E0 270 : Machine Learning
Tutorial

Due: April 02, 2022

1. *SVM.* Suppose we are given a dataset $\mathcal{D} := \{x_i, y_i\}$. What can you say about finding the SVM if the four support vectors X_1, X_2, X_3, X_4 are given to you a priori? Can you show that solving for the SVM weights with the four support vectors is equivalent to finding the weights using the whole dataset?

2. *SVM.*
   (a) Prove or disprove: the function
   $$k(x, y) = Z(x - y)^\top Q Z(x - y)$$
   is a kernel, where Q is a symmetric positive definite matrix, and $Z(\cdot)$ is a vector of monomials upto degree d.
   (b) Let $\mathcal{X} = [0, 2\pi]^2$ (thus every data point $\mathcal{X} \ni X = (\theta, \phi)$, $\theta, \phi \in [0, \pi/2]$). Prove or disprove: the function
   $$k(X, Y) = \cos(\theta_X - \theta_Y)\cos(\phi_X - \phi_Y)$$
   is a valid kernel. Is this kernel valid if we extend $\mathcal{X}$ to $[0, 2\pi]$?
   (c) Suppose we have a kernel that always yields a Gram matrix which has exactly one zero eigenvalue, with the corresponding eigenvector being elementwise positive. Would such a kernel be useful in the hard-margin SVM setting? Why or why not? What about in the soft-margin setting? (hint: look at this problem from the dual)

3. *Convolutional Neural Network.* Consider a convolutional neural network which has L layers. We will consider the l-th and $(l + 1)$th layers. Assume the input dimensions are $n \times n \times M$, and the output dimensions are $P \times P \times O$.
   (a) What is the dimension of the weight tensor?
   (b) Let $Y^{l+1} = [\text{vec}(Y_1), \cdots \text{vec}(Y_O)]$ be the output of the layer, where $\text{vec}(\cdot)$ maps a matrix of size $m \times n$ to a vector of length mn. Let $\hat{W} = \text{vec}(W)$ be the vectorized weights. Can you write the output Y as a matrix, where each column is a dot product of $\hat{W}$ and some function of X?

4. *PCA.* Implement pca for digit classification. In this problem you will be using MNIST handwritten digits dataset. The task is to recognize the digit and identify it as one of 10 classes from 0-9.

   The MNIST database of handwritten digits, have size-normalized and centered fixed-size images. The training and test data sets have 785 columns. The first column consists of the class labels 0-9. The rest of the columns contain the pixel-values of the associated image.

   In this task, you need to do the following:

   - Implement pca algorithm in python using your own code
   - Implement Random forest classifier. The principal components obtained from pca algorithm has to be given as input to Random forest classifier

- Plot a graph of number of principal components (k) versus accuracy of Random forest classifier, where k ranges from $2 - 50$ and $Accuracy = \frac{N_c}{N}$, where N_c is the number of corrected classified data samples, and N is the total number of samples of the test set.