Corrections to "Worst-Case Analysis of Rule Discovery"

Einoshin Suzuki

Electrical and Computer Engineering, Yokohama National University,
79-5 Tokiwadai, Hodogaya, Yokohama 240-8501, Japan
suzuki@dnj.ynu.ac.jp

In the article cited above, which will appear in the Proceedings of Discovery Science 2001, due to an error, the part from "Pr(r_{g} undiscovered) in section 4 to the end of section 4 should be replaced by the following.

$$
\begin{aligned}
&\Pr\ (r_{\mathrm{g}} \text{ undiscovered}) \\
&\leq \mathrm{MAX}\left[\sum_{k=0}^{\lceil m\theta_{\mathrm{S}}\rceil-1} \mathrm{B}(k;m,\Pr(y)),\ \sum_{k=0}^{\lceil m\widehat{\Pr}(y)\theta_{\mathrm{F}}\rceil-1} \mathrm{B}(k;m\widehat{\Pr}(y),\Pr(x|y))\right] && (23)\\
&= \mathrm{MAX}\left[\sum_{k=\phi(m,\theta_{\mathrm{S}})}^{m} \mathrm{B}(k;m,1-\Pr(y)),\right. \\
&\qquad \left.\sum_{k=\phi(m\widehat{\Pr}(y),\theta_{\mathrm{F}})}^{m\widehat{\Pr}(y)} \mathrm{B}(k;m\widehat{\Pr}(y),1-\Pr(x|y))\right] && (24)\\
&< \mathrm{MAX}\left\{\exp\left[-2m\left(\frac{\phi(m,\theta_{\mathrm{S}})}{m}-1+\Pr(y)\right)^2\right],\right. \\
&\qquad \left.\exp\left[-2m\widehat{\Pr}(y)\left(\frac{\phi(m\widehat{\Pr}(y),\theta_{\mathrm{F}})}{m\widehat{\Pr}(y)}-1+\Pr(x|y)\right)^2\right]\right\} && (25)\\
&< \mathrm{MAX}\left\{\exp\left[-2m(-\theta_{\mathrm{S}}+1-\zeta)^2\right],\exp\left[-2m\theta_{\mathrm{S}}(-\theta_{\mathrm{F}}+1-\epsilon)^2\right]\right\}. && (26)
\end{aligned}
$$

where $\phi(m,\theta) \equiv m - \lceil m\theta \rceil + 1$.

Note that we consider separately the cases in which the generality and the accuracy of a good rule are below the respective thresholds in (23). Since their probabilities are unknown, we use the same technique as in subsection 3.2. Note that (24) corresponds to replacement of p by $1-p$ in (5). In (25), the Chernoff bound (6) is employed from (22). Finally in (26), we employ (20) and $\widehat{\Pr}(y) \geq \theta_{\mathrm{S}}$. Note that the last inequality holds in the second term since r_{g} is undiscovered due to apparently low accuracy.

Similarly to subsection 3.2, the following can be obtained as a the number m of examples for discovery in which overlooking a good rule is avoided with a high probability.

$$m \geq \frac{\ln\left(\frac{|R|}{\delta}\right)}{2\mathrm{MIN}\left[\left(-\theta_{\mathrm{S}}+1-\zeta\right)^{2},\ \theta_{\mathrm{S}}\left(-\theta_{\mathrm{F}}+1-\epsilon\right)^{2}\right]} \tag{27}$$

Note that (27) is equivalent to (17), and similar discussions as subsections 3.2 and 3.3 hold. Note that large margins ($1-\zeta-\theta_{\mathrm{S}}$ and $1-\epsilon-\theta_{\mathrm{F}}$ in this case) represent small thresholds in this case, and small thresholds typically result in a large number of candidates of the discovered rule to be inspected. The automatic adjustment of thresholds [11] can be also a realistic measure for this problem.

References

1. E. Suzuki: "Worst-Case Analysis of Rule Discovery", *Proc. Fourth Int'l Conf. of Discovery Science (DS)*, Springer (2001).