

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN THANH NHÂN

**ỨNG DỤNG CHUỖI THỜI GIAN TRONG DỰ BÁO NHU CẦU
PHỤ TẢI ĐIỆN Ở CÔNG TY ĐIỆN LỰC TÂY NINH**

LUẬN VĂN THẠC SỸ KỸ THUẬT

(Theo định hướng ứng dụng)

TP. HỒ CHÍ MINH – NĂM 2022

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN THANH NHÂN

**ỨNG DỤNG CHUỖI THỜI GIAN TRONG DỰ BÁO NHU CẦU
PHỤ TẢI ĐIỆN Ở CÔNG TY ĐIỆN LỰC TÂY NINH**

CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN

MÃ SỐ: 8.48.01.04

LUẬN VĂN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC:

TS. HUỖNH TRỌNG THƯA

TP. HỒ CHÍ MINH – NĂM 2022

LỜI CAM ĐOAN

Tôi xin cam đoan luận văn thạc sĩ hệ thống thông tin “**Ứng dụng chuỗi thời gian trong dự báo nhu cầu phụ tải điện ở Công ty Điện lực Tây Ninh**” do tôi nghiên cứu, tổng hợp và thực hiện.

Toàn bộ nội dung luận văn, những điều được trình bày là của chính cá nhân tôi hoặc là được tham khảo, tổng hợp từ nhiều nguồn tài liệu khác nhau. Tất cả các tài liệu tham khảo, tổng hợp đều được trích xuất với nguồn gốc rõ ràng. Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

TP. HCM, ngày 25 tháng 01 năm 2022

Học viên thực hiện luận văn

Nguyễn Thanh Nhân

LỜI CẢM ƠN

Trước tiên tôi xin gửi lời cảm ơn sâu sắc nhất đến giảng viên hướng dẫn trực tiếp của tôi - **TS. Huỳnh Trọng Thưa**. Cảm ơn Thầy đã luôn lắng nghe những quan điểm cá nhân và đưa ra những nhận xét quý báu, góp ý và dẫn dắt tôi đi đúng hướng trong suốt thời gian thực hiện đề tài luận văn thạc sĩ...

Tôi cũng xin trân trọng cảm ơn các Thầy Cô giáo Khoa Đào tạo Sau đại học - Học viện Công nghệ Bưu chính viễn thông Cơ sở Thành phố Hồ Chí Minh đã truyền đạt những kiến thức chuyên môn sâu rộng trong ngành cho tôi, có được nền tảng kiến thức hỗ trợ rất lớn cho tôi trong quá trình nghiên cứu, hoàn thiện luận văn.

Cuối cùng, tôi xin gửi lời cảm ơn chân thành đến gia đình, bạn bè và quý đồng nghiệp vì đã luôn đồng hành, tạo điều kiện, hỗ trợ và khuyến khích tôi cố gắng trong suốt những năm tháng học tập cũng như quá trình nghiên cứu viết luận văn này.

Xin chân thành cảm ơn!

TP. HCM, ngày 25 tháng 01 năm 2022

Học viên thực hiện luận văn

Nguyễn Thanh Nhân

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT	v
DANH SÁCH BẢNG	vi
DANH SÁCH HÌNH VẼ.....	vii
MỞ ĐẦU	1
1. Lý do chọn đề tài.....	1
2. Tổng quan vấn đề nghiên cứu.....	2
3. Mục tiêu nghiên cứu.....	4
4. Đối tượng và phạm vi nghiên cứu	5
5. Phương pháp nghiên cứu	5
6. Cấu trúc luận văn	6
CHƯƠNG 1: TỔNG QUAN VỀ DỰ BÁO CHUỖI THỜI GIAN.....	8
1.1. Tổng quan về dự báo	8
1.2. Các phương pháp dự báo	9
1.2.1. Các mô hình thống kê.....	9
1.2.2. Các mô hình hiện đại dựa trên máy học, khai phá dữ liệu và trí tuệ nhân tạo	12
1.3. Thực trạng dự báo phụ tải điện ở Công ty Điện lực Tây Ninh	13
CHƯƠNG 2: MÔ HÌNH DỰ BÁO CHUỖI THỜI GIAN VÀ ỨNG DỤNG TRONG DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN.....	16
2.1. Phân tích chuỗi thời gian nhu cầu phụ tải điện.....	16
2.2. Mô hình ARIMA.....	21

2.3. Mô hình SARIMA.....	24
2.4. Mô hình mạng LSTM.....	25
2.5. Mô hình Prophet.....	28
CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN.....	29
3.1. Xây dựng mô hình dự báo	29
3.2. Xây dựng mô hình dự báo SARIMA	31
3.3. Xây dựng mô hình dự báo LSTM.....	33
3.4. Xây dựng mô hình dự báo Prophet	35
CHƯƠNG 4: KẾT QUẢ DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN Ở CÔNG TY ĐIỆN LỰC TÂY NINH.....	37
4.1. Kết quả dự báo	37
4.1.1. Dự báo nhu cầu phụ tải 1 tháng	37
4.1.2. Dự báo nhu cầu phụ tải 12 tháng	43
4.2. So sánh kết quả khi áp dụng các mô hình Prophet, SARIMA, LSTM.	44
4.3. Nhận xét chung	46
CHƯƠNG 5: XÂY DỰNG HỆ THỐNG DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN.....	47
5.1. Thiết kế hệ thống.....	47
5.1.1. Mô hình hệ thống.....	47
5.1.2. Thiết kế cơ sở dữ liệu cho hệ thống	48
5.1.3. Xây dựng Dashboard phân tích, dự báo	50
5.2. Kết quả.....	52
KẾT LUẬN.....	54
DANH MỤC TÀI LIỆU THAM KHẢO	55

DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Tiếng Việt
ANN	Artificial Neural Network	Mạng nơ-ron nhân tạo
LSTM	Long Short Term Memory	Mạng nơ-ron bộ nhớ dài-ngắn
ARIMA	Autoregressive Intergrated Moving Average	Mô hình trung bình trượt tích hợp tự hồi quy
SARIMA	Seasonal ARIMA	ARIMA mùa vụ
SVM	Support Vector Machine	Véc-tơ máy hỗ trợ
RMSE	Root Mean Square Error	Sai số trung bình bình phương gốc (Độ lệch chuẩn của phần dư)
MAPE	Mean Absolute Percentage Error	Sai số phần trăm tuyệt đối trung bình

DANH SÁCH BẢNG

Bảng 3.1: Các giá trị RMSE nhỏ nhất từng mô hình thực hiện tìm kiếm lưới trên dữ liệu 09 Điện lực.....	33
Bảng 4.1: Giá trị RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai dự báo theo mô hình Prophet	37
Bảng 4.2: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 điện lực	38
Bảng 4.3: Độ đo RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai mô hình SARIMA.....	39
Bảng 4.4: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 mô hình tương ứng 09 Điện lực	40
Bảng 4.5: Độ đo RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai mô hình mạng LSTM.....	41
Bảng 4.6: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 mô hình tương ứng 09 Điện lực	42
Bảng 4.7: Các giá trị RMSE của dữ liệu 12 tháng dùng đánh giá và kết quả dự báo của mô hình Prophet tương ứng của 09 Điện lực	43
Bảng 4.8: Các giá trị RMSE của dữ liệu 12 tháng dùng đánh giá và kết quả dự báo của mô hình SARIMA tương ứng của 09 Điện lực	44
Bảng 4.9: So sánh hiệu suất dự báo từng tháng của 3 mô hình SARIMA, Prophet và LSTM.....	44
Bảng 4.10: So sánh hiệu suất dự báo 12 tháng của 2 mô hình SARIMA và Prophet.....	46

DANH SÁCH HÌNH VẼ

Hình 2.1a: Biểu đồ sản lượng điện thương phẩm hàng tháng của một doanh nghiệp có mức tiêu thụ hàng năm trên 1 triệu kWh	17
Hình 2.1b: Biểu đồ sản lượng điện thương phẩm hàng tháng các khách hàng có mức tiêu thụ dưới 1 triệu kWh/năm	18
Hình 2.2: Biểu đồ phân rã các đặc trưng (thành phần) của dữ liệu chuỗi thời gian sản lượng điện thương phẩm hàng tháng chu kỳ 12 tháng của các khách hàng có mức tiêu thụ dưới 1 triệu kWh/năm.....	20
Hình 2.3a: Biểu đồ chu kỳ 12 tháng của chuỗi thời gian sản lượng điện thương phẩm	20
Hình 2.3b: Biểu đồ sai phân bậc 1 của chuỗi thời gian sản lượng điện thương phẩm	21
Hình 2.4: Một mạng LSTM và cấu trúc bên trong của tế bào LSTM.....	26
Hình 2.5: Cấu trúc của một tế bào LSTM	27
Hình 3.1: Minh họa phân chia dữ liệu chuỗi thời gian sản lượng điện thương phẩm trong lựa chọn mô hình và đánh giá các mô hình	29
Hình 3.2: Mã giả thuật toán đánh giá mô hình theo chiến lược kiểm tra gổ đầu được áp dụng trong lựa chọn mô hình và đánh giá hiệu suất các mô hình.....	30
Hình 3.3: Mã giả thuật toán tìm kiếm lưới khám phá bộ tham số SARIMA cho một chuỗi thời gian đơn biến	32
Hình 3.4: Đoạn mã Python xây dựng mô hình LSTM.....	34
Hình 3.5a: Biểu đồ giá trị RMSE các mô hình LSTM theo số neurons, epochs	34
Hình 3.5b: Biểu đồ giá trị RMSE các mô hình LSTM theo số neurons, epochs.....	35
Hình 3.6: Đoạn mã Python mô hình dự báo Prophet	35
Hình 4.1: Biểu đồ so sánh kết quả dự báo từng tháng của 3 mô hình SARIMA, Prophet và mạng LSTM của 9 điện lực.....	45
Hình 5.1: Mô hình kiến trúc tổng thể.....	47

Hình 5.2: Lược đồ ERD cơ sở dữ liệu dự báo	48
Hình 5.4: Dữ liệu mẫu bảng khách hàng.....	49
Hình 5.5: Dữ liệu mẫu bảng dữ liệu phụ tải điện hàng tháng.....	49
Hình 5.6: Dữ liệu mẫu bảng tham số các mô hình.....	50
Hình 5.8: Mã nguồn Dashboard phân tích, dự báo	51
Hình 5.9: Kết quả dự báo 12 tháng phụ tải điện khách hàng thuộc Điện lực Thành phố Tây Ninh từ tháng 2/2021	52
Hình 5.10: Kết quả dự báo 12 tháng phụ tải điện khách hàng thuộc Điện lực Gò Dầu từ tháng 2/2021.....	53

MỞ ĐẦU

1. Lý do chọn đề tài

Dự báo nhu cầu phụ tải điện là hoạt động tính toán, ước lượng và đánh giá xu hướng tăng trưởng, tốc độ tăng trưởng và nhu cầu của phụ tải điện được cung cấp điện từ hệ thống điện trong ngắn hạn, trung hạn và dài hạn [1]. Theo miền thời gian tương lai, dự báo nhu cầu phụ tải điện có thể được phân thành các loại dự báo sau [2], [3]:

- Dự báo ngắn hạn (dưới 1 ngày/tuần) đóng một vai trò quan trọng trong công tác lập phương thức vận hành tại các công ty điện lực.

- Dự báo trung hạn (1 ngày/tuần đến 1 năm) cần thiết cho việc quy hoạch, vận hành nguồn, lưới cũng như đánh giá hiệu quả hoạt động.

- Dự báo dài hạn (trên 1 năm) là một phần quan trọng trong quy hoạch hệ thống điện, quy định biểu giá và phân phối điện năng.

Dự báo nhu cầu phụ tải điện là một hoạt động có vai trò quan trọng trong việc quy hoạch, vận hành hệ thống điện. Công ty Điện lực Tây Ninh là một đơn vị phân phối điện bên cạnh nhu cầu dự báo ngắn hạn phục vụ lập phương thức vận hành (điều độ hệ thống điện) thì nhu cầu dự báo phụ tải điện trung hạn nhằm nâng cao hiệu quả hoạt động, độ tin cậy và đảm bảo cân đối cung cầu phụ tải điện cũng hết sức cần thiết.

Thực hiện chỉ đạo của Tập đoàn Điện lực Việt Nam và triển khai thực hiện Thông tư 19 [1] của Bộ công thương, các Công ty Điện lực phải thực hiện dự báo phụ tải của các nhóm khách hàng, thành phần phụ tải trong hoạt động điều hành, vận hành hệ thống điện hằng tháng. Tuy nhiên việc triển khai thực hiện gặp nhiều khó khăn do chưa có phương pháp dự báo phụ tải phù hợp với nguồn dữ liệu hiện hữu. Hiện nay, chủ yếu các Công ty Điện lực dự báo nhu cầu phụ tải dựa theo phương pháp chuyên

gia hoặc mô hình hàm hồi quy đa biến phụ thuộc vào các yếu tố như thời tiết, nhiệt độ, chỉ số kinh tế, ... cho nên kết quả dự báo còn hạn chế.

Nhằm giải quyết các vấn đề trên, luận văn đề xuất phương pháp nghiên cứu dự báo chuỗi thời gian để dự báo nhu cầu phụ tải điện trung hạn (xu thế, tốc độ phát triển và nhu cầu của phụ tải điện). Từ đó thiết kế xây dựng một hệ thống dự báo nhu cầu phụ tải điện để giải quyết các khó khăn nêu trên cũng như để đáp ứng yêu cầu báo cáo thống kê và hỗ trợ ra quyết định là một nhu cầu cấp thiết đối với Công ty Điện lực Tây Ninh. Điều này giúp Lãnh đạo của Công ty Điện lực Tây Ninh nắm bắt kịp thời tình hình phát triển phụ tải điện và nhanh chóng đưa ra những điều chỉnh, giải pháp kịp thời cũng như các định hướng thích hợp về mặt điều hành hệ thống điện nhằm đạt được mục tiêu nâng cao hiệu quả, độ tin cậy cung cấp điện và đảm bảo cân đối cung cầu phụ tải điện phục vụ nhu cầu sử dụng điện của khách hàng. Từ những lý do trên, tôi xin lựa chọn đề tài nghiên cứu **“Ứng dụng chuỗi thời gian trong dự báo nhu cầu phụ tải điện ở Công ty Điện lực Tây Ninh”**.

2. Tổng quan vấn đề nghiên cứu

Dự báo phụ tải chính xác là điều cần thiết trong hoạt động điều hành, vận hành hệ thống điện. Nếu dự báo sai lệch quá nhiều về khả năng cung cấp và nhu cầu phụ tải điện sẽ dẫn đến giảm hiệu quả hoạt động. Nếu dự báo quá thừa về nhu cầu phụ tải sẽ phải huy động nguồn quá lớn làm tăng vốn đầu tư dẫn đến lãng phí vốn đầu tư và không khai thác hết công suất thiết bị, ngược lại nếu dự báo thiếu nhu cầu phụ tải sẽ dẫn đến cung cấp điện không đủ cho nhu cầu của phụ tải, giảm độ tin cậy cung cấp điện làm giảm hiệu quả hoạt động.

Dự báo chuỗi thời gian là một lớp mô hình quan trọng trong thống kê, kinh tế lượng và máy học. Mô hình là chuỗi thời gian (time series) là mô hình được áp dụng trên các chuỗi đặc thù có yếu tố thời gian. Một mô hình chuỗi thời gian thường dự báo dựa trên giả định rằng các qui luật trong quá khứ sẽ lặp lại ở tương lai. Do đó xây dựng mô hình chuỗi thời gian là đang mô hình hóa mối quan hệ trong quá khứ giữa biến độc lập (biến đầu vào) và biến phụ thuộc (biến mục tiêu). Dựa vào mối quan hệ

này để dự đoán giá trị trong tương lai của biến phụ thuộc. Hiện nay có rất nhiều nghiên cứu dự báo chuỗi thời gian và dự báo phụ tải điện, tuy nhiên có thể phân loại thành 2 nhóm là phương pháp cổ điển dựa trên thống kê và phương pháp hiện đại dựa trên công nghệ trí tuệ nhân tạo.

Nhiều nghiên cứu dự báo phụ tải điện dựa trên phương pháp thống kê có thể kể đến như các mô hình dự báo sử dụng hồi quy đơn biến, đa biến mà đại diện tiêu biểu là mô hình ARIMA (Autoregressive Intergrated Moving Average - Tự hồi quy tích hợp trung bình trượt). Lê Duy Phúc và các đồng sự [4] đã sử dụng phương pháp lọc và mô hình dự báo ARIMA trong dự báo phụ tải điện ngắn hạn lưới điện TP.HCM, kết quả dự báo với độ tin cậy 95%. Nazih Abu-Shikhah và các đồng sự [5] sử dụng phương pháp hồi quy đa biến và hồi quy phi tuyến trong dự báo phụ tải điện trung hạn ở NEPCO (National Electric Power Company of Jordan). Isaac A. Samuel và các đồng sự [6] sử dụng phân tích hồi quy trong dự báo phụ tải điện của trường Covenant University.

Trong những năm gần đây, phương pháp dự báo đã chuyển từ phương pháp cổ điển sang ứng dụng những thành tựu của trí tuệ nhân tạo vào trong dự báo. Các thuật toán học máy như máy vector hỗ trợ và mạng nơ ron nhân tạo được ứng dụng rộng rãi đã cải thiện đáng kể độ chính xác của dự đoán nhu cầu phụ tải điện. Các nghiên cứu [7], [8], [9] sử dụng mạng nơ-ron nhân tạo lan truyền ngược (BP ANN) kết hợp các yếu tố ảnh hưởng như chỉ số tiêu dùng, chỉ số công nghiệp, tốc độ tăng dân số và nhiệt độ để dự báo phụ tải điện trung hạn có độ chính xác và khả năng giải thích tốt hơn các phương pháp truyền thống khác. Lei Guo và cộng sự [10] đã sử dụng mạng nơron nhiều lớp và LSTM để dự đoán phụ tải điện trong thời gian ngắn và kết quả thực nghiệm đã chứng minh tính ưu việt của mạng nơron học sâu. Zhitong Ma và các đồng sự [11] đã sử dụng máy vector hỗ trợ để dự đoán mức tiêu thụ năng lượng của tòa nhà ở Trung Quốc dưới ảnh hưởng các yếu tố thời tiết.

Gần đây, các nghiên cứu dự báo chuỗi thời gian phi tuyến theo mô hình cộng (Additive Model) trong đó các xu thế phi tuyến khớp với tính thời vụ hàng năm, hàng

tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ cũng được sử dụng cho hiệu quả tốt. Sean J. Taylor và Benjamin Letham [12] đã giới thiệu phương pháp phân tích chuỗi thời gian theo mô hình cộng bằng thuật toán Facebook Prophet trong dự báo số sự kiện trên Facebook. Huỳnh Trọng Thừa và các đồng sự [13] đã công bố kết quả nghiên cứu sử dụng thuật toán phân tích chuỗi thời gian Facebook Prophet trong dự báo hoạt động đầu tư ở Khu Công nghệ cao TP.HCM. Zar Zar Oo và Sabai Phyu [14] sử dụng thuật toán Prophet để dự báo nhiệt độ ở tỉnh Myintkyina, Myanmar cho kết quả tốt so với các phương pháp truyền thống. Toni Toharudin và các đồng sự [15] đã công bố báo cáo so sánh kết quả dự báo nhiệt độ không khí hàng ngày trong 5 năm ở Bandung, Taiwan giữa 2 mô hình dự báo bởi LSTM và Facebook Prophet; Các kết quả cho thấy rằng, Prophet cho kết quả tốt hơn đối với nhiệt độ tối đa, trong khi LSTM cho kết quả tốt hơn đối với nhiệt độ tối thiểu; Tuy nhiên, sự khác biệt về giá trị của RMSE không quá lớn đáng kể. Emir Žunić và các đồng sự [16] đề xuất bộ khung ứng dụng thuật toán Prophet trong dự báo doanh số bán hàng trong thực tế để dự báo doanh số hằng tháng, hằng quý sắp tới.

