



COS711 Assignment 1

Due date: 27 August 2023, at 23h30

1 General instructions

This assignment is theoretical, and will test your understanding of the backpropagation algorithm. You have to submit a single pdf document containing your answers to the questions provided. Note: the assignment is designed to test your ability to derive the weight update equations for arbitrary loss and activation functions. Thus, you will lose marks by skipping over steps. Make sure your derivations are readable, notation is correct, and the steps (including simplifications) are clear.

The report will be checked for plagiarism using Turnitin, and should be submitted through the ClickUp system. You are advised but not required to typeset your report in LaTeX.

2 Deriving backpropagation (25 marks)

A feed-forward neural network is set up to have an input layer of size I , a hidden layer of size H , and an output layer of size O . The following activation functions are employed in the neurons on each layer:

- Input layer: $f(x) = x$
- Hidden layer: $f(x) = e^{-x^2} - \frac{1}{2}x$
- Output layer: $f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Hidden and output neurons are summation units, i.e. x in the activation functions above refers to the *net* input signal, thus the output signal of a neuron j is $y_j = f(\text{net}_j) = f(\sum_k w_{kj}y_k)$. Bias signal is sent to all hidden, as well as all output neurons. Assume the objective function E is used, defined for a single output neuron as:

$$E(y_i) = \frac{1}{2}(1 - y_i t_i) + \frac{1}{2}\sqrt{(1 - y_i t_i)^2 + \epsilon}$$

where y_i is the output of the i -th output neuron, t_i is the target output of the i -th output neuron, and ϵ refers to a non-zero constant, respectively.

Answer the questions below:

1. Derive the update rule for the non-bias weights between the hidden and the output layer. Show all steps, including simplifications. (10 marks)
2. Derive the update rule for the non-bias weights between the input and the hidden layer. Show all steps, including simplifications. Make sure to write down the complete update equation with all intermediate variables correctly substituted. (10 marks)
3. Will the bias weight update rules differ from the non-bias weight update rules? Show the update rule for the bias weights associated with the hidden layer. (5 marks)

Show your workings for all of the questions above. No marks will be awarded for final solutions only.