Exercise 8

The calculations used for the code are found on the next page.

The a2 activation is a linear activation and does not use sigmoid. The a1 activation uses sigmoid.

A black screen with white text

AI-generated content may be incorrect.

The mean squared error after epoch 40 is around 0.008 both for training and test sets, which is good. 0 would mean a perfect fit.

The chain tree drawn in the next page assumes a hidden layer and input layer of one unit, this is to simplify the thought process. In the code, since the layers have their data in matrix form, it can easily be extended to support more weights as shown in the code without any significant changes to the logic.

In the next page has the same value as in the code.

A math equations on a white background

AI-generated content may be incorrect.