CAP 5610
Assignment #3 Solution

March 17, 2025

Arman Sayan

---

1 Kernel Computation Cost [30 points]

1. [10 points] Consider we have a two-dimensional input space such that the input vector is $x = (x_1, x_2)^T$. Define the feature mapping $\phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T$. What is the corresponding kernel function, i.e., $K(x, z)$? Do not leave $\phi(x)$ in your final answer.

Ans:
To find the kernel function $K(x, z)$, we use the definition of the kernel function as follows:

$$K(x, z) = \phi(x)^T \phi(z)$$

where $\phi(x)$ represents feature mapping.
Given that $\phi(x) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T$, for another input vector $z = (z_1, z_2)^T$, the feature mapping is $\phi(z) = (z_1^2, \sqrt{2}z_1z_2, z_2^2)^T$.
Now, we can compute the dot product of the feature mappings as follows:

$$\begin{aligned} K(x, z) &= \phi(x)^T \phi(z) \\ &= \begin{bmatrix} x_1^2 & \sqrt{2}x_1x_2 & x_2^2 \end{bmatrix} \begin{bmatrix} z_1^2 \\ \sqrt{2}z_1z_2 \\ z_2^2 \end{bmatrix} \\ &= x_1^2z_1^2 + 2x_1x_2z_1z_2 + x_2^2z_2^2 \end{aligned}$$

With a deduction from the above calculations, we can observe the final expression as

$$K(x, z) = (x_1z_1 + x_2z_2)^2$$

Since $x_1z_1 + x_2z_2$ is the dot product of the input vectors x and z, namely x^Tz, the corresponding kernel function is

$$\boxed{\boldsymbol{K(x, z) = (x^Tz)^2}}$$

2. [20 points] Suppose we want to compute the value of the kernel function $K(x, z)$ from the previous question on two vectors $x, z \in \mathbb{R}^2$. How many additions and multiplications are needed if you
   i. [10 points] Map the input vector to the feature space and then perform the dot product on the mapped features?

   Ans:

   ii. [10 points] Compute through the kernel function you derived in question 1?

   Ans:

2 Activation Functions and Loss Functions [30 points]

1. [20 points] For this assignment, you are encouraged to consult Dr. GOOGLE and Dr. ChatGPT. But please explain things in your own language while you write the answer. For each of the following activation functions, briefly describe the type of non-linearity (if any) it introduces and discuss their pros and cons.
   (a) Linear Activation Function
   (b) Sigmoid Activation Function
   (c) Tanh Activation Function
   (d) ReLU (Rectified Linear Unit) Activation Function

Ans:

2. [10 points] For each of the following loss functions, briefly describe their mathematical formula and discuss for which type of learning task (classification, regression, etc.) they are most appropriate.
   (a) Mean Squared Error Loss
   (b) Binary Cross Entropy Loss
   (c) Hinge Loss
   (d) Softmax Cross Entropy Loss

Ans:

3 Linear SVM Implementation [40 points]

Solution for Q3:

Ans:

Please check the source code and outputs included in the appendix named as

CAP_5610_Assignment_3_Solution_Arman_Sayan.ipynb

for the solution.