Machine Learning II: Assignment
14 performance points (max),
email: PDF+code to jan.nagler@gmail.com
due: Tue, October 22, 2019

1. Use and modify and download the expectation-maximization (EM) Python program for two coins, as developed in the lecture. Generate an *unrepresentative* series of exactly $n = 5 \times 25$ total coin flips (5 times 25 flips with the randomly selected coin), given two coins thrown with equal probability (1/2), but with different heads probabilities, p_A (coin A), and p_B (coin B), respectively. Choose and fix a single parameter combination within $0.1 < p_A < 0.9$ and $0.1 < p_B < 0.9$. This means to generate a series of (H)eads and (T)ails that is virtually incompatible, i.e. highly unlikely, given the *ground truth* $\theta = (p_A, p_B)$ of your choice, yet being a valid realization (instance) of the underlying fair double-coin process. Once this highly unlikely (say, unlucky) realization is found and generated, analyze this given instance with the EM algorithm. The EM steps will show convergence to some (MLE) estimates of p'_A and p'_B, which best represent the unlikely dataset but deviate substantially from p_A and p_B.

   The solution with the largest value of

$$\text{score} = \min[abs(\log(p'_A/p_A)), abs(\log(p'_B/p_B))]$$

   (that you need to compute and print) wins a price, handed over by the lecturer – but only if this value is unique among the submissions. If it is not, the 2nd largest score wins, if unique, and so on. If there is no winner, the present may, sadly, be thrown out of one randomly selected window.

   Good unluck!