

# HỘI VÔ TUYẾN - ĐIỆN TỬ VIỆT NAM

THE RADIO AND ELECTRONICS ASSOCIATION OF VIETNAM



## KỶ YẾU HỘI NGHỊ QUỐC GIA LẦN THỨ XXII

### VỀ ĐIỆN TỬ, TRUYỀN THÔNG VÀ CÔNG NGHỆ THÔNG TIN

#### REV - ECIT 2019

Hà Nội, ngày 07/12/2019



#### ĐƠN VỊ TÀI TRỢ



#### BẢO TRỢ TRUYỀN THÔNG



NHÀ XUẤT BẢN THÔNG TIN VÀ TRUYỀN THÔNG

# KỶ YẾU HỘI NGHỊ QUỐC GIA LẦN THỨ XXII

VỀ ĐIỆN TỬ, TRUYỀN THÔNG VÀ CÔNG NGHỆ THÔNG TIN

REV - ECIT 2019

**Chịu trách nhiệm xuất bản, nội dung**

*Giám đốc - Tổng biên tập:*

TRẦN CHÍ ĐẠT

**Biên tập:**

NGUYỄN TIẾN PHÁT - BÙI HỮU LỘ - PHẠM THỊ THANH

**Trình bày:**

PHẠM VĂN ANH

**Thiết kế bìa:**

PHẠM VĂN ANH



Số xác nhận đăng ký xuất bản: 4933-2019/CXBIPH/1-197/TTTT

Số quyết định xuất bản: 142/QĐ-NXB TTTT ngày 06 tháng 12 năm 2019

Nộp lưu chiểu Quý IV năm 2019

Mã ISBN: 978-604-80-4338-4

# LỜI NÓI ĐẦU

## “Hội nghị Quốc gia về Điện tử, Truyền thông và Công nghệ Thông tin”

*(The National Conference on Electronics, Communications and Information Technology, viết tắt là REV-ECIT)* là Hội nghị khoa học của Hội Vô tuyến - Điện tử Việt Nam (REV) được tổ chức hàng năm. Từ năm 1990 đến nay, Hội nghị REV-ECIT đã trở thành diễn đàn lớn để công bố và trình bày các báo cáo khoa học của giới khoa học ngành Điện tử - Viễn thông cả nước.

Hội nghị REV-ECIT 2019 được tổ chức với chủ đề “**Phát triển công nghệ thông tin di động 5G và ứng dụng**” đã nhận được 66 báo cáo khoa học từ các Viện, Trường và các Tổ chức nghiên cứu khoa học. 54 báo cáo đã được chấp nhận, trong đó 22 báo cáo được trình bày trong 4 tiểu ban là: tiểu ban Truyền thông và Vô tuyến; tiểu ban Kỹ thuật điện tử; tiểu ban Công nghệ thông tin và mạng; tiểu ban xử lý tiên hiệu; 32 báo cáo được trình bày bằng poster.

Cùng với các Tiểu ban báo cáo khoa học, Hội Vô tuyến - Điện tử Việt Nam đã phối hợp với Cục tần số Vô tuyến điện, Bộ Thông tin và Truyền thông tổ chức thêm phiên diễn đàn cấp cao về phát triển thông tin di động 5G và ứng dụng: Tương lai của 5G; Sự sẵn sàng về thương mại của thiết bị 5G; Phát triển các ứng dụng trên nền tảng di động băng rộng 5G; Sự sẵn sàng của VNPT để triển khai 5G tại Việt Nam; Thực tế bước đầu triển khai 5G tại Việt Nam; Việc chuẩn bị tần số cho 5G tại Việt Nam... do lãnh đạo Cục Tần số vô tuyến điện, đại diện các tập đoàn, các công ty công nghệ lớn ở Việt Nam và thế giới trình bày.

Bên lề Hội nghị REV-ECIT 2019 Ban tổ chức đã phối hợp với các doanh nghiệp tổ chức triển lãm và trưng bày một số sản phẩm tiêu biểu trong các lĩnh vực có liên quan do các doanh nghiệp trong nước nghiên cứu và phát triển.

Ban Tổ chức trân trọng giới thiệu “*Kỷ yếu của Hội nghị Quốc gia lần thứ XXII về Điện tử, Truyền thông và Công nghệ thông tin REV-ECIT 2019*”.

