

## Nghiên cứu và xây dựng chương trình giám sát tải không xâm nhập ứng dụng bộ thư viện Scikit - Learn (SKLEARN)

### Implementing Scikit-Learn libraries for developing a Non-Intrusive Load Monitoring program

Đặng Hoàng Anh<sup>1\*</sup>, Đào Văn Dũng<sup>1</sup>, Nguyễn Văn Quang<sup>2</sup>

<sup>1</sup>Trường Điện - Điện tử - Đại học Bách khoa Hà Nội

<sup>2</sup>Trường đại học Công nghiệp Hà Nội

\* Corresponding author E-mail: [anh.danghoang@hust.edu.vn](mailto:anh.danghoang@hust.edu.vn)

#### Abstract

In recent years, load monitoring plays an important role in energy saving and smart grid development. However, detailed monitoring of individual loads requires a huge number of measuring devices, leading to a lot of difficulties in investment, development and management. To solve this problem, non-intrusive load monitoring technique which apply machine learning algorithms allows to identify individual loads consumptions base on total consumption data, thereby significantly reducing the number of measuring devices and investment costs. In this paper, through the application of machine learning algorithms, we analyzed the energy consumption dataset of an apartment with total and break-down consumption data. The result point outs difficulties and potentials of applied machine learning in the energy disaggregation.

**Keywords:** Machine Learning, Energy disaggregation, Non-Intrusive Load Monitoring.

#### Tóm tắt

Những năm gần đây, giám sát sử dụng điện năng đang trở thành yếu tố then chốt trong định hướng tiết kiệm năng lượng và phát triển lưới điện thông minh. Tuy nhiên, để giám sát và theo dõi chi tiết các phụ tải thành phần cần một lượng lớn thiết bị đo đếm, từ đó dẫn tới khó khăn về đầu tư, triển khai và quản lý. Để giải quyết vấn đề này, giải thuật Giám sát tải không xâm nhập (NILM) ứng dụng kỹ thuật học máy (ML) cho phép xác định phụ tải thành phần dựa trên dữ liệu đo tổng tiêu thụ điện, qua đó giảm đáng kể số lượng thiết bị đo đếm và chi phí đầu tư. Trong bài báo này, thông qua ứng dụng các giải thuật Machine learning, nhóm nghiên cứu đã phân tích bộ dữ liệu tiêu thụ điện năng của một căn hộ với các phụ tải tổng và thành phần đa dạng. Kết quả của bài báo đã chỉ ra những khó khăn và tiềm năng trong việc ứng dụng học máy phân tách phụ tải thành phần từ tổng tiêu thụ điện năng.

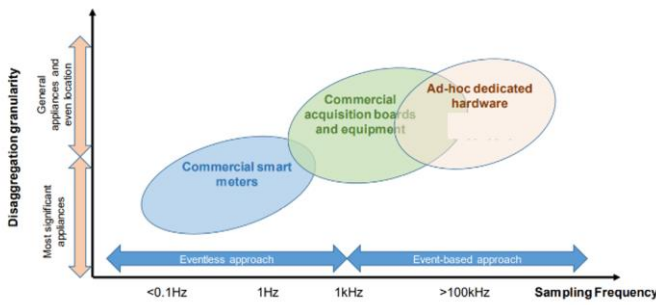
#### 1. Giới thiệu

Nhằm ứng phó với biến đổi khí hậu và sự nóng lên toàn cầu, sử dụng hiệu quả năng lượng ngày càng trở thành một trong các tiêu chí được quan tâm trong các dự án về năng lượng. Những năm gần đây, Chính phủ Việt Nam đã ban hành hàng loạt các quyết định và chương trình quốc gia về sử dụng năng lượng tiết kiệm và hiệu quả. Đặc biệt, các đồng hồ thông minh và hệ thống giám sát năng lượng đóng một vai trò quan trọng trong các mục tiêu về quản lý, hiệu quả năng lượng. Thực tế, việc phản hồi các thông tin về tiêu thụ năng lượng đến người dùng có thể giúp tiết kiệm năng lượng lên đến hơn 14% tổng lượng điện năng tiêu thụ [1]. Tuy nhiên, việc triển khai các hệ

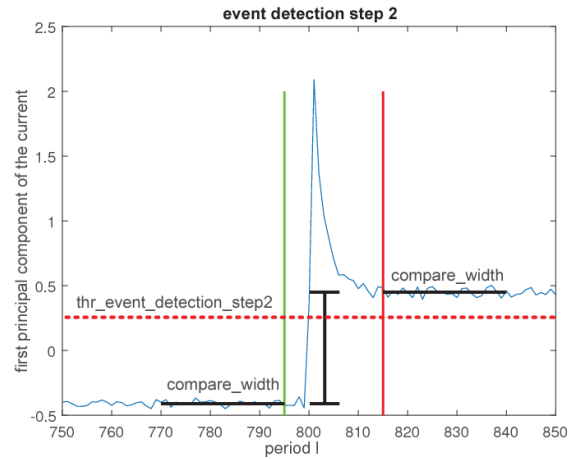
thống giám sát năng lượng theo cách truyền thống cần một số lượng các thiết bị giám sát năng lượng, dẫn đến khó khăn trong đầu tư, triển khai và quản lý. Vì vậy, kỹ thuật Giám sát tải trong không xâm phạm đã nổi lên như một giải pháp tốt nhất để đáp ứng các yếu tố kỹ thuật, đồng thời giải quyết các vấn đề chi phí đầu tư và quản lý nhờ số lượng điểm đo cần triển khai thấp hơn.

Giám sát tải không xâm nhập (NILM) là một quá trình phân tích dữ liệu điện năng tiêu thụ tổng và đưa ra các thông tin tiêu thụ điện của từng thiết bị trong hệ thống điện. Trong giai đoạn đào tạo, một mô hình học máy được lựa chọn để đào tạo với bộ dữ liệu bao gồm dữ liệu tiêu thụ của tải thành phần. Sau đó, mô hình đã được đào tạo sẽ sử dụng để đưa ra các dự đoán về tải tiêu thụ thành phần dựa trên dữ liệu tiêu thụ tổng được cung cấp cho mô hình.

