

**Biau, Gérard**

**Analysis of a random forests model.** (English) Zbl 1283.62127  
*J. Mach. Learn. Res.* 13, 1063-1095 (2012).

Summary: Random forests are a scheme proposed by Leo Breiman in the 2000's for building a predictor ensemble with a set of decision trees that grow in randomly selected subspaces of data. Despite growing interest and practical use, there has been little exploration of the statistical properties of random forests, and little is known about the mathematical forces driving the algorithm. In this paper, we offer an in-depth analysis of a random forests model suggested by *L. Breiman* [see, e.g., *Ann. Stat.* 32, No. 1, 1–11 (2004; [Zbl 1105.62308](#))], which is very close to the original algorithm. We show in particular that the procedure is consistent and adapts to sparsity, in the sense that its rate of convergence depends only on the number of strong features and not on how many noise variables are present.

**MSC:**

[62H30](#) Classification and discrimination; cluster analysis (statistical aspects)  
[62G09](#) Nonparametric statistical resampling methods

Cited in **1** Review  
Cited in **36** Documents

**Keywords:**

[random forests](#); [randomization](#); [sparsity](#); [dimension reduction](#); [consistency](#); [rate of convergence](#)

**Full Text:** [Link](#)