**BỘ GIÁO DỤC VÀ ĐÀO TẠO**



**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM**

**BÁO CÁO CUỐI KỲ**

**PHƯƠNG PHÁP THỰC HIỆN**

**DỰ ĐOÁN ĐIỆN NĂNG TIÊU THỤ CỦA CÁC**

**HỘ GIA ĐÌNH**

# MÃ MÔN HỌC: TRÍ TUỆ NHÂN TẠO

# Lớp: CDT25A

## GVHD: TS. Bùi Hà Đức

Tp. Hồ Chí Minh, tháng 09 năm 2025

# DANH SÁCH NHÓM THAM GIA

Tên đề tài: Phương pháp thực hiện dự đoán điện năng tiêu thụ của các hộ gia đình

|  |  |  |
| --- | --- | --- |
| STT | Họ và Tên | MSSV |
| 1 | Nguyễn Bá Thế | 2530905 |
| 2 | Trần Phước Khang | 2530903 |

Nhận xét của giáo viên:

..................................................................................................................................................

..................................................................................................................................................

..................................................................................................................................................

..................................................................................................................................................

*Ngày 27 tháng 9 năm 2025 Giáo viên chấm điểm*

# PHẦN 1: GIỚI THIỆU

# Trong thời đại chuyển đổi số, việc dự đoán điện năng tiêu thụ của hộ gia đình đóng vai trò quan trọng trong quản lý năng lượng và quy hoạch lưới điện. Báo cáo này trình bày quá trình xây dựng và so sánh ba mô hình dự báo chuỗi thời gian, bao gồm ARIMA, SARIMA và Prophet, trên bộ dữ liệu điện năng tiêu thụ hộ gia đình tại London. Kết quả thực nghiệm cho thấy mô hình Prophet đạt hiệu quả cao nhất. Qua đó, báo cáo khẳng định khả năng áp dụng các mô hình học máy vào quản lý năng lượng, đồng thời đề xuất hướng nghiên cứu mở rộng với dữ liệu thời tiết và đặc điểm nhân khẩu học để nâng cao độ chính xác.

# 1.1 Bối cảnh

# Hiện nay, tại Việt Nam và nhiều quốc gia trên thế giới, các công ty điện lực đã triển khai hệ thống đồng hồ điện thông minh, giúp thu thập dữ liệu điện năng tiêu thụ theo thời gian thực. Tuy nhiên, việc khai thác dữ liệu này để phục vụ dự báo và tối ưu hóa vận hành hệ thống điện vẫn chưa được chú trọng đúng mức.

# 1.2 Vấn đề nghiên cứu

# Làm thế nào để dự đoán trước lượng điện năng tiêu thụ của từng hộ gia đình trong khoảng thời gian từ 1 tuần đến 1 tháng, dựa trên dữ liệu lịch sử, nhằm:

# Hỗ trợ công tác điều độ điện năng.

# Giúp hộ gia đình quản lý và tối ưu hóa mức tiêu thụ.

# Cung cấp nền tảng cho các dịch vụ thông minh trong tương lai.

# 1.3 Mục tiêu nghiên cứu

# Xây dựng mô hình AI dự báo điện năng tiêu thụ.

# So sánh hiệu quả giữa các mô hình ARIMA, SARIMA và Prophet.

# Đánh giá độ chính xác thông qua các chỉ số MAE, RMSE, MAPE.

# (Tuỳ chọn) Khảo sát thêm tác động của yếu tố thời tiết hoặc nhân khẩu học.

# PHẦN 2. CƠ SỞ LÝ THUYẾT

# 2.1 Chuỗi thời gian

# Chuỗi thời gian là tập hợp dữ liệu được sắp xếp theo thứ tự thời gian, thường có tính mùa vụ, xu hướng và nhiễu. Dự báo chuỗi thời gian là một trong những bài toán quan trọng trong khoa học dữ liệu.