**BAN TỔ CHỨC REV-ECIT 2019**

# MỤC LỤC

Nghiên Cứu Và Xây Dựng Phần Mềm Xác Định Trạng Thái Đi Xe Máy Sử Dụng Thuật Toán Học Máy Véc-Tơ Và Cảm Biến Tích Hợp Trên Điện Thoại	1
Các ứng dụng tiềm năng của 5G: Quan điểm và nhận thức của các bên liên quan tại Việt Nam	6
UIT-ODSMS: Hệ thống quản lý hỗ trợ hiển và nhận mô tạng sử dụng công nghệ Blockchain	13
Nghiên cứu, xây dựng phần mềm nâng cao độ chính xác trong đo kiểm đánh giá KPI mạng vô tuyến	19
Phát Hiện Trong Mạng Radar Nhiều Vị Trí Xử Lý Phân Tán Khi Quyết Định Từ Các Đài Thành Phần Tương Quan	24
Phát hiện Trojan phần cứng sử dụng kỹ thuật phân tích đặc tính độ trễ đường truyền tín hiệu	29
Thiết kế chế tạo cân cân bằng ứng dụng trong sản xuất chế phẩm máu ứng dụng tại Viện huyết học và truyền máu trung ương	34
Ứng Dụng Của Học Máy Trong Lĩnh Vực Dự Báo: So Sánh Hiệu Quả Của Mô Hình Dự Báo Truyền Thống ARIMA Và Mô Hình Dự Báo Sử Dụng Mạng Nơ Ron NNAR	38
Thiết Kế Hệ Thống Nhận Diện Cử Chỉ Tay Trên Ảnh Nhiệt Sử Dụng Mô Hình VGG16 và SVM	43
Cải thiện độ chính xác của hệ thống định vị trong nhà PDR sử dụng mã QR	48
Định vị robot dùng bản đồ tín hiệu WiFi	53
Nghiên cứu xây dựng hệ thống kiểm tra mạng ổn định trở kháng đường dây nguồn sử dụng trong các phép thử nghiệm tương thích điện từ trường thuộc tiêu chuẩn quân sự MIL-STD 461	58
Nghiên cứu, thiết kế, chế tạo thiết bị thử nghiệm kín nước dụng cho nhóm trang thiết bị đi kèm người nhái	68
Giám Sát Định Vị, Bản Đồ Hóa Và Điều Hướng Cho Robot Tự Hành Đa Hướng Sử Dụng Hệ Điều Hành Lập Trình ROS	74
Copulas Gauss cho Bài Toán Xác Định Độ Tương Quan và Ứng Dụng Đối Sánh Vector Đặc Trưng	79
Thiết Kế Mạng Tiếp Điện Song Hành Cho Mảng Anten Tuyến Tính Có Yêu Cầu Đặt Dài Rộng Các Điểm Không Trên Giản Đồ Bức Xạ	85
Đề xuất giải pháp chống tấn công blackhole xem xét cân bằng năng lượng cho mạng WSNs	90
Thiết Kế Hệ Thống Điều Khiển Trong Nhà Máy Giấy Sử Dụng Mạng Truyền Thông CC-link	85
Ăng-ten mảng khe phân cực tròn sử dụng cấu trúc tiếp điện bằng hốc cộng hưởng SIW hoạt động ở băng tần 28 GHz cho các ứng dụng 5G	100
Phân Tích Hiệu Năng Mạng Chuyển Tiếp Đa Chặng Sử Dụng NOMA Dưới Sự Ảnh Hưởng Của Giao thoa Đồng Kênh Và Khiếm Khuyết Phần Cứng	106
Concept Ăng ten cho hệ thống vô tuyến khả tri sử dụng tụ điện biến dung dựa trên lớp điện môi mỏng BST	112

Đề xuất giao thức mã hóa giả xác suất có thể chối từ sử dụng mã hóa Vernam kết hợp thuật toán ElGamal-----	122
Thu thập Chỉ Số Nước Tiêu Thụ Tự động Ứng Dụng Công Nghệ Bluetooth Beacon-----	127
Anten đa băng tái cấu hình theo tần số cho các ứng dụng WSN/Wi-Fi/WIMAX/LTE/5G trong các thiết bị đầu cuối IoT-----	132
Hệ Thống Phát Hiện Và Nhận Diện Mặt Người Sử Dụng Mô Hình SQUEEZENET Và SSD-----	138
Điều khiển ổn định robot di động đa hướng-----	144
Nghiên Cứu Các Phương Thức Xử Lý Tín Hiệu Trong Hệ Thống Giám Sát MLAT-----	150
Đánh Giá Hiệu Năng Xác Suất Dừng Mạng Thông Tin Vệ Tinh Chuyển Tiếp Hai Chiều Sử Dụng M. Fountain-----	152
Phân tích dữ liệu số chiều lớn bằng một số phương pháp học máy-----	157
Giải pháp nâng cao hiệu năng hệ thống mã hóa, giải mã tiếng nói dựa trên tính chất thừa của dữ liệu tiếng nói trong miền thời gian-----	163
Giải Pháp Nguồn Năng Lượng Mặt Trời Cho Hệ Thống Aquaponics Trên Nền Tầng Đám Mây-----	167
Thiết Kế Nút Di Động Không Dây Đa Giao Tiếp Vô Tuyến Trong Mô Phòng Mạng Ad hoc-----	172
Thuật Toán Song Song Khai Thác Tập Sinh Tối Thiểu của Tập Phổ Biến Đóng-----	176
Xây dựng hệ thống phân loại sản phẩm sử dụng robot cộng TÁC YUMI-----	182
Đề xuất thuật toán nhận dạng chữ số viết tay sử dụng mạng CNN-----	188
Khai thác dữ liệu nhằm phát hiện vết nứt đường ống nhà máy điện nguyên tử-----	194
Thiết Kế Bộ Lựa Chọn Dải Tần Số Cho VCO Dải Rộng-----	199
Phát hiện Trojan phần cứng bằng phương pháp phân tích kênh bên sử dụng kỹ thuật học máy-----	203
Phân tích, Đánh giá Hiệu năng Mã hóa Video với Chuẩn H.265/HEVC-----	208
Một Phương Pháp Giải Bài Toán Chia Sẻ Bí Mật-----	213
Thiết kế anten mảng pha hoạt động tại băng tần X cho ứng dụng kiểm định chất lượng hệ thống Radar-----	217
Nâng Cao Chất Lượng Hệ Thống Truyền Dẫn Quang Sử Dụng Bộ Liên Hợp Pha Quang OPC Cho Tín Hiệu DP-QPSK Và DP-16QAM-----	222
Đề xuất anten nhỏ sử dụng chấn tử đơn cực và xoắn cầu cho hệ thống nội soi bằng viên nang không dây-----	228
Một giải pháp phối hợp trở kháng cho anten xoắn kích thước nhỏ-----	232
Nghiên cứu tính tương quan không gian cho mô hình kênh MIMO cỡ rất lớn-----	236
Mô hình hóa và phân tích hiệu năng cho các mạng cảm biến không dây hoạt động dựa trên năng lượng thu thập-----	241
Phân Tích Hiệu Năng Mạng Khuếch Đại Chuyển Tiếp Đa Chặng Dưới Sự Ảnh Hưởng Chung Của Nhiều Đồng Kênh Và Nhiều Phần Cứng-----	247
Mạng Noron tích chập Resnet50: Ứng dụng nhận dạng khuôn mặt người hiệu quả trong các điều kiện	

