

**ĐẠI HỌC ĐÀ NẴNG**  
**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT**

**BÁO CÁO TỔNG KẾT**  
**ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG**

**ĐÁNH GIÁ LỰA CHỌN ỨNG DỤNG GOOGLE TREND**  
**TRONG TỐI ƯU HÓA DỰ BÁO NHU CẦU TIÊU THỤ**  
**NĂNG LƯỢNG ĐIỆN DỰA TRÊN CÁC THUẬT TOÁN**  
**TRÍ TUỆ NHÂN TẠO**

**Mã số: T2023-06-23**

**Chủ nhiệm đề tài: ThS. Doãn Văn Khánh**  
**Đơn vị: Khoa Điện - Điện tử**  
**Chương trình đào tạo: Kỹ thuật Điện - Điện tử**

*Đà Nẵng, Tháng 02 Năm 2025*

ĐẠI HỌC ĐÀ NẴNG  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT

**BÁO CÁO TỔNG KẾT**  
**ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG**

**ĐÁNH GIÁ LỰA CHỌN ỨNG DỤNG GOOGLE TREND  
TRONG TỐI ƯU HÓA DỰ BÁO NHU CẦU TIÊU THỤ  
NĂNG LƯỢNG ĐIỆN DỰA TRÊN CÁC THUẬT TOÁN  
TRÍ TUỆ NHÂN TẠO**

**Mã số: T2023-06-23**

Xác nhận của cơ quan chủ trì đề tài

**KT. HIỆU TRƯỞNG  
PHÓ HIỆU TRƯỞNG**



**PGS. TS. Võ Trung Hùng**

Chủ nhiệm đề tài

A blue ink signature, likely belonging to ThS. Doãn Văn Khánh.

**ThS. Doãn Văn Khánh**

## MỤC LỤC

|                                                                                             |           |
|---------------------------------------------------------------------------------------------|-----------|
| DANH MỤC CÁC CHỮ VIẾT TẮT VÀ KÝ HIỆU .....                                                  | i         |
| DANH MỤC CÁC HÌNH ẢNH.....                                                                  | iii       |
| DANH MỤC CÁC BẢNG BIỂU.....                                                                 | iv        |
| THÔNG TIN KẾT QUẢ NGHIÊN CỨU .....                                                          | v         |
| <b>MỞ ĐẦU .....</b>                                                                         | <b>1</b>  |
| 1. Tính cấp thiết của đề tài .....                                                          | 1         |
| 2. Mục tiêu của đề tài .....                                                                | 2         |
| 3. Cách tiếp cận và phương pháp nghiên cứu.....                                             | 3         |
| 4. Đối tượng và phạm vi nghiên cứu .....                                                    | 4         |
| 5. Tính mới của đề tài.....                                                                 | 5         |
| <b>CHƯƠNG 1: TỔNG QUAN VỀ TIÊU THỤ VÀ SỬ DỤNG NĂNG LƯỢNG ĐIỆN ...</b>                       | <b>6</b>  |
| 1.1 Giới thiệu về nhu cầu tiêu thụ điện trong bối cảnh hiện đại .....                       | 6         |
| 1.2 Các yếu tố ảnh hưởng đến nhu cầu tiêu thụ điện .....                                    | 6         |
| 1.3 Xu hướng phát triển của ngành điện và các thách thức trong dự báo .....                 | 9         |
| <b>CHƯƠNG 2 : PHƯƠNG PHÁP LUẬN GIẢI QUYẾT BÀI TOÁN DỰ BÁO TIÊU THỤ NĂNG LƯỢNG ĐIỆN.....</b> | <b>11</b> |
| 2.1 Tổng quan về bài toán dự báo .....                                                      | 11        |
| 2.2 Mô hình dự báo chuỗi thời gian .....                                                    | 13        |
| 2.3 Một số vấn đề với mô hình dự báo dựa trên mạng nơ-ron học sâu .....                     | 25        |
| <b>CHƯƠNG 3: XÂY DỰNG CÁC MÔ HÌNH DỰ BÁO.....</b>                                           | <b>27</b> |
| 3.1 Phương pháp tiếp cận.....                                                               | 27        |
| 3.2 Dữ liệu Google Trend .....                                                              | 28        |
| 3.2.1 Giới thiệu chung .....                                                                | 28        |
| 3.2.2 Tiêu chí chọn từ khóa khi sử dụng Google Trend trong nghiên cứu khoa học ..           | 30        |
| 3.2.3 Từ khóa tìm kiếm trong phạm vi đề tài .....                                           | 31        |
| 3.3 Mô hình Bộ nhớ dài ngắn hạn LSTM .....                                                  | 32        |
| 3.4 Mô hình biến đổi Transformer .....                                                      | 37        |
| <b>CHƯƠNG 4: THU NHẬP, XỬ LÝ DỮ LIỆU .....</b>                                              | <b>40</b> |
| 4.1 Khái quát về thu nhập và xử lý dữ liệu .....                                            | 40        |
| 4.2 Thư viện Python sử dụng trong nghiên cứu.....                                           | 41        |

|                                                                    |           |
|--------------------------------------------------------------------|-----------|
| 4.3 Ứng dụng thu nhập, xử lý dữ liệu trong đề tài nghiên cứu ..... | 42        |
| <b>CHƯƠNG 5: KIỂM CHỨNG THỰC NGHIỆM, ĐÁNH GIÁ KẾT QUẢ.....</b>     | <b>46</b> |
| 5.1 Kiểm chứng thực nghiệm.....                                    | 46        |
| 5.2 Phân tích kết quả .....                                        | 48        |
| 5.3 Kết luận, đề xuất hướng nghiên cứu.....                        | 50        |
| TÀI LIỆU THAM KHẢO                                                 |           |
| PHỤ LỤC                                                            |           |
| THUYẾT MINH ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG                |           |
| MỤC LỤC MINH CHỨNG SẢN PHẨM CỦA ĐỀ TÀI                             |           |

## DANH MỤC CÁC CHỮ VIẾT TẮT VÀ KÝ HIỆU

|       |                                                |
|-------|------------------------------------------------|
| ML    | Machine learning                               |
| AI    | Artificial Intelligence                        |
| ANN   | Artificial Neural Network                      |
| RNN   | Recurrent Neural Network                       |
| FNN   | Feed Forward Neural Network                    |
| LSTM  | Long Short-Term Memory                         |
| MSE   | Means Square Error                             |
| RMSE  | Root Mean Squared Error                        |
| MAE   | Mean Absolute Error                            |
| AR    | Autoregression                                 |
| ADL   | Autoregressive Distributed Lag                 |
| MA    | Moving Average                                 |
| ARIMA | Autoregressive Integrated Moving Average       |
| CNN   | Convolutional Neural network                   |
| NLP   | Natural Language Processing                    |
| TPB   | Theory of Planned Behavior                     |
| RSS   | Residual Sum of Squares                        |
| ADF   | Augmented Dickey-Fuller (ADF) test             |
| KPSS  | Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. |

## DANH MỤC CÁC KÝ HIỆU

|                 |                                                    |
|-----------------|----------------------------------------------------|
| $\hat{y}_t$     | Giá trị dự báo tại thời điểm $t$                   |
| $y$             | Các giá trị quan sát trong quá khứ                 |
| $\alpha, \beta$ | Các tham số làm trơn                               |
| $\varepsilon$   | Giá trị nhiễu trắng                                |
| $\Phi_i$        | Trọng số tương ứng với các giá trị trong quá khứ   |
| $\varepsilon_t$ | Các sai số dự đoán tại thời điểm tương ứng         |
| $\theta_i$      | Trọng số t.u với các sai số trong quá khứ          |
| $h_t(x)$        | Giá trị đầu ra qua các layer tương ứng đầu vào $x$ |
| $w; W$          | Trọng số và Ma trận trọng số                       |
| $b; B$          | Độ chệch và vec-tơ điều chỉnh                      |
| $a$             | Hàm kích hoạt (tanh, sigmoid...)                   |
| $\alpha$        | Tốc độ học                                         |
| $L$             | Hàm mất mát                                        |
| $C_t$           | Trạng thái ô nhớ                                   |
| $\sigma$        | Hàm sigmoid                                        |
| tanh            | Hàm tanh                                           |

## DANH MỤC CÁC HÌNH ẢNH

|           |                                                                                                            |
|-----------|------------------------------------------------------------------------------------------------------------|
| Hình 1.1  | Mức thay đổi phần trăm theo năm trong nhu cầu điện, khu vực Châu Á - Thái Bình Dương, giai đoạn 2019-2026. |
| Hình 2.1  | Đặc trưng xu hướng và chu kỳ của dữ liệu kiểu thời gian                                                    |
| Hình 2.2  | Cấu tạo của tế bào nơron sinh học – tiền đề xây dựng mạng nơron nhân tạo                                   |
| Hình 2.3  | Mô hình mạng nơron nhân tạo trong dự báo chuỗi thời gian                                                   |
| Hình 2.4  | Mô hình mạng nơron nhân tạo đơn giản                                                                       |
| Hình 2.5  | Một số hàm kích hoạt thông dụng                                                                            |
| Hình 2.6  | Dự báo chuỗi thời gian sử dụng mạng nơron kết hợp với các biến can thiệp                                   |
| Hình 3.1  | Tìm kiếm theo từ khóa                                                                                      |
| Hình 3.2  | Mạng nơron hồi quy RNN                                                                                     |
| Hình 3.3  | Thành phần kiến trúc cơ bản của LSTM                                                                       |
| Hình 3.4  | Các ký hiệu trong hình vẽ                                                                                  |
| Hình 3.5  | Trạng thái tế bào                                                                                          |
| Hình 3.6  | Mô tả hàm sigmoid                                                                                          |
| Hình 3.7  | Mô tả hàm sigmoid kết hợp với tanh                                                                         |
| Hình 3.8  | Kết hợp để cho ra đầu ra                                                                                   |
| Hình 3.9  | Kết hợp các tầng để cho ra đầu ra là input tiếp theo của dữ liệu                                           |
| Hình 3.10 | Mô hình minh họa kiến trúc Transformer                                                                     |
| Hình 3.11 | Các lớp chi tiết của kiến trúc Transformer                                                                 |
| Hình 4.1  | Môi trường thực thi Google Colab                                                                           |
| Hình 4.2  | Nhu cầu tiêu thụ điện của khu vực Vương quốc Anh và dữ liệu Google Trends, tương ứng                       |
| Hình 5.1a | Mô hình LSTM sử dụng Google Trends trên tập kiểm thử                                                       |
| Hình 5.1b | Mô hình Transformer sử dụng Google Trends trên tập kiểm thử                                                |
| Hình 5.2a | Mô hình LSTM chỉ sử dụng dữ liệu tiêu thụ điện trên tập kiểm thử                                           |
| Hình 5.2b | Mô hình Transformer chỉ sử dụng dữ liệu tiêu thụ điện                                                      |

**DANH MỤC CÁC BẢNG BIỂU**

|          |                                               |
|----------|-----------------------------------------------|
| Bảng 3.1 | So sánh đặc điểm của ANN, LSTM và Transformer |
| Bảng 5.1 | Các tham số tối ưu của mô hình                |
| Bảng 5.2 | Kiểm tra tính dừng của hai chuỗi thời gian    |
| Bảng 5.3 | Kiểm định mối quan hệ nhân quả Granger        |
| Bảng 5.4 | Kết quả đánh giá hiệu suất mô hình.           |

ĐẠI HỌC ĐÀ NẴNG  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT

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

## THÔNG TIN KẾT QUẢ NGHIÊN CỨU

### 1. Thông tin chung:

- Tên đề tài: Đánh giá lựa chọn ứng dụng Google Trend trong tối ưu hóa dự báo nhu cầu tiêu thụ năng lượng điện dựa trên các thuật toán trí tuệ nhân tạo.
- Mã số: T2023-06-23
- Chủ nhiệm: ThS. Doãn Văn Khánh
- Cơ quan chủ trì: Trường Đại học Sư phạm Kỹ thuật- Đại học Đà Nẵng
- Thời gian thực hiện: 3/2024-2/2025

### 2. Mục tiêu:

- Xây dựng phương pháp dự báo nhu cầu tiêu thụ điện kết hợp các dữ liệu từ Google Trends.
- Đánh giá xem liệu Google Trend có thể cải thiện dự báo tiêu thụ năng lượng điện hay không?

### 3. Tính mới và sáng tạo:

- Tích hợp dữ liệu Google Trends vào dự báo nhu cầu điện : Trong khi hầu hết các nghiên cứu trước đây chỉ tập trung vào các yếu tố truyền thống như dữ liệu lịch sử tiêu thụ điện, thời tiết, hoặc giá điện, nghiên cứu này mở rộng phạm vi bằng cách khai thác Google Trends như một nguồn dữ liệu phản ánh hành vi người tiêu dùng. Đây là một cách tiếp cận mới trong lĩnh vực dự báo điện năng.
- Kết quả từ việc sử dụng dữ liệu Google Trends đặt nền tảng cho việc tích hợp và khai thác hiệu quả hơn các nguồn dữ liệu thay thế, chẳng hạn như phản hồi khách hàng và truy vấn tìm kiếm từ các nền tảng dịch vụ khách hàng (CSKH), vốn đang được các công ty điện lực triển khai rộng rãi nhưng chưa được tận dụng đầy đủ trong dự báo nhu cầu điện.

### 4. Tóm tắt kết quả nghiên cứu:

- Qua quá trình kiểm chứng trên bộ dữ liệu thực tế, có thể thấy rằng việc tích hợp Google Trends vào mô hình dự báo giúp cải thiện đáng kể độ chính xác, thể hiện qua giảm sai số MSE so với các mô hình chỉ sử dụng dữ liệu nhu cầu điện.
- Tuy nhiên, hiệu quả của Google Trends phụ thuộc vào cách dữ liệu được tiền xử lý và tích hợp vào mô hình. Không phải mọi xu hướng tìm kiếm đều có ý nghĩa trong

dự báo điện năng, do đó cần áp dụng phương pháp lọc và lựa chọn từ khóa phù hợp. Các mô hình dự báo cũng cần được huấn luyện nhiều lần để tối ưu hóa tham số và đảm bảo sự tương thích với bộ dữ liệu

#### 5. Tên sản phẩm:

- 01 bài báo đăng trên kỷ yếu hội nghị thuộc IEEE Xplore: Van Khanh Doan; Dong Nguyen Doan; Binh Hau Nguyen, "Improving Energy Demand Prediction Use Deep Learning Network", *9th International Conference on Applying New Technology in Green Buildings (ATiGB)*, 2024, pp.323-327. DOI: 10.1109/ATiGB63471.2024.10717715.

#### 6. Hiệu quả, phương thức chuyển giao kết quả nghiên cứu và khả năng áp dụng:

- Đề tài là tài liệu nghiên cứu trong lĩnh vực dự báo trong hệ thống điện, được chuyển giao trực tiếp cho Bộ môn Hệ Thống Điện, Khoa Điện-điện tử của Trường Đại học Sư phạm Kỹ thuật Đà Nẵng.

TM. Hội đồng Khoa  
Chủ tịch



TS. Trần Hoàng Vũ

Ngày 01 tháng 04 năm 2025  
Chủ nhiệm đề tài



ThS. Doãn Văn Khánh

XÁC NHẬN CỦA TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT  
KT. HIỆU TRƯỞNG

PHÓ HIỆU TRƯỞNG



PGS. TS. Võ Trung Hùng

## INFORMATION ON RESEARCH RESULTS

### 1. General information:

Project title: Evaluating the application of Google Trends in optimizing electricity demand forecasting using artificial intelligence algorithms.

Code number: T2023-06-23

Coordinator: Van Khanh Doan

Implementing institution: The University of Technology and Education, the University of Danang.

Duration: from March 2024 to February 2025

### 2. Objective(s):

Develop a forecasting method for electricity demand by integrating Google Trends data and evaluate its effectiveness in improving forecasting accuracy.

### 3. Creativeness and innovativeness:

While most research has focused on traditional factors such as historical electricity consumption, weather conditions, and electricity prices, this study broadens the scope by incorporating Google Trends as a data source reflecting consumer behavior. This approach introduces a novel perspective in electricity demand forecasting. The findings from utilizing Google Trends data lay the groundwork for the effective integration and utilization of alternative data sources, such as customer feedback and search queries from customer service platforms, which are widely implemented by power utilities but remain underutilized in demand forecasting.

### 4. Research results:

Empirical validation demonstrates that integrating Google Trends data into forecasting models enhances accuracy, as reflected by a reduction in MSE compared to models based solely on electricity demand data. However, its effectiveness is contingent on proper data preprocessing and integration. Not all search trends are relevant, necessitating robust

filtering and keyword selection. Additionally, forecasting models require iterative optimization to fine-tune parameters and ensure alignment with the dataset

### **5. Products:**

01 research paper has been published in proceeding of the IEEE Xplore: Van Khanh Doan, Dong Nguyen Doan, Binh Hau Nguyen, “Improving Energy Demand Prediction Use Deep Learning Network,” 9th International Conference on Applying New Technology in Green Buildings (ATiGB), 2024, pp. 323-327, DOI: 10.1109/ATiGB63471.2024.10717715

### **6. Effects, transfer alternatives of reserach results and applicability:**

The project serves as a research document in the field of forecasting in electric power systems and has been directly transferred to the Department of Electrical Power Systems, Faculty of Electrical and Electronics Engineering at the University of Technology and Education, University of Da Nang.

# MỞ ĐẦU

## 1. Tính cấp thiết của đề tài

Trong bối cảnh sản xuất điện năng ngày càng phụ thuộc vào các nguồn năng lượng tái tạo như gió và mặt trời, sự biến động trong sản xuất điện trở nên nhanh chóng và đáng kể hơn. Điều này làm gia tăng nhu cầu cân bằng năng lượng, đảm bảo sự ổn định giữa sản xuất và tiêu thụ điện. Cân bằng năng lượng đóng vai trò quan trọng trong việc duy trì an ninh năng lượng, và để thực hiện hiệu quả nhiệm vụ này, các phương pháp lập kế hoạch sản xuất điện cần được tối ưu hóa. Dự báo điện năng là một trong những công cụ quan trọng hỗ trợ lập kế hoạch, trong đó khả năng dự báo chính xác giúp nâng cao hiệu quả quản lý năng lượng.

Việc nghiên cứu và áp dụng các mô hình dự báo nhu cầu tiêu thụ điện năng khác nhau nhằm tìm ra phương pháp tối ưu là thực sự cần thiết. Trong bối cảnh trí tuệ nhân tạo (AI) ngày càng được ứng dụng rộng rãi, hai yếu tố then chốt quyết định độ chính xác của dự báo là mô hình dự báo và dữ liệu đầu vào. Các thuật toán AI đã được nghiên cứu và phát triển mạnh mẽ trên thế giới, tạo điều kiện thuận lợi cho việc xây dựng mô hình dự báo. Tuy nhiên, vấn đề tiếp cận và sử dụng dữ liệu đầu vào vẫn là một thách thức lớn, nhất là khi dữ liệu ngày càng trở thành một nguồn tài nguyên quý giá trong nền kinh tế số.

Dữ liệu được ví như "dầu mỏ" mới của thời đại số, đóng vai trò quan trọng trong sự phát triển của AI. Tuy nhiên, việc tiếp cận các nguồn dữ liệu chính thống cũng như tìm kiếm dữ liệu tiềm năng không hề dễ dàng. Đề tài này hướng đến việc khai phá và tận dụng các nguồn dữ liệu phi truyền thống để giải quyết bài toán dự báo nhu cầu tiêu thụ điện năng, trong đó tập trung vào Google Trends như một nguồn dữ liệu tiềm năng.

Với sự phát triển mạnh mẽ của công nghệ, đặc biệt là máy tính và điện thoại thông minh, hành vi tìm kiếm thông tin theo thời gian thực đã trở thành một phần không thể thiếu trong đời sống hiện đại. Người dùng ngày càng dựa vào công cụ tìm kiếm để tra cứu thông tin về sản phẩm, dịch vụ và các vấn đề liên quan đến nhu cầu cá nhân, trong đó có nhu cầu sử dụng điện năng. Google Trends, một nền tảng phân tích dữ liệu tìm kiếm trực tuyến, cung cấp cái nhìn tổng quan về mức độ quan tâm của công chúng đối với

một chủ đề cụ thể theo thời gian và khu vực địa lý. Đã có nhiều nghiên cứu chứng minh rằng dữ liệu từ Google Trends mang lại giá trị cao trong việc dự báo xu hướng của nhiều lĩnh vực khác nhau, từ kinh tế, y tế, tài chính đến tiêu dùng. Đặc biệt, trong bối cảnh nhu cầu điện năng chịu ảnh hưởng từ nhiều yếu tố như biến động thời tiết, sự kiện kinh tế và xu hướng tiêu dùng của người dân, Google Trends nổi lên như một nguồn dữ liệu tiềm năng, phản ánh phần nào hành vi tiêu thụ điện thông qua sự quan tâm của người dùng đối với các thiết bị điện, điều kiện thời tiết và giá điện.

Bằng cách phân tích và khai thác dữ liệu từ Google Trends, có thể phát hiện được những mối liên hệ ẩn giữa hành vi tìm kiếm và nhu cầu sử dụng điện, từ đó cải thiện độ chính xác của các mô hình dự báo. Nghiên cứu này hướng tới việc ứng dụng Google Trends như một yếu tố bổ sung trong dự báo nhu cầu tiêu thụ điện năng, giúp nâng cao hiệu quả quản lý năng lượng, tối ưu hóa vận hành hệ thống điện và hỗ trợ các nhà hoạch định chính sách trong việc đưa ra các quyết định phù hợp nhằm đảm bảo sự cân bằng giữa cung và cầu điện năng trong tương lai.

## **2. Mục tiêu của đề tài**

Nghiên cứu này nhằm phát triển một phương pháp dự báo nhu cầu tiêu thụ điện năng bằng cách kết hợp dữ liệu tìm kiếm từ Google Trends với các nguồn dữ liệu truyền thống như lịch sử tiêu thụ điện. Việc tích hợp Google Trends vào mô hình dự báo được kỳ vọng sẽ giúp khai thác các ‘dữ liệu ẩn’ từ hành vi tìm kiếm trực tuyến của người dùng, qua đó phản ánh xu hướng tiêu thụ điện theo thời gian thực.

Dữ liệu từ Google Trends có tiềm năng bổ sung thông tin quan trọng về sự thay đổi trong hành vi tiêu dùng điện, đặc biệt trong bối cảnh nhu cầu điện bị ảnh hưởng bởi các yếu tố ngoài lưới điện như điều kiện thời tiết bất thường, sự quan tâm đến thiết bị điện hoặc các biến động kinh tế. Vì vậy, nghiên cứu tập trung vào việc xây dựng một khung phương pháp luận giúp tích hợp hợp lý nguồn dữ liệu này vào các mô hình dự báo, từ đó nâng cao khả năng phản ánh chính xác sự thay đổi của nhu cầu điện trong ngắn hạn và trung hạn.

Bên cạnh việc xây dựng phương pháp dự báo, nghiên cứu còn đặt mục tiêu đánh giá vai trò và mức độ ảnh hưởng của Google Trends đối với độ chính xác của mô hình dự báo nhu cầu điện. Cụ thể, nghiên cứu sẽ tiến hành các thí nghiệm thực nghiệm trên tập dữ

liệu thực tế để so sánh hiệu suất dự báo của các mô hình sử dụng dữ liệu truyền thống với các mô hình có tích hợp dữ liệu từ Google Trends. Việc phân tích kết quả sẽ giúp làm rõ liệu dữ liệu tìm kiếm trực tuyến có thể cung cấp thông tin bổ sung có giá trị hay không, cũng như xác định những điều kiện và phương pháp tối ưu để hai thác hiệu quả nguồn dữ liệu này trong bối cảnh dự báo tiêu thụ điện.

Thông qua nghiên cứu này, đề tài kỳ vọng sẽ góp phần mở rộng phạm vi ứng dụng của dữ liệu tìm kiếm trực tuyến trong lĩnh vực quản lý năng lượng, đồng thời cung cấp một hướng tiếp cận mới mang tính liên ngành, kết hợp giữa khoa học dữ liệu, trí tuệ nhân tạo và quản lý hệ thống điện.

### **3. Cách tiếp cận và phương pháp nghiên cứu**

#### **3.1 Cách tiếp cận**

Nghiên cứu được thực hiện dựa trên việc kiểm chứng tính chính xác của các mô hình dự báo thông qua các thí nghiệm thực nghiệm trên bộ dữ liệu thực tế. Quá trình nghiên cứu bao gồm việc xây dựng mô hình, thu thập và xử lý dữ liệu, tiến hành huấn luyện và kiểm tra mô hình trên các tập dữ liệu phù hợp. Các kết quả đạt được sẽ được phân tích, đánh giá nhằm rút ra những kết luận quan trọng về hiệu quả của việc tích hợp dữ liệu Google Trends vào dự báo nhu cầu tiêu thụ điện. Toàn bộ quá trình nghiên cứu cùng với các kết quả, nhận xét và đề xuất sẽ được tổng hợp trong một báo cáo khoa học và một bài báo khoa học công bố kết quả nghiên cứu.

#### **3.2 Phương pháp nghiên cứu**

Phương pháp nghiên cứu tiếp cận trong đề tài này bao gồm :

- Nghiên cứu tài liệu và tổng quan lý thuyết: Thu thập, phân tích các nghiên cứu trước đây liên quan đến dự báo nhu cầu điện năng, các mô hình trí tuệ nhân tạo và ứng dụng của dữ liệu Google Trends trong lĩnh vực năng lượng.
- Xây dựng mô hình dự báo: Phát triển các mô hình dự báo nhu cầu điện năng dựa trên thuật toán bộ nhớ dài ngắn hạn LSTM và mô hình biến đổi Transformer, với và không có sự kết hợp của dữ liệu Google Trends, để so sánh hiệu quả của từng phương pháp.
- Thiết lập kiểm chứng thực nghiệm: Áp dụng các mô hình vào bộ dữ liệu thực tế, thực hiện huấn luyện và kiểm tra trên tập dữ liệu của Vương quốc Anh, sau đó đánh giá hiệu suất dự báo thông qua các chỉ số sai số như MSE, RMSE, MAE.

- Phân tích và đánh giá kết quả: So sánh độ chính xác giữa các mô hình dự báo, rút ra những nhận định về tác động của Google Trends đối với dự báo nhu cầu điện năng, đồng thời đề xuất phương pháp tối ưu trong việc tích hợp nguồn dữ liệu này.

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

##### 4.1 Đối tượng

Nghiên cứu này tập trung vào các đối tượng chính có ảnh hưởng đến dự báo nhu cầu tiêu thụ điện năng, bao gồm:

- Dữ liệu từ Google Trends: Dữ liệu phản ánh xu hướng tìm kiếm của người dùng, được sử dụng để xác định các tín hiệu tiềm ẩn liên quan đến nhu cầu tiêu thụ điện năng.
- Dữ liệu phụ tải điện của Vương quốc Anh: Bao gồm lịch sử nhu cầu tiêu thụ điện theo thời gian, cung cấp thông tin nền tảng cho các mô hình dự báo.
- Mô hình Bộ nhớ dài – ngắn hạn (LSTM): Một dạng mạng nơ-ron hồi quy sâu (RNN) chuyên xử lý dữ liệu chuỗi thời gian, giúp nắm bắt các quan hệ tuần hoàn và xu hướng trong dữ liệu phụ tải điện.
- Mô hình học sâu với Transformers: Một phương pháp tiên tiến có khả năng mô hình hóa mối quan hệ dài hạn trong dữ liệu chuỗi, được đánh giá về hiệu suất so với LSTM trong bài toán dự báo nhu cầu điện năng.

##### 4.2 Phạm vi nghiên cứu

- **Dự báo ngắn hạn:** Nghiên cứu chỉ tập trung vào bài toán dự báo tiêu thụ điện năng trong ngắn hạn, không xem xét các phương pháp dự báo trung hạn và dài hạn.
- **Khu vực nghiên cứu:** Phạm vi địa lý của nghiên cứu là Vương quốc Anh, với các phân tích dựa trên dữ liệu thực tế về tiêu thụ điện tại quốc gia này.

Ngoài ra, một trong những thách thức lớn đối với các nghiên cứu dự báo nhu cầu điện năng, đặc biệt tại các nước đang phát triển như Việt Nam, là vấn đề tiếp cận dữ liệu. Các dữ liệu về tiêu thụ điện thường bị hạn chế hoặc không được công khai đầy đủ, gây khó khăn cho các nhà nghiên cứu muốn xây dựng các mô hình dự báo chính xác. Điều này dẫn đến tình trạng các nghiên cứu bị giới hạn về phạm vi, cũng bởi lẽ

đó đề tài lựa chọn phạm vi nghiên cứu tại Vương quốc Anh với nguồn dữ liệu mở để đảm bảo tính khả thi và độ tin cậy của phân tích.

Việc khai thác các nguồn dữ liệu mới và dễ dàng tiếp cận như Google Trends trở nên đặc biệt quan trọng trong bối cảnh trên. Google Trends không chỉ cung cấp một nguồn dữ liệu phong phú phản ánh xu hướng tìm kiếm của người dùng, mà còn có thể giúp bổ sung những thông tin mà dữ liệu truyền thống chưa cung cấp được. Nếu được tích hợp hiệu quả, nguồn dữ liệu này có thể giúp cải thiện chất lượng dự báo tiêu thụ điện, đặc biệt trong bối cảnh hạn chế dữ liệu chính thống.

## **5. Tính mới của đề tài**

Góc nhìn hành vi tiêu dùng trong dự báo điện : Thay vì chỉ xem xét yếu tố kỹ thuật (công suất, thời tiết, lịch sử tiêu thụ), nghiên cứu đề cập đến cách dữ liệu tìm kiếm trực tuyến phản ánh xu hướng tiêu dùng điện, có thể bao gồm sự quan tâm đến thiết bị điện, thời tiết, và yếu tố kinh tế. Kết quả từ việc sử dụng dữ liệu Google Trends đặt nền tảng cho việc tích hợp và khai thác hiệu quả hơn các nguồn dữ liệu thay thế, chẳng hạn như phản hồi khách hàng và truy vấn tìm kiếm từ các nền tảng dịch vụ khách hàng (CSKH), vốn đang được các công ty điện lực triển khai rộng rãi nhưng chưa được tận dụng đầy đủ trong dự báo nhu cầu điện.

# **CHƯƠNG 1: TỔNG QUAN VỀ TIÊU THỤ VÀ SỬ DỤNG NĂNG LƯỢNG ĐIỆN**

## **1.1 Giới thiệu về nhu cầu tiêu thụ điện trong bối cảnh hiện đại**

Điện là một nguồn năng lượng sạch và an toàn, không trực tiếp tạo ra khí thải ô nhiễm tại điểm sử dụng, khác với các loại nhiên liệu rắn, lỏng hay khí. Từ góc độ sử dụng cuối cùng, việc chuyển đổi từ nhiên liệu hóa thạch hoặc sinh khối sang điện mang lại lợi ích môi trường đáng kể. Ví dụ, hệ thống đường sắt điện phát thải thấp hơn so với phương tiện sử dụng dầu diesel hay than, trong khi xe điện được xem là giải pháp thay thế thân thiện với môi trường so với xe chạy bằng nhiên liệu hóa thạch. Tuy nhiên, lợi ích này chỉ được tối ưu hóa khi quá trình sản xuất và phân phối điện được kiểm soát chặt chẽ nhằm giảm thiểu các tác động tiêu cực đến môi trường. Cùng với sự phát triển của ngành năng lượng trong thập kỷ tới, sẽ có ảnh hưởng sâu rộng đến nhu cầu tiêu thụ điện và môi trường trong dài hạn. Xu hướng gia tăng mức tiêu thụ điện bình quân đầu người, đi kèm với sự mở rộng của các dịch vụ điện khí hóa, đặt ra thách thức trong công tác dự báo nhu cầu tiêu thụ điện. Điều này trở nên đặc biệt quan trọng trong bối cảnh chính sách năng lượng cần dựa trên các mô hình dự báo chính xác để đảm bảo cân bằng cung - cầu và hướng tới phát triển bền vững. Những thách thức trong công tác dự báo nhu cầu tiêu thụ điện gặp nhiều khó khăn do tính không chắc chắn của các yếu tố tác động đến triển vọng năng lượng và điện. Những yếu tố này bao gồm biến động nhân khẩu học, tốc độ tăng trưởng kinh tế, xu hướng xã hội, chính trị, cũng như sự phát triển công nghệ. Trong đó, hành vi tiêu thụ điện là một biến số quan trọng, chịu ảnh hưởng bởi các yếu tố kinh tế, xã hội, công nghệ và chính sách năng lượng. Những thay đổi trong hành vi tiêu dùng có thể làm phức tạp hóa việc dự báo nhu cầu điện, từ đó tác động trực tiếp đến quá trình xây dựng chiến lược phát triển năng lượng. Do đó, để cải thiện độ chính xác trong dự báo, cần xem xét đồng thời cả các yếu tố kỹ thuật và các yếu tố hành vi.

## **1.2 Các yếu tố ảnh hưởng đến nhu cầu tiêu thụ điện**

### **1.2.1 Yếu tố vĩ mô**

- **Yếu tố kỹ thuật:** Chất lượng và hiệu suất của hệ thống lưới điện, công nghệ sản xuất, truyền tải và lưu trữ điện đóng vai trò quan trọng trong việc đảm bảo cung cấp điện ổn định. Các cải tiến công nghệ giúp giảm thất thoát điện và tăng hiệu suất hoạt động của hệ thống, từ đó ảnh hưởng trực tiếp đến mức tiêu thụ điện năng.

- **Yếu tố kinh tế:** Thu nhập của người dân, giá điện, chi phí vận hành và chính sách trợ giá điện đều là những biến số kinh tế then chốt quyết định hành vi tiêu dùng điện. Khi giá điện tăng, người tiêu dùng có xu hướng tìm cách giảm mức tiêu thụ qua việc thay đổi thói quen sử dụng hoặc đầu tư vào các thiết bị tiết kiệm năng lượng.
- **Yếu tố xã hội và văn hóa:** Nhận thức cộng đồng về bảo vệ môi trường, xu hướng tiêu dùng bền vững và các chiến dịch tuyên truyền về tiết kiệm năng lượng có ảnh hưởng trực tiếp đến mức tiêu thụ điện. Các giá trị văn hóa và chuẩn mực xã hội có thể thúc đẩy người dân thay đổi thói quen sử dụng điện, hướng đến việc tiêu dùng điện một cách hợp lý và hiệu quả.
- **Yếu tố môi trường:** Điều kiện khí hậu và biến đổi thời tiết, như mùa hè nóng bức hoặc mùa đông lạnh giá, có tác động rõ rệt đến nhu cầu sử dụng điện cho các mục đích điều hòa, sưởi ấm hoặc làm mát. Những hiện tượng thời tiết cực đoan cũng có thể gây ra những biến động đột ngột trong mức tiêu thụ điện.

