**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

**TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**ĐỒ ÁN**

**HỌC PHẦN MÁY HỌC ỨNG DỤNG**

**Đề tài**

**ƯỚC LƯỢNG MỨC ĐỘ DI CHUYỂN CỦA CÁC PHƯƠNG TIỆN GIAO THÔNG VỚI CENTERNET**

**Nhóm sinh viên thực hiện:**

**1. Nguyễn Sinh Viên - B1709590**

**2.**

**3.**

**4.**

#### Cần Thơ, …/….

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

**TRƯỜNG CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**ĐỒ ÁN**

**HỌC PHẦN MÁY HỌC ỨNG DỤNG**

**Đề tài**

**ƯỚC LƯỢNG MỨC ĐỘ DI CHUYỂN CỦA CÁC**

**PHƯƠNG TIỆN GIAO THÔNG VỚI CENTERNET**

**Giảng viên hướng dẫn: Nhóm sinh viên thực hiện:**

**1. Nguyễn Văn A – B2008603**

**2.**

**3.**

**4.**

#### Cần Thơ, …./…..

**NHẬN XÉT CỦA GIẢNG VIÊN**

---------------------------------------------------------------------------------

---------------------------------------------------------------------------------

-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

-----------------------------------------------------------------------------------------------------------------------------------------------------------------

---------------------------------------------------------------------------------

-----------------------------------------------------------------------------------------------------------------------------------------------------------------

---------------------------------------------------------------------------------

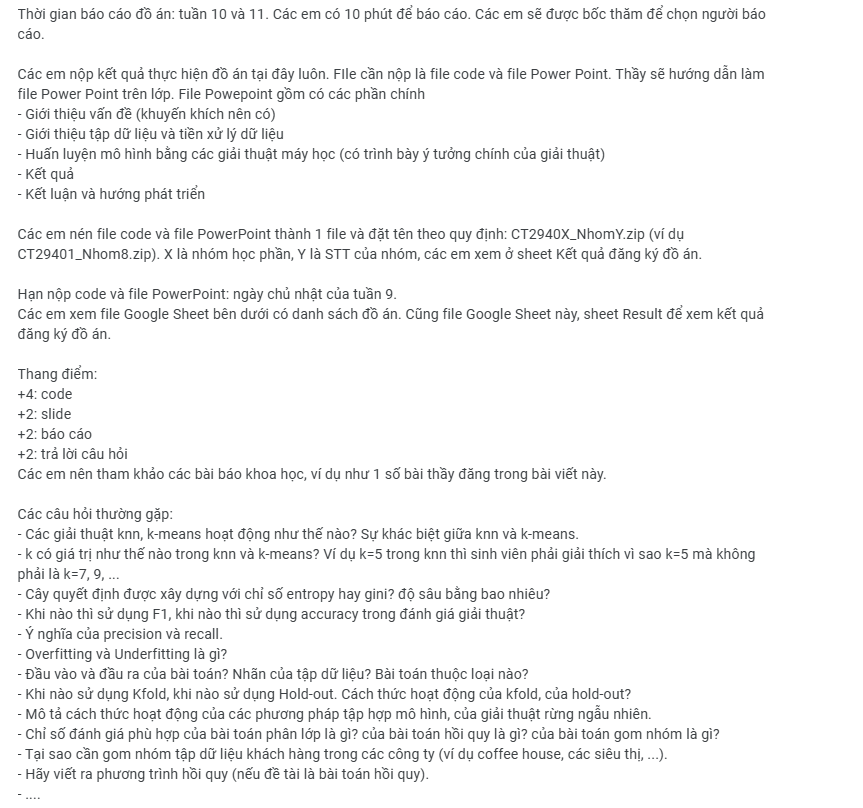
*Cần Thơ, ngày tháng năm*

(Ký và ghi rõ họ tên)

**MỤC LỤC**

**PHÂN CÔNG CÔNG VIỆC**

**PHẦN NỘI DUNG**

****

## Mô tả bài toán

**\*Tên bài toán:** Dự đoán năng lượng tiêu thụ của thiết bị trong nhà.

**\*Bối cảnh bài toán (có thể đổi thành giới thiệu bài toán):** Trong bối cảnh nhà thông minh (smart home), việc theo dõi và dự đoán mức tiêu thụ năng lượng của các thiết bị gia dụng (Appliances) có ý nghĩa quan trọng để:  
 - Tối ưu hóa tiêu thụ điện năng

- Giảm hóa đơn tiền điện

- Cảnh báo sớm các thiết bị tiêu tốn năng lượng bất thường

- Tăng hiệu quả quản lý năng lượng  
Bài toán này sử dụng dữ liệu cảm biến được thu thập từ nhiều vị trí trong ngôi nhà và cả các điều kiện thời tiết bên ngoài để dự đoán mức năng lượng tiêu thụ (Wh) của các thiết bị trong nhà.

**\*Mục tiêu bài toán:** Dự đoán chính xác giá trị tiêu thụ năng lượng thiết bị tại một thời điểm cụ thể dựa trên dữ liệu cảm biến trong và ngoài nhà.