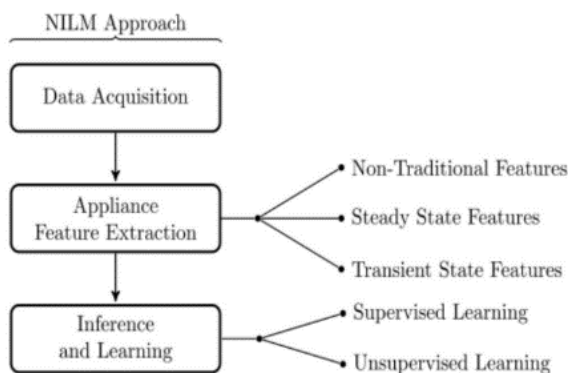
Trong bài báo này, kỹ thuật giám sát tải không xâm nhập được nghiên cứu để dự đoán tiêu thụ điện của một căn hộ tại Hà Nội. Trong phần đầu tiên của bài báo, phương pháp luận của kỹ thuật giám sát tải không xâm nhập được đưa ra. Sau đó, các mô hình NILM sẽ được xây dựng từ bộ thư viện Scikit-learn với các thuật toán học máy thông dụng và được tối ưu hoá các siêu tham số nhằm gia tăng độ chính xác của kết quả dự đoán. Kết quả thu được cho thấy các thách thức cũng như cơ hội trong việc áp dụng học máy phân tách phụ tải điện.



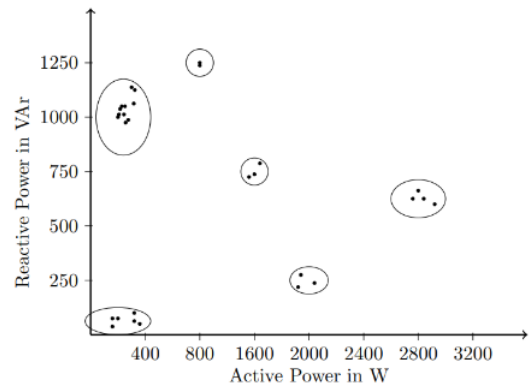
Hình 1: Giám sát tiêu thụ [3]



Hình 2: Xác định sự kiện [3].



Hình 3: Trích xuất đặc trưng [4].



Hình 4: Phân loại tải [5].

## 2. Quá trình thực hiện của giám sát tải không xâm nhập

Kỹ thuật NILM được xác định có 4 bước chính, bao gồm: Giám sát tiêu thụ, xác định sự kiện, trích xuất đặc trưng và phân loại tải [2].

### 2.1. Giám sát tiêu thụ

Bước đầu tiên của kỹ thuật NILM là thực hiện đo và thu thập dữ liệu. Thông tin được thu thập lý tưởng nhất là thông tin về điện áp và dòng điện tiêu thụ với tần số cao. Tuy nhiên, đa phần thiết bị thương mại cho mục đích giám sát và quản lý năng lượng toà nhà đều có tốc độ lấy mẫu thấp (0.2-1Hz), do vậy các thông tin khác như công suất tác dụng, công suất phản kháng, hệ số công suất hay các đặc tính V-I khác cũng sẽ được xem xét để thu thập.

### 2.2. Xác định sự kiện

Mỗi khi thiết bị điện trong hệ thống được thay đổi sang một trạng thái khác (chuyển từ trạng thái tắt sang bật hoặc chuyển sang một trạng thái hoạt động mới) sẽ gây ra các thay đổi trong trạng thái điện. Dựa trên mô tả trong Hình 2, kỹ thuật NILM thực hiện xác định sự kiện trên bằng cách phân tích các thay đổi trong các thông số đo đạc được. Thông thường có hai phương pháp được sử dụng để phát hiện sự kiện [6]:

- Phương pháp phát hiện cạnh là phương pháp xác định sự kiện thông qua những thay đổi trong các thông số điện như công suất, dòng điện, sóng hài,... Phương pháp này đòi hỏi xây dựng những bộ phát hiện sự kiện phức tạp. Ngoài ra bộ dữ liệu tần suất cao được khuyến nghị để tăng mức độ chính xác khi xác định các sự kiện.
- Phương pháp xác suất là phương pháp đơn giản hơn trong việc xác định các sự kiện. Hệ thống thực hiện ghi lại các thay đổi giá trị công suất và tính toán xác suất của thay đổi đó ứng với từng thiết bị cụ thể. Phương pháp này được khuyến nghị sử dụng cho bộ dữ liệu tần suất thấp.

### 2.3. Trích xuất đặc trưng

Mỗi một sự kiện được xác định, các thông số đặc trưng của sự kiện sẽ được phân tích và đưa ra. Như mô tả trong Hình 3, thông số đặc trưng được trích xuất gồm hai loại: Thông số đặc trưng cho quá trình quá độ và trạng thái ổn định [7]. Các thiết bị điện được cấu tạo từ nhiều thành phần khác nhau và gây ra các thay đổi khác nhau trong tín hiệu điện và tạo ra các thông số đặc trưng cho quá trình quá độ. Để có thể lấy được các thông số này, bộ dữ liệu yêu cầu phải được thu thập ở tần suất cao. Các thông số đặc trưng trạng thái quá độ có thể là công suất quá độ, dạng sóng của dòng điện khởi động hay nhiễu điện áp. Ở mặt khác, các thông số của trạng thái ổn định thì trái ngược lại do có thể thu thập được với bộ dữ liệu tần số thấp. Các thông số bao gồm sự thay đổi về công suất tiêu thụ, dòng điện tiêu thụ, nhiễu điện,...

