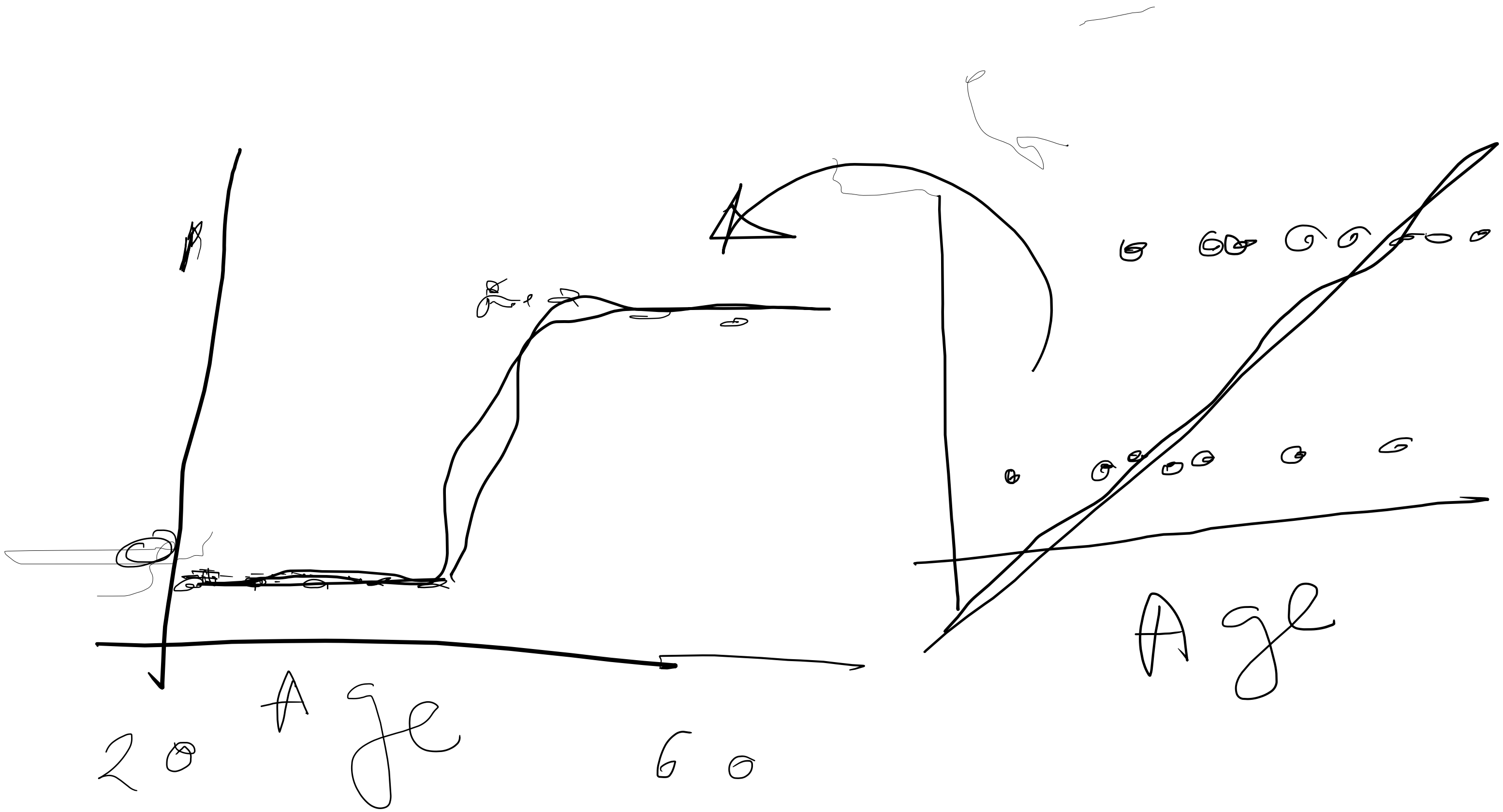
Logistic Regression
logistic classification

Arunkumar Nair



Age
20
60
Age

Case-Study Data

We are provided a sample of 1000 customers.

We need to predict the probability whether a **customer of a Particular Age** will buy (y) a particular magazine or

not.

As we’ve a categorical outcome variable, we’ll use logistic regression.

