CRASH COURSE GATE 2025

Machine Learning

SVM

Q1 What is the main objective of an SVM?
(A) Maximize the margin between classes
(B) Minimize the distance between support vectors
(C) Maximize the number of support vectors
(D) Minimize the number of misclassified points

Q2 Which kernel is often used in SVM to handle non-linear data?
(A) Linear kernel
(B) Polynomial kernel
(C) Radial Basis Function (RBF) kernel
(D) Sigmoid kernel

Q3 What does the term 'support vector' refer to in SVM?

(A) A point on the decision boundary
(B) A point that lies closest to the decision boundary
(C) A point that lies furthest from the decision boundary
(D) Any point that is misclassified

Q4 Which of the following is true about the SVM with a linear kernel?
(A) It can only be used for binary classification
(B) It cannot handle linearly inseparable data
(C) It can find a non-linear decision boundary
(D) It requires a very large dataset

Q5 What is the purpose of the C parameter in SVM?
(A) To control the degree of the polynomial kernel
(B) To control the trade-off between maximizing the margin and minimizing the classification error
(C) To determine the width of the RBF kernel
(D) To control the maximum number of iterations

Q6 In the context of SVM, what is the 'kernel trick?
(A) A method to speed up the convergence of the SVM algorithm
(B) A technique to transform non-linearly separable data into a higher-dimensional space where it becomes linearly separable
(C) A technique to reduce the number of support vectors
(D) A method to minimize the computational cost of training an SVM

Q7 Which of the following statements is true about SVM in high-dimensional spaces?
(A) SVMs tend to overfit the data in high-dimensional spaces
(B) SVMs are particularly effective in high-dimensional spaces
(C) The performance of SVMs does not change with dimensionality
(D) SVMs require more support vectors in high-dimensional spaces

Q8 What is the hinge loss in SVM?

(A) A loss function used to penalize incorrect classifications
(B) A measure of the distance between support vectors
(C) A function used to calculate the margin
(D) A method to reduce the number of features

Q9 Which algorithm is commonly used to solve the optimization problem in SVM?

(A) Gradient Descent
(B) Newton's Method
(C) Sequential Minimal Optimization (SMO)
(D) K-Means

Q10 Given a 2D dataset, consider two support vectors: (1,1) and (3.3). If the separating hyperplane is $x_1 + x_2 - 3 = 0$, what is the margin?

(A) 0.5 (B) 1

(C) $\sqrt{2}$ (D) $\frac{2}{\sqrt{2}}$

Q11 Consider a linear SVM with the weight vector w = [4,3| and bias b = –5. What is the value of y for the point (1,1) in the decision function f(x) = w · x + b?

(A) 5root2 (B) 3root2

(C) 4root2 (D) 2root2

Q12 If an SVM model has a weight vector w = [1, 2] and bias b = – 3, what is the decision boundary equation?

(A) $x_1 + 2x_2 - 3 = 0$

(B) $x_1 + 2x_2 + 3 = 0$

(C) $2x_1 + x_2 - 3 = 0$

(D) $2x_1 + x_2 + 3 = 0$

Answer Key

Q1	(A)
Q2	(C)
Q3	(B)
Q4	(B)
Q5	(B)
Q6	(B)
Q7	(B)
Q8	(A)
Q9	(C)
Q10	(B)
Q11	(D)
Q12	(A)

Hints & Solutions

Q1 Text Solution:

A) **Maximize the margin between classes**: SVM finds the hyperplane that separates classes in such a way that the distance (margin) between the hyperplane and the nearest data points (support vectors) from each class is maximized. This margin maximization helps in achieving better generalization and robustness of the classifier.
B) Minimize the distance between support vectors: This is not the primary objective of SVM. While SVM does use support vectors to define the decision boundary, the focus is on maximizing the margin around the decision boundary, not minimizing the distance between support vectors.
C) Maximize the number of support vectors: SVM aims to find a decision boundary using as few support vectors as possible while still maximizing the margin. Maximizing the number of support vectors is not the main goal.
D) Minimize the number of misclassified points: While SVM does aim to find a decision boundary that separates classes accurately, its primary objective is not specifically to minimize the number of misclassified points. The focus is on margin maximization.

Q2 Text Solution:

Radial Basis Function (RBF) kernel because it is highly effective for handling non-linear data in Support Vector Machines (SVM). The RBF kernel transforms the input data into a higher-dimensional space, allowing the SVM to create complex decision boundaries. This capability enables the SVM to classify data that is not linearly separable in the original feature space. The flexibility of the RBF kernel to adjust to the data's intricacies makes it a popular choice for various applications, providing a good balance between accuracy and computational efficiency.

