Purpose:-

You will be able To face around the world in AI Revolution

Career Opportunities:-

Machine learning enhance career across various Profession with AI knowledge now expected in multiple fields

→ Medical Professional
→ Engineers → Business Analysts

Who will Benefits from This Course?

→ Aspiring AI Professionals
→ ML Practitioners → Interested Students

Who May Not Benefits?

→ no Programming knowledge
→ Weak in Mathematics
→ No Technical Interest

ML:-(AI)

Ability of computers To Think lean and make decision like humans

Purpose:-

You will be able To face around the world in AI Resvolution

Career Opportunities:-

Machine learning enhance career across various profession with AI knowledge now expected in multiple fields.

→ Medical Professional
→ Engineers → Business Analysts

Who will Benefits from This Course?

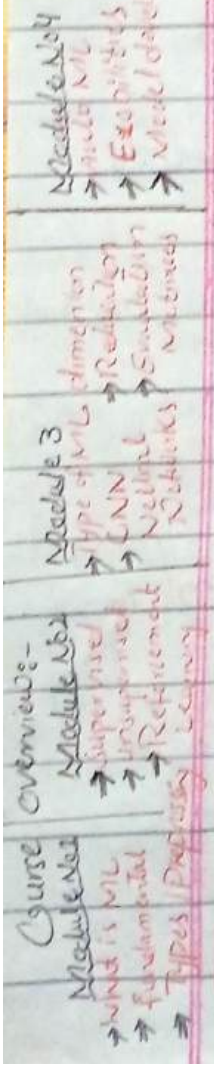
→ ASpiring AI professionals
→ ML Practitioners → Interested Students

Who May Not Benefits?

→ no programming knowledge
→ Weak in Mathematics
→ No Technical Interest

ML:-(AI)

Ability of computers To Think learn and make decision like humans



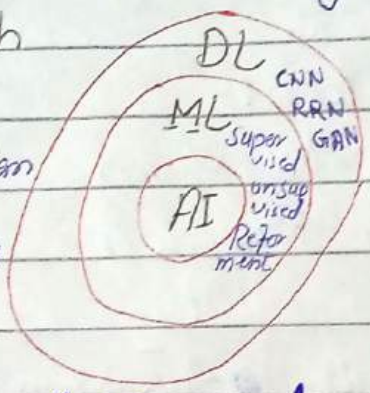
Applications:-

AI application

→ Speech Recognition → Text understanding predictions, → Problem-Solving. → health-care

To Solve Complex Problem throug ML efficiency.

Start with
→ Rolsed Dased System
→ Statistic way
→ Predition inference



Start
→ Research Based Algorithms
→ Solving Complex problems.

Machine Learning

There are manys algorithms of Supervised UnSupervised
It is a Subset of AI.

AI

(ANI) (Artifical Narrow Intelligence)
→ Used in Self driving Cars, Search engineers, farming facteries

(AGI) (Artifical general intelligence).
→ AI Capable of performing all human Tasks.

learns from data hidden patterns.

Machine Learning Traditional System

Answers Rules

→ It's rely on pre-defined Rols
→ Pattern recognition and make decision with explicit programing.

→ ML System analyze past data To from Predictive models inspired by human Brains.

Data Rules

Application :-
Computer Vision
Speech Recognition

ML	DL
→ Required explict feature engineering and domin expertise for effective Tast and classification	→ Autonomously learn hierarchical features from raw data. eliminating featere extraction

Key Components :-

→ Data and Features
→ Loss and optimization
→ Algorithms and Training
→ Evaluation and developments
Hyper Parameters

History :-

1970 Automated ML models
1979 driving car

→ 1642 Pascal first Mechnical Machine
→ 1847 Boolean logic
→ 1952 1st machine learning application
→ 1679 binary coded machines emerge.
→ 1950 Alan Turing develops The Turning Test

Lecture No # 2+3

Topic:-

Types of Machine Learning and Basic Workflow

- The fundamental types of machine learning approaches.

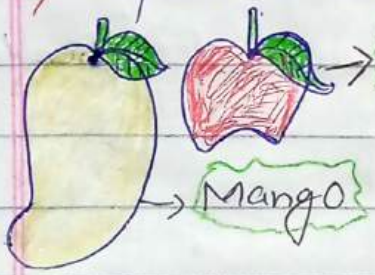
(1) (Supervised Learning):-
Use Labled Data for Training (Input-output Pair)

(2) (Un-Supervised Learning):-
Works with unlabeled Data.

