

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CẦN THƠ  
TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**LUẬN VĂN TỐT NGHIỆP  
NGÀNH KHOA HỌC MÁY TÍNH**

**Đề tài**

**XÂY DỰNG NGÂN HÀNG THÔNG TIN  
GIỐNG CÂY TRỒNG Ở VIỆT NAM CÓ TÍCH HỢP  
TÍNH NĂNG GOM CỤM**

**Sinh viên thực hiện: Châu Bôl  
Mã số: B2017024  
Khóa: 46**

**Cần Thơ, 05/2024**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC CẦN THƠ  
TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**LUẬN VĂN TỐT NGHIỆP  
NGÀNH KHOA HỌC MÁY TÍNH**

**Đề tài**

**XÂY DỰNG NGÂN HÀNG THÔNG TIN  
GIỐNG CÂY TRỒNG Ở VIỆT NAM CÓ TÍCH HỢP  
TÍNH NĂNG GOM CỤM**

**Giáo viên hướng dẫn:  
TS. Lưu Tiến Đạo  
TS. Trương Xuân Việt**

**Sinh viên thực hiện:  
Chau Bôl  
Mã số: B2017024  
Khóa: 46**

**Cần Thơ, 05/2024**

TRƯỜNG ĐẠI HỌC CẦN THƠ  
TRƯỜNG CÔNG NGHỆ THÔNG TIN  
VÀ TRUYỀN THÔNG

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM  
Độc lập – Tự do – Hạnh phúc

## XÁC NHẬN CHỈNH SỬA LUẬN VĂN THEO YÊU CẦU CỦA HỘI ĐỒNG

Tên luận văn (tiếng Việt và tiếng Anh):

Xây dựng ngân hàng thông tin các giống cây trồng ở Việt Nam có tích hợp tính năng gom cụm.

Building an information bank for plant varieties in Vietnam with integrated clustering feature.

Họ tên sinh viên: Chau Bôl

MASV: B2017024

Mã lớp: DI20Z6A2

Đã báo cáo tại hội đồng ngành: Khoa học máy tính

Ngày báo cáo: 08/05/2024

Luận văn đã được chỉnh sửa theo góp ý của Hội đồng.

Cần Thơ, ngày tháng 06 năm 2024

**Giáo viên hướng dẫn**

(Ký và ghi họ tên)

## NHẬN XÉT CỦA GIẢNG VIÊN

This image shows a full page of white paper with horizontal dashed gray lines, typical of primary-ruled notebook paper. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

## LỜI CẢM ƠN

Mỗi sự thành công đều gắn liền với những sự giúp đỡ, hỗ trợ dù ít hay nhiều, dù trực tiếp hay gián tiếp của người khác. Trong suốt khoảng thời gian từ những bước chân đầu tiên đến giảng đường đại học đến ngày hôm nay, em đã nhận được rất nhiều sự quan tâm, giúp đỡ của quý thầy cô, gia đình, bạn bè.

Và để hoàn thành luận văn tốt nghiệp này, em xin bày tỏ lòng biết ơn sâu sắc đến thầy Lưu Tiến Đạo và thầy Trương Xuân Việt, đã tận tình hướng dẫn em trong suốt thời gian thực hiện luận văn. Em xin cảm ơn các thầy rất nhiều vì trong suốt khoảng thời gian vừa qua thầy đã dành nhiều thời gian và tâm huyết để hỗ trợ em hoàn thành tốt luận văn tốt nghiệp của mình. Đồng thời em cũng rất biết ơn quý Thầy Cô Trường Công nghệ thông tin và Truyền thông, Trường Đại học Cần Thơ, đã giúp đỡ và tạo điều kiện cho em có cơ hội củng cố và hoàn chỉnh lại những kiến thức mà em đã được học.

Với điều kiện thời gian cũng như kinh nghiệm còn hạn chế, bài báo cáo này không thể tránh được những thiếu sót. Em rất mong nhận được sự chỉ bảo, đóng góp ý kiến của các quý thầy cô để em có điều kiện bổ sung, nâng cao ý thức của mình.

Em xin chân thành cảm ơn!

Cần Thơ, 06 tháng 05 năm 2024

Sinh viên thực hiện

Chau Bôl

---

## MỤC LỤC

<b>PHẦN GIỚI THIỆU .....</b>	<b>1</b>
1. Đặt vấn đề.....	1
2. Lịch sử giải quyết vấn đề.....	1
3. Mục tiêu đề tài .....	2
4. Đối tượng và phạm vi nghiên cứu .....	2
5. Nội dung nghiên cứu .....	2
5.1 Phương pháp thực hiện .....	2
5.2 Hướng giải quyết: .....	2
6. Kết quả đạt được.....	3
7. Bố cục luận văn .....	3
<b>PHẦN NỘI DUNG .....</b>	<b>4</b>
<b>CHƯƠNG 1.....</b>	<b>4</b>
<b>MÔ TẢ BÀI TOÁN.....</b>	<b>4</b>
1. Mô tả chi tiết bài toán.....	4
2. Vấn đề và giải pháp liên quan đến bài toán.....	4
2.1 Thu thập dữ liệu.....	4
2.2 Thiết kế cơ sở dữ liệu .....	7
<b>CHƯƠNG 2.....</b>	<b>8</b>
<b>CƠ SỞ LÝ THUYẾT .....</b>	<b>8</b>
1. Gom cụm giống cây trồng .....	8
1.1 Giới thiệu về giải thuật gom cụm .....	8
1.2 Giải thuật K-means .....	10
1.3 Phân cụm phân cấp (Hierarchical Clustering).....	14
2. Giới thiệu các công nghệ sử dụng .....	15
2.1 Giới thiệu Node Package Manager.....	15
2.2 Công nghệ Frontend .....	16
2.3 Công nghệ Backend.....	17
2.4 Hệ quản trị cơ sở dữ liệu MySQL .....	17
2.5 Công cụ hỗ trợ và thiết kế hệ thống.....	17
<b>CHƯƠNG 3.....</b>	<b>19</b>

---

<b>THIẾT KẾ VÀ CÀI ĐẶT .....</b>	<b>19</b>
2. Các yêu cầu chức năng của hệ thống.....	19
2.1 Sơ đồ Use case cho Khách vãng lai.....	19
2.2 Sơ đồ Use case cho Người dùng hệ thống.....	19
2.3 Sơ đồ Use case cho Quản trị viên .....	20
2.2 Mô tả các chức năng .....	21
3. Thiết kế cơ sở dữ liệu .....	27
3.1 Thực thể User.....	27
3.2 Thực thể Species.....	28
3.3 Thực thể Genus.....	28
3.4 Thực thể Properties .....	28
3.5 Thực thể Properties_Value .....	29
3.6 Thực thể Detail_Species .....	29
3.7 Thực thể Province .....	29
3.8 Thực thể Distribution.....	30
3.9 Thực thể Image .....	30
3.10 Thực thể Detail_Image .....	30
<b>CHƯƠNG 4.....</b>	<b>31</b>
<b>KẾT QUẢ THỰC NGHIỆM .....</b>	<b>31</b>
1. Giao diện sản phẩm .....	31
1.1 Giao diện trang chủ.....	31
1.2 Giao diện tìm kiếm .....	32
1.3 Giao diện đăng ký và đăng nhập.....	34
1.4 Giao diện chức năng Thêm vùng phân bố .....	35
1.5 Giao diện chức năng Thêm mẫu giống.....	35
1.6 Giao diện chức năng Phân cùm phân cấp.....	37
1.7 Giao diện gom nhóm với Kmeans .....	38
1.8 Giao diện Duyệt mẫu giống mới.....	38
1.9 Giao diện quản lý các thuộc tính .....	39
1.10 Giao diện quản lý danh sách loại giống.....	40
1.11 Giao diện trên thiết bị di động .....	40
2. Kết quả kiểm thử .....	43
1.1 Mục đích kiểm thử .....	43

---

---

1.2	Phạm vi kiểm thử .....	43
1.3	Kế hoạch kiểm thử .....	43
1.4	Quản lý kiểm thử.....	44
1.5	Các chức năng kiểm thử.....	44
<b>PHẦN KẾT LUẬN.....</b>		<b>46</b>
1.	Kết quả đạt được.....	46
2.	Hạn chế .....	46
3.	Hướng phát triển.....	46
<b>TÀI LIỆU THAM KHẢO .....</b>		<b>47</b>



---

## DANH MỤC HÌNH

Hình 1: Thuật toán gom cụm dữ liệu .....	8
Hình 2: Mô phỏng sự phân cụm dữ liệu .....	9
Hình 3: Phương pháp gom cụm dữ liệu .....	9
Hình 4: Dữ liệu sau khi được hiển thị trên mặt phẳng .....	11
Hình 5: Khởi tạo trọng tâm ngẫu nhiên .....	11
Hình 6: So sánh và phân cụm tiếp tục dựa trên thuộc tính chiều dài và chiều rộng hạt lúa .....	12
Hình 7: Cập nhật trọng tâm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa .....	13
Hình 8: Kết quả cuối cùng của phân cụm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa .....	13
Hình 9: Biểu diễn phân cụm trong Hierarchical Clustering .....	14
Hình 10: Sơ đồ Use case cho Khách vãng lai .....	19
Hình 11: Sơ đồ Use case cho Người dùng hệ thống .....	20
Hình 12: Sơ đồ Use case cho Quản trị viên .....	20
Hình 13: Thiết kế cơ sở dữ liệu .....	27
Hình 14: Giao diện trang chủ .....	31
Hình 15: Giao diện các loại giống cây trồng .....	31
Hình 16: Giao diện chi tiết về mẫu giống .....	32
Hình 17: Giao diện tìm kiếm đơn giản .....	32
Hình 18: Giao diện tìm kiếm bằng bộ lọc .....	33
Hình 19: Giao diện kết quả trả về khi tìm kiếm .....	33
Hình 20: Giao diện đăng nhập hệ thống .....	34
Hình 21: Giao diện đăng ký hệ thống .....	35
Hình 22: Giao diện thêm vùng phân bố .....	35
Hình 23: Giao diện thêm mẫu giống thủ công .....	36
Hình 24: Giao diện danh sách mẫu giống đã thêm .....	36
Hình 25: Giao diện thêm mẫu giống từ tập tin Excel .....	37
Hình 26: Giao diện phân cụm phân cấp .....	37
Hình 27: Gom nhóm Kmeans .....	38
Hình 28: Kết quả gom nhóm Kmeans .....	38
Hình 29: Giao diện danh sách giống đang chờ duyệt .....	39
Hình 30: Giao diện danh sách giống đã xóa .....	39
Hình 31: Quản lý các thuộc tính .....	40
Hình 32: Giao diện quản lý danh sách loại giống .....	40
Hình 33: Giao diện trang chủ trên thiết bị di động .....	41
Hình 34: Giao diện danh sách mẫu giống trên thiết bị di động .....	41
Hình 35: Giao diện chi tiết thông tin giống cây trên thiết bị di động .....	42
Hình 36: Giao diện tìm kiếm dựa trên các đặc tính hình, nông sinh học trên thiết bị di động .....	42
Hình 37: Giao diện kết quả tìm kiếm trên thiết bị di động .....	43

---

## DANH MỤC BẢNG

Bảng 1: Kết quả dữ liệu thu thập được .....	5
Bảng 2: Giá trị đặc tính hình thái và nông sinh học của giống ngô .....	5
Bảng 3: Dữ liệu về các mẫu giống lúa .....	10
Bảng 4: Phân cụm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa .....	12
Bảng 7: Mô tả chức năng “Tra cứu giống cây trồng” .....	22
Bảng 8: Mô tả chức năng “Tìm kiếm giống cây trồng” .....	21
Bảng 9: Mô tả chức năng “Tìm kiếm giống cây trồng dựa trên các đặc tính cây trồng”. .....	22
Bảng 11: Mô tả chức năng "Thêm mẫu giống" .....	23
Bảng 12: Mô tả chức năng "Thêm mẫu giống từ file Excel" .....	24
Bảng 13: Mô tả chức năng "Thêm vùng phân bố" .....	24
Bảng 14: Mô tả chức năng "Góm nhóm mẫu giống với Kmeans" .....	25
Bảng 15: Mô tả chức năng "Phân cụm phân cấp" .....	25
Bảng 16: Mô tả chức năng "Duyệt danh sách mẫu giống" .....	26
Bảng 17: Thông tin tài khoản người dùng .....	27
Bảng 18: Thông tin giống cây trồng .....	28
Bảng 19: Thông tin loại giống cây .....	28
Bảng 20: Thông tin đặc tính hình thái, nông sinh học của giống cây .....	29
Bảng 21: Thông tin giá trị của các đặc tính .....	29
Bảng 22: Thông tin chi tiết về giá trị thuộc tính của giống cây .....	29
Bảng 23: Tọa độ GPS theo vĩ độ và kinh độ của các địa điểm .....	30
Bảng 24: Sự phân bố của các giống cây trồng .....	30
Bảng 25: Thông tin và đường dẫn ảnh .....	30
Bảng 26: Thông tin hình ảnh của giống cây .....	30
Bảng 27: Kiểm thử chức năng tìm kiếm .....	44
Bảng 28: Kiểm thử chức năng tìm kiếm dựa trên các thuộc tính .....	44
Bảng 29: Kiểm thử chức năng duyệt danh sách thêm giống mới .....	45
Bảng 30: Kiểm thử chức năng quản lý loại giống .....	45
Bảng 31: Kiểm thử chức năng quản lý thuộc tính giống cây trồng .....	45

---

**DANH MỤC TỪ VIẾT TẮT**

STT	Thuật ngữ/ Từ viết tắt	Ý nghĩa
1	API	Application Programming Interface
2	HTTP	HyperText Transfer Protocol
3	IRRI	International Rice Research Institute
4	NIAS	National Institute of Agrobiological Sciences
5	NPM	Node Package Manager
6	ORM	Object Relational Mapping
7	URL	Uniform Resource Locator

---

## ABSTRACT

Vietnam, a country rich in plant genetic resources, is facing a serious challenge, which is the decline of crop genetic resources. Despite its diverse terrain and climate, which favor agricultural development, the country's crop varieties are under pressure from climate change, environmental pollution, and agricultural market fluctuations. This threatens genetic diversity and impacts the stability and growth of the agricultural sector.

In response to this need, the construction of the " Building an information bank for plant varieties in Vietnam " is proposed to conserve, exploit and use this valuable resource. This information bank will provide information about the varieties to researchers, students, technical specialists, agricultural engineers and also farmers, helping them choose the varieties that suit their needs. At the same time, the application also helps to collect clusters of varieties with similar properties, thereby assisting users in the classification and selection of suitable varieties.

With support from the “Building an information bank for plant varieties” will help strengthen the conservation and development of characteristic varieties, while promoting research and application of technology in agriculture.

---

## TÓM TẮT

Việt Nam, một quốc gia có nguồn gen cây trồng phong phú, đang đối mặt với thách thức nghiêm trọng là sự suy giảm của các nguồn gen giống cây. Mặc dù đất nước này có địa hình và khí hậu đa dạng, tạo điều kiện thuận lợi cho nền nông nghiệp, nhưng các giống cây đang phải đối mặt với áp lực từ biến đổi khí hậu, ô nhiễm môi trường và biến động thị trường nông sản. Điều này đe dọa sự đa dạng gen và ảnh hưởng đến sự ổn định và phát triển của ngành nông nghiệp.

Nhằm đáp ứng nhu cầu này, việc xây dựng "Ngân hàng thông tin giống cây trồng ở Việt Nam" được đề xuất để bảo tồn, khai thác và sử dụng nguồn tài nguyên quý giá này. Ngân hàng thông tin này sẽ cung cấp thông tin về các giống cây cho các nhà nghiên cứu, sinh viên, chuyên gia kỹ thuật, kỹ sư nông nghiệp và cả nông dân, giúp họ lựa chọn giống cây phù hợp với nhu cầu của mình. Đồng thời, ứng dụng còn giúp gom cùm các giống cây có đặc tính tương đồng, từ đó hỗ trợ người dùng trong việc phân loại và tìm kiếm giống cây phù hợp.

Ngân hàng thông tin giống cây trồng sẽ giúp tăng cường bảo tồn và phát triển các giống cây đặc trưng, đồng thời thúc đẩy nghiên cứu và ứng dụng công nghệ trong nông nghiệp.

## PHẦN GIỚI THIỆU

### 1. Đặt vấn đề

Theo những báo cáo những năm gần đây, khí hậu ở nước ta luôn thay đổi thất thường, đặc biệt là vào mùa khô thì tình trạng hạn hán diễn ra trầm trọng. Trong bài báo “Nhận định tình hình ENSO, hạn mặn khu vực Tây Nguyên và Đồng bằng sông Cửu Long năm 2024” [9], Cục Thủy lợi đánh giá rằng nguy cơ xảy ra hạn hán, thiếu nước, xâm nhập mặn ở cấp độ cao (cấp độ 3-4), trên phạm vi rộng, kéo dài ở một số vùng trên cả nước, đặc biệt tại các khu vực Tây Nguyên, Đồng bằng sông Cửu Long, Trung Bộ và Đông Nam Bộ.

Vấn đề cấp bách đặt ra là cần tìm các giải pháp chọn tạo nhanh và chính xác giống cây trồng mới, có năng suất, chất lượng cao, chống chịu các tác nhân sinh học và phi sinh học, đặc biệt là thích ứng với điều kiện khí hậu cực đoan đang diễn ra phức tạp.

Mặc dù Việt Nam là một quốc gia có nguồn gen cây trồng phong phú, nhưng ước tính có trên 80% các giống cây trồng địa phương đã không còn tồn tại trong sản xuất, và con số các loài cây bị đe dọa vẫn không ngừng tăng. Trong xu thế hội nhập kinh tế, do các tác động khác nhau của con người, và do những ảnh hưởng của biến đổi khí hậu, của suy thoái môi trường, các nguồn gen cây trồng của nước ta đã và đang bị suy giảm nghiêm trọng.

Để giải quyết vấn đề trên, chúng ta cần tạo ra một hệ thống có thể cung cấp các thông tin về các đặc tính hình thái và nông sinh học của cây trồng, giúp kỹ sư nông nghiệp cũng như là các nhà nông có những lựa chọn phù hợp cho việc gieo trồng với điều kiện môi trường từng khu vực.

Với những lý do trên, tôi quyết định thực hiện đề tài “Xây dựng ngân hàng thông tin giống cây trồng ở Việt Nam có tích hợp tính năng gom cùm”.

