How to evaluate classification model?

⇒ R2 Score ——> No

It tells us how far/ close we are to actual result

Goal ⇒

Check how many correct values are predicted

Accuracy

EXAMPLE

Understand with example

f

y_i	$\hat{y}_i$
1	1
0	0
1	0
0	0
0	1

n

$$y_i = \hat{y}_i$$

$$\text{Accuracy} = \frac{\text{no. of correct predictions}}{\text{Total number of predictions}}$$

$\Rightarrow$ 3 / 5 = 0.6

Our model accuracy is 60%

Logistic Regression

What are odds?

⇒ Betting example

For example, odds of horse winning race is 4 : 1

Odds is ratio of : $\dfrac{\text{Probability of success}}{\text{probability of failure}}$

P (Winning) = 4 / 4+1 = 4 / 5

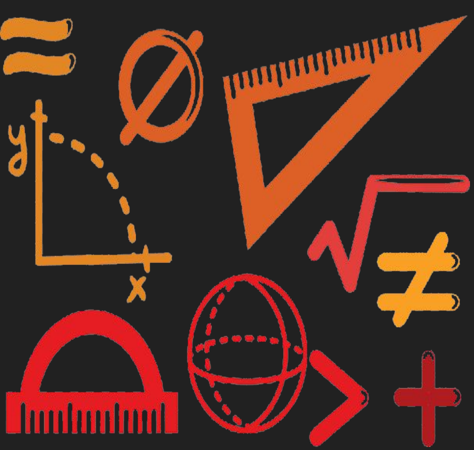
P (Losing) = 1/ 1+4 = 1/5

To sum up :

Therefore odds = P / 1 - P (Odds of class label 1)

P = P (y = 1 / x)

1 - P = P (y = 0 / x)



This is similar to which concept?

$$\sigma(Z)$$

(Also defined probability)

$$p = \sigma(z)$$

$$= \sigma(w^T x + w_0)$$

$$= \frac{1}{1 + e^{-z}} = \frac{1}{1 + \frac{1}{e^z}} = \frac{1}{\frac{e^z + 1}{e^z}}$$

$$\boxed{=> p = \frac{e^z}{1 + e^z}}$$

Similarly , for 1 - P

$$1 - p = 1 - \frac{e^z}{1 + e^z}$$

$$=> 1 - p = \frac{1}{1 + e^z}$$

Now , we know odds = P / 1 - P

$$\frac{p}{1-p}$$

$$= \frac{\frac{e^z}{1+e^z}}{\frac{1}{1+e^z}}$$

$$=> \frac{p}{1-p} = e^z$$

Taking log on both sides,

$$log_e(odds) = log_e(e^z)$$

$$=> log(odds) = log(\frac{p}{1-p}) = z$$

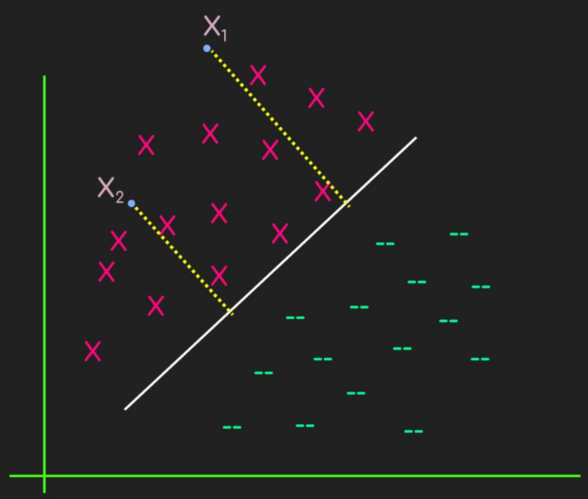
What does this mean geometrically ?

Z : distance of point from hyperplane line

So, log (odds) is distance of point from the line

Therefore, More the distance of point from line increases

Odds of it being in class 1 higher



Odds, of x1 > odds of x2

In class 1

Linear Regression

$$y = w_x^T + w_o$$

(Direct linear relation)

Logistic Regression

$$log(\frac{P}{1-P}) = w_x^T + w_o$$

(Non linear relation due to log , no direct relation)

Impact of Outliers

Remember outliers, how do you think they will impact logistic regression model.

⇒ How can you measure its impact?

