**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 – 12345678**

Lớp **: XXXXX**

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 – 12345678**

Lớp **: XXXXX**

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

Trình bày tóm tắt vấn đề nghiên cứu, các hướng tiếp cận, cách giải quyết vấn đề và một số kết quả đạt được, những phát hiện cơ bản trong vòng 1 -2 trang của **từng Bài toán, mỗi bài toán là 1 đoạn văn bản khác nhau**.

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc387692905)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN i](#_Toc387692906)

[TÓM TẮT i](#_Toc387692907)

[MỤC LỤC 1](#_Toc387692908)

[DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ 1](#_Toc387692909)

[BÀI 1 – THU THẬP DỮ LIỆU 1](#_Toc387692910)

[1.1 Yêu cầu thu thập dữ liệu 1](#_Toc387692911)

[1.2 Dịch và đánh giá câu dịch 1](#_Toc387692912)

[1.3 Sinh thêm dữ liệu 1](#_Toc387692915)

[1.4 Đánh giá dữ liệu sinh thêm 1](#_Toc387692916)

[1.5 Luu bộ dữ liệu 1](#_Toc387692916)

[BÀI 2 – 1](#_Toc387692917)

2[.1 Tiểu mục 1 1](#_Toc387692918)

2[.2 Tiểu mục 2 1](#_Toc387692919)

[BÀI 3 – 1](#_Toc387692917)

3[.1 Tiểu mục 1 1](#_Toc387692918)

3[.2 Tiểu mục 2 1](#_Toc387692919)

LÀM VIỆC NHÓM 2

TỰ ĐÁNH GIÁ 6

DANH MỤC CÁC BẢNG

[Bảng 3.1 Ví dụ cho chèn bảng 1](#_Toc387689363)

[Bảng 3.2 Kết quả 1](#_Toc387689363)

CHƯƠNG 1

**THU THẬP DỮ LIỆU**

* 1. **Yêu cầu thu thập dữ liệu**

Chủ đề mà chúng em chọn là : Thể thao

|  |  |
| --- | --- |
| 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 câu hỏi và câu trả lời bằng tiếng Anh

* 1. **Dịch và đánh giá câu dịch**

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

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 |
| How many players are on a volleyball team? | Có bao nhiêu cầu thủ trong một đội bóng chuyền? | TÍN – ĐẠT – NHÃ | A volleyball team consists of six players on the court at any given time during a match. | Một đội bóng chuyền bao gồm sáu cầu thủ trên sân tại bất kỳ thời điểm nào trong một trận đấu | TÍN – ĐẠT – NHÃ |
| What muscle groups does shuttlecock kicking primarily work? | Đá cầu tác động chủ yếu vào những nhóm cơ nào? | TÍN – ĐẠT – NHÃ | Primarily the legs, core, and lower back, but also improves overall body coordination. | Chủ yếu là chân, lõi và lưng dưới, nhưng cũng cải thiện khả năng phối hợp tổng thể của cơ thể. | TÍN – ĐẠT – NHÃ |
| What is a hat-trick in soccer? | Hat-trick trong bóng đá là gì? | TÍN – ĐẠT – NHÃ | A hat-trick in soccer refers to a player scoring three goals in a single match. | Một hat-trick trong bóng đá đề cập đến một cầu thủ ghi ba bàn thắng trong một trận đấu. | TÍN – ĐẠT – NHÃ |
| How many players are there on a soccer team? | Có bao nhiêu cầu thủ trong một đội bóng đá? | TÍN – ĐẠT – NHÃ | A standard soccer team consists of 11 players on the field at a time, including one goalkeeper. | Một đội bóng đá tiêu chuẩn bao gồm 11 cầu thủ trên sân cùng một lúc, trong đó có một thủ môn. | TÍN – ĐẠT – NHÃ |
| …………… | …….. | ……. | ………… | …………….. | …. |

Bảng dịch câu hỏi và đánh giá câu dịch

**1.3 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

**1.3.1. 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?".

**1.3.2. 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?".

**1.3.3. 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.

**1.3.4. 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.

**1.3.5. 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.

**1.4 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 |  |  |
| Thu hẹp phạm vi |  |  |
| So sánh |  |  |
| Thay đổi chủ đề |  |  |
| Thay đổi góc nhìn |  |  |

**1.5 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

Trình bày chủ đề lựa chọn

Trình bày ít nhất 2 câu hỏi có liên quan đến chủ đề lựa chọn, lý giải tại sao lại đặt ra câu hỏi này, ý nghĩa của câu hỏi đặt ra.