môi trường phức tạp-----	253
Ứng dụng Mạng Nơron chập phân tầng nhiều lớp và mô hình Facenet trong phát hiện và nhận dạng khuôn mặt người-----	259
Nghiên cứu phương pháp giấu tin trong âm thanh sử dụng mật lạ tần số trong biến đổi Fouties rời rạc và phương pháp tối ưu hóa bày đàn PSO-----	265
Xây dựng hệ thống giám sát thiết bị điện ứng dụng truyền thông Zigbee và mô hình điện toán sương mù-----	271
Thiết bị datalogger thu thập, điều khiển và truyền dữ liệu tự động-----	277
Xác suất dừng hệ thống FD-NOMA với nút chuyển tiếp sử dụng công nghệ thu thập năng lượng----	283
Phân biệt nguyên âm dài và ngắn của tiếng Việt sử dụng tham số “giả tần số formant” SSCF-----	289

# Ứng Dụng Của Học Máy Trong Lĩnh Vực Dự Báo: So Sánh Hiệu Quả Của Mô Hình Dự Báo Truyền Thống ARIMA Và Mô Hình Dự Báo Sử Dụng Mạng Nơ Ron NNAR

Dương Đình Tú, Lê Văn Chương, Hồ Sỹ Phương, Tạ Hùng Cường, Mai Thế Anh  
Bộ môn Kỹ thuật điều khiển và tự động hóa, Viện Kỹ thuật và công nghệ, Đại học Vinh  
Email: [duongdinhthu@vinhuni.edu.vn](mailto:duongdinhthu@vinhuni.edu.vn), [lvchuong85@gmail.com](mailto:lvchuong85@gmail.com),  
[hophuong@vinhuni.edu.vn](mailto:hophuong@vinhuni.edu.vn), [tahungcuong3011@gmail.com](mailto:tahungcuong3011@gmail.com), [theanh@vinhuni.edu.vn](mailto:theanh@vinhuni.edu.vn)

**Abstract** - Một ứng dụng quan trọng của lĩnh vực học máy là phát triển các mô hình dự báo, trong đó có mô hình dự báo sử dụng mạng nơ ron tự hồi quy NNAR. Trong bài báo này, chúng tôi tiến hành so sánh hiệu quả của mô hình dự báo NNAR và mô hình dự báo truyền thống ARIMA với một số chuỗi dữ liệu mẫu khác nhau. Kết quả nghiên cứu thu được cho thấy sự vượt trội của mô hình dự báo NNAR so với mô hình dự báo ARIMA. Các kết quả này cũng làm rõ hơn các kết luận chưa rõ ràng về tính hiệu quả của mô hình NNAR so với mô hình ARIMA.

**Keywords** - Dự báo, mô hình ARIMA, mô hình NNAR, lỗi dự báo.

## I. GIỚI THIỆU

Dự báo cho phép đưa ra thông tin trong tương lai của một hiện tượng, quá trình chính xác nhất có thể, dựa trên tất cả các thông tin có sẵn, bao gồm dữ liệu lịch sử và thông tin về các sự kiện liên quan có thể ảnh hưởng đến hiện tượng, quá trình đó. Ngày nay dự báo được sử dụng rộng rãi trong các lĩnh vực kinh tế, kỹ thuật. Trong lĩnh vực tự động hóa các quá trình sản xuất, dự báo là một nhiệm vụ cấp thiết để tối ưu và nâng cao hiệu quả sản xuất sản phẩm. Đối với quá trình sản xuất, dự báo ngắn hạn là cần thiết cho việc lập kế hoạch nhân sự, sản xuất và vận chuyển; dự báo trung hạn cho phép xác định các yêu cầu tài nguyên trong tương lai, để mua nguyên liệu thô, thuê nhân công hoặc mua máy móc thiết bị; dự báo dài hạn được sử dụng trong hoạch định chiến lược.

Một kỹ thuật dự báo phổ biến là dự báo dựa trên phân tích chuỗi thời gian của các giá trị trong quá khứ. Phương pháp dự báo chuỗi thời gian được chia thành hai loại lớn, bao gồm:

- Các phương pháp dự báo cổ điển dựa trên kỹ thuật thống kê;
- Các phương pháp dự báo hiện đại sử dụng kỹ thuật học máy (machine learning), học sâu (deep learning) và trí tuệ nhân tạo (artificial Intelligence).

Trong các phương pháp dự báo cổ điển, phương pháp Box-Jenkins với mô hình tự hồi quy tích hợp trung bình trượt ARIMA (autoregressive integrated moving average) được xem là phương pháp dự báo hiệu quả nhất và đã được sử dụng rộng rãi cho các chuỗi thời gian.

