Cpts570-hw2

Mengxiao-11651502

October 13, 2019

1 Analytical Part

1.1 Problem 1

1. The w is:

$$\frac{1}{\frac{1-(-1)}{C_+ - C_-}} = \frac{1}{\frac{2}{C_+ - C_-}} = \frac{\frac{1}{n_+}\sum_{i:y_i=+1} x_i - \frac{1}{n_-}\sum_{i:y_i=-1} x_i}{2}$$

Since $wx + b$ is the decision boundary, the b is:

$$w(\frac{C_+ + C_-}{2}) + b = 0 \Rightarrow b = \frac{(\frac{1}{n_-}\sum_{i:y=-1} x_i)^2 - (\frac{1}{n_+}\sum_{i:y=+1} x_i)^2}{4}$$

2. Since $w = \sum_{i=1}^{n_+ + n_-} \alpha_i \cdot y_i \cdot x_i$, we can know:

$$w = \sum_{i:y=+1}^{n_+} \alpha_i \cdot x_i - \sum_{i:y=-1}^{n_-} \alpha_i \cdot x_i$$

$$= \frac{\frac{1}{n_+}\sum_{i:y_i=+1} x_i - \frac{1}{n_-}\sum_{i:y_i=-1} x_i}{2}$$

$$\Rightarrow for \ y_i = +1, \alpha_i = \frac{n_-}{2n_+ n_-}$$

$$for \ y_i = -1, \alpha_i = \frac{n_+}{2n_+ n_-}$$

1.2 Problem 2

1. Use the Tylar's Formula on the kernel function $K(x_i, x_j)$, we can know

$$K(x_i, x_j) = K(x_i, x_j) - (x - x_i)\exp^{-\frac{1}{2}(x_i - x_j)} +(x - x_j)\exp^{\frac{1}{2}(x_i - x_j)}$$

$$+\frac{1}{2!}(x - x_i)^2 \frac{1}{2}\exp^{-\frac{1}{2}(x_i - x_j)} -\frac{1}{2!}(x - x_j)^2 \frac{1}{2}\exp^{-\frac{1}{2}(x_i - x_j)}$$

$$-\frac{2}{2!}(x - x_i)(x - x_j)\frac{1}{2}\exp^{-\frac{1}{2}(x_i - x_j)}$$

$$+ \cdots$$

Finally, we can find that this function can be expanded to infinite dimensions.

2.

$$||\Phi(x_i) - \phi(x_j)||^2$$
$$= \Phi(x)^2 + \Phi(y)^2 - 2\Phi(x)\Phi(y)$$
$$= K(x,x)^2 + K(y,y)^2 - 2K(x,y)$$
$$= 1 + 1 - 2 * \exp^{-\frac{1}{2}||x_i - x_j||^2} \leq 2$$

1.3   Problem 3

Since we want to prove $f(x_{far}; \alpha, b) \approx b$,

$$w \cdot \Phi(x_{far}) + b \approx b$$

$$\Rightarrow w \cdot \Phi(x_{far}) \approx 0$$

$$\Rightarrow \sum_{i \in SV} y_i \alpha_i K(x_i, x_{far}) \approx 0$$

Since $K(x_i, x_{far}) = \exp^{-\frac{1}{2}||x_i - x_{far}||^2}$, we know

$$x_i - x_{far} \approx \infty$$

$$\exp^{-\frac{1}{2}||x_i - x_{far}||^2} \approx 0$$

$$K(x_i, x_{far} \approx 0)$$

Finally, we approve it.

1.4   Problem 4

It is not a valid kernel.

We want to show this function is not a kernel, so we need to show a counter-example for symmetry or positivity conditions. It is really a symmetry function, but I will prove that it is not positivity conditions.

If we want the kernel function be positivity, we need to show that for all m data points and all $m \times 1$ vectors, $t^T K t \geq 0$. But we know that $K(x_i, x_j) = - < x_i, x_j >$. So, whatever t we have, we will get $t^T K t \leq 0$. It is not positivity.

1.5 Problem 5

I will just separate the parameter $C \sum_{i=1}^{N}$ in the fomula:

$$\min \frac{1}{2}||w||^2 + C \sum_{i=1}^{N} \xi_i$$

to $C_+ \sum_{i:y=+1}^{N+} \xi_i$ and $C_- \sum_{i:y=-1}^{N-} \xi_i$, so that we can use the extra information.

1.6 Problem 6

a. At first we use the formula in Q5.

$$\min \frac{1}{2}||w||^2 + C_+ \sum_{i:y=+1}^{N+} xi_i + C_- \sum_{i:y=-1}^{N-} \xi_i$$

Then, we change the formulation:
repeat the following:

$$Pick \ K_+ \ and \ K_- \ from \ the \ data$$
$$contstrain \ y_i^{K_i}(w \cdot x_i^{k_i} + b) \leq 1 - \xi \ then$$
$$compute \ \min \frac{1}{2}||w||^2 + C_+ \sum_{i:y=+1}^{N+} xi_i + C_- \sum_{i:y=-1}^{N-} \xi_i$$

b. Just separate the whole parameter to some groups and at first try with only separate to two groups, then try to separate to more groups like four groups if the classifier do not have good appearence on two groups.

c. If the classifier can have good effectiveness we can accept and correct rate we need.

1.7 Problem 7

a. We can sort the SV by α_i(in Dual Form) from high to low, then choose the highest B support vectors to use, since the SV with highest score would effect the weight more.

b. We still need to sort the SV, and choose the first B SV. But now, we need to choose the most effective numbers of SV that compute the score of these B SV and throw those SV without high enough effect (like 5% increase of the whole score).