Q3 Text Solution:

Support Vector Machines (SVM), the term 'support vector' refers to the data points from the training dataset that are closest to the decision boundary between the classes. These points are crucial because they directly influence the position and orientation of the decision boundary. Here's why option B is correct:
B) A point that lies closest to the decision boundary: Support vectors are specifically the points that lie closest to the decision boundary. They are the critical elements in defining the maximum margin (distance) between the classes in SVM. These points essentially "support" the construction of the decision boundary by being the ones that the algorithm relies on to determine the optimal hyperplane that separates the classes with the largest possible margin.

Q4 Text Solution:

Option (B) is correct because SVM with a linear kernel cannot handle linearly inseparable data. The linear kernel in SVM aims to find a linear decision boundary that separates classes in the feature space. If the data is not linearly separable, meaning classes cannot be separated by a straight line, the SVM with a linear kernel will fail to accurately classify such data. In such cases, nonlinear kernels like polynomial or radial basis function (RBF) kernels are used in SVM to map the data into a higher-dimensional space where classes can be separated by a hyperplane. Therefore, SVM with a linear kernel is limited to handling linearly separable data.

Q5 Text Solution:

The correct answer is (B): To control the trade-off between maximizing the margin and minimizing the classification error.
The C parameter in SVM (Support Vector Machine) adjusts the penalty for misclassification of training examples. A smaller C encourages a larger margin, potentially allowing more training examples to be misclassified (if they lie outside the margin), while a

larger C penalizes misclassifications more heavily, potentially resulting in a smaller margin. Thus, C controls the balance between achieving a wider margin (better generalization) and correctly classifying more training examples.

Q6 Text Solution:

The correct answer is (B): A technique to transform non-linearly separable data into a higher-dimensional space where it becomes linearly separable.

The kernel trick allows SVMs to handle non-linearly separable data by implicitly mapping it into a higher-dimensional space where the data points can be separated by a hyperplane. This transformation is achieved without explicitly

Q7 Text Solution:

Option (B) is correct because SVMs are particularly effective in high-dimensional spaces. In high-dimensional spaces, data points are more likely to be linearly separable, which aligns well with SVM's ability to find optimal hyperplanes. This makes SVMs a preferred choice for classification tasks where the number of features (dimensions) is large

Q8 Text Solution:

Option (A) is correct because the hinge loss in SVM is a loss function used to penalize incorrect classifications. It quantifies the loss incurred when a model predicts the wrong class label for an instance. The hinge loss encourages the SVM model to correctly classify instances by increasing the loss as the margin violation increases.

Q9 Text Solution:

Option (C) is correct because Sequential Minimal Optimization (SMO) is commonly used to solve the optimization problem in SVM. SMO is an algorithm specifically designed for training SVMs by iteratively solving two-variable optimization problems, which allows for efficient training of SVM models even with large datasets. It focuses on optimizing the dual problem of SVMs, aiming to find the optimal values of the Lagrange multipliers (also known as support vector coefficients) that define the separating hyperplane.

Q10 Text Solution:

Option (B) is correct because the margin in SVM is defined as the distance from the hyperplane to the nearest support vector.

Given the hyperplane equation (x1 + x2 - 3 = 0), we can rewrite it in the form (x2 = -x1 + 3).

The distance from a point ((x_1, x_2)) to this line (hyperplane) is given by the formula (frac{|Ax1 + By + C|}{sqrt(A^2 + B^2)}

Q11 Text Solution:

To find the value of (y) for the point (1,1) using the decision function f(x) = w x + b?= where (w = [4, 3]^T) and (b = -5), we proceed as follows:

w·x=4·1+3·1=4+3=7

Now, add the bias b:

f(x)=w·x+b=7−5=2

Therefore, the value of y for the point (1,1) in the decision f(x) is 2root2

Q12 Text Solution:

To determine the decision boundary equation from the given SVM weight vector (w = [1, 2]) and bias (b = -3), we use the formula for the decision boundary in SVM, which is w·x+b=0.

Given (w = [1, 2]) and (b = -3):

1. Construct the Decision Boundary Equation:

 1·x1+2·x2−3=0

2. **Simplify the Equation:**

 x1+2x2−3=0

Therefore, the decision boundary equation for this SVM model is A