



EINDHOVEN UNIVERSITY OF TECHNOLOGY

PROJECT AI FOR MIA

8P361

Assignment 2

Group 3:

Niek van Belzen (1675230)

Esmée Janssen (1867334)

Quinten Liu (1842471)

Tijmen Vierwind (1685392)

February 21, 2025

1 Exercise 1

The example neural network classification model in this assignment is relatively simple — it contains a single hidden layer of 64 neurons. Perform a set of experiments with more complex models, e.g. with more layers (deeper models), more neurons per layer or a combination. Describe the set of experiments that you have performed. What is the accuracy of the best model? How did you determine which model is the best?

Using the `keras-tuner` python module, hyper-parameter tuning becomes relatively easy to perform. To make tuning less intensive, 3 Epochs with batch size 16 were performed on each hyper-parameter in a grid search of 3 by 3. The model could have 1,5 or 10 hidden layers, and 64,128 or 256 neurons per layer. The table below shows the results:

Layers	Neurons	Val Loss	Val Accuracy
1	64	0.2089	0.9417
1	128	0.1863	0.9457
1	256	0.1843	0.9463
5	64	0.1177	0.9643
5	128	0.1040	0.9687
5	256	0.0976	0.9692
10	64	0.1585	0.9542
10	128	0.1061	0.9683
10	256	0.1093	0.9662

From this, the conclusion can be drawn that a model with 5 hidden layers and 256 neurons performs the best out of these options, with the highest validation accuracy of 0.9692 and the lowest validation loss.

Training a model with the original default epochs (10) and batch size (32) and the found optimal hyper-parameters gives a test loss of 0.0868 and test accuracy of 0.9721.

2 Exercise 2

Compare the performance of the following three models:

- *Neural network without any hidden layers (the input layer connects directly to the output layer).*

The accuracy for a Neural Network without any hidden layers is 0.9151.

- *Neural network with 3 hidden layers with ReLU activations.*

The accuracy for a Neural Network with 3 hidden layers with ReLU activations is 0.9633.

- *Neural network with 3 hidden layers with linear activations (i.e. without nonlinearities between the layers).*

The accuracy for a Neural Network with 3 hidden layers with Linear activations is 0.9179.

Analyze the performance of the three models. What is the reason behind the difference in performance between the second and third models?

The model with the highest accuracy is the Neural Network with 3 hidden layers with ReLU activations (second model). The difference between the second and third model is the shape of the activation function in the hidden layers. For Neural Network 2, the model uses ReLU which is a non-linear function. Non-linearities are needed to train a more complex model. Without non-linearities, the model would simply be a linear model and will not perform as well for the same data. Model 1 does not use any hidden layer, therefore the input is directly linked to the output and will not perform better than model 2.

3 Exercise 3

Train a neural network model (the specific architecture is up to you) for a four class classification problem derived from MNIST in the following way:

- "vertical digits": 1, 7
- "loopy digits": 0, 6, 8, 9
- "curly digits": 2, 5
- "other": 3, 4

The resulting model was obtained by relabeling the training and test datasets before pre-processing. The distribution of classes in the full dataset can be seen below. The distribution makes sense as there are 2 times more 'loopy digits' than any other class in this classification problem. With 5 hidden layers of 256 neurons, as found optimal in exercise 1, this resulted in a test accuracy of approximately 0.9744 and a test loss of 0.08315.

