Foundations of Artificial Intelligence

Prof. Dr. B. Nebel, Prof. Dr. M. Riedmiller
S. Lange, J. Witkowski, D. Zhang
Summer Term 2010

University of Freiburg
Department of Computer Science

Exercise Sheet 7

Due: Tuesday, June 22, 2010

Exercise 7.1 (Programming Assignment Statistical Learning)
Consider the 'candy manufacturer problem' from the lecture: A candy manufacturer sells five kinds of bags that are indistinguishable from the outside:

h_1**:** 100% cherry

h_2**:** 75% cherry and 25% lime

h_3**:** 50% cherry and 50% lime

h_4**:** 25% cherry and 75% lime

h_5**:** 100% lime

The prior distribution of $h_1, h_2, \ldots, h_5$ is given by $< 0.1, 0.2, 0.4, 0.2, 0.1 >$. Given a sequence C, C, L, L, L, L, L, L, C, L of 'observations' d_i drawn from a single bag of unknown type, with C=cherry and L=lime, solve the following assignments assuming an infinite bag size. You can either do the necessary calculations by hand—then explaining them in detail—or by writing a program that generates the necessary output (strongly preferred). In the latter case, please hand in the source code of your program / script / excel sheet AND your plots and answers—any programming language will be accepted, but don't expect your tutor to debug and / or run your code.

(a) Calculate the probability $P(h|d_i)$ of each hypothesis $h_1, h_2, \ldots, h_5$ after observing the first $i = 1, 2, \ldots, 10$ elements of the sequence and produce a plot similar to the plot depicted on the left of slide 10/8 of the lecture.

(b) Calculate and plot the 'bayes prediction' as well as the MAP prediction after observing the first $i = 1, 2, \ldots, 10$ elements of the sequence (see plot on the right of slide 10/8).

(c) Now repeat the whole procedure in order to produce predictions according to the maximum likelihood hypothesis. Plot the predictions (similar to the plot from (b)) as well as the values of $P(h|d_i)$ (similar to the plot from (a)) as estimated using the maximum-likelihood approach.

(d) Briefly discuss the different results of the bayes prediction, MAP prediction and prediction according to the maximum-likelihood hypothesis. How would these three predictions develop in the limit, e.g. when observing the above sequence for an infinite number of times?

(e) How would the predictions differ *after* observing a random permutation of the *complete* sequence—thus observing the same number of three cherries and seven limes in a different succession?

The exercise sheets may and should be handed in and be worked on in groups of three (3) students. Please fill the cover sheet[1] and attach it to your solution.

[1] `http://www.informatik.uni-freiburg.de/~ki/teaching/ss10/gki/coverSheet-english.pdf`