Long Nguyen

1001705873

1a)

P(< 80) = P(< 80 | Maine)P(Maine) * P(< 80 | Sahara)P(Sahara) = (.80*.05)+(.1*.95) = .135

P(Maine | < 80) = P(< 80 | Maine)*P(Maine) / P(< 80) = (.80 * .05) / .135 = .2963

1b)

P(T2) = P(T2 < 80 | Maine)P(Maine) + P(T2 <80 | Sahara)P(Sahara) = .135

P(T2 < 80 | T1 < 80) = p(T2 < 80, Maine | T1 < 80) + p(T2 < 80, Sahara | T1 < 80) = p(T2 < 80 | T1 < 80, Maine)p(Maine | T1 < 80) + p(T2 < 80 | T1 < 80, Sahara)p(Sahara | T1 < 80) = p(T2 < 80 | Maine)p(Maine | T1 < 80) + p(T2 < 80 | Sahara)p(Sahara | T1 < 80) = .80*.2963 + .10*(1-.2963) = .3074

Since the probability of the second email is not independent of the first email, the probability of the second e-mail being below 80 degrees is about 30.74%.

1c)

P(T1 < 80 and T2 < 80 and T3 < 80) = p(T1 < 80 and T2 < 80 and T3 < 80 | Maine)p(Maine) + p(T1 < 80 and T2 < 80 and T3 < 80 | Sahara)p(Sahara) = .8^3*.05 + .1^3*.95 = 0.02655

The probability of the first three emails indicate daily highs under 80 degrees is about 2.655%

2)

For the function P to be a valid probability function, it must satisfy the following properties:

i) The probability of any event cannot be less than 0 or greater than 1

For this property, the two events we know are P(A) and P(B) which are both under 1 and greater than 0 so they satisfy this property.

ii) The sum of probabilities of all possible atomic events is 1

For this property, the two events we know are P(A) and P(B), which when summed up are under the value of 1. But since we do not know the values of P(C) and P(D), they would have to be (P(C) + P(D) equal to 0.1 when summed up since P(A) + P(B) = 0.9.

Therefore, based on the given information, P is possibly a valid probability function.

3)

For function P to be a valid probability density function, it must satisfy the following properties:

i) P(x) >= 0

For this property, the value P(x) = 0.3 is non-negative for all values 0 to 10, so this condition is partially satisfied since we don't know about the other value outside this interval.

ii) $\int_{-\infty}^{\infty} P(x)dx = 1$

For this property, this states that the total area under the function must be equal to 1 but the integral of P(x) over [0,10] is not equal to 1.

$$\int_{0}^{10} 0.3dx = 3$$

Therefore, function P cannot be a probability density function since it violates property ii

4)

p(F = a) = p(F=a | B=r)p(B=r) + p(F=a | B=b)p(B=b) = 0.55

p(F = o) = p(F=o | B=b)p(B=b) + p(F=o | B=r)p(B=r) = 0.45

p(B=b | F=a) = p(F=a | B=b)p(B=b) / p(F=a) = 0.82

p(B=b | F=o) = p(F=o | B=b)p(B=b) / p(F=o) = 0.33

p(B=r | F=a) = p(F=a | B=r)p(B=r) / p(F=a) = 0.18

p(B=r | F=o) = p(F=o | B=r)p(B=r) / p(F=o) = 0.67

Based on the values above (excluding p(F=a) and p(F=o)), the probability that the classifier will give the correct output for x about 82% if the box is blue and the fruit is an apple, 33% if the box is blue and the fruit is an orange, 18% if the box is red and the fruit is an apple, and 67% if the box is red and the fruit is an orange. This will result in an average correctness of 49%.

5)

Training output:

Class: 1, Attribute: 1, mean = 0.52, std = 0.10

Class: 1, Attribute: 2, mean = 0.54, std = 0.10

Class: 1, Attribute: 3, mean = 0.52, std = 0.07

Class: 1, Attribute: 4, mean = 0.41, std = 0.17

Class: 1, Attribute: 5, mean = 0.50, std = 0.01

Class: 1, Attribute: 6, mean = 0.00, std = 0.01

Class: 1, Attribute: 7, mean = 0.50, std = 0.05

Class: 1, Attribute: 8, mean = 0.24, std = 0.05

Class: 2, Attribute: 1, mean = 0.45, std = 0.11

Class: 2, Attribute: 2, mean = 0.45, std = 0.10

Class: 2, Attribute: 3, mean = 0.53, std = 0.06

Class: 2, Attribute: 4, mean = 0.23, std = 0.11

Class: 2, Attribute: 5, mean = 0.50, std = 0.04

Class: 2, Attribute: 6, mean = 0.00, std = 0.01

Class: 2, Attribute: 7, mean = 0.49, std = 0.06

Class: 2, Attribute: 8, mean = 0.33, std = 0.14

Class: 3, Attribute: 1, mean = 0.43, std = 0.10

Class: 3, Attribute: 2, mean = 0.48, std = 0.11

Class: 3, Attribute: 3, mean = 0.36, std = 0.06

Class: 3, Attribute: 4, mean = 0.22, std = 0.08

Class: 3, Attribute: 5, mean = 0.51, std = 0.05

Class: 3, Attribute: 6, mean = 0.00, std = 0.01

Class: 3, Attribute: 7, mean = 0.51, std = 0.04

Class: 3, Attribute: 8, mean = 0.27, std = 0.09

Class: 4, Attribute: 1, mean = 0.79, std = 0.07

Class: 4, Attribute: 2, mean = 0.76, std = 0.07

Class: 4, Attribute: 3, mean = 0.38, std = 0.06

Class: 4, Attribute: 4, mean = 0.32, std = 0.11

Class: 4, Attribute: 5, mean = 0.50, std = 0.01

Class: 4, Attribute: 6, mean = 0.00, std = 0.01

Class: 4, Attribute: 7, mean = 0.51, std = 0.07

Class: 4, Attribute: 8, mean = 0.27, std = 0.09

Class: 5, Attribute: 1, mean = 0.74, std = 0.16

Class: 5, Attribute: 2, mean = 0.62, std = 0.13

Class: 5, Attribute: 3, mean = 0.42, std = 0.08

Class: 5, Attribute: 4, mean = 0.30, std = 0.12

Class: 5, Attribute: 5, mean = 0.50, std = 0.01

Class: 5, Attribute: 6, mean = 0.00, std = 0.01

Class: 5, Attribute: 7, mean = 0.51, std = 0.06

Class: 5, Attribute: 8, mean = 0.24, std = 0.04

Class: 6, Attribute: 1, mean = 0.54, std = 0.14

Class: 6, Attribute: 2, mean = 0.50, std = 0.12

Class: 6, Attribute: 3, mean = 0.51, std = 0.05

Class: 6, Attribute: 4, mean = 0.24, std = 0.10

Class: 6, Attribute: 5, mean = 0.50, std = 0.01

Class: 6, Attribute: 6, mean = 0.49, std = 0.39

Class: 6, Attribute: 7, mean = 0.51, std = 0.03

Class: 6, Attribute: 8, mean = 0.24, std = 0.05

Class: 7, Attribute: 1, mean = 0.48, std = 0.11

Class: 7, Attribute: 2, mean = 0.47, std = 0.09

Class: 7, Attribute: 3, mean = 0.54, std = 0.06

Class: 7, Attribute: 4, mean = 0.22, std = 0.12

Class: 7, Attribute: 5, mean = 0.50, std = 0.04

Class: 7, Attribute: 6, mean = 0.00, std = 0.03

Class: 7, Attribute: 7, mean = 0.50, std = 0.06

Class: 7, Attribute: 8, mean = 0.26, std = 0.09

Class: 8, Attribute: 1, mean = 0.74, std = 0.11

Class: 8, Attribute: 2, mean = 0.73, std = 0.11

Class: 8, Attribute: 3, mean = 0.49, std = 0.05

Class: 8, Attribute: 4, mean = 0.29, std = 0.07

Class: 8, Attribute: 5, mean = 0.50, std = 0.01

Class: 8, Attribute: 6, mean = 0.00, std = 0.01

Class: 8, Attribute: 7, mean = 0.46, std = 0.08

Class: 8, Attribute: 8, mean = 0.23, std = 0.02

Class: 9, Attribute: 1, mean = 0.55, std = 0.14

Class: 9, Attribute: 2, mean = 0.56, std = 0.16

Class: 9, Attribute: 3, mean = 0.51, std = 0.07

Class: 9, Attribute: 4, mean = 0.20, std = 0.07

Class: 9, Attribute: 5, mean = 0.50, std = 0.01

Class: 9, Attribute: 6, mean = 0.00, std = 0.01

Class: 9, Attribute: 7, mean = 0.53, std = 0.05

Class: 9, Attribute: 8, mean = 0.24, std = 0.05

Class: 10, Attribute: 1, mean = 0.78, std = 0.06

Class: 10, Attribute: 2, mean = 0.73, std = 0.12

Class: 10, Attribute: 3, mean = 0.48, std = 0.11

Class: 10, Attribute: 4, mean = 0.33, std = 0.07

Class: 10, Attribute: 5, mean = 1.00, std = 0.01

Class: 10, Attribute: 6, mean = 0.00, std = 0.01

Class: 10, Attribute: 7, mean = 0.55, std = 0.02

Class: 10, Attribute: 8, mean = 0.23, std = 0.01

Classification output:

classification accuracy = 0.4483