

# XÂY DỰNG MÔ HÌNH PHÂN LOẠI BỆNH DỰA TRÊN ẢNH LÁ LÚA

GVHD : TH.S PHẠM TRỌNG HUYNH

SINH VIÊN :

Nguyễn Hoàng Nam - 1150080149

Trương Văn Quốc Phong - 1150080153

LỚP : 11DHCNPM2



11DHCNPM2

# GIỚI THIỆU ĐỀ TÀI

Tên dự án	Xây dựng mô hình phân loại bệnh lá lúa bằng học sâu (Deep Learning)
Mục tiêu	Phát triển mô hình AI có khả năng tự động phân loại các bệnh phổ biến trên lá lúa dựa trên hình ảnh, hỗ trợ nông dân phát hiện sớm và giảm thiểu thiệt hại mùa vụ.
Phạm vi	Thu thập, gán nhãn và huấn luyện dữ liệu hình ảnh lá lúa; phát triển mô hình CNN/MobileNetV2; xây dựng ứng dụng web demo.
Kết quả mong đợi	Đạt độ chính xác ≥ 90%; tạo bộ dữ liệu hình ảnh có nhãn; có báo cáo, mô hình và demo chạy thực tế.
Thành viên nhóm	Nguyễn Hoàng Nam - Trương Văn Quốc Phong
Công nghệ sử dụng	Python, PyTorch, TensorFlow, OpenCV, ReactJS, Node.js, MongoDB.
Rủi ro chính	Thiếu dữ liệu chất lượng; sai lệch nhãn ảnh; hiệu năng mô hình thấp; thiếu tài nguyên GPU khi huấn luyện.
Thời gian dự kiến	Từ 01/09/2025 đến 25/11 /2026.



# HỌP TRIỂN KHAI ĐỀ TÀI

Tên dự án	Xây dựng mô hình phân loại bệnh lá lúa
Mục tiêu buổi họp	Khởi động dự án, phân công vai trò, thống nhất tiến độ, công cụ và phạm vi công việc.
Thành viên tham gia	Tất cả thành viên nhóm, giảng viên hướng dẫn.
Nội dung chính	Giới thiệu dự án, trình bày mục tiêu, timeline, phân công công việc, thống nhất phương pháp liên lạc và báo cáo tiến độ.
Ngày tổ chức	3/9/2025

# HỢP ĐỒNG NHÓM

Họ và Tên	Ký tên xác nhận
Nguyễn Hoàng Nam	Nguyễn Hoàng Nam
Trương Văn Quốc Phong	Trương Văn Quốc Phong

**Quy tắc ứng xử (Code of Conduct):**  
Là một nhóm dự án, chúng em sẽ:  
Làm việc chủ động, dự đoán các vấn đề tiềm ẩn và tìm cách phòng tránh.  
Giữ cho các thành viên luôn được cập nhật thông tin liên quan đến dự án.  
Tập trung vào những gì tốt nhất cho toàn bộ nhóm và kết quả dự án.

**Tham gia (Participation):**  
Trung thực và cởi mở trong mọi hoạt động của dự án.  
Khuyến khích sự đa dạng và tinh thần hợp tác trong nhóm.  
Tạo cơ hội cho mọi thành viên được tham gia bình đẳng.  
Sẵn sàng học hỏi và tiếp nhận ý tưởng mới.  
Thông báo sớm cho nhóm nếu có thành viên vắng mặt hoặc trễ tiến độ.

**Liên lạc (Communication):**  
Thống nhất phương thức liên lạc qua Zalo, Google Meet hoặc email.  
Gặp mặt trực tuyến ít nhất 2 lần mỗi tuần để trao đổi tiến độ. • Ghi biên bản họp và lưu trữ trên Google Drive.  
Báo cáo tiến độ vào cuối mỗi tuần.

**Giải quyết vấn đề (Problem Solving):**• Tất cả thành viên cùng tham gia thảo luận khi có vấn đề.  
Đưa ra ý kiến mang tính xây dựng, không đổ lỗi.  
Xây dựng giải pháp chung và thống nhất trước khi thực hiện.