Linear to Logistic – (a)

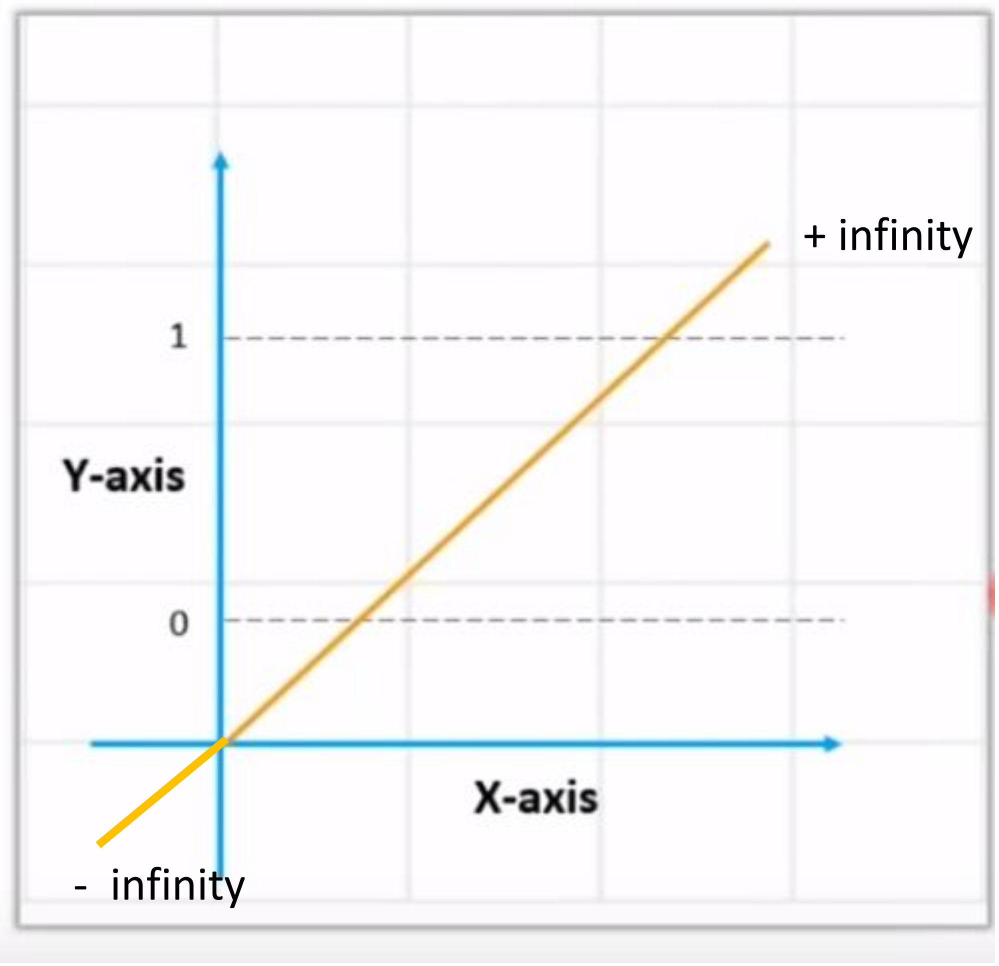
- To start with logistic regression, first write the simple linear regression equation with dependent variable enclosed in a link function:

$$Y = c + m X$$

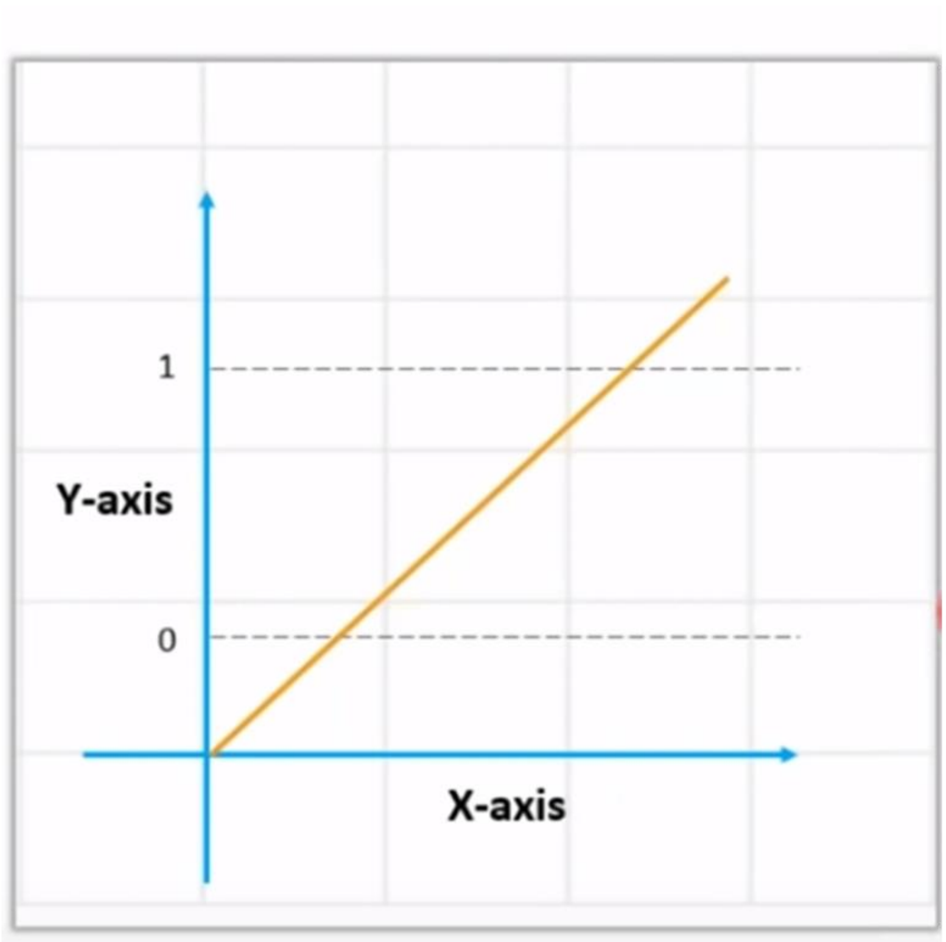
$$g(y) = \beta o + \beta(\text{Age}) \text{——} (a)$$

For understanding, consider '*Age*' as independent variable.

