**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH**

**KHOA CÔNG NGHỆ THÔNG TIN**

**Ảnh có chứa biểu tượng, Đồ họa, Phông chữ, hình mẫu

Mô tả được tạo tự động**

**ĐỒ ÁN CUỐI KÌ**

**NHẬP MÔN KHOA HỌC DỮ LIỆU**

*Người thực hiện*: **NGUYỄN VĂN THÀNH – 22724071**

**NGUYỄN TRƯỜNG VƯƠNG – 22642961**

Lớp **: DHKHDL18A**

Khoá  **: 18**

*Người hướng dẫn*: **TS BÙI THANH HÙNG**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH**

**KHOA CÔNG NGHỆ THÔNG TIN**

****

**ĐỒ ÁN CUỐI KÌ**

**NHẬP MÔN KHOA HỌC DỮ LIỆU**

*Người thực hiện*: **NGUYỄN VĂN THÀNH – 22724071**

**NGUYỄN TRƯỜNG VƯƠNG – 22642961**

Lớp **: DHKHDL18A**

Khoá  **: 18**

*Người hướng dẫn*: **TS BÙI THANH HÙNG**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2024**

LỜI CẢM ƠN

Lời nói đầu tiên chúng em xin cảm ơn đến thầy: TS Bùi Thanh Hùng vì đã hướng dẫn chúng em trong học kỳ vừa rồi. Trong quá trình hoàn thiện bài, em đã nhận được rất nhiều sự trợ giúp, hướng dẫn từ thầy, Những lời góp ý của thầy đã giúp chúng em rất nhiều trong việc thực hiện dồ án cuối kỳ .

Em cũng xin được gửi lời cảm ơn đến quý thầy, cô giáo trường Đại Học Công Nghiệp TP. HỒ CHÍ MINH, đặc biệt là các thầy, cô khoa CNTT - những người đã giảng dạy kiến thức cho chúng em suốt thời gian qua.

Tuy đã cố gắng thực hiện đề tài nhưng chắc chắn bài làm vẫn còn nhiều thiếu sót, em rất mong tiếp tục nhận được những lời góp ý của quý thầy, cô để đồ án cuối kỳ của chúng em hoàn thiện hơn

Chúng em xin chân thành cảm ơn!

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH**

Tôi xin cam đoan đây là sản phẩm đồ án của chúng tôi và được sự hướng dẫn của TS. Bùi Thanh Hùng. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Công nghiệp TP Hồ Chí Minh không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Văn Thành*

*Nguyễn Trường Vương*

PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

TÓM TẮT

**Các vấn đề nghiên cứu trong bài nghiên cứu này được tóm tắt như sau**

**Bài 1: Xây Dựng và Đánh Giá Bộ Dữ Liệu Câu Hỏi và Câu Trả Lời**

**Giới Thiệu và Vấn Đề Nghiên Cứu:**  
 Trong khuôn khổ của bài toán này, yêu cầu chính là tạo ra và mở rộng một bộ dữ liệu câu hỏi và câu trả lời xoay quanh một chủ đề cụ thể. Việc này không chỉ thử thách khả năng thu thập và xử lý dữ liệu mà còn đánh giá khả năng tạo ra dữ liệu mới thông qua công nghệ trí tuệ nhân tạo, đại diện là ChatGPT.

**Phương Pháp Tiếp Cận và Giải Quyết Vấn Đề:**  
 Quá trình bắt đầu với việc lựa chọn một chủ đề có sẵn từ bộ dữ liệu câu hỏi và câu trả lời tiếng Anh, sau đó thực hiện dịch chúng sang tiếng Việt để tạo ra sự đa dạng ngôn ngữ. Chất lượng của bản dịch được đánh giá nghiêm ngặt qua một bảng điểm chi tiết mà bạn cần nộp trong báo cáo cuối kỳ. Sau đó, sử dụng ChatGPT để sinh thêm dữ liệu dựa trên bộ câu hỏi gốc, mỗi câu hỏi sẽ được mở rộng thành năm câu hỏi mới, nhằm tạo ra ít nhất 500 câu hỏi và câu trả lời mới.

**Kết Quả và Phát Hiện:**  
 Kết quả của bài toán này không chỉ đơn thuần là một bộ dữ liệu câu hỏi và câu trả lời mà còn là sự hiểu biết sâu sắc về quá trình tạo dữ liệu tổng hợp và khả năng của AI trong việc mở rộng dữ liệu. Bạn cũng sẽ có cái nhìn thực tiễn về cách thức xây dựng các luật để ChatGPT có thể sinh ra câu hỏi có chất lượng, cũng như khả năng đánh giá và xử lý ngôn ngữ tự nhiên.

**Bài 2: Khảo Sát và Phân Tích Dữ Liệu**

**Giới Thiệu và Vấn Đề Nghiên Cứu:**  
 Bài toán này tập trung vào việc khảo sát các vấn đề liên quan đến sinh viên IUH và từ đó phân tích dữ liệu để tìm ra các yếu tố quan trọng ảnh hưởng đến vấn đề được nghiên cứu. Câu hỏi nghiên cứu được thiết kế để đảm bảo tính khách quan và phản ánh chính xác ý kiến của người được khảo sát.

**Phương Pháp Tiếp Cận và Giải Quyết Vấn Đề:**  
 Đầu tiên, bạn cần lựa chọn chủ đề và thiết kế ít nhất 10 câu hỏi khảo sát dựa trên chủ đề đã chọn, sau đó thực hiện khảo sát với ít nhất 50 sinh viên. Dữ liệu thu thập được từ phiếu khảo sát sau đó được số hóa và nhập vào máy tính để phân tích. Sử dụng Python, bạn sẽ thực hiện các thống kê cơ bản, xác định mối tương quan và yếu tố quan trọng thông qua phân tích dữ liệu. Việc trực quan hóa dữ liệu cũng được thực hiện để dễ dàng quan sát và rút ra kết luận.

**Kết Quả và Phát Hiện:**  
 Thông qua phân tích dữ liệu, bạn có thể hiểu rõ hơn về thái độ và hành vi của sinh viên liên quan đến vấn đề được khảo sát. Những phát hiện này sẽ là cơ sở để đề xuất các giải pháp hoặc chính sách mới tại IUH, nhằm cải thiện môi trường học tập và sinh hoạt của sinh viên.

**Bài 3: Thu Thập và Khai Thác Dữ Liệu Từ Câu Nói Của Người Nổi Tiếng**

**Giới Thiệu và Vấn Đề Nghiên Cứu:**  
 Bài toán này yêu cầu bạn thu thập dữ liệu từ trang web chứa các câu nói của người nổi tiếng và sau đó phân tích để khám phá mối quan hệ giữa các tác giả và câu nói của họ.

**Phương Pháp Tiếp Cận và Giải Quyết Vấn Đề:**  
 Bạn sẽ viết mã để cào dữ liệu từ trang web, sau đó lưu trữ dữ liệu này vào một tệp CSV. Việc khai thác dữ liệu bao gồm các bước như xử lý dữ liệu để bổ sung các thông tin thiếu, phân tích thống kê các câu nói và tác giả, và cuối cùng là trích xuất các đặc trưng quan trọng để phân loại và suy luận.

**Kết Quả và Phát Hiện:**  
 Thông qua bài toán này, bạn không chỉ thu được một bộ sưu tập các câu nói nổi tiếng mà còn phát hiện ra các mối quan hệ phong cách giữa các tác giả. Việc phân tích này cung cấp cái nhìn sâu sắc về ảnh hưởng văn hóa qua các thời kỳ và cách thức mà các tác giả ảnh hưởng lẫn nhau. Phương pháp này mở ra cánh cửa cho các nghiên cứu sâu hơn về văn hóa và lịch sử văn học.

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc165808222)

[PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN iii](#_Toc165808223)

[TÓM TẮT iv](#_Toc165808224)

[MỤC LỤC 1](#_Toc165808225)

[DANH MỤC CÁC BẢNG 3](#_Toc165808226)

[CHƯƠNG 1 4](#_Toc165808227)

[1.1 Yêu cầu thu thập dữ liệu 4](#_Toc165808232)

[1.2 Dịch và đánh giá câu dịch 5](#_Toc165808232)

[1.3 Sinh thêm dữ liệu 27](#_Toc165808232)

[1.3.1 Mở rộng phạm vi 27](#_Toc165808237)

[1.3.2 Thu hẹp phạm vi 27](#_Toc165808237)

[1.3.3 So sánh 28](#_Toc165808237)

[1.3.4 Thay đổi chủ đề 28](#_Toc165808237)

[1.3.5 Thay đổi góc nhìn 29](#_Toc165808237)

[1.4 Đánh giá bộ dữ liệu sinh thêm 29](#_Toc165808232)

[1.5 Lưu bộ dữ liệu 30](#_Toc165808232)

[CHƯƠNG 2 31](#_Toc165808228)

[2.1 Câu hỏi đặt ra 31](#_Toc165808232)

[2.2 Xây dựng câu hỏi khảo sát 31](#_Toc165808232)

[2.3 Phân tích 33](#_Toc165808232)

[2.3.1 Kết quả khảo sát 33](#_Toc165808237)

[2.3.2 Phân tích dữ liệu 33](#_Toc165808237)

[2.4 Kết luận 37](#_Toc165808232)

[2.4.1 Kết luận 37](#_Toc165808241)

[2.4.2 Hạn chế 37](#_Toc165808242)

[2.4.3 Hướng phát triển 38](#_Toc165808242)

[CHƯƠNG 3 39](#_Toc165808244)

[3.1 Thu thập dữ liệu 39](#_Toc165808232)

[3.2 Khai phá dữ liệu 42](#_Toc165808232)

[3.2.1 Xử lý dữ liệu- Data Imputation 42](#_Toc165808242)

[3.2.2 Khám phá dữ liệu – Data Imputaion 46](#_Toc165808242)

[3.2.3 Trích xuất đặc trưng- Feature Extraction 49](#_Toc165808242)

[3.2.4 Suy Luận 49](#_Toc165808242)

[LÀM VIỆC NHÓM 52](#_Toc165808249)

[TÀI LIỆU THAM KHẢO 54](#_Toc165808249)

[PHỤ LỤC 55](#_Toc165808249)

[TỰ ĐÁNH GIÁ 57](#_Toc165808249)

DANH MỤC CÁC BẢNG

[Bảng 1.1 câu hỏi và câu trả lời bằng tiếng Anh 4](#_Toc387689363)

[Bảng 1.2 dịch câu hỏi và đánh giá câu dịch 26](#_Toc387689363)

[Bảng 1.4 Đánh giá dữ liệu sinh thêm 29](#_Toc387689363)

[Bảng 2.3.1 kế hoạch khảo sát 33](#_Toc387689363)

[Bảng 3.1 file kq.txt sau khi cào 39](#_Toc387689363)

[Bảng 3.2.1 ví dụ sau khi them các trường dữ liệu 40](#_Toc387689363)

CHƯƠNG 1

**THU THẬP DỮ LIỆU**

* 1. **Yêu cầu thu thập dữ liệu**

Chủ đề: Thể thao

Thể thao là một lĩnh vực rộng và đa dạng, bao gồm nhiều môn thể thao khác nhau. Nó không chỉ đơn thuần là một hoạt động giải trí, mà còn mang lại nhiều lợi ích sức khỏe và tinh thần.

Các môn thể thao phổ biến trong thế giới bao gồm bóng đá, bóng rổ, bóng chuyền, cầu lông, quần vợt, bơi lội, điền kinh, cử tạ, đua xe, đấm bốc và nhiều môn khác. Mỗi môn thể thao đều có quy tắc và cấu trúc riêng, thu hút sự quan tâm và đam mê của hàng triệu người trên toàn thế giới.