**Nguyên tắc họp (Meeting Guidelines):**  
Hẹp định kỳ vào mỗi Thứ 3 và Thứ 6 hàng tuần.  
Ghi biên bản, nêu rõ quyết định và việc cần làm.  
Gửi biên bản trong vòng 24 giờ sau buổi họp.

**Nội dung công việc chính:**

1. Tìm hiểu về định nghĩa bài toán (mô tả trong ngữ cảnh nhất định, các bộ dữ liệu, phương pháp, độ đo, kết quả hiện tại).  
Nếu không có dữ liệu sẵn thì xây dựng bộ dữ liệu mới.
2. Chọn thuật toán phù hợp cho bài toán.
3. Tìm hiểu tập dữ liệu và thu thập thêm dữ liệu phù hợp cho đa dạng bài toán (nếu có).
4. Gán nhãn cho bộ dữ liệu ảnh thu thập phù hợp với thuật toán đã chọn (nếu có).
5. Xây dựng thực nghiệm để huấn luyện (training) mô hình, kết hợp xây dựng thuật toán và đánh giá kết quả đạt được.

# PHÂN RÃ CÔNG VIỆC

## 1. DẠNG PHÁC THẢO

### 1. DỰ ÁN PHÂN LOẠI BỆNH LÁ LÚA

#### 1.1 KHỞI ĐỘNG (GIAI ĐOẠN KHỞI ĐỘNG)

##### 1.1.1 XÁC ĐỊNH VẤN ĐỀ VÀ MỤC TIÊU

##### 1.1.2 LẬP PROJECT CHARTER

##### 1.1.3 PHÊ DUYỆT DỰ ÁN

#### 1.2 LẬP KẾ HOẠCH (GIAI ĐOẠN LẬP KẾ HOẠCH)

##### 1.2.1 XÁC ĐỊNH PHẠM VI VÀ YÊU CẦU DỮ LIỆU

##### 1.2.2 PHÂN CÔNG NHÓM VÀ VAI TRÒ

##### 1.2.3 HỌP KHỞI ĐỘNG

##### 1.2.4 LÊN KẾ HOẠCH HUẤN LUYỆN VÀ THU THẬP DỮ LIỆU

#### 1.3 THỰC