University of Salzburg — Lecturer: Roland Kwitt

# **Machine Learning** (911.236)

Exercise sheet **E**

**Exercise 1.** 4 P.

Lets assume that in each iteration of AdaBoost, running the weak learner **always** returns a hypothesis with error less than 1/2.

In the $t$-th iteration, the weak learner returns a hypothesis $h_t$. Show, by means of contradiction, that in the $(t+1)$-th iteration the same hypothesis (i.e., $h_t$) cannot be chosen.

*Hint*: In the AdaBoost algorithm, we have written down the expression to compute the error $\epsilon_t$. Use this, in combination with $\alpha_t$ and $Z_t$ to show that choosing $h_t$ would violate the weak learning assumption from above.

**Exercise 2.** 4 P.

Lets assume we do not want an iteration dependent $\alpha_t$, but a fixed one, i.e., just $\alpha$ and let

$$0 < \gamma \leq \frac{1}{2} - \epsilon_t \quad .$$

Take the upper-bound for the empirical error for a hypothesis returned by AdaBoost, i.e.,

$$L_S(h) \leq \prod_{t=1}^{T} Z_t \quad ,$$

and derive the *best* value for $\alpha$. *Hint*: First, formulate an inequality of the form $Z_t \leq r(\alpha)$ and analyze when $r(\alpha)$ is minimized (i.e., taking the derivative, setting it to zero, etc.).

**Exercise 3.** 3 P.

Take the upper-bound for the empirical error of a hypothesis $h$ returned by AdaBoost (i.e., the specical case mentioned in the lecture), i.e.,

$$L_S(h) \leq e^{-2\gamma^2 T}$$

and show that for

$$T > \frac{\log(m)}{2\gamma^2}$$

the hypothesis that is returned is consistent (meaning zero empirical error) for sample size $m$.

**Total #points:** 11 P.