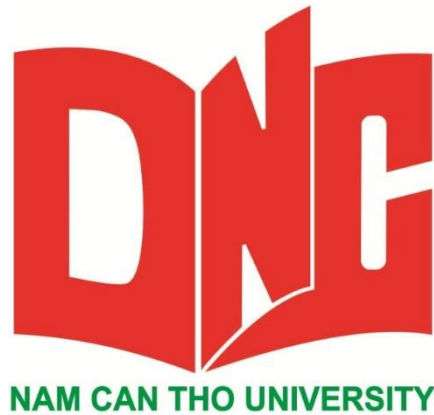


TRƯỜNG ĐẠI HỌC NAM CẦN THƠ
KHOA KỸ THUẬT - CÔNG NGHỆ



LÊ HOÀNG DUY

MSSV: 177799

CÀI ĐẶT GIẢI THUẬT K - NEAREST NEIGHBORS
NHẬN DẠNG CHỮ SỐ VIẾT TAY

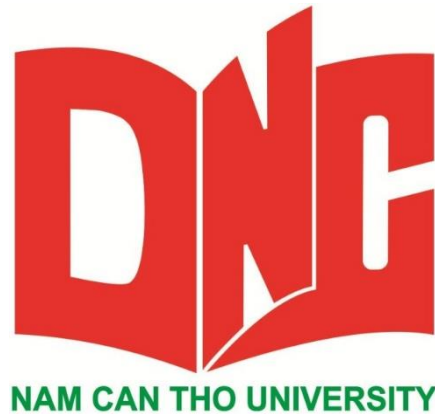
ĐỒ ÁN THỰC TẬP

Ngành Công nghệ Thông tin

Mã số Ngành: 7480201

05 – 2021

**TRƯỜNG ĐẠI HỌC NAM CẦN THƠ
KHOA KỸ THUẬT – CÔNG NGHỆ**



**LÊ HOÀNG DUY
MSSV: 177799**

**CÀI ĐẶT GIẢI THUẬT K-NEAREST NEIGHBORS
NHẠP DẠNG CHỮ SỐ VIẾT TAY**

**ĐỒ ÁN THỰC TẬP
Ngành Công nghệ Thông tin
Mã số Ngành: 7480201**

**GIẢNG VIÊN HƯỚNG DẪN
Th.S PHAN THỊ XUÂN TRANG**

05 - 2021

CHẤP THUẬN CỦA HỘI ĐỒNG

Đồ án thực tập cuối khóa, đề tài “Cài đặt giải thuật K – Nearest Neighborss nhận dạng chữ số viết tay” do sinh viên Lê Hoàng Duy thực hiện dưới sự hướng dẫn của Th.S Phan Thị Xuân Trang. Đồ án thực tập đã báo cáo và được hội đồng chấm đồ án thông qua ngày ... tháng ... năm ...

Ủy viên

Thư ký

.....
(GHI CHÚC DANH, HỌ, TÊN)

.....
(GHI CHÚC DANH, HỌ, TÊN)

Phản biện 1

Phản biện 2

.....
(GHI CHÚC DANH, HỌ, TÊN)

.....
G(HI CHÚC DANH, HỌ, TÊN)

Cán bộ hướng dẫn

Chủ tịch hội đồng

.....
(GHI CHÚC DANH, HỌ, TÊN)

.....
(GHI CHÚC DANH, HỌ, TÊN)

LỜI CẢM ƠN

Lời đầu tiên, cho phép em xin chân thành gửi lời cảm ơn đến tất cả các thầy, cô đã nhiệt tình giúp đỡ em từ khi bắt đầu thực tập cuối khóa đến khi kết thúc để kịp hoàn thành đúng với thời gian đã quy định.

Em xin cảm ơn quý thầy cô Khoa Kỹ thuật - Công nghệ Trường Đại học Nam Cần Thơ đã tận tâm giảng dạy và truyền đạt cho em những kiến thức quý báu để em hoàn thiện chương trình học tại trường. Đồng thời, em cũng rất cảm ơn Chủ tịch Hội đồng quản trị, Ban Giám hiệu Trường Đại học Nam Cần Thơ đã tạo điều kiện về thời gian, không gian để em có thể hoàn thành nhiệm vụ học tập của mình.

Trong thời gian qua thực tập vừa qua, em đã có rất nhiều cố gắng nỗ lực để hoàn thành đồ án nhưng vẫn không tránh khỏi những thiếu sót. Kính mong nhận được sự đóng góp ý kiến của quý thầy cô. Đặc biệt, em xin gửi lời cảm ơn chân thành nhất đến giảng viên – Th.S. Phan Thị Xuân Trang. Trong khoảng thời gian thực hiện thực tập cuối khóa cảm ơn thầy đã quan tâm, hướng dẫn tận tình, đầy đủ chi tiết để em có đủ kiến thức và áp dụng để hoàn thành tốt lần thực tập cuối khóa này để không phụ tấm lòng của thầy đã giảng dạy qua các năm học tại trường..

Em xin chân thành cảm ơn!