### 1.2.2 Hành vi tiêu thụ điện

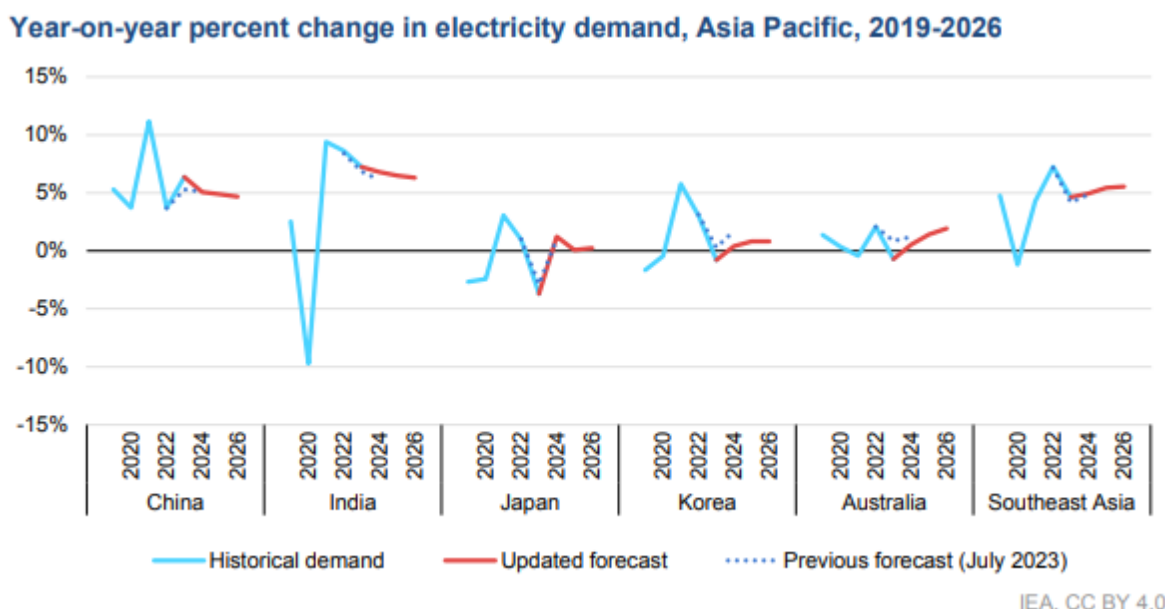
- **Thói quen sử dụng điện:**
  - *Mô hình tiêu thụ theo thời gian:* Mức tiêu thụ điện thường có xu hướng tăng vào buổi sáng khi người dân bắt đầu hoạt động và vào buổi tối khi các hoạt động gia đình được tăng cường. Ở các quốc gia áp dụng mô hình giá điện theo thời gian thực (time-of-use pricing), người tiêu dùng có thể điều chỉnh hoạt động tiêu thụ điện để tối ưu hóa chi phí, qua đó giảm tải cho hệ thống lưới vào các khung giờ cao điểm.
  - *Sử dụng thiết bị điện:* Hành vi sử dụng các thiết bị điện, như việc để đèn sáng qua đêm, sử dụng điều hòa liên tục hay không tắt các thiết bị khi không cần thiết, có thể làm gia tăng mức tiêu thụ điện. Ngược lại, các chiến dịch tiết kiệm điện và nâng cao nhận thức có thể khuyến khích người tiêu dùng sử dụng điện một cách hiệu quả hơn.
- **Tác động của công nghệ:**

- *Thiết bị tiết kiệm năng lượng:* Sự ra đời của các thiết bị điện hiệu suất cao như đèn LED, điều hòa inverter và tủ lạnh tiết kiệm điện đã giúp giảm mức tiêu thụ trên mỗi đơn vị thiết bị. Tuy nhiên, hiện tượng hiệu ứng hồi phục (rebound effect) có thể xảy ra khi chi phí vận hành giảm dẫn đến việc tiêu thụ nhiều hơn, làm giảm bớt hiệu quả tiết kiệm.
- *Hệ thống nhà thông minh và IoT:* Ứng dụng công nghệ thông tin và truyền thông trong hệ thống nhà ở thông qua IoT và các giải pháp nhà thông minh giúp tự động điều chỉnh mức tiêu thụ điện dựa trên nhu cầu sử dụng thực tế, từ đó tối ưu hóa mức tiêu thụ. Mặc dù công nghệ này giúp quản lý điện năng hiệu quả, nhưng cũng đồng thời tạo ra sự phụ thuộc cao hơn vào thiết bị điện và hệ thống tự động hóa.
- **Ảnh hưởng của giá điện và chính sách:**
  - *Biến động giá điện:* Khi giá điện tăng, người tiêu dùng có xu hướng điều chỉnh hành vi tiêu dùng như sử dụng thiết bị tiết kiệm năng lượng hoặc hạn chế tiêu thụ điện vào những thời điểm có giá cao. Các chính sách giá linh hoạt có thể tác động đáng kể đến mô hình tiêu thụ điện của từng hộ gia đình.
  - *Chính sách hỗ trợ và khuyến khích:* Các chương trình trợ cấp cho điện mặt trời, xe điện và các giải pháp năng lượng tái tạo đang dần thay đổi cách thức tiêu thụ điện. Ngoài ra, các chiến dịch nâng cao nhận thức cộng đồng như “Giờ Trái Đất” cũng góp phần thúc đẩy hành vi tiết kiệm điện.
- **Yếu tố xã hội và môi trường:**
  - *Nhận thức về môi trường:* Sự gia tăng nhận thức về bảo vệ môi trường dẫn đến việc nhiều người tiêu dùng ưu tiên lựa chọn nguồn điện từ năng lượng tái tạo. Điều này không chỉ thay đổi cơ cấu tiêu thụ mà còn thúc đẩy việc áp dụng các biện pháp tiết kiệm điện.
  - *Xu hướng đô thị hóa:* Sự gia tăng dân số tại các khu vực đô thị làm thay đổi nhu cầu sử dụng điện, với mức tiêu thụ cao hơn do nhu cầu sinh hoạt và công nghiệp.

### 1.3 Xu hướng phát triển của ngành điện và các thách thức trong dự báo

Sự phát triển của ngành điện trong những thập kỷ qua đã chứng kiến những bước chuyển mình quan trọng, đặc biệt là với sự gia tăng của các nguồn năng lượng tái tạo như điện mặt trời, điện gió và điện sinh khối. Xu hướng này không chỉ góp phần giảm thiểu tác động môi trường của ngành năng lượng mà còn tạo ra nhiều thách thức mới đối với công tác dự báo nhu cầu tiêu thụ điện.

Trước hết, sự gia tăng mức tiêu thụ điện bình quân đầu người, kết hợp với quá trình điện khí hóa các ngành công nghiệp và dịch vụ, đòi hỏi các mô hình dự báo phải linh hoạt và thích ứng với những thay đổi nhanh chóng của thị trường. Các mô hình dự báo truyền thống dựa vào dữ liệu lịch sử và các biến số kinh tế - xã hội có thể không phản ánh kịp thời các yếu tố mới như xu hướng sử dụng công nghệ nhà thông minh hay sự thay đổi trong hành vi tiêu dùng do các chiến dịch tiết kiệm điện.



Hình 1.1 Mức thay đổi phần trăm theo năm trong nhu cầu điện, khu vực Châu Á - Thái Bình Dương, giai đoạn 2019-2026.

Thêm vào đó, các sự kiện toàn cầu như khủng hoảng năng lượng, biến động giá nhiên liệu và những thay đổi về chính sách năng lượng cũng gây ra mức độ không chắc chắn cao trong việc dự báo nhu cầu điện. Những biến động này yêu cầu các mô hình dự báo không chỉ dựa vào các dữ liệu định lượng truyền thống mà còn cần tích hợp thêm các nguồn dữ liệu thời gian thực, giúp cung cấp cái nhìn toàn diện và kịp thời về xu hướng tiêu thụ điện. Trong bối cảnh đó, việc khai thác các nguồn dữ liệu phi truyền thống như

Google Trends được coi là một giải pháp tiềm năng nhằm bổ sung thông tin về hành vi tiêu dùng và mức độ quan tâm của công chúng đối với các vấn đề tiết kiệm năng lượng. Sự kết hợp giữa dữ liệu truyền thống và dữ liệu thời gian thực từ các công cụ phân tích tìm kiếm giúp cải thiện độ chính xác của các mô hình dự báo, từ đó hỗ trợ các nhà hoạch định chính sách đưa ra các chiến lược phù hợp để cân bằng cung – cầu và thúc đẩy phát triển bền vững trong ngành điện.

Tóm lại, xu hướng phát triển của ngành điện và những thách thức trong việc dự báo nhu cầu tiêu thụ điện đòi hỏi một cách tiếp cận đa chiều, tích hợp các yếu tố kỹ thuật, kinh tế, xã hội và môi trường. Việc hiểu rõ các yếu tố ảnh hưởng đến nhu cầu sử dụng điện không chỉ giúp cải thiện các mô hình dự báo mà còn đóng vai trò quan trọng trong việc xây dựng các chính sách năng lượng hiệu quả.

## **CHƯƠNG 2 : PHƯƠNG PHÁP LUẬN GIẢI QUYẾT BÀI TOÁN DỰ BÁO TIÊU THỤ NĂNG LƯỢNG ĐIỆN**

### **2.1 Tổng quan về bài toán dự báo**

Dự báo là sự kết hợp giữa khoa học và nghệ thuật trong việc dự đoán những sự kiện có khả năng xảy ra trong tương lai. Tính khoa học của dự báo thể hiện ở việc quá trình dự báo cần dựa trên các dữ liệu phản ánh tình hình thực tế trong quá khứ và hiện tại, đồng thời dựa trên các xu hướng được phân tích theo nền tảng khoa học để dự đoán những sự kiện có thể xảy ra trong tương lai. Tính nghệ thuật của dự báo nằm ở khả năng phán đoán và vận dụng kinh nghiệm trong những tình huống thiếu thông tin hoặc khi nhu cầu của khách hàng biến động mạnh. Chính từ những yếu tố này, có thể thấy rằng dự báo luôn có sai số, và việc dự đoán chính xác hoàn toàn chỉ xảy ra ngẫu nhiên. Độ chính xác của dự báo cũng giảm dần khi thời gian dự báo kéo dài.

Mục tiêu dự báo có thể khác nhau đối với từng đối tượng, nhưng nhìn chung bao gồm các bước sau:

- Xác định đối tượng cần dự báo: Xác định chỉ số cần dự báo, đồng thời xem xét các giả định liên quan và các biến số áp dụng.
- Thu thập dữ liệu: Bước này bao gồm việc thu thập các dữ liệu cần thiết.
- Chọn phương pháp dự báo: Lựa chọn một kỹ thuật dự báo phù hợp nhất, đồng thời xem xét các biến số, giả định và tập dữ liệu liên quan. Thông thường có 2 phương pháp dự báo chính :
  - Phương pháp định lượng: Sử dụng các mô hình toán học, thống kê, hoặc máy học để dự báo dựa trên dữ liệu lịch sử. Ví dụ: hồi quy tuyến tính, ARIMA, hoặc mạng nơ-ron nhân tạo.
  - Phương pháp định tính: Áp dụng khi dữ liệu lịch sử không đủ hoặc không phù hợp, dựa trên ý kiến chuyên gia, khảo sát, hoặc phân tích tình huống.
- Xây dựng mô hình dự báo: Dữ liệu được phân tích bằng cách sử dụng phương pháp đã chọn, và một dự báo được xây dựng từ phân tích này. Các biến đầu vào quan trọng được lựa chọn và tiền xử lý để đảm bảo tính nhất quán và chất lượng dữ liệu. Sau đó, mô hình được huấn luyện, kiểm định và hiệu chỉnh nhằm tối ưu hóa độ chính xác, trước khi áp dụng vào thực tế để tạo ra dự báo đáng tin cậy

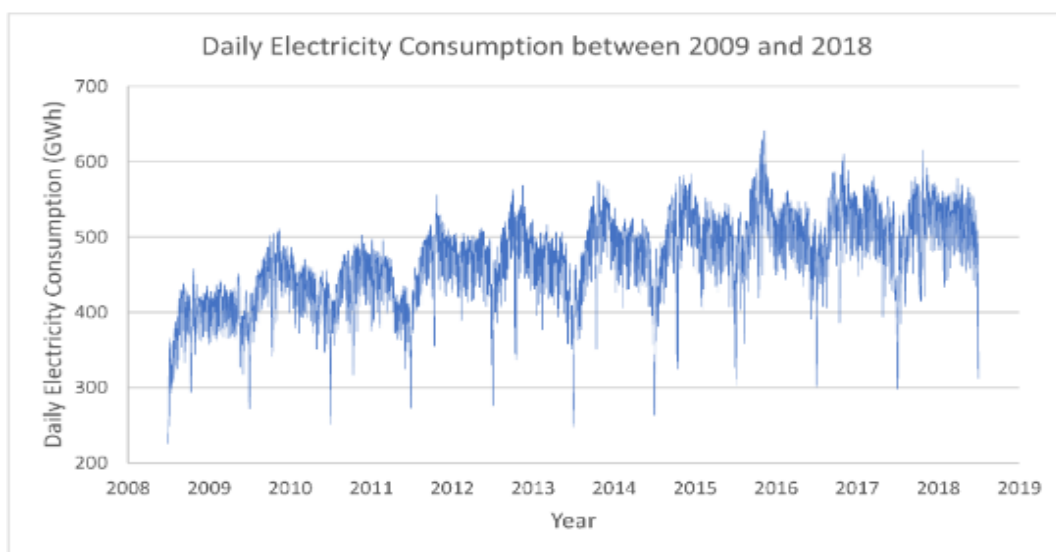
- Xác minh dự báo: Đánh giá độ chính xác của dự báo bằng các chỉ số như MAE (Mean Absolute Error), RMSE (Root Mean Square Error), hoặc MAPE (Mean Absolute Percentage Error)...
- Trình bày dự báo: Có thể sử dụng trực quan hóa dữ liệu để biểu diễn dự báo dưới dạng hình ảnh mà các bên liên quan có thể hiểu rõ hơn và áp dụng trong quá trình ra quyết định.

Với bài toán dự báo nhu cầu tiêu thụ năng lượng điện, mục tiêu cốt lõi là dự báo lượng điện tiêu thụ trong tương lai. Tùy thuộc vào mục đích của dự báo, khoảng thời gian dự báo có thể dao động từ ngắn hạn short-term (theo giờ, theo ngày) phục vụ điều độ lưới điện, trung hạn medium-term (theo tuần, theo tháng) hỗ trợ lập kế hoạch cung ứng, đến dài hạn long-term (theo năm) để định hướng phát triển hạ tầng và chính sách năng lượng. Dữ liệu tiêu thụ điện thu nhập là dữ liệu có tính chất chuỗi thời gian và có sự phụ thuộc vào nhiều yếu tố bên ngoài, (được gọi là các biến ngoại sinh - Exogenous Factors, là những yếu tố như kinh tế, xã hội, hành vi tiêu thụ đề cập chương 1). Những biến ngoại sinh (thời tiết, khí hậu, hành vi tiêu thụ, ..), chính những yếu tố này giải thích tính phi tuyến trong việc tiêu thụ điện của một cá nhân, một tổ chức hay một quốc gia. Bởi lẽ đó, dự báo chính xác lượng điện năng tiêu thụ được coi là một bài toán khó bởi phụ thuộc vào nhiều biến số "không chắc chắn".

Dữ liệu tiêu thụ điện là dữ liệu chuỗi thời gian, có các đặc trưng của một chuỗi thời gian là tính xu hướng và tính chu kỳ. Theo [1-3], dữ liệu thời gian thực hay chuỗi thời gian là một chuỗi các giá trị (dữ liệu dạng số) của một đại lượng được ghi nhận là thời gian, được đánh chỉ mục theo thời gian. Do đó, dự báo lượng điện năng tiêu thụ là đi ước tính các điểm dữ liệu số rời rạc cho các chỉ số thời gian trong tương lai. Các giá trị của chuỗi thời gian của đại lượng  $X$  được kí hiệu là  $X_1, X_2, X_3, \dots, X_t, \dots, X_n$  với  $X_t$  là giá trị của  $X$  tại thời điểm  $t$ . Hình 2-1 là biểu đồ minh họa chuỗi thời gian sản lượng điện tiêu thụ với đặc trưng xu hướng và chu kỳ. Dữ liệu chuỗi thời gian thường chịu ảnh hưởng của nhiều yếu tố như thời tiết, chính sách năng lượng, biến động kinh tế và hành vi tiêu dùng của người dân. Trong đó, xu hướng phản ánh sự thay đổi dài hạn của nhu cầu điện, có thể tăng lên do sự phát triển kinh tế hoặc giảm xuống do các biện pháp tiết kiệm năng

lượng. Bên cạnh đó, tính chu kỳ thể hiện qua các mô hình lặp lại theo ngày, tuần, tháng hoặc theo mùa, phụ thuộc vào đặc điểm sinh hoạt và sản xuất của từng khu vực.

Ngoài xu hướng và chu kỳ, dữ liệu tiêu thụ điện còn có thể chứa các thành phần ngẫu nhiên do các yếu tố bất thường như thiên tai, khủng hoảng kinh tế hoặc thay đổi đột ngột trong chính sách năng lượng. Những yếu tố này có thể làm gián đoạn mô hình tiêu thụ điện, gây ra những biến động khó lường trong dự báo. Do đó, việc lựa chọn mô hình dự báo phù hợp cần cân nhắc đến các thành phần này để đảm bảo độ chính xác cao nhất.



*Hình 2.1 Đặc trưng xu hướng và chu kỳ của dữ liệu kiểu thời gian*

## 2.2 Mô hình dự báo chuỗi thời gian

Phần này sẽ cung cấp một cái nhìn tổng quan về các kỹ thuật nổi bật nhất được sử dụng trong Dự báo Chuỗi Thời gian, trong phần tiếp theo tập trung vào các mô hình được áp dụng chính trong nghiên cứu dự báo tiêu thụ điện năng của phạm vi đề tài.

### **Làm Trơn Hàm Mũ (Exponential Smoothing)**

Làm trơn hàm mũ là một phương pháp được đề xuất bởi Brown, Holt và Winters vào những năm 1950 [1-3]. Ý tưởng đằng sau phương pháp này là dự báo điểm dữ liệu tiếp theo trong chuỗi thời gian bằng cách tính trung bình có trọng số của các quan sát trong quá khứ. Các trọng số được gán sẽ giảm dần theo hàm mũ khi khoảng cách thời gian đến điểm mục tiêu tăng lên. Quan sát càng gần thời điểm hiện tại thì trọng số được gán càng cao. Phương pháp này tương đối đơn giản và có thể nhanh chóng áp dụng cho

nhiều loại chuỗi thời gian khác nhau, dẫn đến tầm quan trọng lớn của nó trong các ứng dụng công nghiệp [4].

Có ba phương pháp áp dụng Làm trơn hàm mũ (Exponential Smoothing) với độ phức tạp tăng dần. Phương pháp đơn giản nhất được gọi là Làm trơn hàm mũ đơn giản (Simple Exponential Smoothing), được sử dụng cho các chuỗi thời gian không có xu hướng hoặc mô hình mùa rõ ràng. Dự báo  $y$  cho chỉ số thời gian  $t$  được tính bằng trung bình có trọng số của các quan sát trong quá khứ với trọng số giảm theo hàm mũ:

$$\hat{y}_t = \alpha y_t + (1-\alpha)y_{t-1} + \alpha(1-\alpha)y_{t-2} + \alpha(1-\alpha)^2 y_{t-3} + \dots \quad (2.1)$$

Trong đó, tham số làm trơn  $0 \leq \alpha \leq 1$  kiểm soát tốc độ giảm của các trọng số. Với  $\alpha$  lớn hơn, mô hình sẽ nhạy cảm hơn với các quan sát gần đây. Tổng này bao gồm tất cả các quan sát trong quá khứ từ  $y_1$  đến  $y_{t-1}$ .

Ngoài ra, các phương trình của *Làm trơn hàm mũ* có thể được viết dưới dạng thành phần (component form), giúp mở rộng sang các dạng phức tạp hơn của làm trơn hàm mũ. Trong cách tiếp cận này, chuỗi thời gian được biểu diễn dưới dạng tổng của các thành phần khác nhau [5]. Đối với *Làm trơn hàm mũ đơn giản*, chỉ có một thành phần duy nhất được xem xét, đó là mức độ  $\ell$ . Hạn chế của *Làm trơn hàm mũ đơn giản* là nó không phù hợp với dữ liệu có xu hướng hoặc biến đổi theo mùa.

Để giải quyết vấn đề này, phương pháp xu hướng tuyến tính của Holt hay còn *làm trơn hàm mũ kép (double exponential smoothing)*, được sử dụng [4]. Phương pháp này mở rộng từ *làm trơn hàm mũ đơn giản* bằng cách thêm một thành phần xu hướng  $b$  và một tham số làm trơn thứ hai  $0 \leq \beta \leq 1$  cho xu hướng này, bên cạnh thành phần mức độ  $\ell$  và tham số  $\alpha$  đã biết.

## ARIMA

Phương pháp tiếp cận của các mô hình *ARIMA* khác với các mô hình *Làm trơn hàm mũ*. Trong *ARIMA*, các thành phần như mức độ, xu hướng và mùa không được mô hình hóa riêng biệt. Thay vào đó, xu hướng và mùa được loại bỏ bằng cách lấy sai phân của dữ liệu, để thu được một chuỗi thời gian được coi là *dừng (stationary)* [6]. Một chuỗi thời gian được gọi là *dừng* nếu các đặc tính thống kê của nó, chẳng hạn như giá trị trung bình (mean) hoặc phương sai (variance), không phụ thuộc vào thời gian. Do đó, một chuỗi

thời gian có xu hướng (trend) hoặc tính thời vụ (seasonality) sẽ không phải là chuỗi dừng.

Một cách để biến một chuỗi thời gian thành dừng là loại bỏ xu hướng và tính thời vụ bằng cách lấy sai phân (*differencing*). Cụ thể:

- Đối với xu hướng, sai phân được tính bằng cách lấy chênh lệch giữa các quan sát liên tiếp.
- Đối với tính thời vụ, sai phân được tính bằng cách lấy chênh lệch giữa các quan sát cách nhau một khoảng thời gian bằng một chu kỳ mùa.

Việc lấy sai phân có thể giúp giảm hoặc thậm chí loại bỏ xu hướng và tính thời vụ, từ đó làm ổn định giá trị trung bình của chuỗi thời gian và khiến nó trở nên (hoặc gần như) dừng. Ngoài ra, các phép biến đổi khác, chẳng hạn như lấy *logarit* của chuỗi thời gian, cũng có thể giúp ổn định phương sai của chuỗi thời gian [5].

Nhìn chung, ARIMA kết hợp ba quá trình thành một mô hình lớn hơn:

1. Tự hồi quy (Autoregression - AR): Sử dụng mối quan hệ giữa các giá trị trong quá khứ để dự đoán giá trị tương lai.
2. Trung bình trượt (Moving Average - MA): Sử dụng sai số dự đoán trong quá khứ để cải thiện dự đoán hiện tại.
3. Lấy sai phân (Differencing - I): Loại bỏ xu hướng và mùa bằng cách lấy sai phân của dữ liệu, sau đó tích hợp (integrate) lại để đưa dữ liệu về dạng ban đầu.

Trong phần mô hình *Tự hồi quy (Autoregression - AR)*, ý tưởng chính là sử dụng một mô hình *Hồi quy Tuyến tính (Linear Regression)* đơn giản với các giá trị trong quá khứ làm biến dự đoán. Do đó, dự đoán cho  $y_t$  là một tổ hợp tuyến tính của các quan sát trong quá khứ của chính chuỗi thời gian đó. Đây là lý do tại sao nó được gọi là *Tự hồi quy*. Đối với hồi quy này, một bậc  $p$  cần được xác định, đại diện cho phạm vi "nhìn lại" (*lookback range*), tức là số lượng các quan sát trong quá khứ sẽ được đưa vào mô hình hồi quy. Một mô hình Tự hồi quy bậc  $p$  thường được ký hiệu là  $AR(p)$  và có thể được viết dưới dạng:

$$\hat{y}_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + \varepsilon_t \quad (2.2)$$

Trong đó:

- $\Phi_i$  là các trọng số (weights) tương ứng với các giá trị trong quá khứ.
- $c$  là một hằng số (constant)

$\varepsilon_t$  là nhiễu trắng (white noise), đại diện cho các yếu tố ngẫu nhiên không thể dự đoán được.

Mô hình AR(p) sử dụng các giá trị trong quá khứ từ  $y_{t-1}$  đến  $y_{t-p}$  để dự đoán giá trị hiện tại  $y_t$ , và các trọng số  $\phi_i$  phản ánh mức độ ảnh hưởng của từng giá trị quá khứ lên giá trị hiện tại.

Ý tưởng của *Trung bình trượt (Moving Average - MA)* tương tự nhưng khác biệt so với *Hồi quy*. Thay vì sử dụng các giá trị trong quá khứ của mục tiêu hồi quy, một mô hình Trung bình trượt bậc  $q$  (hay còn gọi là  $MA(q)$ ) sử dụng các sai số dự đoán trong quá khứ làm biến dự đoán trong mô hình *Hồi quy Tuyến tính*. Điều này dẫn đến phương trình dự báo:

$$\hat{y}_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (2.3)$$

Trong đó:

- $\theta_i$  là các trọng số tương ứng với các sai số dự đoán trong quá khứ.
- $c$  là một hằng số.
- $\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q}$  là các sai số dự đoán tại các thời điểm tương ứng.

Phương trình hồi quy cho một mô hình *ARIMA* có thể được viết như sau:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2.4)$$

Trong đó:

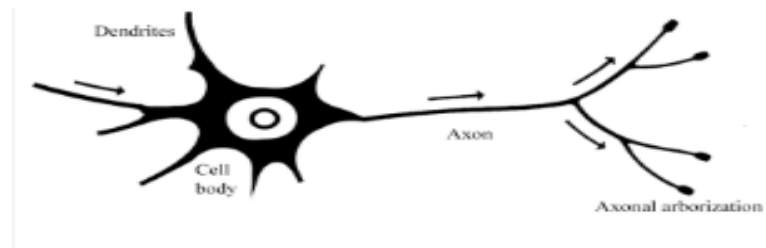
- $y_t'$ : Đại diện cho các giá trị của chuỗi thời gian đã được lấy sai phân.
- $c$ : Hằng số
- $\phi_i$ : Trọng số của các giá trị trong quá khứ (thành phần AR).
- $\theta_i$ : Trọng số của các sai số dự đoán trong quá khứ (thành phần MA).

- $\varepsilon_t$ : Nhiều trắng tại thời điểm  $t$ .

## Mô hình học máy và trí tuệ nhân tạo

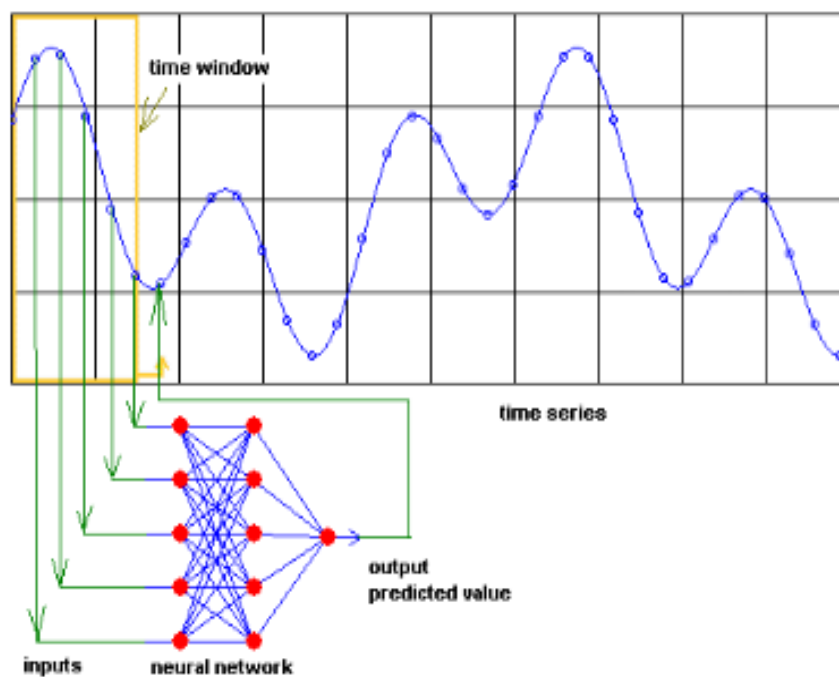
Trong khi hầu hết các mô hình thống kê được xem xét ở trên đều giả định rằng dữ liệu tuân theo một cấu trúc cơ bản có thể được mô tả bằng một mối quan hệ toán học, các mô hình dựa trên *Học máy và Trí tuệ nhân tạo* lại sử dụng ít giả định hơn. Chúng đơn giản mong đợi mô hình sẽ đưa ra các giá trị giống nhau khi nhận được cùng một đầu vào. Vì điều này không bao gồm bất kỳ yêu cầu nào về cách mối quan hệ giữa đầu vào và đầu ra trông như thế nào, hầu hết các mô hình *Học máy và Trí tuệ nhân tạo* có thể xử lý được các quan hệ phi tuyến tính. Dự báo chuỗi thời gian có thể được diễn đạt như một bài toán học có giám sát. Đối với các trạng thái mới chưa biết  $X_t$ , nhãn tương ứng  $y_t$  cần được dự đoán. Các tham số của mô hình được học thông qua *Học máy* bằng cách lặp đi lặp lại việc điều chỉnh giá trị các tham số cho đến khi một hàm mất mát nhất định được tối thiểu hóa [7]. Việc này thường được thực hiện bằng cách chia dữ liệu có sẵn thành ba tập: Tập huấn luyện, tập kiểm định và tập kiểm tra. Mô hình được huấn luyện trên tập huấn luyện và đánh giá trên tập thứ hai - tập kiểm định. Với kết quả đánh giá nhận được, một số quyết định về thiết kế mô hình có thể được đưa ra trước khi thực hiện đánh giá cuối cùng về hiệu suất của mô hình trên tập kiểm tra. Việc tối thiểu hóa hàm mất mát trong quá trình huấn luyện sẽ giúp mô hình hóa chính xác phân phối của dữ liệu, đồng thời độ phức tạp của mô hình nên được giữ ở mức nhỏ để tránh hiện tượng overfitting (quá khớp) dữ liệu huấn luyện. Khi lựa chọn mô hình, nguyên lý "Ockham's razor" cần được áp dụng. Nguyên lý này nói rằng nếu có nhiều giả thuyết được xem xét, thì giả thuyết đơn giản nhất thường là giả thuyết đúng. Các đặc tính phi tuyến tính và phức tạp trong quan hệ giữa các yếu tố ảnh hưởng đến chuỗi thời gian tiêu thụ điện, do đó các mô hình AI, đặc biệt là mạng nơ-ron nhân tạo (Artificial Neural Networks – ANN), đã trở thành một công cụ mạnh mẽ để giải quyết bài toán dự báo. Mạng nơ-ron nhân tạo hoạt động giống như bộ não của con người, được học bởi kinh nghiệm (thông qua việc huấn luyện), lưu giữ các tri thức sau đó sử dụng các tri thức đó trong việc dự đoán các dữ liệu chưa biết. Một mạng nơ-ron bao gồm nhiều nút nối với nhau, mô phỏng mạng nơ-ron thần kinh của não người. Mạng nơ-ron nhân tạo bao gồm ba thành phần cơ bản: mô hình của nơ-ron, cấu trúc và sự liên kết giữa các nơ-ron. Trong một số trường

hợp, mạng nơ-ron nhân tạo là một hệ thống thích ứng, tự thay đổi cấu trúc của mình dựa trên các thông tin bên ngoài hay bên trong chạy qua mạng trong quá trình học.



Hình 2.2 Cấu tạo của tế bào nơron sinh học tiền đề xây dựng mạng nơ-ron nhân tạo

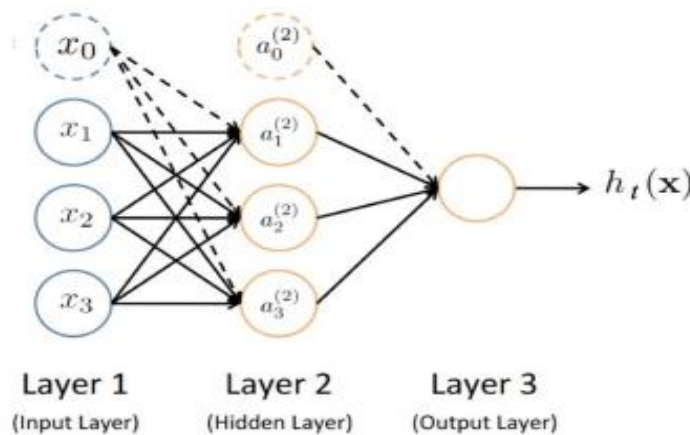
Tiếp theo, sẽ trình bày các lý thuyết cơ bản của mạng nơ-ron trước khi đi sâu vào xây dựng các mô hình mạng nơ-ron nhân tạo tiếp theo ở chương 3, cách chúng có thể được sử dụng để dự báo chuỗi thời gian, cũng như những ưu điểm và thách thức khi áp dụng chúng vào thực tế.



Hình 2.3 Mô hình mạng nơ-ron nhân tạo trong dự báo chuỗi thời gian

Kiến trúc chung của một ANN gồm 3 thành phần đó là lớp đầu vào (input layer), lớp ẩn (hidden layer) và lớp đầu ra (output layer). Dữ liệu từ bên ngoài sẽ đi vào mạng nơ-ron nhân tạo qua lớp đầu vào. Các nút đầu vào xử lý, phân tích, phân loại dữ liệu sau đó chuyển sang lớp tiếp theo, trong đó các lớp trung gian (lớp ẩn) kết nối đầu vào và đầu ra để mô hình hóa các mối quan hệ phi tuyến tính. Mỗi đơn vị trong một lớp sử dụng một cách biểu diễn dữ liệu cụ thể; ví dụ, đối với dữ liệu chuỗi thời gian, lớp đầu vào có thể tương ứng với một vector các giá trị số hoặc một ma trận chứa dữ liệu phụ trợ.

Hình 2.4 Mô hình mạng nơ-ron nhân tạo đơn giản



Với một kiến trúc mạng nơ-ron nhất định, chúng ta cần tìm các trọng số chính xác để kết nối các nơ-ron khác nhau (còn được gọi là nút, tế bào hoặc đơn vị), trong đó các nơ-ron được biểu diễn bằng các vòng tròn như trên hình. Có nhiều phương pháp để huấn luyện mạng nơ-ron, chẳng hạn như lan truyền ngược (backpropagation), thuật toán di truyền (genetic algorithms) hoặc tìm kiếm ngẫu nhiên (random search)...; tất cả các phương pháp này đều nhằm tìm ra các tham số tối ưu cho mạng nơ-ron. Các mô hình học sâu ngày nay phần lớn là sử dụng lan truyền ngược.