3. Mục tiêu nghiên cứu

3.1. Mục đích nghiên cứu

Xây dựng hệ thống dự báo nhu cầu phụ tải điện trung hạn nhằm hỗ trợ ban Lãnh đạo Công ty Điện lực Tây Ninh trong hoạt động điều hành, vận hành hệ thống điện đảm bảo mục tiêu nâng cao hiệu quả, độ tin cậy cung cấp điện và đảm bảo cân đối cung cầu phụ tải điện phục vụ nhu cầu sử dụng điện của khách hàng.

3.2. Mục tiêu tổng quát

Đánh giá thực trạng nhu cầu dự báo phụ tải ở Công ty Điện lực Tây Ninh và nghiên cứu phương pháp dự báo chuỗi thời gian trong dự báo nhu cầu phụ tải điện trung hạn, từ đó đề xuất giải pháp dự báo nhu cầu phụ tải điện trung hạn ở Công ty Điện lực Tây Ninh.

3.3. Mục tiêu cụ thể

- Nghiên cứu các phương pháp dự báo chuỗi thời gian trong dự báo nhu cầu phụ tải điện trung hạn.

- Đánh giá thực trạng dự báo nhu cầu phụ tải điện trung hạn ở Công ty Điện lực Tây Ninh.

- Đề xuất giải pháp dự báo nhu cầu phụ tải điện trung hạn dựa trên phân tích chuỗi thời gian Facebook Prophet ở Công ty Điện lực Tây Ninh.

- Xây dựng hệ thống thực nghiệm và đánh giá hệ thống dự báo nhu cầu phụ tải điện trung hạn ở Công ty Điện lực Tây Ninh.

4. Đối tượng và phạm vi nghiên cứu

4.1. Đối tượng nghiên cứu:

- Thực trạng nhu cầu dự báo phụ tải điện trung hạn ở Công ty Điện lực Tây Ninh.

- Các phương pháp dự báo nhu cầu phụ tải điện trung hạn và phương pháp dự báo đề xuất là dự báo chuỗi thời gian theo mô hình cộng Facebook Prophet.

- Giải pháp hợp lý dự báo nhu cầu phụ tải điện trung hạn ở Công ty Điện lực Tây Ninh.

4.2. Phạm vi nghiên cứu:

Về nội dung: Nhu cầu dự báo phụ tải điện trung hạn và giải pháp đề xuất nhằm dự báo nhu cầu phụ tải trung hạn ở Công ty Điện lực Tây Ninh.

Về thời gian: Các tài liệu nghiên cứu được sử dụng trong luận văn là các tài liệu trong và ngoài nước trong khoảng thời gian từ năm 2010-2020. Dữ liệu nghiên cứu là sản lượng điện tiêu thụ của các nhóm khách hàng, thành phần phụ tải, địa bàn quản lý hằng tháng từ tháng 01/2006 đến tháng 03/2021 ở Công ty Điện lực Tây Ninh.

Về không gian: Nghiên cứu được thực hiện và ứng dụng ở Công ty Điện lực Tây Ninh.

5. Phương pháp nghiên cứu

Đề tài này sử dụng phương pháp nghiên cứu lý thuyết kết hợp với xây dựng ứng dụng thực nghiệm:

- Nghiên cứu lý thuyết dự báo phụ tải điện dài hạn, các thành phần phụ tải điện và các yếu tố ảnh hưởng đến dự báo.

- Thu thập dữ liệu nhu cầu phụ tải điện của các nhóm thành phần phụ tải điện, theo khu vực và nhóm khách hàng: Sản lượng điện tiêu thụ hàng tháng của các nhóm khách hàng trên địa bàn tỉnh Tây Ninh do Công ty Điện lực Tây Ninh quản lý. Đây chính là dữ liệu của các đối tượng nghiên cứu phục vụ cho quá trình phân tích và tìm hiểu xu thế, tốc độ phát triển và nhu cầu của đối tượng nghiên cứu.

- Xây dựng hệ thống thực nghiệm, đánh giá kết quả thực nghiệm, đưa ra hướng phát triển mở rộng của đề tài để đáp ứng những nhu cầu triển khai thực tế.

6. Cấu trúc luận văn

Ngoài phần mở đầu, mục lục, kết luận và kiến nghị, danh mục hình vẽ, danh mục bảng biểu, tài liệu tham khảo, phụ lục, phần chính của luận văn gồm 5 chương như sau:

Chương 1. Tổng quan về dự báo nhu cầu phụ tải điện: Chương này nêu khái quát, hệ thống hóa cơ sở lý luận về vấn đề dự báo, đồng thời tập trung trình bày các nghiên cứu dự báo chuỗi thời gian và ứng dụng trong báo phụ tải đã được sử dụng

Chương 2. Mô hình dự báo chuỗi thời gian và ứng dụng trong dự báo nhu cầu phụ tải điện: Chương này sẽ giới thiệu về các yêu cầu đặt ra đối với bài toán dự báo phụ tải, đồng thời nêu các lý thuyết liên quan đến giải pháp tiền xử lý dữ liệu, các phương pháp dự báo phụ tải sử dụng trong luận văn, qua đó đề xuất các mô hình dự báo phụ tải.

Chương 3. Xây dựng mô hình dự báo nhu cầu phụ tải điện: Chương này trình bày các kết quả phân tích, thiết kế, cài đặt và thử nghiệm các mô hình dự báo của hệ thống.

Chương 4. Kết quả dự báo nhu cầu phụ tải điện ở Công ty Điện lực Tây Ninh: Chương này trình bày thực nghiệm các kết quả dự báo nhu cầu phụ tải ở Công ty Điện lực Tây Ninh và nhận xét chung.

Chương 5. Xây dựng hệ thống dự báo nhu cầu phụ tải điện: Chương này giới thiệu hệ thống dự báo nhu cầu phụ tải điện hàng tháng ứng dụng các mô hình dự báo phụ tải điện đã nghiên cứu..

CHƯƠNG 1: TỔNG QUAN VỀ DỰ BÁO CHUỖI THỜI GIAN

Chương này nêu khái quát, hệ thống hóa cơ sở lý luận về vấn đề dự báo, đồng thời tập trung trình bày các nghiên cứu dự báo chuỗi thời gian và ứng dụng trong báo phụ tải điện đã được sử dụng.

1.1. Tổng quan về dự báo

Dự báo là khoa học và nghệ thuật tiên đoán các sự việc có thể sẽ xảy ra trong tương lai. Tính khoa học của dự báo thể hiện ở chỗ khi tiến hành dự báo người ta phải căn cứ trên các dữ liệu phản ánh tình hình thực tế trong quá khứ và hiện tại, căn cứ vào xu thế trên cơ sở khoa học để dự đoán những sự việc có thể sẽ xảy ra trong tương lai. Tính nghệ thuật trong dự báo thể hiện ở chỗ phải sử dụng tài phán đoán, kinh nghiệm trong những điều kiện thiếu thông tin hoặc nhu cầu của khách hàng biến động mạnh. Từ những đặc điểm này cho thấy dự báo bao giờ cũng có sai số, chỉ ngẫu nhiên nếu chúng ta dự báo đúng hoàn toàn, tính chính xác của dự báo càng thấp khi thời gian dự báo càng dài.

Các phương pháp dự báo cơ bản được phân thành hai loại cơ bản: các phương pháp định tính và định lượng, và việc lựa chọn loại thích hợp chủ yếu phụ thuộc vào dữ liệu có sẵn. Các phương pháp dự báo định tính mang tính chủ quan, sự kiện tương lai được dự đoán chủ quan dựa trên việc sử dụng ý kiến của các chuyên gia; tuy nhiên, chúng không hoàn toàn là phỏng đoán, nhưng chúng là các phương pháp tiếp cận có cấu trúc được phát triển để thu được các dự báo tốt mà không cần sử dụng dữ liệu lịch sử. Do đó, các phương pháp này hữu ích và được triển khai khi dữ liệu lịch sử không có sẵn hoặc khan hiếm. Mặt khác, các phương pháp dự báo định lượng dựa trên toán học và công thức thống kê. Chúng được áp dụng khi có sẵn dữ liệu, nhưng phải thỏa mãn hai điều kiện: thông tin về quá khứ có sẵn và giả định rằng một số các khía cạnh của mô hình trong quá khứ sẽ tiếp tục trong tương lai. Các phương pháp dự báo định lượng liên quan đến một loạt các phương pháp và mỗi phương pháp có các thuộc tính, độ chính xác và chi phí riêng phải được cân nhắc khi lựa chọn một

phương pháp cụ thể trong các ngành cụ thể, cho các mục đích cụ thể. Hầu hết các dự báo định lượng yêu cầu dữ liệu chuỗi thời gian được thu thập trong các khoảng thời gian đều đặn theo thời gian hoặc dữ liệu cắt ngang (dữ liệu được thu thập tại một thời điểm duy nhất).

1.2. Các phương pháp dự báo

Các phương pháp dự báo ban đầu hầu như hoàn toàn dựa theo phương pháp thống kê truyền thống nhưng với sự tiến bộ của khoa học hiện đại, các công nghệ dự báo đã phát triển đáng kể. Gần đây, các phương pháp dự báo dựa trên lý thuyết học máy là ngày càng trở nên phổ biến hơn trong dự báo. Phần sau đây giới thiệu và mô tả các phương pháp dự báo được sử dụng phổ biến nhất theo hai hướng sau: các mô hình thống kê và các mô hình hiện đại dựa trên máy học, khai phá dữ liệu và trí tuệ nhân tạo.

1.2.1. Các mô hình thống kê

Trong cách tiếp cận dự báo theo mô hình thống kê thì các nghiên cứu dự báo chuỗi thời gian theo mô hình ARIMA đơn giản hoặc SARIMA (ARIMA theo mùa) dựa theo phương pháp chuẩn Box-Jenkins được sử dụng một cách khá thường xuyên.

1.2.1.1. Mô hình Tự hồi quy tích hợp trung bình trượt (ARIMA)

Lê Huy Phúc và các đồng sự [4] đã nghiên cứu ứng dụng phương pháp lọc dữ liệu để đánh giá độ tin cậy của nguồn dữ liệu đầu vào nhằm nâng cao chất lượng dự báo phụ tải điện ngắn hạn trên lưới điện TP.HCM. Trong nghiên cứu, nhóm tác giả đã đề xuất áp dụng mô hình ARIMA và ANN để có sự đối chiếu (sau khi loại bỏ các yếu tố bất thường trong dữ liệu), kết quả cho thấy MAPE của mô hình ARIMA nhỏ hơn so với mô hình ANN, tương ứng 4,9% so với 5,87%.

Huỳnh Tấn Nguyên và Nguyễn Văn Lượng [17] nghiên cứu này ứng dụng mô hình ARIMA trong dự báo chỉ số giá tiêu dùng CPI ở Việt Nam trong quý 2 năm 2016. Số liệu được thu thập từ Tổng cục Thống kê Việt Nam giai đoạn tháng 1/2010 đến tháng 03/2016. Kết quả cho thấy mô hình ARIMA (2,1,1) là thích hợp cho việc

dự báo. Kết quả dự báo CPI quý 2 năm 2016 lần lượt là 159.5409%, 159.5447%, 159.5476% cho kết quả dự báo CPI tốt và việc ứng dụng mô hình ARIMA đã được thực hiện ở một số nghiên cứu trên thế giới.

Trịnh Thị Phan Lan [18] sử dụng mô hình trong hình chuyển động tự hồi quy ARIMA để tìm ra mô hình tốt cho việc dự báo biến động tỷ giá tại Việt Nam. Kết quả nghiên cứu cho thấy, mô hình SARIMA cho kết quả dự báo tốt nhất trong các mô hình được nghiên cứu. Sai số trong mô hình không lớn cho thấy, có thể sử dụng mô hình SARIMA để dự báo tỷ giá trong tương lai.

Nghiêm Phúc Hiếu [19] sử dụng phương pháp Box-Jenkins để xây dựng mô hình ARIMA theo mùa (hay còn gọi là SARIMA - Seasonal Autoregressive Integrated Moving Average) nhằm dự báo lượng khách quốc tế đến Việt Nam dựa trên số liệu công bố hàng tháng của Tổng cục Du lịch Việt Nam. Nghiên cứu cũng đưa ra dự báo thử nghiệm về lượng khách quốc tế đến Việt Nam những tháng cuối năm 2017 với mức độ sai số chấp nhận được từ 1.7% đến 12.4%.

Lê Nhị Bảo Ngọc và các đồng sự [20] sử dụng mô hình SARIMA để dự báo giá giao lên tàu (FOB) thực tôm sú ngắn hạn với nguồn số liệu là chuỗi giá tôm sú có kích cỡ 30-40 con/kg theo thời gian từ tháng 1/2011 đến tháng 12/2016. Kết quả nghiên cứu khẳng định mô hình là phù hợp để giải thích được sự biến động giá FOB thực của tôm sú trong giai đoạn nói trên. Đồng thời, mô hình dự báo rất đáng tin cậy, giá trị thực của tháng 1 trong năm 2017 nằm trong khoảng tin cậy 95% và gần bằng với giá trị dự báo với điểm sai số dự báo nhỏ.

1.2.1.2. Mô hình phân rã

Ý tưởng của mô hình phân rã là chuỗi thời gian có thể được phân tách thành ba thành phần: xu hướng (hướng dài hạn), theo mùa (có hệ thống, chuyển động liên quan đến lịch) và bất thường (biến động không hệ thống, ngắn hạn). Mô hình phân rã thông thường chia thành 2 loại: Mô hình cộng và Mô hình nhân

Trong mô hình cộng, chuỗi thời gian quan sát (O_t) được coi là tổng của ba thành phần độc lập: theo mùa S_t , xu hướng T_t và bất thường I_t :

$$O_t = T_t + S_t + I_t \quad (1.1)$$

Trong mô hình nhân, chuỗi thời gian quan sát (O_t) được biểu thị như là tích của các thành phần xu hướng T_t , theo mùa S_t và bất thường I_t :

$$O_t = T_t \times S_t \times I_t \quad (1.2)$$

Gần đây, các nghiên cứu dự báo chuỗi thời gian phi tuyến theo mô hình cộng (Additive Model) trong đó các xu thế phi tuyến khớp với tính thời vụ hàng năm, hàng tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ cũng được sử dụng cho hiệu quả tốt. Sean J. Taylor và Benjamin Letham [12] đã giới thiệu phương pháp phân tích chuỗi thời gian theo mô hình cộng bằng thuật toán Facebook Prophet trong dự báo số sự kiện trên Facebook. Huỳnh Trọng Thừa và các đồng sự [13] đã công bố kết quả nghiên cứu sử dụng thuật toán phân tích chuỗi thời gian Facebook Prophet trong dự báo hoạt động đầu tư ở Khu Công nghệ cao TP.HCM. Zar Zar Oo và Sabai Phyu [14] sử dụng thuật toán Prophet để dự báo nhiệt độ ở tỉnh Myintkyina, Myanmar cho kết quả tốt so với các phương pháp truyền thống. Toni Toharudin và các đồng sự [15] đã công bố báo cáo so sánh kết quả dự báo nhiệt độ không khí hàng ngày trong 5 năm ở Bandung, Taiwan giữa 2 mô hình dự báo bởi LSTM và Facebook Prophet; Các kết quả cho thấy rằng, Prophet cho kết quả tốt hơn đối với nhiệt độ tối đa, trong khi LSTM cho kết quả tốt hơn đối với nhiệt độ tối thiểu; Tuy nhiên, sự khác biệt về giá trị của RMSE không quá lớn đáng kể. Emir Žunić và các đồng sự [16] đề xuất bộ khung ứng dụng thuật toán Prophet trong dự báo doanh số bán hàng trong thực tế để dự báo doanh số hàng tháng, hàng quý sắp tới.

Cách tiếp cận dự báo theo mô hình thống kê, đặc biệt là mô hình dự báo chuỗi thời gian có ưu thế trong dự báo chuỗi thời gian khi mà xu thế, tính mùa vụ được thể hiện rõ ràng.

1.2.2. Các mô hình hiện đại dựa trên máy học, khai phá dữ liệu và trí tuệ nhân tạo

Các mô hình thống kê truyền thống còn hạn chế và đôi khi có thể dẫn đến các giải pháp không đạt yêu cầu. Các lý do là khối lượng tính toán quá lớn dẫn đến tiêu tốn nhiều thời gian để đạt kết quả và độ phức tạp của các mẫu dữ liệu phi tuyến. Do đó, máy học và các kỹ thuật dựa trên trí tuệ nhân tạo cung cấp một giải pháp thay thế đầy hứa hẹn và hấp dẫn.

1.2.2.1. Thuật toán Mạng neural nhân tạo (ANN)

Lei Guo và các đồng sự [10] áp dụng mô hình mạng neural LSTM (Long Short Term Memory) để dự báo phụ tải điện ngắn hạn. Mạng LSTM được xây dựng theo hai hướng, một giờ và một tuần, tương ứng. Mô hình dự đoán trong một giờ là đạt yêu cầu với MAPE 2.09%. Mô hình dự đoán của một tuần không hiệu quả (MAPE 7.03%) do thời gian dự báo quá dài và được cải thiện (MAPE 5.99%) bằng cách thêm đầu vào của thông tin hàng tuần và thông tin dự báo thời tiết, cho thấy tầm quan trọng của thông tin thời tiết để dự báo phụ tải điện ngắn hạn tuần.

E. C. Ashigwuike và các đồng sự [8] đã có nghiên cứu so sánh dự báo phụ tải trung hạn 48 tháng tới của lưới điện phân phối AMAC theo mô hình hồi quy đa biến và mô hình ANN có tính đến ảnh hưởng của nhiệt độ, thời gian, tỷ lệ gia tăng dân số và các hoạt động của các khu vực khác nhau (mô hình NARX-ANN). Các dữ liệu của các giá trị đỉnh hàng tháng đến hàng năm được thu thập trong khoảng thời gian từ năm 2012 đến quý đầu tiên của 2018. Kết quả nghiên cứu cho thấy mô hình NARX-ANN đề xuất hiệu quả hơn mô hình hồi quy đa biến do có tính đến các yếu tố ảnh hưởng đến phụ tải điện.

Trong nước, các nghiên cứu ứng dụng mạng ANN [21], [22] trong dự báo phụ tải điện ở Tiền Giang và Đà Nẵng. Các kết quả nghiên cứu cho thấy việc ứng dụng mạng ANN trong dự báo phụ tải ngắn hạn cho kết quả tốt. Kết quả nghiên cứu [21] phương pháp dự báo dùng mạng truyền thẳng nhiều lớp kết hợp với giải thuật lan

truyền ngược cho kết quả dự báo với sai số nhỏ, với dự báo 24 giờ trong ngày, các ngày trong tuần cho kết quả tương đối chính xác ($< 5\%$) nhưng với dự báo các tháng trong năm thì sai số tương đối lớn ($9,629\% - < 10\%$).