## 2.4. Phân loại tải

Bước cuối cùng của kỹ thuật giám sát tải không xâm nhập là phân loại ra thiết bị nào đã gây ra sự kiện. Dựa trên các dữ liệu được thu thập và với các đặc trưng xác định được, các thuật toán học máy sẽ thông qua các phương pháp khác nhau để dự đoán được loại thiết bị gây ra các sự kiện. Một ví dụ đơn giản là có thể phân cụm các điểm dữ liệu và phân loại thông qua các cụm dữ liệu đó như Hình 4. Với các ứng dụng của NILM, loại thuật toán học máy thường được sử dụng bao gồm hai dạng:

- Học tập có giám sát (Supervised learning algorithms) - thường được sử dụng [8].
- Học tập không giám sát (Unsupervised learning algorithms) [9].

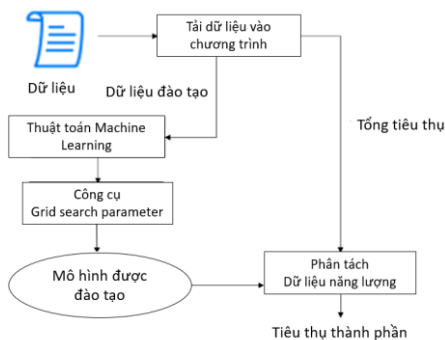
## 3. Chương trình NILM sử dụng thư viện SK-Learn

### 3.1. Xây dựng chương trình

Scikit-learn [10] là một thư viện được viết phần lớn bằng Python, được tích hợp một loạt thuật toán machine learning hiện đại được sử dụng cho các phương pháp học tập có giám sát và học tập không được giám sát ở quy mô trung bình. Bộ thư viện tập trung vào việc đưa machine learning đến những người không phải là chuyên gia bằng cách sử dụng ngôn ngữ bậc cao và dễ dàng tính dễ sử dụng, hiệu suất, tài liệu hướng dẫn và tính nhất quán của API.

Trong nội dung bài báo, thông qua bộ thư viện Scikit-learn, một số thuật toán thông dụng đưa ra và thông qua các công cụ hỗ trợ của thư viện để lựa chọn các tham số tối ưu cho từng thuật toán. Kết quả hướng đến xây dựng các mô hình dự đoán cho kết quả tốt nhất. Các thuật toán được lựa chọn bao gồm:

- Random Forest Regressor
- Gradient Boosting Regressor
- K - Nearest Neighbors Regressor
- Gaussian Process Regressor
- Multi - layer Perceptron Regressor



Hình 5. Chương trình NILM dựa trên thư viện Sklearn

Chương trình hoạt động theo các bước được mô tả trong Hình 5. Bộ dữ liệu được chuẩn bị bao gồm dữ liệu tiêu thụ tổng và dữ liệu thành phần. Sau đó sẽ được chia thành hai bộ dữ liệu dùng cho mục đích huấn luyện mô hình và kiểm tra độ chính xác. Bộ dữ liệu dùng để huấn luyện sẽ được đưa vào mô hình và được tiến hành training. Tuy nhiên, trong quá trình huấn luyện, mô hình sẽ được thay đổi các siêu tham số đầu vào nhằm mục đích gia tăng độ chính xác. Tại bước này, chương

trình sử dụng công cụ GridSearchCV được cung cấp kèm theo thư viện Scikit-learn. Công cụ này hỗ trợ việc thay lần lượt các tham số trong một bộ tham số được cung cấp từ trước và so sánh độ chính xác với nhau.

Mô hình có độ chính xác cao nhất sẽ được lựa chọn cùng với bộ tham số tương ứng. Bộ dữ liệu kiểm tra bao gồm dữ liệu tiêu thụ tổng sẽ được đưa vào để mô hình tiến hành dự đoán và kiểm tra lại với dữ liệu tiêu thụ thành phần, từ đó đánh giá được độ chính xác của mô hình.

### 3.2. Đánh giá độ chính xác

Để đánh giá độ chính xác trong việc ước tính tiêu thụ điện của từng thiết bị, thông số RMSE được sử dụng để đánh giá sai số giữa giá trị công suất dự đoán được với giá trị công suất thực tế.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{P}_i - P_i)^2}{n}} \quad (1)$$

Với:

$\hat{P}_i$ : Công suất tác dụng dự đoán được

$P_i$ : Công suất tác dụng thực tế

Ngoài ra, mục đích của kỹ thuật NILM là dự đoán mức tiêu thụ điện của từng thiết bị, do vậy sai số giữa tổng tiêu thụ điện dự đoán và thực tế sẽ được đưa ra để so sánh.

$$\text{Sai số tiêu thụ điện} = \frac{|\hat{E} - E|}{E} \times 100\% \quad (2)$$

$$\hat{E} = \int_0^T \hat{P}(t) \times dt \quad (3)$$

$$E = \int_0^T P(t) \times dt \quad (4)$$

Với:

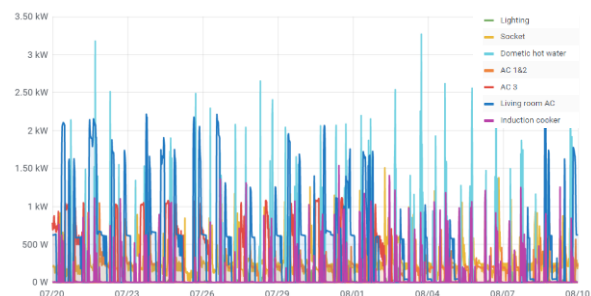
$\hat{E}$ : Tổng điện năng tiêu thụ dự đoán.

$E$ : Tổng điện năng tiêu thụ thực tế.

## 4. Triển khai và phân tích kết quả

### 4.1. Đối tượng thử nghiệm

Chương trình NILM sẽ được thử nghiệm với bộ dữ liệu tiêu thụ điện thực tế của một căn hộ chung cư tại Hà Nội có diện tích 91m<sup>2</sup>, gồm 1 phòng khách, 2 phòng ngủ và 1 phòng làm việc. Dựa trên dữ liệu tiêu thụ tổng, mô hình sẽ dự đoán và đưa ra hoạt động của các phụ tải trong căn hộ.