*Lan truyền ngược* là một thành phần quan trọng trong một phương pháp huấn luyện mạng nơ-ron cơ bản nhưng mạnh mẽ. Trên một lô dữ liệu huấn luyện, sau khi khởi tạo ngẫu nhiên các ma trận trọng số:

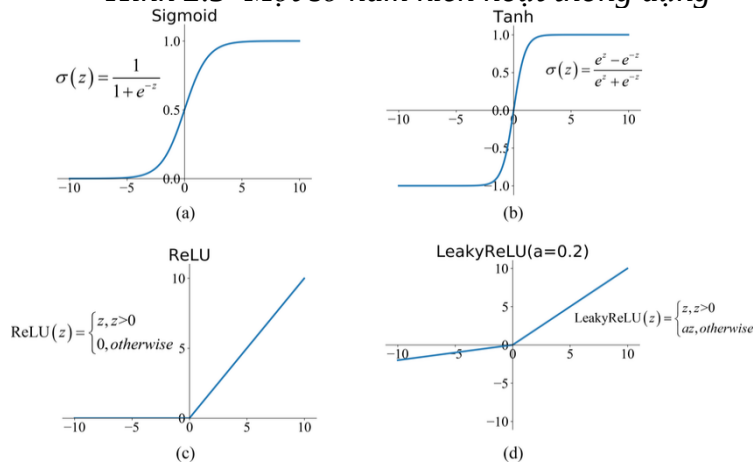
1. Tính toán hàm mất mát bằng lan truyền tiến (feed-forward propagation);
2. Tạo gradient của hàm mất mát bằng lan truyền ngược (backpropagation);
3. Cập nhật trọng số của mạng nơ-ron bằng phương pháp gradient descent;

Hàm mất mát có thể được hiểu là mức độ sai lệch của dự đoán so với thực tế; một mô hình hoàn hảo có hàm mất mát bằng không, trong khi nhiều lỗi dự báo hơn đồng nghĩa với mất mát lớn hơn. Các phương pháp học máy thường hướng đến việc tối thiểu hóa hàm mất mát. Đối với lớp số  $i$  và đơn vị  $j$  (trong mỗi lớp), ký hiệu tiêu chuẩn được sử dụng để tạo trọng số trong mạng nơ-ron như sau. Gọi  $z$  là đầu ra,  $w$  là trọng số và  $b$  là độ chệch (bias). Dưới đây, chúng ta định nghĩa hàm đầu ra thô của mỗi đơn vị trong mỗi lớp như sau:

$$z_j^{(i)} = w_j^{(i)}x + bz_j^{(i)} \quad (2.5)$$

Quá trình lan truyền tiến (forward propagation) diễn ra từ trái sang phải của mạng nơ-ron, như minh họa trong Hình 2.4. Giá trị đầu ra thô được tính cho mỗi đơn vị theo phương trình trên, được hiểu là tổng có trọng số của các giá trị đầu vào của đơn vị đó. Sau đó, một hàm kích hoạt thường được áp dụng để biến đổi phi tuyến tính giá trị đầu ra thô (trong tất cả các lớp trừ lớp đầu vào); điều này có thể được hiểu như một ngưỡng kích hoạt, sao cho giá trị  $z_j^{(i)}$  được "kích hoạt" ở một mức nhất định. Các hàm kích hoạt phổ biến bao gồm hàm tanh, sigmoid, ReLU..

Hình 2.5 Một số hàm kích hoạt thông dụng



Bây giờ định nghĩa hàm kích hoạt:  $a_j^{(i)} = g(z_j^{(i)})$ . Giả sử hàm kích hoạt là sigmoid ta

có 
$$g(z_j^{(i)}) = \frac{1}{1 + e^{-z}}$$

Sau khi hoàn thành một vòng lan truyền tiến, chúng ta thu được giá trị kích hoạt  $a$ , được xem như đầu ra dự đoán  $\hat{y}$ . Giá trị này sẽ được so sánh với đầu ra thực tế  $y$ , được biểu diễn dưới dạng  $h_t(x)$  tại lần lặp  $t$ .

Từ đây, chúng ta tính toán một giá trị duy nhất gọi là *hàm mất mát*, thể hiện mức độ sai lệch giữa dự đoán và thực tế. Có nhiều loại hàm mất mát khác nhau, nhưng đối với các bài toán hồi quy, hàm mất mát trung bình bình phương (*Mean Squared Error - MSE*) là một trong những lựa chọn phổ biến nhất. Công thức của MSE trên  $n$  mẫu dữ liệu như sau

$$L(y, \hat{y}) = \frac{\sum_{i=1}^n ((y_i - \hat{y}_p)^2)}{n} \quad (2.6)$$

Tiếp theo, trong bước lan truyền ngược, các đạo hàm (gradient) được tính ngược lại, từ phải sang trái qua các lớp của mạng nơ-ron, như minh họa trong Hình 2.2. Về thực tế, chúng ta làm việc theo hướng ngược lại để sử dụng kết quả tính toán mất mát trước đó và xác định những sai sót nào trong mạng nơ-ron đã dẫn đến độ chính xác thấp hơn.

Mỗi đạo hàm theo trọng số  $w$  được tính bằng quy tắc chuỗi như sau

$$\frac{\partial L(w)}{\partial w} = \frac{\partial L(w)}{\partial a} \times \frac{\partial a}{\partial z} \times \frac{\partial z}{\partial w} \quad (2.7)$$

Cuối cùng, các trọng số  $w$  của mạng nơ-ron được cập nhật bằng phương pháp gradient descent, được xác định với tốc độ học  $\alpha$  như sau:

$$w \leftarrow w - \alpha \frac{\partial L(w)}{\partial w} \quad (2.8)$$

Chúng ta lặp lại quá trình và cập nhật trọng số cho đến khi mô hình hội tụ. Có nhiều cách để cải thiện từ quy trình cơ bản này. Ví dụ, kỹ thuật *dropout* có thể được sử dụng để giảm hiện tượng *quá khớp* bằng cách loại bỏ ngẫu nhiên một số neurons với một xác suất nhất định.

Tốc độ học  $\alpha$  có thể được cố định hoặc điều chỉnh theo từng bước lặp bằng các phương pháp như Adam, Adagrad, momentum.

Để minh họa rõ hơn cách hoạt động của mạng nơ-ron nhân tạo trong dự báo nhu cầu điện, ta xét một ví dụ cụ thể. Giả sử chúng ta có bộ dữ liệu nhu cầu tiêu thụ điện hàng giờ trong 10 giờ liên tiếp như sau :

| Thời gian<br>(giờ) | Tiêu thụ điện (kW) |
|--------------------|--------------------|
| 1                  | 100                |
| 2                  | 110                |
| 3                  | 120                |
| 4                  | 130                |
| 5                  | 140                |
| 6                  | 150                |
| 7                  | 160                |
| 8                  | 170                |
| 9                  | 180                |
| 10                 | 190                |

Mục tiêu của chúng ta là sử dụng các giá trị tiêu thụ điện trong quá khứ để dự đoán giá trị tiêu thụ điện trong tương lai. Ví dụ, sử dụng 5 giá trị trước đó để dự đoán giá trị tiếp theo (dải cửa sổ time window hay lookback range ).

Tiếp sẽ tạo các cặp dữ liệu đầu vào và đầu ra từ chuỗi thời gian. Với dải cửa sổ trượt là 5, tức là sử dụng 5 giá trị trước đó để dự đoán giá trị tiếp theo, dữ liệu sẽ được chia như sau:

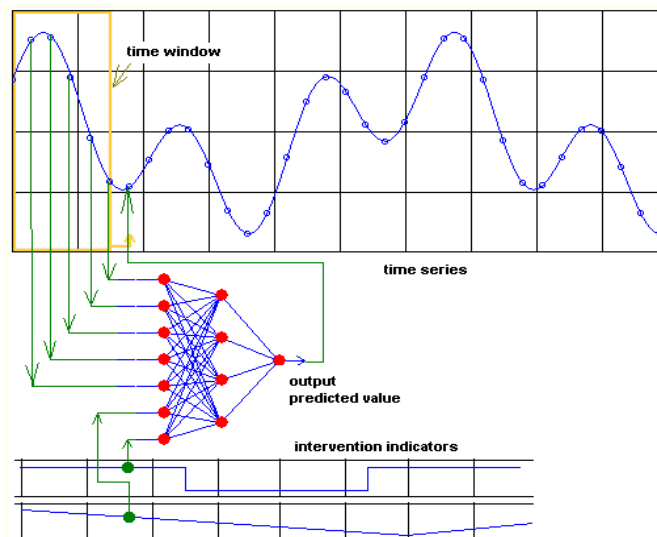
| (X)                       | (y) |
|---------------------------|-----|
| [100, 110, 120, 130, 140] | 150 |
| [110, 120, 130, 140, 150] | 160 |
| [120, 130, 140, 150, 160] | 170 |

| (X)                       | (y) |
|---------------------------|-----|
| [130, 140, 150, 160, 170] | 180 |
| [140, 150, 160, 170, 180] | 190 |

Chúng ta sẽ xây dựng một mạng nơ-ron đơn giản với các lớp sau :

- Lớp đầu vào: Số lượng nút tương ứng với dải cửa sổ  $s = 5$ .
- Lớp ẩn : Một lớp ẩn với 10 nút và hàm kích hoạt ReLU.
- Lớp đầu ra : Một nút đầu ra để dự đoán giá trị tiêu thụ điện

Tiếp đến khởi tạo một ma trận trọng số bất kỳ, độ lệch tương ứng với giá trị đầu vào tương ứng, và tính toán cho từng lớp ẩn dựa trên các bước miêu tả ở trên. Bằng cách huấn luyện với nhiều mẫu, mô hình sẽ học được quy luật biến động của tiêu thụ điện và dự báo giá trị tiếp theo dựa trên dữ liệu lịch sử.



Hình 2.6 Dự báo chuỗi thời gian sử dụng mạng nơ-ron kết hợp với các biến can thiệp

Chúng ta có thể bổ sung thêm thông tin vào chuỗi thời gian, chẳng hạn các biến ngoại sinh. Thông tin này sau đó có thể được sử dụng để tối ưu mô hình, dự báo chính xác hơn. Một cách phổ biến để thêm thông tin quan trọng là sử dụng các biến can thiệp như Hình 2.6 (chỉ báo can thiệp), một dạng đặc biệt của biến ngoại sinh, biểu thị thông tin về chuỗi thời gian hoặc bối cảnh của khoảng thời gian cần dự báo.

Ví dụ, khi dự báo mức tiêu thụ năng lượng, việc biết rằng chúng ta đang dự báo cho thứ Hai hay thứ Bảy có thể cải thiện đáng kể độ chính xác, vì hành vi tiêu thụ điện thay đổi theo ngày trong tuần. Thông tin này không có sẵn trực tiếp từ chuỗi thời gian và cần được bổ sung vào mô hình.

Ngoài ra, các biến ngoại sinh khác như nhiệt độ, thời tiết đã được nghiên cứu và chứng minh có thể giúp cải thiện dự báo, chúng phản ánh những yếu tố bên ngoài tác động đến nhu cầu điện. Do đó, việc kết hợp cả chỉ báo can thiệp và các biến ngoại sinh có thể giúp mô hình dự báo hoạt động hiệu quả hơn.

Ví dụ miêu tả ở trên về cách xây dựng mạng nơ-ron nhân tạo để giải quyết bài toán dự báo chỉ khai triển ở bước khởi tạo. Trên thực tế, việc xây dựng một mạng nơ-ron đòi hỏi nhiều vòng lặp tính toán nhằm tìm ra các tham số tối ưu tương ứng với từng bộ dữ liệu. Quá trình này sẽ rất khó thực hiện trên một bộ dữ liệu lớn, phức tạp và cần thiết các mô hình học sâu tối ưu hơn một mạng nơ-ron nhân tạo thông thường cùng với sự hỗ trợ của các công cụ tính toán mạnh mẽ.

Một số mô hình phổ biến nhất của mạng nơ-ron nhân tạo, trong đó có một số mô hình sẽ được nghiên cứu và thực nghiệm ở phần tiếp theo của đề tài.

. • Các mạng nơ-ron sử dụng nhiều lớp nơ-ron theo kiến trúc cơ sở của Multi-layer perceptron [8] nhằm xử lý các dữ liệu dạng bảng. Bao gồm mạng nơ-ron học sâu truyền thẳng, mạng nơ-ron sâu và rộng hay mạng nơ-ron phân tích ma trận nhân tử...

• Mạng nơ-ron tích chập (CNN ) [8]-11]: Đây là một loại mạng nơ-ron cải tiến được chuyên biệt hóa, thường được sử dụng cho xử lý hình ảnh và video. Nó được thiết kế để tự động xác định các đặc trưng trong hình ảnh, chẳng hạn như cạnh, góc và các cấu trúc khác bằng cách áp dụng bộ lọc tích chập vào dữ liệu đầu vào.

• Mạng nơ-ron hồi quy (RNN ) [12-14]: Mạng nơ-ron tiến hóa được thiết kế để xử lý các đầu vào có dạng dữ liệu tuần tự, chẳng hạn như giọng nói, văn bản và dữ liệu chuỗi thời gian. RNN có khả năng bắt các phụ thuộc thời gian trong dữ liệu đầu vào, làm cho nó phù hợp cho nhiều ứng dụng, chẳng hạn như dịch ngôn ngữ, nhận dạng giọng nói và tạo âm nhạc.

- Mạng nơ-ron biến đổi (Transformer) [15-18]: Kiến trúc mạng Transformer được giới thiệu như một phương án thay thế kiến trúc mạng RNN trong việc xử lý dữ liệu tuần tự, chẳng hạn như ngôn ngữ. Hạn chế của mô hình RNN là không xử lý song song được chuỗi dữ liệu đầu vào và không thể nắm bắt được sự phụ thuộc dài hạn. Thay vì kết nối tuần tự của RNN, Transformer sử dụng lớp tự chú ý cho phép mô hình nắm bắt sự phụ thuộc giữa các thành phần khác nhau trong chuỗi đầu vào. Điều này khiến cho Transformer có thể được thực hiện song song và cho phép nó xử lý các chuỗi đầu vào dài hơn.

## **2.3 Một số vấn đề với mô hình dự báo dựa trên mạng nơ-ron học sâu**

### **1. Thiếu Dữ Liệu Đủ Lớn**

- Một số mạng nơ-ron học sâu yêu cầu lượng dữ liệu lớn để học được các mẫu phức tạp. Nếu dữ liệu lịch sử quá ít, mô hình có thể overfit hoặc không học được xu hướng thực sự.
- Chuỗi thời gian thường có tính tuần hoàn theo năm, nếu chỉ có dữ liệu trong một khoảng thời gian ngắn (ví dụ: vài tháng), mô hình sẽ khó học được xu hướng dài hạn.

### **2. Không Xử Lý Đúng Định Dạng Dữ Liệu**

- Mạng nơ-ron học sâu yêu cầu dữ liệu đầu vào ở định dạng phù hợp. Nếu không xử lý đúng cách (ví dụ: không chuẩn hóa dữ liệu, không loại bỏ outliers), mô hình có thể học sai phân phối thực tế.
- Với các chuỗi thời gian có giá trị thay đổi theo mùa, việc không thực hiện điều chỉnh theo mùa có thể khiến mô hình khó hội tụ.

### **3. Không Sử Dụng Biến Ngoại Sinh Hợp Lý**

- Một số mô hình DL chỉ học từ chuỗi thời gian mà không tận dụng biến ngoại sinh như thời tiết, ngày trong tuần, giá nhiên liệu,...
- Khi không sử dụng các biến này, mô hình có thể bỏ lỡ những tác động quan trọng từ môi trường bên ngoài.

### **4. Chọn Mô Hình Quá Phức Tạp So Với Nhiệm Vụ**

- DL mạnh nhưng không phải lúc nào cũng tốt hơn các mô hình truyền thống như ARIMA, XGBoost.
- Nếu chuỗi thời gian có cấu trúc đơn giản, dùng các mạng học sâu có thể gây overfitting và không mang lại lợi ích đáng kể so với mô hình truyền thống

#### 5. Hiện tượng quá khớp

- Hiện tượng quá khớp: Một số điểm dữ liệu huấn luyện có thể dự đoán quá tốt, khiến mô hình có thể quá khớp với chúng thay vì tổng quát hóa, dẫn đến độ chính xác thấp hơn trên tập kiểm tra (test set).

## CHƯƠNG 3: XÂY DỰNG CÁC MÔ HÌNH DỰ BÁO

### 3.1 Phương pháp tiếp cận

Đề tài tập trung hướng đến chủ đề dự báo nhu cầu điện ngắn hạn trong khoảng thời gian ngày tiếp theo và mô hình dự báo được xây dựng dựa trên các mô hình mạng nơ-ron học sâu. Dự báo nhu cầu điện ngắn hạn có tính ứng dụng rộng rãi đối với các công ty điện lực, thị trường năng lượng và người tiêu dùng.

Nhìn chung, có hai cách tiếp cận để cải thiện tối ưu dự báo:

#### 1. Tìm kiếm dữ liệu tốt hơn

- Với cách tiếp cận này, tập trung vào việc cải thiện chất lượng, số lượng hoặc tính đa dạng của dữ liệu đầu vào cho mô hình dự báo.
- Ví dụ: Thêm các nguồn dữ liệu mới (như Google Trends trong nghiên cứu này), làm sạch dữ liệu, hoặc thu thập thêm dữ liệu hữu ích khác.
- Việc có dữ liệu tốt hơn giúp mô hình học được các mẫu hình phức tạp hơn và đưa ra dự báo chính xác hơn.

#### 2. Tìm kiếm mô hình tốt hơn

- Với cách tiếp cận này, tập trung vào việc phát triển hoặc lựa chọn các thuật toán, kiến trúc mô hình tiên tiến, tối ưu hơn để xử lý dữ liệu.
- Ví dụ: Sử dụng các mô hình học sâu và đánh giá giữa các mô hình để xem mô hình nào hoạt động tối ưu hơn từ đó lựa chọn.
- Mô hình tốt hơn sẽ có khả năng học sâu hơn, nắm bắt được các mối quan hệ phức tạp trong dữ liệu và đưa ra dự báo chính xác hơn.

Trong nghiên cứu này, sẽ tiếp cận song song hai phương án trên. Cụ thể, việc tìm kiếm dữ liệu tốt hơn tập trung vào việc tích hợp dữ liệu Google Trends vào mô hình dự báo nhu cầu tiêu thụ điện. Xem xét nghiên cứu đánh giá hai mô hình Transformer và LSTM. Transformer vốn được sử dụng rộng rãi trong Xử lý Ngôn ngữ Tự nhiên (NLP) để dự đoán dữ liệu chuỗi [18]. Mặc dù các mô hình Transformer có tiềm năng đáng kể trong NLP, việc ứng dụng chúng trong dự báo chuỗi thời gian vẫn còn chưa được ứng dụng rộng.

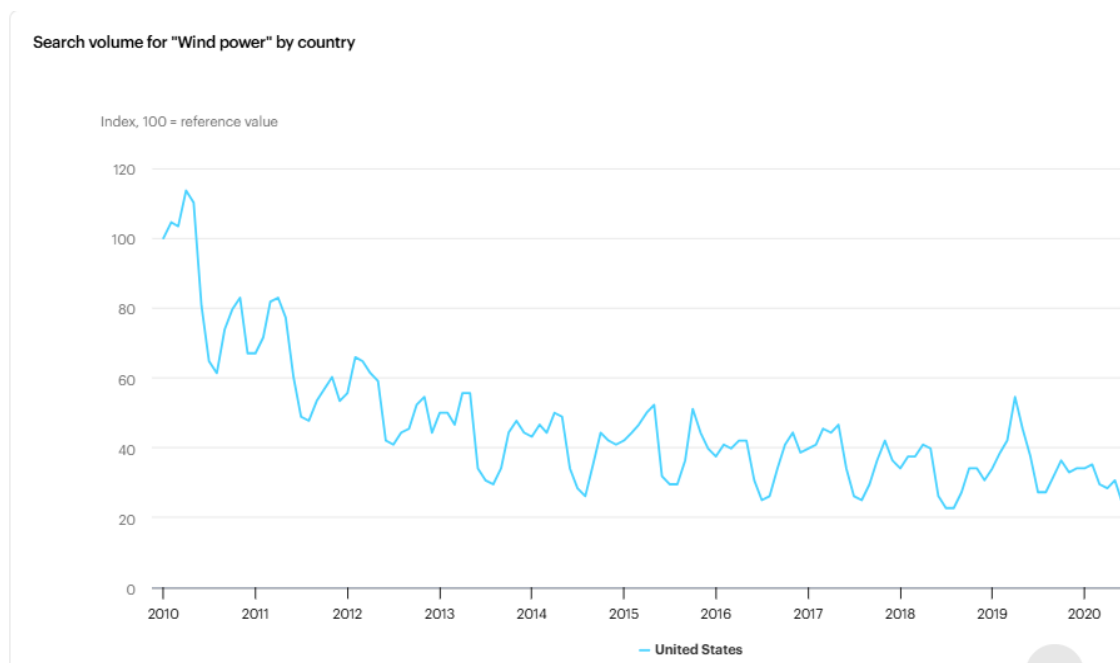
Việc tích hợp dữ liệu Google Trends trong dự báo nhu cầu năng lượng không phải là một khái niệm mới. Google Trends đã được giới thiệu như một công cụ dự báo hữu ích cho các dự đoán của thị trường dầu mỏ [19–22]. Tuy nhiên, trong lĩnh vực thị trường điện, đặc biệt là dự báo nhu cầu tiêu thụ điện, tiềm năng của nó vẫn chưa được khai thác. Trong các ngành công nghiệp sinh lợi như thị trường điện, rõ ràng một cải thiện nhỏ trong độ chính xác dự báo cũng có thể mang lại một tác động đáng kể [23–25].

## 3.2 Dữ liệu Google Trend

### 3.2.1 Giới thiệu chung

Google Trends là một công cụ phân tích trực tuyến do Google Inc. cung cấp, cho phép nghiên cứu và phân tích tần suất cũng như sự phân bố của các truy vấn tìm kiếm trên nền tảng Google Search. Công cụ này cung cấp quyền truy cập vào một tập dữ liệu lớn về khối lượng tìm kiếm, có thể được phân loại theo vị trí địa lý, khoảng thời gian, cũng như danh mục chủ đề cụ thể. Nhờ vào khả năng cung cấp dữ liệu theo thời gian thực với chi phí thấp, Google Trends đã trở thành một nguồn dữ liệu lớn (big data) tiềm năng trong nhiều lĩnh vực nghiên cứu, đặc biệt là trong các mô hình dự báo.

*Hình 3.1 Tìm kiếm theo từ khóa*



Google Trends cung cấp hai phương pháp tiếp cận chính để thu thập và phân tích dữ liệu: (1) tìm kiếm theo từ khóa cụ thể và (2) tìm kiếm theo danh mục. Mỗi phương pháp có ưu điểm và hạn chế riêng, ảnh hưởng trực tiếp đến tính chính xác và mức độ liên quan của dữ liệu thu được.

Trong phương pháp tìm kiếm theo từ khóa cụ thể, hệ thống truy xuất tất cả các truy vấn tìm kiếm có chứa từ khóa đã nhập, bất kể ngữ cảnh hoặc lĩnh vực liên quan. Chẳng hạn, khi tìm kiếm thuật ngữ “wind power”, dữ liệu trả về sẽ bao gồm tất cả các truy vấn có chứa từ này, bất kể người dùng có quan tâm đến khía cạnh năng lượng gió hay đang tìm hiểu về các vấn đề liên quan đến năng lượng gió. Do đó, phương pháp này có thể gặp vấn đề về tính chính xác nếu từ khóa có nhiều nghĩa khác nhau hoặc có sự thay đổi về mặt ngữ nghĩa theo thời gian.

Ngược lại, tìm kiếm theo danh mục giúp khắc phục phần nào vấn đề trên bằng cách giới hạn phạm vi phân tích trong một danh mục cụ thể được định sẵn trong Google Trends. Hiện nay, nền tảng này cung cấp hơn 1.200 danh mục khác nhau, bao gồm nhưng không giới hạn ở các lĩnh vực như Năng lượng, Tài chính, Bất động sản, Làm đẹp & Thể hình, và Công nghệ. Khi một từ khóa được tìm kiếm trong phạm vi của một danh mục, Google Trends sẽ chỉ trả về các truy vấn có liên quan đến danh mục đó, giúp giảm thiểu rủi ro sai lệch ngữ nghĩa theo thời gian. Điều này đặc biệt quan trọng đối với các nghiên cứu đòi hỏi tính nhất quán về bối cảnh và độ chính xác của dữ liệu. Bên cạnh hai phương pháp tìm kiếm trên, Google Trends còn cung cấp hai tính năng quan trọng giúp hỗ trợ quá trình thu thập và xử lý dữ liệu:

1. **Truy vấn liên quan (Related Queries):** Đây là tập hợp các truy vấn tìm kiếm khác có mức độ liên quan cao đến từ khóa ban đầu, giúp mở rộng phạm vi nghiên cứu và cung cấp thêm góc nhìn về hành vi tìm kiếm của người dùng.
2. **Chủ đề liên quan (Related Topics):** Chức năng này tự động nhóm các truy vấn có nội dung tương tự với từ khóa chính, đồng thời xử lý các lỗi chính tả, biến thể ngôn ngữ và cách diễn đạt khác nhau của cùng một thuật ngữ.

Nhờ hai tính năng trên, Google Trends giúp tối ưu hóa việc thu thập dữ liệu bằng cách giảm thiểu các sai sót do chính tả hoặc sự khác biệt trong cách sử dụng ngôn ngữ. Điều này đặc biệt hữu ích khi phân tích dữ liệu trong các lĩnh vực có nhiều thuật ngữ đồng nghĩa hoặc khi nghiên cứu xu hướng tìm kiếm theo thời gian.

Tuy nhiên, một điểm hạn chế đáng kể của hai tính năng này là Google không cung cấp thông tin chi tiết về cách chúng được xây dựng hoặc thuật toán phía sau quá trình phân nhóm dữ liệu. Điều này khiến chúng trở thành một "hộp đen" (black box), khiến các nhà nghiên cứu gặp khó khăn trong việc đánh giá mức độ chính xác và tin cậy của thông tin

thu thập được. Việc không thể xác định rõ liệu một danh mục hoặc một chủ đề liên quan có thực sự phản ánh đúng ý định tìm kiếm của người dùng hay không có thể dẫn đến sai lệch trong diễn giải kết quả.

### **3.2.2 Tiêu chí chọn từ khóa khi sử dụng Google Trend trong nghiên cứu khoa học**

Do Google Trends không được thiết kế chuyên biệt cho mục đích nghiên cứu khoa học, việc đảm bảo tính hợp lệ và mức độ tin cậy của dữ liệu đòi hỏi một quy trình lựa chọn từ khóa nghiêm ngặt. Trong bối cảnh nghiên cứu về xu hướng tìm kiếm liên quan đến năng lượng điện (electricity) hoặc các dạng năng lượng khác khác, việc lựa chọn từ khóa cần đáp ứng các tiêu chí sau:

1. **Mức độ liên quan:** Các từ khóa được lựa chọn phải có mối liên hệ chặt chẽ với chủ đề nghiên cứu. Chẳng hạn, khi nghiên cứu về giá điện, các từ khóa như "tăng giá điện" (electricity price increseas) hoặc "tăng cao" (price strike) có thể được xem xét thay vì chỉ sử dụng từ "price" đơn thuần, để tránh tình trạng dữ liệu bị pha loãng bởi những truy vấn không liên quan.
2. **Tính nhất quán về ngữ nghĩa theo thời gian:** Từ khóa cần duy trì ý nghĩa nhất quán trong suốt khoảng thời gian nghiên cứu. Điều này đặc biệt quan trọng đối với các nghiên cứu dài hạn, nơi sự thay đổi về ngữ nghĩa có thể gây sai lệch trong phân tích dữ liệu.
3. **Khối lượng tìm kiếm đủ lớn:** Từ khóa được lựa chọn cần có lượng tìm kiếm đủ lớn để đảm bảo dữ liệu có ý nghĩa thống kê. Những từ khóa quá chuyên biệt hoặc ít được tìm kiếm có thể dẫn đến dữ liệu không đủ mạnh hoặc có tính biến động cao, gây khó khăn trong việc diễn giải kết quả.
4. **Tránh sự mơ hồ về ngữ nghĩa:** Các từ khóa có nhiều nghĩa khác nhau trong các ngữ cảnh khác nhau cần được loại bỏ để tránh nhầm lẫn trong quá trình phân tích. Ví dụ, từ "load" không chỉ được sử dụng trong lĩnh vực năng lượng (electricity load) mà còn xuất hiện trong các công trình kỹ thuật (structural load-tải trọng), do đó cần xem xét bối cảnh cụ thể để đảm bảo tính chính xác của dữ liệu thu thập.

5. **Bao gồm các từ đồng nghĩa và biến thể:** Để đảm bảo tính toàn diện của dữ liệu, các thuật ngữ đồng nghĩa hoặc các biến thể phổ biến của từ khóa chính cần được đưa vào quá trình tìm kiếm. Ví dụ, khi nghiên cứu về tiết kiệm điện, có thể bổ sung các từ khóa như “tiết kiệm điện” (energy saving, electricity saving), “Hiệu quả năng lượng” (energy efficiency), “Giảm tiêu thụ điện” (reduce electricity consumption), “Giảm hóa đơn điện” (reduce electricity bills), “Sử dụng năng lượng hiệu quả” (efficient energy use)

### 3.2.3 Từ khóa tìm kiếm trong phạm vi đề tài

*Cách tiếp cận dữ liệu Google Trend để nghiên cứu trong đề tài:*

Wilhite et al. (1996) [26] nhấn mạnh rằng nhu cầu tiêu thụ năng lượng không chỉ phụ thuộc vào các yếu tố kỹ thuật hay kinh tế, mà còn bị ảnh hưởng đáng kể bởi bối cảnh văn hóa, khí hậu, cũng như hệ thống hạ tầng của từng khu vực. Trong nghiên cứu này, chúng tôi tập trung vào khu vực Vương quốc Anh và sử dụng dữ liệu từ Google Trends với truy vấn "*Efficient Energy Use*" từ năm 2021 đến 2024 để phản ánh mối quan tâm của công chúng đối với tiết kiệm năng lượng.

Bối cảnh nghiên cứu này đặc biệt quan trọng khi xét đến tác động của cuộc khủng hoảng năng lượng toàn cầu trong những năm gần đây. Kể từ năm 2021, thế giới đã phải đối mặt với tình trạng giá khí đốt tăng cao do xung đột Nga - Ukraine, dẫn đến chi phí năng lượng tăng đáng kể. Điều này đã thúc đẩy sự quan tâm mạnh mẽ đến các biện pháp tiết kiệm điện trên phạm vi toàn cầu. Riêng tại Vương quốc Anh, số lượt tìm kiếm với từ khóa "*save electricity*" đã tăng hơn bốn lần vào năm 2022, phản ánh rõ ràng mối quan tâm ngày càng gia tăng của công chúng.

Hiện tượng này có thể được lý giải dựa trên *Lý thuyết Hành vi Hoạch định* (Theory of Planned Behavior - TPB) [27], trong đó thái độ cá nhân, các chuẩn mực xã hội, và khả năng kiểm soát hành vi đóng vai trò quan trọng trong việc hình thành ý định tiết kiệm điện cũng như thực hiện hành vi này trên thực tế. Cụ thể:

- Thái độ: Một cá nhân có xu hướng tiết kiệm điện khi họ nhận thức được lợi ích của hành vi này, chẳng hạn như giảm chi phí hóa đơn điện hoặc góp phần bảo vệ môi trường.

- Chuẩn mực xã hội: Khi tiết kiệm điện trở thành một chủ đề được quan tâm rộng rãi trong cộng đồng, cá nhân có nhiều khả năng tuân theo xu hướng chung để duy trì sự hòa nhập xã hội.
- Kiểm soát hành vi: Khi công nghệ hỗ trợ tiết kiệm điện (ví dụ: thiết bị tiết kiệm năng lượng hoặc hệ thống giám sát điện thông minh) trở nên phổ biến và dễ tiếp cận hơn, người dân có khả năng cao hơn trong việc thay đổi thói quen tiêu thụ điện của họ.

Các nghiên cứu trước đây cũng chỉ ra rằng yếu tố kinh tế là động lực quan trọng ảnh hưởng đến quyết định tiết kiệm năng lượng [25]. Khi giá điện tăng cao, người tiêu dùng có xu hướng điều chỉnh hành vi của họ nhằm giảm thiểu tác động tài chính. Điều này giải thích lý do tại sao các tìm kiếm liên quan đến tiết kiệm điện lại tăng vọt trong thời kỳ khủng hoảng năng lượng năm 2022 tại Vương quốc Anh nói riêng và các quốc gia EU nói chung.

Trong nghiên cứu này, chúng tôi đưa ra giả thuyết rằng mức độ quan tâm của công chúng đến vấn đề năng lượng, thể hiện qua dữ liệu Google Trends, có tương quan với mức độ tham gia vào các biện pháp tiết kiệm điện, qua đó ảnh hưởng vấn đề tiêu thụ điện. Nếu giả thiết này đúng thì nó sẽ giúp ích trong việc cải thiện vấn đề dự báo tiêu thụ năng lượng điện. Dù việc định lượng mức độ quan tâm của công chúng là một thách thức, nhưng Google Trends có thể được xem như một chỉ báo (proxy) có giá trị, giúp cung cấp thông tin kịp thời về xu hướng tiêu thụ điện và hỗ trợ tối ưu hóa các mô hình dự báo năng lượng.

### 3.3 Mô hình Bộ nhớ dài ngắn hạn LSTM

Mục tiêu chính trong dự báo chuỗi thời gian là dự đoán giá trị tương lai dựa trên dữ liệu trong quá khứ. Tuy nhiên, dữ liệu chuỗi thời gian có một đặc điểm quan trọng: *tính phụ thuộc theo thời gian*, nghĩa là giá trị tại một thời điểm bị ảnh hưởng bởi các giá trị trước đó. Hầu hết các mô hình truyền thống như hồi quy tuyến tính hay mạng nơ-ron truyền thẳng (Feedforward Neural Network) không thể xử lý tốt các mối quan hệ phụ thuộc theo thời gian dài.

LSTM (Long Short-Term Memory) là một loại mạng nơ-ron hồi tiếp RNN, được thiết kế đặc biệt để xử lý dữ liệu có mối quan hệ dài hạn như chuỗi thời gian. Ví dụ: Nếu

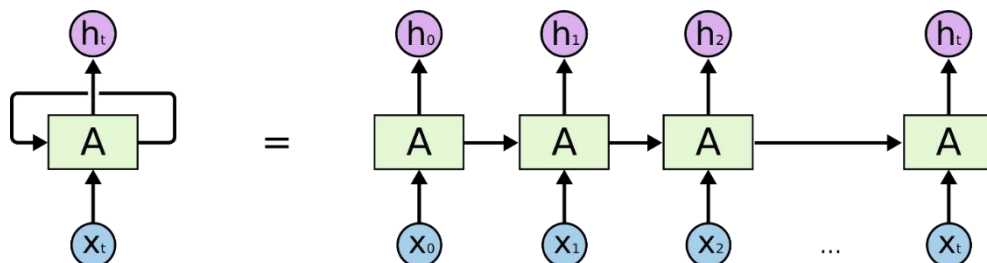
chúng ta muốn dự báo nhu cầu điện năng mỗi giờ. Nhu cầu điện lúc 12 giờ trưa có thể bị ảnh hưởng không chỉ bởi 11 giờ trưa mà còn bởi cả xu hướng từ sáng đến trưa. Một mô hình thông thường có thể chỉ nhìn vào một số giờ gần nhất, nhưng LSTM có thể "ghi nhớ" xu hướng dài hơn, chẳng hạn từ ngày hôm trước. Với RNN, khi mô hình học qua nhiều bước thời gian (ví dụ: 100 giờ trước), các thông tin từ xa dần dần bị mất đi. Hiểu đơn giản, giả sử chúng đang học một bài giảng kéo dài 2 giờ. Nếu giáo viên chỉ nhấn mạnh nội dung của 5 phút trước, nhưng lại quên nhắc đến một khái niệm quan trọng từ 1 giờ trước, thì khi làm bài kiểm tra, chúng ta có thể sẽ bị thiếu thông tin. RNN gặp khó khăn khi cần sử dụng thông tin từ quá khứ xa hơn (chẳng hạn, dự báo điện tiêu thụ dựa vào dữ liệu từ tuần trước). Điều này đã được giải thích trong nghiên cứu của Hochreiter và Bengio, et al. (1994), chính là trong quá trình huấn luyện, chúng ta sử dụng thuật toán lan truyền ngược để cập nhật trọng số. Nếu các giá trị gradient quá nhỏ khi truyền ngược qua nhiều bước thời gian, trọng số gần như không được cập nhật, khiến mô hình quên đi các thông tin quan trọng từ quá khứ xa hơn. Đó là lý do LSTM ra đời để xử lý vấn đề mất mát thông tin (vanish gradient) của RNN.

