on their loan or not based on some demographic and financial data. Here is a sample dataset containing 10 loan applicants and whether they defaulted on their loan or not:

| Applicant ID | Age | Income | Education Level | Defaulted |
|---|---|---|---|---|
| 1 | 25 | 20,000 | High School | No |
| 2 | 35 | 50,000 | Bachelor's | No |
| 3 | 45 | 80,000 | Master's | No |
| 4 | 28 | 22,000 | High School | No |
| 5 | 32 | 45,000 | Bachelor's | Yes |
| 6 | 46 | 70,000 | Master's | No |
| 7 | 24 | 18,000 | High School | Yes |
| 8 | 38 | 60,000 | Bachelor's | No |
| 9 | 32 | 48,000 | Bachelor's | No |
| 10 | 29 | 25,000 | High School | Yes |

| Applicant ID | Age | Income | Education Level | Defaulted |
|---|---|---|---|---|
| 11 | 31 | 55,000 | Bachelor's | ? |

In this example, we have a new applicant who is 31 years old, has an annual income of $55,000, and has a Bachelor's degree. The question mark in the Defaulted column indicates that we do not know whether this applicant will default on their loan or not. We can use our Naive Bayes classifier to predict the value of the Defaulted column for this new applicant based on the values of the other columns.

Prior : $P(buy=yes) = \frac{3}{10}$ $P(buy=No) = \frac{7}{10}$

P( buy | age 30-39 , Income = 40,000 - 59,999 , Education = Bachelor's )

Likelihood

P( buy | age 30-39 , Income = 40,000 - 59,999 , Education = Bachelor's | buy )

p( age 30-39 | buy ) × ( P( income = 40,000 - 59,999) × P( Education = Bachelor's )

$\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

$P(X|C_i)$ $p(X|buy = \text{"yes"}) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = 0.04$

P( No | age 30-39 , Income = 40,000 - 59,999 , Education = Bachelor's | No )

p( age 30-39 | No ) × ( P( income = 40,000 - 59,999 | No) × P( Education = Bachelor's | No)

$\frac{3}{7}$ $\frac{1}{7}$ $\frac{3}{7}$

$P(X|C_i)$ $p(X|buy = \text{"No"}) = \frac{3}{7} \times \frac{1}{7} \times \frac{3}{7} = 0.026$

$P(X|C_i) \times P(C_i)$ Yes = $0.037 \times 0.30 = 0.011$

No = $0.04 \times 0.7 = 0.34$

Ans No