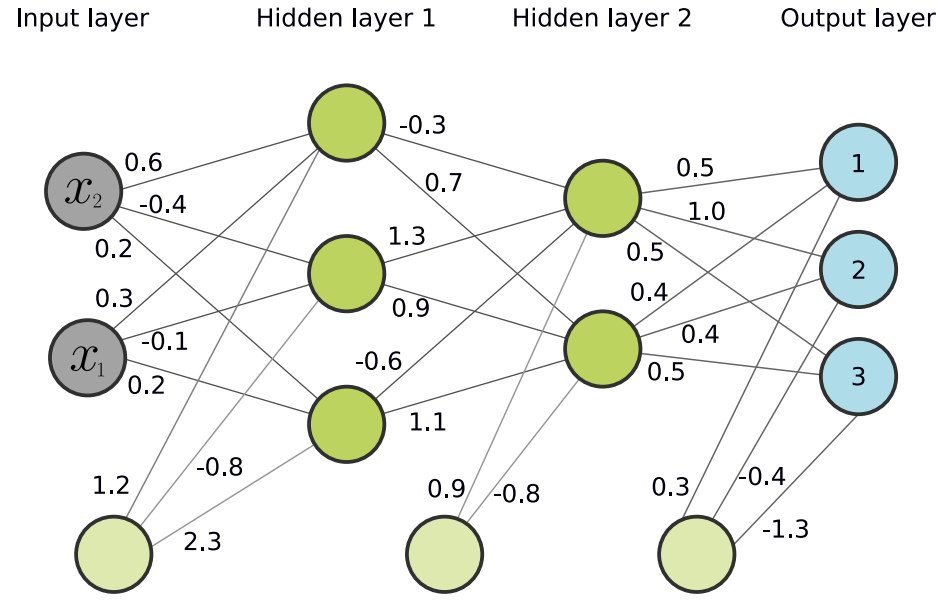
STAT40970 – Machine Learning & A.I. (Online)

Assignment 2

Deadline - Friday 19th April at 23:59pm

Exercise 1

The figure below shows a multiple layer neural network deployed to predict the outcome of a 3-class categorical target variable y. The activation function in both hidden layers is the rectified linear unit (ReLU) and the output layer uses a softmax output activation function. The weight and bias parameters are shown along the edges in the network. The numbers in the output nodes denote the class label corresponding to each node.



1. Consider a generic input observation vector $\mathbf{x} = (x_1, x_2)$. Perform a full forward propagation calculation through the network for the input observation vector $(0.1, -0.7)$.

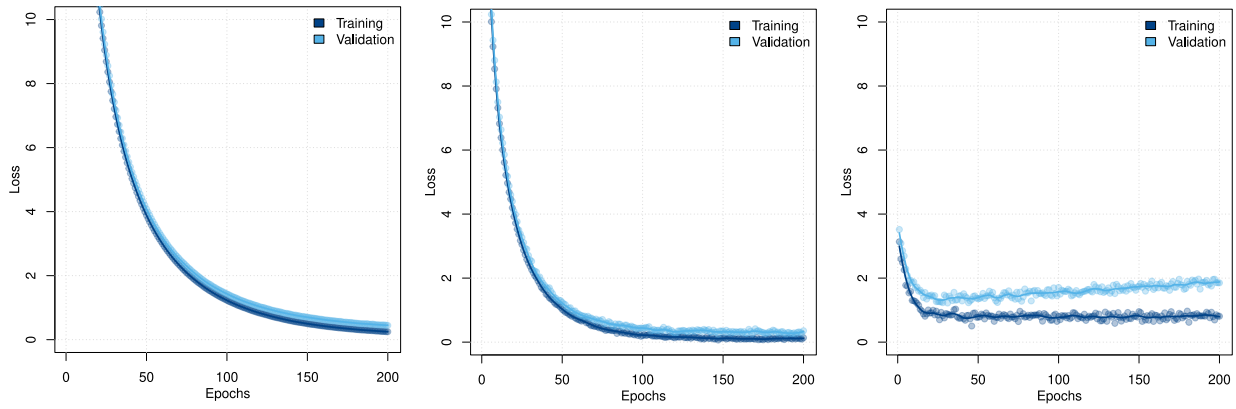
(20 marks)

2. What is the effect of removing the bias parameters from the first layer of this network? Justify your answer.

(10 marks)

3. The network is trained focusing on fine tuning of a particular hyperparameter. The figures below (in the next page) report the loss learning curves corresponding to three different values of this hyperparameter. The value of this hyperparameter is increasing going from left to right. What is most likely this hyperparameter and how does it affect the learning and validation processes? Comment briefly.

(10 marks)



4. The table below reports the outputs produced by the network and the associated labels of the target variable for a sample of 4 observations. Compute the average loss and the classification accuracy for this sample.

Label	1	2	2	3
o_{i1}	0.62	0.12	0.07	0.22
o_{i2}	0.35	0.08	0.40	0.07
o_{i3}	0.03	0.80	0.53	0.71

(20 marks)

Exercise 2

The following R `keras` chunk of code is used to define a convolutional neural network applied to the classification of image data and trained on a sample of N = 15725 training instances.

```
# model definition
model <- keras_model_sequential() %>%
  #
  # convolutional layers
  layer_conv_2d(filters = 64, kernel_size = c(6,6), activation = "relu",
                input_shape = c(256, 256, 3), name = "conv_1") %>%
  layer_max_pooling_2d(pool_size = c(2,2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(4,4), activation = "relu",
                name = "conv_2") %>%
  layer_max_pooling_2d(pool_size = c(2,2)) %>%
  layer_conv_2d(filters = 128, kernel_size = c(2,2), activation = "relu",
                name = "conv_3") %>%
  layer_max_pooling_2d(pool_size = c(2,2)) %>%
  #
  # fully connected layers
  layer_flatten() %>%
  layer_dense(units = 64, activation = "relu",
              kernel_regularizer = regularizer_l2(0.2),
              name = "dense_1") %>%
  layer_dense(units = 15, activation = "softmax") %>%
  #
  # compile
  compile(
    loss = "categorical_crossentropy",
    metrics = "accuracy",
    optimizer = optimizer_adam()
  )
```

```r
# training and validation
fit <- model %>% fit(
  x = x_train, y = y_train,
  validation_data = list(x_val, y_val),
  epochs = 200,
  batch_size = 185,
  callbacks = callback_early_stopping(monitor = "val_loss", patience = 20)
)
```

1. Using the information from the code chunk above, extrapolate the following:

- Type and size of the images in the data.
- Depth of the convolutional layers.
- Size of the filters.
- Number of batches processed in each training epoch.
- Activation functions and output activation.
- Regularization (if any).

(10 marks)

2. Compute the following quantities, showing all steps and calculations.:

- The number of parameters of the second convolutional layer.
- The number of parameters of the first dense layer.

(30 marks)

Submission rules and instructions

- Write a short report and submit it as a single pdf file (approximately max 10 pages).
- Multiple submissions before deadline are allowed and only the latest one will be considered for marking.
- Submission after deadline will incur in penalization as UCD rules (see "Module details" document).
- In general, for full marks you **must explain** concisely and clearly **all reasoning**, as well as **show all steps** and **computations** in your answers. Correct answers alone will not achieve full marks.
- **Plagiarism is strictly prohibited** and it will incur in severe penalties (see "Module details" document and "Information materials" tab).