### **Easy MCQs:**

1. **Which of the following is an example of supervised learning?**a) K-Means Clustering  
   b) Linear Regression  
   c) PCA  
   d) K-NN  
   **Answer:** b) Linear Regression
2. **Which of the following is a classification algorithm?**a) Linear Regression  
   b) K-Means  
   c) Logistic Regression  
   d) PCA  
   **Answer:** c) Logistic Regression
3. **What is the main purpose of the feature engineering process in ML?**a) Improve model performance  
   b) Collect more data  
   c) Create a dataset for validation  
   d) Split the data into training and test sets  
   **Answer:** a) Improve model performance
4. **Which metric is typically used for evaluating classification models?**a) Mean Squared Error  
   b) Accuracy  
   c) R-Squared  
   d) MAPE  
   **Answer:** b) Accuracy
5. **Which of the following is an unsupervised learning algorithm?**a) Logistic Regression  
   b) Decision Trees  
   c) K-Means Clustering  
   d) Naïve Bayes  
   **Answer:** c) K-Means Clustering
6. **Which of the following techniques can help with multicollinearity?**a) Feature scaling  
   b) Lasso regression  
   c) Decision trees  
   d) Normalization  
   **Answer:** b) Lasso regression
7. **In which of the following does a model learn from labeled data?**a) Unsupervised Learning  
   b) Reinforcement Learning  
   c) Supervised Learning  
   d) Semi-Supervised Learning  
   **Answer:** c) Supervised Learning
8. **Which algorithm is most suitable for large datasets and noisy data?**a) Naïve Bayes  
   b) Decision Trees  
   c) K-Nearest Neighbors  
   d) Random Forest  
   **Answer:** d) Random Forest
9. **Which metric is used to evaluate the performance of regression models?**a) Precision  
   b) Accuracy  
   c) Root Mean Squared Error (RMSE)  
   d) Recall  
   **Answer:** c) Root Mean Squared Error (RMSE)
10. **In k-means clustering, what does the 'k' represent?**a) Number of clusters  
    b) Number of attributes  
    c) Number of features  
    d) Number of training instances  
    **Answer:** a) Number of clusters
11. **Which technique is used for reducing the dimensionality of data?**a) Regression  
    b) Principal Component Analysis (PCA)  
    c) Decision Trees  
    d) Naïve Bayes  
    **Answer:** b) Principal Component Analysis (PCA)
12. **Which of the following is an example of a regression problem?**a) Predicting whether a customer will buy a product (Yes/No)  
    b) Predicting the price of a house based on features  
    c) Predicting whether an email is spam  
    d) Clustering customers based on purchasing behavior  
    **Answer:** b) Predicting the price of a house based on features
13. **In Logistic Regression, what is the output variable?**a) Continuous variable  
    b) Binary categorical variable  
    c) Multiclass categorical variable  
    d) Multivariate variable  
    **Answer:** b) Binary categorical variable
14. **Which of the following is an ensemble method?**a) Naïve Bayes  
    b) Support Vector Machine  
    c) Random Forest  
    d) K-Means  
    **Answer:** c) Random Forest
15. **Which of the following is a clustering algorithm?**a) Naïve Bayes  
    b) Decision Trees  
    c) K-Means  
    d) Linear Regression  
    **Answer:** c) K-Means
16. **What does a confusion matrix evaluate?**a) Model’s accuracy  
    b) Model’s bias  
    c) Performance of a classification model  
    d) Performance of a regression model  
    **Answer:** c) Performance of a classification model
17. **Which of the following is a hyperparameter in a decision tree model?**a) Maximum depth  
    b) Data normalization  
    c) Features selection  
    d) Number of iterations  
    **Answer:** a) Maximum depth
18. **Which algorithm is used for dimensionality reduction?**a) Decision Trees  
    b) K-Nearest Neighbors  
    c) Principal Component Analysis (PCA)  
    d) Naïve Bayes  
    **Answer:** c) Principal Component Analysis (PCA)
19. **What is the purpose of cross-validation in machine learning?**a) To split data into training and testing sets  
    b) To evaluate model performance and prevent overfitting  
    c) To choose the best model  
    d) To normalize the data  
    **Answer:** b) To evaluate model performance and prevent overfitting
20. **Which of the following is an example of an outlier detection method?**a) K-Means Clustering  
    b) Decision Trees  
    c) Isolation Forest  
    d) Naïve Bayes  
    **Answer:** c) Isolation Forest

### **Medium MCQs:**

1. **What is the objective of the Naïve Bayes classifier?**a) Minimize error  
   b) Classify instances based on conditional probability  
   c) Maximize error  
   d) Partition data into clusters  
   **Answer:** b) Classify instances based on conditional probability
2. **Which of the following is an advantage of using Random Forests?**a) High interpretability  
   b) It handles missing data well  
   c) It is fast for training  