(3) (Reforment Learning):-
Learns through interaction with enviments.

Examples:-

→ Supervised Learning



Models

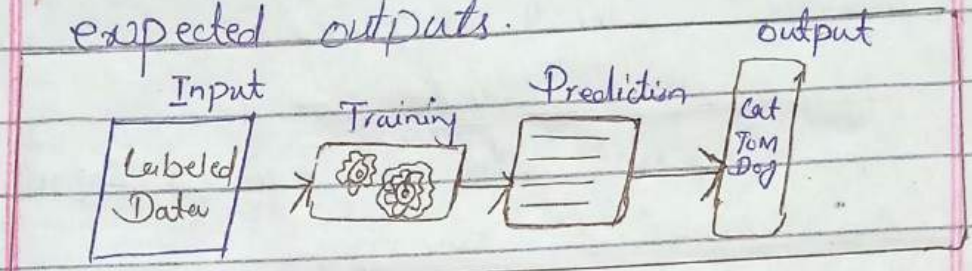
- Training a model with Lableled fruit Images.
- After Training the mode classify New Inges.

Application:-

→ Spam filtering	• Spam vs non-Spam emails
→ Speech Recognition	• Audio To-Text Conversion
→ Machine Translation	• Language Translation
→ Online Advesting	• Predicting Ad clicks.
→ Self-Driving Cars	• Deteding vehicles using Radar

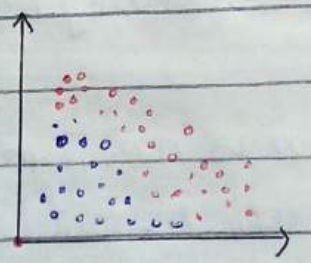
Supervised Learning:-

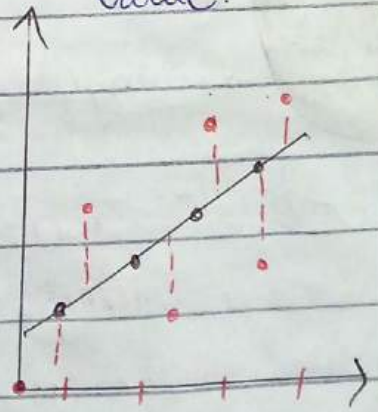
- Using Labeled data with Input output pair
- Learns patterns from example to predict result on new data.
- Easily Measured performance against expected outputs.



Types :-

Classification	Regression
→ categorizing data into predefined Labels (Spame or not Spame) in form of group.	→ Predicted continous values (house price future Sales) discrete value.

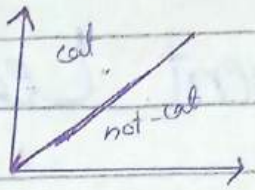




(Types of Classification:-)

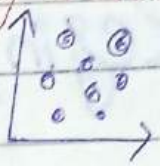
i. Binary classification

→ Two possible outcomes (Yes/No, Spam/No Spam)



ii Multi-class

→ More than 2 outcomes (child, young, adult, elderly)



(Common Algorithms:-)

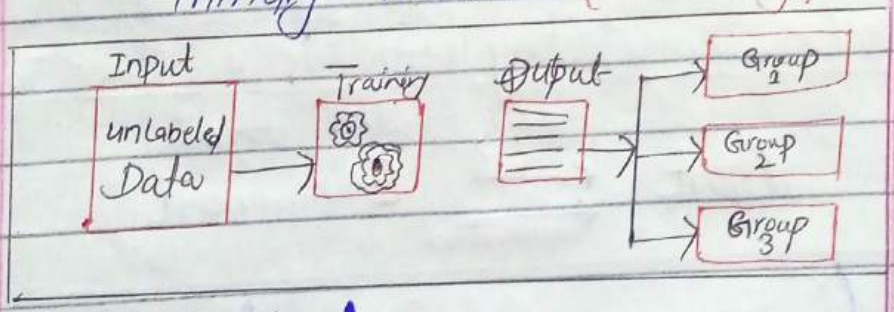
- → K-Nearest Neighbors
- → Support Vector Machines
- → Decision Trees
- → Logistic Regression
- → Linear Regression
- → Random Forest.