Ngày nay với sự phát triển nhanh chóng và mạnh mẽ của lĩnh vực học máy, các mô hình dự báo sử dụng mạng nơ ron nhân tạo (Artificial Neural Network) đã được phát triển, sử dụng trong nhiều nghiên cứu về dự báo. Trong trường hợp mô hình dự báo sử dụng mạng nơ ron tự hồi quy NNAR (neural network autoregression), các giá trị trễ (lagged values) của chuỗi thời gian được sử dụng làm đầu vào cho mô hình và đầu ra là các giá trị dự đoán của chuỗi thời gian.

Đã có nhiều công bố về tính hiệu quả của dự báo khi sử dụng mô hình cổ điển ARIMA và mô hình dự báo hiện đại sử dụng mạng nơ ron NNAR với các kết quả trái ngược nhau. Trong công trình [6] các tác giả nghiên cứu áp dụng mô hình NNAR và ARIMA cho chuỗi thu nhập bình quân đầu người hàng quý của Tây Đức từ 1960 đến 1987. Kết quả cho thấy mô hình ARIMA là phù hợp hơn với chuỗi dữ liệu này và cho kết quả chính xác hơn. Tương tự trong các công bố [7][8], các chuỗi dữ liệu CUPE dọc theo bờ biển đông bắc Ấn Độ và số lượng khách du lịch đến Mauritius đã được nghiên cứu. Kết quả chỉ ra rằng, mặc dù mô hình NNAR có khả năng xử lý các mối quan hệ phi tuyến tính giữa các biến rất tốt, nhưng đối với các chuỗi dữ liệu này, mô hình ARIMA cho kết quả dự báo tốt hơn. Trong công bố [9] về dự báo các đặc tính chất lượng nước của nhà máy xử lý nước Sanandaj, Iran, các tác giả cho rằng mô hình NNAR có phần tốt hơn mô hình ARIMA trong tính toán các lỗi dự báo  $R^2$ . Còn trong bài báo [10] các giả dự báo mực nước ngầm ở thị trấn Rajshahi, Bangladesh, với kết quả mô hình NNAR là vượt trội so với mô hình ARIMA. Tương tự trong các công trình [11][12][13] về dự báo lượng mưa ở thị trấn Bangalore, Ấn Độ; chuỗi các nhu cầu về sản phẩm của cửa hàng thời trang ở Istanbul, Thổ Nhĩ Kỳ và sản lượng xây dựng ở Hồng Kông, kết quả cho thấy mô hình NNAR dự báo tốt hơn rất nhiều so với mô hình ARIMA, cả trong dự báo ngắn hạn và dài hạn. Như vậy việc so sánh hiệu quả của mô hình NNAR và ARIMA vẫn còn chưa rõ ràng.

Trong bài báo này, hiệu quả của các mô hình ARIMA và NNAR được nghiên cứu và so sánh khi dự báo với các chuỗi dữ liệu sản xuất khác nhau. Các kết quả cho thấy, mô hình NNAR đạt được hiệu quả tốt hơn nhiều so với mô hình ARIMA. Điều này càng làm rõ hơn các kết luận của các nghiên cứu trước đây về hiệu quả của mô hình NNAR so với mô hình ARIMA.



Phần còn lại của bài báo được tổ chức như sau: Trong phần II, chúng tôi miêu tả phương pháp nghiên cứu; phần III cung cấp các kết quả nghiên cứu và phân tích lý thuyết. Cuối cùng, chúng tôi kết luận bài báo trong phần IV.

## II. PHƯƠNG PHÁP NGHIÊN CỨU

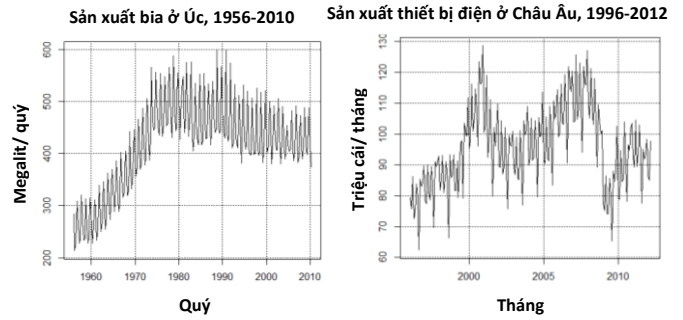
### 1. Dữ liệu đầu vào

Trong bài báo chúng tôi sử dụng các chuỗi thời gian của dữ liệu sản xuất về điện, gas, xăng, thiết bị điện tử, ... của các quốc gia Mỹ, Úc, Châu Âu... đã được nghiên cứu và công bố trong tài liệu [1][17].

Các chuỗi dữ liệu được đưa ra trong bảng 1. Đồ thị chuỗi thời gian của các dữ liệu này được thể hiện trên hình 1. Có thể nhận thấy, hầu hết các chuỗi dữ liệu đều có những biến đổi phức tạp về cả giá trị và xu hướng theo thời gian.

