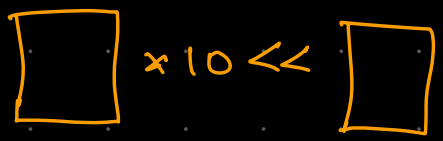
Welcome Everyone

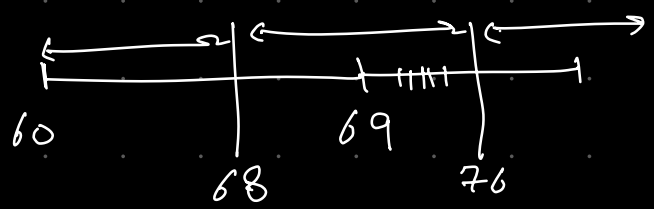
1. Naive Bayes
2. KNN
3. DT - RF - Boosting
4. SVM
5. PCA
6. Unsupervised



Naive Bayes

- Bernaulli — binary
- Multinomial — multinomial — text
- Gaussian

Virat



We assume that our independent variable $(x_i / y = y_j)$ it is having a probability distribution

$\Rightarrow$ Gaussian probability distribution

Normal Distribution Formula

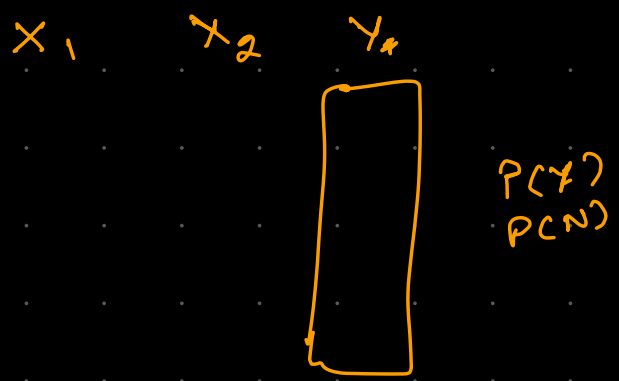
The probability density function of normal or gaussian distribution is given by;

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

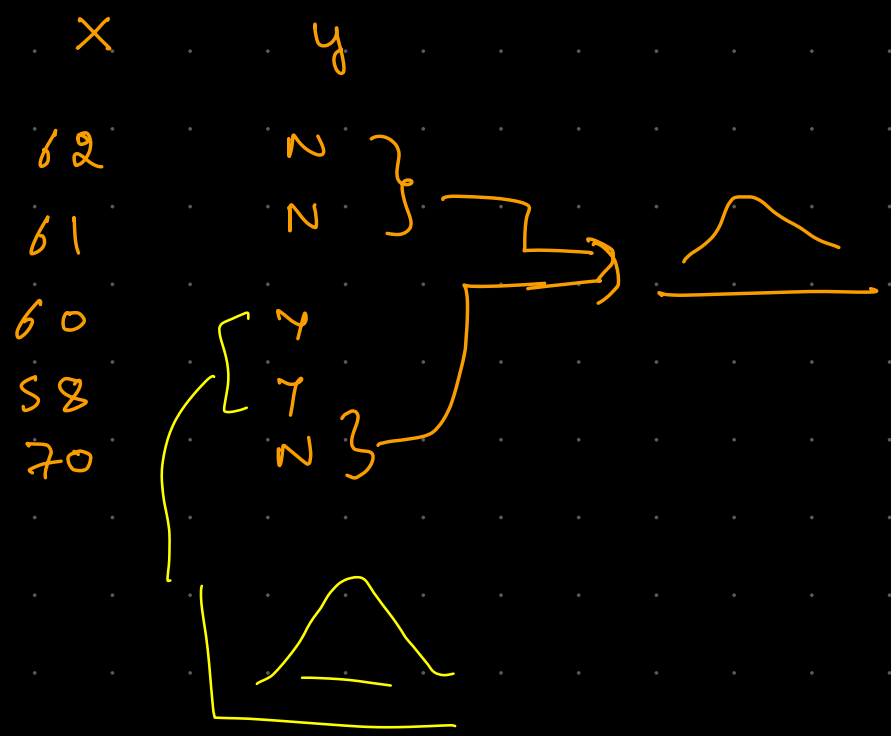
Where,

- x is the variable
- μ is the mean
- σ is the standard deviation

So gaussian eqn/probability distribution :-



We find mean and variance for a particular independent variable and our class.