Un-Supervised Learning:-

No Labeled data

Model identifiers patterns and groups similar data points

Primary method: (Clustering)



(Application):-

- → Clustering — Grouping similar objects
- → dimensionality — Removing unuseful data while retaining key features
- → Anomaly Detection — identifying unsual [behavior fraud detection]

(Common Algorithms:-)

- K-Means clustering
- Principle Component (PCA)
- Hierarchical clustering.
- DBSCAN
- One-class SVM

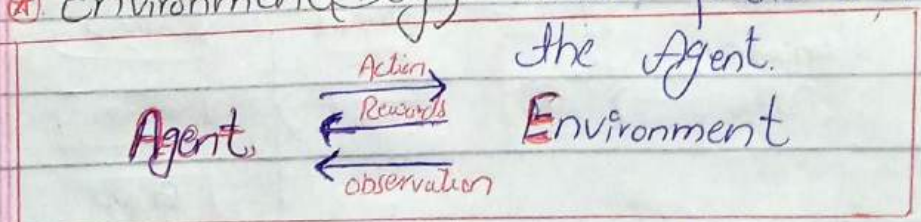
Reginforcement Learning:-

- The model (agent) interacts with an enviment.
- Receives feedback (reward or penalties) based on its action.
- The goal is to maximize cumulative rewards.

Example:-

The learning Process. (boy and dog) Jio

- Agent (Dog):- learns by interacting with the environment.
- Environment (boy):- Provide feedback To the Agent.



(Common Adgorithms:-).

- Q-Learning
- SARSA
- Deep Q-Networks
- Policy Gradient Methods
- Actor-Critic Methods
- Markov Decision Process

Workflow :~

Data Collection	© Gather Relevant data
Data Preprocessing	© Clean & format data
Model Selection	© choose best ML model
Training Model	© Teach the model using data
Evaluation	© Assess model performance
Development	© into integrate Application

Challenges :~

→ Data Imbalance :~

© Some classes have much more data than others

Example :- A class (disease) (1 patient) B class (disease) (9 patient)

→ Features Selection :-

© identifying Relevant features and Removing redundant ones.

Example : Voters data

→ • Overfitting :-

© Model works well on Training data but poorly on new data (Testing) #unseen

Lectures 3 [Attributes]

© All Already have compelcted in "data Science Coures."

Working with datasets in Google Colab

→ Google Colab is a Cloud Based

tool for Running "Python Code" without "Local Setup"

- Open datasets with pd.csv [files.csv]
- Import Libraries like Pandas import pandas as pd.
- Use "Google Drive" to Store and access "datasets"

Lecture No#4

Why Do We need Data Preprocessing:?

We should do befor give To models clean the unnecessary things in our data. like.

Improved Peformance Reduce Complexity
Algorithm Compatibility
Format Suitability. Critical for Sucess.

Data Spliting:- Always used in every ML Model

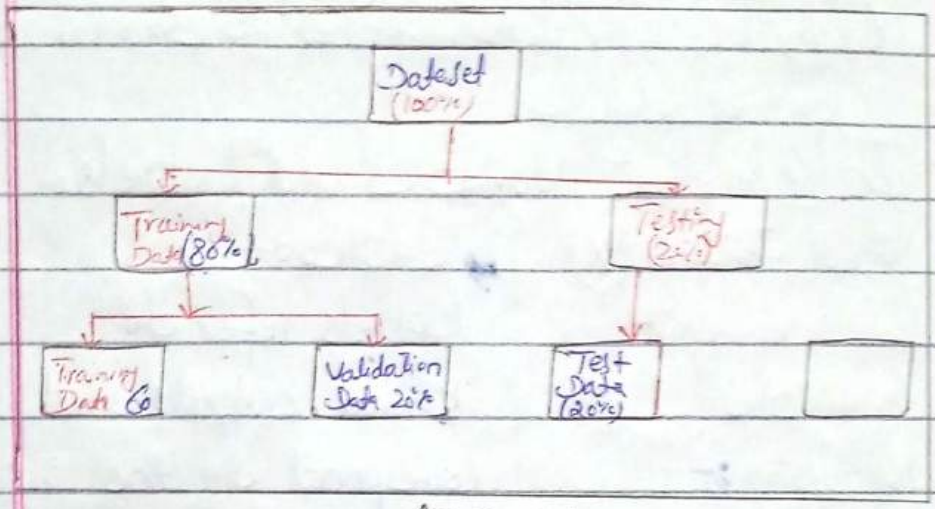
Training data	Testing Data
Used To Train the Models	Used To evaluated the Model Performance

2

Validation Data:-

Used To check for Model overfiting and ensure Accuracy if applicable.

