

Bài tập 1. [Tùy chọn loại bỏ trong bộ phân loại] Trong nhiều bài toán phân loại, người ta có tùy chọn gán  $x$  cho lớp  $j$  hoặc, nếu bạn quá không chắc chắn, chọn từ chối lựa chọn. Nếu chi phí cho việc từ chối nhỏ hơn chi phí cho việc phân loại sai đối tượng, nó có thể là hành động tối ưu. Hãy để một; nghĩa là bạn chọn hành động  $i$ , với  $i = 1: C + 1$ , trong đó  $C$  là số lớp và  $C' + 1$  là hành động từ chối. Đặt  $Y = j$  là trạng thái đúng (nhưng chưa biết) của thiên nhiên. Xác định hàm mất mát như sau:

$$0 \text{ if } i=j \text{ and } i=j \in \{1, \dots, C\}$$

$$\text{Ma } |Y = j| = 4 \text{ A if } i = C + 1 \text{ (1)}$$

Nếu không thì

Nói cách khác, bạn chịu 0 tổn thất nếu phân loại đúng, bạn chịu  $A$ , tổn thất (chi phí) nếu bạn chọn tùy chọn từ chối và bạn phải chịu  $A$ , tổn thất (chi phí) nếu bạn mắc lỗi thay thế (phân loại sai-cation).

1. Chứng minh rằng rủi ro tối thiểu đạt được nếu chúng ta quyết định  $Y = j$  nếu  $p(Y = j|x) > p(Y = k|x)$  với mọi  $k$  (tức là  $j$  là lớp có khả năng xảy ra cao nhất) và nếu  $p(Y = j'') > 1 - \alpha$ : ngược lại quyết định từ chối.

2. Mô tả một cách định tính những gì xảy ra khi  $4./As$  được tăng từ 0 lên 1 (tức là số tương đối

chi phí từ chối tăng lên).