#LifeKoKaroLift

Data Science Certification Program

upGrad

Course : ML

Lecture On : Logistic Regression

Instructor : Shivam Garg

upGrad

Today's Agenda

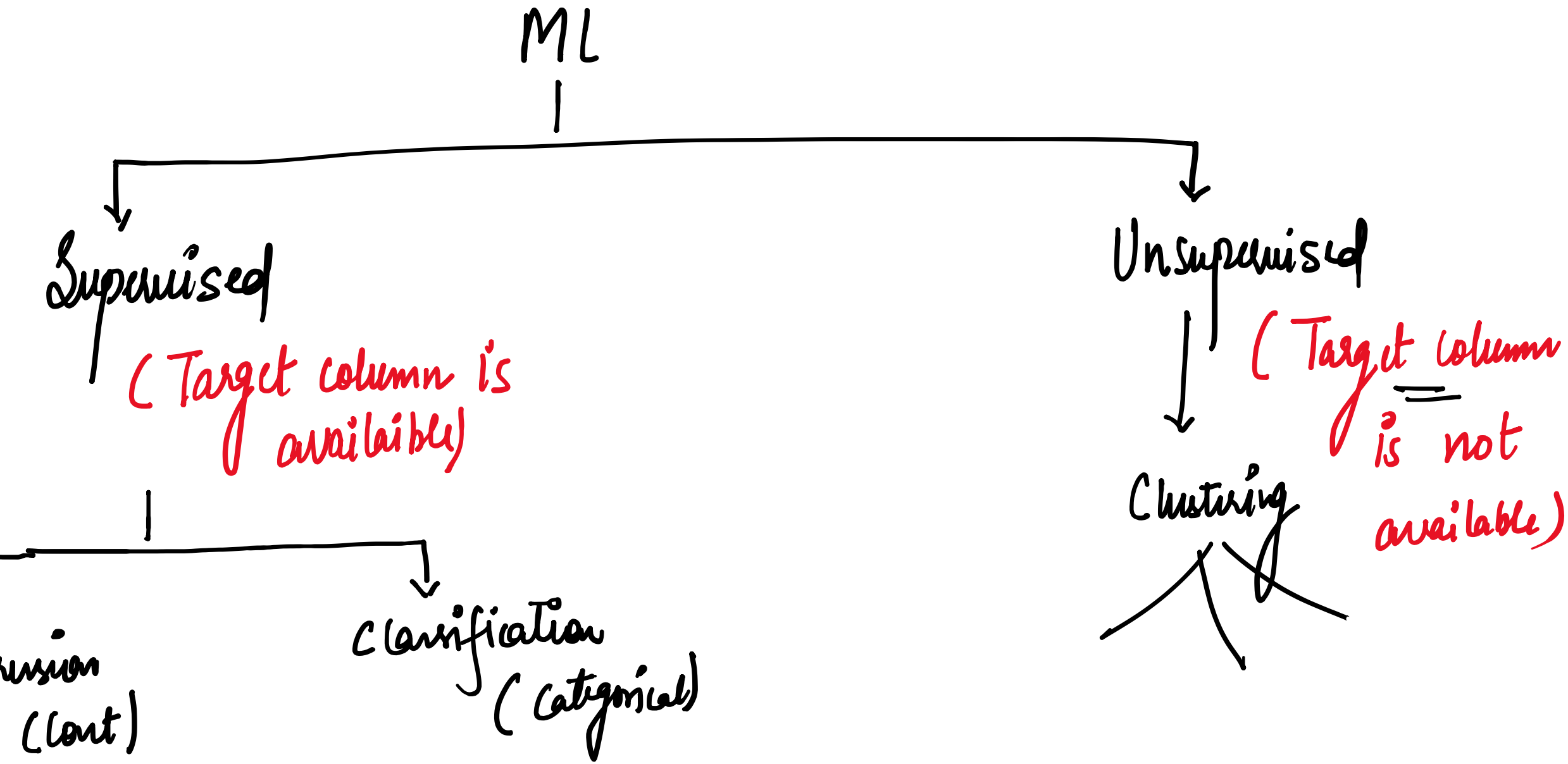
1 Introduction to Classification

2 Binary and Multiclass Classification } Industrial Examples

3 Sigmoid Curve

4 Fundamental Concepts of Logistic Regression }

Evaluation metrics


ML
Supervised
(Target column is available)
Unsupervised
(Target column is not available)
Clustering
Classification
(categorical)
(cont)

Regression v/s Classification

upGrad

Regression

- Used when output is a Continuous Variable
- Want to predict the actual value of something

Classification

- Used when output is a Discrete Variable (Categorical)
- Want to predict different classes

Examples of Regression problems

- Given an image, predict the number of "likes" that the image will get on Facebook
- Given a product on Amazon, how many units will it sell next month?

Examples of Classification problems

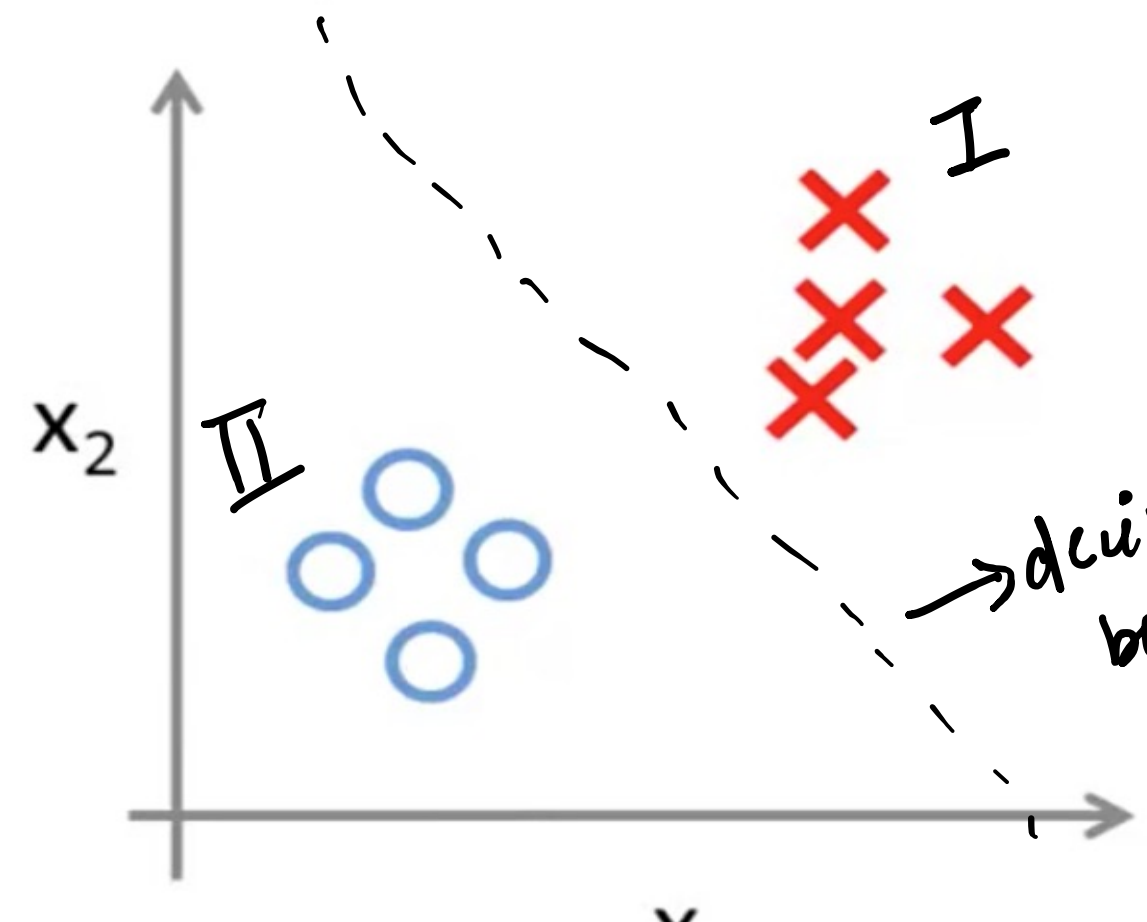
- Given an image, does it contain a "cat" or a "dog"?
- Given a review on Amazon, is the sentiment of the review "positive" or "negative"?