2.2 Xây dựng câu hỏi khảo sát

Trình bày cách thức xây dựng ít nhất 10 câu hỏi câu hỏi khảo sát để trả lời cho 2 câu hỏi đặt ra ở 2.1 trong đó có 07 câu hỏi với câu trả lời là lựa chọn phương án trả lời từ 1 tới 5 và 03 câu hỏi với câu trả lời là text.

Liệt kê ít nhất 10 câu hỏi câu hỏi khảo sát và giải thích chi tiết các câu hỏi khảo sát này

2.3 Phân tích

2.3.1 Kết quả khảo sát

Trình bày kết quả khi tiến hành khảo sát ít nhất 10 câu hỏi khảo sát trên với ít nhất 50 bạn trong trường về các nội dung như: Cách thức tiến hành khảo sát, thời gian, số lượng, đối tượng, kết quả thu được,…

2.3.2 Phân tích dữ liệu

Trình bày cách tiền xử lý dữ lý dữ liệu và

Kết quả phân tích 10 câu hỏi khảo sát ở trên

Chú ý tập trung vào các vấn đề đề bài gợi ý:

Mỗi phân tích đều bao gồm các nội dung theo yêu cầu sau:

* Mô tả về dữ liệu
* Thực hiện các thống kê căn bản
* Tìm mối tương quan giữa các câu hỏi khảo sát và kết quả
* Xác định các yếu tố quan trọng ảnh hưởng đến kết quả
* Trực quan hóa dữ liệu và kết quả

2.4 Kết luận

Trình bày kết luận sau khi thực hiện bài tập này

Những điểm còn hạn chế và hướng phát triển trong tương lai

CHƯƠNG 3

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

3.1 Thu thập dữ liệu

Trình bày cách thu thập dữ liệu và kết quả dữ liệu thu thập được

3.2 Khai phá dữ liệu

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

Trình bày về yêu cầu, cách tiếp cận giải quyết, kết quả dạt được, đánh giá, thảo luận, v.v.phù hợp để bạn hoàn thành yêu cầu này một cách hiệu quả.

***Yêu cầu***:

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

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

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

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

Trình bày về yêu cầu, cách tiếp cận giải quyết, kết quả dạt được, đánh giá, thảo luận, v.v*.* phù hợp để bạn hoàn thành yêu cầu này một cách hiệu quả.

***Yêu cầu***: Bạn cần khám phá dữ liệu để hiển thị một số thông tin thống kê và phân tích của tập dữ liệu đã cho. Chẳng hạn như:

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

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

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

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

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

- Phân tích, trực quan mối quan hệ giữa các tác giả với nhau,…

Trên đây chỉ là một số gợi ý, bạn có thể đề xuất thêm các phân tích, thống kê khác.

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

Trình bày về yêu cầu, cách tiếp cận giải quyết, kết quả dạt được, đánh giá, thảo luận, v.v*.* phù hợp để bạn hoàn thành yêu cầu này một cách hiệu quả.

***Yêu cầu***: Hãy đề xuất cách trích xuất đặc trưng từ bộ dữ liệu đã cho, cung cấp lý do và giải thích cách làm của bạn.

**3.2.4. Suy luận**

Trình bày về yêu cầu, cách tiếp cận giải quyết, kết quả dạt được, đánh giá, thảo luận, v.v*.* phù hợp để bạn hoàn thành yêu cầu này một cách hiệu quả.

***Yêu cầu***: Bạn được yêu cầu phân loại câu nói theo tên người nổi tiếng và tính độ tương đồng phong cách nói giữa các tác giả theo 2 yêu cầu sau:

- 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?

- 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?

