

CLASS TEST  
ML ADVANCED  
DATED- 28/10/2023

1. Which algorithm does Gradient Boosting use as its base learner?  
**a) Decision Trees**  
b) Support Vector Machines  
c) K-Nearest Neighbors
2. CATBoost is known for handling:  
**a) Categorical Features**  
b) Continuous Features  
c) Both a and b
3. Apriori algorithm is used for:  
**a) Association Rule Mining**  
b) Regression Analysis  
c) Image Recognition
4. RFM model is commonly used in:  
**a) Customer Segmentation**  
b) Text Classification  
c) Time Series Analysis
5. In XGBoost, what does the term "boosting" refer to?  
a) Increasing model bias  
**b) Combining weak learners to create a strong learner**  
c) Reducing model complexity
6. ADABOOST assigns:  
a) Equal weights to all data points  
**b) Higher weights to misclassified data points**  
c) Higher weights to correctly classified data points
7. What is the key difference between Gradient Boosting and ADABOOST?  
a) The choice of weak learners  
**b) The way weights are assigned to data points**  
c) The number of iterations
8. CATBoost is particularly useful when dealing with:  
a) Numerical features  
**b) Categorical features**  
c) Binary features
9. LightGBM is known for its:  
**a) Leaf-wise tree growth**  
b) Level-wise tree growth  
c) Balanced tree growth
10. KNN is a:  
**a) Supervised learning algorithm**  
b) Unsupervised learning algorithm

c) Reinforcement learning algorithm

11. What is the main advantage of using Support Vector Machines (SVM)?

- a) Effective in high-dimensional spaces
- b) Handles non-linear relationships well
- c) Both a and b**

12. RFM model stands for:

- a) Recency, Frequency, Monetary**
- b) Regression, Feature selection, Modeling
- c) Random Forest Metrics

13. In XGBoost, what is the purpose of the learning rate parameter?

- a) Controls the step size in gradient descent**
- b) Determines the number of boosting rounds
- c) Adjusts the weights of misclassified instances

14. CATBoost automatically handles:

- a) Missing values
- b) Outliers
- c) Both a and b**

15. LightGBM is efficient in terms of:

- a) Memory usage
- b) Computation speed
- c) Both a and b**

16. Support Vector Machines aim to find the hyperplane that:

- a) Maximizes the margin between classes**
- b) Minimizes the margin between classes
- c) Is perpendicular to the feature space

17. In SVR, what is the role of the epsilon parameter?

- a) Controls the width of the margin**
- b) Controls the trade-off between smoothness and accuracy
- c) Determines the number of support vectors

18. ADABOOST assigns higher weights to:

- a) Misclassified data points**
- b) Correctly classified data points
- c) Outliers

19. LightGBM is designed for:

- a) High-dimensional data**
- b) Low-dimensional data
- c) One-dimensional data

20. KNN is sensitive to:

- a) Feature scaling
- b) Outliers**
- c) Missing values

21. In SVM, the kernel function is used to:

- a) **Project data into a higher-dimensional space**
- b) Reduce the dimensionality of the data
- c) Normalize the data

22. LightGBM uses a histogram-based learning approach, which is efficient for:

- a) Small datasets
- b) **Large datasets**
- c) Balanced datasets

23. KNN is a non-parametric algorithm, meaning:

- a) It makes assumptions about the underlying data distribution
- b) **It does not make assumptions about the underlying data distribution**
- c) It relies on a fixed set of parameters

24. Apriori algorithm is used to discover:

- a) **Hidden patterns in data**
- b) Optimal hyperparameters
- c) Feature importance

25. In XGBoost, what is the significance of the subsample parameter?

- a) **The fraction of samples used for fitting the individual base learners**
- b) The number of features to consider when making a split
- c) The minimum loss reduction required to make a further partition on a leaf node