TRƯỜNG ĐẠI HỌC HỌC VĂN LANG

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỒ ÁN MÔN HỌC**

**CÁC HỆ HỖ TRỢ RA QUYẾT ĐỊNH (71ITAI40303)**

**Hệ thống khuyến nghị nghề nghiệp cho sinh viên**

**GVHD: TS. NGUYỄN THÁI ANH**

**Nhóm sinh viên thực hiện (Họ tên - Mã SV):**

1. Lê Duy Khang - 2274802010374
2. Huỳnh Nhật Khánh Quỳnh– 2274802010742
3. Dư Nguyễn Hà Mi– 2274802010525
4. Phạm Nguyễn Hoài Phương - 2274802010702
5. Trần Hữu Luân - 2274802010520

TP. Hồ Chí Minh – năm 2024

Contents

[**CHƯƠNG I: GIỚI THIỆU** 3](#_Toc182486150)

[**1. Tình hình hiện tại:** 3](#_Toc182486151)

[**2. Mục tiêu của hệ thống:** 3](#_Toc182486152)

[**3. Đối tượng nghiên cứu:** 4](#_Toc182486153)

[**CHƯƠNG II: CƠ SỞ LÝ THUYẾT** 5](#_Toc182486154)

[**1. Các yếu tố ảnh hưởng đến nghề nghiệp:** 5](#_Toc182486155)

[**2. Mô hình KNN:** 6](#_Toc182486156)

[**2.1 Tổng quan về mô hình KNN** 6](#_Toc182486157)

[**2.2 Xây dựng mô hình KNN:** 6](#_Toc182486158)

[**3. Mô hình hệ thống:** 7](#_Toc182486159)

[**3.1 Kiến trúc hệ thống:** 7](#_Toc182486160)

[**3.2 Quy trình hoạt động của hệ thống:** 8](#_Toc182486161)

[**4. Thiết kế giao diện người dùng bằng thư viện Flask:** 8](#_Toc182486162)

[**4.1 Thiết kế giao diện người dùng** 8](#_Toc182486163)

[**4.2 Quy trình hoạt động của giao diện:** 9](#_Toc182486164)

[**CHƯƠNG III: PHƯƠNG PHÁP NGHIÊN CỨU** 9](#_Toc182486165)

[**1. Thu thập dữ liệu:** 9](#_Toc182486166)

[**2. Xử lý dữ liệu:** 10](#_Toc182486167)

[**3. Trực quan hóa dữ liệu:** 10](#_Toc182486168)

[**4. Xuất dữ liệu:** 11](#_Toc182486169)

[**CHƯƠNG IV: KẾT QUẢ** 11](#_Toc182486170)

[**Mã nguồn:** 13](#_Toc182486171)

[**app.py** 13](#_Toc182486172)

[**Database** 14](#_Toc182486173)

[**Information** 16](#_Toc182486174)

[**Mô hình KNN** 19](#_Toc182486175)

[**CHƯƠNG V: KẾT LUẬN** 20](#_Toc182486176)

[**TÀI LIỆU KHAM THẢO** 21](#_Toc182486177)

# CHƯƠNG I: GIỚI THIỆU

## **1. Tình hình hiện tại:**

Trong bối cảnh xã hội ngày nay, nhu cầu lựa chọn nghề nghiệp trở nên cấp thiết đối với sinh viên. Sự phát triển không ngừng của các ngành kinh tế, công nghệ và xã hội đã khiến việc lựa chọn nghề nghiệp trở thành một quyết định phức tạp hơn bao giờ hết. Các ngành nghề ngày càng đa dạng và phức tạp, cùng với sự thay đổi liên tục của thị trường lao động, khiến sinh viên phải đối mặt với nhiều khó khăn trong việc đưa ra quyết định nghề nghiệp phù hợp. Điều này đòi hỏi sinh viên phải xem xét kỹ lưỡng các yếu tố như sở thích cá nhân, năng lực và nhu cầu thị trường lao động.

Tuy nhiên, nhiều sinh viên vẫn gặp khó khăn trong việc tìm ra lựa chọn phù hợp cho bản thân. Chúng ta có thể dễ dàng bắt gặp những tiêu đề báo chí như: “Sai lầm khi đua vào ngành hot, nghề mốt”, “Chọn ngành nghề sai, giới trẻ phải trả giá bằng thanh xuân”... Những dòng tiêu đề này phản ánh thực trạng nhiều năm qua về việc lựa chọn nghề nghiệp chưa phù hợp, dẫn đến sai lầm trong việc định hướng tương lai.

Một trong những thách thức lớn mà sinh viên gặp phải là thiếu thông tin chính xác về các ngành nghề. Nhiều sinh viên còn mơ hồ về việc ngành nghề nào phù hợp với tính cách và sở thích cá nhân của mình. Điều này dẫn đến việc chọn ngành chỉ dựa trên xu hướng xã hội, mong muốn của gia đình hoặc quyết định vội vàng mà không có sự hiểu biết đầy đủ về nghề nghiệp họ sẽ theo đuổi. Hậu quả là nhiều sinh viên phải chuyển ngành, bỏ học giữa chừng, hoặc tốt nghiệp mà không có hứng thú hay động lực để theo đuổi nghề đã học. Tình trạng này không chỉ gây lãng phí thời gian, tiền bạc mà còn ảnh hưởng lớn đến sự phát triển nghề nghiệp sau này.

Bên cạnh đó, nhiều sinh viên còn chưa nắm rõ yêu cầu cụ thể của từng nghề nghiệp cũng như triển vọng việc làm trong tương lai. Điều này dẫn đến việc theo học những ngành không phù hợp với năng lực hoặc nhu cầu cá nhân, gây ra sự lãng phí lớn về thời gian và chi phí, đôi khi còn là sự thất vọng khi họ không đạt được thành công trong sự nghiệp. Vì vậy, sinh viên cần có cái nhìn rõ ràng, toàn diện hơn về các lựa chọn nghề nghiệp để có thể đưa ra những quyết định phù hợp với bản thân, giúp họ vững bước hơn trên con đường phát triển sự nghiệp tương lai.

**2. Mục tiêu của hệ thống:**

Hệ thống khuyến nghị lựa chọn nghề nghiệp được thiết kế nhằm mục đích giải quyết các vấn đề sinh viên thường gặp phải khi đối mặt với việc lựa chọn nghề nghiệp. Mục tiêu chính của hệ thống là giúp sinh viên xác định một cách rõ ràng và chính xác các ngành nghề phù hợp với sở thích, tính cách và năng lực cá nhân, từ đó đưa ra quyết định nghề nghiệp đúng đắn, tránh lãng phí thời gian và nguồn lực vào những con đường không phù hợp.

