The below graphs show the results of a model with sigmoid activation in all neurons, trained for 500 epochs. A final test accuracy of 84.4% was achieved, while only 80.3% was achieved in training. This is the output of task 1.

Chart, line chart

Description automatically generatedChart, line chart

Description automatically generated

Initially from looking at the graphs, it looks like there was still room for improvement with regards to accuracy. As such, I started by increasing the number of epochs I trained the model for. Results from 5000 epochs of training can be seen below. The test accuracy plateaued at around 4000 epochs, while the train accuracy kept improving. This, combined with the graph of the errors, indicates that the model began overfitting. As such, going forward 4000 epochs will be used for training.

Chart, line chart

Description automatically generated with medium confidenceA picture containing shape

Description automatically generated

Following this, I looked at different activation functions. I changed the first layer to have relu activation. This resulted in much higher accuracies much faster. Training accuracy reached 99.1%, with test accuracy reaching 94.4% after very few epochs.

A picture containing graphical user interface

Description automatically generatedShape

Description automatically generated with low confidence

From here, I thought it might be a good idea to lower the learning rate. Dropping the learning rate to 0.002 gave similar accuracies, while keeping the test and train accuracies closer together. This reduces the risk of overfitting.

A picture containing shape

Description automatically generatedShape

Description automatically generated with medium confidence

In a final attempt to improve the accuracy of the model I increased the number of neurons in the hidden layer. Surprisingly I saw little to no difference in performance when I added neurons. I attempted many values, none of which resulted in any useful improvements. Below are the accuracy and error graphs for a model with 20 neurons in the hidden layer.

A picture containing graphical user interface

Description automatically generatedShape

Description automatically generated with medium confidence