Machine Learning 6.867 - Pset 3

November 6, 2015

1 Multi-Class SVM

2 Neural Networks

Neural networks are used in machine learning to make predictions, similar to logistic regression, SVM, or regression. We can represent neural networks using a graph with nodes and edges (see Bishop figure 5.1). Assume that we observe data $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}), i = 1, \ldots, n$, where $\mathbf{x}^{(i)} \in \mathbb{R}^D$ and $\mathbf{y}^{(i)} \in \{0,1\}^K$. Let $(\mathbf{x}, \mathbf{y}) = ([x_1, \ldots, x_D], [y_1, \ldots, y_K])$ be a general observation. We create nodes for each of the features x_i, referred to as *inputs*, and nodes for each of the class labels y_i which are called *outputs*. Next, we introduce a series of nodes in the middle of the graph, called *hidden units*, and we draw edges connecting the *inputs* $\rightarrow$ *hidden units* $\rightarrow$ *outputs*. The key idea in neural networks is that there

2.1 Implementation