@S S Roy, 28th July
Ref :Han and Kamber, 2011

Bayes Classification Methods

✅ Step-by-Step Naïve Bayes Classification Tutorial with Full Example

🔍 What Are Bayesian Classifiers?

Bayesian classifiers predict the probability that a data point (called a **tuple**) belongs to a particular class using **Bayes' Theorem**.

📘 Bayes' Theorem

$$P(H \mid X) = \frac{P(X \mid H) \cdot P(H)}{P(X)}$$

Where:

- $P(H \mid X)$: Posterior probability (probability of hypothesis H given evidence X)
- $P(H)$: Prior probability of hypothesis H
- $P(X \mid H)$: Likelihood (probability of evidence X given that H is true)
- $P(X)$: Prior probability of evidence X

🧠 Naïve Bayes Assumption

Attributes are **conditionally independent** given the class.

📄 Training Data Table (From Uploaded File)

ID	Age	Income	Student	Credit_rating	Buys_compu
1	Youth	High	No	Fair	No
2	Youth	High	No	Excellent	No
3	Middle-aged	High	No	Fair	Yes
4	Senior	Medium	No	Fair	Yes
5	Senior	Low	Yes	Fair	Yes
6	Senior	Low	Yes	Excellent	No
7	Middle-aged	Low	Yes	Excellent	Yes

ID	Age	Income	Student	Credit_rating	Buys_compu
8	Youth	Medium	No	Fair	No
9	Youth	Low	Yes	Fair	Yes
10	Senior	Medium	Yes	Fair	Yes
11	Youth	Medium	Yes	Excellent	Yes
12	Middle-aged	Medium	No	Excellent	Yes
13	Middle-aged	High	Yes	Fair	Yes
14	Senior	Medium	No	Excellent	No

🧪 Predict the Class of a New Tuple:

Tuple X:

`(Age=Youth, Income=Medium, Student=Yes, Credit_rating=Fair)`

We compute for both classes:

- C_1 = Yes
- C_2 = No

🧮 Step 1: Prior Probabilities

From the dataset of 14 instances:

- $P(\text{Yes}) = \frac{9}{14} = 0.643$
- $P(\text{No}) = \frac{5}{14} = 0.357$

🧮 Step 2: Conditional Probabilities

Let's break down each conditional probability from the table.

For Class = "Yes" (9 instances):

- $P(\text{Age=Youth} \mid \text{Yes}) = \frac{2}{9} = 0.222$
- $P(\text{Income=Medium} \mid \text{Yes}) = \frac{4}{9} = 0.444$
- $P(\text{Student=Yes} \mid \text{Yes}) = \frac{6}{9} = 0.667$
- $P(\text{Credit\_rating=Fair} \mid \text{Yes}) = 6/9 = 0.667$

For Class = "No" (5 instances):

- $P(\text{Age=Youth} \mid \text{No}) = \frac{3}{5} = 0.600$
- $P(\text{Income=Medium} \mid \text{No}) = \frac{2}{5} = 0.400$
- $P(\text{Student=Yes} \mid \text{No}) = \frac{1}{5} = 0.200$
- $P(\text{Credit\_rating=Fair} \mid \text{No}) = 2/5 = 0.400$

🧮 Step 3: Compute Likelihoods

For "Yes":

$$P(X \mid \text{Yes}) = 0.222 \times 0.444 \times 0.667 \times 0.667 = 0.044$$

For "No":

$$P(X \mid \text{No}) = 0.600 \times 0.400 \times 0.200 \times 0.400 = 0.019$$

🧮 Step 4: Multiply with Prior

Posterior for Yes:

$$P(\text{Yes} \mid X) \propto 0.044 \times 0.643 = 0.028$$

Posterior for No:

$$P(\text{No} \mid X) \propto 0.019 \times 0.357 = 0.007$$

✅ Final Prediction:

Since:

$$P(\text{Yes} \mid X) > P(\text{No} \mid X)$$

The **Naïve Bayes classifier predicts: "Buys Computer = Yes"**

⚠ What If Any Probability Is Zero?

Use **Laplace Correction**:

If an attribute value never appears for a class:

$$P(x_k \mid C_i) = \frac{\text{count} + 1}{\text{total for } C_i + \text{number of attribute values}}$$

📌 Summary Table of Calculation

Step	Value "Yes"	Value "No"
Prior Probability	9/14 = 0.643	5/14 = 0.357
Age=Youth	2/9 = 0.222	3/5 = 0.600
Income=Medium	4/9 = 0.444	2/5 = 0.400
Student=Yes	6/9 = 0.667	1/5 = 0.200
Credit_rating=Fair	6/9 = 0.667	2/5 = 0.400
Likelihood	0.044	0.019

Step	Value “Yes”	Value “No”
Posterior (Final Score)	0.044 × 0.643 = **0.028**	0.019 × 0.357 = **0.007**
Prediction	✅ **Yes**	❌ No