University of Salzburg — Lecturer: Roland Kwitt

Machine Learning (911.236)

Exercise sheet **B**

Exercise 1. 3 P.

Say a company runs a web service where the response time X (measured in seconds) of a certain API call is a non-negative random variable with the following probability mass function (PMF):

Response time x (in [s])	0	1	2	4	8
Probability $\mathbb{P}(X = x)$	0.1	0.3	0.3	0.2	0.1

Compute the expected value $\mathbb{E}[X]$ and then use Markov's inequality to upper bound $\mathbb{P}(X \geq 5)$. Also, compute the exact value of $\mathbb{P}(X \geq 5)$ from the PMF and compare your results. What does this tell you about the Markov inequality?

Exercise 2. 3 P.

Say you have a learning problem with $n < \infty$ input variables $v_1, \ldots, v_n$ that can only take on binary values, that is $v_i \in \{0, 1\}$; hence $\mathcal{X} = \{0, 1\}^n$. Also, our label set is $\mathcal{Y} = \{0, 1\}$. As an example, take $n = 3$, then $\mathcal{X} = \{0, 1\}^3$ and we have

$$\mathcal{X} = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)\} \ ,$$

and a possible training set S (with $m = 5$ instances) could look like

$$S = ((0, 0, 0), 0), ((0, 1, 0), 1), ((0, 1, 1), 1), ((1, 0, 1), 1), ((1, 1, 1), 0) \ .$$

We consider the hypothesis class $\mathcal{H}$ of *all* boolean functions from $\mathcal{X} \to \mathcal{Y}$. Use our results from the lecture (PAC learnability of finite hypothesis classes) and argue what the problem might be in this example, especially in terms of the number of required samples for any given choice of $\epsilon, \delta \in (0, 1)$.

Total #points: 6 P.