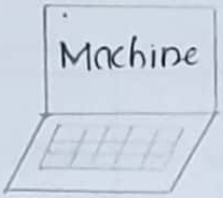
Introduction to Machine learning :-

Height	Weight	Gender
180	90	Male
150	45	Female
175	85	Male
160	50	Female
155	50	Female



→ lookup / search → to do Task

Test

H/cm	W/kg.	G
180	90	Male
150	50	female
156	50	Female

By Range [Expert-system]

Let us Consider we have a data set of height and weight in cm, kgs respectedly for those dimentionality we easily predict the Gender of that data point in test phase because we have data Information of Gender with. r.to given height & Column. So from this in-future also we easily predict Gender. by naturally we do this process but In Machine learning process how we feed & Intellectual to Our machine about this task Or predictions.

So in this Situation we need the terminologies of AI & Data Science.

What is Artificial Intelligence :- We human beings are Naturally Intelligent but a Machinery (like Computer, Car, Robot, etc) are dumb at Initilly, we feed them by Some Input Commonds (like 'task's deffrence and do operations) and get Outputs. ie generally machine mimics the Intelligence of humans is called as Artificial Intelligence.

Now a days Some machineries doing a job of human doing let us.

Scanned with CamScanner

Consider Tesla cars this are self drive Cars this fallows road Signals & System Commands according to that this car drive Selfily.

even in Our mobiles lot of Operations done with ~~cost~~ ~~Cross~~ regarding AI Such as Spam mail detection in gmail, in games like pubg or ludo there is machine play with human beyne, text Speech notation Commonds from google search bar, recommonded System in Amazon shopping, youtube like Several tasks Complited as we human doing. So now a days in every where 'AI' involved. and do Some tasks. (or) Operations like human. So here entire this process we generally put Commonds to machine for doing Some tasks.

So In above example we detect the gender of a person by giving their inputs (H/W) So In generally height & weight of a person are positively Co related with each Other (ie when height Increases then weight also Increases their existing Some positive Correlationship blw them).

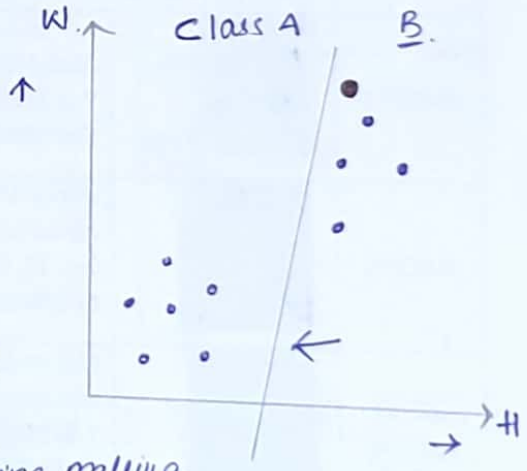
So fist of all we train Our data set to predict in future any given point is given to predict Geneder. In this process we load the data set in Excel, csv or any format and train Our machine with Some Commands. So Initially machine (ie Computer) is dumb we have to train that by Input data & Commonds. after train the data In table 2 from that we test data let in first step we give input H/W to machine predict Gender as Male & female.

Scanned with CamScanner

if we Observe '3' case in this values are Input's diffrent which we didn't defined (ie 156,50) new featured point Comes then how to Solve this task? for this we use expert System (if-else) Condition. after we Use this we get Gender of data point (ie. male/female).

By the above process we Complited Numerical tasks, tasks like in-format of (descrete / continious) for Suppose if we have Catogorical type of data. then how we do predictions (or) Operations in Such Condition.

In this we Can't Use expert System or any other we used in numerical operation. So here we Solved this problems by train a data for Suppose if we have a image of apple we train Our machine to it is a apple, if we have an Image of any animal we train Our machine to this is ~~any~~ animal So how we do this entire process. by using "Supervise data" In this process we train a machine in a Such away to predict catogorical outputs also. for given input, if we plot a Scatterplot graph on the data. We get like-



On X axis we take height & y-axis take 'W' weight according to train the data plot respected points and make point representation in data. now this graph plotted points Observe in this Simply we can plot line making. that line diffrentiate this points in two classes (A,B) as Female, male respectivly. Now if we have any new point we Simply Say as this line representation.

