✱ M/c Learning ⇒

- Branch of AI.
- The process of developing a model & algo that makes computer to learn from data.
- ML teaches system how to think, understand & predict the o/p like human.

✱ Types ⇒ Supervised
⇒ Unsupervised
⇒ Reinforcement
⇒ Semi supervised

✱ Supervised ⇒

- When the model in which it used labelled data to teach m/c.
- Labelled data → I/P
→ O/P

- With every I/P we have an associated o/p with it.

Ex ⇒ Spam email detection.

- 2 categories ⇒ Classification
⇒ Regression

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Classification ⇒ It is a type of supervised learning in which ~~data~~ o/p is categorial (Discrete value).

Eg ⇒ Identify the new mail is spam or not.
⇒ Picture is of dog or cat.

★ Types ⇒ Binary (Dog, cat)
⇒ Multi class (Dog, cat, cow)
⇒ Multi label (Actor into diffrent movie)

★ Eg ⇒ SVM
⇒ Decision tree
⇒ KNN
⇒ Random Forest

Main Ideas, Questions & Summary:

Library / Website Ref.:-

✱ Regression ⇒

- It used to give O/P in continous variables.
- Predict relationship b/w 2 variables
- how one variable on change of other variable.
- Dependent variable ⇒ Depend on other variable
  Undepend variable ⇒ Variable used to predict dependent variable.

- Ex ⇒ House price depend on its size / location.

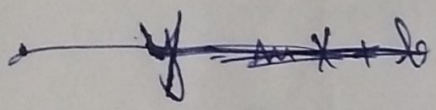
- Ex ⇒ Linear Regression
  ⇒ Polynomial
  ⇒ Decision tree
  ⇒ Random forest

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

✱ Linear Regression ⇒

- It is used for relationship b/w dependent & one or more indepnd variable.
- It make relationship using linear equation
- Predict continuous o/p on basis of independant variable.



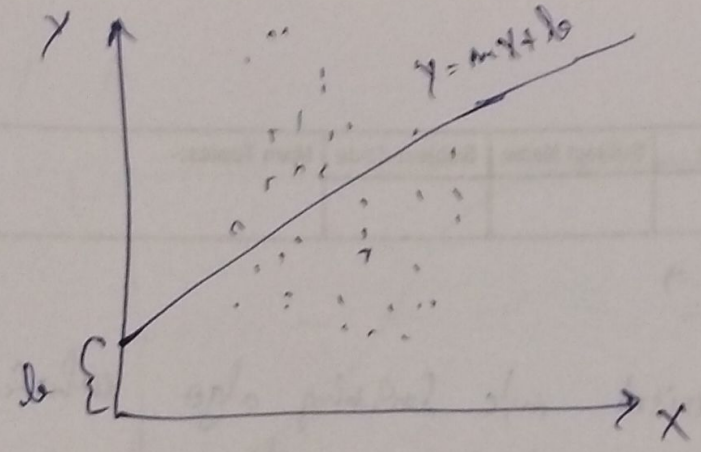
- $Y = mX + b$ → Linear eq

Y = dependent var
X = Independent var
m = slope of line (how much Y changes for unit change in X)
b = intercept (Value of Y when X is 0)

Main Ideas, Questions & Summary:

Library / Website Ref.:-

⇒



* Best fit line is the main aim of Linear Regression.

- It implies that the error b/w the predicted & actual values should be kept minimum.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

