NCTU Introduction to Machine Learning, Homework 4

Deadline: Nov. 29, 23:59

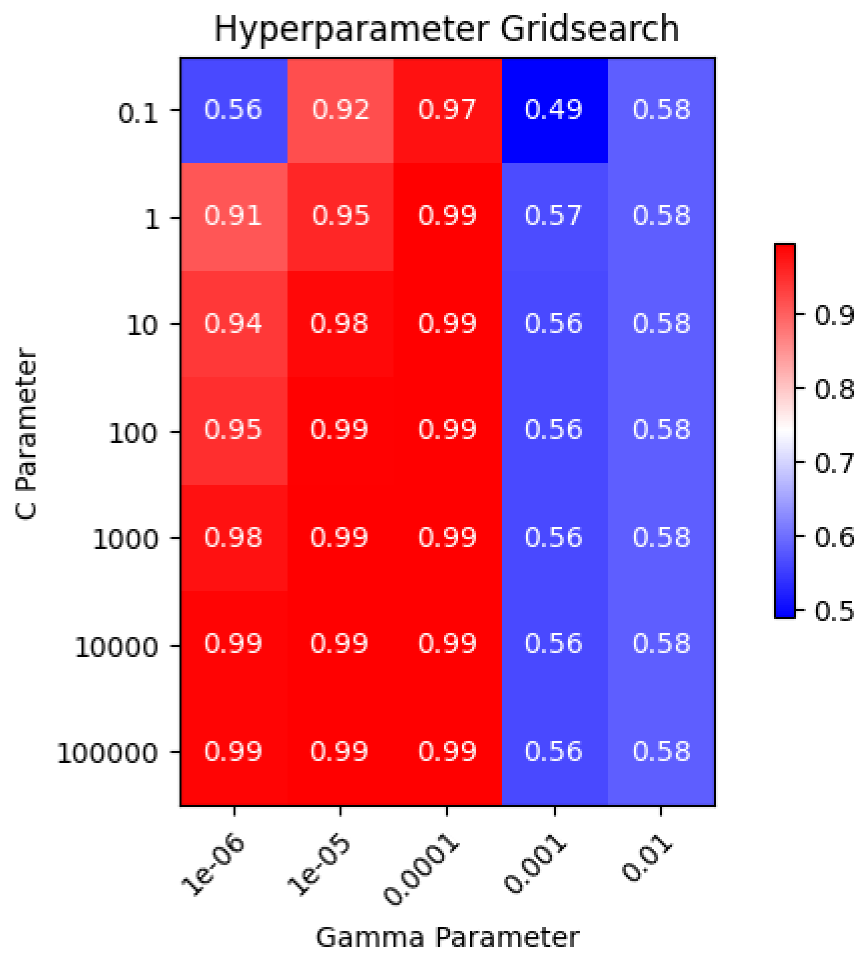
Part. 1, Coding (50%):

1. Print the best hyperparameters you found.

```
0.9932857142857143
[10, 0.0001]
```

2. (10%) Plot the grid search results of your SVM. The x and y represent "gamma" and "C" hyperparameters, respectively. And the color represents the average score of validation folds.

   Note: This image is for reference, not the answer
   Note: matplotlib is allowed to us



Part. 2, Questions (50%):

(10%) Show that the kernel matrix $K = \left[k\left(x_n, x_m\right)\right]_{nm}$ should be positive semidefinite is the necessary and sufficient condition for $k(x, x')$ to be a valid kernel.

① $K = [k(x_n, x_m)]_{nm}$

K is symmetric, Thus we have $K = V\Lambda V^T$

where V is an orthonormal matrix V_t and the diagonal matrix Λ contains the eigenvalues λ_t of K

If K is positive semidefinite, all eigenvalues are non-negative

Consider the feature map: $\phi: x_i \rightarrow (\sqrt{\lambda_t}\, V_{ti})_{t=1}^{n} \in R^n$

We find that $\phi(x_i)^T \phi(x_j) = \sum_{t=1}^{n} \lambda_i V_{ti} V_{tj} = (V\Lambda V^T)_{ij} = K_{ij} = k(x_i, x_j)$

(10%) Given a valid kernel $k_1(x, x')$, explain that $k\left(x,\ x'\right) =\ exp(k_1\left(x, x'\right))$ is also a valid kernel. Your answer may mention some terms like ____ series or ____ expansion.

