

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH**

NGUYỄN NGỌC ÂU

**HỆ THỐNG NHẬN DẠNG THÔNG MINH
CHẨN ĐOÁN NHANH ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN**

LUẬN ÁN TIẾN SỸ

NGÀNH: KỸ THUẬT ĐIỆN

Tp. Hồ Chí Minh – tháng 6/2018.

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN NGỌC ÂU

HỆ THỐNG NHẬN DẠNG THÔNG MINH
CHẨN ĐOÁN NHANH ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN
NGÀNH: KỸ THUẬT ĐIỆN – 62520202

Hướng dẫn khoa học:

1. PGS. TS. Quyền Huy Ánh
2. PGS. TS. Phan Thị Thanh Bình

Số: 246./QĐ-ĐHSPKT-SĐH

Tp. Hồ Chí Minh, ngày 04 tháng 10 năm 2013

QUYẾT ĐỊNH

Về việc giao đề tài luận án và người hướng dẫn nghiên cứu sinh khóa 2013 HIỆU TRƯỞNG TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH

Căn cứ quyết định số 426/QĐ-TTg ngày 27/10/1976 của Thủ tướng Chính phủ quy định về mạng lưới tổ chức và nhiệm vụ của các trường đại học;

Căn cứ chương 6 Điều lệ Trường Đại học ban hành kèm theo quyết định số 58/2010/QĐ-TTg ngày 22 tháng 9 năm 2010 của Thủ tướng Chính phủ;

Căn cứ thông tư số 10/2009/TT-BGDĐT ngày 07/5/2009 của Bộ Giáo dục và Đào tạo về việc Ban hành Quy chế đào tạo trình độ tiến sĩ;

Căn cứ thông tư số 05/2012/TT-BGDĐT ngày 15/02/2012 của Bộ Giáo dục và Đào tạo về việc sửa đổi, bổ sung một số điều của Quy chế đào tạo trình độ tiến sĩ ban hành kèm theo Thông tư số 10/2009/TT-BGDĐT ngày 07/5/2009 của Bộ trưởng Bộ Giáo dục và Đào tạo;

Xét nhu cầu công tác và khả năng cán bộ;

Xét đề nghị của Trưởng phòng Đào tạo,

QUYẾT ĐỊNH

Điều 1: Giao đề tài luận án tiến sĩ và người hướng dẫn cho:

Nghiên cứu sinh : *Nguyễn Ngọc Âu*

Ngành : Kỹ thuật điện

Khoá: 2013 – 2016

Tên luận án : *Hệ thống nhận dạng thông minh chẩn đoán nhanh ổn định động hệ thống điện*

Người HD thứ nhất (HD chính): *PGS.TS Quyền Huy Ánh*

Người HD thứ hai : *PGS.TS Phan Thị Thanh Bình*

Thời gian thực hiện : *01/10/2013 đến 01/10/2016*

Điều 2: Giao cho Phòng Đào tạo quản lý, thực hiện theo đúng Quy chế đào tạo trình độ tiến sĩ của Bộ Giáo dục & Đào tạo đã ban hành.

Điều 3: - Trưởng các đơn vị và các Ông (Bà) có tên ở điều 1 chịu trách nhiệm thi hành quyết định này.

- Quyết định có hiệu lực kể từ ngày ký.

Nơi nhận :

- BGH (để biết);
- Như điều 2, 3;
- Lưu P.ĐT



HIỆU TRƯỞNG

Đỗ Văn Dũng
PGS.TS. Đỗ Văn Dũng

LÝ LỊCH CÁ NHÂN

I. LÝ LỊCH SƠ LƯỢC				
Họ và tên: Nguyễn Ngọc Âu		Giới tính: Nam		
Ngày, tháng, năm sinh: 20/12/1970		Nơi sinh: Tiền Giang		
Quê quán: Tiền Giang		Dân tộc: Kinh		
Học vị cao nhất: Thạc sỹ		Năm, nước nhận học vị: 2003		
Đơn vị công tác: Khoa điện điện tử				
Chỗ ở hiện nay: 114/18/6 Vườn Lài, An Phú Đông, Quận 12, TpHCM.				
Điện thoại liên hệ:		CQ: 08 38968641	ĐD: 0983 79 19 29	
Email: ngocau@hcmute.edu.vn				
II. QUÁ TRÌNH ĐÀO TẠO				
1. Đại học				
Hệ đào tạo: chính quy				
Nơi đào tạo: Đại Học Sư Phạm Kỹ Thuật Tp.HCM				
Ngành học: Điện khí hóa & cung cấp điện				
Nước đào tạo: Việt Nam			Năm tốt nghiệp: 1998	
2. Sau đại học				
Thạc sỹ chuyên ngành: Kỹ thuật điện			Năm cấp bằng: 2003	
Nơi đào tạo: Đại Học Bách Khoa Tp.HCM				
3. Ngoại ngữ			Tiếng Anh: B2	
II. QUÁ TRÌNH CÔNG TÁC CHUYÊN MÔN				
Thời gian		Nơi công tác		Công việc đảm nhận
1998 đến nay		ĐH. SPKT TP.HCM		CBGD
III. QUÁ TRÌNH NGHIÊN CỨU KHOA HỌC				
1. Các đề tài nghiên cứu khoa học đã từng tham gia				
TT	Tên đề tài nghiên cứu	Năm	Đề tài cấp	Vai trò
1	Khảo sát bộ biến đổi điện một pha sang ba pha dùng máy điện quay	2012	Cấp trường T2012-50	Chủ nhiệm
2	Đánh giá ổn định hệ thống điện nhiều máy phát	2012-2013	Cấp trường trọng điểm T2013-02TD	Chủ nhiệm
3	Lựa chọn mẫu trong đánh giá thông minh ổn định động hệ thống điện	2014-2015	Cấp trường trọng điểm T2015-34TD	Chủ nhiệm
4	Ứng dụng mạng thần kinh nhân tạo chẩn đoán ổn định động hệ thống điện nhiều máy	2015-2016	Cấp trường trọng điểm T2016-48TD	Chủ nhiệm
5	Hệ thống nhận dạng đánh giá ổn định động hệ thống điện	2016-2017	Cấp trường trọng điểm T2017-68TD	Chủ nhiệm
2. Các công trình đã công bố				
TT	Tên công trình	Năm công bố	Tên tạp chí	
1.	Ngoc Au Nguyen , Trong Nghia Le, Huy Anh Quyen, Thi Thanh Binh Phan, "Data reduction for dynamic stability classification in power system", IETE Journal of Research, DOI:10.1080/03772063.2017.1417752, ISSN: 0377-2063 (Print) 0974-780X (Online), Jan. 2018.			
2.	Ngoc Au Nguyen , Trong Nghia Le, Huy Anh Quyen, Binh Phan Thi Thanh, Thai Binh Nguyen "Hybrid Classifier Model for Dynamic Stability Prediction in Power System", International Conference-ICSSE 2017, p. 67-70, 2017 IEEE.			

3. **Nguyen Ngoc Au**, Quyên Huy Anh, Phan Thi Thanh Binh, "Dynamic Stability Recognition Of Power System Using Generalized Regression Neural Networks" Journal of Technical Education Science, N^o 40A (2017), ISSN 1859-1272, 2017.
4. **Nguyen Ngoc Au**, Quyên Huy Anh, Phan Thi Thanh Binh, "Feature subset selection in dynamic stability assessment power system using artificial Neuralnetworks" Science & Technology Development Journal, ISSN 1859-0128, Vol.18, p. 15-24, No.K3 – 2015.
5. **Nguyen Ngoc Au**, Quyên Huy Anh, Phan Thi Thanh Binh, "Dynamic stability assessment of power system using Multilayer feedforward Neural networks with Reduced feature selection", The 2nd International Conference on GTSD, 2014 (GTSD'14).
6. **Nguyen Ngoc Au**, Quyên Huy Anh, Phan Viet Thinh, "Feature Selection For Dynamic Stability Prediction Of Power System Using Neural Network", Journal of Technical Education Science, N^o 34(2015), ISSN 1859-1272, 2015.
7. Quyên Huy Anh, **Nguyen Ngoc Au**, Nguyen Vu Phuong Thao, "Design dual input power system stabilizer for multi-machine system based on focused-time-delay Neural network", Journal of Technical Education Science, N^o 25(2013), ISSN 1859-1272.
8. **Nguyễn Ngọc Âu**, Lê Trọng nghĩa, Quyền Huy Ánh, Phan Thị Thanh Bình, 'Sa thải phụ tải dựa trên nhận dạng nhanh ổn định động hệ thống điện', ISSN 1859-1531 Tạp Chí Khoa Học Và Công Nghệ, Đại Học Đà Nẵng-Số 11(129).2017 QUYỂN 2, p. 6-11, 2017.
9. Trong Nghĩa Le, **Ngoc Au Nguyen**, Huy Anh Quyên, "Emergency control of load shedding based on coordination of artificial Neural network and analytic hierarchy process algorithm", IEEE International Conference on System Science and Engineering (ICSSE 2017), p. 67-70, 2017 IEEE.
10. L.T. Nghĩa, T.T. Giang, **N.N. Au**, Q.H. Anh, D.N. An, "Emergency Control of Load Shedding Based on Fuzzy-AHP Algorithm", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 6 Issue 09, p. 185-191, September – 2017.
11. Phan Thị Thanh Bình, Nguyễn Thụy Mai Khanh, **Nguyễn Ngọc Âu**. "Phân tích tình ổn định điện áp khi có máy phát điện gió DFIG", Tạp chí Phát triển KH&CN, Tập 19, Số K5-2016, trang 5-12, 2016.
12. Trong Nghĩa Le, Huy Anh Quyên, **Ngoc Au Nguyen**, "Application of fuzzy-analytic hierarchy process algorithm and fuzzy load profile for load shedding in power systems", Electrical Power and Energy Systems 77 (2016) 178–184, 2016.
13. Quyền Huy Ánh, Nguyễn Phan Thanh, **Nguyễn Ngọc Âu**, Trương Ngọc Hưng, 'Nghiên cứu hiệu quả bảo vệ máy biến áp của thiết bị chống sét van có xét đến các yếu tố ảnh hưởng', Tạp chí phát triển khoa học và công nghệ, tập 12, số 08, 2009.
14. Quyền Huy Ánh, **Nguyễn Ngọc Âu**, 'Mô phỏng thiết bị chống sét lan truyền trên đường nguồn hạ áp', Tạp chí khoa học và công nghệ các trường đại học kỹ thuật, số 42-43, 2003
15. Trương Việt Anh, **Nguyễn Ngọc Âu**, Lê Trọng Nghĩa, 'Giáo trình thực tập cung cấp điện', ISBN: 978-604-73-1695-5, Nhà xuất bản Đại Học Quốc Gia Thành Phố Hồ Chí Minh, 2014.

Xác nhận của cơ quan

Thủ Đức, ngày tháng 6 năm 2018

Người khai ký tên

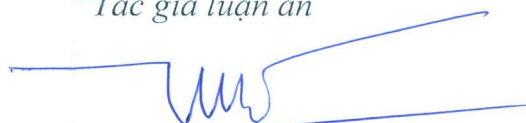


LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của tôi. Các số liệu, kết quả nêu trong luận án là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tp. Hồ Chí Minh, ngày tháng 6 năm 2018.

Tác giả luận án



Nguyễn Ngọc Âu

LỜI CẢM ƠN

Xin chân thành cảm ơn PGS.TS Quyền Huy Ánh, PGS.TS Phan Thị Thanh Bình đã tận tình hướng dẫn NCS trong quá trình thực hiện luận án.

Xin chân thành cảm ơn Ban Giám Hiệu, Phòng Đào Tạo và Bộ Phận Quản Lý Sau Đại Học, Bộ Môn Điện Công Nghiệp, Khoa Điện Điện Tử Trường Đại Học Sư Phạm Kỹ Thuật Thành phố Hồ Chí Minh đã tạo điều kiện tốt nhất cho NCS trong quá trình thực hiện luận án.

Xin chân thành cảm ơn các bạn NCS đã chia sẻ kinh nghiệm nghiên cứu trong quá trình thực hiện luận án.

Cảm ơn gia đình đã ủng hộ và chịu thiệt thòi nhất định trong thời gian NCS thực hiện luận án.

Tác giả

Nguyễn Ngọc Âu

TÓM TẮT

Hệ thống điện hiện đại đối mặt các thách thức mới với rất nhiều thay đổi. Đầu tư nguồn phát không đáp ứng kịp nhu cầu phát triển phụ tải gây áp lực lên hệ thống điện phải vận hành gần giới hạn ổn định. Trong khi đó, hệ thống điện vận hành luôn gặp phải các trường hợp sự cố bất thường. Các sự cố này gây hại đến ổn định động hệ thống điện và có thể dẫn đến tan rã hệ thống điện. Đánh giá ổn định quá độ của hệ thống điện trong những dao động lớn do sự cố gây ra, các phương pháp truyền thống tỏ ra kém hiệu quả. Vì vậy, một nhu cầu là cần hệ thống nhận dạng nhanh mất ổn định động hệ thống điện nhằm cảnh báo sớm thì còn cơ hội để điều khiển đưa hệ thống điện trở về trạng thái ổn định. Thế giới đang chứng kiến sự trỗi dậy một cách mạnh mẽ của trí thông minh nhân tạo ứng dụng vào lĩnh vực khoa học công nghệ trong những năm đầu thế kỷ 21 và thời gian sắp tới. Trong đó, ứng dụng công nghệ tính toán thông minh nhân tạo trong chẩn đoán ổn định hệ thống điện được nhiều nhà khoa học quan tâm. Qua nghiên cứu, tác giả nhận thấy ứng dụng mạng thần kinh nhân tạo cho bài toán chẩn đoán ổn định hệ thống điện gặp phải vấn đề phức tạp của dữ liệu cần xử lý để nâng cao độ chính xác. Luận án đề xuất hướng nghiên cứu ứng dụng hệ thống thông minh vào xây dựng mô hình chẩn đoán ổn định động hệ thống điện gồm ba vấn đề cụ thể cần giải quyết: Một là vấn đề lựa chọn tập biến. Hai là vấn đề giảm không gian mẫu. Ba là cải tiến mô hình mạng nơ-ron để nâng cao độ chính xác nhận dạng.

Tóm tắt các đóng góp chính của luận án như sau:

1. Đề xuất quy trình xây dựng tập biến cho bài toán nhận dạng ổn định động hệ thống điện. Giới thiệu và áp dụng thành công hai giải thuật chọn biến, SFFS và FR, vào giảm không gian biến một cách hiệu quả. Kết quả kiểm tra trên sơ đồ IEEE 39-bus, giải thuật SFFS cho kết quả chọn biến tốt hơn phương pháp FR nhờ vào mở rộng không gian tìm kiếm của nó.

2. Áp dụng thành công giải thuật giảm không gian mẫu K-means lai (Hybrid K-means – HK), cải tiến từ giải thuật phân cụm dữ liệu nổi tiếng K-means, khai phá dữ liệu ổn định động hệ thống điện. Kết quả kiểm tra cho thấy giải thuật HK đã khắc phục được nhược điểm K-means, giúp nâng cao chất lượng phân cụm dữ liệu, giảm không gian mẫu một cách hiệu quả. Việc giảm không gian mẫu làm cho mỗi nhóm mẫu con có một tâm đại diện. Điều này có ý nghĩa rất quan trọng là làm cho mô hình không tăng dung lượng bộ nhớ lưu trữ mẫu mới, linh hoạt trong việc cập nhật làm giàu tri thức mới, và giúp mô hình nâng cao khả năng bao phủ dữ liệu, thích nghi với mẫu mới rất cao.
3. Đề xuất quy trình xây dựng hệ thống nhận dạng thông minh chẩn đoán nhanh ổn định động hệ thống điện dựa trên cơ sở mạng nơ-ron, khai phá dữ liệu. Luận án đã phát triển thành công mô hình mạng nơ-ron song song cải tiến cho bài toán nhận dạng ổn định động hệ thống điện. Mô hình đề xuất đạt được mục tiêu quan trọng là nâng cao độ chính xác phân lớp. Với cách tiếp cận trình bày trong luận án, các khâu thực hiện đã được quy trình hóa, điều này giúp mô hình nhận dạng hoàn toàn có thể mở rộng cho các trường hợp nhiễu lớn khác nhau.
4. Bài toán nhận dạng trong đề tài có thể tổng hợp cho bài toán điều khiển ổn định hệ thống điện chẳng hạn như điều khiển sa thải phụ tải.
5. Xây dựng cách tiếp cận nhận dạng nhanh trạng thái ổn định động hệ thống điện khi có sự cố ngắn mạch nghiêm trọng xảy ra chỉ dựa trên các biến đầu vào là độ thay đổi công suất tác dụng tải, độ thay đổi công suất tác dụng trên các nhánh, và sụt điện áp tại các nút. Kết quả kiểm tra cho độ chính xác nhận dạng cao về trạng thái ổn định hệ thống điện mà không cần giải hệ phương trình vi phân như phương pháp truyền thống.
6. Hệ thống nhận dạng thông minh được xây dựng có thể được sử dụng như công cụ trợ giúp các điều độ viên đề ra chiến lược điều khiển trong những tình huống khẩn cấp, và huấn luyện các điều độ viên hệ thống điện xử lý các tình huống dựa trên các kịch bản sự cố.

7. Việc giảm không gian biến có ý nghĩa rất lớn trong việc giảm chi phí thiết bị đo lường cảm biến, đơn giản hóa sơ đồ đấu nối, và giúp hệ thống tăng tốc độ xử lý. Việc giảm không gian mẫu có ý nghĩa quan trọng trong việc tiết kiệm bộ nhớ lưu trữ mẫu, giảm chi phí thu thập dữ liệu, và giúp mô hình dễ dàng cập nhật dữ liệu mới.

ABSTRACT

Modern power systems are forced to operate under highly stressed operating conditions closer to their stability limits due to the rapid growth of electricity demand. The power system operates in the event of unusual problems. These problems damage the power system and can lead to system blackout. Evaluating the transient stability of the power system in large oscillations caused by the incident, traditional methods are ineffective. The need is to fast identify the unstable power system for early warning, so that the opportunity driving power system into re-stability state can be easier. The world is witnessing a strong rise of artificial intelligence applied to the field of science and technology in the early 21st century and in the coming time. In particular, artificial intelligence applications in the diagnosis of power system stability are many scientists concerned. In study, the author found that the artificial neural network application for the power system diagnostics encountered a complex problem of data processing and classification accuracy. From analyzing, the author proposed three specific issues to solve: feature selection, sample reduction space, and improved classification accuracy.

Summary of theoretical and academic contribution of the dissertation:

1. Propose the process of building the feature set for dynamic power system stability prediction. Successfully applied two reduction feature space algorithms that are SFFS and FR. The tested results on IEEE 39-bus power net showed that the SFFS algorithm gives better results than the FR method because of its expanded search space.
2. Successfully applied Hybrid K-means algorithm in data mining of power system stability, the test results show that the HK algorithm overcomes the K-means disadvantage, improves the quality of clustering, and reduces the sample space efficiently. Reducing the sample space, each cluster has a representative center. This helps the model not increase memory capacity, flexibility in updating knowledge, and improve data coverage.
3. With the approach presented in the thesis, the implementation process has been clearly defined, which makes the model can completely expand to the case of different large noises.
4. The problem of recognition in the thesis can be synthesized for the problem of controlling the power system stability such as the load shedding control.

5. Approach of building fast recognition method for dynamic power system stability prediction is caused by faults. The selected features are variables that characterize at fault-on mode of power system such as voltage drops in the nodes, changes in active power flows in transmission lines, and nodal active powers. The test results have high accuracy classification of the stability status of the power system without solving the differential equation as a traditional method.
6. The built intelligent recognition system can be used as a tool to assist operators to develop control strategies in emergency situations, and a training tool for handling situations based on incident scenarios.
7. Feature reduction space is significant in reducing the cost of sensor measurement, simplifying the connection diagram, and helping the system to work faster. Sample reduction space is important for the system to save sample storage space, reduce data collection cost, and update data.

