TASK 24

SHERAZ BIN TAHIR

MAIL: BAIGS543@GMAIL.COM

UNDERFITTING, OVERFITTING, AND REGULARIZATION

1. Underfitting

Underfitting occurs when a model is too simple to capture the underlying patterns in the data. This often happens when the model has too few parameters or the wrong structure to learn from the training data.

- Symptoms: The model performs poorly on both the training data and the validation/test data, showing high bias and low variance.
- Causes:
  - The model is too simplistic.
  - The model is not trained long enough.
  - Insufficient features or not enough relevant features are used.

2. Overfitting

Overfitting happens when a model learns not just the underlying pattern in the training data but also the noise and details specific to that data. As a result, the model performs very well on the training data but poorly on unseen (validation/test) data.

- Symptoms: The model has a low error rate on the training data but a high error rate on validation/test data, showing low bias and high variance.
- Causes:
  - The model is too complex with too many parameters.
  - The model is trained for too long.
  - The training data is not sufficiently diverse or large.

3. Regularization

Regularization is a technique used to prevent overfitting by adding a penalty to the loss function, discouraging the model from fitting too closely to the training data. This helps to improve the model's generalization to unseen data.

- Types of Regularization:
  - L1 Regularization (Lasso): Adds a penalty equal to the absolute value of the magnitude of coefficients. This can lead to sparse models where some coefficients are exactly zero, effectively selecting features.
  - L2 Regularization (Ridge): Adds a penalty equal to the square of the magnitude of coefficients. This discourages large coefficients but typically does not zero them out.
  - Dropout: In neural networks, dropout is a regularization technique where randomly selected neurons are ignored during training. This prevents the network from becoming overly reliant on specific pathways, improving generalization.
- Effect: Regularization helps balance the bias-variance trade-off, reducing overfitting while avoiding underfitting.