Advanced Machine Learning Homework 4

February 19, 2020

VC Dimension of Neural Networks

We take 0/1 classification problem for data with d dimensional features as an example.

A neural network with one hidden layer can be written as

$$o = \mathbf{w}_2^T \sigma(\mathbf{W}_1 \mathbf{v} + \mathbf{b}_1) + b_2, \tag{1}$$

where $\mathbf{v}$ is the d dimensional input feature of the data, while $\mathbf{W}_1, \mathbf{w}_2, \mathbf{b_1}, b_2$ are the parameters of the neural network model. $\mathbf{W}_1$ is a $n * d$ matrix, $\mathbf{w}_2$ is a d dimensional vector, and $\mathbf{b}_1$ is a n dimensional bias vector while b_2 is the bias.

When $o > 0$ we classify the datum as label 1, while when $o <= 0$ we classify it as label 0. This forms a neural network, or multi-layer-perceptron, with one hidden layer containing n neurons.

Problems

1. Given n, d, calculate the VC dimension of the neural network for the linear activation case, i.e. $\sigma(x) = x$. Prove your result.
2. Given n, d, calculate the VC dimension of the neural network for the ReLU activation case, i.e. $\sigma(x) = \max(0, x)$. Prove your result.