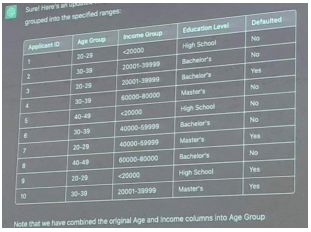
Imagine you work for a bank and you want to predict whether a loan applicant will default on their loan or not based on some demographic and financial data. Here is a sample dataset containing 10 loan applicants and whether they defaulted on their loan or not:

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.

แบ่งช่วงอายุ 10-19, 20-29, 30-39, 40-49 และแบ่งช่วงรายได้ < 20,000, 20,001 - 39,999, 40,000 - 59,999, 60,000 - 80,000



Sure! Here's an updated version of the dataset with the Age and Income columns grouped into the specified ranges:

Applicant ID	Age Group	Income Group	Education Level	Defaulted
1	20-29	<20000	High School	No
2	30-39	20001-39999	Bachelor's	No
3	30-39	20001-39999	Bachelor's	Yes
4	30-39	60000-80000	Master's	No
5	40-49	<20000	High School	No
6	30-39	40000-59999	Bachelor's	No
7	20-29	40000-59999	Master's	Yes
8	40-49	60000-80000	Bachelor's	No
9	20-29	<20000	High School	Yes
10	30-39	20001-39999	Master's	Yes

Note that we have combined the original Age and Income columns into Age Group

$P(\text{Defaulted} \mid \text{Age Group} = 30-39, \text{Income Group} = 40{,}000-59{,}999, \text{Education Level} = \text{Bachelor's})$

Prior

$P(C_i)$: $P(\text{Defaulted} = \text{Yes}) = \frac{4}{10} = 0.4$

$P(\text{Defaulted} = \text{NO}) = \frac{6}{10} = 0.6$

Likelihood

$P(X|C_i)$: $P(\text{Age Group} = 30-39, \text{Income Group} = 40{,}000-59{,}999, \text{Education Level} = \text{Bachelor's} \mid \text{Defaulted})$

$P(\text{Age Group} = 30-39 \mid \text{Defaulted} = \text{Yes}) = \frac{1}{4} = 0.25$

$P(\text{Age Group} = 30-39 \mid \text{Defaulted} = \text{NO}) = \frac{3}{6} = 0.5$

$P(\text{Income Group} = 40{,}000-59{,}999 \mid \text{Defaulted} = \text{Yes}) = \frac{1}{4} = 0.25$

$P(\text{Income Group} = 40{,}000-59{,}999 \mid \text{Defaulted} = \text{NO}) = \frac{1}{6} = 0.167$

$P(\text{Education Level} = \text{Bachelor's} \mid \text{Defaulted} = \text{Yes}) = \frac{1}{4} = 0.25$

$P(\text{Education Level} = \text{Bachelor's} \mid \text{Defaulted} = \text{NO}) = \frac{3}{6} = 0.5$

$X = (\text{Age} = 31, \text{Incom} = 55{,}000, \text{Education Level} = \text{Bachelor'S})$

$P(X|C_i)$: $P(X \mid \text{Defaulted} = \text{Yes}) = 0.25 \times 0.25 \times 0.25 = 0.0156$

$P(X \mid \text{Defaulted} = \text{NO}) = 0.5 \times 0.167 \times 0.5 = 0.0917$

$P(X|C_i) \times P(C_i)$: $P(X \mid \text{Defaulted} = \text{Yes}) \times P(\text{Defaulted} = \text{Yes}) = 0.0156 \times 0.4 = 0.00624$

$P(X \mid \text{Defaulted} = \text{NO}) \times P(\text{Defaulted} = \text{NO}) = 0.0917 \times 0.6 = 0.025$

Therefor, X belongs to class(Defaulted = NO)