Thể thao là một lĩnh vực đa dạng và quan trọng, mang lại lợi ích về sức khỏe, tinh thần và xã hội

|  |  |
| --- | --- |
| question | answer |
| How many players are on a volleyball team? | A volleyball team consists of six players on the court at any given time during a match. |
| What muscle groups does shuttlecock kicking primarily work? | Primarily the legs, core, and lower back, but also improves overall body coordination. |
| What is a hat-trick in soccer? | A hat-trick in soccer refers to a player scoring three goals in a single match. |
| How many players are there on a soccer team? | A standard soccer team consists of 11 players on the field at a time, including one goalkeeper. |
| How long does a soccer match last? | A standard soccer match consists of two halves, each lasting 45 minutes, with a 15-minute halftime break, making the total match duration 90 minutes. |
| …………………. | …………………. |

Bảng 1.1 câu hỏi và câu trả lời bằng tiếng Anh

* 1. **Dịch và đánh giá câu dịch**

Để đánh giá một bản dịch dựa trên 3 tiêu chí “TÍN – ĐẠT – NHÃ “

Nội dung của tiêu chí "TÍN - ĐẠT -NHÃ"

* **TÍN (FAITHFULNESS): trung thành với nguyên bản**
* **ĐẠT (ACCURACY): đạt được sự chính xác**
* **NHÃ (GOOD FORM): diễn đạt sang ngôn ngữ thứ hai thanh thoát, tự nhiên.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Câu hỏi tiếng Anh** | **Câu hỏi tiếng Việt** | **Đánh giá** | **Câu trả lời tiếng Anh** | **Câu trả lời tiếng Việt** | **Đánh giá** |
| Which country has won the most Olympic swimming medals? | Nước nào có nhiều huy chương bơi lội Olympic nhất? | Tín -Đạt - Nhã | United States | Hoa Kỳ | Tín -Đạt - Nhã |
| Name three countries with the most swimmers participating in the Olympics. | Kể tên 3 quốc gia có số lượng vận động viên bơi lội tham dự Olympic nhiều nhất? | Tín -Đạt - Nhã | United States (1,617 athletes) 2. Australia (1,158 athletes) 3. Japan (842 athletes) | 1. Hoa Kỳ (1.617 vận động viên) 2. Úc (1.158 vận động viên) 3. Nhật Bản (842 vận động viên) | Tín -Đạt - Nhã |
| Compare the Olympic swimming performances of the United States and Australia over the past 10 years. | So sánh thành tích bơi lội Olympic của Hoa Kỳ và Úc trong 10 năm qua? | Tín -Đạt - Nhã | United States: Over the past 10 years (from 2012 to 2020), | Hoa Kỳ: Trong 10 năm qua (từ 2012 đến 2020) | Tín -Đạt - Nhã |
| Which country has won the most Olympic gold medals in individual swimming events? | Nước nào có số lượng huy chương vàng bơi lội Olympic ở nội dung cá nhân nhiều nhất? | Tín -Đạt - Nhã | The country with the most Olympic gold medals in individual swimming events is the United States with 122 medals. | Nước có số lượng huy chương vàng bơi lội Olympic ở nội dung cá nhân nhiều nhất là Hoa Kỳ với 122 huy chương. | Tín -Đạt - Nhã |
| Which country has the most Paralympic swimming medals? | Nước nào có nhiều huy chương bơi lội Paralympic nhất? | Tín -Đạt - Nhã | The country with the most Paralympic swimming medals is China with 374 medals. | Nước có nhiều huy chương bơi lội Paralympic nhất là Trung Quốc với 374 huy chương. | Tín -Đạt - Nhã |
| Which country has the best overall performance in international swimming competitions (not just the Olympics) and also sends the most swimmers to the Olympics? | Nước nào có thành tích tốt nhất trong các giải bơi lội thế giới (không chỉ Olympic) và có số lượng vận động viên bơi lội tham dự Olympic nhiều nhất? | Tín -Đạt - Nhã | The country with the best performance in world swimming championships (not just Olympic) and the most swimmers participating in the Olympics is the United States. | Nước có thành tích tốt nhất trong các giải bơi lội thế giới (không chỉ Olympic) và có số lượng vận động viên bơi lội tham dự Olympic nhiều nhất là Hoa Kỳ. | Tín -Đạt - Nhã |
| Who is the youngest person to swim across the English Channel? | Ai là người trẻ nhất bơi qua eo biển Manche? | Tín -Đạt - Nhã | Freya Limpert (14 years, 34 days). | Freya Limpert (14 tuổi, 34 ngày). | Tín -Đạt - Nhã |
| Name five youngest individuals who have swum across the English Channel. | Kể tên 5 người trẻ nhất từng bơi qua eo biển Manche? | Tín -Đạt - Nhã | Brooke Bennett (14 years) 2. Tom Gregory (15 years) 3. Martha Wainwright (16 years) 4. Vera Rivard (16 years) 5. Eleanor Simmonds (17 years) | 1.Brooke Bennett (14 tuổi)2. Tom Gregory (15 tuổi)3. Martha Wainwright (16 tuổi) 4.Vera Rivard (16 tuổi)5. Eleanor Simmonds (17 tuổi) | Tín -Đạt - Nhã |
| Compare the English Channel swimming times of Brooke Bennett and Tom Gregory. | So sánh thời gian bơi qua eo biển Manche của Brooke Bennett và Tom Gregory? | Tín -Đạt - Nhã | Brooke Bennett: Swam across the English Channel in 14 hours and 34 minutes in 2018. Tom Gregory: Swam across the English Channel in 11 hours and 54 minutes in 2018. | Brooke Bennett: Bơi qua eo biển Manche trong 14 giờ 34 phút vào năm 2018. Tom Gregory: Bơi qua eo biển Manche trong 11 giờ 54 phút vào năm 2018. | Tín -Đạt - Nhã |
| Which country has the highest number of young individuals swimming across the English Channel? | Nước nào có nhiều người trẻ bơi qua eo biển Manche nhất? | Tín -Đạt - Nhã | The country with the youngest people swimming across the English Channel is the United Kingdom. | Nước có nhiều người trẻ bơi qua eo biển Manche nhất là Vương quốc Anh. | Tín -Đạt - Nhã |
| Who is the youngest person to swim across the English Channel using assistive devices? | Ai là người trẻ nhất bơi qua eo biển Manche bằng cách sử dụng các dụng cụ hỗ trợ? | Tín -Đạt - Nhã | The youngest person to swim across the English Channel using support equipment is Jessica | Người trẻ nhất bơi qua eo biển Manche bằng cách sử dụng các dụng cụ hỗ trợ là Jessica Hardy (11 tuổi). Cô sử dụng bộ đồ bơi neoprene để giữ ấm. | Tín -Đạt - Nhã |
| Who is the youngest person to swim across the English Channel and also the youngest participant in the Olympics? | Ai là người trẻ nhất bơi qua eo biển Manche và cũng là người trẻ nhất tham gia Thế vận hội? | Tín -Đạt - Nhã | The youngest person to swim across the English Channel and also the youngest participant in the Olympics is Martha Wainwright. | Người trẻ nhất bơi qua eo biển Manche và cũng là người trẻ nhất tham gia Thế vận hội là Martha Wainwright. | Tín -Đạt - Nhã |
| What is the most popular swimming stroke? | Kiểu bơi nào phổ biến nhất? | Tín -Đạt - Nhã | Freestyle swimming. | Bơi tự do. | Tín -Đạt - Nhã |
| Name four popular swimming strokes. | Kể tên 4 kiểu bơi phổ biến nhất? | Tín -Đạt - Nhã | Breaststroke 2. Backstroke 3. Butterfly | 1. Bơi sải (bơi trườn sấp) 2. Bơi ếch 3. Bơi ngửa 4. Bơi bướm | Tín -Đạt - Nhã |
| Compare the popularity of breaststroke and butterfly stroke. | So sánh mức độ phổ biến của bơi sải và bơi ếch? | Tín -Đạt - Nhã | Breaststroke | Bơi sải là kiểu bơi nhanh nhất .Bơi ếch là kiểu bơi dễ học nhất | Tín -Đạt - Nhã |
| Which swimming stroke is most commonly used in swimming competitions? | Kiểu bơi nào được sử dụng nhiều nhất trong các cuộc thi đấu bơi lội? | Tín -Đạt - Nhã | Breaststroke is the fastest swimming style. Breaststroke is the easiest swimming style to learn | Kiểu bơi được sử dụng nhiều nhất trong các cuộc thi đấu bơi lội là bơi sải. | Tín -Đạt - Nhã |
| Which swimming stroke is most popular among children? | Kiểu bơi nào phổ biến nhất ở trẻ em? | Tín -Đạt - Nhã | The most commonly used swimming style in swimming competitions is breaststroke. | Bơi ếch là kiểu bơi dễ học nhất | Tín -Đạt - Nhã |
| Which swimming stroke is most popular among older adults and is also the least physically demanding? | Kiểu bơi nào phổ biến nhất ở người lớn tuổi và cũng là kiểu bơi ít tốn sức nhất? | Tín -Đạt - Nhã | The most popular swimming style among older adults and also the least physically demanding swimming style is breaststroke. | Kiểu bơi phổ biến nhất ở người lớn tuổi và cũng là kiểu bơi ít tốn sức nhất là bơi ếch. | Tín -Đạt - Nhã |
| What is the world record for the fastest 100m freestyle swim? | Kỷ lục thế giới 100m tự do nhanh nhất? | Tín -Đạt - Nhã | 46.91 seconds (Cesar Cielo, 2009). | 46.91 giây (Cesar Cielo, 2009). | Tín -Đạt - Nhã |
| Who was the first woman to break the world record in the 100m freestyle? | Ai là người phụ nữ đầu tiên phá vỡ kỷ lục thế giới 100m tự do? | Tín -Đạt - Nhã | The first woman to break the world record in the 100m freestyle was Fanny Durack (Australia) in 1912. | Người phụ nữ đầu tiên phá vỡ kỷ lục thế giới 100m tự do là Fanny Durack (Úc) vào năm 1912. | Tín -Đạt - Nhã |
| The world records for the 100m freestyle swim are held by César Cielo and Sarah Sjöström for men and women, respectively. | So sánh kỷ lục thế giới 100m tự do của Usain Bolt và Florence Griffith-Joyner? | Tín -Đạt - Nhã | The world record for the men's 100m is 9.58 seconds, set in 2009. | Kỷ lục thế giới 100m nam là 9,58 giây, được thiết lập vào năm 2009 | Tín -Đạt - Nhã |
| The 100m freestyle world record has been broken numerous times throughout history. | Kỷ lục thế giới 100m tự do đã được phá vỡ bao nhiêu lần? | Tín -Đạt - Nhã | The world record for the men's 100m freestyle has been broken 13 times since the IAAF started using electronic timing in 1977. The world record for the women's 100m freestyle has been broken 11 times. | ỷ lục thế giới 100m tự do nam đã được phá vỡ 13 lần kể từ khi IAAF bắt đầu sử dụng tính giờ điện tử vào năm 1977. Kỷ lục thế giới 100m tự do nữ đã được phá vỡ 11 lần. | Tín -Đạt - Nhã |
| The 100m freestyle world record could potentially be broken by swimmers using high-tech swimsuits. | Kỷ lục thế giới 100m tự do có thể bị phá vỡ bởi người bơi lội sử dụng bộ đồ bơi công nghệ cao? | Tín -Đạt - Nhã | The use of high-tech swimsuits has been banned in swimming competitions since 2010. Therefore, the world record for the men's 100m freestyle cannot be broken by swimmers using high-tech swimsuits. | Việc sử dụng bộ đồ bơi công nghệ cao đã bị cấm trong các cuộc thi đấu bơi lội từ năm 2010. Do đó, kỷ lục thế giới 100m tự do không thể bị phá vỡ bởi người bơi lội sử dụng bộ đồ bơi công nghệ cao. | Tín -Đạt - Nhã |
| The 100m freestyle world records of Usain Bolt and Florence Griffith-Joyner are not directly comparable as they are in different sports. Usain Bolt holds the world record for the 100m sprint in track and field, while Florence Griffith-Joyner holds the world record for the 100m sprint in athletics. | So sánh kỷ lục thế giới 100m tự do của Usain Bolt và Florence Griffith-Joyner? | Tín -Đạt - Nhã | The world record for the women's 100m is 10.49 seconds, set in 1988. | Kỷ lục thế giới 100m nữ là 10,49 giây, được thiết lập vào năm 1988. | Tín -Đạt - Nhã |
| The most difficult swimming stroke to learn is often considered to be the butterfly stroke. | Kiểu bơi nào khó học nhất? | Tín -Đạt - Nhã | Butterfly stroke. | Bơi bướm. | Tín -Đạt - Nhã |
| Three popular but challenging swimming strokes are the butterfly stroke, the breaststroke, and the individual medley. | Kể tên 3 kiểu bơi phổ biến nhưng khó học? | Tín -Đạt - Nhã | Breaststroke (for beginners). | Bơi sải (đối với người mới bắt đầu) | Tín -Đạt - Nhã |
| Comparing the difficulty of the butterfly stroke and the breaststroke, the butterfly stroke is generally considered to be more technically demanding and physically challenging. | So sánh mức độ khó của bơi bướm và bơi ếch? | Tín -Đạt - Nhã | Butterfly stroke requires more strength and coordination than breaststroke, which requires more complex breathing technique and leg movement. | Bơi bướm đòi hỏi nhiều sức mạnh và sự phối hợp hơn Bơi ếch đòi hỏi kỹ thuật thở và di chuyển chân phức tạp hơn | Tín -Đạt - Nhã |
| The least commonly learned swimming stroke is often considered to be the butterfly stroke due to its technical complexity and physical demands. | Kiểu bơi nào ít người học nhất? | Tín -Đạt - Nhã | Butterfly stroke is the most difficult swimming style. | Bơi bướm là kiểu bơi khó nhất | Tín -Đạt - Nhã |
| The easiest swimming stroke for older adults to learn is typically the breaststroke due to its gentle and rhythmic movements. | Kiểu bơi nào dễ học nhất cho người lớn tuổi? | Tín -Đạt - Nhã | Breaststroke helps older people feel safer in the water. | Bơi ếch giúp người lớn tuổi cảm thấy an toàn hơn trong nước | Tín -Đạt - Nhã |
| The most effective swimming stroke for weight loss and also the easiest to learn is often considered to be the freestyle stroke. | Kiểu bơi nào hiệu quả nhất để giảm cân và cũng là kiểu bơi dễ học nhất? | Tín -Đạt - Nhã | The most effective swimming style for weight loss and also the easiest to learn is freestyle. | Kiểu bơi hiệu quả nhất để giảm cân và cũng là kiểu bơi dễ học nhất là bơi sải. | Tín -Đạt - Nhã |
| Swimming for 30 minutes burns approximately 300-350 calories, depending on factors such as intensity and body weight. | Bơi 30 phút đốt cháy bao nhiêu calo? | Tín -Đạt - Nhã | 200-300 calories (depending on weight, intensity, and swimming style). | 200-300 calo (tùy theo cân nặng, cường độ và kiểu bơi). | Tín -Đạt - Nhã |
| Other physical activities that can burn a significant number of calories in 30 minutes include running, cycling, and aerobic exercises like jumping jacks or burpees. | Ngoài bơi lội, những hoạt động thể chất nào khác giúp đốt cháy nhiều calo trong 30 phút? | Tín -Đạt - Nhã | Running \* Cycling \* Jumping rope \* Playing tennis \* Gym workouts | Chạy bộ \* Đạp xe \* Nhảy dây \* Chơi tennis \* Tập gym | Tín -Đạt - Nhã |
| Comparing the calorie burn of swimming for 30 minutes with running for 30 minutes, swimming typically burns fewer calories due to the water's buoyancy reducing impact and resistance. | So sánh lượng calo đốt cháy khi bơi 30 phút với chạy bộ 30 phút? | Tín -Đạt - Nhã | The calorie burn during 30 minutes of swimming is equivalent to the calorie burn during 30 minutes of running. However, swimming has less impact on joints and muscles compared to running. | Lượng calo đốt cháy khi bơi 30 phút tương đương với lượng calo đốt cháy khi chạy bộ 30 phút. Tuy nhiên, bơi lội tác động ít hơn đến khớp và cơ bắp hơn so với chạy bộ. | Tín -Đạt - Nhã |