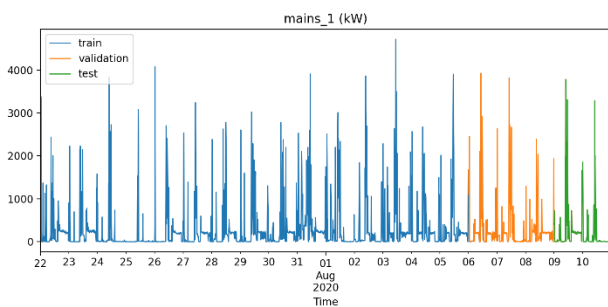


Hình 6: Bộ dữ liệu tiêu thụ điện của đối tượng căn hộ chung cư tại Hà Nội

Bộ dữ liệu tiêu thụ điện trong 20 ngày được thể hiện như Hình 6. Trong đó bình nóng lạnh, điều hoà và bếp từ là thông tin của thiết bị riêng biệt, còn lại thông tin về chiếu sáng bao gồm các đèn được sử dụng trong căn hộ và ổ cắm là tổ hợp nhiều thiết bị được sử dụng thông qua ổ cắm.

#### 4.2. Xây dựng trường hợp thử nghiệm

Bộ dữ liệu được xây dựng từ toàn bộ dữ liệu tiêu thụ điện của căn chung cư. Dữ liệu tiêu thụ điện tổng sẽ được tính tổng của các tải tiêu thụ thành phần. Mục tiêu của trường hợp thử nghiệm nhằm hướng đến đánh giá độ hiệu quả của chương trình với dữ liệu thực tế và tập trung đi sâu vào phân tích các dạng tải riêng lẻ trong công trình như chiếu sáng, điều hoà không khí, bếp từ và bình nóng lạnh. Riêng dạng tải ổ cắm thuộc dạng dùng hỗn hợp của nhiều tải khác nên nhóm tác giả không kể đến dạng tải này. Dữ liệu tiêu thụ tổng sẽ được tính bằng tổng của các tải được kê trên.

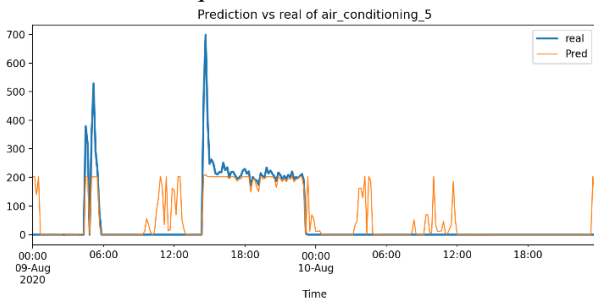


Hình 7: Dữ liệu đào tạo và kiểm tra

Bộ dữ liệu được chia làm ba phần như mô tả trong Hình 7: Dữ liệu từ ngày đầu tiên (ngày 22/7/2020) đến hết ngày 06/08/2020 được sử dụng để làm bộ dữ liệu đào tạo cho các mô hình. Phần dữ liệu tiếp theo đến hết ngày 08/09/2020 (3 ngày) dùng làm dữ liệu kiểm tra phục vụ cho GridSearchCV nhằm tối ưu tham số cho mô hình. Dữ liệu còn lại (1 ngày) dùng để kiểm tra độ chính xác cho mô hình sau khi được tối ưu tham số.

#### 4.3. Đánh giá kết quả thử nghiệm

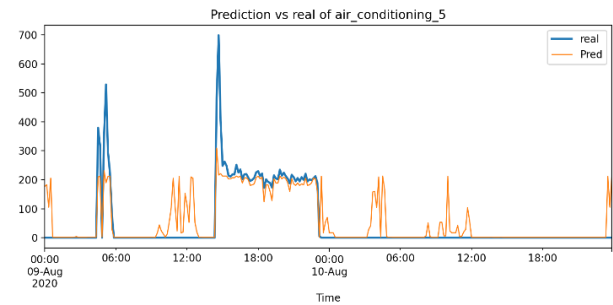
Trong phần này, một số kết quả dự đoán đặc trưng của các mô hình sẽ được đưa ra phân tích, qua đó có thể đánh giá mức độ chính xác và hiệu quả của mô hình sau khi tối ưu các tham số.



Hình 8: Kết quả dự đoán tải điều hoà của mô hình sử dụng thuật toán Random forest

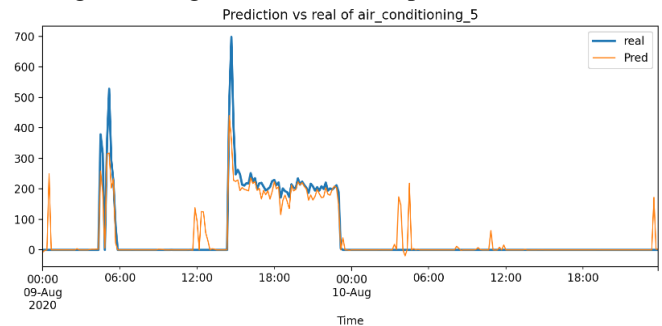
Với kết quả từ mô hình sử dụng thuật toán Random Forest thể hiện trong Hình 8, mô hình mặc dù đã được tối ưu tham số, tuy nhiên kết quả sau khi dự đoán vẫn chưa thực sự chính xác. Trong khoảng thời gian lúc 6:00 và khoảng 13:00 - 23:00 điều

hoà được sử dụng, mô hình cũng đã xác định điều hoà có tiêu thụ điện, tuy nhiên mô hình cũng không dự đoán được các giá trị đỉnh mà chỉ đưa ra được giá trị trung bình trong giai đoạn tiêu thụ. Ngoài ra còn một số thời điểm điều hoà không được sử dụng, tuy nhiên mô hình cũng cho kết luận phụ tải có tiêu thụ.