Types of Classification:

upGrad

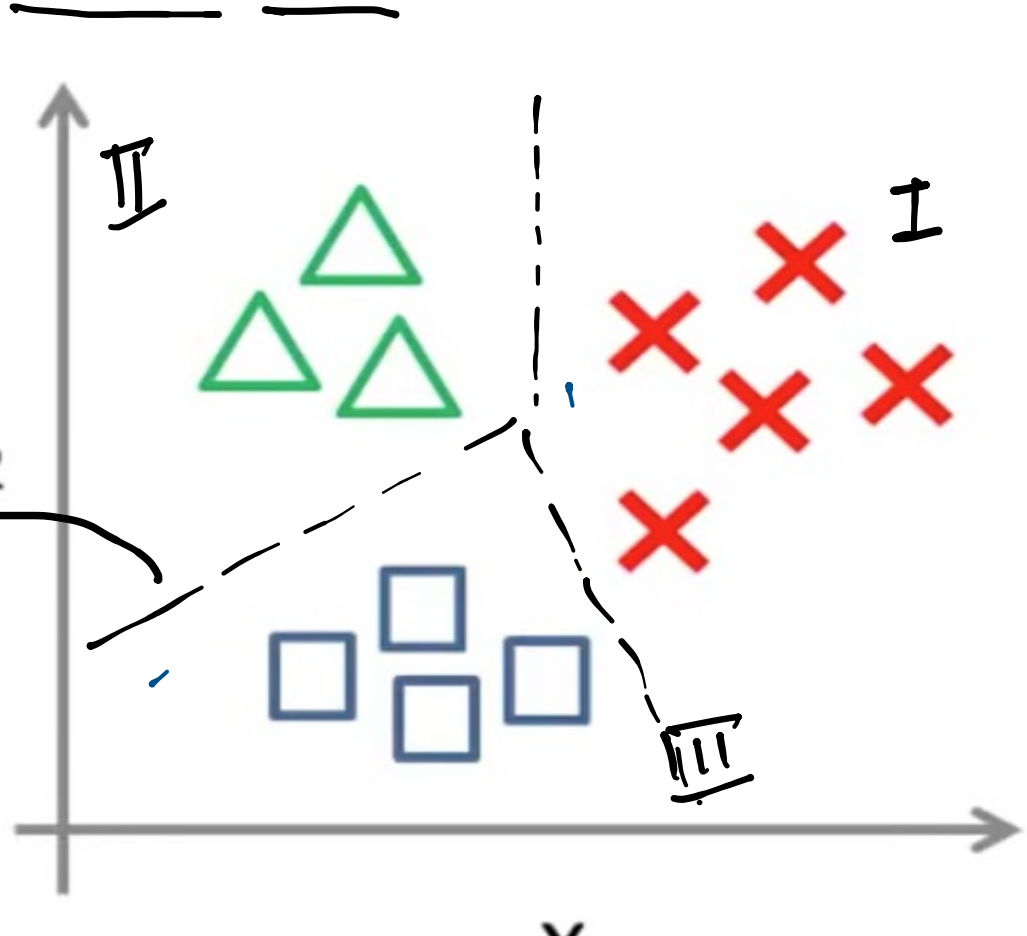
2 classes (categories)

Binary classification:



> 2 classes

Multi-class classification:



Examples of Binary Classification:

