Name: _____ Roll Number: _____

1. In linear regression, what does minimizing the mean squared error (MSE) accomplish?

    A. It minimizes the absolute error between the predicted and actual values.

    B. It minimizes the likelihood of the model given the data.

    C. It minimizes the squared difference between the predicted and actual values, penalizing larger errors more.

    D. It ensures the model has the best possible generalization on unseen data.

    Answer: C. Minimizing MSE helps to find the best-fit curve by reducing the squared differences between predicted and actual values. MSE penalizes larger errors more severely, leading to more accurate predictions in terms of squared error.

2. Which of the following is a characteristic of overfitting in polynomial regression?

    A. The model fits the training data poorly and performs well on unseen data.

    B. The model fits the training data very well but performs poorly on unseen data.

    C. The model does not have enough capacity to capture the patterns in the training data.

    D. The model produces a linear fit, regardless of the degree of the polynomial.

    Answer: B. Overfitting occurs when a model fits the training data too well, capturing noise rather than the underlying relationship. This results in poor performance on new, unseen data.

3. What is the purpose of using regularization in polynomial regression?

    A. To increase the model complexity and improve training performance.

    B. To prevent overfitting by penalizing large model coefficients.

    C. To reduce the error on the training set.

    D. To ensure the model fits the training data perfectly.

    Answer: B. Regularization adds a penalty to large model coefficients, preventing overfitting by encouraging simpler models that generalize better to unseen data. It doesn't necessarily minimize training error but aims for better generalization.

4. In linear regression in a polynomial space, how does increasing the degree of the polynomial affect the model?

    A. It always improves the model's generalization performance.

    B. It increases the model's complexity and risk of overfitting.

    C. It reduces the model's flexibility, resulting in underfitting.

    D. It increases the training error but reduces the test error.

    Answer: B. Increasing the degree of the polynomial increases the model's capacity to fit the training data, but it also raises the risk of overfitting, where the model captures noise rather than the underlying pattern.

5. What is the main purpose of using cross-validation in machine learning?

    A. To maximize the training accuracy of the model.

    B. To reduce the computational cost of training the model.

C. To see how well the model generalizes to unseen data by splitting the dataset into multiple training and validation sets.

D. To ensure that the model's complexity increases with the number of training samples.

Correct Choice C.