

# Quiz 5

Started: Nov 12 at 11:31am

## Quiz Instructions

### Question 1

1 pts

Given  $\mathbf{x} \in \mathbb{R}^3$ , what is the corresponding mapping function  $\phi(\mathbf{x})$  for the kernel function  $K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T \phi(\mathbf{y}) = (1 + \mathbf{x}^T \mathbf{y})^2$ .

- (A)  $\phi(\mathbf{x}) = [1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3]$
- (B)  $\phi(\mathbf{x}) = [1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1^2, x_2^2, x_3^2]$
- (C)  $\phi(\mathbf{x}) = [1, 2x_1, 2x_2, 2x_3, 2x_1x_2, 2x_1x_3, 2x_2x_3, x_1^2, x_2^2, x_3^2]$
- (D)  $\phi(\mathbf{x}) = [1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_3, \sqrt{2}x_1x_2, \sqrt{2}x_1x_3, \sqrt{2}x_2x_3, x_1^2, x_2^2, x_3^2]$

☐ A

☐ B

☐ C

☒ D

### Question 2

1 pts

Which of the following statements about kernels are true?

- (A) After applying the mapping function  $\phi(\mathbf{x}) = [\mathbf{x}, \mathbf{x}^2]$ , the data always become linearly separable.
- (B)  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2, K(\mathbf{x}, \mathbf{y}) = (1 + 8\mathbf{x}^T \mathbf{y})^2$  is a valid kernel.
- (C) Let  $K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^T \phi(\mathbf{y})$ . Computing the inner product between  $\phi(\mathbf{x})$  and  $\phi(\mathbf{y})$  using kernel function  $K(\mathbf{x}, \mathbf{y})$  is always slower than directly estimating from  $\phi(\mathbf{x})^T \phi(\mathbf{y})$ .
- (D) A kernel function may allow us to map feature vectors in to another space where the data is linearly separable.

Select all that apply

☐ A☒ B☐ C☒ D**Question 3****1 pts**

Given a set of training data, Hard SVM returns a model with the hyper-plane  $\mathbf{w}^T \mathbf{x} + b = 0$ , where  $\mathbf{w} = [-1, -1]$  and  $b = 1$ . Which data point(s) **cannot** belong to the training dataset?

(A)  $\mathbf{x} = (0, 0.5), y = 1$ .

(B)  $\mathbf{x} = (0, 0), y = 1$ .

(C)  $\mathbf{x} = (2, 1), y = -1$ .

(D)  $\mathbf{x} = (2, 2), y = 1$ .

select all that apply

☒ A☐ B☐ C☒ D**Question 4****1 pts**

As we see in the lecture, Soft SVM identifies a hyper-plane  $\mathbf{w}^T \mathbf{x} + b = 0$  by solving the following optimization problem.

$$\begin{aligned} \min_{\mathbf{w}, b, \xi_i} \quad & \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \cdot \sum_{i=1}^N \xi_i \\ \text{s.t. } \quad & \forall i, y(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0 \end{aligned}$$

Let  $\mathbf{w} = [-1, -1]$  and  $b = 1$  are the solution of the above optimization problem. Given two positive data points  $\mathbf{x}_1 = (0, 0)$  and  $\mathbf{x}_2 = (1, 0)$ ,  $y_1 = y_2 = 1$ , what are their corresponding slack variables  $\xi_1$  and  $\xi_2$ , respectively?

(A)  $\xi_1 = 0, \xi_2 = 0.5$

(B)  $\xi_1 = 0, \xi_2 = 1$

(C)  $\xi_1 = 1, \xi_2 = 0$

(D)  $\xi_1 = 1, \xi_2 = 0.5$

☐ A

☒ B

☐ C

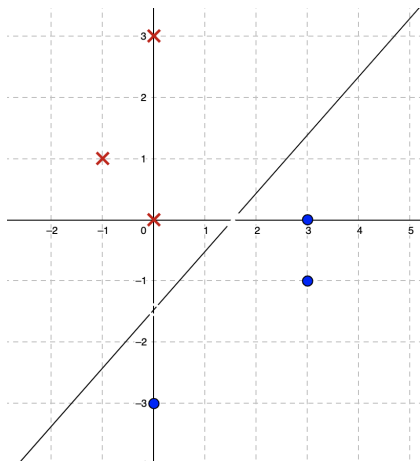
☐ D

## Question 5

1 pts

Assume you have datapoints and labels as  $X = \{[0, 0], [0, 3], [-1, 1], [3, 0], [3, -1], [0, -3]\}$  and  $Y = \{-1, -1, -1, 1, 1, 1\}$ .

We train a hard SVM classifier as shown in below.



What is  $w_1$ ? (round to 2 decimal places)

$w_1 =$  \_\_\_\_\_

1

### Question 6

1 pts

Follow the previous question. What is  $w_2$ ? (round to 2 decimal places)

$w_2 =$  \_\_\_\_\_

-1

### Question 7

1 pts

What is  $b$ ? (round to 2 decimal places)

$b =$  \_\_\_\_\_

-1.5

No new data to save. Last checked at 12:18pm

Submit Quiz