COMP 5630 Fall 2022 Assignment 6

Will Humphlett (wah0028)

October 25, 2022

1. Kernel Computation Cost

   (a) Define the feature mapping

$$\begin{aligned} K(x,z) &= \phi x \cdot \phi z \\ &= (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T \cdot (z_1^2, \sqrt{2}z_1x_2, z_2^2)^T \\ &= x_1^2z_1^2 + \sqrt{2}x_1x_2 \cdot \sqrt{2}z_1z_2 + x_2^2z_2^2 \\ &= x_1^2z_1^2 + 2x_1x_2z_1z_2 + x_2^2z_2^2 \\ &= (x_1z_1 + x_2z_2)^2 \end{aligned} \tag{1}$$

   If $x = (x_1, x_2)^T$ and $z = (z_1, z_2)^T$, then

$$K(x,z) = (x_1z_1 + x_2z_2)^2 = (x \cdot z)^2 \tag{2}$$

   (b) How many additions and subtractions

$$x_1^2z_1^2 + 2x_1x_2z_1z_2 + x_2^2z_2^2 \tag{3}$$

   2 additions, 10 multiplications

$$(x_1z_1 + x_2z_2)^2 \tag{4}$$

   1 addition, 3 multiplications

2. Kernel Functions

   (a) Prove this is a legal kernel
   If $K(x,x')$ is valid, then K is symmetric and $\sum_{i,j}^n a_ia_jK(x_i,x_j) >= 0$

      i. Symmetric

$$K(x,x') = (\phi(x), \phi(x')) = (\phi(x'), \phi(x)) = K(x', x) \tag{5}$$

      ii. $\sum_{i,j}^n a_ia_jK(x_i,x_j) >= 0$

$$\begin{aligned} K(x,x') &= (\textstyle\sum_i^n x_ix_i')(\sum_j^n x_jx_j') \\ &= \textstyle\sum_i^n \sum_j^n x_ix_jx_i'x_j' \\ &= \textstyle\sum_{i,j}^n (x_ix_j)(x_i'x_j') \\ &= \phi(x) \cdot \phi(x') \end{aligned} \tag{6}$$

      if $\phi : x \longrightarrow R^2$, then $\phi(x) = (x_1x_1, x_1x_2, x_2x_1, x_2x_2)$

(b) Justify linear separability

The kernel will separate all data to either 1 or 0, making the data it maps linearly separable.

(c) Why is this a bad idea?

Depending on the data, linearly separating the data can lead to overfitting.

3. Implementing SVM (code at: github.com/wumphlett)

```
Optimal Hyperparameters: Learning Rate: 0.00015, # of Epochs: 10000, Reg Const: 100
Training Accuracy: %94.583504
Validation Accuracy: %93.969231
Test Accuracy: %94.215385
```