### 2. Lịch sử giải quyết vấn đề

Những năm qua, vấn đề nâng cao hiệu quả sản xuất nông nghiệp, nhất là trong lĩnh vực trồng trọt được Nhà nước Việt Nam ưu tiên đầu tư nhằm tạo sản phẩm có giá trị cao, chất lượng, tăng thu nhập cho nhà nông.

Tác giả Vũ Thị Thu Hiền đã dùng phần mềm Excel cùng với phần mềm NTSYSpc để phân tích, đánh giá sự đa dạng di truyền của 41 giống lúa có nguồn gốc khác nhau dựa trên 14 tính trạng kiểu hình [1]. Bốn mươi một giống lúa này được phân thành 10 nhóm cách biệt di truyền với sự sai khác 0,08. Tác giả Đoàn Thị Thùy Linh và Nguyễn Văn Khoa (2013) [2] đã sử dụng các đặc trưng cơ bản của cây lúa như chiều cao cây, số nhánh hữu hiệu, chiều dài và chiều rộng lá đòng, chiều dài bông, chiều dài và chiều rộng hạt gạo, khối lượng 1.000 hạt... để đánh giá sự khác biệt về hình thái của 50 mẫu giống lúa địa phương vùng Tây Bắc được thu thập và trồng khảo nghiệm tại huyện Thuận Châu, tỉnh Sơn La. Các số liệu phân tích thống kê bằng phần mềm Excel kết hợp với phần mềm NTSYSpc.

Bài báo "Ứng dụng giải thuật gom nhóm dữ liệu để nhận diện sự tương đồng giữa các giống lúa" của các tác giả Lưu Tiến Đạo, Âu Tấn Tài, Vũ Anh Pháp và Trần Nguyễn Minh Thư [3], được công bố trong Tạp chí Khoa học Trường Đại học Cần

Thơ năm 2017, là một đóng góp quan trọng trong lĩnh vực nghiên cứu khoa học khi xây dựng được phần mềm ứng dụng cung cấp công cụ hỗ trợ gom nhóm các giống lúa dựa vào nhiều đặc điểm hình thái, đặc tính nông sinh học. Tuy nhiên, hệ thống chỉ tập trung nghiên cứu về các đặc tính của giống lúa mà không mở rộng ra các loại cây trồng khác, điều này giới hạn sự áp dụng của phần mềm trong ngành nông nghiệp và yêu cầu phải có những nghiên cứu tiếp theo để phát triển và mở rộng phạm vi ứng dụng của công nghệ này đối với các loại cây trồng khác nhau.

### 3. Mục tiêu đề tài

Mục tiêu chính của đề tài là xây dựng hệ thống cung cấp các thông tin về đặc điểm hình thái và đặc tính nông sinh học của các giống cây trồng. Thông qua ngân hàng thông tin, các nhà nghiên cứu, sinh viên, chuyên viên kỹ thuật, kỹ sư nông nghiệp và ngay cả nông dân tìm được những giống cây trồng có các đặc tính mong muốn. Bên cạnh đó, có thể thực hiện gom nhóm các giống cây trồng dựa trên các thuộc tính của cây.

### 4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: là thông tin về các mẫu giống cây trồng được thu thập qua các bài báo, các nghiên cứu về đặc tính hình thái và nông sinh học cây trồng. Ngân hàng thông tin phải được thiết kế sao cho thân thiện với người dùng, đặc biệt là người nông dân dễ dàng tìm kiếm thông tin về giống cây phù hợp với mảnh đất của họ.

Phạm vi nghiên cứu: mỗi giống loài sẽ được mô tả bởi các đặc tính hình thái, đặc tính nông sinh học (ví dụ: lúa sẽ được mô tả bởi các đặc tính như góc lá đồng, độ cứng cây, chiều cao cây, thời gian sinh trưởng, ...).

### 5. Nội dung nghiên cứu

#### 5.1 Phương pháp thực hiện

- Phân tích yêu cầu:
  - o Thực hiện khảo sát nhu cầu và thực trạng, phân tích đánh giá nhu cầu và thực trạng, đặc tả yêu cầu xây dựng hệ thống.
  - o Phân tích đối tượng sử dụng, chức năng theo đối tượng sử dụng.
- Thiết kế:
  - o Sơ đồ yêu cầu chức năng.
  - o Giao diện tương ứng với người dùng.
  - o Phân tích và thiết kế các mô hình hệ thống cơ sở dữ liệu, luồng dữ liệu.
- Cài đặt:
  - o Công cụ lập trình NPM
  - o Thư viện ReactJS, React Native, Redux, TailwindCSS, NodeJS, ExpressJS, Sequelize và một số thư viện hỗ trợ khác.
  - o Hệ quản trị cơ sở dữ liệu MySQL.
  - o Trình soạn thảo code (IDE – Integrated Development Environment) là Visual Studio Code cùng một số extensions hỗ trợ khác.

#### 5.2 Hướng giải quyết:

- Cơ sở lý thuyết cần nắm:

- Hiểu biết cơ sở lý thuyết về các đặc tính hình thái, nông sinh học của các giống cây trồng.
- Nắm vững kiến thức về máy học ứng dụng, các phương pháp học không giám sát.
- Lý thuyết phân tích và thiết kế hệ thống thông tin.
- Hệ quản trị cơ sở dữ liệu MySQL.
- Kiến thức xây dựng ứng dụng trên website và thiết bị di động.
- Công nghệ sử dụng:
  - Phía backend: Framework ExpressJS của NodeJS, Sequelize (ORM dành cho NodeJS).
  - Phía frontend: HTML, JavaScripts, ReactJS, React Native, Redux, TailwindCSS.
  - Sử dụng hệ quản trị cơ sở dữ liệu MySQL.
- Công cụ hỗ trợ xây dựng và phát triển hệ thống:
  - Công cụ thiết kế hệ thống: PowerDesigner, StarUML, công cụ trực tuyến draw.io.
  - Công cụ lập trình: Visual Studio Code.
  - Công cụ hỗ trợ thực thi: Postman, Google Chrome, thiết bị di động Android.
  - Công cụ quản lý cơ sở dữ liệu: phpMyAdmin.

## 6. Kết quả đạt được

Xây dựng và thiết kế thành công ngân hàng thông tin các giống cây trồng trên cả nền tảng website và cả thiết bị di động. Bên cạnh đó người dùng còn có thể sử dụng chức năng gom cùm để phân ra các giống cây có các thuộc tính tương đồng.

## 7. Bố cục luận văn

### Phần Giới thiệu

Giới thiệu tổng quát về đề tài.

### Phần Nội dung

**Chương 1:** Mô tả chi tiết bài toán, các vấn đề đặt ra và giải pháp liên quan đến bài toán.

**Chương 2:** Cơ sở lý thuyết.

**Chương 3:** Thiết kế và cài đặt.

**Chương 4:** Kiểm thử hệ thống và đánh giá .

### Phần Kết luận

Trình bày kết quả đạt được và hướng phát triển hệ thống.



---

# PHẦN NỘI DUNG

## CHƯƠNG 1

### MÔ TẢ BÀI TOÁN

#### 1. Mô tả chi tiết bài toán

Với “Ngân hàng thông tin các giống cây trồng”, điều quan trọng nhất của đề tài sẽ là việc thu thập và xây dựng mô hình cơ sở dữ liệu. Quá trình này đòi hỏi sự tỉ mỉ và chính xác để đảm bảo rằng mọi thông tin được thu thập đều đầy đủ và chính xác. Việc thu thập dữ liệu từ các nguồn đáng tin cậy và đa dạng là bước đầu tiên quan trọng nhất. Đồng thời, việc xây dựng một cơ sở dữ liệu phù hợp và có cấu trúc linh hoạt là chìa khóa để tạo ra một hệ thống thông tin hiệu quả và dễ sử dụng.

Trong quá trình thu thập dữ liệu, ta sẽ tập trung vào việc tổng hợp thông tin từ nhiều nguồn khác nhau, bao gồm các tổ chức nghiên cứu, các bài báo khoa học về ngành nông nghiệp, và các nghiên cứu về các đặc tính hình thái, đặc tính nông sinh học của cây trồng. Với mỗi giống loài cây trồng thì các đặc tính về hình thái, đặc tính nông sinh học là khác nhau (ví dụ ở giống lúa sẽ có 69 các đặc tính). Do đó, sau khi thu thập dữ liệu, ta sẽ tiến hành phân tích, đánh giá tổng quát để đưa ra cách thiết kế, xây dựng mô hình cơ sở dữ liệu sao cho thông tin được tổ chức một cách hợp lý (phù hợp với tất cả các loại giống) và dễ dàng truy xuất, từ đó tạo ra một nền tảng vững chắc cho việc phát triển hệ thống thông tin giống cây trồng hiệu quả.

Từ mô hình cơ sở dữ liệu, ta sẽ tiến hành sẽ phát triển một ứng dụng website và di động để cung cấp truy cập dễ dàng và linh hoạt vào thông tin về các giống cây trồng. Ứng dụng này sẽ được thiết kế với giao diện người dùng thân thiện, cho phép người sử dụng tìm kiếm, lựa chọn giống cây phù hợp một cách thuận tiện. Các tính năng như tìm kiếm nâng cao dựa vào các thuộc tính, phân cụm giống cây sẽ được tích hợp để cung cấp trải nghiệm tốt nhất cho người dùng.

#### 2. Vấn đề và giải pháp liên quan đến bài toán

##### 2.1 Thu thập dữ liệu

###### a. Vấn đề

Với “Ngân hàng thông tin giống cây trồng ở Việt Nam” thì điều quan trọng nhất đối với hệ thống chính là dữ liệu về các đặc tính hình thái, đặc tính nông sinh học của các giống cây. Thông thường đối với các tổ chức, các tập dữ liệu sẽ luôn luôn được bảo mật vì đây là nguồn tài nguyên quý giá và cả về công sức của quá trình nghiên cứu. Ngoài ra thì dữ liệu thường phân tán và không đồng nhất, từ các trung tâm nghiên cứu đến các tổ chức cộng đồng về nông nghiệp, mỗi nguồn thông tin đều có định dạng và tiêu chuẩn riêng. Điều này dẫn đến việc thu thập dữ liệu trở nên rất khó khăn.

###### b. Giải pháp

Phương pháp chính để thu thập dữ liệu là dựa vào các bài báo, nghiên cứu về các đặc tính hình thái và đặc tính nông sinh học của các giống cây trồng. Và để giải quyết sự phân tán và không đồng nhất giữa các dữ liệu, việc tổng hợp lại các dữ liệu thu thập từ nhiều nguồn khác nhau là cần thiết. Qua quá trình tổng hợp, các dữ liệu

sẽ được xử lý và thống nhất để tìm ra điểm chung và tiêu chuẩn cho dữ liệu, từ đó tạo ra một cơ sở dữ liệu đồng nhất và dễ quản lý. Bằng cách này, ngân hàng thông tin giống cây trồng sẽ trở thành một nguồn tài nguyên quý giá cho cả nhà nghiên cứu và cộng đồng nông dân, giúp họ tiếp cận và sử dụng thông tin một cách hiệu quả để nâng cao sản xuất và bảo vệ nguồn tài nguyên cây trồng một cách bền vững.

Kết quả dữ liệu đã thu thập sẽ được trình bày trong Bảng 1 sau đây:

**Bảng 1: Kết quả dữ liệu thu thập được**

Số thứ tự	Loại giống	Số mẫu	Số thuộc tính	Cơ quan đánh giá tiêu chuẩn
1	Lúa	42	69	IRRI
2	Ngô	20	15	Viện Nghiên cứu và Phát triển cây trồng, Học viện Nông nghiệp Việt Nam
3	Dưa lưới	30	20	Viện Nghiên cứu và Phát triển cây trồng, Học viện Nông nghiệp Việt Nam
4	Hoa Lan	22	15	Viện Nghiên cứu và Phát triển cây trồng, Học viện Nông nghiệp Việt Nam
5	Đậu Nanh	15	22	Viện Nghiên cứu và Phát triển cây trồng, Học viện Nông nghiệp Việt Nam

Để hình dung kỹ hơn về các đặc tính hình thái và nông sinh học của cây trồng, ta sẽ tham khảo thêm từ Bảng 2 về các giá trị đặc tính của giống ngô dưới đây:

**Bảng 2: Giá trị đặc tính hình thái và nông sinh học của giống ngô**

Số thứ tự	Thuộc tính	Giá trị thuộc tính
1	Sắc tố antoxian của bẹ lá thứ nhất	(1) Không có/rất nhạt (2) Nhạt (3) Trung bình (4) Đậm
2	Hình dạng phần đỉnh lá thứ nhất	(1) Nhọn (2) Nhọn đến tròn
3	Mức xanh của bộ lá	(1) Nhạt (2) Trung bình (3) Đậm
4	Sự gợn sóng của mép lá	(1) Không có/rất ít (2) Trung bình (3) Nhiều
5	Góc giữa phiến lá và thân	(1) Rộng (2) Trung bình (3) Hẹp

Số thứ tự	Thuộc tính	Giá trị thuộc tính
6	Mức độ cong của phiến lá	(1) Rất cong (2) Cong (3) Khá cong (4) Hơi cong (5) Không cong/rất ít cong
7	Chiều rộng phiến lá (lá trên liền kề với bắp trên cùng)	(1) Hẹp (2) Trung bình (3) Rộng
8	Mức độ dích dắc của thân	(1) Không có/rất ít (2) Ít (3) Nhiều
9	Sắc tố antoxian ở rễ chân kiềng	(1) Không có/rất ít (2) Trung bình (3) Đậm
10	Sắc tố antoxian của lông	(1) Không có/rất nhạt (2) Nhạt (3) Trung bình (4) Đậm
11	Sắc tố antoxian ở chân đế mày	(1) Không có/rất nhạt (2) Nhạt (3) Trung bình (4) Đậm
12	Sắc tố antoxian của mày không kể chân đế	(1) Không có/rất nhạt (2) Nhạt (3) Trung bình (4) Đậm
13	Sắc tố antoxian của bao phấn	(1) Không có/rất nhạt (2) Nhạt (3) Trung bình (4) Đậm
14	Mật độ của hoa	(1) Thưa (2) Trung bình (3) Dày
15	Góc giữa trục chính cò và nhánh bên	(1) Rất hẹp (2) Hẹp (3) Trung bình (4) Rộng
16	Thế của nhánh bên cò	(1) Không cong/rất ít (2) Hơi cong (3) Khá cong
17	Số nhánh cò cấp 1	(1) Ít (2) Trung bình (3) Nhiều

Số thứ tự	Thuộc tính	Giá trị thuộc tính
18	Thời gian trổ cờ	(1) Sớm (2) Rất sớm (3) Trung bình (4) Muộn
19	Chiều dài nhánh cờ	(1) Ngắn (2) Trung bình (3) Dài
20	Chiều cao bông cờ	(1) Thấp (2) Trung bình (3) Cao

## 2.2 Thiết kế cơ sở dữ liệu

### a. Vấn đề

Với mỗi mẫu giống cây trồng, dữ liệu về các đặc tính hình thái và đặc tính nông sinh học khá là phong phú và đa dạng. Đặc biệt, mỗi loại giống loài lại có số lượng và loại các đặc tính khác nhau. Ví dụ, một mẫu lúa có đến 69 đặc tính hình thái và nông sinh học, trong khi một mẫu ngô chỉ có 15 đặc tính tương tự. Sự khác biệt này đặt ra thách thức lớn trong việc thiết kế cơ sở dữ liệu, đòi hỏi hệ thống phải đủ linh hoạt và uyển chuyển để có thể quản lý và lưu trữ các đặc tính này một cách hiệu quả.

### b. Giải pháp

Để giải quyết những vấn đề trong việc thiết kế cơ sở dữ liệu cho ngân hàng thông tin các giống cây trồng, tôi đã thực hiện việc phân tích các thông tin dữ liệu để đưa ra mô hình phù hợp và linh hoạt. Dựa trên các yêu cầu về tính đa dạng và phong phú của dữ liệu, cũng như nhu cầu truy xuất nhanh chóng và hiệu quả, tôi đã xây dựng một mô hình cơ sở dữ liệu quan hệ gồm năm bảng chính (chi tiết được trình bày ở Hình 13):

- Genus: Bảng này sẽ lưu trữ thông tin chi tiết về các loại giống.
- Species: Lưu trữ các thông tin chung về tên, xuất xứ của các giống cây trồng.
- Properties: Lưu trữ thông tin về các đặc tính của giống cây.
- PropertiesValue: Lưu trữ các giá trị của các đặc tính hình thái và nông sinh học.
- DetailSpecies: Mô tả chi tiết các thông tin giá trị thuộc tính của cây trồng.

Bằng cách áp dụng mô hình cơ sở dữ liệu này, hệ thống sẽ đảm bảo khả năng lưu trữ và quản lý các thông tin một cách hiệu quả, đồng thời hỗ trợ tốt việc truy xuất và xử lý thông tin về các giống cây trồng với sự đa dạng và phong phú về đặc tính.

## CHƯƠNG 2 CƠ SỞ LÝ THUYẾT

### 1. Gom cụm giống cây trồng

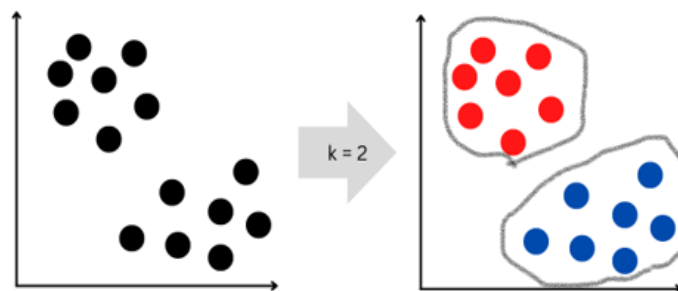
#### 1.1 Giới thiệu về giải thuật gom cụm

Khai phá dữ liệu thuộc quá trình khám phá trí thức. Về bản chất là giai đoạn duy nhất tìm ra được thông tin mới, tiềm ẩn có trong cơ sở dữ liệu chủ yếu phục vụ cho mô tả và dự đoán.

Gom cụm dữ liệu là một kỹ thuật trong khai phá dữ liệu với mục đích chính là khám phá cấu trúc của mẫu dữ liệu để thành lập các nhóm dữ liệu từ tập dữ liệu lớn, cho phép phân tích và nghiên cứu cho từng cụm dữ liệu nhằm khám phá và tìm kiếm các thông tin tiềm ẩn, hữu ích.