upGrad

Fintech

Fraud Detection

X: Bureau Variables

Y: Fraud/Not-Fraud

Type of Document

X: Textual Features

Y: Politics/Sports

Health status

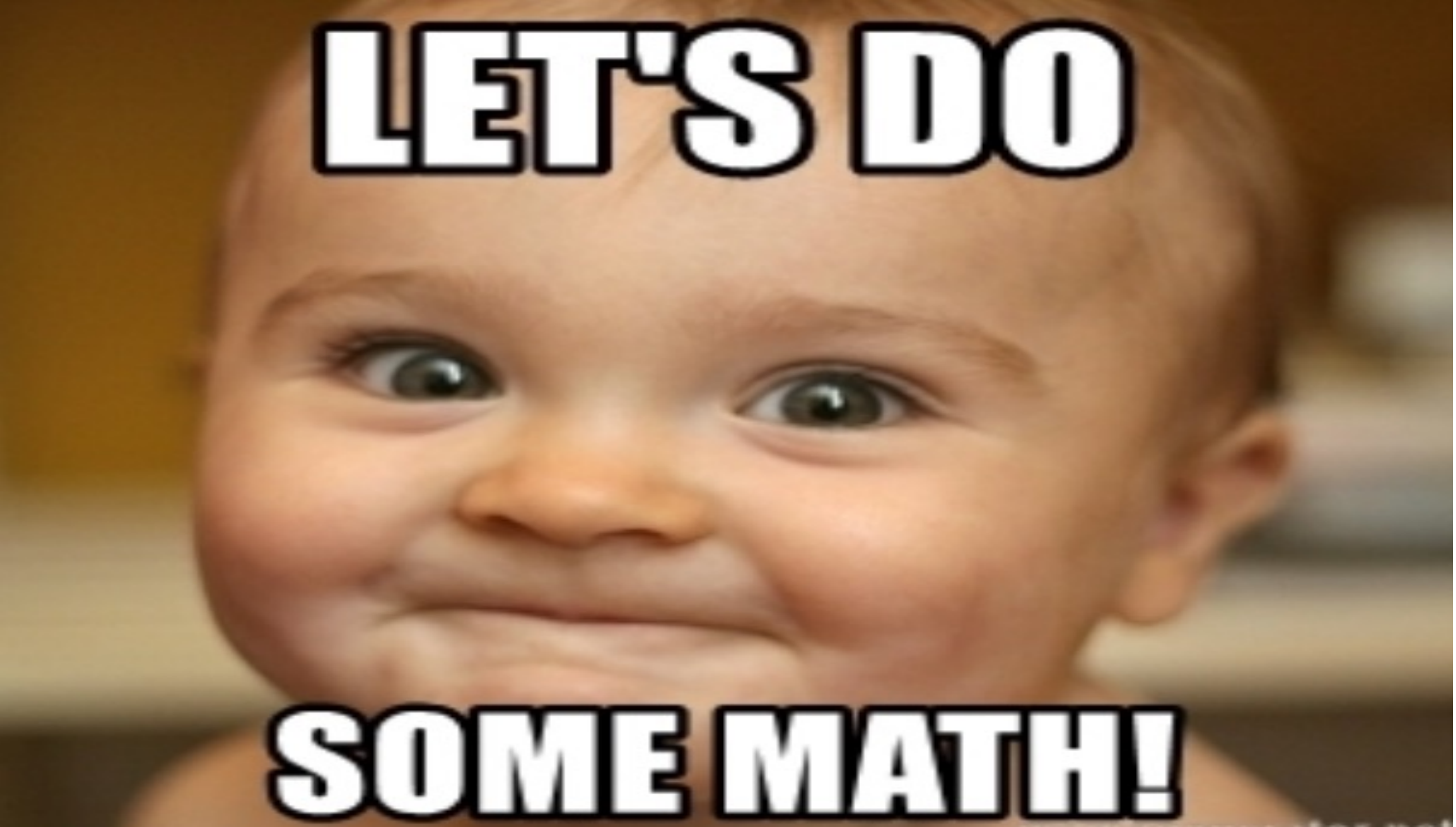
X: Heath features like BP, Sugar level etc.

Y: Healthy/ Un-healthy

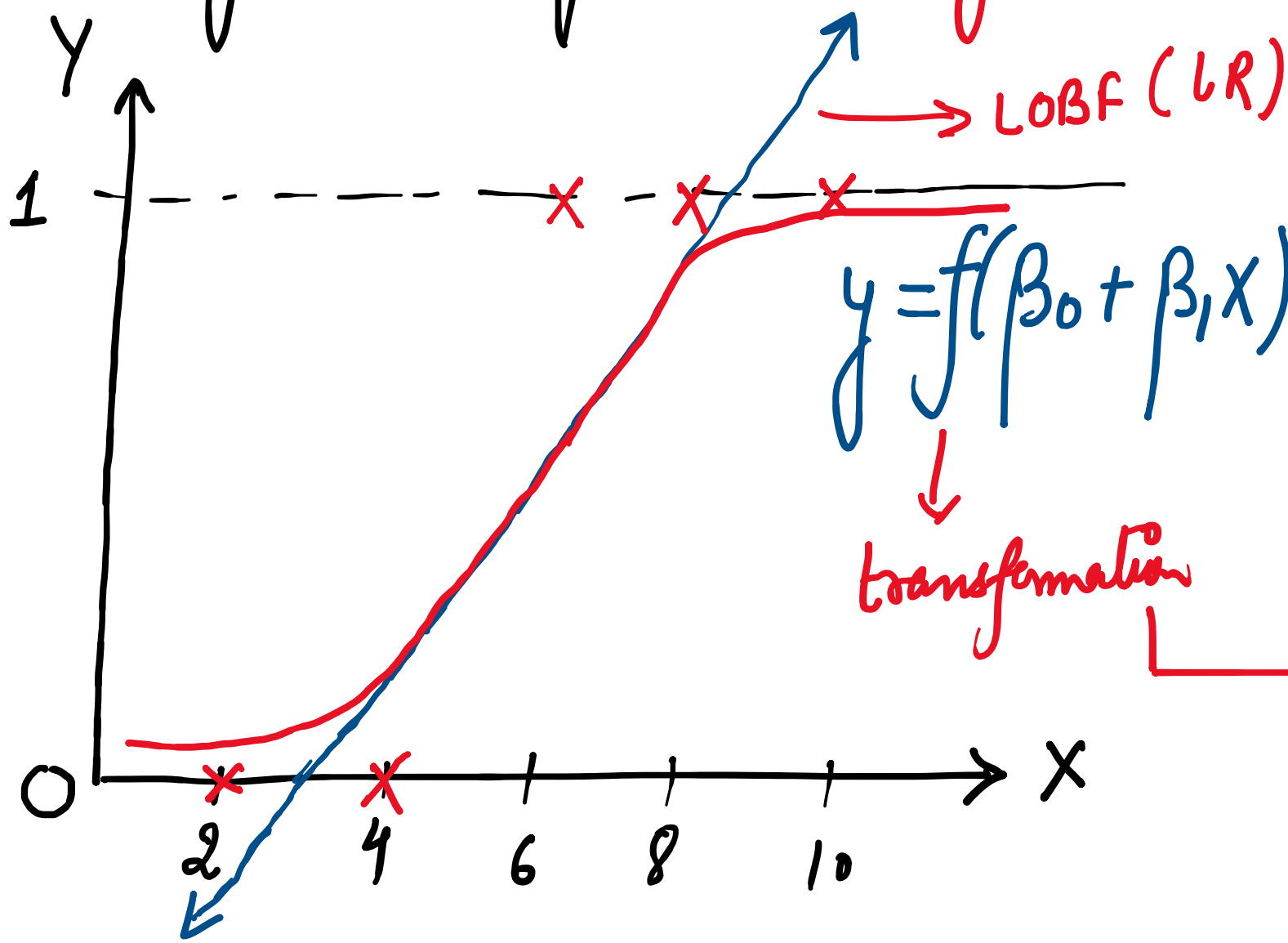
Image Classification

X: Image Features

Y:Cat/Dog


LET'S DO
SOME MATH!

→ Logistic Regression :- Algorithm for classification task.