Common Ratio:- 70-30, 80-20, 75-25 85-15 and 90-10



(3 Parts)

⭐ Purposes of Both Data:-

(Training data)	(Testing data)
→ Used for model learning	Used to check
Should be Accurate	the model learning
To ensure good learning	if the Testing
The model "Learns" from	Result are incorrect, its
this data. (70%)	model not "learn" well (30%)

Validation :- It helps in evaluating models Performance.

→ Model Training
→ Initial Split
→ Performance Testing
→ Fine-Tuning

Target Variable:-

The output of Models is called Target variable. [which is predicated]

Prediction Goal:-

Representation is represented by (y.)
Inputs:- is Represented by (X)
Example:- Predicted wether a Smoker or not Smoker.

Independent Variables and Labels:-

- input features (X) are independent variable
- Mapping:- The model learns To Map input (X) To output (y)

- Lables:- Labels refer to known outputs that correspond to the inputs
- Training data:- Contain Labels data for model learning.

Data Cleaning:-

- Accurate data
- Standardization [e.g Ph.d]
- Removing Issue (duplicate Rows)
- Correction
- Critical Step:- Preparing data for Machine learning. Models

Handing Missing Values:-

Are Common in data set.

Issue	Solution	Examples
→ Missing Text data	Assign default values	unKnown for missing City
→ " " Numeric data	Assign Most frequently value (mode)	Most common Salary in dataSet.
→ Numeric data	Assign Mean/median	Average Age for
→ Excessive missing	Removes Rows like del records with Too many miss fields	Missing Age values.

Handing Outliers :-

is a value that are Significantly different from the rest of data. Can distort model Training and accuray.

Methods To Handiles:-

Removing extreme outliers Capping outilers To a Resonbile Range. Transformer data To Reduce the Impact outliers.

Label Encoding:-

@ Assign a unique number To each category. like "Red" as 0 "Blue" as 1 "Green" as 2 all Assign valued are Numerical.

One-Hot-Encoding

Convert Categorical Variables into binary columns.

Example:- Three cities [Lahore, Karachi Islmabad] three columns.

- Lahore [1, 0, 0]
- Karachi [0, 1, 0]
- Islamabad [0, 0, 1]

Ordinal Encoding:-

Assign numerical values to "ordered" categories.

Example:- The order

High = 2 Medium = 1 Low = 0

One-Hot-Encoding

Convert Categorical Variables into binary columns.

Example:- Three cities [Lahore, Karachi Islmabad] three columns.

- Lahore [1, 0, 0]
- Karachi [0, 1, 0]
- Islamabad [0, 0, 1]

Ordinal Encoding:-

Assign numerical values to "ordered" categories.

Example:- The order

High = 2 Medium = 1 Low = 0

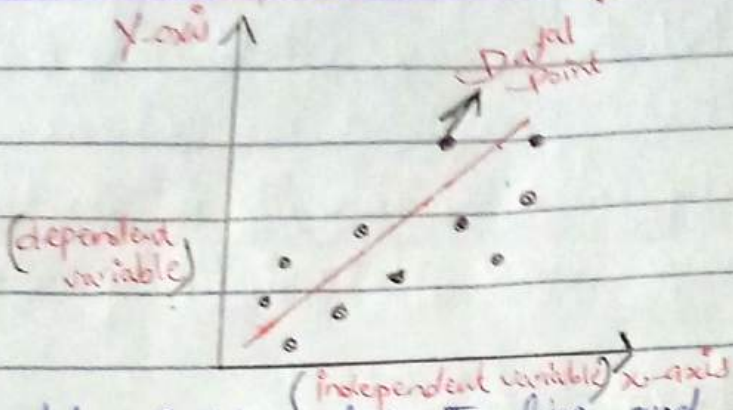
Lectures No# 5+6

Topic:- Data Processing (Part 2) Practical Part

Linear Regression:-

Is a Method that find a Relationship b/w dependent variable (Y) one or more independent variable (x) using a straight line.

