Calvin Chen

Self Check 12

1. Data Probability Table

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Yes | | | | No | | | |
|  | Round | | Square | | Round | | Square | |
|  | Large | Small | Large | Small | Large | Small | Large | Small |
| Blue | 0.000 | 0.000 | 0.000 | 0.000 | 0.067 | 0.067 | 0.000 | 0.067 |
| Green | 0.067 | 0.000 | 0.200 | 0.000 | 0.000 | 0.067 | 0.000 | 0.067 |
| Red | 0.133 | 0.067 | 0.000 | 0.000 | 0.000 | 0.000 | 0.067 | 0.133 |

We can use this data to construct the Bayesian Probability Table:

|  |  |
| --- | --- |
| P(Yes) | 0.467 |
| P(No) | 0.533 |
| P(Round|Yes) | 0.571 |
| P(Round|No) | 0.375 |
| P(Square|Yes) | 0.429 |
| P(Square|No) | 0.625 |
| P(Large|Yes) | 0.85714286 |
| P(Large|No) | 0.25 |
| P(Small|Yes) | 0.14285714 |
| P(Small|No) | 0.75 |
| P(Blue|Yes) | 0 |
| P(Blue|No) | 0.375 |
| P(Green|Yes) | 0.57142857 |
| P(Green|No) | 0.25 |
| P(Red|Yes) | 0.42857143 |
| P(Red|No) | 0.375 |

1. Assuming that the features are all mutually exclusive, we can calculate the normalize probability distribution of Square Large Red. We have the following formula:

P(YES, Square, Large, Red) = P(YES | Square, Large, Red) \* P(Square | Large, Red) \* P(Large | Red) \* P(Red)

Expanding, and then simplifying it using Bayes’ Formula, we are left with:

P(YES, Square, Large, Red) = P(YES | Square, Large, Red) \* P(Square, Large, Red)

Rearranging the terms, we then have:

P(YES | Square, Large, Red) = P(YES, Square, Large, Red) / P(Square, Large, Red)

and its counterpart:

P(No | Square, Large, Red) = P(No, Square, Large, Red) / P(Square, Large, Red)

From the above data:

P(YES, Square, Large, Red): 0

P(NO, Square, Large, Red): 0.067

P(Square, Large, Red): 0.067

Normalizing the probability

P(YES | Square, Large, Red): 0 / 0.067 = 0

P(NO | Square, Large, Red) : 0.067 / 0.067 = 1

Given the provided data, the model would predict that Square Large Red would NOT be safe.