Machine Learning, IUB Fall 2003, Final Exam: Solutions (fragments)

Remark: *this is not a full-fledged solution sheet, because many tasks were "open" and required discussions. Only two technical, selected questions are sketchily answered here.*

Problem 2: b (i): D is the training data. (ii): $p(\theta^c)$ is the prior, a metadistribution in an 40-dim measure space. (iii) $p(D|\theta^c)$ conditional probability density of (training) data given true distribution parameters θ^c. For given θ^c, this is a numerical value. (iv) posterior distribution of distribution parameters, given training data (and prior info). Also 40-dim.

d: Let A_c abbreviate the event "test pattern is in class c". Then the required version of Bayes formula is $P(A_c\,|\mathbf{x}) = P(\mathbf{x}|\,A_c)\,P_c\,/\,(\Sigma_c\,P(\mathbf{x}|\,A_c)\,P_c)$, and
$P(\mathbf{x}\,|\,A_c) = \Pi_{\,k\,=\,1,\,...,\,20}\;\mathcal{N}(\mu^{c\ \mathrm{PME}}_{k},\,\sigma^{c\ \mathrm{PME}}_{k})\,(x_k)$, where $\mathcal{N}(\mu^{c\ \mathrm{PME}}_{k},\,\sigma^{c\ \mathrm{PME}}_{k})$ is the normal distribution with parameters ${}^{c\ \mathrm{PME}}_{k}$, $\sigma^{c\ \mathrm{PME}}_{k}$.

Problem 3: Optimal weights are $w_1 = 2$ and $w_2 = 3$; an easy guess that can be justified by principle of orthogonality (PO) as follows. PO states that (only) for optimal weights, the error is uncorrelated with tap inputs. The error signal here is $y(n) - x(n) = - u(n-2) + v(n)$, which is uncorrelated with tap inputs $u(n)$ and $u(n-1)$ because u is iid. The residual error is the sum of two uncorrelated errors $u(n-2)$ and $v(n)$, whose variances add to $\xi_{\min} = 1.1$.

Compute optimal weights for example p. 88 my old script.