

PRE REQUISITES

Step - 1

Identify class Labels

0 - No

1 - Yes

Step - 2

Calculate Mean (μ)

$$\mu = \frac{\sum_{i=1}^n x_i}{N}$$

$x_i \rightarrow$ Feature Value for each Sample

$N \rightarrow$ Total No of Samples

For Example

class 0 \rightarrow [2.5, 3.5, 3.0, 2.8, 3.2]

$$\mu = \frac{15.0}{5} = 3$$

Step - 3

Calculate Variance (σ^2)

Variance is Provided

For class 0 — 1.0

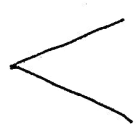
For class 1 — 2.0

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

$$\begin{aligned}\sigma^2 &= \frac{0.25 + 0.25 + 0 + 0.04 + 0.04}{5} \\ &= \frac{0.58}{5} = 0.116\end{aligned}$$

Step - 4

Observed Value (x)

$x = 4.0$  class 0
class 1

Check Value belongs to 0 or 1

step - 5

Calculate Prior

$$P(y) = \frac{\text{No of Samples in class } y}{\text{Total no of Samples}}$$

eg
100 $\begin{cases} \rightarrow 60 \text{ in class 0} \\ \rightarrow 40 \text{ in class 1} \end{cases}$

$$P(0) = 0.6$$

$$P(1) = 0.4$$

Summary

$$\mu_0 = 3.0$$

$$\mu_1 = 5.0$$

$$\sigma_0^2 = 1.0$$

$$\sigma_1^2 = 2.0$$

$$\lambda = 4.0$$

Gaussian Naive Bayes

Class	Mean	Variance	Observed Value	Prior
0	3.0	1.0	4.0	0.6
1	5.0	2.0	4.0	0.4

Aim

- 1) Compute likelihood for
 $P(x/0)$ & $P(x/1)$
- 2) Calculate Posterior Probability
 $P(0/x)$ & $P(1/x)$
By multiplying with the likelihood with prior
- 3) Normalize Posterior Probability
- 4) Predict the class with highest Posterior Probability.

step - 1

Gaussian Density formula

$$P(x/y) = \frac{1}{\sqrt{2\pi \cdot \sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Calculate for class label 0

$$P(x/0) \quad \mu_0 = 3.0 \quad \sigma_0^2 = 1.0 \quad x = 4.0$$

$$\frac{1}{\sqrt{2\pi \cdot 1}} \exp\left(-\frac{(4-3)^2}{2 \cdot 1}\right) \rightarrow \text{Simplify it}$$

$$\text{Approx Values } \sqrt{2\pi} = 2.5066$$

$$\exp(-0.5) = 0.6065$$

$$= 0.2419$$

step - 2

Similarly do it for $P(x/1)$

$$\sqrt{4\pi} = 3.5449$$

$$\exp(-0.25) = 0.7788$$

$$= 0.2197$$

step - 3

Multiply Prior Value for class 0

$$P(0/x) \propto P(x/0) \cdot P(0) = \underline{0.2419 \cdot 0.6}$$
$$= 0.14514$$

: Similarly Calculate for class 1

$$P(1/x) \propto P(x/1) \cdot P(1) = 0.2197 \cdot 0.4$$
$$= 0.0879$$

step - 4

Normalize Posterior Probability

$$P(x) = P(0/x) + P(1/x) = 0.14514 + 0.087$$
$$= 0.233$$

Divide each class by $P(x)$

For class label 0 $\rightarrow P(0/x) = \frac{0.14514}{0.233} = 0.622$

For class label 1 $\rightarrow P(1/x) = \frac{0.0879}{0.233} = 0.377$

step - 5

Make Prediction

$$P(0/x) > P(1/x) \quad \text{Predicted class is 0}$$

↓
Based on Posterior Probability

Bernoulli Naive Bayes

One or Absence it predict Zero of a Condition like word appearing in a Document.

Step - 1

Binarize the features Convert all features Values into binary like

Absence = 0 Presence = 1

Step - 2

Likelihood estimation

Class Conditional Probability



For each feature X Compute the Probability Absence or Presence with class y

Formula

$$P(x_i | y) = \frac{\text{Count of } x_i = 1 \text{ in class} + \alpha}{\text{Total Samples in class} + 2\alpha}$$

α is the Smoothing Parameter which is known as Laplace Smoothing

Multiply Probability for all features
Compute the Product of their respective Probability of the Condition on the class

$$P(x|y) = \prod_{i=1}^n P(x_i | y)$$

Step - 3

Prior Probability
Compute Prior Probability of each class like we did in Gaussian Naive Bayes

Step - 4

Make Prediction

Calculate Posterior Probability like we did in Gaussian Naive Bayes

Multinomial Naive Bayes

Count Based frequency like Everything is Seen with MB.

Example

Counting Repeated words in the Script

Formula

$$P(x_i | y) = \frac{\text{count of } x_i \text{ in class } y + \alpha}{\text{Total count of all features in class } y + \alpha \cdot |V|}$$

α = Smoothing Parameter

$|V|$ = Total no of unique features among all the classes else Same as Bernouli's.