# Assignment 2: answers Group 4

## Exercise 1:

The following models were run for 10 epochs: the default model with one layer of 64 neurons, a model with one layer of 128 neurons, a model with three layers of 64 neurons and a model with three layers of 64, 128, and 64 neurons, respectively. Compared to the default model, the model with one layer of 128 neurons had a higher accuracy for both the training and validation datasets. Furthermore, the loss for both datasets was also lower, which is desirable. For the model with three layers of 64 neurons, the accuracy for both datasets was higher than both the default model and the model with one layer of 128 neurons. Additionally, the loss for both datasets was also lower than both previous models. Finally, the model with three layers of 64, 128, and 64 neurons, respectively, performed similarly to the model with three layers of 64 neurons. The accuracy and loss for each model can be found in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Train accuracy | Validation accuracy | Train loss | Validation loss |
| One layer of 64 neurons | 0.9467 | 0.941 | 0.189 | 0.2033 |
| One layer of 128 neurons | 0.9513 | 0.9435 | 0.1752 | 0.1931 |
| Three layers of 64 neurons | 0.9679 | 0.9613 | 0.109 | 0.1294 |
| Three layers of 64, 128 and 64 neurons, respectively | 0.9671 | 0.9602 | 0.1105 | 0.1378 |

The model with three layers of 64 neurons has both the highest accuracy and lowest loss. This means that it produces the most correct predictions and has the least error in wrong predictions. It should be noted that for all models, after 10 epochs, the validation accuracy is slightly lower than the train accuracy, which may indicate that the models may be slightly overfit to the training data.

## Exercise 2:

The results of the three models are given in the table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Train accuracy | Validation accuracy | Train loss | Validation loss |
| No hidden layers | 0.9098 | 0.9003 | 0.3239 | 0.3574 |
| Three layers of 64 neurons with ReLU activations | 0.9681 | 0.9612 | 0.1073 | 0.1305 |
| Three layers of 64 neurons with linear activations | 0.9245 | 0.9143 | 0.2699 | 0.3023 |

As can be seen in the table, the model with three layers of 64 neurons with ReLU activations has the highest accuracy and lowest loss for the train and validation dataset. The model with three layers, but linear activations, has a slightly lower accuracy and slightly higher loss compared to the previous model. The model with no hidden layers has the lowest accuracy and the highest loss for both datasets.

The difference in performance between the models with ReLU activations and only linear activations is caused by the fact that the model with only linear activations can be reduced to a linear regression model. This makes a neural network without any non-linear activations effectively only one layer deep.