MỤC LỤC

Trang tựa	Trang
Quyết định giao đề tài	
LỜI CAM ĐOAN	iii
LỜI CẢM ƠN	iv
TÓM TẮT	v
MỤC LỤC.....	x
CÁC CHỮ VIẾT TẮT.....	xvi
CÁC KÝ HIỆU	xviii
CÁC THUẬT NGỮ	xix
DANH MỤC HÌNH	xx
CHƯƠNG 1. MỞ ĐẦU	
1.1 Tính cần thiết.....	1
1.2 Mục tiêu của luận án.....	2
1.3 Đối tượng và phạm vi nghiên cứu	2
1.4 Cách tiếp cận và phương pháp nghiên cứu.....	3
1.5 Điểm mới về mặt khoa học của luận án	3
1.6 Ý nghĩa thực tiễn của luận án	4
1.7 Bố cục của luận án.....	4

CHƯƠNG 2. TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU ĐÁNH GIÁ ỔN ĐỊNH HỆ THỐNG ĐIỆN

2.1	Tổng quan	6
2.2	Ổn định động hệ thống điện	6
2.2.1	Các chế độ làm việc hệ thống.....	6
2.2.2	Ổn định hệ thống điện	8
2.3	Các phương pháp phân tích ổn định động hệ thống điện	9
2.3.1	Mô hình toán học hệ thống điện nhiều máy	9
2.3.2	Phương pháp tích phân số	9
2.3.3	Phương pháp diện tích.....	11
2.3.4	Phương pháp trực tiếp	11
2.3.5	Phương pháp mô phỏng theo miền thời gian	14
2.4	Các nghiên cứu khoa học liên quan.....	15
2.5	Tóm tắt chương 2.....	20

CHƯƠNG 3. CÔNG NGHỆ TÍNH TOÁN THÔNG MINH 23

3.1	Mạng nơ-ron nhân tạo	23
3.1.1	Giới thiệu về mạng nơ-ron nhân tạo.....	23
3.1.2	Mô hình nơ-ron sinh học	23
3.1.3	Mô hình nơ-ron nhân tạo.....	24
3.1.3.1	Cấu trúc mô hình một nơ-ron nhân tạo.....	24

3.1.3.2	Cấu trúc mô hình mạng nơ-ron nhân tạo nhiều lớp	26
3.1.4	Các qui tắc học	27
3.1.5	Ánh xạ và lan truyền ngược	30
3.1.6	Mặt lỗi	31
3.1.7	Cực tiểu cục bộ.....	31
3.1.8	Phân loại mạng nơ-ron	32
3.1.9	Mạng Perceptron nhiều lớp	32
3.1.10	Mạng nơ-ron Generalized Regression Nơ-ron Network (GRNN)....	33
3.1.10.1	Hàm cơ sở bán kính RBF	33
3.1.10.2	Kỹ thuật hàm cơ sở bán kính RBF	33
3.1.10.3	Mạng hàm cơ sở bán kính RBF	34
3.1.10.4	Mạng Generalized Regression Nơ-ron Network	35
3.2	Bộ phân lớp K-Nearest Neighbor	36
3.3	Bộ phân lớp máy vector hỗ trợ	37
3.3.1	Bộ phân lớp nhị phân tách lớp tuyến tính	37
3.3.2	Biên tối ưu mềm	39
3.3.3	Hàm nhân.....	41
3.4	Tóm tắt chương 3.....	41

CHƯƠNG 4. LỰA CHỌN THÔNG SỐ ĐẶC TRƯNG CHO HỆ THỐNG NHẬN DẠNG THÔNG MINH TRONG CHẨN ĐOÁN ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

4.1	Xây dựng tập mẫu.....	43
4.2	Lựa chọn tập biến	44
4.2.1	Tiêu chuẩn chọn biến	46
4.2.1.1	Tiêu chuẩn Fisher	46
4.2.1.2	Tiêu chuẩn Divergence.....	46
4.2.1.3	Tiêu chuẩn Scatter Matrix (SM).....	47
4.2.2	Các giải thuật tìm kiếm tập biến.....	48
4.2.2.1	Giải thuật Sequential Forward Floating Selection (SFFS):	48
4.2.2.2	Giải thuật xếp hạng biến (FR-Feature Ranking):	50
4.3	Tóm tắt chương 4.....	50

CHƯƠNG 5. PHƯƠNG PHÁP XÂY DỰNG HỆ THỐNG NHẬN DẠNG THÔNG MINH CHẨN ĐOÁN ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

5.1	Giới thiệu	52
5.2	Xây dựng tập mẫu.....	52
5.3	Nghiên cứu đề xuất quy trình chọn biến đặc trưng	55
5.4	Trích xuất tri thức	57
5.5	Đánh giá.....	58
5.6	Biểu diễn kết quả	59

5.7	Nghiên cứu đề xuất quy trình giảm không gian mẫu	63
5.7.1	Giải thuật KM.....	64
5.7.2	Giải thuật Hybrid K-means (HK).....	66
5.7.3	Đề xuất quy trình giảm không gian mẫu	67
5.8	Nghiên cứu đề xuất mô hình bộ phân lớp cải tiến nâng cao độ chính xác nhận dạng (Advanced parallel Classifier Model-APCM).....	68
5.8.1	Xây dựng mô hình phân lớp lai dựa vào năng lượng mẫu	68
5.8.2	Mô hình HCM đề nghị	69
5.8.3	Ý tưởng xây dựng mô hình APCM dựa vào kỹ thuật phân cụm dữ liệu	70
5.8.3.1	Quy trình xây dựng mô hình APCM.....	72
5.8.3.2	Vận hành mô hình	73
5.9	Tóm tắt chương 5.....	75
CHƯƠNG 6. ỨNG DỤNG MÔ HÌNH APCM CHẨN ĐOÁN NHANH ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN IEEE 39-BUS		
6.1	Giới thiệu sơ đồ hệ thống điện chuẩn IEEE 39-bus	77
6.2	Xây dựng tập mẫu và tập biến ban đầu	78
6.3	Chọn biến đặc trưng	79
6.4	Giảm không gian mẫu.....	83
6.5	Áp dụng mô hình cải tiến	90
6.6	Nhận xét.....	95

6.7	Tóm tắt chương 6.....	97
CHƯƠNG 7. KẾT LUẬN		
7.1	Kết quả đạt được.....	99
7.2	Hướng phát triển.....	101
TÀI LIỆU THAM KHẢO.....		102
PHỤ LỤC		

CÁC CHỮ VIẾT TẮT

1. ANN (Artificial Nơ-ron Network)	Mạng Nơ-ron nhân tạo
2. Advanced parallel Classifier Model (APCM)	Mô hình bộ phân lớp song song cải tiến
3. AccRate	Độ chính xác tổng thể (phân lớp)
4. CCT (Critical Clearing Time)	Thời gian cắt tới hạn
5. CL (Cluster)	Cụm
6. CL(S), CL(U), CL(S,U):	cụm dữ liệu S, U và cụm (S,U)
7. D(S), D(U), D(S,U)	tập dữ liệu S, U và gồm (S,U)
8. FCT (Fault Clearing Time)	Thời gian cắt sự cố
9. FP (false positives)	Sai dương
10. FN (false negatives)	Sai âm
11. F-Score	Chỉ số F-Score
12. FR (Feature Ranking)	Xếp hạng biến
13. GRNN (Generalized Regression Nơ-ron Network)	Mạng Nơ-ron hồi quy tổng quát
14. HK (Hybrid K-means)	Giải thuật K-means lai
15. HTĐ	Hệ Thống Điện
16. IS (Intelligent System)	Hệ thống thông minh
17. K-NN (K-Nearest Neighbor)	K láng giềng gần nhất
18. KM (K-means)	Giải thuật K-means
19. MLP (Multi-layered Feedforward Nơ-ron Network)	Mạng nơ-ron truyền thẳng nhiều lớp
20. MMD (Max-Min Distance)	Khoảng cách Max-Min
21. Pre (Precision)	Độ chính xác
22. Rec (Recall)	Độ hồi tưởng

23. SVM (Support Vecotor Machine)	Vec-tơ máy hỗ trợ
24. SFS (Sequential Forward Selection)	Lựa chọn tiến
25. SFFS (Sequential Forward Floating Selection)	Lựa chọn tiến lùi
26. SSM (Sequential Search Method)	Phương pháp tuần tự
27. S (Stable)	Ổn định
28. TN (true negatives)	Đúng âm
29. TP (true positives)	Đúng dương
30. tr (trace)	Tổng đường chéo của ma-trận
31. U (Unstable)	Không ổn định

CÁC KÝ HIỆU

1. δ Góc của rotor của máy phát
2. ω Tốc độ góc của rotor máy phát
3. P_m Công suất cơ của máy phát
4. P_e Công suất điện của máy phát
5. M Mô-men quán tính của máy phát
6. w Trọng số
7. $\|\cdot\|$ Khoảng cách Euclide
8. m Giá trị trung bình dữ liệu
9. σ Độ lệch chuẩn của dữ liệu
10. $+$ Mẫu lớp S
11. $-$ Mẫu lớp U
12. 1 Nhãn lớp ổn định
13. 0 Nhãn lớp không ổn định
14. Sp (Spread) Thông số độ rộng
15. N_i Số nơ-ron ẩn
16. N_{int} Số lần lặp của giải thuật thuật phân cụm dữ liệu KM
17. D_E Giá trị mục tiêu của giải thuật phân cụm dữ liệu KM

CÁC THUẬT NGỮ

1. Artificial Intelligence: thông minh nhân tạo, là lĩnh vực nghiên cứu tính toán trên máy tính bắt chước hành vi thông minh của con người ⁽¹⁾.
2. Classification: phân lớp, là nhóm, phân loại, hay phân chia theo cùng đặc tính ⁽¹⁾.
3. Intelligence: thông minh, là khả năng có thể học, hiểu và suy luận ⁽¹⁾.
4. Prediction: dự báo hay chẩn đoán, là dự báo hay chẩn đoán điều sẽ xảy ra ⁽¹⁾.
5. Recognition: nhận dạng, là hoạt động nhớ đối tượng hoặc hoạt động nhận biết đối tượng ⁽¹⁾.
6. System: hệ thống, là một nhóm các phần tử, các phần tử của thiết bị,... chúng được kết nối với nhau hoặc làm việc cùng nhau ⁽¹⁾.

⁽¹⁾ Oxford Dictionary

DANH SÁCH CÁC HÌNH

TRANG

Hình 2.1	Trạng thái các điểm vận hành của hệ thống điện.....	7
Hình 2.2	Phân loại ổn định hệ thống điện theo IEEE/CIGRE	8
Hình 2.3	Đường cong dao động $\delta(t)$	10
Hình 3.1	Mô hình nơ-ron sinh học	24
Hình 3.2	Mô hình một nơ-ron nhân tạo	25
Hình 3.3	Cấu trúc mô hình mạng nơ-ron nhân tạo nhiều lớp truyền thẳng	27
Hình 3.4	Mô hình học có giám sát	29
Hình 3.5	Mô hình học củng cố.....	29
Hình 3.6	Mô hình học không có giám sát	30
Hình 3.7	Cấu trúc mạng RBF.....	35
Hình 3.8	Sơ đồ cấu trúc mạng nơ-ron GRNN.....	36
Hình 3.9	Sơ đồ minh họa thuật toán SVM.....	38
Hình 3.10	Minh họa cho trường hợp dữ liệu hai lớp không hoàn toàn tách biệt.....	40
Hình 4.1	Kỹ thuật chọn biến Filter.	45
Hình 4.2	Kỹ thuật chọn biến Wrapper.	46
Hình 5.1	Các khâu cơ bản của quá trình xây dựng hệ thống nhận dạng.....	52
Hình 5.2	Quy trình xây dựng tập mẫu.....	54
Hình 5.3	Quy trình chọn biến đặc trưng F&W	57

Hình 5.4 Sơ đồ thiết kế màn hình quan sát	60
Hình 5.5 Màn hình biểu diễn trạng thái ổn định hệ thống điện	61
Hình 5.6 Sơ đồ thiết kế màn hình giám sát trong không gian hai chiều	63
Hình 5.7 Quy trình giảm không gian dữ liệu	67
Hình 5.8 Quy trình xây dựng mô hình HCM theo năng lượng mẫu	68
Hình 5.9 Mô hình HCM đề nghị	69
Hình 5.10 Minh họa ý tưởng xây dựng mô hình APCM	72
Hình 5.11 Quy trình xây dựng mô hình mạng nơ-ron cải tiến.....	73
Hình 5.12 Mô hình APCM đề nghị.....	73
Hình 6.1 Sơ đồ hệ thống điện IEEE 39 bus	77
Hình 6.2 Xếp hạng biến theo chuẩn Fisher	79
Hình 6.3 Xếp hạng biến theo chuẩn Divergence	80
Hình 6.4 Giá trị khoảng cách tính toán biến bằng giải thuật SFFS theo chuẩn SM .	80
Hình 6.5 Đánh giá chọn tập biến.....	81
Hình 6.6 Các nhóm mẫu của tập S và tập U được rút gọn.....	84
Hình 6.7 Đặc tuyến hội tụ của giải thuật HK và KM tại $S_R=900$	84
Hình 6.8 Đặc tuyến hội tụ của giải thuật HK và KM tại $U_R=400$	85
Hình 6.9 Các khả năng kết hợp hình thành tập mẫu rút gọn.....	85
Hình 6.10 Đánh giá độ chính xác phân lớp, rút dữ liệu với giải thuật KM, 1-NNC	86
Hình 6.11 Đánh giá độ chính xác phân lớp, rút dữ liệu với giải thuật HK, 1-NNC	87

Hình 6.12 Thực nghiệm tìm số nơ-ron ẩn.....	88
Hình 6.13 Thực nghiệm tìm hệ số Spread.....	89
Hình 6.14 Đồ thị độ chính nhận dạng với số Bộ phân lớp song song	92
Hình 6.15 Mô hình APCM gồm 5 bộ phân lớp song song GRNNC	94

DANH SÁCH CÁC BẢNG

	TRANG
Bảng 3.1 Một số hàm phi tuyến thường dùng trong các mô hình nơ-ron.....	26
Bảng 4.1 Giải thuật chọn biến SFS	49
Bảng 4.2 Giải thuật chọn biến SFFS.....	49
Bảng 4.3 Giải thuật xếp hạng biến (FR)	50
Bảng 5.1 Ma trận lẫn lộn kích thước 2x2.....	58
Bảng 5.2 Các chỉ số độ đo đánh giá bộ phân lớp dựa vào ma trận lẫn lộn.....	59
Bảng 5.3 Giải thuật KM.....	65
Bảng 5.4 Giải thuật HK.....	66
Bảng 6.1 Độ chính xác kiểm tra phân lớp 1-NN, d=15 và d=104, kfold=10	82
Bảng 6.2 Các biến được chọn	82
Bảng 6.3 Độ chính xác kiểm tra phân lớp 1-NN với không gian mẫu giảm	87
Bảng 6.4 Độ chính xác phân lớp với MLPC, GRNNC và SVMC.....	90
Bảng 6.5 Độ chính xác, độ hồi tưởng và hệ số F-Score.....	90
Bảng 6.6 Các nhóm mẫu được tách ra từ CL(900,400)	91
Bảng 6.7 Kết quả huấn luyện trung bình gồm 5 bộ MLPC song song	92
Bảng 6.8 Kết quả huấn luyện gồm 5 bộ GRNNC song song.....	93
Bảng 6.9 Kết quả huấn luyện gồm 5 bộ SVMC song song	93
Bảng 6.10 Độ chính xác, độ hồi tưởng và hệ số F-Score của mô hình APCM	93

Bảng 6.11 Tóm tắt kết quả xử lý dữ liệu.....94

Chương 1. MỞ ĐẦU

1.1 TÍNH CẦN THIẾT

Hệ thống điện hiện đại đối mặt các thách thức mới với rất nhiều thay đổi. Đầu tư nguồn phát không đáp ứng kịp nhu cầu phát triển phụ tải gây áp lực lên hệ thống điện phải vận hành gần giới hạn ổn định. Trong khi đó, hệ thống điện vận hành luôn gặp phải các trường hợp bất thường như cắt điện máy phát, cắt điện đường dây, gia tăng tải đột ngột, cắt máy biến áp, và nguy hiểm nhất là ngắn mạch xảy ra, Các sự cố này gây hại đến ổn định động hệ thống điện và có thể dẫn đến tan rã hệ thống điện. Trên thế giới đã xảy ra rất nhiều sự cố tan rã hệ thống điện nghiêm trọng gắn liền với sự mất ổn định của hệ thống điện và đã gây ra những thiệt hại kinh tế vô cùng to lớn. Hệ thống điện ổn định là khả năng hệ thống điện, từ điều kiện vận hành ban đầu, chịu đựng được các sự cố bất thường và trở về trạng thái cân bằng sau sự cố, tiếp tục duy trì liên tục cung cấp điện cho khách hàng. Xử lý quá trình quá độ trong hệ thống điện có một giá trị rất quan trọng để nâng cao độ tin cậy và tính ổn định chế độ vận hành hệ thống điện. Các khó khăn trong việc điều khiển chế độ hệ thống điện liên quan đến tính phức tạp của chế độ điều khiển, khó khăn trong mô tả toán học của quá trình xảy ra trong hệ thống điện. Để đánh giá tính ổn định quá độ của hệ thống điện phức tạp nhiều máy trong những dao động lớn do sự cố gây ra, các phương pháp truyền thống tỏ ra kém hiệu quả và không thuận lợi, đặc biệt trong những điều kiện bất định, thiếu thông tin và yêu cầu khắc khe về thời gian giải. Vì vậy, một nhu cầu là cần hệ thống chẩn đoán nhanh mất ổn định động hệ thống điện nhằm cảnh báo sớm thì còn cơ hội để điều khiển đưa hệ thống điện trở về trạng thái ổn định. Ứng dụng công nghệ trí thức đánh giá ổn định hệ thống điện được nhiều tác giả quan tâm và tập trung nghiên cứu trong những năm gần đây. Tuy nhiên, ứng dụng công nghệ trí thức gặp phải hai vấn đề khó khăn ảnh hưởng đến độ chính xác nhận dạng:

- Một là: biến đầu vào lớn dẫn đến tăng chi phí đo lường cảm biến, và có những biến dư thừa có thể gây nhiễu khi học.

- Hai là: dữ liệu mẫu lớn gây nên rất tốn kém chi phí thu thập dữ liệu, tăng bộ nhớ lưu trữ, gây khó khăn khi học.

Các vấn đề trên ảnh hưởng đến độ chính xác nhận dạng cho mô hình ứng dụng hệ thống thông minh nhân tạo. Vì vậy, nghiên cứu chọn biến đặc trưng và mẫu đại diện, cũng như nghiên cứu xây dựng mô hình nhận dạng thông minh trong chẩn đoán ổn định động hệ thống điện trên cơ sở mạng nơ-ron nhằm nâng cao độ chính xác là rất cần thiết. Đó cũng là động cơ để Nghiên cứu sinh chọn đề tài nghiên cứu ‘**Hệ thống nhận dạng thông minh chẩn đoán nhanh ổn định động hệ thống điện**’.

1.2 MỤC TIÊU CỦA LUẬN ÁN

- Nghiên cứu các phương pháp giảm kích thước không gian dữ liệu.
- Xây dựng hệ thống nhận dạng thông minh dựa trên cơ sở mạng nơ-ron cải tiến có khả năng chẩn đoán nhanh ổn định động hệ thống điện với độ chính xác cao.

1.3 ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

- Đối tượng nghiên cứu:
 - Nghiên cứu quy trình xây dựng hệ thống thông minh chẩn đoán ổn định động hệ thống ứng dụng kỹ thuật khai phá dữ liệu và công nghệ thông minh nhân tạo với độ chính xác cao.
 - Đánh giá hiệu quả của phương pháp đề xuất trên hệ thống điện chuẩn IEEE 39-bus.
- Phạm vi nghiên cứu:

Đề tài tập trung nghiên cứu ứng dụng tính toán thông minh nhân tạo, khai phá dữ liệu vào giải quyết vấn đề chẩn đoán nhanh ổn định động hệ thống điện do các sự cố gây ra. Phạm vi nghiên cứu tập trung vào các vấn đề sau:

 - Nghiên cứu giảm không gian biến.
 - Nghiên cứu giảm không gian mẫu.

- Nghiên cứu cải tiến mô hình mạng nơ-ron nhằm nâng cao độ chính xác nhận dạng.

1.4 CÁCH TIẾP CẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU

- Nghiên cứu các tài liệu, các bài báo, các sách về chẩn đoán ổn định hệ thống điện.
- Nghiên cứu phương pháp chọn biến hay giảm không gian biến.
- Nghiên cứu phương pháp giảm không gian mẫu.
- Xây dựng hệ thống nhận dạng thông minh dựa trên cơ sở mạng nơ-ron, khai phá dữ liệu.
- Đánh giá phương pháp đề xuất trên sơ đồ lưới điện chuẩn IEEE 39-bus.
- Nghiên cứu với sự trợ giúp của các phần mềm Matlab, PowerWorld.

