Machine Learning

GATE

Support Vector Machines

DPP : 01

Q1 Which of following statements is/are correct ?

S_1: Linear support vector machine is a parametric algorithm.

S_2: KNN is a non-parametric algorithm.

(A) Only S_1 (B) Only S_2

(C) Both S_1 and S_2 (D) None of these

Q2 Which of the following statement is/are correct ?

(A) Supervised learning model use labeled data for training.

(B) Unsupervised learning model use labeled data for training.

(C) A supervised learning model gives accurate result.

(D) An unsupervised learning model may be less accurate than supervised learning model.

Q3 In a supervised learning dataset, if there are 100 labeled examples and each example has 10 features, how many labels are there in total ?

(A) 10 (B) 100

(C) 1000 (D) 10000

Q4 You are training an SVM classifier for a binary classification task. During training, the SVM encounters a data point that lies exactly on the decision boundary. What is the impact of this data point on the optimal hyperplane learned by the SVM ?

(A) The SVM will completely ignore this data point as it provides no information for separating the classes.

(B) The presence of this data point will force the SVM to adjust the hyperplane to maximize the margin, even if it slightly misclassifies the point.

(C) The SVM will consider this data point as a support vector and use it to define the final hyperplane.

(D) The presence of this data point will have no effect on the SVM's decision boundary.

Q5 Consider a binary classification problem where you are training a Support Vector Machine (SVM) classifier with a linear kernel. The primal optimization problem for SVM can be formulated as:

Minimize. $\frac{1}{2}||\mathrm{w}||^2$

Subject to yi (wT xi + b) ≥ 1, for i = 1,2, …, n

Where w is the weight vector, b is the bias term, xi are the feature vectors, and yi are the class labels

(either +1 or –1).

Which of the following statements regarding the dual form of the SVM optimization problem are TRUE?

(A) The dual form of the SVM optimization problem involves maximizing a quadratic function subject to equality constraints.

(B) In the dual form, Lagrange multipliers (also known as dual variables) are introduced to represent the importance of each training example in defining the decision boundary.

(C) The dual form allows for the computation of the weight vector directly.

(D) Solving the dual form is typically more efficient than solving the primal form when the number of features is large compared to the number of samples.

Q6 In a classification dataset, there are 400 labeled examples belonging to 5 distinct classes. If each class has an equal number of examples, how many examples are there in each class?

(A) 80 (B) 100

(C) 200 (D) 400

Q7 Which evaluation metrics is commonly used for classification tasks when class imbalance is present ?

(A) Mean squared error

(B) Accuracy

(C) F1-score

(D) R-squared

Q8 How many statements are FALSE?

(A) Support Vector Machines, similar to logistic regression models, provide a probability distribution over the potential labels for a given input example.

(B) In general, as we transition from a linear kernel to higher-order polynomial Kernels, we anticipate the support vectors to largely remain unchanged.

(C) Support Vector Machines construct decision boundaries with maximum margins, and these boundaries generally result in lower generalization errors compared to other linear classifiers.

(D) Any decision boundary obtained from a generative model with class-conditional Gaussian distributions could, in principle, be replicated using an SVM with a polynomial kernel of degree three or lower.

Q9 You are working with a high-dimensional dataset and have trained a Support Vector Machine (SVM) for binary classification. The SVM decision boundary is given by the equation :

$4x_1 - 2x_2 + 3x_3 + 5x_4 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10} - 15 = 0.$

where $x_1, x_2, \ldots\ldots, x_{10}$ are the features.

Now, if a data point is situated on the decision boundary with $x_1 = 1$ and $x_3 = 2$, what is the corresponding value for x_7?

(A) –1.5 (B) –1.25

(C) –1 (D) –0.625

Q10 If the SVM classifier has a decision function $f(x) = 4x_1 - 5x_2 + 1$, and the margin is 2, what is the distance from the origin to the decision boundary?

Q11 For an SVM with a linear decision boundary, if the weight vector is w = [3, – 2] and the bias b = – 1 what is the decision function for a point (4, 2) ?

Q12 Given two points (2, 3) and (4,5) with the same class, and their distances from the decision boundary are 1 and 2 respectively. What is the width of the margin if these points are support vectors?

Q13 If the support vectors of an SVM are (1, 2) and (3, 4) with labels +1 and –1 respectively, and the decision boundary is defined as $2x_1 + 3x_2 - 5 = 0$, calculate the decision value for the point (2, 3).

Q14 In a binary SVM classification with the decision function $f(x) = 3x_1 - 4x_2 + 2$, calculate the value of f(x) for the point (1, –1).