Scanned with CamScanner

as reference to division the two classes from this line we can predict the new data point class. So this is a categorical type of data result, how we come up this result by line making, how to we reach this line making by "machine learning" Technique Under this we used "Supervised machine learning technique" Under that classification type of problem

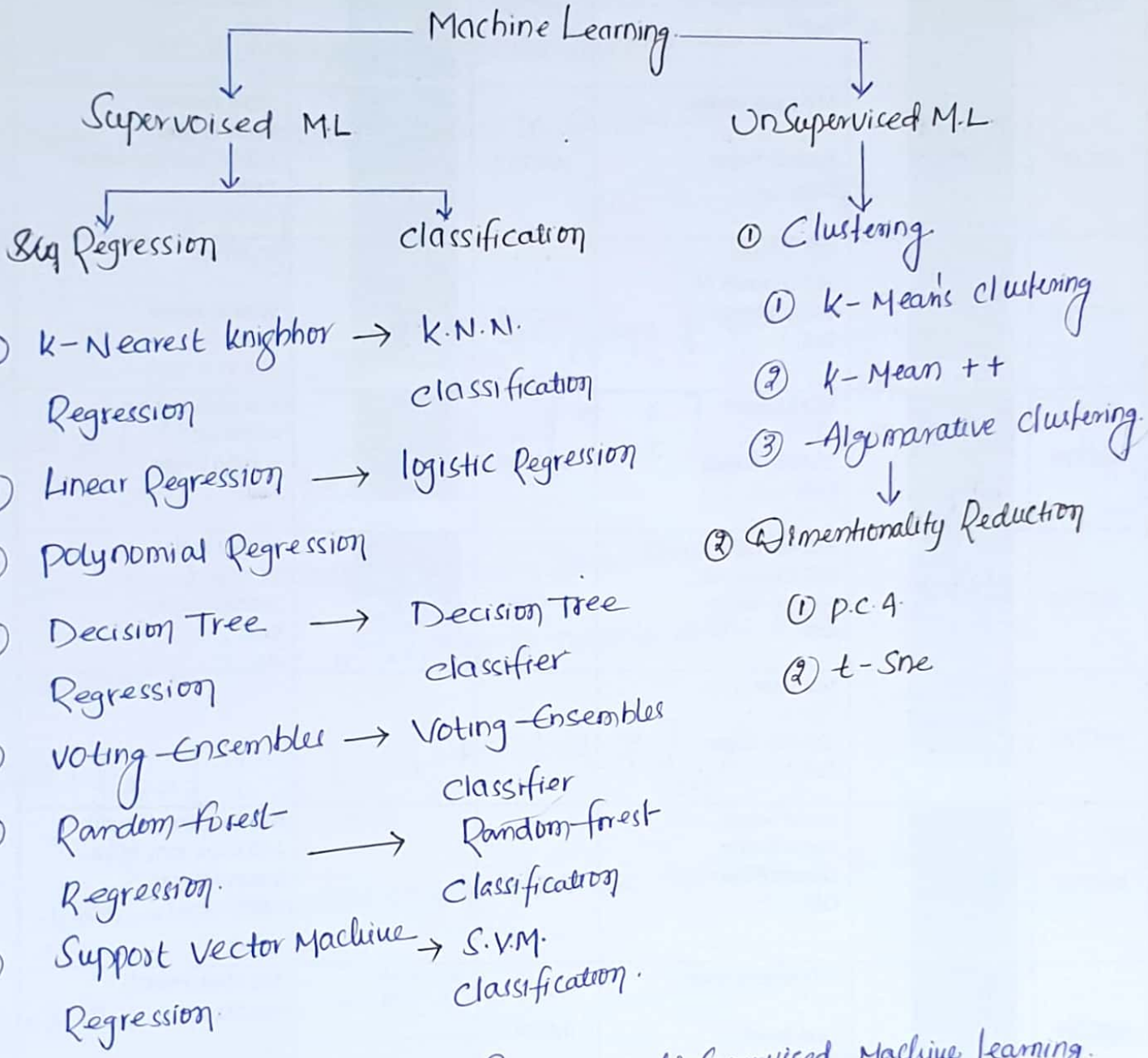
Machine learning :- it is a method of data analysis that automates analytical model building. it is a branch of "Artifical Intelligence" based on the idea that System can learn from data, identify patterns and make decisions with minimal human intervention.

classification of Machine learning :- Classification of M.L based on variable, the given random variable (output target) is which type of variable if it is a descrete type of variable then we go to ~~Regression~~ classification Task in this we use regression related algorithms, if it is a classification type of variable that mean the random variable is Continious type then do ~~classification~~ Regression related algorithms. this whole terminology comes under "Supervised Machine learning" technique.