☆ Logistic function ⇒

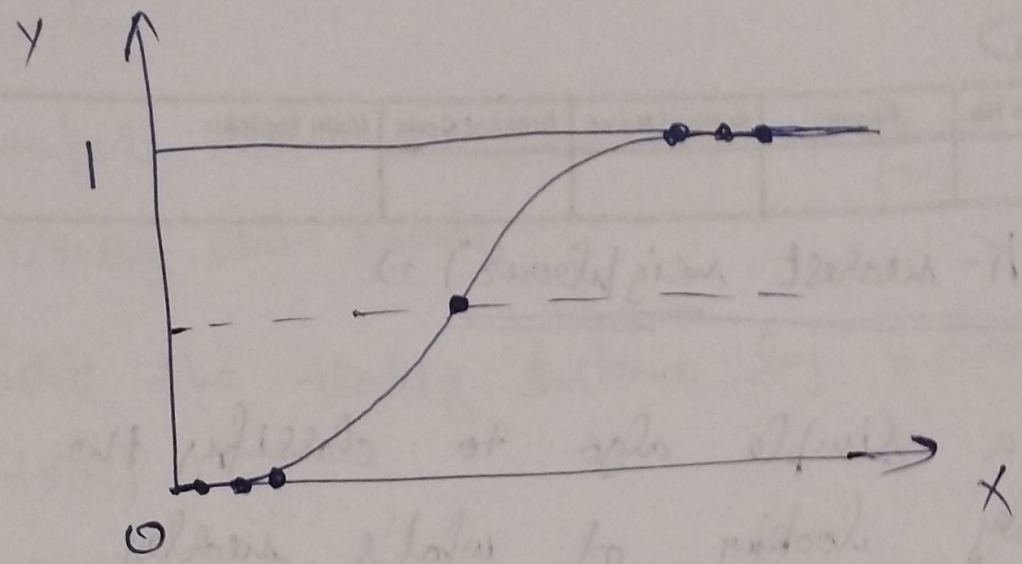
- It is a supervised m/c learning algo where we predict the probability of a instance whether it belong to given class or not.
- Dependent variable is categorical & binary (0,1)
- Sigmoid f(n) is used to give value b/w 0 & 1.
- $$Y = \frac{1}{1 + e^{-(a_0 + a_1 x)}}$$ → Sigmoid f(n)

a_0 = intercept
a_1 = Coefficient
x = Indepet variable

Ex ⇒ If study > 5 Hr pass
else fail

Main Ideas, Questions & Summary:

Library / Website Ref.:-



** Diff Linear & Logistic

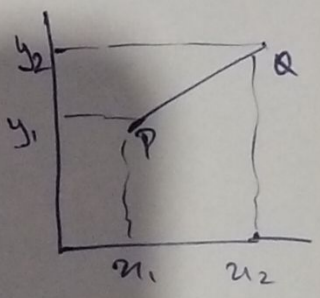
POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ KNN (K-nearest neighbour) ⇒

- It is a simple algo to classify the things by looking at what's nearby.
- Also called lazy learner algo
- It is a method of supervised learning algo used for classification & regression.
- It classify the new instance based on majority class among its k closest neighbors
- Use euclidean distance method.

$$d(p,q) = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}$$



- Eg ⇒ Movie time, Rating
  ↓
  Genre predict

Main Ideas, Questions & Summary:

Library / Website Ref.:-

☆ Steps ⇒

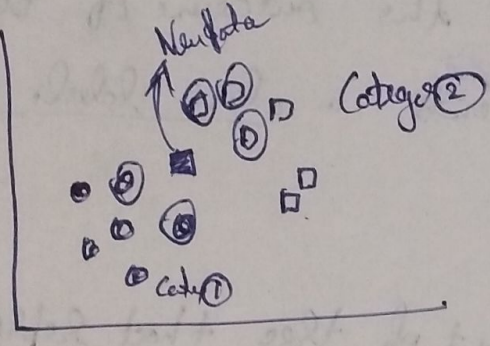
① Compute distance b/w new instance (x) & all training data points.

② Sort the training instance by increasing distance to (x)

③ Select the top K closest instance. (K nearest).

④ Count the class labels among these K neighbour

⑤ Assign the majority class to x.



☆ K is just a no that tells how many nearby point (neighbour) to look while taking decision.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Decision Tree ⇒

- Supervised m/c learning algo used for both classification & Regression.
- It works by splitting the data into subset based on feature values, forming a tree like structure.
- Each branch represent the outcome of test.
- Each leaf node represents a class label. or value.

- Key concepts ⇒

(1) Root Node ⇒ Starting point of tree that represent entire dataset.

(2) Decision Node ⇒ Intermidate node where data is split.

