Homework-5

Question 1

The "dropouts" technique is controlled by a probability parameter p. In this question p is the probability the weight is **dropped**.

1. What is the meaning of selecting $p = 0$?

2. What is the meaning of selecting $p = 1$?

3. Suppose $p = 0.2$, and the dropouts technique is applied to a layer with 100 weights. At the end of the training approximately how many weights do you expect to have 0 values?

4. Suppose $p = 0.2$. If at the end of the training a certain weight value is 3.0, what should be the value of the same weight during testing?

Question 2

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).) Using a linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

(The linear kernel is: $K(u, v) = u^T v$.)

Question 3

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Show that with linear kernel the (dual) optimization problem that needs to be solved in terms of the α_i is:

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2 + \alpha_4)^2 + (\alpha_3 + \alpha_4)^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad \alpha_1 = \alpha_2 + \alpha_3 + \alpha_4 \end{aligned}$$

(The linear kernel is: $K(u, v) = u^T v$.)

B

The solution to the above optimization problem is: $\alpha_1 = 4,\ \alpha_2 = 2,\ \alpha_3 = 2,\ \alpha_4 = 0$.

a. What are the indexes of the support vectors? Circle them below.

Answer: 1 2 3 4

b. What computation needs to be carried out to determine the classification of the point $x_5 = (-1, 0)$ by this SVM.

c. What computation needs to be carried out to determine the classification of the point $x_5 = (-1, 1)$ by this SVM.

d. What computation needs to be carried out to determine the classification of the point $x_5 = (1, 1)$ by this SVM.

Question 4

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	-1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).) Using a linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values? (The linear kernel is: $K(u, v) = u^T v$.)

Question 5

An SVM is trained with the following data:

i	1	2	3	4
x_i	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
y_i	-1	1	1	-1

Let $\alpha_1, \ldots, \alpha_4$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Show that with linear kernel the (dual) optimization problem that needs to be solved in terms of the α_i is:

$$\begin{aligned} \text{Maximize:} \quad & \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2 - \alpha_4)^2 + (\alpha_3 - \alpha_4)^2\right) \\ \text{subject to:} \quad & \alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \geq 0,\ \alpha_4 \geq 0, \quad \alpha_1 = \alpha_2 + \alpha_3 - \alpha_4 \end{aligned}$$

(The linear kernel is: $K(u, v) = u^T v$.)

B

Observe that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = k$ satisfies the constraints for any $k \geq 0$, and that the function to be maximized in terms of k is $4k$. Based on these observations, what is the solution to the above (dual) optimization problem?

Question 6

An SVM is trained with the following data:

i	1	2	3	4	5
x_i	$(0,0)$	$(1,0)$	$(2,0)$	$(3,0)$	$(0,1)$
y_i	-1	1	1	1	-1

Let $\alpha_1, \ldots, \alpha_5$ be the Lagrangian multipliers associated with this data. (α_i is associated with (x_i, y_i).)

A

Using the linear kernel, what (dual) optimization problem needs to be solved in terms of the α_i in order to determine their values?

(The linear kernel is: $K(u, v) = u^T v$.)

Answer

B

The solution to the optimization problem is:

$$\alpha_1 = 2, \quad \alpha_2 = 2, \quad \alpha_3 = 0, \quad \alpha_4 = 0, \quad \alpha_5 = 0.$$

a. Show the computation that needs to be carried out to determine the classification of the point $x = (1, 1)$ by this SVM.

b. Show the computation that needs to be carried out to determine the classification of the point $x = (0, -2)$ by this SVM.

C

To obtain solution with soft margins using the l_1 norm (this is the one described in class), the term $C\sigma_i\zeta_i$ is added to the primal function. Using the value of $C = 10$, repeat parts a, b of B.

a. Show the computation that needs to be carried out to determine the classification of the point $x = (1, 1)$ by this soft margins SVM.

b. Show the computation that needs to be carried out to determine the classification of the point $x = (0, -2)$ by this soft margins SVM.