#1. Which technique is used to reduce the risk of overfitting in decision trees?

☐
Pruning
☐
Grafting
☐
Bagging
☐
Boosting
☐
Random selection of features

#2. What does the term “bias-variance tradeoff” refer to in machine learning?

☐
The balance between model complexity and the amount of training data
☐
The balance between underfitting and overfitting
☐
The tradeoff between the accuracy of a model and its interpretability
☐
The tradeoff between bias (error due to too simple model) and variance (error due to too complex model)
☐
The tradeoff between the number of features and the number of samples in the dataset

#3. What is the purpose of a confusion matrix in classification problems?

☐
To visualize the performance of a model
☐
To display the distribution of classes in a dataset

☐
To evaluate the quality of features
☐
To quantify the performance of a classification model
☐
To calculate the probability of class membership

#4. Which algorithm is commonly used for collaborative filtering in recommendation systems?

☐
Matrix Factorization
☐
K-Nearest Neighbors (KNN)
☐
Support Vector Machines (SVM)
☐
Decision Trees
☐
Random Forest

#5. What is the role of a learning rate in gradient descent optimization?

☐
It controls the size of the steps taken during optimization
☐
It determines the number of training samples
☐
It specifies the number of features in the dataset
☐
It defines the number of iterations in training
☐
It adjusts the complexity of the model

#6. In a neural network, what is the purpose of the activation function?

☐
To introduce non-linearity into the model
☐
To reduce the number of parameters
☐
To improve the interpretability of the model
☐
To speed up the training process
☐
To add noise to the data

#7. Which type of machine learning algorithm is well-suited for recommendation systems?

☐
Collaborative Filtering
☐
Decision Trees
☐
Support Vector Machines (SVM)
☐
K-Means Clustering
☐
Linear Regression

#8. What is the objective of Principal Component Analysis (PCA) in dimensionality reduction?

☐
To project data onto a lower-dimensional subspace while retaining the most important information
☐

To increase the number of features in the dataset
☐
To separate data into distinct clusters
☐
To minimize the training time
☐
To transform data into a non-linear space

#9. What is the main advantage of using an ensemble learning method like Random Forest?

☐
It reduces overfitting and increases model robustness
☐
It simplifies the model and improves interpretability
☐
It speeds up the training process
☐
It eliminates the need for feature engineering
☐
It requires fewer training samples

#10. In which phase of a machine learning project is cross-validation typically applied?

☐
Model evaluation
☐
Data preprocessing
☐
Model training
☐
Model deployment
☐

Data collection

#11. What is the purpose of L1 regularization (Lasso) in linear regression?

☐
It encourages sparsity by penalizing the absolute values of coefficients
☐
It increases the number of features in the model
☐
It adds noise to the data
☐
It reduces the complexity of the model
☐
It enforces non-negativity of coefficients

#12. Which algorithm is used for finding frequent itemsets in association rule mining?

☐
Apriori algorithm
☐
K-Means Clustering
☐
Naive Bayes
☐
Decision Trees
☐
Support Vector Machines (SVM)

#13. What is the objective of batch normalization in deep learning?

☐
To stabilize and speed up the training process
☐
To reduce the number of parameters in the model

☐
To increase the learning rate
☐
To add noise to the data
☐
To simplify the model

#14. Which technique is used to combat the class imbalance problem in classification tasks?

☐
Resampling (e.g., oversampling or undersampling)
☐
Bagging
☐
Pruning
☐
Feature selection
☐
Dimensionality reduction

#15. What is the purpose of the Mean-Shift clustering algorithm?

☐
Identifying dense regions of data points in feature space
☐
Reducing model complexity
☐
Performing image segmentation
☐
Classifying data into predefined categories
☐
Increasing the learning rate

#16. Which algorithm is commonly used for text classification in natural language processing (NLP)?

☐
Support Vector Machines (SVM)
☐
Random Forest
☐
Logistic Regression
☐
Naive Bayes
☐
Decision Trees

#17. What is the purpose of the term “dropout” in deep learning?

☐
To randomly deactivate neurons during training to prevent overfitting
☐
To increase the learning rate
☐
To add noise to the data
☐
To reduce model complexity
☐
To improve the interpretability of the model

#18. What is the primary objective of the Expectation-Maximization (EM) algorithm in unsupervised learning?

☐
Estimating the parameters of a probabilistic model in the presence of hidden variables
☐
Reducing model complexity

☐
Performing feature selection
☐
Adding non-linearity to the model
☐
Performing clustering

#19. Which technique is used for feature importance ranking in Random Forest models?

☐
Gini Impurity or Mean Decrease in Gini
☐
Recursive Feature Elimination (RFE)
☐
Principal Component Analysis (PCA)
☐
Support Vector Machines (SVM)
☐
Singular Value Decomposition (SVD)

#20. What is the purpose of the term “regularization” in machine learning?

☐
To prevent overfitting by adding a penalty term to the loss function
☐
To increase the number of features in the model
☐
To reduce the learning rate
☐
To control the number of samples in the training set
☐
To add noise to the data

Finish

Results