* if column has missing ~~mix~~, no data drop it
* if colum has no variance, little variance, drop it.

All these is ackened by RFE (Recursive feature Elimination

→ from sklearn.feature_selection import RFE

— x —

* Logistic Regression *

Its a classification technique, in supervised ML.

Ex: 1) Classify given $\underbrace{\text{e-mail}}_{\text{input}}$ is $\underbrace{\text{spam or ham}}_{y_i \to \text{targets}}$

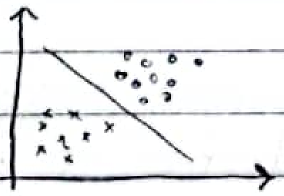
2) Given some handwritten digits in photos. tell what character is written

Here $x_i \to$ photo

$y_i \to \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

3) Given some $\underbrace{\text{credit card data}}_{x_i \text{ input}}$, Detect $\underbrace{\text{fraud}}_{y_i \to \text{fraud, nofraud}}$

4) ~~Given~~



→ Logistic regression classification graphically..

2 catagories → binary classification

3 or more Catagories → multi-class classification

Task: find line that best ~~passes~~ separates points and also classify

Eq^n of a plane in 3D

$w_1 x_1 + w_2 x_2 + w_3 x_3 + w_0 = 0$

2D → $w_1 x_1 + w_2 x_2 + w_0 = 0$ $\{ y = mx + c$

$w_2 x_2 = -w_1 x_1 - w_0$

$x_2 = -\frac{w_1}{w_2} - \frac{w_0}{w_2}$; so, $\frac{w_0}{w_2} \to x$ intercept

$-\frac{w_1}{w_2} \to$ slope/cofficient of x_1

DATE : __/__/__
PAGE : ______

If this line passes through (0,0) origin

$x_2 = -\frac{w_1}{w_2} x_1 - 0$

So, $-\frac{w_0}{w_2} = c = 0$ $\because$ $w_2 \neq 0$

$\therefore -w_0 = 0$

So, whole eq^n of line (0,0) 2D →

$w_1 x_1 + w_2 x_2 = 0 \Rightarrow \pi_2 : w^T x = 0$

3D → $w_1 x_1 + w_2 x_2 + w_3 x_3 = 0$

So,

$\pi_2 : w^T x = 0 \Rightarrow w \cdot x = 0$ dot product

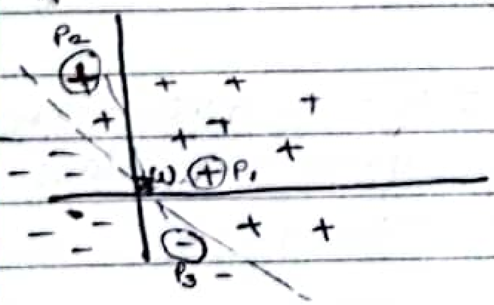
$\|w\| \, \|x\| \underbrace{\cos\theta}_{90^\circ} = 0$

Here 'w' is norm of hyperplane π_2

$\pi_3 : w^T x + w_0 = 0$

w is a norm of π_3

LR Task : Seprating π



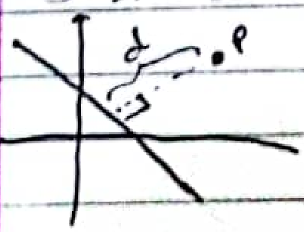
$\vec{w} \cdot \vec{P_1}$

$\|w\| \, \|P\| \cos\theta = +ve$

$\vec{w} \cdot \vec{P_2} = +ve$

$\vec{w} \cdot \vec{P_3} = (-)ve$

* Distance of a point from a line (π)



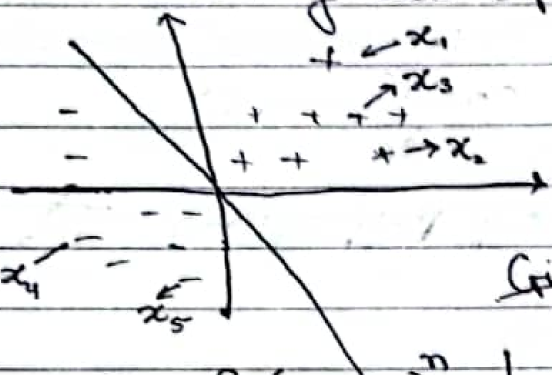
$d = \frac{w^T \cdot P}{\|w\|}$

DATE : __/__/__
PAGE : ______

Task ! To find a hyperplane (π) i.e. w, w_0, such that this π best seperates +ve points from -ve points

Assumptions in LoR

- Classes (+ve & -ve) should be perfectly linearly seperable, as possible
- Best seperation passes through origin $\boxed{w_0 = 0}$
- w is norm of best sepration π $\boxed{\|w\| = 1}$



$$d = \frac{w^T x_i}{\|w\|}$$

~~Given : $D_n = \{(x_i, y_i)\}_{i=1}^{n}$~~

Given : $D_n = \{(x_i, y_i)\}_{i=1}^{n} \mid x_i \in \mathbb{R}^2, \; y_i \in \{+ve, -ve\}$

as for assumption, $\|w\|$, length of w is 1

$d = w^T x_i$

if $d = w \cdot x_i > 0$: $y_i \rightarrow$ +ve class

if $d = w \cdot x_i < 0$: $y_i \rightarrow$ -ve class

for $x_i \rightarrow y_i$ {-ve , -ve}

so, $\{ y_i * d \}$

Case 1 if y_i = +ve ; $d > 0$

so, $y_i * d \rightarrow$ +ve

——— x ———

Evaluation of Linear regression :

→ Mean square error (MSE) good

→ ~~Coff~~ coefficient of determination (R^2) bad

↳ $R^2 = 1 - \dfrac{\text{Residual Sum of square (RSS)}}{\text{Total sum of square (TSS)}}$

i.e.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

y_i = actual output/target

$\hat{y}_i$ = predicted output (

$\bar{y}_i$ = mean of actual target

Case 1: Best LR model → RSS = 0

$$R^2 = 1 - \frac{0}{TSS} = 1$$

Case 2: Average LR model → RSS = TSS

$$R^2 = 1 - \frac{TSS}{TSS} = 1 - 1 = 0$$

Case 3 worst LR model → RSS > TSS

$$R^2 = (-)ve$$

R^2 tries to capture explained variance by own

Is not good