Machine learning revision

**Notes with Answers and Concepts: Linear Regression**

1. **Primary Goal of Linear Regression**:
   * **Answer**: b) To model the relationship between independent and dependent variables.
   * **Concept**: Linear regression aims to find the best-fit line that predicts the dependent variable based on independent variables.
2. **Linear Regression Model**:
   * **Answer**: b) It assumes a linear relationship between features and target.
   * **Concept**: Linear regression assumes that the relationship between the input features and the target variable is linear.
3. **Intercept Term w0​**:
   * **Answer**: c) Intercept (bias term).
   * **Concept**: The intercept represents the value of the dependent variable when all independent variables are zero.
4. **Loss Function**:
   * **Answer**: b) Mean squared error.
   * **Concept**: The mean squared error (MSE) measures the average squared difference between actual and predicted values.
5. **Purpose of Minimizing Loss Function**:
   * **Answer**: b) To find the optimal values of the coefficients w0, w1,…, wn*.*
   * **Concept**: Minimizing the loss function helps in finding the best-fit line by optimizing the coefficients.
6. **Assumption of Linear Regression**:
   * **Answer**: c) Residuals have constant variance.
   * **Concept**: Homoscedasticity assumes that the residuals have constant variance across all levels of the independent variables.
7. **Multicollinearity**:
   * **Answer**: b) The independent variables are highly correlated with each other.
   * **Concept**: Multicollinearity occurs when independent variables are highly correlated, leading to unstable coefficient estimates.
8. **Effect of Outliers**:
   * **Answer**: c) Outliers can significantly affect the model’s coefficients.
   * **Concept**: Outliers can disproportionately influence the model, leading to inaccurate predictions.
9. **Limitation of Linear Regression**:
   * **Answer**: c) It assumes a linear relationship between features and target.
   * **Concept**: Linear regression cannot model non-linear relationships, which limits its applicability.
10. **Homoscedasticity**:
    * **Answer**: b) The residuals have constant variance.
    * **Concept**: Homoscedasticity ensures that the variance of residuals is consistent across all levels of the independent variables.
11. **Normal Equation**:
    * **Answer**: b) A closed-form solution to find the optimal coefficients.
    * **Concept**: The normal equation provides a direct method to compute the optimal coefficients without iterative optimization.
12. **Role of Intercept Term**:
    * **Answer**: b) It represents the value of y when all features are zero.
    * **Concept**: The intercept term accounts for the baseline value of the dependent variable.
13. **Residuals**:
    * **Answer**: a) Residuals are the differences between actual and predicted values.
    * **Concept**: Residuals measure the error in predictions and are used to evaluate the model’s performance.
14. **Effect of Increasing Features**:
    * **Answer**: b) The model becomes more prone to overfitting.
    * **Concept**: Adding more features increases the model’s complexity, which can lead to overfitting.
15. **Relationship Between Features and Overfitting**:
    * **Answer**: b) More features increase the risk of overfitting.
    * **Concept**: Overfitting occurs when the model captures noise in the training data due to excessive complexity.
16. **Purpose of Regularization**:
    * **Answer**: c) To prevent overfitting by penalizing large coefficients.
    * **Concept**: Regularization techniques like Ridge and Lasso add a penalty to the loss function to control model complexity.
17. **Bias-Variance Trade-off**:
    * **Answer**: c) Both a) and b) are true.
    * **Concept**: High bias leads to underfitting, and high variance leads to overfitting. Balancing both is crucial for optimal performance.
18. **Effect of Reducing Features**:
    * **Answer**: b) The model becomes less prone to overfitting.
    * **Concept**: Reducing the number of features simplifies the model, reducing the risk of overfitting.
19. **Difference Between Simple and Multiple Linear Regression**:
    * **Answer**: a) Simple linear regression uses one feature, while multiple linear regression uses multiple features.
    * **Concept**: Simple linear regression models the relationship between one independent variable and the dependent variable, while multiple linear regression models the relationship between multiple independent variables and the dependent variable.
20. **Role of Gradient Descent**:
    * **Answer**: b) To iteratively minimize the loss function.
    * **Concept**: Gradient descent is an optimization algorithm used to minimize the loss function by iteratively adjusting the coefficients.

**Decision Trees**

1. **Primary Goal of Decision Trees**:
   * **Answer**: a) To classify data into categories.
   * **Concept**: Decision trees are used for classification and regression tasks by splitting the data into subsets based on feature values.
2. **Best Split in Decision Trees**:
   * **Answer**: d) All of the above.
   * **Concept**: Entropy, information gain, and Gini index are used to decide the best split in a decision tree.
3. **Entropy Value of 0**:
   * **Answer**: b) The dataset is perfectly pure.
   * **Concept**: An entropy value of 0 indicates that the dataset is perfectly pure, meaning all instances belong to one class.
4. **Effect of Increasing Depth**:
   * **Answer**: b) It increases the risk of overfitting.
   * **Concept**: Increasing the depth of a decision tree makes it more complex, which can lead to overfitting.
5. **Gini Index**:
   * **Answer**: a) It measures the impurity of a dataset.
   * **Concept**: The Gini index measures the impurity of a dataset and is used to decide the best split.
6. **Tree Depth and Complexity**:
   * **Answer**: b) Deeper trees are more complex.
   * **Concept**: Deeper trees capture more complex patterns but are more prone to overfitting.
7. **Limitation of Decision Trees**:
   * **Answer**: c) They are prone to overfitting.
   * **Concept**: Decision trees can overfit the training data if they are too deep.
8. **Purpose of Pruning**:
   * **Answer**: b) To reduce the risk of overfitting.
   * **Concept**: Pruning removes unnecessary branches from the tree to reduce complexity and prevent overfitting.
9. **Information Gain**:
   * **Answer**: a) It measures the reduction in entropy after a split.
   * **Concept**: Information gain measures how much a split reduces entropy and is used to decide the best split.
10. **Effect of Reducing Depth**:
    * **Answer**: b) It reduces the risk of overfitting.
    * **Concept**: Reducing the depth of a decision tree simplifies the model, reducing the risk of overfitting.