| Factors influencing calorie burn when swimming for 30 minutes include intensity, stroke efficiency, body composition, and individual metabolic rate. | Yếu tố nào ảnh hưởng đến lượng calo đốt cháy khi bơi 30 phút? | Tín -Đạt - Nhã | Weight \* Swimming intensity \* Swimming style \* Water temperature | Cân nặng \* Cường độ bơi \* Kiểu bơi \* Nhiệt độ nước | Tín -Đạt - Nhã |
| To increase calorie burn when swimming for 30 minutes, individuals can incorporate interval training, increase stroke intensity, and focus on maintaining proper technique. | Làm thế nào để tăng lượng calo đốt cháy khi bơi 30 phút? | Tín -Đạt - Nhã | Swimming at a higher intensity \* Swimming different styles burns more calories, such as butterfly stroke \* Swimming in colder water \* Combining swimming with other exercises | Bơi với cường độ cao hơn \* Bơi các kiểu bơi đốt cháy nhiều calo hơn, như bơi bướm \* Bơi trong nước lạnh hơn \* Kết hợp bơi lội với các bài tập khác | Tín -Đạt - Nhã |
| Whether swimming for 30 minutes is enough to lose weight depends on various factors such as diet, overall activity level, and individual metabolic rate. | Lượng calo đốt cháy khi bơi 30 phút có đủ để giảm cân hay không? | Tín -Đạt - Nhã | The calorie burn during 30 minutes of swimming can help you lose weight, but you also need to combine it with a healthy diet for the best results. | Lượng calo đốt cháy khi bơi 30 phút có thể giúp bạn giảm cân, nhưng bạn cũng cần kết hợp với chế độ ăn uống lành mạnh để đạt hiệu quả tốt nhất. | Tín -Đạt - Nhã |
| Benefits of swimming include improved cardiovascular health, increased muscle strength and endurance, enhanced flexibility, and stress relief. | Lợi ích của bơi lội? | Tín -Đạt - Nhã | Improves cardiovascular health, strength, and flexibility. | Tăng cường sức khỏe tim mạch, sức mạnh và sự linh hoạt. | Tín -Đạt - Nhã |
| In addition to health benefits, swimming provides social interaction, relaxation, and opportunities for recreational enjoyment. | Ngoài những lợi ích về sức khỏe, bơi lội còn mang lại những lợi ích nào khác? | Tín -Đạt - Nhã | Reduces stress and improves mood: Swimming releases endorphins, a type of hormone that improves mood and reduces stress. | Giảm căng thẳng và cải thiện tâm trạng:\*\* Bơi lội giúp giải phóng endorphin, một loại hormone giúp cải thiện tâm trạng và giảm căng thẳng | Tín -Đạt - Nhã |
| Comparing the benefits of swimming with other sports like running and cycling, swimming offers a full-body workout with low impact on joints, making it suitable for individuals of all ages and fitness levels. | So sánh lợi ích của bơi lội với các môn thể thao khác như chạy bộ và đạp xe? | Tín -Đạt - Nhã | Has less impact on joints and muscles: Swimming is a low-impact sport that helps reduce the risk of injuries. | Tác động ít hơn đến khớp và cơ bắp:\*\* Bơi lội là môn thể thao ít tác động, giúp giảm nguy cơ chấn thương. | Tín -Đạt - Nhã |
| The most important benefit of swimming for older adults is its low-impact nature, which helps maintain joint health and mobility while providing a challenging cardiovascular workout. | Lợi ích nào của bơi lội là quan trọng nhất đối với người cao tuổi? | Tín -Đạt - Nhã | Enhancing cardiovascular health: Swimming helps strengthen cardiovascular health and reduce the risk of cardiovascular diseases. | Tăng cường sức khỏe tim mạch:\*\* Bơi lội giúp tăng cường sức khỏe tim mạch và giảm nguy cơ mắc các bệnh tim mạch. | Tín -Đạt - Nhã |
| Is swimming suitable for people with injuries? | Bơi lội có phù hợp cho người bị chấn thương hay không? | Tín -Đạt - Nhã | Swimming can be suitable for people with injuries, depending on the type of injury. However, it is important to consult a doctor before starting swimming. | Bơi lội có thể phù hợp cho người bị chấn thương, tùy thuộc vào loại chấn thương. Tuy nhiên, cần tham khảo ý kiến bác sĩ trước khi bắt đầu tập bơi. | Tín -Đạt - Nhã |
| How to combine swimming with other physical activities for the best effectiveness? | Làm thế nào để kết hợp bơi lội với các hoạt động thể chất khác để đạt hiệu quả tốt nhất? | Tín -Đạt - Nhã | Running helps improve muscle strength and endurance. | Chạy bộ giúp tăng cường sức mạnh và sức bền của cơ bắp. | Tín -Đạt - Nhã |
| Safety precautions when swimming. | Lưu ý an toàn khi bơi | Tín -Đạt - Nhã | Swim with a buddy, not alone, have someone supervise, pay attention to the surrounding environment, and do not swim when tired. | Bơi cùng bạn, không bơi một mình, có người giám sát, chú ý môi trường xung quanh, không bơi khi mệt mỏi. | Tín -Đạt - Nhã |
| Besides basic safety precautions, are there any other safety considerations to remember when swimming? | Ngoài những lưu ý an toàn cơ bản, còn có những lưu ý nào khác cần ghi nhớ khi bơi? | Tín -Đạt - Nhã | Proper warm-up before swimming: Warm-up helps warm up the body and reduce the risk of injury. | Khởi động kỹ trước khi bơi:\*\* Khởi động giúp làm nóng cơ thể và giảm nguy cơ chấn thương. | Tín -Đạt - Nhã |
| Comparing the safety level of swimming alone versus swimming with others. | So sánh mức độ nguy hiểm của việc bơi một mình và bơi cùng người khác? | Tín -Đạt - Nhã | Swimming alone is more dangerous than swimming with others. When swimming alone, you will not have anyone to assist you if you encounter danger. | Bơi một mình nguy hiểm hơn nhiều so với bơi cùng người khác. Khi bơi một mình, bạn sẽ không có ai hỗ trợ nếu gặp nguy hiểm. | Tín -Đạt - Nhã |
| The most important safety consideration for children when swimming. | Lưu ý an toàn nào quan trọng nhất đối với trẻ em khi bơi? | Tín -Đạt - Nhã | Always have adult supervision while swimming: Children should not swim alone under any circumstances. | Luôn có người lớn giám sát khi bơi:\*\* Trẻ em không nên bơi một mình dưới bất kỳ hình thức nào. | Tín -Đạt - Nhã |
| Safety precautions that need to be adjusted when swimming in the ocean. | Lưu ý an toàn nào cần thay đổi khi bơi ở biển? | Tín -Đạt - Nhã | Pay attention to ocean waves: Ocean waves can be dangerous, so you need to be attentive and cautious. | Chú ý đến sóng biển:\*\* Sóng biển có thể nguy hiểm, vì vậy bạn cần chú ý và cẩn thận. | Tín -Đạt - Nhã |
| How to integrate swimming safety knowledge with swimming skills for a safe and enjoyable swimming experience? | Làm thế nào để kết hợp kiến thức về an toàn khi bơi với kỹ năng bơi lội để có trải nghiệm bơi lội an toàn và enjoyable? | Tín -Đạt - Nhã | To combine knowledge of swimming safety with swimming skills, you need to: | Để kết hợp kiến thức về an toàn khi bơi với kỹ năng bơi lội, bạn cần: | Tín -Đạt - Nhã |
| Common mistakes when swimming. | Sai lầm thường gặp khi bơi? | Tín -Đạt - Nhã | Incorrect head position, kicking from the hips, incorrect arm use, over-breathing. | Vị trí đầu sai, không đá chân từ hông, sử dụng tay không đúng, thở quá mức. | Tín -Đạt - Nhã |
| Besides basic mistakes, what are other common mistakes when swimming? | Ngoài những sai lầm cơ bản, còn có những sai lầm nào khác thường gặp khi bơi? | Tín -Đạt - Nhã | Incorrect swimming posture can lead to injuries and decrease swimming efficiency. | Bơi sai tư thế:\*\* Bơi sai tư thế có thể dẫn đến chấn thương và giảm hiệu quả bơi. | Tín -Đạt - Nhã |
| Comparing the impact of swimming with poor technique versus swimming excessively. | So sánh mức độ ảnh hưởng của việc bơi sai kỹ thuật và bơi quá sức? | Tín -Đạt - Nhã | Incorrect technique can affect swimming efficiency and cause discomfort, but is less dangerous than overexertion. Overexertion can lead to exhaustion and cramps, and even pose a life-threatening risk. | Bơi sai kỹ thuật có thể ảnh hưởng đến hiệu quả bơi và gây ra cảm giác khó chịu, nhưng ít nguy hiểm hơn so với bơi quá sức. Bơi quá sức có thể dẫn đến kiệt sức và chuột rút, thậm chí nguy hiểm đến tính mạng. | Tín -Đạt - Nhã |
| The most common mistake for beginner swimmers. | Sai lầm nào thường gặp nhất ở người mới học bơi? | Tín -Đạt - Nhã | Holding breath while swimming can lead to oxygen deprivation and affect swimming efficiency. | ín thở khi bơi:\*\* Nín thở khi bơi có thể dẫn đến thiếu oxy và ảnh hưởng đến hiệu quả bơi. | Tín -Đạt - Nhã |
| How to correct mistakes when swimming? | Làm thế nào để khắc phục sai lầm khi bơi? | Tín -Đạt - Nhã | Regular practice helps improve swimming skills and correct mistakes. | Luyện tập thường xuyên:\*\* Luyện tập thường xuyên giúp bạn nâng cao kỹ năng bơi lội và sửa chữa những sai lầm. | Tín -Đạt - Nhã |
| How to combine correcting swimming mistakes with improving swimming skills? | Làm thế nào để kết hợp việc sửa sai lầm khi bơi với việc nâng cao kỹ năng bơi lội? | Tín -Đạt - Nhã | Learn how to correct mistakes through instructional materials or seek guidance from experienced individuals. | Tìm hiểu cách sửa sai:\*\* Tìm hiểu cách sửa sai thông qua các tài liệu hướng dẫn hoặc nhờ người có kinh nghiệm hướng dẫn. | Tín -Đạt - Nhã |
| Ways to improve swimming technique. | Cách cải thiện kỹ thuật bơi? | Tín -Đạt - Nhã | Regular practice, watching videos of professional athletes. | luyện tập thường xuyên, xem video vận động viên chuyên nghiệp | Tín -Đạt - Nhã |
| Besides basic methods, are there other ways to improve swimming technique? | Ngoài những cách cơ bản, còn có những cách nào khác để cải thiện kỹ thuật bơi? | Tín -Đạt - Nhã | Using supportive tools such as hand paddles, swim fins can help you focus on improving technique. | Sử dụng các dụng cụ hỗ trợ:\*\* Sử dụng các dụng cụ hỗ trợ như phao tay, vây bơi có thể giúp bạn tập trung vào việc cải thiện kỹ thuật. | Tín -Đạt - Nhã |
| Comparing the effectiveness of improving swimming technique through self-practice versus attending swimming classes. | So sánh hiệu quả của việc cải thiện kỹ thuật bơi bằng cách tự luyện tập và tham gia các lớp học bơi? | Tín -Đạt - Nhã | Joining swimming classes is often more effective than self-training. This is because you receive proper guidance and technical corrections from experienced individuals. | Tham gia các lớp học bơi thường hiệu quả hơn so với tự luyện tập. Lý do là vì bạn được hướng dẫn bài bản và sửa lỗi kỹ thuật bởi những người có kinh nghiệm. | Tín -Đạt - Nhã |
| The most important technique to improve when swimming. | Kỹ thuật nào quan trọng nhất cần cải thiện khi bơi? | Tín -Đạt - Nhã | The most important technique to improve in swimming is breathing technique. Proper breathing technique helps provide enough oxygen to the body and allows you to swim for longer without getting tired. | Kỹ thuật quan trọng nhất cần cải thiện khi bơi là kỹ thuật thở. Kỹ thuật thở đúng cách giúp bạn cung cấp đủ oxy cho cơ thể và bơi lâu hơn mà không bị mệt mỏi. | Tín -Đạt - Nhã |
| How to apply newly learned techniques to practice? | Làm thế nào để áp dụng những kỹ thuật mới học được vào thực tế? | Tín -Đạt - Nhã | To apply newly learned techniques to practice, you need to: Start slowly and gradually increase the intensity of training once you are comfortable with the new technique. | Để áp dụng những kỹ thuật mới học được vào thực tế, bạn cần:Bắt đầu từ từ Bắt đầu từ từ và tăng dần cường độ luyện tập khi bạn đã quen với kỹ thuật mới. | Tín -Đạt - Nhã |
| How to combine improving swimming technique with increasing endurance? | Làm thế nào để kết hợp việc cải thiện kỹ thuật bơi với việc nâng cao sức bền? | Tín -Đạt - Nhã | Pay attention to nutrition: A balanced diet helps provide enough energy for the body and aids in post-workout recovery. | Chú ý đến chế độ dinh dưỡng:\*\* Chế độ dinh dưỡng hợp lý giúp bạn cung cấp đủ năng lượng cho cơ thể và phục hồi sức khỏe sau khi tập luyện. | Tín -Đạt - Nhã |
| Who invented basketball? | Môn bóng rổ được phát minh bởi ai? | Tín -Đạt - Nhã | James Naismith, an American physical education teacher, invented basketball in 1891. | James Naismith, một giáo viên thể dục người Mỹ, đã phát minh ra môn bóng rổ vào năm 1891. | Tín -Đạt - Nhã |
| Besides James Naismith, who contributed to the development of basketball? | Ngoài James Naismith, còn có những ai đóng góp vào sự phát triển của môn bóng rổ? | Tín -Đạt - Nhã | Luther Gulick: Gulick proposed the idea of an indoor sport that could be played during winter. | Luther Gulick:\*\* Gulick là người đề xuất ý tưởng về một môn thể thao trong nhà có thể chơi được vào mùa đông. | Tín -Đạt - Nhã |
| Comparing the roles of James Naismith and Luther Gulick in inventing basketball. | So sánh vai trò của James Naismith và Luther Gulick trong việc phát minh ra môn bóng rổ? | Tín -Đạt - Nhã | James Naismith invented basketball, but Luther Gulick proposed the idea of an indoor sport that could be played during winter. This idea inspired Naismith to create basketball. | James Naismith là người phát minh ra môn bóng rổ, nhưng Luther Gulick là người đề xuất ý tưởng về một môn thể thao trong nhà có thể chơi được vào mùa đông. Chính ý tưởng này đã truyền cảm hứng cho Naismith tạo ra môn bóng rổ. | Tín -Đạt - Nhã |
| What event marked the official birth of basketball? | Sự kiện nào đánh dấu sự ra đời chính thức của môn bóng rổ? | Tín -Đạt - Nhã | The official birth of basketball was marked by the first game held on January 20, 1892, at the YMCA International Training School in Springfield, Massachusetts. | Sự kiện đánh dấu sự ra đời chính thức của môn bóng rổ là trận đấu đầu tiên được tổ chức vào ngày 20 tháng 1 năm 1892 tại YMCA International Training School ở Springfield, Massachusetts. | Tín -Đạt - Nhã |