1.2.2.2. Support Vector Machines (SVMs)

Zhitong Ma và các đồng sự [11] đã sử dụng mô hình SVM để dự đoán mức tiêu thụ năng lượng xây dựng ở Trung Quốc. Mô hình có bảy thông số bao gồm dữ liệu thời tiết như nhiệt độ bầu khô ngoài trời trung bình hàng năm, độ ẩm tương đối và toàn cầu bức xạ mặt trời lấy từ NMIC, CMA và các yếu tố kinh tế như tỷ lệ đô thị hóa, tổng nội địa sản phẩm, mức tiêu dùng hộ gia đình và tổng diện tích cơ cấu thu thập từ NBSC từ năm 2000 đến năm 2014 được lấy làm đầu vào. Mô hình đề xuất có thể được khuyến nghị để dự đoán mức tiêu thụ năng lượng của tòa nhà ở những khu vực mà dữ liệu liên quan bị thiếu hoặc không có sẵn, giúp đưa ra quyết định về tòa nhà tiết kiệm năng lượng.

Kết quả các nghiên cứu ứng dụng mô hình máy học cho thấy, để nâng cao hiệu quả dự báo thì các yếu tố ảnh hưởng đến vấn đề dự báo phải được tính đến và bao gồm trong tham số đầu vào. Đối với dự báo phụ tải điện thì các yếu tố ảnh hưởng đến nhu cầu phụ tải cần phải xem xét là tốc độ tăng trưởng kinh tế, thời tiết, nhiệt độ, độ ẩm, mật độ dân số, ... Tuy nhiên, các yếu tố này lại khó dự đoán cũng làm ảnh hưởng đến kết quả dự báo của các mô hình dựa trên máy học.

1.3. Thực trạng dự báo phụ tải điện ở Công ty Điện lực Tây Ninh

Công ty Điện lực Tây Ninh là đơn vị thành viên trong Tổng công ty Điện lực miền Nam (EVN SPC) thuộc Tập đoàn Điện lực Việt Nam (EVN) thực hiện chức năng quản lý, phân phối và phát triển điện trên địa bàn tỉnh Tây Ninh. Công ty Điện lực Tây Ninh có tổng cộng 882 lao động công tác trên các lĩnh vực sản xuất kinh doanh điện, tư vấn khảo sát thiết kế, giám sát chất lượng thi công xây dựng công trình điện, quản lý vận hành lưới điện trên địa bàn 09 huyện, thị và thành phố trong tỉnh.

Trong các năm qua sản xuất công nghiệp trên địa bàn tỉnh Tây Ninh phục hồi mạnh mẽ, nhiều doanh nghiệp đầu tư mới đưa vào hoạt động. Do đó, sản lượng điện thương phẩm giai đoạn từ năm 2016-2019 tăng trưởng cao đạt khoảng 15,1% (Công ty luôn có tốc độ tăng trưởng thuộc nhóm đầu trong 21 Công ty Điện lực thuộc các tỉnh phía Nam). Lũy kế 12 tháng năm 2020 toàn Công ty thực hiện 4.707,01 triệu kWh, tăng 14,65% so với cùng kỳ năm 2019 (4.105,38 triệu kWh; Điện lực Gò Dầu có tỷ lệ tăng so với cùng kỳ năm 2019 cao nhất (24,97%). Thương phẩm theo 5 thành phần phụ tải 12 tháng năm 2020 so với cùng kỳ 2019: Nông-Lâm-Thủy sản tăng 58,70%; Công nghiệp-xây dựng tăng 15,96%; Thương nghiệp-khách sạn-nhà hàng tăng 13,15%; Quản lý và tiêu dùng dân cư tăng 08,94%; Các hoạt động khác tăng 10,94%. Về tỷ trọng 12 tháng năm 2020 các thành phần phụ tải trong thương phẩm: Nông-Lâm-Thủy sản chiếm 3,11%; Công nghiệp-xây dựng chiếm 74,64%; Thương nghiệp-khách sạn-nhà hàng chiếm 1,57%; Quản lý và tiêu dùng dân cư chiếm 18,79%; Các hoạt động khác chiếm 1,89%.

Với kết quả đó, để có kế hoạch chuẩn bị sẵn sàng đáp ứng nhu cầu phụ tải điện trong tương lai phục vụ công tác điều hành hệ thống điện với mục tiêu hiệu quả, tin cậy thì việc dự báo chính xác xu hướng, tốc độ phát triển và nhu cầu của phụ tải điện trong tương lai trở nên cần thiết.

Bên cạnh đó, với mục tiêu kép là triển khai thực hiện Thông tư số 19/1017/TT-BCT của Bộ công thương [1] và nâng cao hiệu quả hoạt động, độ tin cậy cung cấp điện và đảm bảo cân đối cung cầu giữa các nguồn điện phát và nhu cầu phụ tải điện phục vụ công tác điều hành hàng tháng, Tập đoàn Điện lực Việt Nam (EVN) đã chỉ đạo các Tổng công ty phân phối điện triển khai thực hiện dự báo nhu cầu phụ tải điện (sản lượng điện thương phẩm) ở các Công ty Điện lực với phương pháp dự báo phụ tải bằng việc kết hợp 02 phương pháp dự báo trực tiếp và thống kê đối với từng nhóm khách hàng, cụ thể như sau:

- Nhóm khách hàng lớn (điện thương phẩm $\geq 1.000.000$ kWh/năm): áp dụng phương pháp dự báo trực tiếp

- Nhóm khách hàng còn lại (điện thương phẩm < 1.000.000 kWh/năm): áp dụng phương pháp dự báo thống kê và sử dụng các công cụ để dự báo điện tiêu thụ dự kiến hàng tháng và cả năm của khách hàng.

Trong quá trình triển khai thực hiện công tác dự báo nhu cầu phụ tải điện (sản lượng điện thương phẩm hàng tháng), Công ty Điện lực Tây Ninh sử dụng các công cụ dự báo bằng phương pháp dự báo phụ tải theo mô hình hàm hồi quy đa biến có sẵn trên Microsoft Excel hoặc sử dụng phương pháp chuyên gia để dự báo. Tuy nhiên, các yếu tố đầu vào cho việc dự báo bằng phương pháp hồi quy đa biến như GDP (Tổng thu nhập quốc nội), thời tiết, nhiệt độ và độ ẩm tùy theo khu vực khó thu thập và khó dự báo được chính xác làm ảnh hưởng đến quá trình dự báo. Hoặc dự báo theo phương pháp chuyên gia lại phụ thuộc yếu tố kinh nghiệm, không mang tính bền vững và cũng bị ảnh hưởng bởi các điều kiện phức tạp của dự báo. Ngoài hai phương pháp dự báo đó, Công ty Điện lực Tây Ninh cũng chưa từng có công trình nghiên cứu áp dụng các phương pháp dự báo khác vào công tác dự báo nhu cầu phụ tải điện. Do đó, kết quả dự báo là hạn chế và có sự chênh lệch với kết quả thực tế.

CHƯƠNG 2: MÔ HÌNH DỰ BÁO CHUỖI THỜI GIAN VÀ ỨNG DỤNG TRONG DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN

Chương này sẽ giới thiệu về các yêu cầu đặt ra đối với bài toán dự báo phụ tải, đồng thời nêu các lý thuyết liên quan đến giải pháp tiên xử lý dữ liệu, các phương pháp dự báo phụ tải sử dụng trong luận văn, qua đó đề xuất các mô hình dự báo phụ tải.

2.1. Phân tích chuỗi thời gian nhu cầu phụ tải điện

Trong các bài toán dự báo nói chung và các bài toán dự báo nhu cầu phụ tải điện nói riêng, dữ liệu dùng cho dự báo là dữ liệu chuỗi thời gian. Dữ liệu chuỗi thời gian là một chuỗi các quan sát của một quá trình ngẫu nhiên được lưu trữ theo thứ tự thời gian. Một số ví dụ về chuỗi thời gian như dữ liệu về giá chứng khoán hằng ngày, doanh thu bán hàng hằng ngày, sản lượng điện tiêu thụ hàng tháng của một xưởng sản xuất, ... Ưu điểm của chuỗi thời gian là nó có thể lưu trữ được trạng thái của một trường dữ liệu theo thời gian để từ đó rút trích ra các thông tin quan trọng để dự báo cho các trạng thái trong tương lai. Chính vì thế dữ liệu chuỗi thời gian đóng một vai trò cực kỳ quan trọng đối với sự phát triển.

Đặc điểm của dữ liệu chuỗi thời gian:

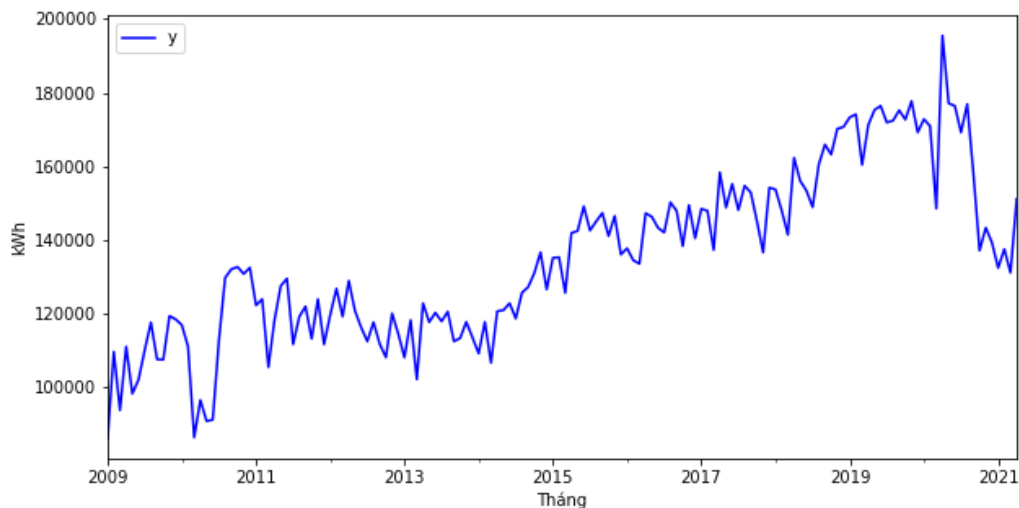
Dữ liệu chuỗi thời gian các đặc trưng là tính xu hướng và tính chu kỳ:

- Tính xu hướng là yếu tố thể hiện xu hướng thay đổi của dữ liệu theo thời gian. Một số ví dụ có thể thấy tính xu hướng thay đổi theo thời gian của chuỗi thời gian dân số thế giới tăng qua các năm, nhiệt độ trung bình trái đất tăng theo thời gian... Đây là đặc trưng thường thấy của rất nhiều dữ liệu chuỗi thời gian.
- Tính chu kỳ là qui luật có tính chất lặp lại của dữ liệu theo thời gian. Bất kỳ sự vận động, phát triển của sự vật hay hiện tượng nào cũng đều chịu ảnh hưởng của chu kỳ và lặp lại theo thời gian như sự thay đổi thời tiết, sự phát

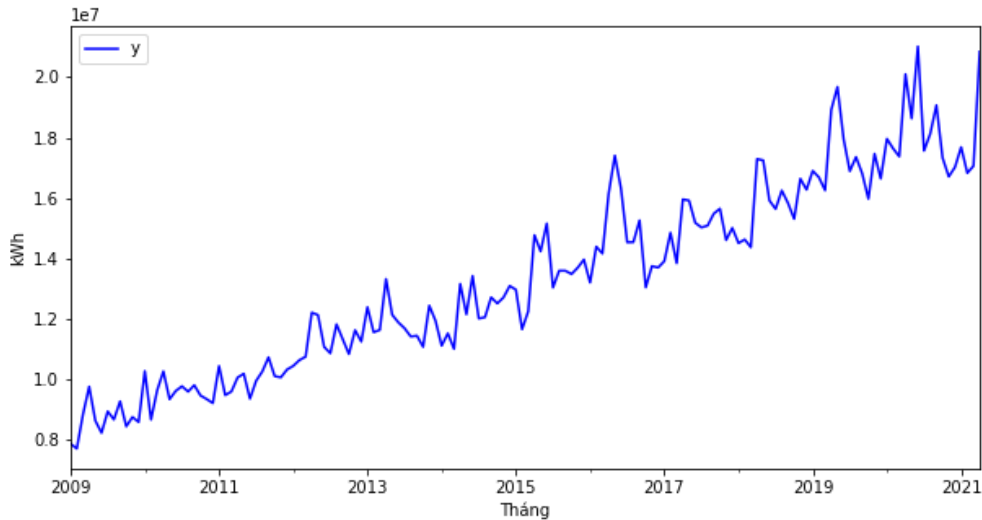
triển của các loài động vật cho tới hành vi mua sắm, tiêu dùng của con người... Chính vì thế tìm ra được yếu tố chu kỳ sẽ giúp ích cho việc dự báo chính xác hơn.

- Tính ngẫu nhiên thể hiện sự biến thiên không đoán trước được xung quanh xu hướng của dữ liệu theo thời gian.

Dữ liệu sản lượng điện thương phẩm hàng tháng của người dân hay các cơ quan tổ chức là dữ liệu chuỗi thời gian, cũng có các đặc trưng của một chuỗi thời gian là tính xu hướng và tính chu kỳ. Theo sự phát triển của xã hội nhu cầu về năng lượng ngày càng tăng, do đó chuỗi thời gian sản lượng điện thương phẩm có xu hướng tăng theo thời gian và cũng có tính chu kỳ là năm (12 tháng). Hình 2-1a, 2-1b là biểu đồ minh họa chuỗi thời gian sản lượng điện thương phẩm hàng tháng của một doanh nghiệp có tổng sản lượng tiêu thụ hàng năm trên 1 triệu kWh và tổng hợp của các khách hàng có mức tiêu thụ hàng tháng dưới 1 triệu kWh trên địa bàn Thành phố Tây Ninh.



Hình 2.1a: Biểu đồ sản lượng điện thương phẩm hàng tháng của một doanh nghiệp có mức tiêu thụ hàng năm trên 1 triệu kWh



Hình 2.1b: Biểu đồ sản lượng điện thương phẩm hàng tháng các khách hàng có mức tiêu thụ dưới 1 triệu kWh/năm

Phân tích khám phá dữ liệu sản lượng điện thương phẩm hàng tháng:

Phân tích chuỗi thời gian là một cách tiếp cận để phân tích dữ liệu có dạng chuỗi thời gian để rút ra các đặc điểm có ý nghĩa của dữ liệu và tạo ra những hiểu biết hữu ích khác. Phân tích khám phá dữ liệu là một cách có hệ thống để khám phá dữ liệu bằng cách sử dụng các phương pháp chuyển đổi và trực quan hóa dữ liệu. Phân tích khám phá dữ liệu chuỗi thời gian sản lượng điện thương phẩm hàng tháng cho chúng ta cái nhìn đầu tiên về dữ liệu, từ đó rút ra các đặc trưng hỗ trợ cho việc lựa chọn các tham số của mô hình dự báo hay tiền xử lý dữ liệu trước khi đưa vào mô hình dự báo. Một số đặc điểm của dữ liệu chuỗi thời gian có thể rút ra khi thực hiện khám phá dữ liệu, cụ thể:

Tính dừng: Một chuỗi thời gian Y_t được xem là dừng nếu như giá trị trung bình và phương sai của chuỗi không thay đổi theo thời gian và giá trị của hiệp phương sai giữa hai thời đoạn chỉ phụ thuộc vào độ trễ về thời gian giữa hai thời đoạn này chứ không phụ thuộc vào thời điểm thực tế mà hiệp phương sai được tính. Cụ thể:

- Trung bình: $\mu = E(Y_t) = \text{hằng số}$
- Phương sai: $\sigma^2 = Var(Y_t) = \text{hằng số}$

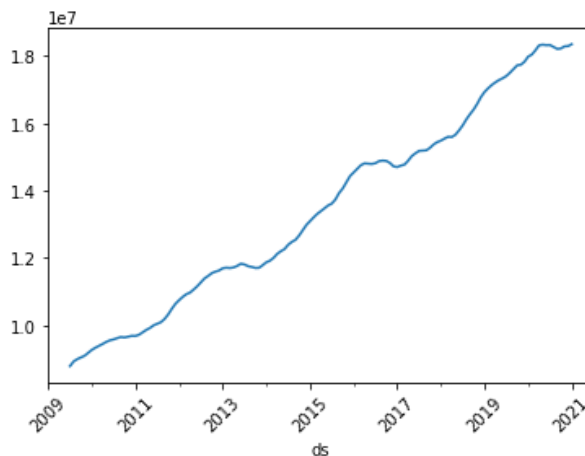
- Hiệp phương sai giữa 2 thời đoạn: $cov(Y_t, Y_{t-k}) = g_k$

Việc khảo sát chuỗi thời gian dừng dễ dàng hơn. Phương pháp chuyển đổi đơn giản nhất chuỗi không dừng sang chuỗi dừng là lấy sai phân thứ d.

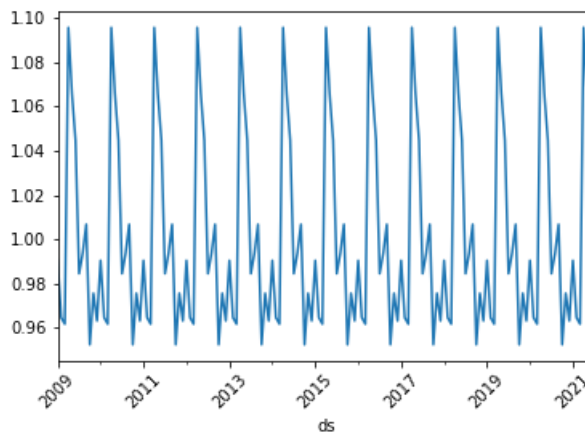
Tính mùa vụ: Tính mùa vụ là hành vi có tính chu kỳ của chuỗi thời gian trên cơ sở năm lịch. Chuỗi thời gian có tồn tại tính mùa vụ sẽ không có tính dừng. Phương pháp chuyển đổi đơn giản nhất để khử tính mùa vụ là lấy sai phân thứ m. Nếu Y_t có tính mùa vụ với chu kỳ m thời đoạn thì chuỗi $Z_t = Y_t - Y_{t-m}$ sẽ được khảo sát thay vì chuỗi Y_t .

Hình 2-2 minh họa khám phá các đặc trưng của chuỗi thời gian sản lượng điện thương phẩm các khách hàng dưới 1 triệu kWh/năm cho thấy xu hướng tăng theo thời gian của chuỗi, tính chu kỳ và biến thiên ngẫu nhiên.

