

# Regression

## Interview Questions

(Practice Project)



## Easy Level Questions

### Q1. What is Simple Linear Regression?

**Ans1:** Simple Linear Regression is a method to model the relationship between a dependent variable ( $Y$ ) and a single independent variable ( $X$ ) using a straight line. The equation is  $Y = \beta_0 + \beta_1 X + \epsilon$  where  $\beta_0$  is the intercept,  $\beta_1$  slope, and  $\epsilon$  is the error term.

### Q2. What are the assumptions of Simple Linear Regression?

**Ans2:** The assumptions are:

- **Linearity:** The relationship between  $X$  and  $Y$  is linear.
- **Independence:** Observations are independent of each other.
- **Homoscedasticity:** Constant variance of errors.
- **Normality:** Errors are normally distributed.

### Q3. What is Gradient Descent?

**Ans3:** Gradient Descent is an optimization algorithm used to minimize the loss function by iteratively updating the model parameters in the direction of the steepest descent.

### Q4. What is the difference between Batch Gradient Descent and Stochastic Gradient Descent (SGD)?

**Ans4:** Batch Gradient Descent uses the entire dataset to compute the gradient, while Stochastic Gradient Descent (SGD) updates the parameters using one training example at a time.

### Q5. What is the purpose of the Learning Rate in Gradient Descent?

**Ans5:** The learning rate controls the step size during the parameter update. A small learning rate ensures slow convergence, while a large learning rate may cause overshooting.

### Q6. What does R-squared represent in a regression model?

**Ans6:** R-squared represents the proportion of variance in the dependent variable that is explained by the independent variables in the model. It ranges from 0 to 1.

### Q7. What is a cost function?

**Ans7:** A cost function is a function that measures the performance of a model for given data. It outputs a single number representing how bad the model's predictions are compared to the actual outcomes.

### Q8. What does the term 'overfitting' mean in machine learning?

**Ans:** Overfitting occurs when a model learns the noise in the training data, leading to poor generalization to new, unseen data.

## Medium Level Questions

### Q9. What is the difference between Simple and Multiple Linear Regression?

**Ans9:** Simple Linear Regression involves one independent variable, while Multiple Linear Regression involves two or more independent variables to predict the dependent variable.

### Q10. Explain the concept of Regularization in machine learning.

**Ans10:** Regularization adds a penalty to the loss function to prevent overfitting by discouraging overly complex models. Common techniques include L1 (Lasso) and L2 (Ridge) regularization.

### Q11. How does the Gradient Descent algorithm find the global minimum of a function?

**Ans11:** Gradient Descent finds the global minimum by iteratively updating the parameters in the direction of the negative gradient of the loss function. The process stops when the gradient is close to zero, indicating a minimum.

### Q12. What is the significance of the bias term in a linear model?

**Ans12:** The bias term (intercept) allows the model to shift the regression line up or down to better fit the data. It represents the value of the dependent variable when all independent variables are zero.

### Q13. What is a learning curve, and why is it useful?

**Ans13:** A learning curve plots the model's performance (e.g., loss or accuracy) over time or training iterations. It helps in diagnosing problems like overfitting or underfitting.

### Q14. What is Ridge Regression, and when is it used?

**Ans14:** Ridge Regression is a type of linear regression that includes L2 regularization. It is used when there is multicollinearity, or to prevent overfitting by penalizing large coefficients.

### Q15. What is the difference between Lasso and Ridge Regression?

**Ans15:** Lasso (L1) regression uses L1 regularization, which can shrink some coefficients to zero, effectively selecting features. Ridge (L2) regression uses L2 regularization, which penalizes large coefficients but doesn't eliminate them.

### Q16. What are the different types of Gradient Descent, and how do they differ?

**Ans16:** The three types are:

- **Batch Gradient Descent:** Uses the entire dataset to compute the gradient.
- **Stochastic Gradient Descent (SGD):** Updates parameters using one training example at a time.
- **Mini-Batch Gradient Descent:** Uses a subset of the data to compute the gradient, combining the benefits of Batch and SGD.

## Hard Level Questions

**Q17. How would you handle multicollinearity in a regression model?**

**Ans17: Multicollinearity can be handled by:**

- Removing highly correlated predictors.
- Using Ridge Regression (L2 regularization) which penalizes large coefficients.
- Applying Principal Component Analysis (PCA) to reduce dimensionality.

**Q18. Explain the difference between Ridge, Lasso, and Elastic Net Regression.**

**Ans18:**

- **Ridge Regression:** L2 regularization, penalizes the square of the coefficients.
- **Lasso Regression:** L1 regularization, penalizes the absolute value of the coefficients and can shrink some to zero.
- **Elastic Net Regression:** Combines both L1 and L2 regularization, allowing for both feature selection and shrinkage.

**Q19. How would you evaluate a regression model's performance if the data has outliers?**

**Ans19:** If the data has outliers, you can evaluate the model using metrics that are less sensitive to outliers, such as the Mean Absolute Error (MAE) or Median Absolute Error, instead of Mean Squared Error (MSE).