Supervised Machine learning :- it is a Subcategory of ML & AI, it is defined by its use of labeled data sets to train algorithms that to classify data or predict Out comes accurately. So here in this process we Supervise our data. to give a good ~~model~~ or best model. (we have output variable)

Scanned with CamScanner

UnSupervised Machine Learning :- In which users do not need to Supervise the model, instead it allows the model to work on its own to discover patterns and information that was previously undetected, it mainly deals with the Unllabeled data (we don't have output)



Here we just focused on Supervised, UnSupervised Machine learning. Techniques Under Supervoiced we learn diff algorithms as shown above graph.

Mainly Machine learning classified as — ① Supervised learning ② UnSupervised learning, ③ Reinforcement learning. Hybrid learning process

Scanned with CamScanner

(4) Semi Supervised Learning, (5) Self Supervised learning.

(6) Multi - Instance learning. Statistical Inference

(7) Inductive learning, (8) Deductive learning.

Matrix :- It is a Set of Numbers arranged in row and column format So as to form a rectangular array. the numbers are called the elements, or entries, of the matrix, represent in Capital alphabet

Ex: $A = [\quad]_{m \times n}$ ⇒ Here m - no of rows, n - no of columns.

Size / Shape of Matrix - $m \times n$

Case - I

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2\times 2} + B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}_{2\times 2} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

Here the both matrix are having same shape of (2×2) matrix So easily we can add the matrix ie addition is possible in this matrix

Case - II $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2\times 2} + B = \begin{bmatrix} 2 & 2 \end{bmatrix}_{1\times 2} \neq [\quad]$

Here Addition is not possible because the shape of the matrix is not similar.

Case - III $\boxed{A+B = B+A}$

We can write or perform "A+B = B+A" when shape of matrix should be similar

Multiplication of matrix can be possible, when matrix $A_{m\times\boxed{n}} * B_{\boxed{n}\times p}$ in-form where no of columns in matrix 'A' should be equal to no of rows in matrix 'B' then only it is possible. then the form of 'B' is anything.

Scanned with CamScanner

$$A_{m \times n} * B_{n \times p} \Longrightarrow \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2 \times 2} * \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{2 \times 1}$$

where → A_m → represent no of rows — B_m → NO of rows

A_n → number of columns B_p → NO of columns

In Process of Multiplication Compulsary the Columns of given matrix should be similar with row of given other matrix then the result would be in the form of $m * p$ at the above example the resultant matrix should be 2×1

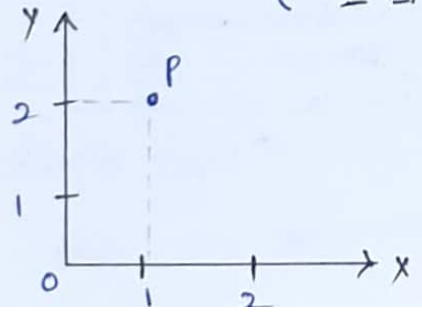
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{2 \times 2}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{2 \times 1} \Rightarrow A * B = \begin{bmatrix} 1 \times 1 + 2 \times 1 \\ 3 \times 1 + 4 \times 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 7 \end{bmatrix}_{2 \times 1}$$

Transpostation of Matrix A^T :- Transpostation of Matrix is nothing but Interchange the Shape of matrix means if we have a matrix of shape $m \times n$, then Interchange the rows in to Columns and columns in to row's this process known as transpost ation of matrix (A^T) after transpost the matrix should be $n \times m$,

Ex: $$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}_{\underset{2 \times 2}{m \times n}} \Rightarrow A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}_{\substack{n \times m \\ 2 \times 2}}$$

from this — $(A+B)^T = A^T + B^T = B^T + A^T$ (addition of Transpost)

$(A_{m \times n} * B_{n \times p})^T = B^T * A^T$ (Multiplication of Transpost)



⇒ P denotes the point In mathematical notation we represent this point (P) in Vector notations.

