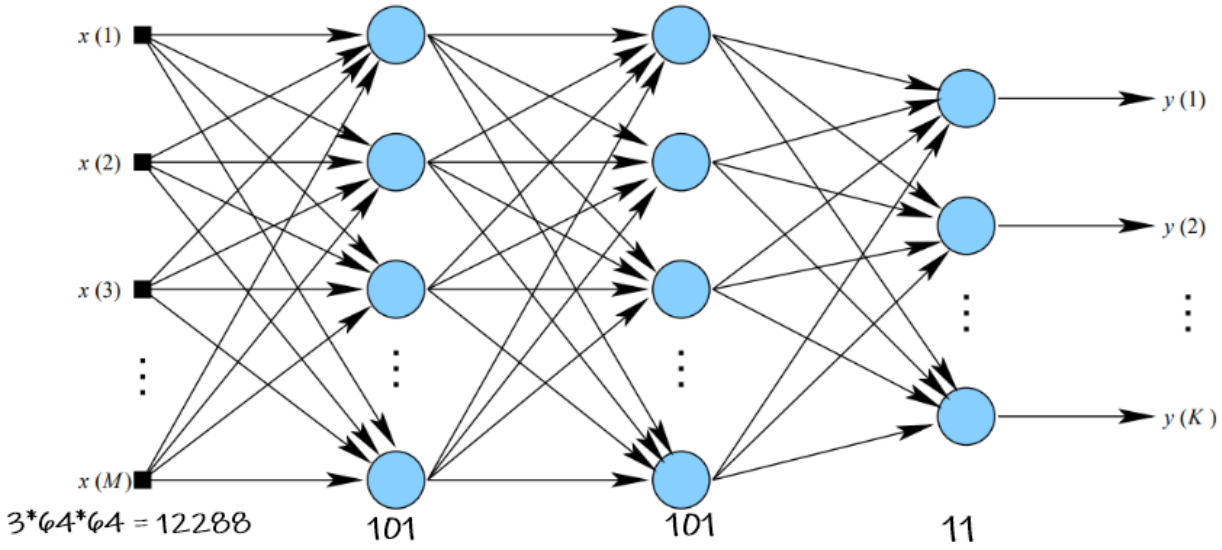
Exercise Set 2 Valtteri Nikkanen

1. Full-connected network
   a. Number of parameters in the network



So the number of parameters can be calculated :
12288*101*101*11 = 1 378 848 768

We have about 1.4 billion weights in the network.

b. If there needs to be 5 times as many training samples as weights there needs to be
1228800000*5 = 6 894 243 840

So, we would need nearly 6.9 billion training samples.

2, 3, 4. Running with the GPU

```
2.6.0
2023-03-26 12:59:41.364355: I tensorflow/core/platform/cpu_feature_guard.cc:142] This TensorFlow binary is optimized
 with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical
 operations:  AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
GPU found
2023-03-26 12:59:42.024673: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1510] Created device /device:GPU:0 with
 6607 MB memory:  -> device: 0, name: NVIDIA GeForce GTX 1080, pci bus id: 0000:01:00.0, compute capability: 6.1
Found 528 images belonging to 2 classes.
Found 132 images belonging to 2 classes.
2023-03-26 12:59:42.070133: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1510] Created device
 /job:localhost/replica:0/task:0/device:GPU:0 with 6607 MB memory:  -> device: 0, name: NVIDIA GeForce GTX 1080, pci bus
  id: 0000:01:00.0, compute capability: 6.1
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten (Flatten)            (None, 12288)             0
_________________________________________________________________
dense (Dense)                (None, 10)                122890
_________________________________________________________________
dense_1 (Dense)              (None, 10)                110
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 22
=================================================================
Total params: 123,022
Trainable params: 123,022
Non-trainable params: 0
_________________________________________________________________
None
2023-03-26 12:59:42.622082: I tensorflow/compiler/mlir/mlir_graph_optimization_pass.cc:185] None of the MLIR
 Optimization Passes are enabled (registered 2)
Epoch 1/10
17/17 [==============================] - 3s 61ms/step - loss: 0.6572 - accuracy: 0.7311
Epoch 2/10
17/17 [==============================] - 1s 63ms/step - loss: 0.4290 - accuracy: 0.8106
Epoch 3/10
17/17 [==============================] - 1s 64ms/step - loss: 0.4157 - accuracy: 0.8333
Epoch 4/10
17/17 [==============================] - 1s 53ms/step - loss: 0.2378 - accuracy: 0.9205
Epoch 5/10
17/17 [==============================] - 1s 46ms/step - loss: 0.2406 - accuracy: 0.9148
Epoch 6/10
17/17 [==============================] - 1s 51ms/step - loss: 0.1899 - accuracy: 0.9280
Epoch 7/10
17/17 [==============================] - 1s 58ms/step - loss: 0.1992 - accuracy: 0.9186
Epoch 8/10
17/17 [==============================] - 1s 53ms/step - loss: 0.3177 - accuracy: 0.8826
Epoch 9/10
17/17 [==============================] - 1s 55ms/step - loss: 0.1750 - accuracy: 0.9337
Epoch 10/10
17/17 [==============================] - 1s 48ms/step - loss: 0.1328 - accuracy: 0.9583

Validation data
5/5 [==============================] - 0s 44ms/step - loss: 0.2041 - accuracy: 0.9167
Validation loss: 0.2041366994380951
Validation accuracy: 0.9166666865348816
```