Linear Regression

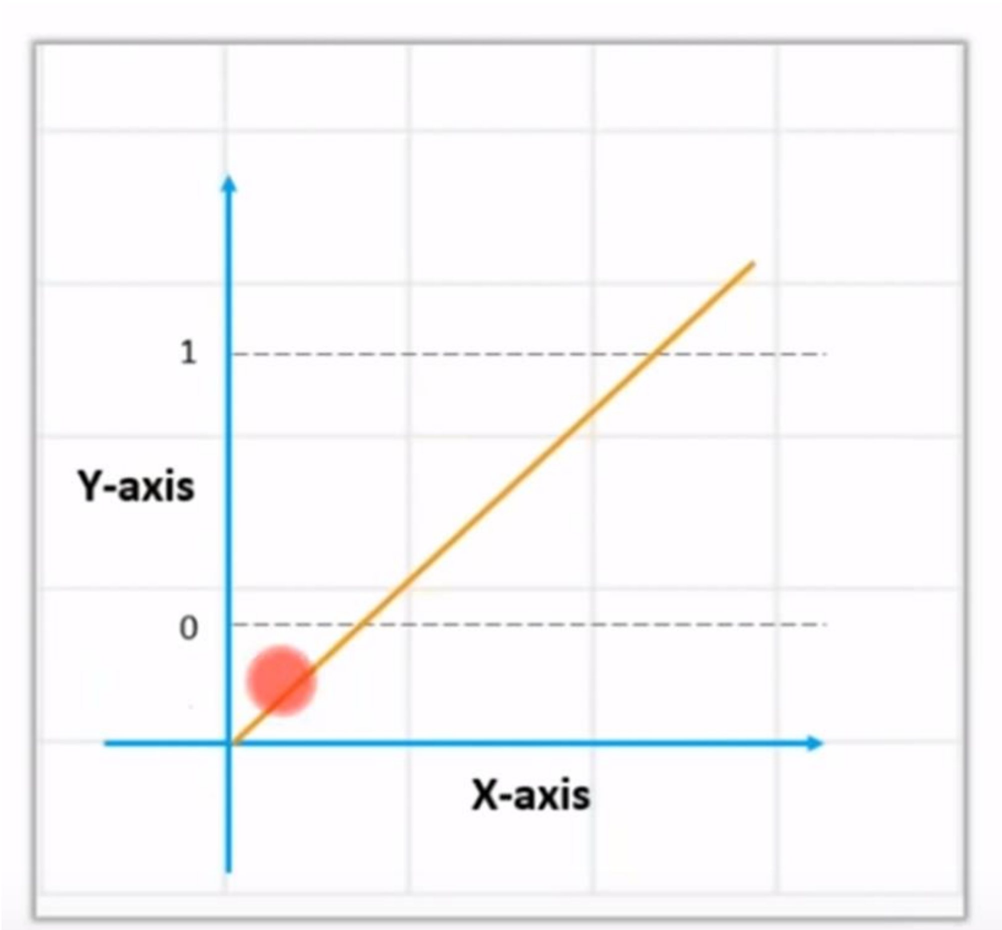


Linear regression equation: $y = \beta 0 + \beta 1X1 + \beta 2X2 \ldots. + \beta nXn$



