

# 朴素贝叶斯算法详解

## 一、核心思想

朴素贝叶斯（Naive Bayes）是一种基于 贝叶斯定理（Bayes Theorem） 的分类算法。它通过先验概率与特征条件概率，计算出后验概率，并以最大后验概率作为分类结果。

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

其中：

$\underbrace{P(Y)}_{\text{先验概率}}$  代表在未观察到特征时，对类别的先验信念。  
 $\underbrace{P(X|Y)}_{\text{似然（条件概率）}}$  代表在类别已知时，观测到特征的可能性。  
 $\underbrace{P(Y|X)}_{\text{后验概率}}$  是我们真正关心的量，用于分类决策。

---

## 二、贝叶斯定理与朴素假设

朴素贝叶斯假设各特征之间条件独立：

$$P(X|Y) = \prod_{i=1}^n P(x_i|Y)$$

因此有：

$$P(Y|X) \propto P(Y) \prod_{i=1}^n P(x_i|Y)$$

最终分类规则为：

$$\hat{Y} = \arg \max_Y P(Y|X)$$


---

### 三、流程图

$$P(Y) \text{ (先验)} + P(X|Y) \text{ (似然)} \xrightarrow{\text{贝叶斯定理}} P(Y|X) \text{ (后验)} \xrightarrow{\text{选择最大后验}} \text{分类结果}$$


---

### 四、例子：邮件分类

设类别为：

$$Y \in \{\text{Spam}, \text{NotSpam}\}$$

特征为某个单词（如 “Win”）是否出现。

#### 1. 先验概率

$$P(\text{Spam}) = 0.4, \quad P(\text{NotSpam}) = 0.6$$

#### 2. 似然概率

$$P(\text{"Win"}|\text{Spam}) = 0.8, \quad P(\text{"Win"}|\text{NotSpam}) = 0.1$$

#### 3. 后验概率比较

根据贝叶斯定理，有：

$$P(\text{Spam}|\text{"Win"}) \propto P(\text{"Win"}|\text{Spam})P(\text{Spam}) = 0.8 \times 0.4 = 0.32$$

$$P(\text{NotSpam}|\text{"Win"}) \propto P(\text{"Win"}|\text{NotSpam})P(\text{NotSpam}) = 0.1 \times 0.6 = 0.06$$

因此：

$$P(\text{Spam}|\text{"Win"}) > P(\text{NotSpam}|\text{"Win"})$$

结论：该邮件被判定为垃圾邮件（Spam）。

## 五、朴素贝叶斯在代码中的体现

```
from sklearn.naive_bayes import GaussianNB

model = GaussianNB()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

- fit() 阶段：计算各类别的先验概率  $P(Y)$  以及特征的条件分布参数（均值与方差）。
- predict() 阶段：使用

$$P(Y|X) \propto P(Y) \prod_i P(x_i|Y)$$

计算每个类别的后验概率，选择最大者作为预测结果。

## 六、总结

朴素贝叶斯使用先验与似然来计算后验，并根据后验进行分类。

简而言之：

- 先验  $P(Y)$ ：还没看数据前对类别的信念；
- 似然  $P(X|Y)$ ：数据告诉我们的信息；

- 后验  $P(Y|X)$ : 综合了两者之后更新的信念;

最终, 朴素贝叶斯根据最大后验概率进行分类:

$$\hat{Y} = \arg \max_Y P(Y|X)$$