Hình vẽ trên mô tả một đoạn của mạng nơ-ron hồi quy A với đầu vào là  $x_t$  và đầu ra là  $h_t$ . Một vòng lặp cho phép thông tin có thể được truyền từ bước này qua bước khác của mạng nơ-ron. Một mạng nơ-ron hồi quy có thể được coi là nhiều bản sao chép của cùng một mạng, trong đó mỗi đầu ra của mạng này là đầu vào của một mạng sao chép khác.

Hình 3.2 Mạng nơ-ron hồi quy RNN

#### Cách hoạt động của LSTM khác biệt:

- LSTM có một thành phần quan trọng gọi là "Cell State"  $C_t$  (trạng thái ô nhớ), có thể

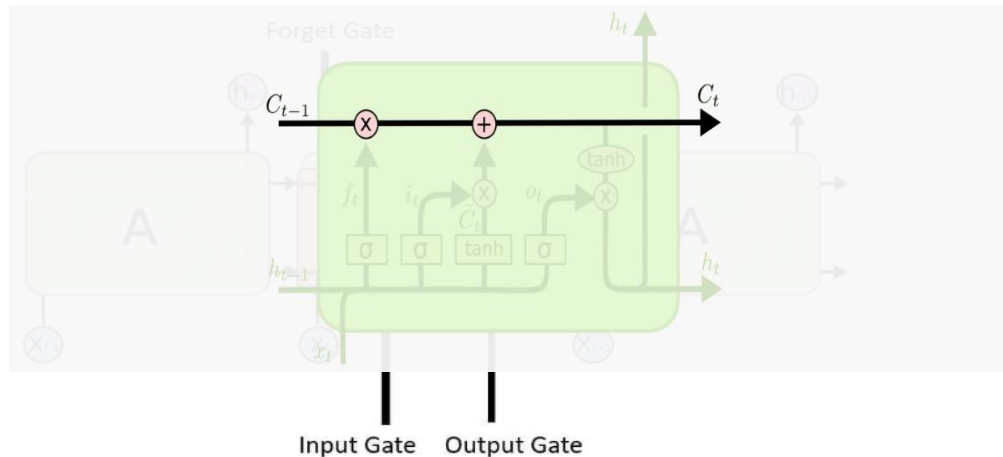


xem như một đường cao tốc thông tin, cho phép thông tin từ quá khứ đi qua nhiều bước thời gian mà không bị mất mát quá nhiều.

- Nhờ cơ chế này, LSTM có thể ghi nhớ thông tin quan trọng từ rất lâu trước đó, ngay cả khi chuỗi dữ liệu kéo dài hàng trăm bước thời gian.

*Hình 3.3 Thành phần kiến trúc cơ bản của LSTM*

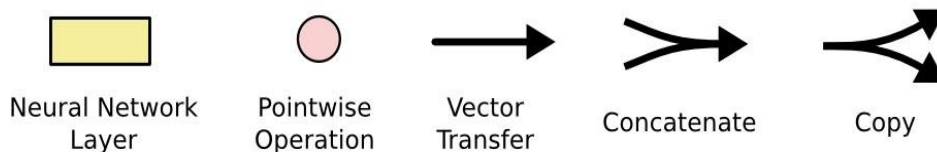
- Forget Gate:** Quyết định thông tin nào cần loại bỏ từ cell state.
- Input Gate:** Quyết định thông tin mới nào cần thêm vào cell state.
- Cập nhật Cell State:** Kết hợp thông tin từ forget gate và input gate để cập nhật cell state.



- Output Gate:** Quyết định thông tin nào từ cell state cần được đưa ra làm đầu ra.
- Hidden state:** Tính toán trạng thái ẩn dựa trên Cell State và Output gate

*Hình 3.4 Các ký hiệu trong hình vẽ*

Ở sơ đồ trên, mỗi một đường mang một véc-tơ từ đầu ra của một nút tới đầu vào của một nút khác. Các hình trong màu hồng biểu diễn các phép toán như phép cộng véc-tơ chẳng hạn, còn các ô màu vàng được sử dụng để học trong các tầng mạng nơ-ron. Các đường hợp nhau kí hiệu việc kết hợp, còn các đường rẽ nhánh ám chỉ nội dung của nó được sao chép và chuyển tới các nơi khác nhau.



Phần quan trọng nhất của mô hình LSTM là trạng thái tế bào -đường chạy thông ngang phía trên của sơ đồ hình vẽ. Trạng thái tế bào là một dạng giống như băng truyền. Nó chạy xuyên suốt tất cả các các nút mạng và chỉ tương tác tuyến tính với nhau một chút. Vì vậy mà các thông tin có thể dễ dàng truyền đi thông suốt mà không sợ bị thay đổi.

### Hình 3.5 Trạng thái tế bào

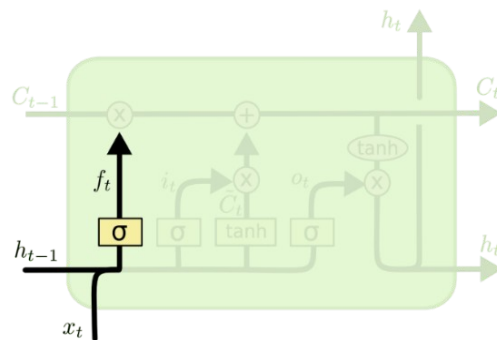
LSTM có khả năng bỏ đi hoặc thêm vào các thông tin cần thiết cho trạng thái tế bào, chúng được điều chỉnh cẩn thận bởi các nhóm được gọi là cổng. Các cổng là nơi sàng lọc thông tin đi qua nó, chúng được kết hợp bởi một tầng mạng sigmoid và một phép nhân.

### Hình 3.6 Mô tả hàm sigmoid

Tầng sigmoid sẽ cho đầu ra là một số trong khoảng  $[0, 1]$ , mô tả có bao nhiêu thông tin có thể được thông qua. Với đầu ra của tầng này là 0 thì có nghĩa là không cho thông tin nào qua cả, còn khi kết quả là 1 thì có nghĩa là cho tất cả các thông tin đi qua nó. Một LSTM gồm có 3 cổng như vậy để duy trì và điều hành trạng thái của tế bào.

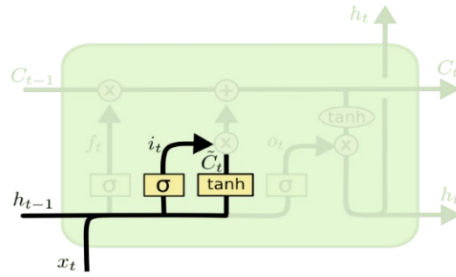
Bước đầu tiên của LSTM là quyết định xem thông tin nào cần bỏ đi từ trạng thái tế bào. Quyết định này được đưa ra bởi tầng sigmoid - gọi là “tầng cổng quên” (forget gate layer). Nó sẽ lấy đầu vào là  $h_{t-1}$  và  $x$  rồi đưa ra kết quả là một số trong khoảng  $[0, 1]$  cho mỗi số trong trạng thái tế bào  $C_{t-1}$ . Đầu ra là 1 thể hiện rằng nó giữ toàn bộ thông tin lại, còn 0 chỉ rằng toàn bộ thông tin sẽ bị bỏ đi.

Bước tiếp theo là quyết định xem thông tin mới nào sẽ lưu vào trạng thái tế bào. Việc này gồm 2 phần. Đầu tiên là sử dụng một tầng sigmoid được gọi là “tầng cổng vào”



$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

(input gate layer) để quyết định giá trị nào sẽ cập nhật. Tiếp theo là một tầng tanh tạo ra một véc-tơ cho giá trị mới nhằm thêm vào cho trạng thái. Trong bước tiếp theo, sẽ kết hợp 2 giá trị đó lại để tạo ra một cập nhật cho trạng thái.

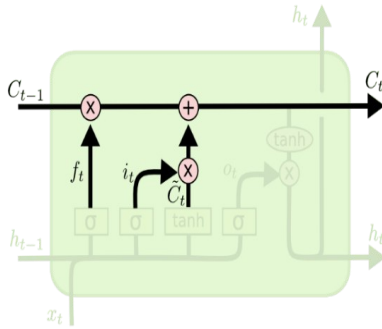


$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Hình 3.7 Mô tả hàm sigmoid kết hợp với tanh

Cập nhật trạng thái tế bào cũ  $C_{t-1}$  thành trạng thái mới  $C_t$ ; ở các bước trước đó đã quyết định những việc cần làm, nên giờ chỉ cần thực hiện là xong. Nhân trạng thái cũ với  $f$  để bỏ đi những thông tin quyết định quên lúc trước. Sau đó cộng thêm  $i, *$ . Trạng thái mới thu được này phụ thuộc vào việc quyết định cập nhập mỗi giá trị trạng thái ra sao.



$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

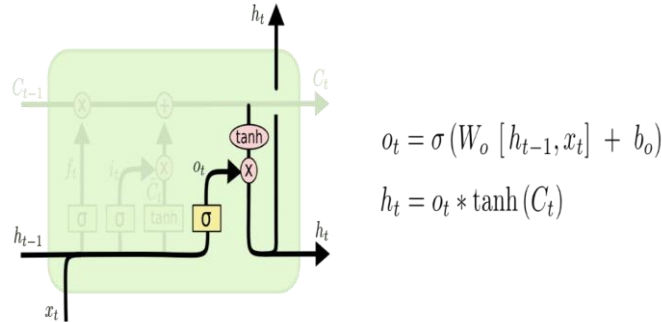
Hình 3.8 Kết hợp để cho ra đầu ra

Cuối cùng, cần quyết định xem muốn đầu ra là gì. Giá trị đầu ra sẽ dựa vào trạng thái tế bào, nhưng sẽ được tiếp tục sàng lọc. Đầu tiên, chạy một tầng sigmoid để quyết định phần nào của trạng thái tế bào muốn xuất ra. Sau đó, đưa nó trạng thái tế bào qua một hàm tanh để co giá trị nó về khoảng  $[-1, 1]$ , và nhân nó với đầu ra của cổng sigmoid để được giá trị đầu ra mong muốn.

Hình 3.9 Kết hợp các tầng để cho ra đầu ra là input tiếp theo của dữ liệu

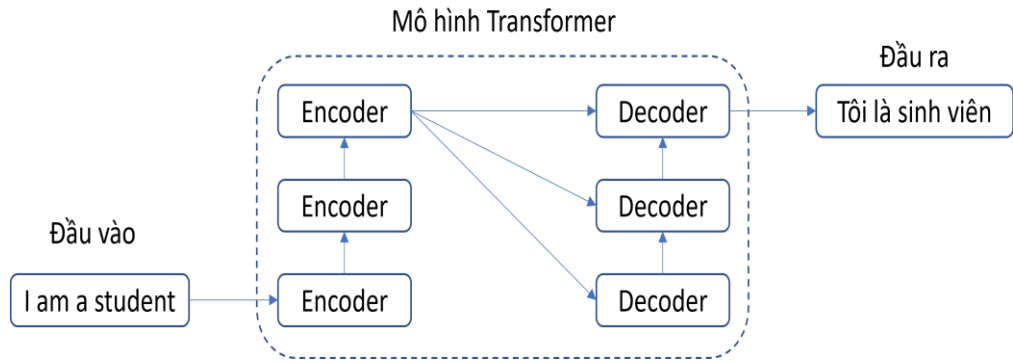
### 3.4 Mô hình biến đổi Transformer

Mô hình biến đổi (*Transformer*) là một mô hình nhúng chuỗi dựa trên cơ chế tự chú ý được giới thiệu nhóm tác giả Google Brain năm 2017 [27]. Giống như các mạng nơ-



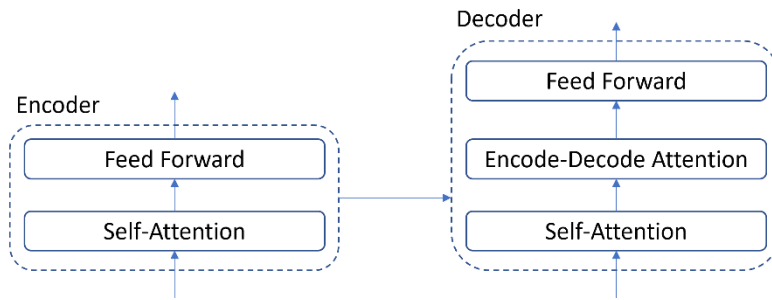
ron hồi quy *RNN*, các *Transformer* được thiết kế để xử lý dữ liệu tuần tự theo dạng chuỗi cho các tác vụ như dịch máy hay tóm tắt tự động. Tuy nhiên khác với mạng *RNN*, các *Transformer* không yêu cầu dữ liệu dạng chuỗi cần xử lý theo thứ tự. Ví dụ, nếu dữ liệu đầu vào là một câu ngôn ngữ tự nhiên, *Transformer* không cần phải xử lý phần đầu câu trước cuối câu. Với ưu điểm này, mô hình *Transformer* hỗ trợ nhiều phép tính toán song song, nhờ vậy giảm thời gian huấn luyện trên các tập dữ liệu lớn. Hiện mô hình *Transformer* và các biến thể của nó đang là mô hình được nhiều sự lựa chọn cho các vấn đề về xử lý ngôn ngữ tự nhiên [27] để thay thế các mô hình *RNN*, ví dụ như biến thể bộ nhớ dài-ngắn hạn, *LSTM*. Mô hình *Transformer* nâng cao sự phát triển của các mô hình "huấn luyện trước" (*pre-training*) như BERT (*Bidirectional Encoder Representations from Transformers*) [28] và GPT (*Generative Pre-trained Transformer*) [29].

Mô hình biến đổi *Transformer* bao gồm hai mô-đun chính là khối mã hóa (*encoder*) và khối giải mã (*decoder*) được mô tả như Hình 3.10 :



Hình 3.10 Mô hình minh họa kiến trúc Transformer

Ở mức cao, kiến trúc *Transformer* tiếp cận khá giống với các mạng nơ-ron học sâu cơ bản vì nó cũng sử dụng kết hợp lớp nhúng và mạng nơ-ron truyền thẳng FNN. Tuy nhiên có 2 điểm khác là (1) kiến trúc *Transformer* sử dụng lớp nhúng theo cơ chế tự chú ý [27] để biến đổi dữ liệu đầu vào theo dạng chuỗi tuần tự, và (2) các khối này được xếp lớp với nhau để xử lý song song được nhiều thuộc tính khác nhau từ chuỗi dữ liệu đầu vào. Mô hình chi tiết của khối mã hóa và giải mã của kiến trúc *Transformer* được mô tả ở Hình 3.11.



Hình 3.11 Các lớp chi tiết của kiến trúc Transformer

### Lớp tự chú ý :

Lớp tự chú ý (*self-attention*) được dùng để tính toán sự phụ thuộc giữa các thành phần trong chuỗi dữ liệu tuần tự đầu vào .

$$Tự\ chú\ ý(attention) = f(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.10)$$

Trong đó:

- Q (Query), K(Key), và V (Value) là các ma trận được thu được từ dữ liệu đầu vào thông qua các ma trận trọng số học được.
- $d_k$  là số chiều của vector chìa khóa (key).
- softmax đảm bảo các trọng số xác suất của từng giá trị trong ma trận attention.

## Mạng nơ-ron truyền thẳng

Ngoài các lớp tự chú ý, mỗi lớp trong khối mã hóa và giải mã đều chứa các mạng nơ-ron truyền thẳng với lớp kết nối đầy đủ. Kiến trúc mạng bao gồm hai phép biến đổi tuyến tính sử dụng hàm kích hoạt *ReLU* như công thức 2.11

$$\text{FNN}(x) = (XW_1 + b_1)W_2 + b_2 \quad (2.11)$$

Trong đó  $X$  là véc-tơ đầu vào của mô hình,  $W_1$  và  $W_2$  là ma trận trọng số và  $b_1$  và  $b_2$  là véc-tơ điều chỉnh.

Tóm lại, có thể dễ thấy rằng LSTM và Transformer đều là các mở rộng của mạng nơ-ron nhân tạo, nhưng chúng được thiết kế đặc biệt để giải quyết các bài toán mà mạng nơ-ron nhân tạo truyền thống không thể xử lý hiệu quả, chẳng hạn như xử lý dữ liệu tuần tự, nắm bắt phụ thuộc dài hạn và tận dụng tính toán song song.

*Bảng 3.1 So sánh đặc điểm của ANN, LSTM và Transformer*

| Đặc điểm              | ANN                | LSTM                        | Transformer                            |
|-----------------------|--------------------|-----------------------------|----------------------------------------|
| Xử lý dữ liệu tuần tự | Không phù hợp      | Phù hợp                     | Phù hợp                                |
| Bộ nhớ                | Không có           | Có (cell state và các cổng) | Không cần (sử dụng cơ chế tự chú ý)    |
| Phụ thuộc dài hạn     | Không nắm bắt được | Nắm bắt được                | Nắm bắt được                           |
| Tính toán song song   | Có thể             | Khó (do tính chất tuần tự)  | Rất tốt (xử lý toàn bộ chuỗi cùng lúc) |

## CHƯƠNG 4: THU NHẬP, XỬ LÝ DỮ LIỆU

### 4.1 Khái quát về thu nhập và xử lý dữ liệu

Dữ liệu đóng vai trò quan trọng trong quá trình đào tạo và tối ưu hóa các mô hình học máy và trí tuệ nhân tạo. Tuy nhiên, việc thu thập toàn bộ dữ liệu là không khả thi do các hạn chế về thời gian và tài nguyên tính toán. Do đó, việc lựa chọn phương pháp thu thập dữ liệu phù hợp, kết hợp với các kỹ thuật xử lý dữ liệu tiên tiến, là yếu tố then chốt nhằm nâng cao độ chính xác và hiệu suất của mô hình. Quá trình này không chỉ đảm bảo tính khả thi của nghiên cứu mà còn giúp mô hình hoạt động trên tập dữ liệu có chất lượng cao nhất. Ba phương pháp thu thập dữ liệu phổ biến bao gồm:

1. **Crowdsourcing:** Thu thập dữ liệu thông qua khảo sát người dùng, phản hồi từ cộng đồng hoặc tương tác trên các nền tảng trực tuyến.
2. **Logging:** Ghi nhận và lưu trữ dữ liệu về hành vi người dùng, bao gồm lịch sử truy cập và tương tác với hệ thống.
3. **Scraping:** Trích xuất dữ liệu từ các trang web hoặc cơ sở dữ liệu có sẵn nhằm thu thập thông tin một cách có hệ thống.

Dữ liệu thô sau khi thu thập thường chứa nhiều nhiễu, sai sót và không đồng nhất, do đó cần trải qua quá trình tiền xử lý để đảm bảo tính chính xác và đồng nhất trước khi đưa vào mô hình. Các bước tiền xử lý dữ liệu chính bao gồm:

- **Chuẩn hóa (Normalization):** Chuyển đổi dữ liệu về phạm vi phù hợp nhằm đảm bảo sự ổn định khi huấn luyện mô hình.
- **Làm sạch dữ liệu (Data Cleaning):** Loại bỏ các giá trị trùng lặp, xử lý dữ liệu bị thiếu hoặc không hợp lệ.
- **Đồng nhất dữ liệu (Homogeneity):** Chuyển đổi dữ liệu về cùng một định dạng để tránh sự không nhất quán giữa các nguồn khác nhau.
- **Đảm bảo tính toàn vẹn (Integrity):** Kiểm tra và xác thực dữ liệu từ các nguồn đáng tin cậy để đảm bảo tính chính xác và đại diện của dữ liệu.
- **Trích xuất đặc trưng (Feature Extraction):** Xác định và lựa chọn các đặc trưng quan trọng nhằm giảm số chiều dữ liệu và tối ưu hóa khả năng học của mô hình.

- **Mã hóa dữ liệu (Encoding):** Biến đổi các biến định tính thành dạng số nhằm đảm bảo khả năng xử lý của mô hình học máy.

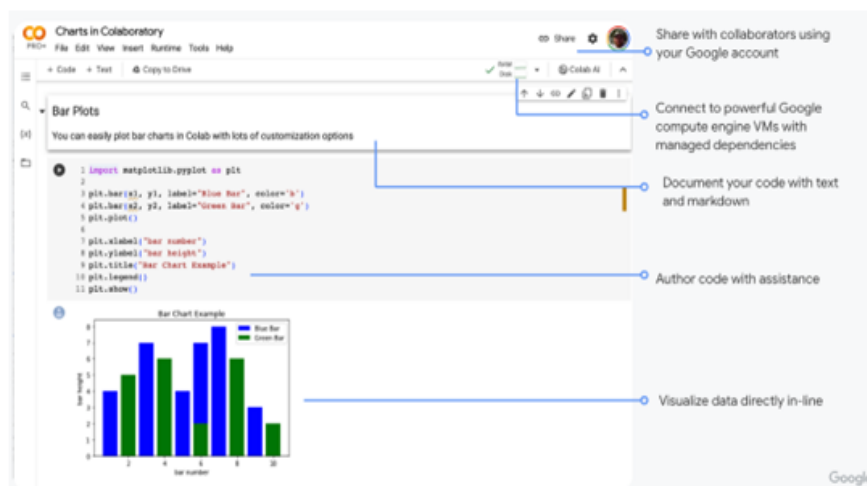
Nhờ quá trình thu thập và tiền xử lý dữ liệu bài bản, mô hình AI/ML có thể khai thác tối đa thông tin từ dữ liệu, từ đó nâng cao hiệu quả dự báo và ra quyết định trong các ứng dụng thực tiễn.

## 4.2 Thư viện Python sử dụng trong nghiên cứu

Trong nghiên cứu này, sẽ sử dụng Google Colab làm môi trường thực thi để xử lý dữ liệu, xây dựng và huấn luyện mô hình dự báo nhu cầu tiêu thụ điện năng. Google Colab cung cấp tài nguyên tính toán mạnh mẽ, bao gồm GPU miễn phí, giúp tăng tốc quá trình huấn luyện mô hình học sâu. Nhiều thư viện Python để hỗ trợ các giai đoạn từ thu thập, xử lý dữ liệu đến xây dựng, huấn luyện và đánh giá mô hình dự báo nhu cầu điện năng. Việc sử dụng các thư viện này giúp đảm bảo tính chính xác, hiệu suất và khả năng mở rộng của mô hình.

*Hình 4.1 Môi trường thực thi Google Colab*

### 1. Xử lý dữ liệu



- **Pandas (import pandas as pd):** Được sử dụng để nhập, thao tác và làm sạch dữ liệu. Pandas hỗ trợ xử lý dữ liệu dạng bảng, giúp dễ dàng tổ chức và chuẩn bị dữ liệu đầu vào cho mô hình.
- **NumPy (import numpy as np):** Cung cấp các công cụ tính toán số học, đặc biệt là xử lý mảng và ma trận, hỗ trợ cho các phép toán đại số tuyến tính cần thiết trong học sâu.

- Scikit-learn (from sklearn.preprocessing import MinMaxScaler, StandardScaler): Hỗ trợ chuẩn hóa dữ liệu bằng các bộ biến đổi như MinMaxScaler và StandardScaler, giúp cải thiện hiệu suất mô hình. Ngoài ra, train\_test\_split được sử dụng để chia tập dữ liệu thành tập huấn luyện và kiểm tra.

## 2. Xây dựng mô hình học sâu

- PyTorch (import torch, import torch.nn as nn): Thư viện chính để xây dựng mô hình học sâu. PyTorch cung cấp các công cụ mạnh mẽ để tạo và huấn luyện mạng nơ-ron, bao gồm cả mô hình LSTM và Transformer. torch.nn chứa các lớp quan trọng như nn.LSTM, nn.Linear và nn.ReLU để xây dựng mạng nơ-ron nhân tạo.

## 3. Đánh giá mô hình

- Scikit-learn (from sklearn.metrics import mean\_squared\_error, r2\_score): Cung cấp các chỉ số đánh giá mô hình, bao gồm:
  - mean\_squared\_error (MSE): Đo lường sai số bình phương trung bình giữa giá trị thực tế và dự báo.
  - r2\_score: Đánh giá mức độ phù hợp của mô hình đối với dữ liệu.

## 4. Trực quan hóa dữ liệu

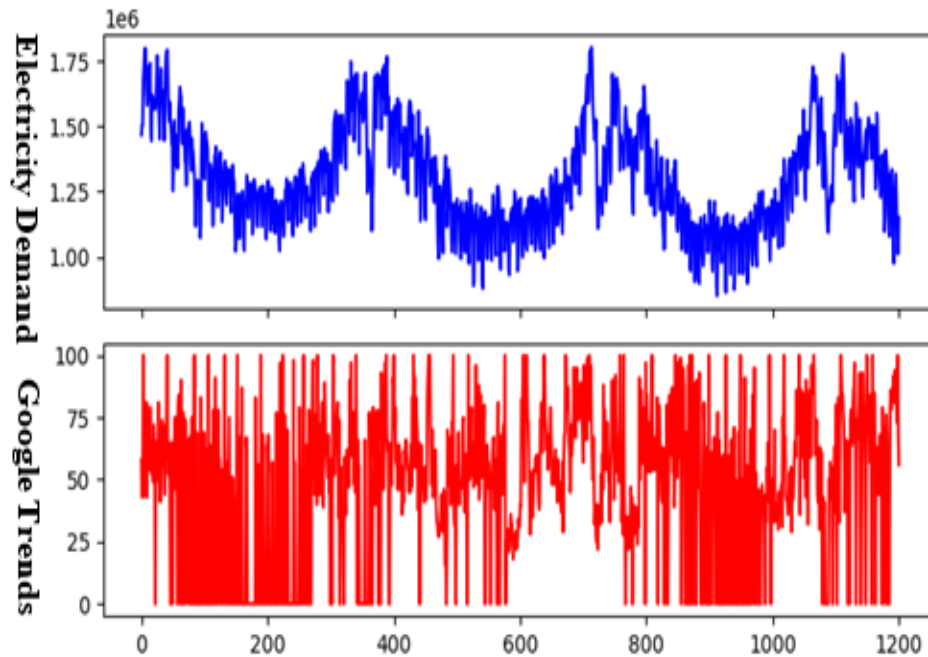
- Matplotlib (import matplotlib.pyplot as plt, from matplotlib.pyplot import figure, import matplotlib.dates as mdates):
  - plt giúp vẽ biểu đồ để phân tích dữ liệu đầu vào và kết quả dự báo.
  - figure giúp tùy chỉnh kích thước đồ thị.
  - mdates hỗ trợ định dạng trục thời gian trong biểu đồ chuỗi thời gian

### 4.3 Ứng dụng thu nhập, xử lý dữ liệu trong đề tài nghiên cứu

Dữ liệu chính được sử dụng trong nghiên cứu này đến từ hai nguồn:

1. Nền tảng ENTSO-E [32] cung cấp dữ liệu về nhu cầu điện của khu vực Vương quốc Anh (EDUK);
2. Dữ liệu Google Trends do Google cung cấp, thể hiện mức độ phổ biến của các cụm từ tìm kiếm theo thời gian [33].

Nhu cầu điện tiêu thụ được biểu thị dưới dạng công suất trung bình hoạt động, đo bằng MW. Dữ liệu được thu thập từ ngày 1 tháng 1 năm 2021 đến ngày 15 tháng 4 năm 2024, với khoảng thời gian lấy mẫu là 30 phút. Đối với dữ liệu Google Trends, như đã trình bày ở mục 3.1.2 cụm từ tìm kiếm "Efficient Energy Use" đã được lựa chọn trong nghiên cứu này.



Hình 4.2 Nhu cầu tiêu thụ điện của khu vực Vương quốc Anh và Dữ liệu Google Trends, tương ứng.

Sau khi chuẩn bị dữ liệu, tập dữ liệu được chuyển đổi sang một định dạng chung để có thể xử lý bằng các mô hình được đề xuất. Điều này yêu cầu một bước xử lý dữ liệu. Sử dụng ngôn ngữ Python trên nền tảng Google Colab và thư viện *StandardScaler* và *MinMaxScaler* của Scikit-learn để chuẩn hóa dữ liệu. Sau đó, tập dữ liệu được chia thành 80% cho tập huấn luyện và 20% cho tập kiểm tra. Mô hình học sâu được hoàn thành thông qua sử dụng thư viện PyTorch.

```
df_train, df_test = train_test_split(df, test_size=0.2,
                                     random_state=42)
df_train.to_csv('train.csv', index=True)
df_test.to_csv('test.csv', index=True)
```

```
nd_scaler = StandardScaler()
gt_scaler = MinMaxScaler()
df['nd_scaled'] = nd_scaler.fit_transform(df[['nd']])
df['gt'] = gt_scaler.fit_transform(df[['gt']])
```

**Xử lý chuỗi thời gian:** Do bài toán dự báo chuỗi thời gian cần sử dụng thông tin từ quá khứ để dự đoán tương lai, dữ liệu được chuyển đổi thành dạng phù hợp với mô hình bằng cách sử dụng một cửa sổ thời gian. Sau khi dữ liệu được xử lý, và cần chuyển đổi thành tensor để sử dụng với mô hình học sâu. Chẳng hạn như mô hình LSTM, Để đảm bảo mô hình hoạt động hiệu quả, dữ liệu đầu vào cần được chuẩn bị đúng định dạng trong PyTorch:

*(batch\_size, sequence\_length, num\_features)*

Trong đó:

- **batch\_size:** Số lượng chuỗi dữ liệu được xử lý trong một lần huấn luyện.
- **sequence\_length:** Độ dài của mỗi chuỗi thời gian đầu vào (số bước thời gian mà mô hình sử dụng để dự báo).
- **num\_features:** Số lượng đặc trưng tại mỗi bước thời gian (ví dụ: công suất tiêu thụ điện, nhiệt độ, v.v.).

Giả sử ta có chuỗi dữ liệu tiêu thụ điện:

[100,102,105,110,120,130,140]

Với cửa sổ thời gian  $T=3$ , ta tạo được tập dữ liệu đầu vào:

$$A = \begin{bmatrix} 100 & 102 & 105 \\ 102 & 105 & 110 \\ 105 & 110 & 120 \end{bmatrix}$$

Mỗi hàng trong ma trận trên tương ứng với một chuỗi (sequence ) đầu vào, và nhãn dự đoán tương ứng là:  $y=[110,120,130,140]$ . Đơn cử, với một bộ dữ liệu chứa hàng ngàn điểm dữ liệu, việc thử nghiệm nhiều giá trị  $T$  ( chưa kể đến tìm kiếm các tham số đầu vào mô hình ) để tìm ra cửa sổ thời gian tối ưu cho dự báo đã có thể rất tốn kém về mặt tính toán. Khi thử nhiều giá trị  $T$ , mỗi lần thay đổi  $T$  cần tiền xử lý lại dữ liệu và huấn luyện lại mô hình, khiến quá trình này mất nhiều thời gian.

Với mô hình Transformer có thể tiếp nhận dữ liệu, tập dữ liệu đầu vào cần được chuyển đổi thành dạng tensor ba chiều với cấu trúc:

$$(B,T,F)$$

Trong đó:

- B (Batch Size): Số lượng mẫu trong mỗi lần huấn luyện
- T (Time Steps): Số bước thời gian (ví dụ: 24 giờ trước đó)
- F (Features): Số lượng đặc trưng (ví dụ, chỉ có 1 cột điện năng tiêu thụ thì  $F=1$ )

Tóm lại, quá trình thu thập và xử lý dữ liệu đóng vai trò then chốt trong việc xây dựng mô hình dự báo nhu cầu điện năng hiệu quả. Việc lựa chọn nguồn dữ liệu, phương pháp tiền xử lý, và định dạng đầu vào phù hợp không chỉ ảnh hưởng đến hiệu suất mô hình mà còn quyết định khả năng áp dụng thực tế của nghiên cứu.

## CHƯƠNG 5: KIỂM CHỨNG THỰC NGHIỆM, ĐÁNH GIÁ KẾT QUẢ

### 5.1 Kiểm chứng thực nghiệm

#### Bước 1: Kiểm định mối quan hệ nhân quả Granger

Mục đích chính của bước này là đánh giá hiệu quả của Google Trends dựa trên mối tương quan của nó với nhu cầu điện tiêu thụ. Do đó, nghiên cứu bắt đầu với kiểm định nhân quả Granger để nghiên cứu các mối quan hệ này.

Kiểm định nhân quả Granger, được đặt theo tên của nhà kinh tế đạt giải Nobel Clive W. Granger, về cơ bản là một khái niệm thống kê được sử dụng để xác định liệu một chuỗi thời gian có thể giúp dự đoán một chuỗi thời gian khác hay không [35]. Phương pháp này được xem xét thông qua mô hình tự hồi quy (AR) và mô hình trễ phân bố tự hồi quy (ADL) như sau:

$$Y_t = a_n + \sum_{i=1}^p (a_i Y_{t-i}) + \epsilon_t \quad (5.1)$$

$$Y_t = \beta_0 + \sum_{i=1}^p (\beta_i Y_{t-i}) + \sum_{i=1}^q (\gamma_i X_{t-i}) + \theta_t \quad (5.2)$$

Mô hình biểu diễn trong phương trình (1) chỉ sử dụng các giá trị quá khứ của Y để dự báo  $Y_t$ . Tuy nhiên, mô hình biểu diễn trong phương trình (2) kết hợp cả các giá trị quá khứ của Y và X lên đến bậc p và q tương ứng. Để xác định liệu  $X_t$  có gây ra  $Y_t$  theo nghĩa Granger hay không, thực hiện chỉ số kiểm định F như sau:

$$F = \frac{(RSS_1 - RSS_2)/q}{RSS_2/(T-2p-q)} \quad (5.3)$$

Ở đây,  $RSS_1$  và  $RSS_2$  là tổng bình phương phần dư (Residual Sum of Squares - RSS) từ các mô hình ở phương trình (1) và (2); T là số lượng quan sát. Nếu thống kê F có ý nghĩa thống kê lớn hơn ở mức đã chọn (thông thường mức 5% được chọn), chúng ta có thể bác bỏ giả thuyết không null hypothesis (Biến X KHÔNG Granger gây ra biến Y" (Tức là: thông tin trong quá khứ của X **không giúp ích** trong việc dự đoán Y) và kết luận rằng việc thêm X vào mô hình dự báo Y giúp cải thiện đáng kể khả năng dự báo của mô hình.