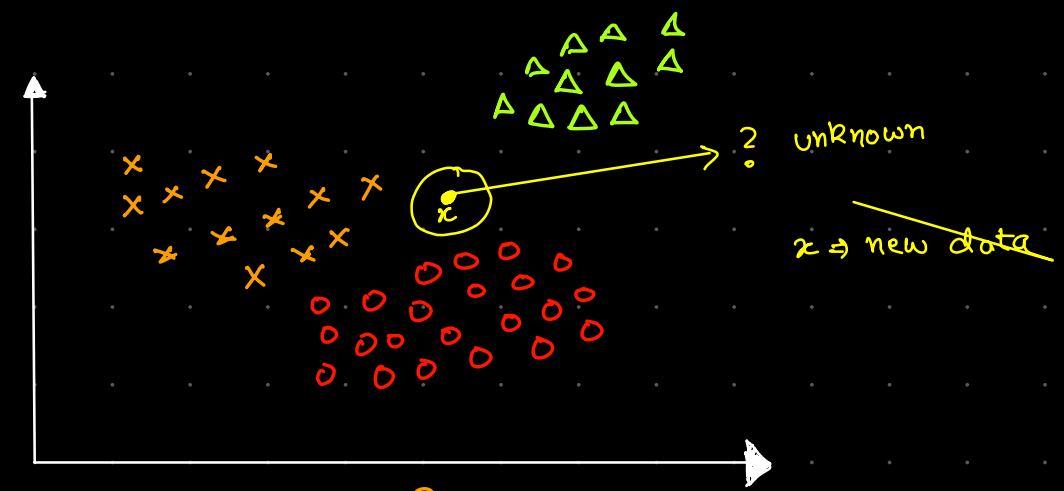


$$P\left(\frac{\text{setosa}}{x_1=4.9,\ x_2=2.8,\ x_3=4.7,\ x_4=1.2}\right)$$

$$= P(\text{setosa}) * P\left(\frac{4.9\ 1.2}{\text{setosa}}\right)$$

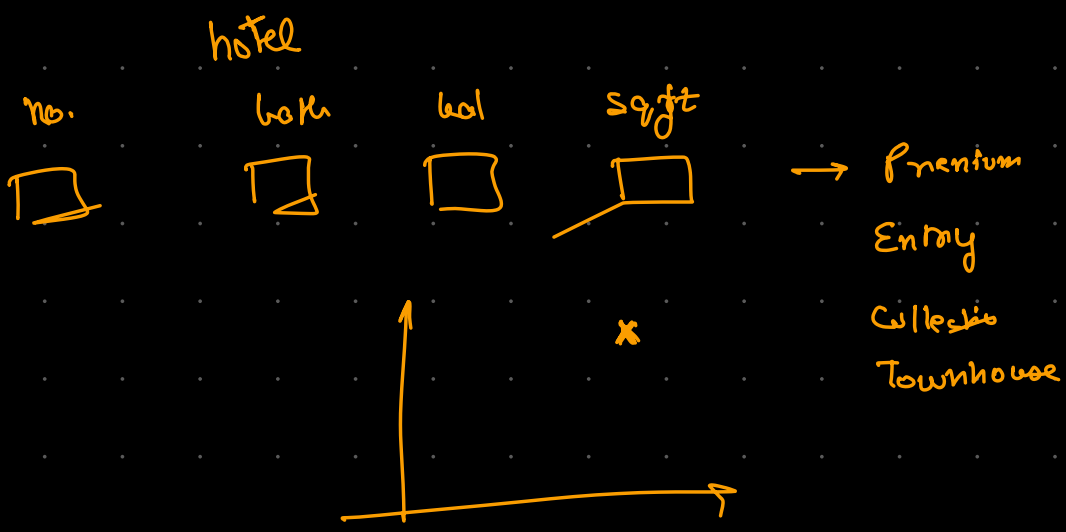
Gaussian distribution

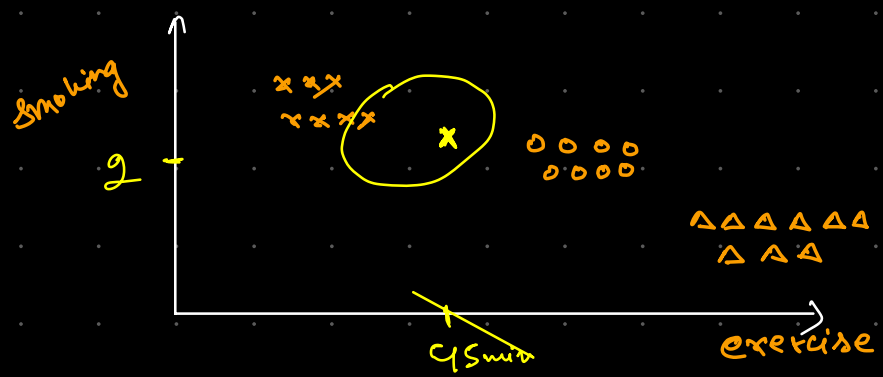
