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:

10-19 , 20-29 , 30-39, 40-49

Applicant ID	Age	Income	Education Level	Defaulted
1	25 20-29	~~20,000~~ < 20,000	High School	No
2	35 30-39	50,000 40,000-59,999	Bachelor's	No
3	45 40-49	80,000 60,000-89,999	Master's	No
4	28 20-29	22,000 20,001-39,999	High School	No
5	32 30-39	45,000 40,000-59,999	Bachelor's	Yes
6	46 40-49	70,000 60,000-89,999	Master's	No
7	24 20-29	18,000 < 20,000	High School	Yes
8	38 30-39	60,000 60,000-89000	Bachelor's	No
9	32 30-39	48,000 40,000-59,999	Bachelor's	No
10	29 20-29	25,000 20,001-39,999	High School	Yes

ตารางทำนาย

Applicant ID	Age	Income	Education Level	Defaulted
11	31 30-39	55,000	Bachelor's	?

40,000 - 59,999

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.

อยากรู้ว่าเขาจะเบี้ยวหนี้ไหม.
ถ้ามีคนเกินเข้ามาใหม่ เก จ. เบี้ยวหนี้ไหม

วิธีทำ

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

Prior $\quad P(\text{Defaulted} = \text{"Yes"}) = \dfrac{3}{10}$

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

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

$$\dfrac{1}{3} \times \dfrac{1}{3} \times \dfrac{1}{3}$$

∴ $P(X \mid C_i) = P(X \mid \text{Defaulted}) = \dfrac{3}{10} \times \dfrac{1}{3} \times \dfrac{1}{3} \times \dfrac{1}{3} = 0.01111$

Pvior $P(\text{Defaulted} = \text{``NO''}) = \dfrac{7}{10}$

Likelihood $P(\text{Age} = 30\text{-}39, \text{ Income} = 40{,}000\text{-}59{,}999, \text{ Education} = \text{Bachelor's} \mid \text{NO})$

$P(\text{Age} = 30\text{-}39 \mid \text{NO}) \times P(\text{Income} = 40{,}000\text{-}59{,}999 \mid \text{NO}) \times P(\text{Education} = \text{Bachelor's} \mid \text{NO})$

$\dfrac{3}{7} \times \dfrac{2}{7} \times \dfrac{3}{7}$

∴ $P(X \mid C_i) = P(X \mid \text{Defaulted}) = \dfrac{7}{10} \times \dfrac{3}{7} \times \dfrac{2}{7} \times \dfrac{3}{7} \approx 0.0367$

ดังนั้นจะได้ว่า คนที่เข้ามาใหม่ที่มีอายุในช่วง 30 - 39 ปี   รายได้ในช่วง 40,000 - 59,000 บาท   และ สถานะการศึกษา เป็น Bachelor's
มีโอกาสอย่าง เป็น ที่จะกู้เงินแล้วกลับไม่เป็นหนี้