National Institute of Technology Calicut

Department of Computer Science and Engineering

CS4044D Machine Learning

Midterm Test 1 | 02/09/2021 | 8:30AM-10:30AM | Total Marks: 15

1. [3 Marks] A clinical test is implemented to detect whether a patient has a certain disease or not, whose test result is either *positive* or *negative*.
For a patient with this disease, the probability of returning *positive* result is 0.98. For patient without this disease, the probability of returning *negative* result is 0.97.
The probability for any person to have this disease is 0.008.
If *positive* result is returned for some person, does he/she has this disease or not? Clearly write each step used in finding your answer.

2. [4 Marks] It is required to build a machine learning model to predict whether a vehicle costs Rs. 50000/- or not, given height of the vehicle. Suppose 1209 vehicles are considered and there are 221 vehicles that cost more than Rs. 50000/-. Assume there is a vehicle with height 1.05m whose price is to be checked greater than Rs. 50000/-. It is also given that the vehicle height spectrum is discretized and there are 46 vehicles with price more than Rs. 50000/- and 59 vehicles with price less than Rs. 50000/- in the interval where the given vehicle height 1.05m falls. Does the vehicle's price is greater than Rs. 50000/-.

3. [2.5 Marks] Consider a simple loss function as given below in the table. For a particular x, it is also given that P(w_1 | x) = 0.01 and P(w_2 | x) = 0.99. For the given x, which action takes least loss? Show all the steps in finding your answer.

Action / Class	α_1="Treatment A"	α_2="Treatment B"	α_3="Treatment C"
w_1 = "Covid-19"	5	50	10000
w_2 = "No Covid-19"	60	3	0

4. [0.5 Mark] Under what assumptions Bayesian classifier with discriminant function is optimal?

5. [2 Marks] Let $w_{max}(x)$ be the state of nature for which $P(w_{max} \mid x) \geq P(w_i \mid x)$ for all i, i = 1, ..., c.
Show that $P(w_{max} \mid x) \geq \frac{1}{c}$.

6. [1.5 Marks] Suppose we have 5 classes. Draw the decision regions that will be formed by a linear machine classifier. Clearly label each decision regions and indicate the hyperplanes also.

7. [1.5 Mark] Derive the *likelihood ratio* for binary classification problem and mention how this can be used for deciding the class? Show all the steps.