Machine Learning
Artur d'Avila Garcez
7th Tutorial – Random Forests

Download and unpack the RandomForestCodeImplementation.zip file from Moodle. This file contains an implementation of random forests, i.e. a collection of decision trees, for regression and classification.

To test the Classification Forest run: mainTreeClfUsage.m

(Remember to add subfolder *lib* to the *path*)

To test the Regression Forest run: mainTreeUsage.m

Change the number of trees in the forest (mainTreeUsage.m calls a number of other .m files. Find out in which file the number of trees can be changed)

Change the maximum depth of the trees

Change the size of the forest

Inspect the code. What is the main difference between the two implementations for regression and classification?

Challenge: Calculate the mean and variance of each leaf node. Plot a sum of Gaussians which can be a measure of the forest's final regression result.