Lecture: 03

→ **Supervised Learning**

uses labeled data with paired inputs and outputs

• classification • Regression] Types of supervised learning.

→ **Unsupervised Learning**

* no labeled data

* model identify patterns and group similar data points

→ **Reinforcement Learning**

* model (agent) interacts with an environment

* Recieve feedback (reward or penalties) based on its actions

* goal is to maximize rewards

→ **Applications of supervised Learning**

Spam filtering - Audio to text conversion
Language translation - predicting ad clicks - Detecting vehicles using radar - Detecting product defects

→ **Types of supervised learning:**

1) classification (categorizing data into predefined labels)

2) Regression (Predicting continuous values)

→ **Types of classification**

1) Binary classification (two possible outcomes)

2) Multi-class classification (More than two outcomes)

→ **Common Supervised Learning Algorithms**

KNN

Logistic Regression

SVM

~~Logi~~ Linear Regression

Random forest

Decision tree

→ **Applications of unsupervised learning**

clustering

Dimensionality Reduction

Anomaly Detection

→ **Common Algorithms of unsupervised Learning**

K - Mean clustering

DBSCAN

Hierarchial clustering

one - class. SVM

principal component Analysis (PCA)

→ **Common Algorithms of Reinforcement learning**

Q - Learning

SARSA

Deep Q-Network

policy Gradient Methods

Actor critic Method

Markov ~~Nantoo~~ Decision process

→ **Machine Learning   work flow**

① Data collection
② Data preprocessing
③ Model selection
④ Training the model
⑤ Evaluation
⑥ Deployment

→ **challenges in ML**

* Data Imbalance (some classes have much more data than others)
* Feature selection
* overfitting (work well on training data but poor on new da

Lecture: 04

• **what is an Attribute?**

An attribute is a distinct characteristic or property of a data object that can be measured, quantified or categorized.

Attribute → a column in database table

Machine Learning → (feature) an independant variable used for prediction or classification

variable → (statistical term) A measurable property that can take different values

➡ Data object

Sample (in statistics)

Observation (in research)

Data point (in visualization)

➡ Types of attributes

Nominal → (categorical data) (no ordering)

↳ can represent using symbols or codes

↳ not supports mathematical operations

↳ mode ~~is~~ can be ~~use~~ calculated.

Binary → only two values

↳ represented using 0 or 1

↳ Symmetric vs Asymmetric

↓   ↓

Both values   one value is more
~~are important~~   important

ordinal → inherant order or ranking exists

↳ (e.g) grading , ranks

↳ Discretization & ordinal attributes (converting numeric
attributes into an ordinal attribute)

↳ operations (mode , median)

→ arithmetic operations are invalid

Numeric → Quantitative attributes

↳ Ratio scale

↳ interval scale

→ **Discrete attributes**
   finite or countable
   (integer values)

→ **continuous Attribute**
   floating point values

practical work on Google colab. →

Lecture: 05

Data Preprocessing in ML

→ **why do we need data preprocessing?**
* improve performance
* Reduced complexity
* Algorithm compatible
* Ensure that data is in format : suitable for
  model interpretation

→ **Data splitting** (training, testing, validation)

Training data	Testing data.
→ used for model learning	→ used for checking the model learning
→ should be accurate to ensure good learning	→ If testing results incorrect, model has not learned well
→ It could be 80-20, 75-25, 70-30, etc	

→ Target variable selection

X → Input data (Independant variable)

y → target variable

→ Training data
→ Testing data

➡ Data cleaning

* Remove duplicate values / illogical values
* remove incorrect data
* handling missing values etc.
* handling outliers

→ Label encoding :- Assign a unique number to each
  category (e.g) Red → 0 , Blue → 1 , pink → 2 etc

→ one-hot encoding :- creates binary column for each
  category

→ ordinal encoding :- It encode categories by meaningful
  order
      Low → 0, median → 1 , high → 2.

practical work on Google colab ----