UNIT-III

Fill in the Blanks

1. Decision trees are used for both classification and _____ tasks.
   Answer: Regression
2. In a decision tree, the topmost node is called the _____ node.
   Answer: Root
3. CART stands for Classification and _____ Trees.
   Answer: Regression
4. Ensemble learning combines multiple _____ to improve performance.
   Answer: Models
5. Bagging helps reduce model _____.
   Answer: Variance
6. Boosting focuses on correcting the errors made by previous _____.
   Answer: Models
7. K-means is an example of _____ learning.
   Answer: Unsupervised
8. In K-means clustering, data points are grouped into _____ clusters.
   Answer: K
9. Nearest Neighbor methods classify samples based on the _____ data points.
   Answer: Closest
10. Gaussian Mixture Models are based on Gaussian _____ distributions.
    Answer: Probability
11. Decision trees split data using _____ rules.
    Answer: Decision
12. In ensemble learning, classifiers are combined to improve _____.
    Answer: Accuracy
13. The K in K-means represents the number of _____.
    Answer: Clusters
14. Basic statistics includes measures such as mean, median, and _____.
    Answer: Mode
15. Boosting mainly reduces _____ errors.
    Answer: Bias
16. In bagging, training samples are selected using _____ sampling.
    Answer: Random
17. Gaussian Mixture Models are commonly abbreviated as _____.
    Answer: GMM
18. The leaf nodes of a decision tree represent the final _____.
    Answer: Decision
19. Nearest Neighbor algorithms are examples of _____ learning methods.
    Answer: Instance-based
20. Unsupervised learning works with _____ data.
    Answer: Unlabeled

Multiple Choice Questions (MCQs)

1. Which algorithm is commonly used for classification and regression?
   a) Decision Tree
   b) Bubble Sort
   c) DFS
   d) Stack
   Answer: a) Decision Tree
2. CART stands for:
   a) Classification and Regression Trees
   b) Clustering and Ranking Trees
   c) Classification and Rule Testing
   d) Combined Analysis Regression Technique
   Answer: a) Classification and Regression Trees
3. Which learning method combines multiple models?
   a) Reinforcement Learning
   b) Ensemble Learning
   c) Supervised Learning
   d) Deep Learning
   Answer: b) Ensemble Learning
4. Bagging mainly helps in reducing:
   a) Bias
   b) Variance
   c) Dimensions
   d) Noise
   Answer: b) Variance
5. Boosting mainly improves weak learners by:
   a) Ignoring errors
   b) Combining outputs randomly
   c) Focusing on previous mistakes
   d) Increasing dimensions
   Answer: c) Focusing on previous mistakes
6. K-means clustering belongs to:
   a) Supervised Learning
   b) Reinforcement Learning
   c) Unsupervised Learning
   d) Semi-supervised Learning
   Answer: c) Unsupervised Learning
7. In K-means, cluster centers are called:
   a) Medians
   b) Centroids
   c) Nodes
   d) Edges
   Answer: b) Centroids
8. Which method classifies data based on nearby examples?
   a) Decision Tree
   b) K-Nearest Neighbor
   c) Linear Regression
   d) Apriori
   Answer: b) K-Nearest Neighbor

9. Gaussian Mixture Models are based on which distribution?
   a) Uniform Distribution
   b) Bernoulli Distribution
   c) Gaussian Distribution
   d) Poisson Distribution
   Answer: c) Gaussian Distribution
10. Which node represents the final output in a decision tree?
   a) Root Node
   b) Parent Node
   c) Leaf Node
   d) Hidden Node
   Answer: c) Leaf Node
11. Which of the following is a clustering algorithm?
   a) K-means
   b) Linear Regression
   c) Perceptron
   d) Naive Bayes
   Answer: a) K-means
12. Ensemble methods improve performance by combining multiple:
   a) Databases
   b) Features
   c) Learners
   d) Processors
   Answer: c) Learners
13. In bagging, models are trained on:
   a) Same dataset only
   b) Random subsets of data
   c) Sorted datasets
   d) Small datasets only
   Answer: b) Random subsets of data
14. Which algorithm is often used in nearest neighbor classification?
   a) KNN
   b) PCA
   c) DBSCAN
   d) Apriori
   Answer: a) KNN
15. Which learning method does NOT require labeled data?
   a) Supervised Learning
   b) Unsupervised Learning
   c) Reinforcement Learning
   d) Active Learning
   Answer: b) Unsupervised Learning
16. The root node of a decision tree contains:
   a) Final output
   b) First decision/split
   c) Hidden values
   d) Errors
   Answer: b) First decision/split
17. Which of the following is an ensemble technique?
   a) Bagging

b) Boosting
c) Random Forest
d) All of the above
Answer: d) All of the above

18. Basic statistics is important in machine learning for:
    a) Data analysis
    b) Data visualization only
    c) Hardware optimization
    d) Gaming
    Answer: a) Data analysis
19. Gaussian Mixture Models are mainly used for:
    a) Clustering
    b) Sorting
    c) Encryption
    d) Networking
    Answer: a) Clustering
20. The objective of K-means is to minimize:
    a) Classification accuracy
    b) Distance within clusters