1.5 ĐIỂM MỚI VỀ MẶT KHOA HỌC CỦA LUẬN ÁN

- Đề xuất quy trình xây dựng tập biến cho bài toán nhận dạng ổn định động hệ thống điện. Giới thiệu và áp dụng thành công hai giải thuật chọn biến, SFFS và FR, vào giảm không gian biến một cách hiệu quả. Kết quả kiểm tra trên sơ đồ IEEE 39-bus, giải thuật SFFS cho kết quả chọn biến tốt hơn phương pháp FR nhờ vào mở rộng không gian tìm kiếm của nó.
- Áp dụng thành công giải thuật giảm không gian mẫu K-means lai (Hybrid K-means – HK), cải tiến từ giải thuật phân cụm dữ liệu nổi tiếng K-means, khai phá dữ liệu ổn định động hệ thống điện. Kết quả kiểm tra cho thấy giải thuật HK đã khắc phục được nhược điểm giải thuật K-means, giúp nâng cao chất lượng phân cụm dữ liệu, giảm không gian mẫu một cách hiệu quả. Việc giảm không gian mẫu làm cho mỗi nhóm mẫu con có một tâm đại diện. Điều này có ý nghĩa rất quan trọng là làm cho mô hình không tăng dung lượng bộ nhớ lưu trữ mẫu mới, linh hoạt trong việc cập nhật làm giàu tri thức mới, và giúp mô hình nâng cao khả năng bao phủ dữ liệu, thích nghi với mẫu mới rất cao.

- Đề xuất quy trình xây dựng hệ thống nhận dạng thông minh chẩn đoán nhanh ổn định động hệ thống điện dựa trên cơ sở mạng nơ-ron, khai phá dữ liệu. Luận án đã phát triển thành công mô hình mạng nơ-ron song song cải tiến cho bài toán nhận dạng ổn định động hệ thống điện. Mô hình đề xuất đạt được mục tiêu quan trọng là nâng cao độ chính xác phân lớp. Với cách tiếp cận trình bày trong luận án, các khâu thực hiện đã được quy trình hóa, điều này giúp mô hình nhận dạng hoàn toàn có thể mở rộng cho các trường hợp nhiễu lớn khác nhau.
- Bài toán nhận dạng trong đề tài có thể tổng hợp cho bài toán điều khiển ổn định hệ thống điện chẳng hạn như điều khiển sa thải phụ tải.

1.6 Ý NGHĨA THỰC TIỄN CỦA LUẬN ÁN

- Xây dựng cách tiếp cận nhận dạng nhanh trạng thái ổn định động hệ thống điện khi có sự cố ngắn mạch nghiêm trọng xảy ra chỉ dựa trên các biến đầu vào là độ thay đổi công suất tác dụng tải, độ thay đổi công suất tác dụng trên các nhánh, và độ sụt điện áp tại các nút. Kết quả kiểm tra cho độ chính xác nhận dạng cao về trạng thái ổn định hệ thống điện mà không cần giải hệ phương trình vi phân như phương pháp truyền thống.
- Hệ thống nhận dạng thông minh được xây dựng có thể được sử dụng như công cụ trợ giúp các điều độ viên đề ra chiến lược điều khiển trong những tình huống khẩn cấp, và huấn luyện các điều độ viên hệ thống điện xử lý các tình huống dựa trên các kịch bản sự cố.
- Việc giảm không gian biến có ý nghĩa rất lớn trong việc giảm chi phí thiết bị đo lường cảm biến, đơn giản hóa sơ đồ đấu nối, và giúp hệ thống tăng tốc độ xử lý. Việc giảm không gian mẫu có ý nghĩa quan trọng trong việc tiết kiệm bộ nhớ lưu trữ mẫu, giảm chi phí thu thập dữ liệu, và giúp mô hình dễ dàng cập nhật dữ liệu.

1.7 BỐ CỤC CỦA LUẬN ÁN

Bố cục của luận án bao gồm 7 chương:

Chương 1. Mở đầu

Chương 2. Tổng quan tình hình nghiên cứu đánh giá ổn định hệ thống điện

Chương 3. Công nghệ tính toán thông minh

Chương 4. Lựa chọn thông số đặc trưng cho hệ thống nhận dạng thông minh trong chẩn đoán ổn định động hệ thống điện.

Chương 5. Xây dựng hệ thống nhận dạng thông minh trong chẩn đoán nhanh ổn định động hệ thống điện

Chương 6. Ứng dụng mô hình mạng nơ-ron cải tiến chẩn đoán ổn định động hệ thống điện IEEE 39-bus

Chương 7. Kết luận

Chương 2. **TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU ĐÁNH GIÁ ỔN ĐỊNH HỆ THỐNG ĐIỆN**

2.1 TỔNG QUAN

Hệ thống điện vận hành luôn đối mặt với các trường hợp bất thường như cắt điện một máy phát, cắt điện đường dây, gia tăng tải đột ngột, cắt máy biến áp, và nguy hiểm nhất là các trường hợp ngắn mạch. Các sự cố này gây hại nghiêm trọng đến ổn định hệ thống điện và trên thế giới đã chứng kiến nhiều vụ tan rã hệ thống điện [1]–[3]. Xử lý quá trình quá độ trong hệ thống điện có một giá trị rất quan trọng để nâng cao độ tin cậy và tính ổn định chế độ vận hành hệ thống điện. Các khó khăn trong việc điều khiển chế độ hệ thống điện liên quan đến tính phức tạp của chế độ điều khiển, khó khăn trong mô tả toán học của quá trình xảy ra trong hệ thống điện. Đánh giá tính ổn định quá độ của hệ thống điện trong những dao động lớn do sự cố gây ra, các phương pháp truyền thống tỏ ra kém hiệu quả và không thuận lợi, đặc biệt trong những điều kiện bất định, thiếu thông tin và yêu cầu khắc khe về thời gian giải. Trong khi, chế độ quá độ dao động lớn do sự cố gây ra cần phải xử lý nhanh thì còn cơ hội điều khiển đưa hệ thống trở về trạng thái ổn định. Sự trỗi dậy mạnh mẽ của trí thông minh nhân tạo và ứng dụng vào lĩnh vực khoa học công nghệ trong những năm đầu thế kỷ 21 và thời gian sắp tới. Vì vậy, ứng dụng công nghệ mới nổi, tính toán thông minh nhân tạo, vào chẩn đoán ổn định hệ thống điện [4] là hướng nghiên cứu hiện đại, được nhiều nhà khoa học quan tâm. Tóm lại, phương pháp tính toán thông minh nhân tạo có khả năng giải quyết bài toán rất phức tạp trong phân tích ổn định động hệ thống điện [5], [6]. Đây cũng là động cơ để tác giả thực hiện luận án này.

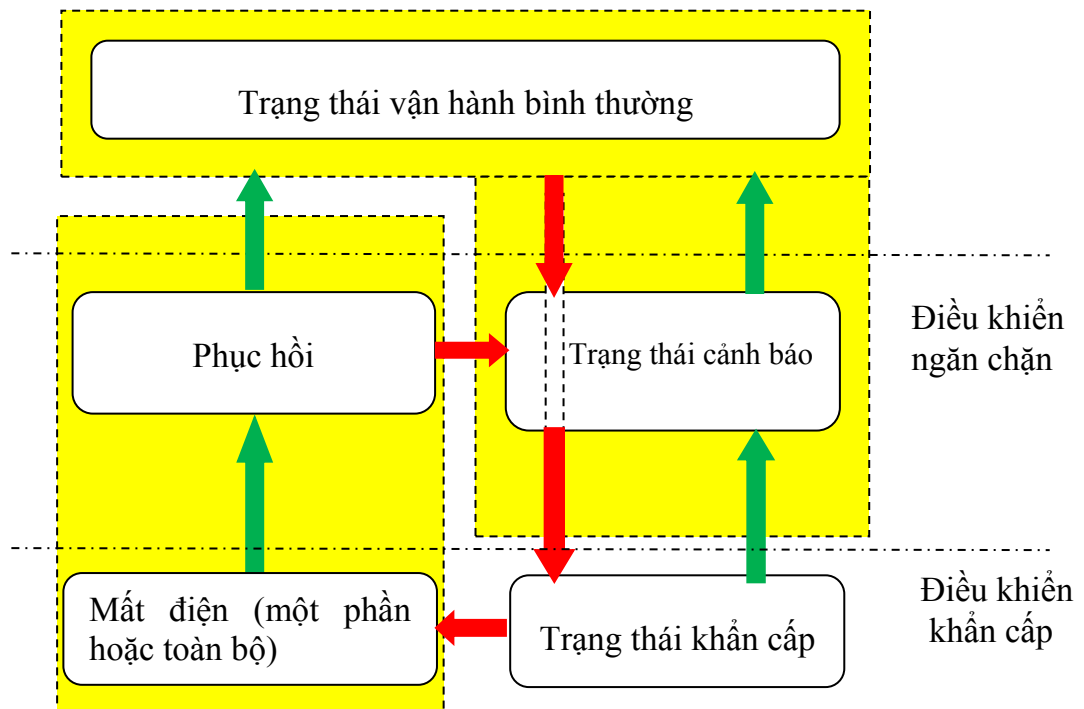
2.2 ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

2.2.1 Các chế độ làm việc hệ thống

Các chế độ làm việc của hệ thống điện được chia làm 2 loại chính: chế độ xác lập và chế độ quá độ [7].

Chế độ xác lập: là chế độ trong đó các thông số của hệ thống không thay đổi hoặc thay đổi trong những khoảng thời gian tương đối ngắn, chỉ biến thiên nhỏ xung quanh các trị số định mức, là chế độ bình thường và lâu dài của hệ thống điện, còn được gọi là chế độ xác lập bình thường. Chế độ sau sự cố hệ thống được phục hồi và làm việc tạm thời cũng thuộc về chế độ xác lập, mà còn được gọi là chế độ xác lập sau sự cố.

Chế độ quá độ: là chế độ trung gian chuyển từ chế độ xác lập này sang chế độ xác lập khác. Chế độ quá độ thường diễn ra sau những sự cố hoặc thao tác đóng cắt các phần tử đang mang công suất mà thường được gọi là các kích động lớn. Chế độ quá độ được gọi là chế độ quá độ bình thường nếu nó tiến đến chế độ xác lập mới. Trong trường hợp này các thông số hệ thống bị biến thiên nhưng sau một thời gian lại trở về trị số gần định mức và tiếp theo ít thay đổi. Ngược lại, có thể diễn ra chế độ quá độ với thông số hệ thống biến thiên mạnh, sau đó tăng trưởng vô hạn hoặc giảm đến 0. Chế độ quá độ đó được gọi là chế độ quá độ sự cố.



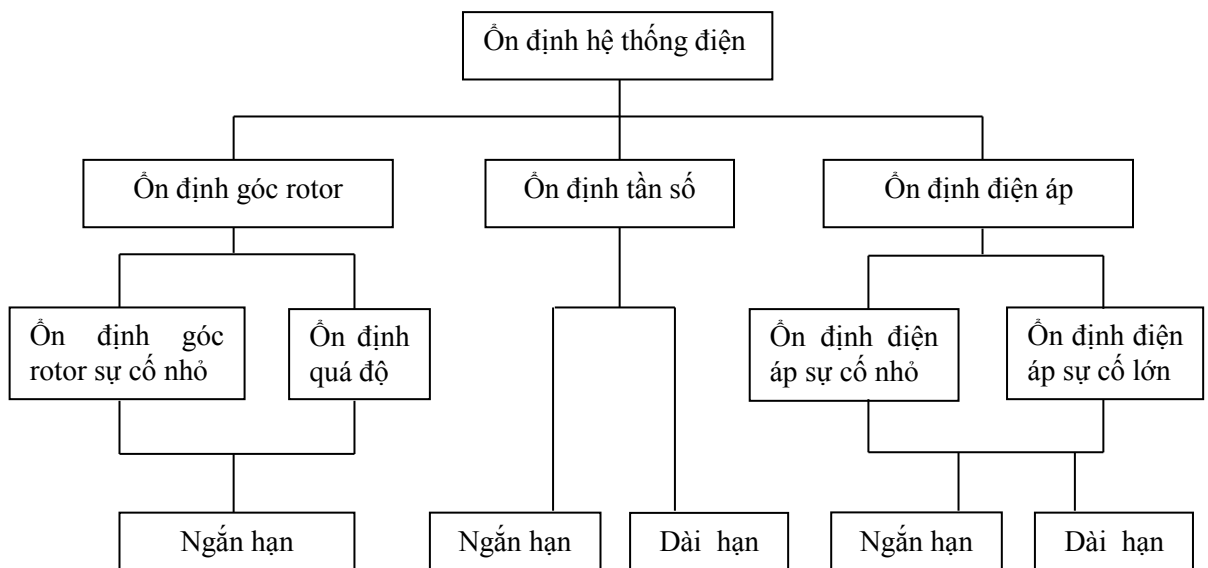
Hình 2.1 Trạng thái các điểm vận hành của hệ thống điện

Hình 2.1 trình bày quá trình thay đổi các trạng thái vận hành của hệ thống điện [8]. Qua đó cho thấy trạng thái vận hành của hệ thống điện luôn phải được theo dõi đánh giá một cách liên tục để đảm bảo vận hành tin cậy, và có biện pháp điều khiển can thiệp kịp thời nếu tình huống xấu xuất hiện.

2.2.2 Ổn định hệ thống điện

Ổn định hệ thống điện đề cập đến khả năng của các máy phát điện dịch chuyển từ một trạng thái vận hành xác lập này đến trạng thái vận hành xác lập khác sau khi bị kích động mà không mất đồng bộ, có hai loại ổn định hệ thống điện : ổn định tĩnh và ổn định động [7], [9]. Ổn định tĩnh là khả năng của hệ thống sau những kích động nhỏ phục hồi được chế độ ban đầu hoặc rất gần với chế độ ban đầu. Ổn định động là khả năng của hệ thống sau những kích động lớn phục hồi được trạng thái vận hành cho phép ban đầu hoặc gần trạng thái vận hành cho phép ban đầu.

Trong [10], [11] ổn định hệ thống điện được phân loại theo IEEE/CIGRE gồm: ổn định góc rotor, ổn định tần số và ổn định điện áp. Phân loại ổn định hệ thống điện được trình bày như Hình 2.2. Ổn định động là ổn định góc rotor hay không ổn định góc rotor sau dao động lớn hay sự cố nghiêm trọng chẳng hạn như ngắn mạch ba pha.



Hình 2.2 Phân loại ổn định hệ thống điện theo IEEE/CIGRE

2.3 CÁC PHƯƠNG PHÁP PHÂN TÍCH ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

Phân tích ổn định động hệ thống điện là xem xét khả năng hệ thống điện chuyển sang làm việc ổn định ở trạng thái xác lập mới sau dao động lớn. Ổn định quá độ do dao động lớn gây ra như sa thải máy phát, sa thải phụ tải, mở đường dây hoặc do sự cố ngắn mạch, nghiêm trọng nhất là ngắn mạch ba pha. Các dao động lớn gây ra mất cân bằng công suất cơ đầu vào và công suất điện ra của máy phát. Khi đó, làm cho rô-tô các máy phát dao động với máy phát khác và gây phá vỡ ổn định hệ thống điện.

Có nhiều phương pháp phân tích ổn định hệ thống điện. Trong phần này trình bày bốn phương pháp áp dụng khảo sát ổn định động hệ thống điện: phương pháp tích phân số, phương pháp diện tích, phương pháp trực tiếp, và phương pháp mô phỏng theo miền thời gian.

2.3.1 Mô hình toán học hệ thống điện nhiều máy

Hành vi động của máy điện phát điện được mô tả theo phương trình vi phân sau [7], [9], [10]:

$$M_i \frac{d^2 \delta_i}{dt^2} = P_{mi} - P_{ei} \quad (2.1)$$

$$\frac{d\delta_i}{dt} = \varpi_i \quad (2.2)$$

Thay (2.2) vào (2.1), khi đó (2.1) trở thành (2.3):

$$M_i \frac{d\varpi_i}{dt} = P_{mi} - P_{ei} \quad (2.3)$$

Trong đó: δ_i là góc của rotor của máy phát thứ i ; ϖ_i là tốc độ góc của rô-tô máy phát thứ i ; P_{mi} : công suất cơ của máy phát thứ i ; P_{ei} : công suất điện của máy phát thứ i ; M_i : mô-men quán tính của máy phát thứ i .

2.3.2 Phương pháp tích phân số

Trong phần này giới thiệu hai phương pháp tích phân số là phương pháp Euler và Phương pháp Runge-Kutta [7], [9], [10].

- Phương pháp Euler:

Cho phương trình vi phân:

$$dx/dt=f(x,t)$$

Ở bước tích phân thứ $n+1$:

$$x_{n+1} = x_n + (dx/dt)_n \cdot \Delta t = x_n + f(x_n) \cdot \Delta t \quad (2.4)$$

- Phương pháp Runge-Kutta:

Phương pháp Runge-Kutta bậc 4 sử dụng biểu thức phức tạp hơn:

$$x_{n+1} = x_n + 1/6 (k_1 + 2k_2 + 2k_3 + k_4) \quad (2.5)$$

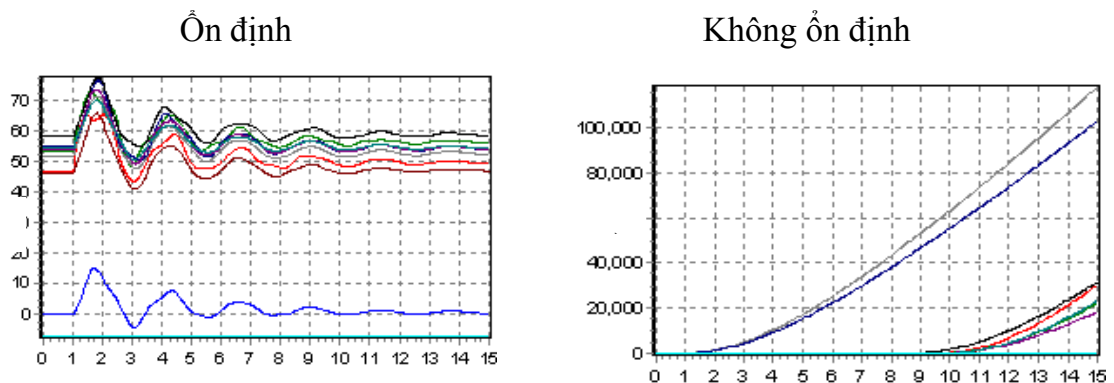
Ở đây: $k_1 = f(x_n, t_n) \cdot \Delta t$;

$$k_2 = f(x_n + k_1/2, t_n + \Delta t/2) \Delta t$$

$$k_3 = f(x_n + k_2/2, t_n + \Delta t/2) \Delta t$$

$$k_4 = f(x_n + k_3, t_n + \Delta t/2) \Delta t$$

Phương pháp tích phân số cho kết quả khá chính xác. Tuy nhiên, trong phương pháp tích phân số cần chú ý vấn đề sai số tích lũy trong các bước giải. Kết quả nhận được từ phương pháp tích phân số sẽ là trị số các góc lệch $\delta(t)$ ở những thời điểm khác nhau. Nếu kết quả tính toán đủ dài thì ở mọi thời điểm có $\delta(t) < 180^\circ$ thì quá trình quá độ ổn định, nếu $\delta(t) \geq 180^\circ$ thì quá trình quá độ không ổn định [7], [12], [13]. Cũng có thể bằng quan sát trực quan dạng đường cong $\delta(t)$ dao động tắt dần hay tăng trưởng vô hạn để kết luận đặc trưng ổn định của hệ thống như Hình 2.3.



Hình 2.3 Đường cong dao động $\delta(t)$

Trạng thái ổn định của hệ thống điện đánh giá theo luật như sau:

$$\left\{ \begin{array}{l} \text{Nếu } \delta_{ij} < 180^0 \text{ thì 'Ổn định'} \\ \text{Nếu } \delta_{ij} \geq 180^0 \text{ thì 'Không ổn định'} \end{array} \right. \quad (2.6)$$

2.3.3 Phương pháp diện tích

Phương trình chuyển động rô-to của hệ thống điện đơn giản [7], [9], [10]:

$$\frac{d^2 \delta}{dt^2} = \frac{P_m - P_e}{T_J} = \frac{P_a}{T_J} \quad (2.7)$$

P_a là công suất tăng tốc.

$$\Leftrightarrow 2 \frac{d\delta}{dt} \frac{d^2 \delta}{dt^2} = 2 \frac{P_a}{T_J} \frac{d\delta}{dt} \quad (2.8)$$

$$\Leftrightarrow \frac{d}{dt} \left(\frac{d\delta}{dt} \right)^2 = \frac{d}{dt} \left(\frac{2}{T_J} \int_{\delta_0}^{\delta} P_a d\delta \right) \Rightarrow \frac{d\delta}{dt} = \sqrt{\left(\frac{2}{T_J} \int_{\delta_0}^{\delta} P_a d\delta \right) + C} \quad (2.9)$$

Với C là hằng số tích phân, $C = 0$ khi $d\delta/dt=0$ và $\delta = \delta_0$, trong đó δ_0 là góc rô-to trước khi có nhiễu. Do đó điều kiện cần thiết để hệ ổn định là tồn tại một góc δ sao cho:

$$\left(\frac{2}{T_J} \int_{\delta_0}^{\delta} P_a d\delta \right) = 0 \quad (2.10)$$

Hệ thống sẽ ổn định nếu diện tích dưới P_a giảm tới không tại một giá trị nào đó của δ (hay diện tích tăng tốc bằng diện tích hãm tốc) nghĩa là phần diện dương bằng diện tích âm.