**LÀM VIỆC NHÓM**

Trình bày tóm tắt cách thức làm việc nhóm

Phân chia công việc của các thành viên trong nhóm

Tổng số lần gặp nhau (tính theo buổi)

Tổng thời gian gặp nhau (tính theo giờ)

**TÀI LIỆU THAM KHẢO**

**Chú ý các tài liệu tham khảo phải được chỉ mục vào trong các Chương nếu có tham khảo, phải có ít nhất từ 5-15 tài liệu tham khảo chính thống: Sách, bài báo Hội nghị, tạp chí - journal, … và tài liệu này phải được đánh chỉ mục từ nhỏ tới lớn và gắn vào trong báo cáo.**

**Các hình ảnh tham khảo hay các phần tham khảo ít thì có thể đánh thành footnote ngay trong trang tham khảo**

1. Quách Ngọc Ân (1992), “Nhìn lại hai năm phát triển lúa lai”, *Di tuyền học ứng dụng*, 98(1), tr. 10-16.
2. Bộ nông nghiệp & PTNT (1996), *Báo cáo tổng kết 5 năm (1992-1996) phát triển lúa lai,* Hà Nội.
3. Nguyễn Hữu Đống, Đào Thanh Bằng, Lâm Quang Dụ, Phan Đức Trực (1997), *Đột biến –* *Cơ sở lý luận và ứng dụng,* Nhà xuất bản nông nghiệp, Viện khoa học kỹ thuật nông nghiệp Việt Nam, Hà Nội.
4. Nguyễn Thị Gấm (1996), *Phát hiện và đánh giá một số dòng bất dục đực cảm ứng nhiệt* *độ,* Luận văn thạc sĩ khoa học nông nghiệp, Viện khoa học kỹ thuật nông nghiệp Việt Nam, Hà Nội.

……….

1. Anderson J.E. (1985), The Relative Inefficiency of Quota, The Cheese Case, *American* *Economic Review*, 75(1), pp. 178-90.
2. Borkakati R. P.,Virmani S. S. (1997), Genetics of thermosensitive genic male sterility in Rice, *Euphytica* 88, pp. 1-7.
3. Boulding K.E. (1955), *Economics Analysis*, Hamish Hamilton, London.
4. Burton G. W. (1988), “Cytoplasmic male-sterility in pearl millet (penni-setum glaucum L.)”, *Agronomic Journal* 50, pp. 230-231.
5. Central Statistical Oraganisation (1995), *Statistical Year Book*, Beijing.
6. FAO (1971), *Agricultural Commodity Projections (1970-1980)*, Vol. II. Rome.
7. Institute of Economics (1988), *Analysis of Expenditure Pattern of Urban Households in* *Vietnam,* Departement pf Economics, Economic Research Report, Hanoi.

**PHỤ LỤC**

Phần này bao gồm những nội dung cần thiết nhằm minh họa hoặc hỗ trợ cho nội dung đồ án như số liệu, biểu mẫu, tranh ảnh. . . . nếu sử dụng những câu trả lời cho một *bảng câu hỏi thì bảng câu hỏi mẫu này phải được đưa vào phần Phụ lục ở dạng nguyên bản* đã dùng để điều tra, thăm dò ý kiến; **không được tóm tắt hoặc sửa đổi**. Các tính toán mẫu trình bày tóm tắt trong các biểu mẫu cũng cần nêu trong Phụ lục của luận văn. Phụ lục không được dày hơn phần chính của đồ án

**MỘT SỐ CHÚ Ý KHI VIẾT BÁO CÁO**

1. Thống nhất kích cỡ chữ, kiểu chữ trong toàn bộ báo cáo. Không tô màu chữ, chỉ dùng màu đen
2. Các công thức phải tự gõ và đánh số theo Chương, ví dụ 1.1, 2.1, 2.2, 2.3
3. Các hình và Bảng phải đánh số theo chương, ví dụ Hình 1.1, Hình 2.1, Bảng 3.1, Bảng 3.2
4. Các hình nếu lấy ở ngoài phải đề footnote chú thích nguồn ở dưới
5. Hình mô hình tổng quát phải tự vẽ bằng Word, không dán hình
6. Các tài liệu tham khảo phải đính vào luận văn theo thứ tự từ nhỏ tới lớn, bắt đầu từ 1, ít nhất phải từ 5-15 tài liệu tham khảo, lựa chọn các tài liệu tham khảo mới
7. Tóm tắt trình bày được các nội dung sau: giới thiệu, phương pháp làm, kết quả, nhận xét (không dùng hình, bảng ở mục này)

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) | | |  |  |