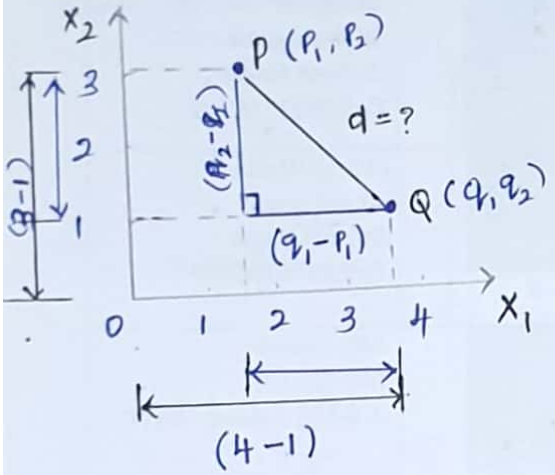
$$\vec{P} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Scanned with CamScanner

Vectors :- By the above graph we represent point in Vector notation

Basically Vector is a row and Column Combination defaultly ie matrix representation when ever we denoting point that should be in Column representation only not in row representation.

Eucledian Distance :- Distance between two points in Euclidian Space is the length of the line Segment between the two points, it Can be Calculated from the Cartesian Coordinates of the points using the pythagorean theorem.



$\Rightarrow$ $P = \begin{bmatrix} 1 \\ 3 \end{bmatrix}^{p_1}_{p_2}$ $Q = \begin{bmatrix} 4 \\ 1 \end{bmatrix}^{q_1}_{q_2}$

According to Pythagorus theorem

$$d^2 = (q_1 - p_1)^2 + (p_2 - q_2)^2$$

$$d = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2}$$

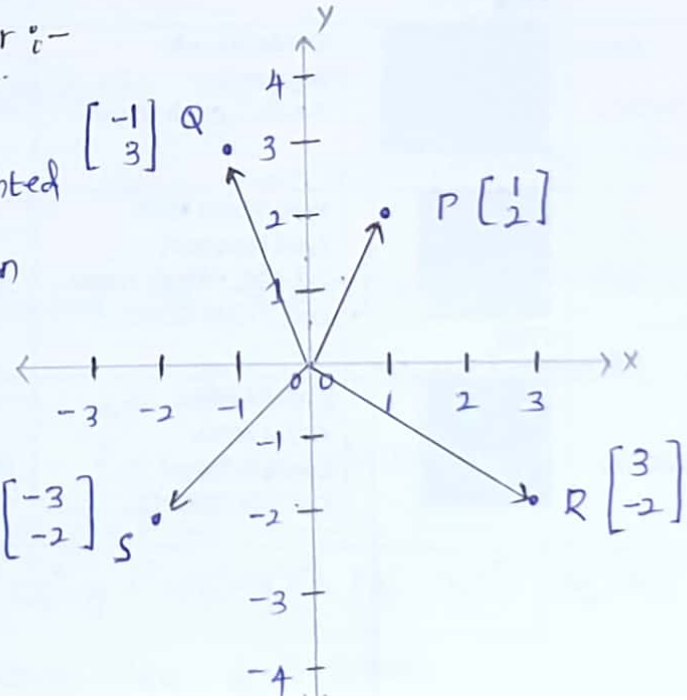
$$\sqrt{(1-4)^2 + (3-1)^2} = \sqrt{(3)^2 + (2)^2}$$

Eucledian Distance $\rightarrow$ $d = \sqrt{13}$ //

Data point Representation using Vector :-

D_1	D_2
1	2
-1	3
3	-2
-3	-2

Data points are represented in vector Notation as shown this representation based on column vector.