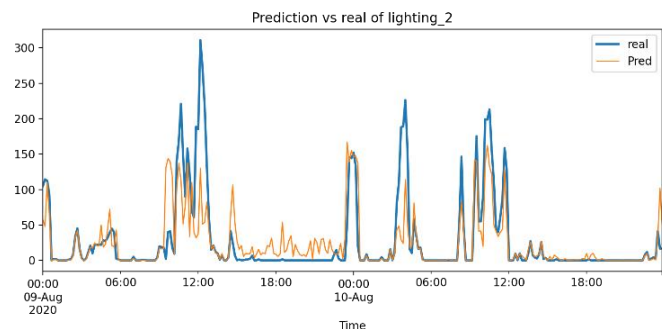
Hình 9: Kết quả dự đoán tải điều hoà của mô hình sử dụng thuật toán Gradient Boosting

Về cơ bản có thể thấy kết quả dự đoán của mô hình sử dụng thuật toán Gradient Boosting (thể hiện trong hình Hình 9) cho ra kết quả tương tự mô hình sử dụng thuật toán Random Forest. Ngoài ra, mô hình được cải thiện một số thiếu sót như đã có thể xác định một vài đỉnh tải mặc dù giá trị dự đoán chưa chính xác hoàn toàn. Một số thời điểm xác định tải tiêu thụ bị sai đã giảm nhưng vẫn chưa thể khắc phục hoàn toàn.



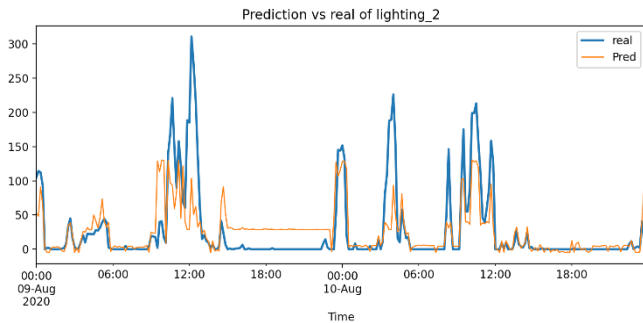
Hình 10: Kết quả dự đoán tải điều hoà của mô hình sử dụng thuật toán Gaussian Process Regressor

Kết quả nhận được từ mô hình Gaussian Process Regressor thể hiện trong Hình 10 khả quan hơn rất nhiều so với các mô hình bên trên. Ngoài việc có thể đưa ra các kết quả dự đoán thời điểm tải được sử dụng khá chính xác và đưa ra các giá trị trung bình, các đỉnh tải cũng được mô hình nhận biết và dự đoán ra với kết quả tương đối chính xác so với thực tế. Ngoài ra, các thời điểm dự đoán sai cũng xảy ra ít hơn rất nhiều so với các mô hình trên.



Hình 11: Kết quả dự đoán tải chiếu sáng của mô hình sử dụng thuật toán K- Nearest Neighbors Regressor (kNN)

Khác với phụ tải điều hòa, phụ tải chiếu sáng có đặc trưng tiêu thụ nhỏ hơn và được sử dụng biến thiên rất nhiều trong quá trình được sử dụng. Kết quả thể hiện trong Hình 11 của mô hình sử dụng thuật toán k-NN cho thấy khoảng thời gian từ 0:00 đến 6:00 khi căn hộ tiêu thụ nhỏ và ít dao động, mô hình cho ra kết quả dự đoán tương đối tốt, gần như bám sát được đường tải tiêu thụ. Các thời điểm khác mô hình cũng dự đoán được việc vận hành của phụ tải. Tuy nhiên, một số thời điểm sử dụng phụ tải có dao động lớn, mô hình vẫn có những sai sót nhất định trong việc dự báo.



**Hình 12:** Kết quả dự đoán tải chiếu sáng của mô hình sử dụng thuật toán Multi-layer Perceptron Regressor (MLP)

Cuối cùng là kết quả của mô hình sử dụng thuật toán Multi-layer Perceptron Regressor thể hiện tại Hình 12 cho thấy khả năng dự đoán tải kém hiệu quả hơn so với mô hình sử dụng thuật toán k-NN. Dù mô hình đã có thể dự đoán được tiêu thụ của phụ tải và cũng có thể xác định được tương đối các đỉnh phụ tải. Tuy nhiên mô hình có một số dự đoán sai khá nghiêm trọng vào thời điểm từ 13:00 đến 23:00 khi dự đoán sai hoàn toàn tiêu thụ của phụ tải. Ngoài ra, các các đỉnh tiêu thụ của tải cũng được mô hình dự đoán kém chính xác hơn.

**Bảng 1** Sai số kết quả dự đoán của các mô hình sau khi được tối ưu tham số

	Chiếu sáng		Bình nóng lạnh		Điều hòa không khí		Bếp từ	
	RMS E Công suất	Sai số điện năng	RMS E Công suất	Sai số điện năng	RMSE Công suất	Sai số điện năng	RMS E Công suất	Sai số điện năng
Random Forest	37	10%	157	1.9 %	70	18%	151	6.3%
Gradient Boosting	36	7.4%	144	3.1 %	66	19%	144	10%
k-NN	38	11%	161	3%	71	20%	156	11%
Gaussian Process	41	34%	174	35%	42	2.9%	167	54%
MLP	41	9%	169	78%	77	6.4%	158	13%

Bảng 1 tổng hợp lại kết quả nhận được từ các mô hình nhằm so sánh được mức độ hiệu quả của các mô hình khác nhau. Có thể thấy các dạng phụ tải phức tạp như chiếu sáng hay điều hòa không khí đều gây ra sai số lớn hơn (10-20%) so với các dạng phụ tải có dạng tiêu thụ đơn giản (3-5%). Ngoài ra một số mô hình mặc dù cho thấy khả năng dự đoán đỉnh tải tốt hơn nhưng xét về tổng quan lại cho sai số tổng thể cao hơn như mô hình Gaussian Process. Tuy nhiên, xuất hiện đột biến khi mô hình này dự đoán tải điều hòa cho kết quả tốt hơn đáng kể (2.9%) so với các mô hình còn lại (~20%).