(3) Leaf Node ⇒ Represent O/P

(4) Splitting ⇒ Divide a node into 2 or more sub nodes

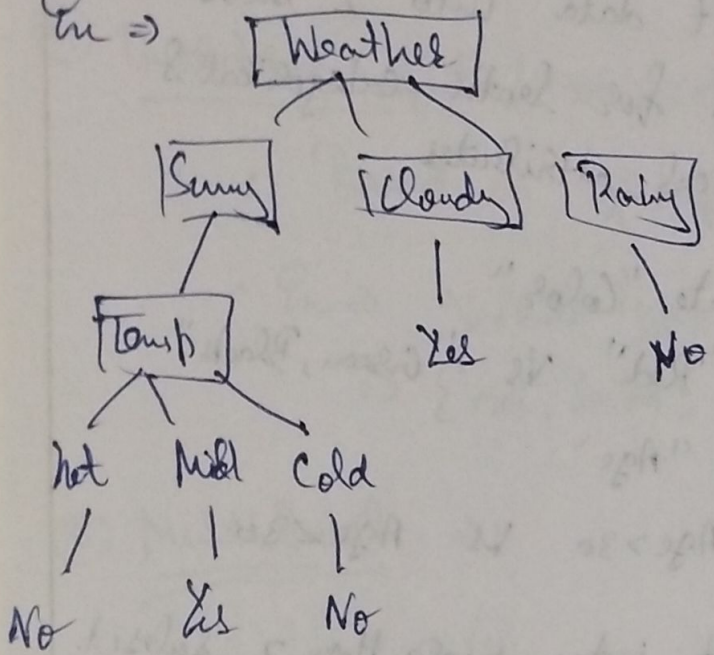
Main Ideas, Questions & Summary:

Library / Website Ref.:-

(5) Entropy ⇒ Criteria to measure the least features to split the data

(6) Info Gain ⇒ Help to select the feature that best splits the data. (Max IG choose)

Ex ⇒



Weather to play or not

✱ Entropy ⇒

$$S = -\sum_{i=1}^{n} P_i \log_2(P_i)$$

✱ IG ⇒

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{n \in \text{Value}(A)} \frac{|S_n|}{|S|} \cdot \text{Entropy}(S_n)$$

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

☆ Various split types ⇒

(1) Binary split ⇒ Split data into 2 subset
⇒ Common for both categorical & numerical attributes.

- Categorial ⇒ Attribute "Color"
  Split "Red" Vs "Green, Black"
- Numeric ⇒ Attribute "Age"
  Split Age > 30 Vs Age < 30

(2) Multiway split ⇒ Split into more than 2 subset
⇒ Categorial Attribute

- Category ⇒ Weather
  ⇒ Sunny Vs Cloudy Vs Rainy
- Numeric (Less use) ⇒ age - - -

Main Ideas, Questions & Summary:

Library / Website Ref.:-

③ Ordered split (For numerical attribute) ⇒

- Split based on threshold value
- Only 2 branches are created
- Ex ⇒ "Salary ≤ 500" Vs "Sal ≥ 500"

④ Unordered split (For categorical attribute) ⇒

- Based on grouping subset of category value
- Ex ⇒ Vechile Type
  ⇒ {car, bike} Vs {Truck, Bus}

⑤ Multilevel ⇒

- Spliting on same attribute at multiple levels in tree.
- Ex ⇒ First split, Age ≥ 30 ≤ 30
  ⇒ later, Age ≥ 50 Vs ≤ 50

⇒ Both cat, numerical.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

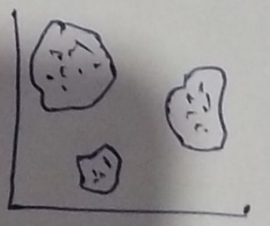
✱ Unsupervised Learning ⇒

- Discover pattern & relationship using unlabelled data.
- Discover hidden pattern, similarities, clusters.
- 2 types ⇒ Clustering
  ⇒ Association

✱ Clustering ⇒ Process of grouping data points into clusters based on their similarity.

⇒ Useful for identifying pattern & relationship in data without need of labelled examples.

⇒ K-means



Main Ideas, Questions & Summary:

Library / Website Ref.:-

* Association => Technique for discovering relationship b/w items in dataset.

