Unit-1 - 2 Marks

1. **Machine learning:-**

Machine learning is a branch of Artificial Intelligence. Where the system can learn and improve from the past experience.
It is concerned with the design and development of algorithms

2. **Different types of Machine learning:**

There are mainly 4 types of Machine learning

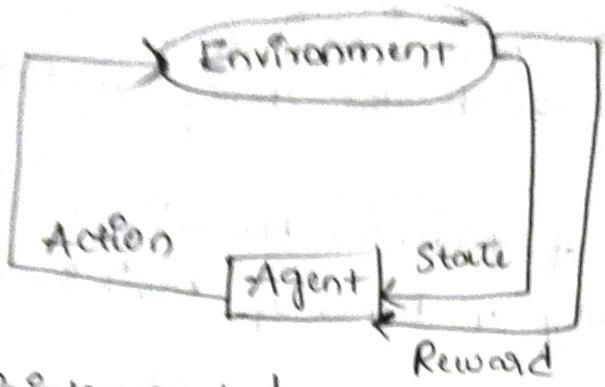
1. Supervised
2. Unsupervised
3. Semi-supervised
4. Reinforcement

3. **Diff b/w Supervised and Unsupervised**

Supervised ML	Unsupervised ML
1. Supervised learning algorithms are trained using labeled data	1. Unsupervised learning algorithm are trained using Unlabeled data
2. SL model takes direct feedback to check if it is predicting correct output or not	2. UL model does not take any feedback
3. SL model predict the output	3. USL model find the hidden patterns in data
4. In SL Input data is provided to the model along with the output	4. In USL only input data is provided to the model
5. SL can be categorized into Classification and Regression	5. USL can be classified in Clustering and Association problems
6. It needs supervision to train the model	6. It does not need any supervision to train the model

4. Reinforcement learning Technique :

Reinforcement learning is a type of machine learning method where an intelligent agent interacts with the environment and learns to act within that



In Reinforcement learning the agent learns automatically using feedbacks without any labeled data

5. Semi-Supervised ML :-

- Semi-supervised learning is a type of machine learning algorithm that represents the intermediate group between Supervised and Unsupervised learning algorithm
- It Uses the combination of labelled and Unlabelled data and the majority will be Unlabelled data

6. Highlight Unsupervised Machine learning techniques :

1. K-means clustering
2. KNN
3. Hierarchal Clustering
4. Anomaly detection
5. Neural Networks
6. Apriori Algorithm
7. principal component Analysis
8. Independent Component Analysis

11- Applications of Machine learning:

1. Online fraud detection
2. Online Customer Support
3. Social media Service
4. Automatic Translation
5. product Recommendation
6. Video surveillance
7. Virtual personal Assistants

12. Three measures Used in Association rule mining:

1. Support
2. Confidence
3. Lift

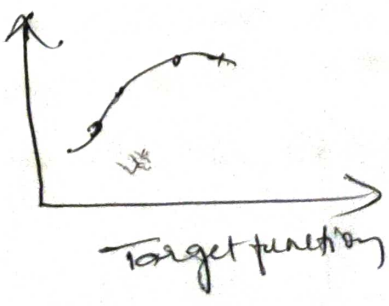
13. Trade-off b/w bias and variance:

Bias: The diff b/w average value predicted by our ML model and correct target value is known as Bias

Variance: The amount of variability in the target function in response to a change in the training data is known as variance

Bias-Variance Trade-off

Bias and variance are complements of each other: the increase of one will result in the decrease of the other vice versa. Hence finding the right balance of values is known as the Bias-Variance Trade-off