Hệ thống sẽ cung cấp các công cụ đánh giá, phân tích sở thích, kỹ năng, và đặc điểm cá nhân của từng sinh viên. Các bài kiểm tra trắc nghiệm, bảng câu hỏi về sở thích nghề nghiệp, cũng như các bài kiểm tra tâm lý về tính cách sẽ được tích hợp vào hệ thống để đưa ra các gợi ý chính xác và cá nhân hóa cho từng người dùng. Mục tiêu là không chỉ giúp sinh viên biết mình phù hợp với những nghề nghiệp nào mà còn giải thích rõ lý do vì sao những nghề đó lại phù hợp với họ, từ đó giúp sinh viên có sự tự tin và hiểu biết sâu sắc hơn về bản thân.

Ngoài việc đánh giá sở thích và năng lực cá nhân, hệ thống cũng cung cấp một lượng thông tin chi tiết và cập nhật về các ngành nghề. Điều này bao gồm mô tả công việc, các yêu cầu về trình độ, kỹ năng cần thiết, cũng như triển vọng phát triển nghề nghiệp trong tương lai. Hệ thống còn tích hợp các dữ liệu về thị trường lao động, giúp sinh viên hiểu rõ hơn về nhu cầu tuyển dụng của từng ngành nghề trong thời điểm hiện tại và trong tương lai, từ đó họ có thể đưa ra quyết định phù hợp với thực tế.

Tóm lại, hệ thống khuyến nghị lựa chọn nghề nghiệp không chỉ giúp sinh viên đưa ra quyết định sáng suốt mà còn đóng vai trò quan trọng trong việc xây dựng kế hoạch nghề nghiệp lâu dài, giúp họ sẵn sàng đối mặt với những thách thức của thị trường lao động và phát triển bền vững trong tương lai.

**3. Đối tượng nghiên cứu:**

Đối tượng nghiên cứu của hệ thống khuyến nghị lựa chọn nghề nghiệp không gì khác ngoài những sinh viên, những người đang trong hành trình khám phá và xác định con đường nghề nghiệp tương lai của bản thân. Trong bối cảnh đầy thử thách này, hệ thống hướng tới một số nhóm đối tượng chính:

Trước tiên, sinh viên năm nhất và năm hai là nhóm đối tượng đặc biệt quan trọng. Đây là giai đoạn mà các bạn thường phải đối mặt với quyết định lớn lao: chọn chuyên ngành nào và hướng đi nghề nghiệp ra sao. Tuy nhiên, không ít sinh viên trong độ tuổi này vẫn chưa rõ ràng về sở thích, khả năng của bản thân hoặc thậm chí thiếu thông tin cần thiết về các ngành nghề. Điều này dẫn đến không ít khó khăn trong việc ra quyết định. Chính vì vậy, hệ thống khuyến nghị sẽ trở thành một người bạn đồng hành, giúp các bạn đánh giá các yếu tố như sở thích, tính cách và kết quả học tập, từ đó đưa ra những gợi ý nghề nghiệp phù hợp nhất.

Tiếp theo, sinh viên năm cuối và mới tốt nghiệp cũng là một nhóm đáng chú ý. Sau khi hoàn thành chương trình học, không ít sinh viên bỗng dưng cảm thấy hoang mang về ngành học mình đã chọn, hoặc nhận ra rằng con đường mình đang đi có thể không thực sự phù hợp với bản thân. Hệ thống này sẽ là công cụ hữu ích giúp các bạn trẻ này đánh giá lại các yếu tố cá nhân, đồng thời mở ra những lựa chọn nghề nghiệp thay thế và định hướng các khóa học bổ sung, nhằm giúp họ tìm được mục tiêu nghề nghiệp mới đúng với bản thân.

Bên cạnh đó, những sinh viên đang cân nhắc chuyển ngành cũng không thể bị bỏ qua. Một số bạn có thể cảm thấy không hài lòng hoặc không phù hợp với ngành học hiện tại, nhưng lại chưa biết nên chuyển sang ngành nào khác. Hệ thống sẽ giúp các bạn đánh giá lại những đặc điểm cá nhân, từ đó đưa ra các đề xuất về những ngành nghề khác có thể phù hợp hơn với khả năng và sở thích của mình.

Cuối cùng, phụ huynh và những người tư vấn hướng nghiệp dù không phải là đối tượng trực tiếp sử dụng hệ thống, nhưng phụ huynh và người tư vấn hướng nghiệp cũng là những người có ảnh hưởng lớn đến quyết định nghề nghiệp của sinh viên. Hệ thống cung cấp công cụ và dữ liệu để họ tham khảo và tư vấn cho sinh viên một cách chính xác và hiệu quả hơn.

Nhìn chung, hệ thống hướng tới việc khuyến nghị tất cả các đối tượng đang gặp khó khăn hoặc có nhu cầu trong việc lựa chọn và định hướng nghề nghiệp, từ đó giúp họ đưa ra quyết định phù hợp với khả năng và đam mê của bản thân.

**CHƯƠNG II: CƠ SỞ LÝ THUYẾT**

**1. Các yếu tố ảnh hưởng đến nghề nghiệp:**

Lựa chọn nghề nghiệp là một hành trình không đơn giản, nó chịu ảnh hưởng từ nhiều yếu tố khác nhau và mang tính cá nhân sâu sắc. Mỗi người đều có những ước mơ, hoài bão, và những đặc điểm riêng, khiến cho quá trình lựa chọn nghề nghiệp trở thành một câu chuyện đầy màu sắc và phức tạp.

Trước hết, **sở thích cá nhân** là một yếu tố vô cùng quan trọng. Chúng ta thường lựa chọn công việc dựa trên những gì mình thích làm, những hoạt động mang lại cho ta niềm vui và sự hứng khởi. Việc hiểu rõ sở thích của bản thân giống như việc ta xác định được phương hướng giữa biển cả rộng lớn của các nghề nghiệp. Các bài kiểm tra như MBTI giúp sinh viên hiểu sâu hơn về tính cách của mình, từ đó đưa ra những định hướng nghề nghiệp phù hợp.

Tiếp theo là **năng lực và kỹ năng** của mỗi người. Không chỉ có đam mê, năng lực và kỹ năng của bản thân là điều kiện cần để ta có thể thành công trong một công việc nào đó. Việc tự đánh giá chính xác năng lực của bản thân giúp sinh viên không chỉ tự tin hơn mà còn biết mình có khả năng đáp ứng những yêu cầu của ngành nghề nào. Đó là chìa khóa giúp họ mở cánh cửa đến với thành công.

**Gia đình và xã hội** cũng để lại những dấu ấn rõ rệt trong quyết định nghề nghiệp của sinh viên. Gia đình là nơi hình thành những quan niệm đầu tiên về công việc và nghề nghiệp, vì vậy, sự kỳ vọng hay mong muốn của cha mẹ có thể ảnh hưởng mạnh mẽ đến sự lựa chọn của con cái. Bên cạnh đó, xã hội và xu hướng chung của thời đại cũng có thể dẫn dắt các bạn trẻ đến những lựa chọn mà đôi khi không thực sự phù hợp với khả năng và đam mê của bản thân. Chính vì vậy, việc hiểu rõ những tác động này giúp sinh viên có cái nhìn khách quan và đưa ra lựa chọn sáng suốt hơn.