This called Simple Linear Regression



Max data - Points clare To cline and fit it (not ideal) Try it

Types:-

→ Simple linear → Mulliple Linear.
→ Polynomial Regression → Ridge Regression
→ Lasso Regression.

Mathemical form:-(SLR)

Least Square Methad

$$SSE = \sum (Y_i - (B_0 + B_1 X_i))^2$$

Minimize SSE To find optimal values for B_0 and B_1

Normal Equations:-

$$B_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}, \quad B_0 = \bar{Y} - B_1 \bar{X}$$

where

→ X_i = independent variable
→ Y_i = predicated B_0, B_1 is constant value

Multiple Linear Regression:

- MLR is a Statistical method used to predict an outcome (dependent variable) based on two or more input features (independent var's)

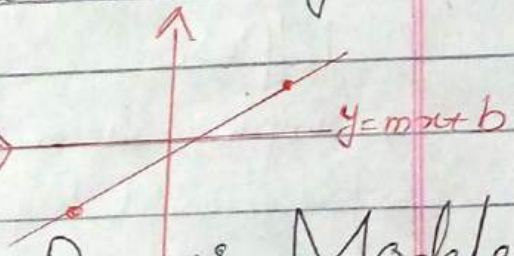
Formula:-

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots b_nX_n + \epsilon$$

Mathematics behind Linear Regression

Slope Equation

$$y = mx + b$$



Evaluating Linear Regression Models

Common Matrics To Access model performance.

★ **Mean Squar Error**

Average of absolute diff b/w prediction and actual values

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

Mean Squared Error

Average of Squared diff b/w predictions and actual values

y_i = Actual, $\hat{y}_i$ = predicted value, Minimize diff b/w

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Root Mean Squared Error

Squar root of MSE provides error in original units

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

R-Squared R^2

Proportion of Variance explained by the model

Making Predictions

The linear Regression prediction process follows these sequential steps.

Train the Model

- fit data To learn Relationships
- Apply Equation
- Use linear Regression Mathematic Equation

Input New data

- Provide unseen values for prediction
- Generate Predictions
- get estimated outcomes

Equaluation Matrix :- Compare prediction with actual values Accuracy.

Making Predictions

The linear Regression prediction process follows these sequential steps

Train the Model	Input New data
◉ fit data To learn Relationships	◉ Provide unseen values for prediction
Apply Equation	Generate Predictions
◉ Use linear Regression Mathematic Equation	◉ Get estimated outcomes

Equvaluation Matrix :- Compare prediction with actual values Accuracy.

"7" is Practical Less

Lecture # 8

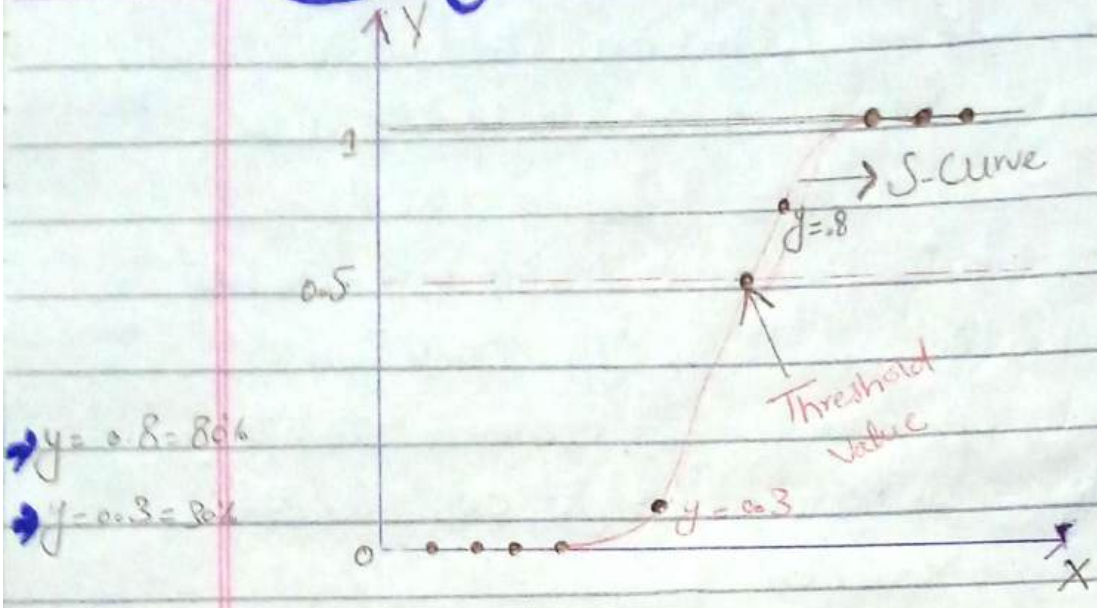
Logistic Regression

◉ Is a Supervised Machine learning algorithm used for classification Task.