Eg => FP growth

=> Apriori Algo

POORNIMA

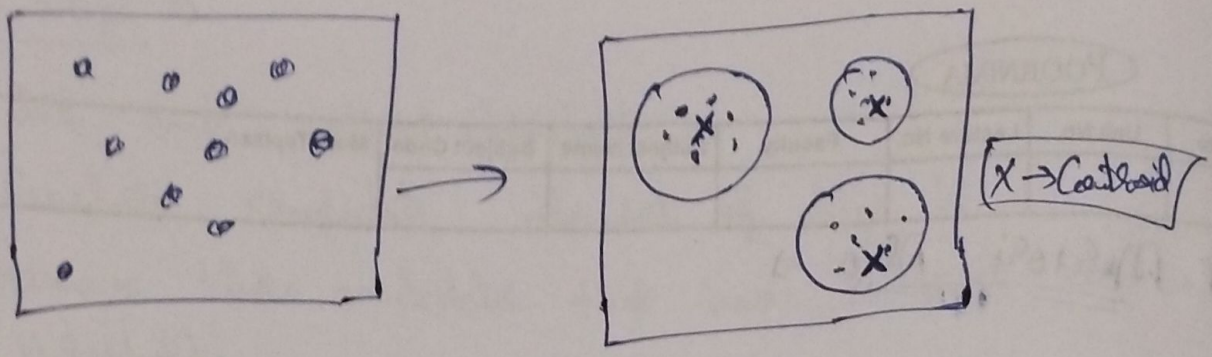
Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ K- means clustering ⇒

- It is a unsupervised m/c learning algo which groups the unlabelled dataset into different clusters.
- Organize data into group based on their similarity.
- Ex ⇒ Grouping customer with same interest.
- Works by first randomly picking some central point called centroids.
- Each data point assigned to closest centroid. forming a cluster.
- After point is assign to cluster then centroid is updated by using euclidean distance

Main Ideas, Questions & Summary:

Library / Website Ref.:-



✡ Steps ⇒

① Choose the no. of cluster (K).

② Randomly select K data point as initial centroid.

③ Each point is assigned to nearest centroid using euclidian dis.

④ Compute new centroids (using mean of all point)

⑤ Repeat [③ & ④] until

⇒ Cluster assignments no longer change

⇒ A maximum no of iterations is reached.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Apriori Algo ⇒

- Algo helps in finding frequent itemset in a dataset & derives rules about item co-occurrence.
- Used in association rule mining & market basket analysis.

★ ① Item Set ⇒ Group of items that appear together frequently in transaction.

② Support ⇒ The frequency of transaction that contain an itemset

③ Confidence ⇒ How often items in Y appear in transaction that contain X ($X \rightarrow Y$).

④ Lift ⇒ Measure how much more likely Y is to appear with X than by chance.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

* Steps =>

① Set min support threshold.
② Generate candidate itemset of size K.
③ Remove those candidates that have support less than threshold.
④ Use frequent itemsets to generate c(K+1).
⑤ Repeat until no more frequent itemsets found.
⑥ Generate association rule from frequent itemset using confidence & lift.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

* FP growth =>

- Frequency Pattern growth algo.
- It is a powerful & efficient method for frequent itemset & association rule learning.
- Often used by market basket analysis to find frequent pattern (item combination) in large dataset.
- Improvement over Apriori algo.

Main Ideas, Questions & Summary:

Q ⇒ Min Sup = 60%

Min Confidence = 80%

T_1	MONKEY
T_2	DONKEY
T_3	MAKE
T_4	MUCKY
T_5	COOKIE

(count once 'O')

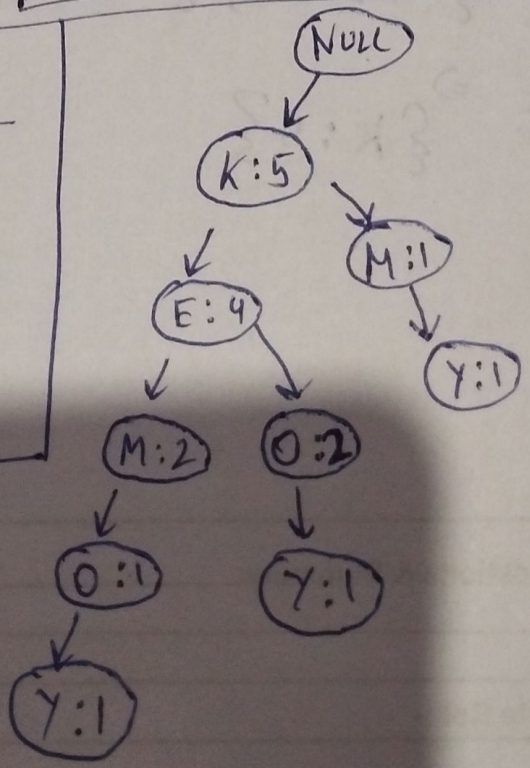
Ans (i) FP Tree

$\frac{60}{100} \times 5 \Rightarrow 3$ → Support

✓ M = 3
✓ O = 3 (count once)
✗ N = 2
✓ K = 5
✓ E = 4
✓ Y = 3
✗ D = 1
✗ A = 1
✗ U = 1
✗ C = 1
✗ I = 1