→ LOBF (LR)

$y = f(\beta_0 + \beta_1 X)$

↓ transformation

→ Sigmoid fn

↓ categorical col^n

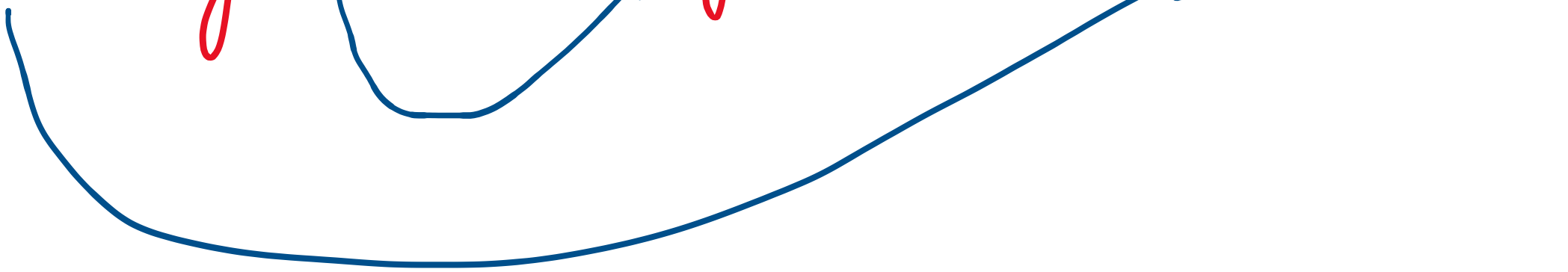
X	Y
2	0
8	1
4	0
6	1
10	1

→ Why can't we Linear regression !

Classification $Y \in [0,1]$

Regression $-\infty < y < \infty$

Logistic Regression = Regression + Sigmoid (logit)



→ Sigmoid Curve :- S Shaped curve.

