

[Lin, Yi; Jeon, Yongho](#)

Random forests and adaptive nearest neighbors. (English) Zbl 1119.62304
J. Am. Stat. Assoc. 101, No. 474, 578-590 (2006).

Summary: We study random forests through their connection with a new framework of adaptive nearest-neighbor methods. We introduce a concept of potential nearest neighbors (k-PNNs) and show that random forests can be viewed as adaptively weighted k -PNN methods. Various aspects of random forests can be studied from this perspective. We study the effect of terminal node sizes on the prediction accuracy of random forests. We further show that random forests with adaptive splitting schemes assign weights to k-PNNs in a desirable way: for the estimation at a given target point, these random forests assign voting weights to the k-PNNs of the target point according to the local importance of different input variables. We propose a new simple splitting scheme that achieves desirable adaptivity in a straightforward fashion. This simple scheme can be combined with existing algorithms. The resulting algorithm is computationally faster and gives comparable results. Other possible aspects of random forests, such as using linear combinations in splitting, are also discussed. Simulations and real datasets are used to illustrate the results.

MSC:

62-XX Statistics
65C60 Computational problems in statistics (MSC2010)

Cited in **35** Documents

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