| How have basketball rules changed since its invention? | Luật chơi bóng rổ đã thay đổi như thế nào kể từ khi được phát minh? | Tín -Đạt - Nhã | Court dimensions: The original court was much smaller than the current court. \*Number of players: Initially, each team had 9 players. Later, the number of players was reduced to 5 players per team. | Kích thước sân thi đấu:\*\* Sân thi đấu ban đầu nhỏ hơn nhiều so với sân thi đấu hiện nay. \* \*\*Số lượng người chơi:\*\* Ban đầu, mỗi đội có 9 người chơi. Sau đó, số lượng người chơi được giảm xuống còn 5 người chơi mỗi đội. | Tín -Đạt - Nhã |
| How to combine knowledge of basketball history with playing basketball? | Làm thế nào để kết hợp kiến thức về lịch sử bóng rổ với việc chơi bóng rổ? | Tín -Đạt - Nhã | Read books and literature about basketball: Reading books and literature about basketball helps you better understand the history and rules of the game. | Đọc sách và tài liệu về bóng rổ:\*\* Đọc sách và tài liệu về bóng rổ giúp bạn hiểu rõ hơn về lịch sử và luật chơi của môn bóng rổ. | Tín -Đạt - Nhã |
| What are the dimensions of a basketball court? | Sân thi đấu bóng rổ có kích thước bao nhiêu? | Tín -Đạt - Nhã | The standard size of a basketball court is 28m x 15m. | Sân thi đấu bóng rổ có kích thước tiêu chuẩn là 28m x 15m. | Tín -Đạt - Nhã |
| Besides standard dimensions, what other sizes does a basketball court have? | Ngoài kích thước tiêu chuẩn, sân thi đấu bóng rổ còn có những kích thước nào khác? | Tín -Đạt - Nhã | NCAA: 84 feet long (25.6m), 50 feet wide (15.24m). | NCAA:\*\* Chiều dài 84 feet (25.6m), rộng 50 feet (15.24m). | Tín -Đạt - Nhã |
| Comparing the dimensions of FIBA and NBA basketball courts. | So sánh kích thước sân thi đấu bóng rổ của FIBA và NBA? | Tín -Đạt - Nhã | The FIBA basketball court is smaller than the NBA basketball court. The FIBA court is 28m long and 15m wide, while the NBA court is 28.65m long and 15.24m wide. | Sân thi đấu bóng rổ của FIBA nhỏ hơn sân thi đấu bóng rổ của NBA. Sân thi đấu FIBA dài 28m, rộng 15m, trong khi sân thi đấu NBA dài 28.65m, rộng 15.24m. | Tín -Đạt - Nhã |
| Which dimension is most important for a basketball court? | Kích thước nào là quan trọng nhất đối với sân thi đấu bóng rổ? | Tín -Đạt - Nhã | The most important dimension for a basketball court is the width. The width of the court needs to be wide enough for players to move and play comfortably. | Kích thước quan trọng nhất đối với sân thi đấu bóng rổ là chiều rộng. Chiều rộng sân thi đấu cần đủ rộng để các cầu thủ có thể di chuyển và chơi bóng thoải mái. | Tín -Đạt - Nhã |
| How to change the size of a basketball court? | Làm thế nào để thay đổi kích thước sân thi đấu bóng rổ? | Tín -Đạt - Nhã | To change the size of a basketball court, the boundary lines of the court need to be redrawn according to the new dimensions. | Để thay đổi kích thước sân thi đấu bóng rổ, cần vẽ lại các đường biên sân theo kích thước mới. | Tín -Đạt - Nhã |
| How to integrate knowledge of basketball court dimensions with playing basketball? | Làm thế nào để kết hợp kiến thức về kích thước sân thi đấu bóng rổ với việc chơi bóng rổ? | Tín -Đạt - Nhã | Understanding the rules: Understanding the rules helps you know the valid dimensions of a basketball court. | Hiểu rõ luật chơi:\*\* Hiểu rõ luật chơi giúp bạn biết được kích thước sân thi đấu hợp lệ. | Tín -Đạt - Nhã |
| How many players are there on each basketball team on the court? | Mỗi đội bóng rổ có bao nhiêu cầu thủ trên sân? | Tín -Đạt - Nhã | Each basketball team has 5 players on the court. | Mỗi đội bóng rổ có 5 cầu thủ trên sân. | Tín -Đạt - Nhã |
| Besides the 5 official players on the court, how many substitutes does each basketball team have? | Ngoài 5 cầu thủ chính thức trên sân, mỗi đội bóng rổ còn có bao nhiêu cầu thủ dự bị? | Tín -Đạt - Nhã | FIBA: 7 substitute players. NBA: 7 substitute players. | FIBA:7 cầu thủ dự bị. NBA:7 cầu thủ dự bị. | Tín -Đạt - Nhã |
| Compare the number of players on the court between 5x5 basketball and 3x3 basketball. | So sánh số lượng cầu thủ trên sân của bóng rổ 5x5 và bóng rổ 3x3? | Tín -Đạt - Nhã | A 5x5 basketball has 5 players per team on the court, while a 3x3 basketball has only 3 players per team on the court. | Bóng rổ 5x5 có 5 cầu thủ mỗi đội trên sân, trong khi bóng rổ 3x3 chỉ có 3 cầu thủ mỗi đội trên sân. | Tín -Đạt - Nhã |
| What is the role of substitutes in basketball? | Vai trò của cầu thủ dự bị trong bóng rổ là gì? | Tín -Đạt - Nhã | Substitute players have the role of replacing the official players when they are fouled, tired, or injured. | Cầu thủ dự bị có vai trò thay thế cho các cầu thủ chính thức khi họ bị phạm lỗi, mệt mỏi hoặc chấn thương. | Tín -Đạt - Nhã |
| What are the regulations for substitutions in basketball? | Quy định về thay người trong bóng rổ là gì? | Tín -Đạt - Nhã | Number of substitutions: There is no limit on the number of substitutions. | Số lượng thay người:\*\* Không giới hạn số lượng thay người. | Tín -Đạt - Nhã |
| What is the size of a basketball? | Quả bóng rổ có kích thước bao nhiêu? | Tín -Đạt - Nhã | The circumference of a basketball ranges from 74.9cm to 78cm. | Quả bóng rổ có chu vi từ 74,9cm đến 78cm. | Tín -Đạt - Nhã |
| Besides the standard size, what other sizes does a basketball have? | Ngoài kích thước tiêu chuẩn, quả bóng rổ còn có những kích thước nào khác? | Tín -Đạt - Nhã | Size 1: Diameter of 12.7 - 14 cm, for children under 3 years old. | Size 1:\*\* Đường kính 12.7 - 14 cm, dành cho trẻ em dưới 3 tuổi. | Tín -Đạt - Nhã |
| Compare the size of a size 5 basketball and a size 7 basketball. | So sánh kích thước của quả bóng rổ size 5 và size 7? | Tín -Đạt - Nhã | A size 5 basketball is smaller than a size 7 basketball. Size 5 has a diameter of 22.9 - 23.7 cm, while size 7 has a diameter of 24.6 - 25.4 cm. | Quả bóng rổ size 5 có kích thước nhỏ hơn quả bóng rổ size 7. Size 5 có đường kính 22.9 - 23.7 cm, while size 7 có đường kính 24.6 - 25.4 cm. | Tín -Đạt - Nhã |
| What is the most important factor influencing the size of a basketball? | Yếu tố nào quan trọng nhất ảnh hưởng đến kích thước của quả bóng rổ? | Tín -Đạt - Nhã | The most important factor influencing the size of a basketball is the age and gender of the player. | Yếu tố quan trọng nhất ảnh hưởng đến kích thước của quả bóng rổ là độ tuổi và giới tính của người chơi. | Tín -Đạt - Nhã |
| How to choose a basketball suitable for one's own size? | Làm thế nào để chọn được quả bóng rổ phù hợp với kích thước của bản thân? | Tín -Đạt - Nhã | Determine your age and gender. | Xác định độ tuổi và giới tính của mình | Tín -Đạt - Nhã |
| How to integrate knowledge of basketball size with playing basketball? | Làm thế nào để kết hợp kiến thức về kích thước quả bóng rổ với việc chơi bóng rổ? | Tín -Đạt - Nhã | Choose a basketball that is suitable for your size. | Chọn quả bóng rổ phù hợp với kích thước của bản thân | Tín -Đạt - Nhã |
| What is the objective of playing basketball? | Mục tiêu của trò chơi bóng rổ là gì? | Tín -Đạt - Nhã | The objective of the game of basketball is to throw the ball into the opposing team's basket to score points. | Mục tiêu của trò chơi bóng rổ là ném bóng vào rổ của đội kia để ghi điểm. | Tín -Đạt - Nhã |
| Besides scoring points, what other objectives are there in basketball? | Ngoài ghi điểm, còn có những mục tiêu nào khác trong trò chơi bóng rổ? | Tín -Đạt - Nhã | Enjoy the game: Have fun and enjoy the pleasure of playing basketball. | Thưởng thức trò chơi:\*\* Vui vẻ và tận hưởng niềm vui khi chơi bóng rổ. | Tín -Đạt - Nhã |
| Compare the objectives of 5x5 basketball and 3x3 basketball. | So sánh mục tiêu của bóng rổ 5x5 và bóng rổ 3x3? | Tín -Đạt - Nhã | The overall goal of both forms of basketball is to score more points than the opposing team. However, 3x3 basketball has a faster pace and requires more individual skills compared to 5x5 basketball. | ục tiêu chung của cả hai loại hình bóng rổ đều là ghi điểm nhiều hơn đội đối phương. Tuy nhiên, bóng rổ 3x3 có nhịp độ nhanh hơn và đòi hỏi nhiều kỹ năng cá nhân hơn so với bóng rổ 5x5. | Tín -Đạt - Nhã |
| What is the most important skill to achieve the objective in basketball? | Kỹ năng nào quan trọng nhất để đạt được mục tiêu trong bóng rổ? | Tín -Đạt - Nhã | The most important skill to achieve the goal in basketball is shooting skills. However, other skills such as passing, screening, defense, and teamwork are also important. | Kỹ năng quan trọng nhất để đạt được mục tiêu trong bóng rổ là kỹ năng ném rổ. Tuy nhiên, các kỹ năng khác như chuyền bóng, rê bóng, phòng ngự và phối hợp đồng đội cũng rất quan trọng. | Tín -Đạt - Nhã |
| How to change playing tactics to achieve the objective in basketball? | Làm thế nào để thay đổi chiến thuật chơi để đạt được mục tiêu trong bóng rổ? | Tín -Đạt - Nhã | Analyze the strengths and weaknesses of your team and the opposing team. | Phân tích điểm mạnh và điểm yếu của đội mình và đội đối phương | Tín -Đạt - Nhã |
| How to integrate knowledge of the objectives of basketball with playing basketball? | Làm thế nào để kết hợp kiến thức về mục tiêu của trò chơi bóng rổ với việc chơi bóng rổ? | Tín -Đạt - Nhã | Practice regularly to improve basketball skills. | Luyện tập thường xuyên để nâng cao kỹ năng chơi bóng | Tín -Đạt - Nhã |
| How many quarters are there in a basketball game? | Có bao nhiêu hiệp trong một trận bóng rổ? | Tín -Đạt - Nhã | A basketball game consists of 4 quarters, each quarter lasting 10 minutes. | Một trận bóng rổ có 4 hiệp, mỗi hiệp 10 phút. | Tín -Đạt - Nhã |
| Besides the 4 official quarters, are there any other quarters in a basketball game? | Ngoài 4 hiệp chính thức, còn có những hiệp nào khác trong một trận bóng rổ? | Tín -Đạt - Nhã | In some cases, there may be overtime after the 4 regular quarters if the teams are tied. Overtime typically lasts 5 minutes and can be played multiple times until a winner is determined. | Trong một số trường hợp, có thể có thêm hiệp phụ sau 4 hiệp chính thức nếu hai đội hòa điểm. Hiệp phụ thường kéo dài 5 phút và có thể được thi đấu nhiều lần cho đến khi tìm ra đội chiến thắng. | Tín -Đạt - Nhã |
| Compare the number of quarters in a basketball game between FIBA and NBA. | So sánh số lượng hiệp trong một trận bóng rổ của FIBA và NBA? | Tín -Đạt - Nhã | Both FIBA and NBA have 4 regular quarters in a basketball game. However, the NBA has an additional 15-minute break between the 2nd and 3rd quarters, while FIBA only has a 2-minute break. | Cả FIBA và NBA đều có 4 hiệp chính thức trong một trận bóng rổ. Tuy nhiên, NBA có thêm 15 phút nghỉ giữa hiệp 2 và hiệp 3, trong khi FIBA chỉ có 2 phút nghỉ. | Tín -Đạt - Nhã |
| What is the purpose of dividing quarters in basketball? | Mục đích của việc chia hiệp trong bóng rổ là gì? | Tín -Đạt - Nhã | Reduce fatigue for players. | Giảm bớt sự mệt mỏi cho các cầu thủ. | Tín -Đạt - Nhã |
| How to change the number of quarters in a basketball game? | Làm thế nào để thay đổi số lượng hiệp trong một trận bóng rổ? | Tín -Đạt - Nhã | The number of quarters in a basketball game is usually determined by the tournament organizers. | Việc thay đổi số lượng hiệp trong một trận bóng rổ thường được quy định bởi ban tổ chức giải đấu. | Tín -Đạt - Nhã |
| How to integrate knowledge of the number of quarters in a basketball game with playing basketball? | Làm thế nào để kết hợp kiến thức về số lượng hiệp trong một trận bóng rổ với việc chơi bóng rổ? | Tín -Đạt - Nhã | Allocate energy appropriately for each quarter. | Phân bổ sức lực hợp lý cho từng hiệp. | Tín -Đạt - Nhã |
| How to score in basketball? | Làm thế nào để ghi điểm trong bóng rổ? | Tín -Đạt - Nhã | Points can be scored by throwing the ball into the basket. A successful basket from within the 3-point line is worth 2 points, while a successful basket from outside the 3-point line is worth 3 points. Successful free throws are worth 1 point. | Có thể ghi điểm bằng cách ném bóng vào rổ. Ném rổ thành công từ trong vạch 3 điểm sẽ được 3 điểm, ném rổ thành công từ ngoài vạch 3 điểm sẽ được 2 điểm, ném phạt thành công sẽ được 1 điểm. | Tín -Đạt - Nhã |
| Besides shooting, what other ways are there to score in basketball? | Ngoài ném rổ, còn có những cách nào khác để ghi điểm trong bóng rổ? | Tín -Đạt - Nhã | Dunking: Throwing the ball into the basket from above with one hand, usually done after jumping. A successful dunk is worth 2 points. | Ném úp rổ:\*\* Ném bóng vào rổ từ trên cao bằng một tay, thường được thực hiện sau khi nhảy cao. Ném úp rổ thành công được tính 2 điểm | Tín -Đạt - Nhã |
| Compare the scoring methods between 5x5 basketball and 3x3 basketball. | So sánh cách thức ghi điểm giữa bóng rổ 5x5 và bóng rổ 3x3? | Tín -Đạt - Nhã | The scoring rules in 5x5 basketball and 3x3 basketball are similar. However, 3x3 basketball has additional "inside" (from within the 3-point line) and "outside" (outside the 3-point line) rules to add excitement to the game. | Cách thức ghi điểm trong bóng rổ 5x5 và bóng rổ 3x3 tương tự nhau. Tuy nhiên, bóng rổ 3x3 có thêm quy tắc "nội dung" (từ vạch 3 điểm trở ra) và "ngoại dung" (ngoài vạch 3 điểm) để tăng tính hấp dẫn cho trận đấu. | Tín -Đạt - Nhã |
| What is the most important skill to score in basketball? | Kỹ năng nào quan trọng nhất để ghi điểm trong bóng rổ? | Tín -Đạt - Nhã | The most important skill to score points in basketball is shooting skills. However, other skills such as passing, screening, movement, and teamwork are also important. | Kỹ năng quan trọng nhất để ghi điểm trong bóng rổ là kỹ năng ném rổ. Tuy nhiên, các kỹ năng khác như chuyền bóng, rê bóng, di chuyển và phối hợp đồng đội cũng rất quan trọng. | Tín -Đạt - Nhã |
| How to change scoring tactics in basketball? | Làm thế nào để thay đổi chiến thuật ghi điểm trong bóng rổ? | Tín -Đạt - Nhã | Analyze the strengths and weaknesses of my team and the opposing team. | Phân tích điểm mạnh và điểm yếu của đội mình và đội đối phương. | Tín -Đạt - Nhã |

