TRƯỜNG ĐẠI HỌC HỌC VĂN LANG

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỒ ÁN MÔN HỌC**

**HỌC MÁY**

NGÀNH: CÔNG NGHỆ THÔNG TIN

***Đề tài:***

**ÁP DỤNG THUẬT TOÁN …**

**SVTH: Nguyễn Văn B**

**MSSV: 207CT..**

**GVHD: TRẦN NGỌC VIỆT**

THÀNH PHỐ HỒ CHÍ MINH – NĂM 2025

**LỜI CẢM ƠN**-------------

Lời đầu tiên, nhóm em xin gửi lời cảm ơn chân thành đến thầy Trần Ngọc Việt. Trong quá trình học tập và tìm hiểu bộ môn Nhập môn học máy, chúng em đã được nhận sự quan tâm giúp đỡ, hướng dẫn tận tình, tâm huyết của thầy. Thầy đã giúp

**Nhóm thực hiện báo cáo**

**MỤC LỤC**

[**CHƯƠNG 1. CƠ SỞ LÝ THUYẾT – THUẬT TOÁN NAIVE BAYES CLASSIFIER** 1](#_Toc139835835)

[**1.1. Lý do chọn đề tài** 1](#_Toc139835836)

[**1.2. Cơ sở lý thuyết Naive Bayes Classifier** 1](#_Toc139835837)

[**A.** **Naive Bayes C** 1](#_Toc139835838)

[**B.** **Định lý Bayes** 1](#_Toc139835839)

[**C.** **Một số kiểu mô hình**  2](#_Toc139835840)

[**D.** **Ứng dụng thuật toán**  3](#_Toc139835841)

[**1.3 Kết luận** 3](#_Toc139835842)

[**CHƯƠNG 2. ÁP DỤNG THUẬT TOÁN** 4](#_Toc139835843)

[**2.1 Phát biểu bài toán:** 4](#_Toc139835844)

[**2.2 Minh họa bài toán** 4](#_Toc139835845)

[**2.3 Mã nguồn** 6](#_Toc139835846)

[**2.4 Kết luận** 7](#_Toc139835847)

[**CHƯƠNG 3. KẾT LUẬN** 8](#_Toc139835848)

[**TÀI LIỆU THAM KHẢO** 9](#_Toc139835849)

# CHƯƠNG 1. CƠ SỞ LÝ THUYẾT – THUẬT TOÁN NAIVE BAYES CLASSIFIER

## **1.1. Lý do chọn đề tài**

Với đề tài Naive Bayes Classifier đây là một thuật toán đơn giản và nhanh chóng. Vì thuật toán không

## **1.2. Cơ sở lý thuyết Naive Bayes Classifier**

### **Naive Bayes Classifier là gì?**

Naive Bayes

**Định lý Bayes**

Định lý Bayes cho phép tính xác suất xảy ra của một sự kiện ngẫu nhiên A khi biết sự kiện liên quan B đã xảy ra. Xác suất xảy ra A của riêng nó, không quan tâm đến B. Kí hiệu là P(A).

* Xác suất xảy ra B của riêng nó, không quan tâm đến A. Kí hiệu là P(B).
* Xác suất xảy ra B khi biết A xảy ra. Kí hiệu là P(B|A). Đại lượng này gọi là khả năng (likelihood) xảy ra B khi biết A đã xảy ra.

Ở trên ta có thể thấy xác suất xảy ra của sự kiện A phụ thuộc và xác suất của sự kiện B, nhưng trong thực tế xác suất A có thể phụ thuộc vào xác suất của các giả thuyết khác có thể là B1,B2,B3,…. Bn.

### **Một số kiểu mô hình Naive Bayes**

Hình 1. 2: Định lý Bayes mở rộng

* **Multinomial Naive Bayes**
* Mô hình này chủ yếu được sử dụng trong phân loại văn bản. Đặc trưng đầu vào ở đây chính là tần suất xuất hiện của từ trong văn bản đó.
* **Bernoulli Naive Bayes**
* Mô hình này được sử dụng khi các đặc trưng đầu vào chỉ nhận giá trị nhị phân 0 và 1 (phân bố Bernoulli).

Hình 1. 3: Gaussian Naïve Bayes

### **Ứng dụng thuật toán Naïve Bayes**

Thuật toán Naïve Bayes Classifier được áp dụng vào các loại ứng dụng sau:

* ***Real time Prediction: NBC chạy khá nhanh nên nó thích hợp áp dụng ứng dụng nhiều vào các ứng dụng chạy thời gian thực, như hệ thống cảnh báo phát hiện sự cố...***
* ***Multi class Prediction:*** Nhờ vào định lý Bayes mở rộng ta có thể ứng dụng vào các loại 3 ứng dụng đa dự đoán, tức là ứng dụng có thể dự đoán nhiều giả thuyết mục tiêu.
* ***Text classification/ Spam Filtering/ Sentiment Analysis***: NBC cũng rất thích hợp cho các hệ thống phân loại văn bản hay ngôn ngữ tự nhiên vì tính chính xác của nó lớn hơn các thuật toán khác. Ngoài ra các hệ thống chống thư rác cũng rất ưu chuộng thuật toán này. Và các hệ thống phân tích tâm lý thị trường cũng áp dụng NBC để tiến hành phân tích tâm lý người dùng ưu chuộng hay không ưu chuộng các loại sản phẩm nào từ việc phân tích các thói quen và hành động của khách hàng.
* ***Recommendation System***: Naive Bayes Classifier được sử dụng rất nhiều để xây dựng hệ thống gợi ý

## **1.3 Kết luận**

Naive Bayes mô hình là đơn vị phân tích mô hình dễ

# CHƯƠNG 2. ÁP DỤNG THUẬT TOÁN NAÏVE BAYES CLASSIFIER CHO BÀI TOÁN DỰ ĐOÁN MÓN ĂN CỦA MIỀN BẮC NAM

## **2.1 Phát biểu bài toán:**

* Bài toán phân loại món ăn của hai miền Bắc và Nam dựa trên Naïve Bayes Classifier là một bài toán trong xử lý

## **2.2 Minh họa bài toán**

|

Class N:

d = |V| = 8

* NN = 8

16 = NN + |V|

* Test:

E6 = [1, 2, 1, 1, 0, 0, 1, 0]

P(B|x6) P(B) d­i=1 P(xi / B) = \* \*

=2,626390145. **10-6**

=1,430511475. **10-7**

* **P(e6|B) > P(e6|N) nên e6 thuộc lớp B**

Xác suất:

E6 sẽ thuộc miền Bắc

Với kết quả này sẽ tương đồng với kết quả bài toán đưa ra trong đoạn mã ở phần bên dưới là phần mã nguồn

## **2.3 Mã nguồn**

Phần này sẽ áp dụng thuật toán Naïve Bayes Classifier trên về món ăn 2 miền Bắc Nam thi những đoạn mã Code:

import math

from sklearn.naive\_bayes import MultinomialNB

import numpy as np

# Xác định dữ liệu

e1 = [2, 0, 1, 1, 1, 1, 0, 0]     # dữ liệu đầu vào được chuyển hóa 0,1,2

e2 = [1, 1, 0, 1, 1, 0, 1, 1]

e3 = [0, 1, 0, 0, 0, 0, 1, 1]

e4 = [0, 0, 0, 0, 1, 1, 1, 1]

e5 = [0, 0, 1, 0, 1, 1, 1, 0]

train\_data = np.array([e1, e2, e3, e4, e5]) #dl đầu vào

ket\_qua = np.array(['B', 'B', 'B', 'N', 'N']) #gán nhãn dl

# Xác định dữ liệu thử nghiệm

e6 = np.array([[1,2,1,1,0,0,1,0]])         #Dữ liệu kiểm tra giá trị đặc trưng

# mô hình tính toán

ml = MultinomialNB(alpha=1)

# Huấn luyện bộ phân loại Naive Bayes

ml.fit(train\_data, ket\_qua)

# Dự đoán xác xuất của e6

print('Probability of e6:', ml.predict\_proba(e6))

# Dự đoán lớp e6

print('Predicting class of e6:', str(ml.predict(e6)[0]))

## **2.4 Kết luận**

Mô hình Naive Bayes là mô hình phân lớp đơn giản dễ cài đặt, có tốc độ xử lý nhanh. Tuy nhiên có

# CHƯƠNG 3. KẾT LUẬN

Mô hình Naive Bayes là mô hình phân lớp đơn giản dễ cài đặt, có tốc độ xử lý nhanh. Tuy nhiên có nhược điểm lớn là yêu cầu các

# TÀI LIỆU THAM KHẢO

1. [Hướng dẫn Naive Bayes: Phân loại Naive Bayes trong Python](https://helpex.vn/article/huong-dan-naive-bayes-phan-loai-naive-bayes-trong-python-5c6b22c7ae03f628d053c334) : [https://helpex.vn/article/huong-dan-naive-bayes-phan-loai-naive-bayes-trong-python-](https://helpex.vn/article/huong-dan-naive-bayes-phan-loai-naive-bayes-trong-python-5c6b22c7ae03f628d053c334) [5c6b22c7ae03f628d053c334](https://helpex.vn/article/huong-dan-naive-bayes-phan-loai-naive-bayes-trong-python-5c6b22c7ae03f628d053c334) Ngày 30/6/2023.
2. Naïve Bayes Algorithm: Everything You Need to Know: <https://www.kdnuggets.com/2020/06/naive-bayes-algorithm-everything.html> Ngày 30/6/2023
3. Thuật toán phân loại Naive Bayes và ứng dụng: <https://whitehat.vn/threads/thuat-toan-phan-loai-naive-bayes-va-ung-dung.13775/> Ngày 1/7/2023