Naive Bayes

Naive Bayes is Based on Bayes' Theorem, which tells us how to compute the Probability of a class given some features.

1. Bayes' Theorem

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}$$

Where:

- C = class
- X = features
- P(C) = prior probability of class
- P(X|C) = likelihood
- P(C|X) = posterior (final score we compare)

Since P(X) is same for all classes, we only compare the numerator:

$$P(C|X) \propto P(X|C)\,P(C)$$

2. Naive Assumption

Naive Bayes assumes all features are independent:

$$P(X|C) = \prod_{j=1}^{n} P(x_j|C)$$

This is why it's called naive, but it makes calculations very simple.

3. Gaussian Likelihood

For numeric features, we assume each feature follows a Normal distribution:

$$P(x|C) = \frac{1}{\sqrt{2\pi\sigma_C^2}} \exp\left(-\frac{(x-\mu_C)^2}{2\sigma_C^2}\right)$$

Where:

- μ_C = mean of features for class C
- σ_C^2 = variance.

4. Posterior Score for a Class

For each class, we compute:

$$\log P(C|X) = \log P(C) + \sum_{j=1}^{n} \log P(x_j|C)$$

Why logs?

Because multiplying tiny probabilities can underflow - logs turn multiplication into addition.

---

5. Prediction

Choose the class with the highest posterior:

$$\hat{C} = \arg\max_{C} \left[\log P(C) + \sum \log P(x_j|C) \right]$$