

Machine Learning

ECE 4332 / ECE 5332

Project 3

1. Generate the data set \mathcal{D} as follows:
 - a. $L = 100$
 - b. $N = 25$
 - c. X contains samples from a uniform distribution $\mathcal{U}(0,1)$.
 - d. $t = \sin(2\pi X) + \varepsilon$, where ε contains samples from a Gaussian distribution $\mathcal{N}(0, \sigma=0.3)$.
2. Select a set of permissible values for the regularization parameter λ .
3. For each value of λ , use the method of “linear regression with non-linear models” to fit Gaussian basis functions to each of the datasets. Use $s = 0.1$.
4. Produce the plot as shown below, where

$$\bar{f}(x) = \frac{1}{L} \sum_{l=1}^L f^{(l)}(x)$$
$$(\text{bias})^2 = \frac{1}{N} \sum_{n=1}^N [\bar{f}(x^{(n)}) - h(x^{(n)})]^2$$
$$\text{variance} = \frac{1}{N} \sum_{n=1}^N \frac{1}{L} \sum_{l=1}^L [f^{(l)}(x^{(n)}) - \bar{f}(x^{(n)})]^2$$

5. The test error curve is the average error for a test data set of 1000 points.

Upload your .m or .py file to Blackboard prior to the deadline.

Name your files as Lastname_Firstname_Project3.*

