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 and one or more independent variables by fitting a linear equation to observed data.

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

y=β0​+β1​x+ϵ  
where yyy is the dependent variable, xxx is the independent variable, β0\beta\_0β0​ is the intercept, β1\beta\_1β1​ is the slope (coefficient), and ϵ\epsilonϵ is the error term.

Q3. What is the purpose of the cost function in linear regression? Answer:

The cost function measures how well the regression model fits the data by quantifying the difference between the predicted values and actual values. It guides the optimization process to find the best model parameters by minimizing this error, commonly using Mean Squared Error (MSE).

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 increase in the corresponding independent variable, assuming all other variables remain constant.

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

 Linearity: Relationship between independent and dependent variables is linear.

 Independence: Observations are independent of each other.

 Homoscedasticity: Constant variance of errors across all levels of independent variables.

 Normality: Errors are normally distributed.

 No multicollinearity: Independent variables are not highly correlated.



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:

θ:=θ−α⋅∇θ​J(θ)

where θ\thetaθ is the parameter vector, α\alphaα is the learning rate, and ∇θJ(θ)\nabla\_\theta J(\theta)∇θ​J(θ) is the gradient of the cost function with respect to θ\thetaθ.

Q8. What is the role of the learning rate in Gradient Descent? Answer:

The learning rate controls the size of the steps taken toward the minimum of the cost function during each iteration. It balances convergence speed and stability.



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.

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, longer training time.

 Too large: Can cause overshooting the minimum or divergence, leading to unstable training.



Section 3: Bias & Variance

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

* **Bias:** Error due to overly simplistic assumptions in the model, causing it to miss relevant relations (underfitting).
* **Variance:** Error due to model sensitivity to small fluctuations in the training data, causing overfitting.

Q12. What is the Bias-Variance tradeoff?

Answer:

The tradeoff between bias and variance describes the balance where reducing bias increases variance and vice versa. The goal is to find a model complexity 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:** When a model is too simple to capture the underlying pattern of data, leading to poor performance on training and test data.

 **Overfitting:** When a model learns the noise or details in the training data too well, performing well on training but poorly on unseen data.

Q15. How can you reduce overfitting in a model?

Answer:  Use regularization (L1, L2).

 Increase training data or use data augmentation.

 Use dropout (in neural networks).

 Simplify the model architecture.

 Use early stopping during training.