We achieved roughly 92% accuracy with the neural network with sigmoid output layer activation.

```
2.6.0
2023-03-26 13:01:08.331230: I tensorflow/core/platform/cpu_feature_guard.cc:142] This TensorFlow binary is optimized
 with oneAPI Deep Neural Network Library (oneDNN) to use the following CPU instructions in performance-critical
 operations:  AVX AVX2
To enable them in other operations, rebuild TensorFlow with the appropriate compiler flags.
GPU found
2023-03-26 13:01:08.903455: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1510] Created device /device:GPU:0 with
 6607 MB memory:  -> device: 0, name: NVIDIA GeForce GTX 1080, pci bus id: 0000:01:00.0, compute capability: 6.1
Found 528 images belonging to 2 classes.
Found 132 images belonging to 2 classes.
2023-03-26 13:01:08.950823: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1510] Created device
 /job:localhost/replica:0/task:0/device:GPU:0 with 6607 MB memory:  -> device: 0, name: NVIDIA GeForce GTX 1080, pci bus
  id: 0000:01:00.0, compute capability: 6.1
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten (Flatten)            (None, 12288)             0
_________________________________________________________________
dense (Dense)                (None, 10)                122890
_________________________________________________________________
dense_1 (Dense)              (None, 10)                110
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 22
=================================================================
Total params: 123,022
Trainable params: 123,022
Non-trainable params: 0
_________________________________________________________________
None
2023-03-26 13:01:09.374319: I tensorflow/compiler/mlir/mlir_graph_optimization_pass.cc:185] None of the MLIR
 Optimization Passes are enabled (registered 2)
Epoch 1/10
17/17 [==============================] - 2s 55ms/step - loss: 0.6858 - accuracy: 0.6780
Epoch 2/10
17/17 [==============================] - 1s 80ms/step - loss: 0.6308 - accuracy: 0.8068
Epoch 3/10
17/17 [==============================] - 1s 46ms/step - loss: 0.5777 - accuracy: 0.8314
Epoch 4/10
17/17 [==============================] - 1s 49ms/step - loss: 0.5403 - accuracy: 0.8769
Epoch 5/10
17/17 [==============================] - 1s 48ms/step - loss: 0.5074 - accuracy: 0.8883
Epoch 6/10
17/17 [==============================] - 1s 44ms/step - loss: 0.4899 - accuracy: 0.8883
Epoch 7/10
17/17 [==============================] - 1s 45ms/step - loss: 0.4301 - accuracy: 0.9375
Epoch 8/10
17/17 [==============================] - 1s 45ms/step - loss: 0.4148 - accuracy: 0.9337
Epoch 9/10
17/17 [==============================] - 1s 46ms/step - loss: 0.3937 - accuracy: 0.9621
Epoch 10/10
17/17 [==============================] - 1s 43ms/step - loss: 0.3774 - accuracy: 0.9451

Validation data
5/5 [==============================] - 0s 36ms/step - loss: 0.3759 - accuracy: 0.9242
Validation loss: 0.3759040832519531
Validation accuracy: 0.9242424368858337
```

Seems like we achieved about 92% accuracy with our neural network with softmax as our output layer activation.

Running with CPU

```
2.6.0
No GPU found
Found 528 images belonging to 2 classes.
Found 132 images belonging to 2 classes.
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten (Flatten)            (None, 12288)             0
_________________________________________________________________
dense (Dense)                (None, 10)                122890
_________________________________________________________________
dense_1 (Dense)              (None, 10)                110
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 22
=================================================================
Total params: 123,022
Trainable params: 123,022
Non-trainable params: 0
_________________________________________________________________
None
2023-03-26 12:56:53.386428: I tensorflow/compiler/mlir/mlir_graph_optimization_pass.cc:185] None of the MLIR
 Optimization Passes are enabled (registered 2)
Epoch 1/10
17/17 [==============================] - 1s 43ms/step - loss: 0.6564 - accuracy: 0.6686
Epoch 2/10
17/17 [==============================] - 1s 41ms/step - loss: 0.4871 - accuracy: 0.7973
Epoch 3/10
17/17 [==============================] - 1s 42ms/step - loss: 0.3753 - accuracy: 0.8220
Epoch 4/10
17/17 [==============================] - 1s 58ms/step - loss: 0.3044 - accuracy: 0.8750
Epoch 5/10
17/17 [==============================] - 1s 45ms/step - loss: 0.2716 - accuracy: 0.8845
Epoch 6/10
17/17 [==============================] - 1s 43ms/step - loss: 0.2115 - accuracy: 0.9280
Epoch 7/10
17/17 [==============================] - 1s 40ms/step - loss: 0.2019 - accuracy: 0.9223
Epoch 8/10
17/17 [==============================] - 1s 41ms/step - loss: 0.3287 - accuracy: 0.8655
Epoch 9/10
17/17 [==============================] - 1s 43ms/step - loss: 0.1962 - accuracy: 0.9318
Epoch 10/10
17/17 [==============================] - 1s 45ms/step - loss: 0.1704 - accuracy: 0.9451

Validation data
5/5 [==============================] - 0s 36ms/step - loss: 0.1458 - accuracy: 0.9545
Validation loss: 0.14576727151870728
Validation accuracy: 0.9545454382896423
```

Using sigmoid activation. Seems that there isn't that much of a difference with the run times but that might be because the network is of a manageable size still with this few neurons.