

Assignment 1:

Machine Learning Basics:

1. Supervised Machine learning with example

Answer: Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.

In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. It applies the same concept as a student learns in the supervision of the teacher.

Example:

Collect and Prepare the Data

Suppose you have a dataset of houses with features such as:

- Size (in square feet)
- Number of bedrooms
- Number of bathrooms
- Age of the house (in years)

Size (sq ft)	Bedrooms	Bathrooms	Age	Price
1000	3	2	20	300,000
2000	4	3	25	450,000
1500	3	2	70	350,000
1200	2	1	20	200,000
2200	4	3	20	500,000

2. Unsupervised machine learning with example

Answer: Unsupervised learning is a branch of machine learning that deals with unlabeled data. Unlike supervised learning, where the data is labeled with a specific category or outcome, unsupervised learning algorithms are tasked with finding patterns and relationships within the data without any prior knowledge of the data's meaning. This

makes unsupervised learning a powerful tool for exploratory data analysis, where the goal is to understand the underlying structure of the data.

Example:

A is Mall data that contains information about its clients that subscribe to them.

Customer Id	Age	Gender	Salary	Score
1	20	M	2000	C
2	22	M	3000	B
3	23	F	2000	C
4	43	F	3000	B
5	55	F	3500	A

3. Reinforcement learning with example

Answer: Reinforcement learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment. The agent learns to achieve a goal (maximize rewards) through trial and error, where it receives feedback in the form of rewards or penalties.

Example:

Marketing and Advertising

Both brands and consumers can use reinforcement learning to their benefit. For brands selling to target audiences, they can use real-time bidding platforms, A/B testing, and automatic ad optimization. This means that they can place a series of advertisements in the marketplace and the host will automatically serve the best-performing ads in the best spots for the lowest prices. Although brands post and set up the campaigns themselves, Marketing and Advertising Platform are also learning which types of ads are resonating with audiences and will display those ads more frequently and prominently.

4. Classification vs regression vs clustering

Answer:

Classification

Definition: Classification is a supervised learning task where the goal is to predict the categorical class labels of new observations based on past observations with known labels. To classify data points into predefined categories or classes.

Example:

- Spam detection in emails (classify emails into spam or not spam).
- Image recognition (classify images into categories like cats, dogs, cars, etc.).
- Sentiment analysis (classify text as positive, negative, or neutral).

Regression

Definition: Regression is a supervised learning task where the goal is to predict continuous numerical values based on input features. To predict a continuous outcome variable (dependent variable) based on one or more input features (independent variables).

Examples:

Predicting house prices based on features like size, number of rooms, location, etc.

Predicting stock prices based on historical data and market indicators.

Estimating the temperature based on time of day, season, and location.

Clustering

Definition: Clustering is an unsupervised learning task where the goal is to identify groups of similar instances (data points) in the data. To group data points into clusters such that instances within the same cluster are more similar to each other than to those in other clusters.

Examples:

- Customer segmentation based on purchasing behavior.
- Grouping news articles based on their content for topic modeling.
- Segmenting market regions based on demographic data.