Assignment 4 ML

J L Gouws
19G4436

April 14, 2022

1. The weights from the first layer to the second layer(first hidden layer) are supposed to represent the contribution of the pixels to little line segments. The sencond layer is supposed to pick up on small “edges" or line segments.

2. The weights from the second layer to the third layer represent collection of the “edges" making up larger structures. These weights are supposed to represent the contribution of the small line segments to lines and circles. The weights for input into the final layer are supposed to represent how the lines and circles get added together to form the numbers.

3. The weights are not coded or worked out by hand to represent this. A computer works out the weights. The network needs some starting values for the weights, these are set randomly for ease of use. Since the network starts randomly there is no reason for the training network to follow a human’s semantic structure. The learning tries to find patterns that are natural for the network to recognize, which might not be natural/sensical to humans.

1. Machine learning is the study of finding or making a function f that takes input x from a set D and gives output y, where nothing is known about the relationship between x and y a priori. Supervised learning is the case where y is known for some subset of D. Unsupervised learning is the case where nothing is known about the ouput of the function. An example would be taking a picture, x of a dog and applying f to classify, y, the picture as a picture of a dog.

2. (a) Feature scaling is changing the values of features so that they conform to some restriction, without changing the information encoded within the features.

   (b) Distance based learning algorithms.

   - K Nearest Neighbors
   - Neural networks

   (c) Tree based algorithms

   - Random forest.
   - Regression problems.

3. (a)

   (b) **Advantages** It gives a better estimation of the average generalization of the model.

   Disadvantages It takes a long time to run. This is a less significant problem, but more data is generally needed.