Simply Data set it is having only two dimentionality. (2D)

Scanned with CamScanner

This Data set can ber represent in matrix format also called as "Data Matrix". the given data set having only two Colomns less dimentionality

$$\text{Data Matrix} = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 3 & -2 \\ -3 & -2 \end{bmatrix}_{4\times 2}$$

↗ Dimentionality (# columns)

↳ No of Data points

But this Data matrix is not in Correct format becouse in point vector representation of point is only in column notation. (defaultly). So we have to change row as column wise So we have to do Simply Transport.

$$\begin{bmatrix} [1\ 2]^T \\ [-1\ 3]^T \\ [3\ -2]^T \\ [-3\ -2]^T \end{bmatrix} \Rightarrow \text{point representation of Data Matrix}$$

$$p = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_{2\times 1} \quad \varphi = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{3\times 1} \quad p = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}_{4\times 1} \Rightarrow \underline{x}_i \Rightarrow n \text{ \# of } D.$$

Point representation in 2D, 3D, 4D, ---- an D (Dimentionaility)

By visually we can Seen 2-Dimentionality only, and Some imaginary Past vision also we can Seen (ie 3-D) above three dimentionality we Can't Seen point representation. but we perform one place a point by mathemati-cally but not we visualize it.

i = index of column / dimention

j = no of Data points.

x_1	x_2	x_3	x_4
x_1^0	x_2^0	x_3^0	x_4^0
x_1^1	x_2^1	x_3^1	x_4^1

$\Rightarrow X_i^j \Rightarrow$

$$d = \sqrt{(x_1^0 - x_1^1)^2 + (x_2^0 - x_2^1)^2 + (x_3^0 - x_3^1)^2 + (x_4^0 - x_4^1)^2}$$

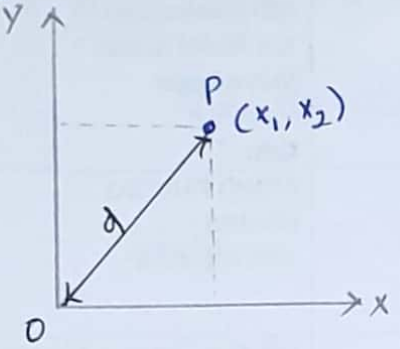
↳ Distance b/w 4-D points.

Scanned with CamScanner

Length of a vector :- A point which having Magnitude & Direction is called Vector, represent in $(\vec{v})$. from this magnitude tells us the length of a vector; In Simply Distance between Origin to the point is called as "length of a vector". In other terms "Eulidian Distance" also called as

Length of a vector represent as "$\|\vec{p}\|$"

$$\vec{p} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \|\vec{p}\| = \sqrt{x_1^2 + x_2^2} \quad \text{2 Dimention}$$

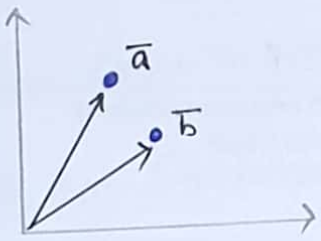


$$\vec{q} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \Rightarrow \|\vec{q}\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$$

↳ length of a vector.

[Origin] ⇒ d represent Eulidian Distance aswell length of a vector

Addition of a vector :-



$$\vec{a} + \vec{b} \Rightarrow \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \begin{bmatrix} a_1+b_1 \\ a_2+b_2 \end{bmatrix}_{2\times 1}$$

Multiplication of a vector :- Basically we multiply the vectors by two methods they are 1) Cross product 2) Dot product

Dot product :-

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \Rightarrow \boxed{\vec{a}.\vec{b} = a_1 \times b_1 + a_2 \times b_2}$$

Note :- The resultant should be a Scalar value.. (point)

Ex: $\vec{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \vec{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix} \Rightarrow \vec{a}.\vec{b} = 1\times 5 + 3\times 6 \Rightarrow 23$ → Scalar

Simply we can say Dotproduct ⇒ $\boxed{\vec{a}.\vec{b} = \sum_{i=1}^{n} a_i b_i}$ n-Dimention

Scanned with CamScanner

for Suppose if we have (4-D) point vectors and we have to multi-ply them as Dot product method.

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

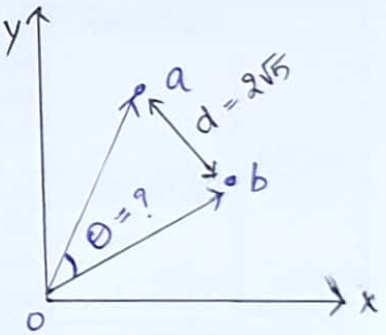
$\vec{a}.\vec{b} = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4 \Rightarrow \vec{a}^T * \vec{b}$ also written

from the above eqn $[a_1\ a_2\ a_3]_{m\times n} \times \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}_{n\times p}$ can be written,