   d) It always performs better than SVM  
   **Answer:** b) It handles missing data well
3. **Which of the following techniques can help to improve the performance of k-Nearest Neighbors (K-NN)?**a) Cross-validation  
   b) Feature scaling  
   c) Removing outliers  
   d) All of the above  
   **Answer:** d) All of the above
4. **What does the Gini index measure in a decision tree?**a) Model accuracy  
   b) Impurity of a node  
   c) Number of nodes  
   d) Entropy of the dataset  
   **Answer:** b) Impurity of a node
5. **Which of the following statements is true about the K-Means algorithm?**a) It works well with non-convex data  
   b) The number of clusters is fixed after initialization  
   c) It does not require the number of clusters to be predefined  
   d) It is sensitive to the initial placement of centroids  
   **Answer:** d) It is sensitive to the initial placement of centroids
6. **Which algorithm is commonly used for classification tasks where classes are linearly separable?**a) Decision Trees  
   b) k-Nearest Neighbors  
   c) Support Vector Machine  
   d) Naïve Bayes  
   **Answer:** c) Support Vector Machine
7. **What is the main advantage of using L1 regularization (Lasso) over L2 regularization (Ridge)?**a) L1 regularization improves prediction accuracy  
   b) L1 regularization can lead to sparse models by setting some coefficients to zero  
   c) L2 regularization prevents overfitting  
   d) L1 regularization works better for large datasets  
   **Answer:** b) L1 regularization can lead to sparse models by setting some coefficients to zero
8. **What is a typical use case of a decision tree model?**a) Clustering data points  
   b) Time series prediction  
   c) Predicting categorical outcomes  
   d) Regressing continuous variables  
   **Answer:** c) Predicting categorical outcomes
9. **Which of the following methods is most suitable for classification with missing data?**a) Linear Regression  
   b) Decision Trees  
   c) K-Means  
   d) Naïve Bayes  
   **Answer:** b) Decision Trees
10. **What does the term "overfitting" refer to in machine learning?**a) The model is too simple  
    b) The model performs well on unseen data  
    c) The model fits the training data too closely and fails to generalize  
    d) The model uses too few features  
    **Answer:** c) The model fits the training data too closely and fails to generalize
11. **Which technique can be used for model evaluation and validation?**a) Cross-validation  
    b) Bootstrapping  
    c) Holdout method  
    d) All of the above  
    **Answer:** d) All of the above
12. **Which ensemble method works by combining the outputs of multiple models to improve predictive performance?**a) Random Forest  
    b) Naïve Bayes  
    c) K-Means  
    d) Linear Regression  
    **Answer:** a) Random Forest
13. **Which of the following is true about Support Vector Machines?**a) SVM is a supervised learning algorithm  
    b) SVM performs better for regression than classification  
    c) SVM is not suitable for linearly separable data  
    d) SVM uses decision trees for classification  
    **Answer:** a) SVM is a supervised learning algorithm
14. **What does the term "model bias" refer to?**a) The model’s ability to generalize  
    b) The error introduced due to overly simplistic assumptions  
    c) The model's ability to fit complex data  
    d) The model’s complexity  
    **Answer:** b) The error introduced due to overly simplistic assumptions
15. **Which of the following is not a clustering technique?**a) DBSCAN  
    b) K-Means  
    c) Decision Trees  
    d) Hierarchical Clustering  
    **Answer:** c) Decision Trees
16. **What is the objective of unsupervised learning?**a) To learn from labeled data  
    b) To predict continuous values  
    c) To find hidden patterns and groupings in data  
    d) To improve the model's accuracy  
    **Answer:** c) To find hidden patterns and groupings in data
17. **Which of the following is not a characteristic of K-Nearest Neighbors?**a) Instance-based learning  
    b) It requires labeled data  
    c) It can be used for both classification and regression  
    d) It builds a decision tree  
    **Answer:** d) It builds a decision tree
18. **What is the purpose of regularization in linear regression?**a) Reduce bias  
    b) Prevent overfitting by penalizing large coefficients  
    c) Improve accuracy on the test set  
    d) Increase variance  
    **Answer:** b) Prevent overfitting by penalizing large coefficients
19. **Which of the following algorithms is used for regression tasks?**a) K-Nearest Neighbors  
    b) Naïve Bayes  
    c) Linear Regression  
    d) DBSCAN  
    **Answer:** c) Linear Regression
20. **Which ensemble technique combines weak models to create a strong model by iteratively improving predictions?**a) Bagging  
    b) Boosting  
    c) Random Forest  