THI (GIAI ĐOẠN THỰC THI)

##### 1.3.1 THU THẬP VÀ GÁN NHÃN DỮ LIỆU

##### 1.3.2 HUẤN LUYỆN MÔ HÌNH CNN/MOBILENETV2

##### 1.3.3 ĐÁNH GIÁ KẾT QUẢ VÀ TINH CHỈNH

##### 1.3.4 XÂY DỰNG ỨNG DỤNG DEMO

#### 1.4 KIỂM SOÁT (GIAI ĐOẠN KIỂM SOÁT)

##### 1.4.1 THEO DÕI TIẾN ĐỘ VÀ CHẤT LƯỢNG MÔ HÌNH

##### 1.4.2 QUẢN LÝ RỦI RO VÀ CẬP NHẬT KẾ HOẠCH

#### 1.5 KẾT THÚC (GIAI ĐOẠN KẾT THÚC)

##### 1.5.1 VIẾT BÁO CÁO TỔNG KẾT

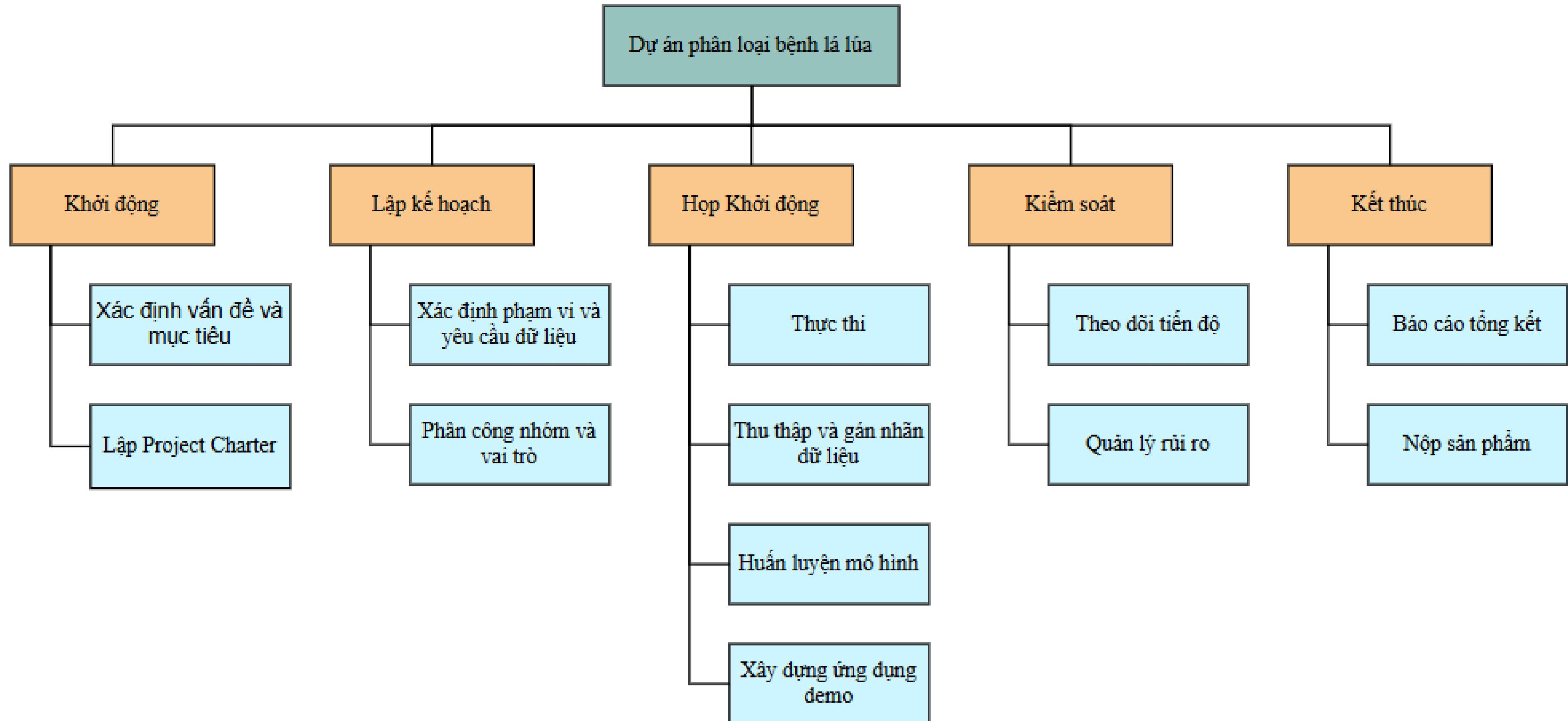
##### 1.5.2 NỘP SẢN PHẨM VÀ TÀI LIỆU

##### 1.5.3 HỌP TỔNG KẾT VÀ RÚT KINH NGHIỆM

# BẢNG PHÂN RÃ CÔNG VIỆC

Cấp độ	Mã WBS	Tên hạng mục
1		1 Dự án phân loại bệnh lá lúa
2		11 Khởi động
3		111 Xác định vấn đề và mục tiêu
3		112 Lập Project Charter
2		12 Lập kế hoạch
3		121 Xác định phạm vi và yêu cầu dữ liệu
3		122 Phân công nhóm và vai trò
3		123 Họp Khởi động
2		13 Thực thi
3		131 Thu thập và gán nhãn dữ liệu
3		132 Huấn luyện mô hình
3		133 Xây dựng ứng dụng demo
2		14 Kiểm soát
3		141 Theo dõi tiến độ
3		142 Quản lý rủi ro
2		15 Kết thúc
3		151 Báo cáo tổng kết
3		152 Nộp sản phẩm

# TREE STRUCTURE VIEW



# MICROSOFT PROJECT 2016

File Task Resource Report Project View Format Tell me what you want to do... Sign in X