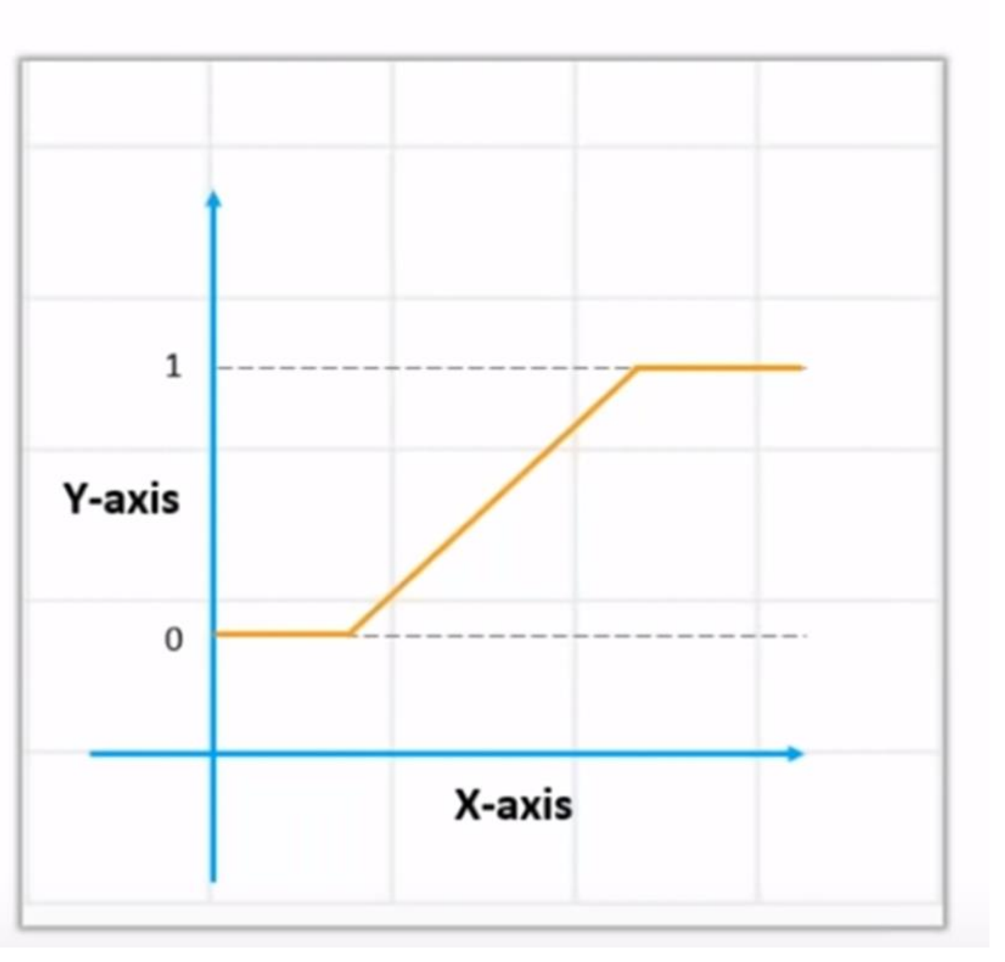
Since our value of Y will be between 0 and 1, the linear line has to be clipped at 0 and 1.

Value of Y – between 0 and 1



Since our value of Y will be between 0 and 1, the linear line has to be clipped at 0 and 1.

How to get the value of 0 and 1



How to get the value of 0 and 1?
Use Sigmoid

- We Apply sigmoid function on the linear regression equation to get the S-curve so that it lies between 0 and 1

Sigmoid function: $\boldsymbol{p = 1 / 1 + e^{-y}}$

- A sigmoid function is a mathematical function/equation having a characteristic "S"-shaped curve or sigmoid curve.

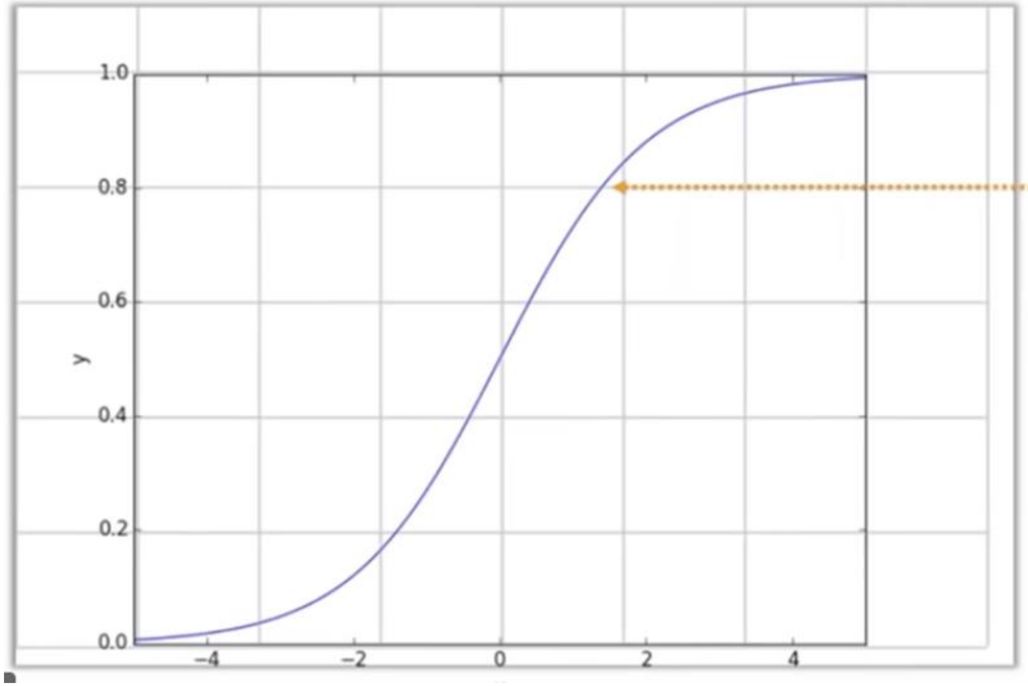
Convert Linear to Logistics

- Linear regression equation: **$y = \beta 0 + \beta 1X1 + \beta 2X2 \ldots. + \beta nXn$**
- Sigmoid function: **$p = 1 / 1 + e^{-y}$**

  e^{-y} y is replaced

- Logistic Regression equation: **$p = 1 / 1 + e^{-(\beta 0 + \beta 1X1 + \beta 2X2 \ldots. + \beta nXn)}$**

