

# Modern Machine Learning

## Homework #5

1. Consider a Lasso regression model with the cost function  $J(\boldsymbol{\beta})$  given by the expression

$$J(\boldsymbol{\beta}) = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\|\boldsymbol{\beta}\|_1$$

Suppose that the error on the training set  $J_{train}(\boldsymbol{\beta})$  is low, and that the error on the testing set  $J_{val}(\boldsymbol{\beta})$  is significantly higher.

- (a) Is the model overfitting or underfitting?
  - (b) Provide two possible solutions to improve the performance of the regression model.
  - (c) If we replace the Lasso regression by a ridge regression, will the testing error decrease? Explain your answer.
2. Consider an  $L$ -layer neural network and a training set of the form  $\{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), (\mathbf{x}^{(2)}, \mathbf{y}^{(2)}), \dots, (\mathbf{x}^{(n)}, \mathbf{y}^{(n)})\}$ . The cost function  $J(\boldsymbol{\Theta})$  is given by the expression

$$J(\boldsymbol{\Theta}) = \sum_{k=1}^n \|\mathbf{y}^{(k)} - \mathbf{a}^{(k)(L)}\|^2 + \frac{\lambda}{2} \sum_l \sum_i \sum_j \left(\theta_{ij}^l\right)^2,$$

where  $\theta_{ij}^l$  is the weight from the  $j$ th neuron in layer  $l - 1$  to the  $i$ th neuron in layer  $l$ . Assume that  $i$  and  $j$  are fixed.

- (a) Provide two possible solutions to improve performance if the model is overfitting the data. Explain your answer.
  - (b) Provide two possible solutions to improve performance if the model is underfitting the data. Explain your answer.
3. An alternate approach to regulation called dropout has received much attention and is viewed as the preferred regulation practice, at least in deep neural networks.
  - (a) Provide the definition of dropout and explain how it works, advantages, and disadvantages. How is the level of regulation controlled?
  - (b) Provide a graphic representation of two neural networks, one without dropout and another with dropout.