Mathematically,

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

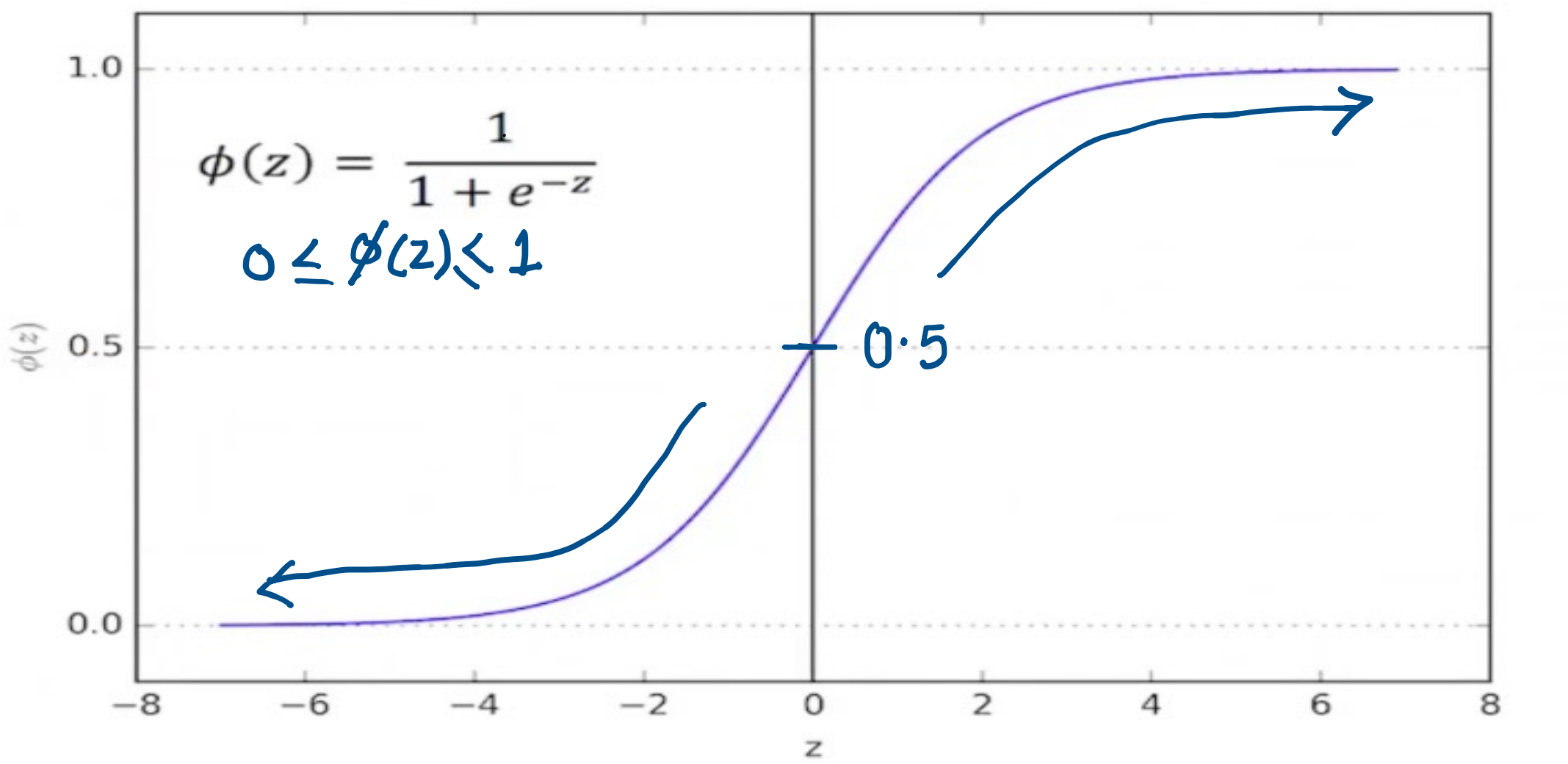
$$0 \leq f(z) \leq 1$$

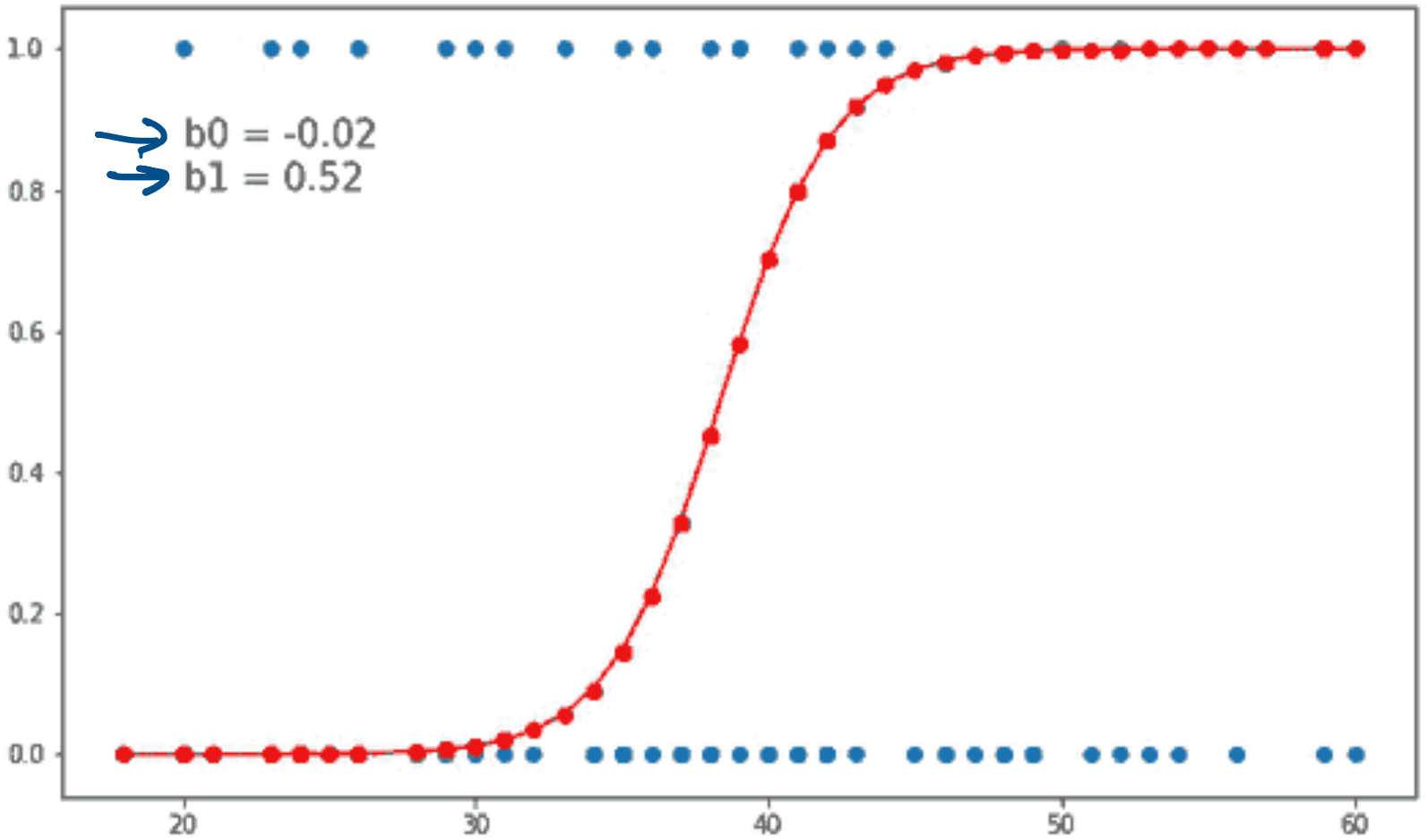
Eqⁿ of logistic Regression :

$$y = \frac{1}{1+e^{-(\beta_0+\beta x)}}$$

Sigmoid Curve:

upGrad




b0 = -0.02
b1 = 0.52
1.0
0.8
0.6
0.4
0.2
0.0
20
30
40
50
60

Poll Question-1:

upGrad

Ques-1: What is the output of Sigmoid function if z=1000?

(A) 1

(B) -1

(C) 0

(D)Infinity

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

$$f(1000) = \frac{1}{1+e^{-1000}} \approx 0$$

≈ 1

Poll Answer-1:

upGrad

Ques-1: What is the output of Sigmoid function if z=1000?

(A) 1

(B) -1

(C) 0

(D)Infinity

Poll Question-2:

upGrad

Ques-2: What is the output of Sigmoid function if z=-1000?

(A) 1

(B) -1

(C) 0

(D)Infinity

$$f(-1000) = \frac{1}{1+e^{-(-1000)}}$$

$$= \frac{1}{1+e^{1000}}$$

$$= \frac{1}{\text{Very Big Num}}$$

$$\approx 0$$

Poll Question-2:

upGrad

Ques-5: What is the output of Sigmoid function if z=-100?

(A) 1

(B) -1

(C) 0

(D)Infinity

Poll Question-3:

upGrad

Ques-3: Fraud Detection, Image Classification, Diagnostic, and Customer Retention are applications in which of the following?

(A) Unsupervised Learning: Regression

(B) Supervised Learning: Classification

(C) Unsupervised Learning: Clustering

(D) Reinforcement Learning

Poll Answer-2:

upGrad

Ques-3: Fraud Detection, Image Classification, Diagnostic, and Customer Retention are applications in which of the following?