## 5. Kết luận

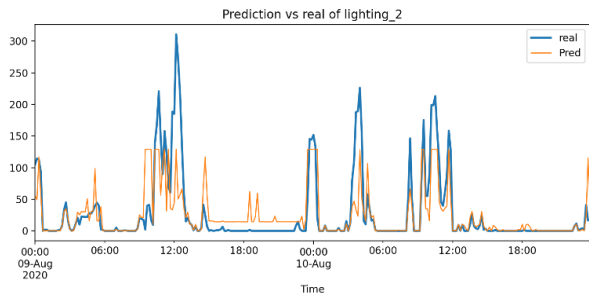
Trong bài báo này, việc dự đoán tiêu thụ của tải thành phần dựa trên dữ liệu đo tổng của một căn hộ chung cư tại Hà Nội đã được triển khai thử nghiệm với việc ứng dụng các thuật toán machine learning thông dụng được cung cấp bởi thư viện scikit-learn. Các thuật toán được lựa chọn để thử nghiệm bao gồm Random Forest, Gradient Boosting Regressor, K- Nearest Neighbors Regressor, Gaussian Process Regressor và Multi - layer Perceptron Regressor. Từ so sánh các kết quả nhận được tiếp tục cho thấy các thuật toán đều có các ưu nhược điểm khác nhau. Các mô hình như Random Forest, Gradient Boosting hoặc k-NN cho kết quả sai số tổng quan khá ấn tượng khi sai số của kết quả thấp hơn của các thuật toán còn lại mặc dù việc dự đoán ra các đỉnh tải thấp kém hơn. Mặt còn lại, thuật toán Gaussian Process Regressor và Multi-layer Perceptron Regressor mặc dù cho các kết quả dự đoán được các đỉnh phụ tải tốt hơn, tuy nhiên khi xét về các sai số tổng quan lại cho ra sai số cao hơn các thuật toán còn lại. Như vậy, có thể thấy việc đưa ra các trọng số của kết quả hướng đến sẽ có tính quyết định khá lớn khi muốn lựa chọn thuật toán nào trong việc dự đoán các tải thành phần. Trong các nghiên cứu sâu hơn về sau, các công việc tiếp theo sẽ tập trung vào nâng cao độ chính xác của việc dự đoán bao gồm sai số và việc dự đoán ra các đặc tính phụ tải. Ngoài ra, việc đưa vào các biến phụ trợ nhằm gia tăng độ chính xác của kết quả dự đoán sẽ được cân nhắc và ứng dụng cho các thuật toán.

## Tài liệu tham khảo

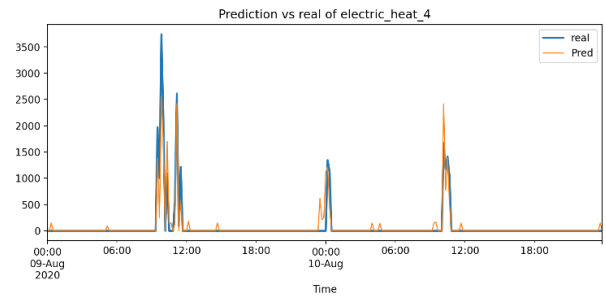
- [1] K. Ehrhardt-Martinez, K. A. Donnelly, J. A. Laitner, Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities, ACEE E105, (2010)
- [2] S. Makonin, Approaches To Non-Intrusive Load Monitoring (NILM) In The Home, A depth report of Simon Fraser University, (2010).
- [3] A. Ruano, A. Hernandez, J. Ureña, M. Ruano, J. Garcia, NILM Techniques for Intelligent Home Energy Management and Ambient Assisted Living: A Review, A research report of Energies 2019. 12. 2203, (2019).
- [4] C. Puente, R. Palacios, Y. González-Arechavala, E. F. Sánchez-Úbeda., Non-Intrusive Load Monitoring (NILM) for Energy Disaggregation Using Soft Computing Techniques, A research report of Energies 2020, 13(12), 3117, (2020).
- [5] C. Klemenjak, P. Goldsborough, Non-Intrusive Load Monitoring: A Review and Outlook, arXiv:1610.01191, (2016).
- [6] K. D. Anderson, M. E. Berges, A. Ocneanu, D. Benitez, J. M.F. Moura, Event detection for Non Intrusive load monitoring, 38th Annual Conference on IEEE Industrial Electronics Society, (2012) 3312-3317
- [7] M. Jürgel, J. Mass, Non-Intrusive Load Monitoring in the OpenHAB smart home framework, A bachelor's thesis of University of Tartu, (2019).
- [8] J. Kelly, W. Knottenbelt, Neural NILM: Deep Neural Networks Applied to Energy Disaggregation, The 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environment, (2015) 55-64
- [9] N. Batra, J. Kelly, O. Parson, H. Dutta, W. Knottenbelt, A. Rogers, A. Singh, M. Srivastava, Demo Abstract: NILMTK v0.2: A Non-intrusive Load Monitoring Toolkit for Large Scale Data Sets, arXiv:1409.5908, (2014) 182-183
- [10] Scikit-learn: machine learning in Python, [Online]. Available: <https://scikit-learn.org/stable/>.

### Phụ lục: Kết quả dự báo của các mô hình

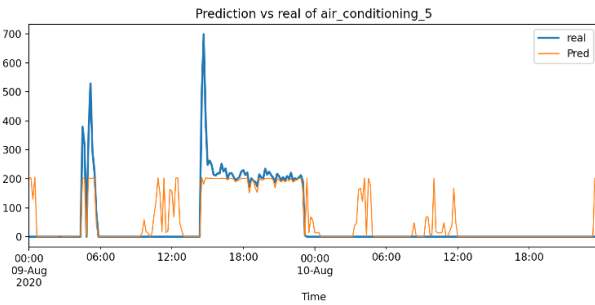
Kết quả của mô hình Machine learning sử dụng thuật toán Random Forest



