



Aprendizagem 2022
Lab 7-8: Neural Networks

Practical exercises

I. Multi-layer Perceptron

1. Consider a network with three layers: 5 inputs, 3 hidden units and 2 outputs where all units use a sigmoid activation function.
 - a) Initialize connection weights to 0.1 and biases to 0. Using the squared error loss do a stochastic gradient descent update (with learning rate $\eta=1$) for the training example $\{\mathbf{x} = [1 \ 1 \ 0 \ 0 \ 0]^T, \mathbf{z} = [1 \ 0]^T\}$
 - b) Compute the MLP class for the query point $\mathbf{x}_{new} = [1 \ 0 \ 0 \ 0 \ 1]^T$
2. Consider a network with four layers with the following numbers of units 4, 4, 3, 3. Assume all units use the *hyperbolic tangent* activation function.
 - a) Initialize all connection weights and biases to 0.1. Using the squared error loss do a *stochastic gradient descent* update (with learning rate $\eta=0.1$) for the training example: $\{\mathbf{x} = [1 \ 0 \ 1 \ 0]^T, \mathbf{z} = [0 \ 1 \ 0]^T\}$
 - b) Reusing the computations from the previous exercise do a *gradient descent update* (with learning rate $\eta=0.1$) for the batch with the training example from the a) and the following: $\{\mathbf{x} = [0 \ 0 \ 1 \ 0]^T, \mathbf{z} = [0 \ 0 \ 1]^T\}$
 - c) Consider the learned MLPs from a) and b). Which has smallest squared error?
Which model has better classification accuracy?
 - d) Compute the MLP class for the query point $\mathbf{x}_{new} = [1 \ 1 \ 1 \ 0]^T$
3. Repeat the exact same exercise, but this time with following adaptations:
 - the output units have a *softmax* activation function
 - the error function is *cross-entropy*What are the major differences between using squared error and cross-entropy?

Programming quest

Resources: https://scikit-learn.org/stable/modules/neural_networks_supervised.html
as well as *Classification*, *Regression* and *Evaluation* notebooks

4. Consider a 10-fold CV, and MLPs with a single hidden layer with 5 nodes. Using *sklearn*:
 - a) assess the classification accuracy of the MLP on the *iris* data using a cross-entropy loss
 - b) assess the MAE of the MLP on the *housing* data using a squared error loss