**2. Mô hình KNN:**

**2.1 Tổng quan về mô hình KNN**

KNN viết tắt của K-Nearest Neighbors, là một phương pháp học máy dùng để phân loại hoặc hồi quy, dựa trên nguyên tắc rằng các đối tượng có đặc điểm tương tự sẽ ở gần nhau trong không gian đặc trưng. Trong KNN, việc phân lớp hoặc dự đoán giá trị của một đối tượng mới được thực hiện bằng cách tính toán khoảng cách giữa đối tượng đó và tất cả các đối tượng trong dữ liệu huấn luyện. Trong thế giới của học máy, KNN không cần phải huấn luyện phức tạp, nó chỉ cần quan sát những gì xung quanh và đưa ra câu trả lời gần gũi nhất. KNN có khả năng sử dụng trực tiếp dữ liệu đã có mà không cần phải xây dựng mô hình dự đoán trước đó.

Về nguyên lý hoạt động khi một điểm dữ liệu mới xuất hiện, KNN không thực hiện việc học theo cách thông thường như các mô hình khác mà thay vào đó nó dựa vào việc tìm kiếm những điểm dữ liệu tương tự trong tập dữ liệu huấn luyện. Những điểm này được chọn dựa trên khoảng cách, và thuật toán sẽ dự đoán giá trị của điểm mới này bằng cách "tham khảo ý kiến" của các điểm lân cận.

Điểm mạnh của KNN là sự đơn giản, không cần phải xây dựng những mô hình phức tạp hay lập trình các thuật toán cầu kỳ. Chỉ cần chọn một giá trị thích hợp và tính khoảng cách, KNN sẽ giúp dự đoán dữ liệu đã có sẵn. Đặc biệt KNN rất hiệu quả trong các tập dữ liệu nhỏ, nơi mà sự đơn giản của nó trở thành một lợi thế. Trong các trường hợp này KNN có thể cho kết quả tốt mà không cần phải tinh chỉnh quá nhiều siêu tham số. Các ứng dụng của KNN cũng rất đa dạng từ việc phân loại hình ảnh, nhận diện giọng nói, đến các hệ thống gợi ý sản phẩm và thậm chí là dự đoán giá trị như bất động sản hay chứng khoán.

Một trong những khó khăn lớn nhất là khi dữ liệu trở nên lớn và phức tạp. Khi cần so sánh điểm dữ liệu mới với hàng ngàn hoặc hàng triệu điểm dữ liệu cũ, KNN có thể trở nên rất chậm chạp. Việc tính toán khoảng cách cho tất cả các điểm này làm tốn rất nhiều tài nguyên.

Dù có những hạn chế KNN vẫn là một trong những thuật toán cơ bản và dễ hiểu nhất trong học máy. Nó không chỉ giúp người mới học dễ dàng tiếp cận, mà còn mang lại hiệu quả bất ngờ trong nhiều ứng dụng thực tiễn.

**2.2 Xây dựng mô hình KNN:**

Trước hết việc thu thập dữ liệu là điều kiện tiên quyết. Dữ liệu này bao gồm thông tin cá nhân của sinh viên như: điểm số, kỹ năng, sở thích, và tính cách. Đồng thời thu thập thông tin liên quan đến các ngành nghề bao gồm yêu cầu kỹ năng, sở thích tương ứng, và cơ hội nghề nghiệp trong tương lai. Dữ liệu lịch sử về các sinh viên trước đây cùng với các lựa chọn nghề nghiệp của họ và mức độ thành công trong công việc cũng sẽ là nguồn dữ liệu quan trọng để mô hình học hỏi.

Khi đã có dữ liệu, việc làm sạch và tiền xử lý là bước tiếp theo. Để mô hình KNN hoạt động hiệu quả cần chuẩn hóa các thuộc tính khác nhau ví dụ như là điểm số hay kỹ năng, về cùng một thang giá trị. Điều này giúp đảm bảo tính đồng nhất khi tính toán khoảng cách giữa các đối tượng. Ngoài ra, những thông tin dạng phân loại như là sở thích hay ngành học cũng cần được chuyển đổi thành dạng số hoặc biến nhị phân để dễ dàng xử lý.

Sau khi dữ liệu đã được chuẩn bị, quá trình xây dựng mô hình KNN sẽ bắt đầu. Trước hết thì cần xác định tham số **K**, tức số lượng "láng giềng gần nhất" mà mô hình sẽ xem xét khi đưa ra dự đoán. Tiếp đó, mô hình KNN sẽ tính khoảng cáchgiữa sinh viên cần được tư vấn và các sinh viên khác trong dữ liệu đã thu thập. **Khoảng cách Euclid** là một phương pháp phổ biến để đo lường sự khác biệt giữa các điểm dữ liệu. Sau khi tính toán, KNN sẽ tìm ra **K sinh viên gần nhất** với sinh viên cần tư vấn dựa trên sự tương đồng về điểm số, sở thích, hoặc kỹ năng. Cuối cùng với **bài toán phân loại nghề nghiệp**, mô hình sẽ dựa vào những ngành nghề đã được chọn bởi các sinh viên tương đồng. Ngành nghề nào xuất hiện nhiều nhất trong số K sinh viên gần nhất sẽ được đề xuất cho sinh viên mới.

Sau khi mô hình KNN đã được tinh chỉnh và đánh giá, bước cuối cùng là tích hợp nó vào hệ thống. Khi hệ thống tiếp nhận thông tin về học sinh, mô hình KNN sẽ phân tích và tư vấn ngành nghề phù hợp dựa trên dữ liệu từ các sinh viên trước đây có những điểm chung về sở thích và năng lực. Mô hình KNN sẽ giúp hệ thống đưa ra những gợi ý nghề nghiệp một cách sát thực tế nhờ vào việc so sánh thông tin của học sinh với các đối tượng tương tự đã có trong dữ liệu. Điều này không chỉ giúp học sinh lựa chọn ngành nghề phù hợp mà còn giúp họ tránh sai lầm trong việc chọn ngành không đúng với khả năng và sở thích của mình.

**3. Mô hình hệ thống:**

Hệ thống khuyến nghị lựa chọn nghề nghiệp cho học sinh, sinh viên là một công cụ định hướng nghề nghiệp giúp các bạn trẻ khám phá và xác định con đường phù hợp dựa trên sở thích, kỹ năng và đặc điểm tính cách cá nhân. Hệ thống này bao gồm các thành phần chính như: giao diện người dùng, cơ sở dữ liệu lưu trữ thông tin và mô hình KNN. Thứ nhất, giao diện người dùng là nơi sinh viên có thể nhập các thông tin cần thiết để hệ thống phân tích và đưa ra gợi ý. Thứ hai là cơ sở dữ liệu lưu trữ toàn bộ thông tin về sinh viên, ngành nghề, và dữ liệu lịch sử; giúp mô hình KNN có thể đối chiếu, so sánh với những lựa chọn trước đó. Cuối cùng là mô hình KNN chính là trái tim của hệ thống, đảm nhiệm vai trò tính toán khoảng cách, tìm ra những sinh viên có đặc điểm tương đồng và đưa ra gợi ý nghề nghiệp phù hợp nhất.

