INSTITUT MONTEFIORE
Département Electricité, Electronique et Informatique
Méthodes stochastiques
Prof. Louis WEHENKEL

Liège, June 2005.

PhD Thesis project
'Kernel properties of Random Tree Models'

Random tree models have been extensively developed in the field of Machine Learning in the recent years. Instances of such methods are for example bagging, random subspace, random forests, and extra-trees.
Random trees can be generated efficiently and the combination of large sets of random trees generally leads to accurate models. Under some hypothesis on the random distribution of trees, it is possible to characterize analytically the approximation produced by an infinite ensemble of trees. These models can also be interpreted as Kernel interpolators where the kernel is a mixture of randomized piece-wise constant kernels. Different random distribution of the tree models lead to different geometrical and statistical properties of the resulting models and kernels.

The subject of this thesis project concerns the theoretical study of the properties of various Random Tree Models in order to improve our basic understanding of these methods and create new algorithms with pre-specified properties.

Subject areas: machine learning, random processes, statistics, geometry and functional analysis

References:

1. Zhao, G., `A new perspective on classification'. Ph.D. thesis, Utah State University, Department of Mathematics and Statistics, 2000.
2. Breiman, L., `Some infinity theory for predictor ensembles'. Technical Report 579, University of California, Department of Statistics, 2000.
3. Lin, Y. and Y. Jeon, `Random forests and adaptive nearest neighbors'. Technical Report 1055, University of Wisconsin, Department of Statistics, 2002.
4. Breiman, L., `Consistency for a simple model of random forests', Technical Report 579, University of California, Department of Statistics, 2004
5. P. Geurts, 'Contributions to decision tree induction: bias/variance tradeoff and time series classification', PhD Thesis, University of Liège, Department of Electrical Engineering and Computer Science', 2002.

Sart-Tilman, B28
B – 4000 LIEGE 1 / Belgique

☎ +32–(0)4–366.26.84
Fax +32–(0)4–366.29.84

N° 48, stop Montefiore
Parking 32

e-mail : L.Wehenkel@ulg.ac.be