

---

PPGEE2249 – Aprendizado de Máquina – PPGEE  
Prof. Daniel Guerreiro e Silva

---

*Assignment 4*

- 1) Choose a toolbox and use it to train a Feedforward Neural Network for a **multiclass classification** task. Split your data into a training set, a validation set to define the training stopping condition (number of epochs) and a test set, for the final evaluation of the model. Softmax neurons must be used in the output layer, one for each class, and the training criterion must be cross-entropy minimization. Describe your dataset, the network architecture (hidden units, dimension and number of layers) and discuss the overall performance (dataset splitting strategy, classification accuracy, confusion matrix). Suggestion of dataset repository: Kaggle, UC Irvine Machine Learning Repository — <http://archive.ics.uci.edu/ml/>
- 2) Choose a toolbox (e.g., Python [scikit-learn](#) or [WEKA](#)) and use it to train a decision tree **over the same classification problem**, previously chosen in question 1. Present the resulting tree (dimension, rule-based description) and compare its performance with the results of question 1.
- 3) Pick a library/toolbox at your own choice, for an experiment testing different kernel functions and parameter setups in order to obtain the best SVM classifier for a **two-class classification task**, whose data is chosen by you. Don't forget to discuss the results.