KNN (K-nearest neighbour)

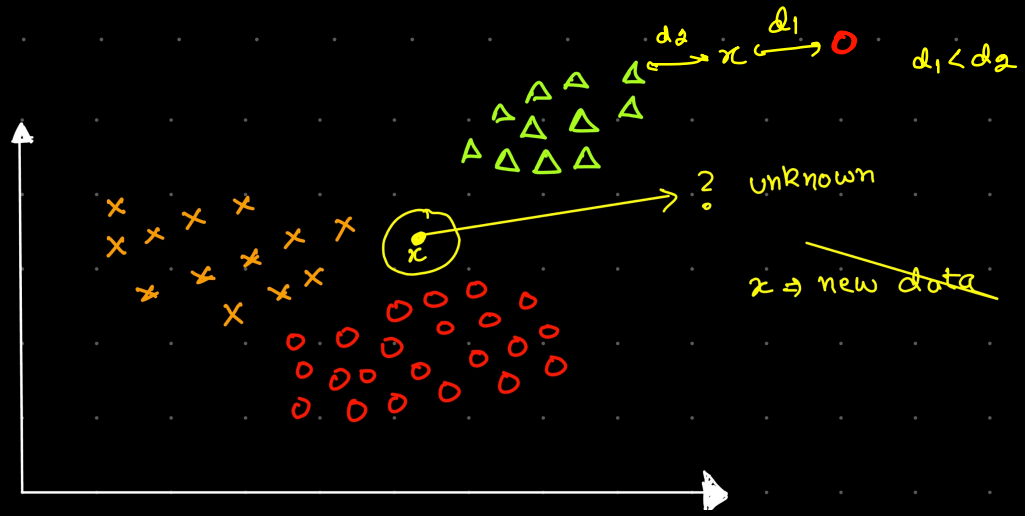


→ classification & Regression

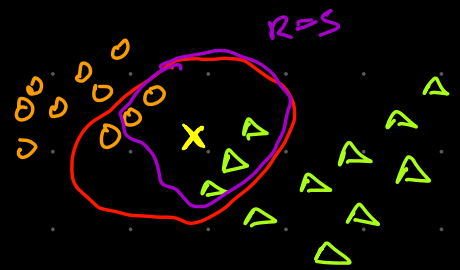
K is very important & its a hyperparameter







K will/should be an odd no.



what will x be?

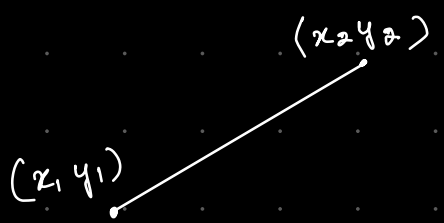
k = 6

k=1 ? what if we take K = 1 ?

