AUM Computer Science

Machine Learning, Spring 2023, Dr. Olcay Kursun

Bayes Rule Practice Question

Suppose that we have a single variable, x, to use for a binary classification problem with classes 1 and 2 with prior probabilities 75% and 25%, respectively. The average and the standard deviation of x values for class-1 examples are 200 and 50; and for class-2, 250 and 100. Show how Bayes Rule can be used for classification (show the formula and steps)? What is the predicted class (and posterior probabilities) of a test example with x=225? Assume that x is normally distributed.

**Solution:** Bayes Rule states that

If they were not given in the question, we could estimate the distribution parameters (assuming a Gaussian distribution) of feature x as follows:

(mean)

(standard deviation)

(probability density function for Gaussian)

P(x = 225| class = 1) = = **0.00704 (likelihood of x=225 for class-1)**

P(x = 225| class = 2) =**= 0.00387 (likelihood of x=225 for class-2)**

P(x = 225) = P(x = 225 | class = 1) \* P(class = 1) + P(x = 225 | class = 2) \* P(class = 2)

= 0.00704 \* 0.75 + 0.00387 \* 0.25 = 0.0062475, which is the weighted sum of the likelihoods found for each class above, which adds up to “evidence” (likelihood of observing x=225 overall). Relative to each other among classes (which gives us the posterior probabilities 0.85 vs 0.15, so we classify the test example as class-1):

P(class = 1 | x = 225) = 0.00704 \* 0.75 / 0.0062475 **= 0.8451 (~ 85%)**

P(class = 2 | x = 225) = 0.00387 \* 0.25 / 0.0062475 = 0**.1549 (~ 15%)**