TU Delft Delft University of Technology

Exercise Sheet - Machine learning

First two exercises are obtained from Russell & Norvig

Exercise 1. Consider the noisy-OR model for fever described in Russell & Norvig Section 14.3. Explain how to apply maximum-likelihood learning to fit the parameters of such a model to a set of complete data. (Hint: use the chain rule for partial derivatives.)

Exercise 2. Consider the application of EM to learn the parameters for the network in Figure 20.13(a), given the true parameters in Equation (20.7):

a) Explain why the EM algorithm would not work if there were just two attributes in the model rather than three.

b) Show the calculations for the first iteration of EM starting from Equation (20.8).

Bonus (hard)

Exercise 3. Look at the EM transition parameter update:

$$\theta_{y\to v}^{trans(k+1)} = \frac{\sum_t Q^{(k+1)}(S_{t-1}=y, S_t=v)}{\sum_t Q^{(k+1)}(S_{t-1}=y)}$$

and its derivation (in the lecture slides, see also https://www.worldscientific.com/doi/pdfplus/10.1142/S0218001401000836 (An introduction to hidden Markov models and Bayesian networks of Z Ghahramani), but note the notation is quite dissimilar from R&N).

For the observation probabilities, the update rule is:

$$\theta_{y\to w}^{obs} = \frac{\sum_{t \text{ s.t. } o_t=w} Q^{(k+1)}(S_t=y)}{\sum_t Q^{(k+1)}(S_t=y)}.$$

Derive this update step.