Name: Pratik Jasani

Course: EN 605.645

**Bayesian Network Self Check**

Suppose we have the following training data where Shape, Size and Color are the features (attributes)

and Safe? is the class label:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **#** | **Shape** | **Size** | **Color** | **Safe?** |
| 1 | Round | Large | Blue | No |
| 2 | Square | Large | Green | Yes |
| 3 | Square | Small | Red | No |
| 4 | Round | Large | Red | Yes |
| 5 | Square | Small | Blue | No |
| 6 | Round | Small | Blue | No |
| 7 | Round | Small | Red | Yes |
| 8 | Square | Small | Green | No |
| 9 | Round | Large | Green | Yes |
| 10 | Square | Large | Green | Yes |
| 11 | Square | Large | Red | No |
| 12 | Square | Large | Green | Yes |
| 13 | Round | Large | Red | Yes |
| 14 | Square | Small | Red | No |
| 15 | Round | Small | Green | No |

1. Calculate the probabilities required for a Naive Bayes Classifier.

P(safe = yes) = 7/15 = 0.467

P(safe = no) = 8/15 = 0.533

|  |  |  |  |
| --- | --- | --- | --- |
| Features | Domain Vals | Prob(Yes) | Prob(No) |
| Shape | Round | (3+1)/(7+1) = 0.5 | (4+1)/(8+1) = 0.556 |
| Square | (3+1)/(7+1) = 0.5 | (5+1)/(8+1) = 0.667 |
| Size | Large | (6+1)/(7+1) =0.875 | (2+1)/(8+1) = 0.333 |
| Small | (6+2)/(7+1) = 1 | (1+1)/(8+1) = 0.222 |
| Color | Red | (3+1)/(7+1) = 0.5 | (3+1)/(8+1) = 0.444 |
| Green | (4+1)/(7+1) = 0.625 | (2+1)/(8+1) = 0.333 |
| Blue | (0 +1)/(7+1) = 0.125 | (3+1)/(8+1) = 0.444 |

2. What is the normalized probability distribution over the possible class labels for the example:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 16 | Square | Large | Red | ? |

P\_no = P(shape = Square | class = no) \* P(size = Large | class = no) \* P(color = Red | class = no)\*P(class=no)

= 0.667 \* 0.333 \* 0.444 \* 0.533 = 0.05256

Normalize => 0.05256/(0.05256+ 0.10215) = 0.33973

P\_yes = P(shape = Square | class = yes) \* P(size = Large | class = yes) \* P(color = Red | class = yes) \* P(safe=yes)

= 0.5 \* 0.875 \* 0.5 \* 0.467 = 0.10215

Normalize = 0.10215/(0.05256+ 0.10215) = 0.6603

The predicted label is yes.