2010 Paper 8 Question 2

Artificial Intelligence II

Consider the following learning problem in which we wish to classify inputs, each consisting of a single real number, into one of two possible classes C_1 and C_2. There are three potential hypotheses where $\Pr(h_1) = 3/10$, $\Pr(h_2) = 5/10$ and $\Pr(h_3) = 2/10$. The hypotheses are the following functions

$$h_i(x) = x - \frac{i-1}{5}$$

and the likelihood for any hypothesis h_i is

$$\Pr(x \in C_1 | h_i, x) = \sigma(h_i(x))$$

where $\sigma(y) = 1/(1+\exp(-y))$. You have seen three examples: $(0.9, C_1)$, $(0.95, C_2)$ and $(1.3, C_2)$, and you now wish to classify the new point $x = 1.1$.

(a) Explain how in general the *maximum a posteriori (MAP)* classifier works. [3 marks]

(b) Compute the class that the MAP classifier would predict in this case. [10 marks]

(c) The preferred alternative to the MAP classifier is the Bayesian classifier, computing $\Pr(x \in C_1 | x, \mathbf{s})$. where $\mathbf{s}$ is the vector of examples. Show that

$$\Pr(x \in C_1 | x, \mathbf{s}) = \sum_{h_i} \Pr(x \in C_1 | h_i, x) \Pr(h_i | \mathbf{s})$$

What are you assuming about independence in deriving this result? [3 marks]

(d) Compute the class that the Bayesian classifier would predict in this case. [4 marks]