Bảng 1.2 dịch câu hỏi và đánh giá câu dịch

* 1. **Sinh thêm dữ liệu**

Trình bày chi tiết ít nhất 05 luật 5 Luật tạo câu hỏi: Mở rộng, thu hẹp, so sánh, thay đổi góc nhìn, thay đổi chủ đề

* + 1. **Mở rộng phạm vi**

**Mục đích:** Giúp bạn khám phá thêm thông tin, ý tưởng mới liên quan đến chủ đề đang bàn.

**Cách thức áp dụng:**

* **Đặt câu hỏi bao quát hơn:** Thay vì hỏi về một khía cạnh cụ thể, hãy đặt câu hỏi bao quát về toàn bộ chủ đề.
* **Tìm kiếm mối liên hệ:** Xác định các khái niệm, sự kiện, con người có liên quan đến chủ đề chính.
* **Phân tích bối cảnh:** Xem xét các yếu tố xung quanh chủ đề như thời gian, địa điểm, văn hóa, xã hội.

**Ví dụ:**

* Thay vì hỏi "Tại sao bầu trời có màu xanh?", hãy hỏi "Sự xuất hiện của các màu sắc trong tự nhiên phụ thuộc vào những yếu tố nào?".
  + 1. **Thu hẹp phạm vi**

**Mục đích:** Giúp bạn tập trung vào một khía cạnh cụ thể của chủ đề, đi sâu vào chi tiết.

**Cách thức áp dụng:**

* **Đặt câu hỏi cụ thể:** Chia nhỏ chủ đề thành các phần nhỏ hơn và đặt câu hỏi về từng phần.
* **Xác định yếu tố quan trọng:** Xác định những yếu tố then chốt ảnh hưởng đến chủ đề chính.
* **Loại bỏ thông tin không liên quan:** Loại bỏ những chi tiết không cần thiết để tập trung vào vấn đề cốt lõi.

**Ví dụ:**

* Thay vì hỏi "Làm thế nào để cải thiện sức khỏe?", hãy hỏi "Chế độ dinh dưỡng nào phù hợp cho người muốn giảm cân?".
  + 1. **So sánh**

**Mục đích:** Giúp bạn nhận diện sự giống nhau và khác nhau giữa hai hoặc nhiều đối tượng, khái niệm.

**Cách thức áp dụng:**

* **Xác định điểm chung:** Tìm kiếm những điểm tương đồng giữa các đối tượng, khái niệm cần so sánh.
* **Phân tích điểm khác biệt:** Xác định những điểm khác biệt về đặc điểm, tính chất, chức năng, ...
* **Sử dụng từ ngữ so sánh:** Sử dụng các từ ngữ như "giống nhau", "khác nhau", "tương tự", "trái ngược”, ...

**Ví dụ:**

* So sánh sự giống nhau và khác nhau giữa hai loại hình thể thao: bóng đá và bóng rổ.
  + 1. **Thay đổi góc nhìn**

**Mục đích:** Giúp bạn nhìn nhận vấn đề từ nhiều khía cạnh khác nhau, đa dạng hóa cách tiếp cận.

**Cách thức áp dụng:**

* **Đặt mình vào vị trí người khác:** Xem xét vấn đề từ góc nhìn của những người có liên quan.
* **Tưởng tượng các khả năng khác nhau:** Suy nghĩ về những tình huống, kết quả có thể xảy ra.
* **Tham khảo ý kiến từ nhiều nguồn:** Tham khảo ý kiến của những người có chuyên môn, kinh nghiệm khác nhau.

**Ví dụ:**

* Nhìn nhận vấn đề ô nhiễm môi trường từ góc nhìn của người dân, doanh nghiệp và chính phủ.
  + 1. **Thay đổi chủ đề**

**Mục đích:** Giúp bạn chuyển hướng sang một chủ đề mới liên quan hoặc có liên quan gián tiếp đến chủ đề hiện tại.

**Cách thức áp dụng:**

* **Tìm kiếm mối liên kết:** Xác định mối liên hệ giữa chủ đề hiện tại và chủ đề mới.
* **Sử dụng câu chuyển tiếp:** Sử dụng các từ ngữ như "liên quan đến", "gây ra", "ảnh hưởng đến”, để kết nối hai chủ đề.
* **Đưa ra ví dụ:** Sử dụng ví dụ cụ thể để minh họa cho việc chuyển đổi chủ đề.

**Ví dụ:**

* Sau khi thảo luận về tác hại của ô nhiễm môi trường, hãy chuyển sang chủ đề về các giải pháp bảo vệ môi trường.
  1. **Đánh giá dữ liệu sinh thêm**

Ví dụ: Từ câu hỏi góc “Nước nào có nhiều huy chương bơi lội Olympic nhất?

” chúng ta có thể dùng luật để tạo ra các câu hỏi khác như:

|  |  |  |
| --- | --- | --- |
| **Sử dụng luật** | **Câu hỏi mới được tạo ra** | **Đánh giá** |
| Mở rộng phạm vi | Kể tên 3 quốc gia có số lượng vận động viên bơi lội tham dự Olympic nhiều nhất? | Thỏa mãn yêu cầu của luật, không trùng với các câu hỏi khác |
| Thu hẹp phạm vi | Nước nào có nhiều huy chương bơi lội Paralympic nhất? | Thỏa mãn yêu cầu của luật, không trùng với các câu hỏi khác |
| So sánh | So sánh thành tích bơi lội Olympic của Hoa Kỳ và Úc trong 10 năm qua? | Thỏa mãn yêu cầu của luật, không trùng với các câu hỏi khác |
| Thay đổi chủ đề | Nước nào có số lượng huy chương vàng bơi lội Olympic ở nội dung cá nhân nhiều nhất? | Thỏa mãn yêu cầu của luật, không trùng với các câu hỏi khác |
| Thay đổi góc nhìn | Nước nào có thành tích tốt nhất trong các giải bơi lội thế giới (không chỉ Olympic) và có số lượng vận động viên bơi lội tham dự Olympic nhiều nhất? | Thỏa mãn yêu cầu của luật, không trùng với các câu hỏi khác |

Bảng 1.4 Đánh giá dữ liệu sinh thêm

* 1. **Lưu bộ dữ liệu**

Lưu bộ dữ liệu đã sinh ra ở trên thành 2 bộ dữ liệu tương ứng: Data, DChatGPT. Mỗi bộ dữ liệu đều có câu hỏi/trả lời tương ứng. Nộp bộ dữ liệu lên LMS

CHƯƠNG 2

**NHẬN THỨC VÀ GIẢI PHÁP VỀ TÌNH TRẠNG KẸT XE TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HỒ CHÍ MINH (IUH)**

2.1 Câu hỏi đặt ra

*CHỦ ĐỀ: Kẹt xe tại Đại học Công nghiệp Thành phố Hồ Chí Minh (IUH)*

Kẹt xe là vấn đề phổ biến ở các trường đại học trên toàn thế giới. Kẹt xe là vấn đề phổ biến ở các trường đại học trên toàn thế giới. Sự gia tăng sử dụng xe hơi cá nhân và sự suy giảm của hệ thống giao thông công cộng và không sử dụng động cơ đã tạo ra vấn đề này. Đại học trở thành một nguồn phát sinh lưu lượng giao thông lớn, nhưng theo cách phức tạp1. Phần lớn các chuyến đi liên quan đến việc di chuyển của sinh viên từ và đến lớp.

Về tình trạng kẹt xe tại trường Đại học Công nghiệp TP.HCM (IUH), gây ra nhiều hệ lụy phiền toái cho sinh viên, ảnh hưởng đến chất lượng học tập và cuộc sống.

***Hai câu hỏi liên quan đến chủ đề***

1.Nguyên nhân gây ra kẹt xe ở IUH là gì?

2.Biện pháp nào có thể giảm thiểu tình trạng kẹt xe ở khu vực này?

