

## Study Exercises: Overfitting

1. Describe bias and variance
2. What is overfitting in the context of machine learning?
3. Explain the bias-variance trade-off and its relevance to overfitting.
4. How does underfitting differ from overfitting in terms of model performance?
5. Why is it important to have a separate validation set in machine learning?
6. Explain cross-validation and describe what is the role of cross-validation in detecting overfitting.
7. Describe the concept of model generalization and its relationship to overfitting.
8. How do regularization techniques, such as L1 and L2 regularization, help prevent overfitting?
9. What is the purpose of dropout in neural networks, and how does it combat overfitting?
10. Explain how early stopping can be used to mitigate overfitting during model training.
11. How does data augmentation address overfitting, particularly in computer vision applications?
12. Why is feature selection relevant to controlling overfitting?
13. Describe the impact of biased or noisy training data on the likelihood of overfitting.
14. What is a validation curve, and how does it reveal overfitting in a model?
15. Imagine you have a machine learning model with a very high training accuracy but a significantly lower validation accuracy. Is this a problem? How would you address it?
16. Consider a scenario where you have a small dataset with very few samples. How would you approach this situation to prevent overfitting?
17. Discuss the trade-offs between early stopping and the risk of underfitting. When would you choose one over the other?
18. Imagine you have a dataset with many features, and you suspect overfitting. How would you approach feature selection to mitigate overfitting?
19. How can you determine if your model's performance improvements are statistically significant when implementing various techniques to combat overfitting?
20. Describe a real-world example where overfitting can have severe consequences.
21. Explain regularization.
22. How does the choice of the regularization strength  $\lambda$  affect the degree of regularization and the model's performance?

23. How can regularization techniques be combined to strike a balance between L1 and L2 regularization in a model?
24. Explain the impact of the choice of loss function on overfitting in machine learning models.
25. How can you visualize the effect of regularization on a model's parameters, and what does this reveal about the regularization's impact?
26. In what situations might regularization be less effective in preventing overfitting, and what alternative approaches can be considered?
27. How can the choice of activation functions in a neural network impact overfitting?
28. How does the choice of optimization algorithm in training a machine learning model influence the risk of overfitting?
29. Explain the differences between batch normalization and dropout as techniques to prevent overfitting in neural networks.
30. What is the relationship between the number of layers in a deep neural network and the risk of overfitting, and how can this relationship be managed?