Mathematics > Probability

The Convergence Rate of Majority Vote under Exchangeability

Miles E. Lopes

(Submitted on 4 Mar 2013)

Majority vote plays a fundamental role in many applications of statistics, such as ensemble classifiers, crowdsourcing, and elections. When using majority vote as a prediction rule, it is of basic interest to ask "How many votes are needed to obtain a reliable prediction?" In the context of binary classification with Random Forests or Bagging, we give a precise answer: If err_t denotes the test error achieved by the majority vote of t \geq 1 classifiers, and err* denotes its nominal limiting value, then under basic regularity conditions, err_t = err* + c/t + o(1/t), where c is a constant given by a simple formula. More generally, we show that if V_1,V_2,... is an exchangeable Bernoulli sequence with mixture distribution F, and the majority vote is written as M_t=median (V_1,...,V_t), then 1-\E[M_t] = F(1/2)+ (F"(1/2)/8)(1/t)+o(1/t) when F is sufficiently smooth.

Comments:	11 pages
Subjects:	**Probability (math.PR)**; Social and Information Networks (cs.SI); Statistics Theory (math.ST); Machine Learning (stat.ML)
Cite as:	**arXiv:1303.0727 [math.PR]**
	(or **arXiv:1303.0727v1 [math.PR]** for this version)

Submission history

From: Miles Lopes [view email]
[v1] Mon, 4 Mar 2013 15:16:00 GMT (404kb)

Which authors of this paper are endorsers?