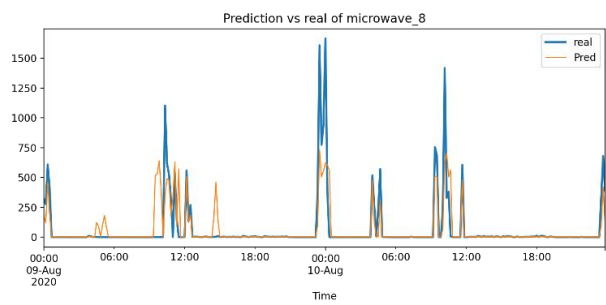
a, Tài chiếu sáng



b, Tài bình nóng lạnh

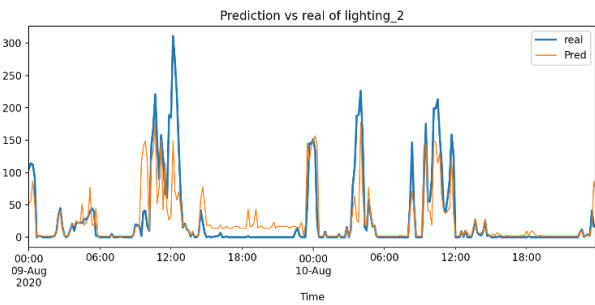


c, Tài điều hoà

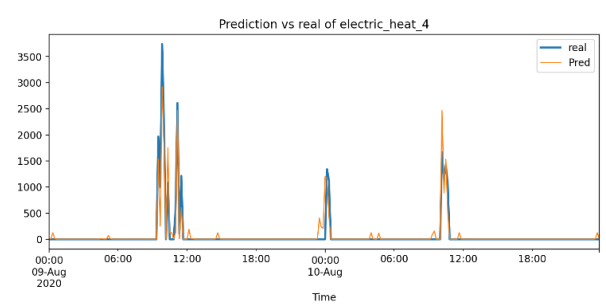


d, Tài bếp từ

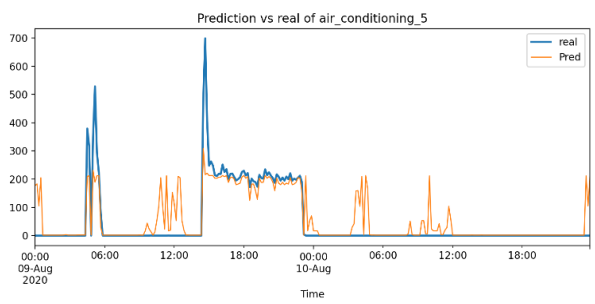
Kết quả của mô hình Machine learning sử dụng thuật toán Gradient Boosting



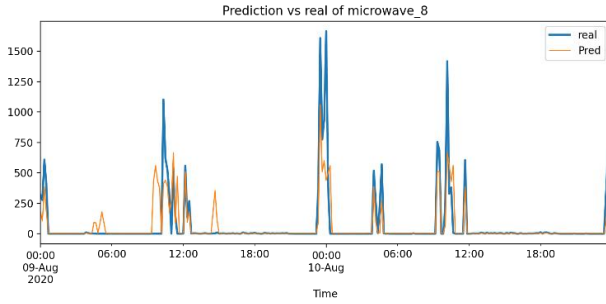
a, Tài chiếu sáng



b, Tài bình nóng lạnh

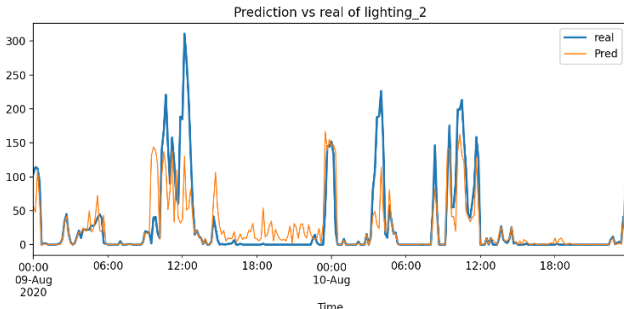


c, Tài điều hoà

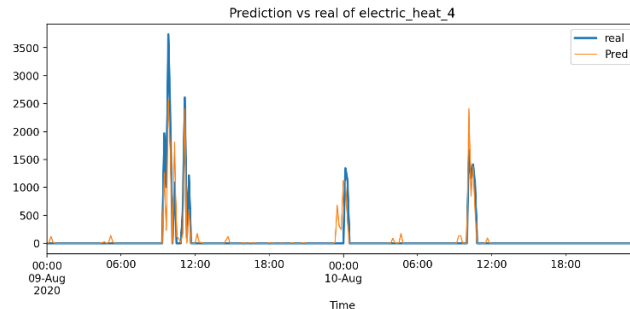


d, Tài bếp từ

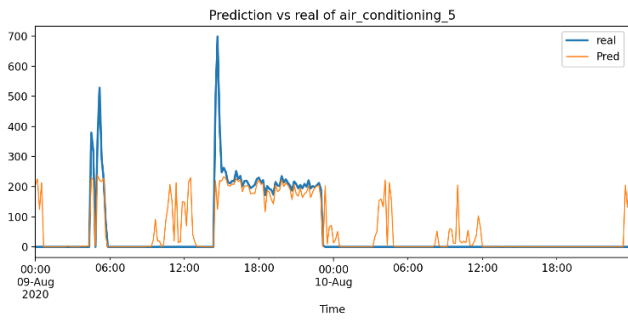
Kết quả của mô hình Machine learning sử dụng thuật toán k-NN



