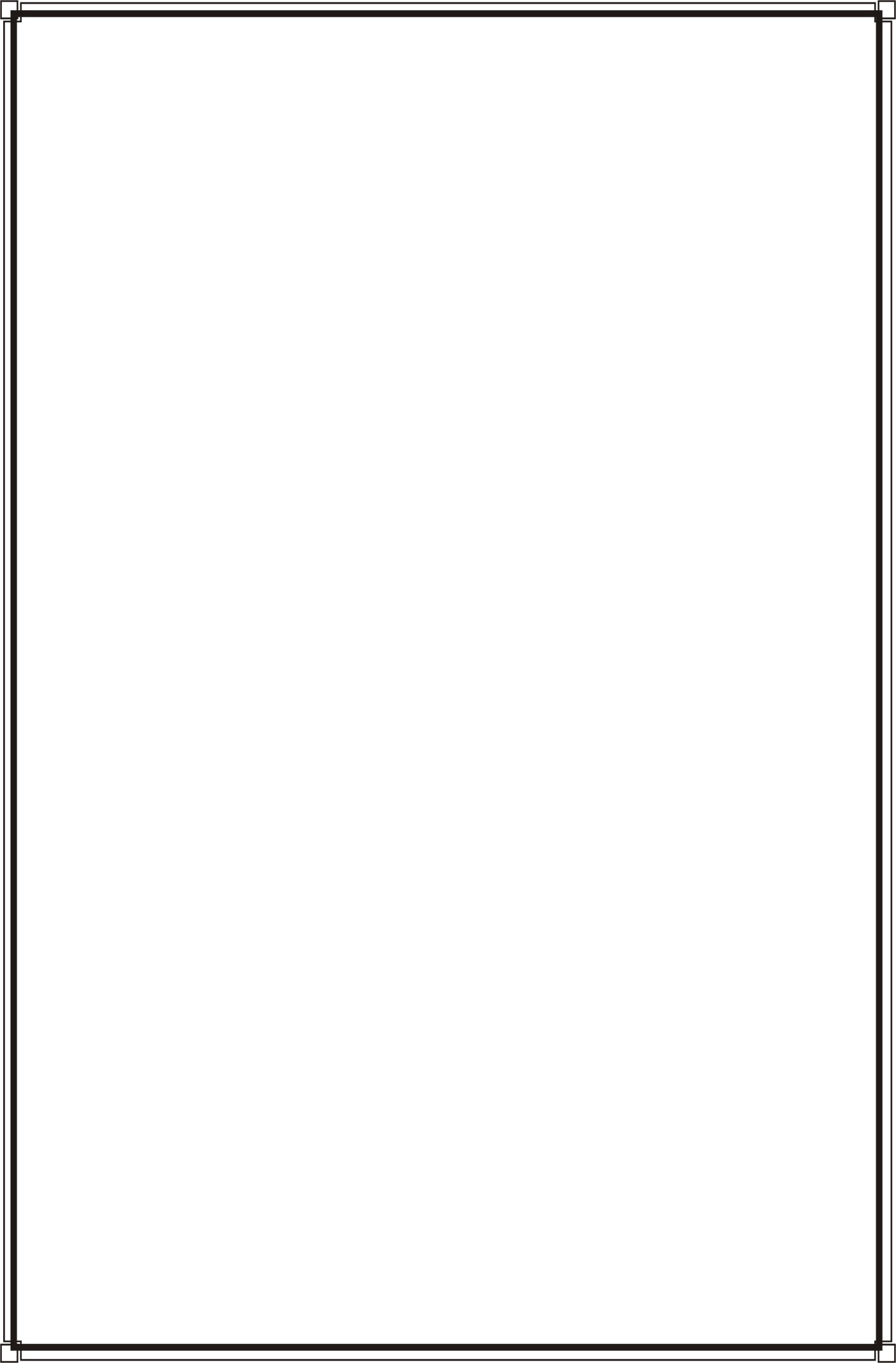
**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP. HCM**