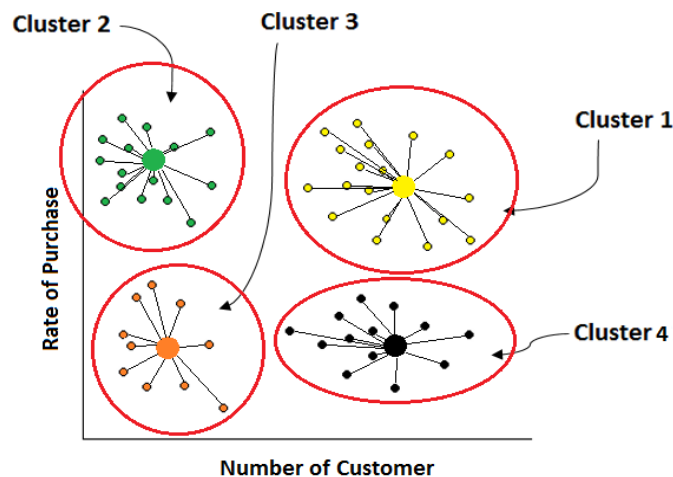
##### 1.1.1 Khái niệm và mục đích của gom cụm dữ liệu

Bài toán gom cụm dữ liệu là một nhánh ứng dụng chính của lĩnh vực học không giám sát, mà dữ liệu mô tả trong bài toán là không được gán nhãn. Trong trường hợp này, thuật toán sẽ tìm cách phân cụm dữ liệu thành từng nhóm có đặc điểm tương tự nhau, nhưng đồng thời đặc tính giữa các nhóm đó lại phải càng khác biệt càng tốt. Số các cụm dữ liệu có thể được xác định trước theo kinh nghiệm hoặc có thể tự động xác định theo thuật toán. Ta có thể hình dung đơn giản cách thức hoạt động qua Hình 1:



Hình 1: Thuật toán gom cụm dữ liệu

Độ tương tự được xác định dựa trên giá trị các thuộc tính mô tả đối tượng. Thông thường, phép đo khoảng cách thường được sử dụng để đánh giá độ tương tự hay phi tương tự. Vấn đề gom cụm có thể minh họa như Hình 2 dưới đây :

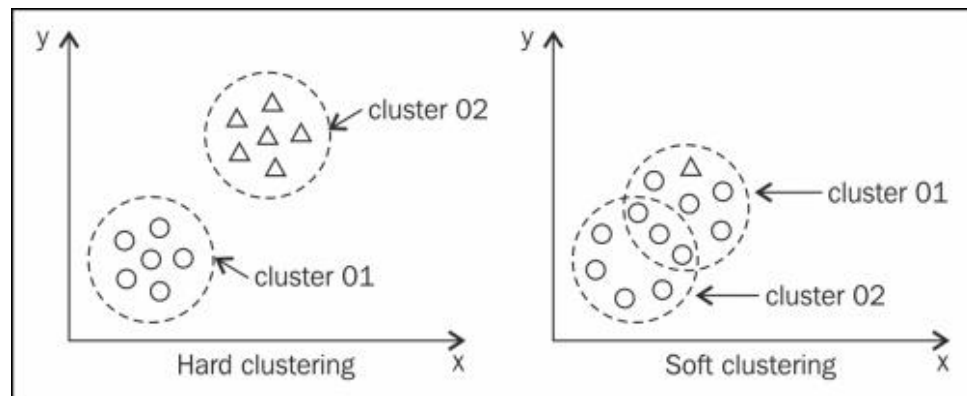


Hình 2: Mô phỏng sự phân cụm dữ liệu

### 1.1.2 Phương pháp gom cụm dữ liệu

Gom cụm dữ liệu được chia làm hai loại là gom cụm dữ liệu cứng và gom cụm dữ liệu mềm (Hình 3):

- Gom cụm dữ liệu cứng (hay phân cụm rõ) là phương pháp gán mỗi đối tượng vào một và chỉ một cụm và xác định rõ ranh giới giữa các cụm. Một số thuật toán: Thuật toán K-Means, thuật toán K-Medoids,...
- Gom cụm dữ liệu mềm (hay phân cụm mờ) là phương pháp cho phép mỗi đối tượng có thể thuộc một hoặc nhiều cụm dữ liệu và có sự mờ hồ hoặc mờ ranh giới giữa các cụm: Thuật toán Fuzzy C-mean,...



Hình 3: Phương pháp gom cụm dữ liệu

Tùy theo đặc điểm về tính tương đồng của các đối tượng trong bài toán đang xét, có nhiều cách tiếp cận cho thuật toán gom cụm. Các kỹ thuật gồm:

- Phân cụm phân cấp (Hierarchical Data Clustering)
- Phân cụm phân hoạch (Partition Based Data Clustering)
- Phân cụm dựa trên mật độ (Density Based Data Clustering)
- Phân cụm dựa trên lưới (Grid Based Data Clustering)

## 1.2 Giải thuật K-means

### 1.2.1 Phân cụm phân hoạch

Phân hoạch một tập dữ liệu có  $n$  phần tử thành  $k$  nhóm đến khi xác định được số lượng phần tử xác định trong mỗi cụm. Trong quá trình phân hoạch, tâm của các cụm được sinh ra ngẫu nhiên và số lượng tâm tương ứng với số lượng nhóm của tập dữ liệu. Các đối tượng không phải tâm sẽ được gán vào nhóm của cụm trung tâm gần nhất. Sau mỗi lần như vậy vị trí tâm sẽ được cập nhật lại cho đến khi kết quả phân cụm đạt tối ưu. Ý tưởng chính của thuật toán này là sử dụng chiến lược tham ăn (Greedy) để tìm kiếm nghiệm.

### 1.2.2 Giải thuật K-means

K-means là một giải thuật tiêu biểu của phương pháp phân cụm phân hoạch, và đây cũng là giải thuật được sử dụng phổ biến trong kỹ thuật gom cụm. Ý tưởng chính của thuật toán K-means là tìm cách gom nhóm các đối tượng đã cho vào  $k$  cụm ( $k$  là số các cụm được xác định trước,  $k$  nguyên dương) sao cho tổng bình phương khoảng cách giữa các đối tượng đến tâm nhóm (centroid) là nhỏ nhất.

Quá trình gom cụm K-means sẽ được thực hiện như sau:

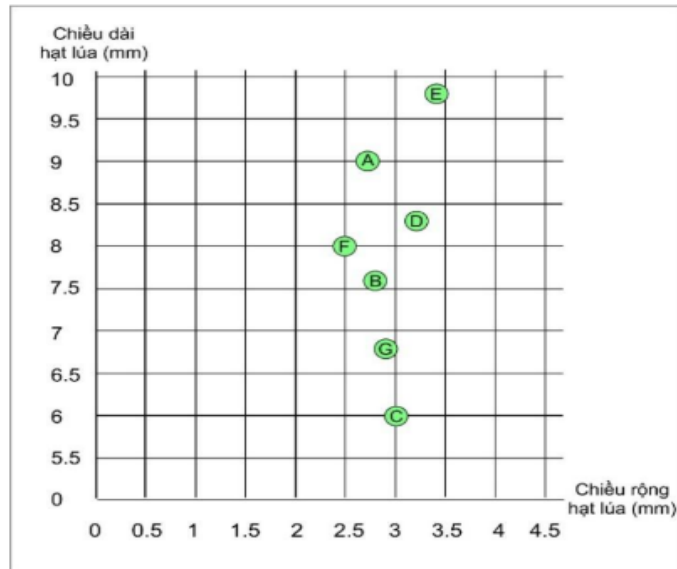
- Từ tập dữ liệu ban đầu gồm  $n$  phần tử và số cụm xác định là  $k$ .
- Chọn  $k$  đối tượng di ( $i=1, \dots, k$ ) làm tâm của  $k$  cụm từ tập dữ liệu ban đầu.
- Đối với mỗi đối tượng không phải là tâm, tính khoảng cách (Euclidean, Mahattan, ...) từ nó đến trọng tâm của các cụm còn lại. Xác định trọng tâm gần nhất cho mỗi đối tượng tức là xác định nhóm cho mỗi đối tượng trong tập dữ liệu dựa vào khoảng cách tính được.
- Cập nhật lại trọng tâm cho mỗi cụm bằng cách tính trung bình cộng vector của các đối tượng dữ liệu trong mỗi cụm.
- Lặp lại các bước trên cho đến khi các trọng tâm của cụm không còn thay đổi.

Ví dụ: chúng ta có bảng dữ liệu về các mẫu giống lúa.

**Bảng 3: Dữ liệu về các mẫu giống lúa**

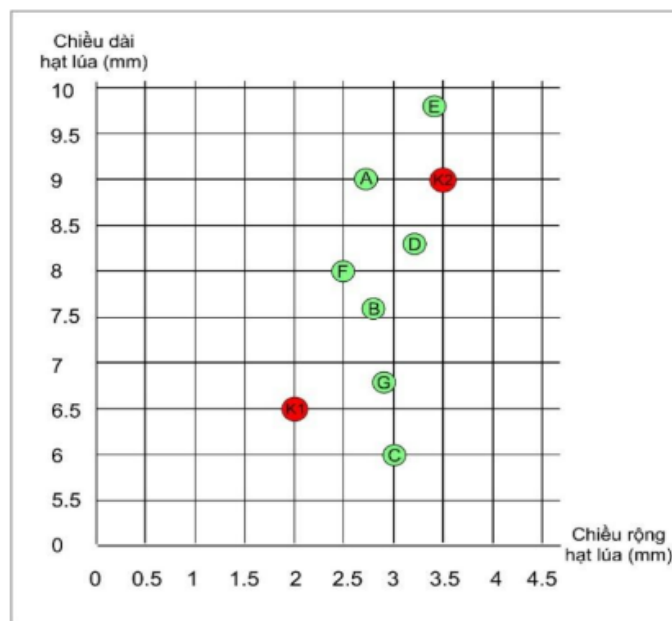
Giống lúa	Chiều dài hạt lúa (mm)	Chiều rộng hạt lúa (mm)
A	9	2.7
B	7.6	2.8
C	6	3
D	8.3	3.2
E	9.8	3.4
F	8	2.5
G	6.8	2.9

Dữ liệu sau khi được hiển thị trên mặt phẳng như Hình 4:



Hình 4: Dữ liệu sau khi được hiển thị trên mặt phẳng

- Bước 1: Giả sử chúng ta muốn chia các mẫu giống lúa thành 2 nhóm, đầu tiên sẽ thực hiện khởi tạo trọng tâm ngẫu nhiên K1(6.5, 2), K2(9, 3.5) (Hình 5):



Hình 5: Khởi tạo trọng tâm ngẫu nhiên

- Bước 2: Gán các điểm vào cụm có trọng tâm tương ứng dựa vào độ tương tự giữa các giống lúa với nhau (khoảng cách Euclidean được sử dụng để tính toán dựa trên thuộc tính chiều dài hạt lúa và chiều rộng hạt lúa trong ví dụ này).

$$d(A, K1) = \sqrt{(9 - 6.5)^2 + (2.7 - 2)^2} = 2.6$$



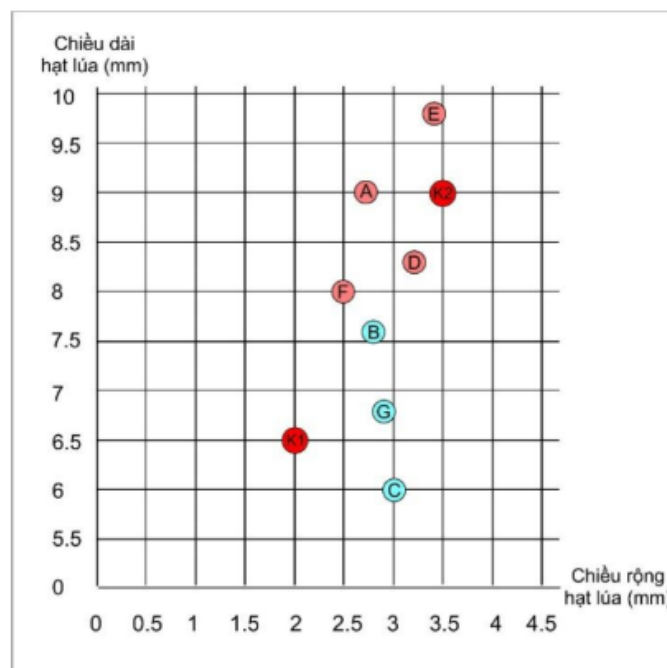
$$d(A, K2) = \sqrt{(9.2 - 9)^2 + (2.7 - 3.5)^2} = 0.8$$

- Ta thấy  $d(A, K1) > d(A, K2)$  nên mẫu A sẽ thuộc nhóm K2
- Tương tự với các điểm còn lại ta có được bảng sau (Bảng 2):

**Bảng 4: Phân cùm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa**

	K1(6.5, 2)	K2(9, 3.5)	
A(9, 2.7)	2.6	0.8	A ∈ K2
B(7.6, 2.8)	1.36	1.84	B ∈ K1
C(6, 3)	1.12	3.86	C ∈ K1
D(8.3, 3.2)	2.16	1.04	D ∈ K2
E(9.8, 3.4)	3.58	0.51	E ∈ K2
F(8, 2.5)	1.58	1.41	F ∈ K2
G(6.8, 2.9)	0.95	2.28	G ∈ K1

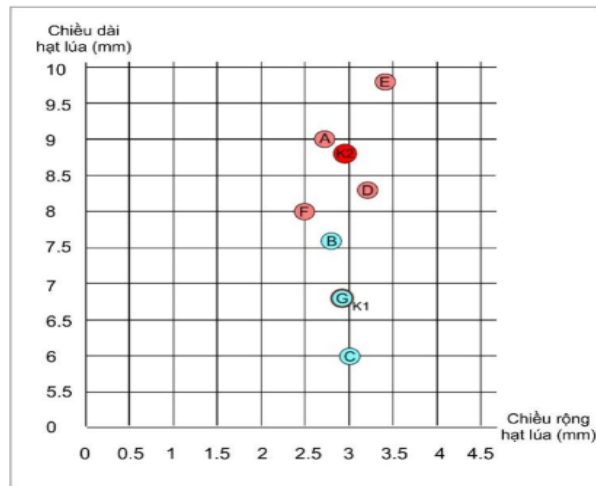
So sánh và gán các điểm còn lại vào mỗi cùm ta được kết quả như Hình 6 sau đây:



**Hình 6: So sánh và phân cùm tiếp tục dựa trên thuộc tính chiều dài và chiều rộng hạt lúa**

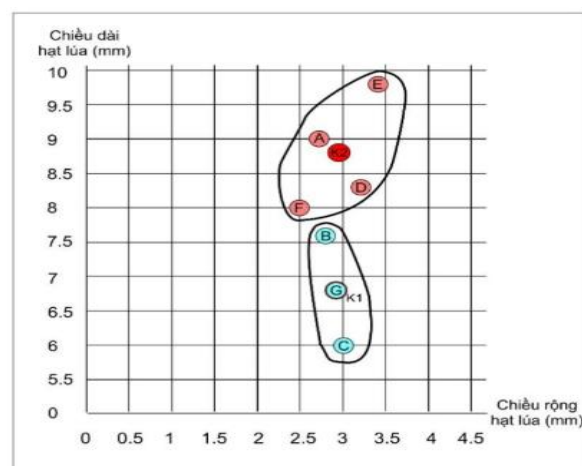
- Bước 3: Cập nhật lại trọng tâm. Tọa độ mới của trọng tâm là trung bình tổng tọa độ của các điểm thuộc nhóm chứa trọng tâm đó.

- $K1_{\text{Chiều dài}} = \frac{7.6+6+6.8}{3} = 6.8$
- $K1_{\text{Chiều rộng}} = \frac{2.8+3+2.9}{3} = 2.9$
- Tọa độ mới K1(6.8, 2.9). Tương tự với trọng tâm K2 sẽ là (8.8, 2.95) (Hình 7).



Hình 7: Cập nhật trọng tâm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa

- Quay lại bước 2 cho đến khi vị trí trọng tâm không còn thay đổi sau các lần cập nhật (giải thuật kết thúc).
- Ta được kết quả cuối cùng của quá trình gom nhóm (Hình 8):

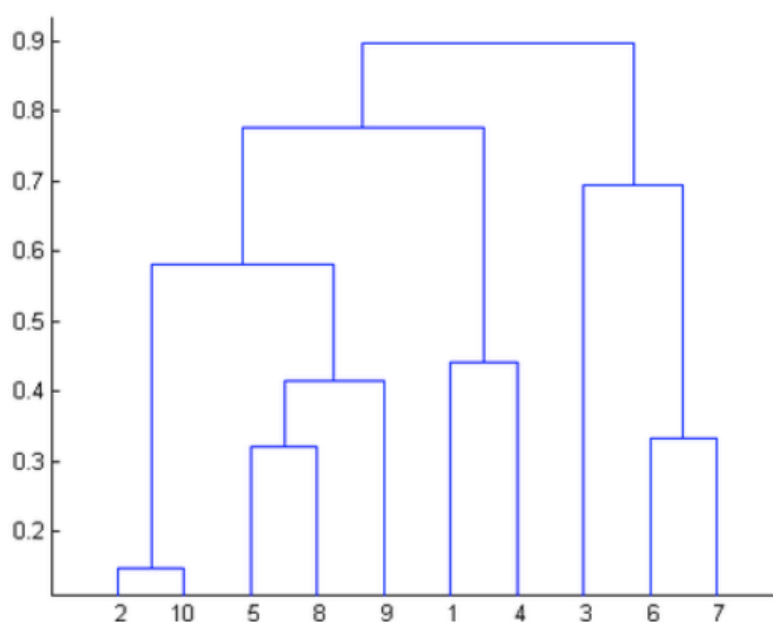


Hình 8: Kết quả cuối cùng của phân cụm dựa trên thuộc tính chiều dài và chiều rộng hạt lúa

### 1.3 Phân cụm phân cấp (Hierarchical Clustering)

#### 1.3.1 Giới thiệu về Hierarchical Clustering

Thuật toán phân cụm K-means cho thấy cần phải cấu hình trước số lượng cụm cần phân chia. Ngược lại, phương pháp phân cụm phân cấp (Hierarchical Clustering) không yêu cầu khai báo trước số lượng cụm. Thay vào đó, thuật toán chỉ yêu cầu xác định trước thước đo về sự khác biệt giữa các cụm (không giao nhau), dựa trên sự khác biệt từng cặp giữa các quan sát trong hai cụm. Theo phương pháp này, chúng tạo ra những biểu diễn phân cấp trong đó các cụm ở mỗi cấp của hệ thống phân cấp được tạo bằng cách hợp nhất các cụm ở cấp độ thấp hơn bên dưới. Ở cấp thấp nhất, mỗi cụm chứa một quan sát. Ở cấp cao nhất, chỉ có một cụm chứa tất cả dữ liệu. Các cấp của biểu diễn phân cụm được thể hiện trong đồ thị cây phân cấp bên dưới (Hình 9).



