HOMEWORK - 5

9XS230001

Question 1 :-

1.) Selecting $p=0$ means no weights are dropped

All weights are retained during training.

2.) Selecting $p=1$ means all weights are dropped.

essentially, the layer is deactivated during training.

3.) $p=0.2$ and the dropouts technique is applied to a layer with 100 weights. At the end of the training. Expected number of weights with zero values $= 0.2 \times 100 = 20$ weights

4.) During testing, the weight is multiplied by $1-p$, so the value would be $3.0 \times 0.8 = 2.4$.

Question 2 :-

$x_1 = (0,0)$ $x_2 = (0,1)$, $x_3 = (1,0)$, $x_4 = (1,1)$

$y_1 = -1$ $y_2 = 1$ $y_3 = 1$ $y_4 = 1$

Linear kernel $K(u,v) = u^T v$

$$\text{Maximize} = \sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{4} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$\sum_{i=1}^{4} \alpha_i y_i = 0$$

$$\text{Maximize} \sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{4} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

Maximize $= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left((\alpha_2+\alpha_4)^2 + (\alpha_3+\alpha_4)^2\right)$

This leads to an optimizing problem involving the α_i's for the given training data.

Question 3:-

Part A $x_1 = (0,0)$ $x_2 = (0,1)$ $x_3 = (1,0)$ $x_4 = (1,1)$

$x_1 = (0,0)$, $x_2 = (0,1)$ $x_3 = (1,0)$ $x_4 = (1,1)$

dot product of linear kernel are:-

$x_1^T x_1 = 0$ $x_1^T x_2 = 0$ $x_1^T x_3 = 0$

0, 1, 0, 1, 1, 1, 2

Maximize $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2}\left[(\alpha_2+\alpha_4)^2 + (\alpha_3+\alpha_4)^2\right]$

$\alpha_1 \geq 0$, $\alpha_2 \geq 0$ $\alpha_3 \geq 0$ $\alpha_4 \geq 0$ $\alpha_1 = \alpha_2 + \alpha_3 + \alpha_4$

Part B:-

$\alpha_1 = 4$ $\alpha_2 = 2$ $\alpha_4 = 0$

a) Support vectors are the data points when $\alpha_i > 0$

Therefore the support vectors are the point corresponding to $\alpha_1, \alpha_2, \alpha_3$ i.e x_1 x_2 x_3 indent (1, 2, 3)

b) $x_4 = (-1, 0)$

$$f(x_5) = \sum_{i=1}^{4} \alpha_i y_i K(x_i, x_5) + b$$

$x_5 = (-1, 0)$

$K(x_1, x_5) = 0$

$K(x_2, x_5) = 0$

$K(x_3, x_5) = -1$

$K(x_4, x_5) = -1$

$f(x_5) = -2 + b$

$b = 0$

$f(x_5) = -2$ $\quad f(x_5) < 0$

the result is $\boxed{-1}$.

c) $f(x_5) = -2 + b$

$f(x_5) = -2$

$b = 0$

result is $\boxed{-1}$

d) $x_5 = (1, 1)$

$x_2 = (1, 1)$

$f(x_5) = -4 + b$

$b = 0$ $\quad f(x_5) = -4$

$\boxed{\text{result} = -1}$

Q4 SVM optimization with Linear Kernel

$x_1 = (0,0)$ $x_2 = (0,1)$ $x_3 = (1,0)$ $x_4 = (1,1)$

$y_1 = -1$ $y_2 = 1$ $y_3 = 1$ $y_4 = -1$

2.1 dual formulation of the SVM is -

Maximize $\sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{4} \alpha_i \alpha_j y_i y_j (x_i, x_j)$

Maximize $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - \frac{1}{2} \left[(\alpha_2 - \alpha_4)^2 + (\alpha_3 - \alpha_4)^2 \right]$

$\alpha_1 \geq 0$ $\alpha_2 \geq 0$ $\alpha_3 \geq 0$ $\alpha_4 \geq 0$ $\cancel{\alpha_1 = \alpha_2 + \alpha_3}$

$\alpha_1 = \alpha_2 + \alpha_3 - \alpha_4$

Part B

$\alpha_1 = 4$ $\alpha_2 = 2$ $\alpha_3 = 2$ $\alpha_4 = 0$

$f(x_5) = \sum_{i=1}^{4} \alpha_i y_i K(x_i, x_5) + b$

$x_5 = (-1, 0)$:

If $f(x_5) > 0$ classify as 1 else -1

2) $x_5 = (-1, 1)$

Repeat same Computation $K(x_1, x_5), K(x_2, x_5)$
$K(x_3, x_5)$ $K(x_4, x_5)$

3) $x_5 = (1, 1)$

Similarly repeat Computation and decision function to classify x_5

Q5 1) Dual formulation

$$K(x_i, x_j) = x_i^T x_j$$

$$\text{Maximum} \sum_{i=1}^{4} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{4} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$K(x_i, x_j) = x_i^T x_j$$

$$\text{Maximize } \alpha_1 + \alpha_3 + \alpha_3 + \alpha_4 - \frac{1}{2}\left[(\alpha_2 - \alpha_4)^2 + (\alpha_3 - \alpha_4)^2\right]$$

$\alpha_1 \geq 0$ $\alpha_2 \geq 0$ $\alpha_3 \geq 0$ $\alpha_4 \geq 0$ $\alpha_1 = \alpha_2 + \alpha_3 - \alpha_4$

Solution $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4$

$K = 0$ $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$

Part B.

$$\text{Maximize } 4K - \frac{1}{2}\left[(K-K)^2 + (K-K)^2\right)$$

$$4K = 0 = 4K$$

$\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 = 0$ $K = 0$ $\alpha_i = 0$

This means that there are no non zero Lagrange multipliers, indicating that the SVM model has no support vectors for this particular case. (i.e. the optimization has no contributing vectors)

Question 6

Part A :-

$x_1 = (0,0)$ $x_2 = (1,0)$ $x_3 = (2,0)$ $x_4 = (3,0)$ $x_5 = (0,1)$

labels $y_1 = -1$ $y_2 = 1$ $y_3 = 1$ $y_4 = 1$ $y_5 = -1$

Maximize $\sum_{i=1}^{5} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{5} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$

α_i are the lagrangian multipliers

Maximize $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 - \frac{1}{2} \left(\sum_{i,j=1}^{5} \alpha_i \alpha_j y_i y_j x_i x_j \right)$

Part B :

$\alpha_1 = 2$ $\alpha_4 = 0$

$\alpha_2 = 2$ $\alpha_5 = 0$

$\alpha_3 = 0$

a) $x = (1,1)$

$x = (1,1)$

$f(x) = \sum_{i=1}^{5} \alpha_i y_i K(x_i, x) + b$

$K(x_i, x)$:

$K(x_1, x) = 1$ $\quad$ $K(x_3, x) = 2$

$K(x_2, x) = 1$ $\quad$ $K(x_4, x) = 3$

$K(x_5, x) = 1$

$f(x) = 2(-1) \cdot 1 + 2(1) \cdot 1 + 0.2 + 0.3 + 0.1 + b$

$f(x) = -2 + \cancel{2} + b = b$

if $b > 0$ $\quad$ 1 otherwise -1

b) $x = (0, -2)$

$K(x_i, x)$ $\quad x = (0, -2)$

$K(x_1, x) = 0$ $\quad$ $K(x_3, x) = 0$

$K(x_2, x) = 0$ $\quad$ $K(x_4, x) = 0$

$K(x_5, x) = 1$

$f(x) = \cancel{0}\, 2(-1) \cdot 0 + 2(1) \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 0.1 + b$

$f(x) = 0 + b$

if $b > 0$ classify 1 otherwise -1

Part C: Classification $x = (1, 1)$

Cl ~~follow~~ it has same Computation as in Part B

$x = (1, 1)$

$K(x_1, x) = 1$

$K(x_2, x) = 1$

$K(x_3, x) = 2$

$K(x_4, x) = 3$

$K(x_5, x) = 1$

$f(x) = 2(-1) \cdot 1 + 2(1) \cdot 1 + 0 \cdot 2 + 0 \cdot 3 + 0 \cdot 1 + b = -2 + 2 + b = b$

b) $x = (0, -2)$

$K(x_1, x) = 0$

$K(x_2, x) = 0$

$K(x_3, x) = 0$

$K(x_4, x) = 0$

$K(x_5, x) = 1$

$f(x) = 2(-1) \cdot 0 + 2(1) \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 0 \cdot 1 + b = 0 + b$

The regularisation term will influence b but doesn't change the [illegible] of this computation.