(A) Unsupervised Learning: Regression

(B) **Supervised Learning: Classification**

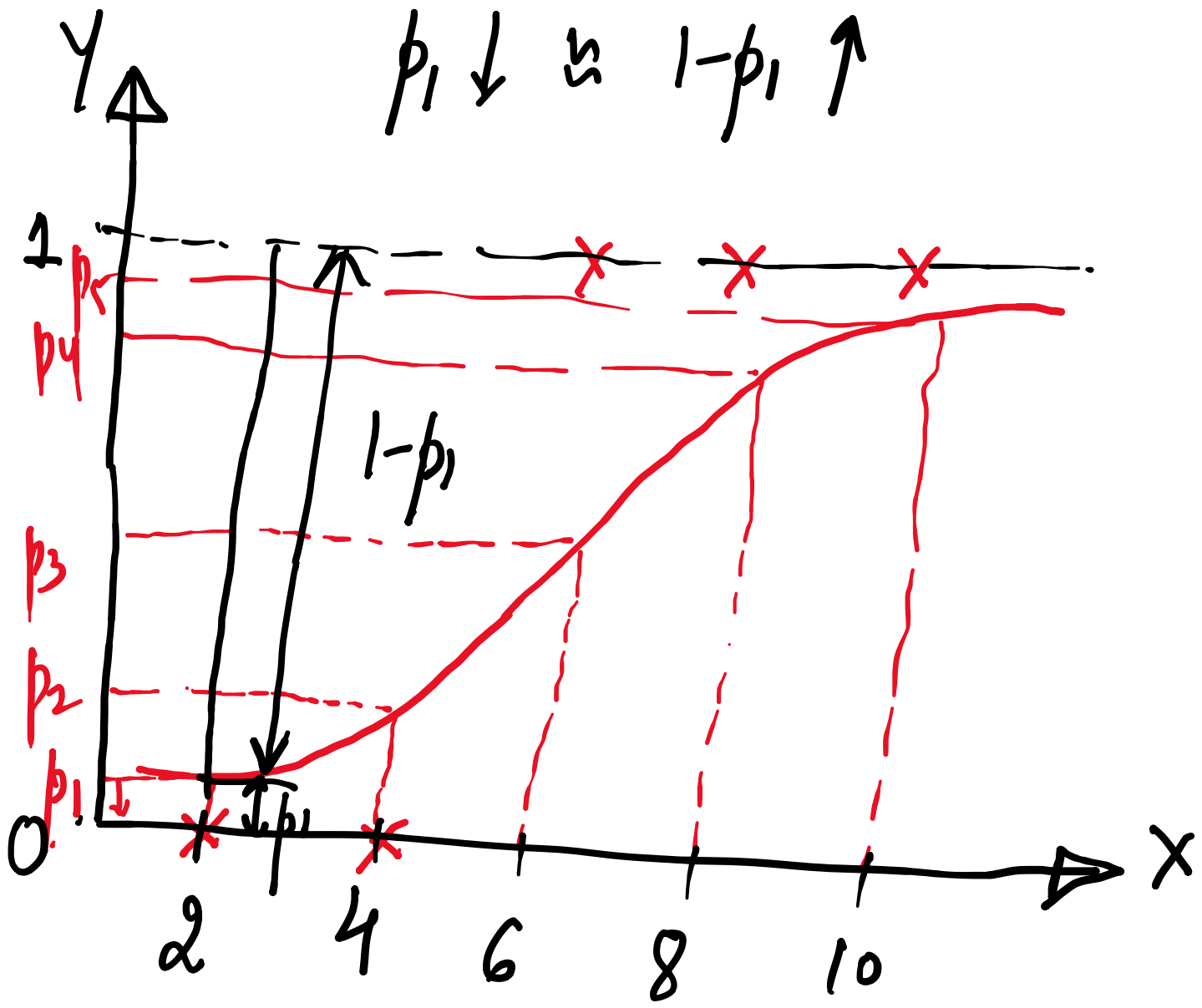
(C) Unsupervised Learning: Clustering

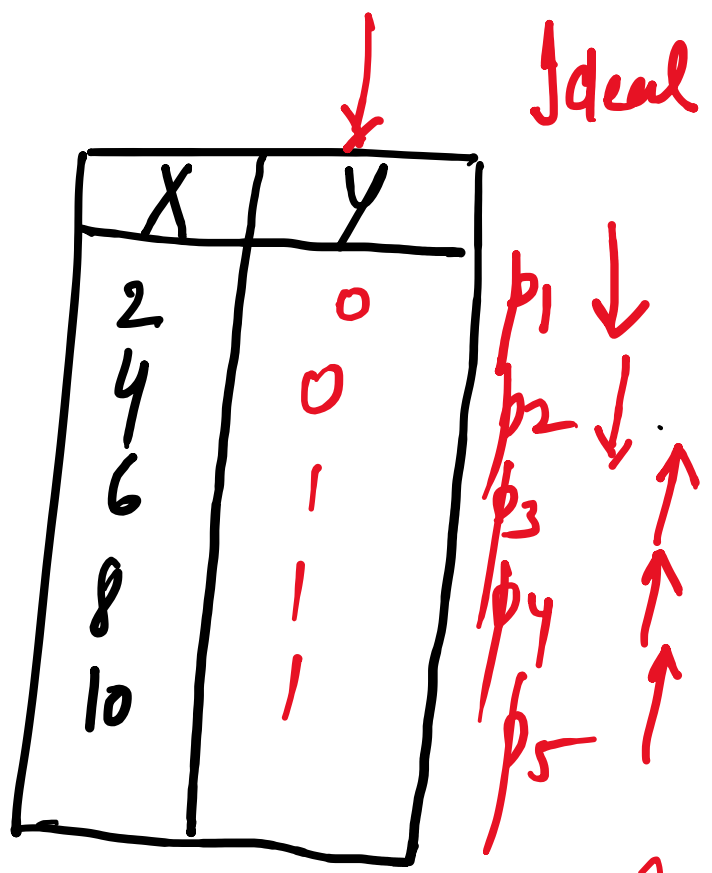
(D) Reinforcement Learning

→ Cost function :- This is the error/loss function which is to be minimized to get the optimal value of coefficients (β_i) → $\beta_0, \beta_1, \beta_2$ - - -

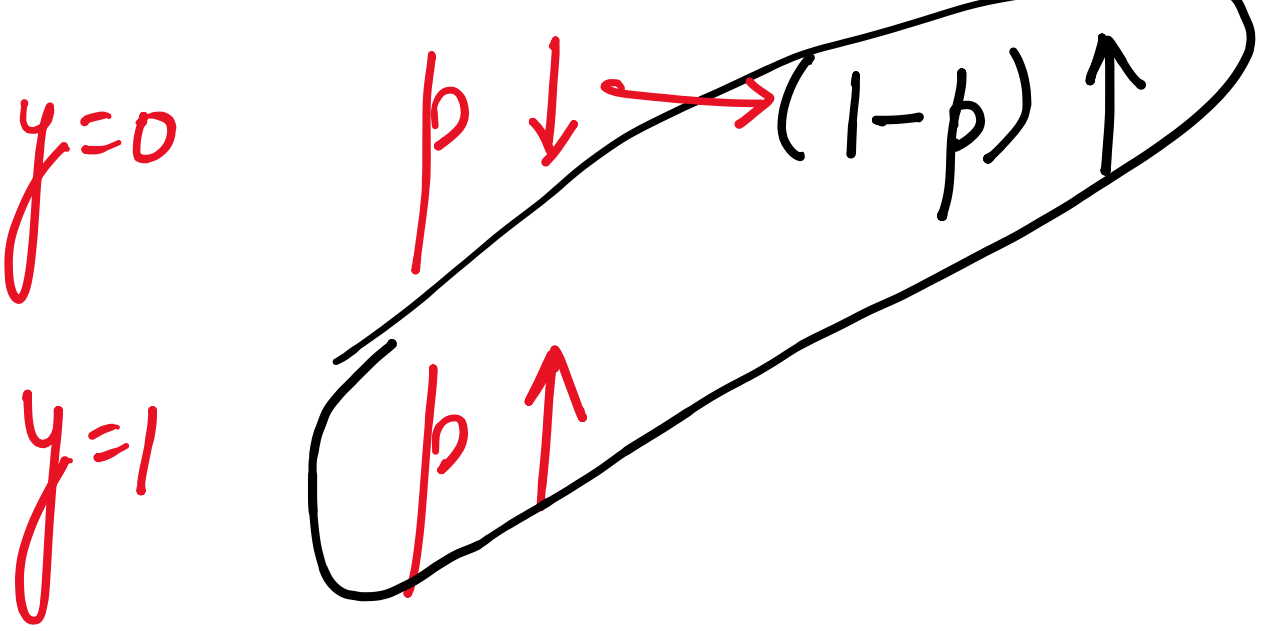
* Cost f^n of linear Reg : Mean Square Error (MSE)