**3.1 Kiến trúc hệ thống:**

Về kiến trúc hệ thống, đầu tiên phía bên người dùng nơi sinh viên có thể nhập dữ liệu và gửi yêu cầu thông qua giao diện người dùng. Tiếp theo là phía bên máy chủ chịu trách nhiệm chính trong việc lưu trữ mô hình KNN và kết nối với cơ sở dữ liệu. Máy chủ sẽ tiếp nhận thông tin, xử lý qua mô hình KNN và trả về kết quả gợi ý ngành nghề phù hợp. Một phần quan trọng của máy chủ là **Database Server**, chịu trách nhiệm lưu trữ và quản lý dữ liệu, đảm bảo dữ liệu được bảo mật và có thể truy xuất nhanh chóng khi cần.

**3.2 Quy trình hoạt động của hệ thống:**

Quy trình hoạt động của hệ thống diễn ra theo các bước khoa học và nhất quán, từ khi nhận thông tin đến khi đưa ra gợi ý nghề nghiệp cho sinh viên. Đầu tiên, sinh viên truy cập vào giao diện người dùng và nhập thông tin cá nhân như sở thích, kỹ năng, điểm số, và các đặc điểm tính cách. Sau đó, dữ liệu sẽ được chuyển đến hệ thống xử lý dữ liệu để được làm sạch, mã hóa và chuẩn hóa. Tiếp theo, mô hình KNN sẽ so sánh thông tin của sinh viên với các dữ liệu có sẵn, sử dụng khoảng cách Euclid hoặc các phương pháp tương tự để tìm ra những sinh viên có đặc điểm tương đồng nhất. Dựa trên kết quả phân tích, hệ thống sẽ đưa ra các ngành nghề được đánh giá là phù hợp nhất và hiển thị kết quả gợi ý ngay trên giao diện người dùng để sinh viên tham khảo. Nếu sinh viên chấp nhận hoặc từ chối gợi ý, hệ thống có thể lưu trữ kết quả này để tiếp tục cải thiện mô hình, từ đó nâng cao độ chính xác và hiệu quả cho những lần tư vấn tiếp theo.

Hệ thống này không chỉ đóng vai trò là công cụ hỗ trợ, mà còn giúp sinh viên hình dung và xây dựng con đường nghề nghiệp rõ ràng hơn, dựa trên cơ sở dữ liệu và khoa học, góp phần định hướng tương lai một cách vững chắc và có cơ sở.

**4. Thiết kế giao diện người dùng bằng thư viện Flask:**

Flask là một framework web Python đơn giản và nhẹ nhàng, giúp xây dựng các ứng dụng web nhanh chóng và hiệu quả. Trong báo cáo này, chúng em sẽ thiết kế giao diện người dùng cho một ứng dụng web bằng cách sử dụng Flask

**4.1 Thiết kế giao diện người dùng**

Giao diện người dùng của hệ thống được thiết kế nhằm tối ưu hóa sự đơn giản và dễ sử dụng, giúp học sinh sinh viên dễ dàng tương tác và nhận gợi ý nghề nghiệp phù hợp dựa trên tính cách và sở thích cá nhân. Giao diện tập trung vào các yếu tố thân thiện, trực quan và hướng tới người dùng, đảm bảo rằng quy trình nhập liệu và nhận kết quả diễn ra một cách thuận tiện và nhanh chóng.

Giao diện bao gồm các thành phần chính như sau. Đầu tiên là phần tiêu đề “Hệ Thống Gợi Ý Nghề Nghiệp,” giúp người dùng nhận diện rõ ràng mục đích của hệ thống ngay từ cái nhìn đầu tiên. Bên dưới tiêu đề là các trường nhập liệu cho phép người dùng chọn tính cách và sở thích của mình. Các lựa chọn này giúp cá nhân hóa hệ thống, từ đó đưa ra các gợi ý nghề nghiệp phù hợp nhất cho từng người. Để người dùng thực hiện yêu cầu, hệ thống có nút “Gửi” sau khi nhấn dữ liệu sẽ được gửi đến máy chủ để xử lý và phân tích.

Kết quả sẽ được hiển thị ngay sau khi hệ thống hoàn tất xử lý yêu cầu. Phần hiển thị kết quả được trình bày rõ ràng, gồm nhóm tính cách mà người dùng thuộc về và danh sách các nghề nghiệp đề xuất. Cách trình bày này giúp người dùng có cái nhìn toàn diện và trực quan về nhóm tính cách của mình cũng như những hướng đi nghề nghiệp tiềm năng phù hợp với đặc điểm cá nhân.

**4.2 Quy trình hoạt động của giao diện:**

Quy trình hoạt động của giao diện bao gồm ba bước chính. Đầu tiên là bước lựa chọn, người dùng lựa chọn thông tin về tính cách và sở thích vào các trường nhập liệu được cung cấp. Giao diện này được thiết kế để đảm bảo tính dễ dàng và thân thiện, giúp người dùng hoàn thành việc nhập thông tin nhanh chóng và chính xác. Sau khi nhấn nút “Gửi,” dữ liệu sẽ được chuyển đến server Flask. Tại đây, các thuật toán phân tích sẽ xử lý thông tin cá nhân và phân loại người dùng vào các nhóm tính cách nhất định, từ đó đưa ra danh sách nghề nghiệp phù hợp. Cuối cùng, hệ thống sẽ trả kết quả về cho người dùng và hiển thị trên màn hình. Người dùng sẽ thấy nhóm tính cách của mình và các nghề nghiệp đề xuất, giúp họ hiểu rõ hơn về bản thân và có thể đưa ra lựa chọn nghề nghiệp một cách chính xác.

**CHƯƠNG III: PHƯƠNG PHÁP NGHIÊN CỨU**

Trong chương này, chúng ta sẽ cùng khám phá các phương pháp được áp dụng để nghiên cứu và phát triển hệ thống khuyến nghị sinh viên lựa chọn nghề nghiệp. Các phương pháp này bao gồm quá trình thu thập dữ liệu, xử lý, trực quan hóa và xuất dữ liệu. Những bước này đóng vai trò quan trọng trong việc xây dựng hệ thống hiệu quả, giúp sinh viên nhận được những gợi ý nghề nghiệp chính xác và phù hợp với bản thân.

**1. Thu thập dữ liệu:**

Để thực hiện nghiên cứu một cách hiệu quả, phương pháp nghiên cứu được chia thành nhiều giai đoạn, bắt đầu từ việc thu thập dữ liệu. Đây là bước quan trọng đầu tiên, nơi mà thông tin sẽ được tổng hợp từ nhiều nguồn khác nhau, nhằm đảm bảo tính đa dạng và độ chính xác.