**Conceptual Questions**

1. **What is the role of the root node in a decision tree?**  
   a) It represents the final decision.  
   b) It is the first node where the splitting begins.  
   c) It is used to prune the tree.  
   d) It represents the leaf nodes.

**Answer**: b) It is the first node where the splitting begins.

1. **Which of the following is true about leaf nodes in a decision tree?**  
   a) They represent the final decision or classification.  
   b) They are used to split the data further.  
   c) They are the same as the root node.  
   d) They are used to calculate entropy.

**Answer**: a) They represent the final decision or classification.

1. **What is the effect of using a very small dataset to train a decision tree?**  
   a) The tree will be shallow.  
   b) The tree will overfit the data.  
   c) The tree will underfit the data.  
   d) The tree will perform well on unseen data.

**Answer**: b) The tree will overfit the data.

1. **Which of the following is true about the Gini index?**  
   a) It ranges from -1 to 1.  
   b) A Gini index of 0 indicates maximum impurity.  
   c) A Gini index of 1 indicates maximum purity.  
   d) It is the same as entropy.

**Answer**: c) A Gini index of 1 indicates maximum purity.

1. **What is the main advantage of using decision trees?**  
   a) They are easy to interpret and visualize.  
   b) They are computationally expensive.  
   c) They are prone to overfitting.  
   d) They cannot handle categorical data.

**Answer**: a) They are easy to interpret and visualize.

**Numerical Questions**

1. **Calculate the Gini index for a dataset with [6+, 4-].**

Gini=1−((610)2+(410)2)=1−(0.36+0.16)=0.48Gini=1−((106​)2+(104​)2)=1−(0.36+0.16)=0.48

**Answer**: 0.48

1. **Calculate the entropy for a dataset with [8+, 2-].**

Entropy=−810log⁡2(810)−210log⁡2(210)≈0.72Entropy=−108​log2​(108​)−102​log2​(102​)≈0.72

**Answer**: ≈ 0.72

1. **Given the following dataset, calculate the information gain for splitting on feature X*X*:**
   * Before split: [5+, 5-]
   * After split:
     + Left branch: [4+, 1-]
     + Right branch: [1+, 4-]

Entropy before split=1

Entropy before split= 1

Entropy left= −4/5log2(4/5) – 1/5log2(1/5) ~ 0.72

Entropy right = -1/5log2(1/5) – 4/5log2(4/5) ~ 0.72

Weighted entropy after split =5/10 \* 0.72 + 5/10 \* 0.72 = 0.72

Information gain=1−0.72=0.28

**Answer**: 0.28

**K-Nearest Neighbors (K-NN)**

1. **Primary Goal of K-NN**:
   * **Answer**: a) To classify data into categories.
   * **Concept**: K-NN is used for classification and regression by finding the K nearest neighbors to a given data point.
2. **Effect of Small K*K***:
   * **Answer**: b) The model becomes sensitive to noise and may overfit.
   * **Concept**: A small K makes the model sensitive to noise and outliers, leading to overfitting.
3. **Rule of Thumb for Choosing K*K***:
   * **Answer**: b) K<n​, where n is the number of training examples.
   * **Concept**: A common rule of thumb is to choose K less than the square root of the number of training examples.
4. **Distance Metric in K-NN**:
   * **Answer**: d) All of the above.
   * **Concept**: Euclidean, Manhattan, and Chebyshev distances are commonly used in K-NN.
5. **Effect of Increasing K*K***:
   * **Answer**: c) The model becomes more robust to noise.
   * **Concept**: Increasing K*K* makes the model more robust to noise but may lead to underfitting.
6. **Disadvantage of Large K*K***:
   * **Answer**: b) It increases computation complexity.
   * **Concept**: A large K*K* increases computation complexity as more neighbors need to be considered.
7. **K-NN as a Lazy Learner**:
   * **Answer**: c) It is a lazy learning algorithm.
   * **Concept**: K-NN is a lazy learner because it does not require training before making predictions.
8. **Relationship Between K*K* and Complexity**:
   * **Answer**: a) A smaller K*K* leads to a more complex model.
   * **Concept**: A smaller K*K* leads to a more complex model that is sensitive to noise.
9. **Effect of Outliers in K-NN**:
   * **Answer**: c) Outliers can significantly affect the model’s predictions.
   * **Concept**: Outliers can disproportionately influence the predictions in K-NN.
10. **Choice of K for Noisy Data**:
    * **Answer**: b) A larger K*K* is preferred for noisy data.
    * **Concept**: A larger K*K* smooths out the effect of noise, making the model more robust.
11. **What is the primary disadvantage of using a small K*K* in K-NN?**  
    a) The model becomes more robust to noise.  
    b) The model becomes sensitive to noise and may overfit.  
    c) The model becomes computationally expensive.  
    d) The model becomes less accurate.

**Answer**: b) The model becomes sensitive to noise and may overfit.

1. **Which of the following is true about the choice of K*K* in K-NN?**  
   a) A smaller K*K* is preferred for noisy data.  
   b) A larger K*K* is preferred for noisy data.  
   c) K*K* has no effect on the model’s performance.  
   d) K*K* should always be set to 1.

**Answer**: b) A larger K*K* is preferred for noisy data.

1. **What is the effect of using a large K*K* in K-NN?**  
   a) The model becomes more sensitive to noise.  
   b) The model becomes less computationally expensive.  
   c) The model becomes more robust to noise.  
   d) The model becomes more likely to overfit.

**Answer**: c) The model becomes more robust to noise.

1. **Which of the following is true about the distance metric in K-NN?**  
   a) Euclidean distance is the only metric used.  
   b) The choice of distance metric does not affect the model’s performance.  
   c) Different distance metrics can be used depending on the data.  
   d) Distance metrics are not used in K-NN.

**Answer**: c) Different distance metrics can be used depending on the data.

1. **What is the effect of increasing the number of features in K-NN?**  
   a) The model becomes simpler.  
   b) The model becomes more prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model becomes more prone to overfitting.

**Numerical Questions**

1. **Given the following dataset, classify the new point (3,4)(3,4) using K-NN with K=3*K*=3:**
   * Data points: (1,2)(1,2) → Class A, (2,3)(2,3) → Class A, (4,5)(4,5) → Class B, (5,6)(5,6) → Class B
   * Distance metric: Euclidean distance.