#### Bước 2: Kiểm chứng mô hình trên tập dữ liệu thực

Quá trình kiểm chứng mô hình trên tập dữ liệu thực (dữ liệu nhu cầu điện tiêu thụ tại Vương quốc Anh) được thiết kế theo hai kịch bản khác nhau để đánh giá tác động của dữ liệu Google Trends đối với hiệu suất dự báo nhu cầu năng lượng. Hai kịch bản kiểm chứng gồm:

1. Sử dụng duy nhất tập dữ liệu nhu cầu điện để huấn luyện và kiểm tra mô hình, không sử dụng bất kỳ biến ngoại sinh nào.
2. Kết hợp tập dữ liệu nhu cầu điện với dữ liệu Google Trends như một biến ngoại sinh nhằm cải thiện hiệu suất dự báo.

Hiệu suất dự báo của các mô hình (hàm mất mát) LSTM và Transformer được đánh giá bằng chỉ số lỗi trung bình bình phương (MSE). Chỉ số này đo lường mức độ sai lệch

trung bình giữa giá trị dự báo và giá trị thực tế, trong đó giá trị nhỏ hơn cho thấy mô hình có hiệu suất tốt hơn.

$$MSE = \frac{\sum_{i=1}^n ((y_i - \hat{y}_p)^2}{n} \quad (5.4)$$

### Tối ưu hóa tham số mô hình:

Quá trình tối ưu hóa tham số được thực hiện thông qua kỹ thuật tìm kiếm lưới (grid search) để xác định cấu hình tốt nhất trên tập huấn luyện. Các tham số được tối ưu như sau:

- **LSTM:**
  - Số lớp ẩn: 1-3 lớp
  - Số đơn vị ẩn: 32-256
  - Dropout: 0.1-0.3
- **Transformer:**
  - Số lượng đầu self-attention: 2, 4, 8, 16
  - Số neuron trong các lớp tuyến tính: 32, 64, 128
  - Tốc độ học:
  - Số epoch: 100 (sử dụng dừng sớm - early stopping để tránh overfitting)
  - Kích thước batch: 256

Ngoài ra, áp dụng phương pháp cửa sổ để biểu diễn dữ liệu thời gian và dự báo các giá trị trong tương lai. Dải giá trị cửa sổ được xác định từ 1 đến 20, và giá trị tối ưu được tìm thấy là 14.

Bảng 5.1 trình bày chi tiết các giá trị tham số tối ưu được xác định cho cả hai mô hình, kết hợp và không kết hợp dữ liệu Google Trends.

*Bảng 5.1 Các tham số tối ưu của mô hình*

| Mô hình                                              | LSTM              | Transformer          |
|------------------------------------------------------|-------------------|----------------------|
| (Đơn biến) Chỉ sử dụng dữ liệu nhu cầu tiêu thụ điện | Số lớp chồng: 1   | Số lượng đầu: 4      |
|                                                      | Số đơn vị ẩn: 256 | Số lượng neuron: 128 |
|                                                      | Dropout: 0.2      |                      |
| (Đa biến) Kết hợp với dữ liệu Google Trends          | Số lớp chồng: 2   | Số lượng đầu: 16     |
|                                                      | Số đơn vị ẩn: 128 | Số lượng neuron: 128 |
|                                                      | Dropout: 0.1      |                      |

## 5.2 Phân tích kết quả

Để thực hiện kiểm định nhân quả Granger, trước tiên, đề tài kiểm tra tính dừng của cả hai chuỗi thời gian (Google Trends cho từ khóa “*Efficient Energy Use*” và nhu cầu điện) bằng kiểm định Augmented Dickey-Fuller (ADF) và kiểm định Kwiatkowski-Phillips-Schmidt-Shin (KPSS). Kết quả, được trình bày trong *Bảng 5.2*, cho thấy cả hai chuỗi đều có tính dừng. Tiếp theo, kiểm định Granger được thực hiện và kết quả được báo cáo trong *Bảng 5.3*. Kết quả chỉ ra rằng Google Trends là một yếu tố dự báo hữu ích cho việc dự báo nhu cầu điện ở khu vực Vương quốc Anh.

*Bảng 5.2 Kiểm tra tính dừng của hai chuỗi thời gian*

| <i>Ở dạng nguyên mẫu</i>                               |                 |                  |
|--------------------------------------------------------|-----------------|------------------|
| <b>Chuỗi thời gian</b>                                 | <b>ADF test</b> | <b>KPSS test</b> |
| <i>Nhu cầu tiêu thụ điện</i>                           | 0.0126          | 0.0289           |
| <i>Google Trend với từ khóa “efficient energy use”</i> | 0.0164          | 0.0002           |

Cả hai kiểm định ADF và KPSS đều cho thấy chuỗi thời gian của nhu cầu điện và Google Trends cho cụm từ “*Efficient Energy Use*” là dừng. Cụ thể, kiểm định ADF cho kết quả giá trị p-value thấp- **xác suất** để kiểm tra tính hợp lý của giả thuyết không  $H_0$  ( $p$ -value  $=0 < 5\%$  đã chọn), bác bỏ giả thuyết không về tính dừng, trong khi kiểm định KPSS cho thấy thống kê kiểm định thấp, không đủ bằng chứng để bác bỏ giả thuyết không về tính dừng.

Như thể hiện trong *Bảng 5.3* giá trị F-statistic là 7.2452, cho thấy mô hình có bao gồm các giá trị quá khứ của Google Trends đóng góp đáng kể vào việc dự báo nhu cầu điện. Hơn nữa, giá trị p cực thấp cung cấp bằng chứng mạnh mẽ để bác bỏ giả thuyết không.

*Bảng 5.3 Kiểm định mối quan hệ nhân quả Granger*

|                                                                                                                        |        |
|------------------------------------------------------------------------------------------------------------------------|--------|
| <b><i>H0: Từ khóa “Efficient Energy Use” có quan hệ nhân quả với với nhu cầu tiêu thụ điện tại Vương quốc Anh.</i></b> |        |
| <b>F-stat (lag=8)</b>                                                                                                  | 7.2452 |
| <b>p-value (lag=8)</b>                                                                                                 | 0      |

Tiến hành đánh giá hiệu suất của hai mô hình dự báo chuỗi thời gian, bao gồm LSTM và Transformer, trên tập dữ liệu nhu cầu điện không sử dụng dữ liệu kết hợp với Google Trends, kết quả tổng hợp được tóm gọn ở Bảng 5.4. Chỉ số đánh giá hiệu suất MSE, cho thấy rằng mô hình Transformer hoạt động tốt hơn đáng kể so với mô hình LSTM. Điều này được thể hiện rõ hơn trong Hình 5.1 b, Hình 5.2b, nơi mô hình Transformer có khả năng theo sát xu hướng biến động của dữ liệu tốt hơn, trong khi LSTM thể hiện mức độ sai lệch lớn hơn.

*Bảng 5.4 Kết quả đánh giá hiệu suất mô hình.*

| <b>Mô hình</b>                                   | <b>MSE</b> |
|--------------------------------------------------|------------|
| LSTM (không tích hợp dữ liệu Google Trend )      | 1.1625     |
| Transformer (không tích hợp dữ liệu Google Trend | 1.1342     |
| LSTM tích hợp dữ liệu Google Trend               | 1.0955     |
| Transformer tích hợp dữ liệu Google Trend        | 1.0928     |

Một trong những lý do giúp Transformer đạt được hiệu suất cao hơn là khả năng mô hình hóa các quan hệ dài hạn trong dữ liệu tốt hơn so với LSTM. Trong khi LSTM sử dụng cơ chế cổng (gating mechanism) để duy trì thông tin qua các bước thời gian, Transformer sử dụng cơ chế tự chú ý (self-attention), cho phép nó xác định mối quan hệ giữa các thời điểm khác nhau trong chuỗi dữ liệu một cách hiệu quả. Điều này đặc biệt quan trọng trong bài toán dự báo nhu cầu điện, nơi các mẫu dữ liệu có sự tương quan theo mùa (seasonality) và theo xu hướng dài hạn.

Tuy nhiên, một trong những hạn chế của Transformer mà chúng tôi quan sát được trong quá trình thí nghiệm là yêu cầu tính toán cao hơn đáng kể khi số lượng đầu (heads) trong cơ chế tự chú ý tăng lên. Khi số lượng heads được đặt ở giá trị cao hơn, lượng thông tin cần xử lý trong mỗi lần tính toán cũng tăng lên, dẫn đến sự gia tăng đáng kể về thời gian và tài nguyên tính toán. Điều này đặt ra một thách thức quan trọng khi triển khai Transformer trên các hệ thống có giới hạn về phần cứng.

Đối với LSTM, chúng tôi nhận thấy rằng thời gian huấn luyện cũng bị ảnh hưởng đáng kể khi tăng số lượng lớp chồng (stacked layers) lên 3 hoặc tăng kích thước ẩn

(hidden size) lên 256. Khi mô hình trở nên phức tạp hơn, số lượng tham số cần tối ưu hóa tăng lên, khiến quá trình huấn luyện mất nhiều thời gian hơn và có nguy cơ dẫn đến hiện tượng quá khớp (overfitting) nếu không có các biện pháp điều chỉnh phù hợp, chẳng hạn như dropout hoặc regularization.

Ngoài ra, kích thước cửa sổ (window size) đóng vai trò quan trọng trong quá trình dự báo. Chúng tôi đã thử nghiệm với nhiều giá trị khác nhau cho kích thước cửa sổ và nhận thấy rằng khi cửa sổ quá nhỏ (từ 1 đến 5), cả hai mô hình đều không thể dự báo chính xác xu hướng dữ liệu. Điều này là do lượng thông tin trong cửa sổ quá hạn chế, khiến mô hình không thể học được các quy luật quan trọng trong chuỗi thời gian. Ngược lại, khi kích thước cửa sổ quá lớn (từ 17 đến 20), hiệu suất dự báo cũng không được cải thiện đáng kể. Điều này có thể do việc mở rộng cửa sổ quá mức dẫn đến mô hình phải xử lý một lượng dữ liệu quá lớn trong mỗi lần dự báo, làm tăng độ phức tạp tính toán mà không mang lại lợi ích đáng kể về mặt chính xác.

Kết quả thí nghiệm cho thấy rằng việc lựa chọn kích thước cửa sổ phù hợp là rất quan trọng trong việc tối ưu hóa hiệu suất của mô hình. Trong nghiên cứu này, chúng tôi tìm thấy rằng kích thước cửa sổ tối ưu là 14, một giá trị cân bằng giữa việc nắm bắt đủ thông tin quá khứ và duy trì độ phức tạp mô hình ở mức hợp lý.

Tóm lại, kết quả thực nghiệm cho thấy rằng mô hình Transformer có hiệu suất vượt trội hơn so với LSTM trong cả hai kịch bản, dù chỉ sử dụng dữ liệu nhu cầu điện hay kết hợp với Google Trends. Việc tích hợp dữ liệu ngoại sinh giúp cải thiện độ chính xác dự báo đáng kể. Tuy nhiên, Transformer yêu cầu tài nguyên tính toán lớn hơn, trong khi LSTM có thể bị chậm lại khi số lớp chồng hoặc kích thước ẩn tăng lên. Kích thước cửa sổ cũng là một tham số quan trọng cần được lựa chọn cẩn thận để tối ưu hóa hiệu suất của mô hình.

## **5.3 Kết luận, đề xuất hướng nghiên cứu**

### **5.3.1 Kết luận chung**

Dựa trên kết quả phân tích, nghiên cứu này đã cung cấp một đánh giá chi tiết về việc sử dụng Google Trends để tối ưu hóa dự báo nhu cầu điện năng bằng cách ứng dụng các thuật toán trí tuệ nhân tạo. Các kết luận chính bao gồm:

- Google Trends có ý nghĩa trong dự báo điện năng, giúp mô hình nhận diện được xu hướng tiêu thụ điện dựa trên hành vi tìm kiếm trực tuyến.
- Transformer là mô hình tối ưu khi kết hợp Google Trends, vượt trội hơn LSTM do khả năng khai thác thông tin từ dữ liệu thời gian dài hơn.
- Việc lựa chọn kích thước cửa sổ hợp lý rất quan trọng: Nếu quá ngắn, mô hình không nắm bắt được xu hướng; nếu quá dài, hiệu quả dự báo không cải thiện đáng kể.
- Tài nguyên tính toán cần được cân nhắc: Transformer đòi hỏi nhiều tài nguyên hơn khi tăng số lượng heads, trong khi LSTM cũng gặp hạn chế khi mở rộng kích thước ẩn hoặc số lớp.

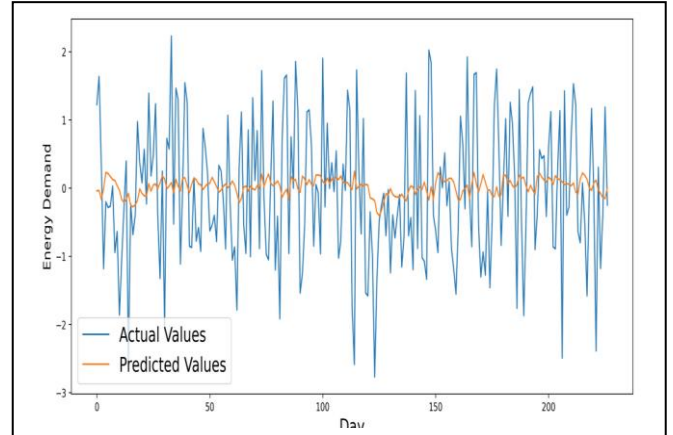
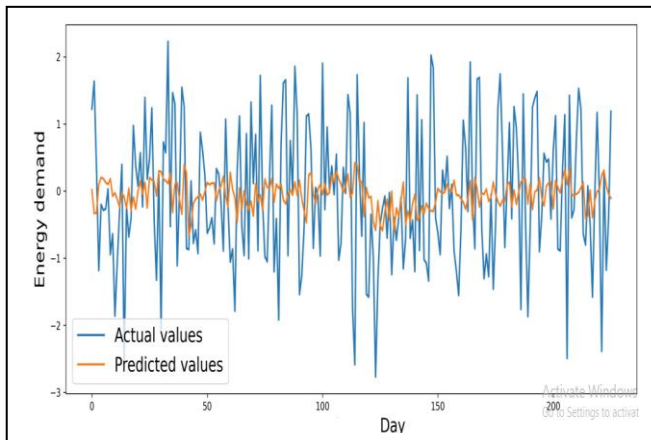
Thông qua nghiên cứu này, đã chứng minh rằng Google Trends có thể trở thành một nguồn dữ liệu hữu ích trong dự báo tiêu thụ điện, nếu được khai thác một cách hợp lý. Điều này mở ra hướng nghiên cứu mới trong việc kết hợp dữ liệu tìm kiếm trực tuyến với AI để nâng cao độ chính xác của các hệ thống dự báo năng lượng.

### 5.3.2 Đề xuất hướng nghiên cứu

Dựa trên những phát hiện trong nghiên cứu, có một số hướng phát triển tiềm năng có thể được tiếp tục khai thác:

1. Mở rộng phạm vi nghiên cứu với các biến ngoại sinh khác: Ngoài Google Trends, có thể thử nghiệm với các yếu tố ngoại sinh khác như giá điện, điều kiện thời tiết theo thời gian thực, hoặc dữ liệu kinh tế. Việc kết hợp nhiều nguồn dữ liệu có thể giúp mô hình dự báo trở nên toàn diện hơn.
2. Kết quả từ việc sử dụng dữ liệu Google Trends đặt nền tảng cho việc tích hợp và khai thác hiệu quả hơn các nguồn dữ liệu thay thế, chẳng hạn như phản hồi khách hàng và truy vấn tìm kiếm từ các nền tảng dịch vụ khách hàng (CSKH), vốn đang được các công ty điện lực triển khai rộng rãi nhưng chưa được tận dụng đầy đủ trong dự báo nhu cầu điện.
3. Phát triển mô hình AI tối ưu hơn để khai thác dữ liệu Google Trends :

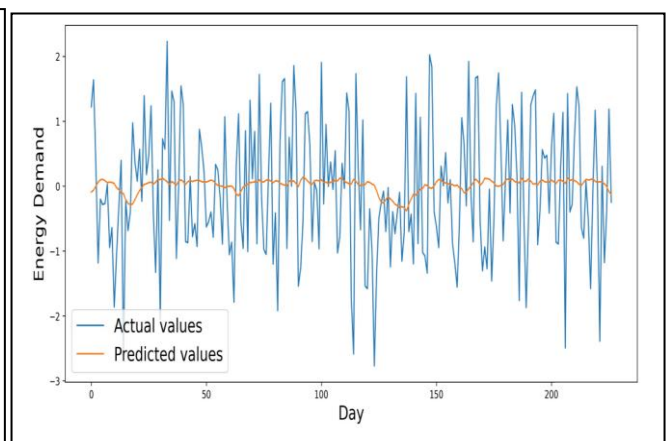
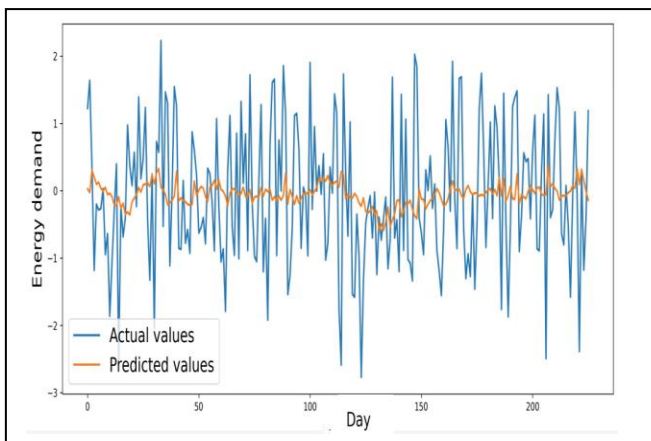
- Áp dụng mô hình lai giữa LSTM và Transformer để tận dụng thế mạnh của cả hai phương pháp.



**a, Transformer**

**b, LSTM**

Hình 5.1 Mô hình Transforme. LSTM sử dụng Google Trends trên tập kiểm thử



**a, Transformer**

**b, LSTM**

Hình 5.2 Mô hình Transformer, LSTM chỉ sử dụng dữ liệu tiêu thụ điện trên tập kiểm thử

## TÀI LIỆU THAM KHẢO

- [1] Brown, R. G. (1959). Statistical forecasting for inventory control. McGraw/Hill.
- [2] Holt, C. C. (1957). Forecasting seasonals and trends by exponentially weighted averages (O.N.R. Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA.
- [3] Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324–342.
- [4] Gardner, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1), 1–28
- [5] Gardner, E. S. (2006). Exponential smoothing: The state of the art — Part II. *International Journal of Forecasting*, 22, 637–666.
- [6] P. R. Winters. “Forecasting sales by exponentially weighted moving averages”. In: *Management science* 6.3 (1960), pp. 324–342.
- [7] Mario Figueiredo, Yuri F. Saporito, Forecasting the term structure of commodities future prices using machine learning, *Digital Finance*, 10.1007/s42521-022-00069-3, 5, 1, (57-90), (2022)
- [8] C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006
- [9] Y. LeCun, Y. Bengio, and G. Hinton. “Deep learning”. In: *nature* 521.7553 (2015), pp. 436–444.
- [10] H. S. Hippert, C. E. Pedreira, and R. C. Souza. “Neural networks for short-term load forecasting: A review and evaluation”. In: *IEEE Transactions on power systems* 16.1 (2001), pp. 44–55
- [11] T. Zufferey, A. Ulbig, S. Koch, and G. Hug. “Forecasting of smart meter time series based on neural networks”. In: *International workshop on data analytics for renewable energy integration*. Springer. 2016, pp. 10–21.
- [12] R. J. Hyndman and G. Athanasopoulos. *Forecasting: principles and practice*. 2nd ed. OTexts: Melbourne, Australia, 2018. url: <http://OTexts.com/fpp2> (visited on 03/10/2021)
- [13] S. B. Taieb, J. W. Taylor, and R. J. Hyndman. “Hierarchical probabilistic forecasting of electricity demand with smart meter data”. In: *Journal of the American Statistical Association* (2020), pp. 1–17.
- [14] S. B. Taieb, J. W. Taylor, and R. J. Hyndman. “Hierarchical probabilistic forecasting of electricity demand with smart meter data”. In: *Journal of the American Statistical Association* (2020), pp. 1–17.
- [15] Y. Goldberg. “A primer on neural network models for natural language processing”. In: *Journal of Artificial Intelligence Research* 57 (2016), pp. 345–420
- [16] D. Bahdanau, K. Cho, and Y. Bengio. “Neural machine translation by jointly learning to align and translate”. In: *arXiv preprint arXiv:1409.0473* (2014).
- [17] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah. “A review on time series forecasting techniques for building energy consumption”. In: *Renewable and Sustainable Energy Reviews* 74 (2017), pp. 902–924.
- [18] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017)

- [19] A. P. Douglas, A. M. Breipohl, F. N. Lee, and R. Adapa: Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, Volume 35, Issue 1, Pages 213-223, 2019.
- [20] A. P. Douglas, A. M. Breipohl, F. N. Lee, and R. Adapa: Online big data-driven oil consumption forecasting with Google trends. *International Journal of Forecasting*, Volume 35, Issue 1, Pages 213-223, 2019.
- [21] Dean Fantazzini; Nikita Fomichev: Forecasting the real price of oil using Donline search data. *International Journal of Computational Economics and Econometrics (IJCEE)*, Vol. 4, No. 1/2, Pages 1-31, 2014
- [22] Jian-Feng Guo, Qiang Ji:How does market concern derived from the Internet affect oil prices?. Volume 112, Pages 1536-1543, 2013
- [23] Xin Li, Jian Ma, Shouyang Wang, Xun Zhang:How does Google search affect trader positions and crude oil prices?. *Economic Modelling*, Volume 49, Pages 162-171, 2015.
- [24] D. W. Bunn, E. D. Farmer: Comparative models for electrical load forecasting. John Wiley & Sons, New York (1985)
- [25] Lean Yu, Yaqing Zhao, Ling Tang, Zebin Yang: Risk due to load forecast uncertainty in short term power system planning. *IEEE Trans. Power Systems*, Vol. 13, No. 4, pp. 1493-1499, 1998
- [26] Wilhite, H. et al: A cross-cultural analysis of household energy use behaviour in Japan and Norway. In: *Energy Policy* vol. 24, no. 9, pp. 795–803.
- [27] Icek Ajzen: The theory of planned behavior. In: *Organizational Behavior and Human Decision Processes*, Volume 50, Issue 2, Pages 179-211, 1991, ISSN 0749-5978.
- [28] Vassileva, I., Campillo, J.: Increasing energy efficiency in low-income households through targeting awareness and behavioral change. In: *Renewable energy*, 67, 59-63. doi:10.1016/j.renene.2013.11.046
- [29] A. Vaswani, N. Shazeer, N. Parmar, et al., “Attention Is All You Need”, *CoRR*, vol. abs/1706.03762, 2017. arXiv: [1706 . 03762](https://arxiv.org/abs/1706.03762). [Online]. Available: <http://arxiv.org/abs/1706.03762>.
- [30] T. Wolf, L. Debut, V. Sanh, et al., “Transformers: State-of-the-Art Natural Language Processing”, in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Online: Association for Computational Linguistics, Oct. 2020, pp. 38–45. doi: [10. 18653/v1/2020.emnlp- demos.6](https://doi.org/10.18653/v1/2020.emnlp-demos.6). [Online]. Available: <https://aclanthology.org/2020.emnlp-demos.6>
- [31] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever, et al., “Improving language understanding by generative pre-training”, 2018.
- [32] ENTSO-E. Transparency Platform
- [33] <https://transparency.entsoe.eu>
- [34] <https://trends.google.com.vn/trends/?geo=GB>
- [35] M. Ding, Y. Chen, and S. L. Bressler, “Granger causality: Basic theory and application to neuroscience,” *Handbook of time series analysis*, vol. 1, pp.

437–460, 2006.

# **PHỤ LỤC**

**Kiểm tra quan hệ nhân quả Granger -Casuality**  
**Xây dựng mô hình dự báo sử dụng ngôn ngữ Python**



```

min_p_value = np.min(p_values)
df.loc[r, c] = min_p_value
df.columns = [var + '_x' for var in variables]
df.index = [var + '_y' for var in variables]
return df

```

Start coding or `grangers_causation_matrix` with AI.

Start coding or `grangers_causation_matrix` with AI.

```

grangers_causation_matrix(df, variables = df.columns)

```

```

nd_x  gt_x
nd_y  1.0  0.0
gt_y  0.0  1.0

```

```

import numpy as np
import pandas as pd
from statsmodels.tsa.stattools import grangercausalitytests

def grangers_causation_matrix(data, variables, maxlag, test='ssr_chi2test', stat='p_value', verbose=False):
    """
    Check Granger Causality of all possible combinations of the time series.
    The rows are the response variable, columns are predictors. The values in the table
    are the selected statistic (P-Values or F-Statistics). P-Values lesser than the significance level (0.05)
    imply that the null hypothesis that the coefficients of the corresponding past values are zero
    (i.e., X does not cause Y) can be rejected.
    """

```

Parameters:

- `data`: pandas dataframe containing the time series variables.
- `variables`: list containing names of the time series variables.
- `maxlag`: int, maximum number of lags to check for causality.
- `test`: str, the test statistic to use (default is 'ssr\_chi2test').
- `stat`: str, the statistic to return ('p\_value' or 'f\_stat') (default is 'p\_value').
- `verbose`: bool, if True, prints the selected statistic values of the tests.

Returns:

A pandas dataframe containing the selected statistic (P-Values or F-Statistics) of the Granger causality tests.

```

"""
df = pd.DataFrame(np.zeros((len(variables), len(variables))), columns=variables, index=variables)
for c in df.columns:
    for r in df.index:
        test_result = grangercausalitytests(data[[r, c]], maxlag=maxlag, verbose=False)
        if stat == 'p_value':
            values = [round(test_result[i+1][0][test][1], 4) for i in range(maxlag)]
        elif stat == 'f_stat':
            values = [round(test_result[i+1][0][test][0], 4) for i in range(maxlag)]
        else:
            raise ValueError("Invalid stat parameter. Choose 'p_value' or 'f_stat'.")
        if verbose:
            print(f'Y = {r}, X = {c}, Values = {values}')
        min_value = np.min(values)
        df.loc[r, c] = min_value
df.columns = [var + '_x' for var in variables]
df.index = [var + '_y' for var in variables]
return df

```

```

grangers_causation_matrix(df, variables = df.columns)

```

```

nd_x  gt_x
nd_y  1.0  0.0
gt_y  0.0  1.0

```

Distributions



2-d distributions



Time series



Values



```

grangers_causation_matrix(df, variables = df.columns)

```

```

nd_x  gt_x
nd_y  1.0  0.0
gt_y  0.0  1.0

```

```

import numpy as np
import pandas as pd
from statsmodels.tsa.stattools import grangercausalitytests

```

```

def grangers_causation_matrix(df, variables, maxlag, test='ssr_chi2test', stat='p_value', verbose=False):
    """

```

```

elastic matrix:
      id_3    g1_3
0 0000 24.4139
1 3443 0.0000

```

```

def __init__(self, d_model: int, dropout: float = 0.1, max_len: int = 5000) -> None:
    super().__init__()
    self.dropout = torch.nn.Dropout(dropout)
    self.positional_encoding = torch.nn.Parameter(torch.zeros(1, d_model, max_len))
    self.pe = torch.nn.Parameter(torch.zeros(1, d_model, max_len))
    self.register_buffer("pe", self.pe)
    self.register_buffer("pos_enc", self.positional_encoding)

def forward(self, x: torch.Tensor) -> torch.Tensor:
    # Add positional encoding to the given tensor.
    x = x + self.pe
    x = self.dropout(x)
    return x

class TransformerBlock(torch.nn.Module):
    def __init__(self, in_dim: int, out_dim: int, embed_dim: int, num_heads: int, num_layers: int):
        super().__init__()
        self.in_dim = in_dim
        self.out_dim = out_dim
        self.embed_dim = embed_dim
        self.num_heads = num_heads
        self.num_layers = num_layers
        self.attention = torch.nn.MultiheadAttention(torch.nn.Linear(in_dim, embed_dim),
                                                    torch.nn.Linear(in_dim, embed_dim),
                                                    torch.nn.Linear(in_dim, embed_dim))
        self.feed_forward = torch.nn.Linear(embed_dim, out_dim)
        self.dropout = torch.nn.Dropout(dropout)
        self.layer_norm1 = torch.nn.LayerNorm(embed_dim)
        self.layer_norm2 = torch.nn.LayerNorm(out_dim)

    def forward(self, src: torch.Tensor, tgt: torch.Tensor) -> torch.Tensor:
        # Forward function of the model.
        src = self.layer_norm1(src)
        tgt = self.layer_norm1(tgt)
        x = self.attention(src, tgt, tgt)
        x = self.dropout(x)
        x = self.layer_norm2(x)
        x = self.feed_forward(x)
        return x

class Transformer(torch.nn.Module):
    def __init__(self, in_dim: int, out_dim: int, embed_dim: int, num_heads: int, num_layers: int):
        super().__init__()
        self.in_dim = in_dim
        self.out_dim = out_dim
        self.embed_dim = embed_dim
        self.num_heads = num_heads
        self.num_layers = num_layers
        self.encoder = torch.nn.LSTM(embed_dim, embed_dim, batch_first=True)
        self.decoder = torch.nn.LSTM(embed_dim, embed_dim, batch_first=True)
        self.attention = torch.nn.MultiheadAttention(torch.nn.Linear(embed_dim, embed_dim),
                                                    torch.nn.Linear(embed_dim, embed_dim),
                                                    torch.nn.Linear(embed_dim, embed_dim))
        self.feed_forward = torch.nn.Linear(embed_dim, out_dim)
        self.dropout = torch.nn.Dropout(dropout)
        self.layer_norm1 = torch.nn.LayerNorm(embed_dim)
        self.layer_norm2 = torch.nn.LayerNorm(out_dim)

    def forward(self, src: torch.Tensor, tgt: torch.Tensor) -> torch.Tensor:
        # Forward function of the model.
        src = self.layer_norm1(src)
        tgt = self.layer_norm1(tgt)
        x = self.encoder(src)
        x = self.decoder(x, tgt)
        x = self.attention(x, x, x)
        x = self.dropout(x)
        x = self.layer_norm2(x)
        x = self.feed_forward(x)
        return x

def train(model, device, train_loader, valid_loader, num_epochs=10):
    # If self.train:
    # Add noise to decoder inputs during training
    # tgt = tgt + torch.normal(0, 0.1, size=tgt.shape).to(device)
    # Embed encoder input and add positional encoding.
    # [bs, src_seq_len, embed_dim]
    src = self.encoder.embedding(src)
    src = self.positional_encoding(src)
    # Generate mask to avoid attention to future outputs.

```

```

        # [tgt_seq_len, tgt_seq_len]
        tgt_mask = torch.nn.Transformer.generate_square_subsequent_mask(tgt.shape[1])
        # Embed decoder input and add positional encoding.
        # [bs, tgt_seq_len, embed_dim]
        tgt = self.decoder_embedding(tgt)
        tgt = self.positional_encoding(tgt)

        # Get prediction from transformer and map to output dimension.
        # [bs, tgt_seq_len, embed_dim]
        pred = self.transformer(src, tgt, tgt_mask=tgt_mask)
        pred = self.output_layer(pred)

    return pred

def infer(self, src: torch.Tensor, tgt_len: int) -> torch.Tensor:
    """Runs inference with the model, meaning: predicts future values
    for an unknown sequence.
    For this, iteratively generate the next output token while
    feeding the already generated ones as input sequence to the decoder.

    Args:
        src: input to the encoder [bs, src_seq_len, num_features]
        tgt_len: desired length of the output

    Returns:
        torch.Tensor: inferred sequence
    """
    output = torch.zeros((src.shape[0], tgt_len + 1, src.shape[2])).to(src.device)
    output[:, 0] = src[:, -1]
    for i in range(tgt_len):
        output[:, i + 1] = self.forward(src, output[:, :i])

    return output[:, 1:]

```

start coding or generate with AI.

**THUYẾT MINH ĐỀ TÀI KHOA HỌC VÀ  
CÔNG NGHỆ CẤP TRƯỜNG**

ĐẠI HỌC ĐÀ NẴNG  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT

THUYẾT MINH ĐỀ TÀI  
KHOA HỌC & CÔNG NGHỆ CẤP TRƯỜNG

|                                                                                                                                                                                                                                                                                                                                                                       |           |                                        |                                                                                                                                                    |                                 |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|--|
| <b>1. TÊN ĐỀ TÀI</b><br>Đánh giá lựa chọn ứng dụng Google Trend trong tối ưu hóa dự báo nhu cầu tiêu thụ năng lượng điện dựa trên các thuật toán trí tuệ nhân tạo.                                                                                                                                                                                                    |           |                                        | <b>2. MÃ SỐ</b><br>T2023-06-23                                                                                                                     |                                 |  |
| <b>3. LĨNH VỰC NGHIÊN CỨU</b><br>Tự nhiên <input type="checkbox"/> Kỹ thuật <input checked="" type="checkbox"/> Môi trường <input type="checkbox"/><br>Kinh tế: <input type="checkbox"/> Nông Lâm <input type="checkbox"/> ATLĐ <input type="checkbox"/><br>Giáo dục <input type="checkbox"/> Y Dược <input type="checkbox"/> Sở hữu trí tuệ <input type="checkbox"/> |           |                                        | <b>4. LOẠI HÌNH NGHIÊN CỨU</b><br>Cơ bản <input type="checkbox"/> Ứng dụng <input checked="" type="checkbox"/> Triển Khai <input type="checkbox"/> |                                 |  |
| <b>5. THỜI GIAN THỰC HIỆN:</b> 12 tháng (Từ tháng 03 năm 2024 đến tháng 02 năm 2025)                                                                                                                                                                                                                                                                                  |           |                                        |                                                                                                                                                    |                                 |  |
| <b>6. CƠ QUAN CHỦ TRÌ ĐỀ TÀI</b><br>Tên cơ quan: Trường Đại học Sư phạm Kỹ thuật<br>Điện thoại: 0236.3822571<br>E-mail: dhspktdn@ute.udn.vn<br>Địa chỉ: 48 Cao Thắng, Quận Hải Châu, TP. Đà Nẵng<br>Họ và tên thủ trưởng cơ quan chủ trì: PGS.TS. Phan Cao Thọ                                                                                                        |           |                                        |                                                                                                                                                    |                                 |  |
| <b>7. CHỦ NHIỆM ĐỀ TÀI</b><br>Họ và tên: Doãn Văn Khánh<br>Chức danh khoa học:<br>Địa chỉ cơ quan: 48 Cao Thắng, Quận Hải Châu, TP. Đà Nẵng<br>Điện thoại cơ quan:<br>Di động: 0396794322<br>E-mail: dvkhanh@ute.udn.vn<br>Học vị: Thạc sĩ<br>Năm sinh:<br>Địa chỉ nhà riêng: 35 Mai Am, Đà Nẵng<br>Điện thoại nhà riêng :<br>Fax:                                    |           |                                        |                                                                                                                                                    |                                 |  |
| <b>8. NHỮNG THÀNH VIÊN THAM GIA NGHIÊN CỨU ĐỀ TÀI</b>                                                                                                                                                                                                                                                                                                                 |           |                                        |                                                                                                                                                    |                                 |  |
| TT                                                                                                                                                                                                                                                                                                                                                                    | Họ và tên | Đơn vị công tác và lĩnh vực chuyên môn | Nội dung nghiên cứu cụ thể được giao                                                                                                               | Chữ ký                          |  |
| <b>9. ĐƠN VỊ PHỐI HỢP CHÍNH</b>                                                                                                                                                                                                                                                                                                                                       |           |                                        |                                                                                                                                                    |                                 |  |
| Tên đơn vị trong và ngoài nước                                                                                                                                                                                                                                                                                                                                        |           | Nội dung phối hợp nghiên cứu           |                                                                                                                                                    | Họ và tên người đại diện đơn vị |  |

## 10. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU THUỘC LĨNH VỰC CỦA ĐỀ TÀI Ở TRONG VÀ NGOÀI NƯỚC

10.1. Ngoài nước (Phân tích, đánh giá tình hình nghiên cứu thuộc lĩnh vực của đề tài trên thế giới, liệt kê danh mục các công trình nghiên cứu, tài liệu có liên quan đến đề tài được trích dẫn khi đánh giá tổng quan)

Dự báo nhu cầu tiêu thụ năng lượng điện là một trong những vấn đề thách thức, nguyên nhân chính là do có nhiều yếu tố không chắc chắn về nhu cầu sử dụng điện như điều kiện thời tiết, điều kiện kinh tế, thay đổi công nghệ, thói quen sử dụng...etc.. Vậy nên dự báo chính xác được coi là một bài toán khó với nhiều biến "không chắc chắn". Tuy nhiên, với sự phát triển của hệ thống quản lý năng lượng cũng như các thiết bị điện thông minh, ngành điện lực đã tạo ra một lượng lớn dữ liệu tiêu thụ điện với nhiều tiềm năng. Bởi lẽ thông qua nó có thể phản ánh hành vi/đặc điểm của người tiêu dùng. Do đó việc xây dựng các mô hình dự báo dựa trên trí tuệ nhân tạo với đầu vào kết hợp các nguồn dữ liệu này đã mang lại những hiệu quả nhất định.