Hình 9: Biểu diễn phân cụm trong Hierarchical Clustering

Trục hoành thể hiện chỉ số của các quan sát trong nhóm được phân vào một cụm, trong khi trục tung là giá trị thước đo sự khác biệt giữa các cụm. Một cụm được đại diện bởi một nút mà toàn bộ các quan sát khác nếu thuộc cụm thì đều liên kết tới nút đó. Như vậy chúng ta có thể nhận thấy rằng các cụm có sự phân cấp dựa vào cấp độ của nút. Khi kẻ một đường thẳng nằm ngang cắt toàn bộ các đường thẳng đứng ta sẽ thu được các cụm tương ứng với các nút nằm gần nhất bên dưới đường thẳng. Bất kì hai cụm nào trong số chúng sẽ không chồng lấn nhau.

Trong Hierarchical Clustering, có hai chiến lược phân chia chính phụ thuộc vào chiều di chuyển trên biểu đồ cây phân cấp:

- Chiến lược hợp nhất: Chiến lược này sẽ đi theo chiều bottom-up (từ dưới lên trên). Quá trình phân cụm bắt đầu ở dưới cùng tại các nút lá (còn gọi là leaf node hoặc terminal node). Ban đầu mỗi quan sát sẽ được xem là một cụm tách biệt được thể hiện bởi một nút lá. Ở mỗi cấp độ chúng ta sẽ tìm cách hợp một

cặp cụm thành một cụm duy nhất nhằm tạo ra một cụm mới ở cấp độ cao hơn tiếp theo. Cụm mới này tương ứng với các nút quyết định (non-leaf node). Như vậy sau khi hợp cụm thì số lượng cụm ít hơn. Một cặp được chọn để hợp nhất sẽ là những cụm trung gian không giao nhau.

- Chiến lược phân chia: Chiến lược này sẽ thực hiện theo chiều top-down. Tức là phân chia bắt đầu từ nút gốc của đồ thị. Nút gốc bao gồm toàn bộ các quan sát, tại mỗi cấp độ chúng ta phân chia một cách đệ quy các cụm đang tồn tại tại cấp độ đó thành hai cụm mới. Phép phân chia được tiến hành sao cho tạo thành hai cụm mới mà sự tách biệt giữa chúng là lớn nhất.

### 1.3.2 Phương pháp hợp nhất giữa các cụm

Trong hierarchical clustering, các phương pháp hợp nhất (agglomerative) được sử dụng để xây dựng cấu trúc phân cụm bằng cách kết hợp các điểm dữ liệu vào các cụm ngày càng lớn. Các phương pháp này đều bắt đầu với mỗi điểm dữ liệu là một cụm độc lập và sau đó hợp nhất chúng dựa trên một tiêu chí cụ thể.

Có nhiều cách để hợp nhất giữa hai cụm lại, tùy theo từng loại mà có thể cho ra các kết quả gom cụm khác nhau:

- Single linkage: Phương pháp này đo lường sự khác biệt giữa hai cụm bằng cách lấy ra cặp điểm gần nhất giữa hai cụm.
- Complete linkage: Phương pháp này đo lường sự khác biệt giữa hai cụm bằng cách lấy ra hai cặp điểm xa nhau nhất giữa hai cụm.
- Average linkage: Phương pháp này sẽ lấy trung bình toàn bộ khoảng cách giữa các cặp điểm được lấy từ hai cụm.

Các phương pháp này đều giúp tạo ra một thước đo về sự không tương đồng hay chính là khoảng cách giữa hai cụm. Tuy nhiên phương pháp Single linkage và Complete linkage thường bị ảnh hưởng bởi những điểm dữ liệu ngoại lai. Chẳng hạn hai cụm rất cách xa nhau nhưng do hai điểm ngoại lai của chúng lại rất gần nhau có thể trả về một khoảng cách theo Single linkage rất bé. Một tình huống khác, khi hai cụm rất gần nhau nhưng do hai điểm ngoại lai của chúng rất xa nên khoảng cách được đo theo Complete linkage lại rất lớn. Và để kết quả của bài toán khả quan hơn nên ta sẽ quyết định lựa chọn Average linkage để làm phương pháp hợp nhất.

## 2. Giới thiệu các công nghệ sử dụng

### 2.1 Giới thiệu Node Package Manager

NPM – Node Package Manager: là một công cụ tạo và quản lý các thư viện lập trình Javascript cho Node.js.

Công dụng của NPM:

- NPM sẽ giúp cho công việc quản lý đơn giản hơn rất nhiều, đa số các thư viện đều đã có sẵn. Vì vậy, chúng ta chỉ cần tải về và việc include các thư viện sẽ trở nên dễ dàng hơn.
- Khi sử dụng công cụ NPM, người quản lý thư viện không mất nhiều thời gian và công sức để thực hiện được mọi việc vì mỗi đoạn code được sử dụng sẽ phụ thuộc nhiều mã nguồn mở khác nhau.

Phương thức hoạt động của NPM: hoạt động dựa trên 2 vai trò chính:

- NPM cung cấp một kho lưu trữ trực tuyến cho người dùng để chia sẻ và xuất bản các dự án mã nguồn mở của Node.js. Người dùng có thể tải lên và chia sẻ các gói JavaScript của mình trên NPM để người khác dễ dàng tìm kiếm, cài đặt và sử dụng.
- NPM sẽ trở thành một công cụ dòng lệnh hỗ trợ người dùng tương tác với kho lưu trữ NPM và quản lý các gói cũng như dependency của dự án. Người dùng có thể sử dụng các lệnh như gỡ cài đặt, cài đặt, cập nhật và kiểm tra phiên bản của các gói thông qua công cụ. NPM cũng cho phép người dùng tạo và quản lý các tập tin package.json (chứa các metadata, thông tin về dự án và những dependency).

## 2.2 Công nghệ Frontend

**HTML (Hypertext Markup Language – Ngôn ngữ đánh dấu siêu văn bản):** là xương sống của một trang web, được sử dụng để tạo và cấu trúc các thành phần trong trang web hoặc ứng dụng, phân chia các đoạn văn, heading, titles, ... và HTML không phải là ngôn ngữ lập trình. Một tài liệu HTML được hình thành bởi các phần tử HTML (HTML Elements) được quy định bằng các cặp thẻ (tag và attributes). Các cặp thẻ này được bao bọc bởi một dấu ngoặc nhọn và thường được khai báo thành một cặp.

**CSS (Cascading Styles Sheets):** là một ngôn ngữ được sử dụng để tìm và định dạng lại các phần tử được tạo bởi các ngôn ngữ đánh dấu (HTML). Nói ngắn gọn hơn là ngôn ngữ tạo phong cách cho trang web. Phương thức hoạt động của CSS là nó sẽ dựa vào các vùng chọn, vùng chọn có thể là tên một thẻ HTML, tên một ID, class hay nhiều kiểu khác. Sau đó nó sẽ áp dụng các thuộc tính cần thay đổi lên vùng chọn đó.

**TailwindCSS:** là một utility-first framework CSS giúp xây dựng trang web một cách hiện đại mà không cần rời khỏi file HTML. TailwindCSS cung cấp các class với những nhiệm vụ khác nhau, sử dụng class để thêm vào HTML. Nó cũng giống như Bootstrap, nó có những class built-in mà chúng ta có thể dễ dàng.

**Javascript:** là một ngôn ngữ lập trình kịch bản dựa vào đối tượng phát triển có sẵn hoặc tự định nghĩa, Javascript được sử dụng rộng rãi trong các ứng dụng Website, được hỗ trợ hầu như trên tất cả các trình duyệt như Firefox, Google Chrom, ... trên máy tính lẫn điện thoại.

**ReactJS:** là một thư viện viết bằng Javascript, dùng để xây dựng giao diện người dùng. ReactJS được sử dụng rộng rãi và có hệ sinh thái đa dạng, phong phú. Ở lĩnh vực Web, ReactJS có thể đem lại trải nghiệm tốt cho người dùng, cũng như khả năng Hot Reload giúp việc lập trình nhanh hơn.

**React Native:** là một framework phát triển ứng dụng di động được xây dựng bởi Facebook, cho phép xây dựng ứng dụng di động đa nền tảng với giao diện người dùng tương tự như ứng dụng native.. React Native cho phép lập trình viên sử dụng cùng một codebase để phát triển ứng dụng cho cả hệ điều hành iOS và Android, giúp tiết kiệm thời gian và công sức. Nó sử dụng các thành phần UI được xây dựng lại từ React, một thư viện JavaScript phổ biến để xây dựng giao diện người dùng web.

**Redux:** là một “predictable state management tool” cho các ứng dụng viết bằng Javascript, có tác dụng tạo ra một lớp quản lý mọi trạng thái, giúp ứng dụng có thể

hoạt động một cách nhất quán và có thể hoạt động trong nhiều môi trường khác nhau như client, server, native, ...

### 2.3 Công nghệ Backend

**NodeJS:** là một Javascript runtime được build dựa trên engine Javascript V8 của Google Chrome. NodeJS sử dụng kiến trúc hướng sự kiện event-driven, mô hình non-blocking I/O làm cho nó nhẹ và hiệu quả hơn. Hệ thống nén của NodeJS – npm là hệ thống thư viện nguồn mở lớn nhất thế giới.

**ExpressJS:** là một framework được xây dựng trên nền tảng của Node.js. Nó cung cấp các tính năng mạnh mẽ để phát triển web hoặc mobile. Express hỗ trợ các method HTTP và middleware tạo ra API vô cùng mạnh mẽ và dễ sử dụng. Bao gồm một số chức năng chính như sau:

- Thiết lập các lớp trung gian để trả về các HTTP request.
- Define router cho phép sử dụng với các hành động khác nhau dựa trên phương thức HTTP và URL.
- Cho phép trả về các trang HTML dựa vào các tham số.

**Sequelize:** là một ORM TypeScript và Node.js hiện đại dành cho Oracle, Postgres, MySQL, MariaDB, SQLite và SQL Server, .... Cùng với các tính năng như relations, transaction, replication, eager và lazy loading.

**JSON Web Token (JWT):** là một tiêu chuẩn mở (RFC 7519) nhằm xác minh thông tin an toàn giữa các bên Client-Server dưới dạng JSON object. Thông tin này có thể được xác minh và tin cậy vì nó được ký điện tử - digitally signed. JWT có thể được ký bằng cách sử dụng một secret (với thuật toán HMAC) hoặc cặp public/private key dùng chuẩn RSA hoặc ECDSA. Signed tokens có thể xác minh tính toàn vẹn của các claim có trong đó, trong khi encrypted tokens ẩn các claim từ các bên khác. Khi token được đăng ký bởi các cặp public/private keys, signature cũng xác nhận rằng chỉ có bên giữ private keys là nơi đã đăng ký nó.

### 2.4 Hệ quản trị cơ sở dữ liệu MySQL

**MySQL:** là hệ quản trị cơ sở dữ liệu tự do nguồn mở phổ biến nhất thế giới và được các nhà phát triển rất ưa chuộng trong quá trình phát ứng dụng. Vì MySQL là hệ quản trị cơ sở dữ liệu tốc độ cao, ổn định và dễ sử dụng, có tính khá chuyên, hoạt động trên nhiều hệ điều hành cung cấp một hệ thống lớn các hàm tiện ích rất mạnh.

### 2.5 Công cụ hỗ trợ và thiết kế hệ thống

**Visual Studio Code:** là một trình soạn thảo, biên tập code hoàn toàn miễn phí được Microsoft phát triển cho các lập trình viên và có mặt trên hầu hết các hệ điều hành phổ biến. Là một sự kết hợp độc đáo – đỉnh cao giữa IDE và Code Editor.

PowerDesigner và StarULM: phần mềm hỗ trợ đắc lực trong việc thiết kế các mô hình quan hệ: CDM (Conceptual Data Model – Mô hình dữ liệu mức quan niệm), PDM (Physical Data Model – Mô hình dữ liệu mức vật lý), Use Case (Use Case Diagram – Sơ đồ chức năng),...

**Xampp:** là một phần mềm cho phép giả lập môi trường server hosting ngay trên máy tính. XAMPP được viết tắt của X + Apache + MySQL + PHP + Perl vì nó được tích hợp sẵn Apache, MySQL, PHP, FTP server, Mail Server. Còn X thể hiện cho sự đa nền tảng của XAMPP vì nó có thể dùng được cho 4 hệ điều hành khác nhau: Windows, MacOS, Linux và Solaris.

---

**Postman:** là một ứng dụng mã nguồn mở dùng để phát triển và kiểm thử các API (Application Programming Interface). Công nghệ cung cấp môi trường cho các nhà phát triển thực hiện hoạt động tạo, chia sẻ, kiểm thử và quản lý các API. Nền tảng Postman có các tính năng như tạo yêu cầu HTTP, kiểm thử tự động, quản lý biến môi trường và biến toàn cục. Những tiện ích này sẽ giúp nhà phát triển tiết kiệm thời gian trong việc phát triển và kiểm thử API. Đồng thời, Postman còn có chế độ tương tác thông minh với API, giúp người dùng hiểu rõ hơn về cách API hoạt động và phản hồi của chúng.

## CHƯƠNG 3

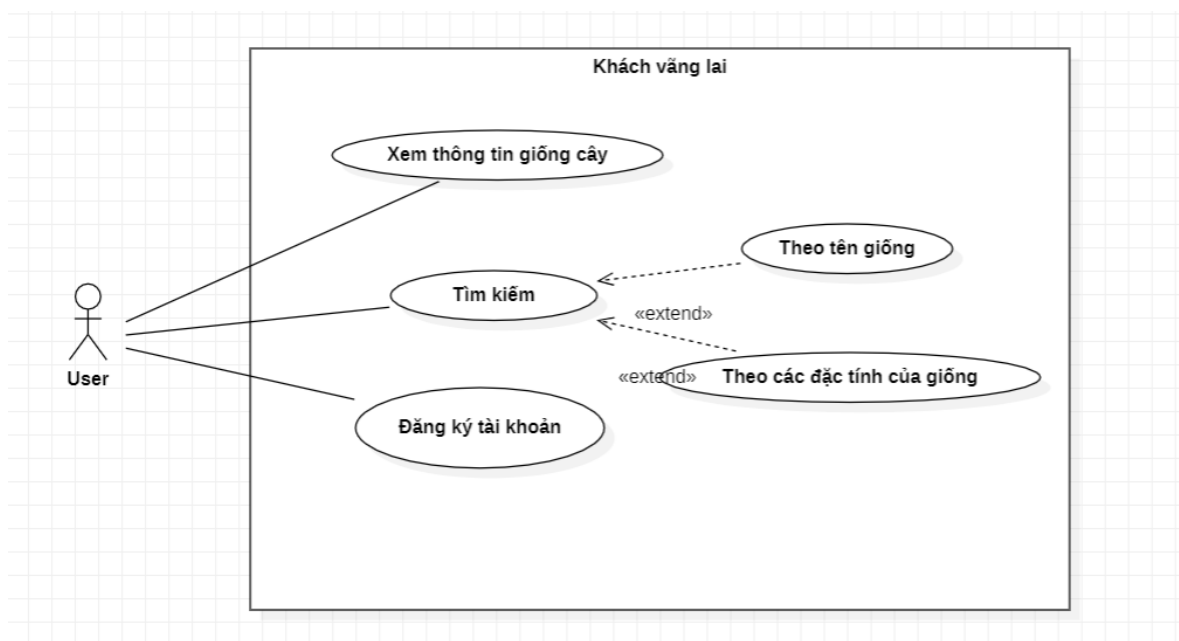
### THIẾT KẾ VÀ CÀI ĐẶT

#### 2. Các yêu cầu chức năng của hệ thống

Đối với “Ngân hàng thông tin giống cây trồng”, chúng ta có ba đối tượng sử dụng chính là: khách vãng lai, người dùng hệ thống và quản trị viên.

##### 2.1 Sơ đồ Use case cho Khách vãng lai

Đối với khách vãng lai thì có thể thực hiện các chức năng như là xem chi tiết thông tin giống cây trồng, thực hiện tìm kiếm thông tin giống cây theo tên hoặc tìm kiếm với các thuộc tính của giống cây, đăng ký tài khoản người dùng. Chúng ta có thể hình dung các chức năng đó qua sơ đồ use case trong Hình 10 dưới đây.

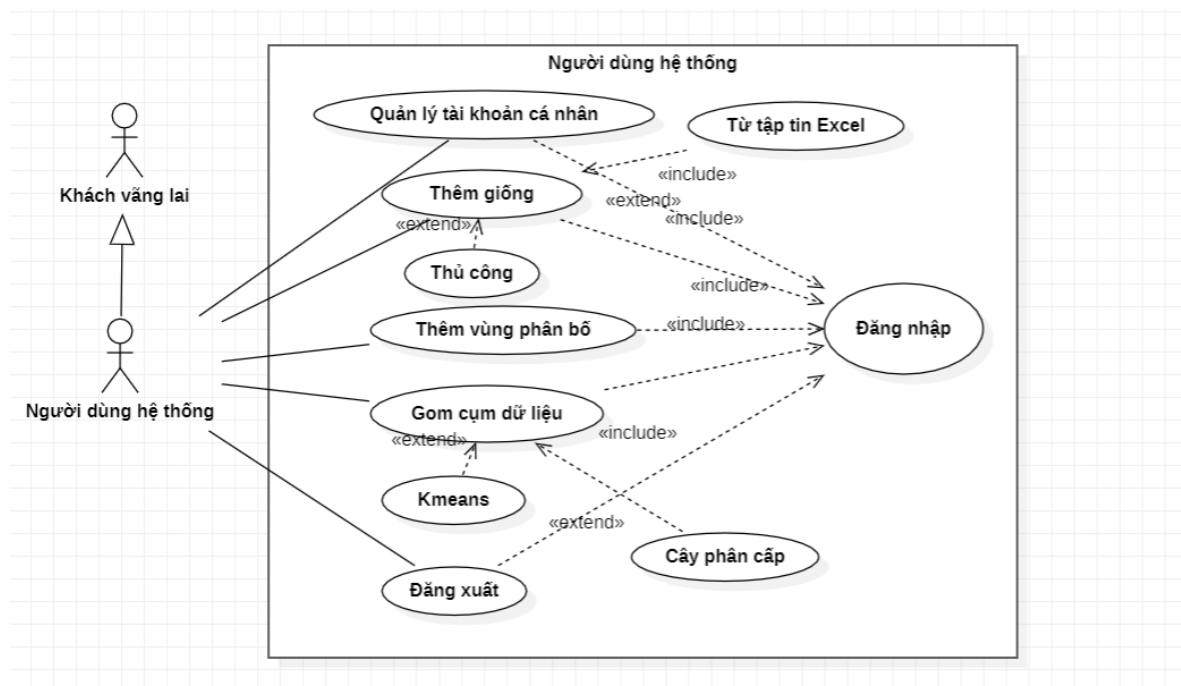


Hình 10: Sơ đồ Use case cho Khách vãng lai

##### 2.2 Sơ đồ Use case cho Người dùng hệ thống

Đối với người dùng hệ thống, ngoài các chức năng của khách vãng lai còn có thể thực hiện các chức năng sau khi đăng nhập thông tin tài khoản: quản lý tài khoản, thực hiện thêm giống cây mới bằng thủ công hoặc trực tiếp từ tệp tin Excel, thực hiện thêm vùng phân bố cho giống cây, các kỹ sư nông nghiệp cũng có thể phân tích sự tương đồng các mẫu giống với gom cùm dữ liệu. Chúng ta có thể hình dung các chức năng đó qua sơ đồ use case trong Hình 11 dưới đây.