Cut Copy Calibri 11 Mark on Track Respect Links Manually Schedule Auto Schedule Inspect Move Mode Task Summary Notes Milestone Details Add to Timeline Find Clear Scroll to Task Fill

Gantt Chart Paste Format Painter B I U Font Clipboard Duration Start Finish Predecessors Resource Names

Sep 14, '25 Sep 21, '25 Sep 28, '25 W T F S S M T W T F S S M T W T F S S M T W T F

		Task Mode	Task Name	Duration	Start	Finish	Predecessors	Resource Names
1		➡	Khởi động dự án	37 days	Thu 9/11/25	Sun 11/2/25		Trương Văn Quốc
2		➡	Lập kế hoạch	0 days	Sun 11/2/25	Sun 11/2/25		Nguyễn Hoàng Nam
3		➡	Xác định yêu cầu	1 day	Thu 9/11/25	Thu 9/11/25	2	Nguyễn Hoàng Nam
4		➡	Phân công nhiệm vụ	1 day	Fri 9/12/25	Fri 9/12/25	3	Trương Văn Quốc I
5		➡	Nghiên cứu & Thu thập dữ liệu	1 day	Mon 9/15/25	Mon 9/15/25		Nguyễn Hoàng Nam[50%]
6		➡	Tìm hiểu lý thuyết	1 day	Tue 9/16/25	Tue 9/16/25	5	Trương Văn Quốc I
7		➡	Thu thập dữ liệu ảnh	1 day	Fri 9/19/25	Fri 9/19/25		Nguyễn Hoàng Nam
8		➡	Tiền xử lý dữ liệu	1 day	Mon 9/22/25	Mon 9/22/25	7	Trương Văn Quốc I
9		➡	Gán nhãn dữ liệu	1 day	Tue 9/23/25	Tue 9/23/25	8	Nguyễn Hoàng Nam
10		➡	▪ Mô hình CNN	13 days	Tue 9/23/25	Thu 10/9/25		
11		➡	Xây dựng mô hình CNN	1 day	Tue 9/23/25	Tue 9/23/25		Trương Văn Quốc I
12		➡	Chọn kiến trúc mô hình	1 day	Fri 9/26/25	Fri 9/26/25		Nguyễn Hoàng Nam
13		➡	Cài đặt mô hình	1 day	Mon 9/29/25	Mon 9/29/25	12	Trương Văn Quốc I
14		➡	Huấn luyện mô hình	1 day	Tue 9/30/25	Tue 9/30/25		Nguyễn Hoàng Nam
15		➡	Tinh chỉnh tham số	1 day	Wed 10/1/25	Wed 10/1/25	14	Trương Văn Quốc I
16		➡	Đánh giá & Kiểm thử mô hình	1 day	Fri 10/3/25	Fri 10/3/25		Nguyễn Hoàng Nam[50%]
17		➡	Tạo tập kiểm thử	1 day	Sun 10/5/25	Sun 10/5/25		Trương Văn Quốc I
18		➡	Kiểm thử kết quả	1 day	Mon 10/6/25	Mon 10/6/25	17	Nguyễn Hoàng Nam
19		➡	Phân tích lỗi	1 day	Thu 10/9/25	Thu 10/9/25		Trương Văn Quốc I
20		➡	Phát triển ứng dụng minh họa	1 day	Fri 10/10/25	Fri 10/10/25		Nguyễn Hoàng Nam[50%]
21		➡	Thiết kế giao diện	1 day	Sun 10/12/25	Sun 10/12/25		Nguyễn Hoàng Nam
22		➡	Tích hợp mô hình	1 day	Mon 10/13/25	Mon 10/13/25	21	Trương Văn Quốc I
23		➡	Kiểm thử ứng dụng	1 day	Fri 10/17/25	Fri 10/17/25		Nguyễn Hoàng Nam
24		➡	Báo cáo & Hoàn thiện	1 day	Mon	Mon		Trương Văn Quốc

