Machine Learning Theory Exam – 2A

1. Describe shortly the process of creating a machine learning model.

2. List and briefly explain the most common evaluation metrics for regression problems.

3. Why do we need to split the data into training and testing sets?

4. What is unsupervised learning in machine learning?

5. Mention one regression algorithm and briefly describe how it works.

6. Explain the concept of overfitting in machine learning. How can we detect and avoid them?

7. Which of the following is a supervised learning algorithm?  
 - A) Linear Regression  
 - B) K-Means Clustering  
 - C) Decision Trees  
 - D) Support Vector Machines

8. In a classification problem, what does a low recall indicate?  
 - A) Low false negatives  
 - B) Low false positives  
 - C) High true negatives

9. What is the main goal of a classification model?  
 - A) Categorizing data into classes  
 - B) Predicting a continuous value  
 - C) Reducing the size of data  
 - D) Increasing the dimensionality

10. Cancer Detection - Real-Life Scenario

Total patients: 200

Patients with cancer: 20

Model predicted 16 have cancer (12 correct, 4 wrong)

Missed 8 real cancer cases

Calculate: accuracy, precision and recall (positive – has cancer)