Multiple Choice Questions:

1. What is Machine Learning?

A. A type of artificial intelligence that enables machines to learn from data without being explicitly programmed

B. A form of manual data analysis

C. A set of instructions given to a computer to solve a problem

D. A method of downloading information from the internet

2. Which of the following is a Supervised Learning algorithm?

A. K-Means

B. Support Vector Machine (SVM)

C. Apriori

D. DBSCAN

3. Which of these is an example of Unsupervised Learning?

A. Logistic Regression

B. Random Forest

C. K-Means Clustering

D. Linear Regression

4. In which scenario would you use Classification rather than Regression?

A. Predicting the price of a house based on area

B. Predicting whether an email is spam or not

C. Predicting a student’s score in an exam

D. Predicting the number of cars in a parking lot

5. Which of the following statements is false regarding Machine Learning?

A. Machine learning models need large amounts of data to perform well

B. Machine learning models are always 100% accurate

C. Machine learning can be used to make predictions based on patterns in data

D. Machine learning requires some form of data preprocessing

6. What is Overfitting in machine learning?

A. A model that performs well on the training data but poorly on new, unseen data

B. A model that generalizes well to new data

C. A model that has not been trained enough

D. A model with fewer parameters than required

7. Which of the following is a dimensionality reduction technique?

A. Support Vector Machine (SVM)

B. K-Means

C. Principal Component Analysis (PCA)

D. Decision Tree

8. Which of the following is an advantage of using K-Nearest Neighbors (KNN)?

A. It’s computationally efficient for large datasets

B. It is a non-parametric algorithm, so it doesn’t make strong assumptions about the data

C. It automatically selects the best number of neighbors (K)

D. It doesn’t require any distance metric

9. Which evaluation metric is used to measure the accuracy of a regression model?

A. Confusion Matrix

B. Mean Squared Error (MSE)

C. F1-Score

D. Recall

10. Which type of machine learning problem is the Apriori algorithm used for?

A. Classification

B. Regression

C. Association Rule Learning

D. Clustering

True/False Questions:

11. Linear Regression is a supervised learning algorithm.

True / False

12. The more complex a model is, the better it will generalize to unseen data.

True / False

13. Unsupervised learning requires labeled data.

True / False

14. Gradient Descent is an optimization algorithm commonly used in machine learning to minimize the loss function.

True / False

15. Random Forest is an ensemble learning method that combines multiple decision trees.

True / False

Short Answer Questions:

16. What is the difference between supervised and unsupervised learning?

17. What is a hyperparameter, and how does it differ from a model parameter?

18. Give an example of a real-world application of machine learning in healthcare.

19. What is the purpose of splitting a dataset into training and testing sets?

20. Explain the concept of cross-validation in machine learning.