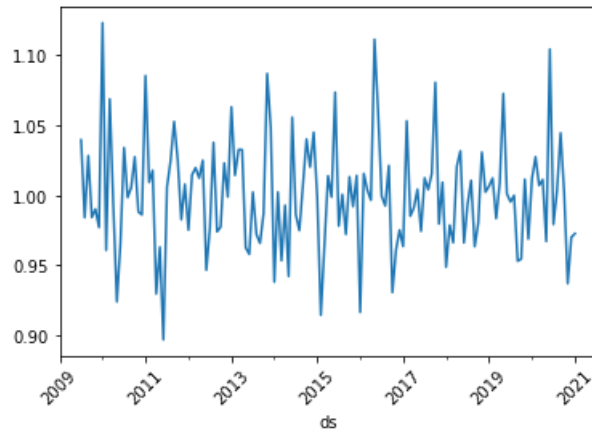
Thành phần xu hướng



Thành phần chu kỳ

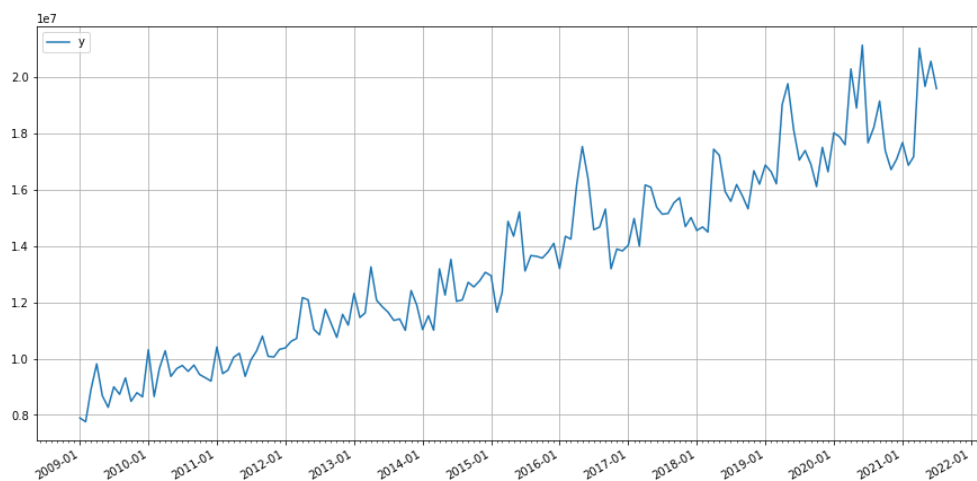


Thành phần ngẫu nhiên

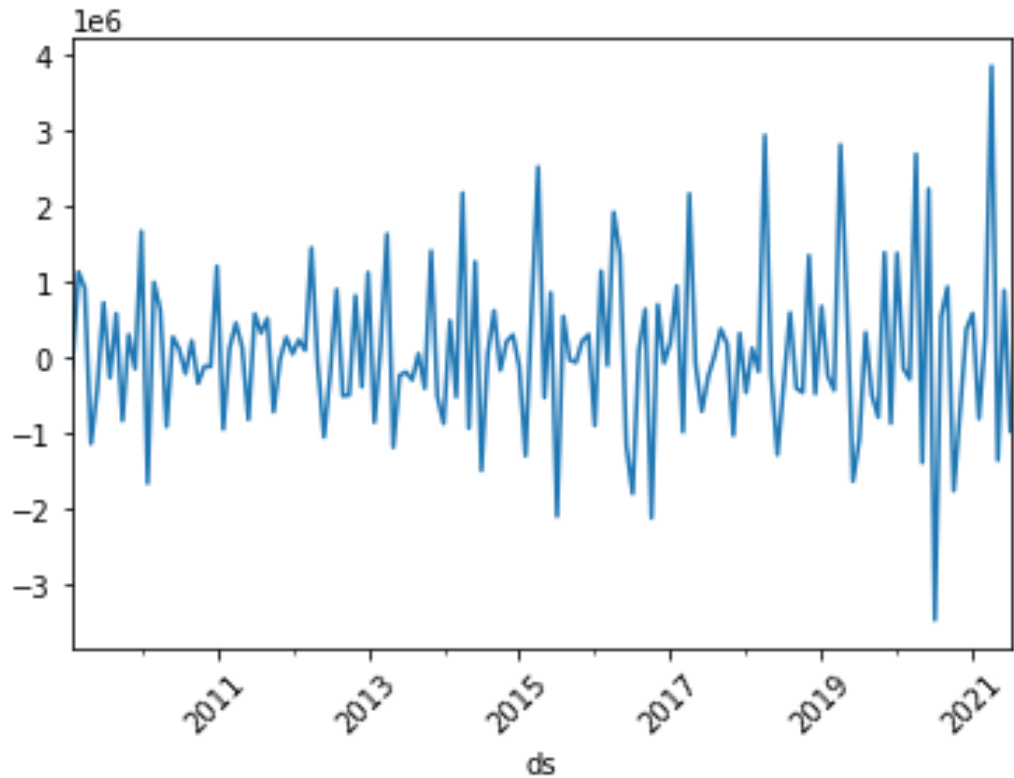


Hình 2.2: Biểu đồ phân rã các đặc trưng (thành phần) của dữ liệu chuỗi thời gian sản lượng điện thương phẩm hàng tháng chu kỳ 12 tháng của các khách hàng có mức tiêu thụ dưới 1 triệu kWh/năm

Một cách trực quan để khảo sát tính mùa vụ và tính dừng của chuỗi thời gian là vẽ biểu đồ theo chu kỳ và biểu đồ sai phân của nó. Hình 2-3a là biểu đồ sản lượng điện thương phẩm hàng tháng của các khách hàng có mức tiêu thụ dưới 1 triệu kWh/năm cho thấy chu kỳ mùa vụ của chuỗi là 12 tháng và chuỗi là không dừng (xu hướng tăng theo thời gian). Hình 2-3b là biểu đồ sai phân bậc 1 của chuỗi thời gian chuỗi thời gian sản lượng điện thương phẩm, qua biểu đồ cho thấy chuỗi dừng khi chuyển đổi sang chuỗi sai phân bậc 1.



Hình 2.3a: Biểu đồ chu kỳ 12 tháng của chuỗi thời gian sản lượng điện thương phẩm



Hình 2.3b: Biểu đồ sai phân bậc 1 của chuỗi thời gian sản lượng điện thương phẩm

Như vậy, qua phân tích khám phá dữ liệu chuỗi thời gian sản lượng điện thương phẩm cho thấy chuỗi có xu hướng tăng theo thời gian gian và chu kỳ mùa vụ là 12 tháng, chuỗi dừng khi tính sai phân bậc 1.

2.2. Mô hình ARIMA

Mô hình trung bình trượt tích hợp tự hồi quy (Autoregressive Integrated Moving Average - ARIMA) là một phương pháp rất phổ biến trong việc lập mô hình chuỗi thời gian, thường được gọi là phương pháp Box-Jenkins (1976). Mô hình ARIMA phân tích tính tự tương quan trong dữ liệu quan sát để đưa ra mô hình dự báo. Mô hình ARIMA được kết hợp bởi 3 thành phần chính: AR (Autoregressive - thành phần tự hồi quy), MA (Moving Average - thành phần trung bình trượt) và I (Integrated – thành phần tích hợp, thực hiện quá trình sai phân để tạo chuỗi dừng).

Mô hình tự hồi quy (Autoregressive) AR(p): là quá trình trong đó giá trị dừng hiện tại y_t phụ thuộc tuyến tính dựa trên p giá trị trong quá khứ của nó. Ý tưởng của mô hình AR(p) là hồi quy dựa trên sự tồn tại mối quan hệ phụ thuộc tuyến tính của các phân tử chuỗi dừng vào p phân tử của nó trong quá khứ cộng với nhiễu trắng không tương quan với nó:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (2.1)$$

Trong đó:

y_t : giá trị quan sát dừng hiện tại.

$y_{t-1}, y_{t-2}, \dots, y_{t-p}$: là các giá trị quan sát dừng trong quá khứ.

$\phi_1, \phi_2, \dots, \phi_p$: là p hệ số tự hồi quy được lựa chọn để sinh ra giá trị hồi quy phù hợp nhất cho y_t qua các giá trị quá khứ $y_{t-1}, y_{t-2}, \dots, y_{t-p}$.

ϵ_t : là sai số dự báo ngẫu nhiên tại thời điểm t . ϵ_t là nhiễu trắng có trung bình bằng 0.

Mô hình trung bình trượt (Moving Average) MA(q): giá trị quan sát dừng hiện tại y_t là hàm phụ thuộc tuyến tính q sai số dự báo quá khứ và sai số ngẫu nhiên hiện tại:

$$y_t = \theta_0 + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t \quad (2.2)$$

Trong đó:

y_t : giá trị quan sát dừng hiện tại.

$\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$: là q sai số dự báo quá khứ.

$\theta_1, \theta_2, \dots, \theta_q$: là q hệ số hồi quy trung bình trượt được lựa chọn phù hợp nhất cho y_t qua các giá trị quá khứ $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}$.

ϵ_t : là sai số dự báo ngẫu nhiên tại thời điểm t . ϵ_t là nhiễu trắng có trung bình bằng 0.

Tích hợp $I(d)$: Thông thường các chuỗi thời gian không dừng, hầu hết đều tăng hoặc giảm theo thời gian, do vậy cần biến đổi chuỗi thời gian trở thành chuỗi dừng. Phương pháp chủ yếu thường được sử dụng là lấy sai phân để chuỗi biến đổi thu được là chuỗi dừng và chuỗi này được gọi là chuỗi tích hợp. Khi lấy sai phân bậc d để chuỗi cuối cùng thu được là chuỗi dừng thì được gọi là chuỗi tích hợp bậc d : $I(d)$.

Quá trình sai phân bậc d của một chuỗi y_t là:

- Sai phân bậc 1: $I(1) = \Delta y_t = y_t - y_{t-1}$
- Sai phân bậc d : $I(d) = \Delta^d y_t = \underbrace{\Delta(\dots \Delta(\Delta y_t) \dots)}_{d \text{ lần}}$

Trên thực tế, rất ít chuỗi thời gian là chuỗi tích hợp bậc lớn hơn 1. Một chuỗi sẽ dừng sau quá trình sai phân bậc 0 hoặc 1.

Như vậy, phương trình trung bình trượt tích hợp tự hồi quy ARIMA (p, d, q) của chuỗi y_t có thể được biểu diễn dưới dạng:

$$\Delta^d y_t = \phi_1 \Delta^d y_{t-1} + \dots + \phi_p \Delta^d y_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (2.3)$$

Trong đó:

- d là bậc sai phân để chuỗi y_t là chuỗi dừng
- p là bậc tự hồi quy
- q là bậc trung bình trượt
- $\Delta^d y_t$ là giá trị sai phân bậc d
- ϵ_t là giá trị nhiễu trắng

Mô hình ARIMA được thực hiện theo phương pháp Box-Jenkins với bốn bước lập:

- Nhận dạng mô hình

- Ước lượng tham số và lựa chọn mô hình
- Kiểm định mô hình
- Dự báo

Các bước xây dựng mô hình ARIMA để dự báo chuỗi thời gian:

Nhận dạng mô hình: Nhận dạng mô hình ARIMA(p,d,q) là tìm các giá trị thích hợp của p, d, q, với d là bậc sai phân của chuỗi thời gian được khảo sát, p là bậc tự hồi quy và q là bậc trung bình trượt.

Ước lượng tham số và lựa chọn mô hình: Sau khi đã nhận dạng các giá trị thích hợp của p và q, bước tiếp theo là ước lượng các thông số của các số hạng tự hồi quy và trung bình trượt trong mô hình.

Kiểm định mô hình: Sau khi đã lựa chọn mô hình ARIMA cụ thể và ước lượng các tham số của nó, ta tìm hiểu xem mô hình lựa chọn có phù hợp với dữ liệu ở mức chấp nhận hay không bởi vì có thể một mô hình ARIMA khác cũng phù hợp với dữ liệu. Nếu mô hình ARIMA không phù hợp thì lựa chọn mô hình khác phù hợp hơn. Như vậy, phương pháp Box-Jenkins là một quá trình lặp lại.

Dự báo: Ở bước cuối cùng này, sau khi kiểm định mô hình dự báo, khi mô hình phù hợp với dữ liệu sẽ được sử dụng vào việc dự báo.

2.3. Mô hình SARIMA

Mô hình trung bình trượt tích hợp tự hồi quy theo mùa (SARIMA hoặc ARIMA theo mùa), là một mô hình mở rộng của ARIMA hỗ trợ dữ liệu chuỗi thời gian đơn biến với thành phần theo mùa. SARIMA bổ sung thêm ba tham số mới là thành phần tự hồi quy (AR), tích hợp (I) và trung bình động (MA) cho thành phần theo mùa của chuỗi thời gian đơn biến, cũng như một tham số xác định khoảng thời gian theo mùa. Việc xác định bộ tham số của mô hình SARIMA yêu cầu chọn các tham số cho cả các yếu tố xu hướng và theo mùa của chuỗi.

Ba tham số xu hướng giống như mô hình ARIMA:

- p: Bậc tự hồi quy xu hướng.
- d: Bậc tích hợp xu hướng
- q: Bậc trung bình động xu hướng.

Bốn tham số theo mùa không phải là một phần của ARIMA phải được xác định:

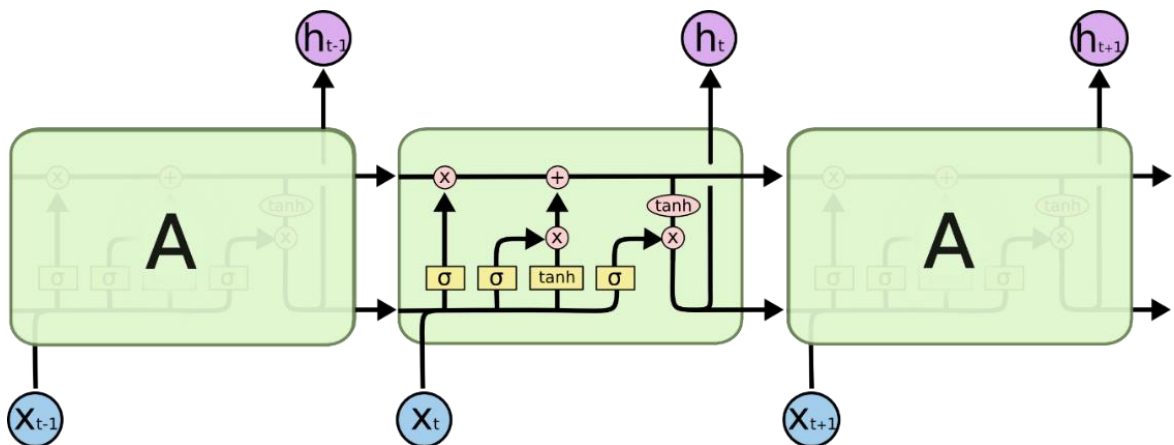
- P: Bậc tự hồi quy theo mùa.
- D: Bậc tích hợp theo mùa.
- Q: Bậc trung bình động theo mùa.
- m: Số bước thời gian (kỳ) trong một mùa.

Như vậy, tùy theo giá trị của các tham số mà mô hình SARIMA có thể là mô hình ARIMA, ARMA, AR hay là MA.

2.4. Mô hình mạng LSTM

Mạng LSTM (Long Short Term Memory) là một biến thể của mạng nơ-ron hồi quy (Recurrent Neural Network - RNN) được thiết kế để giải quyết các bài toán về phụ thuộc xa trong RNN do bị ảnh hưởng của vấn đề triệt tiêu gradient (gradient vanishing). LSTM lần đầu tiên được đề xuất bởi Hochreiter và Schmidhuber và được cải tiến bởi nhiều nhà nghiên cứu khác.

Hình 2-4 biểu diễn cấu trúc của một mạng LSTM gồm nhiều tế bào liên kết với nhau thành một chuỗi. Mỗi tế bào LSTM có 4 tầng (tương ứng với 4 hàm kích hoạt sigmoid, tanh) tương tác với nhau để tạo ra trạng thái tế bào (cell state). Hình 2-6 là cấu trúc bên trong của một tế bào LSTM.

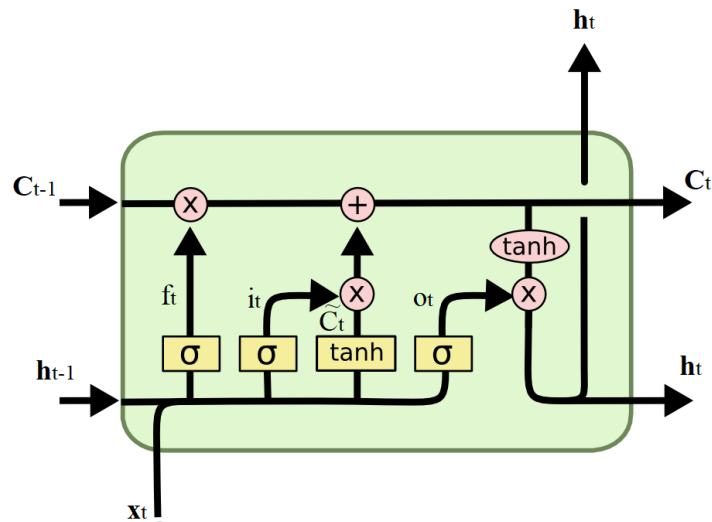


Hình 2.4: Một mạng LSTM và cấu trúc bên trong của tế bào LSTM

(Nguồn: <https://colah.github.io/>)

Mỗi tế bào có thể ghi nhớ hoặc quên thông tin cần thiết thông qua cơ chế sàng lọc thông tin của các cổng (gate). Một tế bào LSTM có 3 cổng như vậy để duy trì và điều hành trạng thái của tế bào:

- *Cổng quên (forget gate)*: Có nhiệm vụ loại bỏ những thông tin không cần thiết nhận được khỏi trạng thái bên trong tế bào.
- *Cổng vào (input gate)*: Có nhiệm vụ chọn lọc những thông tin cần thiết nào được thêm vào trạng thái bên trong tế bào.
- *Cổng ra (output gate)*: Có nhiệm vụ xác định những thông tin nào từ trạng thái bên trong tế bào được sử dụng như đầu ra.



Hình 2.5: Cấu trúc của một tế bào LSTM

(Nguồn: <https://colah.github.io/>)

Về mặt toán học, cơ chế hoạt động của một tế bào LSTM như sau:

- Đầu tiên, tế bào LSTM nhận trạng thái bên trong của bộ nhớ trước đó C_{t-1} và xác định có quên trạng thái trước đó hay không qua phương trình:

$$C'_t = C_{t-1} * f_t \quad (2.4)$$

Trong đó, công quên f_t trả về giá trị trong đoạn $[0,1]$, xác định mức độ quên trạng thái bên trong của tế bào trước.

- Song song đó, trạng thái bên trong tế bào hiện tại C_t được cập nhật thêm thông tin đầu vào qua :

$$C_t = C'_t + (i_t * \check{C}_t) \quad (2.5)$$

Trong đó, i_t chứa giá trị kích hoạt cổng vào, xác định mức độ sàng lọc thông tin đầu vào và \check{C}_t là giá trị trạng thái tiềm năng của tế bào hiện tại.

- Như vậy, trạng thái bên trong của tế bào hiện tại truyền sang tế bào kế tiếp là:

$$C_t = C_{t-1} * f_t + (i_t * \check{C}_t) \quad (2.6)$$

- Cuối cùng, trạng thái đầu ra của tế bào (h_t) được tính dựa theo phương trình:

$$h_t = o_t * \tanh(C_t) \quad (2.7)$$

Trong đó, o_t là giá trị kích hoạt cổng ra, xác định mức độ sàng lọc thông tin đầu ra của trạng thái bên trong tế bào.

2.5. Mô hình Prophet

Prophet là một mô hình dự báo dữ liệu chuỗi thời gian dựa trên mô hình cộng (Additive model) trong đó các xu hướng phi tuyến tính phù hợp với thời vụ hàng năm, hàng tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ. Cơ sở của mô hình Prophet là phân rã chuỗi thời gian thành 4 thành phần đại diện cho xu hướng, tính chu kỳ, ảnh hưởng của ngày lễ và sai số mô hình theo phương trình:

$$y(t) = g(t) + s(t) + h(t) + \epsilon t \quad (2.8)$$

Trong đó:

- $g(t)$ là hàm đại diện cho xu hướng, sự thay đổi không có tính chu kỳ của chuỗi thời gian.
- $s(t)$ là hàm đại diện cho sự thay đổi có tính chu kỳ của chuỗi thời gian (ví dụ: hàng tuần, hàng tháng, hàng năm)
- $h(t)$ là hàm đại diện cho sự ảnh hưởng của ngày nghỉ (do người dùng cung cấp) xảy ra theo lịch một hoặc vài ngày.
- ϵt là sai số mang tính ngẫu nhiên không xác định được của mô hình.