K > E > M > O > Y

T_1	KEMOY
T_2	KEOY
T_3	KEM
T_4	KMY
T_5	KEO



POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

(iii) Conditional Pattern Base ⇒

K = X (don't take as attach to NULL)
(If other than NULL attach then take alle along with

E = 4
M = 3
O = 3
Y = 1

- For Y ⇒ {K, E, M, O : 1} {K, E, O : 1} {K, M : 1}
- For O ⇒ {K, E, M : 1} {K, E : 2}
- For M ⇒ {K, E : 2} {K : 1}
- For E ⇒ {K : 4}

Main Ideas, Questions & Summary:

Library / Website Ref.:-

(iii) Conditional FP tree ⇒

Seperate if different branches.

- For Y ⇒

{K:3 (E:2, M:2, O:2)}

× Min supp = 3

- For O ⇒

{K:3, E:3, (M:1)}

×

- For M ⇒ {K:3, (E:2)}

×

- For E ⇒

{K:4}

(iv) Frequent Pattern generation ⇒

- For Y ⇒

{K,Y:3}

- For O ⇒

{K,O:3} {E,O:3} {K,E,O:3}

- For M ⇒

{K,M:3}

- For E ⇒

{K,E:4}

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

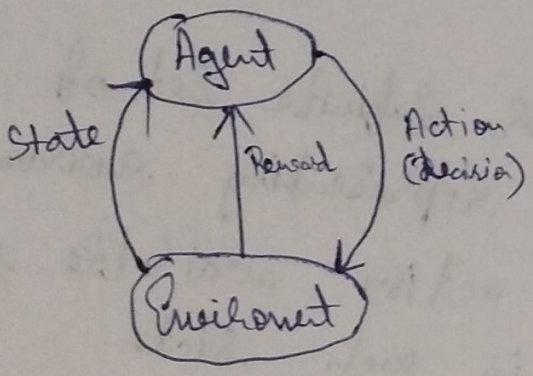
☆ Reinforcement Learning ⇒

- It is a learning algo that interacts with environment by producing actions & discovering errors.
- Trial, error, delay are characteristic of it.
- Model increase performance using Reward feedback.
- Ex ⇒ Q-Learning
  - ⇒ SARSA
  - ⇒ Deep Q-learning
- Ex ⇒ AI with chess.
- Make Agent explore different moves & recieve +, - feedback.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

- Agent ⇒ Decision maker that perform actions
- Environment ⇒ World / System in which agent operates.
- State ⇒ Situation / Condition the agent is currently in.
- Action ⇒ Possible moves or decision the agent can make.
- Reward ⇒ Feedback or result from environment based on agent's action.



POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

✱ Markov decision process (MDP) ⇒

- It is a mathematical framework used to describe an environment in decision-making scenarios where the o/p is under control of decision maker.

- Components ⇒

① S (State Space) ⇒ Finite or infinite set of states representing environment

② A (Action Space) ⇒ Set of action available to agent in each state

③ P (Transition Probability) ⇒ $P(S' | S, a)$
⇒ Prob of moving to state S' from state S by taking action a

④ R (Reward fn) ⇒ The reward after transitioning from S to S' via a
⇒ $R(S, a, S')$

⑤ γ (Gamma) ⇒ Value range $[0,1]$ that determine importance of future reward.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

* Markov Property ⇒ States that next state depends only on the current state & action, not on past states.

$$P(s' \mid s, a, s_{t-1}) = P(s' \mid s, a)$$

ॐ Jai Baba ki ॐ

* Recommendation System ⇒ Sub class of m/c
  ⇒ Use for predicting the rating or ranking of a item that user.
  ⇒ It is used to predict rating a user might given to specific item.
  ⇒ Eg ⇒ Google, Insta
  ⇒ Spotify suggest new song on the basis of our interest & likes.

