📘 Machine Learning Unit II

🎯 PYQ-Based Important Questions – Exam Ready Answers

Q1. Explain Decision Tree using ID3 Algorithm

Definition

Decision Tree is a supervised learning algorithm used for classification. It represents decisions in the form of a tree, where internal nodes represent attributes, branches represent conditions, and leaf nodes represent class labels.

Easy Explanation

Decision Tree works like a flowchart. It asks questions step by step and finally gives an output class.

Example:

```
Is weather sunny?
      ↓
Yes → Play
No  → Don't Play
```

ID3 Algorithm

ID3 stands for **Iterative Dichotomiser 3**. It builds a decision tree using **Information Gain**.

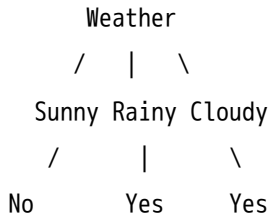
Important Terms

Term	Meaning
Entropy	Measures impurity or uncertainty
Information Gain	Measures how useful an attribute is
Root Node	First decision attribute
Leaf Node	Final output class

Steps of ID3 Algorithm

1. Calculate entropy of the complete dataset.
2. Calculate information gain for each attribute.
3. Select the attribute with highest information gain as root node.
4. Split the dataset based on selected attribute.
5. Repeat the process until all examples are classified.

Diagram



Advantages

- Easy to understand
- Works well for classification
- Requires less data preprocessing
- Can handle categorical data

Disadvantages

- May cause overfitting
- Small data change can change tree
- Not always best for complex data

Applications

- Medical diagnosis
- Loan approval
- Student result prediction
- Email classification

Conclusion

Decision Tree is a simple and powerful classification algorithm. ID3 builds a tree using information gain and selects the best attribute at each step.

Q2. Explain Support Vector Machine / SVM

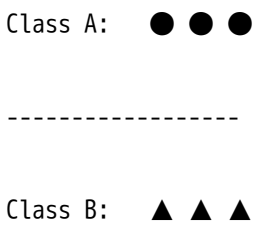
Definition

Support Vector Machine is a supervised learning algorithm used for classification and regression. It finds the best separating line or hyperplane between different classes.

Easy Explanation

SVM tries to draw the best boundary between two classes so that the distance between classes is maximum.

Diagram



Important Terms

Term	Meaning
Hyperplane	Decision boundary
Support Vectors	Nearest data points to hyperplane
Margin	Distance between hyperplane and support vectors
Kernel	Converts non-linear data into higher dimension

Working of SVM

1. Plot data points in feature space.
2. Find possible separating lines.
3. Choose the line with maximum margin.
4. New data is classified based on which side of hyperplane it falls.

Types of SVM

Type	Meaning
Linear SVM	Used when data is linearly separable
Non-linear SVM	Used when data is not linearly separable

Kernel Functions

Kernel functions help SVM handle non-linear data.

Common kernels:

```
Linear Kernel
Polynomial Kernel
RBF Kernel
Sigmoid Kernel
```

Advantages

- High accuracy
- Works well with high-dimensional data
- Effective for classification problems
- Kernel trick handles non-linear data

Disadvantages

- Slow for large datasets
- Difficult to choose proper kernel
- Less suitable for noisy data

Applications

- Face detection
- Text classification
- Handwriting recognition
- Medical diagnosis

Conclusion

SVM is a powerful classification algorithm that finds the optimal hyperplane with maximum margin. It is highly useful for both linear and non-linear classification.

Q3. Explain Neural Networks

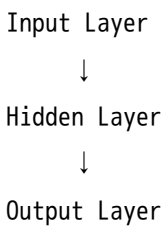
Definition

Neural Network is a machine learning model inspired by the human brain. It consists of connected units called neurons that process information.

Easy Explanation

Just like the human brain learns from experience, neural networks learn from data.

Basic Structure



Components of Neural Network

Component	Meaning
Neuron	Basic processing unit
Weight	Importance of input
Bias	Extra value added to improve learning
Activation Function	Decides output of neuron
Layers	Groups of neurons

Working

1. Input data is given to input layer.
2. Weights are applied to inputs.
3. Hidden layer processes data.
4. Activation function produces output.
5. Final result is generated by output layer.

Example

For image recognition:

```
Image Pixels → Neural Network → Cat/Dog Prediction
```

Advantages

- Learns complex patterns
- Useful for large datasets
- High accuracy
- Works well for image and speech recognition

Disadvantages

- Requires large data
- Training takes time
- Difficult to interpret
- Needs high computation

Applications

- Image recognition
- Speech recognition
- Chatbots
- Medical prediction

- Self-driving cars

Conclusion

Neural Networks are powerful learning models inspired by the brain. They are useful for solving complex real-world problems.

Q4. Explain K-Nearest Neighbors / KNN

Definition

K-Nearest Neighbors is a supervised learning algorithm used for classification and regression. It classifies a new data point based on the majority class of its nearest neighbors.

Easy Explanation

KNN follows the idea:

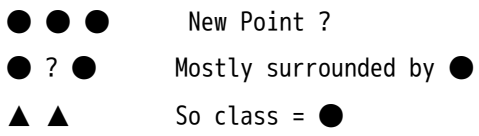
```
A person is known by the company he keeps.
```

If a new point is surrounded mostly by Class A points, it is classified as Class A.

Working of KNN

1. Choose value of K.
2. Calculate distance between new point and all training points.
3. Select K nearest neighbors.
4. Count majority class among neighbors.
5. Assign that class to new point.