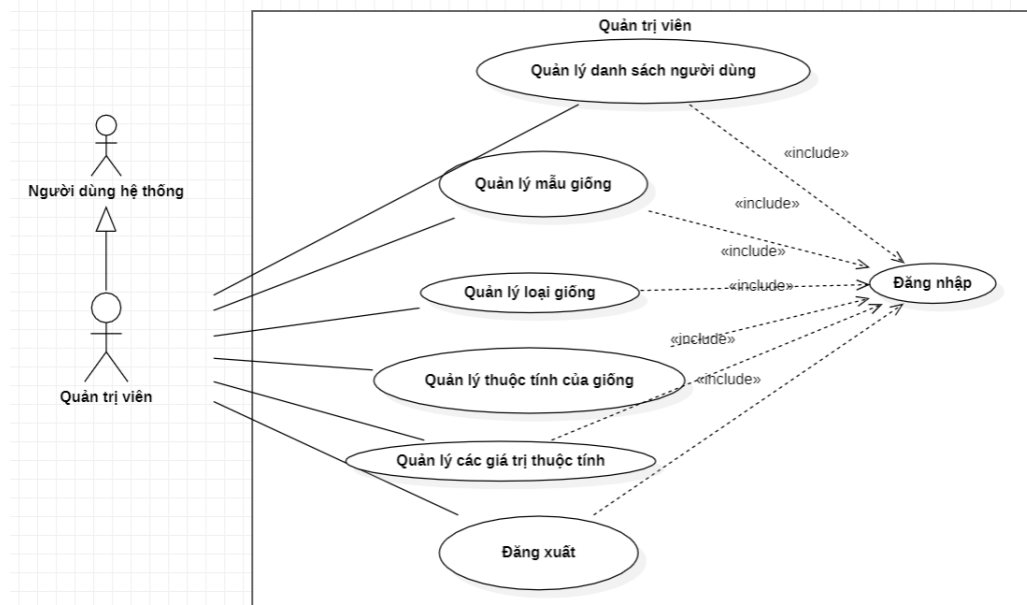




Hình 11: Sơ đồ Use case cho Người dùng hệ thống

### 2.3 Sơ đồ Use case cho Quản trị viên

Đối với quản trị viên, thực hiện quản lý các thông tin như: quản lý thông tin các giống cây, danh sách người dùng, duyệt danh sách các giống lúa được thêm, quản lý loại giống, các thuộc tính và các giá trị thuộc tính của cây trồng (Hình 12).



Hình 12: Sơ đồ Use case cho Quản trị viên

## 2.2 Mô tả các chức năng

Với ngân hàng thông tin, tôi sẽ chỉ đề cập đến các chức năng chính của hệ thống và bỏ qua các chức năng cơ bản như đăng nhập, đăng ký, quản lý tài khoản cá nhân.

### 2.2.1 Chức năng “Tìm kiếm giống cây trồng”

Chức năng “tìm kiếm giống cây trồng” là một trong những chức năng của tất cả actor trong hệ thống. Cho phép người dùng tìm kiếm tra cứu thông tin về các giống cây trồng. Các thông tin cụ thể về chức năng này bao gồm các kịch bản được mô tả cụ thể ở trong **Bảng 5** dưới đây:

**Bảng 5: Mô tả chức năng “Tìm kiếm giống cây trồng”**

Tên use case	Tìm kiếm giống cây trồng
Tóm tắt	Cho phép người dùng tra cứu các giống cây trồng.
Actor	Người dùng.
Điều kiện tiên quyết	Không có.
Kịch bản thường	<ol style="list-style-type: none"><li>Người dùng chọn thanh tìm kiếm, nhập tên mẫu giống cần tra cứu.</li><li>Nhấn nút tìm kiếm để hệ thống trả kết quả tìm kiếm. <b>A1</b> – Không thể tìm thấy các thông tin giống với từ khóa tìm kiếm mong muốn.</li><li>Hệ thống hiển thị thông tin theo yêu cầu người tìm kiếm.</li></ol>
Kịch bản thay thế	<p><b>A1</b> – Không thể tìm thấy các thông tin giống với từ khóa tìm kiếm mong muốn. Chuỗi A1 bắt đầu ở bước 2 của kịch bản thường.</p> <ol style="list-style-type: none"><li>Hệ thống hiển thị thông báo cho biết không tìm thấy kết quả phù hợp. Trở về bước 1 của kịch bản thường.</li></ol>
Kết quả	Thao tác tìm kiếm giống cây trồng thành công.

### 2.2.2 Chức năng “Tra cứu giống cây trồng”

Chức năng “tra cứu giống cây trồng” là một trong những chức năng của tất cả actor trong hệ thống. Cho phép người dùng tra cứu thông tin về các giống cây trồng. Các thông tin cụ thể về chức năng này bao gồm các kịch bản được mô tả cụ thể ở trong **Bảng 6** dưới đây:

**Bảng 6: Mô tả chức năng “Tra cứu giống cây trồng”**

Tên use case	Tra cứu giống cây trồng
Tóm tắt	Cho phép người dùng tra cứu các giống cây trồng.
Actor	Người dùng.
Điều kiện tiên quyết	Không có.
Kịch bản thường	1. Người dùng chọn Giống cây hoặc các hình ảnh về loại giống muốn tra cứu. 2. Chọn mẫu giống mà người dùng muốn tra cứu. 3. Hệ thống hiển thị thông tin theo yêu cầu.
Kết quả	Thao tác tra cứu giống cây trồng thành công.

### 2.2.3 Chức năng “Tìm kiếm giống cây trồng dựa trên các đặc tính cây trồng”

Chức năng “tìm kiếm giống cây trồng dựa trên các đặc tính cây trồng” là một trong những chức năng của tất cả actor trong hệ thống. Cho phép người dùng tìm kiếm tra cứu thông tin về các giống cây trồng dựa trên các đặc tính hình thái, nông sinh học. Các thông tin cụ thể về chức năng này bao gồm các kịch bản được mô tả cụ thể ở trong **Bảng 7** dưới đây:

**Bảng 7: Mô tả chức năng “Tìm kiếm giống cây trồng dựa trên các đặc tính cây trồng”.**

Tên use case	Tìm kiếm giống cây trồng dựa trên các đặc tính cây trồng
Tóm tắt	Cho phép người dùng tra cứu các giống cây trồng.
Actor	Người dùng.
Điều kiện tiên quyết	Không có.
Kịch bản thường	1. Người dùng chọn bộ lọc, chọn loại giống mà mình muốn kiếm. 2. Với các đặc tính hình thái, nông sinh học thì chọn các giá trị tương ứng. 3. Nhấn nút Search để hệ thống trả kết quả tìm kiếm. <b>A1</b> – Không thể tìm thấy các thông tin giống với các đặc tính đưa ra. 4. Hệ thống hiển thị các mẫu giống thỏa các điều kiện người dùng chọn.

<b>Kịch bản thay thế</b>	<b>A1</b> – Không thể tìm thấy các thông tin giống với giá trị đặc tính được chọn. Chuỗi A1 bắt đầu ở bước 2 của kịch bản thường. <b>4.</b> Hệ thống hiển thị thông báo cho biết không tìm thấy kết quả phù hợp. Trở về bước 1 của kịch bản thường.
<b>Kịch bản lỗi</b>	
<b>Kết quả</b>	Thao tác tìm kiếm giống cây theo các đặc tính cây trồng thành công.

#### 2.2.4 Chức năng “Thêm mẫu giống”

Chức năng “Thêm mẫu giống” là một trong chức năng của actor người dùng, cho phép người dùng thêm mẫu giống mới. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 8** dưới đây:

**Bảng 8: Mô tả chức năng "Thêm mẫu giống"**

Tên use case	Thêm mẫu giống
<b>Tóm tắt</b>	Cho phép người dùng thêm mẫu giống mới vào hệ thống, quản trị viên sẽ phê duyệt yêu cầu của người dùng được gửi lên.
<b>Actor</b>	Người dùng đã đăng nhập.
<b>Điều kiện tiên quyết</b>	Người dùng cần đăng nhập vào tài khoản cá nhân.
<b>Kịch bản thường</b>	<ol style="list-style-type: none"><li>1. Người dùng tiến hành thêm mẫu giống mới trên hệ thống, mẫu giống sẽ được đưa vào danh sách mẫu đã thêm.</li><li>2. Quản trị viên tiến hành phê duyệt mẫu giống và trả lại trạng thái duyệt.</li><li>3. Người dùng được hệ thống gửi thông báo mẫu giống đã xét duyệt. Hệ thống lưu lại thông tin mẫu giống này.</li></ol>
<b>Kết quả</b>	Quá trình thêm mẫu giống thành công

#### 2.2.5 Chức năng “Thêm mẫu giống từ file Excel”

Chức năng “Thêm mẫu giống từ file Excel” là một trong chức năng của actor người dùng, cho phép người dùng thêm mẫu giống mới. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 9** dưới đây:

**Bảng 9: Mô tả chức năng "Thêm mẫu giống từ file Excel"**

Tên use case	Thêm mẫu giống từ file Excel
Tóm tắt	Cho phép người dùng thêm mẫu giống mới từ file exce, vào hệ thống, quản trị viên sẽ phê duyệt yêu cầu của người dùng được gửi lên.
Actor	Người dùng đã đăng nhập.
Điều kiện tiên quyết	Người dùng cần đăng nhập vào tài khoản cá nhân.
Kịch bản thường	<ol style="list-style-type: none"><li>1. Người dùng tiến hành upload file Excel chứa các thông tin giống cây lên trên hệ thống, các mẫu giống sẽ được đưa vào danh sách giống đã thêm.</li><li>2. Quản trị viên tiến hành phê duyệt mẫu giống và trả lại trạng thái duyệt.</li><li>3. Người dùng được hệ thống gửi thông báo mẫu giống đã xét duyệt. Hệ thống lưu lại thông tin mẫu giống này.</li></ol>
Kết quả	Quá trình thêm mẫu giống từ file Excel thành công.

#### 2.2.6 Chức năng “Thêm vùng phân bố”

Chức năng “Thêm vùng phân bố” là một trong chức năng của actor người dùng, cho phép người dùng thêm vùng phân bố cho mẫu giống. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 10** dưới đây:

**Bảng 10: Mô tả chức năng "Thêm vùng phân bố"**

Tên use case	Thêm vùng phân bố
Tóm tắt	Cho phép người dùng thêm vùng phân bố của mẫu giống vào hệ thống.
Actor	Người dùng đã đăng nhập.
Điều kiện tiên quyết	Người dùng cần đăng nhập vào tài khoản cá nhân.
Kịch bản thường	<ol style="list-style-type: none"><li>1. Người dùng tiến hành chọn mẫu giống.</li><li>2. Người dùng chọn các địa điểm phân bố cho mẫu giống, bấm nút thực hiện.</li></ol>
Kết quả	Quá trình thêm vùng phân bố thành công

### 2.2.7 Chức năng “Gom nhóm mẫu giống với Kmeans”

Chức năng “Gom nhóm mẫu giống với Kmeans” là một trong chức năng của actor người dùng, cho phép người gom cụm các mẫu giống từ các đặc tính hình thái, nông sinh học. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 11** dưới đây:

**Bảng 11: Mô tả chức năng "Gom nhóm mẫu giống với Kmeans"**

Tên use case	Gom nhóm mẫu giống với Kmeans
Tóm tắt	Cho phép người dùng gom cụm được các mẫu giống cây dựa vào các đặc tính hình thái, nông sinh học.
Actor	Người dùng đã đăng nhập.
Điều kiện tiên quyết	Người dùng cần đăng nhập vào tài khoản cá nhân.
Kịch bản thường	<ol style="list-style-type: none"><li>1. Người dùng tiến hành upload file các mẫu giống.</li><li>2. Người dùng tiến hành chọn số nhóm, các đặc tính của mẫu giống để gom cụm.</li><li>3. Hệ thống sẽ đưa ra kết quả các cụm mẫu giống và thông tin chi tiết về mỗi cụm.</li></ol>
Kịch bản thay thế	
Kịch bản lỗi	
Kết quả	Quá trình gom nhóm với Kmeans thành công.

### 2.2.8 Chức năng “Phân cụm phân cấp”

Chức năng “Phân cụm phân cấp” là một trong chức năng của actor người dùng, cho phép người phân cụm các mẫu giống. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 12** dưới đây:

**Bảng 12: Mô tả chức năng "Phân cụm phân cấp"**

Tên use case	Phân cụm phân cấp
Tóm tắt	Cho phép người dùng phân cụm được các mẫu giống cây.
Actor	Người dùng đã đăng nhập.
Điều kiện tiên quyết	Người dùng cần đăng nhập vào tài khoản cá nhân.

<b>Kịch bản thường</b>	<ol style="list-style-type: none"><li>1. Người dùng tiến hành upload file các mẫu giống.</li><li>2. Bấm nút “Xác nhận” để thực hiện.</li><li>3. Hệ thống sẽ đưa ra hình ảnh cây phân cấp phân cùm các mẫu giống đã upload.</li></ol>
<b>Kết quả</b>	Quá trình phân cùm phân cấp thành công.

### 2.2.9 Chức năng “Duyệt danh sách mẫu giống”

Chức năng “Duyệt danh sách mẫu giống” là một trong chức năng của người quản trị viên, cho phép người quản trị viên duyệt hoặc từ chối duyệt các mẫu giống mà người dùng đã thêm. Các thông tin cụ thể của chức năng này được thể hiện cụ thể trong **Bảng 13** dưới đây:

**Bảng 13: Mô tả chức năng "Duyệt danh sách mẫu giống"**

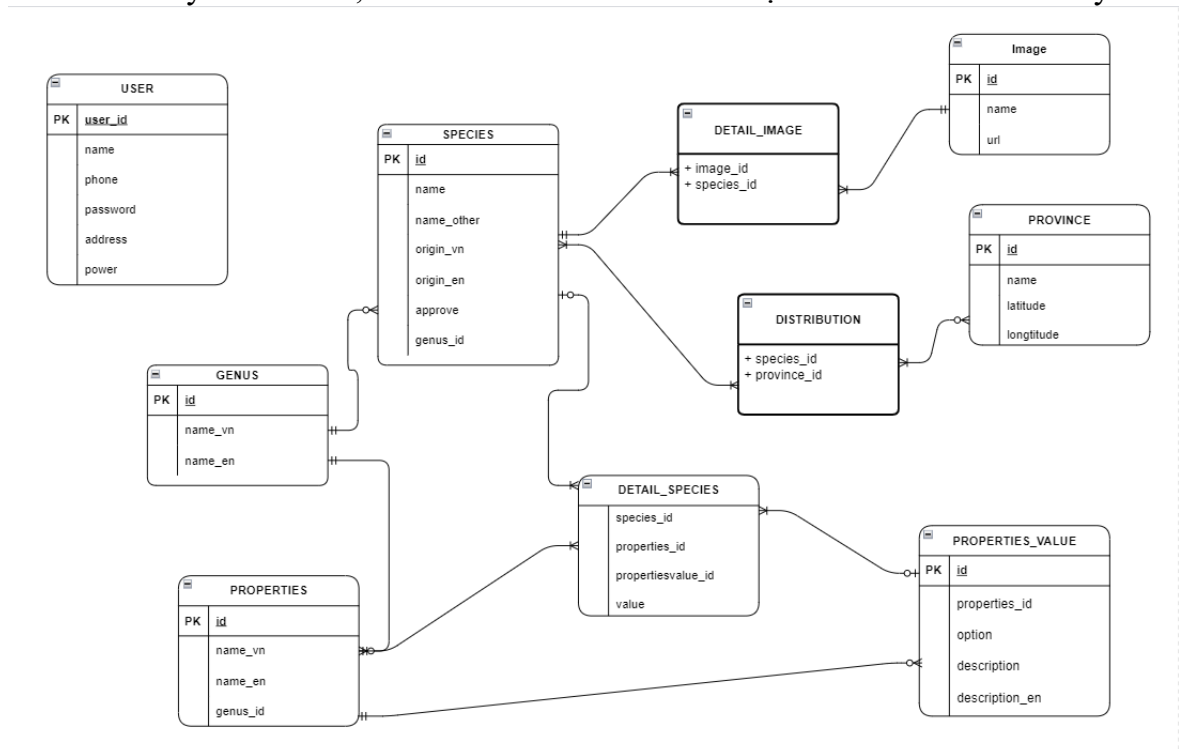
Tên use case	Duyệt danh sách mẫu giống
<b>Tóm tắt</b>	Cho phép người quản trị viên thực hiện duyệt hoặc từ chối duyệt các mẫu giống mà người dùng đã thêm.
<b>Actor</b>	Người quản trị viên.
<b>Điều kiện tiên quyết</b>	Người quản trị viên.
<b>Kịch bản thường</b>	<ol style="list-style-type: none"><li>1. Dựa vào danh sách các mẫu giống mà người dùng đã thêm, người quản trị tiến duyệt hoặc xóa. <b>Có thể nhảy đến:</b> <b>A1</b> – Người quản trị thực hiện duyệt. <b>A2</b> – Người quản trị thực hiện xóa.</li><li>2. Hệ thống cập nhật lại danh sách chờ duyệt.</li></ol>
<b>Kịch bản thay thế</b>	<b>A1</b> – Người quản trị thực hiện duyệt. Hệ thống cập nhật lại trạng thái của mẫu giống thành đã được duyệt. <b>A2</b> – Người quản trị thực hiện xóa. Hệ thống cập nhật lại trạng thái của mẫu giống thành bị xóa và đưa vào danh sách mẫu giống bị xóa.
<b>Kịch bản lỗi</b>	
<b>Kết quả</b>	Quá trình phân cùm phân cấp thành công.