*Lý do đặt câu hỏi, ý nghĩa*

1.*“Nguyên nhân gây ra kẹt xe ở IUH là gì?”* - Câu hỏi này quan trọng vì nó giúp chúng ta hiểu rõ hơn về nguồn gốc của vấn đề. Kẹt xe không chỉ gây phiền toái cho sinh viên và giáo viên, mà còn ảnh hưởng đến hiệu suất học tập và giảng dạy. Việc xác định nguyên nhân gây ra tình trạng này sẽ giúp chúng ta tìm ra giải pháp hiệu quả để khắc phục.

2*.“Biện pháp nào có thể giảm thiểu tình trạng kẹt xe ở khu vực này?”* - Câu hỏi này cũng rất quan trọng vì nó tập trung vào việc tìm kiếm giải pháp cho vấn đề. Một khi chúng ta đã hiểu rõ nguyên nhân gây ra tình trạng kẹt xe, chúng ta có thể bắt đầu tìm kiếm các biện pháp để giảm thiểu hoặc loại bỏ hoàn toàn vấn đề này. Điều này không chỉ cải thiện chất lượng cuộc sống cho sinh viên và giáo viên, mà còn giúp cải thiện hiệu suất học tập và giảng dạy

2.2 Xây dựng câu hỏi khảo sát

*Cách thức xây dựng các câu hỏi khảo sát*

Cách thức xây dựng 10 câu hỏi này dựa trên hai yếu tố chính: nguyên nhân gây ra tình trạng kẹt xe và biện pháp giảm thiểu tình trạng này. Các câu hỏi được thiết kế để thu thập thông tin về cảm nhận cá nhân của người trả lời về vấn đề này, từ đó giúp hiểu rõ hơn về mức độ nghiêm trọng của vấn đề và tìm ra các giải pháp phù hợp. Cụ thể, các câu hỏi tập trung vào các yếu tố như ảnh hưởng của sự gia tăng sinh viên và phương tiện lưu thông, cơ sở hạ tầng, việc mở rộng cơ sở vật chất, điều chỉnh lịch học, và việc tăng cường giao thông công cộng.

*10 câu hỏi khảo sát và giải thích*

1. Bạn cảm nhận mức độ nghiêm trọng của tình trạng kẹt xe tại IUH như thế nào?

* Câu hỏi này được xây dựng để thu thập thông tin về cảm nhận cá nhân của người trả lời về tình trạng kẹt xe tại IUH. Điều này giúp hiểu rõ hơn về mức độ nghiêm trọng của vấn đề từ góc nhìn của người trả lời.

1. Bạn đánh giá mức độ nghiêm trọng của ảnh hưởng từ sự gia tăng sinh viên và phương tiện lưu thông tại IUH đến tình trạng kẹt xe như thế nào?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về ảnh hưởng của sự gia tăng sinh viên và phương tiện lưu thông đến tình trạng kẹt xe.

1. Bạn đánh giá mức độ ảnh hưởng của cơ sở hạ tầng như bãi đậu xe và các lối vào trường đến tình trạng kẹt xe như thế nào?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về ảnh hưởng của cơ sở hạ tầng đến tình trạng kẹt xe.

1. Bạn nhận thấy hiệu quả của việc mở rộng cơ sở vật chất như bãi đậu xe và các lối vào trường trong việc giảm bớt tình trạng kẹt xe ra sao?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về hiệu quả của việc mở rộng cơ sở vật chất trong việc giảm bớt tình trạng kẹt xe.

1. Theo bạn, việc điều chỉnh lịch học để giảm bớt sinh viên và phương tiện tại trường cùng một lúc có hiệu quả như thế nào?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về hiệu quả của việc điều chỉnh lịch học để giảm bớt sinh viên và phương tiện tại trường cùng một lúc.

1. Bạn đánh giá mức độ ảnh hưởng của việc tăng cường giao thông công cộng đến tình trạng kẹt xe như thế nào?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về ảnh hưởng của việc tăng cường giao thông công cộng đến tình trạng kẹt xe.

1. Bạn đánh giá mức độ nghiêm trọng của việc khuyến khích sử dụng phương tiện công cộng để giảm lượng xe cá nhân như thế nào?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về mức độ nghiêm trọng của việc khuyến khích sử dụng phương tiện công cộng để giảm lượng xe cá nhân.

1. Bạn có đề xuất cụ thể nào khác để giải quyết tình trạng kẹt xe tại IUH không?

* Câu hỏi này nhằm thu thập thông tin về các đề xuất cụ thể của người trả lời để giải quyết tình trạng kẹt xe tại IUH.

1. Theo bạn, đâu là nguyên nhân chính dẫn đến tình trạng kẹt xe tại IUH?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về nguyên nhân chính dẫn đến tình trạng kẹt xe tại IUH.

1. Bạn nghĩ gì về việc sử dụng phương tiện công cộng như xe buýt hay xe đạp điện để giảm lượng xe cá nhân?

* Câu hỏi này nhằm thu thập thông tin về cảm nhận của người trả lời về việc sử dụng phương tiện công cộng như xe buýt hay xe đạp điện để giảm lượng xe cá nhân.

2.3 Phân tích

2.3.1 Kết quả khảo sát

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phiếu khảo sát** | **Cách thức tiến hành khảo sát** | **Thời gian** | **Số lượng** | **Đối tượng** | **Kết quả thu được** |
| Phiếu khảo sát về tình trạng kẹt xe tại IUH | Phân phối biểu mẫu khảo sát trực tuyến | 1 tuần | 50 | Sinh viên IUH | 50/50 |

Bảng 2.3.1 kế hoạch khảo sát

2.3.2 Phân tích dữ liệu

Tiền xử lý dữ liệu

* Chuẩn bị tệp dữ liệu: Sau khi tiến hành khảo sát, kết quả sẽ được lưu vào file xlsx
* Dọc dữ liệu : đọc dữ liệu từ file đã thu được bằng công cụ như python
* Làm sạch dữ liệu: Loại bỏ các giá trị ngoại lệ, điểm dữ liệu trùng lặp, và sửa chữa các lỗi nhập liệu.
* Chuẩn hóa dữ liệu: Đưa các biến về cùng một thang đo để không có biến nào bị chi phối trong quá trình phân tích.
* Mã hóa dữ liệu phân loại: Chuyển đổi các biến phân loại thành dạng số để có thể xử lý trong mô hình học máy.
* Phân chia tập dữ liệu: Tách tập dữ liệu thành tập huấn luyện và tập kiểm tra để đánh giá mô hình.
* Tính năng mở rộng quy mô: Áp dụng các phương pháp như Min-Max Scaling hoặc Standardization để điều chỉnh phạm vi của dữ liệu.

**Kết quả khảo sát**

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Nhiều màu sắc

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Nhiều màu sắc

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Mô tả được tạo tự động

Ảnh có chứa văn bản, biểu đồ, ảnh chụp màn hình, Phông chữ

Mô tả được tạo tự động

Ảnh có chứa văn bản, ảnh chụp màn hình, biểu đồ, Phông chữ

Mô tả được tạo tự động

2.4 Kết luận

2.4.1 Kết luận

Quá trình thu thập dữ liệu và phân tích dữ liệu thành biểu đồ pie là một phương pháp hiệu quả để đánh giá kết quả và trực quan hóa thông tin.

Đầu tiên, quá trình thu thập dữ liệu đảm bảo rằng chúng ta có đủ thông tin và dữ liệu chính xác từ nguồn tài liệu đáng tin cậy.

Tiếp theo, việc phân tích dữ liệu được thực hiện bằng cách sử dụng công cụ và phần mềm thích hợp để xử lý và biểu diễn dữ liệu.

Biểu đồ pie là một cách hiệu quả để trực quan hóa dữ liệu và đánh giá kết quả. Nó cho phép chúng ta nhìn thấy tỷ lệ phần trăm của mỗi phần tử trong dữ liệu và so sánh chúng với nhau.

Việc sử dụng biểu đồ pie giúp chúng ta nhanh chóng nhận ra các xu hướng, sự phân bố và mối quan hệ giữa các phần tử trong dữ liệu.

Từ biểu đồ pie, chúng ta có thể rút ra những kết luận và đánh giá về kết quả. Ví dụ, chúng ta có thể nhìn thấy tỷ lệ phần trăm của các nhóm dữ liệu và đưa ra nhận xét về sự đa dạng hoặc tập trung của dữ liệu.

2.4.2 Hạn chế

Các câu hỏi khảo sát còn gặp một số hạn chế như :

* Chủ quan: Câu hỏi dựa trên cảm nhận cá nhân có thể không phản ánh chính xác mức độ nghiêm trọng của tình trạng kẹt xe, vì mỗi người có thể cảm nhận khác nhau.
* Thiếu thông tin cụ thể: Câu hỏi không cung cấp dữ liệu cụ thể về số lượng sinh viên tăng lên hay lượng phương tiện lưu thông, khiến người trả lời khó đánh giá mức độ ảnh hưởng.
* Giả định: Câu hỏi giả định rằng việc mở rộng cơ sở vật chất sẽ có hiệu quả trong việc giảm bớt tình trạng kẹt xe mà không đề cập đến các yếu tố khác có thể ảnh hưởng.
* Thiếu đối chiếu: Câu hỏi không so sánh việc điều chỉnh lịch học với các biện pháp khác, khiến cho việc đánh giá hiệu quả trở nên thiếu cơ sở.
* Không đề cập đến tần suất và chất lượng: Câu hỏi không đề cập đến tần suất và chất lượng của giao thông công cộng, là những yếu tố quan trọng ảnh hưởng đến việc sử dụng phương tiện này.
* Thiếu đa dạng quan điểm: Câu hỏi không khuyến khích người trả lời xem xét nhiều nguyên nhân khác nhau có thể dẫn đến tình trạng kẹt xe.
* Thiếu thông tin về phương tiện công cộng: Câu hỏi không cung cấp thông tin về tình trạng hiện tại của phương tiện công cộng, như tần suất, độ tin cậy, và sự tiện lợi.

**2.4.3 Hướng phát triển trong tương lai**

Trong tương lai, chúng tôi sẽ tiếp tục nghiên cứu và phát triển các mô hình dự đoán để hiểu rõ hơn về các yếu tố ảnh hưởng đến tình trạng kẹt xe tại IUH. Chúng tôi dự định sẽ phân tích sâu hơn về mối liên hệ giữa số lượng sinh viên, phương tiện lưu thông, và cơ sở hạ tầng với mức độ kẹt xe. Bằng cách này, chúng tôi hy vọng sẽ xác định được các giải pháp cụ thể và hiệu quả để giảm thiểu tình trạng này, đồng thời cải thiện chất lượng cuộc sống và hiệu quả giao thông trong khu vực trường học

CHƯƠNG 3

**KHAI PHÁ DỮ LIỆU TỪ CÂU NÓI CỦA NGƯỜI NỔI TIẾNG**

3.1 Thu thập dữ liệu

Cách thu thập dữ liệu

**Khởi tạo các biến**: Mã khởi tạo các biến như base\_url, url, và count. base\_url là URL cơ sở của trang web muốn cào dữ liệu. url là URL hiện tại mà mã đang cào, và count là số lượng trích dẫn đã được cào.

**Mở tệp để ghi:** Mã mở một tệp tên kq.txt để ghi dữ liệu đã cào.

Vòng lặp while: Mã chạy một vòng lặp while, cào dữ liệu từ mỗi trang cho đến khi không còn trang nào hoặc đã cào đủ 40 trích dẫn.

**Gửi yêu cầu GET**: Mã gửi một yêu cầu GET đến URL hiện tại và nhận về phản hồi.

**Phân tích HTML:** Mã sử dụng BeautifulSoup để phân tích HTML từ phản hồi.

Tìm tất cả các trích dẫn: Mã tìm tất cả các phần tử div có class là quote. Mỗi phần tử như vậy chứa một trích dẫn.

**Lặp qua mỗi trích dẫn:** Mã lặp qua mỗi trích dẫn, tìm nội dung trích dẫn và tác giả, sau đó ghi chúng vào tệp.

**Tìm trang tiếp theo:** Mã tìm liên kết đến trang tiếp theo và cập nhật url để cào trang tiếp theo trong lần lặp tiếp theo của vòng lặp while.

