

SC4000/CZ4041/CE4041: Machine Learning

Lecture 2 Tutorial Question Sets

Question 1: Suppose A , B and C are three variables of binary values (0 or 1). Given the probabilities $P(A = 1, B = 0) = 0.4$, $P(A = 0) = 0.3$, and $P(A = 1, B = 1, C = 1) = 0.1$, compute the following probabilities:

1. $P(B = 1|A = 1)$.
2. $P(C = 0|B = 1, A = 1)$.

Question 2: Suppose that if a person has lung cancer, his/her probability of having gene X is 0.9, and if a person does not have lung cancer, his/her probability of having gene X is 0.2. The probability of a person having lung cancer is 0.01. Now, we know that a patient A has gene X .

1. Use Bayesian decision theory with 0/1 loss to predict whether the patient A has lung cancer or not.
2. Consider that costs of misclassification are different. Assume that the cost for correct decisions is 0, the cost of misclassifying a person who does not have lung cancer to be a patient with lung cancer is 0.007, and the cost of misclassifying a person who has lung cancer to be a healthy person is 1. Please use Bayesian decision theory with the predefined loss to predict whether the patient A has lung cancer or not.