Subject Classification $\sum_{i=1} y_i (w^T x + b)$ max.

fails at outliers

$y_i (w^T x + b) \geq 0$

for max point

Convert this dist function to Probabi use sigmoid function

$$\frac{1}{1+e^{-(w^T x + b)}}$$

Logistic Regression

1. It is supervised learning Algo.
2. Mainly used in binary classification

$y = f(x)$

- y — dependent Variable (Categorized data)
- f — Sigmoid fun
- x — independent Variable (i/p Value)

Classification Predictive Model

Train ← Training Data

odd → Probabi Success / Prob failure

Model ← Test

↓

Predicted o/p $(\hat{y})$ → [Categorical Data T/F]

↳ Evaluation (Error Calculation)

logodds = $w^T x + b$

$\beta_0 + \beta_1 x_1 + \beta_2$ Error = $(y - \hat{y})$

- low Bias & low variance (Good Model)
- low Bias & High variance (Overfitting)
- High Bias & High variance (Underfitting)

Overfitting, Underfitting → Retraining

Predictive Modelling Metrics

The following Performance Metrics are used to Evaluate the Classification Predictive Model.

- Precision
- Recall
- Accuracy
- Error-Rate
- F1 score
- Sensitivity
- Specificity

To Cal. above measures, following are required
i.e ① True Positive (TP)
② False Positive (FP)
③ True Positive (TN)
④ False Negative (FN)

eg// i/p sample → Classifier model → Predicted o/p {Covid Present (or) NOT} Binary Class set → (+ve) Covid(+) / (-ve) Covid(-ve)

↑ Train with dataset (samples) Covid Patients samples

True Positive (TP) :- i/p sample → Classifier model → Predicted o/p Covid Present
{sample with actual o/p Covid Present}

False Positive (FP) :- i/p sample → Classifier model → Predicted o/p Covide present
(sample with actual o/p Covide NO)

True ~~Positive~~ Negative (TN) :- i/p sample → Classifier model → Predicted o/p Predict (no)
(sample with actual o/p ~~False~~)

False Negative (FN) i/p sample → Classifier model → Predicted o/p (NO)
Sample with actual o/p (True)

Confusion Matrix

→ Describes the TP, FP, TN & FN values based on the given training Dataset & the associated Predicted O/P

→ Status of Confusion Matrix Measure Confusion Matrix.

→ Confusion Matrix size depends on training dataset class labels

ie: (1) Binary Class training dataset, Confusion matrix is (2×2) Matrix

(2) Multi Class training Dataset, Confusion Matrix is (n×n) Matrix

n : # of class lables

: Confusion Matrix (2×2) : #2 class lables

$FPR = \frac{FP}{\text{actual false}}$ $\quad \frac{FP}{FP+TN}$

$TPR = \frac{TP}{TP+FN}$ (Recall) Sensitive

	Sample actual Class label is +ve	Sample actual class label is -ve	
Sample Predicted class label +ve	+ve TP +ve	-ve FP +ve	Σ total no of sample Predicted +ve
Sample Predicted Class lable -ve	+ve FN -ve	-ve TN -ve	Σ total no of sample Predicted negative
	Σ total no of samples actual class +ve	Σ total sample actual class Negative	

$\frac{TP}{TP+FP}$

Precision : $\frac{\text{no of samples predicted true \& actual true}}{\text{total no of samples predicted true}}$

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

eg Test Model is trained with 20 samples & also tested with 20 samples. 16 samples as positive sample 8 samples are true positive

$\Rightarrow$ Precision $= \frac{TP}{TP+FP}$ (total no of sample predicted True) $= \frac{8}{16}$

eg Model trained with 20 samples 24 samples actually positive → Model evaluated training dataset 20 positive samples in that 16 samples are True positive

Precision $= \frac{\text{no of samples actually true \& predicted true (TP)}}{\text{\# of samples predicted positive (TP+FP)}} = \frac{16}{20} = 0.8$

Conclusion : Precision is measure of how many samples are correctly predicted as a true class among total # of samples predicted as a true class

False Positive ... precision ... rate

Accuracy: Ratio b/w the no. of samples correctly predicted & total # of samples evaluated on model

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

Error Rate: $\text{Error Rate} = \boxed{1 - \text{Accuracy}}$

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

F1 Score:

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

Sensitivity (Recall) $= \frac{TP}{TP + FN}$ → focused on true prediction

(TPR) True Positive Rate

Specificity = Ratio b/w the number of samples predicted as a true Negative & the total # of sample actually class lable "-ve"

$$= \frac{TN}{TN + FP} \Rightarrow \text{Focused on Negative Class prediction}$$

$$1 - \frac{FP}{\text{Act neg}} = 1 - FPR$$

Logistic Regression

- Supervised ~~ka~~ learning Algo.
- Classification Prediction Modelling Algo.
- Mainly Used for Binary Classification
- Produces O/P in a Categorized format based on the linear Representation of a data points.
- In both of the linear & logistic regression Algo. straight line eq. is used to train the Model But linear Regression produces continuous value as an O/P & logistic regression produces the categorized data as an O/P.