Cần Thơ, ngày ... tháng ... năm 2021

Người thực hiện

LÊ HOÀNG DUY

LỜI CAM ĐOAN

Tôi xin cam kết báo cáo thực tập này được hoàn thành dựa trên các kết quả nghiên cứu của tôi và các kết quả nghiên cứu này chưa được dùng cho bất cứ báo cáo thực tập cùng cấp nào khác.

Cần Thơ, ngày ... tháng ... năm 2021

Người thực hiện

LÊ HOÀNG DUY

MỤC LỤC

CHƯƠNG 1: GIỚI THIỆU NỘI THỰC TẬP.....	5
1.1 Đặt vấn đề nghiên cứu.....	5
1.2 Giới thiệu về công ty.....	6
1.2.1 Thông tin về công ty.....	6
1.2.2 Trụ sở chính.....	6
1.2.3 Người đại diện theo pháp luật:	6
1.2.4 Thông tin liên hệ.....	6
1.2.5 Ngành nghề kinh doanh.....	7
1.2.6 Sản phẩm và dịch vụ đang kinh doanh.....	8
1.3 Mục tiêu nghiên cứu.....	11
1.4 Giới thiệu về bộ cơ sở dữ liệu	11
CHƯƠNG 2: CƠ SỞ LẬP LUẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU	13
2.1 Cơ sở lý luận	13
2.2 Phương pháp nghiên cứu.....	13
2.2.1 Phương pháp nghiên cứu lý thuyết.....	13
2.2.2 Phương pháp nghiên cứu thực nghiệm	13
2.2.3 Phương pháp điều tra	13
2.3 Giới thiệu về giải thuật K – Nearest Neighbors.....	14
2.3.1 Thuật toán KNN.....	14
2.3.2 Tiếp cận học xếp hạng K-NN cho tìm kiếm chuyên gia.....	15
2.3.3 Mô hình học xếp hạng K-NN	15
2.3.4 Khoảng cách trong không gian của vector	15
2.3.5 Một số norm thường dùng.....	16
2.3.6 Kiểm thử độ tin cậy của giải thuật K – Nearest Neighbors...17	
2.4 Giới thiệu về ngôn ngữ Python.....	18
CHƯƠNG 3: GIẢI THUẬT K – NEAREST NEIGHBORS TRONG	
NHẬN DẠNG CHỮ SỐ VIẾT TAY	21
3.1 Phương pháp nhận dạng K – Nearest Neighbors.....	21
3.2 Giới thiệu chung	21
3.3 Quá trình nhận dạng chữ số viết tay	23
3.3.1 Đưa ảnh vào	23
3.3.2 Tiền xử lý.....	23
3.3.3 Sử dụng KNN để nhận dạng.....	23
CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU.....	26
4.1 Kết quả nghiên cứu bộ models	26
4.2 Kết quả nghiên cứu tỉ lệ đúng của một model	26
4.3 Giao diện.....	27
4.4 Hướng dẫn sử dụng	29
4.5 Hướng dẫn cài đặt	36
CHƯƠNG 5: KẾT LUẬN.....	37

DANH SÁCH BẢNG

Bảng 1.1 Ngành nghề kinh doanh	7
Bảng 1.2: Sản phẩm phần mềm	8
Bảng 1.3: Bảng dịch vụ	10
Bảng 4.1: bảng Parameter	35

DANH SÁCH HÌNH

Hình 1.1 Giới thiệu bộ dữ liệu mnist	11
Hình 1.2 Giới thiệu ảnh pixel trong dataset_minist	12
Hình 2.1 Ví dụ minh họa về thuật toán KNN	14
Hình 2.2 Mô hình học xếp hạng C4.4-KNN cho tìm kiếm chuyên gia	16
Hình 2.3 Norm 1 và Norm 2 trong không gian 2 chiều	17
Hình 2.4 Biểu đồ thể hiện độ chính xác của KNN	18
Hình 3.1: Sơ đồ use case của chương trình	25
Hình 4.1 Kết quả nghiên cứu cục bộ models.....	26
Hình 4.2 Kết quả nghiên cứu tỉ lệ đúng với model.....	27
Hình 4.3 Giao diện Setup train file.csv, xuất models và test kết quả	28
Hình 4.4 Giao diện ứng dụng nhận dạng chữ số viết tay.....	28
Hình 4.5: Sơ đồ Use case hướng dẫn sử dụng.....	29
Hình 4.6 Chọn file_train.csv	30
Hình 4.7 Chọn đường dẫn lưu List models	30
Hình 4.8 Nhập số lượng models cần xuất ra	31
Hình 4.9 Nhập giá trị N_Neighbors	31
Hình 4.10 Chọn giá trị N_Jobs.....	32
Hình 4.11 Chọn file_test.csv	32
Hình 4.12 Chọn List models đã train.....	33
Hình 4.13 Test và hiện kết quả.....	33
Hình 4.14 Chọn model từ List Models.....	34
Hình 4.15 Chọn model tiến hành nhận dạng chữ số viết tay.....	34

DANH MỤC TỪ VIẾT TẮT

KNN	K-Nearest Neighborss
CNTT	Công nghệ thông tin
MMH	Maximum Marginal Hyperplane
KKT	Karush Kuhn Tucker
OCR	Optical Character Recognition
MNIST	Modified National Institute of Standards and Technology