⇒ Diffent way to build recomdn system → Algorithmic approach
→ probabilistic approach

* Collaborative filtering system ⇒ It is used to find the similar users with the similar interests & recommend what they will like.
⇒ If A & B have similar taste in a product X. A & B are likely to have same taste in other prod

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:

⇒ Approaches of Collaborative filtering
↳ Memory based
↳ Model based

① Memory based ⇒ It is a method of making recommendation by analyzing the relationship b/w user & items based on their past interaction.

• 2 types

↳ User based ⇒ It uses the target user's past rating & other user & then predict recommendation

↳ Item based ⇒ It used rating which are positively given by the user & recomend new items.

② Model based ⇒ It uses m/c learning algorithms to find the preferences & recommend the new preferences.
⇒ Include decision tree, rule based approach
⇒ It used dataset & predict o/p.
⇒ Ex ⇒ YT

Main Ideas, Questions & Summary:

Library / Website Ref.:-

- Adv ⇒ Simple
  ⇒ It capture suitable character only

- Dis Adv ⇒ Not friendly

☆ Content Based filtering ⇒ It works by taking the data from the user.
⇒ It takes data explicitly (rating) or implicitly (clicking on link).
⇒ By the help of data, we can create user profile which is used to give suggestion to user.
⇒ Engine become more accurate, when user give more action on recommendation

- Require following data Sources ⇒

① User Profile ⇒ In this, we create vector that describe the user preference.
⇒ In this, we create utility matrix which describe the relationship b/w user & item.

② Item profile ⇒ In this, we create profile for each item which represent the important characteristic of that item.

- Adv ⇒ Good, if user give more rating
- Dis Adv ⇒ Ineffective for new user.

Ex ⇒ Amazon product feedback.
⇒ Music recommendation

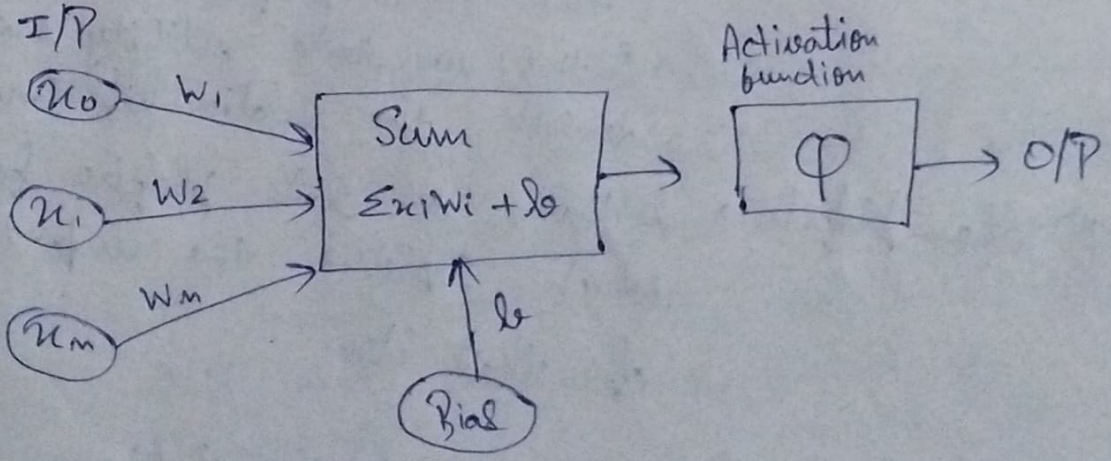
Types of Activation fn

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Artifical Neural Network (ANN) ⇒