In thông báo hoàn thành: Khi vòng lặp while kết thúc, mã in ra thông báo “Hoàn thành cào dữ liệu!”.

kết quả sau khi cào

|  |  |
| --- | --- |
| Albert Einstein | “The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking.” |
| J.K. Rowling | “It is our choices, Harry, that show what we truly are, far more than our abilities.” |
| Albert Einstein | “There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle.” |
| Jane Austen | “The person, be it gentleman or lady, who has not pleasure in a good novel, must be intolerably stupid.” |
| Marilyn Monroe | “Imperfection is beauty, madness is genius and it's better to be absolutely ridiculous than absolutely boring.” |
| Albert Einstein | “Try not to become a man of success. Rather become a man of value.” |
| André Gide | “It is better to be hated for what you are than to be loved for what you are not.” |
| Thomas A. Edison | “I have not failed. I've just found 10,000 ways that won't work.” |
| Eleanor Roosevelt | “A woman is like a tea bag; you never know how strong it is until it's in hot water.” |
| Steve Martin | “A day without sunshine is like, you know, night.” |
| Marilyn Monroe | “This life is what you make it. No matter what, you're going to mess up sometimes, it's a universal truth. But the good part is you get to decide how you're going to mess it up. Girls will be your friends - they'll act like it anyway. But just remember, some come, some go. The ones that stay with you through everything - they're your true best friends. Don't let go of them. Also remember, sisters make the best friends in the world. As for lovers, well, they'll come and go too. And baby, I hate to say it, most of them - actually pretty much all of them are going to break your heart, but you can't give up because if you give up, you'll never find your soulmate. You'll never find that half who makes you whole and that goes for everything. Just because you fail once, doesn't mean you're gonna fail at everything. Keep trying, hold on, and always, always, always believe in yourself, because if you don't, then who will, sweetie? So keep your head high, keep your chin up, and most importantly, keep smiling, because life's a beautiful thing and there's so much to smile about.” |
| J.K. Rowling | “It takes a great deal of bravery to stand up to our enemies, but just as much to stand up to our friends.” |
| Albert Einstein | “If you can't explain it to a six-year-old, you don't understand it yourself.” |
| Bob Marley | “You may not be her first, her last, or her only. She loved before she may love again. But if she loves you now, what else matters? She's not perfect—you aren't either, and the two of you may never be perfect together but if she can make you laugh, cause you to think twice, and admit to being human and making mistakes, hold onto her and give her the most you can. She may not be thinking about you every second of the day, but she will give you a part of her that she knows you can break—her heart. So don't hurt her, don't change her, don't analyze and don't expect more than she can give. Smile when she makes you happy, let her know when she makes you mad, and miss her when she's not there.” |
| Dr. Seuss | “I like nonsense, it wakes up the brain cells. Fantasy is a necessary ingredient in living.” |
| Douglas Adams | “I may not have gone where I intended to go, but I think I have ended up where I needed to be.” |
| Elie Wiesel | “The opposite of love is not hated, it's indifference. The opposite of art is not ugliness, it's indifference. The opposite of faith is not heresy, it's indifference. And the opposite of life is not death, it's indifference.” |
| Friedrich Nietzsche | “It is not a lack of love, but a lack of friendship that makes unhappy marriages.” |
| Mark Twain | “Good friends, good books, and a sleepy conscience: this is the ideal life.” |
| Allen Saunders | “Life is what happens to us while we are making other plans.” |
| Pablo Neruda | “I love you without knowing how, or when, or from where. I love you simply, without problems or pride: I love you in this way because I do not know any other way of loving but this, in which there is no I or you, so intimate that your hand upon my chest is my hand, so intimate that when I fall asleep your eyes close.” |
| Ralph Waldo Emerson | “For every minute you are angry you lose sixty seconds of happiness.” |
| Mother Teresa | “If you judge people, you have no time to love them.” |
| Garrison Keillor | “Anyone who thinks sitting in church can make you a Christian must also think that sitting in a garage can make you a car.” |
| Jim Henson | “Beauty is in the eye of the beholder, and it may be necessary from time to time to give a stupid or misinformed beholder a black eye.” |
| Dr. Seuss | “Today you are You, that is truer than true. There is no one alive who is Youer than You.” |
| Albert Einstein | “If you want your children to be intelligent, read them fairy tales. If you want them to be more intelligent, read them more fairy tales.” |
| J.K. Rowling | “It is impossible to live without failing at something, unless you live so cautiously that you might as well not have lived at all - in which case, you fail by default.” |
| Albert Einstein | “Logic will get you from A to Z; imagination will get you everywhere.” |
| Bob Marley | “One good thing about music, when it hits you, you feel no pain.” |
| Dr. Seuss | “The more that you read, the more things you will know. The more that you learn, the more places you'll go.” |
| J.K. Rowling | “Of course it is happening inside your head, Harry, but why on earth should that mean that it is not real?” |
| Bob Marley | “The truth is, everyone is going to hurt you. You just got to find the ones worth suffering for.” |
| Mother Teresa | “Not all of us can do great things. But we can do small things with great love.” |
| J.K. Rowling | “To the well-organized mind, death is but the next great adventure.” |
| Charles M. Schulz | “All you need is love. But a little chocolate now and then doesn't hurt.” |
| William Nicholson | “We read to know we're not alone.” |
| Albert Einstein | “Any fool can know. The point is to understand.” |
| Jorge Luis Borges | “I have always imagined that Paradise will be a kind of library.” |
| George Eliot | “It is never too late to be what you might have been.” |

Bảng 3.1 file kq.txt sau khi cào

3.2 Khai phá dữ liệu

**3.2.1. Xử lý dữ liệu- Data Imputation**

***Bạn hãy thêm vào Trường STT và điền tự động dữ liệu của trường này?***

* Cách tiếp cận giải quyết:
* Trước tiên, ta khai báo biến stt = 1 để làm biến đếm cho trường STT.
* Trong vòng lặp for quote in quotes, sau khi thu thập thông tin của mỗi tác giả, ta sẽ thêm giá trị của biến stt vào danh sách data như sau: data. append ([stt, author\_name, author\_link, author\_birthday, famous\_quote, age]).
* Sau mỗi vòng lặp, ta tăng giá trị của biến stt lên 1: stt += 1.
* Kết quả đạt được:
* Trường STT sẽ được thêm vào và các tác giả sẽ có giá trị STT tương ứng.
* Đánh giá:
* Phương pháp sử dụng biến đếm trong vòng lặp để đánh số thứ tự cho các tác giả là phù hợp và hiệu quả. Nó cho phép ta tự động điền dữ liệu vào trường STT mà không cần phải thao tác thủ công.
* Thảo luận:
* Cách tiếp cận này có thể được áp dụng cho việc thêm các trường dữ liệu khác vào tệp CSV nếu cần.

***Một số giá trị của dữ liệu Trường ngày sinh chưa có, bạn hãy đề xuất cách điền?***

* Cách tiếp cận giải quyết:
* Trước khi thu thập thông tin ngày sinh của mỗi tác giả, ta kiểm tra xem giá trị ngày sinh có rỗng hay không hợp lệ.
* Nếu giá trị ngày sinh là rỗng hoặc không hợp lệ, ta gán giá trị "Unknown" hoặc "N/A" cho trường ngày sinh.
* Kết quả đạt được:
* Các giá trị trường ngày sinh sẽ được điền bằng giá trị "Unknown" hoặc "N/A" cho các tác giả mà thông tin ngày sinh chưa có sẵn.
* Đánh giá:
* Phương pháp sử dụng giá trị "Unknown" hoặc "N/A" để điền vào các giá trị trống là một cách đơn giản và dễ thực hiện. Nó giúp đảm bảo rằng thông tin không có sẵn được đại diện một cách rõ ràng trong tệp CSV.
* Thảo luận:
* Cách tiếp cận này có thể được sử dụng cho các trường dữ liệu khác trong tệp CSV nếu không có thông tin tương ứng.

***Bạn hãy thêm vào Trường Tuoi (Tuổi) và đề xuất cách điền tuổi của các tác giả?***

* Cách tiếp cận giải quyết:
* Sau khi thu thập thông tin ngày sinh của mỗi tác giả, ta tính tuổi bằng cách lấy năm hiện tại trừ đi năm sinh.
* Ta khai báo biến current\_year để lưu trữ năm hiện tại.
* Trong vòng lặp for quote in quotes, ta tính tuổi của tác giả bằng cách lấy current\_year - birth\_year.
* Sau khi tính được giá trị tuổi của tác giả, ta thêm giá trị tuổi này vào danh sách data.
* Kết quả đạt được:
* Mỗi tác giả sẽ có một giá trị tuổi tương ứng trong trường Tuổi.
* Đánh giá:
* Phương pháp tính tuổi bằng cách lấy năm hiện tại trừ đi năm sinh là một cách đơn giản và chính xác để điền vào trường Tuổi. Nó giúp ta có thông tin về tuổi của tác giả mà không cần phải thao tác thủ công.
* Thảo luận:
* Cách tiếp cận này có thể được sử dụng để tính toán và điền vào các trường dữ liệu khác liên quan đến tuổi của các đối tượng

**Kết quả sau khi tiến hành**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STT | Tên tác giả | Đường link của tác giả | Ngày tháng năm sinh | Câu nói nổi tiếng của tác giả | Tuổi |
| 1 | Albert Einstein | http://quotes.toscrape.com//author/Albert-Einstein | March 14, 1879 | “The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking.” | 145 |
| 2 | J.K. Rowling | http://quotes.toscrape.com//author/J-K-Rowling | July 31, 1965 | “It is our choices, Harry, that show what we truly are, far more than our abilities.” | 59 |
| 3 | Albert Einstein | http://quotes.toscrape.com//author/Albert-Einstein | March 14, 1879 | “There are only two ways to live your life. One is as though nothing is a miracle. The other is as though everything is a miracle.” | 145 |
| 4 | Jane Austen | http://quotes.toscrape.com//author/Jane-Austen | December 16, 1775 | “The person, be it gentleman or lady, who has not pleasure in a good novel, must be intolerably stupid.” | 249 |
| 5 | Marilyn Monroe | http://quotes.toscrape.com//author/Marilyn-Monroe | June 01, 1926 | “Imperfection is beauty, madness is genius and it's better to be absolutely ridiculous than absolutely boring.” | 98 |
| 6 | Albert Einstein | http://quotes.toscrape.com//author/Albert-Einstein | March 14, 1879 | “Try not to become a man of success. Rather become a man of value.” | 145 |
| … | … | ………… | ……. | ……. | … |

Bảng 3.2.1 ví dụ sau khi thêm các trường dữ liệu

**3.2.2. Khám phá dữ liệu- Data Exploration**

***Thống kê về tác giả và câu nói nổi tiếng có trong bộ dữ liệu,***

* Cách tiếp cận giải quyết:
* Đầu tiên, chúng ta sử dụng thư viện pandas để đọc dữ liệu từ file CSV.
* Tiếp theo, chúng ta sử dụng hàm nunique() để tính số lượng tác giả duy nhất trong cột 'Tên tác giả' và in kết quả.
* Sau đó, chúng ta sử dụng hàm nunique() để tính số lượng câu nói nổi tiếng duy nhất trong cột 'Câu nói nổi tiếng của tác giả' và in kết quả.
* Cuối cùng, chúng ta sử dụng phương thức groupby() và hàm count() để tính số lượng câu nói nổi tiếng trung bình mỗi tác giả và in kết quả.
* Kết quả đạt được:
* Kết quả thống kê về tác giả và câu nói nổi tiếng bao gồm số lượng tác giả duy nhất, số lượng câu nói nổi tiếng duy nhất và số lượng câu nói nổi tiếng trung bình mỗi tác giả.
* Đánh giá và thảo luận:
* Kết quả thống kê giúp chúng ta hiểu được sự đa dạng về tác giả và câu nói nổi tiếng trong bộ dữ liệu.
* Số lượng tác giả duy nhất cho biết về sự đa dạng của nguồn cung tác giả.
* Số lượng câu nói nổi tiếng duy nhất cho biết về sự độc đáo của các câu nói trong bộ dữ liệu.
* Số lượng câu nói nổi tiếng trung bình mỗi tác giả cho biết về sự phân bổ câu nói nổi tiếng giữa các tác giả

***Thống kê về năm sinh và độ tuổi của các tác giả,***

* Cách tiếp cận giải quyết:
* Sử dụng thư viện pandas để đọc dữ liệu từ file CSV.
* Sử dụng hàm min () và max () để tính năm sinh nhỏ nhất và lớn nhất của các tác giả.
* Sử dụng hàm mean () để tính độ tuổi trung bình của các tác giả.
* Kết quả đạt được:
* Năm sinh nhỏ nhất của các tác giả: [giá trị]
* Năm sinh lớn nhất của các tác giả: [giá trị]
* Độ tuổi trung bình của các tác giả: [giá trị]
* Đánh giá và thảo luận:
* Kết quả thống kê giúp chúng ta hiểu được phạm vi năm sinh và độ tuổi của các tác giả trong bộ dữ liệu.
* Năm sinh nhỏ nhất và lớn nhất cho biết về khoảng thời gian hoạt động của các tác giả.
* Độ tuổi trung bình cho biết về độ trẻ hoặc già của các tác giả trong bộ dữ liệu

***Thống kê về các câu nói nổi tiếng như: câu dài nhất, ngắn nhất, số từ, …***

* Cách tiếp cận giải quyết:
* Sử dụng thư viện pandas để đọc dữ liệu từ file CSV.
* Tạo một cột mới 'Số từ' bằng cách sử dụng hàm split () để tách các từ trong mỗi câu và đếm số lượng từ.
* Sử dụng hàm max () và min () để tìm câu dài nhất và câu ngắn nhất dựa trên số lượng từ.
* Sử dụng hàm mean () để tính số từ trung bình trong các câu.
* Kết quả đạt được:
* Câu dài nhất: [câu]
* Câu ngắn nhất: [câu]
* Số từ trung bình trong các câu: [giá trị]
* Đánh giá và thảo luận:
* Kết quả thống kê cho biết về độ dài và số từ trong các câu nói nổi tiếng trong bộ dữ liệu.
* Câu dài nhất và câu ngắn nhất cho thấy phạm vi độ dài của các câu nói.
* Số từ trung bình giúp chúng ta hiểu về độ phong phú hay ngắn gọn của các câu trong bộ dữ liệu.

***Thống kê về các từ được sử dụng trong các câu nói,***

* Cách tiếp cận giải quyết:
* Sử dụng thư viện pandas để đọc dữ liệu từ file CSV.
* Tạo một danh sách chứa tất cả các từ trong các câu nói bằng cách sử dụng hàm split () để tách các từ trong mỗi câu.
* Sử dụng hàm Counter () để đếm số lần xuất hiện của từng từ trong danh sách.
* Sắp xếp danh sách theo thứ tự giảm dần của số lần xuất hiện để lấy ra các từ phổ biến nhất.
* Kết quả đạt được:
* Danh sách các từ phổ biến nhất: [danh sách]
* Đánh giá và thảo luận:
* Kết quả thống kê giúp chúng ta hiểu được sự phân bố và sự đa dạng của các từ trong các câu nói trong bộ dữ liệu.
* Các từ phổ biến nhất cho biết về các từ được sử dụng nhiều nhất trong các câu nói.
* Có thể thực hiện việc xử lý và lọc các từ không cần thiết hoặc stopwords để tăng tính chính xác của kết quả.