Sigmoid – S-curve



The Sigmoid "S" Curve

Sigmoid

- Sigmoid curve converts any value from

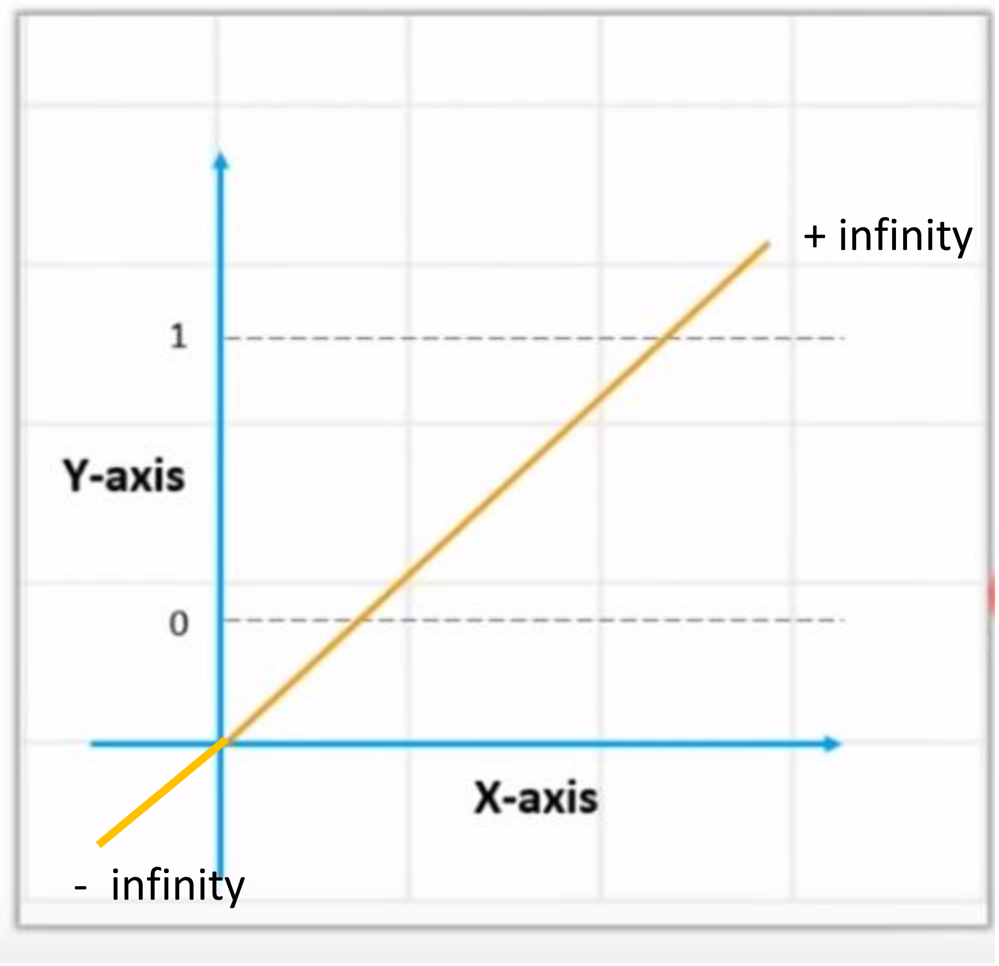
-infinity to +infinity to (0 to 1)

- Sigmoid will output:

