**Supervised Learning - Test 1**:

### **Section 1: Linear Regression**

**Q1. What is Linear Regression?**  
**Answer:** Linear Regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data.

**Q2. Write the formula for simple linear regression.**  
**Answer:**

y = \beta\_0 + \beta\_1x + \varepsilon

* = dependent variable
* = independent variable
* = intercept
* = slope coefficient
* = error term

**Q3. What is the purpose of the cost function in linear regression?**  
**Answer:** The cost function measures how well the model’s predictions match the actual data. In linear regression, it is typically the Mean Squared Error (MSE), which is minimized to find the best-fitting line.

**Q4. How do you interpret the coefficients in a multiple linear regression model?**  
**Answer:** Each coefficient represents the expected change in the dependent variable for a one-unit change in the corresponding independent variable, holding all other variables constant.

**Q5. What are the assumptions of Linear Regression?**  
**Answer:**

1. Linearity
2. Independence of errors
3. Homoscedasticity (constant variance of errors)
4. Normality of errors
5. No multicollinearity among independent variables

### **Section 2: Gradient Descent**

**Q6. What is Gradient Descent?**  
**Answer:** Gradient Descent is an optimization algorithm used to minimize the cost function by iteratively adjusting the model parameters in the direction of the steepest descent (negative gradient).

**Q7. Write the formula for parameter update in Gradient Descent.**  
**Answer:**

\theta := \theta - \alpha \cdot \frac{\partial J(\theta)}{\partial \theta}

* = parameter
* = learning rate
* = cost function

**Q8. What is the role of the learning rate in Gradient Descent?**  
**Answer:** The learning rate determines the step size at each iteration while moving toward a minimum of the cost function. It controls how fast or slow the model learns.

**Q9. What is the primary purpose of regularization in machine learning models?**  
**Answer: C)** To prevent overfitting by penalizing large coefficients in the model.

**Q10. What happens if the learning rate is too small or too large?**  
**Answer:**

* Too small: Slow convergence, may take a long time to reach the minimum.
* Too large: May overshoot the minimum or diverge altogether.

### **Section 3: Bias & Variance**

**Q11. Define Bias and Variance in the context of machine learning models.**  
**Answer:**

* **Bias** is the error due to overly simplistic assumptions in the model.
* **Variance** is the error due to sensitivity to small fluctuations in the training set.

**Q12. What is the Bias-Variance tradeoff?**  
**Answer:** It is the balance between bias and variance where decreasing one typically increases the other. The goal is to find a model with low total error (bias² + variance + irreducible error).

**Q13. How does increasing the complexity of a model affect bias and variance?**  
**Answer:** Increasing complexity usually decreases bias but increases variance, potentially leading to overfitting.

**Q14. What is underfitting and overfitting in machine learning?**  
**Answer:**

* **Underfitting:** Model is too simple, fails to capture the underlying pattern (high bias).
* **Overfitting:** Model is too complex, captures noise as well as signal (high variance).

**Q15. How can you reduce overfitting in a model?**  
**Answer:**

* Use regularization (L1, L2)
* Cross-validation
* Pruning (for trees)
* Reduce model complexity
* More training data
* Early stopping (for iterative models)