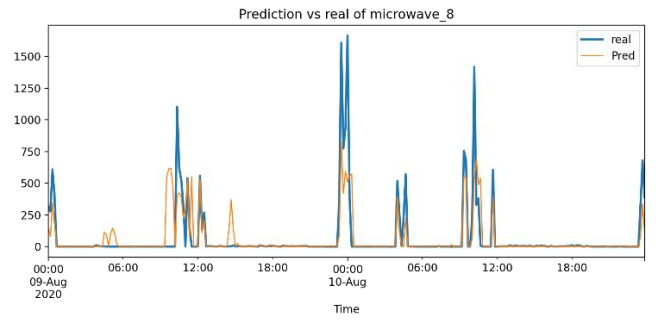
a, Tài chiếu sáng



b, Tài bình nóng lạnh

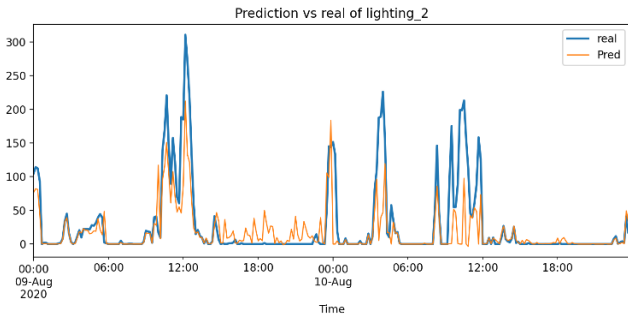


c, Tài điều hoà

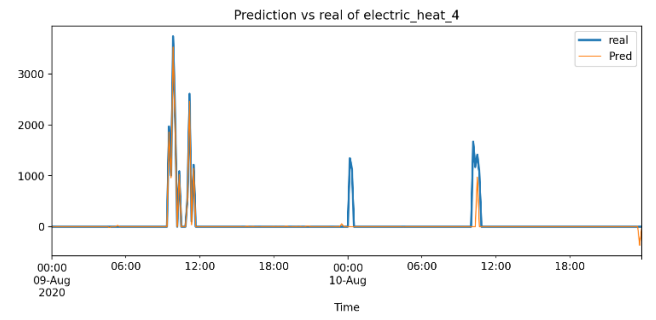


d, Tài bếp từ

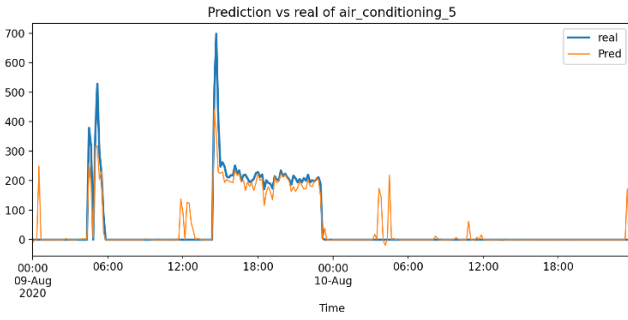
Kết quả của mô hình Machine learning sử dụng thuật toán Gaussian Process



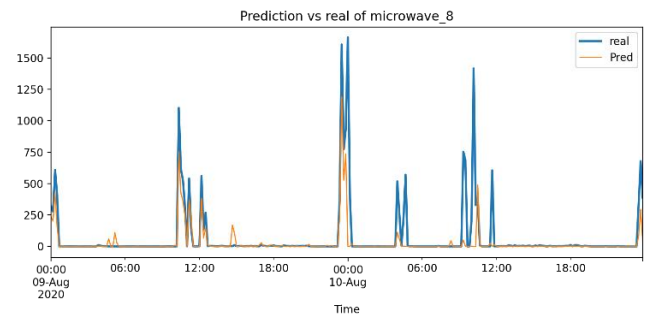
a, Tài chiếu sáng



b, Tài bình nóng lạnh

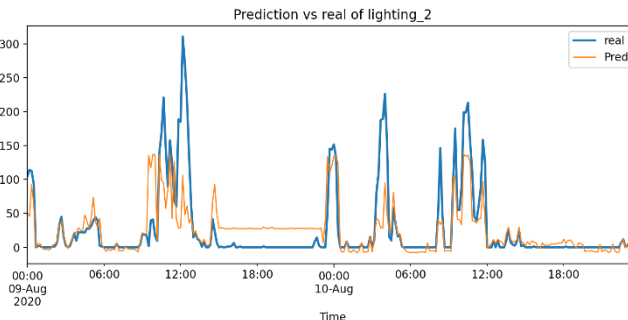


c, Tài điều hoà

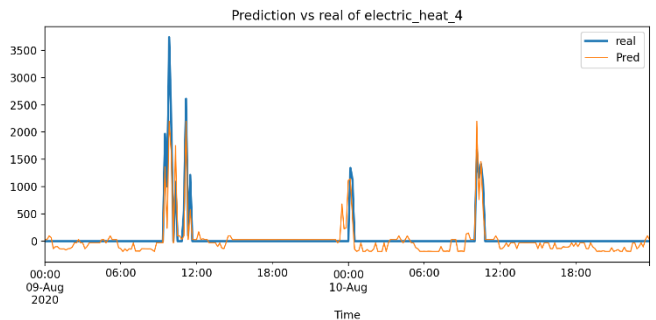


d, Tài bếp từ

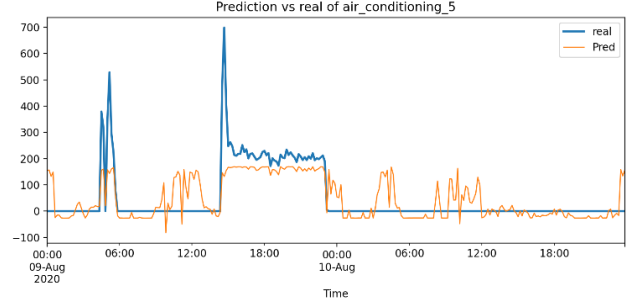
Kết quả của mô hình Machine learning sử dụng thuật toán dựng thuật toán MLP



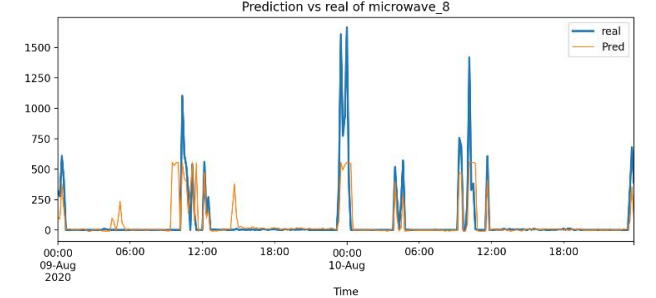
a, Tài chiếu sáng



b, Tài bình nóng lạnh



c, Tài điều hoà



d, Tài bếp từ