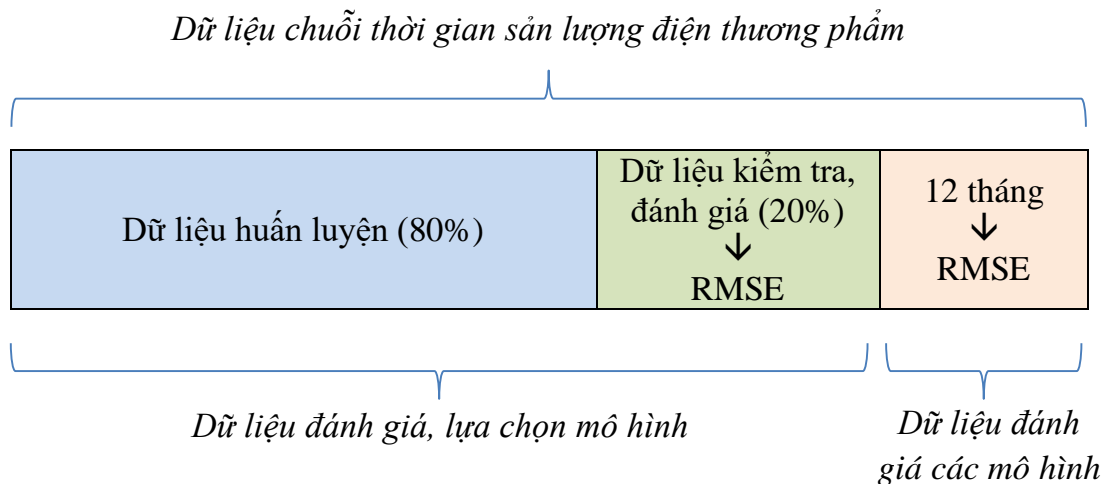
CHƯƠNG 3: XÂY DỰNG MÔ HÌNH DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN

Chương này trình bày các kết quả phân tích, thiết kế, cài đặt và thử nghiệm các mô hình dự báo của hệ thống.

3.1. Xây dựng mô hình dự báo

Để xây dựng mô hình, chúng tôi sử dụng chiến lược kiểm tra dự báo góidầu, hay còn gọi là chiến lược kiểm tra góidầu (walk forward validation) để kiểm tra, đánh giá mô hình dự báo cũng như đánh giá hiệu suất giữa các mô hình.

Trước tiên, toàn bộ dữ liệu đầu vào sản lượng điện thương phẩm (chuỗi dữ liệu thời gian nhu cầu phụ tải điện) được phân chia thành 2 phần, phần dữ liệu đầu là đánh giá, lựa chọn mô hình được phân chia thành dữ liệu huấn luyện và kiểm tra theo tỷ lệ 80% dữ liệu huấn luyện và 20% dữ liệu kiểm tra, đánh giá mô hình thử nghiệm; phần dữ liệu đánh giá các mô hình (12 tháng cuối dùng) dùng đánh giá hiệu suất giữa các mô hình dự báo SARIMA, mạng LSTM và Prophet. Hình 3-1 minh họa việc phân chia dữ liệu chuỗi thời gian sản lượng điện thương phẩm trong việc đánh giá lựa chọn các tham số của mô hình thử nghiệm và đánh giá hiệu suất giữa các mô hình thử nghiệm.



Hình 3.1: Minh họa phân chia dữ liệu chuỗi thời gian sản lượng điện thương phẩm trong lựa chọn mô hình và đánh giá các mô hình.

Kế tiếp, chiến lược kiểm tra góidầu được áp dụng trên các dữ liệu kiểm tra, đánh giá. Hình 3-2 minh họa mã giả áp dụng chiến lược kiểm tra góidầu để lựa chọn các tham số mô hình theo độ đo RMSE, mô hình có RMSE nhỏ nhất được lựa chọn.

Chiến lược kiểm tra góì đầu được áp dụng tương tự trên dữ liệu 12 tháng cuối cùng (dữ liệu đánh giá các mô hình) để đánh giá hiệu suất giữa các mô hình dự báo thử nghiệm.

Mã giả chiến lược kiểm tra góì đầu (walk forward validation)

Lấy tất cả dữ liệu đầu vào X_i ($i = 1, 2, 3, \dots, N$)

Phân chia dữ liệu huấn luyện/kiểm tra theo tỷ lệ 70/30, $T = \text{INT}(N * 0.7)$

Train = $[X_i]$ ($i = 1, 2, \dots, T$)

Test = $[X_i]$ ($i = T + 1, \dots, N$)

Forecast = []

Lặp $j = 1$ to $(N - T)$:

Huấn luyện mô hình trên dữ liệu Train

Dự báo 1 kỳ kế tiếp

Ghi nhận kết quả dự báo vào Forecast

Thêm Test[j] vào Train

Tính hiệu suất mô hình theo RMSE (Test, Forecast)

Hình 3.2: Mã giả thuật toán đánh giá mô hình theo chiến lược kiểm tra góì đầu được áp dụng trong lựa chọn mô hình và đánh giá hiệu suất các mô hình.

Lựa chọn tham số mô hình:

Các mô hình dự báo SARIMA, mạng LSTM và Prophet được xây dựng với vài tham số. Mô hình SARIMA, các tham số có thể lựa chọn để tinh chỉnh hiệu suất của mô hình như p, d, q, P, D, Q và m . Mô hình mạng LSTM, các tham số có thể lựa chọn để tinh chỉnh hiệu suất của mô hình như `neurons`, `batch_size`, `epochs`, ... Mô hình Prophet, các tham số có thể lựa chọn để tinh chỉnh hiệu suất của mô hình như `seasonality_mode`, `seasonality_prior_scale`, ...

Chi tiết việc đánh giá lựa chọn các tham số từng mô hình SARIMA, mạng LSTM và Prophet được nêu rõ trong các phần sau của chương này.

Độ đo đánh giá mô hình:

Độ đo áp dụng trong kiểm tra, đánh giá lựa chọn các mô hình thử nghiệm là độ đo RMSE (root mean squared error):

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad (3.1)$$

Trong đó, \hat{y}_i là kết quả dự báo của mô hình ở thời điểm i , y_i là giá trị quan sát (giá trị thực) ở thời điểm i .

Độ đo RMSE được áp dụng trong đánh giá lựa chọn các tham số mô hình và trong đánh giá hiệu suất giữa các mô hình thử nghiệm SARIMA, mạng LSTM và Prophet.

3.2. Xây dựng mô hình dự báo SARIMA

Các tham số theo xu hướng (p, d, q) và theo mùa (P, D, Q) của mô hình có thể được xác định bằng cách phân tích các đồ thị tự tương quan và tự tương quan một phần (ACF, PACF), và điều này đòi hỏi phải có kỹ năng phân tích chuyên môn. Cách tiếp cận phổ biến thường gặp khi phân tích chuỗi thời gian đơn biến theo mô hình ARIMA là cách tiếp cận Box-Jenkins, thường được gọi là phương pháp Box-Jenkins (1976). Một cách tiếp cận thay thế là tìm kiếm theo lưới (grid search) với một bộ các tham số mô hình và khám phá bộ tham số nào hoạt động tốt nhất cho một chuỗi thời gian đơn biến cụ thể. Với hệ thống máy tính hiện đại ngày, tốc độ xử lý nhanh hơn, cách tiếp cận này có thể khả thi và thực hiện tìm kiếm nhanh hơn, kết quả đạt được có thể khác so với quy trình phân tích theo phương pháp Box-Jenkins nhưng sai số dự báo thấp hơn, tuy không rõ ràng.

Hình 3-3 sau đây là mã giả của thuật toán tìm kiếm lưới của một chuỗi thời gian đơn biến, khám phá bộ tham số của mô hình SARIMA cho kết quả với độ đo RMSE (giữa giá trị thực và kết quả dự báo của mô hình trên dữ liệu kiểm tra) là thấp nhất. Trước tiên, dữ liệu của chuỗi thời gian được chia thành 2 phần dùng huấn luyện và đánh giá theo tỷ lệ 80%, 20%. Kế đến, phương pháp đánh giá hiệu suất bộ tham số mô hình là sử dụng chiến lược kiểm tra dự báo gộp đầu, hay còn gọi là chiến lược

kiểm tra góid đầu (walk forward validation) để kiểm tra, đánh giá các mô hình dự báo theo tổ hợp 7 tham số của mô hình SARIMA. Mô hình với bộ tham số cho kết quả đánh giá hiệu suất theo độ đo RMSE (root mean squared error – căn của trung bình sai số bình phương) thấp nhất được lựa chọn.

Mã giả thuật toán tìm kiếm lưới bộ tham số SARIMA

Train = 80% dữ liệu đầu của chuỗi

Test = 20% dữ liệu cuối của chuỗi, dùng đánh giá mô hình

Scores = []

Configs = Tạo tổ hợp các bộ tham số (p,d,q)(P,D,Q)m

For cfg in Configs:

RMSE = WalkforwardValidate(cfg, Train, Test)

Scores.Append([RMSE, cfg])

Sắp xếp Scores với RMSE tăng dần

Return Scores[0]

Hình 3.3: Mã giả thuật toán tìm kiếm lưới khám phá bộ tham số SARIMA cho một chuỗi thời gian đơn biến

Phương pháp tìm kiếm lưới thực chất là vét cạn tổ hợp các 7 tham số (p,d,q)(P,D,Q)m của mô hình SARIMA. Do đó, tốc độ tìm kiếm là một điểm yếu của phương pháp nhưng ưu điểm là không đòi hỏi kỹ năng phân phân tích.

Thực hiện tìm kiếm lưới để lựa các tham số (p,d,q)(P,D,Q)m của mô hình SARIMA trên dữ liệu điện thương phẩm của 9 Điện lực. Kết quả cho thấy, đối với dữ liệu Điện lực TPTN bộ tham số mô hình SARIMA [(1, 1, 1), (2, 0, 2, 12)] cho kết quả RMSE thấp nhất, đối với dữ liệu Điện lực Gò Dầu bộ tham số mô hình SARIMA [(2, 0, 2), (2, 0, 2, 12)] cho kết quả RMSE thấp nhất, đối với dữ liệu Điện lực Tân Châu bộ tham số mô hình SARIMA [(1, 0, 2), (2, 1, 1, 12)] cho kết quả RMSE thấp nhất...

Bảng 3.1: Các giá trị RMSE nhỏ nhất từng mô hình thực hiện tìm kiếm lưới trên dữ liệu 09 Điện lực

TPTN		TANCHAU		TANBIEN	
Tham số	RMSE	Tham số	RMSE	Tham số	RMSE
[(1, 1, 1), (2, 0, 2, 12)]	783165	[(1, 0, 2), (2, 1, 1, 12)]	819366	[(1, 1, 0), (2, 1, 2, 12)]	328105
[(2, 1, 1), (2, 1, 1, 12)]	793720	[(2, 0, 1), (2, 1, 2, 12)]	823334	[(1, 0, 2), (2, 1, 2, 12)]	329235
[(2, 1, 2), (2, 1, 1, 12)]	794562	[(1, 0, 2), (1, 1, 2, 12)]	829812	[(1, 0, 1), (2, 1, 2, 12)]	330301
[(2, 1, 1), (2, 1, 0, 12)]	796303	[(2, 0, 1), (2, 1, 1, 12)]	848946	[(2, 1, 0), (2, 1, 2, 12)]	337615
[(2, 1, 2), (2, 1, 0, 12)]	796786	[(1, 0, 1), (2, 1, 2, 12)]	850222	[(2, 1, 2), (2, 1, 2, 12)]	339748
GODAU		CHAUTHANH		HOATHANH	
Tham số	RMSE	Tham số	RMSE	Tham số	RMSE
[(2, 0, 2), (2, 0, 2, 12)]	478690	[(2, 0, 1), (2, 0, 2, 12)]	483713	[(2, 1, 0), (2, 0, 2, 12)]	551554
[(2, 0, 1), (2, 1, 2, 12)]	487561	[(1, 0, 2), (2, 0, 2, 12)]	515515	[(2, 1, 0), (1, 1, 2, 12)]	556595
[(1, 0, 1), (2, 1, 2, 12)]	488999	[(0, 0, 0), (2, 0, 0, 12)]	526794	[(2, 0, 1), (1, 1, 2, 12)]	562009
[(1, 1, 0), (2, 1, 2, 12)]	489245	[(1, 1, 2), (2, 1, 2, 12)]	528735	[(2, 0, 2), (1, 1, 2, 12)]	564761
[(1, 0, 2), (2, 1, 2, 12)]	491401	[(2, 1, 2), (2, 1, 2, 12)]	536713	[(2, 1, 2), (2, 0, 2, 12)]	565911
TRANGBANG		DUONGMINHCHAU		BENCAU	
Tham số	RMSE	Tham số	RMSE	Tham số	RMSE
[(1, 1, 1), (2, 1, 2, 12)]	472138	[(2, 0, 2), (1, 0, 2, 12)]	496186	[(0, 0, 0), (1, 0, 2, 12)]	306545
[(1, 1, 2), (2, 1, 2, 12)]	483052	[(2, 0, 2), (2, 0, 2, 12)]	496191	[(0, 0, 0), (2, 0, 1, 12)]	306754
[(2, 1, 0), (2, 1, 2, 12)]	483916	[(1, 0, 2), (2, 0, 2, 12)]	499870	[(0, 0, 0), (2, 0, 2, 12)]	306844
[(2, 0, 0), (2, 1, 2, 12)]	486558	[(1, 0, 2), (1, 0, 2, 12)]	499961	[(0, 0, 0), (2, 0, 0, 12)]	307877
[(1, 1, 0), (2, 1, 2, 12)]	489117	[(0, 0, 2), (2, 0, 1, 12)]	500448	[(0, 0, 0), (1, 0, 1, 12)]	312251

3.3. Xây dựng mô hình dự báo LSTM

Thư viện Keras được sử dụng để xây dựng mô hình dự báo chuỗi thời gian bằng mạng LSTM. Đoạn mã Python sau đây là mô hình LSTM được xây dựng để dự báo nhu cầu điện thương phẩm trong luận văn. Mô hình sử dụng 1 tầng ẩn LSTM với số lượng tế bào là neurons. Trong đó, các tham số neurons, epochs là các tham số cần phải được xác định.

Mã Python xây dựng mô hình LSTM

```

model = Sequential()

model.add(LSTM(neurons,

                activation='tanh',

                batch_input_shape=(1, 1, 1),

                stateful=True))

```

```

model.add(Dense(1))

model.compile(loss='mean_squared_error', optimizer='adam')

for i in range(epochs):

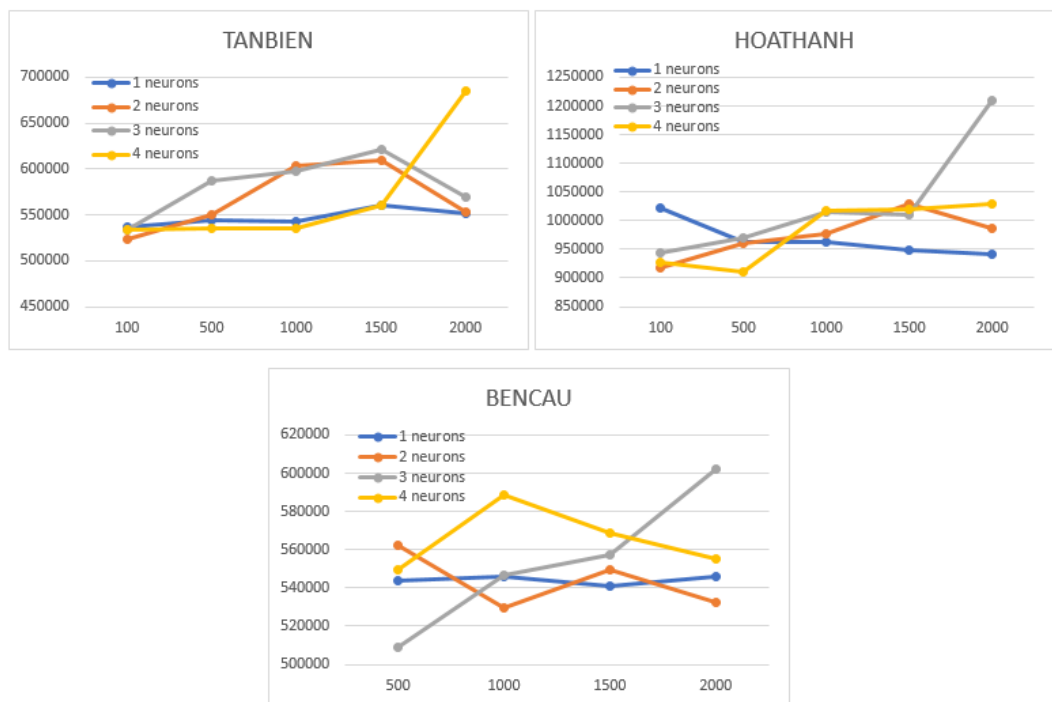
    model.fit(X, y, epochs=1, batch_size=1, shuffle=False)

    model.reset_states()

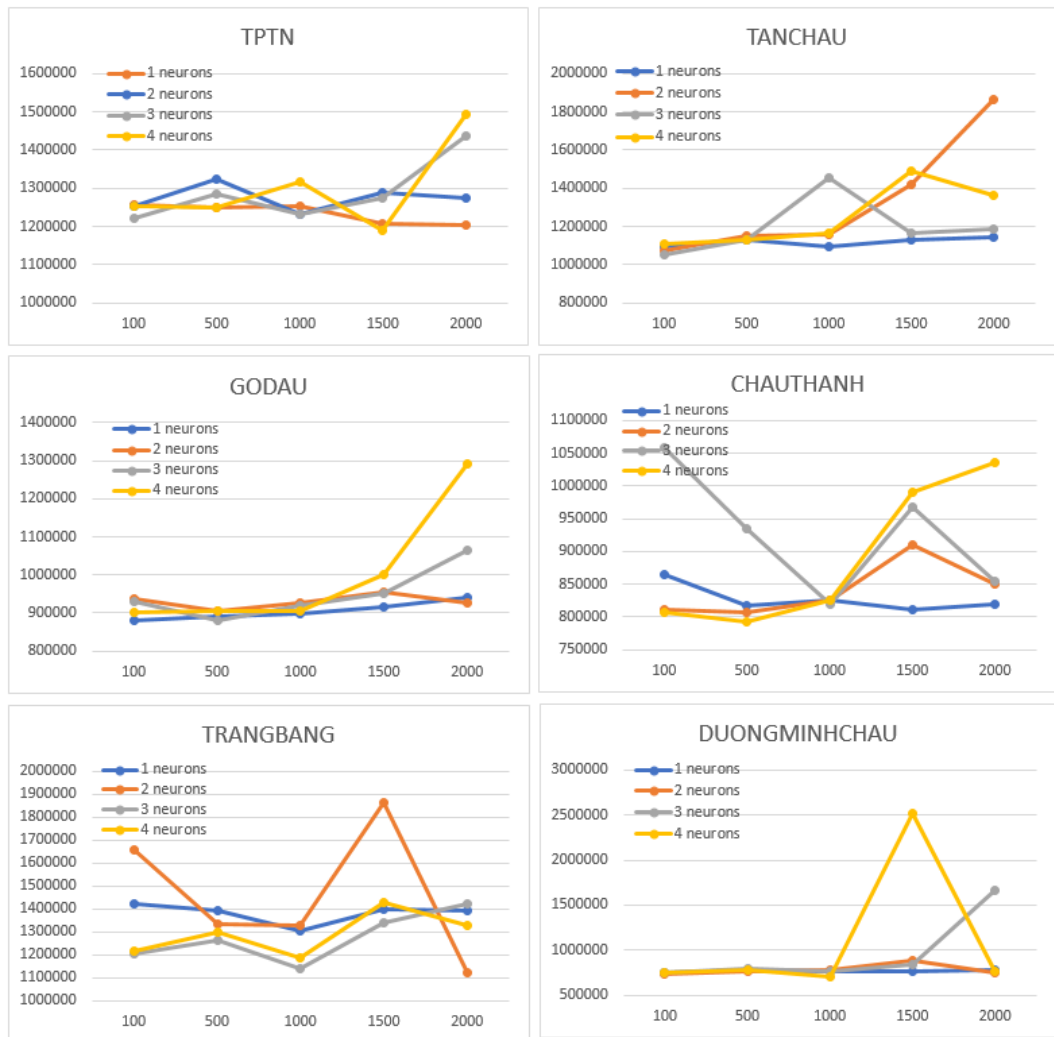
```

Hình 3.4: Đoạn mã Python xây dựng mô hình LSTM

Thuật toán tìm kiếm lưới được áp dụng để tìm kiếm giá trị neurons, epochs của mô hình LSTM phù hợp với dữ liệu của 9 Điện lực. Hình 3-5 là biểu đồ đánh giá hiệu suất mô hình LSTM tương ứng với giá trị neurons, epochs theo độ đo RMSE của 9 Điện lực. Mô hình mạng LSTM cho dữ liệu TPTN với số neurons là 4, số epochs là 1500 cho kết quả RMSE là thấp nhất, tương tự số neurons, epochs của mô hình mạng LSTM cho dữ liệu GODAU là 1, 100, TRANGBANG là 2, 2000, TANCHAU là 3, 100, CHAUTHANH là 4, 500, DUONGMINHCHAU là 4, 1000, TANBIEN là 2, 100, HOATHANH là 4, 500 và BENCAU là 3, 500.