GANNT CHART

```
graph LR; T12[Task 12] --> T13[Task 13]; T12 --> T10[Task 10]; T13 --> T14[Task 14]; T10 --> T15[Task 15]; T15 --> T16[Task 16]; T16 --> T17[Task 17]; T17 --> T18[Task 18]; T18 --> T19[Task 19]; T19 --> T20[Task 20]; T20 --> T21[Task 21]; T21 --> T22[Task 22]; T22 --> T23[Task 23]; T23 --> T24[Task 24];
```

# 1.0 NHẬN DIỆN DỰ ÁN

## Xây dựng mô hình phân loại bệnh dựa trên ảnh lá lúa

Tên	Xây dựng mô hình phân loại bệnh dựa trên ảnh lá lúa
Mô tả	Thiết kế, phát triển và triển khai mô hình học máy sử dụng kỹ thuật xử lý hình ảnh để phân loại các bệnh phổ biến ở cây lúa dựa trên ảnh lá. Bao gồm nghiên cứu vấn đề, lựa chọn thuật toán, chuẩn bị tập dữ liệu, huấn luyện mô hình và đánh giá hiệu suất.
Nhà tài trợ	Giảng viên hướng dẫn: Phạm Trọng Huynh / Khoa Công nghệ Thông tin
Quản lý dự án	Trương Văn Quốc Phong/1150080153
Tài nguyên đội ngũ dự án	Thành viên 1: [Trương Văn Quốc Phong/1150080153] ] - Chịu trách nhiệm nghiên cứu và thu thập tập dữ liệu. - Thành viên 2: [Nguyễn Hoàng Nam/1150080149] - Chịu trách nhiệm lựa chọn thuật toán và triển khai mô hình.

## 2.0 LÝ DO KINH DOANH CHO DỰ ÁN

Cải thiện thực tiễn nông nghiệp bằng cách phát hiện sớm bệnh trên cây lúa, giúp nông dân giảm thiểu tổn thất mùa màng và tối ưu hóa sử dụng thuốc trừ sâu trong bối cảnh thực tế.

Áp dụng các khái niệm công nghệ phần mềm và AI học được trong khóa học vào vấn đề thực tế, thể hiện kỹ năng học máy trong phân loại hình ảnh.

Đóng góp vào nông nghiệp bền vững ở các khu vực như Việt Nam, nơi lúa là cây trồng chính, bằng cách xây dựng công cụ có thể mở rộng cho ứng dụng di động hoặc sử dụng thực địa.

Phù hợp với mục tiêu khóa học về quản lý dự án, làm việc nhóm và triển khai kỹ thuật trong dự án phần mềm.

Nâng cao hồ sơ cá nhân của sinh viên với dự án AI thực hành tập trung vào dữ liệu thực và kết quả đo lường được.

### 3.0 MỤC TIÊU DỰ ÁN (MỤC ĐÍCH)

Tổng thể, tạo ra mô hình phân loại đáng tin cậy để xác định bệnh trên lá lúa (ví dụ: bệnh đạo ôn, đốm nâu, bệnh bạc lá) với độ chính xác khá cao sử dụng mạng nơ-ron tích chập hoặc kỹ thuật tương tự.

Nghiên cứu và hiểu vấn đề, bao gồm tập dữ liệu hiện có, phương pháp (ví dụ: CNN như ResNet hoặc mô hình tùy chỉnh), độ đo đánh giá (ví dụ: accuracy, precision, recall, F1-score) và kết quả hiện tại (ví dụ: mô hình đạt 90%+ trên tập dữ liệu chuẩn).

Lựa chọn và triển khai thuật toán phù hợp, ưu tiên mô hình dựa trên PyTorch cho phân loại hình ảnh.

Thu thập, gán nhãn và tiền xử lý tập dữ liệu để đảm bảo đa dạng và phù hợp cho huấn luyện.

Thực hiện thực nghiệm để huấn luyện, xác thực và kiểm tra mô hình, nhằm đến cải thiện so với baseline.

Phát triển kỹ năng làm việc nhóm, lập kế hoạch và tài liệu hóa cho sinh viên.