Bảng 1. Các chuỗi dữ liệu được sử dụng để dự báo [1][17]

TT	Dữ liệu	Tính theo thời gian	Số quan sát
1	Tổng sản lượng điện lưới sản xuất ở Mỹ từ 1/1973 đến 6/2013, tỉ KW	Hàng tháng	486
2	Tổng khối lượng xăng xe máy được sản xuất ở Mỹ từ 2/2/1991 đến 20/01/2017, triệu thùng	Hàng tuần	1355
3	Tổng sản lượng điện sản xuất ở Úc từ 1956 đến 2010, tỉ KW	Hàng quý	218
4	Tổng sản lượng sản xuất gas ở Úc từ 1956 đến 2010, petajoules	Hàng quý	218
5	Tổng sản lượng bia sản xuất ở Úc từ 1956-2010, mega lít	Hàng quý	218
6	Tổng số lượng thiết bị điện tử, máy tính, chiếu sáng được sản xuất ở châu Âu (Euro Zone - 17 quốc gia) từ 1/1996 đến 3/2012, triệu cái	Hàng tháng	195



Hình 1. Đồ thị chuỗi thời gian của các dữ liệu trong bảng 1 [1][17]

### 2. Mô hình ARIMA

Mô hình tự hồi quy tích hợp trung bình trượt lần đầu tiên được đưa ra bởi Box & Jenkins. Mô hình ARIMA được phân chia thành mô hình ARIMA không có tính mùa vụ và mô hình ARIMA có tính mùa vụ. Mô hình ARIMA(p,d,q) đầy đủ có thể được viết [1]:

$$y'_t = c + \varphi_1 y'_{t-1} + \dots + \varphi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t, (1)$$

Ở đây, - p: hệ số của thành phần hồi quy; - q: hệ số của thành phần trung bình trượt; - d: mức độ lấy sai phân để biến đổi về chuỗi thời gian dừng (stationary); -  $y_t$ : chuỗi thời gian; -  $\varphi_{1-p}$ : các tham số của thành phần hồi quy; -  $\theta_{1-q}$ : các tham số của thành phần trung bình trượt; -  $\varepsilon_t$ : tham số lỗi.

Ngoài ra, một mô hình ARIMA theo mùa vụ có thể được kí hiệu là ARIMA(p,d,q)(P,D,Q)<sub>m</sub>, trong đó - (p,d,q): các thành phần không có tính mùa vụ, - (P,D,Q): các thành phần có tính mùa vụ, - m: tần số của tính mùa vụ.

Ngày nay, các kỹ thuật tính toán, nhận dạng và dự báo bằng mô hình ARIMA đã được hoàn thiện. Trong các công bố [2][3][4][5] các tác giả đã đưa ra một quy trình tối ưu để nhận dạng tham số và dự báo bằng mô hình ARIMA, gồm các bước sau:

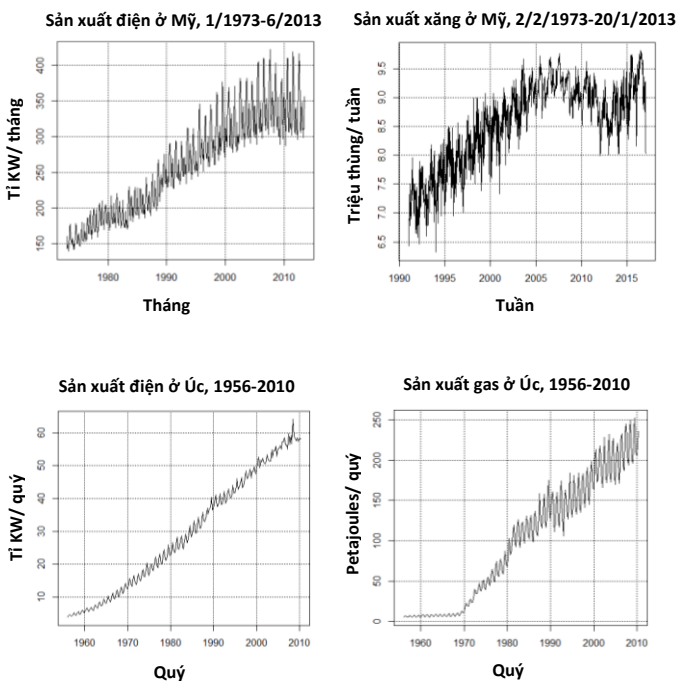
1. Dựa vào đồ thị chuỗi thời gian, xác định sơ bộ loại chuỗi thời gian (có tính dừng hay không, có tính mùa vụ hay không) và các điểm bất thường (nếu có).

2. Nếu có các điểm dữ liệu bất thường, sử dụng chuyển đổi Box-Cox để ổn định phương sai.

3. Kiểm tra tính dừng của chuỗi thời gian bằng các tiêu chuẩn Dickey-Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Schin (KPSS) và Phillips-Perron (PP). Nếu chuỗi thời gian chưa có tính dừng, cần lấy sai phân chuỗi thời gian đó. Lặp lại các bước 1-3 cho đến khi nhận được chuỗi thời gian dừng.

4. Kiểm tra hàm tự tương quan ACF (autocorrelation function) và hàm tự tương quan từng phần PACF (partial autocorrelation) của chuỗi thời gian. Lựa chọn các mô hình đầu tiên dựa trên sự phân tích hai hàm này.

5. Áp dụng các mô hình này cho dữ liệu mẫu. Dựa vào các tiêu chí BIC (Bayesian information criterion), AIC (Akaike information criterion) và AICc (AIC với sự điều chỉnh kích thước mẫu nhỏ), đánh giá và lựa chọn các mô hình phù hợp nhất.



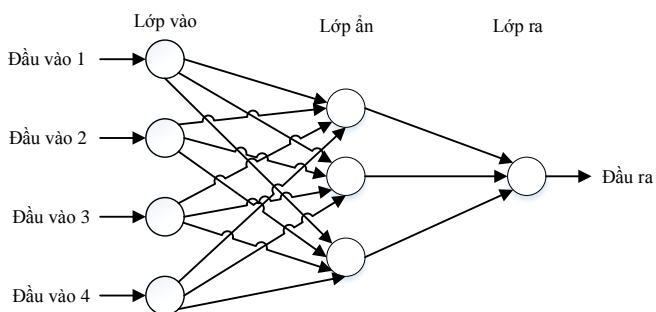


6. Áp dụng các mô hình này cho dữ liệu mẫu một lần nữa. Dựa vào các lỗi “ướm thử” mô hình để đưa ra mô hình phù hợp nhất.

