SMAI-2020-Homework

September 2020

1 Objective Question and a one line justification: [1 mark]

True or False? If we train a Naive Bayes classifier using infinite training data that satisfies all of its modeling assumptions, then it will achieve zero training error over these training examples. Please justify your answer in one sentence.

2 Programming Question: [2 marks]

Go through the notebook on Optimal Bayes Classifier. Complete the exercise section of the notebook. In that section you are required to compare the optimal Bayes classifier and the Naïve Bayes classifier for the case where the covariance is non-zero, and (a) Plot the decision boundary; and (b) print the accuracy achieved by the two methods.

3 Subjective Question: [3 marks]

Consider the following medical diagnosis problem in which there are two alternative hypothesis: (1) The patient is COVID-19 positive, and (2) that the patient does not. The available data is from a particular rapid antigen based test kit with two possible outcomes: $\oplus$ (positive), and $\ominus$ (negative). We have the prior knowledge that over the entire population only 0.08 has this disease. Furthermore, the lab test is only an imperfect indicator of the disease. The test returns a correct positive result in only 90% of the cases in which the disease is actually present, and a correct negative result in only 97% of the cases in which the disease is not present. In other cases, the test returns the opposite result. Suppose we observe a new patient for whom the lab returns a positive result. Given this information,

1. Should we diagnose the person as having COVID-19 or not? (What is the probability of the person having COVID-19?)
2. Assuming that the prior knowledge that the entire population has this disease is 0.6, what is the probability of the person having the disease?