26 Sep
2023

Tue

Classification Model

ML

Supervised   Unsupervised

Regression   Classification

1) Predict   1) Predict
   the Car   Pass / Fail

2) Predict   2) Day of thy exh.
   Light Output   walk

3) Predict HDD   3) Predict the
                          Speed of
                          Plants / a

Continuous
Labu / output   Discret / finite
                      Labu / output

DS Project

1) Problem Statement
2) Data Extraction / collection ⎤
3) Data transformation / clean ⎦ EDA

↗ ⟶ 4) Choose & build model

5) Model Evaluation
6) Model deployment.

7) Model Govern
   - Model monitoring
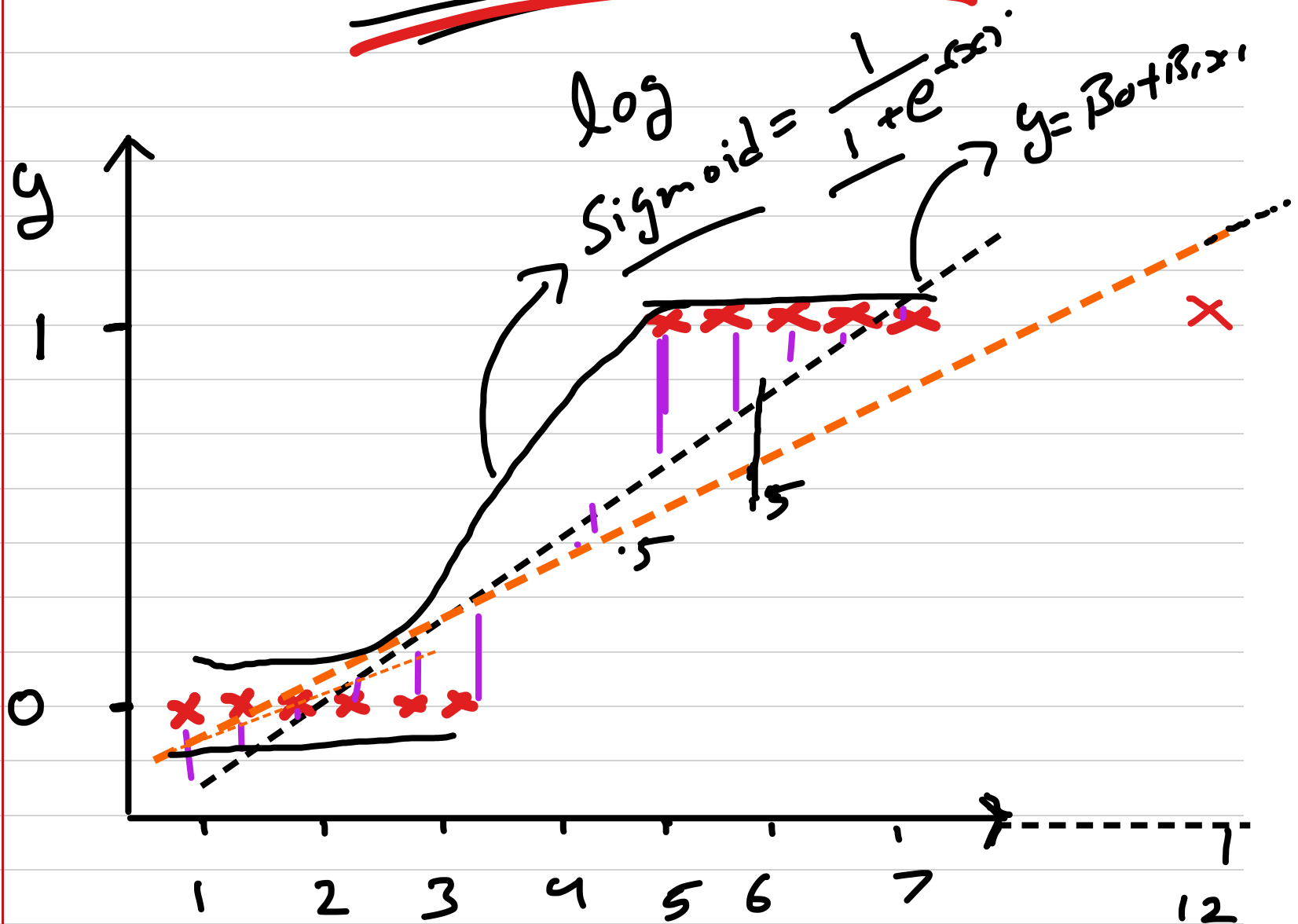   - model hyperParametisation
   - model version'z

   MLOPS   ⟶ MLFlow Free

   - Class tolk

Logistic Regression

Q) Based on Studing
   nours Predict
   the Student would
   ## Pass or Fail

$$\log \quad \text{Sigmoid} = \frac{1}{1+e^{-\beta x_i}} \qquad y = \beta_0 + \beta_1 x_i$$

