TRƯỜNG ĐẠI HỌC KINH TẾ

**KHOA THỐNG KÊ – TIN HỌC**

–––––––––––––––––––––––––––––––

**BÁO CÁO THỰC TẬP NGHỀ NGHIỆP**

**TÊN ĐỀ TÀI**

**PHÂN TÍCH DỮ LIỆU VỀ BỆNH NHÂN COVID19 TẠI VIỆT NAM**

**Sinh viên thực hiện:** Nguyễn Thị Thúy

**Giảng viên hướng dẫn:** Cao Thị Nhâm

**Đà Nẵng, 6/2021**

**LỜI CẢM ƠN**

**LỜI CAM ĐOAN**

# MỤC LỤC

[MỤC LỤC ii](#_Toc76780098)

[LỜI MỞ ĐẦU 1](#_Toc76780099)

[CHƯƠNG 1. TỔNG QUAN VỀ CÔNG NGHỆ VÀ CÔNG CỤ SỬ DỤNG 2](#_Toc76780100)

[1.1. Tổng quan về công nghệ 2](#_Toc76780101)

[1.1.1. Thư viện Pandas 2](#_Toc76780102)

[1.1.2. Mục 1.1.2 2](#_Toc76780103)

[1.2. Tổng quan về Windows Forms 2](#_Toc76780104)

[1.3. Tổng quan về hệ quản trị cơ sở dữ liệu Microsoft SQL Server 3](#_Toc76780105)

[CHƯƠNG 2. phân tích và thiết kế ứng dụng 4](#_Toc76780106)

[2.1. Mô tả bài toán 4](#_Toc76780107)

[2.2. Thiết kế các chức năng 4](#_Toc76780108)

[2.3. Thiết kế cơ sở dữ liệu 4](#_Toc76780109)

[CHƯƠNG 3. PHÁT TRIỂN ỨNG DỤNG 5](#_Toc76780110)

[3.1. Sơ đồ chức năng 5](#_Toc76780111)

[3.2. Thông tin chi tiết các form làm việc 5](#_Toc76780112)

[3.2.1. Form X 5](#_Toc76780113)

[3.2.2. Form Y 5](#_Toc76780114)

[CHƯƠNG 4. KẾT QUẢ 6](#_Toc76780115)

[KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 7](#_Toc76780116)

[TÀI LIỆU THAM KHẢO 8](#_Toc76780117)

[PHỤ LỤC (nếu có) 9](#_Toc76780118)

DANH MỤC HÌNH ẢNH

[Hình 1.1 Ngôn ngữ lập trình Python 2](#_Toc74235469)

[Hình 4.1 Kiến trúc của mô hình RNN 5](#_Toc74235470)

DANH MỤC BẢNG BIỂU

[Bảng 4.1 Kiến trúc 5](#_Toc74235471)

DANH MỤC CÁC TỪ VIẾT TẮT

# LỜI MỞ ĐẦU

1. **Mục tiêu của đề tài**
2. **Nhiệm vụ của đề tài**

* Nghiên cứu ngôn ngữ lập trình Python trong phân tích dữ liệu
* Nghiên cứu thư viện Pandas trong Python dùng để trực quan hóa dữ liệu
* Phân tích và xử lý dữ liệu, thực hiện trực quan hóa dữ liệu về bệnh nhân Covid19 tại Việt Nam.

1. **Kết cấu của đề tài**

Đề tài được tổ chức gồm 3 chương nội dung và phần kết luận.

* **Chương 1**: Tổng quan về công nghệ và công cụ sử dụng
* **Chương 2**: Phân tích dữ liệu
* **Chương 3:** Kết quả
* Kết luận

# TỔNG QUAN VỀ CÔNG NGHỆ VÀ CÔNG CỤ SỬ DỤNG

## Tổng quan về công nghệ

### Thư viện Pandas

* Pandas là một thư viện mã nguồn mở được phát triển bởi Wes McKinney vào năm 2008. Pandas được sử dụng chủ yếu để thao tác, phân tích và dọn dẹp dữ liệu. Pandas cung cấp rất nhiều cấu trúc dữ liệu cũng như các phép tính hỗ trợ thao tác dữ liệu số và dữ liệu thời gian(time series). Pandas nhanh, mạnh và hiệu quả.
* Có thể xử lý tập dữ liệu khác nhau về định dạng: chuỗi thời gian, bảng không đồng nhất, ma trận dữ liệu
* Khả năng import dữ liệu từ nhiều nguồn khác nhau như CSV, DB/SQL, Excel,…
* Xử lý dữ liệu mất mát theo ý người dùng mong muốn: bỏ qua hoặc chuyển sang 0
* Xử lý, phân tích dữ liệu tốt như mô hình hoá và thống kê
* Tích hợp tốt với các thư viện khác của python

1.1.2

## Công cụ sử dụng

Sử dụng công cụ google colab(Colaboratory) dùng để xử lý và phân tích dữ liệu.



## Tổng quan về hệ quản trị cơ sở dữ liệu Microsoft SQL Server

Trình bày nội dung ở đây

# phân tích và thiết kế ứng dụng

## Mô tả bài toán

Diễn đạt lại nội dung mô tả bài toán ở đây

## Thiết kế các chức năng

Trình bày các chức năng của ứng dụng ở đây

## Thiết kế cơ sở dữ liệu

Trình bày các bảng cơ sở dữ liệu ở đây

# PHÁT TRIỂN ỨNG DỤNG

## Sơ đồ chức năng

Trong này vẽ sơ đồ các form có trong ứng dụng

## Thông tin chi tiết các form làm việc

### Form X

Trong này giới thiệu chức năng, các mục có trong màn hình

### Form Y

Trong này giới thiệu chức năng, các mục có trong màn hình

# KẾT QUẢ

Phần này trình bày kết quả đạt được

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Đề tài đã thực hiện được …

# TÀI LIỆU THAM KHẢO

1. Phan, D.-V., et al., *Liver cancer prediction in a viral hepatitis cohort: A deep learning approach.* International Journal of Cancer, 2020. **147**(10): p. 2871-2878.

2. Phan, D.V., et al., *Deep learning approaches for sleep disorder prediction in an asthma cohort.* J Asthma, 2020: p. 1-9.

Tên tác giả, Tên tác giả… *Tên công trình*. Nơi xuất bản, năm, trang

# PHỤ LỤC (nếu có)