* Cost f^n of logistic Reg : Maximum likelihood Estimation (MLE)





$y=0$ $p \downarrow$ $\quad y=1$ $p \uparrow$ }



$$MLE = \left[\text{Max} \left[(1-p_1)(1-p_2) \cdots \right]_{y=0} \; p_3 \times p_4 \times p_5 \times \cdots p_n \right]_{y=1}$$

→ Evaluation Metrics :-

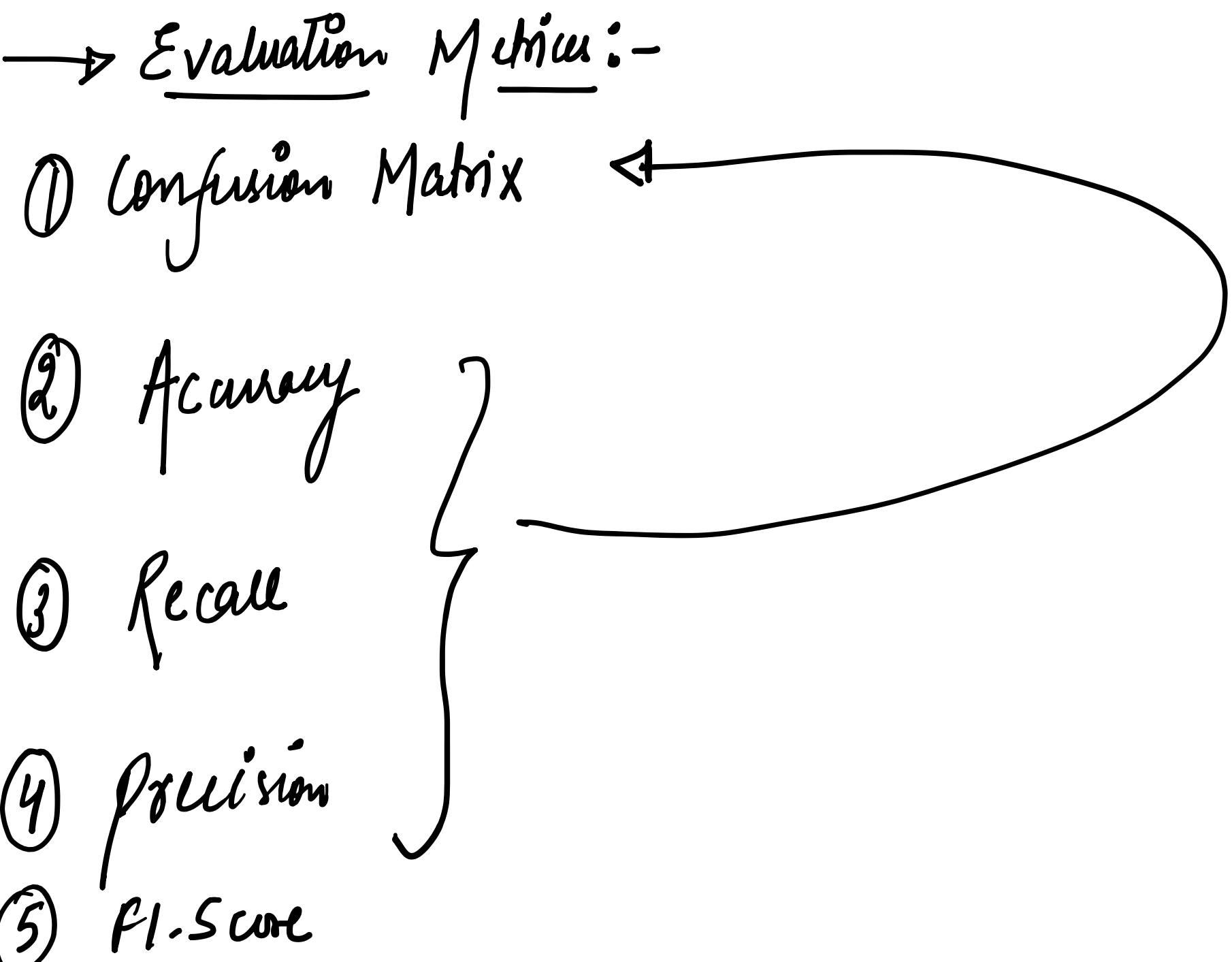
① Confusion Matrix

② Accuracy

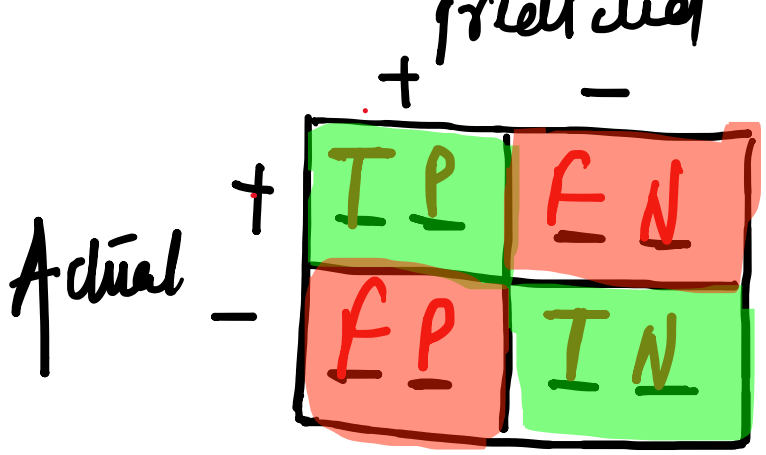
③ Recall

④ Precision

⑤ F1-Score



→ Confusion Matrix :-



+ve → 1

−ve → 0

T/F P/N

→ True / False

→ Actual = predicted

→ Actual ≠ predicted

→ Positive / Negative = predicted class

TP & TN ↑

FP & FN ↓

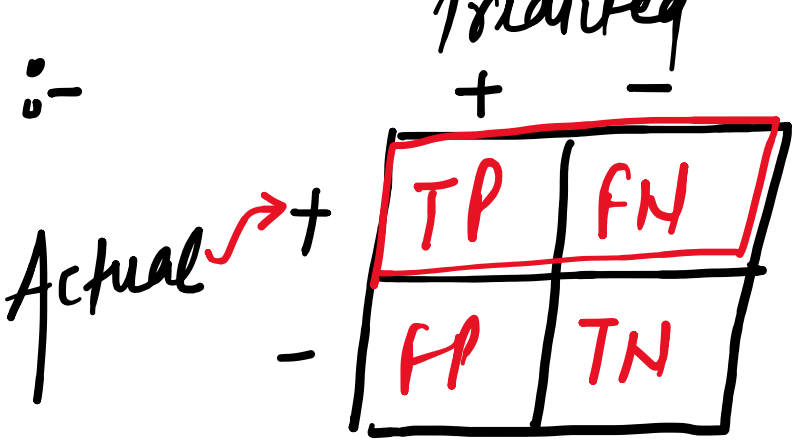
→ Accuracy :-

$$Acc = \frac{\text{Correct Predictions}}{\text{Total predictions}}$$

$$= \frac{TP + TN}{TP + FP + FN + TN}$$

→ Recall / Sensitivity / True positive Rate :-

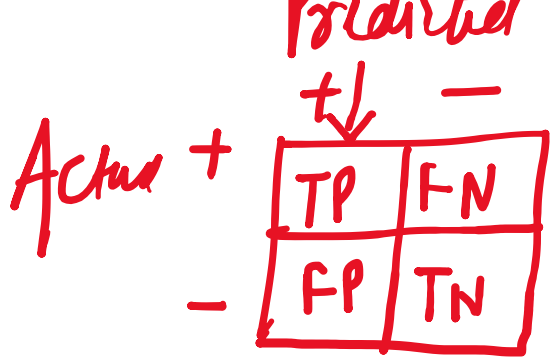
Accuracy of +ve class



