1. What is Logistic Regression, and how does it differ from Linear Regression?

Logistic Regression is a supervised learning algorithm used for classification tasks. Unlike
Linear Regression, which predicts continuous values, Logistic Regression predicts probabilities
that map to discrete classes using the Sigmoid function. It is primarily used for binary classification but can be extended to multiclass problems.

2. What is the mathematical equation of Logistic Regression?

The equation for Logistic Regression is:

$$P(Y=1|X)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_nX_n)}}$$

where $\beta_0,\beta_1,...\beta_n$ are the regression coefficients.

3. Why do we use the Sigmoid function in Logistic Regression?

The Sigmoid function converts any real-valued number into a probability between 0 and 1,making it useful for classification tasks.

$$\sigma(z)=\frac{1}{1+e^{-z}}$$

4. What is the cost function of Logistic Regression?

Logistic Regression uses the Log Loss (Cross-Entropy) function:

$$J(\theta)=-\frac{1}{m}\sum_{i=1}^{m}[y_i\log(h_i)+(1-y_i)\log(1-h_i)]$$

where h_i is the predicted probability.

5. What is Regularization in Logistic Regression? Why is it needed?

Regularization prevents overfitting by adding a penalty term to the cost function. L1 (Lasso) and
L2 (Ridge) regularization control model complexity.

6. Explain the difference between Lasso, Ridge, and Elastic Net regression.

Lasso (L1): Shrinks coefficients and can eliminate some features.
Ridge (L2): Shrinks coefficients without setting any to zero.
Elastic Net: A mix of L1 and L2, balancing feature selection and coefficient shrinkage.

7. When should we use Elastic Net instead of Lasso or Ridge?

Elastic Net is useful when features are highly correlated, as it inherits the strengths of both L1
and L2 regularization.

8. What is the impact of the regularization parameter (λ) in Logistic
Regression?

High $\lambda\lambda$: Stronger regularization, reducing model complexity.

Low $\lambda\lambda$: Weaker regularization, allowing the model to capture more patterns.

9. What are the key assumptions of Logistic Regression?

No multicollinearity
Linearity between independent variables and log-odds
Large sample size

10. What are some alternatives to Logistic Regression for classification
tasks?

Decision Trees
Random Forest
Support Vector Machines (SVM)
Neural Networks

11. What are Classification Evaluation Metrics?

Accuracy
Precision
Recall
F1-score
ROC-AUC

12. How does class imbalance affect Logistic Regression?

Class imbalance skews predictions. Techniques like class weighting, SMOTE (Synthetic Minority Oversampling), and threshold tuning help mitigate this.

13. What is Hyperparameter Tuning in Logistic Regression?

It optimizes model parameters like regularization strength (C) and solver choice to improve
Performance.

14. What are different solvers in Logistic Regression? Which one should be used?

liblinear: Small datasets, L1/L2 regularization
saga: Large datasets, Elastic Net
lbfgs: Multiclass problems

15. How is Logistic Regression extended for multiclass classification?

Using One-vs-Rest (OvR) or Softmax Regression (Multinomial Logistic Regression).

16. What are the advantages and disadvantages of Logistic Regression?

Advantages: Simple, interpretable, probabilistic output
Disadvantages: Assumes linearity, sensitive to multicollinearity

17. What are some use cases of Logistic Regression?

Spam detection
Medical diagnosis
Credit scoring

18. What is the difference between Softmax Regression and Logistic Regression?

Softmax Regression is a generalization of Logistic Regression for multiclass problems.

19. How do we choose between One-vs-Rest (OvR) and Softmax for multiclass classification?

OvR: When classes are imbalanced
Softmax: When all classes need equal consideration

20. How do we interpret coefficients in Logistic Regression?

Coefficients represent the log-odds change for a unit change in an independent variable.