### 3. Thiết kế cơ sở dữ liệu

Mục đích của việc tạo cơ sở dữ liệu cho "Ngân hàng thông tin các giống cây trồng" là để tổ chức và lưu trữ thông tin về các giống cây trồng một cách cấu trúc và có thể truy xuất dễ dàng. Mỗi loại giống cây có thể có một số lượng đặc tính khác nhau, bao gồm các đặc tính hình thái, đặc tính nông sinh học. Do đó, thiết kế cơ sở dữ liệu cần phải linh hoạt để có thể lưu trữ được nhiều loại giống cây với số lượng đặc tính khác nhau.

Để đảm bảo tính linh hoạt, cơ sở dữ liệu được thiết kế cần thỏa mãn yêu cầu việc nhà quản trị dễ dàng thêm, bớt, và hiệu chỉnh thuộc tính của các loại giống cây, cũng như triển khai cho nhiều loại cây trồng khác nhau.

Với các yêu cầu trên, ta sẽ có mô hình cơ sở dữ liệu như Hình 11 dưới đây:



Hình 13: Thiết kế cơ sở dữ liệu

#### 3.1 Thực thể User

Lưu thông tin tài khoản người dùng

Bảng 14: Thông tin tài khoản người dùng

Thuộc tính	Kiểu dữ liệu	Mô tả
user_id (Khóa chính)	string	Mã số tài khoản
name	string	Tên người dùng
phone	string	Số điện thoại người dùng
password	string	Mật khẩu
address	string	Địa chỉ người dùng
gender	boolean	Giới tính của người dùng: + True: Nam



		+ False: Nữ
power	boolean	Quyền hạn tài khoản có các giá trị: + True: Quản trị viên + False: Người dùng bình thường

### 3.2 Thực thể Species

Lưu thông tin giống cây trồng

**Bảng 15: Thông tin giống cây trồng**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số giống cây trồng
name	string	Tên giống cây trồng
name_other	string	Tên gọi khác của giống cây trồng
origin_vn	string	Nguồn gốc của giống lúa
origin_en	string	Nguồn gốc của giống lúa (tiếng Anh)
approve	int	Trạng thái giống cây khi người dùng thêm mới hoặc sửa. Có các giá trị như sau: + 0: là đang đợi duyệt. + 1: là đã được duyệt và được hiển thị ra. + -1: là đã bị xóa. Nếu muốn khôi phục lại thì vào chức năng Danh sách đã xóa.
genus_id (Khóa ngoại)	string	Mã số loại giống cây

### 3.3 Thực thể Genus

Lưu thông tin loại giống cây

**Bảng 16: Thông tin loại giống cây**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số loại giống cây
name_vn	string	Tên loại giống cây
name_en	string	Tên loại giống cây (tiếng Anh)

### 3.4 Thực thể Properties

Lưu thông tin đặc tính hình thái, nông sinh học của giống cây

**Bảng 17: Thông tin đặc tính hình thái, nông sinh học của giống cây**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số đặc tính hình thái, nông sinh học của giống cây
name_vn	string	Tên gọi của đặc tính
name_en	string	Tên gọi của đặc tính (tiếng Anh)

### 3.5 Thực thể Properties\_Value

Lưu thông tin giá trị của các đặc tính

**Bảng 18: Thông tin giá trị của các đặc tính**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số giá trị đặc tính
properties_id (Khóa ngoại)	string	Mã số đặc tính hình thái, nông sinh học của giống cây
option	string	Giá trị của đặc tính (Mỗi đặc tính sẽ có nhiều giá trị và được xem là các lựa chọn)
description	string	Mô tả giá trị đặc tính
description_en	string	Mô tả giá trị đặc tính (tiếng Anh)

### 3.6 Thực thể Detail\_Species

Lưu thông tin chi tiết về giá trị thuộc tính của giống cây

**Bảng 19: Thông tin chi tiết về giá trị thuộc tính của giống cây**

Thuộc tính	Kiểu dữ liệu	Mô tả
species_id	string	Mã số giống cây trồng
properties_id	string	Mã số đặc tính hình thái, nông sinh học của cây trồng
propertiesValue_id	string	Mã số giá trị của đặc tính
value	float	Giá trị của đặc tính

### 3.7 Thực thể Province

Lưu tọa độ GPS theo vĩ độ và kinh độ của các địa điểm

**Bảng 20: Tọa độ GPS theo vĩ độ và kinh độ của các địa điểm**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số địa điểm
name	string	Tên địa điểm
latitude	float	Vĩ độ
longitude	float	Kinh độ

### 3.8 Thực thể Distribution

Lưu sự phân bố của các giống cây trồng

**Bảng 21: Sự phân bố của các giống cây trồng**

Thuộc tính	Kiểu dữ liệu	Mô tả
species_id	string	Mã số giống cây trồng
province_id	string	Mã số địa điểm

### 3.9 Thực thể Image

Lưu thông tin và đường dẫn ảnh

**Bảng 22: Thông tin và đường dẫn ảnh**

Thuộc tính	Kiểu dữ liệu	Mô tả
id (Khóa chính)	string	Mã số của ảnh
name	string	Mô tả về ảnh
url	string	Đường dẫn của ảnh

### 3.10 Thực thể Detail\_Image

Lưu thông tin hình ảnh của giống cây

**Bảng 23: Thông tin hình ảnh của giống cây**

Thuộc tính	Kiểu dữ liệu	Mô tả
species_id	string	Mã số giống cây trồng
image_id	string	Mã số hình ảnh

## CHƯƠNG 4

### KẾT QUẢ THỰC NGHIỆM

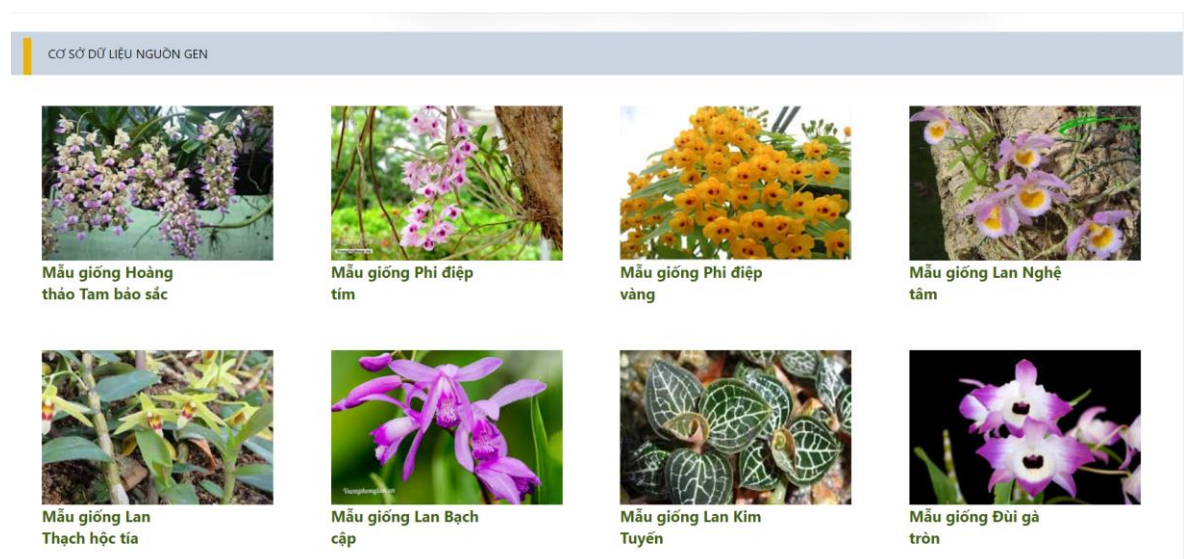
#### 1. Giao diện sản phẩm

##### 1.1 Giao diện trang chủ



Hình 14: Giao diện trang chủ

Ở giao diện trang chủ của website (Hình 14) sẽ hiển thị hình ảnh các loại giống cây trồng có trong dữ liệu. Người dùng nhấp chuột vào các hình ảnh hoặc đưa chuột vào “Giống cây” ở thanh điều hướng rồi chọn các loại giống mình muốn. Sau khi chọn, hệ thống sẽ hiển thị ra danh sách các mẫu giống cây của loại giống đó (Hình 15).



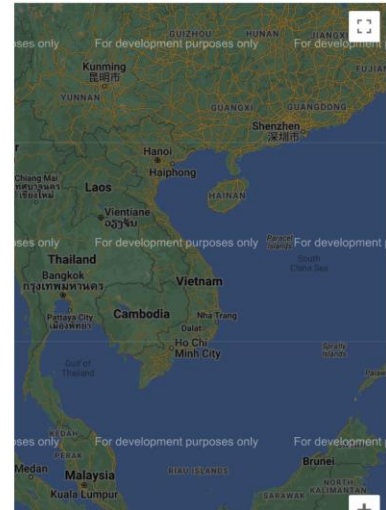
Hình 15: Giao diện các loại giống cây trồng

Với danh sách các mẫu giống cây trồng, để xem chi tiết thông tin các đặc tính hình thái, đặc tính nông sinh học thì thực hiện nhấp chuột vào ảnh của mẫu. Hệ thống sẽ chuyển qua giao diện hiển thị bảng thông tin chi tiết và nơi phân bố của mẫu giống (Hình 16).

**BẢNG CHI TIẾT VỀ ĐẶC ĐIỂM HÌNH THÁI VÀ NÔNG HỌC CỦA MẪU GIỐNG**

Mẫu giống : Phi điệp tím	Tên trước kia : không có	Xuất xứ : Miền Bắc
<b>Đặc tính cụ thể của mẫu giống</b>		
Loại hình thân	Đa thân	
Khả năng phân nhánh	Nhiều	
Hình dạng thân	Tròn	
Độ dài thân chính	Trung bình	
Màu sắc thân	Màu xanh đậm	
Độ cứng thân	Mềm, rủ	
Dài lông (giới hạn giữa hai lá)	Trung bình (1,5-2,5cm)	
Màu sắc lá	Màu xanh đậm	
Hình dạng lá	Tròn	
Độ dày lá	Trung bình	
Phân bố lá	So le	
Dạng rễ	Rễ chùm	
Dạng sống	Khí sinh	
Đặc điểm nở hoa	Tập trung	

**CÁC ĐỊA ĐIỂM GIEO TRỒNG**



Hình 16: Giao diện chi tiết về mẫu giống

## 1.2 Giao diện tìm kiếm

Ngân hàng thông tin trang bị hai cách thức tìm kiếm dành cho người dùng: tìm kiếm đơn giản và tìm kiếm nâng cao dựa vào các đặc tính của mẫu (bộ lọc).

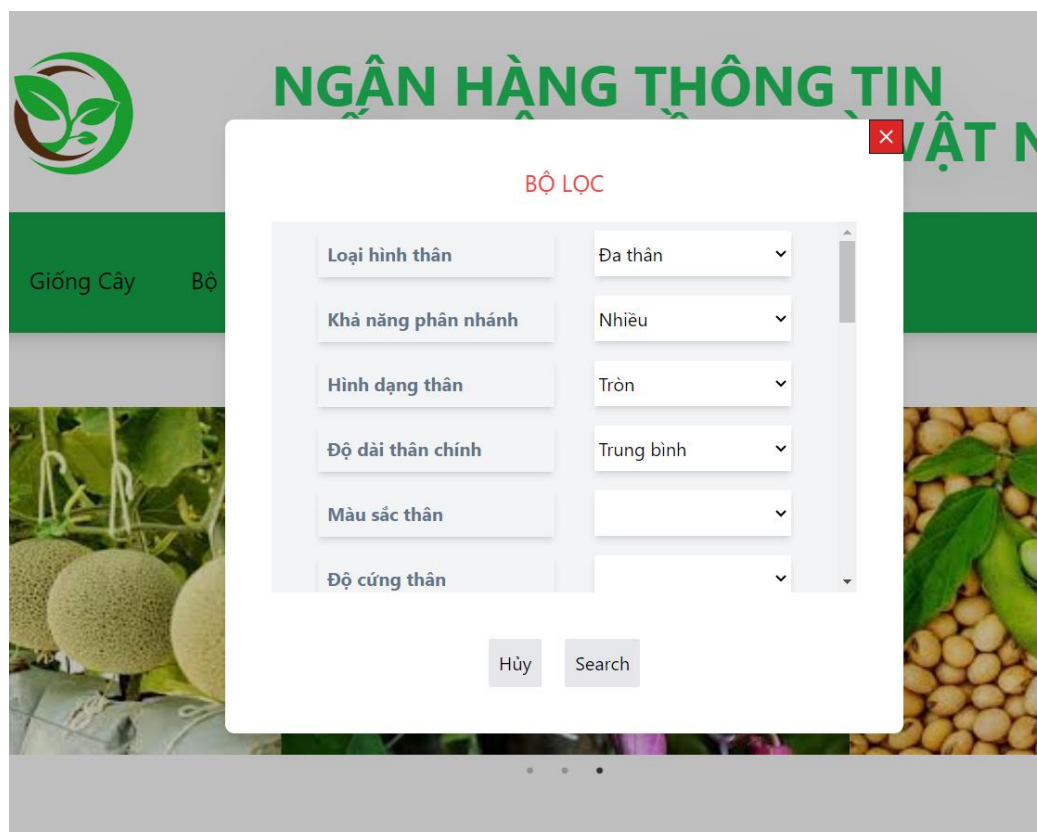
Tìm kiếm đơn giản: người dùng chỉ cần nhập tên giống cây ở ô bên phải, góc trên website. Nếu nhập đúng tên của mẫu giống, hệ thống sẽ chuyển sang giao diện chi tiết thông tin của mẫu, ngược lại sẽ thông báo với người dùng không tìm thấy mẫu giống ở góc bên phải màn hình (Hình 17).



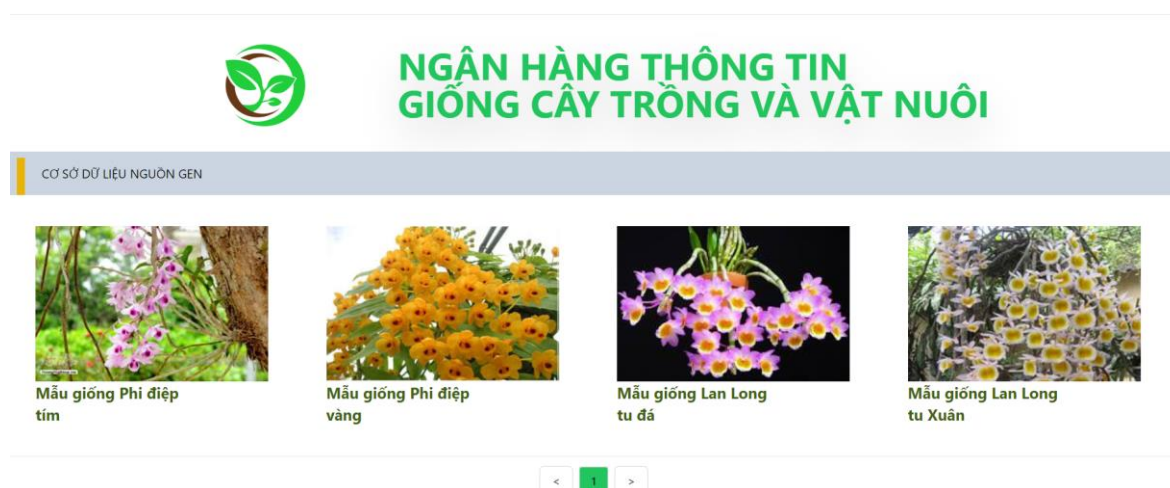
Hình 17: Giao diện tìm kiếm đơn giản



Với chức năng tìm kiếm dựa vào các đặc tính của giống cây, người dùng sẽ nhấp chọn vào “Bộ Lọc” ở thanh điều hướng và chọn loại giống mình muốn tìm. Hệ thống sẽ hiển thị các đặc tính hình thái, đặc tính nông sinh học cũng như các giá trị của các đặc tính (Hình 18). Sau khi bấm nút “Search”, hệ thống hiển thị ra danh sách các mẫu giống thỏa các điều kiện các thuộc tính được chọn (Hình 19).



