ERRATA for paper "Minimax Bounds for Active Learning" - In COLT 2007

Rui M. Castro and Robert D. Nowak

It is virtually impossible to get anything exactly right — Carl de Boor, University of Wisconsin

It was noted by the authors that Theorem 3 in [2] is not entirely correct, and it is not valid in all the generality it was stated. This glitch does not affect the conclusions in [2], but it is important to issue a correction. The following theorem is adapted from [1] (page 85, theorem 2.5).

Theorem (Tsybakov, 2004)**.** *Let $\mathcal{F} \subseteq \Xi$ be a class of models. Associated with each model $f \in \mathcal{F}$ we have a probability measure P_f defined on a common probability space. Let $M \geq 2$ be an integer and let $d_\Delta(\cdot,\cdot) : \Xi \times \Xi \to \mathbb{R}$ be a semi-distance. Suppose we have $\{f_0, \ldots, f_M\} \in \mathcal{F}$ such that*

i) $d_\Delta(f_j, f_k) \geq 2\,a > 0, \quad \forall_{0 \leq j,k \leq M}$,

ii) $P_{f_0} \ll P_{f_j}, \quad \forall_{j=1,\ldots,M}$,

iii) $\frac{1}{M}\sum_{j=1}^{M} KL(P_{f_j} \| P_{f_0}) \leq \gamma \log M$,

where $0 < \gamma < 1/8$. The following bound holds.

$$\inf_{\widehat{f}} \sup_{f \in \mathcal{F}} P_f\left(d_\Delta(\widehat{f}, f) \geq a\right) \geq \frac{\sqrt{M}}{1+\sqrt{M}}\left(1 - 2\gamma - 2\sqrt{\frac{\gamma}{\log M}}\right) > 0 \ ,$$

where the infimum is taken with respect to the collection of all possible estimators of f (based on a sample from P_f), and KL denotes the Kullback-Leibler divergence.

Furthermore if we have $d(\cdot,\cdot) : \Xi \times \Xi \to \mathbb{R}$ such that for all $f \in \Xi$ and $i \in \{0, \ldots, M\}$

$$d(f, f_i) \geq h(d_\Delta(f, f_i)),$$

where $h : \mathbb{R} \to \mathbb{R}$ is a monotonically non-decreasing function, then

$$\inf_{\widehat{f}} \sup_{f \in \mathcal{F}} P_f\left(d(\widehat{f}, f) \geq h(a)\right) \geq \frac{\sqrt{M}}{1+\sqrt{M}}\left(1 - 2\gamma - 2\sqrt{\frac{\gamma}{\log M}}\right) > 0 \ .$$

The revised proof of Theorem 1 in [2] proceeds by showing that for the collection of distributions $P_{\boldsymbol{X}Y}$ constructed we have

$$d_\Delta(G_i^*, G_j^*) \geq L\|h\|_1\, m^{-\alpha} \ ,$$

and

$$d(G, G_i^*) \triangleq R_i(G) - R_i(G_i^*) \geq \frac{4c}{\kappa 2^\kappa} d_\Delta^\kappa(G, G_i^*) \ ,$$

for any set $G \subseteq [0,1]^d$. Applying the above stated theorem yields the final result of Theorem 1 in [2]. The detailed proof is available in technical report [3] (see Appendix B - "Proof of Theorem 3").

References

[1] Alexander B. Tsybakov. *Introduction à l'estimation non-paramétrique.* Mathématiques et Applications, 41. Springer, 2004.

[2] Rui M. Castro and Robert D. Nowak. Minimax bounds for active learning. In *Twentieth Annual Conference on Learning Theory (COLT)*, 2007.

[3] Rui M. Castro and Robert D. Nowak Minimax bounds for active learning. Technical report, ECE Dept., University of Wisconsin - Madison (2007) (available at http://homepages.cae.wisc.edu/∼rcastro/ECE-07-3.pdf).