**KHOA KHOA HỌC ỨNG DỤNG**

**-----🙠☯🙢-----**

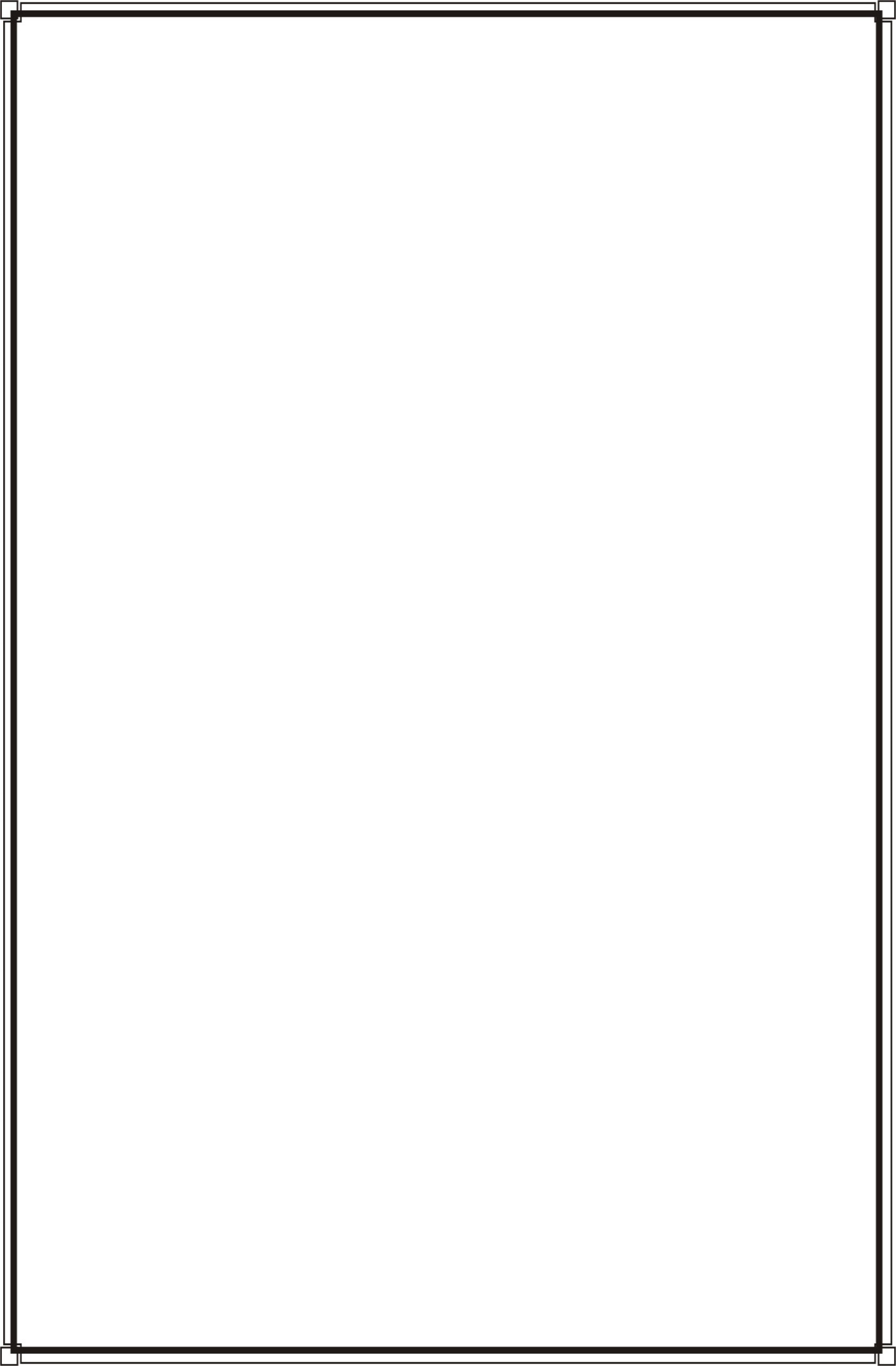


**NHÓM 7**

**PHƯƠNG PHÁP HỌC KHÔNG CÓ GIÁM SÁT VÀ HỌC CÓ GIÁM SÁT (TÌM HIỂU 2 THUẬT TOÁN ĐẠI DIỆN LÀ KMEAN VÀ KNN) VÀ ỨNG DỤNG.**

**BÀI TẬP NHÓM MÔN:TRÍ TUỆ NHÂN TẠO**

**TP. HCM, NĂM 2022**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP. HCM**

**KHOA KHOA HỌC ỨNG DỤNG**

**-----🙠☯🙢-----**



**NHÓM 7**

**PHƯƠNG PHÁP HỌC KHÔNG CÓ GIÁM SÁT VÀ HỌC CÓ GIÁM SÁT (TÌM HIỂU 2 THUẬT TOÁN ĐẠI DIỆN LÀ KMEAN VÀ KNN) VÀ ỨNG DỤNG.**

**BÀI TẬP NHÓM MÔN:TRÍ TUỆ NHÂN TẠO**

**GVHD:Trần Đình Toàn**

**NHÓM THỰC HIỆN:**

1.Vũ Minh Nghĩa - 2001206981

**TP.HCM, NĂM 2022**

# BẢNG ĐÁNH GIÁ KẾT QUẢ THỰC HIỆN CÔNG VIỆC NHÓM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **STT** | **Họ và tên** | **Công việc đảm nhận** | **Nhóm đánh giá kết quả** | **Ghi chú** |
| 1 | Vũ Minh Nghĩa | Tìm Hiểu Thuật Toán, Nguyên lý thuật toán | 100% |  |
| 2 | Vũ Minh Nghĩa | Viết Chương Trình Minh Hoạ Python | 100% |  |
| 3 | Vũ Minh Nghĩa | Viết Báo Cáo Và Thuyết Trình | 100% |  |

# Lời cam đoan

Chúng em xin cam đoan đề tài *“Tìm hiểu tìm hiểu giải thuật Robinson, Và viết chương trình minh hoạ”* do nhóm 7 nghiên cứu và thực hiện, không sao chép của bất kỳ nhóm hay của ai khác.

Chúng em đã kiểm tra dữ liệu theo quy định hiện hành.

Kết quả bài làm đề tài “*Robinson, Và viết chương trình ứng dụng của thuật toán*” do nhóm 7 nghiên cứu và thực hiện là trung thực và không sao chép từ bất kỳ bài tập của nhóm khác.

Các tài liệu được sử dụng trong tiểu luận có nguồn gốc, xuất xứ rõ ràng.

**(Ký và ghi rõ họ tên)**

Nghĩa

