SUPERVISED LEARNING - TEST 1 (20 to 30 mins)

15-Question Test on Linear Regression, Gradient Descent, Bias & Variance

Topics:

 Linear Regression:

 Gradient Descent:

 Bias & Variance:



Section 1: Linear Regression

Q1. What is Linear Regression?   
Answer: Linear Regression is a supervised learning algorithm used to model the relationship between a dependent variable (target) and one or more independent variables (features) by fitting a linear equation to the observed data.

Q2. Write the formula for simple linear regression.

Answer: y=β0​+β1​x+ε

Where:

* yyy = predicted value
* xxx = input feature
* β0\beta\_0β0​ = intercept
* β1\beta\_1β1​ = slope (coefficient)
* ε\varepsilonε = error term

Q3. What is the purpose of the cost function in linear regression? Answer: The cost function (usually Mean Squared Error) measures how well the model's predictions match the actual values. It guides the optimization process to minimize prediction errors.

Q4. How do you interpret the coefficients in a multiple linear regression model?

Answer: Each coefficient represents the expected change in the target variable for a one-unit change in the corresponding feature, assuming all other features remain constant.

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

 Linearity

 Independence of errors

 Homoscedasticity (constant variance of errors)

 Normal distribution of errors

 No multicollinearity among features



Section 2: Gradient Descent

Q6. What is Gradient Descent?

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

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

θ:=θ−α⋅∂θ∂J(θ)​

Where:

* θ\thetaθ = model parameter
* α\alphaα = learning rate
* J(θ)J(\theta)J(θ) = cost function

Q8. What is the role of the learning rate in Gradient Descent? Answer: The learning rate controls the size of each update step. A small learning rate results in slow convergence, while a large learning rate may cause overshooting or divergence.



Q9. What is the primary purpose of regularization in machine learning models?

A) To increase the complexity of the model to fit the training data better.

B) To minimize the training error without regard to generalization.

C) To prevent overfitting by penalizing large coefficients in the model.

D) To ensure that all features are included in the final model regardless of their importance.

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 and long training time.

**Too large:** The algorithm may overshoot the minimum or diverge entirely.



Section 3: Bias & Variance

Q11. Define Bias and Variance in the context of machine learning models. Answer:  **Bias:** Error due to overly simplistic assumptions. High bias causes underfitting.

 **Variance:** Error due to high sensitivity to training data. High variance causes overfitting.

Q12. What is the Bias-Variance tradeoff?

Answer: It’s the balance between bias and variance. Increasing model complexity reduces bias but increases variance. The goal is to find a balance that minimizes total error.

Q13. How does increasing the complexity of a model affect bias and variance?

Answer: Increasing complexity typically **decreases bias** but **increases variance**.

Q14. What is underfitting and overfitting in machine learning? Answer:  **Underfitting:** The model is too simple and performs poorly on training and test data.

 **Overfitting:** The model is too complex and performs well on training data but poorly on test data.

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

 Use regularization (L1/L2)

 Use simpler models

 Gather more training data

 Apply dropout (in neural networks)

 Use cross-validation