**Bayes' Theorem**

* Formula: P(C∣X)=P(X∣C)⋅P(C)/P(X)
  + P(C∣X): Posterior probability of class C given predictor X.
  + P(X∣C): Likelihood of predictor X given class C.
  + P(C): Prior probability of class C.
  + P(X): Prior probability of predictor X.

**Assumptions**

* Features are independent given the class label.
* This assumption is often violated in real-world data, but the classifier still performs well.

**Types of Naive Bayes Classifiers**

1. **Gaussian Naive Bayes**: Assumes continuous values follow a Gaussian (normal) distribution.
2. **Multinomial Naive Bayes**: Used for discrete counts, suitable for text classification.
3. **Bernoulli Naive Bayes**: Used for binary/boolean features.

**Gaussian Naive Bayes**

* Assumes data from each class is drawn from a Gaussian distribution.
* Formula: P(xi∣Ck)=1/√2πσ2k.(exp(−(xi−μk)2/2σk2))
  + μk ​: Mean of the k-th class.
  + σk:Variance of the k-th class.

**Advantages**

* Simple and easy to implement.
* Efficient in terms of storage and computation.
* Handles large datasets and many features well.
* Works well even with less data.

**Disadvantages**

* Assumption of feature independence is often unrealistic.
* Not suitable for datasets with highly correlated features.

**Applications**

* **Text Classification**: Spam filtering, sentiment analysis, document categorization.
* **Medical Diagnosis**: Predicting diseases based on symptoms.
* **Recommendation Systems**: Predicting user preferences.

**Example: Iris Dataset**

* **Iris Dataset**: Classic dataset for classification with 150 instances, 4 features, and 3 classes (setosa, versicolor, virginica).