② $k(x, x') = exp(k_1(x, x'))$

since we have $exp = \lim_{i \to \infty} (1 + \cdots + x^i/i!)$

$\Rightarrow k(x, x') = \lim_{i \to \infty} k_i(x, x') = \sum_{i=0}^{\infty} \frac{x^i (x')^i}{i!} = \sum_{i=0}^{\infty} \frac{x^i}{\sqrt{i!}} \cdot \frac{(x')^i}{\sqrt{i!}}$

$\Rightarrow \phi\ ) = \frac{x^i}{\sqrt{i!}}$, $\phi(x') = \frac{(x')^i}{\sqrt{i!}}$

(20%) Given a valid kernel $k_1(x, x')$, prove that the following proposed functions are or are not valid kernels. If one is not a valid kernel, give an example of $k(x, x')$ that the corresponding K is not positive semidefinite and show its eigenvalues.

a. $k\left(x,\ x'\right) =\ k_1\left(x,\ x'\right) +\ 1$

b. $k\left(x,\ x'\right) =\ k_1\left(x, x'\right) -\ 1$

c. $k\left(x,\ x'\right) =\ k_1\left(x, x'\right)^2 +\ exp\left(\|x\|^2\right) *\ exp(\|x'\|^2)$

d. $k(x, x') = k_1(x, x')^2 + exp(k_1(x, x')) - 1$

3

a. $k(x,x') = k_1(x,x') + 1$

$k_2(x,x') = 1 \Rightarrow k = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{n \times m}$ => eigenvalue = n or 0 => positive semidefinite

=> $k_2(x,x')$ is a valid kernel

By $k(\mathbf{x}, \mathbf{x}') = k_1(\mathbf{x}, \mathbf{x}') + k_2(\mathbf{x}, \mathbf{x}')$ (6.17)

Proof $k(x,x') = k_1(x,x') + 1$ is a valid kernel

b. $k(x,x') = k_1(x,x') - 1$

if $k_1(x,x') = 0 \Rightarrow k(x,x') = -1$

$$\Rightarrow K = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}_{n \times n} \Rightarrow \text{eigenvalue} < 0 \text{, not positive semidefinite}$$

$\Rightarrow k(x,x')$ is not a valid kernel

c. $k(x,x') = k_1(x,x')^2 + \exp(\|x\|^2) \times \exp(\|x'\|^2)$

By $k(\mathbf{x}, \mathbf{x}') = f(\mathbf{x}) k_1(\mathbf{x}, \mathbf{x}') f(\mathbf{x}')$ (6.14)

let $f(x) = \exp(\|x\|^2)$, $k_1(x,x') = 1$, $f(x') = \exp(\|x'\|^2)$

prove $k_2(x,x') = \exp(\|x\|^2) \times \exp(\|x'\|^2)$ is a valid kernel

By $k(\mathbf{x}, \mathbf{x}') = k_1(\mathbf{x}, \mathbf{x}') + k_2(\mathbf{x}, \mathbf{x}')$ (6.17)

prove $k(x,x') = k_1(x,x')^2 + \exp(\|x\|^2) \times \exp(\|x'\|^2)$ is a valid kernel

d. $k(x,x') = k_1(x,x')^2 + \exp(k_1(x,x')) - 1$

$$k(x,x') = k_1(x,x')^2 + \lim_{i \to \infty} \sum_{i=0}^{\infty} \frac{x'^i (x)^i}{i!} - 1$$

$$= k_1(x,x')^2 + \sum_{n=1}^{\infty} \frac{k(x,x')^n}{n!}$$

By $k(\mathbf{x}, \mathbf{x}') = k_1(\mathbf{x}, \mathbf{x}') + k_2(\mathbf{x}, \mathbf{x}')$ (6.17)

$k(\mathbf{x}, \mathbf{x}') = k_1(\mathbf{x}, \mathbf{x}') k_2(\mathbf{x}, \mathbf{x}')$ (6.18)

prove $\exp(k_1(x,x')) - 1$ is a valid kernel, $k_1(x,x')^2$ is a valid kernel

thus $k(x,x') = k_1(x,x')^2 + \exp(k_1(x,x')) - 1$ is a valid kernel

(10%) Consider the optimization problem

$$minimize\ (x - 2)^2$$
$$subject\ to\ (x + 3)(x - 1) \leq 3$$

State the dual problem.

④ minimize $(x-2)^2$ subject to $(x+3)(x-1) \le 3$

$f_1(x) = (x-2)^2$

$f_2(x) = (x+3)(x-1) - 3$

$\Rightarrow$ minimize $f_1(x)$ subject to $f_2(x) \le 0$

$L(x, \lambda) = f_1 + \lambda f_2 = (x-2)^2 + \lambda(x+3)(x-1) - 3\lambda$

$\frac{\partial L}{\partial x} = 2(x-2) + \lambda(x-1) + \lambda(x+3)$

$= \lambda(2x+2) + 2(x-2)$

$= 2\lambda x + 2\lambda + 2x - 4$

$= (2\lambda + 2)x + 2\lambda - 4 = 0 \Rightarrow x = \frac{4-2\lambda}{2\lambda+2} = \frac{2-\lambda}{\lambda+1}$

$\phi(x) = \left(\frac{2-\lambda}{\lambda+1}\right)^2 + \lambda\left(\frac{2-\lambda}{\lambda+1} + 3\right)\left(\frac{2-\lambda}{\lambda+1} - 1\right) - 3\lambda$

dual problem

maximize $\left(\frac{2-\lambda}{\lambda+1}\right)^2 + \lambda\left(\frac{2-\lambda}{\lambda+1} + 3\right)\left(\frac{2-\lambda}{\lambda+1} - 1\right) - 3\lambda$

subject to $\lambda \le 0$