**Question 1 (8 marks)**

**(a) A decision tree is constructed for a classification problem. Given the following dataset, use the Gini impurity to determine the best feature for the first split. (4 marks)**

* Dataset:

| **Feature 1** | **Feature 2** | **Class** |
| --- | --- | --- |
| 1 | 5 | A |
| 2 | 6 | A |
| 3 | 7 | B |
| 4 | 8 | B |
| 5 | 9 | B |

**(b) Explain the concept of pruning in decision trees. How does pruning help to improve the model? (4 marks)**

**Question 2 (7 marks)**

**(a) Explain the process of gradient descent. Why is it important to choose an appropriate learning rate? (3 marks)**

**(b) Compare and contrast batch gradient descent, stochastic gradient descent, and mini-batch gradient descent. (4 marks)**

**Question 3 (5 marks)**

**(a) Given a linear regression model y=θ0+θ1x1+θ2x2y = \theta\_0 + \theta\_1x\_1 + \theta\_2x\_2y=θ0​+θ1​x1​+θ2​x2​, describe the Normal Equation used to find the optimal parameters θ\thetaθ. (2 marks)**

**(b) Discuss one advantage and one disadvantage of using the Normal Equation for training a linear regression model. (3 marks)**

**Question 4 (10 marks)**

**(a) Describe the k-nearest neighbors (k-NN) algorithm. How is the value of kkk chosen, and what effect does it have on the model’s performance? (4 marks)**

**(b) Given a 2D dataset, sketch the decision boundary you would expect from a k-NN classifier with k=3k = 3k=3. Explain your reasoning. (3 marks)**

**(c) What is the impact of having imbalanced datasets on k-NN classification? How can this issue be mitigated? (3 marks)**

**Question 5 (6 marks)**

**(a) What is the logistic function used in logistic regression? Write down the function and explain its significance. (3 marks)**

**(b) Explain how the parameters of a logistic regression model are estimated. (3 marks)**

**Question 6 (8 marks)**

**(a) Define bias and variance in the context of machine learning models. Provide an example of a high bias model and a high variance model. (4 marks)**

**(b) Describe the bias-variance tradeoff and its importance in machine learning. (4 marks)**

**Question 7 (5 marks)**

**(a) What is regularization in machine learning? Why is it used? (2 marks)**

**(b) Compare ridge regression and lasso regression. How does each method prevent overfitting? (3 marks)**

**Question 8 (6 marks)**

**(a) Explain how learning curves can be used to diagnose bias and variance problems in machine learning models. (3 marks)**

**(b) Given training and validation errors for different training set sizes, interpret the following learning curves and diagnose the model’s performance: (3 marks)**

* Diagram: (Provide a learning curve graph where students can analyze the training and validation errors)

**Question 9 (9 marks)**

**(a) Describe how a Support Vector Machine (SVM) finds the optimal separating hyperplane in a binary classification problem. (4 marks)**

**(b) What is the kernel trick in SVMs, and why is it useful? (3 marks)**

**(c) Explain the difference between hard margin and soft margin classification in SVMs. (2 marks)**

**Question 10 (6 marks)**

**(a) What is the purpose of cross-validation in machine learning? Describe how k-fold cross-validation is performed. (4 marks)**

**(b) Why is cross-validation important for model selection and hyperparameter tuning? (2 marks)**