$\therefore$ Simply — $\boxed{\vec{a}.\vec{b} = a^T * b = \sum_{i=1}^{n} a_i b_i}$

if we have data set and from that data. if we plot a Scatterplot and the point representation in form of vector then we find the length of vector from Origin. in that process we calculate distance from Origin to every point in that Situation if angle making in between two points let Consider. then we know the how much angle making w.r.t to two vectors then we Simply do dot product ie multiplication as shown in below

$$\boxed{\vec{a}.\vec{b} = a^T.b \Rightarrow \sum_{i=1}^{n} a_i b_i = \|\vec{a}\|.\|\vec{b}\|.\cos\theta_{a,b}}$$



Data

C1	C2
1	2
5	-1
6	3
5	0

$\Rightarrow \vec{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \vec{b} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \Rightarrow \vec{a}.\vec{b} = 1*5 + 2*0 = 5$

E.Distance $= \sqrt{(1-5)^2 + (2-0)^2} = \sqrt{4^2+2^2} = \sqrt{20} = 2\sqrt{5}$ "

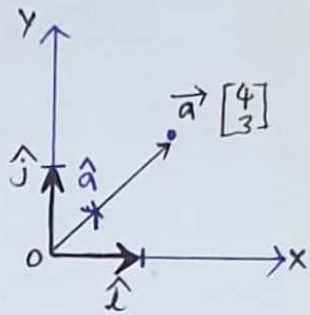
$\Rightarrow \vec{a}.\vec{b} = \|\vec{a}\|.\|\vec{b}\|.\cos\theta$

$1*5 + 2*0 = \sqrt{1^2+2^2} * \sqrt{5^2+0^2}.\cos\theta$

$5 = \sqrt{5}.5.\cos\theta \Rightarrow \theta = \cos^{-1}\left(\frac{1}{\sqrt{5}}\right)$ "

Scanned with CamScanner

Unit Vector :- A vector has both magnitude & direction, A vector that has magnitude of exactly '1' is called as "Unit vector" it is also called as direction vector represented as- $\hat{a} \Rightarrow \wedge$



$$(\text{Unit.v})\ \hat{a} = \frac{\vec{a}}{\|\vec{a}\|} \quad \begin{matrix}(\text{vector}) \\ (\text{length of vector})\end{matrix}$$

Ex: $\vec{a} = \begin{bmatrix}4\\3\end{bmatrix} \Rightarrow \|\vec{a}\| = \sqrt{4^2+3^2} = \sqrt{25}$

$= 5$,,

$\hat{a}$ (unit vector) $= \frac{\vec{a}}{5} \Rightarrow \frac{1}{5}\begin{bmatrix}4\\3\end{bmatrix}$

$$\hat{a} = \begin{bmatrix}4/5\\3/5\end{bmatrix},,$$

* Unit vector has a propertly that the magnitude should be 1,,

Lets proove it, Magnitude is nothing but length of a Unit vector

$$\|\hat{a}\| = \sqrt{\left(\frac{4}{5}\right)^2+\left(\frac{3}{5}\right)^2} = \sqrt{\frac{16}{25}+\frac{9}{25}} = \sqrt{\frac{25}{25}} = \sqrt{1}$$

$\|\hat{a}\| = 1$,, (Magnitude),,

distances calculated in diffrent-forms also-there are several methods-to calculate distances. Such as -

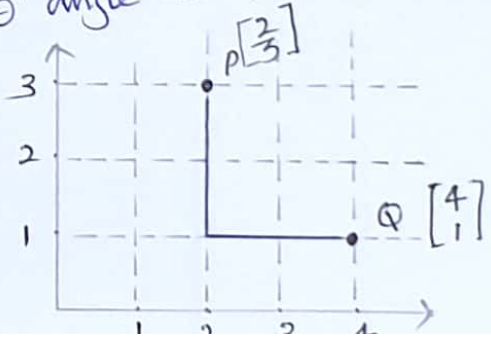
① Euclidian Distance → Accoding-to pythagorus

② Manhatten Distance → it is a name of city in Newyork where the representation done in block wise.

③ Minkowski Distance

④ Cosine Distance & etc. ⇒ when "θ" angle known

When we use Manhatten we represent points in blocked distances.



Scanned with CamScanner