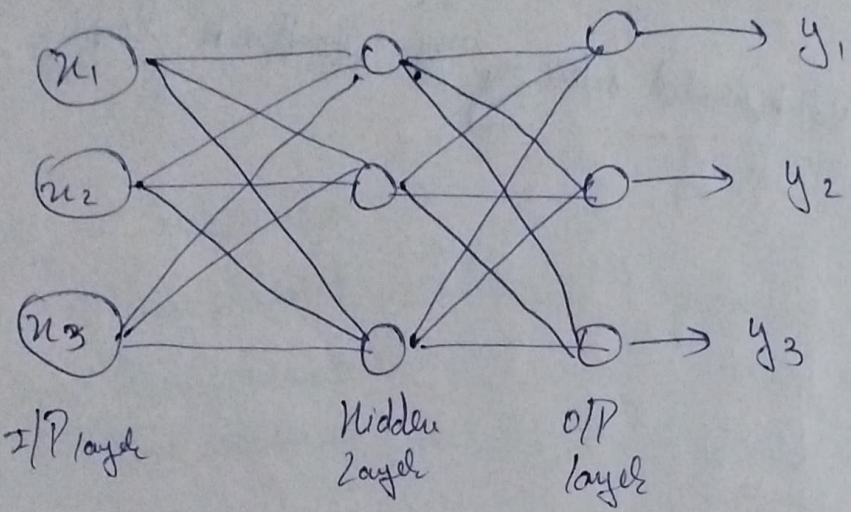
- It is derived from biological neural n/w that develop the structure of human brain.
- In brain neurons are interconnected same in ANN, these neurons are called nodes / units.
- These units are connected to each other & form ANN.
- A layer can have million of units.
- 3 layer ⇒ Input layer (Recieve data from outside world)
  ⇒ hidden layer (Perform calculation to find hidden feature & pattern)
  ⇒ output layer (Give data which is relevable)
- Data have to go through one or multiple hidden layers.
- Diff layer connected to each other which have weights.



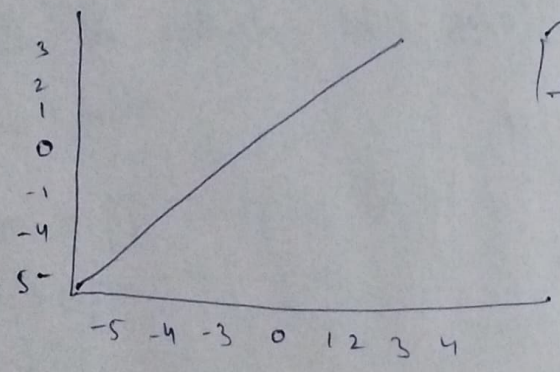
Main Ideas, Questions & Summary:

Library / Website Ref.:-

- Weighted links ⇒ Also known as Synapses links (determine which factor is important)
- Activation function ⇒ These are mathematical $f(x)$ that determine the O/P of a neural network.
  ⇒ They decide whether the neuron should be activated or not.



✡ Linear function ⇒

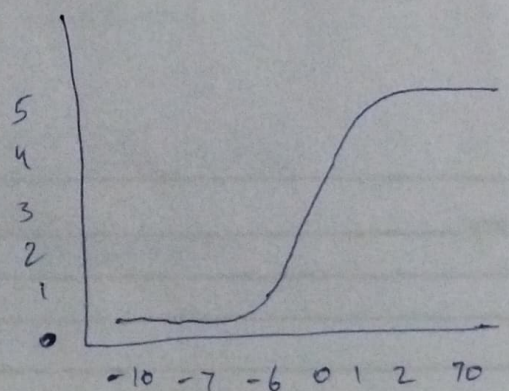


$$z = f(x) = x$$

✡ Non linear ⇒

① Sigmoid $f(x)$ ⇒ logistic $f(x)$

It's in S shape



$$f(x) = \frac{1}{1 + e^{-z}}$$

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

* Types of ANN ⇒

(1) Feed forward n/w ⇒ Information flow from I/P layer to O/P layer without any feedback loop.

(2) Feed Backward n/w

Adv/Dis Adv ANN

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Types of ANN →

(1) Feed forward n/w → Information flow from I/P layer to o/p layer without any feedback loops.

(2) Feed Backward n/w

POORNIMA

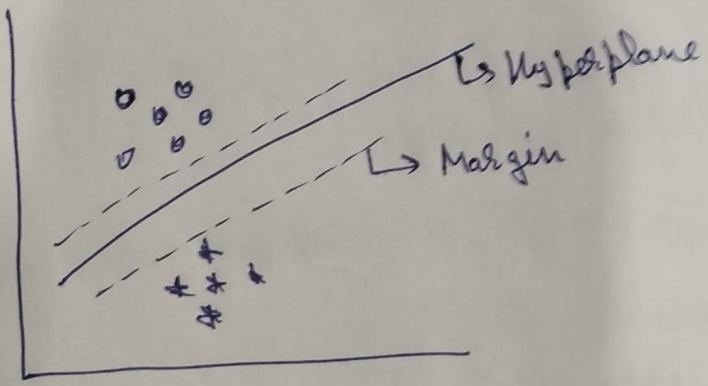
Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

☆ Support Vector m/c (SVM) ⇒

- It is supervised m/c learning algo used for classification & regression

- Hyperplane ⇒ It is a decision boundary that separates data points of diff classes.