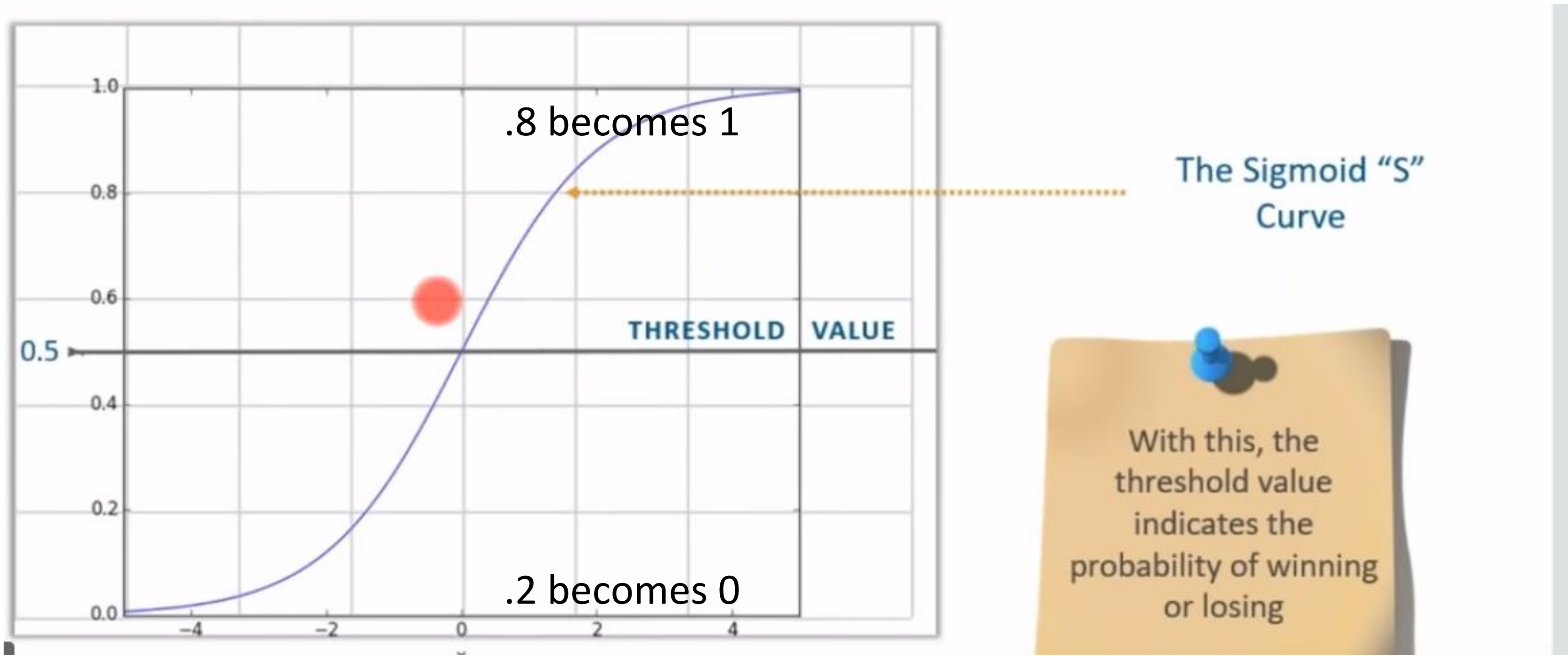
‘0’ as x approaches $-\infty$

‘1’ as x approaches $+\infty$

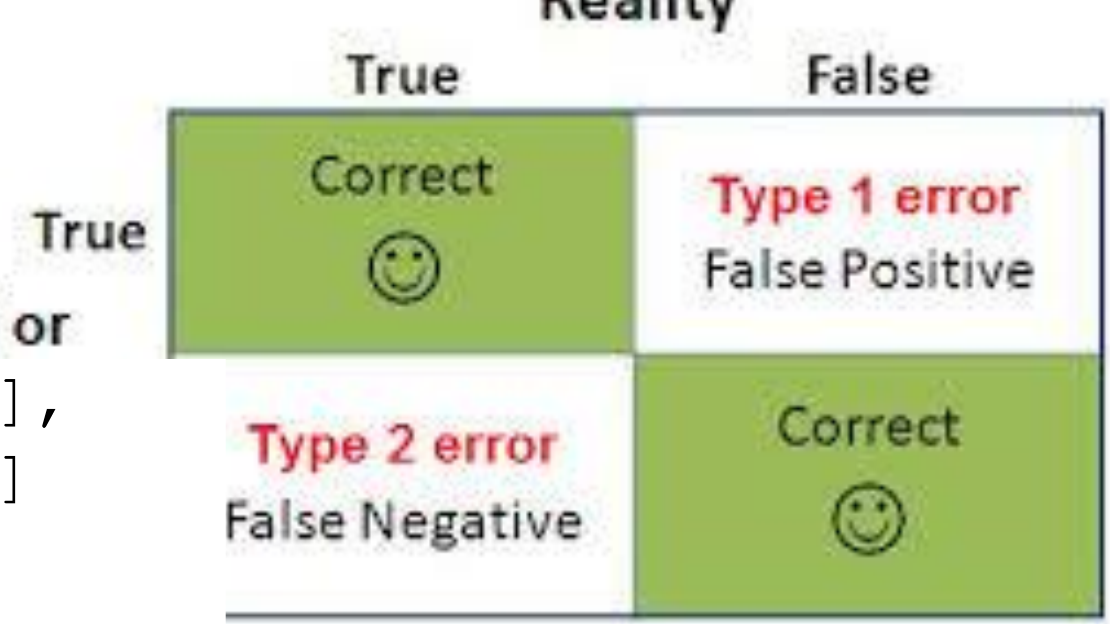
Linear Regression



Probability values for the answers

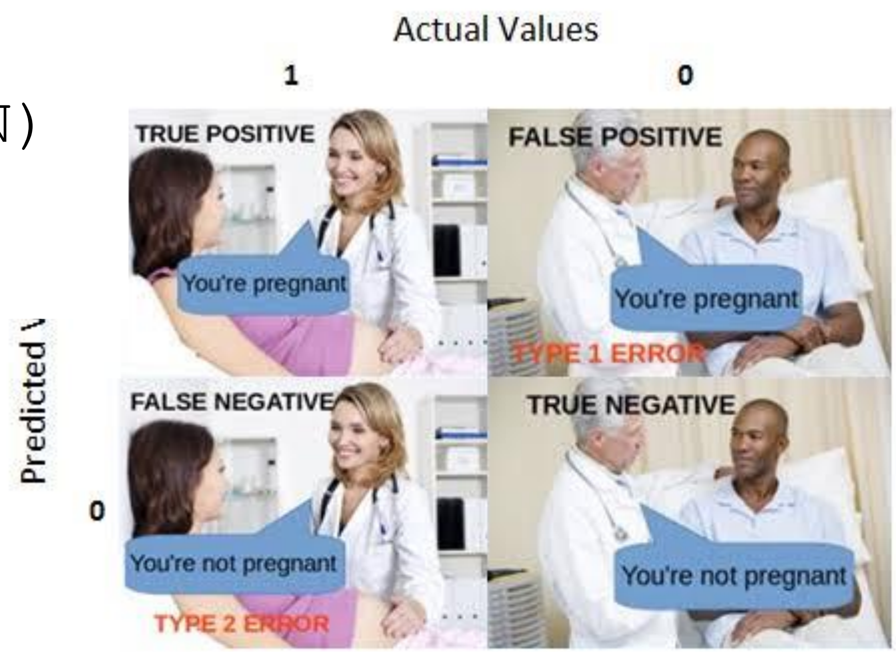


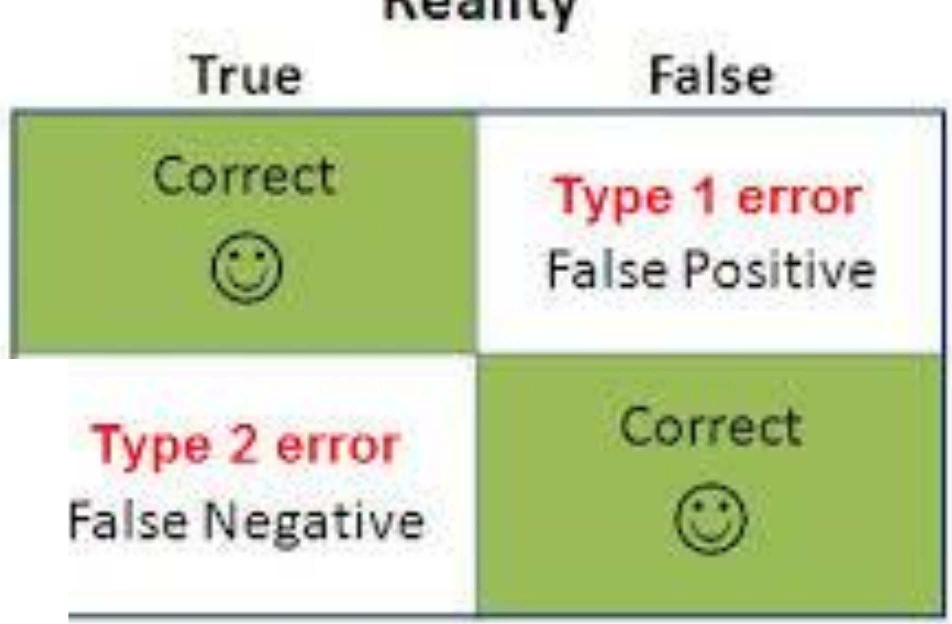