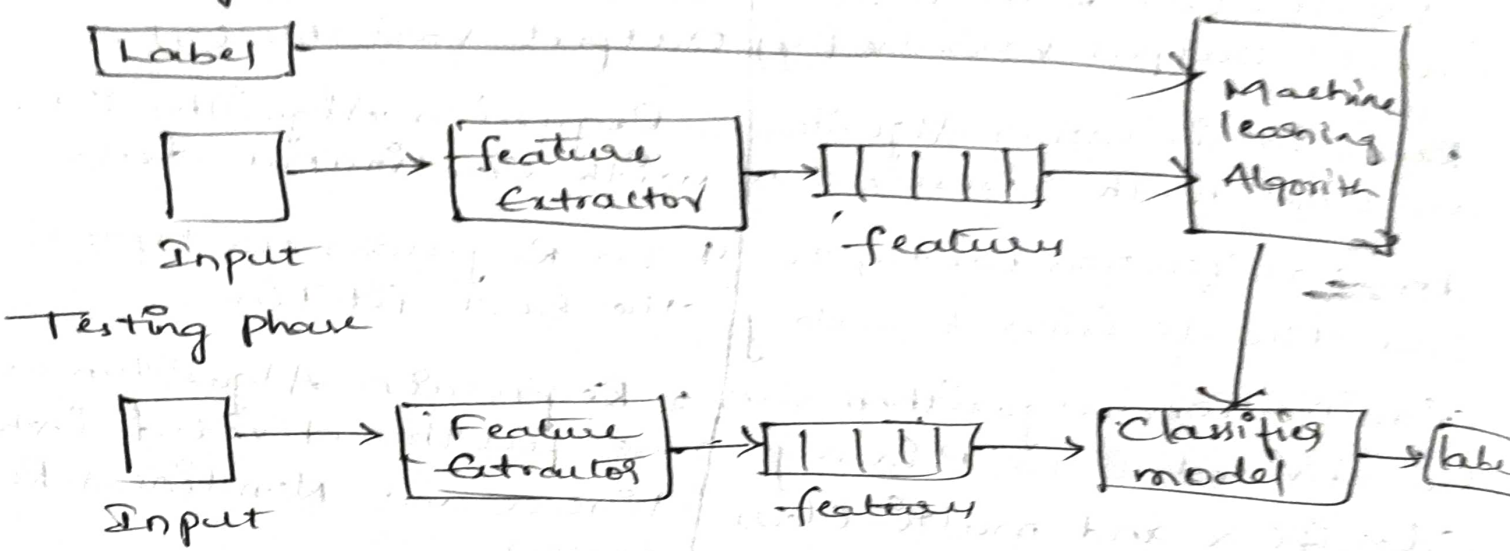


9. Training and Testing phases.

→ Training is the process of making the system able to learn.

Training set and testing set come from the same distribution

Need to make some assumptions or bias



10.

AI	ML	DL

11. Applications of Machine learning:

1. Online - fraud detection
2. Online Customer Support
3. Social media Survey
4. Automatic Translation
5. product Recommendation
6. Video surveillance
7. Virtual personal Assistants

12. Three measures Used in Association rule mining:

1. Support
2. Confidence
3. Lift

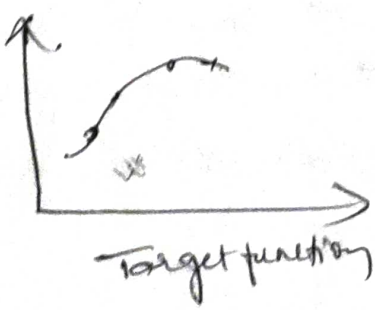
13. Trade-off b/w bias and variance:

Bias: The diff b/w average value predicted by our ML model and correct target value is known as Bias

Variance: The amount of variability in the target function in response to a change in the training data is known as variance

Bias - Variance Trade-off

Bias and variance are complements of each other: The increase of one will result in the decrease of the other vice-versa. Hence finding the right balance of values is known as the Bias - Variance Trade-off



14. What is Overfitting and how can you avoid it?

Overfitting occurs when the model fits more da
than required and it tries to capture each
and every datapoint fed to it.
Overfitted model doesn't perform accurately
It has low bias & high variance.

15. To prevent overfitting several ways that can
be used:
1. Regularization
2. Cross-validation
3. Early stopping
4. feature selection

15. PAC learning:

PAC learning is a framework used for mathemat
-ical analysis
A PAC learner tries to learn a concept by
selecting a hypothesis from a set of hypothesis
that has low generalization error
A good learner should have
high performance
low Error

Unit-2

1. **Bayes theorem :**

Bayes theorem, named after 18th century British mathematician thomas Bayes, is a mathematical formula for determining conditional probability. The theorem provides a way to revise existing prediction or theories given new or additional evidence

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)\,P(B|A)}{P(B)}$$

2. How to Calculate support and confidence with an example:

$$\text{Support}(A) = \frac{\text{No. of transaction in which A appears}}{\text{Total no. of transactions}}$$

$$\text{Confidence}(A \to B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)}$$

3. **Naive Bayes Algorithm:**

It is a supervised algorithm which is based on Bayes theorem and used for solving classification problem

It is one of the simple & most effective Classification algorithm

Naive: It assumes that occurence of a certain feature is independent of the other features

Bayes: It depends on Bayes theorem

4. Parametric methods are used in several ways:
Parametric methods uses a fixed number of parameters to build the model.
It is applicable only for variables
It always considers strong Assumptions about data
It is assumed to be a normal distribution
It have more statistical power.

5- Basic assumptions of the linear regression Algorithm:-
linearity: The relation b/w x and the mean of y is linear
Homoscedasticity: The variance of residual is the same for any value of x
Independence: Observation are independent of each other
Normality: for any fixed value of x, y is normally distributed

6. Tuning complexity:
Tuning is the process of maximizing a model's performance without overfitting or creating too high of a variance in ML.
this is accomplished by selecting hyperparameters

7. list out the necessity enhanced in dimensionality reduction:-
Dimensionality reduction brings many advantages to your machine learning data
Including fewer features mean less complexity

you will need less storage space bcz you have fewer data.

8. purpose of multidimensional scaling:

The purpose of multidimensional scaling is to map the relative location of objects using data that show how the object differ

9. Linear Discriminant Analysis (LDA):

LDA is one of the commonly used dimensionality reduction technique in ML to solve more than two-class classification problems It is also known as Normal discriminant Analysis It is also considered a pre-processing step for modelling differences in ML

10. Quadratic Discriminant

Quadratic discriminant analysis is quite similar to LDA. Except we relaxed the assumption that the mean and Covariance of all the classes were equal.

It is a classical and flexible classification approach which allows diff. b/w groups not only due to mean vectors but also covariance matrices;