Minimax-Optimal Bounds for Detectors Based on Estimated Prior Probabilities

Jiantao Jiao, Lin Zhang, Robert Nowak

(Submitted on 29 Jul 2011 (v1), last revised 28 Feb 2012 (this version, v2))

In many signal detection and classification problems, we have knowledge of the distribution under each hypothesis, but not the prior probabilities. This paper is aimed at providing theory to quantify the performance of detection via estimating prior probabilities from either labeled or unlabeled training data. The error or {\em risk} is considered as a function of the prior probabilities. We show that the risk function is locally Lipschitz in the vicinity of the true prior probabilities, and the error of detectors based on estimated prior probabilities depends on the behavior of the risk function in this locality. In general, we show that the error of detectors based on the Maximum Likelihood Estimate (MLE) of the prior probabilities converges to the Bayes error at a rate of $n^{-1/2}$, where n is the number of training data. If the behavior of the risk function is more favorable, then detectors based on the MLE have errors converging to the corresponding Bayes errors at optimal rates of the form $n^{-(1+\alpha)/2}$, where $\alpha>0$ is a parameter governing the behavior of the risk function with a typical value $\alpha = 1$. The limit $\alpha \rightarrow \infty$ corresponds to a situation where the risk function is flat near the true probabilities, and thus insensitive to small errors in the MLE; in this case the error of the detector based on the MLE converges to the Bayes error exponentially fast with n. We show the bounds are achievable no matter given labeled or unlabeled training data and are minimax-optimal in labeled case.

Comments:	Submitted to IEEE Transactions on Information Theory
Subjects:	**Information Theory (cs.IT)**; Machine Learning (stat.ML)
Cite as:	**arXiv:1107.6027 [cs.IT]**
	(or **arXiv:1107.6027v2 [cs.IT]** for this version)

Submission history

From: Jiantao Jiao [view email]

[v1] Fri, 29 Jul 2011 17:33:59 GMT (178kb)

[v2] Tue, 28 Feb 2012 01:13:43 GMT (179kb)