Lab 1: Exercises

ML701: Machine Learning

MBZUAI

August 30, 2023

Problem 1. Suppose $\mathbb{P}[A], \mathbb{P}[A'], \mathbb{P}[B|A]$, and $\mathbb{P}[B|A']$ are known, where A' is a complement of A, i.e., $A' = \mathcal{X} \setminus A$. Find an expression for $\mathbb{P}[A|B]$ in terms of these four probabilities.

Problem 2. Assume the probability of having tuberculosis (TB) is 0.0005, and a test for TB is 99% accurate. What is the probability one has TB if one tests positive for the disease?

Problem 3. [**Coupon Collector's Problem**]
Imagine you're collecting trading cards from cereal boxes. There are n distinct cards to collect, and each cereal box contains one card, randomly chosen with equal probability. The question is: on average, how many cereal boxes do you need to buy in order to collect all n distinct cards?
Hint: Try to use linearity of expectation.

Problem 4. [**Decision Rule Based on Posterior Probabilities and Joint Densities**]
Consider a binary prediction problem where the outcome Y can take values in the set {0,1}. Suppose we're making decisions based on the ratio:

$$\frac{\mathbb{P}[Y=1|X=x]}{\mathbb{P}[Y=0|X=x]}$$

Prove that any decision rule using the above ratio is equivalent to considering the ratio of the joint densities:

$$\frac{p(x, Y=1)}{p(x, Y=0)}$$

Problem 5. [**Optimal Binary Prediction with Symmetric Loss and MAP Rule**]
Consider a binary prediction problem, where the outcome Y can take values in the set {0,1}. Suppose we have a symmetric loss function, i.e.,

$$\text{loss}(y, y') = \text{loss}(y', y)$$

for all y, y' in $\{0, 1\}$.

1. Show that for every such symmetric loss, the optimal binary predictor is equivalent to the Maximum A Posteriori (MAP) rule.
2. Furthermore, if $\mathbb{P}[Y=1] = \mathbb{P}[Y=0] = \frac{1}{2}$, demonstrate that the likelihood estimator yields the same result as the MAP rule.