**Giai đoạn thu thập dữ liệu** tập trung vào ba khía cạnh chính. Đầu tiên là **sở thích và tính cách của học sinh**, thông qua các bài kiểm tra trắc nghiệm như MBTI, giúp xác định những đặc điểm nổi bật của mỗi em. Thứ hai là **điểm số và năng lực học tập**, với thông tin được thu thập từ hồ sơ học tập và các bài kiểm tra thành tích cá nhân, cho phép hệ thống hiểu rõ hơn về khả năng học tập của học sinh. Cuối cùng, dữ liệu về **các ngành nghề** sẽ được lấy từ các cơ sở dữ liệu, báo cáo về thị trường lao động và các khảo sát nghề nghiệp, đảm bảo rằng hệ thống có đủ thông tin cần thiết để kết nối học sinh với những ngành nghề phù hợp.



**2. Xử lý dữ liệu:**

Khi dữ liệu đã được thu thập, bước tiếp theo là **xử lý dữ liệu**. Tại giai đoạn này, chúng ta sẽ đảm bảo rằng thông tin được chuẩn bị đầy đủ và chính xác. Các kỹ thuật xử lý bao gồm việc **xử lý dữ liệu thiếu**, khi mà những thông tin còn thiếu sẽ được bổ sung hoặc loại bỏ các bản ghi không đầy đủ. **Chuẩn hóa dữ liệu** cũng là một phần quan trọng, giúp các giá trị dữ liệu được đưa về cùng một thang đo và định dạng, đảm bảo tính đồng nhất giữa các loại dữ liệu khác nhau. Đặc biệt, **loại bỏ dữ liệu nhiễu** là cần thiết để tăng cường độ chính xác của mô hình, tránh những thông tin không liên quan làm ảnh hưởng đến kết quả cuối cùng.

**3. Trực quan hóa dữ liệu:**

Sau khi xử lý, **trực quan hóa dữ liệu** sẽ diễn ra, giúp dễ dàng phân tích và hiểu rõ hơn về mối liên hệ giữa các yếu tố đầu vào và đầu ra, cụ thể là ngành nghề phù hợp. Những biểu đồ và đồ thị sinh động sẽ được sử dụng để thể hiện nhiều khía cạnh khác nhau. Chẳng hạn, ta có thể thấy được **mối liên hệ giữa sở thích và ngành nghề**; những học sinh yêu thích nghệ thuật có thể sẽ phù hợp hơn với các lĩnh vực sáng tạo như thiết kế hay marketing. Hay việc **phân bố dữ liệu điểm số** sẽ cho thấy năng lực học tập của các em so với yêu cầu đầu vào của từng ngành nghề, từ đó giúp họ nhận ra lựa chọn nào là hợp lý nhất. Hơn nữa, **tính cách và nghề nghiệp** cũng sẽ được minh họa, cho thấy sự phân bố của các tính cách và những ngành nghề tương ứng, tạo ra một cái nhìn tổng quan và dễ hiểu cho học sinh.

**4. Xuất dữ liệu:**

Cuối cùng, khi tất cả các bước đã được hoàn tất, hệ thống sẽ tiến hành **xuất dữ liệu**. Những gợi ý ngành nghề phù hợp sẽ được trình bày một cách rõ ràng và dễ tiếp cận. Có thể có **báo cáo chi tiết** gồm danh sách các ngành nghề phù hợp dựa trên sở thích, tính cách và năng lực của từng học sinh. Ngoài ra, một **bảng gợi ý ngành nghề** sẽ được tạo ra để giúp học sinh nhanh chóng nắm bắt các lựa chọn tốt nhất. Kết quả cuối cùng có thể được xuất ra dưới dạng file Excel, PDF hoặc các định dạng khác, thuận lợi cho học sinh, sinh viên cũng như phụ huynh và giáo viên có thể tham khảo và sử dụng sau này.

Thông qua những phương pháp nghiên cứu này, hệ thống không chỉ giúp học sinh tự tin hơn trong quyết định nghề nghiệp mà còn mở ra những cơ hội mới, định hướng cho các em những bước đi vững chắc trên con đường tương lai.

