# Random Decision Forests Tin Kam Ho AT&T Bell Laboratories[1]

But trees derived with traditional methods often cannot be grown to arbitrary complexity for possible loss of generalization accuracy on unseen data.

The essence of the method is to build multiple trees in randomly selected subspaces of the feature space. Trees in different subspaces generalize their classification in complementary ways, and their combined classification can be monotonically improved.

In either method the stopping rule is until all the terminal nodes (leaves) contain points of a single class, or until it is impossible to split further. Since we do not want to lose any accuracy on classifying the training data, we do not consider methods to prune back the tree.

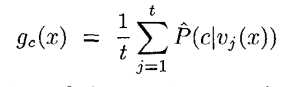
Our method to create multiple trees is to construct trees in randomly selected subspaces of the feature space. For a given feature space of m dimensions, there are 2" subspaces in which a decision tree can be constructed.

A decision tree is constructed in each selected subspace using the entire training set and the algorithms given in the previous section. Notice that each of these trees classifies the training data 100% correctly. Thus each tree generalizes its classification in a different way.

## Discriminant function

For a point 2, let v3(x) be the terminal node that x is assigned to when it descends down tree T3 (j = 1,2, ..., t). Given this, let the posterior probability that x belongs to class c (c = 1,2, ..., n) be denoted by P(ClU3 (x)).

The discriminant function is defined as :



the decision rule is to assign x to class c for which gc(x) is the maximum.

## Single Trees

First we show the results when single trees are constructed in the full feature space, as in the conventional practice. We tested both vectors fi and f2 and both of the tree growing methods.

Two heuristics used & compared: Central axis projection vs Perceptron training

## Random Forest

F1 has 400 components. F2 has 852 components

The subspaces were restricted to 100 or 200 dimensions in the experiments, and the resultant differences in classification accuracy are clear from the figures.

# Shape Quantization and Recognition with Randomized Trees Yali Amit and Donald Geman [2]

* Letter Recognition
* Looks Well Explained

# Random Forests, LEO BREIMAN, Statistics Department, University of California, Berkeley[3]

* Long and mathematical

# Letter Recognition Using Holland-Style Adaptive Classifiers, Frey & Slate[4]

References

[1] Tin Kam Ho, “Random decision forests,” in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, Que., Canada, 1995, vol. 1, pp. 278–282, doi: 10.1109/ICDAR.1995.598994.

[2] Y. Amit and D. Geman, “Shape Quantization and Recognition with Randomized Trees,” *Neural Comput.*, vol. 9, no. 7, pp. 1545–1588, Oct. 1997, doi: 10.1162/neco.1997.9.7.1545.

[3] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[4] P. W. Frey and D. J. Slate, “Letter recognition using Holland-style adaptive classifiers,” *Mach. Learn.*, vol. 6, no. 2, pp. 161–182, Mar. 1991, doi: 10.1007/BF00114162.