    d) K-Means  
    **Answer:** b) Boosting

### **Hard MCQs:**

1. **Which of the following is a major advantage of using Principal Component Analysis (PCA) for dimensionality reduction?**a) It reduces the number of features while retaining most of the information  
   b) It increases the dimensionality of the data  
   c) It helps in classifying the data  
   d) It removes noise from the data  
   **Answer:** a) It reduces the number of features while retaining most of the information
2. **Which algorithm is not sensitive to the scaling of data?**a) K-Nearest Neighbors  
   b) Support Vector Machine  
   c) Decision Trees  
   d) Logistic Regression  
   **Answer:** c) Decision Trees
3. **Which metric is used in evaluating clustering algorithms?**a) F1 Score  
   b) Homogeneity  
   c) Precision  
   d) R-Squared  
   **Answer:** b) Homogeneity
4. **Which of the following is the primary drawback of the k-Nearest Neighbors (K-NN) algorithm?**a) It is computationally expensive for large datasets  
   b) It does not work well with high-dimensional data  
   c) It is not suitable for regression problems  
   d) It cannot be used for classification  
   **Answer:** a) It is computationally expensive for large datasets
5. **What is the purpose of using decision trees for classification problems?**a) To reduce the dimensionality of data  
   b) To model non-linear relationships between variables  
   c) To predict continuous variables  
   d) To perform regression  
   **Answer:** b) To model non-linear relationships between variables
6. **In the context of Naïve Bayes, which of the following assumptions is made about the features?**a) Features are dependent on each other  
   b) Features are mutually exclusive  
   c) Features are independent of each other given the class  
   d) Features are correlated  
   **Answer:** c) Features are independent of each other given the class
7. **Which of the following methods can be used to improve the performance of decision trees?**a) Cross-validation  
   b) Bagging  
   c) Pruning  
   d) All of the above  
   **Answer:** d) All of the above
8. **Which of the following ensemble methods is based on bootstrapping?**a) AdaBoost  
   b) Random Forest  
   c) Gradient Boosting  
   d) XGBoost  
   **Answer:** b) Random Forest
9. **Which algorithm can automatically select the optimal number of clusters in a dataset?**a) K-Means  
   b) DBSCAN  
   c) Hierarchical Clustering  
   d) Gaussian Mixture Model  
   **Answer:** b) DBSCAN
10. **What does the confusion matrix's 'True Positive' represent?**a) Correctly predicted negative cases  
    b) Incorrectly predicted negative cases  
    c) Correctly predicted positive cases  
    d) Incorrectly predicted positive cases  
    **Answer:** c) Correctly predicted positive cases
11. **Which of the following is an advantage of using the Ridge regression technique?**a) It can be used for categorical target variables  
    b) It helps to reduce overfitting in models with high multicollinearity  
    c) It allows for all coefficients to be set to zero  
    d) It is suitable for data with missing values  
    **Answer:** b) It helps to reduce overfitting in models with high multicollinearity
12. **Which of the following methods is used for handling missing values in the dataset?**a) Cross-validation  
    b) Mean imputation  
    c) Feature scaling  
    d) Gradient boosting  
    **Answer:** b) Mean imputation
13. **Which of the following models works best for large and unstructured data?**a) K-Nearest Neighbors  
    b) Support Vector Machines  
    c) Random Forests  
    d) Naïve Bayes  
    **Answer:** c) Random Forests
14. **Which of the following algorithms is used for classification and regression tasks in machine learning?**a) Decision Trees  
    b) K-Nearest Neighbors  
    c) Naïve Bayes  
    d) Principal Component Analysis  
    **Answer:** a) Decision Trees
15. **In a regression model, what is the purpose of the R-squared value?**a) It evaluates how well the model generalizes to unseen data  
    b) It measures the strength of the relationship between the independent and dependent variables  
    c) It indicates how much the variance in the independent variable is explained by the dependent variable  
    d) It measures the model’s bias  
    **Answer:** b) It measures the strength of the relationship between the independent and dependent variables
16. **Which of the following is a disadvantage of decision trees?**a) Prone to overfitting  
    b) They cannot handle categorical data  
    c) They do not scale well with large datasets  
    d) They are slow to train  
    **Answer:** a) Prone to overfitting
17. **Which of the following is the best approach to handle imbalanced datasets?**a) Use cross-validation  
    b) Use class weighting  
    c) Use decision trees  
    d) Use feature scaling  
    **Answer:** b) Use class weighting