**Solution**:

* + Distance to (1,2)(1,2): (3−1)2+(4−2)2=4+4=8≈2.83(3−1)2+(4−2)2​=4+4​=8​≈2.83
  + Distance to (2,3)(2,3): (3−2)2+(4−3)2=1+1=2≈1.41(3−2)2+(4−3)2​=1+1​=2​≈1.41
  + Distance to (4,5)(4,5): (3−4)2+(4−5)2=1+1=2≈1.41(3−4)2+(4−5)2​=1+1​=2​≈1.41
  + Distance to (5,6)(5,6): (3−5)2+(4−6)2=4+4=8≈2.83(3−5)2+(4−6)2​=4+4​=8​≈2.83
  + Nearest neighbors: (2,3)(2,3) (Class A), (4,5)(4,5) (Class B), (1,2)(1,2) (Class A)
  + Majority class: Class A  
    **Answer**: Class A

1. **Calculate the Euclidean distance between the points (1,2,3)(1,2,3) and (4,5,6)(4,5,6).**

Distance=(4−1)2+(5−2)2+(6−3)2=9+9+9=27≈5.20Distance=(4−1)2+(5−2)2+(6−3)2​=9+9+9​=27​≈5.20

**Answer**: ≈ 5.20

**Training vs Testing Error vs Cross-validation, Overfitting**

1. Training vs Testing Error:
   * Training Error: Measures the model’s performance on the training data.
   * Testing Error: Measures the model’s performance on unseen data.
   * Cross-Validation: Evaluates the model’s performance by splitting the data into multiple folds.
2. Underfitting:
   * Occurs when the model is too simple.
   * Results in high training and testing errors.
3. Overfitting:
   * Occurs when the model is too complex.
   * Results in low training error but high testing error.
4. Bias-Variance Trade-off:
   * High Bias: Leads to underfitting.
   * High Variance: Leads to overfitting.
5. Techniques to Reduce Overfitting:
   * Use cross-validation.
   * Reduce model complexity.
   * Increase the amount of training data.

1. **What is the primary purpose of splitting data into training and testing sets?**  
   a) To increase the model’s complexity.  
   b) To evaluate the model’s performance on unseen data.  
   c) To reduce the number of features.  
   d) To increase the training error.

**Answer**: b) To evaluate the model’s performance on unseen data.

1. **Which of the following is true about training error?**  
   a) It measures the model’s performance on the training data.  
   b) It measures the model’s performance on the testing data.  
   c) It is always lower than the testing error.  
   d) It is not affected by overfitting.

**Answer**: a) It measures the model’s performance on the training data.

1. **What does a high testing error indicate?**  
   a) The model is underfitting the data.  
   b) The model is overfitting the data.  
   c) The model is performing well on unseen data.  
   d) The model is too simple.

**Answer**: a) The model is underfitting the data.

1. **Which of the following is true about cross-validation?**  
   a) It is used to increase the training error.  
   b) It is used to evaluate the model’s performance on unseen data.  
   c) It is used to reduce the number of features.  
   d) It is used to increase the model’s complexity.

**Answer**: b) It is used to evaluate the model’s performance on unseen data.

1. **What is the effect of increasing the model’s complexity?**  
   a) It reduces the risk of overfitting.  
   b) It increases the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It reduces the training error but increases the testing error.

**Answer**: b) It increases the risk of overfitting.

1. **Which of the following is true about underfitting?**  
   a) It occurs when the model is too complex.  
   b) It occurs when the model is too simple.  
   c) It results in low training error and low testing error.  
   d) It is not affected by the choice of hyperparameters.

**Answer**: b) It occurs when the model is too simple.

1. **What is the relationship between bias and variance in model performance?**  
   a) High bias leads to overfitting.  
   b) High variance leads to underfitting.  
   c) High bias leads to underfitting, and high variance leads to overfitting.  
   d) High bias and high variance both lead to overfitting.

**Answer**: c) High bias leads to underfitting, and high variance leads to overfitting.

1. **Which of the following techniques can help reduce overfitting?**  
   a) Increasing model complexity.  
   b) Using cross-validation.  
   c) Adding more features.  
   d) Reducing the amount of training data.

**Answer**: b) Using cross-validation.

1. **What is the effect of reducing the depth of a decision tree?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases training error but reduces testing error.

**Answer**: b) It reduces the risk of overfitting.

1. **Which of the following is true about the bias-variance trade-off?**  
   a) A model with high bias has low variance.  
   b) A model with high variance has low bias.  
   c) Both a) and b) are true.  
   d) Neither a) nor b) is true.

**Answer**: c) Both a) and b) are true.

1. **What is the primary goal of using a validation set?**  
   a) To train the model.  
   b) To test the model on unseen data.  
   c) To tune hyperparameters.  
   d) To reduce the size of the training set.

**Answer**: c) To tune hyperparameters.

1. **Which of the following is true about overfitting?**  
   a) It occurs when the model is too simple.  
   b) It occurs when the model is too complex.  
   c) It results in high training error and high testing error.  
   d) It is not affected by the choice of hyperparameters.

**Answer**: b) It occurs when the model is too complex.

1. **What is the effect of increasing the amount of training data?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases the training error but reduces the testing error.

**Answer**: b) It reduces the risk of overfitting.

1. **Which of the following is true about the bias-variance trade-off?**  
   a) High bias leads to overfitting.  
   b) High variance leads to underfitting.  
   c) High bias leads to underfitting, and high variance leads to overfitting.  
   d) High bias and high variance both lead to overfitting.

**Answer**: c) High bias leads to underfitting, and high variance leads to overfitting.

1. **What is the effect of reducing the number of features in a model?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases training error but reduces testing error.

**Answer**: b) It reduces the risk of overfitting.

1. **Which of the following is true about cross-validation?**  
   a) It is used to increase the training error.  
   b) It is used to evaluate the model’s performance on unseen data.  
   c) It is used to reduce the number of features.  
   d) It is used to increase the model’s complexity.

**Answer**: b) It is used to evaluate the model’s performance on unseen data.

1. **What is the effect of increasing the model’s complexity?**  
   a) It reduces the risk of overfitting.  
   b) It increases the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It reduces the training error but increases the testing error.