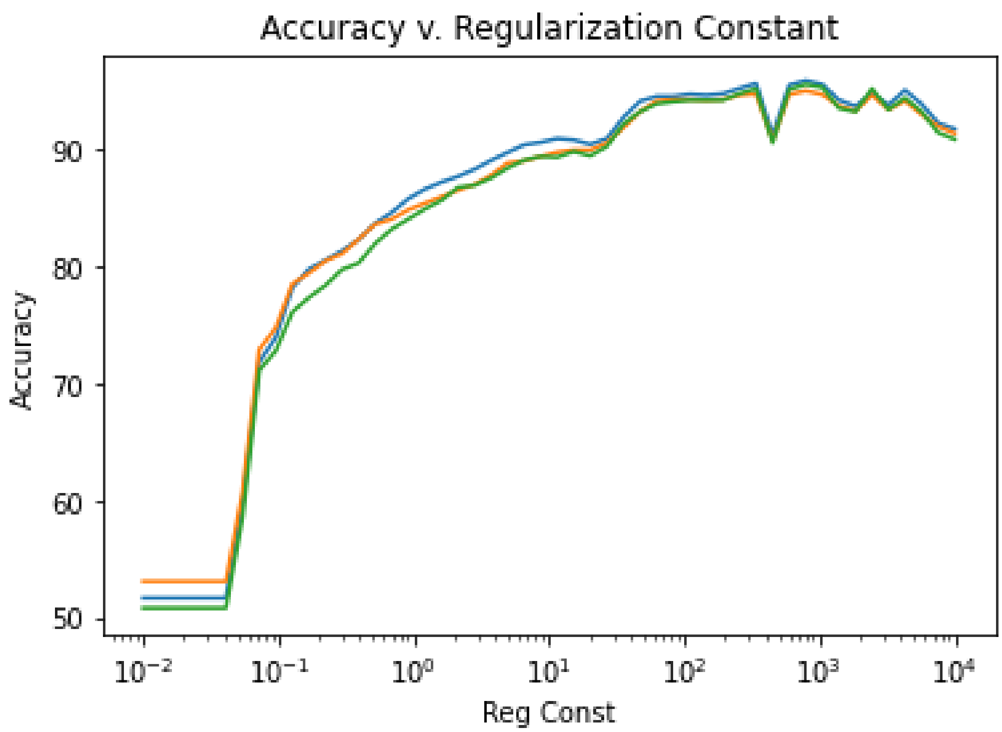

Figure 1: Reg Constant (lr fixed at .00015, epochs 10,000)

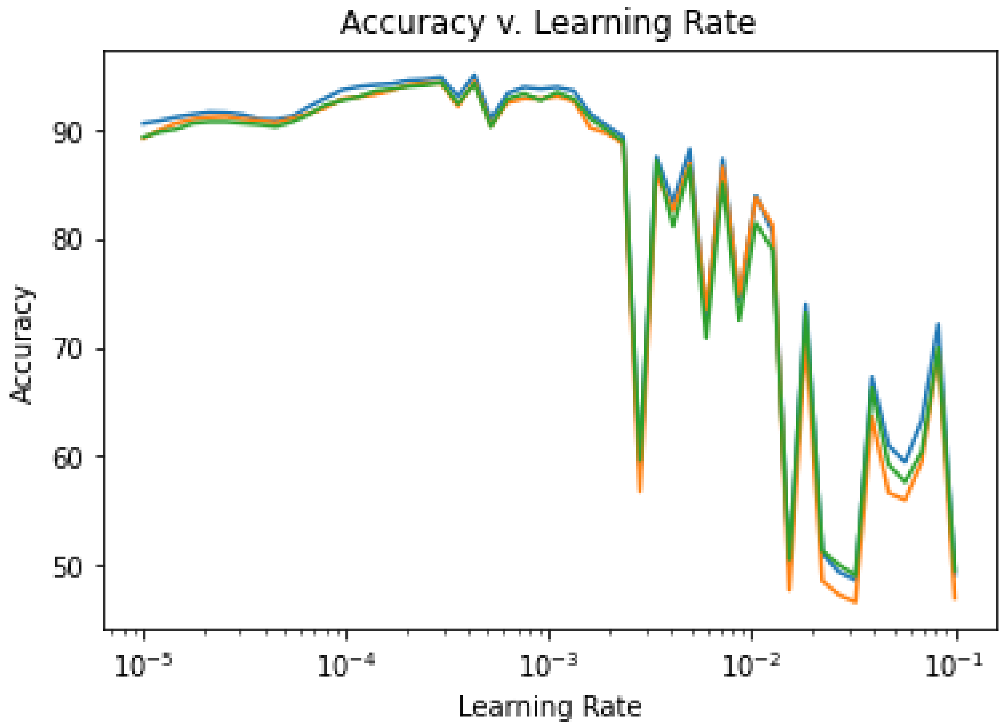

Figure 2: Learning Rate (reg const fixed at 100, epochs fixed at 10,000)