2 things very very

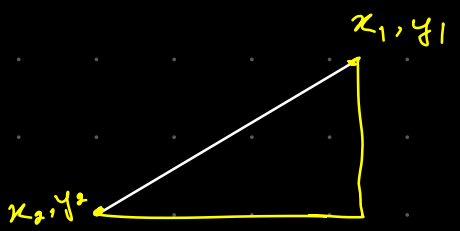
1. K value
2. Distance metric

1. euclidean



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

2. Manhattan



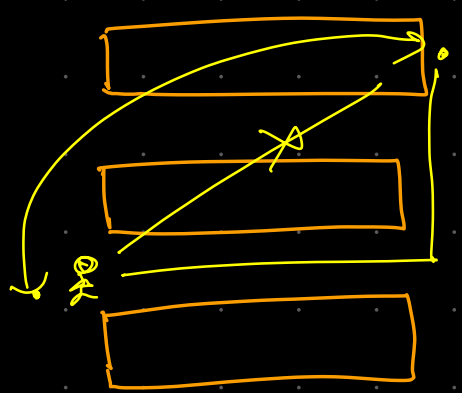
$$|(x_2 - x_1)| + |(y_2 - y_1)|$$

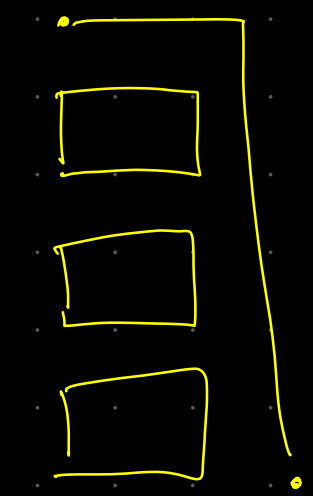
Sum of absolute differences

x_1 y_1 z_1

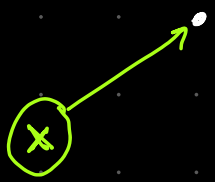
x_2 y_2 z_2

$$|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|$$





$\left(\frac{1}{dist}\right)$



Feature Scaling is very very important in KNN.

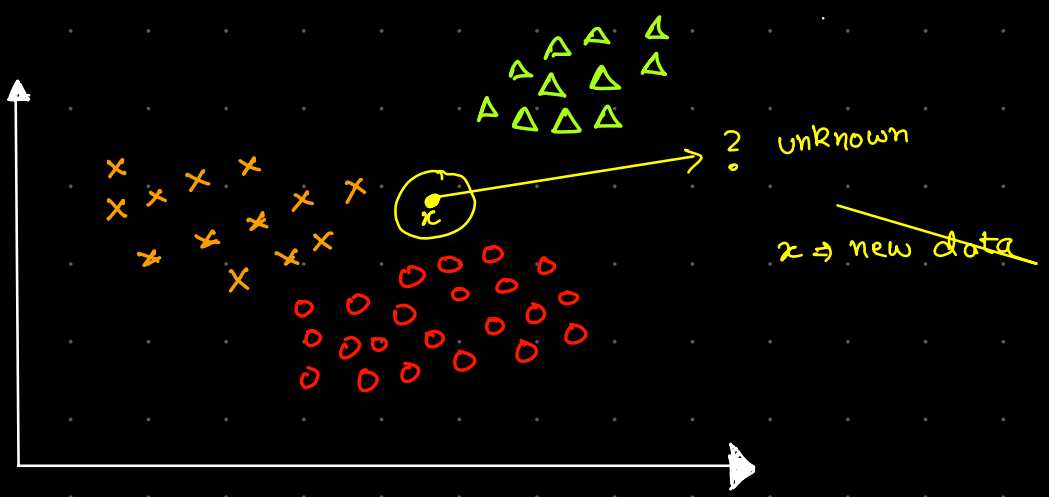
100
200
300

100

10
11
12

1

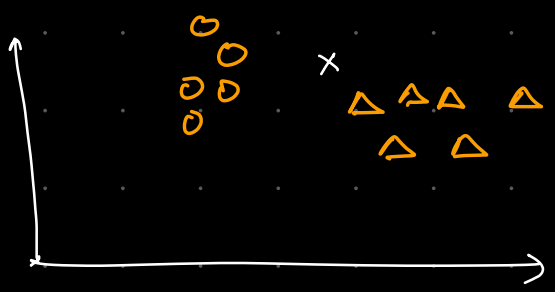
Regression.



We are gonna take avg of all the points.

MSE

How to select K value ?



