ML-With-Python

Day-6

किफायती
Date.....
Pages.....
Rituraj

Regression

- Dependent
- Independent

Independent variable ⇒ Only number of variable

Tea Shop

price of tea → price of milk, price of shop, price of tea power

price of tea ⇒ Dependent variable (only one variable)

when Data is

Numerical ⇒ Regression (linear Regression, Knn, rej, dtr, abr, gbr, xgb)

categorical ⇒ classification (logistic, nb, Knn, dtc, abc, rfc, xgb)

Ex.

IDV

(1) Based upon the number of hours of Study, predict the marks scored by a student.

DV → number (Regression)

IDV

(2) Based upon the number of hours of Study, predict the result of a student.

DV → (Pass/fail) → classification

Titanic -> whether the passenger is survived or not -> classification
Hotel booking -> is cancelled (0 or 1) -> 0 not cancel / cancel -> classification
auto mpg -> relationship b/w mpg and other attributes -> DV (mpg) - Regression
Concrete -> Strength of a concrete -> Numeric -> Regression

Regression

x (idv)	y (dv)
1	5
2	10
3	15
4	20
5	25

x	y
1	5
2	10.7
3	15.2
4	20.1
5	18

x	2	3	y
1	2	1	10
2	2.1	3	10.5

$y = 5x$
$x = 10$, $y = 50$

Supervised learning
labelled data

train -> 70 to 85%.
test -> remaining

Training 80-1.

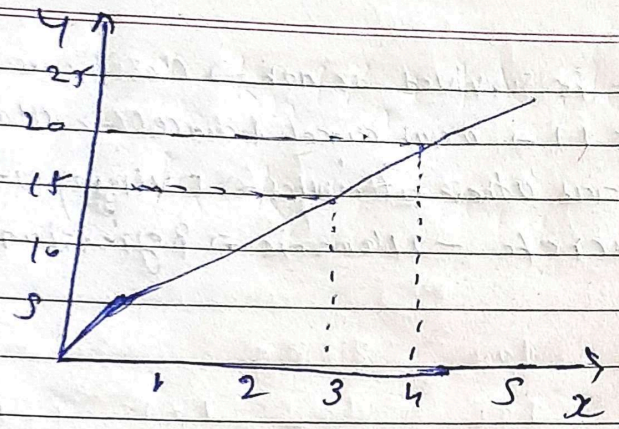
$\hat{y} = 3x$	$(y - \hat{y})^2$
3	4
6	16
9	36
12	64
15	100
	220/5
	44

$\hat{y} = 2.92 \quad (4 - \hat{y})^2 \quad \hat{y} = 3.42$

$\rightarrow$ 48 39

$\hat{y} = 3.42$

किफायती
Date
Pages
Rituraj



$y = mx + c$

slope $= \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$

$\frac{20 - 15}{4 - 3} = 5$

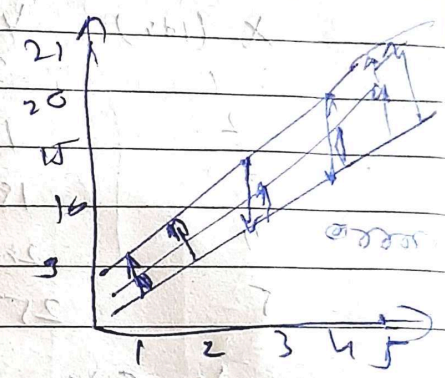
$c = 0$

$y = 5x + 0$ $\quad y = 5x$

$\hat{y} = 5x$

Infinite ~~to~~

Best fit line $\begin{cases} \text{error} = 0 \\ \text{error} = \text{min} \end{cases}$



Step ⇒

1. Understand the problem statement
2. Analytical approch (regression or classification)
3. Data Collection
4. Data understanding
5. Data preparation (Data cleaning)
6. Modeling
7. evaluation
8. ~~Dep~~ Deployment