**Answer**: b) It increases the risk of overfitting.

1. **Which of the following is true about underfitting?**  
   a) It occurs when the model is too complex.  
   b) It occurs when the model is too simple.  
   c) It results in low training error and low testing error.  
   d) It is not affected by the choice of hyperparameters.

**Answer**: b) It occurs when the model is too simple.

1. **What is the relationship between bias and variance in model performance?**  
   a) High bias leads to overfitting.  
   b) High variance leads to underfitting.  
   c) High bias leads to underfitting, and high variance leads to overfitting.  
   d) High bias and high variance both lead to overfitting.

**Answer**: c) High bias leads to underfitting, and high variance leads to overfitting.

1. **Which of the following techniques can help reduce overfitting?**  
   a) Increasing model complexity.  
   b) Using cross-validation.  
   c) Adding more features.  
   d) Reducing the amount of training data.

**Answer**: b) Using cross-validation.

**Numerical Questions**

1. **Given the following training and testing errors for different model complexities, identify the point of overfitting:**
   * Model Complexity: Low, Training Error: 0.5, Testing Error: 0.6
   * Model Complexity: Medium, Training Error: 0.3, Testing Error: 0.4
   * Model Complexity: High, Training Error: 0.1, Testing Error: 0.5

**Answer**: Overfitting occurs at high model complexity, where the training error is low (0.1) but the testing error is high (0.5).

1. **Calculate the cross-validation error for a model with the following errors across 5 folds:**
   * Fold 1: 0.2, Fold 2: 0.3, Fold 3: 0.25, Fold 4: 0.35, Fold 5: 0.3

Cross-Validation Error=(0.2+0.3+0.25+0.35+0.3)/5=0.28

**Answer**: 0.28

1. **Given the following training and testing errors, identify if the model is underfitting, overfitting, or well-fitted:**
   * Training Error: 0.4, Testing Error: 0.45

**Answer**: The model is well-fitted, as both training and testing errors are close and relatively low.

1. **Calculate the bias and variance for a model with the following performance:**
   * Training Error: 0.1, Testing Error: 0.5

**Answer**: High bias (underfitting) and high variance (overfitting).

**Perceptron**

**Perceptron: 20 Questions**

**Conceptual Questions**

1. **What is the primary goal of a perceptron?**  
   a) To classify data into categories.  
   b) To model the relationship between independent and dependent variables.  
   c) To reduce the dimensionality of data.  
   d) To cluster data into groups.

**Answer**: a) To classify data into categories.

1. **Which of the following is true about the perceptron?**  
   a) It can model non-linear decision boundaries.  
   b) It can only model linearly separable data.  
   c) It is used for regression tasks.  
   d) It is a type of unsupervised learning algorithm.

**Answer**: b) It can only model linearly separable data.

1. **What is the activation function used in a perceptron?**  
   a) Sigmoid  
   b) ReLU  
   c) Step function  
   d) Tanh

**Answer**: c) Step function

1. **What happens if the data is not linearly separable in a perceptron?**  
   a) The perceptron will still converge.  
   b) The perceptron will fail to converge.  
   c) The perceptron will switch to a non-linear model.  
   d) The perceptron will reduce the learning rate.

**Answer**: b) The perceptron will fail to converge.

1. **Which of the following is true about the learning rule in a perceptron?**  
   a) It updates the weights based on the gradient of the loss function.  
   b) It updates the weights based on the difference between the predicted and actual output.  
   c) It does not update the weights.  
   d) It updates the weights randomly.

**Answer**: b) It updates the weights based on the difference between the predicted and actual output.

1. **What is the effect of increasing the learning rate in a perceptron?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: a) The model converges faster.

1. **Which of the following is a limitation of the perceptron?**  
   a) It can model non-linear decision boundaries.  
   b) It can only model linearly separable data.  
   c) It is robust to outliers.  
   d) It works well with highly correlated features.

**Answer**: b) It can only model linearly separable data.

1. **What is the geometric interpretation of the perceptron’s decision boundary?**  
   a) A hyperplane separating classes in the feature space.  
   b) A curved surface for non-linear separation.  
   c) A probabilistic boundary.  
   d) A clustering boundary.

**Answer**: a) A hyperplane separating classes in the feature space.

1. **Why does the perceptron algorithm fail for non-linearly separable data?**  
   a) It lacks a learning rate.  
   b) It cannot minimize the error function to zero.  
   c) It assumes linear separability by design.  
   d) It uses a non-linear activation function.

**Answer**: b) It cannot minimize the error function to zero.

1. **What is the role of the bias term w0*w*0​ in the perceptron?**  
   a) It shifts the decision boundary away from the origin.  
   b) It controls the learning rate.  
   c) It reduces the model’s complexity.  
   d) It measures the error.

**Answer**: a) It shifts the decision boundary away from the origin.

1. **Which of the following is true about the perceptron’s loss function?**  
   a) It uses cross-entropy loss.  
   b) It minimizes the number of misclassified points directly.  
   c) It is differentiable everywhere.  
   d) It penalizes misclassifications quadratically.

**Answer**: b) It minimizes the number of misclassified points directly.

1. **In Bishop’s framework, how does the perceptron relate to linear discriminant functions?**  
   a) It is a special case of a linear discriminant function with a step activation.  
   b) It generalizes linear discriminant functions to non-linear boundaries.  
   c) It uses probabilistic outputs like logistic regression.  
   d) It is unrelated to discriminant functions.

**Answer**: a) It is a special case of a linear discriminant function with a step activation.

1. **What is the effect of outliers in the perceptron?**  
   a) Outliers have no effect on the model.  
   b) Outliers improve the model’s performance.  
   c) Outliers can significantly affect the model’s coefficients.  
   d) Outliers are automatically removed by the model.

**Answer**: c) Outliers can significantly affect the model’s coefficients.

1. **What is the primary advantage of the perceptron algorithm?**  
   a) It can model non-linear decision boundaries.  
   b) It is computationally efficient.  
   c) It works well with highly correlated features.  
   d) It is robust to outliers.