# 2.2 Các mô hình dự báo

# ARIMA (Auto-Regressive Integrated Moving Average): mô hình tuyến tính, phù hợp cho chuỗi không dừng.

# SARIMA (Seasonal ARIMA): mở rộng ARIMA với thành phần mùa vụ.

# Prophet: mô hình của Facebook, phù hợp cho dữ liệu có xu hướng, mùa vụ phức tạp và có khả năng xử lý outlier tốt.

# 2.3 Các thước đo đánh giá

# MAE (Mean Absolute Error): trung bình sai số tuyệt đối.

# RMSE (Root Mean Square Error): căn bậc hai trung bình bình phương sai số.

# MAPE (Mean Absolute Percentage Error): tỉ lệ phần trăm sai số trung bình.

# PHẦN 3. DỮ LIỆU

# 3.1 Nguồn dữ liệu

# Bộ dữ liệu “Smart meters in London” từ Kaggle.

# Dữ liệu gốc: chỉ số điện năng tiêu thụ theo nửa giờ của 5,567 hộ gia đình.

# Thời gian thu thập: 2011–2014.

# 3.2 Tiền xử lý

# Gộp dữ liệu từ mức nửa giờ → ngày → tuần → tháng.

# Lọc dữ liệu cho một hộ gia đình cụ thể (ví dụ: MAC000002).

# Xử lý dữ liệu thiếu và chuyển đổi định dạng thời gian (tstp).

# PHẦN 4. PHƯƠNG PHÁP NGHIÊN CỨU

# 4.1 Quy trình thực hiện

# Khám phá và tiền xử lý dữ liệu.

# Huấn luyện mô hình ARIMA trên chuỗi thời gian theo ngày.

# Huấn luyện mô hình SARIMA với thành phần mùa vụ (theo tuần).

# Huấn luyện mô hình Prophet với dữ liệu (ds, y).

# Đánh giá và so sánh mô hình bằng MAE, RMSE, MAPE.

# 4.2 Công cụ

# Python (pandas, statsmodels, prophet, matplotlib, scikit-learn).

# Jupyter Notebook trên Kaggle.

# PHẦN 5. KẾT QUẢ

# 5.1 Dự báo bằng ARIMA

# Biểu đồ: đường xanh (thực tế) và đường đỏ (dự báo).

# 

# Sai số:

# MAE: 4.457

# RMSE: 5.413

# MAPE: 66.12%

# 5.2 Dự báo bằng SARIMA

# Có tính đến thành phần mùa vụ.

# 

# Sai số:

# MAE: 3.797

# RMSE: 4.676

# MAPE: 53.03%

# 5.3 Dự báo bằng Prophet

# Khả năng bắt xu hướng tốt, mô hình mềm dẻo.

# 

# Sai số:

# MAE: 2.729

# RMSE: 4.150

# MAPE: 39.05%

# 5.4 So sánh kết quả

| Mô hình | MAE | RMSE | MAPE |
| --- | --- | --- | --- |
| ARIMA | 4.457 | 5.413 | 66.12% |
| SARIMA | 3.797 | 4.676 | 53.03% |
| Prophet | 2.729 | 4.150 | 39.05% |

# PHẦN 6. THẢO LUẬN

# Prophet vượt trội hơn nhờ khả năng mô hình hóa xu hướng và mùa vụ linh hoạt.

# ARIMA đơn giản, nhanh nhưng khó nắm bắt chu kỳ dài.

# SARIMA cải thiện ARIMA nhờ tính mùa vụ, nhưng chưa bằng Prophet.

# Độ chính xác vẫn có thể nâng cao bằng cách:

# Kết hợp thêm dữ liệu thời tiết, ngày nghỉ lễ.

# Áp dụng các mô hình học sâu như LSTM (nếu tài nguyên tính toán cho phép).

# PHẦN 7. KẾT LUẬN

# Đã xây dựng và so sánh ba mô hình dự báo chuỗi thời gian trên dữ liệu điện năng hộ gia đình.

# Prophet cho kết quả tốt nhất với MAPE ~39%.

# Kết quả này khẳng định khả năng ứng dụng AI trong dự báo điện năng và mở ra hướng nghiên cứu sâu hơn về dữ liệu đa nguồn.

# PHẦN 8. HƯỚNG PHÁT TRIỂN

# Mở rộng mô hình sang nhiều hộ gia đình để đánh giá tính tổng quát.

# Tích hợp thêm dữ liệu thời tiết (nhiệt độ, độ ẩm).

# Kết hợp thông tin nhân khẩu học để phân tích yếu tố ảnh hưởng.

# PHẦN 9. TÀI LIỆU THAM KHẢO

# Kaggle – Smart meters in London dataset:

# <https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london>

# Hyndman, R.J., Athanasopoulos, G. *Forecasting: principles and practice*, OTexts, 2018.

# Taylor, S.J., Letham, B. *Forecasting at scale*, The American Statistician, 2018.

# Box, G.E.P., Jenkins, G.M. *Time Series Analysis: Forecasting and Control*, Wiley, 2015.