Logistic Regression →

Logistic regression is a Supervised classification algorithm that only takes discrete value as input.

Logistic Regression is used when the dependent variable is Categorical.

Eg:

Suppose we have a website, and How long is the user active on that website and whether ads on that website clicked or not.

Time (x)	clicked on Ads (y)
68.95	No
80.23	NO
69.45	NO
74.15	NO
50.0	yes
55.5	yes
80.0	NO
70.5	NO

NO → Categorical data

True / false

Positive / Negative

Pass / fail

So in Categorical data we will use logistic regression.

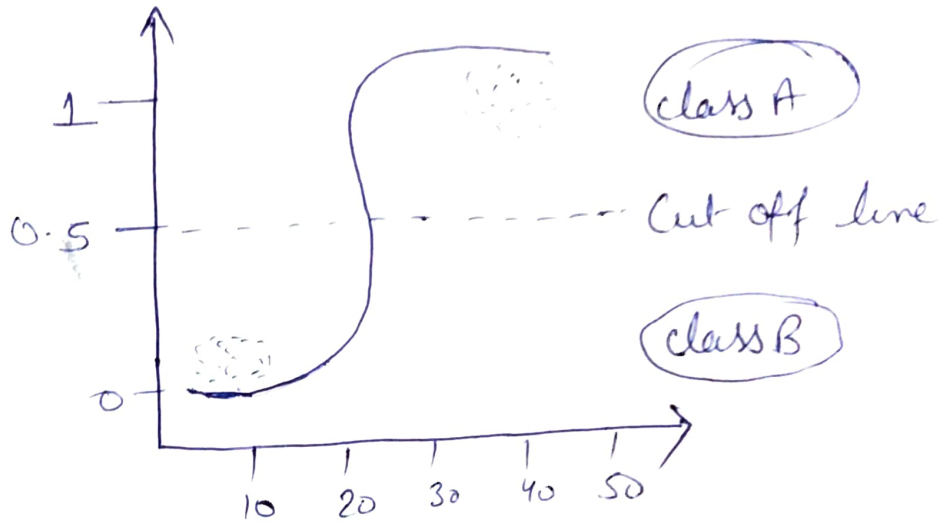
output → $Y = \frac{1}{1+e^{-x}}$ → independent variable.

↓

Sigmoid function

$e = 2.718$ → Euler's number
Euler's Constant

Sigmoid function simply trying to convert the independent variable in to a Expression of Probability that ranges between 0 & 1 with respect to the dependent variable.



Some data Points are available above the cut off line and some are below the cut-off line. these are divided in to <u>class A</u> & <u>class B</u>

Probability → In y axis → between 0 & 1

0 → there is No Posibility / No Probability of occurance

1 → there is Certain Posibility of occurance.

data Set with 0.5 Probability that are positioned on Cut off line and they have No class.

Precisely

When data set that are on the cut off line known as unclassifible data set.

unclassifible is in rare case.

Application of logistic regression

* fraud detection
* disease diagnosis
* Emergency detection
* Spam , No Spam (Mail)

When you are using logistic regression Some Points are important.

1) Data set that you considered it Should be free of Missing value.

2) How Much data points are available in data set for using logistic regression.
   * 30 to 50 data Point for each output.
   * for binary logistic regression.
     Binary – Means — 2 class — 2 output variable

30 to 50 double → 50

60 — 100 data Points . total in this data set.