Một số các nghiên cứu ngoài nước như nghiên cứu của tác giả J. F. Torres và đồng sự đã sử dụng thuật toán bộ nhớ ngắn hạn có trễ (Long Short Term Memory LSTM)-một thuật toán phổ biến của trí tuệ nhân tạo như để giải quyết bài toán dự báo nhu cầu tiêu thụ điện của Tây Ban Nha. Kết quả dự đoán thu được có độ chính xác cao đạt sai số dưới 1,5%. Ngoài ra, mạng LSTM được đề xuất đã thu được các sai số nhỏ nhất khi so sánh với các thuật toán hồi quy tuyến tính (Linear Regression), cây quyết định (K-Nearest Neighbor).

Để cải thiện tính chính xác của mô hình dự đoán, nền tảng GoogleTrend cũng được đưa vào nghiên cứu như một công cụ nâng cao khả năng dự báo. Trong nghiên cứu của Mavragani và các đồng sự đã kết luận rằng "Google Trend mang đến cho lĩnh vực dữ liệu lớn những cơ hội mới vì nó đã được chứng minh là hợp lệ và đã được chứng minh là có giá trị, chính xác và có lợi cho việc dự báo" [1]. Gần đây nhất, các nghiên cứu của tác giả Lean Yu và các đồng sự về sử dụng các dữ liệu từ Google Trend trong việc dự báo nhu cầu tiêu thụ dầu khí. Dựa trên nghiên cứu, nhóm tác giả đã chỉ ra rằng việc giới thiệu dữ liệu trực tuyến hữu ích, dưới dạng xu hướng (Trend) của Google, cải thiện đáng kể khả năng của các mô hình trong việc dự đoán cả xu hướng và giá trị tiêu thụ dầu. Thêm vào đó, việc sử dụng Google Trend, phản ánh chính xác các yếu tố liên quan khác nhau dựa trên vô số kết quả tìm kiếm, khiến mô hình dự báo dựa trên dữ liệu lớn trực tuyến này được đề xuất trở thành một công cụ đầy hứa hẹn để dự đoán mức tiêu thụ dầu.

1. Torres, José & Martínez-Álvarez, Francisco & Troncoso, Alicia. (2022). A deep LSTM network for the Spanish electricity consumption forecasting. Neural Computing and Applications. 34. 1-13. 10.1007/s00521-021-06773-2
2. M. Q. Raza and A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," Renewable and Sustainable Energy Reviews, vol. 50, pp. 1352-1372, 2015/10/01/ 2015.
3. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural computation 9:1735-80. <https://doi.org/10.1162/neco.1997.9.8.1735>.
4. Park, Sungjun & Kim, Jinsoo. (2018). The Effect of Interest in Renewable Energy on US Household Electricity Consumption: An Analysis Using Google Trends Data. Renewable Energy. 127. 10.1016/j.renene.2018.05.044.
5. Lean Yu, Yaqing Zhao, Ling Tang, Zebin Yang. Online big data-driven oil consumption forecasting with Google trends]. International Journal of Forecasting, 2017

[1] Mavragani A, Ochoa G and Tsagarakis KP (2018) Assessing the methods, tools, and statistical approaches in Google Trends research: Systematic review. Journal of Medical Internet Research 20(11): e270.

10.2. Trong nước (Phân tích, đánh giá tình hình nghiên cứu thuộc lĩnh vực của đề tài ở Việt Nam, liệt kê danh mục các công trình nghiên cứu, tài liệu có liên quan đến đề tài được trích dẫn khi đánh giá tổng quan)

Một số nghiên cứu trong nước sử dụng các thuật toán trí tuệ nhân tạo để giải quyết bài toán dự báo phụ tải phải kể đến như nghiên cứu của tác giả Hoàng Ngọc Hoài Quang và các đồng sự sử dụng thuật toán bộ nhớ dài-ngắn hạn (Long Short-Term Memory - LSTM) áp dụng cho dự báo phụ tải tỉnh Thừa Thiên Huế, với các dữ liệu đầu vào là phụ tải điện và nhiệt độ, lượng mưa. Hay nghiên cứu của tác giả Vũ Thị Anh Thơ sử dụng thuật toán bộ nhớ dài-ngắn hạn (LSTM) dự báo phụ tải ngắn hạn ở Đài Loan sử dụng các dữ liệu đầu vào là phụ tải điện và nhiệt độ. Các nghiên cứu này đều cho thấy mức độ chính xác trong dự báo khi so sánh với giá trị đo thực tế.

Nhìn chung, các nghiên cứu trong nước về dự báo phụ tải điện sử dụng mô hình mạng nơ-ron dựa trên thuật toán bộ nhớ dài-ngắn hạn và chỉ dừng lại ở việc sử dụng các dữ liệu đầu vào là nhiệt độ hoặc thời tiết kết hợp phụ tải để dự báo, chưa có nghiên cứu nào (dựa trên hiểu biết của chủ nhiệm đề tài) sử dụng các dữ liệu từ GoogleTrend trong việc cải thiện hiệu suất của chủ nhiệm đề tài) sử dụng các dữ liệu từ GoogleTrend trong việc cải thiện vào đó, các mô hình mạng nơ-ron dựa trên thuật toán bộ nhớ dài-ngắn hạn thường gặp phải một số hạn chế do việc xử lý tuần tự dữ liệu đầu vào và những thách thức liên quan đến lan truyền ngược theo thời gian, đặc biệt là khi xử lý các bộ dữ liệu có thời gian phụ thuộc dài. Với ưu thế khắc phục các nhược điểm từ mô hình LSTM, mô hình học sâu với Transformer cũng đã được nghiên cứu và ứng dụng rộng rãi. Transformer là một mô hình học sâu được ra mắt vào năm 2017 bởi một nhóm tác giả tại Google Brain được dùng chủ yếu ở lĩnh vực xử lý ngôn ngữ tự nhiên và thị giác máy tính. Đặc biệt trong những năm gần đây mô hình Transformers được áp dụng trong dự báo chuỗi thời gian.

[1] Hoàng Ngọc Hoài Quang, Hồ Nghĩa, Hoàng Xuân Thiện, Lê Nguyễn Ngọc Vinh, "Nghiên cứu ứng dụng công nghệ AI vào dự báo phụ tải để đánh giá TTĐN lưới điện trung áp Tỉnh TT-Huế giai đoạn 2021 – 2025", Kỷ yếu hội nghị khoa học và công nghệ điện lực toàn quốc năm 2022 chuyển đổi số và nâng cao hiệu quả vận hành hệ thống điện quốc gia, tập 2.

[2] Tác giả Vũ Thị Anh Thơ sử dụng thuật toán bộ nhớ ngắn hạn có trễ (Long Short-Term Memory - LSTM) dự báo phụ tải ngắn hạn ở Đài Loan sử dụng các dữ liệu đầu vào là phụ tải điện và nhiệt độ.

[3] Nguyễn Hồng Quân, Lê Thị Phương "Dự báo chuỗi thời gian với mô hình Transformers", Khoa Học Và Công Nghệ Qui.

10.3. Danh mục các công trình đã công bố thuộc lĩnh vực của đề tài của chủ nhiệm và những thành viên tham gia nghiên cứu

## 11. TÍNH CẤP THIẾT CỦA ĐỀ TÀI

Với xu hướng sản xuất điện năng phụ thuộc lớn vào thời tiết dựa trên năng lượng gió và hơn và đáng kể hơn trong tương lai. Điều này sẽ làm tăng nhu cầu cân bằng năng lượng. Cân bằng năng lượng đề cập đến việc sản xuất điện có sẵn nhanh chóng trong trường hợp sản xuất bị gián đoạn hoặc cần thiết để cân bằng mức tiêu thụ cao điểm - nó đảm bảo sự cân bằng giữa sản xuất và tiêu thụ điện. Cân bằng giữa sản xuất và tiêu thụ điện là một trong những nhiệm vụ cấp bách đảm bảo đến vấn đề an ninh năng lượng. Việc thực hiện nó phần lớn phụ thuộc vào phương tiện và phương pháp lập kế hoạch sản xuất điện. Dự báo là một tính hợp lệ của các quyết định quản lý năng lượng. Do đó việc áp dụng và nghiên cứu các mô hình dự báo nhu cầu tiêu thụ năng lượng điện khác nhau nhằm tìm ra mô hình dự báo cân bằng năng lượng. Với các mô hình dự báo được xây dựng trên nền tảng của trí tuệ nhân tạo thì mô hình dự báo và dữ liệu đầu vào là 2 yếu tố quan trọng hàng đầu quyết định đến tính chính xác của kết quả dự báo. Với các thuật toán trí tuệ nhân tạo để xây dựng các mô hình, với các dữ liệu đầu vào thì khác. Trong thời đại/kỷ nguyên số, dữ liệu được ví như một loại "dầu mỏ" mới, oxy của trí tuệ nhân tạo, việc truy cập hay sử dụng các dữ liệu để giải quyết các đề bài đưa ra cũng trở nên thách thức hơn khi mà đang dần hình thành nên ngành kinh tế dữ liệu. Việc tiếp cận các nguồn dữ liệu chính thống cũng như tìm kiếm các nguồn dữ liệu tiềm năng để giải quyết bài toán về dự báo nhu cầu tiêu thụ điện năng bên cạnh các nguồn dữ liệu truyền thống.

Với sự phát triển của máy tính và điện thoại thông minh, việc tìm kiếm theo thời gian thực trở nên dễ dàng và việc thu thập thông tin thông qua tìm kiếm trên Internet ngày càng phổ biến. Google Trends, một nền tảng truy vấn lượng tìm kiếm, cho phép thấy được mức độ phổ biến tương đối của các truy vấn tìm kiếm. Google Trend được xem như là một nguồn dữ liệu lớn, nhiều tiềm năng. Google Trend đã được chứng minh là hợp lệ và đã được chứng minh là có giá trị, chính xác và có lợi cho việc dự báo. Xuất phát từ vấn đề nêu trên, tác giả tiếp cận nghiên cứu ứng dụng Google Trend trong việc dự đoán nhu cầu tiêu thụ, sử dụng năng lượng điện.

## 12. MỤC TIÊU ĐỀ TÀI

- Xây dựng phương pháp dự báo nhu cầu tiêu thụ điện kết hợp các dữ liệu từ Google Trend.
- Đánh giá xem liệu Google Trend có thể cải thiện dự báo tiêu thụ năng lượng điện hay không?

## 13. ĐỐI TƯỢNG, PHẠM VI NGHIÊN CỨU

### 13.1. Đối tượng nghiên cứu

- Các dữ liệu từ Google Trend
- Dữ liệu phụ tải điện của Vương quốc Anh
- Mô hình Bộ nhớ dài – ngắn hạn (LSTM)
- Mô hình học sâu với Transformers

### 13.2. Phạm vi nghiên cứu

- Mô hình dự báo đang xét là mô hình dự báo ngắn hạn, không xét cho các dự báo trung hạn và dài hạn.
- Dự báo nhu cầu điện năng tại Vương quốc Anh

#### 14. CÁCH TIẾP CẬN, PHƯƠNG PHÁP NGHIÊN CỨU

##### 14.1. Cách tiếp cận

Trên nền tảng tính chính xác mô hình dự báo, nhóm tiến hành thiết lập thực nghiệm, kiểm chứng trên bộ dữ liệu thực. Các kết quả, nhận xét, kiến nghị được tổng kết trong một quyển báo cáo và một bài báo khoa học.

##### 14.2. Phương pháp nghiên cứu

- Xem xét các đề tài nghiên cứu liên quan, đề xuất ý tưởng và cách thực hiện
- Xây dựng các mô hình dự báo
- Thiết lập kiểm chứng, đánh giá trên các dữ liệu thực

#### 15. NỘI DUNG NGHIÊN CỨU VÀ KẾ HOẠCH NGHIÊN CỨU

##### 15.1. Nội dung nghiên cứu (Trình bày dưới dạng đề cương nghiên cứu chi tiết)

Nội dung 1. Tổng quan về nhu cầu tiêu thụ sử dụng năng lượng điện

Nội dung 2. Phương pháp luận giải quyết bài toán dự báo tiêu thụ năng lượng

Nội dung 3. Xây dựng các mô hình dự báo.

Nội dung 4. Thu nhập, xử lý dữ liệu

Nội dung 5. Kiểm chứng thực nghiệm, đánh giá kết quả

Nội dung 6. Hoàn thiện báo cáo tổng kết đề tài và nghiệm thu

##### 15.2. Kế hoạch nghiên cứu

| STT | Nội dung công việc                                                            | Sản phẩm         | Thời gian<br>(tháng/năm bắt đầu - tháng/năm kết thúc) | Người thực hiện<br>(ghi cụ thể họ và tên) – chèn cột số ngày công |    |
|-----|-------------------------------------------------------------------------------|------------------|-------------------------------------------------------|-------------------------------------------------------------------|----|
| 1.  | Nghiên cứu lý thuyết, kết quả thực nghiệm của các vấn đề liên quan đến đề tài | Báo cáo          | 3/2024-4/2024                                         | Doãn<br>Vân<br>Khánh                                              | 22 |
| 2.  | Phương pháp luận giải quyết bài toán dự báo tiêu thụ năng lượng điện          | Báo cáo          | 5/2024-6/2024                                         |                                                                   | 11 |
| 3.  | Xây dựng mô hình dự báo                                                       | Báo cáo          | 7/2024-9/2024                                         |                                                                   | 22 |
| 4.  | Thu nhập, xử lý dữ liệu                                                       | Bộ dữ liệu       | 10/2024-11/2024                                       |                                                                   | 11 |
| 5.  | Kiểm chứng thực nghiệm, đánh giá kết quả                                      | Báo cáo          | 12/2024-01/2025                                       |                                                                   | 11 |
| 6.  | Hoàn thiện báo cáo tổng kết đề tài và nghiệm thu                              | Báo cáo tổng kết | 02/2025                                               |                                                                   | -  |

## 16. SẢN PHẨM

### 16.1. Sản phẩm khoa học

- Bài báo đăng hoặc chấp nhận đăng trên tạp chí có tên trong danh mục Scopus/SCIE ☐
- Bài báo đăng hoặc chấp nhận đăng trên tạp chí / kỷ yếu được tính điểm trong danh mục HECDSNN ☒
- Sản phẩm khác:
  - Tài liệu hướng dẫn thí nghiệm ☐
  - Tài liệu hướng dẫn thực hành ☐
  - Bài thực hành/bài thí nghiệm ☐
  - Giáo trình ☐
  - Tài liệu tham khảo ☐

### 16.2. Sản phẩm đào tạo:

- Đào tạo Cao học ☐
- Đào tạo NCS ☐
- Hướng dẫn sinh viên bảo vệ đề tài/đồ án tốt nghiệp ☐
- Hướng dẫn sinh viên nghiên cứu khoa học ☐

### 16.3. Sản phẩm ứng dụng

- |                       |                          |                    |                          |                 |                          |
|-----------------------|--------------------------|--------------------|--------------------------|-----------------|--------------------------|
| Chương trình máy tính | <input type="checkbox"/> | Tiêu chuẩn         | <input type="checkbox"/> | Chiến lược      | <input type="checkbox"/> |
| Website               | <input type="checkbox"/> | Phương pháp        | <input type="checkbox"/> | Đề án           | <input type="checkbox"/> |
| Thiết bị máy móc      | <input type="checkbox"/> | Tài liệu dự báo    | <input type="checkbox"/> | Bản quy hoạch   | <input type="checkbox"/> |
| Dây chuyền công nghệ  | <input type="checkbox"/> | Luận chứng kinh tế | <input type="checkbox"/> | Vật liệu        | <input type="checkbox"/> |
| Mô hình               | <input type="checkbox"/> | Quy phạm           | <input type="checkbox"/> | Mẫu             | <input type="checkbox"/> |
| Sơ đồ, bản thiết kế   | <input type="checkbox"/> | Bản kiến nghị      | <input type="checkbox"/> | Giống cây trồng | <input type="checkbox"/> |
| Quy trình công nghệ   | <input type="checkbox"/> | Báo cáo phân tích  | <input type="checkbox"/> | Giống vật nuôi  | <input type="checkbox"/> |

### 16.4. Tên sản phẩm, số lượng và yêu cầu khoa học đối với sản phẩm

| TT | Tên sản phẩm                                                                                        | Số lượng | Yêu cầu khoa học                                                                                                                     |
|----|-----------------------------------------------------------------------------------------------------|----------|--------------------------------------------------------------------------------------------------------------------------------------|
| 1. | Bài báo khoa học đăng hoặc chấp nhận đăng trên tạp chí/kỷ yếu được tính điểm trong danh mục HECDSNN | 01       | Chủ nhiệm đề tài là tác giả thứ nhất hoặc tác giả liên hệ và đơn vị công tác thuộc trường Đại học Sư phạm Kỹ thuật – Đại học Đà Nẵng |
| 2. | Báo cáo tổng kết                                                                                    | 01       |                                                                                                                                      |

## 17. HIỆU QUẢ (giáo dục và đào tạo, kinh tế - xã hội)

- Làm tài liệu cho các học phần Lưới điện thông minh, Phương pháp tính trong Hệ thống điện, Lưới điện thông minh.
- Đề tài đóng góp 1 bài báo khoa học trong thành tích chung của trường Đại học Sư phạm kỹ thuật –ĐH Đà Nẵng.

## 18. PHƯƠNG THỨC CHUYỂN GIAO KẾT QUẢ NGHIÊN CỨU VÀ ĐỊA CHỈ ỨNG DỤNG

- Phương thức chuyển giao kết quả nghiên cứu: Tài liệu nghiên cứu chuyển giao trực tiếp cho Bộ môn Hệ Thống Điện, Khoa Điện-điện tử của Trường Đại học Sư phạm Kỹ thuật Đà Nẵng.
- Địa chỉ ứng dụng: Khoa Điện – Điện tử, bộ môn Hệ thống điện, ứng dụng cho Học phần Ứng dụng Tin học trong Hệ thống điện. Phương pháp tính trong Hệ thống điện.

# 19. KINH PHÍ THỰC HIỆN ĐỀ TÀI VÀ NGUỒN KINH PHÍ

Tổng kinh phí: 17.000.000 đồng

Bằng chữ: Mười bảy triệu đồng chẵn

Trong đó: Nguồn từ Trường: 17.000.000 Các nguồn khác:..... đồng (ĐVT: VNĐ)

| ST<br>T | Khoản chi, nội dung chi                                    | Thời<br>gian<br>thực<br>hiện | Tổng kinh<br>phí  | Nguồn kinh phí        |                       | Ghi<br>chú |
|---------|------------------------------------------------------------|------------------------------|-------------------|-----------------------|-----------------------|------------|
|         |                                                            |                              |                   | Kinh phí<br>từ Trường | Các nguồ<br>n<br>khác |            |
| 1       | Chi tiền công lao động trực tiếp                           | 03/2024-<br>02/2025          | 15.984.000        | 15.984.000            | 0                     |            |
| 2       | Chi mua vật tư, nguyên, nhiên,<br>vật liệu                 |                              | 0                 | 0                     | 0                     |            |
| 3       | Chi sửa chữa, mua sắm tài sản cố<br>định                   |                              | 0                 | 0                     | 0                     |            |
| 4       | Chi hội thảo khoa học, công tác<br>phí                     |                              | 0                 | 0                     | 0                     |            |
| 5       | Chi trả dịch vụ thuê ngoài phục<br>vụ hoạt động nghiên cứu |                              | 0                 | 0                     | 0                     |            |
| 6       | Chi điều tra, khảo sát thu thập số<br>liệu                 |                              | 0                 | 0                     | 0                     |            |
| 7       | Chi văn phòng, phẩm, thông tin<br>liên lạc, in ấn          | 03/2024-<br>02/2025          | 166.000           | 166.000               | 0                     |            |
| 8       | Chi họp hội đồng đánh giá,<br>nghiệm thu cấp cơ sở         |                              | 0                 | 0                     | 0                     |            |
| 9       | Chi quản lý chung                                          | 03/2024-<br>02/2025          | 850.000           | 850.000               | 0                     |            |
| 10      | Chi khác                                                   |                              | 0                 | 0                     |                       |            |
|         | <b>Tổng cộng</b>                                           |                              | <b>17.000.000</b> | <b>17.000.000</b>     | <b>0</b>              |            |

Ngày 28 tháng 03 năm 2024

**TM. HỘI ĐỒNG KHOA**

(ký, họ và tên)

*[Signature]*  
Đỗ Hoàng Vũ

Ngày 26 tháng 03 năm 2024

**Chủ nhiệm đề tài**

ThS. Doãn Văn Khánh

Đà Nẵng, ngày 28 tháng 03 năm 2024

**Cơ quan Chủ trì duyệt**

**KT. HIỆU TRƯỞNG**

**PHÓ HIỆU TRƯỞNG**



**PGS. TS. Võ Trung Hùng**

**BẢNG 1: BẢNG DỰ TRÙ KINH PHÍ THỰC HIỆN ĐỀ TÀI**

| STT | Khoản chi, nội dung chi                                 | Thời gian thực hiện | Tổng kinh phí     | Trong đó           |                |
|-----|---------------------------------------------------------|---------------------|-------------------|--------------------|----------------|
|     |                                                         |                     |                   | Kinh phí từ Trường | Các nguồn khác |
| 1   | Chi tiền công lao động trực tiếp                        | 03/2024-02/2025     | 15.984.000        | 15.984.000         | 0              |
| 2   | Chi mua vật tư, nguyên, nhiên, vật liệu                 |                     | 0                 | 0                  | 0              |
| 3   | Chi sửa chữa, mua sắm tài sản cố định                   |                     | 0                 | 0                  | 0              |
| 4   | Chi hội thảo khoa học, công tác phí                     |                     | 0                 | 0                  | 0              |
| 5   | Chi trả dịch vụ thuê ngoài phục vụ hoạt động nghiên cứu |                     | 0                 | 0                  | 0              |
| 6   | Chi điều tra, khảo sát thu thập số liệu                 |                     | 0                 | 0                  | 0              |
| 7   | Chi văn phòng, phẩm, thông tin liên lạc, in ấn          | 03/2024-02/2025     | 166.000           | 166.000            | 0              |
| 8   | Chi họp hội đồng đánh giá, nghiệm thu cấp cơ sở         |                     | 0                 | 0                  | 0              |
| 9   | Chi quản lý chung                                       | 03/2024-02/2025     | 850.000           | 850.000            | 0              |
| 10  | Chi khác                                                |                     | 0                 | 0                  |                |
|     | <b>Tổng cộng</b>                                        |                     | <b>17.000.000</b> | <b>17.000.000</b>  | <b>0</b>       |

**Bảng chữ : Mười bảy triệu đồng chẵn**

Ngày 26 tháng 03 năm 2024

**Chủ nhiệm đề tài**



ThS. Doãn Văn Khánh

**BẢNG 1A: BẢNG TÍNH TIỀN CÔNG LAO ĐỘNG PHÂN THEO KẾ HOẠCH NGHIÊN CỨU**

**Định mức thù lao tháng của chủ nhiệm (DMCN)**

đồng

2.700.000

| Số TT | Nội dung công việc                                                            | Chức danh khoa học | Hệ số lao động khoa học | Số người trong nhóm chức danh | Định mức thù lao tháng của chủ nhiệm (DMCN) (ĐVT: VNĐ) | Số ngày công (ĐVT: ngày) | Tổng số tháng quy đổi của chức danh/nhóm chức danh (ĐVT: tháng) | Tổng thù lao thực hiện nhiệm vụ (ĐVT: VNĐ) |
|-------|-------------------------------------------------------------------------------|--------------------|-------------------------|-------------------------------|--------------------------------------------------------|--------------------------|-----------------------------------------------------------------|--------------------------------------------|
| (1)   | (2)                                                                           | (2.1)              | (3)                     | (4)                           | (5)                                                    | (6)                      | (7) = 6/22 ngày                                                 | 8 = 3x5x7                                  |
| 1     | Nội dung nghiên cứu                                                           |                    |                         |                               |                                                        |                          | 8,30                                                            | 15.984.0000                                |
| 1     | Nghiên cứu lý thuyết, kết quả thực nghiệm của các vấn đề liên quan đến đề tài |                    |                         |                               |                                                        |                          | 1,0                                                             | 2.160.000                                  |
|       | Thù lao 01 thành viên chính                                                   |                    | 0,8                     | 1                             | 2.700.000                                              | 22                       | 1,0                                                             | 2.160.000                                  |
|       | Doãn Văn Khánh                                                                | Thành viên chính   | 0,8                     | 1                             | 2.700.000                                              | 22                       | 1,0                                                             | 2.160.000                                  |
| 2     | Phương pháp luận giải quyết bài toán dự báo tiêu thụ năng lượng điện          |                    |                         |                               |                                                        |                          | 0,50                                                            | 1.080.000                                  |
|       | Thù lao 01 thành viên chính                                                   |                    | 0,8                     | 1                             | 2.700.000                                              | 11                       | 0,50                                                            | 1.080.000                                  |

|   |                                                                                             |                  |     |   |           |    |                   |           |
|---|---------------------------------------------------------------------------------------------|------------------|-----|---|-----------|----|-------------------|-----------|
|   | Doãn Văn Khánh                                                                              | Thành viên chính | 0,8 | 1 | 2.700.000 | 11 | 0,50              | 1.080.000 |
| 3 | <b>Xây dựng mô hình dự báo</b>                                                              |                  |     |   |           |    | 1,0               | 2.160.000 |
|   | <i>Thù lao 01 thành viên chính</i>                                                          |                  | 0,8 | 1 | 2.700.000 | 22 | 1,0               | 2.160.000 |
|   | Doãn Văn Khánh                                                                              | Thành viên chính | 0,8 | 1 | 2.700.000 | 22 | 1,0               | 2.160.000 |
| 4 | <b>Thu nhập, xử lý dữ liệu để đánh giá mô hình</b>                                          |                  |     |   |           |    | 0,50              | 1.080.000 |
|   | <i>Thù lao 01 thành viên chính</i>                                                          |                  | 0,8 | 1 | 2.700.000 | 11 | 0,50              | 1.080.000 |
|   | Doãn Văn Khánh                                                                              | Thành viên chính | 0,8 | 1 | 2.700.000 | 11 | 0,50              | 1.080.000 |
| 5 | <b>Kiểm chứng thực nghiệm, đánh giá kết quả</b>                                             |                  |     |   |           |    | 0,50              | 1.080.000 |
|   | <i>Thù lao 01 thành viên chính</i>                                                          |                  | 0,8 | 1 | 2.700.000 | 11 | 0,50              | 1.080.000 |
|   | Doãn Văn Khánh                                                                              | Thành viên chính | 0,8 | 1 | 2.700.000 | 11 | 0,50              | 1.080.000 |
| 6 | <b>Thù lao của chủ nhiệm vụ:</b><br>( $TL_{CN} = 1,0 \times DM_{CN} \times 20\% \times T$ ) |                  | 1   | 1 | 2.700.000 |    | 2,40              | 6.480.000 |
|   |                                                                                             |                  |     |   |           |    | (=20% x 12 tháng) |           |

|     |                                                       |  |                                                                                                                                                                       |   |           |    |                                    |            |
|-----|-------------------------------------------------------|--|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|---|-----------|----|------------------------------------|------------|
| 7   | Thù lao của thư ký Khoa học                           |  | 0,3                                                                                                                                                                   | 1 | 2.700.000 |    | 2,40                               | 1.944.000  |
|     | $(TL_{TK} = 0,3 \times DM_{CN} \times 20\% \times T)$ |  |                                                                                                                                                                       |   |           |    | $(= 20\% \times 12 \text{ tháng})$ |            |
| III | Xây dựng báo cáo tổng kết                             |  | Nội dung công việc này do chủ nhiệm và thư ký khoa học của nhiệm vụ thực hiện, đã được tính trong thù lao của chủ nhiệm và thư ký khoa học, không tính thù lao riêng. |   |           |    |                                    |            |
|     | TỔNG CỘNG                                             |  |                                                                                                                                                                       |   |           | 77 | 8,30                               | 15.984.000 |

**Bằng chữ:** Mười lăm triệu chín trăm tám mươi bốn ngàn đồng.

**BẢNG 1B: BẢNG TÍNH TIỀN CÔNG LAO ĐỘNG PHÂN THEO THÀNH VIÊN THAM GIA NGHIÊN CỨU**

| STT | Họ và tên      | Chức danh        | Hệ số | Số ngày<br>(nếu có) | Tổng số tháng quy đổi của chức danh (DVT: tháng) | Định mức thù lao tháng của chủ nhiệm (DMCN) (DVT: VND) | Thành tiền (DVT: VND) |
|-----|----------------|------------------|-------|---------------------|--------------------------------------------------|--------------------------------------------------------|-----------------------|
| -1  | -2             | -3               | -4    | -5                  | -6                                               | -7                                                     | -8                    |
| 1   | Doãn Văn Khánh | Chủ nhiệm        | 1     |                     | 2,4                                              | 2.700.000                                              | 6.480.000             |
| 2   | Doãn Văn Khánh | Thành viên chính | 0,8   | 77                  | 3,5                                              | 2.700.000                                              | 7.560.000             |
| 4   | Doãn Văn Khánh | Thư ký khoa học  | 0,3   |                     | 2,4                                              | 2.700.000                                              | 1.944.000             |
|     | Tổng           |                  |       | 77                  | 8,3                                              | -                                                      | 15.984.000            |

- Mục 2Chi sửa vật tư, nguyên vật liệu:Không
- Mục 3Chi sửa chữa, mua sắm tài sản cố định:Không
- Mục 4Chi hội thảo khoa học, công tác phí:Không
- Mục 5Chi trả dịch vụ thuê ngoài phục vụ hoạt động nghiên cứu:Không
- Mục 6Chi điều tra, khảo sát thu nhập số liệu:Không

Mục 7. Chi văn phòng phẩm, thông tin liên lạc, in ấn:

| TT | Nội dung chi | Đơn vị tính | Số lượng | Đơn giá | Tổng kinh phí | Nguồn kinh phí (đồng) |            |
|----|--------------|-------------|----------|---------|---------------|-----------------------|------------|
|    |              |             |          |         |               | Kinh phí từ Trường    | Nguồn khác |
| -1 | -2           | -3          | -4       | -5      | (6) =(4)x(5)  | -7                    | -8         |
| 1  | In báo cáo   | Báo cáo     | 5        | 33.200  | 166.000       | 166.000               | 0          |
|    | Tổng cộng    |             |          |         | 166.000       | 166.000               | 0          |

Mục 8 Chi họp hội đồng đánh giá, nghiệm thu cấp cơ sở: Không

Mục 9 Chi quản lý chung: 5% \* 17.000.000=850.000 đồng

Mục 10 Chi khác: Không

Chủ nhiệm đề tài  
(ký, ghi họ và tên)



ThS. Doãn Văn Khánh

Đà Nẵng, ngày 28 tháng 03 năm 2024

Cơ quan chủ trì

KIỂM TRA TRƯỞNG  
PHÒNG KẾ TÍNH TOÁN



PGS. TS. Võ Trung Hùng

**HỢP ĐỒNG TRIỂN KHAI DỰ TÀI**

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

Ngày 01 tháng 04 năm 2024

**HỢP ĐỒNG TRIỂN KHAI THỰC HIỆN**  
**ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP TRƯỜNG**  
**Số: 23/HĐ-KHCNCS-2023**

- Căn cứ Bộ luật dân sự ngày 24 tháng 11 năm 2015;
- Căn cứ Luật khoa học và công nghệ ngày 18 tháng 6 năm 2013;
- Căn cứ Quyết định số 1749/QĐ-TTg ngày 08 tháng 11 năm 2017 của Thủ tướng Chính phủ về việc thành lập Trường Đại học Sư phạm Kỹ thuật thuộc Đại học Đà Nẵng;
- Căn cứ Thông tư số 03/2023/TT-BTC ngày 10 tháng 01 năm 2023 của Bộ trưởng Bộ Tài chính về việc quy định lập dự toán, quản lý sử dụng và quyết toán kinh phí ngân sách nhà nước thực hiện nhiệm vụ khoa học và công nghệ;
- Căn cứ Quyết định số 216/QĐ-ĐHSPKT ngày 09 tháng 3 năm 2018 của Hiệu trưởng Trường Đại học Sư phạm Kỹ thuật về việc ban hành Quy định về quản lý đề tài khoa học và công nghệ cấp cơ sở (cấp Trường) do Trường Đại học Sư phạm Kỹ thuật quản lý;
- Căn cứ Quyết định số 271/QĐ-ĐHSPKT ngày 29 tháng 03 năm 2024 của Hiệu trưởng Trường Đại học Sư phạm Kỹ thuật về việc phê duyệt kinh phí và giao nhiệm vụ thực hiện đề tài khoa học và công nghệ cấp Trường;

Sau khi xem xét mục tiêu, nội dung nghiên cứu và sản phẩm khoa học của đề tài KH&CN (sau đây gọi tắt là “đề tài”):

**“ĐÁNH GIÁ LỰA CHỌN ỨNG DỤNG GOOGLE TREND TRONG TỐI ƯU HÓA DỰ BÁO NHU CẦU TIÊU THỤ NĂNG LƯỢNG ĐIỆN DỰA TRÊN CÁC THUẬT TOÁN TRÍ TUỆ NHÂN TẠO”**

Mã số: T2023-06-23

**Bên A: Trường Đại học Sư phạm Kỹ thuật**

Đại diện: **PGS.TS. Phan Cao Thọ** Chức vụ: Hiệu trưởng

Địa chỉ: 48 Cao Thắng, Quận Hải Châu, TP. Đà Nẵng

Điện thoại: 0236.3822571

Fax: 0236.3894884

Số Tài khoản: 3716.1.1055693.00000 tại Kho bạc Nhà nước Đà Nẵng

**2. Bên B: Ông (Bà): Doãn Văn Khánh** là Chủ nhiệm đề tài

- Số CCCD: 044194011700 ngày cấp: 22/11/2022 nơi cấp: Cục Cảnh sát Quản lý Hành chính về trật tự xã hội.

