Algorithmic Stability and Generalization Performance

Olivier Bousquet and Andrej Elisseeff

Or

PAC based on Stability

Standard PAC

Definitions:

1. D = example distribution

2. $l(h(x), y) \in [0, 1]$ = loss function

3. $e_D(h)$ = true error = $E_D l(h(x), y)$

4. $e_S(h)$ = empirical error = $\frac{1}{|S|} \sum_i l(h(x_i), y_i)$

Hypothesis complexity PAC bound

$$\Pr_D(\exists h \in H : e_D(h) - e_S(h) > \epsilon) \leq |H| e^{-2|S|\epsilon^2}$$

Standard PAC bound useful if:

$$\ln |H| \leq \hat{O}(|S|\epsilon^2)$$

Stability based PAC

1. h_S = hypothesis learned on set S

2. $S_{\bar{i}} = S$ with ith example removed

3. β-stability $= |l(h_S(x), y) - l(h_{S_i}(x), y)| \leq \beta$

Stability based PAC bounds:

$$\Pr_D(e_D(h_S) - e_S(h_S) > \epsilon) \leq \frac{64\beta}{\epsilon^2} + \frac{8}{m\epsilon^2}$$

$$\Pr_D(e_D(h_S) - e_S(h_S) > \epsilon) \leq 2e^{-\frac{m\epsilon^2}{2(m\beta+1)^2}}$$

Stability based PAC bounds useful if:

$$\beta < \hat{O}(\epsilon^2)$$