Week_10_mlt_questions

2024 Apr 28 ET

1. **(MSQ)** Given a two-dimensional data set where points from class 1 are:

$$\{(-2,3),\ (-1,1),\ (-1,2),\ (-1,4)\}$$

And points from class 0 are:

$$\{(1,3),\ (1,4),\ (2,4),\ (2,2)\}$$

Which of the following statements are true?

(a) The given data points from classes 1 and 0 can be linearly separated using a Hard-margin SVM.

(b) A perceptron model and a hard margin SVM can give different decision boundary for this dataset.

(c) A Soft-margin SVM would be a more robust choice than a Hard-margin SVM for this dataset as the dataset is not linearly separable.

(d) The width of the separation between the two supporting hyperplanes is 4. (Hint: Calculate width using formulae $\frac{2}{\|\mathbf{w}\|}$)

[Ans: a, b]

2024 Sep 01 ET

2. A hard-margin SVM is trained on a linearly separable dataset with a positive margin. The features are in $\mathbb{R}^2$. The optimal weight vector is $\begin{bmatrix}3\\4\end{bmatrix}$. Find the distance between the two supporting hyperplanes.

[Ans: 0.4]

3. **Comprehension:**
Consider a linearly separable dataset with a positive margin. The symbol α_i^* in the context of SVMs has its usual meaning. Are the following statements true or false?

i. The weight vector output by the perceptron algorithm on this dataset can be expressed as a linear combination of the data-points where the coefficients of the linear combination are integers.

[Ans: TRUE]

ii. In the case of a hard-margin SVM, if $\alpha_i^* \geq 0$, the point $\mathbf{x}_i$ is a support vector.

[Ans: TRUE]

iii. If a soft-margin SVM is trained on this dataset, the optimal weight vector it returns will be the same as the one returned by a hard-margin SVM, irrespective of the value of the hyperparameter C.

[Ans: FALSE]

2024 Dec 22 ET

4. Consider the following optimization problem involving the variable $\mathbf{x} \in \mathbb{R}^2$:

$$\min_{\mathbf{x}} x_1^2 + 2x_2 \quad \text{subject to} \quad x_1 - x_2 \geq 3.$$

What is the output of the following expression for $\mathbf{x}' = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$? Here, $L(\mathbf{x}', \lambda)$ is the Lagrangian function for this optimization problem evaluated at $\mathbf{x}'$.

$$\max_{\lambda \geq 0} L(\mathbf{x}', \lambda)$$

(a) $(x_1')^2 + 2(x_2')$

(b) ∞

(c) $-\infty$

(d) 0

(e) None of these

[Ans: b]

5. If we remove all the non-support vectors from the dataset, what will be the impact on the model using SVM algorithm?

(a) Model will overfit.

(b) The model will not be changed.

(c) Model will underfit.

(d) Accuracy of the model will increase.

[Ans: b]

2025 Aug 31 ET

6. Consider a dataset $D = \{(x_i, y_i) : (-1, -1), (-2, -1), (2, +1), (1, +1)\}$, where $x_i \in \mathbb{R}$ denotes the features and $y_i \in \{+1, -1\}$ denotes the class label. Suppose we fit a hard-margin linear SVM for this dataset, then which of the following points are likely to be support vectors?

(a) $(-1, -1)$ and $(2, +1)$

(b) $(-1, -1)$ and $(1, +1)$

(c) $(-2, -1)$ and $(2, +1)$

(d) $(-2, -1)$ and $(1, +1)$

[Ans: b]

7. **(MSQ)** Given a two-dimensional dataset where points from class 1 are: $\{(-2, 3), (-1, 1)\}$ and points from class 0 are: $\{(1, 3), (1, 4)\}$. Which of the following statements are true?

(a) The given data points from classes 1 and 0 can be linearly separated using a hard-margin SVM.

(b) A perceptron model and a hard-margin SVM can give different decision boundary for this dataset.

(c) A soft-margin SVM would be a more robust choice than a hard-margin SVM for this dataset as the dataset is not linearly separable.

