Erratum for Average-Case Active Learning with Costs

Andrew Guillory
Computer Science and Engineering
University of Washington
guillory@cs.washington.edu

Jeff Bilmes
Electrical Engineering
University of Washington
bilmes@ee.washington.edu

March 28, 2015

Our publication Average-Case Active Learning with Costs [1] as well as the extended tech report of the same name [2] contain an error in the proof of Theorem 3. This erratum gives a correction for the proof. Thank you to Feng Nan for bringing this error to our attention.

The error is at the end of page 9 of the paper (page 7 of the technical report). We have the inequality

$$C(T_g^S, \pi_S) \leq c_i + 12C^*(S) \ln \frac{\pi(S)}{\min(S)} - 12C^*(S)(1 - \pi_S(S^+))$$

The paper then states that because $\pi_S(S^+) \geq \sum_{j \in A} \pi_S(S^j)^2$ we have

$$C(T_g^S, \pi_S) \leq c_i + 12C^*(S) \ln \frac{\pi(S)}{\min(S)} - 12C^*(S)(1 - \sum_{j \in A} \pi_S(S^j)^2)$$

This is incorrect. In order for the inequality to follow we would need to instead have $\pi_S(S^+) \leq \sum_{j \in A} \pi_S(S^j)^2$ which is not the case.

The correction to the proof changes only the term involving $\pi_S(S^+)$ so for clarity we omit the other terms. Starting on the fourth formula line above the bottom of page 9, before applying the $\ln(1 - x) < -x$ bound, we substitute

$$\ln \pi_S(S^+) = \frac{\ln(\pi_S(S^+)^2)}{2}$$

Now applying $\ln(1 - x) < -x$ bound as before we have

$$\frac{\ln(\pi_S(S^+)^2)}{2} \leq \frac{-(1 - \pi_S(S^+)^2)}{2}$$
Now we substitute
\[
\frac{-(1 - \pi_S(S+)^2)}{2} \leq \frac{-(1 - \sum_j \pi_S(S_j)^2)}{2}
\]
because \(\pi_S(S+)^2 \leq \sum_j \pi_S(S_j)^2\) (we are adding more positive numbers to the sum). The rest of the proof proceeds as before except with an extra factor of \(1/2\) not in the original proof. This factor doesn’t affect the final bound as it is lost on the third formula line of page 10 where we have
\[
c_i - 6(1 - \text{CP}(\pi_S))c_i \leq 0
\]
(in the original proof there was a factor of 12 there).

References
