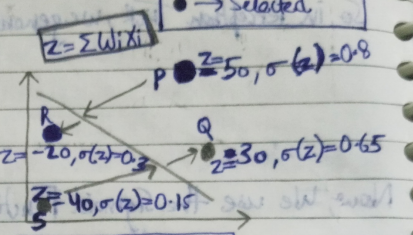
● → Not Selected
● → Selected

⊙ Impact of Sigmoid Function

$W_{new} = W_{old} + \eta(Y - \hat{Y})X_i$

Task-1:- Not Making the $(Y - \hat{Y}) = 0$

$Y_i = \sum W_i X_i \Rightarrow [Z]$



Now, we can see that $(Y_i - \hat{Y}_i)$ is never '0'.

So for Point 'P'. [Right Classified Point]

Now, Equation will be:-

Y_i	$\hat{Y}_i$	$Y_i - \hat{Y}_i$
1	0.8	0.2
0	0.65	-0.65
1	0.8	0.7
0	0.15	-0.15

$W_{new} = W_{old} + \eta * 0.2 * X_i$ Pushing the line...

"Note:- $\sigma(z)$ these all values are Assuming."

So, for Point 'Q'. [Mis Classified Point]

Equation is:-

$W_{new} = W_{old} - \eta * 0.65 * x_i$ Pulling the Line....

Point 'R' [Mis Classified Point]

Equation will be:- $W_{new} = W_{old} + \eta 0.7 * X_i$ Pulling the line....

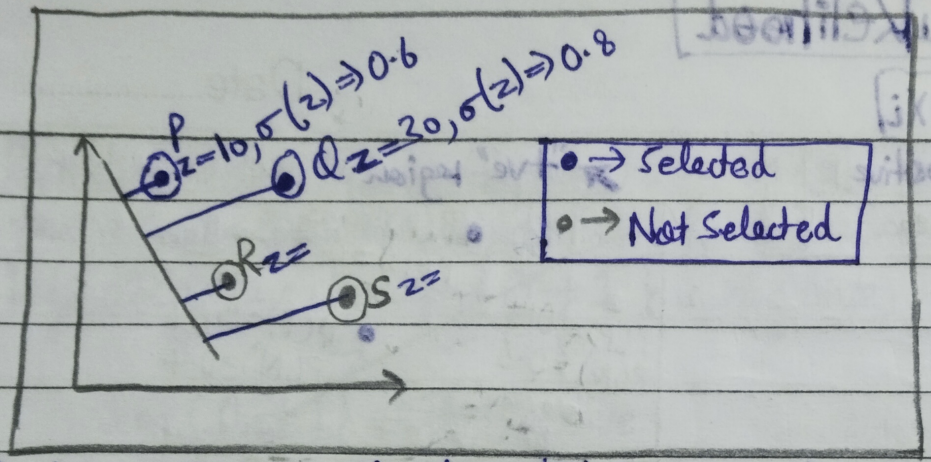
Point 'S' [Right Classified Point]

Equation will be:- $W_{new} = W_{old} - \eta * 0.15 * X_i$ Pushing the Line...

Task-2:- Finding How is the Magnitude Working on the Points.

For this, let's take an example:-

Date.....



Let's say I take, the point 'P'. Assuming that z $[\Sigma W_i X_i] = 10$. So, it's value after Sigmoid Function will be :- 0.6.

Equation ⇒ $\boxed{W_n = W_{old} + \eta * 0.4 * X_i = X_1}$ — ① Pushing the line....

Q Point ⇒ z = 30, σ(z) = 0.8.

	Y_i	$\hat{Y}_i$	$Y_i - \hat{Y}_i$
P	1	0.6	0.4
Q	0	0.8	0.2
	1		
	0		

Equation ⇒ $\boxed{W_{new} = W_{old} + \eta * 0.2 * X_i = X_2}$ — ② Pushing the line....

From ① and ②

Let's say Magnitude = "M"

$\{ M_{X_1} > M_{X_2} \}$ it means Magnitude of X_1 is more than Magnitude of X_2.....

• Similarly Point 'R' and 'S' Pulling the line where $\boxed{M_R < M_S}$.

Hence Proved the Magnitude...