(d) The width of the separation between the two supporting hyperplanes is 4.

[Ans: a, b]

2025 Dec 18 ET

8. Consider an SVM classifier with separating hyperplane defined by $w^T x = 0$, where the weight vector is $w^T = \begin{bmatrix} 2 & -3 & 1 \end{bmatrix}$. Compute the perpendicular distance from the point $x^T = \begin{bmatrix} 1 & 2 & -1 \end{bmatrix}$ to the hyperplane. Enter your answer correct to three decimal places.

[Ans: 1.333 to 1.339]

2026 May 06 ET

9. Consider a hard-margin SVM trained on a linearly separable dataset. The optimal Lagrange multipliers are $\alpha = \begin{bmatrix} 2 & 0 & 3 & 0 \end{bmatrix}^T$. The training dataset ($\{(\mathbf{x}_i, y_i)\}_{i=1}^4 \subset \mathbb{R}^2$) is given by

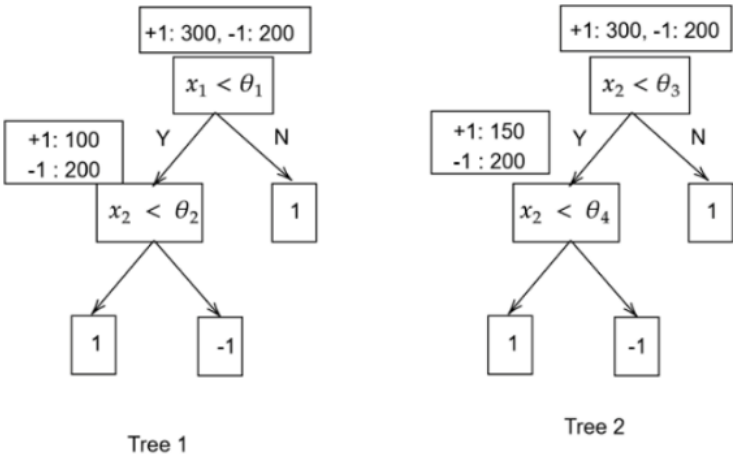
$$D = \{((1,2),+1),\ ((2,2),+1),\ ((2,-1),-1),\ ((1,-2),-1)\}.$$

Given the test point $\mathbf{x}_{\text{test}} = (1,1)$, determine the predicted label $\hat{y} \in \{+1,-1\}$. Enter the answer as either 1 or -1.

Hint: The prediction for a test point $\mathbf{x}_{\text{test}}$ is given by $\hat{y} = \text{sign}\big((\mathbf{w}^*)^T \mathbf{x}_{\text{test}}\big)$.

[Ans: 1]

10. Consider the two decision trees shown in the figure.



Which of the following statements is correct?

(a) Tree 1 has higher information gain than Tree 2

(b) Tree 2 has higher information gain than Tree 1

(c) Both trees have equal information gain

(d) Cannot be determined without further splits

[Ans: a]

2026 Sep 13 ET

11. Consider a scenario where we use leave-one-out cross-validation (LOOCV) with a Support Vector Machine (SVM) for binary classification. Given a dataset $\mathcal{D} = \{(x_1, y_1), \ldots, (x_N, y_N)\}$, the cross-validation error for a sample (x_n, y_n) is $e_n = 1$ if $\hat{y}_n \neq y_n$ and 0 otherwise, where $\hat{y}_n$ is the prediction for x_n produced by the SVM trained on $\mathcal{D} \setminus \{(x_n, y_n)\}$. The overall cross-validation error is

$$E_{\text{CV}} = \frac{1}{N} \sum_{n=1}^{N} e_n.$$

Let K be the number of support vectors obtained by training on the full dataset $\mathcal{D}$. Which statement is true regarding the relationship between E_{CV} and K?

(a) $E_{\text{CV}} \leq \dfrac{K}{N}$

(b) $E_{\text{CV}} \geq \dfrac{K}{N}$

(c) $E_{\text{CV}} = \dfrac{K}{N}$

(d) E_{CV} is independent of K

[Ans: a]