**Section A (20 question)**

**1. Which supervised learning algorithm is used for regression tasks?**

a) Linear Regression b) Logistic Regression

c) Decision Trees d) Support Vector Machines

**2. What is the primary evaluation metric for regression models?**

a) Accuracy b) Precision c) R-squared d) Confusion Matrix

**3. Which library can be used for implementing linear and multi-linear regression in Python?**

a) TensorFlow b) Keras c) scikit-learn d) PyTorch

**4. What is the purpose of logistic regression?**

a) Regression analysis b) Binary classification c) Multiclass classification d) Clustering

**5. Which evaluation metric is used for classification models to measure the model's performance in terms of true positives and true negatives?**

a) R-squared b) Mean Squared Error c) Confusion Matrix d) Precision

**6. Support Vector Machines (SVM) can be used for:**

a) Classification only b) Regression only

c) Both Classification and Regression d) Clustering

**7. What is a decision tree commonly used for?**

a) Regression b) Clustering c) Classification d) Dimensionality Reduction

**8. Which metric measures the proportion of correctly identified positive cases from all actual positives in a classification model?**

a) Precision b) Recall c) Accuracy d) F1-score

**9. Which of the following is NOT an evaluation metric for classification models?**

a) Mean Squared Error b) Precision c) Recall d) Accuracy

**10. In linear regression, what is the goal of the model?**

a) Minimize the sum of squared residuals b) Maximize the likelihood function

c) Minimize the classification error d) Maximize the R-squared value

**11. Logistic Regression is a type of:**

a) Linear model b) Non-linear model c) Ensemble model d) Dimensionality reduction model

**12. Which of the following is used to evaluate the goodness of fit in regression models?**

a) F1-score b) R-squared c) Recall d) Precision

**13. SVM aims to find the:**

a) Line that best separates classes b) Curve that best fits the data

c) Hyperplane that best separates classes d) Decision boundary with maximum margin

**14. Decision trees are prone to:**

a) Overfitting b) Underfitting

c) Both overfitting and underfitting d) Neither overfitting nor underfitting

**15. The primary goal of decision tree algorithms is to:**

a) Maximize accuracy b) Minimize impurity

c) Minimize training time d) Maximize the number of leaves

**16. Which of the following metrics is calculated as the square of the correlation between the predicted and actual values in regression?**

a) R-squared b) Mean Absolute Error

c) Root Mean Squared Error d) Mean Squared Error

**17. What does the confusion matrix in classification models represent?**

a) Predicted versus actual class labels b) Sum of true positives and true negatives

c) Percentage of correctly classified instances d) Mean Squared Error values

**18. What is the significance of the term "kernel" in Support Vector Machines?**

a) It represents the support vectors b) It maps the original data into a higher-dimensional space

c) It represents the margin of the hyperplane d) It is used to regularize the SVM model

**19. Which of the following is NOT a common type of kernel function in SVM?**

a) Linear b) Polynomial c) Radial Basis Function (RBF) d) Exponential

**20. What is the purpose of pruning in decision trees?**

a) To increase the depth of the tree b) To decrease the depth of the tree

c) To reduce overfitting d) To increase computational complexity

**Section B(15 questions)**

**1. Which of the following metrics is calculated as the square of the correlation between the predicted and actual values in regression?**

a) R-squared b) Mean Squared Error

c) Root Mean Squared Error d) Mean Absolute Error

**2. What does the confusion matrix in classification models represent?**

a) Predicted versus actual class labels b) Sum of true positives and true negatives

c) Percentage of correctly classified instances d) Mean Squared Error values

**3. How is the precision calculated in a classification model?**

a) True Positives / (True Positives + False Positives)

b) True Positives / (True Positives + False Negatives)

c) True Positives / (True Positives + True Negatives)

d) True Negatives / (True Negatives + False Positives)

**4. What is the significance of the term "kernel" in Support Vector Machines?**

a) It represents the support vectors b) It maps the original data into a higher-dimensional space

c) It represents the margin of the hyperplane d) It is used to regularize the SVM model

**5. Which of the following is NOT a common type of kernel function in SVM?**

a) Linear b) Polynomial c) Radial Basis Function (RBF) d) Exponential

**6. What is the purpose of pruning in decision trees?**

a) To increase the depth of the tree b) To decrease the depth of the tree

c) To reduce overfitting d) To increase computational complexity

**7. How does logistic regression handle multicollinearity among predictor variables?**

a) It drops one of the correlated variables randomly

b) It assigns equal weights to correlated variables

c) It is not affected by multicollinearity

d) It adjusts the coefficients to account for multicollinearity

**8. What is the formula for the Mean Squared Error (MSE) in regression?**

a) 1/n ∑(yi - ŷi)^2 b) 1/n ∑|yi - ŷi|

c) √(1/n ∑(yi - ŷi)^2) d) 1/n ∑log(yi - ŷi)^2

**9. Which of the following is a disadvantage of decision tree algorithms?**

a) They are computationally expensive b) They cannot handle missing values

c) They are not interpretable d) They are not prone to overfitting

**10. Which metric combines precision and recall into a single value for evaluating classification models?**

a) F1-score b) R-squared c) Mean Absolute Error d) Root Mean Squared Error

**11. How does regularization help in logistic regression?**

a) It simplifies the decision boundary b) It prevents overfitting by penalizing large coefficients

c) It increases model complexity d) It improves convergence speed

**12. In SVM, what is the role of the margin?**

a) It represents the distance between support vectors

b) It measures the width of the decision boundary

c) It is the hyperplane that maximizes the margin

d) It is the distance between the decision boundary and the support vectors

**13. What is the main advantage of using decision trees for classification?**

a) They handle non-linear relationships well b) They are computationally efficient

c) They are robust to outliers d) They can handle high-dimensional data

**14. Which of the following is a measure of the homogeneity of the nodes in a decision tree?**

a) Gini impurity b) Entropy c) Information gain d) Mean Squared Error

**15. How does the complexity parameter affect the decision tree model?**

a) It controls the maximum depth of the tree b) It determines the number of leaf nodes

c) It regulates the trade-off between bias and variance

d) It adjusts the weights of the predictor variables