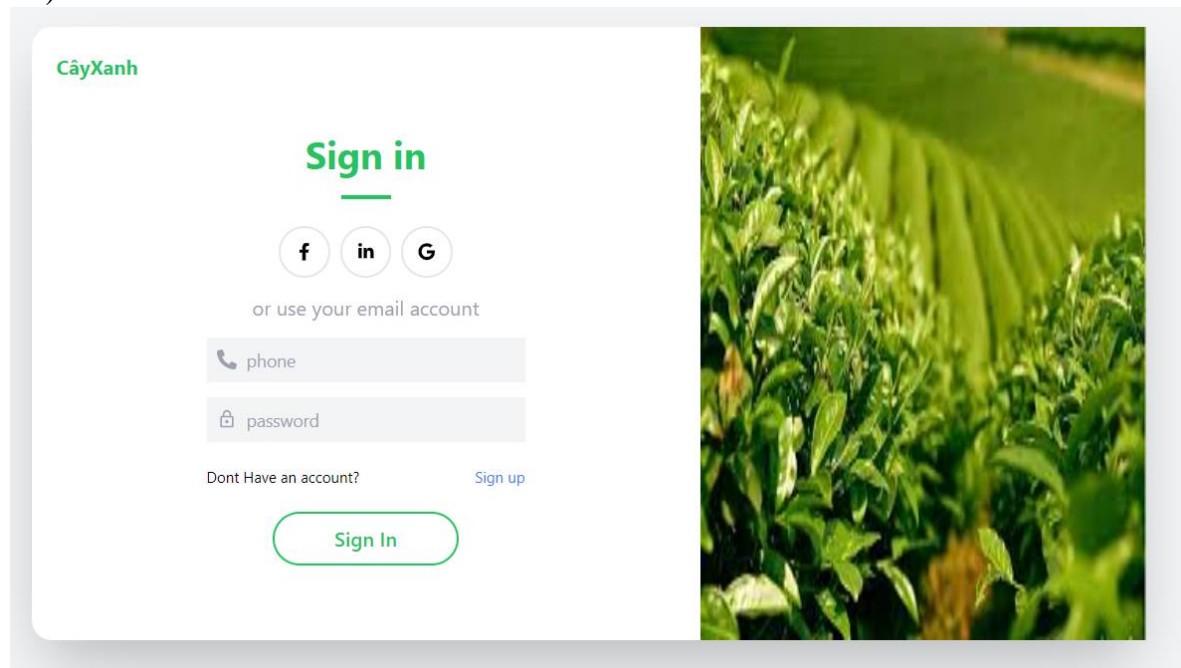
Hình 18: Giao diện tìm kiếm bằng bộ lọc



Hình 19: Giao diện kết quả trả về khi tìm kiếm

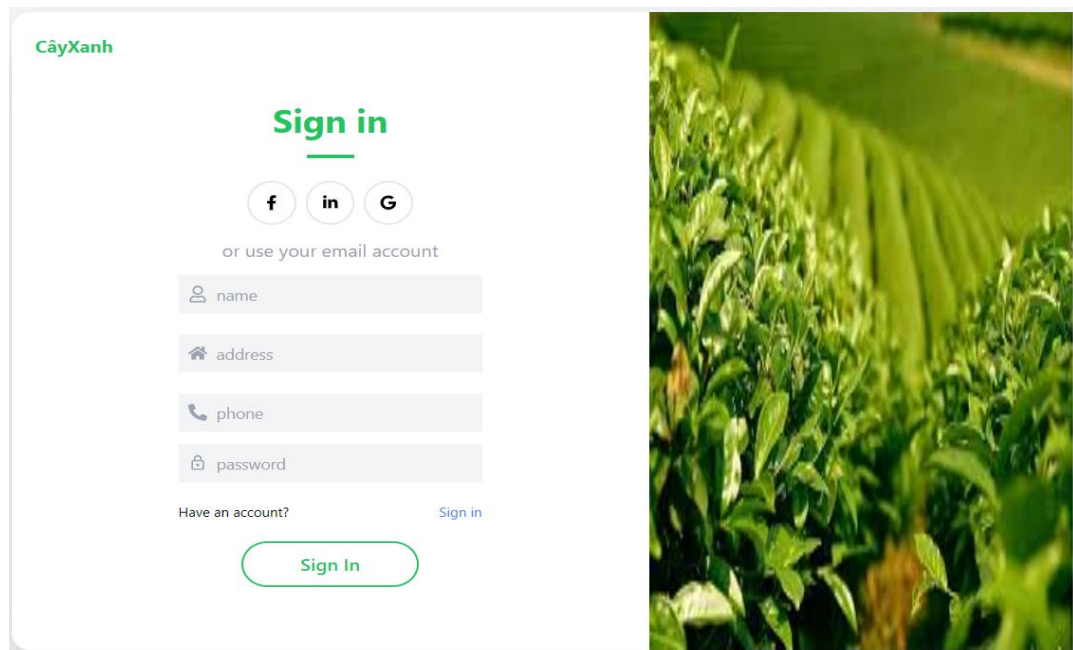
### 1.3 Giao diện đăng ký và đăng nhập

Ở góc bên phải của thanh điều hướng, người dùng thực hiện bấm vào nút “Đăng nhập” để thực đăng nhập tài khoản cá nhân. Hệ thống sẽ hiển thị biểu mẫu thông tin để người dùng nhập vào và nhấn “Sign in” để đăng nhập vào hệ thống (Hình 20).



Hình 20: Giao diện đăng nhập hệ thống

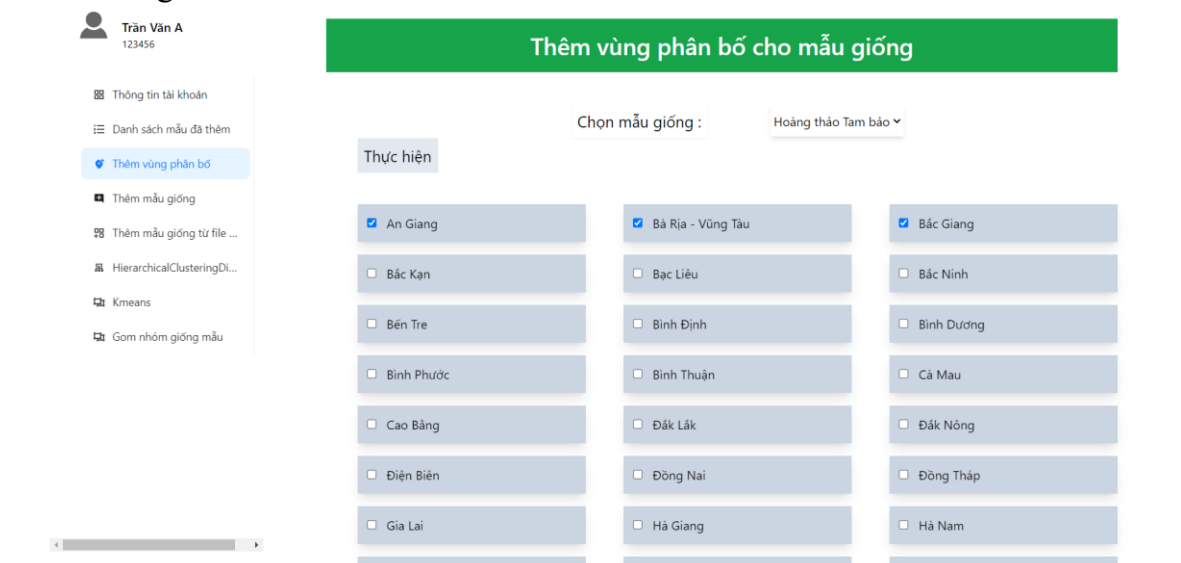
Nếu người dùng chưa có tài khoản, sẽ thực hiện bấm vào nút “Sign up” để hiển thị biểu mẫu đăng ký. Người dùng tiến hành điền các thông tin vào biểu mẫu và nhấn nút “Sign up” để đăng ký tài khoản (Hình 21).



Hình 21: Giao diện đăng ký hệ thống

#### 1.4 Giao diện chức năng Thêm vùng phân bố

Người dùng tiến hành chọn mẫu giống và các địa điểm phân bố của mẫu giống, sau đó bấm nút “Thực hiện” để cập nhật vùng phân bố của mẫu đó (Hình 22). Hiện tại hệ thống chỉ mới cài đặt 63 tỉnh thành ở Việt Nam.



Hình 22: Giao diện thêm vùng phân bố

#### 1.5 Giao diện chức năng Thêm mẫu giống

Với chức năng thêm mẫu giống thủ công, người dùng thực hiện chọn loại giống cây mình muốn thêm và điền đầy đủ các thông tin về mẫu giống cây trồng (Hình 23).



Thêm mẫu giống

Bạn muốn thêm mẫu giống nào?

Đầu nành

Tên giống: Đầu Nành An Giang

Tên gọi khác: Đầu Nành

Xuất xứ: An Giang

Xuất xứ (Tiếng Anh): An Giang

Màu hoa: trắng

Trục hạ diệp: tím

Màu vỏ hạt: vàng

Màu tế: nâu

Dạng hạt (dựa trên tỉ lệ rộng/dài và tỉ lệ dày/rộng của hạt): tủa cầu (0.95 và <=)

Màu trái: nâu hơi đậm

Hình 23: Giao diện thêm mẫu giống thủ công

Sau khi điền đầy đủ thông tin về mẫu giống cây, người dùng tiến hành nhấn nút “Thêm”. Mẫu giống cây mới này sẽ được đưa vào danh sách mẫu giống đang chờ duyệt của quản trị để người quản trị tiến hành xem xét duyệt. Mẫu giống này cũng được đưa vào Danh sách mẫu đã thêm để xem trạng thái duyệt của mẫu (Hình 24).

NGÂN HÀNG THÔNG TIN GIỐNG CÂY TRỒNG VÀ VẬT NUÔI

Danh sách giống đã thêm

Tên giống	Tên gọi khác	Xuất xứ	Chi tiết	Trạng thái
Đầu Nành	DN220	Ở Lâm		Đang chờ duyệt
Đầu 112	DN 112	Canh Ranh		Đã duyệt
1	chưa biết	chưa biết		Đã duyệt
2	chưa biết	chưa biết		Đã duyệt
Đầu Nành An Giang	Đầu Nành	An Giang		Đang chờ duyệt

Hình 24: Giao diện danh sách mẫu giống đã thêm

Ngoài ra người dùng còn có thể thêm nhiều mẫu giống mới một lượt bằng cách tải lên hệ thống tập tin Excel chứa các thông tin về các mẫu mới như Hình 25.

Trần Văn A  
123456

- Thông tin tài khoản
- Danh sách mẫu đã thêm
- Thêm vùng phân bố
- Thêm mẫu giống
- Thêm mẫu giống từ file ...
- HierarchicalClusteringDL...
- Kmeans
- Gom nhóm giống mẫu

### Thêm mẫu giống từ file excel

CHON FILE ĐỂ TẢI LÊN

Choose File ds\_lua.xlsx

Tải file

Xác nhận ✓

STT	N1	N2	N3	N4	N5	N4_En	N5_En	N6	N7	N8	N9	N10	N1
1	1	1	chưa biết	chưa biết	chưa biết	unknown	unknown	-1	-1	-1	-1	-1	1
2	2	2	chưa biết	chưa biết	chưa biết	unknown	unknown	-1	-1	-1	-1	-1	2
3	3	3	chưa biết	chưa biết	chưa biết	unknown	unknown	-1	-1	-1	-1	-1	2
4	4	4	chưa biết	chưa biết	chưa biết	unknown	unknown	-1	-1	-1	-1	-1	2

Hình 25: Giao diện thêm mẫu giống từ tập tin Excel

## 1.6 Giao diện chức năng Phân cụm phân cấp

Người dùng sẽ tải lên tập tin Excel chứa thông tin các mẫu giống cây trồng, hệ thống sẽ thực hiện xử lý và tính toán. Kết quả sẽ là hình ảnh cây phân cấp của các mẫu giống như Hình 26 sau đây.



## NGÂN HÀNG THÔNG TIN GIỐNG CÂY TRỒNG VÀ VẬT NUÔI

Trần Văn A  
123456

- Thông tin tài khoản
- Danh sách mẫu đã thêm
- Thêm vùng phân bố
- Thêm mẫu giống
- Thêm mẫu giống từ file ...
- HierarchicalClusteringDL...
- Kmeans
- Gom nhóm giống mẫu

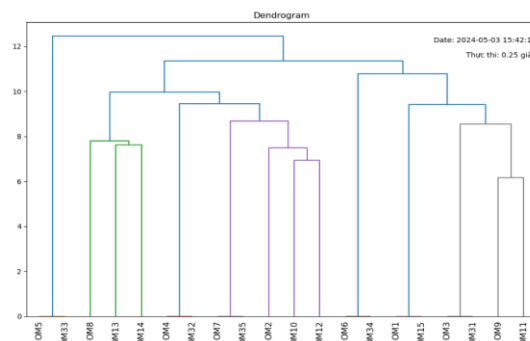
### Hierarchical Clustering Diagram

TẢI FILE LÊN Ở ĐÂY

Choose File text1.xlsx

Tải file

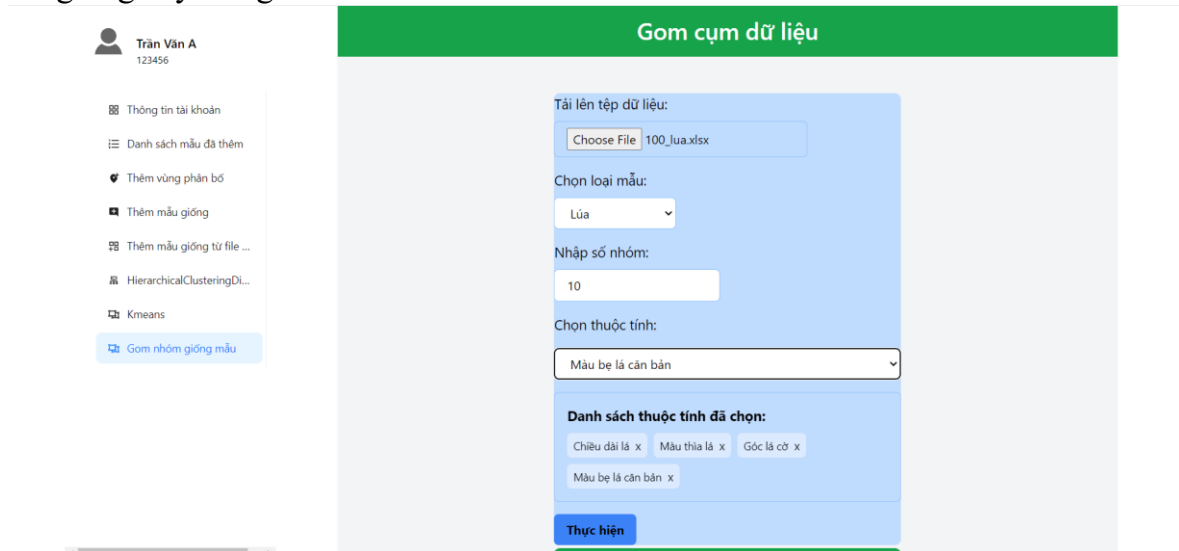
✓ Xác nhận



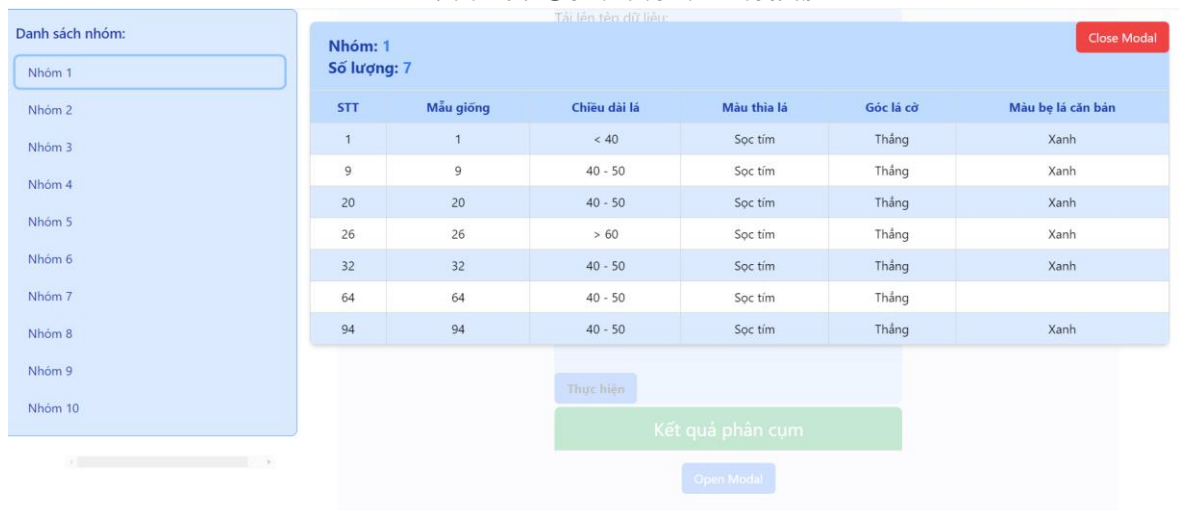
Hình 26: Giao diện phân cụm phân cấp

### 1.7 Giao diện gom nhóm với Kmeans

Tương tự như chức năng phân cụm phân cấp, người dùng tiến hành tải lên tập tin Excel chứa các thông tin mẫu giống cây trồng, sau đó điền đầy đủ thông tin về loại giống, số nhóm và các thuộc tính hình thái, nông sinh học mà mình muốn gom lại (Hình 27). Kết quả sau khi thực hiện sẽ là các cụm nhóm chứa các mẫu giống cây trồng như Hình 28.



Hình 27: Gom nhóm Kmeans



STT	Mẫu giống	Chiều dài lá	Màu thìa lá	Góc lá cờ	Màu bẹ lá cân bản
1	1	< 40	Sọc tím	Thẳng	Xanh
9	9	40 - 50	Sọc tím	Thẳng	Xanh
20	20	40 - 50	Sọc tím	Thẳng	Xanh
26	26	> 60	Sọc tím	Thẳng	Xanh
32	32	40 - 50	Sọc tím	Thẳng	Xanh
64	64	40 - 50	Sọc tím	Thẳng	Xanh
94	94	40 - 50	Sọc tím	Thẳng	Xanh

Hình 28: Kết quả gom nhóm Kmeans

### 1.8 Giao diện Duyệt mẫu giống mới

Sau khi người dùng đã thực hiện thêm giống cây trồng mới, các mẫu giống mới sẽ được đưa vào Danh sách giống đang chờ duyệt (Hình 29). Người quản trị sẽ tiến hành xem chi tiết mẫu giống và đưa ra xem xét về việc duyệt hay xóa mẫu giống ra khỏi danh sách. Nếu quản trị viên thực hiện duyệt thì mẫu giống sẽ được đưa vào danh sách mẫu giống hiển thị trên hệ thống và cập nhật lại trạng thái ở Danh sách thêm của người dùng. Ngược lại thì sẽ đưa vào danh sách mẫu giống đã xóa (Hình 30), tuy nhiên thì các mẫu giống đã xóa vẫn có thể được duyệt lại.