Hình 3.5a: Biểu đồ giá trị RMSE các mô hình LSTM theo số neurons, epochs



Hình 3.5b: Biểu đồ RMSE mô hình LSTM theo số neurons, epochs

3.4. Xây dựng mô hình dự báo Prophet

Thư viện fbprophet được sử dụng để xây dựng mô hình dự báo.

Đoạn mã Python xây dựng mô hình dự báo

```

model = Prophet(
    seasonality_mode = 'multiplicative',
    seasonality_prior_scale = 0.05)
model.fit(data)

```

Hình 3.6: Đoạn mã Python mô hình dự báo Prophet

Prophet có rất ít tham số có thể tinh chỉnh để nâng cao hiệu suất của mô hình, các tham số hầu hết được Prophet xác định tự động. Tuy nhiên, Prophet vẫn cung cấp một vài tham số để hiệu chỉnh mô hình phù hợp với dữ liệu, trong đó có 2 tham số liên quan đến tính chu kỳ theo mùa của dữ liệu chuỗi thời gian đó là:

- *seasonality mode*: có 2 giá trị là additive, multiplicative tương ứng với sự thay đổi của thành phần theo mùa $s(t)$ là tuyến tính hay phi tuyến theo thời gian.
- *seasonality prior scale*: giá trị chỉ mức độ ảnh hưởng thành phần theo mùa $s(t)$ của mô hình. Thông thường có giá trị trong đoạn $[0.01, 10]$

Thuật toán tìm kiếm lưới được áp dụng để tìm kiếm giá trị phù hợp cho 2 tham số này đối với từng chuỗi thời gian. Bảng 3-2 liệt kê kết quả của quá trình thực hiện tìm kiếm lưới trên dữ liệu từ tháng 07/2020 về trước của 9 Điện lực để tìm mô hình Prophet phù hợp với từng dữ liệu.

Bảng 3.1: Các giá trị RMSE khi đánh giá lựa chọn mô hình tương ứng dữ liệu 09 Điện lực

TPTN			TANCHAU			TANBIEN		
seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE
multiplicative	0.05	758970.9	multiplicative	0.05	784137.1	multiplicative	0.05	359507.8
multiplicative	0.1	771802.1	multiplicative	0.1	793558	multiplicative	0.1	362851.4
multiplicative	0.5	795136.7	multiplicative	0.05	827322.5	multiplicative	10	368051.2
multiplicative	10	797153.4	additive	0.1	836975.6	multiplicative	0.5	370449.1
multiplicative	0.01	801381.1	additive	0.5	838103	additive	1	370684.1
GODAU			CHAUTHANH			HOATHANH		
seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE
multiplicative	0.05	530512.6	multiplicative	0.05	584272	multiplicative	0.01	597855.6
multiplicative	0.01	539559.1	multiplicative	0.1	591331.4	multiplicative	0.05	602941.1
multiplicative	5	552324.8	multiplicative	0.01	610002.8	multiplicative	0.1	610043.4
multiplicative	1	552771.6	multiplicative	0.05	637404.8	multiplicative	0.5	631204.3
multiplicative	0.5	556422	multiplicative	0.5	644871.7	multiplicative	1	631279.1
TRANGBANG			DUONGMINHCHAU			BENCAU		
seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE	seasonality mode	seasonality prior scale	RMSE
additive	0.05	1165554	multiplicative	0.05	605918.8	multiplicative	0.05	346294.9
additive	0.5	1180200	multiplicative	0.1	612524.9	multiplicative	0.1	356346.1
additive	0.1	1181139	multiplicative	0.01	629625.8	multiplicative	0.01	358076.2
additive	1	1184882	multiplicative	0.5	629873	multiplicative	0.05	364669.3
multiplicative	10	1189206	multiplicative	10	633805.8	multiplicative	0.01	375704.8

CHƯƠNG 4: KẾT QUẢ DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN Ở CÔNG TY ĐIỆN LỰC TÂY NINH

Chương này trình bày thực nghiệm các kết quả dự báo nhu cầu phụ tải ở Công ty Điện lực Tây Ninh và nhận xét chung.

4.1. Kết quả dự báo

4.1.1. Dự báo nhu cầu phụ tải 1 tháng

4.1.1.1. Mô hình Prophet

Sử dụng mô hình Prophet tìm được để thực hiện dự báo điện thương phẩm của 9 Điện lực từng tháng từ tháng 8/2020 đến tháng 7/2021. Bảng 4-1 là sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo tương ứng; Bảng 4-2 là kết quả RMSE của dữ liệu đánh giá và kết quả dự báo.

Bảng 4.1: Giá trị RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai dự báo theo mô hình Prophet.

Điện lực	RMSE	MAPE
TPTN	872,883	4.3%
GODAU	495,233	2.8%
TRANGBANG	826,265	4.6%
TANCHAU	929,143	6.0%
CHAUTHANH	734,048	3.9%
DUONGMINHCHAU	516,715	3.9%
TANBIEN	408,598	3.9%
HOATHANH	592,019	3.8%
BENCAU	448,668	3.4%

Bảng 4.2: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 điện lực

TPTN				TANCHAU			TANBIEN		
ts	actual	forecast	%	actual	forecast	%	actual	forecast	%
2020-08	18207082	18291356	0.46	12173579	12678891	4.15	8481386	8689160	2.45
2020-09	19142134	18482938	-3.44	12892779	12817088	-0.59	9044726	8739016	-3.38
2020-10	17383526	17715748	1.91	11550918	11995113	3.85	8015818	8232981	2.71
2020-11	16709685	18403618	10.14	11447254	13123573	14.64	7865159	8577322	9.05
2020-12	17086007	18110154	5.99	11820068	13425913	13.59	7896196	8409948	6.51
2021-01	17671275	18397274	4.11	13458661	14366592	6.75	8270063	8642332	4.50
2021-02	16865523	18161468	7.68	12730158	13638046	7.13	8226755	8273196	0.56
2021-03	17164028	18014600	4.96	12701612	13413122	5.60	8463106	8307946	-1.83
2021-04	21021675	20708698	-1.49	15790972	14386571	-8.89	9832594	9386973	-4.53
2021-05	19665283	20266346	3.06	13293759	13300975	0.05	9171984	9324639	1.66
2021-06	20556615	19979122	-2.81	13187777	12901761	-2.17	9545982	9395662	-1.57
2021-07	19586019	18611316	-4.98	12945538	12394651	-4.26	9694522	8895002	-8.25
GODAU				CHAUTHANH			HOATHANH		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
14062419	14054283	-0.06	12939924	12877421	-0.48	14690038	14373148	-2.16	
14497508	14062921	-3.00	13460402	13120362	-2.53	15317514	14533722	-5.12	
13568176	13478036	-0.66	12326056	12389135	0.51	13699521	14067359	2.69	
13204714	13942261	5.59	12198107	12811103	5.03	13660617	14556545	6.56	
13309134	13814234	3.80	12444879	12819391	3.01	13980696	14265798	2.04	
13399435	13631921	1.74	12749806	13146413	3.11	13978412	14491178	3.67	
12765172	12882527	0.92	12506674	12596762	0.72	12898776	13836359	7.27	
12467077	12900983	3.48	12062610	12760440	5.79	13062727	13618585	4.26	
16070697	15352422	-4.47	15336431	14115022	-7.96	15991700	15697893	-1.84	
15254942	15421500	1.09	14132649	14070319	-0.44	15469510	15725517	1.65	
16198569	15638492	-3.46	15299573	14286297	-6.62	16472396	15743098	-4.43	
15528940	14615996	-5.88	14830655	13202298	-10.98	15575638	14995034	-3.73	
TRANGBANG				DUONGMINHCHAU			BENCAU		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
16323277	16868294	3.34	10963565	10519935	-4.05	8343666	8012104	-3.97	
16618422	16683059	0.39	11370927	10610269	-6.69	8684419	8655597	-0.33	
15469593	16509531	6.72	10394945	10385958	-0.09	7786101	7747093	-0.50	
15585036	16381276	5.11	10154700	10951484	7.85	7560714	7908721	4.60	
15504467	16370734	5.59	10636059	10987780	3.31	7809335	7862100	0.68	
15552472	16267010	4.59	11057796	10982800	-0.68	8319252	8374961	0.67	
14309219	14654667	2.41	10775014	11004738	2.13	8149380	8271789	1.50	
14437510	15651243	8.41	10568107	11202179	6.00	8060942	8511260	5.59	
17820239	17388974	-2.42	13011822	12237610	-5.95	9697892	9389158	-3.18	
18032939	17314894	-3.98	11827368	12328543	4.24	9127568	9264995	1.51	
19150348	17847060	-6.81	12021614	11881602	-1.16	10178943	9602848	-5.66	
18177277	17216924	-5.28	11599403	11016961	-5.02	10090874	8861041	-12.19	

4.1.1.2. Mô hình SARIMA

Thực hiện áp dụng mô hình SARIMA tìm được để dự báo điện thương phẩm từng tháng từ tháng (từ tháng 8/2020 đến tháng 07/2021) thuộc 9 Điện lực. Bảng 4-3 là kết quả RMSE của dữ liệu đánh giá và kết quả dự báo, Bảng 4-4 là sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo tương ứng.

Bảng 4.3: Độ đo RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai mô hình SARIMA.

Điện lực	RMSE	MAPE
TPTN	1,003,489	4.3%
GODAU	573,863	3.4%
TRANGBANG	698,122	3.4%
TANCHAU	722,059	4.1%
CHAUTHANH	679,315	4.0%
DUONGMINHCHAU	678,385	4.7%
TANBIEN	471,773	4.5%
HOATHANH	633,737	3.8%
BENCAU	541,012	4.9%

Bảng 4.4: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 mô hình tương ứng 09 Điện lực.

TPTN				TANCHAU			TANBIEN		
ts	actual	forecast	%	actual	forecast	%	actual	forecast	%
2020-08	18207082	19142276	5.14	12173579	12113056	-0.50	8481386	8407014	-0.88
2020-09	19142134	18279396	-4.51	12892779	12515401	-2.93	9044726	8223434	-9.08
2020-10	17383526	18175360	4.56	11550918	11602373	0.45	8015818	8431913	5.19
2020-11	16709685	18746306	12.19	11447254	12401652	8.34	7865159	8612644	9.50
2020-12	17086007	17579844	2.89	11820068	12619588	6.76	7896196	8068946	2.19
2021-01	17671275	17666968	-0.02	13458661	13748411	2.15	8270063	7963083	-3.71
2021-02	16865523	17883926	6.04	12730158	13348657	4.86	8226755	8488751	3.18
2021-03	17164028	17309016	0.84	12701612	12851016	1.18	8463106	8220742	-2.86
2021-04	21021675	19608066	-6.72	15790972	14150202	-10.39	9832594	9525580	-3.12
2021-05	19665283	19558528	-0.54	13293759	13613688	2.41	9171984	9471844	3.27
2021-06	20556615	20655988	0.48	13187777	13146051	-0.32	9545982	9696048	1.57
2021-07	19586019	18047014	-7.86	12945538	11817458	-8.71	9694522	8799671	-9.23
GODAU				CHAUTHANH			HOATHANH		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
18207082	18910556	3.86	12939924	12889094	-0.39	14690038	15123373	2.95	
19142134	18014492	-5.89	13460402	12976749	-3.59	15317514	14580953	-4.81	
17383526	17449704	0.38	12326056	12762646	3.54	13699521	14142590	3.23	
16709685	18969738	13.53	12198107	12943175	6.11	13660617	14844982	8.67	
17086007	17477430	2.29	12444879	12386365	-0.47	13980696	13770106	-1.51	
17671275	18344668	3.81	12749806	13252428	3.94	13978412	14129969	1.08	
16865523	18094460	7.29	12506674	12744439	1.90	12898776	13556491	5.10	
17164028	17477366	1.83	12062610	12790541	6.03	13062727	13345775	2.17	
21021675	20560926	-2.19	15336431	14578661	-4.94	15991700	15236701	-4.72	
19665283	19631726	-0.17	14132649	14648813	3.65	15469510	15913072	2.87	
20556615	20669408	0.55	15299573	14906293	-2.57	16472396	16014951	-2.78	
19586019	17956580	-8.32	14830655	13182139	-11.12	15575638	14609155	-6.21	
TRANGBANG				DUONGMINHCHAU			BENCAU		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
16323277	16417853	0.58	10963565	11119923	1.43	8343666	8072990	-3.24	
16618422	16369578	-1.50	11370927	10601951	-6.76	8684419	8455745	-2.63	
15469593	16265432	5.14	10394945	10368020	-0.26	7786101	7778158	-0.10	
15585036	16037673	2.90	10154700	11543133	13.67	7560714	8431789	11.52	
15504467	14560685	-6.09	10636059	10836644	1.89	7809335	7952574	1.83	
15552472	15544982	-0.05	11057796	10900278	-1.42	8319252	8769214	5.41	
14309219	14363963	0.38	10775014	11325481	5.11	8149380	8819692	8.23	
14437510	15927886	10.32	10568107	10802856	2.22	8060942	8517618	5.67	
17820239	17229582	-3.31	13011822	11821433	-9.15	9697892	9561350	-1.41	
18032939	17025730	-5.59	11827368	12416399	4.98	9127568	9646742	5.69	
19150348	18665928	-2.53	12021614	12232250	1.75	10178943	10053663	-1.23	
18177277	17701488	-2.62	11599403	10735278	-7.45	10090874	8889543	-11.91	

4.1.1.4. Mô hình LSTM

Bảng dữ liệu sau đây liệt kê kết quả thực hiện đánh giá 09 mô hình tìm được trên dữ liệu 12 tháng cuối (dữ liệu đánh giá) của 09 Điện lực tương ứng. Bảng 4-2 là kết quả RMSE của dữ liệu đánh giá và kết quả dự báo, Bảng 4-3 là sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo tương ứng.

Bảng 4.5: Độ đo RMSE giữa dữ liệu đánh giá và dự báo của 09 Điện lực khi triển khai mô hình mạng LSTM.

Điện lực	RMSE	MAPE
TPTN	1,245,483	4.65%
GODAU	1,008,592	5.54%
TRANGBANG	1,260,033	5.60%
TANCHAU	1,125,511	6.07%
CHAUTHANH	1,280,931	7.09%
DUONGMINHCHAU	731,094	4.56%
TANBIEN	576,918	5.08%
HOATHANH	932,367	5.08%
BENCAU	542,478	5.24%

Bảng 4.6: Sai số (%) giữa dữ liệu 12 tháng và kết quả dự báo từng tháng tương ứng của 09 mô hình tương ứng 09 Điện lực.

TPTN				TANCHAU			TANBIEN		
ts	actual	forecast	%	actual	forecast	%	actual	forecast	%
2020-08	18207082	18156858	-0.28	12173579	12095455	-0.64	8481386	8704347	2.63
2020-09	19142134	17945988	-6.25	12892779	12485506	-3.16	9044726	8585652	-5.08
2020-10	17383526	18259126	5.04	11550918	12885686	11.56	8015818	8792464	9.69
2020-11	16709685	17610752	5.39	11447254	12140912	6.06	7865159	8501581	8.09
2020-12	17086008	17399418	1.83	11820068	11891497	0.60	7896196	8280783	4.87
2021-01	17671276	17187046	-2.74	13458661	12064042	-10.36	8270063	8118396	-1.83
2021-02	16865524	17324690	2.72	12730158	12911975	1.43	8226755	8161307	-0.80
2021-03	17164028	17456218	1.70	12701612	12763318	0.49	8463106	8215272	-2.93
2021-04	21021676	17271780	-17.84	15790972	12792859	-18.99	9832594	8322193	-15.36
2021-05	19665284	20456330	4.02	13293759	14314246	7.68	9171984	9361946	2.07
2021-06	20556616	20087094	-2.28	13187777	13617370	3.26	9545982	9274419	-2.84
2021-07	19586020	20003442	2.13	12945538	13448621	3.89	9694522	9382151	-3.22
GODAU				CHAUTHANH			HOATHANH		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
14062419	13382058	-4.84	12939924	12398052	-4.19	14690038	14728414	0.26	
14497508	14158083	-2.34	13460402	12707919	-5.59	15317514	14667629	-4.24	
13568176	14082355	3.79	12326056	13123605	6.47	13699521	14584205	6.46	
13204714	13773066	4.30	12198107	12691901	4.05	13660617	14333994	4.93	
13309134	13863221	4.16	12444879	12610718	1.33	13980696	14357791	2.70	
13399435	13813432	3.09	12749806	12562463	-1.47	13978412	13961574	-0.12	
12765172	13527809	5.97	12506674	12608893	0.82	12898776	13812665	7.09	
12467077	13253608	6.31	12062610	12583537	4.32	13062727	13870869	6.19	
16070697	13215044	-17.77	15336431	12441652	-18.88	15991700	13735303	-14.11	
15254942	16393966	7.47	14132649	15495553	9.64	15469510	16197980	4.71	
16198569	15840376	-2.21	15299573	15772069	3.09	16472396	15261621	-7.35	
15528940	16138426	3.92	14830655	12157694	-18.02	15575638	15805918	1.48	
TRANGBANG				DUONGMINHCHAU			BENCAU		
actual	forecast	%	actual	forecast	%	actual	forecast	%	
16323277	16225126	-0.60	10963565	10795496	-1.53	8343666	7824480	-6.22	
16618422	16698544	0.48	11370927	10937358	-3.81	8684419	8555495	-1.48	
15469593	16592269	7.26	10394945	10903313	4.89	7786101	8539423	9.68	
15585036	15844151	1.66	10154700	10596757	4.35	7560714	7938899	5.00	
15504467	15525153	0.13	10636059	10344985	-2.74	7809335	8043190	2.99	
15552472	15363715	-1.21	11057796	10622557	-3.94	8319252	7848610	-5.66	
14309219	15300289	6.93	10775014	10749287	-0.24	8149380	8074312	-0.92	
14437510	14680741	1.68	10568107	10829715	2.48	8060942	8521193	5.71	
17820240	14372201	-19.35	13011822	10839051	-16.70	9697892	8455919	-12.81	
18032940	15853174	-12.09	11827368	12340752	4.34	9127568	9528923	4.40	
19150348	16732131	-12.63	12021614	12637265	5.12	10178943	9609097	-5.60	
18177276	17577970	-3.30	11599403	11406738	-1.66	10090874	9978063	-1.12	

4.1.2. Dự báo nhu cầu phụ tải 12 tháng

4.1.2.1. Mô hình Prophet

Áp dụng mô hình Prophet dự báo nhu cầu điện thương phẩm 12 tháng tới (từ tháng 08/2020 đến tháng 07/2021) trên dữ liệu 9 Điện lực. Theo đó, dữ liệu 12 tháng (từ tháng 08/2020 đến tháng 07/2021) được dùng làm dữ liệu kiểm tra, đánh giá mô hình và dữ liệu trước tháng 08/2020 dùng làm dữ liệu huấn luyện mô hình. Bảng 4-10 liệt kê kết quả triển khai dự báo theo mô hình Prophet điện thương phẩm 12 tháng tới các khách hàng dưới 12 triệu kWh/năm thuộc 09 Điện lực và độ đo RMSE của dữ liệu đánh giá và kết quả dự báo.

