

Module 14 – Quiz and Explanation

1. What is the primary purpose of Logistic Regression?

- a) Predict continuous values
- b) Predict binary class labels
- c) Perform clustering
- d) Reduce dimensionality

Ans: b) Predict binary class labels

ব্যাখ্যা: Logistic Regression মূলত binary classification (0/1) এর জন্য ব্যবহার হয়। এটি probability estimate করে এবং threshold অনুযায়ী class নির্ধারণ করে।

2. Which function is used to convert linear output into probability?

- a) ReLU
- b) Sigmoid
- c) Softmax
- d) Tanh

Ans: b) Sigmoid

ব্যাখ্যা: Sigmoid function output-কে 0 থেকে 1 এর মধ্যে নিয়ে আসে, তাই এটি probability হিসেবে interpret করা যায়। Logistic Regression-এর মূল অংশ এটি।

3. The decision boundary in Logistic Regression is formed when probability equals:

- a) 0
- b) 1
- c) 0.5

d) -1

Ans: c) 0.5

ব্যাখ্যা: সাধারণত $\text{probability} \geq 0.5$ হলে class 1 ধরা হয়। এটিই decision boundary—যেখানে model দুই class-এর মধ্যে boundary তৈরি করে।

4. In logistic regression, log-odds is defined as:

a) $\log(p)$

b) $\log(1 - p)$

c) $\log(p / (1 - p))$

d) $\log(p \times (1 - p))$

Ans: c) $\log(p / (1 - p))$

ব্যাখ্যা: Log-odds হলো probability ratio-এর logarithm। এটি linear model-এর সঙ্গে সরাসরি relate করা যায়, তাই logistic regression-এর mathematical intuition এখানেই।

5. Which cost function is commonly used in Logistic Regression?

a) Mean Squared Error

b) Hinge Loss

c) Binary Cross-Entropy (Log Loss)

d) L1 Loss

Ans: c) Binary Cross-Entropy (Log Loss)

ব্যাখ্যা: Logistic Regression-এর training-এ Log Loss minimize করা হয়। এটি probability-based error measurement যা classification-এর জন্য উপযুক্ত।

6. What optimization method is generally used to minimize the logistic regression cost function?

a) Random Forest

b) Gradient Descent

c) K-Means

d) PCA

Ans: b) Gradient Descent

ব্যখ্যা: Gradient Descent iterativeভাবে parameters update করে cost কমায়। Logistic Regression-এ এটি standard optimization method।

7. Which metric helps evaluate how many actual positives were correctly predicted?

a) Accuracy

b) Recall

c) Precision

d) F1 Score

Ans: b) Recall

ব্যখ্যা: $\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$ । এটি actual positive-এর কতটা model পকড়তে পেরেছে তা বোঝায়। Medical diagnosis-এর মতো ক্ষেত্রে Recall গুরুত্বপূর্ণ।

8. A confusion matrix includes which set of values?

a) Precision and recall

b) TP, FP, TN, FN

c) Accuracy and error rate

d) Loss and gradient

Ans: b) TP, FP, TN, FN

ব্যখ্যা: Confusion Matrix classification-এর মৌলিক structure, যা model-এর prediction performance চারটি অনুযায়ী দেখায়—TP, FP, TN এবং FN।

9. What effect does L2 regularization have on logistic regression?

- a) Forces many weights to become exactly zero
- b) Penalizes large parameter values
- c) Increases model variance
- d) Removes bias term

Ans: b) Penalizes large parameter values

ব্যাখ্যা: L2 regularization weight-গুলিকে বড় হতে দেয় না। এতে model কম complex হয় এবং overfitting কমে—stability বৃদ্ধি পায়।

10. When implementing logistic regression on a real dataset, what is the FIRST step?

- a) Train the model immediately
- b) Split dataset into train and test sets
- c) Add regularization
- d) Visualize decision boundary

Ans: b) Split dataset into train and test sets

ব্যাখ্যা: Train-test split করা খুব গুরুত্বপূর্ণ যাতে model unbiased evaluation পাওয়া যায়। এরপর preprocessing ও training করা হয়।