Problem
$\begin{cases} \circ \text{ Outlier } \checkmark \\ \circ < 1 \text{ on } > 0 \checkmark \end{cases}$

$$y = \beta_0 + \beta_1 x_1$$

$$\parallel$$

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

$$= \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

$$= y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$

Sigmoid

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

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

Sigmoid is a number

0.5
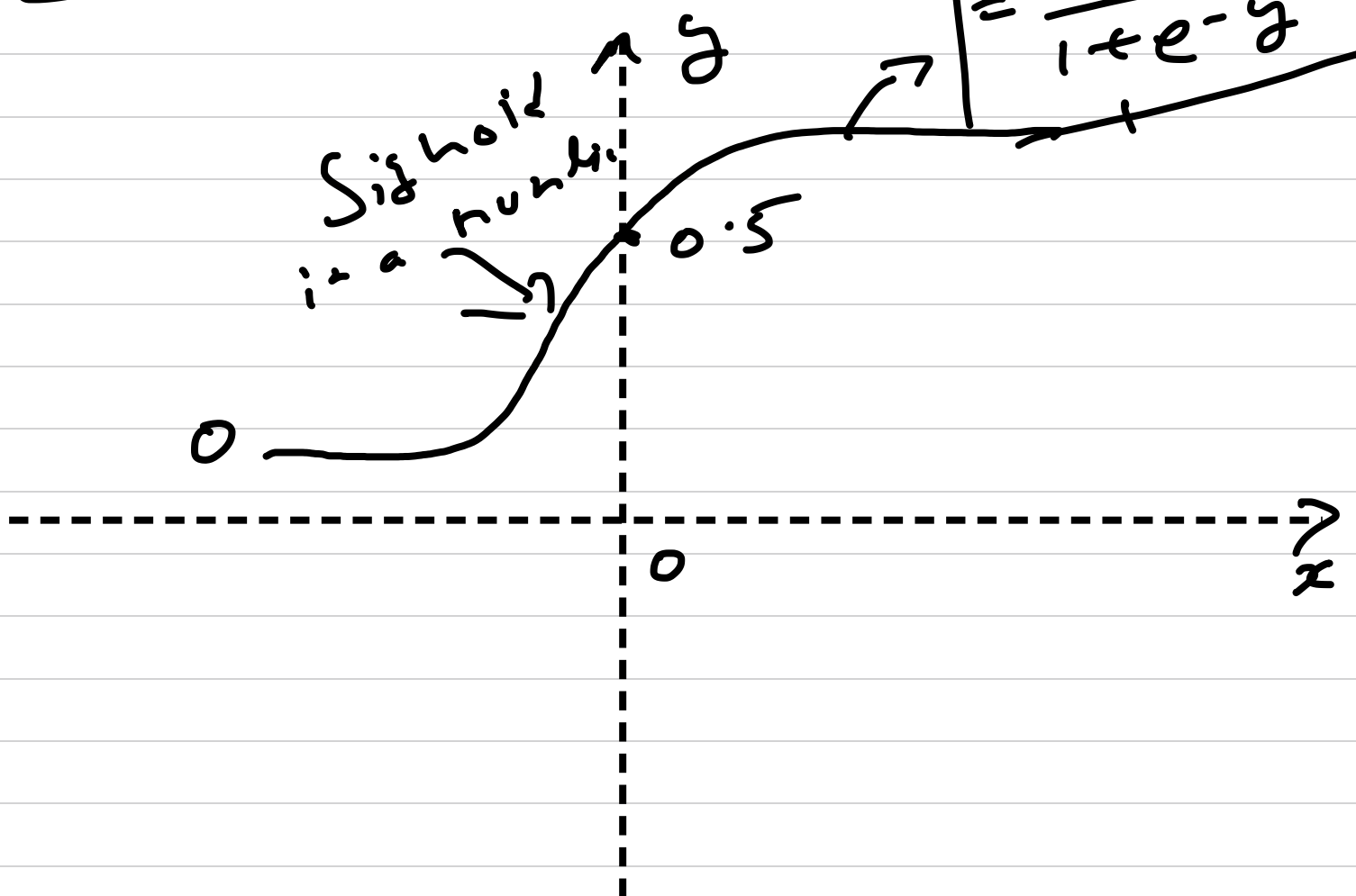
Revision

- ML tree

- Supervised - Classification

- Proble - Y(output)
$\rightarrow$
  Discrete
  Finite

- Model
  $\rightarrow$

  Logistic Regression
  $\rightarrow$

  $\rightarrow$
  Sigmoid $\rightarrow$ LR.

  $$\hat{y} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}}$$

→ Build the model

→ Evaluate the model.