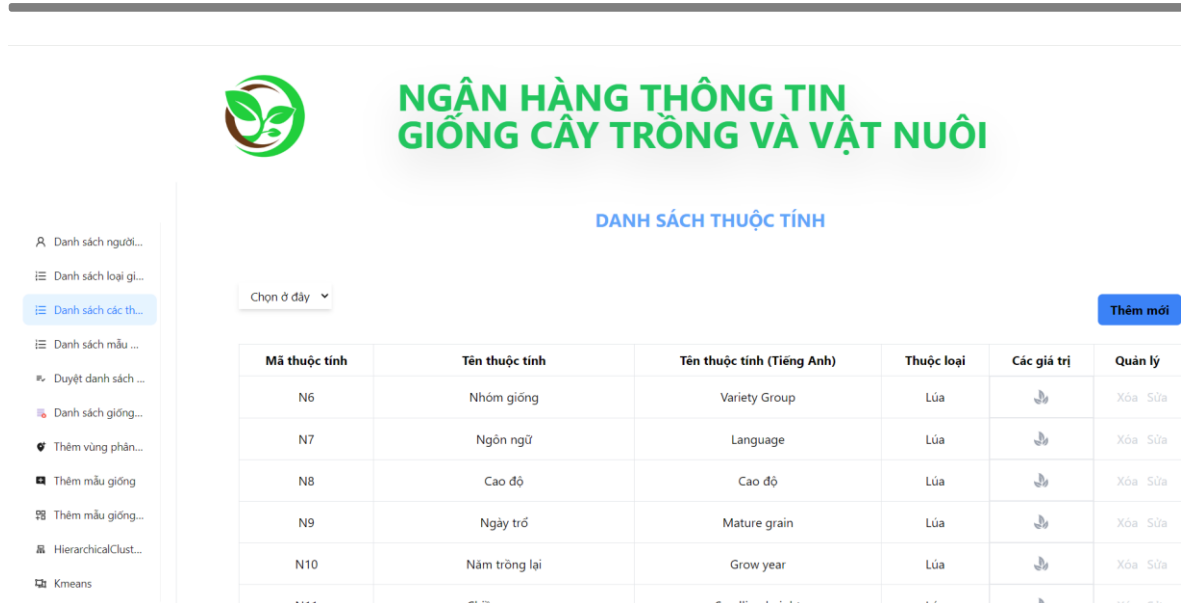
Hình 29: Giao diện danh sách giống đang chờ duyệt



Hình 30: Giao diện danh sách giống đã xóa

### 1.9 Giao diện quản lý các thuộc tính

Với sự linh hoạt của cơ sở dữ liệu, quản trị viên sẽ dễ dàng thực hiện sửa, xóa cũng như thêm mới một đặc tính hình thái, nông sinh học của loại giống nào đó. Giao diện thực hiện sẽ như Hình 31 ở dưới đây:



Hình 31: Quản lý các thuộc tính

### 1.10 Giao diện quản lý danh sách loại giống

Quản trị viên có thể thực hiện việc quản lý danh sách loại giống như thêm, sửa hoặc xóa một loại giống ở trong danh sách. Giao diện thực hiện như Hình 32 dưới đây:



Hình 32: Giao diện quản lý danh sách loại giống

### 1.11 Giao diện trên thiết bị di động

Ngoài ứng dụng trên website, “Ngân hàng thông tin giống cây trồng ở Việt Nam” cũng được xây dựng trên ứng dụng thiết bị di động. Tuy nhiên thì ở trên thiết bị di động thì chỉ có các chức năng cơ bản của người dùng chưa đăng ký, vẫn chưa hoàn thiện các chức năng mà người dùng có tài khoản đã có như trên website.

- Trang chủ trên thiết bị di động như Hình 33 dưới đây:



Hình 33: Giao diện trang chủ trên thiết bị di động  
- Giao diện danh sách mẫu giống (Hình 34):



Hình 34: Giao diện danh sách mẫu giống trên thiết bị di động

- Giao diện chi tiết thông tin về mẫu giống cây trồng khi người dùng bấm vào hình ảnh mẫu (Hình 35).



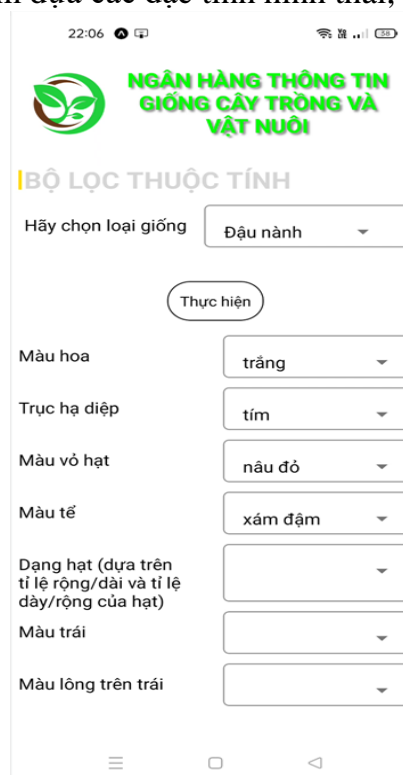
**NGÂN HÀNG THÔNG TIN GIỐNG CÂY TRỒNG VÀ VẬT NUÔI**

**BẢNG CHI TIẾT VỀ ĐẶC ĐIỂM HÌNH THÁI VÀ NÔNG HỌC CỦA MẪU GIỐNG**

Mẫu giống :	Tên trước kia :	Xuất xứ :
NGO13	NGO13	Trung Quốc
Đặc tính cụ thể của mẫu giống		
Sức sống cây mầm	Tốt	
Sắc tố antoxian của bề lá thứ nhất	Không có/rất nhạt	
Hình dạng phần đỉnh lá thứ nhất	Nhọn	
Mức xanh của bộ lá	Trung bình	
Sự gợn sóng của mép lá	Không có/rất ít	
Góc giữa phiến lá và thân	Trung bình	
Mức độ cong của phiến lá	Cong	
Chiều rộng phiến lá (lá trên liền kề với bắp trên cùng)	Hẹp	

Hình 35: Giao diện chi tiết thông tin giống cây trên thiết bị di động

- Giao diện tìm kiếm dựa các đặc tính hình thái, nông sinh học (Hình 36):



**NGÂN HÀNG THÔNG TIN GIỐNG CÂY TRỒNG VÀ VẬT NUÔI**

**BỘ LỌC THUỘC TÍNH**

Hãy chọn loại giống: Đậu nành

Thực hiện

Màu hoa: trắng

Trục hạ diệp: tím

Màu vỏ hạt: nâu đỏ

Màu tế: xám đậm

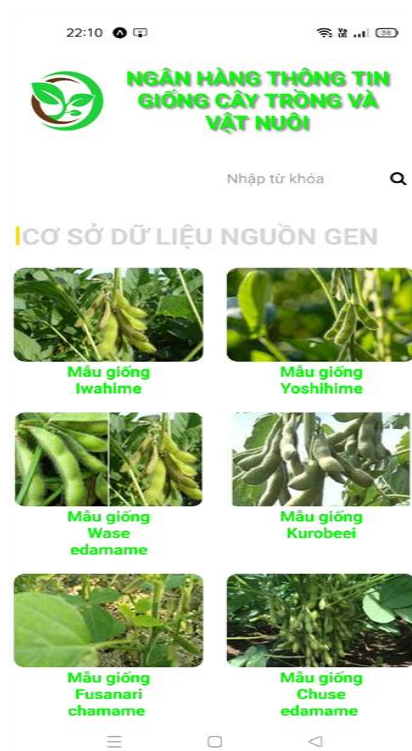
Dạng hạt (dựa trên tỉ lệ rộng/dài và tỉ lệ dày/rộng của hạt):

Màu trái:

Màu lông trên trái:

Hình 36: Giao diện tìm kiếm dựa trên các đặc tính hình, nông sinh học trên thiết bị di động

Và kết quả sau khi thực hiện như Hình 37 sau:



Hình 37: Giao diện kết quả tìm kiếm trên thiết bị di động

## 2. Kết quả kiểm thử

### 1.1 Mục đích kiểm thử

- Tìm các lỗi, sai sót trong chương trình, giúp nhận diện và ghi nhận các lỗi này để có thể sửa chữa trước khi triển khai.
- Giúp đánh giá mức độ chất lượng của sản phẩm và ngăn ngừa các lỗi có thể xảy ra trong tương lai.
- Kiểm tra hệ thống hoạt động đúng với những yêu cầu trong tài liệu đặc tả.
- Làm tài liệu cho giai đoạn bảo trì.

### 1.2 Phạm vi kiểm thử

Kiểm thử một vài chức năng chính trong tài liệu đặc tả và tài liệu thiết kế nhằm đảm bảo hệ thống đáp ứng được các yêu cầu tối thiểu đặt ra.

### 1.3 Kế hoạch kiểm thử

#### 1.3.1 Các chức năng kiểm thử

- Tìm kiếm
- Tìm kiếm theo thuộc tính
- Thêm giống cây mới
- Duyệt danh sách thêm giống mới
- Quản lý loại giống
- Quản lý thuộc tính giống cây trồng

#### 1.3.2 Cách tiếp cận

Hệ thống sẽ được kiểm thử theo cách truyền thống. Mỗi chức năng sẽ được kiểm thử theo thứ tự từ trên xuống.

#### 1.3.3 Tiêu chí đánh giá kiểm thử thành công/thất bại



- Tiêu chí kiểm thử thành công khi kết quả kiểm thử giống hoặc tương tự với kết quả mong muốn.
- Tiêu chí kiểm thử thất bại khi kết quả kiểm thử khác hoàn toàn với kết quả mong muốn.

#### 1.3.4 Tiêu chí đình chỉ và yêu cầu bắt đầu lại

- Tiêu chí đình chỉ khi có một chức năng thông báo lỗi trong quá trình thực hiện kiểm thử.
- Tiêu chí bắt đầu lại khi chức năng bị đình chỉ được sửa lỗi.

### 1.4 Quản lý kiểm thử

#### 1.4.1 Quá trình kiểm thử

- Lập kế hoạch kiểm thử
- Tạo các trường hợp kiểm thử
- Tiến hành kiểm thử
- Ghi nhận, đánh giá và báo cáo kết quả kiểm thử

#### 1.4.2 Môi trường kiểm thử

- **Phần cứng:**
  - o Vi xử lý: Intel(R) Core(TM) i7-7500U
  - o Ram: 8GB
  - o Cấu hình mạng: có kết nối Internet
- **Phần mềm:**
  - o Hệ điều hành Windows 10
  - o Công cụ lập trình: Visual Studio Code
  - o Hệ quản trị CSDL: MySQL
  - o Trình duyệt kiểm thử: Google Chrome

### 1.5 Các chức năng kiểm thử

#### 1.5.1 Chức năng tìm kiếm

**Bảng 24: Kiểm thử chức năng tìm kiếm**

Trường hợp	Kết quả kỳ vọng	Kết quả kiểm thử	Tình trạng
Để trống dữ liệu cần nhập	Thông báo: không tìm thấy mẫu giống	Thông báo: không tìm thấy mẫu giống	Đạt
Nhập chuỗi không phải tên giống cây	Thông báo: không tìm thấy mẫu giống	Thông báo: không tìm thấy mẫu giống	Đạt
Nhập đúng tên mẫu giống	Thông báo: hiển thị thông tin chi tiết của mẫu giống	Thông báo: hiển thị thông tin chi tiết của mẫu giống	Đạt

#### 1.5.2 Chức năng tìm kiếm dựa trên các đặc tính

**Bảng 25: Kiểm thử chức năng tìm kiếm dựa trên các thuộc tính**

Trường hợp	Kết quả kỳ vọng	Kết quả kiểm thử	Tình trạng
Để trống các giá trị thuộc tính	Danh sách mẫu giống rỗng	Danh sách mẫu giống rỗng	Đạt

Chọn các giá trị thuộc tính (lần 1)	Hiện thị danh sách mẫu giống phù hợp	Hiện thị danh sách mẫu giống phù hợp	Đạt
Chọn các giá trị thuộc tính (lần 2)	Danh sách mẫu giống rỗng	Danh sách mẫu giống rỗng	Đạt

### 1.5.3 Duyệt danh sách thêm giống mới

**Bảng 26: Kiểm thử chức năng duyệt danh sách thêm giống mới**

Trường hợp	Kết quả kỳ vọng	Kết quả kiểm thử	Tình trạng
Bấm duyệt một mẫu giống	Thông báo: cập nhật thành công.	Thông báo: cập nhật thành công.	Đạt
Bấm xóa một mẫu giống	Thông báo: cập nhật thành công. Chuyển mẫu giống sang danh sách đã xóa.	Thông báo: cập nhật thành công. Chuyển mẫu giống sang danh sách đã xóa, phù hợp	Đạt

### 1.5.4 Quản lý loại giống

**Bảng 27: Kiểm thử chức năng quản lý loại giống**

Trường hợp	Kết quả kỳ vọng	Kết quả kiểm thử	Tình trạng
Thêm một loại giống cây mới	Thêm thành công	Thêm thành công	Đạt
Sửa thông tin một loại giống	Cập nhật thông tin thành công	Cập nhật thông tin thành công	Đạt
Xóa một loại giống cây trồng	Xóa loại giống ra khỏi danh sách	Xóa loại giống ra khỏi danh sách	Đạt

### 1.5.5 Quản lý thuộc tính giống cây trồng

**Bảng 28: Kiểm thử chức năng quản lý thuộc tính giống cây trồng**

Trường hợp	Kết quả kỳ vọng	Kết quả kiểm thử	Tình trạng
Thêm một đặc tính cho một loại giống	Thêm thành công	Thêm thành công	Đạt
Sửa thông tin một thuộc tính	Cập nhật thông tin thành công	Cập nhật thông tin thành công	Đạt
Xóa một thuộc tính cây trồng	Xóa thuộc tính ra khỏi danh sách	Xóa thuộc tính ra khỏi danh sách	Đạt

---

## PHẦN KẾT LUẬN

### 1. Kết quả đạt được

"Ngân hàng thông tin các giống cây trồng ở Việt Nam" đã tuân thủ đúng quy trình phát triển phần mềm và đáp ứng các yêu cầu chức năng cơ bản đã được đề ra. Hệ thống này không chỉ cung cấp khả năng quản trị hệ thống thông qua việc cấp tài khoản người dùng và quyền truy cập, mà còn cho phép người dùng tra cứu thông tin về các giống cây trồng theo các đặc tính mong muốn.

Đặc biệt, việc tích hợp giải thuật Kmeans và phân cụm phân cấp giúp người dùng dễ dàng gom nhóm các giống cây dựa vào các đặc tính mong muốn. Sự linh hoạt trong việc xây dựng ứng dụng trên cả website và thiết bị di động cũng giúp tăng tính tiện dụng và sẵn sàng của hệ thống, đồng thời đáp ứng nhu cầu tra cứu của người dùng một cách linh hoạt và hiệu quả. Như vậy, "Ngân hàng thông tin các giống cây trồng ở Việt Nam" không chỉ là một công cụ hữu ích cho cộng đồng nghiên cứu và người quản lý trong lĩnh vực nông nghiệp, mà còn đóng góp tích cực vào sự phát triển các công nghệ áp dụng vào lĩnh vực nông nghiệp sau này.

### 2. Hạn chế

- Với ngân hàng thông tin các giống cây trồng thì nguồn dữ liệu về các đặc tính hình thái, đặc tính nông sinh học của các giống cây là quan trọng. Do các dữ liệu ở các tổ chức đều được bảo mật nên việc thu thập dữ liệu còn hạn chế, thông tin mà hệ thống cung cấp về số lượng loại giống không được đa dạng, phong phú.

- Hiện nay, tiêu chuẩn về các đặc tính hình thái, đặc tính nông sinh học mà các tổ chức đưa ra vẫn chưa có sự thống nhất nên thông tin trên hệ thống chỉ là sự tổng hợp khách quan.

- Các chức năng ở trên thiết bị di động vẫn chưa được hoàn thiện như trên website. Giao diện hệ thống còn đơn giản và thiếu thẩm mỹ, không đảm bảo sự hấp dẫn và thu hút người dùng.

### 3. Hướng phát triển

- Thực hiện thu thập thêm dữ liệu về các loại giống cây trồng để tăng tính đa dạng, giúp người dùng có nhiều thông tin tra cứu hơn.

- Hoàn thiện các chức năng trên thiết bị di động, cải thiện giao diện người dùng để tạo ra trải nghiệm sử dụng tốt hơn và tăng cường sự tương tác giữa người dùng và hệ thống.

- Ngoài dữ liệu thông tin các giống cây trồng, có thể phát triển thêm với các nguồn gen tài nguyên khác như các giống vật nuôi,...

- Mở rộng các chức năng phân tích dữ liệu, bao gồm việc sử dụng các thuật toán máy học và học sâu để phân tích và dự đoán các xu hướng trong ngành nông nghiệp.

---

## TÀI LIỆU THAM KHẢO

- [1]. Vũ Thị Thu Hiền, 2012. Đa dạng di truyền dựa trên đặc điểm hình thái của các mẫu giống lúa có nguồn gốc khác nhau. Tạp chí Khoa học và Phát triển. Tập 10, số 6: 844–852.
- [2]. Đoàn Thị Thùy Linh và Nguyễn Văn Khoa, 2013. Đa dạng di truyền một số mẫu giống lúa địa phương vùng Tây Bắc dựa trên đặc điểm hình thái. Hội nghị khoa học toàn quốc về sinh thái và tài nguyên sinh vật lần thứ 5, 18/10/2013, Hà Nội, Việt Nam. Nhà xuất bản Nông nghiệp. Hà Nội, 1132–1139.
- [3]. Lưu Tiến Đạo, Âu Tấn Tài, Vũ Anh Pháp và Trần Nguyễn Minh Thư, 2017. Ứng dụng giải thuật gom nhóm dữ liệu để nhận diện sự tương đồng giữa các giống lúa. Tạp chí Khoa học Trường Đại học Cần Thơ. 140-144.
- [4]. I. R. R. Institute., “Standard evaluation system for rice.” IRRI, International Rice Research Institute, Manila, Philippines, 1996.
- [5]. M. Takeya, F. Yamasaki, S. Uzuhashi, T. Aoki, H. Sawada, T. Nagai, K. Tomioka, N. Tomooka, T. Sato, and M. Kawase, “NIASGBdb: NIAS Genebank databases for genetic resources and plant disease information,” Nucleic Acids Res., vol. 39, no. Database, pp. D1108–D1113, Jan. 2011.
- [6]. Đoàn Thanh Quỳnh, Nguyễn Thị Hảo, Vũ Thị Thu Hiền và Trần Văn Quang, 2016. Đánh giá đa dạng di truyền nguồn gen lúa nếp địa phương dựa trên kiểu hình và chỉ thị phân tử. Tạp chí Khoa học Nông nghiệp Việt Nam. Tập 14, số 4: 527–538.
- [7]. Lương Thế Anh, Nguyễn Thái Nghe và Nguyễn Chí Ngôn, 2014. Xây dựng hỗ trợ khuyến nông trên cây lúa qua mạng thông tin di động. Tạp chí Khoa học Trường Đại học Cần Thơ. 33: 9-21.
- [8]. Nguyễn Lộc Hiền, Trần Thanh Xuyên, Trần Thị Bích Phương và Tadashi Yoshihashi, 2010. Sự đa dạng di truyền của các giống đậu nành rau Nhật Bản. Tạp chí Khoa học Trường Đại học Cần Thơ. 16a: 51–59.
- [9]. “Nhận định tình hình ENSO, hạn mặn khu vực Tây Nguyên và Đồng bằng sông Cửu Long năm 2024”, Trang thông tin điện tử Tổng cục Khí tượng Thủy Văn, ngày 08 tháng 04 năm 2024. Truy cập tại: <http://vmha.gov.vn/du-bao-107/nhan-dinh-tinh-hinh-enso-han-man-khu-vuc-tay-nguyen-va-dong-bang-song-cuu-long-nam-2024-16614.html>