Interpretation Questions – Machine Learning Analysis

Walter Monteiro

1 Question 1 – Bias and Variance Trade-off

Figure 1 presents different hypothetical model fits over a sparse dataset.

The distribution that best balances bias and variance is the one that captures the overall structure of the data without excessively oscillating between individual data points.

A model exhibiting high bias underfits the data, typically represented by an overly simple decision boundary that fails to follow the data distribution. Conversely, a high-variance model overfits the data, adapting excessively to noise and local fluctuations.

The optimal distribution is therefore characterized by:

- Smooth approximation of the data trend
- Limited sensitivity to individual outliers
- Adequate flexibility without excessive complexity

This balance ensures strong generalization performance on unseen samples.

2 Question 2 – Model Evaluation Curve

Figure 2 represents a **Receiver Operating Characteristic (ROC) curve**.

The purpose of this graph is to evaluate the discriminative ability of a classification model across different threshold values. The ROC curve plots:

True Positive Rate (TPR) vs. False Positive Rate (FPR).

The dashed diagonal line represents the performance of a **random classifier**, where predictions are equivalent to chance. Any model whose curve lies close to this diagonal demonstrates little or no discriminative power.

Between the two curves (red and green), the better-performing model is the one whose curve:

- Lies consistently above the other curve
- Is closer to the top-left corner of the plot
- Encloses a larger Area Under the Curve (AUC)

A higher AUC indicates better separation between positive and negative classes across threshold variations.

3 Question 3 – Training and Test Evaluation Analysis

Figure 3 presents the training evolution (accuracy and loss) along with a confusion matrix obtained from test samples.

If training accuracy steadily increases and training loss decreases, but the confusion matrix reveals misclassification patterns in test data, this indicates that the model may not generalize well.

This phenomenon is commonly associated with **overfitting**, where the model:

- Learns detailed patterns from the training data
- Fails to maintain equivalent performance on unseen data

Even if training performance appears strong, poor class separation in the confusion matrix suggests insufficient generalization.

This highlights the importance of validation strategies, regularization techniques, and careful monitoring of generalization performance.