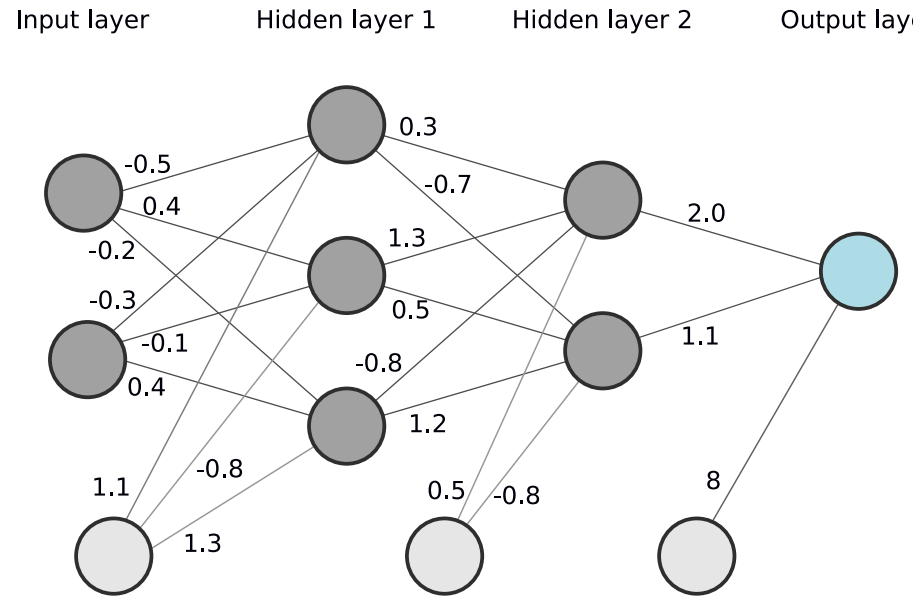
STAT40970 – Machine Learning & A.I. (Online)

Assignment 2

Deadline - Monday 3 April at 17:00

Exercise 1

The figure below shows a multiple layer neural network deployed to predict the outcome of a numerical continuous target variable y. The activation function in the first hidden layer is the sigmoid function, while the one in the second hidden layer is the Rectified Linear Unit (ReLU). The output layer uses an identity output function. The weight and bias parameters are shown along the edges in the network.



1. Comment briefly on the architecture and on the number of parameters of this network. *(5 marks)*

2. Perform a forward propagation calculation through the network for the input observation vector $\mathbf{x} = (-0.5, 0.3)$. What is the value of the output unit corresponding to this input vector? *(15 marks)*

3. The value of the target variable associated to the input vector $\mathbf{x} = (-0.5, 0.3)$ is $y = 7$. Calculate the loss associated with this training instance using an appropriate loss function. *(10 marks)*

Exercise 2

The following R-Keras chunk of code is used to train a deep neural network applied to a sample of $N = 27200$ training instances.

```
# model definition
model <- keras_model_sequential() %>%
  layer_dense(units = 256, input_shape = 2048, name = "layer_1",
              kernel_regularizer = regularizer_l2(l = 0.1)) %>%
  layer_activation_leaky_relu(alpha = 0.01) %>%
  layer_dense(units = 128, name = "layer_2",
              kernel_regularizer = regularizer_l2(l = 0.1)) %>%
  layer_activation_leaky_relu(alpha = 0.01) %>%
  layer_dense(units = 128, name = "layer_3",
              kernel_regularizer = regularizer_l2(l = 0.1)) %>%
  layer_activation_leaky_relu(alpha = 0.01) %>%
  layer_dense(units = 10, activation = "softmax", name = "layer_out") %>%
  compile(loss = "categorical_crossentropy", metrics = "accuracy",
          optimizer = optimizer_sgd(learning_rate = 0.001)
  )

# training and validation
fit <- model %>% fit(
  x = x_train, y = y_train,
  validation_data = list(x_val, y_val),
  epochs = 100,
  batch_size = 544,
  callbacks = callback_early_stopping(monitor = "val_loss", patience = 10)
)
```

Using the information from the code chunk above, indicate the following:

1. Number of input units. *(2 marks)*
2. Number of batches processed in each epoch. *(3 marks)*
3. Type of task for which the network is employed. *(2 marks)*
4. Type of regularization used. *(3 marks)*

Exercise 3

The file `data_epileptic.RData` contains data concerning electroencephalography (EEG) measurements of a sample composed by patients with epilepsy and healthy subjects. The task is to predict the type of EEG activity using the signal features recorded under different conditions and different areas of the brain. More information is available here: `https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition`.

The target variable is a categorical variable indicating the following EEG activities related to different subjects, conditions, and areas of the brain: `1` - Epileptic seizure activity, `2` - Patient with tumor formation, EEG activity recorded on healthy area during epilepsy-free interval, `3` - Patient with tumor formation, EEG activity recorded on tumor location area during epilepsy-free interval, `4` - EEG activity of healthy subject with eyes closed, `5` - EEG activity of healthy subject with eyes open.

The file includes the data matrices `x`, `y`, `x_test`, and `y_test`. The data matrices `x` and `y` contain the data which will be used for training and validation of the model, while the data matrices `x_test` and `y_test` will be used for testing. The data matrices denoted with "x" correspond to the input data, while those denoted with "y" correspond to the target class labels.

1. Deploy a neural network with 2 hidden layers and a neural network with 3 hidden layers. Compare appropriately the performances of these two networks at predicting the type of condition (and area of the brain) an EEG signal belongs to. *(40 marks)*
2. Does adding one extra hidden layer to the network architecture lead to an improvement in the predictive performance? Discuss briefly. *(10 marks)*
3. Evaluate the test accuracy of the two deep neural networks and comment briefly on their performance. Which of the two networks provide the best accuracy at predicting EEG signals of patients with tumor formations on the test data? Motivate. *(10 marks)*

Instructions:

- Use at least one approach of your choice for regularization.
- Submitting only code does not provide full marks.

Submission rules

- Write a short report and submit it as a single pdf file (approximately max 6-8 pages, code excluded).
- Include the R code used for analysis in the report. The report can be produced using R Markdown, with the code included in the main text or as an appendix. **The code must be working and the analysis must be reproducible in all parts**.
- 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).
- **Plagiarism is strictly prohibited** (see "Module details" document and "Information materials" tab).