Logistic Regression:-

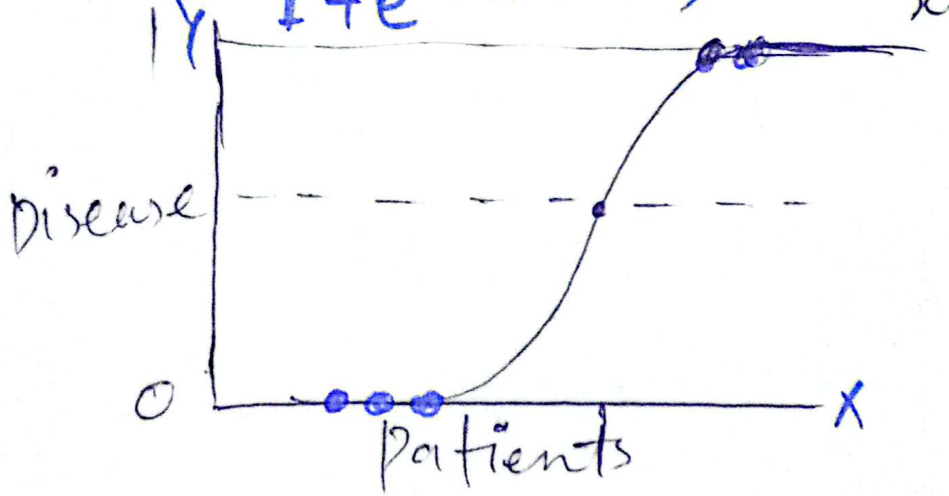
- Supervised classification model
- dependent variable is categorical and binary (0 or 1)

Patient ID	Disease
16061	0
20201	1
30101	0
70000	0
10111	1

Sigmoid

$$y = \frac{1}{1+e^{-(a_0+a_1x)}}$$

a_0 = y-intercept
a_1 = Coefficient of x depend y.



Linear Regression	Logistic Regression
→ Regression model	Classification model
→ $y = a_0 + a_1 x$	$y = \frac{1}{1 + e^{-(a_0 + a_1 x)}}$
→ dependent variable in form of continuous	→ dependent variable in form of Categorial & binary
→ Predict value by an integer number	→ predict value by 1 or 0.

Classification (predicting Classes)

→ k-nearest neighbors (K-NN)

IMDb Rating	Duration	Genre	Distance
8	2.5	Action	0.78
6.2	2.6	Action	0.85
7.2	2.4	Comedy	0.72
8.2	2.5	Comedy	0.78

0.25
0.36
0.16
0.25

IMDb = 7.4
Duration = 2.00

Steps

→ Calculate distance by Eulcidean distance.

$$D = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2}$$