**CHƯƠNG IV: KẾT QUẢ**

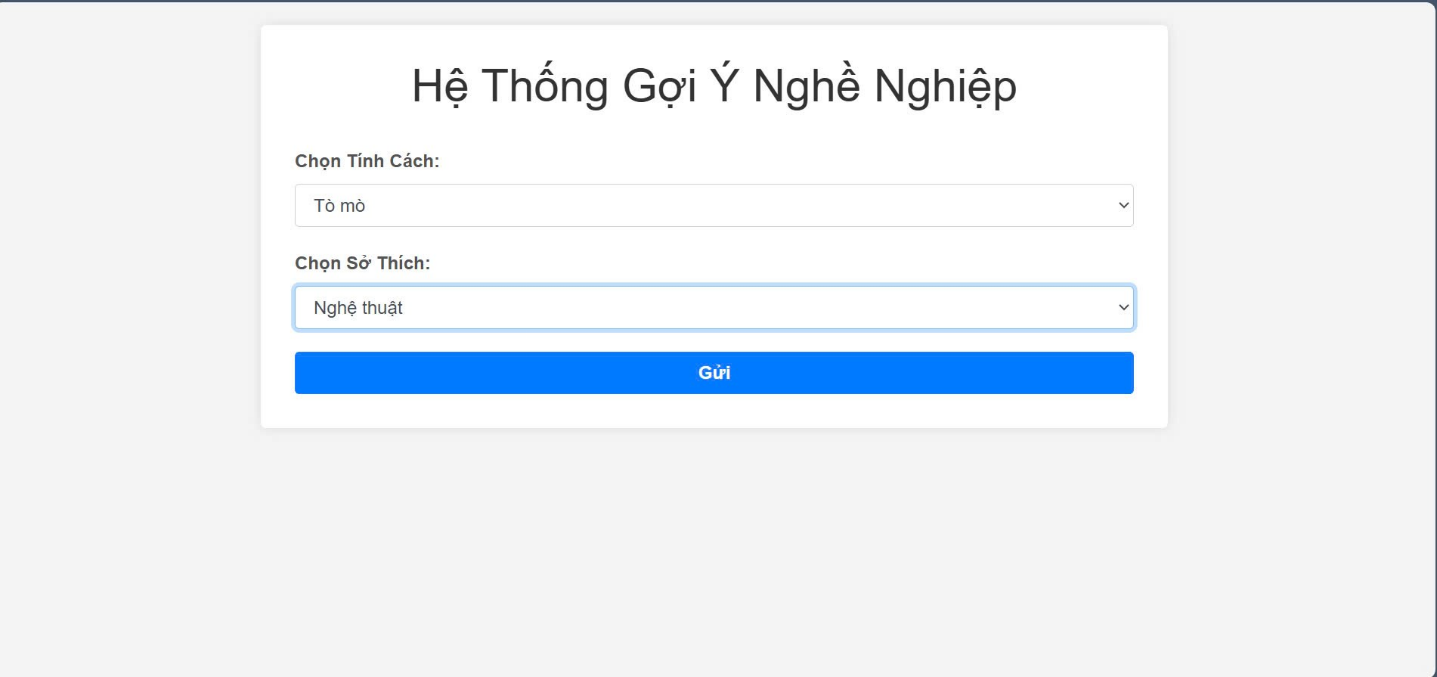
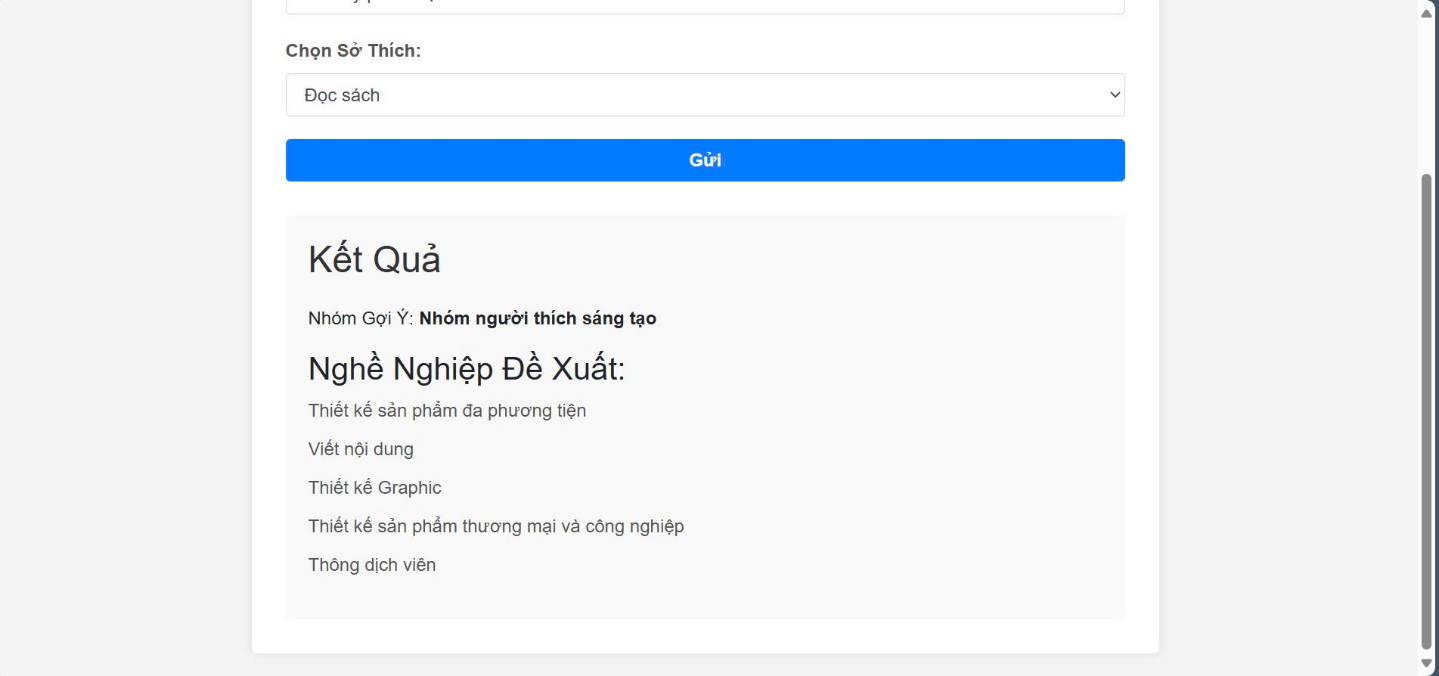
Sau một hành trình dài nghiên cứu và phát triển, hệ thống khuyến nghị lựa chọn nghề nghiệp đã mang lại những kết quả đáng tự hào. Từ những bước đầu thu thập dữ liệu, xây dựng mô hình đến khi hệ thống hoàn thiện và hoạt động ổn định, chúng em đã chứng kiến sự chuyển mình của ý tưởng thành một công cụ thực tế, hữu ích cho sinh viên trong việc định hướng tương lai nghề nghiệp.

Kết quả đầu tiên phải kể đến là khả năng gợi ý nghề nghiệp chính xác và mang tính cá nhân hóa cao. Hệ thống không chỉ đơn thuần đưa ra những ngành nghề phổ biến mà còn dựa vào sở thích, kỹ năng và tính cách của từng người dùng để đưa ra các gợi ý phù hợp nhất. Nhờ vậy, sinh viên có thể nhận được những định hướng rõ ràng, phù hợp với bản thân, từ đó tự tin hơn khi đưa ra quyết định cho con đường sự nghiệp.

Bên cạnh đó, mô hình KNN đã chứng minh được hiệu quả vượt mong đợi. Trong quá trình kiểm thử, mô hình đã hoạt động mượt mà, cho ra những kết quả chính xác, sát với kỳ vọng. Điều này khẳng định rằng lựa chọn mô hình và cách triển khai của nhóm là đúng đắn, tạo nền tảng vững chắc cho hệ thống hoạt động hiệu quả.

Mô hình KNN đã đóng vai trò cốt lõi trong việc xác định và gợi ý nghề nghiệp phù hợp cho sinh viên. Với nguyên lý dựa trên sự tương đồng giữa các cá nhân trong tập dữ liệu, mô hình đã cho thấy hiệu quả vượt trội trong việc phân loại và dự đoán. KNN tỏ ra đặc biệt hữu ích trong việc xử lý dữ liệu đa chiều, bao gồm sở thích cá nhân, kỹ năng, và các đặc điểm tính cách. Nhờ khả năng tính toán khoảng cách giữa các điểm dữ liệu một cách linh hoạt, mô hình đã có thể xác định chính xác mức độ tương đồng giữa sinh viên cần được tư vấn và các trường hợp mẫu trong tập dữ liệu. Kết quả là những gợi ý nghề nghiệp không chỉ sát thực tế mà còn mang tính cá nhân hóa cao.