→ Where the goal is to predict the probility that an instance belong to a given class or not. Logistic regression is a statistical algorithms

→ which analyze the Relationship b/w Two data factors.

Diagram:-



Use for Binary Classification

- is Type of classification Task where There are only Two possible outcomes (Yes/No, 0/1, True/False)

→ Example:-
Predicting whether an email is Spam (1) or not Spam (0).
cat(1) non cat (0), Pass (1), Fail(0)

Application:-

- Medical diagonasis diseases present (1) or absent (0)
0.5 Threshold for pass (Fail prediction.

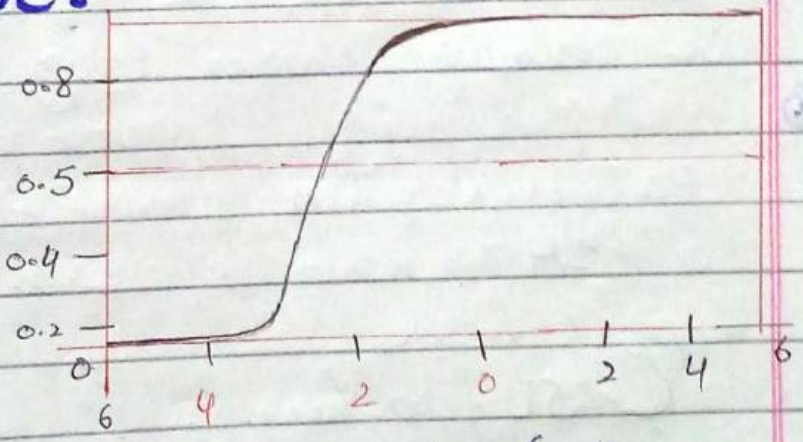
Types of Logistic Regression

Binomial	Multinomial	Ordinal
Two possible outcomes	More than Two outcomes E.g classifying animals as cat, dog, sheep	More Than Two possible outcome but with Some degree.
Pass (1) and fail (0).		Rating as Low, Medium, or High.

Sigmoid function :-

- The Sigmoid function is a mathematical function used To Map the Predicted values To probability The S-form curve is called (Sigmoid function)

figure :-



- We use the concept of the threshold value. Probability (0, 1) above threshold value is (1) and below value is ('0').

- Purpose :- Convert linear Regression output Probability

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

Building a Logistic Regression

→ Import Libraries:- (SKlearn, pandas, numpy)
→ Load and Process data:- handling Missing values, encode categorical variable
→ Split data:- (Train, Test split)
→ Making Predictions:- (predict()) on Test data.
→ Train Model:- Logistic Regression()
→ Evalue Model:- (Accuracy, Precision Recall, F1-score).
→ Optimize & Debop (Tune, hyperparameters handle imbalanced save model).

Challenges:-

- (imbalanced dataset) use Resampling Techniques
- Multicollinearity Features Selection or PCA Two or then Independent variable
- *Overfitting Apply Regularization unseen data pr model Performere is poor.

Cost function

function is used To Minimum distance in

→ Linear Regression
- Mean Squared Error

→ Logistic Regression
- Log Loss Cross Binary Entrop loss

⊛ Confusion Matrix :-

As the name suggests give us a matrix as output and describes the complete performance of the Models.

Predicted values \ Actual values	Positive (1)	Negative (0)
Positive (1)	TP	FP
Negative (0)	FN	TN

→ (True Positive) - The case in Real-world which we predicted Yes and the actual output was also Yes.

Example "Fraud detection" high precision

→ (True Negative) In this case we predicted No and the actual output was No.

To avoid FP (Predicted actually not Fraud) Medical diagnosis

→ (False Negative): The case in which we predicted No and the actual output was Yes.

High recall To min FN) Predicted healthy but actually diseases

→ (False Positive) - The cases in which we predicted Yes and actual output was No.