		Reality	
		True	False
Measured or Perceived	True	Correct ☺	**Type 1 error** False Positive
	False	**Type 2 error** False Negative	Correct ☺



[[119, 11],
 [26, 36]]

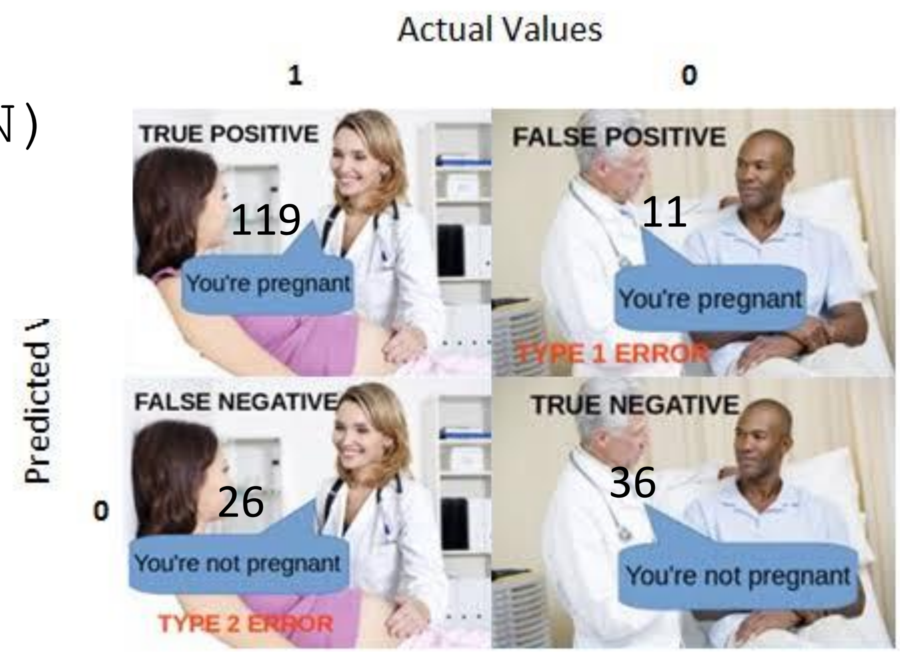
Accuracy=
(TP + TN)/
(TP+TN+FP+FN)