7. Sử dụng mô hình nhận được để dự báo.

### 3. Mô hình NNAR

Dự báo sử dụng mạng nơ ron là phương pháp dự báo dựa trên các mô hình toán học của bộ não. Chúng cho phép xử lý các mối quan hệ phi tuyến phức tạp giữa các giá trị có sẵn để thực hiện dự báo. Mạng nơ-ron nhân tạo có khả năng “học” và xử lý song song. Nó có thể tính toán và dự báo giá trị của biến đầu ra với một tập hợp các thông tin của biến đầu vào được cho trước. Mô hình mạng nơ ron sẽ được “huấn luyện” để có thể “học” từ những thông tin quá khứ. Từ đó, mạng có thể đưa ra kết quả dự báo dựa trên những gì đã được học. Quá trình này sẽ được tiến hành bằng các thuật toán huấn luyện mạng, phổ biến là thuật toán lan truyền ngược (back-propagation algorithm). Mạng thần kinh nhân tạo truyền thẳng đa lớp (multilayer feed-forward network) được sắp xếp gồm: Lớp đầu vào, các lớp ẩn và lớp đầu ra. Lớp đầu vào sẽ là nơi nhận các tín hiệu đầu vào. Các tín hiệu này có thể là một hằng số, dữ liệu thô hoặc cũng có thể là đầu ra của một mạng nơ-ron khác. Các giá trị này sẽ tác động đến các nơ-ron lớp ẩn. Tại lớp ẩn, tín hiệu của lớp vào sẽ được xử lý bằng một hàm kích hoạt (activate function), sau đó tín hiệu sẽ được truyền qua lớp ra. Các lớp ẩn liên kết giữa lớp đầu vào và lớp đầu ra, điều này làm cho mạng thần kinh nhân tạo có khả năng mô phỏng mối tương quan phi tuyến tốt hơn.



Hình 2. Kiến trúc một mạng nơ ron bao gồm 4 đầu vào, 1 lớp ẩn và 1 đầu ra

Ví dụ đối với một mô hình mạng nơ ron đơn giản như ở hình 2, các đầu vào lớp ẩn  $j$  được kết hợp tuyến tính và nhận được:

$$z_j = b_j + \sum_{i=1}^4 w_{i,j} x_i, \quad (2)$$

Trong lớp ẩn, giá trị này sẽ được sửa đổi bằng cách sử dụng hàm phi tuyến như sigmoid để đưa đến đầu ra:

$$s(z) = \frac{1}{1+e^{-z}}, \quad (3)$$

Điều này có xu hướng làm giảm ảnh hưởng của các giá trị đầu vào cực đoan, do đó làm cho mô hình dự báo có phần mạnh mẽ hơn đối với các chuỗi dữ liệu thực tế phức tạp và biến động lớn.

Đối với mô hình NNAR, các giá trị trễ (lagged values) của chuỗi thời gian được sử dụng làm đầu vào cho mô hình và đầu ra là các giá trị dự báo của chuỗi thời gian. Dự báo sử dụng mô hình này gồm 2 bước chính [1]:

1. Đầu tiên, thứ tự tự hồi quy (the order of auto regression) được xác định cho chuỗi thời gian. Thứ tự tự hồi quy cho biết số

lượng giá trị trước đó mà giá trị hiện tại của chuỗi thời gian phụ thuộc vào.

2. Trong bước thứ hai, mạng nơ ron được đào tạo với một bộ huấn luyện được chuẩn bị bằng cách xem xét thứ tự tự hồi quy. Số lượng các nút đầu vào được xác định từ thứ tự tự hồi quy và các đầu vào cho mạng nơ ron là các quan sát bị trễ trước đó (lagged values) trong dự báo chuỗi thời gian đơn biến. Các giá trị dự báo là đầu ra của mô hình mạng nơ ron.

### 4. Các tiêu chí đánh giá hiệu quả của mô hình dự báo

Hiệu quả của mô hình dự báo được đánh giá bằng cách tính toán các lỗi dự báo. Có nhiều phương pháp tính toán lỗi dự báo, được chia ra thành các nhóm chính:

- Lỗi dự báo phụ thuộc vào quy mô chuỗi thời gian (scale-dependent measures);
- Lỗi dự báo dựa trên sai số phần trăm (measures based on percentage errors);
- Lỗi dự báo dựa trên sai số tương đối (measures based on relative errors);
- Lỗi dự báo tỷ lệ (scaled errors).

Ở đây chúng tôi sử dụng ba cách đánh giá lỗi dự báo phổ biến thường được sử dụng trong việc đánh giá hiệu quả của mô hình dự báo [14], được đưa ra trong bảng 2:

Bảng 2. Các lỗi dự báo phổ biến

TT	Lỗi dự báo	Công thức tính toán
1	Sai số trung bình tuyệt đối (Mean Absolute Error)	$MAE = \text{mean}_{i=1,n}  e_i $
2	Sai số phần trăm trung bình tuyệt đối (Mean Absolute Percentage Error)	$MAPE = \text{mean}_{i=1,n} (100 p_i )$
3	Căn bậc hai sai số trung bình (Root Mean Square Error)	$RMSE = \sqrt{\text{mean}_{i=1,n}  e_i^2 }$

Trong đó  $e_i = y_i - f_i$ ,  $y_i$  – giá trị thực tế tại thời điểm  $t$ ,  $f_i$  – giá trị dự báo tại thời điểm  $t$ ;  $p_i = \frac{|e_i|}{y_i}$ . Một mô hình dự báo tốt hơn sẽ đưa các lỗi dự báo nhỏ hơn.