Vũ Minh Nghĩa

# LỜI CẢM ƠN

Trong quá trình học tập và thực hiện bài tiểu luận chúng em xin chân thành cảm ơn Khoa Công Nghệ Thông Tin,Trường Đại Học Công Nghiệp Thực Phẩm Tp.Hồ Chí Minh.

Trước hết nhóm em cảm ơn thầy Trần Đình Toàn là giảng viên đứng lớp môn Trí Tuệ Nhân Tạo của bọn em, dù dịch bệnh khó khăn vẫn tạo điều kiện cho bọn em làm bài của mình một cách tốt nhất, luôn nhiệt tình với bọn em dù trong hoàn cảnh nào đi nữa.

Mặc dù nhóm em đã có nhiều cố gắng thực hiện đề tài, có thể còn nhũng hạn chế và thiếu sót.Chúng em rất mong nhận được sự thông cảm,đóng góp ý kiến và giúp đỡ từ quý thầy cô và các bạn.

Chúng em xin cảm ơn!

Nhóm 7 thực hiện.

# MỤC LỤC

[BẢNG ĐÁNH GIÁ KẾT QUẢ THỰC HIỆN CÔNG VIỆC NHÓM i](#_Toc118665533)

[Lời cam đoan ii](#_Toc118665534)

[LỜI CẢM ƠN iii](#_Toc118665535)

[MỤC LỤC iv](#_Toc118665536)

[PHẦN MỞ ĐẦU 1](#_Toc118665537)

[1. Thuật toán GTS 1](#_Toc118665538)

[2. Mã giả 1](#_Toc118665539)

[3. Các bước giải thuật toán 4](#_Toc118665540)

[4. Ví dụ 5](#_Toc118665541)

[5. Ứng dụng 6](#_Toc118665542)

[6. Thuật toán tô màu 6](#_Toc118665543)

[KẾT LUẬN 9](#_Toc118665544)

[TÀI LIỆU THAM KHẢO 10](#_Toc118665545)

# PHẦN MỞ ĐẦU

1. Giới thiệu thuật toán KNN

Thuật toán tiệm cận, hay thuật toán phân loại K-Nearest Neighbor (KNN, K-NearestNeighbor) là một trong những phương pháp đơn giản nhất trong công nghệ phân loại khai thác dữ liệu. Cái gọi là K hàng xóm gần nhất có nghĩa là K hàng xóm gần nhất, có nghĩa là mỗi mẫu có thể được đại diện bởi K hàng xóm gần nhất của nó. Thuật toán hàng xóm gần nhất là một phương pháp phân loại từng bản ghi trong tập dữ liệu.

