

Neural Networks

Suppose that a training set contains only a single example, repeated 100 times. In 80 of the 100 cases, the single output value is 1; in the other 20, it is 0. What will a back-propagation network predict for this example, assuming that it has been trained on all the training examples and reaches a global optimum? (Hint: To find the global optimum, differentiate the error function and set to zero.)

If we train a neural network for 1000 epochs, does it make a difference whether we present all training examples in turn for 1000 times or whether we first present the first training example 1000 times, then the second training example for 1000 times, and so on? Why?

What are the differences between a single perceptron and a network of perceptrons in terms of a) expressivity and b) ease of solving inductive learning problems?