***Phân tích, trực quan mối quan hệ giữa giữa tác giả và câu nói nổi tiếng,***

* Cách tiếp cận giải quyết:
* Sử dụng thư viện pandas để đọc dữ liệu từ file CSV.
* Tạo một bảng tần số (frequency table) đếm số lần xuất hiện của mỗi tác giả và câu nói nổi tiếng.
* Sử dụng biểu đồ cột (bar chart) hoặc biểu đồ tròn (pie chart) để trực quan hóa mối quan hệ giữa tác giả và câu nói nổi tiếng.
* Kết quả đạt được:
* Biểu đồ cột hoặc biểu đồ tròn trực quan hóa mối quan hệ giữa tác giả và câu nói nổi tiếng.
* Đánh giá và thảo luận:
* Kết quả phân tích và trực quan hóa giúp chúng ta hiểu rõ hơn về phân phối và sự tương quan giữa tác giả và câu nói nổi tiếng trong bộ dữ liệu.
* Biểu đồ cột hoặc biểu đồ tròn giúp chúng ta dễ dàng nhìn thấy tác giả nào có số lượng câu nói nổi tiếng cao nhất.

**3.2.3. Trích xuất đặc trưng- Feature Extraction**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Yêu cầu** | **Cách tiếp cận** | **Kết quả đạt được** | **Đánh giá** | **Thảo luận** |
| Đề xuất cách trích xuất đặc trưng từ bộ dữ liệu. | Phân tích nội dung câu nói:  Tính số lượng từ trong mỗi câu nói.  Xác định các từ khóa chính sử dụng phân tích tần suất từ (TF-IDF).  Phân tích thông tin tác giả:  Trích xuất năm sinh và tính tuổi của tác giả.  Phân loại tác giả theo thời kỳ hoạt động (ví dụ: hiện đại, cổ điển).  Phân tích liên kết:  Xác định số lượng liên kết ngoại lai từ trang web của tác giả | Tập hợp các đặc trưng bao gồm số lượng từ, từ khóa, tuổi tác giả, thời kỳ hoạt động, và số lượng liên kết ngoại lai. | Các đặc trưng này có thể hỗ trợ trong việc phân loại câu nói và tác giả, cũng như trong việc xây dựng các mô hình học máy. | Việc lựa chọn đặc trưng phụ thuộc vào mục tiêu cụ thể của việc phân tích dữ liệu, như phân loại, clustering, hoặc khuyến nghị. |

**3.2.4. Suy luận**

***Hãy dự đoán tên của người nổi tiếng theo câu nói dựa trên các đặc trưng bạn trích xuất ở trên và đánh giá trên bộ dữ liệu đã cho với tỉ lệ Train/Test và các độ đo phù hợp?***

* Cách tiếp cận giải quyết:
* Đầu tiên, chúng ta import các thư viện cần thiết như pandas, TfidfVectorizer, train\_test\_split, accuracy\_score, f1\_score, LogisticRegression, MultinomialNB, LinearSVC và RandomForestClassifier.
* Tiếp theo, chúng ta định nghĩa một lớp Model để huấn luyện và đánh giá các mô hình phân loại.
* Sau đó, chúng ta đọc dữ liệu từ file csv bằng pandas và lưu trữ văn bản vào biến x và nhãn vào biến y.
* Tiến hành tiền xử lý dữ liệu bằng cách sử dụng TfidfVectorizer với các tham số như lowercase=True, strip\_accents='unicode', stop\_words='english'.
* Chia dữ liệu thành tập huấn luyện và tập kiểm tra bằng hàm train\_test\_split.
* Tiếp theo, chúng ta huấn luyện và đánh giá các mô hình phân loại (Multinomial Naive Bayes, Linear Support Vector Machine, Random Forest Classifier và Logistic Regression) bằng cách sử dụng phương thức fit của lớp Model.
* Cuối cùng, chúng ta dự đoán tên của người nổi tiếng dựa trên câu nói và in kết quả dự đoán ra màn hình
* Kết quả đạt được:
* Huấn luyện các mô hình phân loại trên tập huấn luyện.
* Đánh giá các mô hình bằng các độ đo phù hợp như accuracy score và F1-score.
* Dự đoán tên của người nổi tiếng dựa trên câu nói và in ra kết quả dự đoán và tên thực tế của tác giả.
* Đánh giá:
* Đánh giá các mô hình phân loại dựa trên độ chính xác và F1-score để đánh giá hiệu suất dự đoán.
* Kết quả dự đoán có thể được so sánh với tên thực tế của tác giả để xem mô hình có đạt được kết quả chính xác không.
* Thảo luận:
* Kết quả dự đoán có thể được sử dụng để tìm hiểu về sự tương đồng giữa các câu nói và tác giả.
* Có thể thử nghiệm với các mô hình phân loại khác nhau hoặc các phương pháp xử lý dữ liệu khác để cải thiện kết quả dự đoán.

***Hãy đề xuất cách tính độ tương đồng phong cách nói giữa các tác giả và tìm ra các tác giả có phong cách nói tương đồng nhau nhất?***

* Cách tiếp cận giải quyết:
* Sử dụng mô hình Naive Bayes để phân loại câu nói theo tên người nổi tiếng.
* Sử dụng vector TF-IDF để biểu diễn câu nói và tính độ tương đồng phong cách nói giữa các tác giả bằng độ đo cosine similarity.
* Kết quả đạt được:
* Kết quả là mô hình có thể phân loại câu nói theo tên người nổi tiếng và xác định được độ tương đồng phong cách nói giữa các tác giả.
* Kết quả cụ thể là danh sách các tác giả có phong cách nói tương đồng nhau nhất, được sắp xếp theo độ tương đồng giảm dần.
* Đánh giá:
* Đánh giá mô hình được thực hiện bằng cách tính độ chính xác (accuracy) của mô hình Naive Bayes trong việc dự đoán tên người nổi tiếng dựa trên câu nói.
* Độ chính xác của mô hình có thể được sử dụng để đánh giá hiệu quả của quá trình phân loại câu nói.
* Thảo luận:
* Cần thảo luận và đánh giá thêm về phương pháp tính độ tương đồng phong cách nói và cách phân loại câu nói dựa trên tên người nổi tiếng.
* Có thể xem xét các phương pháp khác để tính độ tương đồng phong cách nói, như sử dụng mô hình học máy khác hoặc các phương pháp xử lý ngôn ngữ tự nhiên khác.
* Thảo luận về các cách cải thiện mô hình, như tăng cường dữ liệu huấn luyện, tinh chỉnh siêu tham số, hoặc sử dụng các biểu diễn ngôn ngữ sâu hơn.

**LÀM VIỆC NHÓM**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Thành Viên** | **Nhiệm vụ** | **Bài 1** | **Bài 2** | **Bài 3** | **Số lần gặp (lần)** | **Thời gian gặp(h)** | **Hình thức** |
| Nguyễn Văn Thành  Nguyễn Trường Vương | Tiến hành lên ý tưởng chủ về chủ đề | x |  |  | 1 | 2 | Trực tiếp |
| Nguyễn Văn Thành  Nguyễn Trường Vương | Tiến hành thực hiện các yêu cầu còn lại | x |  |  | 1 | 4 | Trực tuyến qua ZOOM |
| Nguyễn Văn Thành  Nguyễn Trường Vương | Làm việc lên ý tưởng vè chủ đề khảo sát |  | x |  | 1 | 1 | Trực tiếp |
| Nguyễn Trường Vương | Tiến hành thực hiện các câu hỏi và làm phiếu khảo sát |  | x |  |  |  |  |
| Nguyễn Văn Thành  Nguyễn Trường Vương | Tiến hành khảo sát |  | x |  | 2 | 4 | Trực tiếp |
| Nguyễn Văn Thành | Thực hiện các yêu cầu bài 2 |  | x |  |  |  |  |
| Nguyễn Trường Vương | Tổng hợp lại toàn bộ bài 1 và 2 | x | x |  |  |  | Trực tuyến qua github |
| Nguyễn Văn Thành | Thực hiện code các yêu cầu của bài |  |  | x |  |  |  |
| Nguyễn Trường Vương | Check lại Code và hoàn thiện |  |  | x |  |  | Trực tuyến qua github |
| Nguyễn Văn Thành  Nguyễn Trường Vương | Tiến hành check lại toàn bộ bài | x | x | x | 4 | 8 | Trực tiếp và qua zoom |

**TÀI LIỆU THAM KHẢO**

* 1. PGS.TS. Nguyễn Hữu Dũng (2012) Quan điểm của Ông về dịch thuật
  2. Xây dựng nhóm: <https://www.greenhand.vn/suc-khoe-cong-dong/cach-xay-dung-doi-nhom-hieu-qua-chi-voi-5-nguyen-tac>
  3. Các bước xử lý dữ liệu : <https://vi.wikipedia.org/wiki/X%E1%BB%AD_l%C3%AD_d%E1%BB%AF_li%E1%BB%87u>
  4. <https://scikit-learn.org/stable/modules/preprocessing.html>
  5. <http://rpubs.com/phamdinhkhanh/398690>
  6. <https://phamdinhkhanh.github.io/deepai-book/ch_ml/FeatureEngineering.html>

**PHỤ LỤC**

**Phiếu Khảo Sát Nhận Thức và Giải Pháp Về Tình Trạng Kẹt Xe Tại Trường Đại Học Công Nghiệp Hồ Chí Minh (IUH)**

**PHẦN I. THÔNG TIN CHUNG**

**Tên của bạn:………………………………………………………......................**

**Ngành học:………………………...……Lớp:…………………........................**

**PHẦN 2. KHẢO SÁT**

**Phần khảo sát (từ câu 1 đến 7) bao gồm các đáp án đánh giá mức độ nghiêm trọng của câu hỏi từ “Không nghiêm trọng” đến “Rất nghiêm trọng” bao gồm:**

1. **Không nghiêm trọng**

**2. Hơi nghiêm trọng**

**3. Trung bình**

**4. Nghiêm trọng**

**5. Rất nghiêm trọng**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** |
| 1. 1. Bạn cảm nhận mức độ nghiêm trọng của tình trạng kẹt xe tại IUH như thế nào? |  |  |  |  |  |
| 1. 2. Bạn đánh giá mức độ nghiêm trọng của ảnh hưởng từ sự gia tăng sinh viên và phương tiện lưu thông tại IUH đến tình trạng kẹt xe như thế nào? |  |  |  |  |  |
| 1. 3. Bạn đánh giá mức độ nghiêm trọng của ảnh hưởng từ sự gia tăng sinh viên và phương tiện lưu thông tại IUH đến tình trạng kẹt xe như thế nào? |  |  |  |  |  |
| 1. 4. Bạn nhận thấy hiệu quả của việc mở rộng cơ sở vật chất như bãi đậu xe và các lối vào trường trong việc giảm bớt tình trạng kẹt xe ra sao? |  |  |  |  |  |
| 1. 5. Theo bạn, việc điều chỉnh lịch học để giảm bớt sinh viên và phương tiện tại trường cùng một lúc có hiệu quả như thế nào? |  |  |  |  |  |
| 1. 6 Bạn đánh giá mức độ ảnh hưởng của việc tăng cường giao thông công cộng đến tình trạng kẹt xe như thế nào ? |  |  |  |  |  |
| 1. 7. Bạn đánh giá mức độ nghiêm trọng của việc khuyến khích sử dụng phương tiện công cộng để giảm lượng xe cá nhân như thế nào? |  |  |  |  |  |

**Câu hỏi với câu trả lời là text**

1. Bạn có đề xuất cụ thể nào khác để giải quyết tình trạng kẹt xe tại IUH không?

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Theo bạn, đâu là nguyên nhân chính dẫn đến tình trạng kẹt xe tại IUH?

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

1. Bạn nghĩ gì về việc sử dụng phương tiện công cộng như xe buýt hay xe đạp điện để giảm lượng xe cá nhân?

……………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

TỰ ĐÁNH GIÁ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Câu | Nội dung | Điểm chuẩn | Tự chấm | Ghi chú |
| Chương 1  (30 điểm) | 1.1 | 6 |  |  |
| 1.2 | 6 |
| 1.3 | 10 |
| 1.4 | 6 |
| 1.5 | 2 |
| Chương 2  (40 điểm) | 2.1 Câu hỏi đặt ra | 5 |  |  |
| 2.2 Xây dựng câu hỏi khảo sát | 10 |  |  |
| 2.3 Phân tích |  |  |  |
| 2.3.1 Kết quả khảo sát | 10 |  |  |
| 2.3.2 Phân tích dữ liệu | 15 |  |  |
| Chương 3  (30 điểm) | 3.1 Thu thập dữ liệu | 5 |  |  |
| 3.2 Khai phá dữ liệu |  |  |  |
| 3.2.1. Xử lý dữ liệu | 2 |  |  |
| 3.2.2. Khám phá dữ liệu | 10 |  |  |
| 3.2.3. Trích xuất đặc trưng | 3 |  |  |
| 3.2.4. Suy luận | 10 |  |  |
| Báo cáo | (chú ý các chú ý 2,3,4,6 ở trang trước, nếu sai sẽ bị trừ điểm nặng) | 10đ |  |  |
| Làm việc nhóm | Chú ý trả lời đúng 4 yêu cầu trong phần làm việc nhóm | 10đ |  |  |
| Tổng điểm (120) | | |  |  |
| Đôi qua thang điểm 10 (120=10 điểm) | | |  |  |