K=1

K=3

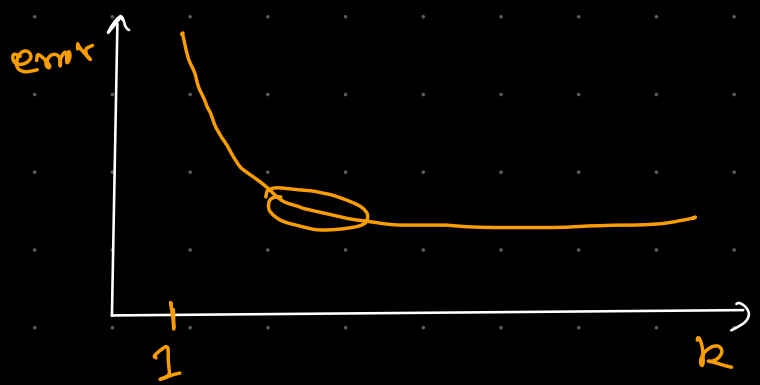
K=5

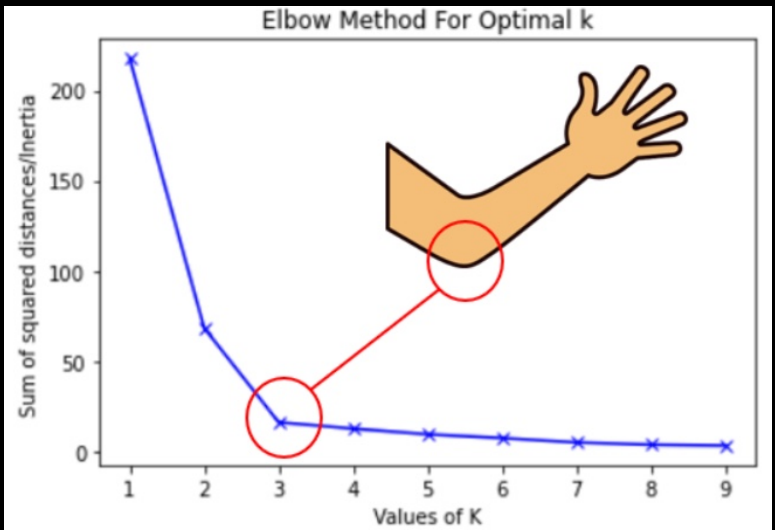
Check accuracy.

//

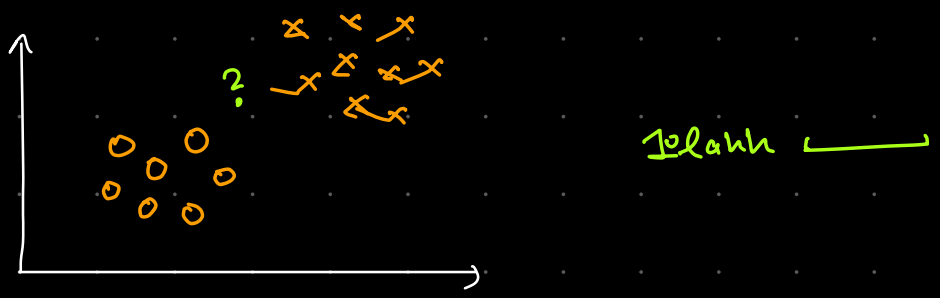
//

K= 104 ⇒ Y ← 100Y 50N





Line plot between K and inertia



(+)

1. Easy to understand / implement
2. Few ~~parameters~~ hyperparameters
3. Adapts easily.

(−)

1. Diff. to scale (lazy algorithm)
2. Suffers from curse of multidimensionality
3. Prone to overfitting.

Decision Tree

multiple if-else statement

① Classification ② Regression

Nested if-else clause

if (study hr > 5)

elif (sh > 10

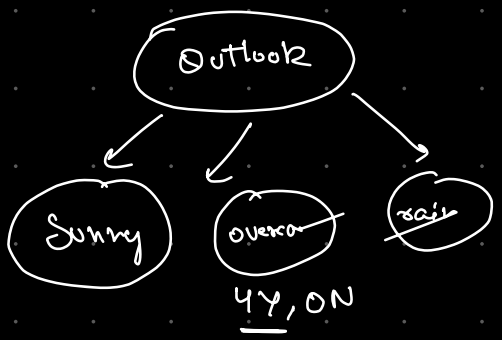
elif (____ < 3)

elif (w____ < 6)
(sleep < ___)

If we have above condⁿ properly made, then we can predict

if a student will pass or not.

Play Tennis



2 technique

1. ID3
2. CART