Programming Exercise 2

Due on 27th November

Pick a binary classification dataset from the LIBSVM repository:

`http://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/`

Tasks:

1. **Decision Trees**: Partition the dataset into a training and a testing set. Run a decision tree learning algorithm usign the training set. Test the decision tree on the testing dataset and report the total classification error (i.e. 0/1 error). Repeat the experiment with a different partition. Plot the resulting trees. Are they very similar, or very different? Explain why.
   Advice: it can be convenient to set a maximum depth for the tree.

2. **Support Vector Machines**: Run SVM to train a classifier, using radial basis as kernel function. Apply cross-validation to evaluate different combinations of values of the model hyper-parameters (box constraint C and kernel parameter γ). How sensitive is the cross-validation error to changes in C and γ? Choose the combination of C and γ that minimizes the cross-validation error, train the SVM on the entire dataset and report the total classification error.
   Advice: use a logaritmic range for γ.

3. **Neural Networks**: Train a Multi-Layer perceptron using the cross-entropy loss with ℓ-2 regularization (weight decay penalty). In other words, the activation function equals the logistic function. Plot curves of the training and validation error as a function of the penalty strength α. How do the curves behave? Explain why.
   Advice: use a logaritmic range for hyper-parameter α. Experiment with different sizes of the training/validation sets and different model parameters (network layers).

Attach the source code for each section. You are free to use the programming language/library of your choice. We recommend, scikit-learn `https://scikit-learn.org/stable/index.html`