Giới hạn của phương pháp: Đối với HTĐ lớn do có nhiều máy phát tương tác phức tạp trên hệ thống thì tiêu chuẩn diện tích sẽ không còn phù hợp.

2.3.4 Phương pháp trực tiếp

Phương pháp trực tiếp bỏ qua khâu giải hệ phương trình vi phân được trình bày trong tài liệu [10]. Đối với một HTĐ quan hệ giữa vector dòng điện và vector điện áp của máy phát được mô tả phương trình sau:

$$I_G = Y_R E_G \quad (2.11)$$

Trong đó: Y_R là ma trận tổng dẫn nút rút gọn; E_G, I_G lần lượt là vector điện áp và dòng điện của các máy phát.

Với một hệ thống có n máy phát, công suất tác dụng cung cấp bởi máy phát thứ i được cho bởi:

$$\dot{E} = E_i \angle \delta_i \quad (2.12)$$

Và tổng dẫn:

$$y_{ij} = G_{ij} + j.B_{ij} \quad (2.13)$$

Với hệ thống n máy phát, công suất tác dụng được cung cấp bởi máy phát thứ i được cho bởi:

$$\begin{aligned} P_i &= \text{Re}(\dot{E}_i \dot{I}_i^*) = \text{Re} \dot{E}_i \left(\sum_{j=1}^n \dot{y}_{ij}^* \cdot \dot{E}_j^* \right) \\ &= E^2 G_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i \cdot E_j \left[B_{ij} \sin(\delta_i - \delta_j) + G_{ij} \cos(\delta_i - \delta_j) \right] \end{aligned} \quad (2.14)$$

Vị trí tâm của quán tính góc (Centre of Inertia-COI) COI được xác định theo phương trình:

$$\delta_{COI} = \frac{1}{H_T} \sum_{i=1}^n H_i \delta_i \quad (2.15)$$

Trong đó: H_T là tổng các hằng số quán tính của tất cả các máy phát trong hệ thống, δ_i là góc lệch roto của máy phát thứ i . Chuyển động của COI được xác định bởi:

$$2H_T p(\Delta\omega_{COI}) = P_{COI} = \sum_{i=1}^n (P_{mi} - P_{ei}) \quad (2.16)$$

Trong đó:

$$\begin{aligned}
 P'_{mi} &= P_m - E_i^2 G_{ii} \\
 P_{ei} &= \sum_{\substack{j=i+1 \\ j \neq i}}^n \left[C_{ij} \sin(\delta_i - \delta_j) + D_{ij} (\delta_i - \delta_j) \right] \\
 C_{ij} &= E_i E_j B_{ij} \\
 D_{ij} &= E_i E_j G_{ij}
 \end{aligned} \tag{2.17}$$

Với: P_{mi} là công suất cơ của máy phát thứ i ; ω_0 là tốc độ đồng bộ tính bằng rad/s; $\Delta\omega_{COI}$ là sự thay đổi tốc độ của COI quanh tốc độ đồng bộ.

Sự chuyển động của máy phát thứ i theo COI có thể được biểu diễn như sau:

$$\theta_i = \delta_i - \delta_{COI} \text{ (rad)} \tag{2.18}$$

$$\text{Và } \varpi_i = \frac{\theta_i}{\varpi_0} = \left(\frac{\delta_i}{\varpi_0} - \Delta\varpi_{COI} \right) \text{ (pu)} \tag{2.19}$$

Hàm năng lượng V biểu diễn tổng năng lượng quá độ của hệ thống sau khi xảy ra sự cố được xác định:

$$\begin{aligned}
 V &= \frac{1}{2} \sum_{i=1}^n J_i \omega_i^2 - \sum_{i=1}^n P'_{mi} (\theta_i - \theta_i^s) \\
 &\quad - \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left[C_{ij} (\cos \theta_{ij} - \cos \theta_{ij}^s) - \int_{\theta_i^s + \theta_j^s}^{\theta_i + \theta_j} D_{ij} \cos \theta_{ij} d(\theta_i + \theta_j) \right]
 \end{aligned} \tag{2.20}$$

Trong đó: θ_i^s là góc ở nút thứ i tại điểm cân bằng ổn định và tại thời điểm sự cố; $J_i = 2H_i\omega_0$ là mômen quán tính của máy phát thứ i (đơn vị tương đối).

Hàm năng lượng bao gồm 4 thành phần sau:

- a) $\frac{1}{2} \sum J_i \omega_i^2$: biểu diễn sự biến đổi động năng rô-to của tất cả các máy phát trong hệ quy chiếu của COI.
- b) $\sum P'_{mi} (\theta_i - \theta_i^s)$: biểu diễn sự biến đổi thế năng rô-to của tất cả các máy phát trong cơ sở COI,
- c) $\sum \sum C_{ij} (\cos \theta_{ij} - \cos \theta_{ij}^s)$: biểu diễn sự thay đổi năng lượng từ tích trữ trong tất cả các nhánh của hệ thống.

d) $\sum \sum \int D_{ij} \cos \theta_{ij} d(\theta_i + \theta_j)$: biểu diễn sự thay đổi của năng lượng tiêu tán trong tất cả các nhánh của hệ thống.

Thành phần trong mục (a) được gọi là động năng (V_{ke}) và là hàm duy nhất theo tốc độ của các máy phát. Tổng 3 thành phần ở các mục (b), (c), (d) được gọi là thế năng (V_{pe}) và là hàm duy nhất của góc các máy phát.

Các bước tính toán đánh giá ổn định quá độ:

1. Tính toán năng lượng tới hạn (V_{cr}).
2. Tính toán năng lượng tổng của hệ thống tại thời điểm giải trừ sự cố (V_{cl}).
3. Tính chỉ số ổn định: ($V_{cr} - V_{cl}$).
4. Đánh giá: hệ thống sẽ ổn định nếu chỉ số ổn định dương.

Giới hạn của phương pháp: Phương pháp trực tiếp hay phương pháp năng lượng tính toán ổn định hệ thống điện có ưu điểm là bỏ qua khâu giải hệ phương trình vi phân nhưng cần nhiều tính toán để xác định chỉ số ổn định quá độ.

2.3.5 Phương pháp mô phỏng theo miền thời gian

Phân tích ổn định động cho một hệ thống điện đặc biệt đối với hệ thống điện nhiều máy là một bài toán khó và rất phức tạp. Phương pháp phân tích ổn định cho hệ thống điện nhiều máy bằng mô phỏng không cần phải biến đổi tương đương thành hệ thống một máy nối với hệ thống vô cùng lớn. Phương pháp này được phân tích dựa trên góc lệch của các máy phát so với máy phát chuẩn để từ đó có thể xác định được hệ thống là ổn định hay mất ổn định. Khi hệ thống bị kích động bởi các nhiễu lớn thì tất cả các thông số và trạng thái của hệ thống cũng thay đổi theo. Các thông số đầu vào thực hiện cho quá trình phân tích chính là các kết quả của bài toán phân bố công suất và các kết quả này không thể được tính toán một cách trực tiếp mà phải thực hiện tính nhiều bước lặp.

Bằng phương pháp mô phỏng miền thời gian, quá trình mô phỏng được lặp đi lặp lại nhiều lần cho đến khi góc lệch của một trong các máy phát so với máy phát chuẩn là tăng vô hạn thì quá trình lặp sẽ dừng lại và khi đó xác định được thời gian cắt tới hạn. Thời gian cắt tới hạn CCT (Critical Clearing Time) là thời gian cắt sự

có tối đa cho phép để hệ thống điện vẫn giữ được ổn định, tương ứng với đó là góc công suất cắt tới hạn.

2.4 CÁC NGHIÊN CỨU KHOA HỌC LIÊN QUAN

Phương pháp miền thời gian tốn nhiều thời gian giải so với phương pháp trực tiếp [14], [15]. Phương pháp trực tiếp hay phương pháp năng lượng, phương pháp diện tích bỏ qua khâu giải hệ phương trình vi phân nên có ưu điểm hơn phương pháp mô phỏng miền thời gian, nhưng các phương pháp này trở nên phức tạp khi xem xét đến mô hình chi tiết của các phần tử, sơ đồ lớn nhiều máy [15]–[18]. Phương pháp mô phỏng miền thời gian cho kết quả chính xác trong phân tích ổn định quá độ hệ thống điện, nhưng trở ngại lớn nhất của phương pháp này là vấn đề thời gian giải [14]. Như vậy, các phương pháp truyền thống không đáp ứng yêu cầu về thời gian trong chế độ điều khiển khẩn cấp. Các quá độ dao động lớn do sự cố gây ra mất ổn định cần phát hiện nhanh giúp đưa ra quyết định điều khiển khẩn cấp nhằm tránh hiện tượng tan rã lưới điện. Sa thải phụ tải được xem là một trong những phương pháp được áp dụng trong các tình huống khẩn cấp giúp tránh mất ổn định. Các nghiên cứu sa thải phụ tải sử dụng rơ-le dưới tần số (UFLS) nhằm ngăn ngừa tần số giảm sâu sau khi sự cố xảy ra, và khôi phục tần số về giá trị cho phép [19], [20]. Do tính phức tạp của HTĐ cho nên trong chế độ điều khiển khẩn cấp việc tính toán gặp khó khăn về mặt thời gian, hoặc việc phải sa thải phụ tải một cách thụ động sau khi tần số giảm xuống dưới giá trị ngưỡng cho phép, với nhiều bước thực hiện [21], điều này gây chậm trễ trong việc ra quyết định dẫn đến mất ổn định HTĐ. Để tối ưu hóa lượng công suất sa thải phụ tải, tính toán nhanh trong điều khiển thì các phương pháp tính toán thông minh thể hiện sự ưu điểm vượt trội so với phương pháp truyền thống [3]. Trong [22], tác giả đề nghị áp dụng mạng nơ-ron nhân tạo (ANN) để đánh giá ổn định điện áp với tín hiệu đầu vào là giá trị điện áp và góc pha điện áp được đo lường từ thiết bị PMU (Phase Measurement Unit). PMU là thiết bị đo lường có khả năng giám sát thời gian thực hệ thống điện. Trong trường hợp yêu cầu tính toán nhanh thì công nghệ tính toán thông minh cụ thể là mạng nơ-ron nhân tạo cung cấp giải quyết vấn đề một cách hiệu quả, nhanh, và chính xác cho bài toán phi tuyến cao [6], [23]–[25]. Trong

những năm gần đây, nhiều nghiên cứu ứng dụng mạng nơ-ron nhân tạo để xây dựng hệ thống thông minh chẩn đoán ổn định hệ thống điện [26]–[30]. Để xây dựng hệ thống thông minh thì thông tin ban đầu về chế độ và sơ đồ được xử lý trước, việc chẩn đoán trực tiếp chế độ ổn định hệ thống điện được xác định nhờ vào khâu trích xuất tri thức hay học quan hệ vào ra thích hợp theo những thông số về chế độ bình thường và sự cố, mà không cần giải hệ phương trình vi phân mô tả hệ thống.

Trong nghiên cứu ổn định quá độ có hai vấn đề quan trọng cần quan tâm là đánh giá ổn định và chẩn đoán ổn định. Chỉ số quan trọng cần tìm trong đánh giá ổn định hệ thống điện theo truyền thống là CCT. Trong chẩn đoán ổn định hệ thống điện thì chỉ số CCT không cần quan tâm, câu hỏi mấu chốt cần trả lời ở đây là sau dao động quá độ kết quả chẩn đoán hệ thống điện ‘ổn định’ hay ‘không ổn định’ [25], [26], [31], [32].

Trong các bài báo [26], [27], [33]–[36], các tác giả đã đề nghị ứng dụng ANN đánh giá ổn định hệ thống điện với kết quả đầu ra là chỉ số CCT. Trong các bài báo [26], [32], [37]–[40], các tác giả đã đề nghị ứng dụng kỹ thuật nhận dạng để chẩn đoán ổn định hệ thống điện với kết quả đầu ra là ‘ổn định’ hay ‘không ổn định’. Biến đầu vào là biến trạng thái các chế độ hệ thống điện.

Đánh giá ổn định hệ thống điện theo chỉ số CCT thì từ sơ đồ hệ thống điện thực hiện mô phỏng off-line để tìm CCT, qua đó thu thập được tập cơ sở dữ liệu để huấn luyện mạng nơ-ron. CCT là giá trị quan trọng để căn cứ vào đó cài đặt, chỉnh định thiết bị bảo vệ. CCT cũng được sử dụng để xác định giá trị biên về thời gian ổn định hệ thống điện, chỉ số biên thời gian ổn định [27], [33]. Trong đánh giá quá độ hệ thống điện theo chỉ số CCT thì còn thông số quan trọng xem xét là thời gian cắt sự cố FCT (Fault Clearing Time). Nếu $CCT > FCT$ thì hệ thống ổn định, nếu $FCT > CCT$ thì hệ thống không ổn định [10], [27].

Chẩn đoán ổn định động hệ thống điện với kết quả đầu ra là ‘ổn định’ hay ‘không ổn định’ [26], [37]. Mô phỏng off-line để tạo cơ sở dữ liệu huấn luyện cho mô hình nhận dạng. Trong mô phỏng off-line, kết quả dạng sóng góc đồng bộ giữa

các máy phát lớn hơn 180° tức hệ thống điện không ổn định, nhỏ hơn 180° tức hệ thống điện ổn định.

Trong chẩn đoán ổn định hệ thống điện, biến được lựa chọn chứa thông tin ở hai chế độ: một là chế độ tiền sự cố (pre-fault) và hai là chế độ sự cố (fault-on).

Biến tiền sự cố là biến chứa thông tin chế độ hệ thống điện vận hành bình thường, hay là biến chứa thông tin hệ thống điện ở trạng thái xác lập. Các biến này gồm công suất các máy phát, công suất trên các đường dây, điện áp các bus, công suất các tải,... Biến tiền sự cố được lựa chọn cho trường hợp chẩn đoán mang tính phòng ngừa với các trường hợp sự cố giả định xảy ra. Kết quả đầu ra của đánh giá ổn định hệ thống điện với dạng biến này giúp cung cấp thông tin để ra quyết định điều khiển ngăn chặn (preventive control). Vấn đề này được áp dụng trong [26], [37] thuộc về bài toán chẩn đoán an ninh hệ thống điện.

Biến sự cố là biến chứa thông tin trạng thái sự cố của hệ thống điện như độ sụt điện áp tại các nút, độ thay đổi công suất trên các nhánh, độ thay đổi công suất các tải, độ thay đổi công suất các máy phát, độ thay đổi các góc phát,... Biến sự cố khác biến tiền sự cố ở đặc điểm là nó vừa mang thông tin của chế độ xác lập hay tiền sự cố, và vừa mang thông tin của chế độ sự cố. Trong chẩn đoán với biến sự cố thì kết quả đầu ra cho biết trạng thái ổn định sắp tới của hệ thống điện do sự cố gây ra là ‘ổn định’ hay ‘không ổn định’. Vấn đề này áp dụng trong [32], [38]–[41] là thuộc về bài toán chẩn đoán ổn định động hệ thống điện. Kết quả đầu ra của chẩn đoán ổn định hệ thống điện với dạng biến sự cố giúp cung cấp thông tin để ra quyết định điều khiển khẩn cấp (emergency control).

Các nghiên cứu chẩn đoán ổn định sớm, biến tiền sự cố, sự cố chưa hoặc không xảy ra, có ý nghĩa là kết quả chẩn đoán nhằm cảnh báo trước để giúp điều khiển ngăn chặn với tình huống xấu giả định có thể xảy ra, thực chất sự cố có thể không xảy ra. Hướng nghiên cứu chẩn đoán với biến tiền sự cố thường tập trung giám sát ở một vài đường dây mang tải lớn gần giới hạn ổn định của hệ thống điện, hay các đường dây liên kết giữa các vùng địa lý khác nhau, hay một hoặc nhóm máy phát với nhóm máy phát còn lại trong hệ thống điện. Kết quả chẩn đoán

này có thể giúp các nhà vận hành hệ thống điện điều chỉnh lịch trình phát, đưa ra chiến lược điều khiển vận hành của các nhà máy điện một cách an toàn, hẳn nhiên là phải chấp nhận tổn thất chi phí cho việc điều chỉnh này mặc dù sự cố có thể không xảy ra ở chế độ vận hành tại thời điểm đó [37], [42], [43]. Tuy nhiên, sự cố xảy ra ở nhiều vị trí khác nhau trên hệ thống điện ở các mức tải khác nhau vẫn gây ra mất ổn định động như ngắn mạch tại các bus, các bus ngay máy phát điện từ nhà máy điện và một số đường dây khác nhau. Trường hợp như vậy là vô cùng nguy hại đến ổn định hệ thống điện và là nguyên nhân dẫn đến tan rã lưới điện nếu không phát hiện kịp lúc. Việc phát hiện nhanh mất ổn định hệ thống điện do sự cố gây ra nhằm có chiến lược điều khiển khẩn cấp để đưa hệ thống điện trở về trạng thái ổn định, cứu hệ thống điện khỏi nguy cơ tan rã lưới. Đây là các trường hợp nghiên cứu cho bài toán chẩn đoán ổn định động được các tác giả công bố trong các bài báo [43]–[46].

Hệ thống nhận dạng thông minh (Intelligent System-IS) cần thông tin đặc trưng quan trọng làm dữ liệu đầu vào, việc giảm không gian dữ liệu đầu vào giúp cho hệ thống xử lý dữ liệu một cách nhanh chóng, chính xác, [34], [44]. Để giảm chi phí tính toán, chi phí đo lường cảm biến, cải thiện tính năng hệ thống nhận dạng thì số biến đầu vào càng ít càng tốt nhưng độ chính xác phải cao, đạt giá trị kỳ vọng.

Về vấn đề giảm không gian biến được tác giả quan tâm nghiên cứu với nhiều công trình công bố. Trong [37], [48], [49] áp dụng chuẩn Fisher, trong [37], [50] áp dụng chuẩn Divergence để chọn biến. Tuy nhiên, vấn đề giảm không gian mẫu cho bài toán chẩn đoán ổn định HTĐ vẫn còn là khoảng trống. Vấn đề nhận dạng ổn định quá độ HTĐ đối mặt với vấn đề khó khăn về kích thước dữ liệu lớn và phức tạp. Việc thu thập và lưu trữ tất cả mọi mẫu là không khả thi. Vấn đề đặt ra là chọn những mẫu đại diện cho không gian mẫu nhằm giảm gánh nặng thu thập dữ liệu, cũng như tiết kiệm tài nguyên bộ nhớ máy tính, giảm chi phí và nâng cao độ chính xác nhận dạng trong chẩn đoán cần quan tâm nghiên cứu giải quyết.

Trong những năm gần đây rất nhiều nghiên cứu tập trung vào hướng nghiên cứu ứng dụng công nghệ tính toán thông minh nhân tạo chẩn đoán ổn định hệ thống điện [40]–[43]. Trong [29] tác giả đã áp dụng mạng nơ-ron nhân tạo để đánh giá ổn định điện áp. Trong [51] mạng nơ-ron nhân tạo được đề xuất áp dụng để ngăn chặn tan rã lưới điện trong mạng micro-grid. Bài báo này tác giả cũng chỉ ra ANN một khi đã được huấn luyện xong thì cho đáp ứng nhanh, hiệu quả trong giải quyết vấn đề ổn định hệ thống điện. Tuy nhiên, bài báo này chưa đề cập đến vấn đề giảm không gian biến và không gian mẫu. Trong [52], bài báo đã đề nghị áp dụng Vec-tơ hỗ trợ máy (SVM) và cây quyết định (Decision Tree –DT) chẩn đoán ổn định HTĐ. Trong công bố này nhóm tác giả đã giới thiệu các nhóm biến và bộ phân lớp kết hợp SVM và DT vào đánh giá ổn định quá độ hệ thống điện. Trong [29], nhóm tác giả chỉ ra rằng chiến lược hệ thống thông minh (IS) là tiếp cận thành công trong đánh giá ổn định hệ thống điện thời gian thực. Trong công trình này, các tác giả đã áp dụng giải thuật Relief để chọn biến và giải thuật này cũng đã được các tác giả đã công bố trước đó [37]. Bài báo [37] cũng là công trình giới thiệu quy trình chọn biến một cách tổng quát, và giới thiệu tiếp cận vấn đề chọn biến một cách có hệ thống. Trong [3], các tác giả cũng phân tích chỉ ra khả năng tính toán nhanh của công nghệ tính toán thông minh trong điều kiện khẩn cấp, quá trình xảy ra nhanh và phức tạp trong hệ thống điện. Trong [53], [54] đã áp dụng DT trong đánh giá ổn định hệ thống điện.