Q15 Suppose the decision boundary of an SVM is $x_1 + 2x_2 - 3 = 0$. What is the margin if the norm of the weight vector w is 1?

Q16 If an SVM model has a decision boundary $f(x) = 0.5\ x_1 + 5x_2 - 1 = 0$, and the support vectors are at a distance of 1 from the decision boundary, what is the margin?

Q17 For an SVM with kernel function $K(x, y) = (x \cdot y)^2$, if x = (1, 2) and y = (2, 1), compute K(x , y).

Q18 Suppose the SVM has a decision function $f(x) = 2x_1 - 3x_2 + 4$ For the point (1, 2), calculate the decision function value.

Q19 In a SVM with a linear kernal, the weight vector w = [1, –1] and bias b = 2. For the point (2, 1), find the decision function value.

Q20 Calculate the decision boundary equation for an SVM with support vectors at (1, 1) and (2,2) with lables + 1 and – 1, respectively.

Answer Key

Q1	(C)
Q2	(A, C, D)
Q3	(B)
Q4	(C)
Q5	(B, C)
Q6	(A)
Q7	(C)
Q8	(A, B, C)
Q9	(D)
Q10	1
Q11	7
Q12	2
Q13	8
Q14	9
Q15	2
Q16	2
Q17	16
Q18	0
Q19	3
Q20	– 1

Hints & Solutions

Q1 Text Solution:

S_1: Linear support vector machine is a parametric because it can only produce linear boundaries.

So, S_1 is true.

S_2: KNN is a non-parametric algorithm because it avoids a priori assumption about the shape of the class boundary and can thus adapt more closely to non-linear boundaries as the amount of training data increases.

So, S_2 is true.

Q3 Text Solution:

In supervised learning, each example in the dataset is associated with a label. Since there are 100 labeled examples in the dataset, there are a total of 100 labels.

Q4 Text Solution:

(A) Incorrect: Even data points on the boundary can provide information about the class separation.

They can potentially become support vectors if they are closest to the decision boundary.

(B) Incorrect: The SVM aims to maximize the margin, but it won't misclassify points to achieve this.

A data point on the boundary already maximizes the margin.

(C) Correct: Support vectors are the data points closest to the decision boundary on either side.

A data point exactly on the boundary is a perfect candidate for a support vector and will directly influence the hyperplane's position.

(D) Incorrect: The data point on the boundary will have an impact on the hyperplane's location since it

defines the closest point to the boundary.

Q5 Text Solution:

(A) Correct: The dual form involves maximizing a quadratic function (the dual objective) subject to equality constraints (the Karush-Kuhn-Tucker conditions).

(B) Correct: In the dual form, Lagrange multipliers (dual variables) are introduced to represent the importance of each training example in defining the decision boundary. These multipliers are used to compute the weight vector w.

(C) Incorrect: The dual form does not allow for the direct computation of the weight vector w. Instead, w is expressed as a linear combination of the support vectors.

(D) Correct: Solving the dual form is typically more efficient than solving the primal form when the number of features is large compared to the number of samples, as the dual form involves optimization over the number of samples rather than the number of features.

This question tests understanding of the optimization involved in SVM and the concept of the dual form, including the role of Lagrange multipliers and the efficiency of solving the dual form compared to the primal form.

Q6 Text Solution:

Since each class has an equal number of examples,

The total number of examples divided by the number of classes gives the number of examples per class:

$\frac{400}{5} = 80.$

Q7 Text Solution:

A common evaluation metric for classification tasks with class imbalance is the F1 score.

Q8 Text Solution:

Support Vector Machines, similar to logistic regression models, provide a probability distribution over the potential labels for a given input example.

Answer with Explanation:

The statement is incorrect. Support Vector Machines (SVMS) are not inherently probabilistic models like logistic regression. While SVMs can be used for classification tasks. Their primary objective is to find the hyperplane that best separates different classes. SVMs do not naturally provide a probability distribution over class labels.

However, there is a variant called "Support Vector Machines for Classification" (SVC) that can be used in a probabilistic manner. By enabling the 'probability' parameter in some implementations of SVM (e.g., scikit-learn's SVC), the model can output probability estimates using techniques like Platt scaling. This allows SVM to provide probability scores, but it's important to note that SVMS are not originally designed for probabilistic predictions.

In general, as we transition from a linear Kernel to higher-order polynomial Kernels, we anticipate the support vectors to largely remain unchanged.

Answer with Explanation:

The statement suggests that when moving from a linear Kernel to higher-order polynomial kernels. the set of support vectors is expected to remain relatively consistent.

Explanation:

- Support vectors are data points that play a crucial role in determining the decision boundary in SVM. They are the instances closest to the decision boundary and contribute to the definition of the margin.
- When using a linear Kernel, the decision boundary is a hyperplane in the original feature space.

