**BÁO CÁO: THUẬT TOÁN NAIVE BAYES**

**1. Giải thuật Naive Bayes hoạt động như thế nào**

Naive Bayes là một **thuật toán phân loại dựa trên xác suất**, được xây dựng từ **Định lý Bayes**.  
Thuật toán được gọi là **“ngây thơ” (naive)** vì giả định rằng **các đặc trưng (features)** trong dữ liệu **độc lập với nhau** trong cùng một lớp (class).

**Định lý Bayes được biểu diễn như sau:**

**Trong đó:**

* : Xác suất hậu nghiệm – xác suất mẫu X thuộc lớp C.
* : Xác suất xuất hiện của X khi biết trước lớp C.
* : Xác suất tiên nghiệm của lớp C.
* : Xác suất xảy ra của dữ liệu X.

**Giả định “ngây thơ”:**

Mô hình chọn lớp có xác suất hậu nghiệm cao nhất để dự đoán:

**2. Các loại mô hình Naive Bayes và khi nào sử dụng**

|  |  |  |
| --- | --- | --- |
| **Mô hình** | **Dữ liệu đầu vào** | **Khi sử dụng** |
| Gaussian Naive Bayes | Dữ liệu liên tục (continuous) | Khi đặc trưng có phân phối gần chuẩn |
| Multinomial Naive Bayes | Dữ liệu rời rạc, đếm tần suất | Phân loại văn bản, đếm số lần xuất hiện từ |
| Bernoulli Naive Bayes | Dữ liệu nhị phân (0/1) | Khi dữ liệu biểu diễn sự có/không (ví dụ: từ khóa xuất hiện) |

**3. Tại sao Naive Bayes được gọi là “ngây thơ”**

Thuật toán được gọi là **“ngây thơ”** vì giả định rằng **các đặc trưng là độc lập** với nhau trong cùng một lớp.  
Trong thực tế, điều này **hiếm khi đúng**, nhưng giả định giúp **đơn giản hóa việc tính toán**,  
và mô hình vẫn **hoạt động tốt trong nhiều bài toán thực tế**, đặc biệt là trong **phân loại văn bản**.

**4. Ưu điểm và hạn chế của Naive Bayes**

**Ưu điểm:**

* Dễ hiểu và dễ triển khai.
* Tốc độ huấn luyện và dự đoán nhanh, phù hợp với dữ liệu lớn.
* Hiệu quả tốt trong các bài toán **phân loại văn bản**, **lọc spam**, **phân tích cảm xúc**.

**Hạn chế:**

* Giả định độc lập giữa các đặc trưng **thường không đúng** trong thực tế.
* Kém hiệu quả khi **các đặc trưng có tương quan mạnh**.
* Không phù hợp với dữ liệu có **mối quan hệ phi tuyến phức tạp**.

**5. Ví dụ code Gaussian Naive Bayes với Scikit-Learn**

from sklearn.datasets import load\_iris

from sklearn.model\_selection import train\_test\_split

from sklearn.naive\_bayes import GaussianNB

from sklearn.metrics import accuracy\_score, classification\_report

# 1. Tải dữ liệu

X, y = load\_iris(return\_X\_y=True)

# 2. Chia dữ liệu train/test

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# 3. Huấn luyện mô hình

model = GaussianNB()

model.fit(X\_train, y\_train)

# 4. Dự đoán và đánh giá

y\_pred = model.predict(X\_test)

print("Độ chính xác:", accuracy\_score(y\_test, y\_pred))

print(classification\_report(y\_test, y\_pred))

**6. Xử lý dữ liệu phân loại trước khi áp dụng Multinomial Naive Bayes**

Multinomial Naive Bayes yêu cầu dữ liệu đầu vào ở dạng **số** (đếm, tần suất hoặc TF-IDF).  
Nếu dữ liệu là **chuỗi phân loại (categorical)**, cần mã hóa trước bằng LabelEncoder hoặc OneHotEncoder.

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()

df['Gender'] = le.fit\_transform(df['Gender'])

**7. Naive Bayes trong phân loại văn bản (Text Classification)**

Naive Bayes (đặc biệt là **MultinomialNB**) thường được dùng trong bài toán **phân loại văn bản**,  
như **phát hiện email spam**, **phân tích cảm xúc**, **phân loại tin tức**, v.v.

Ví dụ minh họa:

from sklearn.model\_selection import train\_test\_split

from sklearn.feature\_extraction.text import CountVectorizer

from sklearn.naive\_bayes import MultinomialNB

from sklearn.metrics import accuracy\_score

texts = ["Win money now", "Limited offer", "Hi, how are you?", "Let's meet tomorrow"]

labels = [1, 1, 0, 0] # 1 = spam, 0 = không spam

# 1. Chuyển văn bản thành ma trận đếm từ

vectorizer = CountVectorizer()

X = vectorizer.fit\_transform(texts)

# 2. Chia dữ liệu train/test

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, labels, test\_size=0.5, random\_state=42)

# 3. Huấn luyện mô hình

model = MultinomialNB()

model.fit(X\_train, y\_train)

# 4. Dự đoán và đánh giá

y\_pred = model.predict(X\_test)

print("Độ chính xác:", accuracy\_score(y\_test, y\_pred))