Bài báo (XII), tác giả đã áp dụng fuzzy-AHP (fuzzy-analytic hierarchy process) để xây dựng nhóm chiến lược sa thải phụ tải có xét đến tầm quan trọng của phụ tải làm giảm thiệt hại về mặt kinh tế. Tuy nhiên, do tính phức tạp của HTĐ cho nên trong chế độ điều khiển khẩn cấp việc tính toán gặp khó khăn về mặt thời gian, hoặc việc phải sa thải phụ tải một cách thụ động sau khi tần số dưới ngưỡng cho phép, điều này gây chậm trễ trong việc ra quyết định dẫn đến mất ổn định HTĐ. Trong công trình (IX) tác giả đã đề nghị phối hợp mạng thần kinh nhân tạo và giải thuật AHP, trong (X) áp dụng fuzzy-AHP cho trường hợp điều khiển sa thải phụ tải khẩn cấp. Bài báo (VIII), tác giả đã đề xuất chiến lược sa thải phụ tải

trên cơ sở nhận dạng nhanh chế độ không ổn định HTĐ ứng dụng mạng thần kinh nhân tạo.

Tóm lại, các công trình công bố liên quan đến nhận dạng ổn định hệ thống điện trong những năm gần đây đã chỉ ra công nghệ tính toán thông minh nhân tạo, kỹ thuật máy học cho khả năng tính toán nhanh, chính xác, và khả năng khai phá tiềm năng thông tin hữu ích đối với dữ liệu phức tạp. Tuy nhiên, các công trình này chỉ tập vào vấn đề chọn biến mà chưa đề cập đến vấn đề giảm không gian mẫu.

Về vấn đề bộ phân lớp hay bộ nhận dạng, qua khảo cứu các công trình đã công bố thì mô hình chẩn đoán gồm có hai dạng mô hình: Thứ nhất, phổ biến, là mô hình đơn với bộ chẩn đoán đơn hay mô hình gồm một mạng nơ-ron [26], [31], [37], [48]. Thứ hai là mô hình song song với bộ chẩn đoán gồm các mạng nơ-ron con song song. Trong [27], tác giả đề xuất mô hình song song, mỗi mô hình con làm nhiệm vụ chẩn đoán với một dạng sự cố có thể gây mất ổn định gồm một trong các dạng sự cố như ba pha, một pha chạm đất, hai pha, hai pha chạm đất. Mô hình này tác giả đã đề xuất áp dụng trong chẩn đoán an ninh hệ thống điện với biến tiền sự cố, và chỉ tập trung giám sát một số ít đường dây, bus như đã trình bày ở phần trên. Tuy nhiên, mô hình này áp dụng chẩn đoán ở nhiều vị trí bus hoặc nhiều đường dây khi bị sự cố gây ra mất ổn định hệ thống điện làm cho số lượng mô hình con rất lớn, phức tạp. Vì vậy, nghiên cứu đề xuất xây dựng mô hình nhận dạng thông minh dựa trên cơ sở mạng nơ-ron cải tiến chẩn đoán nhanh ổn định động hệ thống điện nhằm nâng cao độ chính xác là điều cần thiết.

2.5 TÓM TẮT CHƯƠNG 2

Trong chương này tác giả đã trình bày tổng quan về lý thuyết cơ bản ổn định hệ thống điện và các phương pháp phân tích ổn định hệ thống điện. Tác giả đã phân tích các công trình nghiên cứu liên quan, từ đó chỉ ra những khoảng trống cần đào sâu nghiên cứu và tìm giải pháp giải quyết, và rút ra các vấn đề sau:

- Để đánh giá ổn định quá độ của hệ thống điện trong các dao động lớn do sự cố gây ra, các phương pháp truyền thống tỏ ra kém hiệu quả và không thuận lợi,

đặc biệt trong những điều kiện bất định, thiếu thông tin và yêu cầu khắc khe về thời gian giải.

- Trong trường hợp yêu cầu tính toán nhanh thì công nghệ tính toán thông minh cụ thể là mạng nơ-ron nhân tạo cung cấp giải quyết vấn đề một cách hiệu quả, nhanh, và chính xác cho bài toán phi tuyến cao.
- Hệ thống nhận dạng thông minh cần thông tin đặc trưng làm dữ liệu đầu vào, việc giảm không gian dữ liệu đầu vào giúp cho hệ thống xử lý dữ liệu một cách nhanh chóng, chính xác.
- Lựa chọn biến là vấn đề phức tạp, đặc biệt là không gian biến lớn với hệ thống điện nhiều máy. Thu thập mẫu là công việc nặng nhọc, tốn nhiều chi phí, và thu thập mọi mẫu là không khả thi. Không gian biến và không gian mẫu cần rất tinh gọn, giúp mô hình nhận dạng làm việc nhanh, giảm gánh nặng thu thập dữ liệu, tiết kiệm tài nguyên bộ nhớ máy tính, giảm chi phí.
- Dựa trên những phân tích trên, tác giả đã chỉ ra ba vấn đề cần giải quyết khi xây dựng hệ thống thông minh chẩn đoán ổn định động HTĐ nhằm làm cho hệ thống làm việc nhanh, tiết kiệm bộ nhớ và nâng cao độ chính xác nhận dạng. Các vấn đề cần giải quyết cụ thể là:
 - ✓ Giảm không gian biến.
 - ✓ Giảm không gian mẫu.
 - ✓ Cải tiến mô hình bộ phân lớp ứng dụng mạng nơ-ron để nâng cao độ chính xác nhận dạng.

Từ đó, tác giả đã nghiên cứu và đề xuất các giải pháp như sau:

1. Đối với vấn đề lựa chọn tập biến, tác giả đã nghiên cứu và trình bày các kỹ thuật chọn biến, giải thuật tìm kiếm tập biến và vấn đề này được trình bày cụ thể ở chương 4, quy trình chọn biến trình bày chi tiết ở chương 5. Phương pháp đề nghị áp dụng của tác giả cho thấy hiệu quả trong giảm không gian biến và được công bố trong các bài báo (I-VI, VIII).
2. Đối với vấn đề giảm không gian mẫu, tác giả đã nghiên cứu và trình bày chi tiết quy trình xây dựng tập mẫu ở chương 5. Trong đó, tác giả đã giới thiệu áp dụng giải thuật K-means và giải thuật K-means lai (Hybrid K-

means). So sánh và đánh giá hiệu quả của giải thuật giảm không gian mẫu đề nghị áp dụng được trình bày ở chương 6. Vấn đề giảm không gian dữ liệu được công bố trong bài báo (I)

3. Đối với bộ phân lớp ứng dụng công nghệ tính toán thông minh, tác giả đã nghiên cứu và trình bày lý thuyết cơ sở của các bộ phân lớp ở chương 3. Ý tưởng đề xuất mô hình bộ phân lớp song song cải tiến, và quy trình xây dựng mô hình được trình bày ở chương 5. Có hai ý tưởng đề xuất được áp dụng là dựa vào năng lượng mẫu và dựa vào giải thuật phân cụm dữ liệu để tách mẫu. Xây dựng mô hình Bộ phân lớp song song dựa vào năng lượng mẫu được trình bày ở chương 5 và công bố ở bài báo (II). Kết quả kiểm tra hiệu quả của mô hình mạng nơ-ron cải tiến dựa trên ý tưởng phân cụm dữ liệu được trình bày chi tiết ở chương 6.

Chương 3. CÔNG NGHỆ TÍNH TOÁN THÔNG MINH

3.1 MẠNG NƠ-RON NHÂN TẠO

3.1.1 Giới thiệu về mạng nơ-ron nhân tạo

Mạng nơ-ron nhân tạo (Artificial Nơ-ron Networks – ANN) được xây dựng dựa trên cấu trúc của bộ não con người, giúp đưa ra một phương pháp mới trong lĩnh vực tiếp cận hệ thống thông tin [55]–[59]. Mạng nơ-ron nhân tạo có thể thực hiện các bài toán nhận dạng, phân loại và điều khiển cho các đối tượng tuyến tính và phi tuyến đạt hiệu quả cao hơn so với các phương pháp tính toán truyền thống.

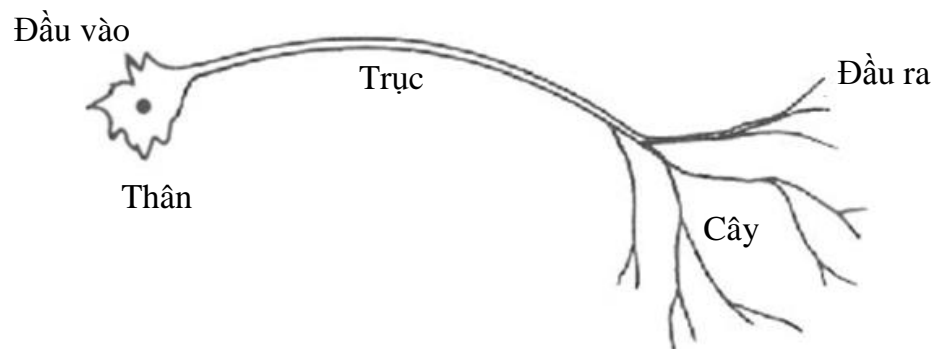
Mạng nơ-ron nhân tạo gồm nhiều nơ-ron nhân tạo liên kết với nhau thành mạng. Nó có hành vi tương tự như bộ não con người với các khả năng học (Learning), gọi lại và tổng hợp thông tin từ sự luyện tập của các mẫu dữ liệu.

Mạng nơ-ron đã có một lịch sử lâu dài. Từ năm 1943, McCulloch và Pitts đã đề xuất một số liên kết cơ bản của mạng nơ-ron. Năm 1949, Hebb đã đưa ra các luật thích nghi trong mạng nơ-ron. Năm 1958, Rosenblatt đưa ra cấu trúc Perception. Năm 1969, Minsky và Papert phân tích sự đúng đắn của Perception, chứng minh các tính chất và chỉ rõ các giới hạn của một số mô hình. Năm 1976, Grossberg dựa vào tính chất sinh học đã đưa ra một số cấu trúc của hệ động học phi tuyến với các tính chất mới. Năm 1982, Hopfield đã đưa ra mạng hồi quy một lớp Hopfield. Năm 1986, Rumelhart đưa ra mô hình xử lý song song và một số kết quả của thuật toán. Năm 1986, thuật toán học lan truyền ngược (Backpropagation) được Rumelhart, Hinton và Williams đề xuất thực hiện luyện mạng nơ-ron truyền thẳng nhiều lớp. Những năm gần đây, có nhiều công trình nghiên cứu đề xuất các cấu trúc, luật học cho nhiều loại mạng nơ-ron truyền thẳng và hồi quy mới có nhiều ưu điểm. Mạng nơ-ron hiện đang được ứng dụng trong nhiều lĩnh vực kinh tế, kỹ thuật.

3.1.2 Mô hình nơ-ron sinh học

Bộ não con người có khoảng 10^{11} nơ-ron sinh học ở nhiều dạng khác nhau [57], [58]. Mô hình của một dạng nơ-ron sinh học được mô tả trên hình 3.1. Cấu

trúc chung của một nơ-ron sinh học gồm có ba phần chính là: thân, cây và trục. Cây gồm các dây thần kinh liên kết với thân. Trục có cấu trúc đơn, dài liên kết với thân. Phần cuối của trục có dạng phân nhánh. Trong mỗi nhánh có một cơ cấu nhỏ là khớp thần kinh, từ đây nơ-ron sinh học này liên kết bằng tín hiệu tới các nơ-ron khác. Sự thu nhận thông tin của nơ-ron sinh học được thực hiện từ cây hoặc từ thân của nó. Tín hiệu thu, nhận ở dạng các xung điện.



Hình 3.1 Mô hình nơ-ron sinh học

Một tế bào thần kinh có ba phần, Hình 3.1:

-Phần đầu có nhiều khớp thần kinh (Dendrite) là nơi tiếp xúc với các với các điểm kết nối (Axon Terminal) của các tế bào thần kinh khác.

-Nhân của tế bào thần kinh (Nucleus) là nơi tiếp nhận các tín hiệu điện truyền từ Dendrite. Sau khi tổng hợp và xử lý các tín hiệu nhận được nó truyền tín hiệu kết quả qua trục nơ-ron (Axon) đến các điểm kết nối (Axon Terminal) ở đuôi.

-Phần đuôi có nhiều điểm kết nối (Axon Terminal) để kết nối với các tế bào thần kinh khác.

3.1.3 Mô hình nơ-ron nhân tạo

3.1.3.1 Cấu trúc mô hình một nơ-ron nhân tạo

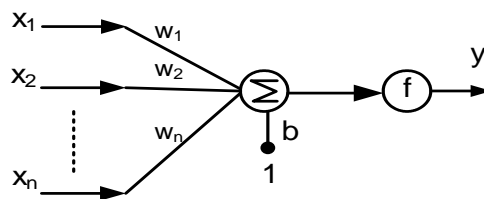
Mạng nơ-ron nhân tạo được thiết kế để mô phỏng tính chất của mạng nơ-ron sinh học, tức là mô phỏng hoạt động dựa theo cách thức của bộ não con người [55]–[58]. Các đặc tính của mạng nơ-ron là:

- Tri thức của mạng được nắm bắt bởi các nơ-ron thông qua quá trình học.
- Trọng số kết nối các nơ-ron đóng vai trò lưu giữ thông tin.

- Mạng nơ-ron tính toán rất nhanh.

Nơ-ron nhân tạo nhận một số các ngõ vào (từ dữ liệu gốc, hay từ ngõ ra của các nơ-ron khác trong mạng). Mỗi kết nối đến ngõ vào có một trọng số (weight) w đặc trưng cho mức độ liên kết giữa các nơ-ron. Trọng số liên kết có giá trị dương tương ứng với khớp thần kinh bị kích thích, ngược lại có giá trị âm tương ứng với khớp thần kinh bị kiềm chế. Mỗi nơ-ron có một giá trị ngưỡng (threshold), ngưỡng này được đưa qua thành phần dịch chuyển (bias) b .

Tín hiệu được truyền qua hàm kích hoạt hay hàm chuyển đổi f tạo giá trị ngõ ra nơ-ron. Hình 3.2 trình bày mô hình nơ-ron với n phần tử ngõ vào, mỗi ngõ vào được kết nối với trọng số w . Tổng ngõ vào được trọng hóa với ngưỡng đưa tới ngõ ra qua hàm chuyển đổi f như biểu thức (3.1).



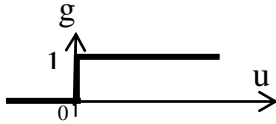
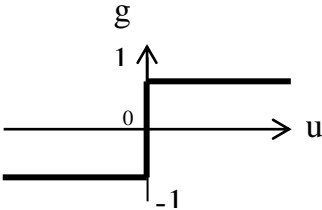
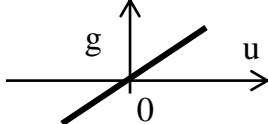
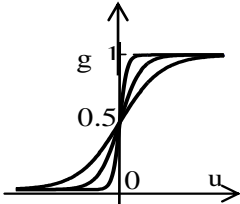
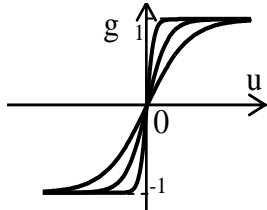
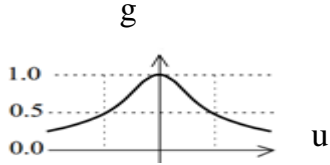
Hình 3.2 Mô hình một nơ-ron nhân tạo

Trong đó: Biến ngõ vào $x = [x_1, x_2, \dots, x_n]$ là vector ngõ vào có n phần tử, trọng số kết nối $W = [w_1, w_2, \dots, w_n]$. Ngõ ra mạng được tính theo công thức:

$$y = f\left(b + \sum_{i=1}^n x_i w_i\right) \quad (3.1)$$

Hàm chuyển đổi (Transfer Function) hay hàm kích hoạt (Activation Function) có nhiệm vụ biến đổi tổng trọng (hoặc net) thành tín hiệu ngõ ra y . Các nơ-ron có thể sử dụng các hàm chuyển đổi khác nhau để tạo ra tín hiệu ngõ ra. Các hàm chuyển đổi phổ biến như hàm hard limit, logsig, tansig, purelin, hàm bán kính, Bảng 3.1.

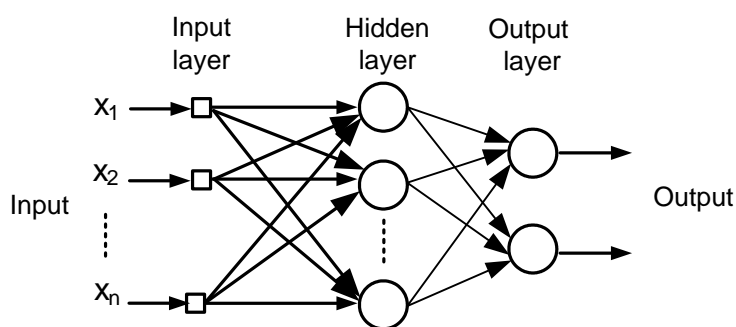
Bảng 3.1 Một số hàm phi tuyến thường dùng trong các mô hình nơ-ron

Tên hàm	Công thức	Đặc tính
Hàm bước nhảy đơn vị (hard limit)	$g(u) = 1$ nếu $u \geq 0$ $g(u) = 0$ nếu $u < 0$	
Hàm bước nhảy lưỡng cực	$g(u) = 1$ nếu $u \geq 0$ $g(u) = -1$ nếu $u < 0$	
Hàm tuyến tính (purelin)	$g(u) = a.u$	
Hàm sigmoid đơn cực (log-sigmoid (logsig))	$g(u) = \frac{1}{1 + e^{-\lambda u}}$	
Hàm sigmoid lưỡng cực (tan-sigmoid (tansig))	$g(u) = \frac{2}{1 + e^{-\lambda u}} - 1$	
Hàm bán kính	$g(u) = e^{-u}$	

3.1.3.2 Cấu trúc mô hình mạng nơ-ron nhân tạo nhiều lớp

Cấu trúc của mạng nơ-ron nhân tạo mô phỏng ý tưởng của mạng nơ-ron sinh học. Tuy nhiên, mạng nơ-ron nhân tạo có cấu trúc đơn giản hơn nhiều, về cả số

lượng nơ-ron và cả cấu trúc mạng. Mạng nơ-ron nhân tạo các nơ-ron được kết nối sao cho có thể dễ dàng biểu diễn bởi một mô hình toán học nào đó. Mạng nơ-ron nhiều lớp gồm một lớp đầu vào và một lớp đầu ra, các lớp nằm giữa lớp đầu vào và lớp đầu ra gọi là các lớp ẩn (Hidden layers). Mạng nơ-ron truyền thẳng là mạng hai hay nhiều lớp mà tín hiệu truyền theo một hướng từ đầu vào đến đầu ra, các nơ-ron được phân thành nhiều lớp, các nơ-ron chỉ được kết nối với các nơ-ron ở lớp liền trước hoặc liền sau lớp nó.



