University of Salzburg | Lecturer: Roland Kwitt

Machine Learning (911.236)

Exercise sheet **C**

Exercise 1. 3 P.

Use the Höffding inequality to show that with probability of at least $1-\delta$ (over the choice of S of size m), we have for a $h \in \mathcal{H}$

$$L_D(h) \leq L_S(h) + \sqrt{\frac{1}{2m}\log\left(\frac{2}{\delta}\right)}\ ,$$

where $L_D(h)$ is the generalization error of the hypothesis h and $L_S(h)$ is the empirical error of h on the training set $S = ((x_1, y_1), \ldots, (x_m, y_m))$. Remember that $L_S(h_S) = 1/m \sum_i \mathbf{1}_{h(x_i)\neq y_i}$. **Hint**: Use the right form of the Höffding inequality we had in the lecture.

Exercise 2. 3 P.

Say we have a finite hypothesis class $\mathcal{H}$, i.e., $|\mathcal{H}| < \infty$ and let $h_S = \arg\min_{h\in\mathcal{H}} L_S(h)$ be the hypothesis with the smallest empirical error on the training set S (i.e., an empirical risk minimizer). Show that with probability of at least $1-\delta$ (over the choice of S of size m), we have

$$L_D(h_S) \leq L_S(h_S) + \sqrt{\frac{1}{2m}\log\left(\frac{2|\mathcal{H}|}{\delta}\right)}\ .$$

Total #points: 6 P.