**Answer**: b) It is computationally efficient.

1. **What is the primary disadvantage of the perceptron algorithm?**  
   a) It can only model linearly separable data.  
   b) It is computationally expensive.  
   c) It is robust to outliers.  
   d) It works well with highly correlated features.

**Answer**: a) It can only model linearly separable data.

1. **What is the effect of reducing the learning rate in a perceptron?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: b) The model converges slower.

1. **What is the effect of increasing the number of features in a perceptron?**  
   a) The model becomes simpler.  
   b) The model becomes more prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model becomes more prone to overfitting.

1. **What is the effect of reducing the number of features in a perceptron?**  
   a) The model becomes more complex.  
   b) The model becomes less prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes more sensitive to outliers.

**Answer**: b) The model becomes less prone to overfitting.

1. **What is the effect of increasing the number of training examples in a perceptron?**  
   a) The model becomes more complex.  
   b) The model becomes less prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes more sensitive to outliers.

**Answer**: b) The model becomes less prone to overfitting.

1. **What is the effect of reducing the number of training examples in a perceptron?**  
   a) The model becomes more complex.  
   b) The model becomes more prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model becomes more prone to overfitting.

**Logistic Regression**

1. **What is the primary goal of logistic regression?**  
   a) To classify data into categories.  
   b) To model the relationship between independent and dependent variables.  
   c) To reduce the dimensionality of data.  
   d) To cluster data into groups.

**Answer**: a) To classify data into categories.

1. **Which of the following is true about logistic regression?**  
   a) It assumes a linear relationship between features and target.  
   b) It assumes a non-linear relationship between features and target.  
   c) It is used for regression tasks.  
   d) It is a type of unsupervised learning algorithm.

**Answer**: a) It assumes a linear relationship between features and target.

1. **What is the activation function used in logistic regression?**  
   a) Sigmoid  
   b) ReLU  
   c) Step function  
   d) Tanh

**Answer**: a) Sigmoid

1. **What is the range of the sigmoid function?**  
   a) (−∞,∞)(−∞,∞)  
   b) [0,1][0,1]  
   c) [−1,1][−1,1]  
   d) [0,∞)[0,∞)

**Answer**: b) [0,1][0,1]

1. **Which of the following is the loss function used in logistic regression?**  
   a) Mean squared error  
   b) Log loss (cross-entropy loss)  
   c) Hinge loss  
   d) Absolute error

**Answer**: b) Log loss (cross-entropy loss)

1. **What is the effect of increasing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: b) The model becomes less complex.

1. **Which of the following is true about the decision boundary in logistic regression?**  
   a) It is non-linear.  
   b) It is linear.  
   c) It is circular.  
   d) It is polynomial.

**Answer**: b) It is linear.

1. **What is the effect of outliers in logistic regression?**  
   a) Outliers have no effect on the model.  
   b) Outliers improve the model’s performance.  
   c) Outliers can significantly affect the model’s coefficients.  
   d) Outliers are automatically removed by the model.

**Answer**: c) Outliers can significantly affect the model’s coefficients.

1. **What is the probabilistic interpretation of the sigmoid function in logistic regression?**  
   a) It represents the likelihood of the data given the parameters.  
   b) It maps the linear combination of features to a probability between 0 and 1.  
   c) It measures the entropy of the model.  
   d) It penalizes misclassifications.

**Answer**: b) It maps the linear combination of features to a probability between 0 and 1.

1. **Why is log loss preferred over mean squared error (MSE) in logistic regression?**  
   a) Log loss is computationally cheaper.  
   b) Log loss is convex, ensuring a unique global minimum.  
   c) MSE cannot handle probabilities.  
   d) Log loss works better with non-linear data.

**Answer**: b) Log loss is convex, ensuring a unique global minimum.

1. **What does the likelihood function in logistic regression represent?**  
   a) The probability of observing the training data given the model parameters.  
   b) The prior distribution of the features.  
   c) The regularization term.  
   d) The entropy of the model.

**Answer**: a) The probability of observing the training data given the model parameters.

1. **How does L2 regularization affect logistic regression?**  
   a) It increases the magnitude of the weights.  
   b) It reduces the magnitude of the weights.  
   c) It has no effect on the weights.  
   d) It makes the decision boundary non-linear.

**Answer**: b) It reduces the magnitude of the weights.

1. **What is the effect of increasing the number of features in logistic regression?**  
   a) The model becomes simpler.  
   b) The model becomes more prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model becomes more prone to overfitting.

1. **What is the effect of reducing the number of features in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes more sensitive to outliers.

**Answer**: b) The model becomes less prone to overfitting.

1. **What is the effect of increasing the number of training examples in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes more sensitive to outliers.

**Answer**: b) The model becomes less prone to overfitting.

1. **What is the effect of reducing the number of training examples in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes more prone to overfitting.  
   c) The model’s performance always improves.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model becomes more prone to overfitting.

1. **What is the effect of increasing the learning rate in logistic regression?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: a) The model converges faster.

1. **What is the effect of reducing the learning rate in logistic regression?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: b) The model converges slower.

1. **What is the effect of increasing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: b) The model becomes less complex.

1. **What is the effect of reducing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: a) The model becomes more complex.

**Performance Measures**

1. **Accuracy**: Measures the proportion of correctly classified instances.
2. **Precision**: Measures the proportion of true positives out of all predicted positives.
3. **Recall (Sensitivity)**: Measures the proportion of true positives out of all actual positives.
4. **F1 Score**: Harmonic mean of precision and recall, useful for imbalanced datasets.
5. **What does accuracy measure in a classification model?**  
   a) The proportion of true positives out of all predicted positives.  
   b) The proportion of correctly classified instances out of all instances.  
   c) The proportion of true positives out of all actual positives.  
   d) The balance between precision and recall.

**Answer**: b) The proportion of correctly classified instances out of all instances.

1. **Which of the following is true about precision?**  
   a) It measures the proportion of true positives out of all predicted positives.  
   b) It measures the proportion of true positives out of all actual positives.  
   c) It measures the proportion of correctly classified instances out of all instances.  
   d) It measures the balance between precision and recall.