Bảng 4.7: Các giá trị RMSE của dữ liệu 12 tháng dùng đánh giá và kết quả dự báo của mô hình Prophet tương ứng của 09 Điện lực.

Điện lực	RMSE	MAPE
TPTN	916,689	4.3%
GODAU	472,559	2.8%
TRANGBANG	923,371	5.0%
TANCHAU	964,092	6.2%
CHAUTHANH	728,420	3.9%
DUONGMINHCHAU	497,566	3.8%
TANBIEN	400,620	3.9%
HOATHANH	578,297	3.7%
BENCAU	445,886	3.3%

4.1.2.2. Mô hình SARIMA

Áp dụng mô hình SARIMA dự báo nhu cầu điện thương phẩm 12 tháng tới (từ tháng 08/2020 đến tháng 07/2021) trên dữ liệu 9 Điện lực. Theo đó, dữ liệu 12 tháng (từ tháng 08/2020 đến tháng 07/2021) được dùng làm dữ liệu kiểm tra, đánh giá mô hình và dữ liệu trước tháng 08/2020 dùng làm dữ liệu huấn luyện mô hình. Bảng 4-8 liệt kê kết quả triển khai dự báo theo mô hình SARIMA điện thương phẩm 12 tháng tới các khách hàng dưới 12 triệu kWh/năm thuộc 9 Điện lực và các giá trị RMSE của dữ liệu đánh giá và kết quả dự báo 12 tháng tới.

Bảng 4.8: Các giá trị RMSE của dữ liệu 12 tháng dùng đánh giá và kết quả dự báo của mô hình SARIMA tương ứng của 09 Điện lực.

Điện lực	RMSE	MAPE
TPTN	1,488,488	7.5%
GODAU	541,559	3.1%
TRANGBANG	667,959	3.0%
TANCHAU	739,631	4.4%
CHAUTHANH	635,139	3.5%
DUONGMINHCHAU	617,175	4.7%
TANBIEN	410,202	3.6%
HOATHANH	830,750	5.1%
BENCAU	537,921	4.9%

4.2. So sánh kết quả khi áp dụng các mô hình Prophet, SARIMA, LSTM

So sánh kết quả dự báo nhu cầu điện thương phẩm 12 tháng tới của 2 mô hình SARIMA, Prophet trên dữ liệu 9 Điện lực. Với dữ liệu TPTN, mô hình Prophet có giá trị RMSE là 872882.90 và giá trị MAPE là 4.10%, mô hình SARIMA có giá trị RMSE là 1003489.30 và giá trị MAPE là 4.25%, mô hình LSTM có giá trị RMSE là 1245483.5 và giá trị MAPE là 4.65%. 6/9 điện lực có kết quả mô hình Prophet cho kết quả tốt hơn mô hình SARIMA, mô hình LSTM cho kết quả thấp nhất. 3/9 điện lực cho kết quả SARIMA tốt hơn mô hình Prophet.

Bảng 4.9: So sánh hiệu suất dự báo từng tháng của 3 mô hình SARIMA, Prophet và LSTM

Điện lực	SARIMA		PROPHET		LSTM	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
TPTN	1003489.30	4.25%	872882.90	4.10%	1245483.5	4.65%
GODAU	573862.80	3.42%	495233.00	2.86%	1008592.5	5.54%
TRANGBANG	698121.84	3.39%	826265.30	4.51%	1260033.3	5.60%
TANCHAU	722059.40	4.14%	929143.10	5.65%	1125511.0	6.07%
CHAUTHANH	679314.90	4.08%	734047.90	4.09%	1280931.0	7.09%
DUONGMINHCHAU	678384.90	4.66%	516714.70	3.94%	731094.1	4.56%
TANBIEN	471772.80	4.55%	408598.30	3.89%	576917.5	5.08%
HOATHANH	633737.00	3.82%	592018.60	3.73%	932366.9	5.08%
BENCAU	541011.70	4.83%	448667.50	3.51%	542477.7	5.24%

E



Hình 4.1: Biểu đồ so sánh kết quả dự báo từng tháng của 3 mô hình SARIMA, Prophet và mạng LSTM của 9 điện lực

So sánh kết quả dự báo nhu cầu điện thương phẩm 12 tháng tới của 2 mô hình SARIMA, Prophet trên dữ liệu 9 Điện lực. Kết quả cho thấy 2 mô hình cho kết quả không khác nhiều so với dự báo từng tháng. Mô hình Prophet vẫn cho kết quả dự báo

12 tháng tốt hơn mô hình SARIMA tại 7/9 điện lực, mô hình SARIMA cho kết quả dự báo tốt hơn mô hình Prophet tại 2/9 điện lực.

Bảng 4.10: So sánh hiệu suất dự báo 12 tháng của 2 mô hình SARIMA và Prophet

Điện lực	SARIMA		PROPHET	
	RMSE	MAPE	RMSE	MAPE
TPTN	1488487.6	7.5%	916688.5	4.3%
GODAU	541558.6	3.1%	472558.5	2.8%
TRANGBANG	667958.7	3.0%	923371.0	5.0%
TANHAU	739631.3	4.4%	964092.0	6.2%
CHAUHANH	635139.3	3.5%	728420.3	3.9%
DUONGMINHCHAU	617175.1	4.7%	497565.9	3.8%
TANBIEN	410201.9	3.6%	400619.9	3.9%
HOATHANH	830750.1	5.1%	578297.2	3.7%
BENCAU	537921.2	4.9%	445886.1	3.3%

4.3. Nhận xét chung

Kết quả thử nghiệm mô hình SARIMA, Prophet và mạng LSTM dự báo điện thương phẩm 12 tháng của dữ liệu điện thương phẩm thuộc 9 điện lực cho thấy mô hình Prophet cho kết quả dự báo với hiệu suất tốt hơn mô hình SARIMA, và mô hình mạng LSTM cho kết quả với hiệu suất thấp nhất (giá trị RMSE lớn nhất).

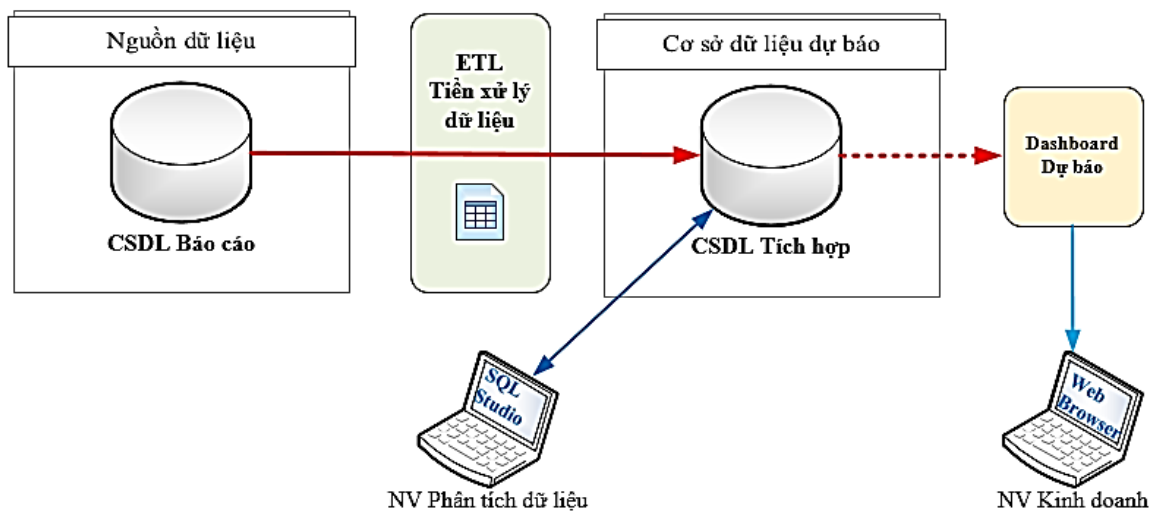
CHƯƠNG 5: XÂY DỰNG HỆ THỐNG DỰ BÁO NHU CẦU PHỤ TẢI ĐIỆN

Chương này giới thiệu hệ thống dự báo nhu cầu phụ tải điện hàng tháng ứng dụng các mô hình dự báo phụ tải điện đã nghiên cứu.

5.1. Thiết kế hệ thống

5.1.1. Mô hình hệ thống

Kiến trúc tổng thể mô hình hệ thống dự báo phụ tải điện khách hàng tại Công ty Điện lực Tây Ninh được thể hiện như Hình 5.1 dưới đây. Trong đó, nguồn dữ liệu lấy từ cơ sở dữ liệu (CSDL) Báo cáo nơi lưu trữ dữ liệu hóa đơn tất cả các khách hàng sử dụng điện. Dữ liệu được trích xuất từ nguồn qua tiến trình tiền xử lý được tổng hợp thành dữ liệu phụ tải điện khách hàng hàng tháng lưu vào CSDL Tích hợp phục vụ chức năng phân tích, dự báo phụ tải điện khách hàng hàng tháng.



Hình 5.1: Mô hình kiến trúc tổng thể

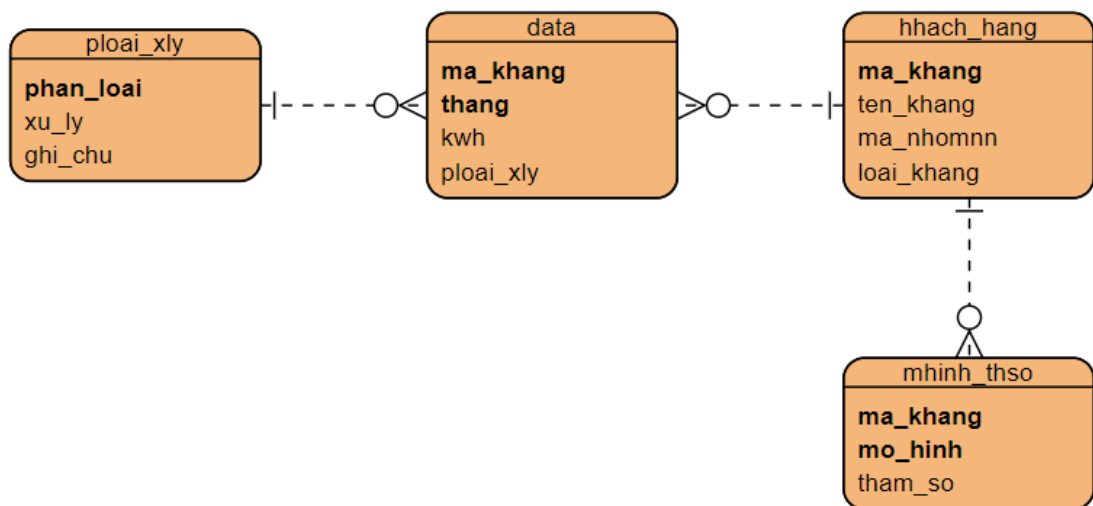
Hàng tháng (từ ngày 15-20 hàng tháng) nhân viên kinh doanh tại các Điện lực truy cập Dashboard tiến hành phân tích, dự báo phụ tải điện khách hàng tháng tới. Nhân viên phân tích dữ liệu chịu trách nhiệm thực hiện tiến trình ETL dữ liệu và đánh giá mô hình, cập nhật các bộ thông số các mô hình đối với từng dữ liệu phụ tải điện định kỳ, chu kỳ đánh giá lại là hàng năm.

Ngôn ngữ, thư viện và khung phần mềm được sử dụng trong hệ thống dự báo phụ tải điện khách hàng:

- Ngôn lập trình: Python
- Cơ sở dữ liệu: SQL Lite 3
- Thư viện xây dựng các mô hình dự báo (nền tảng Python): SARIMAX, Keras LSTM và Facebook Prophet.
- Khung xây dựng Dashboard phân tích, dự báo: Streamlit.

5.1.2. Thiết kế cơ sở dữ liệu cho hệ thống

Hệ quản trị cơ sở dữ liệu sử dụng để xây dựng hệ thống dữ liệu khách hàng là SQL Lite (version 3), đáp ứng đủ nhu cầu lưu trữ và truy xuất dữ liệu. Hình 5.2 là lược đồ ERD của cơ sở dữ liệu tích hợp.



Hình 5.2 Lược đồ ERD cơ sở dữ liệu dự báo

Dữ liệu mẫu của các bảng:

a) Bảng Phân loại xử lý (ploai_xly):

	phan_loai	xu_ly	ghi_chu
1	3	-1	Khách hàng lớn không còn mua điện (tác động đến dữ liệu dự báo)
2	2	1	Nhập chi
3	1	-1	Tách chi
4	0	1	Dữ liệu gốc

Hình 5.3: Dữ liệu bảng phân loại xử lý

b) Bảng Khách hàng (khach_hang):

	ma_khang	ten_khang	ma_nhomn	loai_khang
2	PB0501	Điện lực Thành phố Tây Ninh	NULL	0
3	PB0502	Điện lực Gò Dầu	NULL	0
4	PB0503	Điện lực Trảng Bàng	NULL	0
5	PB0504	Điện lực Tân Châu	NULL	0
6	PB0505	Điện lực Châu Thành	NULL	0
7	PB0506	Điện lực Dương Minh Châu	NULL	0
8	PB0507	Điện lực Tân Biên	NULL	0
9	PB0508	Điện lực Hòa Thành	NULL	0
10	PB0509	Điện lực Bến Cầu	NULL	0

Hình 5.4: Dữ liệu mẫu bảng khách hàng

c) Bảng Dữ liệu phụ tải điện (data):

	ma_khang	thang	kwh	phan_loai
1	PB0501	2009-01	7895062	0
2	PB0501	2009-02	7757958	0
3	PB0501	2009-03	8894204	0
4	PB0501	2009-04	9815589	0
5	PB0501	2009-05	8679552	0
6	PB0501	2009-06	8270427	0
7	PB0501	2009-07	8997132	0
8	PB0501	2009-08	8732236	0
9	PB0501	2009-09	9315298	0
10	PB0501	2009-10	8483531	0
11	PB0501	2009-11	8788366	0
12	PB0501	2009-12	8642656	0
13	PB0501	2010-01	10314525	0
14	PB0501	2010-02	8654880	0
15	PB0501	2010-03	9652584	0
16	PB0501	2010-04	10280127	0

Hình 5.5: Dữ liệu mẫu bảng dữ liệu phụ tải điện hàng tháng

d) Bảng Tham số các mô hình (mhinh_thso):

	ma_khang	mo_hinh	tham_so
1	PB0501	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
2	PB0501	SARIMA	{"order": [1, 1, 1], "seasonal_order": [2, 0, 2, 12]}
3	PB0501	LSTM	{"epochs": 1500, "neurons": 4}
4	PB0502	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
5	PB0502	SARIMA	{"order": [2, 0, 2], "seasonal_order": [2, 0, 2, 12]}
6	PB0502	LSTM	{"epochs": 100, "neurons": 1}
7	PB0503	PROPHET	{"seasonality_mode": "additive", "seasonality_prior_scale": 0.05}
8	PB0503	SARIMA	{"order": [1, 1, 1], "seasonal_order": [2, 1, 2, 12]}
9	PB0503	LSTM	{"epochs": 2000, "neurons": 2}
10	PB0504	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
11	PB0504	SARIMA	{"order": [1, 0, 2], "seasonal_order": [2, 1, 1, 12]}
12	PB0504	LSTM	{"epochs": 100, "neurons": 3}
13	PB0505	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
14	PB0505	SARIMA	{"order": [2, 0, 1], "seasonal_order": [2, 0, 2, 12]}
15	PB0505	LSTM	{"epochs": 500, "neurons": 4}
16	PB0506	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
17	PB0506	SARIMA	{"order": [2, 0, 2], "seasonal_order": [1, 0, 2, 12]}
18	PB0506	LSTM	{"epochs": 1000, "neurons": 4}
19	PB0507	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
20	PB0507	SARIMA	{"order": [1, 1, 0], "seasonal_order": [2, 1, 2, 12]}
21	PB0507	LSTM	{"epochs": 100, "neurons": 2}
22	PB0508	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.01}
23	PB0508	SARIMA	{"order": [2, 1, 0], "seasonal_order": [2, 0, 2, 12]}
24	PB0508	LSTM	{"epochs": 500, "neurons": 4}
25	PB0509	PROPHET	{"seasonality_mode": "multiplicative", "seasonality_prior_scale": 0.05}
26	PB0509	SARIMA	{"order": [0, 0, 0], "seasonal_order": [1, 0, 2, 12]}
27	PB0509	LSTM	{"epochs": 500, "neurons": 3}