### III. KẾT QUẢ NGHIÊN CỨU

Đối với mỗi chuỗi dữ liệu, chúng tôi chia ra làm 2 phần:

- Phần dữ liệu mẫu: bao gồm 100 quan sát đầu tiên của chuỗi;
- Phần dữ liệu kiểm tra: là dữ liệu từ quan sát thứ 101 của chuỗi cho đến quan sát cuối cùng của chuỗi.

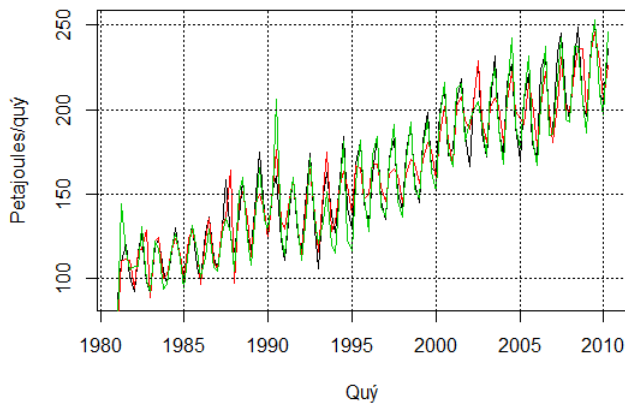
Dự báo sẽ được thực hiện cho 1 quan sát tiếp theo (dự báo giá trị - ngắn hạn). Dữ liệu ban đầu bao gồm 100 quan sát sẽ được sử dụng để nhận dạng mô hình dự báo và dự báo cho quan sát thứ 101 của chuỗi. Lỗi dự báo sẽ được tính toán, ghi lại bằng cách so sánh giá trị dự báo vừa nhận được và giá trị thực tế tại quan sát thứ 101. Các lỗi dự báo được dùng để đánh giá mô hình dự báo hiệu quả nhất được chúng tôi trình bày ở phần 2.4. Sau đó, giá trị thực tế của quan sát thứ 101 sẽ được thêm vào dữ liệu mẫu ban đầu, tạo thành chuỗi dữ liệu mẫu mới gồm 101 quan sát và thực hiện dự báo cho quan sát thứ 102 của chuỗi. Quá trình dự báo được thực hiện tương tự cho đến quan sát cuối cùng của chuỗi. Việc thực hiện dự báo với độ dài chuỗi dữ liệu mẫu khác

nhau và tăng dần sẽ cho phép kiểm tra tính hiệu quả của mô hình ARIMA và mô hình NNAR cho các chuỗi dữ liệu với độ dài khác nhau. Mỗi lần thực hiện dự báo sẽ đưa ra một mô hình dự báo với các tham số mô hình là khác nhau.

Chúng tôi xây dựng các chương trình dự báo trên ngôn ngữ R [15], với sự hỗ trợ của của gói “dự báo” được phát triển bởi Rob Hyndman và các cộng sự [16].

Các kết quả dự báo trên các chuỗi thời gian khác nhau cho thấy cả mô hình ARIMA và mô hình NNAR đều cho những kết quả dự báo tốt. Tuy nhiên mô hình NNAR cho kết quả dự báo chính xác hơn so với mô hình ARIMA. Ví dụ trên hình 3 thể hiện kết quả dự báo sản xuất gas ở Úc theo quý, từ 1981 đến 2010. Mô hình NNAR thể hiện sự “bám” tốt hơn so với mô hình ARIMA ở những quan sát có sự thay đổi lớn và phức tạp.

**Dự báo sản xuất gas ở Úc, 1956-2010**



Hình 3. Dự báo sản xuất gas ở Úc từ 1981-2010: đường màu đen – dữ liệu gốc; đường màu đỏ – dự báo với mô hình ARIMA; đường màu xanh – dự báo với mô hình NNAR

Bảng 3 đưa ra kết quả đánh giá lỗi dự báo (ở bảng 2) cho các mô hình ARIMA và NNAR cho các chuỗi dữ liệu ở bảng 1. Rõ ràng là, mô hình NNAR tỏ ra hiệu quả hơn so với mô hình ARIMA. Đối với các chuỗi dữ liệu số 1, 4, 6 lỗi dự báo của mô hình ARIMA là lớn hơn rất nhiều so với lỗi dự báo của mô hình NNAR. Quan sát ở hình 1 có thể thấy, đây là những chuỗi dữ liệu mà các giá trị của chuỗi biến đổi phức tạp, có tính chất phi tuyến mạnh. Như vậy, mô hình NNAR là vượt trội hơn so với mô hình ARIMA trong việc giải quyết bài toán phi tuyến của chuỗi dữ liệu.