$$\text{Recall} = \frac{\text{Correctly predicted +ve class}}{\text{Total actual +ve class}}$$

$$= \boxed{\text{Recall} = \frac{TP}{TP + FN}}$$

→ Precision :-



$$\text{Precision} = \frac{\text{Correctly predicted true class}}{\text{Total predicted true class}}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

→ F1 Score :- Combined performance of Precision & Recall

$$F1\ Score = \frac{2 \times P \times R}{(P+R)}$$

P = Precision

R = Recall

Model.predict ⟶ $0 < y < 1$

By default = 0.5

cutoff / threshold

$y_p \geq 0.5 = 1$

$y_p < 0.5 = 0$

0.8 ⟶ 1

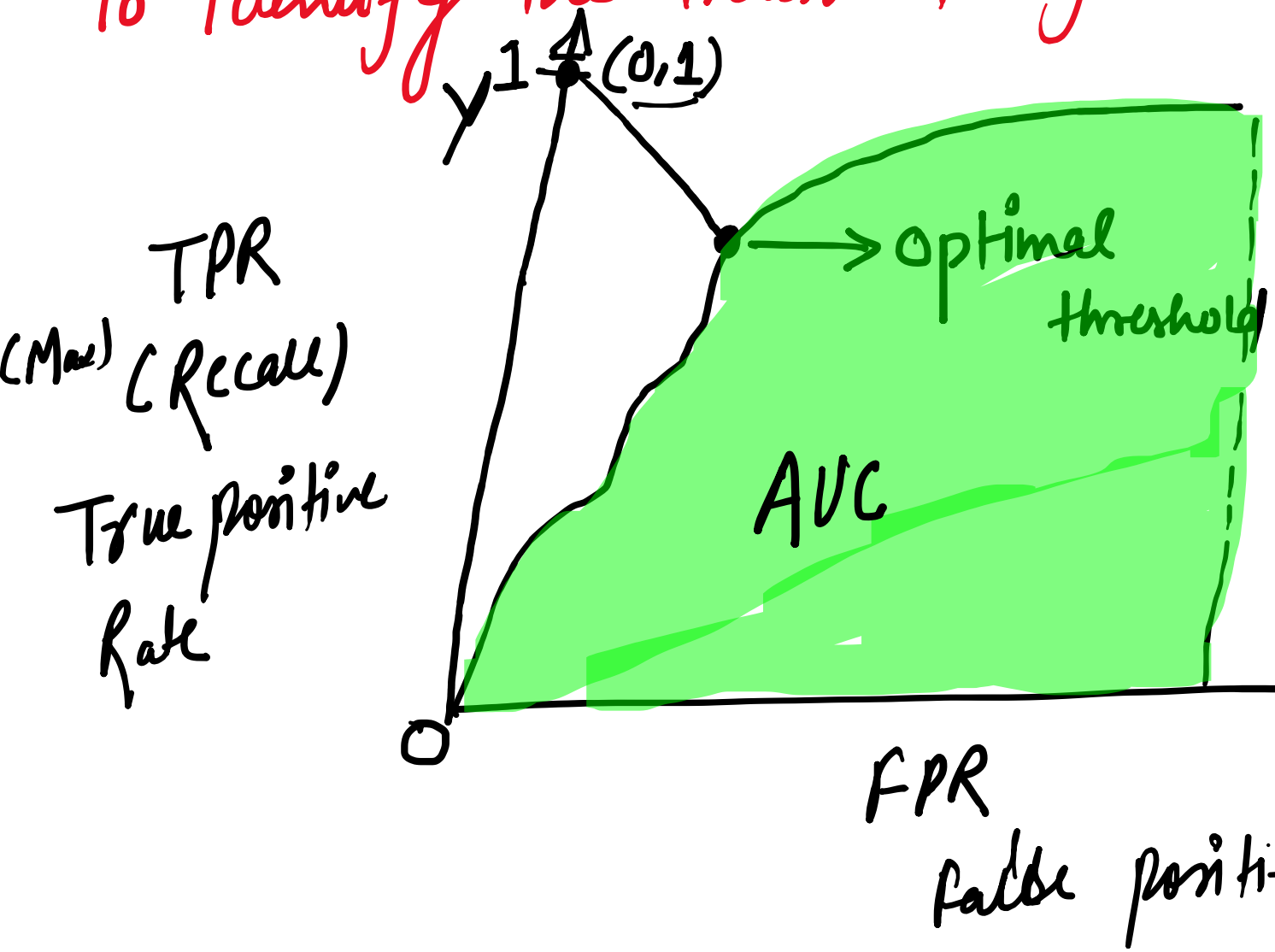
⟶ 0

0.45 ⟶ 1

⟶ 0

→ {ROC :- Receiver Operating Characteristics Curve [ROC-AUC

To identify the threshold for Classification.



	TPR	FPR
0.1 →	–	–
0.2	–	–
0.3	–	–
0.4	–	–
	–	

AUC = Area Under curve

0.1 0.2 0.3 — — — — 0.9 1

0.1 0.15 0.2 — — — — — 0.95 1

0.01 .02 .03 0.4 — — — — 0.98 0.99 1

Doubts:

upGrad