- Số Tài khoản: 5611629911 tại ngân hàng BIDV

- Mã số thuế: 8864198743

- Đơn vị: Khoa Điện-điện tử., Trường Đại học Sư phạm Kỹ thuật, Đại học Đà Nẵng

- Điện thoại: 0396794322. Email: dvkhanh@ute.udn.vn

Cùng thỏa thuận và thống nhất ký kết Hợp đồng thực hiện đề tài KH&CN cấp Trường (sau đây gọi tắt là Hợp đồng) với các điều khoản sau:

**Điều 1. Đặt hàng và nhận đặt hàng thực hiện đề tài KH&CN cấp Trường kinh phí từ Quỹ Khoa học và Công nghệ của trường Đại học Sư phạm Kỹ thuật (sau đây gọi tắt là Quỹ KH&CN)**

Bên A đặt hàng và Bên B nhận đặt hàng thực hiện đề tài KH&CN cấp trường (sau đây gọi là đề tài) theo các nội dung trong Thuyết minh đề tài đã được phê duyệt (sau đây gọi tắt là Thuyết minh).

Thuyết minh là bộ phận không tách rời của Hợp đồng.

**Điều 2. Thời gian thực hiện Hợp đồng**

Thời gian thực hiện đề tài là 12 tháng, từ tháng 03 năm 2024 đến tháng 02 năm 2025.

**Điều 3. Kinh phí đề tài:**

**3.1. Kinh phí thực hiện:**

- Tổng kinh phí đề tài được cấp từ Quỹ KH&CN là 17.000.000.đồng (Bằng chữ : Mười bảy triệu đồng chẵn.).

- Kinh phí đề tài thực hiện theo phương thức khoán chi từng phần. Các phần công việc được giao khoán và không được giao khoán quy định cụ thể tại Điều 7, Thông tư 27/2015/TTLT-BKHCN-BTC ngày 30 tháng 12 năm 2015 của Bộ Khoa học và Công nghệ - Bộ Tài chính về Quy định khoán chi nhiệm vụ khoa học và công nghệ sử dụng ngân sách nhà nước.

**3.2. Tạm ứng và thanh quyết toán kinh phí**

- Tạm ứng đợt 1: Ngay sau khi ký kết hợp đồng, bên A sẽ tạm ứng cho bên B tối đa 50% giá trị hợp đồng khi có yêu cầu tạm ứng của bên B.

- Tạm ứng đợt 2: Bên A cấp tạm ứng tối đa 40% tổng giá trị còn lại của hợp đồng khi có đề nghị của bên B và sau khi bên B phải hoàn thành hồ sơ thanh toán tối thiểu bằng 50% mức kinh phí đã tạm ứng các đợt trước đó trên cơ sở đề tài được đánh giá hoàn thành tiến độ thực hiện đợt 1. Bên A thực hiện thanh toán cho bên B theo đúng quy định hiện hành, đồng thời thu hồi cho đến hết số tiền đã tạm ứng. Số tiền tạm ứng đợt 1 chưa được thu hồi thì sẽ được thu hồi trong lần thanh toán cuối cùng.

- Thanh toán đợt cuối: Sau khi đề tài được nghiệm thu chính thức, bên B phải hoàn thành và nộp hồ sơ thanh quyết toán kinh phí tất cả các nguồn kể từ ngày kết thúc nhiệm vụ theo quyết định phê duyệt nhiệm vụ và hợp đồng đã ký kết và trong phạm vi niên độ tài chính. Bên A thanh toán hết giá trị còn lại của hợp đồng từ nguồn do bên A quản lý theo đúng quy định hiện hành.

- Tùy theo kết quả thực hiện hợp đồng mà bên B chịu những xử lý tài chính trong quy định tại điều 6 hợp đồng này.

#### **Điều 4. Quyền và nghĩa vụ của các bên**

##### **1. Quyền và nghĩa vụ của Bên A**

- a) Cung cấp các thông tin cần thiết cho việc triển khai, thực hiện Hợp đồng;
- b) Bố trí cho Bên B số kinh phí từ Trường ĐHSPKT, ĐHĐN quy định tại Điều 3 Hợp đồng này theo tiến độ kế hoạch, tương ứng với các nội dung nghiên cứu được phê duyệt;
- c) Tổ chức phê duyệt kế hoạch đầu thầu, mua sắm máy móc, thiết bị, nguyên vật liệu và dịch vụ của đề tài bằng kinh phí do Bên A cấp (nếu có) theo quy định;
- d) Trước mỗi đợt cấp kinh phí, trên cơ sở báo cáo tình hình thực hiện đề tài của Bên B, Bên A căn cứ vào sản phẩm, khối lượng công việc đã hoàn thành theo Thuyết minh đề cấp tiếp kinh phí thực hiện Hợp đồng. Bên A có quyền thay đổi tiến độ cấp hoặc ngưng cấp kinh phí nếu Bên B không hoàn thành công việc đúng tiến độ, đúng nội dung công việc được giao;
- đ) Kiểm tra đột xuất để đánh giá tình hình Bên B thực hiện đề tài theo Thuyết minh. Trường ĐHSPKT, ĐHĐN thực hiện kiểm tra định kỳ đánh giá tình hình thực hiện đề tài theo quy định hiện hành về quản lý đề tài khoa học và công nghệ cấp Trường;
- e) Kịp thời xem xét, giải quyết theo thẩm quyền hoặc trình cấp có thẩm quyền giải quyết kiến nghị, đề xuất của Bên B về điều chỉnh nội dung chuyên môn, kinh phí và các vấn đề phát sinh khác trong quá trình thực hiện đề tài;
- g) Tổ chức đánh giá, nghiệm thu kết quả thực hiện đề tài của Bên B theo các yêu cầu, chỉ tiêu trong Thuyết minh;
- h) Có trách nhiệm cùng Bên B tiến hành thanh lý Hợp đồng theo quy định hiện hành;
- i) Phối hợp cùng Bên B xử lý tài sản được mua sắm bằng ngân sách Trường ĐHSPKT, ĐHĐN hoặc được tạo ra từ kết quả nghiên cứu của đề tài sử dụng ngân sách Trường ĐHSPKT, ĐHĐN (nếu có) theo quy định của pháp luật;
- k) Tiếp nhận kết quả thực hiện đề tài, bàn giao kết quả thực hiện đề tài cho tổ chức đề xuất đặt hàng hoặc tổ chức triển khai ứng dụng sau khi được nghiệm thu;
- l) Có trách nhiệm hướng dẫn việc trả thù lao cho tác giả nếu có lợi nhuận thu được từ việc ứng dụng kết quả của đề tài và thông báo cho tác giả việc bàn giao kết quả thực hiện đề tài (nếu có);
- m) Ủy quyền cho Bên B tiến hành đăng ký bảo hộ quyền sở hữu trí tuệ đối với kết quả thực hiện đề tài (nếu có) theo quy định hiện hành;
- n) Thực hiện các quyền và nghĩa vụ khác theo quy định của Luật KH&CN và các văn bản liên quan.

##### **2. Quyền và nghĩa vụ của Bên B**

- a) Tổ chức triển khai đầy đủ các nội dung nghiên cứu của đề tài đáp ứng các yêu cầu chất lượng, tiến độ và chi tiêu theo Thuyết minh;
- b) Cam kết thực hiện và bàn giao sản phẩm cuối cùng đáp ứng đầy đủ các tiêu chí đã được phê duyệt;
- c) Được quyền tự chủ, tự quyết định việc sử dụng phần kinh phí để thực hiện đề tài theo dự toán kinh phí đề tài, được quyền ký kết hợp đồng thuê khoán chuyên môn với các thành viên tham gia thực hiện; và các hợp đồng khác gồm: mua vật tư, nguyên, nhiên, vật liệu, sửa chữa, mua sắm tài sản cố định, dịch vụ thuê ngoài phục vụ nghiên cứu, văn phòng phẩm, thông tin liên lạc, in ấn có giá trị dưới 20 triệu đồng;
- d) Yêu cầu Bên A cung cấp thông tin cần thiết để triển khai thực hiện Hợp đồng;
- đ) Kiến nghị, đề xuất điều chỉnh các nội dung chuyên môn, kinh phí và thời hạn thực hiện Hợp đồng khi cần thiết;
- e) Yêu cầu Bên A cấp đủ kinh phí theo đúng tiến độ quy định trong Hợp đồng khi hoàn thành đầy đủ nội dung công việc theo tiến độ cam kết. Đảm bảo huy động đủ nguồn kinh phí khác theo cam kết. Sử dụng kinh phí đúng mục đích, đúng chế độ hiện hành và có hiệu quả;
- g) Đối với các hợp đồng khác được quy định tại điểm c, khoản 2, Điều 4 hợp đồng này có giá trị từ 20 triệu trở lên: Chủ nhiệm đề tài xây dựng kế hoạch mua sắm để gửi Bên A phê duyệt và thực hiện mua sắm theo quy định của pháp luật;
- h) Chấp hành các quy định pháp luật trong quá trình thực hiện Hợp đồng. Tạo điều kiện thuận lợi và cung cấp đầy đủ thông tin cho các cơ quan quản lý trong việc giám sát, kiểm tra, thanh tra đối với đề tài theo quy định của pháp luật;
- i) Thực hiện việc tự đánh giá, nghiệm thu cấp cơ sở theo quy định hiện hành khi kết thúc đề tài. Sau khi đánh giá, nghiệm thu cấp cơ sở hoàn chỉnh lại hồ sơ theo kết luận của Hội đồng đánh giá cấp cơ sở, Bên B có trách nhiệm chuyển cho Bên A các hồ sơ để Bên A tiến hành việc đánh giá, nghiệm thu theo quy định;
- k) Có trách nhiệm quản lý tài sản được mua sắm bằng ngân sách của Trường ĐHSPKT, ĐHĐN hoặc được tạo ra từ kết quả nghiên cứu của đề tài sử dụng ngân sách Trường ĐHSPKT, ĐHĐN (nếu có). Chủ nhiệm đề tài có trách nhiệm bàn giao tài sản được mua sắm bằng ngân sách Trường ĐHSPKT, ĐHĐN hoặc được tạo ra từ kết quả nghiên cứu của đề tài cho cơ quan chủ quản đề tài để quản lý và sử dụng.
- l) Có trách nhiệm cùng Bên A tiến hành thanh lý Hợp đồng theo quy định;
- m) Thực hiện việc đăng ký bảo hộ quyền sở hữu trí tuệ theo ủy quyền của Bên A đối với kết quả nghiên cứu (nếu có);
- n) Chủ nhiệm đề tài giao nộp kết quả thực hiện đề tài cho bộ phận lưu giữ thông tin của cơ quan chủ quản đề tài. Cơ quan chủ quản đề tài xác nhận việc giao nộp kết quả thực hiện đề tài cho chủ nhiệm đề tài.

o) Các sản phẩm khoa học công bố được tính khi có lời ghi nhận Trường ĐHSPT, ĐHĐN, cụ thể như sau:

+ Đối với bài báo khoa học viết bằng tiếng Anh: "This research is funded by The University of Danang – University of Technology and Education under project number T2023-06-23";

+ Đối với bài báo khoa học viết bằng tiếng Việt: "Nghiên cứu này được tài trợ bởi Trường Đại học Sư phạm Kỹ thuật, Đại học Đà Nẵng trong đề tài có mã số T2023-06-23";

Đối với sản phẩm khoa học có lời ghi nhận Trường ĐHSPT, ĐHĐN và các nguồn kinh phí trong nước khác, Trường ĐHSPT, ĐHĐN sẽ tính mức độ hoàn thành của sản phẩm đó như sau: Phần công trình được tính cho sản phẩm đề tài này được chia theo tỷ lệ đóng góp của mỗi nguồn kinh phí thực hiện. Trong trường hợp không xác định được cụ thể tỷ lệ đóng góp của mỗi nguồn kinh phí thì phần công trình được tính cho sản phẩm đề tài này sẽ bằng một (01) chia cho số nguồn kinh phí trong nước. Tổng số phần các công trình được tính phải đáp ứng yêu cầu về số lượng theo thuyết minh đề tài đã phê duyệt.

Thời gian thực hiện sản phẩm của đề tài phải trong khoảng thời gian thực hiện đề tài.

p) Chủ nhiệm đề tài cùng với các cá nhân trực tiếp sáng tạo ra kết quả nghiên cứu khoa học và phát triển công nghệ được đứng tên tác giả trong đề tài và hưởng quyền tác giả bao gồm cả các lợi ích thu được (nếu có) từ việc khai thác thương mại các kết quả thực hiện đề tài theo quy định pháp luật và các thỏa thuận khác (nếu có);

q) Có trách nhiệm trực tiếp hoặc tham gia triển khai ứng dụng kết quả nghiên cứu khoa học và phát triển công nghệ theo yêu cầu của Bên A hoặc tổ chức, cá nhân được Bên A giao quyền sở hữu, sử dụng kết quả thực hiện đề tài;

r) Thực hiện bảo mật các kết quả thực hiện đề tài theo quy định về bảo vệ bí mật của nhà nước;

s) Thực hiện các quyền và nghĩa vụ khác theo quy định Luật KH&CN và các văn bản liên quan.

## **Điều 5. Chấm dứt Hợp đồng**

Hợp đồng này chấm dứt trong các trường hợp sau:

1. Đề tài đã kết thúc và được nghiệm thu.
2. Bên B bị chấm dứt hợp đồng thực hiện đề tài khi có đề nghị thanh lý Hợp đồng của Hội đồng thanh lý đề tài cấp Trường.

## **Điều 6. Xử lý tài chính khi chấm dứt Hợp đồng**

1. Đối với đề tài đã kết thúc và được nghiệm thu:

a) Đề tài đã kết thúc và đánh giá nghiệm thu từ mức “Đạt” trở lên thì Bên A thanh toán đầy đủ kinh phí cho Bên B theo quy định tại Hợp đồng này.

b) Đề tài đã kết thúc, nhưng nghiệm thu mức “không đạt” thì Bên A xem xét, quyết toán kinh phí cho Bên B trên cơ sở kết luận của Hội đồng đánh giá về nguyên nhân, trách nhiệm và những nội dung công việc mà Bên B đã thực hiện có sản phẩm thực tế được đánh giá.

Bên B có trách nhiệm hoàn trả toàn bộ số kinh phí ngân sách nhà nước đã cấp nhưng chưa sử dụng. Đối với khoản kinh phí đã sử dụng được áp dụng xử lý đối với trường hợp đề tài không hoàn thành được quy định tại Điều 16 Thông tư liên tịch số 27/2015/TTLT-BKHCN-BTC ngày 30/12/2015 của Bộ Khoa học và Công nghệ - Bộ Tài chính quy định khoản chi thực hiện nhiệm vụ khoa học và công nghệ sử dụng ngân sách Nhà nước.

2. Đối với đề tài chấm dứt khi có căn cứ khẳng định không còn nhu cầu thực hiện:

a) Trường hợp Đề tài chấm dứt khi có căn cứ khẳng định không còn nhu cầu thực hiện thì hai bên cùng nhau xác định khối lượng công việc Bên B đã thực hiện để làm căn cứ thanh toán số kinh phí Bên B đã sử dụng nhằm thực hiện đề tài và thu hồi số kinh phí còn lại đã cấp cho Bên B.

b) Trường hợp hai bên thỏa thuận ký Hợp đồng mới để thay thế và kết quả nghiên cứu của Hợp đồng cũ là một bộ phận cấu thành kết quả nghiên cứu của Hợp đồng mới thì số kinh phí đã cấp cho Hợp đồng cũ được tính vào kinh phí cấp cho Hợp đồng mới và được tiếp tục thực hiện với Hợp đồng mới.

3. Đối với Đề tài bị đình chỉ theo quyết định của cơ quan có thẩm quyền hoặc Hợp đồng bị chấm dứt do Bên B không nộp hồ sơ để đánh giá, nghiệm thu Đề tài theo quy định pháp luật thì Bên A sẽ tiến hành thủ tục thanh lý đề tài theo quy định.

4. Đối với Đề tài không hoàn thành do lỗi của Bên A dẫn đến việc chấm dứt Hợp đồng thì Bên B không phải bồi hoàn số kinh phí đã sử dụng để thực hiện Đề tài, nhưng vẫn phải thực hiện việc quyết toán kinh phí theo quy định của pháp luật.

## **Điều 7. Xử lý tài sản khi chấm dứt Hợp đồng**

1. Khi chấm dứt Hợp đồng, việc xử lý tài sản được mua sắm hoặc được hình thành bằng ngân sách Trường ĐHSPKT, ĐHQĐN cấp cho đề tài được thực hiện theo quy định pháp luật.

2. Các sản phẩm vật chất của Đề tài sử dụng ngân sách Trường ĐHSPKT, ĐHQĐN: nguồn thu khi các sản phẩm này được tiêu thụ trên thị trường sau khi trừ các khoản chi phí cần thiết, hợp lệ, được phân chia theo quy định pháp luật.

## **Điều 8. Điều khoản chung**

1. Trong quá trình thực hiện Hợp đồng, nếu một trong hai bên có yêu cầu sửa đổi, bổ sung nội dung hoặc có căn cứ để chấm dứt thực hiện Hợp đồng thì phải thông báo cho bên kia ít nhất là 15 ngày làm việc trước khi tiến hành sửa đổi, bổ sung hoặc chấm dứt thực hiện

Hợp đồng, xác định trách nhiệm của mỗi bên và hình thức xử lý. Các sửa đổi, bổ sung (nếu có) phải lập thành văn bản có đầy đủ chữ ký của các bên và được coi là bộ phận của Hợp đồng và là căn cứ để nghiệm thu kết quả của đề tài.

2. Khi một trong hai bên gặp phải trường hợp bất khả kháng dẫn đến việc không thể hoặc chậm thực hiện nghĩa vụ đã thỏa thuận trong Hợp đồng thì có trách nhiệm thông báo cho Bên kia trong 10 ngày làm việc kể từ ngày xảy ra sự kiện bất khả kháng. Hai bên có trách nhiệm phối hợp xác định nguyên nhân và báo cáo cơ quan quản lý nhà nước có thẩm quyền để giải quyết theo quy định của pháp luật.

3. Hai bên cam kết thực hiện đúng các quy định của Hợp đồng và có trách nhiệm hợp tác giải quyết các vướng mắc phát sinh trong quá trình thực hiện. Bên vi phạm các cam kết trong Hợp đồng phải chịu trách nhiệm theo quy định pháp luật.

4. Mọi tranh chấp phát sinh trong quá trình thực hiện Hợp đồng do các bên thương lượng hòa giải để giải quyết. Trường hợp không hòa giải được thì một trong hai bên có quyền đưa tranh chấp ra để giải quyết theo quy định của pháp luật.

### Điều 9. Hiệu lực của Hợp đồng

Hợp đồng này có hiệu lực từ ngày ký. Hợp đồng này được lập thành 06 bản (01 bản gốc và 05 bản chính) và có giá trị như nhau, Bên A giữ 02 bản, Bên B giữ 04 bản có giá trị pháp lý như nhau.

**BÊN A**

**CƠ QUAN CHỦ QUẢN**



**PGS.TS. Phan Cao Thọ**

**BÊN B**

**CHỦ NHIỆM ĐỀ TÀI**

**Doãn Văn Khánh**

# MỤC LỤC MINH CHỨNG SẢN PHẨM CỦA ĐỀ TÀI

Họ và tên chủ nhiệm đề tài: Doãn Văn Khánh

Mã số đề tài: T2023-06-23

| STT                      | Tên minh chứng                                                                                                                                                                                              |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>SẢN PHẨM KHOA HỌC</b> |                                                                                                                                                                                                             |
| <input type="checkbox"/> | * Bài báo đăng trên tạp chí/kỷ yếu được tính điểm trong danh mục HĐCDGSNN                                                                                                                                   |
|                          | Improving Energy Demand Prediction Use Deep Learning Network”, 9th International Conference on Applying New Technology in Green Buildings (ATiGB), 2024, pp.323-327. DOI: 10.1109/ATiGB63471.2024.10717715. |
| -                        | Trang bìa tạp chí/kỷ yếu                                                                                                                                                                                    |
| -                        | Mục lục                                                                                                                                                                                                     |
| -                        | Toàn văn bài báo                                                                                                                                                                                            |
| -                        | Minh chứng chứng minh tạp chí/kỷ yếu thuộc danh mục tính điểm khoa học của Hội đồng chức danh giáo sư nhà nước.                                                                                             |
|                          |                                                                                                                                                                                                             |



IEEE  
**SMC**

Systems, Man, and Cybernetics Society



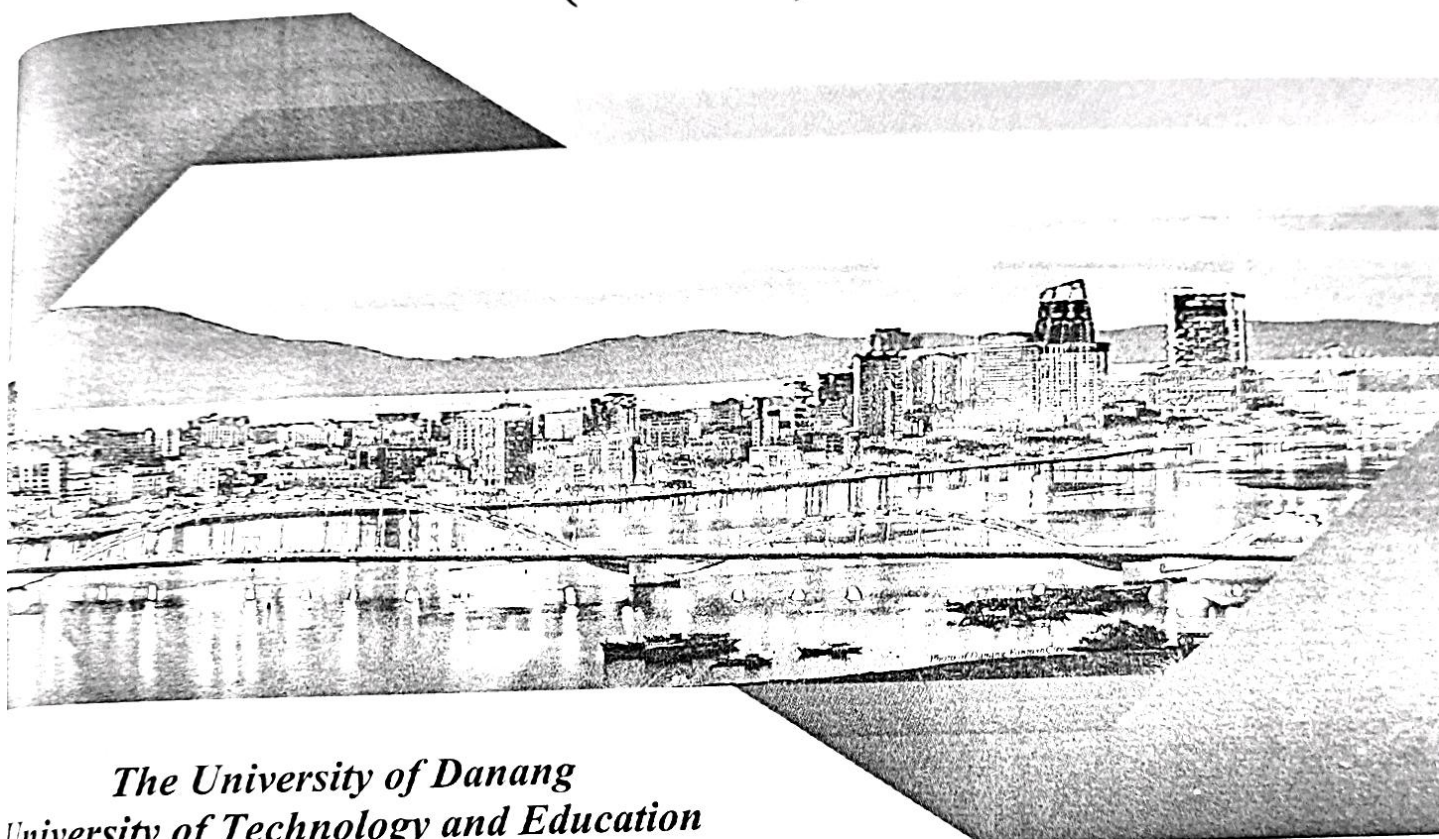
**IEEE**



**VINIF**

Powered by VinBigdata

# 2024 9<sup>th</sup> International Conference on Applying New Technology in Green Buildings (ATiGB)



*The University of Danang*  
*University of Technology and Education*  
20 Thang Street, Hai Chau District, Danang City, Vietnam

*August 30<sup>th</sup> – 31<sup>st</sup>, 2024*

# Table of Contents

|                                                                                                                                                                                                                                     |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Copyright Page                                                                                                                                                                                                                      | ix   |
| Organizing Committee Information                                                                                                                                                                                                    | x    |
| Authors Index                                                                                                                                                                                                                       | xiii |
| Effect Of Mimosa Pudica Linn Extract On Excision Wound Healing<br><i>Chiuyen Phan, Tantai Vo, Dinhtrung Tran, Huongthi Cao, Chithien Nguyen, Huynhthong Nguyen, Anhkhoa Le, Minhkhue Tran Tran and Thiphuong Nguyen</i>             | 1    |
| Building A Multilingual Guide System on Mobile Devices at Danang Museum<br><i>Trung Hung Vo and Thi Le Thuyen Phan</i>                                                                                                              | 6    |
| Higher Order Mode Cluster (HOMC) Generation By A Quantitative Laser Ultrasound Visualization (QLUV) Technique<br><i>Le Van Nguyen, Che Hua Yang and He Thong Bui</i>                                                                | 12   |
| Study of Residual Stress on the Copper Plate Using Laser Ultrasound Technique<br><i>Che Hua Yang, Le Van Nguyen and He Thong Bui</i>                                                                                                | 16   |
| An In Silico Evaluation Of The Multifunctional Activity Of Crocetin Against Neurodegenerative Disease<br><i>Le Duc Loc, Nguyen Minh Thong, Quan V. Vo, Mai Van Bay, Adam Mechler, Pham Cam Nam, Dinh Tien Hieu and Dinh Van Tac</i> | 20   |
| The Use of Mitochondrial Cytochrome C Oxidase Subunit I Gene in Identifying Pila Species in The Mekong Delta<br><i>Trương Văn Xạ and Trần Kim Thoa</i>                                                                              | 28   |
| Analyzing Bearing Capacity of A Shallow Foundation on Cohesion-Less Soils Under Combined Loads<br><i>Quang Pham Ngoc, Vinh Pham Ngoc and Hoa Hoang Phuong</i>                                                                       | 32   |
| Evaluating Static and Kinetic Friction Coefficients of Fabrics through Experimental Model<br><i>Tuan Anh Nguyen, Thanh Hau Nguyen, Thu Ha Le Dinh and Bao Ngoc Tran</i>                                                             | 38   |
| Examine The Effectiveness Of Transmission Laser Welding In Bonding Copper To Glass Joints<br><i>Hoai Nguyen, Minh-Tien Nguyen, Le Chau Thanh Nguyen, Minh-Tung Phung and Huu Phuoc Trang Nguyen</i>                                 | 43   |
| Robust Sliding Mode Control for Brushless DC Motor with Consideration of Load Torque<br><i>Duc Tai Luong Cu, Van Nam Giap and Van Dong Doan</i>                                                                                     | 48   |
| High-Frequency Response Of Double-Layer Varying Thickness FG Plates Considering Sliding Motions<br><i>Thi Hai Van Nguyen and Tam Ta Duc</i>                                                                                         | 53   |

|                                                                                                                                                                                                                                |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Applying Lean Tools in The Shaped Plastic Products Industry: A Case Study<br><i>Thi-Anh-Tuyet Nguyen and Phi-Trung Nguyen</i>                                                                                                  | 269 |
| Design And Simulation Of Vertical Agricultural Farm<br><i>Phi Trung Nguyen and Thi-Anh-Tuyet Nguyen</i>                                                                                                                        | 275 |
| Design and Simulation Industrial Wood Furniture Manufacturing Plants<br><i>Phi Trung Nguyen and Thi-Anh-Tuyet Nguyen</i>                                                                                                       | 280 |
| Study on Al-Cu Brazing Using TIG Welding Technology<br><i>Nam Hoai Nguyen, Nghia Trong Hoang, Khiem Gia Thai, Phoi Vinh Nguyen, Doan Van Nguyen and Long Nhut-Phi Nguyen</i>                                                   | 285 |
| Study On Some Properties Of Pectin/Carboxymethyl Cellulose/Nanocurcumin Films<br><i>Thi Minh Phuong Ngo and Thanh Nguyen Nguyen</i>                                                                                            | 289 |
| Enhancing Prosthetic Accessibility in Vietnam: A Novel Digital Design Process for Amputation Sockets<br><i>Thai Duong Nguyen, Nguyen Duy Minh Phan and Dinh Son Nguyen</i>                                                     | 294 |
| Research on Hazardous Sludge Treatment by Hardening Method for Da Nang Mechanical and Plating Factory<br><i>Thi Xuan Thuy Le, Dinh Long Pham, Hoang Tri Vũ and Thi Phu Tran</i>                                                | 300 |
| Improving The Accuracy of Robotic Arm Using A Calibration Method Based On The Inverse Kinematics Problem<br><i>Phi Nguyen-Thanh, Hung Nguyen-Quoc, Toan Pham-Bao, Luan Vuong-Cong, Hoang Truong-Viet and Vinh Nguyen-Quang</i> | 306 |
| Investigation Of Antibiotic Resistance Characteristics Of Lactobacillus Species Isolated From Meat-Fermented Food<br><i>Thi Dong Phuong Nguyen and Le Hien Hoa Nguyen</i>                                                      | 312 |
| Control Strategy for Liquid Transfer Using a Four-wheel Mecanum Mobile Robot Platform<br><i>Viet Khanh Nguyen, Hue Luu Thi, Duc Thanh Cao, Tung Lam Nguyen, Duc Minh Do and Thanh Ha Vo</i>                                    | 316 |
| Improving Energy Demand Prediction Use Deep Learning Network<br><i>Van Khanh Doan</i>                                                                                                                                          | 323 |
| Study On The Ball Screw Feed Drive System Identification By Using Unscented Kalman Filter Considering Noise Effect<br><i>Hung Nguyen-Quoc, Toan Pham-Bao, Vinh Nguyen-Quang, Huy Huynh-Hoang, Nhi Ngo-Kieu and Tri Truong</i>  | 328 |
| Numerical Design of a 1:3 Optical Power Splitter Based on a 2×3 RI-MMI Coupler Using Silicon Waveguides for C-Band<br><i>Thuy Tran Thi Thanh, Hien Nguyen Trung, Hung Nguyen Tan and Dung Truong Cao</i>                       | 334 |
| Continuous Wearable-based Fall Detection using Tiny Machine Learning<br><i>Duan Luong Cong, Bien Nguyen Quang, Duc Truong Minh, Dung Truong Cao and Minh Nguyen Ngoc</i>                                                       | 339 |

# Improving Energy Demand Prediction Use Deep Learning Network

1<sup>st</sup> Van Khanh Doan

Faculty of Electrical and Electronics  
Engineering

The University of Danang, University  
of Technology and Education

DaNang, Vietnam

dvkhanh@ute.udn.vn

2<sup>nd</sup> Dong Nguyen Doan

Faculty of Applied Science  
International School, Vietnam

National University

Hanoi, Vietnam

dongnd@vnu.edu.vn

3<sup>rd</sup> Binh Hau Nguyen

Faculty of Electrical and Electronics  
No. 2 Posts and Telecommunications

Institute of Technology

Ho Chi Minh, Vietnam

nguyenbinhau@ptithcm.edu.vn

**Abstract**— In recent years, the convergence of deep learning and energy demand forecasting, particularly for short-term power demand, has become a burgeoning field of research. As societies become increasingly reliant on electricity for various aspects of life, accurate forecasting has become critical for maintaining a stable and efficient electricity supply. In this study, we evaluate the performance of various deep learning algorithms to enhance the accuracy of short-term electricity demand forecasting. Additionally, we consider the impact of public concern, as measured by Google Trends, as an exogenous variable that affects electricity demand forecasts. The experimental results show that the hybrid model combining Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM) architecture outperforms both the stand-alone LSTM and Sequence to Sequence model with Attention (Seq2Seq-Attention), providing a better prediction in both cases, whether including or excluding Google Trends data.

**Keywords**—Electricity demand forecasting, LSTM, CNN-LSTM, Seq2Seq-Attention, Google Trends.

## I. INTRODUCTION

The transition to the centralization of energy production has brought about profound economic, technological, social, and environmental changes, while also presenting significant challenges. Among these challenges is the need for accurate electricity demand forecast to balance the grid in real-time. Demand forecasting is an integral component in grid adequacy planning. Specifically, short-term forecasting of electricity is highly relevant for energy market trading and optimizing energy grid usage.

Electricity demand forecasting, however, is a broad field and exhibits variation in terms of scales, timeframes, modeling frameworks, and overall aims. In general, forecasting approaches range from minutes ahead to decades ahead, covering scales from single buildings to country or regional levels. In this study, we focus on short-term forecasts of electricity demand at the regional level, excluding approaches aimed at local analyses (e.g., single buildings, residential areas). This approach considers country-specific peculiarities in data collection and availability, energy system infrastructure, and economic conditions affecting energy consumption. In fact, in some regions, especially in developing countries, some relevant data, such as real-time grid operation data, may not be readily accessible to external researchers due to privacy and security concerns. Additionally, the limited infrastructure for data collection challenges the ability to develop accurate forecasting models. Therefore, in this paper, besides studying various deep learning models and their performance, we also investigate Google Trends data and its impact on the output forecast. The

free accessibility and long-term trends available through Google Trends can complement traditional datasets, offering valuable insights into various aspects of public interest and behavior [1]–[6]. Sungjun Park & Jinsoo Kim, in their research [2], demonstrate the effectiveness of Google Trends as a helpful predictor for residential electricity consumption prediction. Their research used a static empirical model that assumed the relationships between variables are stable over the period being analyzed. However, external environmental or societal influences also made significant contributions to electricity demand besides internal factors. Thus, the nature of electricity demand is nonlinear and dynamic. To address the nonlinear nature of power demand, numerous deep learning models have been proposed and have shown good performance in real-world applications [7–9]. Some variants of Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM) network and Gated Recurrent Unit network (GRU), Sequence to Sequence (Seq2Seq) are commonly applied to deep learning models for time series forecasting [10–16]. The advantages of LSTM and GRU compared with traditional RNN in the direction of power load prediction have been validated in the literature [17]. In addition, the approach of incorporating various deep learning models, such as combining GRU-LSTM, Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM) architecture, provides powerful tools for capturing complex patterns in data as well as improving their performance in forecasting electricity demand [12], [14], [18]. In recent years, the introduction of attention mechanisms have also brought significant benefits to time series forecasting. The attention mechanism can also integrate into Seq2Seq improved the ability of Seq2Seq model to handle long sentences by focusing on relevant parts of the input sequence dynamically [20]. By mean, each model framework has its own advantage and drawbacks in terms of complexity in training, performance, therefore the forecasting accuracy can change significantly among different samples.

