Machine Learning Question Paper

Regression

1. Find the m and c values of Linear Regression using gradient descent.

2. A company wants to predict house prices based on square footage using a Linear Regression model. They collected data on 500 houses and trained a model. However, they observed that the model performs poorly on new data.
   (a) What could be the possible reasons?
   (b) How can they improve their model's performance?

3. Write the polynomial regression equation for two features with degree 2. what are the parameters need to be estimated in this model?

4. Consider a machine learning model trained to predict house prices based on features such as square footage, number of rooms, and location. Provide an example of how an overfitting and an underfitting model would behave in this scenario. How would you improve the model?

Cross-validation

5. Explain the difference between overfitting and underfitting in machine learning models. How can cross-validation help to detect and address these issues?

6. What is k-fold cross-validation and how does it help in model evaluation? Compare it with train-test split and explain why cross-validation is preferred in certain cases.

7. Explain hyperparameter tuning using polynomial regression.

8. Explain hyperparameter tuning in the case of decision trees.

9. What is singular value decomposition (SVD) and how does it relate to PCA?

10. Explain the Curse of Dimensionality in machine learning. How does increasing the number of features impact model performance and what strategies can be used to mitigate its effects?

Evaluation Metrics and Model Performance

11. Explain the difference between precision, recall, and F1-score in classification problems.
12. Why do we use cross-validation in machine learning? Briefly describe k-fold cross-validation.
13. How does the ROC curve help in evaluating a classifier?
14. What is the difference between the R-squared and the mean squared error (MSE) in the regression evaluation?
15. For the following data set.
    (a) Create confusion matrix
    (b) Calculate accuracy, precision, recall, and F1-score for class 1 (positive class).
16. Describe the concept of the confusion matrix and its significance in evaluating classification models.
17. Compute the Mean Absolute Error (MAE) and Compute the Mean Squared Error (MSE).
18. Explain the significance of the ROC curve in classification models. How does it help in evaluating a model's performance? How does an ideal ROC curve look like?
19. Given the two following ROC curves for model A and model B, which is better?

Bayes' Theorem and Probability

20. A rare disease affects 1 in 1,000 people in a population. A test for this disease is 90% accurate in detecting the disease when a person has it (true positive rate), but it also incorrectly identifies 5% of healthy people as having the disease (false positive rate). If a randomly selected person tests positive, what is the probability that they actually have the disease?

DT and SVM

21. Given the following data, construct a decision tree.
22. Why does Linear SVM fail for non-linearly separable data?
23. Explain the concept of margin in SVM and given the optimization framework for Hard SVM

24. List different kernels in SVM and their corresponding hyperparameters

25. Given the covariance matrix of a dataset, the following eigenvalues and corresponding eigenvectors were computed:

    - **Eigenvalues:**
    $$\lambda_1 = 4.5, \quad \lambda_2 = 2.3, \quad \lambda_3 = 0.7$$
    - **Eigenvectors:**
    $$v_1 = \begin{bmatrix} 0.8 \\ 0.5 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -0.4 \\ 0.9 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0.2 \\ -0.1 \end{bmatrix}$$

    identify PC1 and PC2

    Given the following dataset:
    $$X = \begin{bmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{bmatrix}$$
    and the principal component vectors:
    $$PC1 = \begin{bmatrix} 0.8 \\ 0.6 \end{bmatrix}, \quad PC2 = \begin{bmatrix} -0.6 \\ 0.8 \end{bmatrix}$$

    1. Compute the projection of X onto the PC1 and PC2 axis

26. Given a neural network with the following structure:

    - **Input layer:** Two neurons x_1 and x_2
    - **Hidden layer:** Two neurons with activation function σ, weights $w_{11}^{(1)}$, $w_{12}^{(1)}$, $w_{21}^{(1)}$, $w_{22}^{(1)}$, and biases $b_1^{(1)}$, $b_2^{(1)}$
    - **Output layer:** One neuron with activation function f, weights $w_1^{(2)}$, $w_2^{(2)}$, and bias $b^{(2)}$
    - **Loss function:** Mean Squared Error (MSE), given by
    $$L = \frac{1}{2}(y - \hat{y})^2$$

    where $\hat{y}$ is the predicted output and y is the ground truth, compute the partial derivatives of the loss function L with respect to $w_{11}^{(2)}$ and $w_{21}^{(2)}$