Problem 1. 3.5

Propose a three-level cascade where when one level rejects, the next one is used as in equation 3.10. How can we fix the λ on different levels?

$$\lambda_{ik} = \begin{cases} 0 & i = k \\ \lambda & i = K+1 \\ 1 & otherwise \end{cases} \tag{1}$$

Solution

In this example, there is some loss imposed at each tier of rejection. In the event that there are three cascading classes, C_1, C_2, C_3, the following outcomes are possible:

(a) The item is of class C_1. Loss is zero.

(b) The item is rejected from C_1 (incurring loss λ_1) and is of class C_2. Loss is λ_1.

(c) The item is rejected from C_1 and C_2, incurring a loss of $\lambda_1 + \lambda_2$

If $\lambda_1 + \lambda_2 \geq 1$ then there will be a preference to miscategorize into C_2 over correctly categorizing in C_3. Similarly, if $\lambda_1 > 1$ then there will be a preference to miscategorize into C_1 over correctly categorizing into C_2 or C_3. The values of λ must be set such that items that are correctly classified incur a loss less than if they are incorrectly classified.

Problem 2. 3.9

Show that as we move an item from the consequent to the antecedent, confidence can never increase: confidence(ABC D) confidence(AB CD).

Solution

$$
\begin{aligned}
confidence(A, B, C \to D) &\geq confidence(A, B \to C, D) \\
P(D|A, B, C) &\geq P(C, D|A, B) \\
\frac{P(A, B, C, D)}{P(A, B, C)} &\geq \frac{P(A, B, C, D)}{P(A, B)} \\
\frac{1}{P(A, B, C)} &\geq \frac{1}{P(A, B)} \\
\frac{1}{P(C) * P(A, B|C)} &\geq \frac{1}{P(A, B)} \\
\frac{P(C)}{P(C) * P(C|A, B) * P(A, B)} &\geq \frac{1}{P(A, B)} \\
\frac{1}{P(C|A, B)} &\geq 1
\end{aligned}
$$

Since $P(C|A, B)$ is a probability it is bound in the range $[0, 1]$. This is trivially true except in the case when $P(C|A, B) = 0$.

Problem 3. 3.10

Associated with each item sold in basket analysis, if we also have a number indicating how much the customer enjoyed the product, for example, on a scale of 0 to 10, how can you use this extra information to calculate which item to propose to a customer?

Solution

Option 1: A complicated Bayesian inference could be drawn by assigning each product, i a rating value from 0 to 12 and calculating

$$P(r_i|r_{j\neq i})$$

to predict the likely rating for product i. Products with high predicted ratings would be recommended. Ratings of 11 and 12 would correspond to "product not purchased" and "product purchased but not rated".

Option 2: Alternatively, we could treat the rating for each product as a pseudo-continuous distribution and measure cross-corellation between product ratings. This would allow us to recommend products with ratings strongly correlated to products our customer had rated highly in the past.