**Interview Questions**

**1. How does logistic regression differ from linear regression?**

* **Linear regression** predicts continuous values (e.g., price, age).
* **Logistic regression** predicts **categorical outcomes** (e.g., yes/no, 0/1).
* It uses the **sigmoid function** to convert output into probabilities between 0 and 1.

**2. What is the sigmoid function?**

The **sigmoid function** is an S-shaped curve that maps any real number to a value between 0 and 1.  
[  
\sigma(x) = \frac{1}{1 + e^{-x}}  
]  
It helps logistic regression interpret outputs as **probabilities**.

**3. What is precision vs recall?**

* **Precision** = Correct positive predictions ÷ All predicted positives  
  → Measures how accurate positive predictions are.
* **Recall** = Correct positive predictions ÷ All actual positives  
  → Measures how well the model identifies all actual positives.

**4. What is the ROC-AUC curve?**

* **ROC (Receiver Operating Characteristic)** curve shows the trade-off between **True Positive Rate** and **False Positive Rate**.
* **AUC (Area Under Curve)** measures model performance — a higher AUC means better classification ability.

**5. What is the confusion matrix?**

A **confusion matrix** shows model predictions vs actual values:

|  | **Predicted Positive** |  | **Predicted Negative** |
| --- | --- | --- | --- |
| **Actual Positive** | True Positive (TP) |  | False Negative (FN) |
| **Actual Negative** | False Positive (FP) |  | True Negative (TN) |

* + It helps evaluate model performance.

**6. What happens if classes are imbalanced?**

If one class dominates, the model may **ignore minority classes**, leading to poor recall.  
Solutions:

* Use **resampling** (oversampling minority / under sampling majority)
* Use **class weights** or **F1-score** for evaluation.

**7. How do you choose the threshold?**

By default, logistic regression uses **0.5** as a probability threshold.  
You can adjust it using **ROC curve** or **Precision-Recall curve** to balance **precision** and **recall** based on business needs.

**8. Can logistic regression be used for multi-class problems?**

Yes.  
It can be extended using:

* **One-vs-Rest (OvR)** – trains one classifier per class.
* **Multinomial (Softmax)** – predicts probabilities for all classes at once.