Hình 3.3 Cấu trúc mô hình mạng nơ-ron nhân tạo nhiều lớp truyền thẳng

3.1.4 Các qui tắc học

Thông thường mạng nơ-ron được điều chỉnh hoặc được huấn luyện để hướng các đầu vào riêng biệt đến đích ở đầu ra. Ở đây, hàm trọng của mạng được điều chỉnh trên cơ sở so sánh đầu ra với đích mong muốn (target) đến khi đầu ra mạng phù hợp với đích. Những cặp vào/đích (input/target) được dùng để giám sát cho huấn luyện mạng. Ở đó mỗi giá trị vào được gửi đến mạng và giá trị ra tương ứng được thực hiện bằng mạng là sự xem xét và so sánh với giá trị mong muốn. Thông thường tồn tại một sai số vì giá trị mong muốn không hoàn toàn khớp với giá trị thực. Sau mỗi lần thực hiện có tổng bình phương của tất cả các sai số, sai số này được sử dụng để xác định các hàm trọng mới. Mỗi lần thực hiện, hàm trọng của mạng được sửa đổi với đặc tính tốt hơn tương ứng với đặc tính mong muốn. Từng cặp giá trị vào/ra phải được kiểm tra và trọng được điều chỉnh. Sự thay đổi các hàm trọng của mạng được dừng lại nếu tổng các bình phương sai số nhỏ hơn một giá trị đặt trước hoặc đã lặp đủ một số lần thực thi xác định. Có hai kiểu học là học thông số và học cấu trúc.

Học thông số (parameter learning) là tìm ra biểu thức cập nhật các thông số về trọng số, cập nhật kết nối giữa các nơ-ron.

Học cấu trúc (structure learning) là sự biến đổi cấu trúc của mạng nơ-ron gồm số lượng nút (node) và các mẫu liên kết.

Nhiệm vụ của việc học thông số là bằng cách nào đó tìm được ma trận chính xác mong muốn từ ma trận giả thiết ban đầu với cấu trúc của mạng nơ-ron có sẵn. Vì vậy, mạng nơ-ron sử dụng các trọng số điều chỉnh với nhiều phương pháp học khác nhau có thể tính toán gần đúng ma trận trọng số W cần tìm đặc trưng cho mạng. Giả sử có n phần tử xử lý trong một mạng nơ-ron nhân tạo và mỗi phần tử xử lý có m hệ số trọng lượng thích nghi (adaptive weights), thì ma trận hệ số trọng lượng (ma trận kết nối) được định nghĩa bởi:

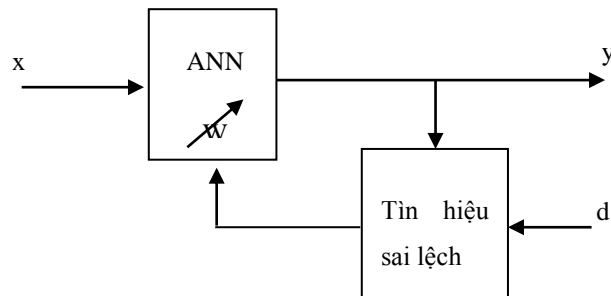
$$w = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n^T \end{bmatrix} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1m} \\ w_{21} & w_{22} & \dots & w_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nm} \end{bmatrix} \quad (3.1)$$

Trong đó w_{ij} là hệ số trọng lượng liên kết từ phần tử xử lý thứ i (nút nguồn) đến phần tử xử lý thứ j (nút đích), $i=[1,n]$ và $j=[1,m]$. Có ba qui tắc học [55]–[57]:

- Học giám sát (supervised learning).
- Học củng cố (reinforcement learning).
- Học không giám sát (unsupervised learning).

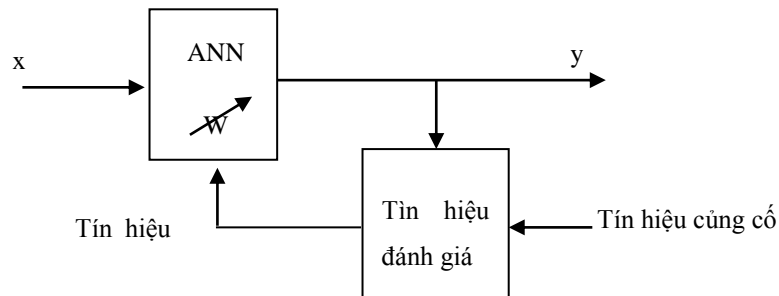
Học có giám sát: là quá trình học ở mỗi thời điểm thứ i khi tín hiệu x_i được đưa vào mạng nơ-ron, tương ứng sẽ có đáp ứng mong muốn d_i của đầu ra cho trước ở thời điểm đó. Trong quá trình học có giám sát, mạng nơ-ron được cung cấp liên tục các cặp số liệu mong muốn vào/ra ở từng thời điểm (x_1, d_1) , (x_1, d_2) , $(x_k, d_k), \dots$ khi cho tín hiệu vào thực là x_k sẽ tương ứng có tín hiệu đầu ra là y_k sẽ lệch với d_k . Sai lệch đó sẽ được truyền ngược tới đầu vào để điều chỉnh thông số mạng nơ-ron là ma trận trọng số W . Quá trình tiếp tục lặp lại sao cho sai lệch giữa tín hiệu ra mong muốn và tín hiệu ra thực tế nằm trong phạm vi cho phép, kết quả nhận được ma trận trọng số với các phần tử W_{ij} đã được điều chỉnh phù hợp với

đặc điểm của đối tượng hay hàm số mạng nơ-ron cần học. Mô hình học có giám sát được minh họa như hình 3.4.



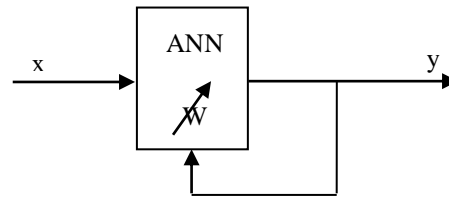
Hình 3.4 Mô hình học có giám sát

Học củng cố: là phương pháp học trong đó tín hiệu d được đưa từ bên ngoài nhưng không được đầy đủ mà có thể chỉ đưa đại diện 1 bit để có tính chất kiểm tra ở một mẫu vào/ra của quá trình học giám sát bị cho rằng không tin cậy. Bit tín hiệu đó được gọi là tín hiệu củng cố để khẳng định quá trình học đúng hay sai. Mô hình học củng cố được minh họa như hình 3.5.



Hình 3.5 Mô hình học củng cố

Học không có giám sát: Kiểu học này hoàn toàn không có tín hiệu từ bên ngoài, giá trị mục tiêu điều khiển không được cung cấp cũng không được tăng cường. Mạng phải tự khám phá các mẫu, các nét đặc trưng, tính tương thích, phân loại dữ liệu vào và mã hoá dữ liệu ra. Mô hình học không có giám sát được minh họa như hình 3.6.



Hình 3.6 Mô hình học không có giám sát

3.1.5 Ánh xạ và lan truyền ngược

Các mẫu (x, y) với $x = (x_1, x_2, \dots, x_n) \in X$, $y = (y_1, y_2, \dots, y_m) \in Y$ được cho trước. Như vậy, học có thể được xem như việc xấp xỉ một ánh xạ $X \rightarrow Y$, trong đó X là tập các vấn đề và Y là tập các lời giải tương ứng cho vấn đề đó.

Mạng lan truyền là một hàm phi tuyến có thể xấp xỉ gần đúng nhất một hàm đích được cho qua một số mẫu trong tập mẫu. Một mạng lan truyền tổng quát là một mạng có n ($n > 2$) lớp: lớp thứ nhất gọi là lớp nhập, lớp thứ n gọi là lớp xuất, và $(n-2)$ lớp ẩn. Số nút của lớp nhập và của lớp xuất do bài toán qui định, còn số nút của lớp ẩn do người thiết kế mạng qui định. Trong mạng lan truyền, mỗi nút của lớp thứ i ($0 < i < n$) liên kết với mọi nút ở lớp thứ $(i+1)$, và các nút trong cùng một lớp không liên kết với nhau. Ngoài ra, còn có một số cung liên kết trực tiếp từ các nút lớp nhập đến các nút trong lớp xuất, mạng này là mạng lan truyền có nối trực tiếp. Mỗi cung trong mạng có gắn một trọng số $w \in \mathbb{R}$.

Phương pháp lan truyền ngược là một quá trình lặp đi lặp lại nhiều lần hai tiến trình chính: ánh xạ và lan truyền ngược sai số. Hai tiến trình này được áp dụng trên một tập mẫu xác định, được gọi là luyện mạng hay là học. Các bước như sau:

- Bước 1: Xây dựng cấu trúc thích hợp cho mạng nơ-ron, chẳng hạn có $(n+1)$ nơ-ron vào (n nơ-ron cho biến vào và 1 nơ-ron cho ngưỡng), m nơ-ron đầu ra và khởi tạo các trọng số liên kết của mạng.
- Bước 2: Đưa một vector x trong tập mẫu huấn luyện X vào mạng
- Bước 3: Tính vector đầu ra 0 của mạng
- Bước 4: So sánh vector đầu ra mong muốn y (là kết quả được cho trong tập huấn luyện) với vector đầu ra 0 do mạng tạo ra; nếu có thể thì đánh

giá lỗi.

- Bước 5: Hiệu chỉnh các trọng số liên kết theo một cách nào đó sao cho ở lần tiếp theo khi đưa vector x vào mạng, vector đầu ra 0 sẽ giống với y hơn.
- Bước 6: Lặp lại các bước từ 2 đến 5 cho tới khi mạng đạt tới trạng thái hội tụ. Việc đánh giá sai số có thể thực hiện theo nhiều cách, cách dùng nhiều nhất là sai số trung bình bình phương (MSE - Mean-Square Error):

$$E = \sum (0 - y)^2 / 2 \quad (3.2)$$

3.1.6 Mặt lỗi

Học là quá trình tìm các trọng số của mạng sao cho ánh xạ nơ-ron khớp nhất với bộ dữ liệu chứa các mẫu của hàm đích và đó chính là tập mẫu. Sai số trung bình bình phương thường được sử dụng để đo lường sự trùng khớp giữa ánh xạ cần xây dựng với hàm đích cho trước qua tập mẫu. Về mặt hình học có thể xem E như là một mặt lỗi. Mặt lỗi là một siêu phẳng trong đó mỗi điểm của nó tương ứng với một điểm trong không gian trọng. Chiều cao trên không gian trọng của mỗi điểm trong mặt lỗi biểu diễn sai số của mô hình ứng với các trọng tương ứng đó. Điểm thấp nhất trong mặt lỗi cho mô hình có sai số nhỏ nhất.

3.1.7 Cực tiểu cục bộ

Nếu mặt lỗi nhấp nhô, nghĩa là có nhiều điểm trũng với độ cao thật khác nhau, các điểm trũng này chưa chắc là điểm trũng thấp nhất. Về mặt hình thức, cực tiểu cục bộ chính là điểm trên mặt lỗi mà tại đó cập nhật trọng số theo bất kỳ hướng nào cũng sẽ làm sai số tăng lên. Tìm cách tránh rơi vào cực tiểu cục bộ là một bài toán khó khăn. Hornik (1990) đã chứng minh được rằng một mạng với số nút ẩn thích hợp có thể xấp xỉ một hàm đích bất kỳ với sai số bất kỳ.

Dù ta không thể tính toán được dạng của toàn mặt lỗi, nhưng có thể tính toán được chiều cao và độ dốc mặt lỗi tại bất kỳ điểm nào trong không gian trọng. Trong toán học độ dốc mặt lỗi được xác định qua đạo hàm riêng theo từng trọng số trong mạng nơ-ron. Khi đã biết được độ dốc của mặt lỗi, thông tin này có thể được sử

dụng theo nhiều cách khác nhau để cập nhật bộ trọng số. Các phương pháp này được gọi là qui tắc học.

3.1.8 Phân loại mạng nơ-ron

Tùy theo cách kết nối giữa các nơ-ron xác định nên cấu trúc mạng. Trọng lượng các kết nối được điều chỉnh hay được huấn luyện để đạt được các yêu cầu của mạng bởi thuật toán huấn luyện. Mạng nơ-ron có thể được phân loại theo cấu trúc và thuật toán huấn luyện.

Theo cấu trúc, mạng nơ-ron phân loại thành mạng truyền thẳng và mạng phản hồi. Mạng truyền thẳng nhiều lớp gồm có lớp vào, các lớp ẩn và lớp ra. Lớp vào sẽ nhận tín hiệu trực tiếp từ bên ngoài. Lớp ra gồm các nơ-ron có tín hiệu ra, đưa ra bên ngoài mạng. Mạng phản hồi thì đầu ra của một số nơ-ron hồi tiếp ngược lại chính nó hay các nơ-ron trong lớp kế tiếp. Vì vậy, các tín hiệu có thể đi theo hai hướng truyền tới (forward) hay truyền lùi (backward).

Theo phương pháp huấn luyện có thể chia mạng theo các nhóm có phương pháp học tập có giám sát, phương pháp học tập không giám sát và phương pháp học tập củng cố hay học tăng cường.

3.1.9 Mạng Perceptron nhiều lớp

Mạng Perceptron nhiều lớp (Multilayer Perceptron – MLP) với thuật toán truyền ngược được sử dụng giải quyết nhiều vấn đề trong nhiều lĩnh vực khác nhau. Mạng gồm có một lớp ngõ vào, một hoặc nhiều lớp ẩn và một lớp ngõ ra. Số lượng nơ-ron trong lớp ngõ vào bằng với số lượng biến đầu vào. Có thể nói rằng thuật toán truyền ngược là thuật toán học quan trọng nhất trong lĩnh vực mạng nơ-ron. Thuật toán được phát triển một cách độc lập bởi Werbos vào năm 1974, Parker vào năm 1985, Rumelhart, Hinton và Williams vào năm 1986, thuật toán truyền ngược đã được nhiều nhà nghiên cứu cải thiện nhằm cải tiến tốc độ hội tụ cũng như tăng cường tính ổn định của thuật toán. Về cơ bản, thuật toán truyền ngược là một thuật toán tìm theo chiều gradient âm và có giám sát. Thuật toán truyền ngược bao gồm hai quá trình: quá trình lan truyền tiến và quá trình lan truyền ngược sai số. Trong quá trình lan truyền tiến, tín hiệu nhập được đưa vào

mạng và tín hiệu truyền này xuyên qua mạng từ lớp này đến lớp khác. Cuối cùng một tập các tín hiệu xuất được tạo ở ngõ ra. Trong suốt quá trình lan truyền tiến các trọng số của mạng không thay đổi. Trong quá trình truyền ngược, một tín hiệu sai số hay lỗi được truyền ngược xuyên qua mạng. Trong quá trình này các trọng số được hiệu chỉnh sao cho đáp ứng thực tế của mạng ngày càng đến gần với đáp ứng mong muốn.

Quy trình huấn luyện phổ biến nhất là học có giám sát với giải thuật lan truyền ngược. Giải thuật lan truyền ngược thực thi phương pháp giảm Gradient nhằm cực tiểu sai số trung bình bình phương giữa ngõ ra và ngõ ra mong muốn. Có nhiều giải thuật huấn luyện, trong đó thuật toán Levenberg-Marquardt được khuyến cáo áp dụng nhờ vào tốc độ hội tụ nhanh và cho độ chính xác nhận dạng cao [59]. Thông số mạng ban đầu là ngẫu nhiên, quá trình huấn luyện dừng khi đạt một trong các điều kiện như số vòng lặp đạt giá trị tối đa cài đặt, thời gian huấn luyện vượt giá trị tối đa, sai số đạt một mức cho phép,...

3.1.10 Mạng nơ-ron Generalized Regression Nơ-ron Network (GRNN)

3.1.10.1 Hàm cơ sở bán kính RBF

Hàm cơ sở bán kính được giới thiệu bởi Powell để giải quyết bài toán nội suy hàm nhiều biến năm 1987. Trong lĩnh vực mạng nơ-ron, mạng nơ-ron RBF được đề xuất bởi Lowe năm 1988 cho bài toán nội suy (xấp xỉ) hàm nhiều biến. Bài toán nội suy hàm nhiều biến được phát biểu như sau [57]: Cho một tập gồm N điểm dữ liệu trong không gian thực kích thước n $\{x_i \in R^n \mid i=1,2,\dots,N\}$ và tập đáp ứng N số thực $\{d_i \in R^1 \mid i=1,2,\dots,N\}$. Tìm hàm $F: R^N \rightarrow R^1$ thỏa mãn điều kiện nội suy:

$$F(x_i) = d_i \quad i=[1,2,\dots, N] \quad (3.3)$$

3.1.10.2 Kỹ thuật hàm cơ sở bán kính RBF

Hàm F có dạng:

$$F(x) = \sum_{i=1}^N w_i \varphi(\|x-x_i\|) \quad (3.4)$$

Trong đó $\{\varphi(\|x-x_i\|) \mid i=1,2,\dots,N\}$ là tập N hàm bán kính RBF, $\|\cdot\|$ là ký hiệu hàm chuẩn Euclidean. $x_i, i=1,2,\dots, N$ là các tâm của các hàm bán kính RBF.

Kết hợp (3.3) và (3.4) được một tập các phương trình quan hệ tuyến tính:

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \dots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (3.5)$$

Trong đó: $\varphi_{ij} = \|x_i - x_j\|$, $i, j = 1, 2, \dots, N$ (3.6)

Ký hiệu vector đáp ứng ngõ ra và vector trọng số, N là số mẫu huấn luyện:

$$d = [d_1, d_2, \dots, d_N]^T \quad (3.7)$$

$$w = [w_1, w_2, \dots, w_N]^T \quad (3.8)$$

Ký hiệu Φ là ma trận nội suy kích thước $N \times N$ với các phần tử φ_{ij} :

$$\Phi = \{\varphi_{ij}\}_{i,j=1}^N \quad (3.9)$$

Phương trình (3.5) được viết lại:

$$\Phi w = d \quad (3.10)$$

Giả thiết Φ là ma trận khả đảo thì:

$$w = \Phi^{-1} \cdot d \quad (3.11)$$

Hàm bán kính RBF thường dùng là hàm Gauss có dạng:

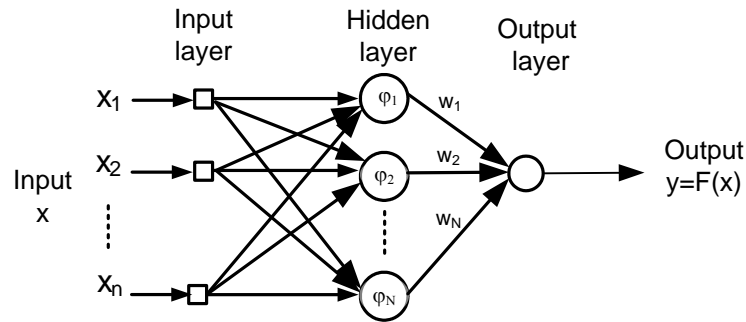
$$\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad \sigma > 0, r \in R \quad (3.12)$$

3.1.10.3 Mạng hàm cơ sở bán kính RBF

Mạng nơ-ron RBF thường được áp dụng cho bài toán phân lớp mẫu hay hồi quy phi tuyến, là mạng truyền thẳng có 3 lớp:

- Lớp vào gồm có n nút cho tín hiệu ngõ vào $x \in R^n$.
- Lớp ẩn là lớp bán kính cơ sở, gồm N nơ-ron, mỗi nơ-ron được mô tả bởi hàm bán kính như sau: $\varphi_j = \varphi(\|x - x_j\|)$, $j = 1, 2, \dots, N$. Trong đó nút dữ liệu thứ j của x_j là tâm của hàm RBF, vector x là vector mẫu ngõ vào. Như vậy mạng RBF thì kết nối từ đầu vào tới nơ-ron ẩn là kết nối trực tiếp.
- Lớp ra là hàm kích hoạt tuyến tính.

Hình 3.7 trình bày cấu trúc mạng RBF với n ngõ vào, N nơ-ron ẩn, và một ngõ ra tuyến tính.



Hình 3.7 Cấu trúc mạng RBF

Trong ứng dụng người ta thường sử dụng hàm Gauss như biểu thức (3.13). Trong đó σ là độ rộng của hàm Gauss. x_j là tâm của hàm nơ-ron thứ j , $j=1,2,\dots,N$. N là số mẫu huấn luyện.

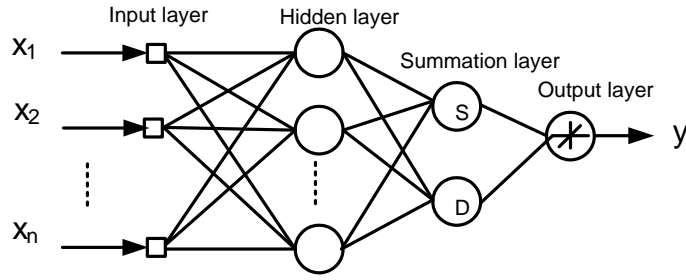
$$\varphi_j = \varphi(x - x_j) = \exp\left(-\frac{1}{2\sigma^2} \|x - x_j\|^2\right), \quad j = 1, 2, \dots, N \quad (3.13)$$

3.1.10.4 Mạng Generalized Regression Nơ-ron Network

Mạng RBF có biến thể là Generalized Regression Nơ-ron Network (GRNN) được đề nghị trong [60]. GRNN được xếp một dạng của mạng nơ-ron xác suất (Probabilistic Nơ-ron Network – PNN). Mạng GRNN cho độ chính xác nhận dạng cao, học nhanh và hội tụ tối ưu với dữ liệu học lớn. Trong trường hợp mẫu nhỏ thì GRNN vẫn có khả năng cho kết quả dự báo tốt.

