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 occurs when a machine learning model learns not only the underlying patterns in the training data but also the noise and outliers. This results in high accuracy on the training set but poor generalization to new, unseen data.**

** Underfitting happens when the model is too simple to capture the underlying structure of the data. It performs poorly on both the training and test datasets**

Q2. What are the signs that a model is overfitting?   
Answer: ** High accuracy on training data but low accuracy on validation/test data**

** Large gap between training and validation loss**

** Model performs well on known data but poorly on new or real-world data**

** Increasing model complexity (e.g., more layers or parameters) worsens generalization**

Q3. What are some common methods to prevent overfitting? Answer:  ** Cross-validation to evaluate model performance on unseen data**

** Regularization techniques like L1 and L2**

** Early stopping during training**

** Pruning in decision trees or neural networks**

** Dropout in neural networks**

** Simplifying the model by reducing features or parameters**

** Using more training data**

Section 2: Regularization

Q4. What is regularization, and why is it used in machine learning models? Answer: **Regularization is a technique used to reduce model complexity and prevent overfitting by adding a penalty term to the loss function. This discourages the model from fitting the noise in the training data and helps it generalize better to new data.**

Q5. Explain the difference between L1 and L2 regularization. Answer: ** L1 Regularization (Lasso): Adds the absolute value of the coefficients to the loss function. It can shrink some coefficients to zero, effectively performing feature selection.  
*Penalty term:* λ \* Σ|wᵢ|**

** L2 Regularization (Ridge): Adds the square of the coefficients to the loss function. It shrinks coefficients more evenly, but rarely sets them to zero.  
*Penalty term:* λ \* Σwᵢ²**

Q6. How does regularization affect the bias-variance tradeoff? Answer**: Regularization increases the bias slightly while reducing the variance.**

* **By adding a penalty to large coefficients, regularization simplifies the model.**
* **This helps the model generalize better to unseen data (lower variance) but may slightly reduce its ability to fit the training data (higher bias).**
* **Overall, it helps achieve a better balance in the bias-variance tradeoff and prevents overfitting**

Section 3: Ridge Regression

Q7. What is Ridge Regression, and how does it differ from standard linear regression? Answer: ** Ridge Regression is a type of linear regression that includes L2 regularization, which adds the sum of the squares of the model coefficients to the loss function.**

** It differs from standard linear regression by penalizing large weights, thereby preventing overfitting.**

** The Ridge loss function is:  
Loss = MSE + λ \* Σwᵢ²**

** Standard linear regression only minimizes the Mean Squared Error (MSE) without any penalty.**

Q8. What is the effect of the regularization parameter (λ) in Ridge Regression? Answer: ** λ (lambda) controls the strength of the regularization:**

* **λ = 0: Ridge regression becomes standard linear regression.**
* **Small λ: Slight regularization, little impact on model complexity.**
* **Large λ: Strong regularization, shrinks weights significantly, possibly leading to underfitting.**

** Choosing an appropriate λ helps balance model complexity and generalization.**

Section 4: Lasso Regression

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

Answer: ** Lasso Regression is linear regression with L1 regularization, which adds the sum of the absolute values of the coefficients to the loss function:  
Loss = MSE + λ \* Σ|wᵢ|**

** Unlike Ridge, Lasso can shrink some coefficients exactly to zero, effectively removing less important features.**

** This makes Lasso useful not only for preventing overfitting but also for automatic feature selection**.

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

Answer: **Lasso Regression is preferred when:**

* **Feature selection is important, as it can reduce the number of features by setting some coefficients to zero.**
* **There are many irrelevant or less important features in the dataset.**
* **You expect that only a few features are truly significant (i.e., sparse solutions are desired).  
  In contrast, Ridge is preferred when all features are potentially useful and multicollinearity is a concern.**

Section 5: Logistic Regression

Q11. Explain the purpose of Logistic Regression.

Answer: **Logistic Regression is used for classification problems, especially binary classification.  
It predicts the probability that a given input belongs to a certain class (e.g., class 1 vs class 0) by applying the sigmoid (logistic) function to the linear combination of inputs.  
The output is a value between 0 and 1, representing the estimated probability of the positive class**

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

Answer: ** Linear Regression uses Mean Squared Error (MSE) as its cost function.**

** Logistic Regression uses the Log Loss (also known as Binary Cross-Entropy), which is more suitable for classification tasks.  
The cost function penalizes wrong confident predictions more heavily and ensures output probabilities stay between 0 and 1.**

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.

Answer: **B) Lasso Regression can set some coefficients to exactly zero, performing feature selection**

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.

Answer: **B) It does not perform feature selection.  
Ridge Regression reduces the magnitude of coefficients but does not eliminate any features completely.**

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

Answer:

**A confusion matrix is a performance measurement tool for classification models like Logistic Regression. It is a table that summarizes the model's predictions against the actual outcomes.**

**For a binary classification problem, the confusion matrix is a 2x2 table with the following components:**

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

**Components:**

* **True Positive (TP): Correctly predicted positive cases.**
* **False Positive (FP): Incorrectly predicted as positive (Type I error).**
* **False Negative (FN): Incorrectly predicted as negative (Type II error).**
* **True Negative (TN): Correctly predicted negative cases.**

**Significance:**

**The confusion matrix helps evaluate various performance metrics, such as:**

* **Accuracy = (TP + TN) / (TP + TN + FP + FN)**
* **Precision = TP / (TP + FP)**
* **Recall (Sensitivity) = TP / (TP + FN)**
* **F1-Score = 2 × (Precision × Recall) / (Precision + Recall)**

**These metrics provide a deeper understanding of the model's performance, especially when dealing with imbalanced datasets. Unlike accuracy alone, they help identify if the model is biased toward one class.**

Top of Form

Bottom of Form