18. **Which of the following models can be used to predict probabilities of an event?**a) Decision Trees  
    b) Linear Regression  
    c) Logistic Regression  
    d) K-Means  
    **Answer:** c) Logistic Regression
19. **Which of the following statements is true about Gradient Boosting?**a) It creates an ensemble of weak learners iteratively  
    b) It combines predictions from multiple decision trees by averaging  
    c) It is based on the concept of bagging  
    d) It does not require any hyperparameter tuning  
    **Answer:** a) It creates an ensemble of weak learners iteratively
20. **In which of the following situations is the Naïve Bayes classifier most effective?**a) When features are strongly correlated  
    b) When the data is normally distributed  
    c) When the dataset is large and contains noisy data  
    d) When features are independent  
    **Answer:** d) When features are independent
21. **What is the purpose of the cost function in machine learning algorithms?**a) To measure the accuracy of the model  
    b) To optimize the model’s parameters during training  
    c) To split the data into training and test sets  
    d) To calculate the training time  
    **Answer:** b) To optimize the model’s parameters during training
22. **What is the main advantage of using an ensemble method like Random Forest over a single decision tree?**a) It reduces bias  
    b) It increases interpretability  
    c) It reduces variance and overfitting  
    d) It speeds up training time  
    **Answer:** c) It reduces variance and overfitting
23. **Which of the following algorithms does not require feature scaling?**a) K-Nearest Neighbors  
    b) Support Vector Machines  
    c) Linear Regression  
    d) Decision Trees  
    **Answer:** d) Decision Trees
24. **What is the typical approach to improve the performance of a machine learning model in the presence of outliers?**a) Removing outliers from the data  
    b) Increasing the model complexity  
    c) Using a robust model like Random Forest  
    d) Decreasing the number of features  
    **Answer:** a) Removing outliers from the data
25. **What does the term "Bias-Variance Tradeoff" refer to in machine learning?**a) Increasing the model complexity reduces both bias and variance  
    b) Reducing model complexity increases bias but reduces variance  
    c) High bias leads to overfitting, and high variance leads to underfitting  
    d) Bias and variance have no effect on model performance  
    **Answer:** b) Reducing model complexity increases bias but reduces variance
26. **Which evaluation metric is preferred for imbalanced classification problems?**a) Accuracy  
    b) Precision  
    c) F1 Score  
    d) ROC AUC  
    **Answer:** c) F1 Score
27. **Which of the following can be a result of multicollinearity in linear regression?**a) Increased prediction accuracy  
    b) Unstable and unreliable coefficient estimates  
    c) Reduced R-Squared value  
    d) Reduced model complexity  
    **Answer:** b) Unstable and unreliable coefficient estimates
28. **In the context of a Naïve Bayes classifier, what assumption is made about the features?**a) Features are linearly correlated  
    b) Features are independent given the class  
    c) Features have a non-Gaussian distribution  
    d) Features have equal importance  
    **Answer:** b) Features are independent given the class
29. **Which of the following is true about Support Vector Machines (SVM)?**a) SVM can only be used for binary classification  
    b) SVM uses a loss function called Hinge loss  
    c) SVM works well with categorical features without preprocessing  
    d) SVM is not effective with high-dimensional data  
    **Answer:** b) SVM uses a loss function called Hinge loss
30. **What is the purpose of the hyperparameter tuning process in machine learning?**a) To minimize the training time  
    b) To find the best configuration of model parameters  
    c) To perform feature selection  
    d) To reduce the size of the dataset  
    **Answer:** b) To find the best configuration of model parameters
31. **Which of the following methods can be used for dimensionality reduction?**a) Principal Component Analysis (PCA)  
    b) K-Nearest Neighbors (K-NN)  
    c) Random Forest  
    d) Support Vector Machines (SVM)  
    **Answer:** a) Principal Component Analysis (PCA)
32. **What is the main purpose of regularization in linear regression?**a) To make the model more complex  
    b) To penalize large coefficients and prevent overfitting  
    c) To speed up the training process  
    d) To handle multicollinearity  
    **Answer:** b) To penalize large coefficients and prevent overfitting
33. **Which of the following is the best metric to evaluate a classification model when the classes are imbalanced?**a) Accuracy  
    b) Precision  
    c) Recall  
    d) ROC AUC  
    **Answer:** d) ROC AUC