Dưới đây là ảnh trang web mà chúng em đạt được:



Hình 1. Màn hình lựa chọn

Hình 2. Màn hình kết quả

**Mã nguồn:**

**app.py**

from flask import Flask, render\_template, request

import pandas as pd

from KNN import le\_traits, le\_interests, le\_groups, knn

app = Flask(\_\_name\_\_)

# Tải dữ liệu Excel để cung cấp tùy chọn cho biểu mẫu

data = pd.read\_excel('data.xlsx')

traits\_options = data['Tính cách'].unique()

interests\_options = data['Sở thích'].unique()

@app.route("/", methods=["GET", "POST"])

def index():

    result = None

    if request.method == "POST":

        # Nhận đầu vào từ người dùng

        user\_trait = request.form.get("trait")

        user\_interest = request.form.get("interest")

        # Mã hóa đầu vào

        trait\_encoded = le\_traits.transform([user\_trait])[0]

        interest\_encoded = le\_interests.transform([user\_interest])[0]

        # Dự đoán nhóm

        group\_encoded = knn.predict([[trait\_encoded, interest\_encoded]])[0]

        group = le\_groups.inverse\_transform([group\_encoded])[0]

        # Lọc các nghề nghiệp phù hợp

        careers = data[data["People"] == group]["Nghề nghiệp phù hợp"].tolist()

        result = {

            "group": group,

            "careers": careers

        }

    return render\_template("index.html", traits=traits\_options, interests=interests\_options, result=result)

if \_\_name\_\_ == "\_\_main\_\_":

    app.run(debug=True)

**Database**

****

****

****

**Information**

import pandas as pd

# Dữ liệu tổng hợp

data = {

    "Nhóm người": [

        "Nhóm người thích sáng tạo",

        "Nhóm người thích sáng tạo",

        "Nhóm người thích sáng tạo",

        "Nhóm người thích sáng tạo",

        "Nhóm người thích sáng tạo",

        "Nhóm người thích suy nghĩ",

        "Nhóm người thích suy nghĩ",

        "Nhóm người thích suy nghĩ",

        "Nhóm người thích suy nghĩ",

        "Nhóm người thích suy nghĩ",

        "Nhóm người thích tổ chức",

        "Nhóm người thích tổ chức",

        "Nhóm người thích tổ chức",

        "Nhóm người thích tổ chức",

        "Nhóm người thích tổ chức",

        "Nhóm người thích hành động",

        "Nhóm người thích hành động",

        "Nhóm người thích hành động",

        "Nhóm người thích hành động",

        "Nhóm người thích hành động",

        "Nhóm người thích đàm phán",

        "Nhóm người thích đàm phán",

        "Nhóm người thích đàm phán",

        "Nhóm người thích đàm phán",

        "Nhóm người thích đàm phán",

        "Nhóm người thích giúp đỡ",

        "Nhóm người thích giúp đỡ",

        "Nhóm người thích giúp đỡ",

        "Nhóm người thích giúp đỡ",

        "Nhóm người thích giúp đỡ",

    ],

    "Tính cách": [

        "Tò mò", "Mở lòng", "Độc lập", "Sáng tạo", "Chấp nhận rủi ro",

        "Tư duy phản biện", "Sâu sắc", "Phân tích", "Nhạy bén", "Độc lập",

        "Chi tiết", "Có tổ chức", "Quyết đoán", "Lãnh đạo", "Có kế hoạch",

        "Năng động", "Thích phiêu lưu", "Quyết đoán", "Tích cực", "Thích thử thách",

        "Khéo léo", "Thuyết phục", "Nhạy bén", "Giao tiếp tốt", "Linh hoạt",

        "Nhân ái", "Cảm thông", "Kiên nhẫn", "Hòa đồng", "Tích cực"

    ],

    "Sở thích": [

        "Nghệ thuật", "Âm nhạc", "Viết lách", "Thử nghiệm", "Du lịch",

        "Đọc sách", "Nghiên cứu", "Thảo luận", "Giải quyết vấn đề", "Tham gia hội thảo",

        "Lên kế hoạch sự kiện", "Quản lý dự án", "Tổ chức cộng đồng", "Điều phối hoạt động", "Giao tiếp hiệu quả",

        "Thể thao", "Du lịch mạo hiểm", "Tham gia hoạt động ngoài trời", "Thực hiện các dự án thực tế", "Tham gia các hoạt động nhóm",

        "Tham gia các cuộc thảo luận", "Đàm phán hợp đồng", "Tham gia vào các hoạt động thương mại", "Giải quyết xung đột", "Thuyết trình",

        "Tình nguyện", "Giúp đỡ người khác", "Tham gia các hoạt động cộng đồng", "Hỗ trợ giáo dục", "Tham gia các tổ chức từ thiện"

    ]

}

# Tạo DataFrame từ dữ liệu

df = pd.DataFrame(data)

# Xuất dữ liệu ra file Excel

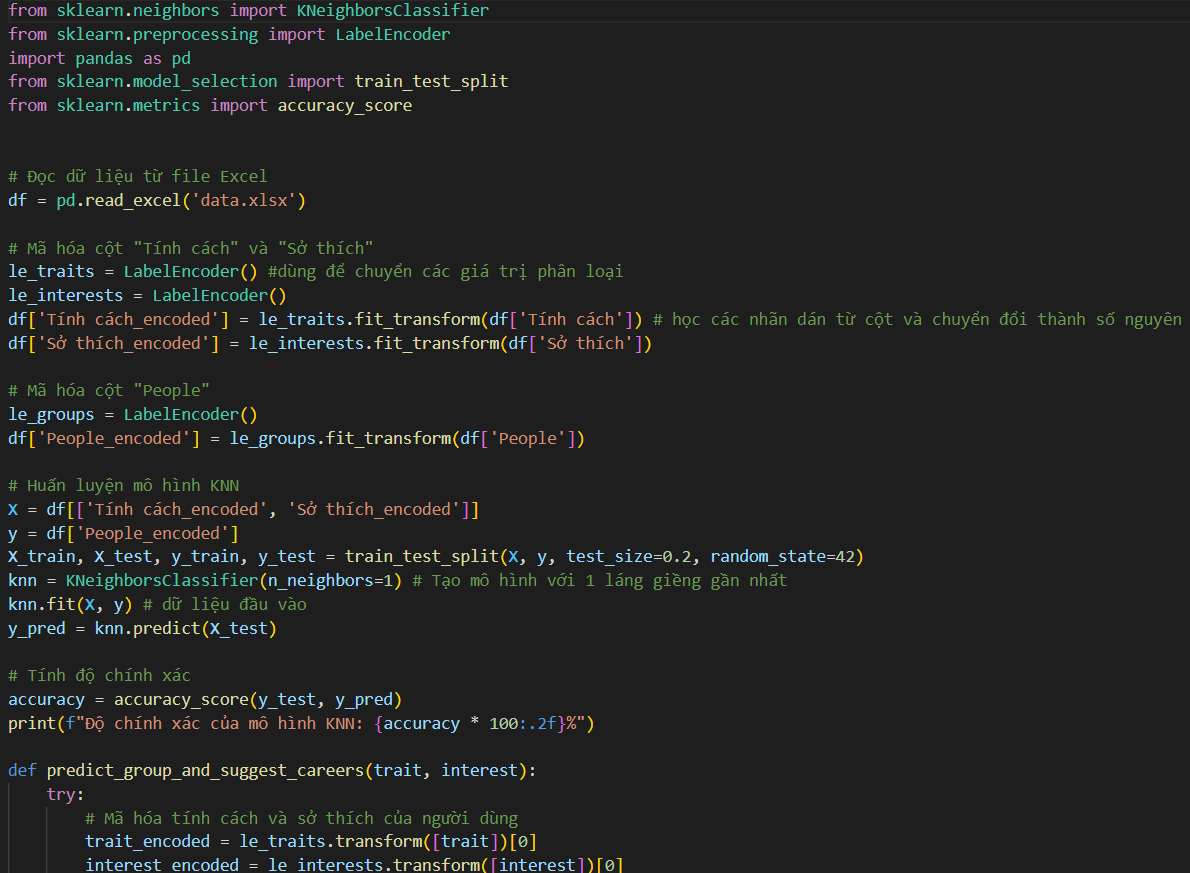
file\_name = "information.xlsx"