Model is $$P = \frac{1}{1+e^{-z}}$$ $e = 2.71$ {natural log}

P: Probability of a dependent Variable (y) which belong to the specific class

Z: linear Combination of an Independent Variable (x)
{Becoz of this Model, to name as regression}

If we used single linear Regression in the logistic Regression then Model is,

$$P = \frac{1}{1+e^{-(mx+c)}}$$

only one Independent Variable (x) present

If we use Multiple linear Regression in the logistic Regression then Model is

$$P = \frac{1}{1+e^{-(m_1x_1+m_2x_2 \ldots m_nx_n+c)}}$$

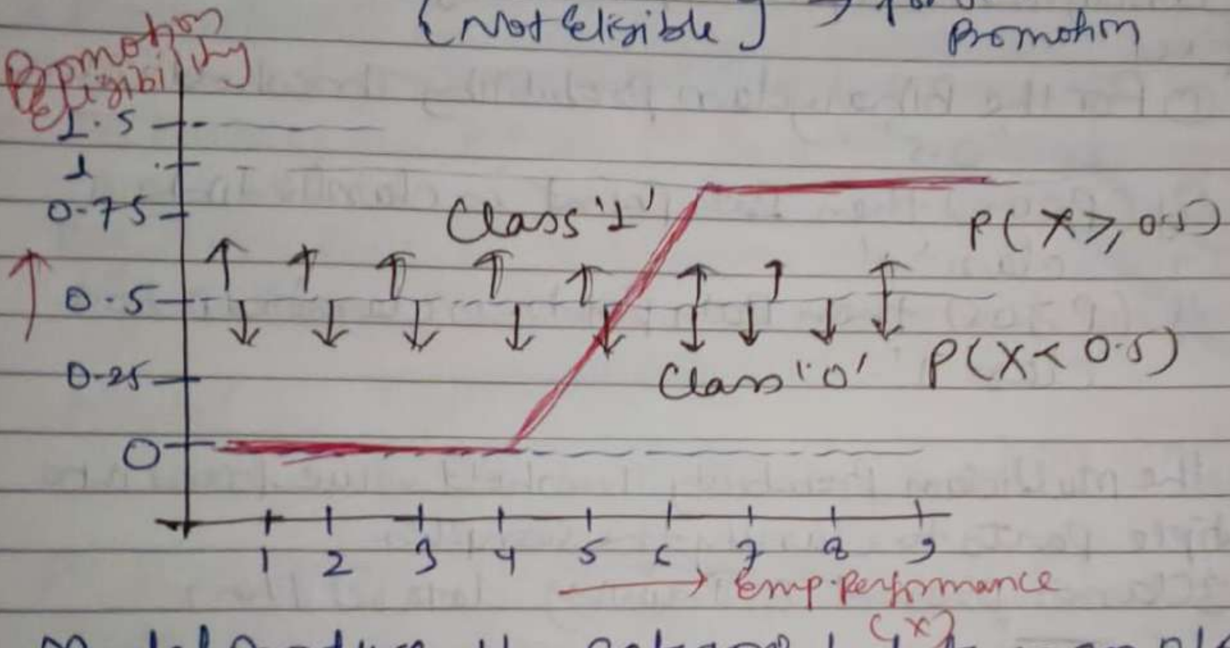
~~Two or more~~ Independent Variable

Ex

logistic Regression

- Company (org^n) wants to give the promotion to the employee based on their performance

{ eligible, Not eligible } ⇒ for the promotion



Model produce the categorical data as an O/P So it is a classification predictive model.

- If $y = P(x)$ then it covers data range [0 to 1]
- If we take odds of success then it covers more data

ie $\text{odds}(O) = \dfrac{P(x)}{1-P(x)}$ → Event occur / → Not Event occur

If $P(x) = 0$ then $\text{odds}(O) = 0$

$P(x) = 1$ then $\text{odd}(O) = \infty$

So, odds of success covers the data range $(0, \infty)$