**Answer**: a) It measures the proportion of true positives out of all predicted positives.

1. **What does recall (sensitivity) measure?**  
   a) The proportion of true positives out of all predicted positives.  
   b) The proportion of true positives out of all actual positives.  
   c) The proportion of correctly classified instances out of all instances.  
   d) The balance between precision and recall.

**Answer**: b) The proportion of true positives out of all actual positives.

1. **What is the F1 score?**  
   a) The arithmetic mean of precision and recall.  
   b) The harmonic mean of precision and recall.  
   c) The geometric mean of precision and recall.  
   d) The weighted average of precision and recall.

**Answer**: b) The harmonic mean of precision and recall.

1. **Which of the following is true about the F1 score?**  
   a) It is useful when the classes are balanced.  
   b) It is useful when the classes are imbalanced.  
   c) It is the same as accuracy.  
   d) It ignores false negatives.

**Answer**: b) It is useful when the classes are imbalanced.

1. **What is the formula for accuracy?**  
   a) TP+TNTP+TN+FP+FNTP+TN+FP+FNTP+TN​  
   b) TPTP+FPTP+FPTP​  
   c) TPTP+FNTP+FNTP​  
   d) 2⋅Precision⋅RecallPrecision+Recall2⋅Precision+RecallPrecision⋅Recall​

**Answer**: a) TP+TN/TP+TN+FP+FN ​

1. **What is the formula for precision?**  
   a) TP+TNTP+TN+FP+FNTP+TN+FP+FNTP+TN​  
   b) TPTP+FPTP+FPTP​  
   c) TPTP+FNTP+FNTP​  
   d) 2⋅Precision⋅RecallPrecision+Recall2⋅Precision+RecallPrecision⋅Recall​

**Answer**: b) TP/TP+FP ​

1. **What is the formula for recall (sensitivity)?**  
   a) TP+TNTP+TN+FP+FNTP+TN+FP+FNTP+TN​  
   b) TPTP+FPTP+FPTP​  
   c) TPTP+FNTP+FNTP​  
   d) 2⋅Precision⋅RecallPrecision+Recall2⋅Precision+RecallPrecision⋅Recall​

**Answer**: c) TP/TP+FN ​

1. **What is the formula for the F1 score?**  
   a) TP+TNTP+TN+FP+FNTP+TN+FP+FNTP+TN​  
   b) TPTP+FPTP+FPTP​  
   c) TPTP+FNTP+FNTP​  
   d) 2⋅Precision⋅RecallPrecision+Recall2⋅Precision+RecallPrecision⋅Recall​

**Answer**: d) 2⋅Precision⋅RecallPrecision+Recall2⋅Precision+RecallPrecision⋅Recall​

1. **Which of the following is true about precision and recall?**  
   a) Precision focuses on the quality of positive predictions, while recall focuses on the completeness of positive predictions.  
   b) Precision focuses on the completeness of positive predictions, while recall focuses on the quality of positive predictions.  
   c) Precision and recall are the same.  
   d) Precision and recall are not related.

**Answer**: a) Precision focuses on the quality of positive predictions, while recall focuses on the completeness of positive predictions.

1. **What is the effect of increasing the threshold in a classification model?**  
   a) Precision increases, recall decreases.  
   b) Precision decreases, recall increases.  
   c) Both precision and recall increase.  
   d) Both precision and recall decrease.

**Answer**: a) Precision increases, recall decreases.

1. **What is the effect of decreasing the threshold in a classification model?**  
   a) Precision increases, recall decreases.  
   b) Precision decreases, recall increases.  
   c) Both precision and recall increase.  
   d) Both precision and recall decrease.

**Answer**: b) Precision decreases, recall increases.

1. **Which of the following is true about the F1 score?**  
   a) It is the arithmetic mean of precision and recall.  
   b) It is the harmonic mean of precision and recall.  
   c) It is the geometric mean of precision and recall.  
   d) It is the weighted average of precision and recall.

**Answer**: b) It is the harmonic mean of precision and recall.

1. **What is the effect of a high F1 score?**  
   a) The model has high precision and high recall.  
   b) The model has high precision and low recall.  
   c) The model has low precision and high recall.  
   d) The model has low precision and low recall.

**Answer**: a) The model has high precision and high recall.

1. **Which of the following is true about accuracy?**  
   a) It is a good metric for imbalanced datasets.  
   b) It is a good metric for balanced datasets.  
   c) It is the same as precision.  
   d) It is the same as recall.

**Answer**: b) It is a good metric for balanced datasets.

1. **What is the effect of a high precision score?**  
   a) The model has a low false positive rate.  
   b) The model has a low false negative rate.  
   c) The model has a high false positive rate.  
   d) The model has a high false negative rate.

**Answer**: a) The model has a low false positive rate.

1. **What is the effect of a high recall score?**  
   a) The model has a low false positive rate.  
   b) The model has a low false negative rate.  
   c) The model has a high false positive rate.  
   d) The model has a high false negative rate.

**Answer**: b) The model has a low false negative rate.

1. **Which of the following is true about the F1 score?**  
   a) It is useful when the cost of false positives and false negatives is similar.  
   b) It is useful when the cost of false positives is much higher than the cost of false negatives.  
   c) It is useful when the cost of false negatives is much higher than the cost of false positives.  
   d) It is not useful for classification tasks.

**Answer**: a) It is useful when the cost of false positives and false negatives is similar.

1. **What is the effect of a high accuracy score?**  
   a) The model has a low false positive rate.  
   b) The model has a low false negative rate.  
   c) The model has a high false positive rate.  
   d) The model has a high false negative rate.

**Answer**: a) The model has a low false positive rate.

1. **Which of the following is true about the F1 score?**  
   a) It is the arithmetic mean of precision and recall.  
   b) It is the harmonic mean of precision and recall.  
   c) It is the geometric mean of precision and recall.  
   d) It is the weighted average of precision and recall.

**Answer**: b) It is the harmonic mean of precision and recall.

Miscellaneous questions:

1. **A decision tree trained on a dataset with highly correlated features is likely to:**  
   a) Perform better due to increased information gain.  
   b) Perform worse due to redundant splits.  
   c) Have no effect on performance.  
   d) Automatically remove correlated features.

