Logistic Regression

* Classification $\longrightarrow$ binary classification
(two categories)

* Core is sigmoid fn $\longrightarrow$

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

Range from $\underline{0 \text{ to } 1}$
↓
Probability

General

We predict new point
using sigmoid fn.

$O/P > 0.5 \Rightarrow$ Class A

$O/P < 0.5 \Rightarrow$ Class B

$O/P = 0.5 \Rightarrow$ Uncertain
(Not classifiable)
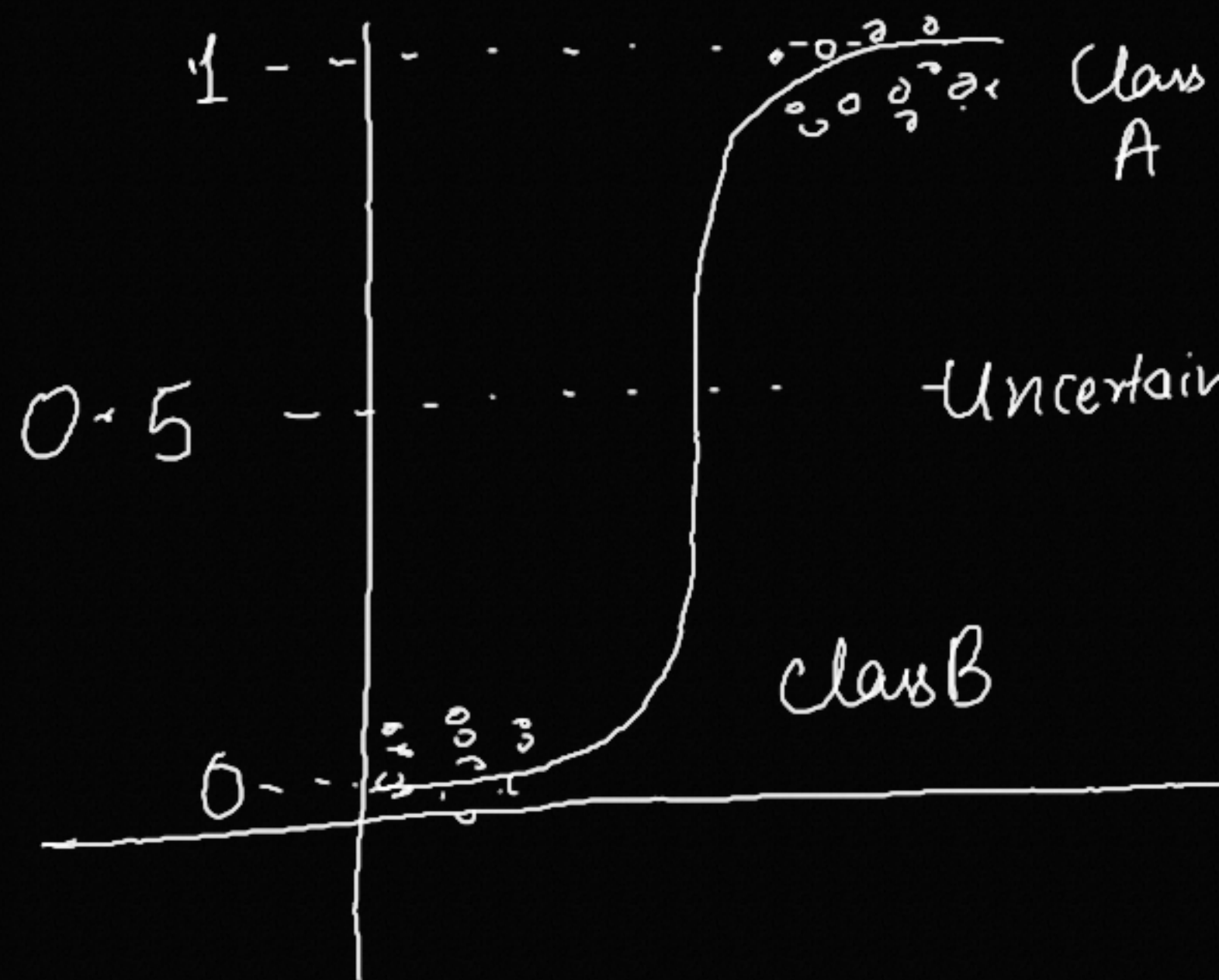
Mathematical

Hours of study (X_1)	Practice set taken (X_2)	Pass/fail (Y)
2	1	0
4	1	0
6	2	1
8	2	1

• Calculate Conditional Probability using sigmoid fn.

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

= Pars

$$P(y=1 \mid X_1, X_2) = \frac{1}{1+e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2)}}$$

Suppose we get,

$$\beta_0 = -5 \qquad\qquad \beta_2 = 2$$
$$\beta_1 = 1$$

*

$$\boxed{\begin{aligned} \text{Model} &= \frac{1}{1 + e^{-(-5 + 1.x_1 + 2.x_2)}} \\ (Y) & \end{aligned}}$$

* <u>Prediction</u> $x_1 = 5$, $x_2 = 2$

$$y = \frac{1}{1 + e^{-(-5 + 1.5 + 2.2)}}$$

$$y = \frac{1}{1 + e^{-4}} = \frac{1}{1.018} = \underline{0.98} \quad (\underline{Pars})$$

$O/P > 0.5 \rightarrow$ Pass

$O/P < 0.5 \rightarrow$ fail

$O/P = 0.5 \rightarrow$ Uncertain

- Advantages

① Simple & easy to understand

② Less computational power

③ Well with linear data

④ You can do feature selection using L1 penalty
(Lasso)

is
You can apply any type of regularization ,

- **Disadvantage**

① Binary classification only.

② Not good with non-linear data

③ Sensitive to outliers

④ Not good with multicollinearity in data.