# TRƯỜNG ĐẠI HỌC THỦ DẦU MỘT

# Giới thiệu - Đại học Thủ Dầu MộtKHOA KỸ THUẬT CÔNG NGHỆ

# BÀI TẬP LỚN MÔN PHÂN TÍCH THỐNG KÊ

# Đề tài

# PHÂN TÍCH ĐIỂM CƠ SỞ NGÀNH CÔNG NGHỆ THÔNG TIN

# SVTH: HOÀNG KIM TUYẾN

# Mã SV: 1824801040043

# Lớp: D18HT01

# Tháng 07/2020

# LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi và được sự hướng dẫn khoa học 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 báo cáo 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 báo cáo của mình.** Trường Đại học Thủ Dầu Một không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*Bình Dương, ngày 23 tháng 07 năm 2020*

*Người thực hiện  
 (ký tên và ghi rõ họ tên)*

(Mẫu)

**MỤC LỤC**

( Bold, size 13)

Trang

**Phần 1**

1.1 Bài 1

1.2 Bài 2

1.3 Bài 3

**Phần 2**

2.1. Mục đích đề tài

2.2. Câu hỏi nghiên cứu

2.3. Thu thập dữ liệu

2.4. Phân tích dữ liệu

2.5. Dự đoán

**Phần 3**

**DANH MỤC TÀI LIỆU THAM KHẢO** 20

***Ghi chú:***

* In đậm và in hoa tiêu đề của các chương, mục lớn
* Chữ số thứ nhất chỉ thứ tự chương
* Chử số thứ 2 chỉ thứ tự mục trong chương
* Chữ số thứ 3,…,chỉ thứ tự các tiểu mục

(Mẫu)

**DANH MỤC CÁC BẢNG, SƠ ĐỒ, HÌNH (Nếu có)**

(In đậm, in hoa, size 13)

BÁNG 1.1 (size 13)...................................................................................................

BẢNG 1.2.................................................................................................................

……..........................................................................................................................

……..........................................................................................................................

……..........................................................................................................................

SƠ ĐỒ 1.1

……..........................................................................................................................

……..........................................................................................................................

……..........................................................................................................................

HÌNH 1.1

…….........................................................................................................................

…….........................................................................................................................

***Ghi chú:***

**-** Xếp sau trang Mục lục

* Chữ số thứ nhất chỉ tên chương
* Chữ số thứ hai chỉ thứ tự bảng biểu, sơ đồ, hình,…trong mỗi chương
* Ở cuối mỗi bảng biểu, sơ đồ, hình,…trong mỗi chương phải có ghi chú, giải thích, nêu rõ nguồn trích hoặc sao chụp,…

(Mẫu)

**Trình bày mỗi trang của Bài tập lớn môn Phân tích thống kê**

Bắt đầu đánh số trang từ Phần 1, in đậm tiêu đề các chương, các mục

**( đánh số trang trên header, tên đề cương và Sinh viên thực hiện trên Footer) size 10**

**Giãn cách dòng 1.5, size chữ 13, hết mỗi phần sang trang mới**

**PHẦN 1**

1.1 Bài 1

1.2 Bài 2

1.3 Bài 3

**PHẦN 2**

2.1. **Mục đích đề tài**

Đối với các sinh viên, khi bắt đầu theo học đại học, tất cả sinh viên đều phải học qua các môn cơ sở ngành của mình. Vì thế, trong đề tài này, em quyết định thu thập điểm các môn cơ sở ngành của các sinh viên CNTT – Trường Đại Học Thủ Dầu Một.  
Đề tài này trả lời các câu hỏi sau đây: các sinh viên ngành CNTT có học tốt các môn cơ sở ngành không, có nên có thái độ nghiêm túc trong việc học các môn cơ sở ngành hay không, , mức độ quan trọng của các môn cơ sở ngành đối với sinh viên CNTT sau khi học các môn chuyên ngành, tỉ lệ sinh viên CNTT học tốt các môn cơ sở ngành

2.2. Câu hỏi nghiên cứu

* Các sinh viên ngành CNTT có học tốt các môn cơ sở ngành không?
* Có nên có thái độ nghiêm túc trong việc học các môn cơ sở ngành hay không?
* Có nên có thái độ nghiêm túc trong việc học các môn cơ sở ngành hay không?
* Mức độ quan trọng của các môn cơ sở ngành đối với sinh viên CNTT sau khi học các môn chuyên ngành?
* Tỉ lệ sinh viên CNTT học tốt các môn cơ sở ngành?

2.3. Thu thập dữ liệu

Trong thời gian nhận được yêu cầu đề tài, em đã thu thập dữ liệu bằng cách tạo google Form, sau đó gửi cho các bạn sinh viên trên Facebook, nhờ các bạn trong lớp tiến hành khảo sát. Trong vòng 2 tuần em thu thập được hơn 150 dòng dữ liệu

Bộ dữ liệu bao gồm

2.4. Phân tích dữ liệu

Phân tích giá trị Lớn nhất, Nhỏ nhất, Trung bình, Phần tử xuất hiện nhiều nhất trong mỗi cột, viết code để sinh ra kết quả tự động và lưu kết quả vào file excel.

Trình bày công thức và ý nghĩa của 2 độ đo Pearson và Spearman, Phân tích mối tương quan giữa các dữ liệu theo 2 độ đo tự viết code: Pearson (ra đồ thị Heatmap) và Spearman.

2.5. Dự đoán

Trình bày tổng quan về giải thuật dự đoán Bayes

Dự đoán bằng phương pháp Bayes tự viết và đánh giá bằng 3 độ đo (Accuracy, Confusion Matrix score và F1 score)

**PHẦN 3: TỰ CHẤM**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nội dung | Ý | Thang điểm | Điểm chấm | Ghi chú |
| Phần 1 | Bài 1 | 1 điểm |  |  |
| Bài 2 | 1 điểm |  |  |
| Bài 3 | 1 điểm |  |  |
| Phần 2 | Dữ liệu | 1 điểm |  |  |
|  | Phân tích | 2 điểm |  |  |
|  | Độ tương quan | 2 điểm |  |  |
|  | Dự báo | 2 điểm |  |  |
| **Tổng** | | 10 |  |  |

**DANH MỤC TÀI LIỆU THAM KHẢO**