Q9 Text Solution:

The SVM decision boundary equation is given by :

$4x_1 - 2x_2 + 3x_3 + 5x_4 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10} - 15 = 0$.

Now, substitute $x_1 = 1$ and $x_3 = 2$ into the equation :

$4(1) - 2x_2 + 3(2) + 5x_4 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10} - 15 = 0$.

Simplify the equation :

$4 - 2x_2 + 6 + 5x_4 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10} = 15$

Combine like terms :

$-2x_2 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10} = 5$

Now, isolate x_7 :

$-8x_7 = 5 - (-2x_2 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10})$

$x_7 = -\frac{1}{8}(5 - 2x_2 - 7x_5 + 6x_6 - 8x_7 - 2x_8 - x_9 + 10x_{10})$

Now, substitute $x_2 = 0, x_5 = 0, x_6 = 0, x_8 = 0, x_9 = 0, x_{10} = 0$,

(making the corresponding terms 0) to simplify the expression.

$$x_7 = -\frac{1}{8}(5 - 0 + 0 + 0 + 0 + 0) = -\frac{1}{8}(5)$$

$$x_7 = -\frac{5}{8} = -0.625$$

Q10 Text Solution:

The distance from the origin to the decision boundary can be calculated using the margin. For SVMs, the margin is the distance between the decision boundary and the support vectors.

In SVMs, the margin is given by:

Margin = $\frac{2}{||w||}$

where w is the weight vector of the decision boundary.

Given the margin is 2:

$$2 = \frac{2}{||w||}$$

$$||w|| = 1$$

The distance from the origin to the decision boundary is the margin divided by 2:

Distance = $\frac{\text{Margin}}{2} = 1$.

Q11 Text Solution:

The decision function is given by:

$$f(x) = w \cdot x + b$$

Substitute w = [3, – 2] x = [4, 2], and b = – 1

f(x) = $(3 \cdot 4) + (-2 \cdot 2) - 1$

f(x) = 12 – 4 – 1 = 7.

Q12 Text Solution:

In an SVM, the width of the margin is twice the distance of a support vector from the decision boundary. Here, the distances of the support vectors are 1 and 2. The margin width is the distance between these support vectors:

Margin Width = 2 × (Minimum Distance)
Margin Width = 2 × 1 = 2.

Q13 Text Solution:

The decision function is :
$f(x) = 2x_1 + 3x_2 - 5$
Substitute $x_1 = 2$ and $x_2 = 3$
$f(x) = 2 \cdot 2 + 3 \cdot 3 - 5$
$f(x) = 4 + 9 - 5 = 8$

Q14 Text Solution:

Substitute $x_1 = 1$ and $x_2 = -1$ into the decision function :
$f(x) = 3 \cdot 1 - 4 \cdot (-1) + 2$
$f(x) = 3 + 4 + 2 = 9.$

Q15 Text Solution:

For an SVM, the margin is calculated as :
Margin = $\frac{2}{||w||}$
Given $||w|| = 1$:
Margin = $\frac{2}{1} = 2.$

Q16 Text Solution:

In SVM, the margin is twice the distance of the support vectors from the decision boundary:
Margin = 2 × (Distance of Support Vector)
Given the distance of the support vectors is 1:
Margin = 2 × 1 = 2.

Q17 Text Solution:

Using the kernal function:
$K(x, y) = (x \cdot y)^2$
Calculate $x \cdot y$:
$x \cdot y = (1 \cdot 2) + (2 \cdot 1) = 2 + 2 = 4$
Then :
$K(x, y) = 4^2 = 16.$

Q18 Text Solution:

Substitute $x_1 = 1$ and $x_2 = 2$ into the decision function :
$f(x) = 2 \cdot 1 - 3 \cdot 2 + 4$
$f(x) = 2 - 6 + 4 = 0$
$f(x) = 4 - 4 = 0.$

Q19 Text Solution:

The decision function is:
$f(x) = w \cdot x + b$
Substitute $w = [1, -1]$, $x = [2, 1]$, and $b = 2$:
$f(x) = (1 \cdot 2) + (-1 \cdot 1) + 2$
$f(x) = 2 - 1 + 2 = 3.$

Q20 Text Solution:

For an SVM, the decision boundary equation is :
$w \cdot x + b = 0$
Using support vectors to determine w and b :
$w = [2 - 1, 2 - 1] = [1, 1]$
To find b, use one of the support vectors say (1, 1) and the label + 1 :
$1 \cdot 1 + 1 \cdot 1 + b = 1$
$2 + b = 1$
$b = 1 - 2 = -1.$