**Bảng 3. So sánh lỗi dự báo đối với hai mô hình ARIMA và NNAR**

TT	Chuỗi		RMSE	MAE	MAPE
1	Tổng sản lượng điện lưới sản xuất ở Mỹ từ 1/1973 đến 6/2013	ARIMA	20,04	15,29	5,35
		NNAR	11,48	8,57	3,06
2	Tổng khối lượng xăng xe máy được sản xuất ở Mỹ từ 2/2/1991 đến 20/01/2017	ARIMA	0,35	0,26	2,9
		NNAR	0,29	0,22	2,6
3	Tổng sản lượng điện sản xuất ở Úc từ 1956 đến 2010	ARIMA	1,44	1,12	2,65
		NNAR	1,22	0,96	2,14

4	Tổng sản lượng sản xuất gas ở Úc từ 1956 đến 2010	ARIMA	12,26	9,45	5,91
		NNAR	9,51	6,82	4,58
5	Tổng sản lượng bia sản xuất ở Úc từ 1956-2010	ARIMA	21,13	16,18	3,52
		NNAR	19,96	15,53	3,37
6	Tổng số lượng thiết bị điện tử, máy tính, chiếu sáng được sản xuất ở châu Âu (Euro Zone – 17 quốc gia) từ 1/1996 đến 3/2012	ARIMA	8,86	7,08	7,30
		NNAR	5,08	4,02	4,05

#### IV. KẾT LUẬN

Báo cáo này trình bày các kết quả dự báo chuỗi thời gian của mô hình cổ điển ARIMA và mô hình hiện đại NNAR đối với các chuỗi dữ liệu sản xuất. Các kết quả nghiên cứu của chúng tôi chỉ ra rằng, mô hình dự báo sử dụng mạng nơ ron NNAR là vượt trội so với mô hình tự hồi quy tích hợp trung bình trượt ARIMA, đặc biệt đối với những chuỗi dữ liệu có tính chất phi tuyến mạnh. Điều này làm rõ hơn kết luận chưa rõ ràng của các nghiên cứu của nhiều tác giả trước về tính hiệu quả của mô hình ARIMA và mô hình NNAR. Mặc dù kết quả nghiên cứu là không đồng nhất với một số nghiên cứu trước, tuy nhiên trong những công việc sắp tới, các tác giả sẽ tiếp tục thử nghiệm trên nhiều loại dữ liệu khác nhau để khẳng định lại kết quả này.

#### TÀI LIỆU THAM KHẢO

- [1] Rob Hyndman, George Athanasopoulos, “forecasting principles and practice”, OTexts: Melbourne, Australia, 2018. <https://otexts.com/fpp2/>
- [2] Suvorov D.N., Duong Dinh Tu, “A method of demand forecasting for precast concrete products in factories”, Vestnik MADI, vol. 2, no. 45, pp. 77-81, 2016.
- [3] Suvorov D.N., Duong Dinh Tu, “A method of demand forecasting for precast concrete products in factories with seasonal adjustment”, Vestnik MADI, vol. 1, no. 48, pp. 106-110, 2017.
- [4] Suvorov D.N., Duong Dinh Tu, “Optimal control of precast concrete production in the conditions of fluctuating demand in Vietnam”, Vestnik MADI, vol. 3, no. 50, pp. 114-119, 2017.
- [5] Duong Dinh Tu, “Management effectiveness evaluation of the precast concrete production process in the conditions of fluctuating demand”, Modern science: actual problems of theory and practice, vol. 1, pp. 21-26, 2018.
- [6] Debasish Sena, Naresh Kumar Nagwani, “A neural network autoregression model to forecast per capita disposable income”, ARPN Journal of Engineering and Applied Sciences, vol. 11, no. 4, pp. 13123-13128, november 2016.
- [7] K. G. Mini, Somy Kuriakose and T. V. Sathianandan. “Modeling CPUE series for the fishery along northeast coast of India: A comparison between the HoltWinters, ARIMA and NNAR models”, Journal of the Marine Biological Association of India Vol. 57, No.2, pp. 76-82, Jul-Dec 2015.
- [8] Ruben Thoplan, “Simple v/s Sophisticated Methods of Forecasting for Mauritius Monthly Tourist Arrival Data”, International Journal of Statistics and Applications 2014, 4(5): pp. 217-223.
- [9] Maleki, Afshin, Nasser Simin, Mehri Solaimany Aminabad, Hadi Mahd, “Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics”, KSCE Journal of Civil Engineering, vol. 22, issue 9, pp 3233–3245, 2018.
- [10] Md. Abdul Khalek, Md. Ayub Ali, “Comparative Study of Wavelet-SARIMA and Wavelet- NNAR Models for Groundwater Level in Rajshahi District”, Vol. 10, Issue 7, pp. 01-15, 2016.

- [11] Lam, K. and Oshodi, O. (2016), "Forecasting construction output: a comparison of artificial neural network and Box-Jenkins model", *Engineering, Construction and Architectural Management*, Vol. 23 No. 3, pp. 302-322, 2016.
- [12] M.A. Ekmiş, M. Hekimoğlu, B. Atak Bülbül , "Revenue forecasting using a feed-forward neural network and ARIMA model", *Sigma J Eng & Nat Sci* 8 (2), 129-134, 2017.
- [13] S. Bhavyashree and Banjul Bhattacharyya, "A comparative study on ARIMA and ANN for rainfall pattern of Bangalore rural district", *RASHI* 3 (2), pp. 45 - 49, 2018.
- [14] Rob Hyndman, Anne Koehler, "Another look at measures of forecast accuracy", *International Journal of Forecasting*, Vol. 22, Issue 4, pp. 679-688, October–December 2006.
- [15] The R Project for Statistical Computing: <https://www.r-project.org/>
- [16] Package "forecast" in R: <https://cran.r-project.org/web/packages/forecast/forecast.pdf>
- [17] Package "fpp2" in R: <https://cran.r-project.org/web/packages/fpp2/fpp2.pdf>