[[119, 11],
 [26, 36]

Accuracy=
(TP + TN)/
(TP+TN+FP+FN)



False Positive Rate (caused by Type I Error)

Type1 Error- not to worry

Sometimes, this error might translate to a simple case where a person is predicted to have some bacterial infection while actually that might not be the case. The medication to treat simple bacterial infections might not be very dangerous and is believed to have very mild or no side effects on the patient.

So, in such cases, we might not worry much about the Type I error..

False Positive Rate (caused by Type I Error)

Type1 Error- if it is cancer then we Need to worry

Things can get complicated and serious if Type 1 error happens in a scenario where a person *not* suffering from cancer is diagnosed to have cancer. This can be really dangerous and sometimes fatal due to the high doses of radiation and chemotherapy that a patient can be exposed to.

True Negative Rate (or Specificity)

True Negative Rate (or Specificity)is a metric that tells us how often the model predicts ‘no’ for an actual ‘no’. It is equivalent to 1 minus False Positive Rate.

False Negative Rate (caused by Type II Error):

Number of items the model wrongly predicted ‘no’ out of the total actual ‘yes’. This metric is especially important in most binary classification problems, as it tells us the frequency with which a positive instance is wrongly identified as negative. For example, if a cancer patient is wrongly diagnosed as not having cancer, that individual would either go undiagnosed or misdiagnosed. Similarly, identifying a fraudulent transaction as non-fraudulent can cause several serious repercussions for a bank. Hence, whenever we intend our model to be a diagnostic aid, we would always want this metric to be as low as possible.

Precision — Out of all the examples that predicted as positive, how many are really positive?
positive predictive value

		Real	
		Positive	Negative
Predicted	Positive	True Positive (tp)	False Positive (fp)
	Negative	False Negative (fn)	True Negative (tn)

$$precision = \frac{tp}{tp + fp}$$

$$recall = \frac{tp}{tp + fn}$$

$$specificity = \frac{tn}{tn + fp}$$

Recall — Out of all the positive examples, how many are predicted as positive?

Specificity — Out of all the people that do not have the disease, how many got negative results?

Recall and Sensitivity are the same.

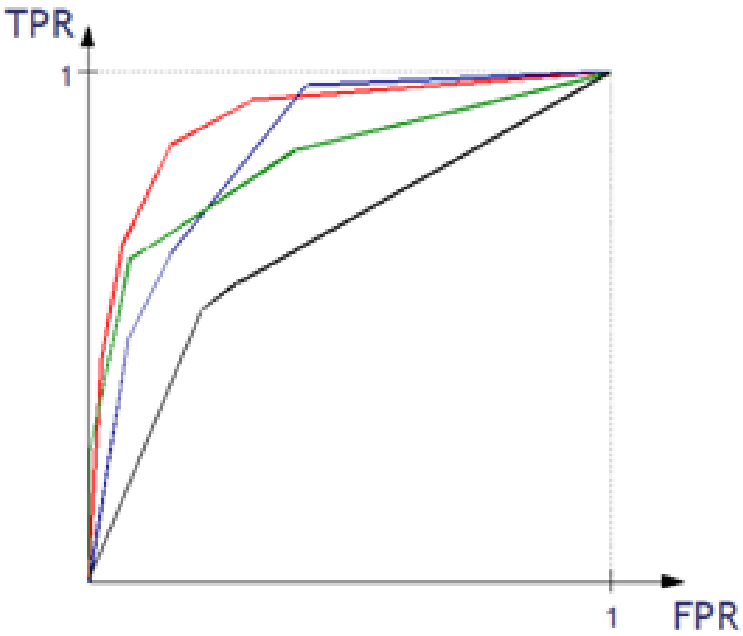
What is the ROC curve?

The Receiver Operating Characteristic (ROC) curve is one of the methods for visualizing classification quality, which shows the dependency between TPR (True Positive Rate) and FPR (False Positive Rate).

predicted→ real↓	*Class_pos*	*Class_neg*
Class_pos	TP	FN
Class_neg	FP	TN

$$\text{TPR (sensitivity)} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

$$\text{FPR (1-specificity)} = \frac{\text{FP}}{\text{TN}+\text{FP}}$$



Thanks

End