Model Evaluation :-

ROC-AUC → F1 Score

- Evaluates model performance across diff Threshold.

- Balance precision recall.

$$F1\text{-}Score = \frac{Precision \cdot recall \times 2}{precision + recall}$$

Precision & Recall :-

- Precision :- Predicated positive that are correct $= \frac{TP}{TP+FP}$
- Recall :- Actual Positive Correctly Predicated

$$Recall = \frac{TP}{TP+FN}$$

Accuracy :-

- Overall correctness of the Model.

$$= \frac{TP+TN}{TP+FP+TN+FN}$$

ROC-AUC

- Evaluates model performance across diff Threshold.

F1 Score

- Balance percision recall.

$$\text{F1-Score} = \frac{\text{Precision} \cdot \text{recall} \times 2}{\text{precision} + \text{recall}}$$

Precision & Recall :-

- Precision :- Predicated positive that are correct $= \frac{TP}{TP+FP}$
- Recall :- Actual Positive Correctly Predicated

$$\text{Recall} = \frac{TP}{TP+FN}$$

Accuracy :-

- Overall correctness of the Model.

$$= \frac{TP+TN}{TP+FP+TN+FN}$$

Lectures No# 9+10

Topic :- K-Nearest Neighbor (KNN) Algorithm..

The Algorithm depend upon your Neighbors. Which has Two Terms. Use Bias Variance Tradeoff.

(Bias)	(Variance)
It is the different	Viavance as the model

b/w Actual value & Predicted value	Sensitivity To fluctuations in data
→ when the biase is high, assumptions the model can't capture the important features of our data	→ When Variance is high, our model will capture all features of given data
→ The intence, where model cannt find pattern in our Training set and hence fails for both seen and unseen data is (called underfitting)	→ will tune itself to the data, predicted it very well but when given new data, It cannot predicted on it as is Too Specific To Training data called (Overfitting).

* What is KNN?

→ Simple Superised ML Algorithms. work for both classification and Regression problem.

→ Instance-based learning (no explicit Training phase).

→ Makes predictions based on similarity to known data points

* Basic workflow (Step 1):-

(Initialization)

→ Given a new data point To classify

(Existing Clases)

→ Three existing classes shown (class1, class2, class3)

(Diagram)

* Distance Calculation (Step 2):-

Key Pts: Measure distance To all points
Two Common distance Matrics.

→ Euclidean distance

→ $d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$

Straight line dis b/w pts.

→ Manhattan distance:-

Sum of absolute difference

$$d(x,y) = \sum_{i=1}^{n} |(x_i - y_i)|$$

Value of 'K' must be Taken by odd numbers (3,5,7--)

Sorting Neighbor (Step 3)

→ Sort All points by distance.
Create ordered List from nearest to farthest

→ Examples:-
$d_1, d_2, d_3 \ldots d_n$
etc
Select Top K-neighbors from this Sorted List.

Classification With (KNN)

Topic:- Decision Tree

→ It is most commonly used algorithms in Machine Learning for Supervised Learning. Used To make decision based on the values of different features It operates by spliting data into Subset making a series of decision at each node in th tree. ”

Machine Learning Models

★ Parameteric or Non-Parametric.

→ Parametric

Are those that Assume a specific form or structure for the function. that map into the output. They allowed a fixed num of Parametric Regardless Size of dataset (Linear, Logistic Regression)

→ Non-Parameteric

Are Models that do not assum a fixed number of parametric Instead they allow num of paramters To grow with the amount of Training data (KNN, DT)

★ Entropy:

→ Measures the impurity or Random ness in the data. The lower the entropy the pure the data.

★ Gain information.

→ The iamount of inform gained by Spliting the dataset at a specific features it is calculate (-) entropy of child node from entropy of parent node.

(ID3 or CART) Algorithm.

- Organize decision rules in hierarchical tree structure.

- **Goals:-**
  Predicted target values by learning rules from training data.

- **Benefits:-**
  Easy to learn and good for visualization.

Terminology:-

- **Root Nate**
  Beginning of tree
- **Decision Node**
  Sub-node that can be further split
- **Pruning:-**
  Removing sub-nodes To prevent (overfitting)
- **Spliting**
  Divided nodes into sub nodes.
- **Leaf/Terminal Node**
  find node with no further splits