Hình 5.6: Dữ liệu mẫu bảng tham số các mô hình

5.1.3. Xây dựng Dashboard phân tích, dự báo

Dashboard phân tích dự báo được xây dựng trên khung phần mềm Streamlit sử dụng ngôn ngữ lập trình Python. Hình 5.7 là đoạn mã nguồn Python truy xuất dữ liệu, thực hiện dự báo và hiển thị kết quả dự báo trực tiếp trên trình duyệt web thông dụng.

```

from datetime import date
import pandas as pd
import streamlit as st
import matplotlib.pyplot as plt
import matplotlib.dates as dates
import matplotlib.cbook as cbook
import json
import data
import prophet_model as prophet
import sarima_model as sarima
import lstm_model as lstm
import reporting_util as rp

don_vis = data.get_khang()

st.sidebar.subheader("DỰ BÁO PHỤ TẢI")
ngay_lv = st.sidebar.date_input("Ngày làm việc:", value=date(2021, 7, 1))
dv = st.sidebar.selectbox("Chọn đơn vị:",
                          don_vis,
                          format_func=lambda opt: don_vis[don_vis['ma_khang'] == opt].values[0, 1])
mh = st.sidebar.selectbox("Chọn mô hình:", ['PROPHET', 'SARIMA', 'LSTM'], index=0)
thang_db = st.sidebar.number_input("Số tháng dự báo:", min_value=1, max_value=50, step=1)
predict = st.sidebar.button("THỰC HIỆN")

thso_mhinh = data.get_thso_mhinh(mh, dv)

if thso_mhinh is not None:
    thso_mhinh = json.loads(thso_mhinh)

st.sidebar.write('THAM SỐ MÔ HÌNH:')
st.sidebar.write(thso_mhinh)

thang_lv = ngay_lv.strftime('%Y-%m')

raw_data = data.get_data(dv)
data = raw_data[raw_data['ds'] < thang_lv]
future_data = raw_data[raw_data['ds'] >= thang_lv]

st.title('DỰ BÁO PHỤ TẢI ĐIỆN KHÁCH HÀNG')
st.header('Biểu đồ phụ tải điện khách hàng')

fig, ax = plt.subplots(figsize=(15, 7.5))
ax.plot(data.index, data['y'] / 1000, label='MWh')
ax.xaxis.set_major_formatter(dates.DateFormatter('%Y-%m'))
ax.xaxis.set_major_locator(dates.MonthLocator(interval=6))
ax.xaxis.set_minor_locator(dates.MonthLocator())
ax.set_ylabel('Sàn lượng')
ax.grid(True)
fig.autofmt_xdate()
plt.legend(loc="upper left")
st.pyplot(fig)

with st.expander("Xem dữ liệu"):
    st.subheader('Dữ liệu lịch sử')
    st.dataframe(pd.DataFrame(data.values, columns=('Tháng', 'Sàn lượng (kWh)')))
    st.subheader('Dữ liệu tương lai')
    st.dataframe(pd.DataFrame(future_data.values, columns=('Tháng', 'Sàn lượng (kWh)')))

```

```

if predict:
    forecast = list()

    st.header('Kết quả dự báo')

    if mh == 'PROPHET':
        model = prophet.create_model(data, thso_mhinh)
        forecast = prophet.forecast(model, periods=thang_db)

    if mh == 'SARIMA':
        model = sarima.create_model(data, thso_mhinh)
        forecast = sarima.forecast(model, periods=thang_db)

    if mh == 'LSTM':
        model = lstm.create_model(data, thso_mhinh)
        forecast = lstm.forecast(model, periods=thang_db)

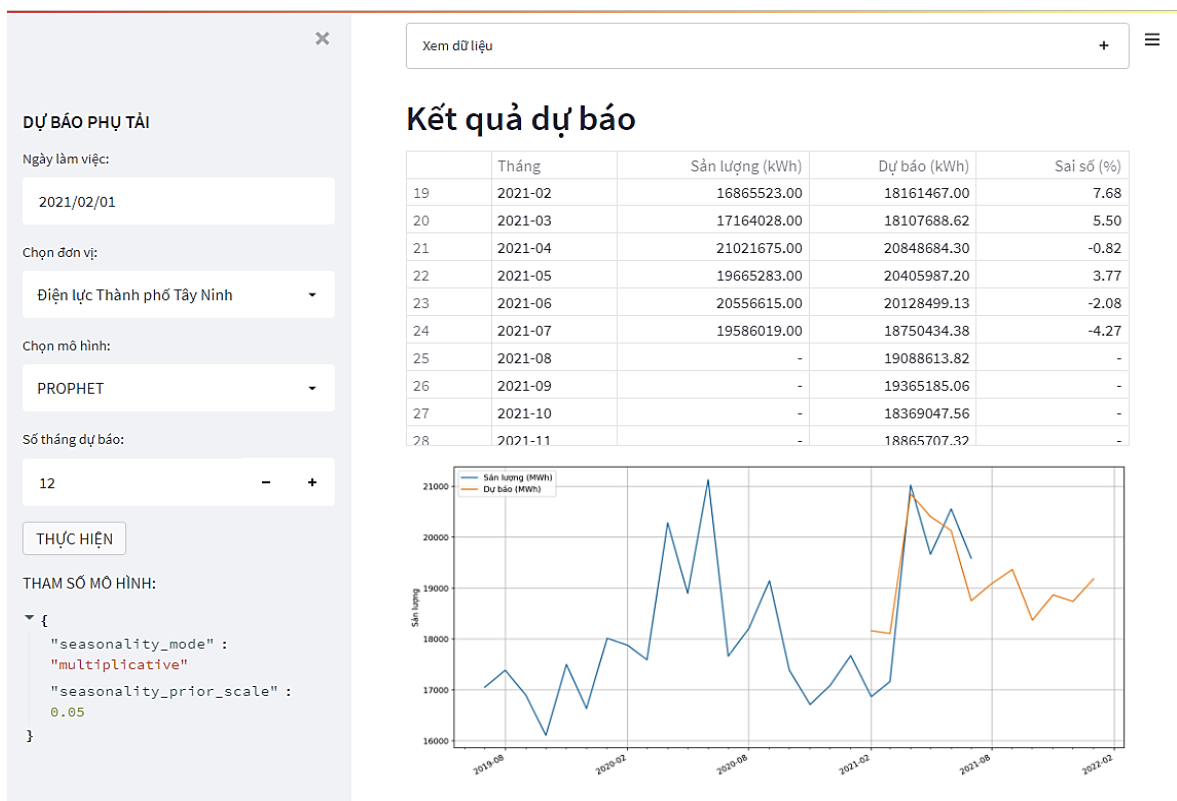
rp.show_report(st, data, future_data, forecast)

```

Hình 5.8: Mã nguồn Dashboard phân tích, dự báo

5.2. Kết quả

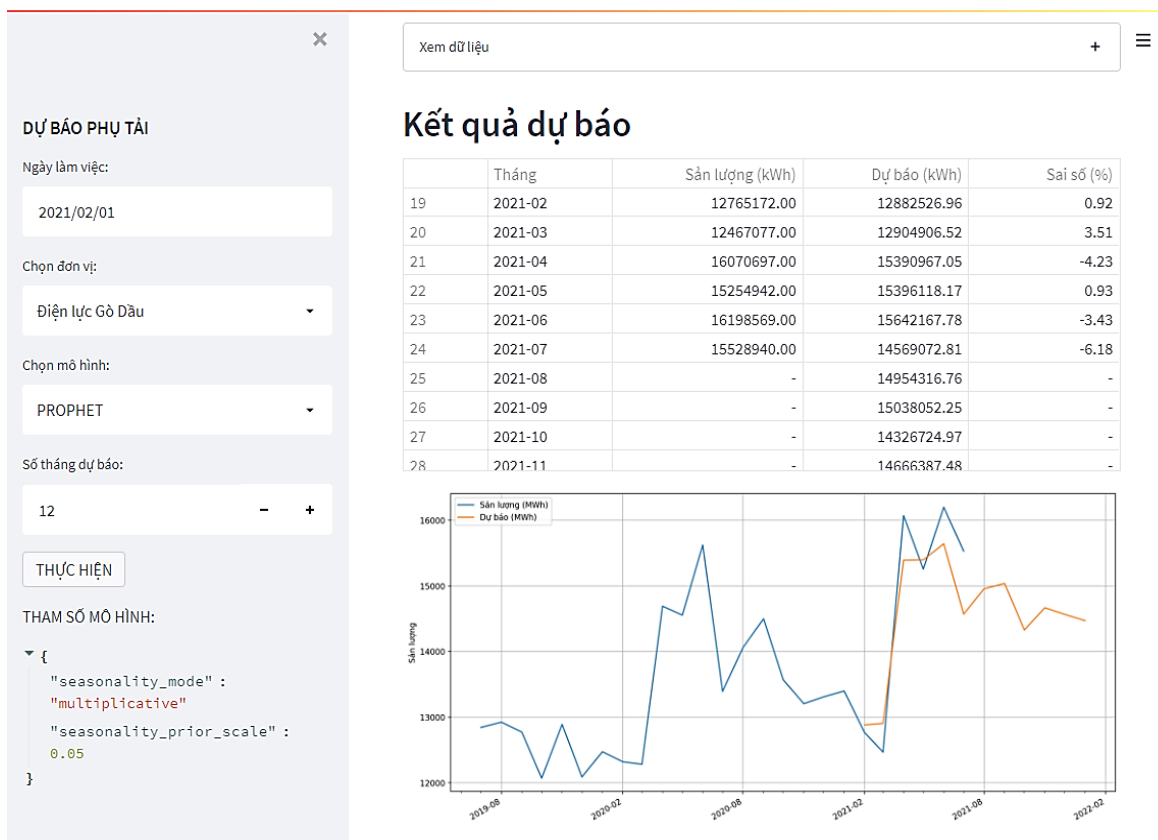
Thực hiện dự báo phụ tải điện khách hàng 12 tháng tới trên Dashboard với dữ liệu phụ tải điện khách hàng thuộc Điện lực Thành phố Tây Ninh từ tháng 02/2021 sử dụng mô hình dự báo Facebook Prophet. Hình 5.9 là kết quả dự báo khi thực hiện trên hệ thống dự báo.



Hình 5.9: Kết quả dự báo 12 tháng phụ tải điện khách hàng thuộc Điện lực Thành phố Tây Ninh từ tháng 2/2021

Sai số tuyệt đối trung bình của 6 tháng (từ 2/2021 đến 7/2021) trên dữ liệu của Điện lực Thành phố Tây Ninh là 4.02%.

Tương tự, thực hiện dự báo phụ tải điện khách hàng 12 tháng tới trên Dashboard với dữ liệu phụ tải điện khách hàng thuộc Điện lực Gò Dầu từ tháng 02/2021 sử dụng mô hình dự báo Facebook Prophet. Hình 5.10 là kết quả dự báo khi thực hiện trên hệ thống dự báo. Sai số tuyệt đối trung bình của 6 tháng (từ 2/2021 đến 7/2021) là 3.2%.



Hình 5.10: Kết quả dự báo 12 tháng phụ tải điện khách hàng thuộc Điện lực Gò Dầu từ tháng 2/2021

KẾT LUẬN

Hiện nay có rất nhiều hướng nghiên cứu để triển khai hệ thống dự báo phụ tải điện đáp ứng nhu cầu điều hành, vận hành hệ thống điện. Luận văn trình bày một hướng tiếp cận có hiệu quả trong việc dự báo nhu cầu phụ tải điện trung hạn phục vụ tốt công tác quy hoạch, điều hành hệ thống điện đáp ứng mục tiêu nâng cao hiệu quả hoạt động, độ tin cậy cung cấp điện và đảm bảo cân đối cung cầu nhằm đáp ứng nhu cầu sử dụng điện của khách hàng, đó là nghiên cứu xây dựng hệ thống dự báo phụ tải điện dựa trên phương pháp dự báo chuỗi thời gian dựa trên thuật toán Facebook Prophet, SARIMA và LSTM. Kết quả thực nghiệm 3 mô hình dự báo trên dữ liệu phụ tải điện khách hàng của 9 Điện lực cho thấy mô hình dự báo chuỗi thời gian dựa trên thuật toán Facebook Prophet cho kết quả dự báo với hiệu suất tốt nhất, kế đến là mô hình SARIMA và sau cùng là mô hình LSTM.

Hướng nghiên cứu mở rộng tiếp theo của đề tài là nghiên cứu ứng dụng thuật toán phân tích chuỗi thời gian Facebook Prophet trong dự báo phụ tải ngắn hạn phục vụ công tác lập phương thức vận hành hệ thống điện (Điều độ hệ thống điện).

DANH MỤC TÀI LIỆU THAM KHẢO

- [1] Bộ Công thương, “Quy định nội dung, phương pháp và trình tự thực hiện nghiên cứu phụ tải,” Thông tư số 19/2017/TT-BCT, 2017.
- [2] H.K. Alfares, M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23-34, 2002.
- [3] D.J. Pedregal, J.R. Trapero, "Mid-term hourly electricity forecasting based on a multi-rate approach," *Energy Conversion and Management*, vol. 51, no. 1, pp. 105-111, 2010.
- [4] LD Phúc, BM Dương, PA Duy, NT Hoan, BĐ Hoài, NM Tùng, NM Khôi, ĐN Minh, NV Dũng, “Ứng dụng phân tích thống kê để đánh giá độ tin cậy của nguồn dữ liệu đầu vào nhằm nâng cao chất lượng dự báo phụ tải điện ngắn hạn trên lưới điện TP.HCM,” *Tạp chí Phát triển Khoa học và Công nghệ – Kỹ thuật và Công nghệ*, tập 2, số 4, pp. 223-239, 2019.
- [5] Nazih Abu-Shikhah, Fawwaz Elkarmi, Osama M. Aloquili, "Medium-Term Electric Load Forecasting Using Multivariable Linear and Non-Linear Regression," *Smart Grid and Renewable Energy*, vol. 2, pp. 126-135, 2011.
- [6] Isaac A. Samuel, Felly-Njoku Chihurumanya F., Adewale Adeyinka A. and Ayokunle A. Awelewa, "Medium-Term Load Forecasting Of Covenant University Using The Regression Analysis Methods," *Journal of Energy Technologies and Policy*, vol. 4, no. 4, pp. 4-16, 2014.
- [7] Pituk Bunnoon, "Mid-Term Load Forecasting Based on Neural Network Algorithm: a Comparison of Models," *International Journal of Computer and Electrical Engineering*, vol. 3, no. 4, pp. 600-605, 2011.
- [8] E. C. Ashigwuike, A. R. A. Aluya, J. E. C. Emechebe and S. A. Benson, "Medium term electrical load forecast of Abuja Municipal Area Council using Artificial Neural Network method," *Nigerian Journal of Technology*, vol. 39, no. 3, pp. 860-870, 2020.

- [9] Saeed M. Badran, Ossama B. Abouelatta, "Forecasting Electrical Load using ANN Combined with Multiple Regression method," *The Research Bulletin of Jordan ACM*, vol. 2, pp. 52-58, 2020.
- [10] Lei Guo, Linyu Wang and Hao Chen, "Electrical Load Forecasting Based on LSTM Neural Networks," *International Conference on Big Data, Electronics and Communication Engineering*, vol. 94, pp. 107-111, 2019.
- [11] Zhitong Ma, Cantao Ye, Huashan Li, Weibin Ma, "Applying support vector machines to predict building energy consumption in China," *Energy Procedia*, vol. 152, pp. 780-786, 2018.
- [12] Sean J. Taylor and Benjamin Letham, "Forecasting at Scale," *The American Statistician*, vol. 72, no. 1, pp. 37-45, 2018.
- [13] HT Thừa, T Hạnh, LNK Thư, HT Trữ, LB Loan, HL Minh, LTK Vân, "Phân tích nhu cầu và dự báo hoạt động đầu tư tại Khu công nghệ cao Thành phố Hồ Chí Minh," 2021.
- [14] Zar Zar Oo and Sabai Phyu, "Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina," *International Journal of Applied Mathematics, Electronics and Computers*, vol. 8, no. 4, pp. 263-267, 2020.
- [15] Toni Toharudin , Resa Septiani Pontoh , Rezzy Eko Caraka , Solichatus Zahroh , Youngjo Lee & Rung Ching Chen, "Employing Long Short-Term Memory and Facebook Prophet Model in Air Temperature Forecasting," *Communications in Statistics - Simulation and Computation*, 2021.
- [16] Emir Žunić, Kemal Korjenić, Kerim Hodžić and Dženana Đonko, "Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-World Data," *International Journal of Computer Science & Information Technology*, vol. 12, no. 2, pp. 23-36, 2020.
- [17] Huỳnh Tấn Nguyên, Nguyễn Văn Lượng, "Ứng dụng mô hình Arima để dự báo chỉ số giá tiêu dùng ở Việt Nam," *Tạp chí Khoa học và Công nghệ Đại học Quảng Bình*, tập 12, 2016.

- [18] Trịnh Thị Phan Lan, “Dự báo biến động tỷ giá bằng mô hình ARIMA,” *Tài chính*, tập 2, số 2, pp. 45-47, 2018.
- [19] Nghiêm Phúc Hiếu, “Ứng dụng mô hình SARIMA dự báo lượng khách quốc tế đến Việt Nam tới năm 2020,” *Kinh tế và Dự báo*, 2017.
- [20] Lê Nhị Bảo Ngọc, Lê Quang Thông và Thái Anh Hòa, “Mô hình dự báo giá tôm sú xuất khẩu Việt Nam,” *Tạp chí Khoa học Trường Đại học Cần Thơ*, tập 54, số 6D, pp. 188-195, 2018.
- [21] LH Phúc, “Dự báo phụ tải điện tỉnh Tiền Giang trên cơ sở Neural Network,” *Đại học Sư phạm Kỹ thuật Thành phố HCM*, 2014.
- [22] LTT Hải, “Nghiên cứu ứng dụng mạng nơon nhân tạo trong dự báo phụ tải ngắn hạn cho Thành phố Đà Nẵng,” *Đại học Đà Nẵng*, 2011.

BẢN CAM ĐOAN

Tôi cam đoan đã thực hiện việc kiểm tra mức độ tương đồng nội dung luận văn “**Ứng dụng chuỗi thời gian trong dự báo nhu cầu phụ tải điện ở Công ty Điện lực Tây Ninh**” qua phần mềm DoIT một cách trung thực và đạt kết quả mức độ tương đồng 13% toàn bộ nội dung luận văn. Bản luận văn kiểm tra qua phần mềm là bản cứng luận văn đã nộp để bảo vệ trước hội đồng. Nếu sai tôi xin chịu các hình thức kỷ luật theo quy định hiện hành của Học viện.

TP.HCM, ngày 25 tháng 01 năm 2022

Học viên thực hiện luận văn

Nguyễn Thanh Nhân



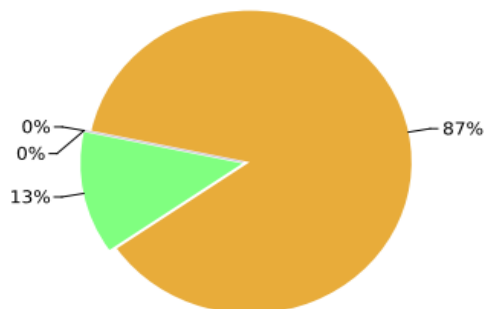
Hệ thống hỗ trợ nâng cao chất lượng tài liệu

KẾT QUẢ KIỂM TRA TRÙNG LẬP TÀI LIỆU

THÔNG TIN TÀI LIỆU

Tác giả	Nguyễn Thanh Nhân
Tên tài liệu	Luận văn thạc sĩ Ứng dụng chuỗi thời gian trong dự báo nhu cầu phụ tải điện ở Công ty Điện lực Tây Ninh
Thời gian kiểm tra	08-12-2021, 10:03:52
Thời gian tạo báo cáo	08-12-2021, 10:05:28

KẾT QUẢ KIỂM TRA TRÙNG LẬP



- Câu (đoạn) người dùng phản hồi
- Phần trăm câu (đoạn) hệ thống không kiểm tra
- Phần trăm câu (đoạn) không trùng lặp
- Phần trăm câu (đoạn) trùng lặp

Điểm	13
Nguồn trùng lặp tiêu biểu	[text.123doc.org, 123doc.org, vi.wikipedia.org]

HỌC VIÊN THỰC HIỆN LUẬN VĂN NGƯỜI HƯỚNG DẪN KHOA HỌC

Nguyễn Thanh Nhân

TS. Huỳnh Trọng Thừa