**Answer**: b) Perform worse due to redundant splits.

1. **If a decision tree splits a node based on a feature with very low variance, what is the most likely outcome?**  
   a) The split will improve the model’s performance.  
   b) The split will have little impact on impurity reduction.  
   c) The tree will overfit the data.  
   d) The tree will underfit the data.

**Answer**: b) The split will have little impact on impurity reduction.

1. **In a decision tree, if the Gini index of a node is 0.5, what does this imply about the class distribution?**  
   a) The node is perfectly pure.  
   b) The node is perfectly impure.  
   c) The classes are equally distributed.  
   d) The node contains only one class.

**Answer**: c) The classes are equally distributed.

1. **Why might a decision tree with very high depth perform poorly on unseen data, even if it performs well on the training data?**  
   a) It captures noise in the training data, leading to overfitting.  
   b) It fails to capture important patterns in the data.  
   c) It is too simple to model the data.  
   d) It has high bias.

**Answer**: a) It captures noise in the training data, leading to overfitting.

1. **If a decision tree is pruned aggressively, what is the most likely effect on the model?**  
   a) It will overfit the data.  
   b) It will underfit the data.  
   c) It will have high variance.  
   d) It will perform well on unseen data.

**Answer**: b) It will underfit the data.

**K-Nearest Neighbors (K-NN)**

1. **If you increase the value of K*K* in K-NN, how does it affect the model’s sensitivity to noise?**  
   a) The model becomes more sensitive to noise.  
   b) The model becomes less sensitive to noise.  
   c) The model’s sensitivity to noise remains unchanged.  
   d) The model becomes more computationally expensive.

**Answer**: b) The model becomes less sensitive to noise.

1. **Why might K-NN perform poorly on high-dimensional data?**  
   a) Distances between points become less meaningful (curse of dimensionality).  
   b) K-NN cannot handle categorical features.  
   c) K-NN requires a large amount of training data.  
   d) K-NN is computationally expensive.

**Answer**: a) Distances between points become less meaningful (curse of dimensionality).

1. **What is the effect of using a very small K*K* in K-NN on the decision boundary?**  
   a) The decision boundary becomes smoother.  
   b) The decision boundary becomes highly irregular.  
   c) The decision boundary becomes linear.  
   d) The decision boundary becomes circular.

**Answer**: b) The decision boundary becomes highly irregular.

1. **How does the choice of distance metric (e.g., Euclidean vs. Manhattan) affect K-NN?**  
   a) It has no effect on the model’s performance.  
   b) It affects how the model measures similarity between points.  
   c) It determines the value of K*K*.  
   d) It reduces the computational complexity of the model.

**Answer**: b) It affects how the model measures similarity between points.

1. **If K-NN is applied to a dataset with imbalanced classes, what is the most likely issue?**  
   a) The model will always predict the majority class.  
   b) The model will perform well on both classes.  
   c) The model will have high precision but low recall.  
   d) The model will have low precision but high recall.

**Answer**: a) The model will always predict the majority class.

**Linear Regression**

1. **If the residuals of a linear regression model are not normally distributed, what does this imply?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The assumption of normality is violated.  
   b) The model is performing well.

**Answer**: c) The assumption of normality is violated.

1. **What is the effect of multicollinearity in linear regression?**  
   a) It improves the model’s performance.  
   b) It makes the model’s coefficients unstable.  
   c) It reduces the model’s complexity.  
   d) It has no effect on the model.

**Answer**: b) It makes the model’s coefficients unstable.

1. **If a linear regression model has a high R-squared value but poor performance on unseen data, what is the most likely issue?**  
   a) The model is underfitting.  
   b) The model is overfitting.  
   c) The model has high bias.  
   d) The model has low variance.

**Answer**: b) The model is overfitting.

1. **What is the effect of adding a polynomial feature (e.g., x2*x*2) to a linear regression model?**  
   a) The model becomes more complex and may overfit.  
   b) The model becomes simpler and may underfit.  
   c) The model’s performance remains unchanged.  
   d) The model’s coefficients become unstable.

**Answer**: a) The model becomes more complex and may overfit.

1. **If the residuals of a linear regression model show a pattern (e.g., a curve), what does this imply?**  
   a) The model is capturing all the patterns in the data.  
   b) The model is missing important non-linear patterns.  
   c) The model is overfitting.  
   d) The model has high variance.

**Answer**: b) The model is missing important non-linear patterns.

**Perceptron**

1. **If the perceptron algorithm fails to converge, what is the most likely reason?**  
   a) The learning rate is too high.  
   b) The data is not linearly separable.  
   c) The model is overfitting.  
   d) The model is underfitting.

**Answer**: b) The data is not linearly separable.

1. **What is the effect of increasing the learning rate in the perceptron algorithm?**  
   a) The model converges faster but may overshoot the optimal solution.  
   b) The model converges slower but is more accurate.  
   c) The model becomes more robust to noise.  
   d) The model becomes less sensitive to outliers.

**Answer**: a) The model converges faster but may overshoot the optimal solution.

1. **If the perceptron algorithm is applied to a dataset with non-linear decision boundaries, what is the most likely outcome?**  
   a) The model will perform well.  
   b) The model will fail to converge.  
   c) The model will overfit the data.  
   d) The model will underfit the data.

**Answer**: b) The model will fail to converge.

1. **What is the effect of using a very small learning rate in the perceptron algorithm?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model converges slower.

1. **If the perceptron algorithm is applied to a dataset with imbalanced classes, what is the most likely issue?**  
   a) The model will always predict the majority class.  
   b) The model will perform well on both classes.  
   c) The model will have high precision but low recall.  
   d) The model will have low precision but high recall.

**Answer**: a) The model will always predict the majority class.

**Logistic Regression**

1. **If the logistic regression model predicts probabilities close to 0.5 for most instances, what does this imply?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The model is performing well.  
   d) The model has high variance.

**Answer**: b) The model is underfitting.

1. **What is the effect of increasing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: b) The model becomes less complex.

1. **If the logistic regression model has high precision but low recall, what is the most likely issue?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The model is predicting too many false positives.  
   d) The model is predicting too many false negatives.

**Answer**: d) The model is predicting too many false negatives.

1. **What is the effect of using a very small learning rate in logistic regression?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less sensitive to outliers.

