

## TỔNG HỢP KẾT QUẢ

1. Thành viên: Nguyễn Văn Tuấn Đạt – 19120472

Khó khăn gặp phải:

- Chưa làm quen với sử dụng git, github.

Học được:

- Cách sử dụng git, github.
- Ứng dụng quy trình khoa học dữ liệu để làm đồ án.
- Nắm thêm các mô hình học máy.
- Các kĩ thuật tiền xử lí trước khi đưa vào mô hình học máy.

2. Thành viên: Hoàng Mạnh Khiêm – 19120543

Khó khăn gặp phải:

- Còn đôi chút khó khăn khi sử dụng Git và GitHub.
- Lần đầu sử dụng Selenium còn chưa quen.

Học được:

- Cách sử dụng Git, GitHub.
- Quy trình tiền xử lí, khám phá, trực quan hóa dữ liệu.
- Các mô hình học máy và tiền xử lí dữ liệu trước khi đưa vào mô hình.

3. Thành viên: Hà Nguyễn Thảo Vy - 20120237

Khó khăn gặp phải:

- Chưa biết cách sử dụng git.
- Tốn nhiều thời gian crawl dữ liệu.
- Câu hỏi bản thân đặt ra còn đơn giản.

Học được:

- Cách crawl dữ liệu hiệu quả.
- Các kĩ thuật khám phá và tiền xử lí dữ liệu cho mô hình.
- Học được thêm các mô hình.

4. Nhóm sẽ làm gì nếu có thêm thời gian:

- Thử nghiệm các mô hình deep learning.
- Xử lí cột “Weather Type” thành các nhóm categorical nhỏ hơn thay vì dùng one-hot vector.

# BẢNG PHÂN CÔNG CHI TIẾT

- Link Notion: <https://spark-handbell-b2d.notion.site/b5456d90866f48a49870239e0edb2423?v=3d3c31cf17c048ef8f8cf64c42764247>

Hide description

## Project Plans

Project Introduction to Data Science  
Topic: Temperature Prediction Ho Chi Minh City

All Lessons By Status +

Filter Sort 🔍 ... New

Name	Content	Date	Status	Job Recipient	Github	Note	
Stage 1	Crawl Data from web (A)	November 15, 2022 → November 26, 2022	Complete	Vy, Khiêm	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.git">https://github.com/tuandat-github/Project-Final-Introduction--DS.git</a>	3/12, cập nhật và sửa chữa file crawl data	
Stage 2	Part 1 (B, C, D)	November 26, 2022 → December 12, 2022	Complete	Hào, Vy, Đạt	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.git">https://github.com/tuandat-github/Project-Final-Introduction--DS.git</a>		
Stage 3	Create Model and Evaluation	November 28, 2022 → December 6, 2022	Complete	Đạt	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.git">https://github.com/tuandat-github/Project-Final-Introduction--DS.git</a>		
Stage 4	Write report, Write breakdown	December 8, 2022 → December 12, 2022	Complete	Vy, Đạt			

+ New

## 2. Stage 1

### Stage 1

Content	Crawl Data from web (A)
Date	November 15, 2022 → November 26, 2022
Github	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.git">https://github.com/tuandat-github/Project-Final-Introduction--DS.git</a>
Job Recipient	Vy, Khiêm
Note	3/12, cập nhật và sửa chữa file crawl data
Status	Complete
+ Add a property	

Add a comment...

Job detail	Performer
Thu thập các trường dữ liệu: Loại thời tiết, nhiệt độ trung bình, nhiệt độ cao nhất, nhiệt độ thấp nhất, tốc độ gió.	Khiêm
Thu thập các trường dữ liệu: lượng mưa, độ ẩm, độ che phủ của mây, áp suất không khí. Lưu dữ liệu thu thập dưới định dạng CSV.	Vy
Chỉnh sửa phần thu thập bị sai sót và đưa file lưu trữ dữ liệu vào thư mục. Thêm file mô tả thông tin các trường dữ liệu.	Khiêm

### 3. Stage 2

#### Stage 2

Content	Part 1 (B, C, D)
Date	November 26, 2022 → December 12, 2022
Github	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.git">https://github.com/tuandat-github/Project-Final-Introduction--DS.git</a>
Job Recipient	Hào, Vy, Đạt
Note	Empty
Status	Complete
+ Add a property	

 Add a comment...

↔ Options ▾ ⋮

Job detail	Performer
Doing Part B	Đạt
Doing Part C	Vy, Khiêm
Complete Part C and B in file notebook	Đạt

### 4. Stage 3


#### Stage 3

Content	Create <u>Model</u> and Evaluation
Date	November 28, 2022 → December 6, 2022
Github	<a href="https://github.com/tuandat-github/Project-Final-Introduction--DS.gi">https://github.com/tuandat-github/Project-Final-Introduction--DS.gi</a>
Job Recipient	Đạt
Note	Empty
Status	Complete
+ Add a property	

 Add a comment...








Job detail	Performer
Doing Part 2: Modelling Data and Evaluation Data	Đạt

5. Stage 4

 Add cover



Stage 4

-  Content
- Write report, Write breakdown
-  Date
- December 8, 2022 → December 12, 2022
-  Github
- Empty
-  Job Recipient
- Vy, Đạt
-  Note
- Empty
-  Status
- Complete
-  Add a property

 Add a comment



Options



Job detail	Performer
Write slide report	Vy
Write breakdown	Đạt