INFO7375

HW9

Jiani Lyu

1. **What are underfitting and overfitting?**

Underfitting: when a model is too simple to capture the underlying pattern of the data, leading to poor performance on both training and unseen data.

Overfitting: when a model is too complex, capturing noise in the training data as if it were a pattern, resulting in high performance on training data but poor generalization to new, unseen data.

2. **What nay cause an dearly stopping of the gradient descent optimization process?**

Early stopping of the gradient descent optimization process may be caused by:

Convergence to a Minimum: The algorithm stops if it converges to a minimum, where further iterations do not significantly decrease the loss.
Predefined Threshold: Reaching a predefined threshold for the loss or accuracy.
No Improvement: A specified number of epochs pass without noticeable improvement in validation loss or accuracy.
Resource Constraints: Limitations in computational resources or time.
Overfitting Prevention: To prevent overfitting, training may stop when validation performance starts to degrade.

3. **Describe the recognition bias vs variance and their relationship.**

Bias refers to the error due to overly simplistic assumptions in the model, leading to underfitting and poor performance on both training and unseen data. Variance refers to the error due to too complex model assumptions, capturing noise as if it were true pattern, leading to overfitting and poor performance on unseen data. The relationship between bias and variance is a trade-off: decreasing bias typically increases variance, and vice versa. The goal is to find a balance that minimizes total error.

4. **Describe regularization as a method and the reasons for it.**

Regularization is a method used to prevent overfitting in machine learning models by adding a penalty on the magnitude of model parameters. This penalty discourages the

model from becoming too complex by either shrinking the coefficients (in linear models) or constraining the weights (in neural networks). The reasons for regularization include improving the model's generalization ability on unseen data, reducing the risk of overfitting, and often leading to simpler, more interpretable models. Common forms of regularization include L1 (lasso), L2 (ridge), and dropout (for neural networks).

5. **Describe dropout as a method and the reasons for it.**

Dropout is a regularization technique used in neural networks to prevent overfitting. It works by randomly deactivating a proportion of neurons in a layer during each training iteration, forcing the network to learn more robust features that don't rely on a few neurons. This enhances the model's generalization ability on unseen data.