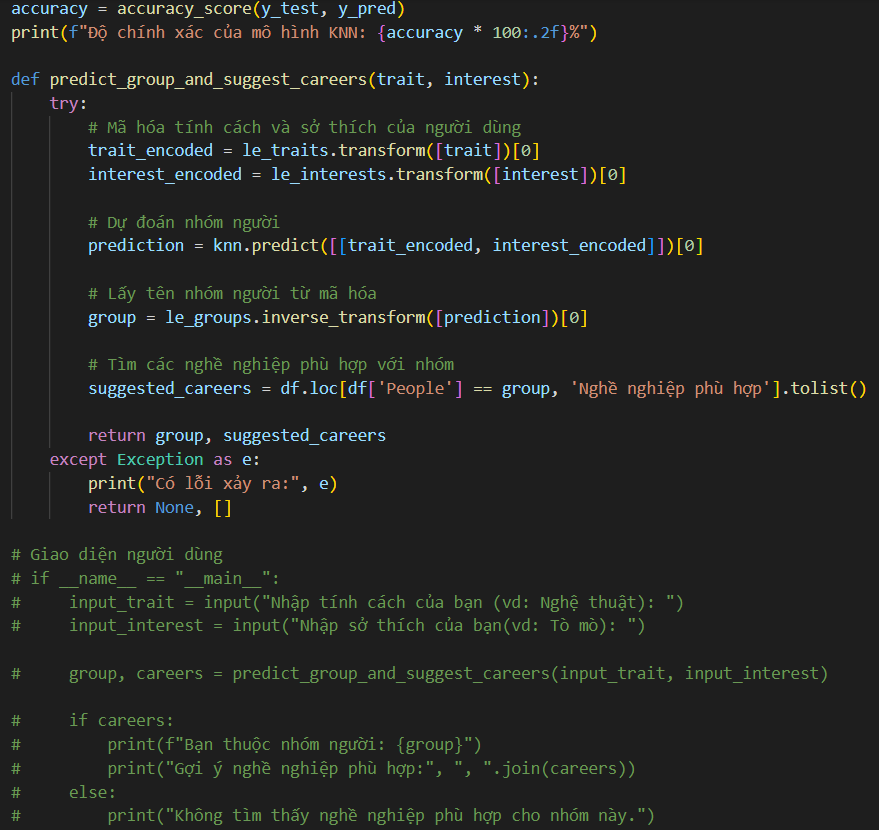
df.to\_excel(file\_name, index=False)

print(f"Dữ liệu đã được xuất vào file {file\_name}")

****

**Mô hình KNN**

****

****

Độ chính xác của mô hình KNN phụ thuộc vào số lượng láng giềng, số lượng càng ít thì độ chính xác càng cao.









# **CHƯƠNG V: KẾT LUẬN**

Sau một quá trình nghiên cứu và phát triển, hệ thống khuyến nghị lựa chọn nghề nghiệp đã hoàn thiện và mang lại nhiều kết quả tích cực. Đây không chỉ là thành quả của sự nỗ lực từ nhóm thực hiện mà còn là sự kết hợp giữa lý thuyết khoa học và ứng dụng thực tiễn trong công nghệ thông tin.

Hệ thống đã chứng minh được khả năng hỗ trợ hiệu quả trong việc định hướng nghề nghiệp cho sinh viên. Thông qua việc phân tích sở thích, kỹ năng và tính cách cá nhân, hệ thống cung cấp những gợi ý nghề nghiệp phù hợp, giúp sinh viên hiểu rõ hơn về bản thân và có cơ sở để đưa ra các quyết định sáng suốt cho tương lai. Mô hình KNN được tích hợp đã thể hiện sự chính xác và khả năng cá nhân hóa cao, phù hợp với nhu cầu đa dạng của người dùng.

Bên cạnh những thành công đạt được, dự án cũng gặp phải một số khó khăn, đặc biệt là trong việc thu thập và xử lý dữ liệu đa dạng từ nhiều nguồn khác nhau. Tuy nhiên, những hạn chế này đã được khắc phục thông qua việc áp dụng các phương pháp khoa học và kỹ thuật xử lý dữ liệu hiện đại. Nhóm nhận thấy rằng hệ thống vẫn còn tiềm năng phát triển, đặc biệt là trong việc mở rộng dữ liệu đầu vào và cải thiện mô hình phân tích để nâng cao độ chính xác và tính năng động.

Trong tương lai, hệ thống có thể được tích hợp thêm các công cụ học máy tiên tiến hơn để cải thiện khả năng dự đoán và mở rộng phạm vi ứng dụng, chẳng hạn như việc tư vấn lộ trình học tập hoặc các khóa học nâng cao kỹ năng. Nhóm cũng mong muốn hệ thống có thể phục vụ nhiều đối tượng hơn, không chỉ giới hạn ở sinh viên, mà còn mở rộng đến những người đang tìm kiếm cơ hội chuyển đổi nghề nghiệp hoặc nâng cao trình độ chuyên môn.

Tóm lại, hệ thống khuyến nghị lựa chọn nghề nghiệp đã hoàn thành sứ mệnh của mình trong việc cung cấp một công cụ hữu ích cho sinh viên và các đối tượng khác trong việc định hướng tương lai. Đây là một bước đi quan trọng, góp phần giúp người dùng tự tin hơn trên hành trình khám phá và phát triển bản thân.

# **TÀI LIỆU KHAM THẢO**

[1] [Nên chọn ngành nghề như thế nào?](file:///C:\Users\M-SI\OneDrive\Tài%20liệu\Nên%20chọn%20ngành%20nghề%20như%20thế%20nào%3f)

[2] [Cách chọn ngành nghề phù hợp với tính cách](file:///C:\Users\M-SI\OneDrive\Tài%20liệu\Cách%20chọn%20ngành%20nghề%20phù%20hợp%20với%20tính%20cách)

[3] [KNN là gì?](file:///C:\Users\M-SI\OneDrive\Tài%20liệu\KNN%20là%20gì%3f)

[4] [Đặc điểm của người sáng tạo](file:///C:\Users\M-SI\OneDrive\Tài%20liệu\Đặc%20điểm%20của%20người%20sáng%20tạo)

[5] [Nhiều thí sinh chọn sai ngành, sai trường, rất cần tư vấn tuyển sinh](file:///C:\Users\M-SI\OneDrive\Tài%20liệu\Nhiều%20thí%20sinh%20chọn%20sai%20ngành,%20sai%20trường,%20rất%20cần%20tư%20vấn%20tuyển%20sinh)