The Annals of Statistics
2006, Vol. 34, No. 6, 2672–2676
DOI: 10.1214/009053606000001046
Main article DOI: 10.1214/009053606000001019

DISCUSSION: LOCAL RADEMACHER COMPLEXITIES AND ORACLE INEQUALITIES IN RISK MINIMIZATION

BY STÉPHAN CLÉMENÇON, GÁBOR LUGOSI[1] AND NICOLAS VAYATIS

Universités Paris 6 and Paris 10, ICREA and Universitat Pompeu Fabra and Université Paris 6

In this magnificent paper, Professor Koltchinskii offers general and powerful performance bounds for empirical risk minimization, a fundamental principle of statistical learning theory. Since the elegant pioneering work of Vapnik and Chervonenkis in the early 1970s, various such bounds have been known that relate the performance of empirical risk minimizers to combinatorial and geometrical features of the class over which the minimization is performed. This area of research has been a rich source of motivation and a major field of applications of empirical process theory.

The appearance of advanced concentration inequalities in the 1990s, primarily thanks to Talagrand's influential work, provoked major advances in both empirical process theory and statistical learning theory and led to a much deeper understanding of some of the basic phenomena. In the discussed paper Professor Koltchinskii develops a powerful new methodology, *iterative localization*, which, with the help of concentration inequalities, is able to explain most of the recent results and go significantly beyond them in many cases.

The main motivation behind Professor Koltchinskii's paper is based on classical problems of statistical learning theory such as binary classification and regression in which, given a sample (X_i, Y_i), $i = 1, \dots, n$, of independent and identically distributed pairs of random variables (where the X_i take their values in some feature space $\mathcal{X}$ and the Y_i are, say, real-valued), the goal is to find a function $f : \mathcal{X} \to \mathbb{R}$ whose risk, defined in terms of the expected value of an appropriately chosen loss function, is as small as possible.

In the remaining part of this discussion we point out how the performance bounds of Professor Koltchinskii's paper can be used to study a seemingly different model, motivated by nonparametric *ranking* problems, which has received increasing attention both in the statistical and machine learning literature. Indeed, in several applications, such as the search engine problem or credit risk screening, the goal is to learn how to rank—or to score—observations rather than just classify them. In this case, performance measures involve pairs of observations, as can be seen, for instance, with the AUC (Area Under an ROC Curve) criterion. In this

Received March 2006.

[1]Supported by the Spanish Ministry of Science and Technology and FEDER Grant BMF2003-03324 and by the PASCAL Network of Excellence under EC Grant 506778.