SL TEST 2 (20 to 30 mins)



15-Question Test on Overfitting & Underfitting,   
Regularization, Ridge Regression, Lasso Regression, and Logistic Regression



Section 1: Overfitting & Underfitting

Q1. Define overfitting and underfitting in the context of machine learning. Answer:

Overfitting : When model performs well on training data but fails to perform well on test data.

Underfitting : When model does not perform well on both training and test data.

Q2. What are the signs that a model is overfitting?   
Answer: Training accuracy will be high but model will have less testing accuracy.

Q3. What are some common methods to prevent overfitting? Answer: Regularization, feature engineering can be used to prevent overfitting.

Section 2: Regularization

Q4. What is regularization, and why is it used in machine learning models? Answer: Regularization is a technique in Machine Learning used to prevent overfitting.

Q5. Explain the difference between L1 and L2 regularization. Answer:L2 regularization adds penalty to cost function which is equal to square of magnitude of coefficients while L1 regularization adds penalty to cost function which is equal to absolute value of magnitude of coefficients.

Q6. How does regularization affect the bias-variance tradeoff? Answer: Regularization penalizes coefficients which can be used to avoid overfitting.

Section 3: Ridge Regression

Q7. What is Ridge Regression, and how does it differ from standard linear regression? Answer:Standard linear regression may not be able to find the underlying structure of data leading to overfitting of model where as Ridge regression adds a penalty equal to square of magnitude of coefficients to prevent overfitting.

Q8. What is the effect of the regularization parameter (λ) in Ridge Regression? Answer: As the value of regularization parameter increases, slope of line decreases.

Section 4: Lasso Regression

Q9. What is Lasso Regression, and how does it perform feature selection?

Answer:Lasso regression also called as L1 regularization is a type of regulation to prevent overfitting by adding penalty to cost function which is equal to absolute value of magnitude of coefficients. It performs automatic feature selection by making some coefficients to zero.

Q10. In which scenarios is Lasso Regression preferred over Ridge Regression?

Answer: Lasso regression is preferred to build a simpler model with automatic feature selection while Ridge regression is preferred when all or most of the features are having relevance and multicollinearity is observed.

Section 5: Logistic Regression

Q11. Explain the purpose of Logistic Regression.

Answer: Logistic regression is a supervised machine learning model used for classification problems.

Q12. How does the cost function in Logistic Regression differ from that in Linear Regression?

Answer:

Section 6: Multiple Choice Questions

Q13. Which of the following statements about Lasso Regression is true?

A) Lasso Regression always includes all features in the final model.

B) Lasso Regression can set some coefficients to exactly zero, performing feature selection. C) Lasso Regression is not affected by the choice of the regularization parameter.

D) Lasso Regression can only be applied to linear models.

Q14. What is the main disadvantage of using Ridge Regression compared to Lasso Regression?

A) It cannot handle multicollinearity.

B) It does not perform feature selection.

C) It requires more computational resources.   
D) It can only be used for binary classification.

Q15. Explain the concept of the confusion matrix and its significance in evaluating the performance of classification models like Logistic Regression.

Answer:

Confusion matrix provides information about TP, TN, FP, FN values predicted by the model.

This data is useful to evaluate model performance in terms of precision and recall and F1 score.

Higher the values of TP, TN, better the model performance.