Gaussian Naïve Bayes

- Gaussian Naive Bayes assumes that features follow a normal (Gaussian) distribution.
- Let’s take a sample dataset

Feature 1	Feature 2	Class
6.0	3.0	0
2.0	2.0	0
3.0	1.0	0
8.0	3.0	1
7.0	2.0	1
9.0	3.0	1

Example

- We want to classify a new data point **(Feature 1 = 4.0, Feature 2 = 2.0)**

Calculate the prior probabilities

The prior probability $P(Class = 0)$ and $P(Class = 1)$ can be calculated as the proportion of each class in the dataset.

- $P(Class = 0)$: There are 3 samples in Class 0 and 6 total samples.

$$P(Class = 0) = \frac{3}{6} = 0.5$$

- $P(Class = 1)$: Similarly, there are 3 samples in Class 1.

$$P(Class = 1) = \frac{3}{6} = 0.5$$

Calculate the mean and variance

- Mean

$$\text{Mean}(\mu) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Standard Deviation

$$\text{Standard Deviation}(\sigma) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2}$$

- Variance

$$\text{Variance}(\sigma^2) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

Calculate mean and variance (For Class 0)

- For each class, we calculate the **mean** and **variance** of each feature, assuming Gaussian (normal) distribution.

For Class 0:

- **Feature 1:**

$$\mu_{0,1} = \frac{6.0 + 2.0 + 3.0}{3} = 3.67$$

$$\sigma_{0,1}^2 = \frac{(6.0 - 3.67)^2 + (2.0 - 3.67)^2 + (3.0 - 3.67)^2}{3} = 2.89$$

- **Feature 2:**

$$\mu_{0,2} = \frac{3.0 + 2.0 + 1.0}{3} = 2.0$$

$$\sigma_{0,2}^2 = \frac{(3.0 - 2.0)^2 + (2.0 - 2.0)^2 + (1.0 - 2.0)^2}{3} = 0.67$$

Calculate mean and variance (For Class 1)

For Class 1:

- **Feature 1:**

$$\mu_{1,1} = \frac{8.0 + 7.0 + 9.0}{3} = 8.0$$

$$\sigma^2_{1,1} = \frac{(8.0 - 8.0)^2 + (7.0 - 8.0)^2 + (9.0 - 8.0)^2}{3} = 0.67$$

- **Feature 2:**

$$\mu_{1,2} = \frac{3.0 + 2.0 + 3.0}{3} = 2.67$$

$$\sigma^2_{1,2} = \frac{(3.0 - 2.67)^2 + (2.0 - 2.67)^2 + (3.0 - 2.67)^2}{3} = 0.22$$

Compute the likelihood

- For a new data point (4.0,2.0), we need to calculate the likelihood $P(x_i|Class_k)$ using the Gaussian probability density function:

$$P(x_i|Class_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

For class 0

For Class 0:

- **Feature 1:**

$$P(4.0|Class = 0) = \frac{1}{\sqrt{2\pi \cdot 2.89}} \exp\left(-\frac{(4.0 - 3.67)^2}{2 \cdot 2.89}\right)$$

$$= \frac{1}{2.68} \exp\left(-\frac{0.11}{5.78}\right) = 0.15$$

- **Feature 2:**

$$P(2.0|Class = 0) = \frac{1}{\sqrt{2\pi \cdot 0.67}} \exp\left(-\frac{(2.0 - 2.0)^2}{2 \cdot 0.67}\right)$$

$$= \frac{1}{2.05} \exp(0) = 0.49$$

For class 1

For Class 1:

- **Feature 1:**

$$P(4.0|Class = 1) = \frac{1}{\sqrt{2\pi \cdot 0.67}} \exp\left(-\frac{(4.0 - 8.0)^2}{2 \cdot 0.67}\right)$$

$$= \frac{1}{2.05} \exp\left(-\frac{16.0}{1.34}\right) \approx 4.95 \times 10^{-6}$$

- **Feature 2:**

$$P(2.0|Class = 1) = \frac{1}{\sqrt{2\pi \cdot 0.22}} \exp\left(-\frac{(2.0 - 2.67)^2}{2 \cdot 0.22}\right)$$

$$= \frac{1}{1.18} \exp\left(-\frac{0.45}{0.44}\right) \approx 0.27$$

Calculate the posterior probabilities

Now calculate the posterior probabilities for each class by multiplying the prior with the likelihoods.

- **For Class 0**:

$$P(Class = 0|4.0, 2.0) \propto P(Class = 0) \cdot P(4.0|Class = 0) \cdot P(2.0|Class = 0)$$

$$= 0.5 \cdot 0.15 \cdot 0.49 = 0.03675$$

- **For Class 1**:

$$P(Class = 1|4.0, 2.0) \propto P(Class = 1) \cdot P(4.0|Class = 1) \cdot P(2.0|Class = 1)$$

$$= 0.5 \cdot 4.95 \times 10^{-6} \cdot 0.27 \approx 6.68 \times 10^{-7}$$

Results

Since $P(Class = 0|4.0, 2.0)$ is larger than $P(Class = 1|4.0, 2.0)$, we predict **Class 0** for the new data point $(4.0, 2.0)$.