6.825 Recitation Problems: Lec. 18

Solutions

December 12, 2001

Let n be the number of data points (in this example n is 8). Let j be the index of the data point with the hidden value. Let D_i be the ith data point.

$$
\begin{aligned}
logPr(D|M) &= log(Pr(D, H = 0|M) + Pr(D, H = 1|M)) \\
&= log(Pr(D|H = 0, M) * Pr(H = 0|M) + Pr(D|H = 1, M) * Pr(H = 1|M)) \\
&= log((\Pi_{i, i \neq j}^{n} Pr(D_i|M)) * Pr(D_j|H = 0, M) * Pr(H = 0|M) + (\Pi_{i, i \neq j}^{n} Pr(D_i|M) * Pr(D_j|H = 0, M) * Pr(H = \\
&= log((\Pi_{i, i \neq j}^{n} Pr(D_i|M)) * (Pr(D_j|H = 0, M) * Pr(H = 0|M) + Pr(D_j|H = 0, M) * Pr(H = 1|M))) \\
&= log((\Pi_{i, i \neq j}^{n} Pr(D_i|M)) * (Pr(D_j, H = 0|M) + Pr(D_j, H = 0|M))) \\
&= \Sigma_{i, i \neq j}^{n} log(Pr(D_i|M)) + log(Pr(D_j, H = 0|M) + Pr(D_j, H = 0|M)) \\
&= 3 log 0.429 + 2 log 0.143 + 2 log 0.285 + log(0.429 + 0.143)
\end{aligned}
$$

To do this, we had to make one assumption, and that is, for all data points D_i, $i \neq j$, D_i and H are conditionally independent given M. (So we went from $Pr(D_i|H = 0, M)$ to $Pr(D_i|M)$.) This makes sense and makes the calculation simpler.