

COMP9444 Neural Networks and Deep Learning

Session 2, 2018

Exercises 2: Backpropagation

This page was last updated: 07/31/2018 10:45:10

1. Identical Inputs

Consider a degenerate case where the training set consists of just a single input, repeated 100 times. In 80 of the 100 cases, the target output value is 1; in the other 20, it is 0. What will a back-propagation neural network predict for this example, assuming that it has been trained and reaches a global optimum? (Hint: to find the global optimum, differentiate the error function and set to zero.)

2. Linear Transfer Functions

Suppose you had a neural network with linear transfer functions. That is, for each unit the activation is some constant c times the weighted sum of the inputs.

- a. Assume that the network has one hidden layer. We can write the weights from the input to the hidden layer as a matrix \mathbf{W}^{HI} , the weights from the hidden to output layer as \mathbf{W}^{OH} , and the bias at the hidden and output layer as vectors \mathbf{b}^{H} and \mathbf{b}^{O} . Using matrix notation, write down equations for the value \mathbf{O} of the units in the output layer as a function of these weights and biases, and the input \mathbf{I} . Show that, for any given assignment of values to these weights and biases, there is a simpler network with no hidden layer that computes the same function.
- b. Repeat the calculation in part (a), this time for a network with any number of hidden layers. What can you say about the usefulness of linear transfer functions?

Make sure you try answering the Exercises yourself, before checking the [Sample Solutions](#)