Phương pháp k-láng giềng gần nhất là một phương pháp phân loại và hồi quy cơ bản, và là một phương pháp thường được sử dụng trong các phương pháp học có giám sát. Thuật toán k-láng giềng gần nhất giả định rằng một tập dữ liệu huấn luyện được cung cấp và các thể loại thể hiện trong đó đã được xác định. Khi phân loại, một phiên bản mới được dự đoán bằng phương pháp biểu quyết đa số theo k loại phiên bản đào tạo láng giềng gần nhất của nó.

2. Ý tưởng cốt lõi của thuật toán KNN

Ý tưởng cốt lõi của thuật toán KNN là nếu hầu hết K mẫu láng giềng gần nhất của một mẫu trong không gian đặc trưng thuộc về một danh mục nhất định, thì mẫu đó cũng thuộc danh mục này và có các đặc điểm của các mẫu trong danh mục này. Khi xác định quyết định phân loại, phương pháp này chỉ xác định loại mẫu được phân chia theo loại của một hoặc một số mẫu gần nhất. Phương pháp KNN chỉ liên quan đến một số lượng rất nhỏ các mẫu liền kề khi đưa ra quyết định về danh mục. Vì phương pháp KNN chủ yếu dựa vào các mẫu giới hạn xung quanh hơn là phương pháp phân biệt miền lớp để xác định danh mục mà nó thuộc về, nên phương pháp KNN chính xác hơn các phương pháp khác để phân chia các tập hợp mẫu khi miền lớp giao nhau hoặc chồng lên nhau nhiều hơn.cho phù hợp.

3. Ba yếu tố của thuật toán KNN

Ba yếu tố của phương pháp k-láng giềng gần nhất: thước đo khoảng cách, lựa chọn giá trị k và quy tắc quyết định phân loại. Các thước đo khoảng cách thường được sử dụng là khoảng cách Euclide và khoảng cách pL tổng quát hơn. Khi giá trị của k nhỏ, mô hình k-láng giềng gần nhất phức tạp hơn và có xu hướng bị thừa; khi giá trị của k lớn, mô hình k-láng giềng gần nhất đơn giản hơn và có xu hướng bị thiếu. Do đó, việc lựa chọn giá trị k sẽ có tác động đáng kể đến kết quả phân loại. Việc lựa chọn giá trị k phản ánh sự đánh đổi giữa sai số xấp xỉ và sai số ước lượng, và k tối ưu thường được chọn bằng xác thực chéo.

4. Ưu nhược điểm của thuật toán KNN

thuận lợi

1. Đơn giản, dễ hiểu, dễ thực hiện, không cần ước lượng tham số, không cần đào tạo 2. Thích hợp để phân loại các sự kiện hiếm gặp 3. Đặc biệt phù hợp với bài toán đa phân loại (đa phương thức, đối tượng có nhiều nhãn loại ), kNN so với hiệu suất SVM tốt hơn.

sự thiếu sót

1. Nhược điểm chính của thuật toán này trong phân loại là khi các mẫu không cân bằng, chẳng hạn như cỡ mẫu của một lớp lớn, trong khi cỡ mẫu của các lớp khác nhỏ, có thể dẫn đến khi một mẫu mới được nhập vào, the sample's Các mẫu của lớp số lượng lớn trong số K láng giềng chiếm đa số. 2. Một nhược điểm khác của phương pháp này là khối lượng tính toán lớn, vì để phân loại từng văn bản, phải tính khoảng cách đến tất cả các mẫu đã biết để có được K hàng xóm gần nhất của nó.

# KẾT LUẬN

Qua đề tài nghiên cứu này, chúng em đã rút ra được nhiều bài học về ngôn ngữ lập trình Python, các giải thuật Robinson và. Tìm hiểu về cấu tạo của giải thuật Robinson trên ngôn ngữ Python. Tuy còn nhiều mặt hạn chế về số lượng thành viên nhưng em đã cố gắng hoàn thành tốt bài tập tuần lần này. Từ đó rút ra được nhiều bài học quan trọng để phát triển tốt hơn trong học tập.

# TÀI LIỆU THAM KHẢO