    ↳ Metrics/Parameter

      o Accuracy

        Score

Classification metrics

$y = (1, 0)$

Actual Values

	Positive (1)	Negative (0)	
Positive (1)	TP	FP Type I	
Negative (0)	FN Type II	TN	

Predicted Values

Confusion Matrix

Q) Based the number of
hours study before exam

X	y	$\hat{y}$
5	1	1
6	1	1
3	0	1 → FP

T.P   FP   T I

T.N

FN

T II

↗ Limit the
Type II Error
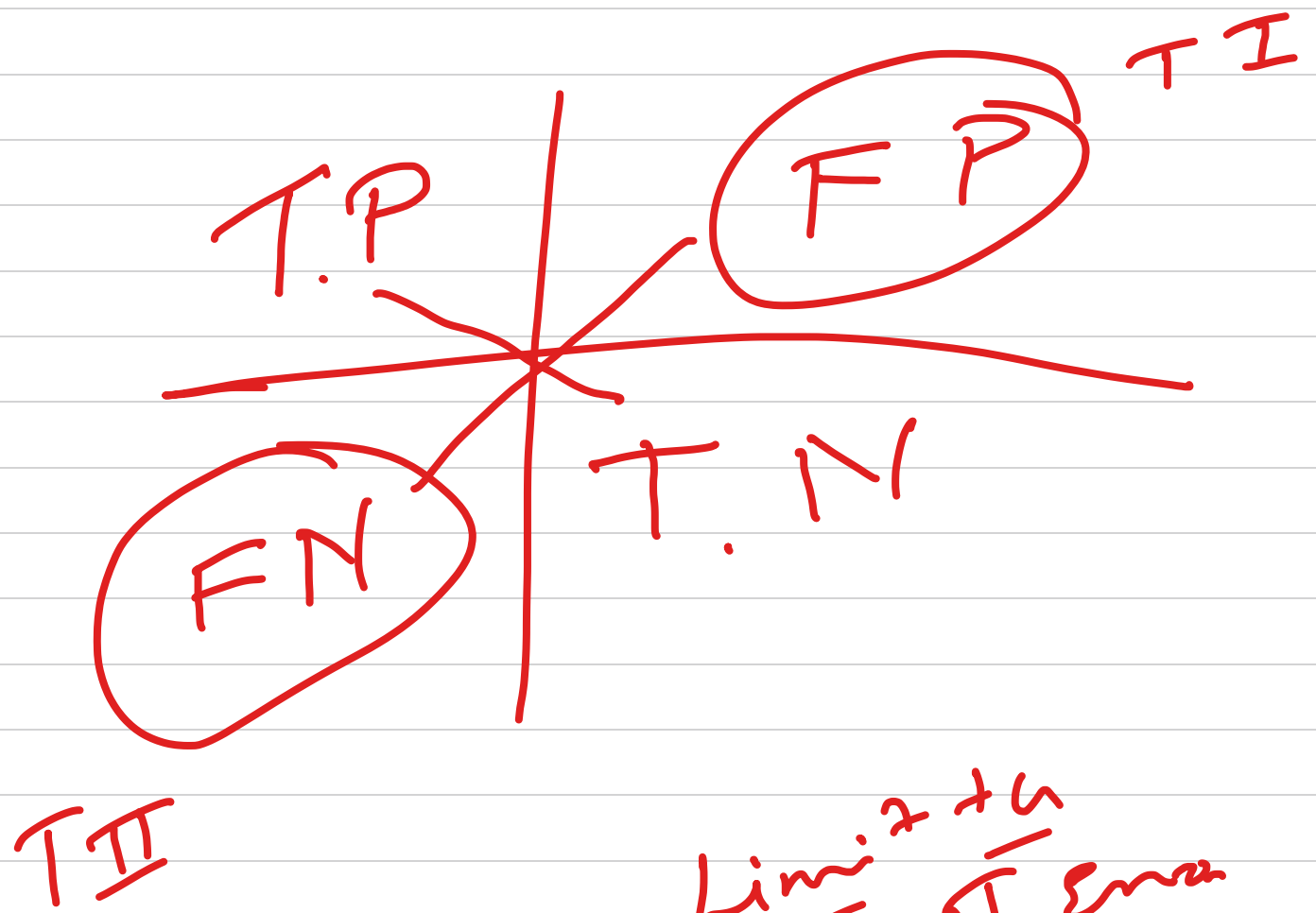
S1 — Testing a Patient win Covid ✓

S2 — CC fraud flagging ↗ Limit
                           ✓ FN

S3 — Judge decidin whather ✓
        Someon is guilty of a Crime

                              ↳ Limit FP

Accuracy Score

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

$$= AC$$

$$
\begin{array}{c|c}
 & \overset{40}{FP} \cdot 1.0 \\
\overset{100}{TP} & \\
\hline
FN & TN \; 40 \\
1.0 &
\end{array}
$$

Precision

$$= \frac{TP}{TP + FP} = \frac{40}{40 + 19} = \frac{4}{5} = 0.8$$

Precision is useful in the case where False Positives a high concern the False Negative

EX:- Spam →Imp Email →Spam

Spam

Recall (sensitivity)

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

Useful in the case when FN is of high concern than FP

Ex:- Very dangerous good item gets a FSC

- Pregnancy :-
- Airport - Security Screen.
- A fraud transaction gets unnoticed.

F1 Scons

Harmonic mean

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

AUC Curve { Graphical }

True Positi

False Positive