### 4.0 PHẠM VI DỰ ÁN

Bao gồm nghiên cứu phân loại bệnh lúa, chuẩn bị dữ liệu (sử dụng nguồn hiện có như PlantVillage hoặc xây dựng tập nhỏ 500-1000 ảnh nếu cần), phát triển mô hình và đánh giá cơ bản.

Giải quyết các thành phần cốt lõi: định nghĩa vấn đề trong bối cảnh nông nghiệp, lựa chọn thuật toán (ví dụ: biến thể CNN), xử lý dữ liệu, pipeline huấn luyện và phân tích kết quả.

Bao gồm hướng dẫn triển khai Python/PyTorch nhưng không bao gồm triển khai sản xuất (ví dụ: không tích hợp ứng dụng web/di động).

Không bao gồm yêu cầu bắt buộc cho tính năng nâng cao như xử lý thời gian thực hoặc tích hợp IoT.

Khung, tóm tắt nghiên cứu, hướng dẫn và công cụ được phát triển và triển khai trong thời gian học kỳ (ví dụ: đến cuối khóa học).

## 5.0 CÁC SẢN PHẨM DỰ ÁN CHÍNH

Tên	Mô tả
Khung	Lộ trình tổng thể phác thảo các thành phần chính của quy trình phân loại bệnh, bao gồm luồng dữ liệu, kiến trúc mô hình và bước đánh giá.
Bản tôn chỉ dự án	Tài liệu này phác thảo cấu trúc, mục tiêu và kế hoạch dự án.
Mô hình logic	Phác thảo các hoạt động chính (nghiên cứu, chuẩn bị dữ liệu, huấn luyện), kết quả (mô hình đã huấn luyện, báo cáo độ chính xác) và biện pháp (độ đo như F1-score).
Nền tảng dự án	Tài liệu bổ sung cho bản tôn chỉ với định nghĩa bệnh lúa, bối cảnh nông nghiệp và kết quả mong đợi.
Tóm tắt nghiên cứu	Tóm tắt tài liệu hiện có, tập dữ liệu (ví dụ: Rice Leaf Disease Dataset), phương pháp (ví dụ: CNN, transfer learning) và kết quả hiện tại (ví dụ: độ chính xác 95% trên benchmark).
Hướng dẫn	Hướng dẫn tổng quát cho thu thập/gán nhãn dữ liệu và huấn luyện mô hình; hướng dẫn kỹ thuật cho triển khai PyTorch.
Công cụ và tài nguyên	Mẫu cho chú thích dữ liệu, script code cho tiền xử lý/huấn luyện, tập dữ liệu mẫu, notebook đánh giá và mẹo gỡ lỗi.
Kế hoạch truyền thông/triển khai	Kế hoạch họp nhóm, cập nhật tiến độ cho giảng viên và trình bày cuối cùng; bao gồm kiểm soát phiên bản (ví dụ: GitHub).
Website/Kho lưu trữ	Kho GitHub chứa code, dữ liệu, báo cáo và tài liệu để chia sẻ trong nhóm và nộp bài.

## 7.0 CÁC VẤN ĐỀ CHÍNH

Mức độ nghiêm trọng	Mô tả
Cao	Các mức độ chuyên môn khác nhau giữa thành viên nhóm; một số quen với PyTorch, những người khác không.
Trung bình	Tính sẵn có và chất lượng tập dữ liệu; dữ liệu công khai có thể thiếu đa dạng cho giống lúa Việt Nam.
Cao	Cam kết từ thành viên nhóm để chịu trách nhiệm nhiệm vụ và tuân thủ thời hạn.
Trung bình	Duy trì động lực trong kỳ nghỉ hoặc kỳ thi.
Cao	Thời gian phối hợp gán nhãn dữ liệu với huấn luyện mô hình.
Trung bình	Tích hợp phản hồi từ tham vấn giảng viên.
Thấp	Nhận thức của sinh viên và giảng viên về độ phức tạp dự án.
Thấp	Nhận thức công chúng (không áp dụng, vì đây là dự án học thuật).