Diagram



Distance Formula

Most commonly used distance is Euclidean distance:

$$\text{Distance} = \sqrt{[(x2 - x1)^2 + (y2 - y1)^2]}$$

Example

If K = 3 and nearest neighbors are:

2 students passed

1 student failed

Then new student is classified as:

Pass

Advantages

- Simple and easy to understand
- No training phase
- Works well for small datasets
- Useful for classification

Disadvantages

- Slow for large datasets
- Sensitive to noisy data
- Choosing correct K is important

Applications

- Recommendation systems
- Pattern recognition
- Medical diagnosis
- Image classification

Conclusion

KNN is a simple classification algorithm that predicts output based on nearest neighbors.

Q5. Explain Linear Regression

Definition

Linear Regression is a supervised learning algorithm used to predict continuous values using a linear relationship between input and output.

Easy Explanation

Linear Regression draws a best-fit line through data points and uses that line for prediction.

Formula

```
y = mx + c
```

Where:

Symbol	Meaning
y	Predicted output
x	Input value
m	Slope
c	Intercept

Diagram

```
Price
  |
  |       *
  |    *
  |  *
  |_________________ Area
```

Example

House price prediction:

```
Input: Area of house
Output: Price of house
```

As area increases, price usually increases.

Working

1. Collect training data.
2. Plot data points.
3. Find best-fit line.
4. Use line equation to predict new values.

Advantages

- Simple to understand
- Easy to implement
- Good for prediction problems
- Fast computation

Disadvantages

- Works only for linear relationships
- Sensitive to outliers
- Not suitable for complex data

Applications

- House price prediction
- Sales forecasting
- Salary prediction
- Stock trend analysis

Conclusion

Linear Regression predicts continuous output using a straight-line relationship between input and output.

Q6. Explain Logistic Regression

Definition

Logistic Regression is a supervised learning algorithm used for classification problems. It predicts categorical output such as Yes/No or 0/1.

Easy Explanation

Although its name contains regression, Logistic Regression is mainly used for classification.

Example

```
Disease Prediction → Disease / No Disease
Email → Spam / Not Spam
```

Output

Logistic Regression gives probability between 0 and 1.

```
If probability > 0.5 → Class 1
If probability < 0.5 → Class 0
```

Working

1. Input data is given.
2. Model calculates probability.
3. Sigmoid function converts value between 0 and 1.

4. Final class is predicted.

Difference Between Linear and Logistic Regression

Linear Regression	Logistic Regression
Predicts continuous value	Predicts class label
Output can be any number	Output is between 0 and 1
Used for regression	Used for classification
Example: price prediction	Example: spam detection

Advantages

- Simple classification algorithm
- Gives probability output
- Easy to implement
- Works well for binary classification

Disadvantages

- Not good for complex non-linear data
- Sensitive to irrelevant features
- Assumes relationship between input and output

Applications

- Disease prediction
- Spam detection
- Fraud detection
- Customer churn prediction

Conclusion

Logistic Regression is an important classification algorithm that predicts probability-based class labels.

Q7. Explain CART

Definition

CART stands for **Classification and Regression Trees**. It is a decision tree algorithm used for both classification and regression problems.

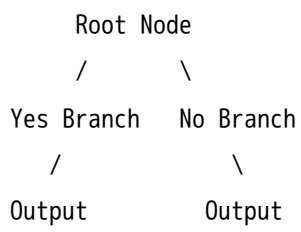
Easy Explanation

CART creates a tree by splitting data into two parts at each step. This is called binary splitting.

Features of CART

Feature	Explanation
Full Form	Classification and Regression Tree
Split Type	Binary split
Used For	Classification and Regression
Measure	Gini Index

CART Structure



CART vs ID3

ID3	CART
Uses Information Gain	Uses Gini Index
Mainly classification	Classification + Regression
Multi-way split possible	Binary split
Works with categorical data	Works with categorical and numerical data

Working

1. Select best attribute using Gini Index.
2. Split data into two parts.
3. Repeat splitting for child nodes.
4. Stop when final class or value is reached.

Advantages

- Works for classification and regression
- Easy to interpret
- Handles numerical data
- Binary split makes structure simple

Disadvantages

- Can overfit data
- Sensitive to small data changes
- May create large tree

Applications

- Loan approval
- Medical prediction
- Sales prediction
- Risk analysis

Conclusion

CART is a powerful decision tree method that supports both classification and regression using binary splitting.

Q8. Explain Perceptron Learning Algorithm

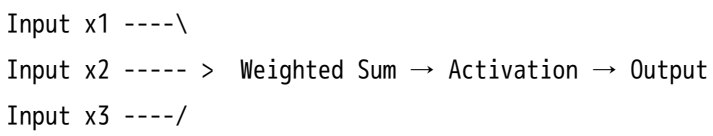
Definition

Perceptron is the simplest type of artificial neural network used for binary classification.

Easy Explanation

A perceptron takes inputs, multiplies them with weights, adds them, and produces an output.

Diagram



Components

Component	Meaning
Inputs	Data values
Weights	Importance of inputs
Bias	Extra adjustment value
Activation Function	Decides final output

Working

1. Inputs are given to perceptron.
2. Each input is multiplied by weight.
3. Weighted values are added with bias.
4. Activation function gives output 0 or 1.
5. If output is wrong, weights are updated.
6. Process repeats until correct classification.

Example

```
Input: Marks, Attendance
Output: Pass / Fail
```

Advantages

- Simple neural network model
- Easy to implement
- Useful for binary classification
- Foundation of neural networks

Disadvantages

- Cannot solve non-linear problems
- Works only for linearly separable data
- Limited learning capacity

Applications

- Binary classification
- Pattern recognition

- Basic neural network learning
- Simple prediction tasks

Conclusion

Perceptron is the basic building block of neural networks. It performs binary classification using weights, bias and activation function