**Answer**: b) The model converges slower.

1. **If the logistic regression model is applied to a dataset with imbalanced classes, what is the most likely issue?**  
   a) The model will always predict the majority class.  
   b) The model will perform well on both classes.  
   c) The model will have high precision but low recall.  
   d) The model will have low precision but high recall.

**Answer**: a) The model will always predict the majority class.

**Performance Metrics**

1. **If a classification model has high accuracy but poor F1 score, what is the most likely issue?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The model is performing well on the majority class but poorly on the minority class.  
   d) The model has high variance.

**Answer**: c) The model is performing well on the majority class but poorly on the minority class.

1. **What is the effect of increasing the threshold in a classification model?**  
   a) Precision increases, recall decreases.  
   b) Precision decreases, recall increases.  
   c) Both precision and recall increase.  
   d) Both precision and recall decrease.

**Answer**: a) Precision increases, recall decreases.

1. **If a model has high recall but low precision, what is the most likely issue?**  
   a) The model is predicting too many false positives.  
   b) The model is predicting too many false negatives.  
   c) The model is overfitting.  
   d) The model is underfitting.

**Answer**: a) The model is predicting too many false positives.

1. **What is the effect of using the F1 score as a performance metric?**  
   a) It balances precision and recall.  
   b) It focuses only on precision.  
   c) It focuses only on recall.  
   d) It ignores false negatives.

**Answer**: a) It balances precision and recall.

1. **If a model has high precision but low recall, what is the most likely issue?**  
   a) The model is predicting too many false positives.  
   b) The model is predicting too many false negatives.  
   c) The model is overfitting.  
   d) The model is underfitting.

**Answer**: b) The model is predicting too many false negatives.

**Entropy and Gini Index**

1. **If the entropy of a dataset is 1, what does this imply about the class distribution?**  
   a) The dataset is perfectly pure.  
   b) The dataset is perfectly impure.  
   c) The classes are equally distributed.  
   d) The dataset contains only one class.

**Answer**: c) The classes are equally distributed.

1. **What is the effect of using the Gini index instead of entropy in a decision tree?**  
   a) The model becomes more computationally expensive.  
   b) The model becomes less computationally expensive.  
   c) The model becomes more sensitive to noise.  
   d) The model becomes less sensitive to noise.

**Answer**: b) The model becomes less computationally expensive.

1. **If the Gini index of a node is 0, what does this imply about the class distribution?**  
   a) The node is perfectly pure.  
   b) The node is perfectly impure.  
   c) The classes are equally distributed.  
   d) The node contains only one class.

**Answer**: a) The node is perfectly pure.

1. **What is the effect of using entropy instead of the Gini index in a decision tree?**  
   a) The model becomes more computationally expensive.  
   b) The model becomes less computationally expensive.  
   c) The model becomes more sensitive to noise.  
   d) The model becomes less sensitive to noise.

**Answer**: a) The model becomes more computationally expensive.

1. **If the entropy of a dataset is 0, what does this imply about the class distribution?**  
   a) The dataset is perfectly pure.  
   b) The dataset is perfectly impure.  
   c) The classes are equally distributed.  
   d) The dataset contains only one class.

**Answer**: a) The dataset is perfectly pure.

**Bias-Variance Tradeoff**

1. **If a model has high bias, what is the most likely issue?**  
   a) The model is too complex.  
   b) The model is too simple.  
   c) The model has low training error.  
   d) The model has low testing error.

**Answer**: b) The model is too simple.

1. **If a model has high variance, what is the most likely issue?**  
   a) The model is too complex.  
   b) The model is too simple.  
   c) The model has high training error.  
   d) The model has high testing error.

**Answer**: a) The model is too complex.

1. **What is the effect of increasing the model’s complexity?**  
   a) It reduces the risk of overfitting.  
   b) It increases the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It reduces the training error but increases the testing error.

**Answer**: b) It increases the risk of overfitting.

1. **What is the effect of reducing the model’s complexity?**  
   a) It reduces the risk of overfitting.  
   b) It increases the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It reduces the training error but increases the testing error.

**Answer**: a) It reduces the risk of overfitting.

1. **What is the effect of increasing the amount of training data?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases the training error but reduces the testing error.

**Answer**: b) It reduces the risk of overfitting.

**Overfitting and Underfitting**

1. **If a model performs well on the training data but poorly on the testing data, what is the most likely issue?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The model has high bias.  
   d) The model has low variance.

**Answer**: a) The model is overfitting.

1. **If a model performs poorly on both the training and testing data, what is the most likely issue?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The model has high variance.  
   d) The model has low bias.

**Answer**: b) The model is underfitting.

1. **What is the effect of using cross-validation?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases the training error but reduces the testing error.

**Answer**: b) It reduces the risk of overfitting.

1. **What is the effect of reducing the number of features in a model?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases the training error but reduces the testing error.

**Answer**: b) It reduces the risk of overfitting.

1. **What is the effect of increasing the regularization parameter in a model?**  
   a) It increases the risk of overfitting.  
   b) It reduces the risk of overfitting.  
   c) It has no effect on model performance.  
   d) It increases the training error but reduces the testing error.

**Answer**: b) It reduces the risk of overfitting.

**Miscellaneous**

1. **If the residuals of a linear regression model are not normally distributed, what does this imply?**  
   a) The model is overfitting.  
   b) The model is underfitting.  
   c) The assumption of normality is violated.  
   d) The model is performing well.

**Answer**: c) The assumption of normality is violated.

1. **What is the effect of increasing the learning rate in gradient descent?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: a) The model converges faster.

1. **What is the effect of reducing the learning rate in gradient descent?**  
   a) The model converges faster.  
   b) The model converges slower.  
   c) The model becomes more robust to noise.  
   d) The model becomes less accurate.

**Answer**: b) The model converges slower.

1. **What is the effect of increasing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: b) The model becomes less complex.

1. **What is the effect of reducing the regularization parameter in logistic regression?**  
   a) The model becomes more complex.  
   b) The model becomes less complex.  
   c) The model becomes more sensitive to outliers.  
   d) The model becomes less accurate.

**Answer**: a) The model becomes more complex.