We derived the loss function of logistic regression

$$L = -[y \log \hat{y} + (1-y)log(1-\hat{y})]$$

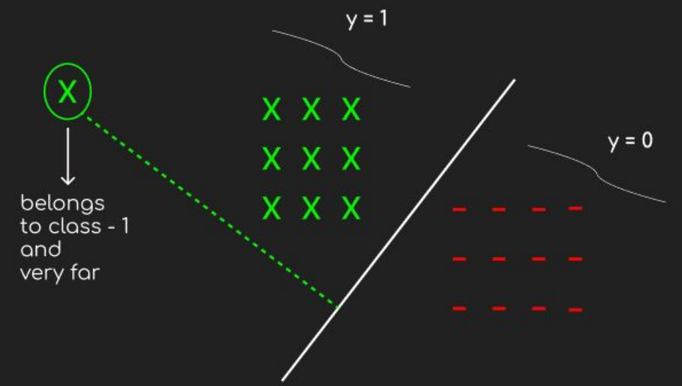
Let's see how it will vary with outliers

Case 1 : When outlier is on correct side

Now, we know:

$$L = -[y \log \hat{y} + (1-y)log(1-\hat{y})]$$

y = 1 0



Since the outlier belongs to class 1 , L becomes

$$L = -log\ \hat{y}$$

Now,

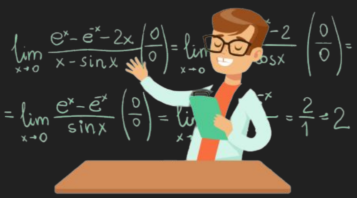
$$\hat{y} = \sigma(z^i)$$

$$(z^i) is\ significantly\ large$$

$$\sigma(z^i) \longrightarrow 1$$

$$log\ \hat{y} \longrightarrow 0$$

Very low impact !

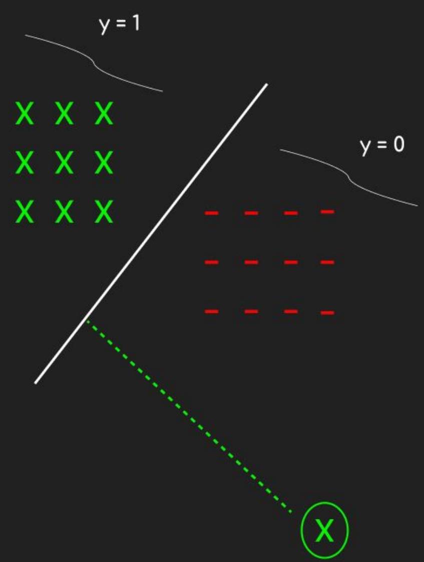


Case 2 : When outlier is on opposite/ incorrect side

Now, we know:

$$L = -[y \log \hat{y} + (1-y) log(1-\hat{y})]$$

y = 1



Since the outlier belongs to class 1 , L becomes

$$L = -log\ \hat{y}$$

Now,

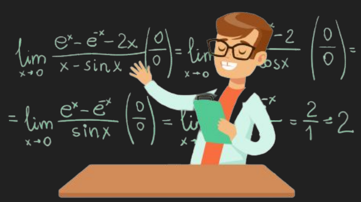
$$\hat{y} = \sigma(z^i)$$

$$(z^i) is\ large\ in\ negative\ i.e,\ -4.3$$

$$\sigma(z^i) \longrightarrow 0$$

$$log\ \hat{y} \longrightarrow \text{large}$$

Very high impact !



Multi class classification

⇒ We dealt with 2 classes so far.
But, what if there are more than 2 classes ??

Suppose :

Orange

Apple

Grape

⇒ 3 classes

Create 3 logistic regression models

M1 : Orange or not

M2 : Apple or not

M3 : Grape or not

3 binary classification models

But we can't use the same database.

Since it has 3 classes as output

⇒ Modify the data

Make all other categories as 0 (Same)

M1 Data

x	y
	1
	0
	0
	1
	0
	0

If $y^{(i)}$ = orange
= 1

else 0

M1 ⇒ Log regression 1

Make all other categories as 0 (Same)

M2 Data

x	y
	0
	1
	0
	0
	1
	0

If $y^{(i)}$ = apple

= 1

else 0

M2 ⇒ Log regression 2

Similarly, M3 ⇒ Log regression 3

How to predict now?