Mạng GRNN có cấu trúc gồm lớp vào, lớp mẫu (lớp ẩn), lớp tổng, và lớp ra, Hình 3.8 [59]. Mạng GRNN giống như mạng RBF nhưng có chút ít thay đổi hơi đặc biệt là ở lớp tổng (summation layer).

Lớp đầu vào kết nối với lớp mẫu, số nơ-ron lớp vào bằng số biến của bài toán, và đầu ra của nó là khoảng cách của biến đầu vào đến các tâm. Chỉ có lớp mẫu (hay lớp ẩn) có hệ số ngưỡng. Các nơ-ron lớp mẫu có nhớ thông tin quan hệ giữa nơ-ron lớp vào và đặc tính đáp ứng ngõ ra của lớp mẫu. Hàm chuyển đổi của nơ-ron lớp mẫu là dạng hàm Gauss như biểu thức (3.13).



Hình 3.8 Sơ đồ cấu trúc mạng nơ-ron GRNN

Mỗi nơ-ron của lớp mẫu kết nối với hai nơ-ron của lớp đặc biệt, là nơ-ron tổng S (S-summation) và nơ-ron tổng D (D-summation). Nơ-ron S tính toán tổng đầu ra của lớp mẫu với trọng số, trong khi nơ-ron D tính tổng đầu ra của lớp mẫu với trọng số kết nối đơn vị. Trọng số kết nối của nơ-ron lớp mẫu và nơ-ron D là giá trị ngõ ra mục tiêu tương ứng với mẫu đầu vào. Giá trị tính toán của lớp ra là tỉ số của mỗi đầu ra S với mỗi đầu ra D, biểu thức (3.14). Hệ số độ rộng (spread) được xác định qua thực nghiệm.

$$y_i = \frac{\sum_{i=1}^n w_i \varphi_i}{\sum_{i=1}^n \varphi_i} \quad (3.14)$$

3.2 BỘ PHÂN LỚP K-NEAREST NEIGHBOR

Thuật toán K-Nearest Neighbor (K-NN) [61], [62] tìm ra các láng giềng gần nhất của mẫu học và quy về các nhãn lớp của chúng dựa trên các nhãn đa số, điều đó có nghĩa là các mẫu được quy về cùng lớp khi chúng là lân cận của nhau. Kỹ thuật này cho rằng vị trí trong không gian đặc trưng hàm ý một quan hệ họ hàng gần gũi ở giữa các nhãn lớp. K-NN phân lớp vec-tơ biến x chưa biết dựa vào các mẫu gần nhất cần xếp lớp và tất cả các mẫu huấn luyện. Giải thuật đơn giản nhất là K=1 hay 1-NN.

Các bước của thuật toán K-NN:

Bước 1. Xác định số K, K được chọn phải là số lẻ cho trường hợp có 2 lớp hay số lớp chẵn.

Bước 2. Tính khoảng cách giữa đối tượng cần phân lớp với tất cả các mẫu huấn luyện. Khoảng cách Euclide thường được sử dụng.

Bước 3. Xếp khoảng cách theo thứ tự tăng dần

Bước 4. Xác định K láng giềng gần nhất với lớp.

Bước 5. Chỉ định vec-tơ x thuộc vào lớp C_i với số mẫu K gần nhất là lớn nhất.

3.3 BỘ PHÂN LỚP MÁY VECTOR HỖ TRỢ

Thuật toán bộ phân lớp máy vector hỗ trợ (Support Vecotor Machine – SVM) là bộ phân lớp mẫu dựa trên phương pháp tiếp cận lý thuyết học thống kê do Vanipk và Chervonenkis đề nghị [63], [64]. SVM nhằm cực tiểu hóa độ phân lớp sai với một đối tượng dữ liệu mới thông qua cực đại hóa biên giữa siêu phẳng phân cách và dữ liệu. Ý tưởng cơ bản của các SVM là xây dựng một siêu phẳng như là một mặt phẳng quyết định. Mặt phẳng này tách biệt lớp dương (+1) và lớp âm (-1) với biên lớn nhất.

3.3.1 Bộ phân lớp nhị phân tách lớp tuyến tính

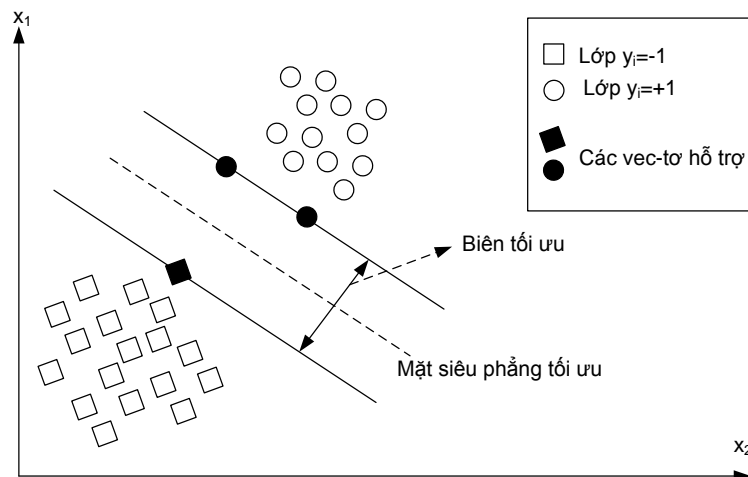
Xét tập dữ liệu huấn luyện $T = \{x_i, y_i\}_{i=1}^N$. Trong đó, x_i là vector dữ liệu đầu vào kích thước n biên và N mẫu, và $y_i \in \{+1, -1\}$ là nhãn lớp của mẫu x_i . Các SVM thực thi cho bài toán phân lớp có hai lớp được trình bày như Hình 3.9.

Siêu phẳng tối ưu phân tách các mẫu dương và các mẫu âm của hai lớp với độ tách biệt cực đại. Độ tách biệt hay là biên xác định bằng khoảng cách giữa các mẫu dương và mẫu âm gần mặt siêu phẳng nhất, Hình 3.9.

Mặt siêu phẳng trong không gian mẫu có phương trình:

$$g(x) = w^T \cdot x + b = 0 \quad (3.10)$$

Mục đích của huấn luyện là tìm ra một siêu phẳng tách biệt dữ liệu huấn luyện tốt nhất.



Hình 3.9 Sơ đồ minh họa thuật toán SVM

Mặt siêu phẳng tách rời các mẫu âm và dương thỏa điều kiện (3.11) và (3.12).

$$w^T \cdot x_i + b \geq 0, \text{ nếu } y_i = +1 \quad (3.11)$$

$$w^T \cdot x_i + b \leq 0, \text{ nếu } y_i = -1 \quad (3.12)$$

Kết hợp (3.11), (3.12), và (3.10) thành dạng thu gọn như (3.13).

$$y_i \cdot (w^T \cdot x_i + b) \geq 1 \quad (3.13)$$

Siêu phẳng tối ưu như phương trình (3.10) là siêu phẳng duy nhất tách tập dữ liệu học với biên cực đại. Nó xác định hướng $w/|w|$ mà khoảng cách từ hình chiếu các vector học của 2 lớp là lớn nhất. Khoảng cách từ một điểm đến siêu phẳng được tính theo phương trình (3.14), với biên là $\frac{2}{\|w\|}$.

$$z = \frac{|g(x)|}{\|w\|} \quad (3.14)$$

Với mỗi điểm x_i tương ứng thuộc nhãn lớp y_i , cần xác định tập thông số (w, b) của siêu phẳng theo (3.15).

$$\begin{cases} \text{Min}_w & \Phi(w) = \frac{1}{2} \|w\|^2 \\ & y_i(w^T x_i + b) \geq 1 \end{cases} \quad (3.15)$$

Đây là bài toán tối ưu phi tuyến, theo điều kiện Krush-Kuhn-Tucker để cực tiểu (3.15) thì thỏa các phương trình sau:

$$\frac{\partial}{\partial w} L(w, b, \lambda) = 0 \quad (3.16)$$

$$\frac{\partial}{\partial b} L(w, b, \lambda) = 0 \quad (3.17)$$

$$\lambda_i \geq 0; i=[1, N] \quad (3.18)$$

$$\lambda_i [y_i \cdot (w^T \cdot x_i + b) - 1] = 0; i=[1, N] \quad (3.19)$$

Trong đó: λ là vector hệ số nhân Lagrange. $L(w, b, \lambda)$ là hàm Lagrange như (3.20).

$$L(w, b, \lambda) = \frac{1}{2} w^T w - \sum_1^N \lambda_i [y_i \cdot (w^T \cdot x_i + b) - 1] \quad (3.20)$$

$$w = \sum_{i=1}^N \lambda_i y_i x_i \quad (3.21)$$

$$\sum \lambda_i y_i = 0 \quad (3.22)$$

3.3.2 Biên tối ưu mềm

Trong trường hợp tập mẫu huấn luyện có thể phân chia tuyến tính được nhưng có nhiễu, Hình 3.10. Hình 3.10 là minh họa cho trường hợp hai lớp dữ liệu không hoàn toàn tách biệt. Trong trường hợp này ràng buộc (3.13) trở thành ràng buộc (3.23).

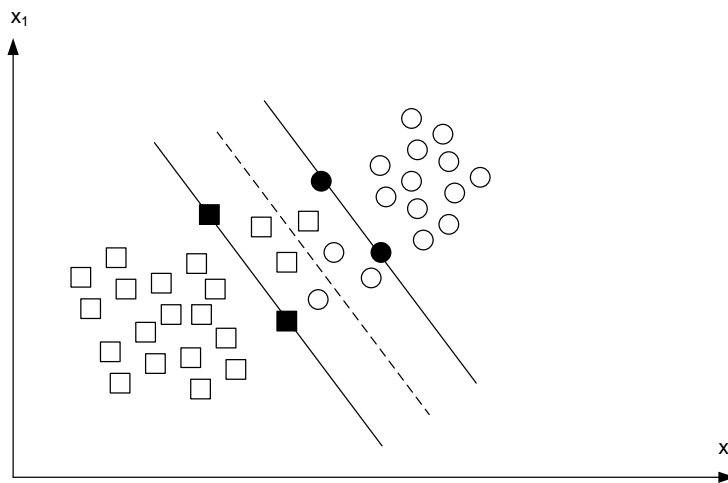
$$y_i \cdot (w^T \cdot x_i + b) \geq 1 - \xi_i \quad (3.23)$$

Trong đó, ξ_i là biến nới lỏng (slack variable), $\xi_i \geq 0$.

Khi đó, bài toán (3.15) trở thành (3.24).

$$\left\{ \begin{array}{l} \text{Min}_w \quad \Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ y_i (w^T x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{array} \right. \quad (3.24)$$

Trong đó, w là trọng số của mặt siêu phẳng. C là hệ số xác định trước, là hệ số phạt định nghĩa giá trị ràng buộc.



Hình 3.10 Minh họa cho trường hợp dữ liệu hai lớp không hoàn toàn tách biệt
Tuy nhiên, trong trường hợp dữ liệu phức tạp phi tuyến thì không thể phân chia tuyến tính được. Trong trường hợp này, giả sử tồn tại ánh xạ từ không gian đầu vào 1 chiều vào không gian k chiều, mà ở đó có thể thỏa mãn điều kiện tách biệt bởi một siêu phẳng.

$$\Phi: R^l \rightarrow R^k \quad (3.25)$$

Khi đó, vector x_i trong không gian R^l tương ứng với vector $\Phi(x_i)$ trong không gian R^k , (3.24) được viết lại thành (3.26).

$$\left\{ \begin{array}{l} \text{Min}_w \quad \Phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{array} \right. \quad (3.26)$$

3.3.3 Hàm nhân

Việc tính toán trực tiếp hàm $\Phi(x_i)$ là việc rất khó khăn. Nếu biết hàm nhân (Kernel function) $K(x_i, x_j)$, để tính tích vô hướng $\Phi(x_i)\Phi(x_j)$ trong không gian k chiều thì việc thực thi tính toán không cần thực hiện trực tiếp với ánh xạ $\Phi(x_i)$ [61], [65]. Trong pha huấn luyện, giải thuật chỉ cần hàm nhân $K(x_i, x_j)$ và hàm $\Phi(x_i)$ không cần thiết phải biết, bởi vì nó đã được xác định qua hàm nhân như (3.27).

$$K(x_i, x_j) = \Phi(x_i)\Phi(x_j) \quad (3.27)$$

Các hàm nhân ứng dụng trong bài toán nhận dạng mẫu thường gồm các dạng sau:

- Hàm đa thức: $K(x, x_i) = (x^T x_i + 1)^d, d > 0$
- Hàm cỡ sở bán kính: $K(x, x_i) = \exp(-\|x - x_i\|^2 / 2\sigma^2)$
- Hàm Hyperbolic Tangent (sigmoid): $K(x, x_i) = \beta x^T x_i + \gamma$
 β, γ là hệ số do người dùng chọn.

3.4 TÓM TẮT CHƯƠNG 3

Tác giả đã nghiên cứu và trình bày lý thuyết cơ sở của các bộ phân lớp gồm:

- Bộ phân lớp mạng perceptron nhiều lớp (MLP).
- Bộ phân lớp mạng hồi quy tổng quát (GRNN).
- Bộ phân lớp K-NN.
- Bộ phân lớp máy vec-tơ hỗ trợ (SVM).

Bộ phân lớp MLP được đề xuất áp dụng trong các công trình đã công bố, cụ thể là trong bài báo (IV-VII). Bộ phân lớp GRNN được đề xuất áp dụng trong bài báo (II, III, VI). Vấn đề huấn luyện mạng nơ-ron tốn nhiều thời gian với nhiều lần thực hiện huấn luyện để chọn thông số mạng phù hợp cho kết quả độ chính xác nhận dạng cao. K-NN (1-NN, $K=1$) được đề xuất áp dụng trong bài báo (I). Chương 6, tác giả đã đề nghị áp dụng bộ phân lớp 1-NN tham gia vào giai đoạn đầu của quy trình xây dựng mô hình nhận dạng sử dụng mạng nơ-ron như chọn

biến, đánh giá chọn mẫu trong giải thuật phân cụm dữ liệu. Với kết quả tập biến và tập mẫu được chọn nhờ đánh giá từ bộ phân lớp 1-NN, lúc này dữ liệu đã qua xử lý nên tinh gọn hơn, và được sử dụng cho bước tiếp theo trong quy trình huấn luyện xây dựng các bộ phân lớp giúp giảm gánh nặng thực thi trong quá trình huấn luyện xây dựng mô hình nhận dạng. Bộ phân lớp máy vec-tơ hỗ trợ (SVM) là công cụ mạnh áp dụng cho bài toán phân lớp gồm hai lớp. Các bộ phân lớp (MLP, GRNN, SVM) được sử dụng để xây dựng mô hình nhận dạng, so sánh đánh giá độ chính xác nhận dạng. Kết quả xây dựng mô hình và so sánh kết quả nhận dạng với các bộ phân lớp được trình bày ở chương 6.

Chương 4. LỰA CHỌN THÔNG SỐ ĐẶC TRƯNG CHO HỆ THỐNG NHẬN DẠNG THÔNG MINH TRONG CHẨN ĐOÁN ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

Trong bài toán chẩn đoán ổn định động hệ thống điện trên cơ sở công nghệ tính toán thông minh nhân tạo, trạng thái hệ thống điện được biểu hiện bằng vec-tơ $x = \{x_{ij}\}$, $i = [1, N]$, $j = [1, n]$, có n chiều trong không gian quan sát R^n . Thành phần của vec-tơ x bao gồm các dấu hiệu đặc trưng cho trạng thái tiền sự cố và trạng thái sự cố hệ thống điện có n biến và N mẫu được trình bày dạng ma trận dữ liệu như sau:

$$\begin{array}{c}
 \text{N mẫu} \left\{ \begin{array}{c} \overbrace{\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Nn} \end{bmatrix}}^{n \text{ biến}} \end{array} \right.
 \end{array}$$

4.1 XÂY DỰNG TẬP MẪU

Xây dựng hệ thống nhận dạng bắt đầu bằng việc xác định tập dữ liệu hay tập mẫu mang tính đặc trưng thể hiện trạng thái chế độ hệ thống điện. Tập mẫu được chọn đại diện cho tất cả các chế độ vận hành của hệ thống điện, đó là các trạng thái vận hành hệ thống tiêu biểu về quan điểm ổn định hệ thống điện. Trạng thái ổn định và không ổn định của hệ thống điện được đánh giá qua từng sự cố để tri thức về ổn định quá độ của hệ thống điện bao quát các dạng sự cố.

Các thông số đặc trưng cho chế độ tiền sự cố bao gồm dòng công suất trên các nhánh, điện áp tại các nút của hệ thống điện. Chú ý các đường dây tải nặng, điện áp tải ở các nút tải, máy phát có công suất lớn, độ sụt áp tại các nút tải khi xảy ra sự cố,...

Các thông số đặc trưng cho chế độ sự cố hệ thống điện bao gồm độ thay đổi công suất trên các nút tải lớn, trên các đường dây truyền tải,... Trong đó, sự dao

động công suất tác dụng, độ sụt điện áp mang tính thông tin cao liên quan đến trạng thái ổn định hệ thống điện.

Dữ liệu có thể gồm dữ liệu lịch sử lưu trữ trong hệ thống thu thập xử lý dữ liệu, mô phỏng off-line. Vì điều kiện vận hành hệ thống điện có dây rộng, mô phỏng off-line được thực hiện để tạo cơ sở dữ liệu ban đầu với dữ liệu cấu thành đầy đủ đại diện của các thông số chế độ sự cố của hệ thống điện.

Mô phỏng off-line để thu thập dữ liệu là công việc vô cùng nặng nhọc, trong khi mô phỏng với tất cả các sự cố, tại tất cả các nút, mọi vị trí trên đường dây là công việc đầy khó khăn và không khả thi trong thực hiện. Tuy nhiên, trên cơ sở kết quả khảo sát của các trường hợp điển hình nhất việc phân tích các tình huống sự cố trong điều kiện thực tế và sự hiểu biết bản chất vật lý của quá trình quá độ có thể giới hạn số lượng các sự cố cần xem xét cũng như số điểm xuất hiện. Điều này cho phép xem xét loại trừ các sự cố không gây mất ổn định hệ thống điện và việc lấy mẫu tập trung vào vùng nguy hiểm khi xuất hiện sự cố sẽ gây mất ổn định hay là ‘vùng quan sát’.

Khi xây dựng tập học nên tính đến ý kiến chuyên gia vận hành hệ thống điện. Cần xem xét thận trọng ở chế độ tải cực đại, đây là chế độ dễ gây hại đến tính ổn định hệ thống điện khi xuất hiện sự cố, và số mẫu mất ổn định tập trung phần lớn là ở chế độ tải cực đại.

4.2 LỰA CHỌN TẬP BIẾN

Đây là bước chọn biến đầu vào và biến đầu ra cho mô hình nhận dạng. Biến đầu vào là vec-tơ biến chứa thông số trạng thái hệ thống. Biến đầu ra là giá trị thông số đại diện cho trạng thái ổn định động hệ thống điện. Có hai dạng biến cho chế độ hệ thống điện là biến chứa dữ liệu ở chế độ xác lập và ở chế độ sự cố.

Biến ở chế độ xác lập hay tiền sự cố (Pre-fault) là thông tin giúp hệ thống nhận dạng biết tính trạng mất ổn định xảy ra để phát tín hiệu đến hệ thống điều khiển ngăn chặn trước (preventive control), giúp vận hành viên hoặc hệ thống điều khiển ra quyết định đưa hệ thống điện vào vùng vận hành an toàn, phòng khi sự cố

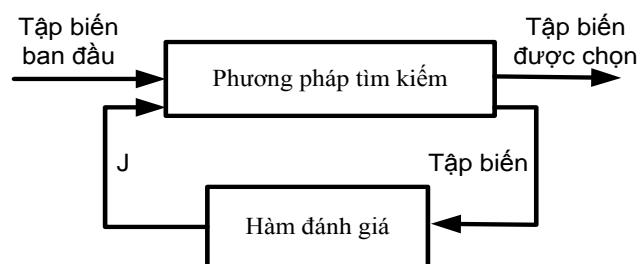
xảy ra gây mất ổn định hệ thống điện, mặc dù thực tế là sự cố có thể không xảy ra. Hẳn nhiên, việc điều khiển này phải chịu chi phí tổn thất cho việc điều chỉnh lại lịch trình phát của các máy phát điện. Đặc trưng cho chế độ tiên sự cố là các điểm vận hành hệ thống điện gồm điện áp tại các bus, công suất các máy phát điện, dòng công suất trên các nhánh, công suất của các tải,...

