**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC NÔNG LÂM TP HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

****

TIỂU LUẬN TỐT NGHIỆP

**NGHIÊN CỨU MÔ HÌNH HỌC MÁY NHẬN DẠNG LOẠI BỆNH TRÊN CÂY LÚA**

**Ngành : Công nghệ thông tin**

**Niên khoá : 2021-2025**

**Lớp : DH21DTC**

**Sinh viên thực hiện : Kiều Thị Ngọc Qúy**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC NÔNG LÂM TP HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

****

TIỂU LUẬN TỐT NGHIỆP

**NGHIÊN CỨU MÔ HÌNH HỌC MÁY NHẬN DẠNG LOẠI BỆNH TRÊN CÂY LÚA**

**CÁN BỘ HƯỚNG DẪN SINH VIÊN THỰC HIỆN**

TS. Nguyễn Thị Phương Trâm Kiều Thị Ngọc Qúy (MSSV: 21130504)

TP. HỒ CHÍ MINH, tháng 4 năm 2024

# DANH SÁCH CHỮ VIẾT TẮT

# DANH MỤC HÌNH ẢNH

# DANH MỤC BẢNG

# TÓM TẮT

**MỤC LỤC**

# MỞ ĐẦU

## LÝ DO CHỌN ĐỀ TÀI

## MỤC TIÊU VÀ PHẠM VI NGHIÊN CỨU

## Ý NGHĨA KHOA HỌC VÀ THỰC TIỄN CỦA ĐỀ TÀI

# TỔNG QUAN ĐỀ TÀI

## Các công trình nghiên cứu liên quan

## Những vấn đề còn tồn tại

## Nội dung nghiên cứu

# PHƯƠNG PHÁP VÀ NỘI DUNG NGHIÊN CỨU

## Cơ sở lý thuyết

## Mô hình ..

## Mô hình lý thuyết của giải pháp đã đề xuất sử dụng trong tiểu luận

## Đánh giá mô hình

Mô hình được đánh giá dựa trên các chỉ số được tính từ confusion matrix (ma trận nhầm lẫn).

Confusion matrix là một bảng được dùng để đánh giá hiệu suất của mô hình phân loại nhị phân hoặc đa lớp.



Hình 1: Confusion Matrix

Với:

* **TP**: số lượng mẫu dương được dự đoán là dương.
* **TN**: số lượng mẫu âm được dự đoán là âm.
* **FP**: số lượng mẫu âm được dự đoán là dương.
* **FN**: số lượng mẫu dương được dự đoán là âm.

Mô hình được đánh giá dựa trên:

* **Accuracy**: đo tỷ lệ dự đoán đúng trên toàn bộ mẫu dữ liệu.
* **Precision**: đo tỷ lệ dự đoán dương đúng trên tổng số dự đoán dương.
* **Recall**: đo tỷ lệ dự đoán dương đúng trên tổng số thực tế dương.
* **F1-Score**: còn gọi là thước đo F, cho biết sự cân bằng giữa precision và recall.
* **Đường cong ROC**: là biểu đồ được vẽ bằng cách tính tỷ lệ dương thực tế (recall) và tỷ lệ dương sai ().
* **AUC**: là diện tích dưới đường cong ROC. AUC đo lường khả năng phân biệt giữa các lớp của mô hình. AUC nằm trong khoảng từ 0 đến 1, mô hình có thể được xem là tốt khi AUC từ 0.8 trở lên.

Vẽ biểu đồ huấn luyện: Sử dụng Matplotlib để vẽ biểu đồ theo dõi độ chính xác (accuracy) và độ mất mát (loss) trên tập train và tập validation. Biểu đồ trực quan hóa quá trình huấn luyện của mô hình, giúp phát hiện các vấn đề trong quá trình huấn luyện như overfitting.

# GIẢI PHÁP CHO BÀI TOÁN

## Phát biểu bài toán

## Giải pháp cụ thể để nhận dạng bệnh

### Giải pháp tiền xử lý dữ liệu

### Xây dựng mô hình

### Huấn luyện và tối ưu mô hình

## Hiện thực giải pháp

### Môi trường triển khai thực nghiệm

Ngôn ngữ lập trình: Python.

Môi trường chạy: Google Colab với RAM, tích hợp với Google Drive để lưu trữ dữ liệu và kết quả mô hình.

Thông tin tập dữ liệu

Bảng 3.1: Thông tin tập dữ liệu

|  |  |
| --- | --- |
| **Thông tin** | **Mô tả** |
| Tên dataset |  |
| Số lượng ảnh |  |
| Định dạng ảnh |  |
| Số lớp |  |
| Số lượng ảnh trong tập huấn luyện |  |
| Số lượng ảnh trong tập xác thực |  |
| Số lượng ảnh trong tập kiểm tra |  |
| Tên các lớp: số lượng ảnh |  |

### Kết quá thực nghiệm [[1]](#footnote-1)

### Phân tích, đánh giá và nhận xét kết quả

Dựa vào chỉ số Accuracy, Precision, Recall, F1-Score của từng mô hình rút ra được kết luận sau:

# KẾT QUẢ, KẾT LUẬN VÀ KIẾN NGHỊ

## Kết quả

## Kết luận

## Kiến nghị

# TÀI LIỆU THAM KHẢO

# PHỤ LỤC

1. GitHub Repository: Notebook Colab cho tiểu luận. Truy cập tại: https://github.com/Kiwi-0102/tieuLuan [↑](#footnote-ref-1)