In summary, this study performs different deep learning models, such as LSTM, hybrid CNN-LSTM, and Seq2Seq-Attention, using energy demand data from a specific region, thereby, the feasibility of deep neural network frameworks is verified. Furthermore, we studied proposed models that incorporate Google Trends data to enhance prediction accuracy.

## II. METHODOLOGY

This section outlines a methodology for forecasting electricity demand, emphasizing the integration of public

concern indicators derived from Google Trends into the forecast models. Generally, two key steps are involved in implementing the model:

- Step 1: Casuality test

The main purpose of this step is to assess the effectiveness of Google Trends in terms of its correlation with electricity demand. Therefore, we begin with Granger causality test to investigate these relationships.

- Step 2: Forecast improvement

In this step, the performance of various deep learning algorithms LSTM, hybrid LSTM-CNN, and Seq2Seq-Attention, will be evaluated, considering both the inclusion and exclusion of the exogenous variable as a public concern captured by Google Trends.

#### A. Casuality test

Granger Causality, named after Nobel Laureate Clive W. Granger, is essentially a statistical concept used to determine if one time series can help predict another time series [19]. Considered by an autoregressive (AR) model and an autoregressive distributed lag (ADL) model as follow:

$$Y_t = a_n + \sum_{i=1}^p (a_i Y_{t-i}) + \epsilon_t \quad (1)$$

$$Y_t = \beta_0 + \sum_{i=1}^p (\beta_i Y_{t-i}) + \sum_{i=1}^q (\gamma_i X_{t-i}) + \theta_t \quad (2)$$

Model (1) uses only the past values of Y to predict Yt however, model (2) incorporates both past values of Yt and past values of an external variable X, offering a more comprehensive approach to forecasting when external factors are relevant. In order to determine whether Xt Granger-causes Yt, we perform an F-test as follows:

$$F = \frac{(RSS_1 - RSS_2)/q}{RSS_2/(T - 2p - q)} \quad (3)$$

Here, RSS1 and RSS2 are the residual sums of squares (RSS) from models (1) and (2); T is the number of observations. If the F-statistic is significant at the chosen level (commonly value 5% was chosen to this study), we can reject the null hypothesis and then conclude that adding X in a model for predicting Y significantly improves the model's ability to predict Y.

#### B. Forecast improvement

Deep learning has been applied to time series forecasting for a long time and has achieved many successes. In this work, standalone LSTM, hybrid CNN-LSTM, and Seq2Seq-Attention models are proposed to predict future electricity demand.

##### 1) LSTM

Initially developed to address the shortcomings of vanilla recurrent neural networks, LSTM networks excel in handling long-term dependencies and patterns within sequential data, making them particularly suited for applications in time series prediction. A typical LSTM cell is shown in Fig 1, where it handles vectors for long-term memory  $c_{t-1}$ , short-term memory  $h_{t-1}$  and the input  $x_t$  at for time t. Point-wise operations (multiplication or addition) are shown by circular markers, while activations (tanh or sigmoid functions) are represented by elongated markers.

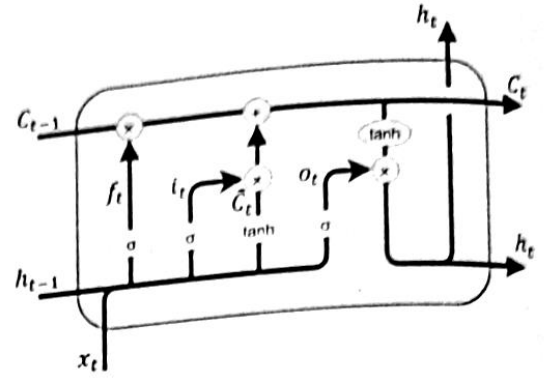
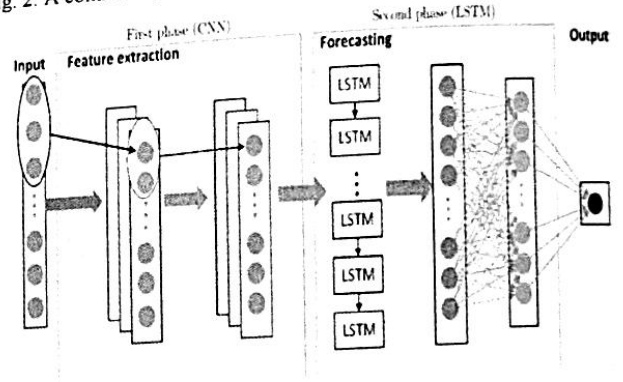


Fig. 1. Long Short-Term Memory block with one cell

##### 2) CNN-LSTM

The CNN-LSTM architecture takes advantage of CNNs for feature extraction on input data, while incorporating LSTMs to assist sequence prediction. This enables the model to effectively handle tasks that involve both spatial and temporal data. As can be seen from Fig.2, the process begins with an input sequence that undergoes feature extraction through multiple convolutional layers. These extracted features are then processed by a series of LSTM layers to capture temporal dependencies and generate the final forecasted output.

Fig. 2. A common hybrid CNN-LSTM network



##### 3) Seq2Seq with Attention

Seq2Seq or (Sequence-to-Sequence) is a structure of deep learning models. Seq2Seq was originally developed for machine translation [20]. Recently, this architecture has also been applied to tackle time series forecasting problems such as low SNR (signal to noise ratio), low autocorrelation, etc. The structure of Seq2seq includes an encoder and a decoder. The encoder compresses inputs into latent vectors, and the decoder decodes the latent vectors into appropriate outputs. The internal structures of Seq2Seq can contain convolutional layers [21] or recurrent layers [16], among others. The attention mechanism can also be integrated into Seq2Seq [20][22] to improve performance. There has been much research using Seq2Seq with attention to increase prediction accuracy, such as the work in [23-24]. This model can pay more attention to the relevant states within the encoder's hidden states.

### III. EXPERIMENTAL STUDY

This section presents an experimental study on day ahead electricity demand prediction, aiming to test the proposed forecast models. Section A describes the data, while section B reports the results and related discussion.

#### A. Data description

The main data utilized in this study comes from two sources: 1) the ENTSO-E transparency platform [25] for electricity demand of the United Kingdom (EDUK) region; 2) Google Trends data is provided by Google, and it shows the popularity of search terms over time [26]. The electricity demand represents average active power measured in MW. The data were collected from January 1, 2021, to April 15, 2024, with a time interval of 30 minutes. For the Google Trends data, the query 'Efficient in Energy Use' was selected for this study.

After data preparation, the dataset are transformed into a more usable and structured format. This require a data preprocessing step. In this study, we use Scikit-learn's StandardScaler to standardize the data. Subsequently, the dataset is split into an 80% training set and a 20% testing set.

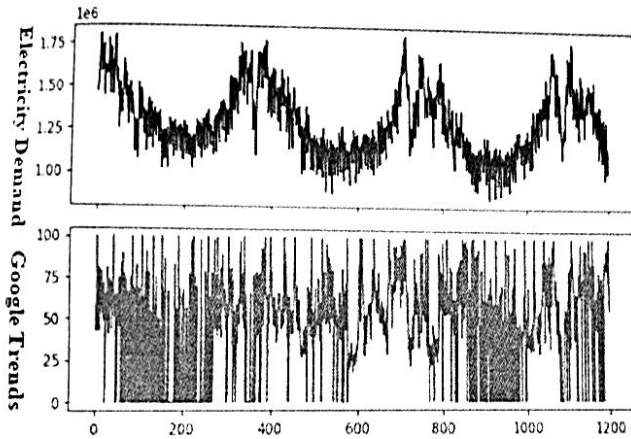


Fig. 3. (Above) National electricity demand of UK region and (Below) Google Trends data, respectively.

#### 3. Result analysis

##### 1) Causality test

In order to perform the Granger causality test, we first test both time series (Google Trends for "efficient energy use" and electricity demand) for stationarity using the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The results, shown in Table 1, indicate that both series are stationary. Second, the Granger test was conducted, and the results are reported in Table 2. The results indicate that Google Trends is a helpful predictor for recasting electricity demand.

TABLE I. RESULT OF THE COINTEGRA TEST

| Time series                            | At the original level |           |
|----------------------------------------|-----------------------|-----------|
|                                        | ADF test              | KPSS test |
| Electricity demand                     | 0.0126                | 0.0289    |
| Google Trend of "efficient energy use" | 0.0164                | 0.0002    |

Both the ADF and KPSS tests suggest that the time series for electricity demand and Google Trends for "efficient energy use" are stationary. This is indicated by the low p-values from the ADF test, which reject the null hypothesis of a unit root, and the low-test statistics from the KPSS test, which fail to reject the null hypothesis of stationarity (since the significant level chosen at 5% for both tests).

TABLE II. RESULT OF THE GRANGER CASUALITY TEST

| H0: Google Trend of "efficient energy use" does Granger cause electricity demand |        |
|----------------------------------------------------------------------------------|--------|
| F-stat (lag=8)                                                                   | 7.2452 |
| p-value (lag=8)                                                                  | 0      |

As shown in Table II, the F-statistic of 7.2452 is relatively high, and the extremely low p-value indicates that including past values of Google Trends provides a significant amount of explanatory power for predicting electricity demand.

##### 2) Model configuration

To achieve an optimal configuration of these models, we apply grid search techniques. The details of optimal parameters are described as following:

Models without trends:

LSTM: 1 LSTM layer with 256 neurons and 1 full-connected layers with 256 neurons

CNN-LSTM: 16 filters with ReLU activation function, 1 Maxpooling layer, and 01 fully connected layer with 32 neurons.

Seq2Seq: 2 GRU layers with 16 neurons and 1 fully connected layer with 16 neurons

The batch-size value is 32 and the learning rate is 1e-04

Model with Google Trends:

LSTM: 2 LSTM layers with 64 neurons and 1 full-connected layer with 128 neurons.

CNN-LSTM: 32 filters with ReLU activation function, 01 Maxpooling layer, and 1 fully connected layer with 128 neurons.

Seq2Seq-Attn: 2 GRU layers with 32 neurons and 01 fully connected layer with 32 neurons

##### 3) Model prediction result

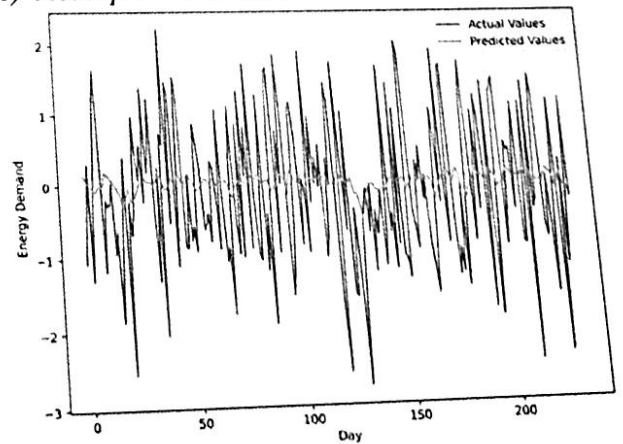


Fig. 4. LSTM models without Google Trends on testing set

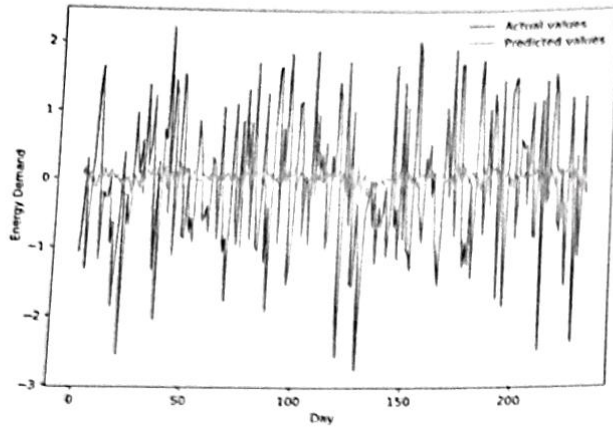


Fig. 5. CNN-LSTM model without Google Trends on testing set

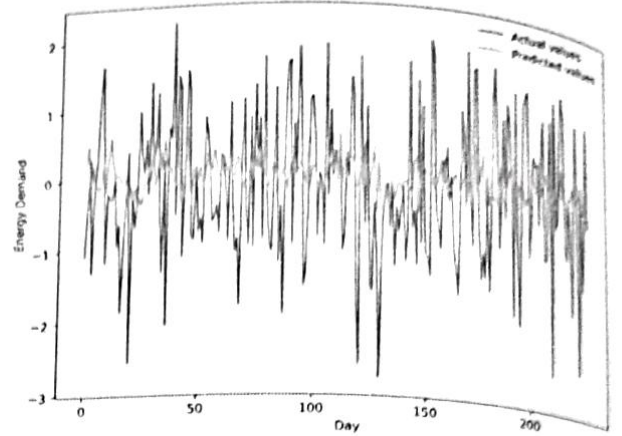


Fig. 8. CNN-LSTM with Google Trends on testing set

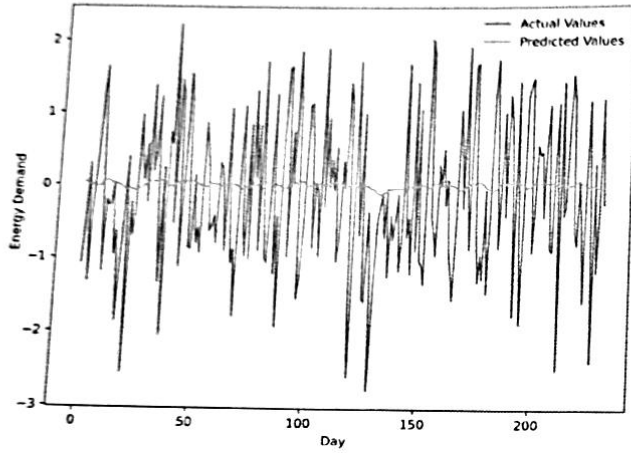


Fig. 6. Seq2Seq-Attention without Google Trends on testing set

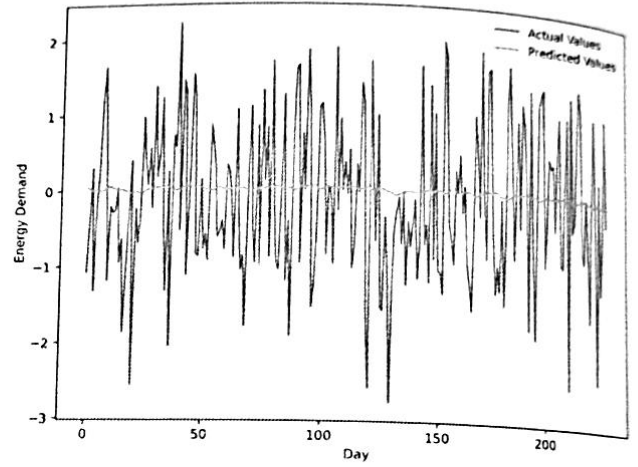


Fig. 9. Seq2Seq-Attn with Google Trends on testing set

Figures 4 to 6 describe the predictions of all models on the testing set without Google Trends. The LSTM-CNN model catches up quite well with the trend of the data. The results of the Seq2Seq-Attention show that the prediction line is very smooth.

Figures 7 to 9 show the predictions of all models on the testing datasets with Google Trends. Similarly to the previous experiments, the LSTM-CNN model has a better prediction compared to the others. The Seq2Seq-Attention predictions seem to be quite poor because it is very smooth.

Table below shows the performance of all models on the testing set regarding the Mean Squared Errors (MSE) metric. The CNN-LSTM has the lowest value with using Google Trends datasets. All of the model performances are better, proving that the Google Trends data is meaningful and helps to improve the performance of the model.

TABLE III. COMPARISON OF THE RESULTS OF DIFFERENT FORECASTING MODEL IN TERM MSE

| Model                 | LSTM  | CNN-LSTM | Seq2seq-Attn |
|-----------------------|-------|----------|--------------|
| without Google Trends | 1.162 | 1.157    | 1.167        |
| with Google Trends    | 1.095 | 1.093    | 1.155        |

#### IV. CONCLUSION

Our objective in this study is to implement various deep learning architectures to accurately gauge electricity demand in our specific case study, the EDUK. Throughout the tests, we found that the hybrid CNN-LSTM architecture achieved better results compared to the stand-alone LSTM and Seq2Seq-Attention architectures. Our models, when

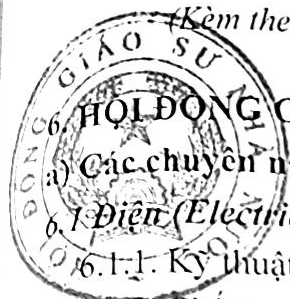
augmented with Google Trends data, showed improved performance across all tested architectures. The CNN-LSTM model, in particular, achieved the lowest Mean Squared Error (MSE) both with and without the inclusion of Google Trends data, highlighting its robustness and effectiveness in capturing complex patterns in electricity demand.

#### ACKNOWLEDGMENT

This research is funded by The University of Danang – University of Technology and Education under project number T2023-06-23.

#### REFERENCES

- [1] Mavragani A, Ochoa G and Tsagarakis KP (2018) Assessing the methods, tools, and statistical approaches in Google Trends research: Systematic review. *Journal of Medical Internet Research* 20(11): e270.
- [2] Park, Sungjun & Kim, Jinsoo. (2018). The Effect of Interest in Renewable Energy on US Household Electricity Consumption: An Analysis Using Google Trends Data. *Renewable Energy*. 127. 10.1016/j.renene.2018.05.044.
- [3] I Yu, Lean & Zhao, Yaqing & Tang, Ling & Yang, Zebin, 2019. "Online big data-driven oil consumption forecasting with Google trends," *International Journal of Forecasting*, Elsevier, vol. 35(1), pages 213-223.
- [4] Jian-Feng Guo, Qiang Ji, "How does market concern derived from the Internet affect oil prices?", *Applied Energy*, Volume 112, 2013, Pages 1536-1543.
- [5] Chun Fu, Clayton Miller, "Using Google Trends as a proxy for occupant behavior to predict building energy consumption", *Applied Energy*, Volume 310, 2022, 118343, ISSN 0306-2619.
- [6] Kadir Amasyali, Nora M. El-Gohary, "A review of data-driven building energy consumption prediction studies", *Renewable and Sustainable Energy Reviews*, Volume 81, Part 1, 2018, Pages 1192-1205.
- [7] Z. Deng, B. Wang, Y. Xu, T. Xu, C. Liu, and Z. Zhu, "Multi-Scale Convolutional Neural Network With Time-Cognition for Multi-Step Short-Term Load Forecasting," *IEEE Access*, vol. 7, pp. 88058–88071, 2019, doi: 10.1109/ACCESS.2019.2926137
- [8] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," *Energies*, vol. 10, no. 1, p. 3, 2016.
- [9] Almalaq, A., & Edwards, G., "A Review of Deep Learning Methods Applied on Load Forecasting," 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). doi:10.1109/icmla.2017.0-110
- [10] M. R. Islam, A. Al Mamun, M. Sohel, M. L. Hossain, and M. M. Uddin, "Lstm-based electrical load forecasting for chattogram city of bangladesh," in 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), 2020, pp. 188–192
- [11] J. F. Torres, F. Mart'inez-Alvarez, and A. Troncoso, "A deep lstm network for the spanish electricity consumption forecasting," *Neural Computing and Applications*, vol. 34, no. 13, pp. 10 533–10 545, Jul 2022. [Online]. Available: <https://doi.org/10.1007/s00521-021-06773-2>
- [12] S. Kumar, L. Hussain, S. Banarjee, and M. Reza, "Energy load forecasting using deep learning approach-lstm and gru in spark cluster," in 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT), 2018, pp. 1–4.
- [13] X. Shao, C.-S. Kim, and P. Sontakke, "Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion cnn-lstm," *Energies*, vol. 13, no. 8, 2020.
- [14] H. Kuang, Q. Guo, S. Li, and H. Zhong, "Short-term power load forecasting method in rural areas based on cnn lstm," in 2021 IEEE 4th International Electrical and Energy Conference (CIEEC), 2021, pp. 1–5.
- [15] H. Eskandari, M. Imani, and M. P. Moghaddam, "Convolutional and recurrent neural network based model for short-term load forecasting," *Electric Power Systems Research*, vol. 195, p. 107173, 2021.
- [16] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27
- [17] Tang, Y et al, "Question detection from acoustic features using recurrent neural network with gated recurrent unit" In *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, 20–25 March 2016; pp. 6125–6129.
- [18] L. Wu, C. Kong, X. Hao, and W. Chen, "A short-term load forecasting method based on gru-cnn hybrid neural network model," *Mathematical Problems in Engineering*, vol. 2020, p. 1428104, Mar 2020.
- [19] M. Ding, Y. Chen, and S. L. Bressler, "Granger causality: Basic theory and application to neuroscience," *Handbook of time series analysis*, vol. 1, pp. 437–460, 2006.
- [20] Gehring, J., Auli, M., Grangier, D., Yarats, D., & Dauphin, Y. N. (2017, July). Convolutional sequence to sequence learning. In *International conference on machine learning* (pp. 1243-1252). PMLR
- [21] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- [22] Shih, S. Y., Sun, F. K., & Lee, H. Y. (2019). Temporal pattern attention for multivariate time series forecasting. *Machine Learning*, 108, 1421-1441
- [23] Abbasimehr, H., & Paki, R. (2022). Improving time series forecasting using LSTM and attention models. *Journal of Ambient Intelligence and Humanized Computing*, 13(1), 673-691.
- [24] Hu, Y., & Xiao, F. (2022). Network self attention for forecasting time series. *Applied Soft Computing*, 124, 109092
- [25] ENTSO-E. Transparency Platform: <https://transparency.entsoe.eu>
- [26] <https://trends.google.com/trends>



(Kèm theo Quyết định số: 25/QĐ-HĐGSNN ngày 05 tháng 7 năm 2024)

## **HỘI ĐỒNG GIÁO SƯ LIÊN NGÀNH ĐIỆN – ĐIỆN TỬ – TỰ ĐỘNG HÓA**

### **a) Các chuyên ngành (Specialization):**

#### **6.1 Điện (Electricity)**

6.1.1. Kỹ thuật điện (Electric Engineering)

6.1.2. Chế tạo máy điện và thiết bị điện (Manufacturing of electro-machine and electric devices)

6.1.3. Hệ thống điện (Electric System)

6.1.4. Điện khí hóa và cung cấp điện (Electrification and electricity supply)

6.1.5 Công nghệ và kỹ thuật nhiệt (Thermal Technique and Technology)

#### **6.2. Điện tử (Electronics)**

6.2.1 Kỹ thuật điện tử (electronic engineering/electronic technique and technology)

6.2.2. Vô tuyến điện và truyền thông (Radio – electronics and telecommunication)

6.2.3. Chế tạo thiết bị và cấu kiện điện tử (Manufacturing of equipment and electronic components)

6.2.4. Kỹ thuật máy tính và viễn tin (Computer engineering and Tele-information)

6.2.5. Hệ thống điện tử và truyền thông (Electronic and communication systems)

#### **6.3. Tự động hóa (Automation)**

6.3.1. Điều khiển học kỹ thuật (Engineering automatics/Control engineering)

6.3.2. Điều khiển tự động (Automatic control)

6.3.3. Cơ điện tử (Mechatronics)

6.3.4. Đo lường và cảm biến (Measuring and sensors)

6.3.5. Kỹ thuật robot (Robotics)

### **b) Danh mục tạp chí được tính điểm:**

| Tên tạp chí/Báo cáo khoa học                                                                                        | Chỉ số ISSN | Loại    | Cơ quan xuất bản                                                                                                   | Điểm công trình                         |
|---------------------------------------------------------------------------------------------------------------------|-------------|---------|--------------------------------------------------------------------------------------------------------------------|-----------------------------------------|
| Bài báo khoa học được công bố trên tạp chí khoa học quốc tế thuộc danh mục Web of Science/Scopus/Scimago Q4 trở lên |             | Tạp chí | WoS/Scopus/Scimago (Q4 trở lên)<br>IF $\geq 3.0$ hoặc Citation $\geq 300$                                          | 0 – 2,0<br>Cộng tối đa 1,0 điểm         |
| Bài báo khoa học được công bố trên các tạp chí khoa học quốc tế khác có mã số ISSN                                  |             | Tạp chí | Do HĐGS liên ngành xác định cụ thể từng trường hợp                                                                 | 0 – 1,0 Online<br>0 – 0,75 Không online |
| Báo cáo khoa học tại hội nghị khoa học quốc tế                                                                      |             | Kỷ yếu  | Đăng toàn văn trong kỷ yếu, có phản biện khoa học, có mã số ISBN và thuộc Web of Science/Scopus/Scimago Q4 trở lên | 0 – 1,25                                |
| Báo cáo khoa học tại hội nghị khoa học quốc tế viết bằng tiếng Anh                                                  |             | Kỷ yếu  | Đăng toàn văn trong kỷ yếu, có phản biện khoa học, có mã số ISBN                                                   | 0 – 1,0                                 |

(Kèm theo Quyết định số: 25/QĐ-HDGSNN ngày 05 tháng 7 năm 2024)

|     |                                                                                                                                                   |                                                                |                            |                                                                  |                                                                  |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|----------------------------|------------------------------------------------------------------|------------------------------------------------------------------|
| 5.  | Báo cáo khoa học tại các hội nghị khoa học quốc gia                                                                                               |                                                                | Kỷ yếu                     | Đăng toàn văn trong kỷ yếu, có phản biện khoa học, có mã số ISBN | 0 – 0,5                                                          |
| 6.  | Các công trình nghiên cứu, phát triển và ứng dụng và công nghệ thông tin và truyền thông (Journal on Information technologies and Communications) | 1859-3526-V<br>1859-3534-E<br>(0866-7039)                      | Tạp chí                    | Bộ Thông tin và Truyền thông                                     | 0 – 0,75 từ quý II/2013<br>0 – 1,0 từ 2024                       |
| 7.  | Chuyên san Đo lường, Điều khiển và Tự động hóa (Measurement Control and Automation)                                                               | 1859- 0551                                                     | Tạp chí                    | Hội Tự động hóa Việt Nam                                         | 0 – 0,75 từ 2014<br>0 – 1,0 từ 2024                              |
| 8.  | Khoa học công nghệ (Journal of Science & Technology)                                                                                              | 1859-3585<br>e-2615-9619                                       | Tạp chí                    | Trường Đại học Công nghiệp Hà Nội                                | 0 – 0,5 từ quý III/2015<br>0 – 0,75 từ 2024                      |
| 9.  | Khoa học Công nghệ năng lượng                                                                                                                     | 1859-4557                                                      | Tạp chí                    | Trường Đại học Điện lực                                          | 0 – 0,5 từ quý III/2015                                          |
| 10. | Khoa học Công nghệ Thông tin và Truyền thông                                                                                                      | 2525-2224                                                      | Tạp chí                    | Học viện Công nghệ Bưu chính Viễn thông                          | 0 – 0,5 từ quý III/2017<br>0 – 0,75 từ 2024                      |
| 11. | Khoa học công nghệ: Khoa học tự nhiên – Kỹ thuật - Công nghệ                                                                                      | 2734-9098<br>e-2615-9562<br>(1859-2171)                        | Tạp chí ACI                | Đại học Thái Nguyên                                              | 0 – 0,5 từ quý II/2013<br>0 – 0,75 từ 2021                       |
| 12. | Khoa học Đại học Sài Gòn (Scientific Journal of Saigon University)                                                                                | 1859-3208                                                      | Tạp chí                    | Trường Đại học Sài Gòn                                           | 0 – 0,25 từ quý III/2015<br>0 – 0,5 từ Quý III/2017<br>0 từ 2020 |
| 13. | Khoa học ĐHQGHN: Khoa học tự nhiên và Công nghệ (VNU Journal of Science: Natural Science and Technology)                                          | p-2615-9317<br>e-2588-1140<br>0866-8612                        | Tạp chí                    | Đại học Quốc gia Hà Nội                                          | 0 – 0,75<br>0 – 1,0 từ 2024                                      |
| 14. | Khoa học Giáo dục Kỹ thuật                                                                                                                        | 2615-9740<br>1859-1272                                         | Tạp chí                    | Trường ĐH Sư phạm Kỹ thuật TP Hồ Chí Minh                        | 0 – 0,25<br>0 – 0,5 từ quý III/2017                              |
| 15. | Khoa học kỹ thuật (Science & Technology)                                                                                                          | 1859-0209                                                      | Tạp chí                    | Học viện Kỹ thuật Quân sự (Le Quy Don Technical University)      | 0 – 0,75<br>0 – 1,0 từ 2024                                      |
| 16. | - Khoa học Đại học Cần Thơ (Can Tho University Journal of Science)<br>- CTU Journal of Innovation and Sustainable Development                     | 1859-2333-V<br>e-2815-5599-V<br>p-2615-9422-E<br>e-2815-5602-E | Tạp chí ACI, Bản tiếng Anh | Trường Đại học Cần Thơ                                           | 0 – 0,25 từ 2020<br>0 – 0,5 từ 2022<br>0 – 0,75 từ 2023          |

(Kèm theo Quyết định số: 25/QĐ-HDGSNN ngày 05 tháng 7 năm 2024)

|                                                                                                                                              |                                                  |                          |                                                                                                                                |                                                         |
|----------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| Khoa học và Công nghệ                                                                                                                        | 2525-2267                                        | Tạp chí                  | Trường Đại học Công nghiệp TP Hồ Chí Minh                                                                                      | 0 – 0,25 từ 2021<br>0 – 0,5 từ 2024                     |
| Khoa học và Công nghệ ( <i>Journal of Science &amp; Technology</i> )                                                                         | 2525-2518<br>e-2815-5874<br>(0866-708X)          | Tạp chí<br>ACI<br>Scopus | Viện Hàn lâm Khoa học và Công nghệ Việt Nam                                                                                    | 0 – 1,0<br>0 – 1,5 từ 2023<br>(bản 2525-2518)           |
| Khoa học và Công nghệ ( <i>Science &amp; Technology</i> )                                                                                    | 1859-1531                                        | Tạp chí                  | Đại học Đà Nẵng                                                                                                                | 0 – 0,5<br>0 – 0,75 từ quý III/ 2017                    |
| Khoa học và Công nghệ các trường Đại học kỹ thuật<br>- Engineering and Technology for Sustainable Development<br>- Smart Systems and Devices | 0868-3980<br>2354-1083<br>2734-9381<br>2734-9373 | Tạp chí                  | ĐHBKHN; ĐH Đà Nẵng; Trường ĐH KTCN, ĐH Thái Nguyên; Trường ĐHBK, ĐHQG TP. HCM; Trường ĐH SPKT TP. HCM; Học viện Công nghệ BCVT | 0 – 0,75<br>0 – 1,0 từ 2024                             |
| Năng lượng nhiệt (Cũ: Khoa học và công nghệ nhiệt)                                                                                           | 0868-3336                                        | Tạp chí                  | Hội Khoa học Công nghệ nhiệt                                                                                                   | 0 – 0,5 đến Quý III/2020<br>0 từ 2021                   |
| Nghiên cứu Khoa học và Công nghệ quân sự                                                                                                     | 1859-1043                                        | Tạp chí                  | Viện Khoa học và Công nghệ Quân sự                                                                                             | 0 – 0,75                                                |
| Khoa học và Công nghệ trong lĩnh vực an toàn thông tin ( <i>Journal of Science and Technology on Information security</i> )                  | 2615-9570                                        | Tạp chí                  | Ban Cơ yếu Chính phủ                                                                                                           | 0 – 0,25 từ 2020<br>0 – 0,5 từ 2022<br>0 – 0,75 từ 2024 |
| Phát triển Khoa học và công nghệ ( <i>Science &amp; Technology Development</i> )                                                             | 1859-0128<br>2588-106X                           | Tạp chí                  | Đại học Quốc gia Thành phố Hồ Chí Minh                                                                                         | 0 – 0,75<br>0 – 1,0 từ 2024                             |
| REV Journal on Electronics and Telecommunications                                                                                            | 1859-378X                                        | Tạp chí                  | Hội Vô tuyến điện tử Việt Nam                                                                                                  | 0 – 1,0 từ quý II/2013                                  |
| Tin học và Điều khiển học, <i>Journal of Computer Science and Cybernetics</i>                                                                | 1813-9663                                        | Tạp chí                  | Viện Hàn lâm Khoa học và Công nghệ Việt Nam                                                                                    | 0 – 0,5 từ quý III/2017<br>0 – 75 từ 2021               |
| Khoa học ( <i>Journal of Science</i> )                                                                                                       | 1859-0357                                        | Tạp chí                  | Trường Đại học Quy Nhơn                                                                                                        | 0 – 0,25 từ 202<br>0 – 0,5 từ 202                       |
| Nghiên cứu khoa học                                                                                                                          | 1859-4190                                        | Tạp chí                  | Trường Đại học Sao Đỏ                                                                                                          | 0 – 0,25 từ 20                                          |
| Khoa học ĐHQGHN: Công nghệ Thông tin-Truyền thông (VNU Journal of Science: Computer Science and Communication Engineering)                   | 2615-9260<br>e-2588-1086                         | Tạp chí                  | Đại học Quốc gia Hà Nội                                                                                                        | 0 – 0,5 từ 20<br>0 – 0,75 từ 2                          |

(Kèm theo Quyết định số: 25/QĐ-HĐGSNN ngày 05 tháng 7 năm 2024)

|     |                                                                                                                                                                |                                                              |         |                                             |                                     |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|---------|---------------------------------------------|-------------------------------------|
| 30. | - Khoa học và Công nghệ Việt Nam bản B (Vietnam Journal of Science and Technology – MOST series B)<br>- Vietnam Journal of Science, Technology and Engineering | 1859-4794-V<br>e-2615-9929-V<br>2525-2461-E<br>e-2615-9937-E | Tạp chí | Bộ Khoa học và Công nghệ                    | 0 – 0,5 từ 2023<br>0 – 0,75 từ 2024 |
| 31. | Khoa học Đại học Công thương (Tên cũ Khoa học và Công nghệ thực phẩm)                                                                                          | 3030-4113<br>e-3030-413X<br>0866-8132                        | Tạp chí | Trường Đại học Công thương TP Hồ Chí Minh   | 0 – 0,25 từ 2023<br>0 – 0,5 từ 2024 |
| 32. | Khoa học và Công nghệ Giao thông (Journal of Science and Transport Technology)                                                                                 | 2734-9942-V<br>2734-9950                                     | Tạp chí | Trường Đại học Công nghệ Giao thông vận tải | 0 – 0,25 từ 2024                    |
| 33. | Khoa học và Công nghệ Trường ĐH Thành Đông (Thanh Dong University Journal of Science and Technology)                                                           | 2734-9500                                                    | Tạp chí | Trường Đại học Thành Đông                   | 0 – 0,25 từ 2024                    |

#### c) Tác giả chính, sách phục vụ đào tạo

- Tác giả chính của bài báo được xác định là tác giả thứ nhất (first author), tác giả chịu trách nhiệm (corresponding author) duy nhất.

- Sách phục vụ đào tạo phải được xuất bản có chỉ số ISBN, được hội đồng chuyên môn thẩm định đạt và được cơ sở đào tạo xác nhận mục đích sử dụng.