- Eq ⇒ $\boxed{w.x + b = 0}$



- Margin ⇒ It is b/w Hyperplane & nearest data point of either class.

  ⇒ SVM tries to maximize this, to improve classification.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

- Support Vectors ⇒ Data points that are closest to the hyperplane.
  ⇒ Influence position of hyperplane
- Hard Margin ⇒ Margin hyperplane that Seperate the class accurately / perfectly.
  ⇒ Straight line
- Soft Margin ⇒ Margin that can't seperate data points
  ⇒ Curve.
- Kernal f(x) ⇒ Function that convert dimension to higher) to make data points linearly seperatable.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

★ Adv of Naive Bayes ⇒

1. Easy to implement
2. Fast & Flexible
3. Work well with large data
4. Handle missing data
5. Work well with categorical I/P
6. Require Less training data

Summary:

✱ KNN (Categorical, not numerical) ⇒

- Distance is measured using Similarity measures instead of euclidean.
- Hamming distance or Simple matching method used
- Dis is measured ⇒

  0 if value are same

  1 if value are different

Record	Color	Shape
A	Red	Circle
B	Blue	Circle
C	Green	Rectangle

  new ⇒ Color = Red
  Shape = Circle

(I) Red = Red (0)
Circle = Circle (0) ⇒ 0+0 ⇒ 0

(II) Blue ≠ Red (1)
Circle = Circle (0) ⇒ 1+0 ⇒ 1

(III) Red ≠ Green (1)
Rect ≠ Cir (1) ⇒ 1+1 ⇒ 2

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

☆ Confusion Matrix ⇒

- It is a performance measurement tool used for evaluating the classification model.
- Used in Supervised learning.
- Check performance by checking the actual & predicted O/P.

☆ Terms ⇒

- True +Ve ⇒ Model predict (correctly) the +Ve class.
- True -Ve ⇒ Model correctly predict the -Ve class.
- False +Ve ⇒ Model incorrectly predict the +Ve class.
- False -Ve ⇒ Incorrectly predict the -Ve class.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

✡ Overfitting ⇒

- It occur in m/c learning algo when the model is well trained with high accuracy data but failed when the new data is given.
- Give poor performance when using in real world data.
- Ex ⇒ Training a data with few data that are similar but the new outside world data is different from the trained data.

✡ Solutions ⇒

① Train model with more data ⇒ Give variation of data for training.

② Use simpler model ⇒ Don't use complex model.

③ Stop training early ⇒ If accuracy start decreasing stop the training.

④ Avoid learning from noise.

POORNIMA

Date	Unit No.	Lecture No.	Faculty	Subject Name	Subject Code	Main Topics:-

✱ Optimal policy in MDP (Markov Decision Process) ⇒

- In MDP, policy defines the action to taken in each state.
- Optimal policy is one that maximize the total reward over time.
- It tells agent the best action from every state to achieve the highest long term reward.

✱ Finding Optimal Policy ⇒

1. Q- Learning
2. SARSA
3. Linear Programming
4. Deep Q Network.
5. Monte Carlo tree Search
6. Value iteration
7. Policy iteration

Ex ⇒ If light is Red, best action is Stop
" " " green, " " " Go.

Main Ideas, Questions & Summary:

Library / Website Ref.:-

☆ Value iteration	Policy iteration
① Updates value $f(n)$	Updates policy
② Focus on finding optimal value	Focus on finding optimal policy.
③ Less stable	More stable
④ Faster (Less iteration)	Slower (more iteration)
⑤ Less complex	More complex

☆ On policy	off policy
① SARSA	Q Learning
② More stable	Less stable
③ Follow same policy it learns	Learn from different policy
④ Slow learning	Fast learning