Homework 5
Due: May 22, 2020, 11:59PM PT

Student Name: *Instructor Name: John Lipor*

Problem 1   Feature Representations (10 pts)

In our study of model selection, we discussed several options of what to do when learning fails. Perhaps the most opaque of these was to "change the feature representation of the data." In our upcoming study of kernels, we will see how to implicitly work with different feature representations. In this problem, we will explicitly work with a new feature representation known as *random Fourier features*, described in the seminal work here.

   You will work with the MNIST dataset using your multiclass ridge regression implementation from Homework 1. Transform each vector $x_i \in \mathbb{R}^d$ via the equation

$$\bar{x}_i = \cos\left(Gx_i + b\right),$$

where $G \in \mathbb{R}^{p \times d}$ is a Gaussian random matrix whose elements are drawn iid from the normal distribution with mean zero and variance $\sigma^2 = 0.1$, and the elements of $b \in \mathbb{R}^p$ are chosen iid from the uniform distribution on $[0, 2\pi]$.

   To select the new feature dimension p, perform a training/validation split of 80%/20% on your training data and use the validation set to perform model selection. **Turn in** the test error for your selected value of p, as well as a plot showing your training and validation error as a function of p.
Notes:

   1. You should use the same G to transform the entire dataset.

   2. You may run into memory issues for $p > 6000$.

   3. You may use $\lambda = 10^{-4}$ as in Homework 1, but feel free to perform model selection on λ and/or σ^2 as well.

Problem 2   Strong Convexity (10 pts)

Prove that $f(w) = \lambda \|w\|^2$ is 2λ-strongly convex.

Problem 3   SLT (10 pts)

(*SLT rules apply.*) UML, Ch. 12, Exercise 3. State how long you worked on the problem before looking at the solution.

Problem 4   SLT (5 pts)

(*SLT rules apply.*) UML, Ch. 13, Exercise 1, first bullet only. State how long you worked on the problem before looking at the solution.