**\*Loại bài toán Machine Learning**: Đây là bài toán hồi quy (regression) vì biến mục tiêu (Appliances) là giá trị liên tục.

**\*Đầu vào (Input)**:  
 - Nhiệt độ và độ ẩm từ nhiều phòng trong nhà (T1 … T9, RH\_1 … RH\_9)  
 - Điều kiện thời tiết bên ngoài (T\_out, RH\_out, Windspeed, Visibility, Tdewpoint, Press\_mm\_hg)  
 - Thời gian (date)  
 - Lượng điện tiêu thụ cho hệ thống chiếu sáng (lights)  
 - Các biến ngẫu nhiên (nếu cần): rv1, rv2

**\*Đầu ra (Output):** Một giá trị duy nhất là sản lượng tiêu thụ năng lượng của thiết bị (Appliances) tại thời điểm đó, tính bằng Wh.

**\*Chỉ số đánh giá mô hình:** - MAE (Mean Absolute Error): Sai số tuyệt đối trung bình.  
 - RMSE (Root Mean Square Error): Căn bậc 2 của sai số bình phương trung bình.  
 - R2 Score (Coefficient of Determination): Hệ số xác định.

## 2. Mô tả dữ liệu, ý nghĩa của dữ liệu

## \*Thông tin thời gian và tiêu thụ điện:

## - date: Thời gian ghi nhận dữ liệu (định dạng: yyyy-mm-dd HH:MM:SS). - Appliances: Mức tiêu thụ năng lượng (Wh) của các thiết bị trong nhà. - lights: Mức tiêu thụ năng lượng (Wh) của hệ thống đèn.

\*Nhiệt độ và độ ẩm các khu vực:

## - T1 -> T9: Nhiệt độ trong các phòng. - RH\_1 -> RH\_9: Độ ẩm trong các phòng.

## \*Điều kiện môi trường bên ngoài:

## - T\_out: Nhiệt độ bên ngoài trời (oC). - Press\_mm\_hg: Áp suất trong không khí (mmHg). - RH\_out: Độ ẩm tương đối ngoài trời (%). - Windspeed: Tốc độ gió (m/s). - Visibility: Tầm nhìn xa (km). - Tdewpoint: Nhiệt độ điểm sương (oC).

## \*Các biến ngẫu nhiên phục vụ mô hình hóa:

## - rv1, rv2: Biến ngẫu nhiên nhân tạo – có thể dùng để kiểm tra ảnh hưởng của nhiễu dữ liệu trong mô hình học máy

## 3. Phân tích dữ liệu và lựa chọn mô hình

- Dữ liệu không có giá trị bị thiếu. Các đặc trưng liên quan đến nhiệt độ, độ ẩm, thời tiết và thời gian đều có mối quan hệ tiềm năng đến mức tiêu thụ điện năng. Vì biến mục tiêu (Appliances) là giá trị liên tục, bài toán thuộc loại hồi quy (regression).

- Sau khi phân tích tương quan, nhận thấy nhiều biến thời tiết và điều kiện môi trường bên trong nhà có tương quan nhất định với mức tiêu thụ điện, đặc biệt là lights, T1, RH\_1, hour, T\_out,...

=> Do đó, các mô hình học máy được lựa chọn là Linear Regression (hồi quy tuyến tính), <abc>, <xyz> để bắt đầu kiểm thử mô hình cơ bản.

1. **Cấu hình máy tính**

- Hệ điều hành: Windows 11

- RAM: 16 GB

- CPU: Intel Core i5-12th

- Phần mềm: Python 3.11, thư viện pandas, scikit-learn, numpy

**5. Huấn luyện mô hình**

\*Tiền xử lý:

- Trích xuất thông tin thời gian từ cột date (hour, day, weekday)  
 - Chuẩn hóa dữ liệu (StandardScaler)  
 - Loại bỏ các cột không cần thiết như rv1, rv2

\*Chia dữ liệu: 80% để huấn luyện, 20% để kiểm tra.

\*Mô hình sử dụng:  
 - Hồi quy tuyến tính (Linear Regression)  
 - abc  
 - xyz

**6. Đánh giá mô hình**

**6.1 Đánh giá mô hình phân lớp / Regression / Cluster**

**\*Linear Regression:**

- MAE: 52.58 Wh

- RMSE: 91.11 Wh  
 - R² Score: 0.17

\*abc:

-  
 -  
 -  
\*xyz:

-  
 -  
 -

**6.2. Nhận xét kết quả thực nghiệm**

- Mô hình Linear Regression đưa ra sai số trung bình khoảng 52.58 Wh.

- Tuy nhiên, hệ số xác định R2=0.17 khá thấp → mô hình mới giải thích được khoảng 17% phương sai trong dữ liệu.

- Điều này cho thấy Linear Regression là một mô hình khởi đầu đơn giản, cần thử thêm các mô hình phi tuyến mạnh hơn như Random Forest, XGBoost, hoặc LightGBM để tăng độ chính xác.

//này là nhận xét riêng cho linear regression

**PHẦN KẾT LUẬN**

## 1. Kết quả đạt được

## 2. Hướng phát triển

# TÀI LIỆU THAM KHẢO

[1].

[2].