Biến ở chế độ động hay chế độ sự cố (Fault-on) là thông tin cho bộ phân loại nhận biết tình trạng mất ổn định xảy ra để phát tín hiệu đến hệ thống điều khiển khẩn cấp (emergency control) và lệnh điều khiển ngay lập tức phải được thi hành để đưa hệ thống trở về trạng thái ổn định, cứu hệ thống điện khỏi nguy cơ tan rã. Đặc trưng chế độ sự cố gồm độ lệch điện áp tại các bus, độ lệch công suất phát, độ lệch công suất trên các nhánh, độ lệch công suất các tải,... Trong thực tế phân tích ổn định hệ thống điện thì độ sụt điện áp và độ thay đổi công suất tác dụng chứa thông tin rất cao và liên quan trực tiếp đến trạng thái ổn định hệ thống điện. Đây là các biến có chứa thông tin cao được lựa chọn làm biến đầu vào cho bộ nhận dạng.

Biến đầu ra đại diện cho điều kiện ổn định động của hệ thống điện được gán nhãn biến nhị phân cho lớp ổn định và lớp không ổn định.

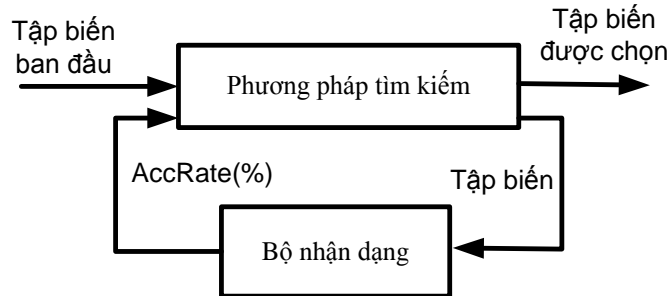
Kỹ thuật chọn biến có hai kỹ thuật chính là kỹ thuật lọc (Filter) và kỹ thuật chọn biến bao gói (Wrapper) [66].

Kỹ thuật Filter áp dụng các giải thuật tìm kiếm với hàm đánh giá hay hàm mục tiêu để hướng chiến lược tìm kiếm tập biến đạt giá trị mục tiêu (J) tốt nhất được trình bày ở Hình 4.1.



Hình 4.1 Kỹ thuật chọn biến Filter.

Kỹ thuật Wrapper áp dụng kỹ thuật tìm kiếm với giá trị mục tiêu là độ chính nhận dạng đạt giá trị cao nhất được trình bày ở Hình 4.2.



Hình 4.2 Kỹ thuật chọn biến Wrapper.

4.2.1 Tiêu chuẩn chọn biến

4.2.1.1 Tiêu chuẩn Fisher

Hàm khoảng cách Fisher [37], [62], [67] hay chuẩn Fisher tìm kiếm độ tách biệt tối ưu tuyến tính giữa 2 lớp dữ liệu. Một tập dữ liệu D gồm n biến x_1, x_2, \dots, x_n với n_1 mẫu trong lớp C_1 và n_2 mẫu trong lớp C_2 , tác vụ là tìm kiếm ánh xạ cực đại $y = W^T X$. Điều này tương ứng với việc tìm kiếm đường thẳng khi chiếu tập dữ liệu lên đường thẳng cho độ tách biệt tối đa. Các biến có giá trị F lớn hơn thì khả năng tách biệt dữ liệu giữa 2 lớp cao hơn.

$$F(w) = \frac{|m_1 - m_2|^2}{\sigma_1^2 + \sigma_2^2} \quad (4.1)$$

Trong đó: m_i là giá trị trung bình dữ liệu của lớp C_i và σ_i^2 là phương sai của dữ liệu của lớp C_i .

4.2.1.2 Tiêu chuẩn Divergence

Khoảng cách Divergence [37], [62], [66] là tiêu chuẩn đo lường độ phân tán của hai lớp dữ liệu. Khoảng cách D_{ij} giữa hai lớp được trình bày ở công thức (4.2). Khoảng cách D_{ij} càng lớn là thì dữ liệu hai lớp càng phân tán.

$$D_{ij} = \frac{1}{2} \text{tr} \left[(C_i - C_j)(C_j^{-1} - C_i^{-1}) \right] + \frac{1}{2} \text{tr} \left[(C_j^{-1} + C_i^{-1})(m_i - m_j)(m_i - m_j)^T \right] \quad (4.2)$$

Trong đó: tr là tổng đường chéo của ma trận và bằng giá trị tổng các giá trị riêng của ma trận; C_i là ma trận hiệp phương sai của lớp i kích thước $[n \times n]$; C_j là ma trận hiệp phương sai của lớp j kích thước $[n \times n]$; m_i vector trung bình của lớp i kích thước $[n \times 1]$; m_j là vector trung bình của lớp j kích thước $[n \times 1]$; $(m_i - m_j)^T$ là hoán vị của $(m_i - m_j)$ và n là số biến.

4.2.1.3 Tiêu chuẩn Scatter Matrix (SM)

Với mẫu n dữ liệu kích thước $x_1, x_2, \dots, x_N, n=[1, N]$. Ma trận hiệp phương sai của toàn bộ mẫu theo công thức (4.3) [61], [62], [66].

$$S_m = \frac{1}{N} \sum_{n=1}^N (x_n - m)(x_n - m)^T \quad (4.3)$$

Trung bình của toàn bộ mẫu:

$$m = \frac{1}{N} \sum_{n=1}^N x_n \quad (4.4)$$

Trung bình mẫu của nhóm thứ c_i :

$$m_i = \frac{1}{N_i} \sum_{x_n \in c_i} x_n \quad (4.5)$$

Ở đây: c là số lớp; N_i là số mẫu của nhóm c_i ; N là tổng số mẫu.

Ma trận nhóm trong (within-class scatter Matrix) là:

$$S_w = \frac{1}{N} \sum_{i=1}^c \frac{N_i}{N} S_i \quad (4.6)$$

$$S_i = \frac{1}{N_i} \sum_{x_n \in c_i} (x_n - m_i)(x_n - m_i)^T \quad (4.7)$$

Ma trận nhóm giữa (between-class scatter matrix), biểu thị độ phân tán dữ liệu của trung bình nhóm quanh trung bình tổng của toàn bộ mẫu dữ liệu.

$$S_b = \sum_{i=1}^c \frac{N_i}{N} (m_i - m)(m_i - m)^T \quad (4.8)$$

Tổng đường chéo (trace) của ma trận phân tán là tổng các phương sai của biến quanh giá trị trung bình. Độ tách biệt lớp của tập biến được đo lường bởi hàm mục tiêu:

$$J = \text{tr}\{S_w^{-1}S_m\} \quad (4.9)$$

Giá trị J có giá trị càng lớn thì các lớp càng tách biệt.

4.2.2 Các giải thuật tìm kiếm tập biến

Việc lựa chọn biến đòi hỏi phải áp dụng chiến lược tìm kiếm để tìm ra tập biến đặc trưng. Chiến lược tìm kiếm được chia ra làm tìm kiếm toàn cục và tìm kiếm cục bộ. Chiến lược tìm kiếm toàn cục có ưu điểm lớn là cho kết quả tối ưu nhưng tốn kém thời gian tính toán, và đôi khi không tính toán được do khó khăn trong tính toán giá trị ma trận hiệp phương sai của tập biến kích thước lớn. Vì vậy, chiến lược tìm kiếm tối ưu không phù hợp khi số biến đầu vào lớn. Trong trường hợp biến đầu vào lớn chiến lược tìm kiếm tối ưu cục bộ có ưu điểm về thời gian tìm kiếm, đơn giản trong tính toán, khả thi trong thực hiện do quá trình tìm kiếm không xuyên qua toàn bộ không gian tìm kiếm.

Luận án giới thiệu giải thuật SFFS với tiêu chuẩn chọn biến SM như biểu thức (4.9), và Bảng 4.2 trình bày giải thuật SFFS. Để thực hiện giải thuật SFFS thì bước đầu cần khởi tạo 2 biến nhờ vào giải thuật SFS (Sequential Forward Selection). Giải thuật SFS được trình bày ở Bảng 4.1

4.2.2.1 Giải thuật Sequential Forward Floating Selection (SFFS):

SFFS và SBFS [61], [62] (Sequential Backward Floating Selection) thuộc giải thuật FSA (Floating Search Algorithm). Chiến lược tìm kiếm SFFS bắt đầu tìm kiếm với tập biến rỗng và áp dụng chiến lược tiến thêm từng biến một để tạo tập biến. Trong mỗi bước tiến giải thuật kết hợp chiến lược tìm lùi để tháo bỏ từng biến một và lùi lại vị trí tập biến có hàm mục tiêu tốt hơn. Quá trình tìm kiếm dừng khi số biến đạt giá trị d mong muốn. Giải thuật SBFS bắt đầu với tập đầy đủ biến và áp dụng chiến lược lùi để tháo bỏ từng biến một để lựa chọn tập biến.

Trong mỗi bước lùi giải thuật kết hợp với chiến lược tiến để thêm một biến vào tập biến và chọn tập biến có hàm mục tiêu tốt hơn. Quá trình tìm kiếm dừng khi số biến đạt giá trị d mong muốn. Giải thuật SBFS tốn rất nhiều thời gian tính toán do phải thực thi với số biến ban đầu đầy đủ. Do vậy, luận án này chỉ giới thiệu và áp dụng giải thuật SFFS.

Bảng 4.1 Giải thuật chọn biến SFS

Giải thuật SFS

Đầu vào: Tập biến đầy đủ Y có D biến

Đầu ra: Tập biến X_k có $k=d$ biến

Bước 1. Tập biến ban đầu rỗng, tiến hành thêm từng biến x_j từ Y vào X_k

$$X_{k+1} = \{X_k \cup x_j\}, \text{ tính } J(k+1), j=[1,D];$$

Bước 2. Chọn tập biến X_{k+1} thỏa điều kiện $\max J(k+1)$.

$$Y=Y \setminus x_j, D=D-1; k=k+1;$$

Bước 3. $k < d$ quay về bước 1. Giải thuật dừng khi $k > d$.

Bảng 4.2 Giải thuật chọn biến SFFS

Giải thuật Giải thuật SFFS

Đầu vào: Tập biến đầy đủ Y có D biến

Đầu ra: Tập biến X_k có $k=d$ biến

Bước 1: Khởi tạo 2 biến ban đầu với giải thuật SFS, $k=2$. Chọn tập X_k có giá trị hàm mục tiêu lớn nhất $J(k)$.

Bước 2: Lựa chọn x_j từ $Y-X_k$, biến x_j làm tăng giá trị hàm mục tiêu J lớn nhất được đưa vào tập biến hiện hành: $X_{k+1} = X_k + x_j$. Giải thuật dừng khi $k > d$.

Bước 3: Tìm biến, x_r , trong tập biến hiện hành, X_{k+1} , biến làm giảm giá trị hàm mục tiêu J ít nhất. Nếu biến x_r là x_j thì $J_{k+1}=J(X_{k+1})$, $k=k+1$ quay về bước 2; ngược lại thực hiện loại biến này ra khỏi tập biến hiện hành $X_k'=X_{k+1}-x_r$, thực hiện bước 4.

Bước 4: Tiếp tục loại biến yếu từ tập biến X'_k khi đó tập biến hình thành là X'_{k-1} trong khi $J(X'_{k-1})>J_{k-1}$ thì $k=k-1$ tiếp tục cho đến khi $k=2$ quay về bước 2.

4.2.2.2 Giải thuật xếp hạng biến (FR-Feature Ranking):

Đây là phương pháp đơn giản và tốn ít thời gian tính toán. Bằng cách tính giá trị mục tiêu từng biến đơn, sắp xếp giá trị này theo giá trị ưu tiên từ quan trọng đến ít quan trọng, d biến nằm trong nhóm đầu ưu tiên được chọn. Giải thuật FR được trình bày ở Bảng 4.3.

Bảng 4.3 Giải thuật xếp hạng biến (FR)

Giải thuật FR

Đầu vào: Tập biến đầy đủ Y có D biến.

Đầu ra: Tập X_k có $k=d$ biến.

Bước 1: Tính giá trị tiêu chuẩn J cho từng biến trong D biến theo (4.1) hoặc (4.2).

Bước 2: Xếp hạng các biến của D biến có giá trị hàm mục tiêu J từ lớn đến nhỏ.

Bước 3: Xây dựng tập con biến có số biến từ d_1 đến d_2 biến, $d_1 \leq d \leq d_2 \leq D$.

Bước 4: Đánh giá độ chính xác nhận dạng của các tập con biến ở bước 3.

Bước 5: Chọn tập con biến X_k có độ chính xác cao nhất trong vùng khảo sát.

4.3 TÓM TẮT CHƯƠNG 4

Trong chương này, tác giả đã phân tích tính đặc trưng trạng thái chế độ hệ thống điện. Các thông số đặc trưng cho chế độ sự cố hệ thống điện chứa thông tin ở chế độ trước và ngay sau sự cố xảy ra. Xây dựng tập mẫu là công việc phức tạp

và nặng nhọc, để giảm gánh nặng trong xây dựng tập mẫu thì việc lấy mẫu cần tập trung vào vùng nguy hiểm khi xuất hiện sự cố gây mất ổn định hay ‘vùng quan sát’.

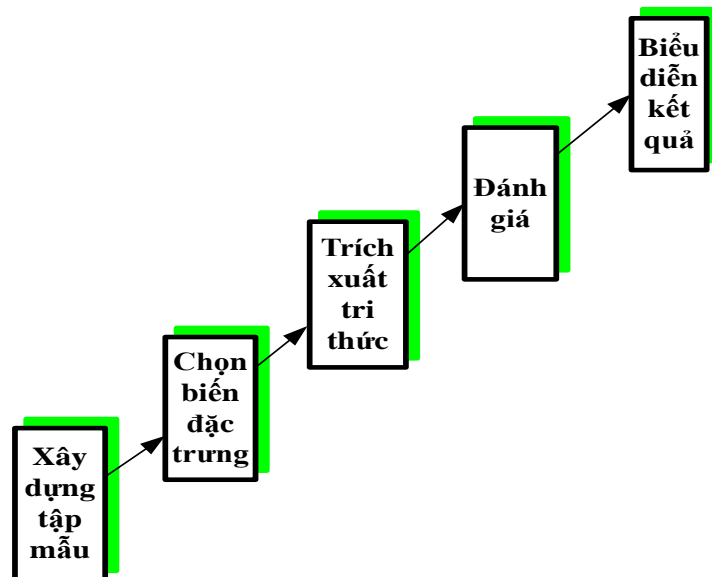
Phân tích chỉ ra độ dao động công suất tác dụng của tải, độ dao động công suất tác dụng trên đường dây, độ sụt điện áp tại các nút là các biến chứa thông tin cao liên quan trực tiếp đến trạng thái ổn định hệ thống điện khi sự cố. Đây là các biến có chứa thông tin cao được lựa chọn làm biến đầu vào cho bộ nhận dạng.

Tác giả đã nghiên cứu tổng kết một cách tổng quát kỹ thuật chọn biến gồm kỹ thuật lọc và kỹ thuật bao gói, và trình bày vấn đề một cách rõ ràng, dễ hiểu. Ba tiêu chuẩn chọn biến để áp dụng vào chọn biến cụ thể là chuẩn Fisher, Divergence, và Scatter Matrix. Luận án trình bày chi tiết các bước của hai phương pháp chọn biến là phương pháp xếp hạng (FR) và phương pháp tìm kiếm tiến lùi (SFBS). Kết quả nghiên cứu trình bày cụ thể ở chương 6 và các bài báo được công bố.

Chương 5. PHƯƠNG PHÁP XÂY DỰNG HỆ THỐNG NHẬN DẠNG THÔNG MINH CHẨN ĐOÁN ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN

5.1 GIỚI THIỆU

Mô hình hệ thống nhận dạng dựa trên cơ sở công nghệ tính toán thông minh nhân tạo chẩn đoán ổn định động hệ thống điện được xây dựng gồm các khâu cơ bản được trình như hình 5.1. Trong đó, phát triển hệ thống thông minh gồm 5 khâu như sau: xây dựng tập mẫu điển hình, lựa chọn biến đặc trưng, trích xuất tri thức, đánh giá kết quả, và biểu diễn kết quả.



Hình 5.1 Các khâu cơ bản của quá trình xây dựng hệ thống nhận dạng.

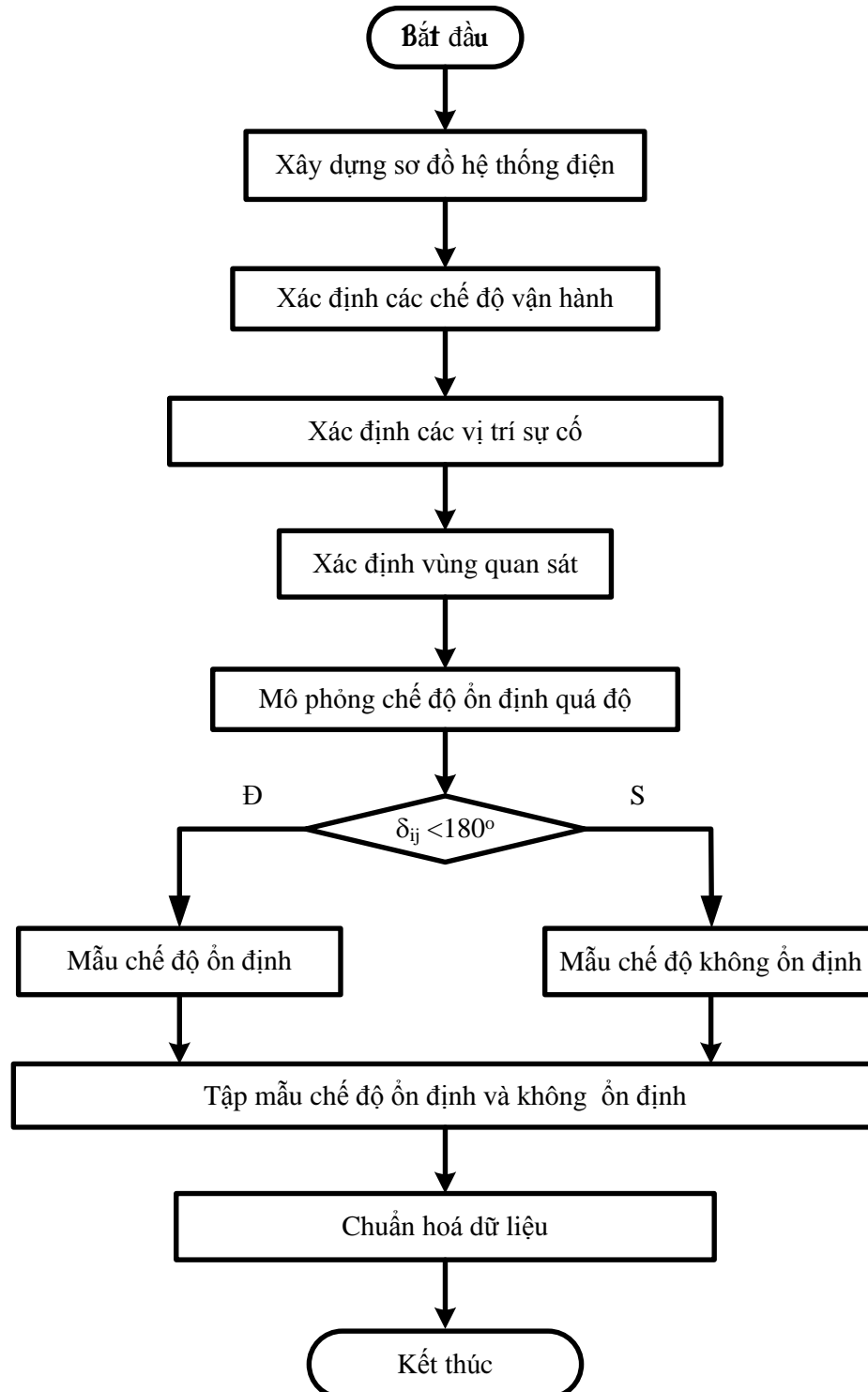
5.2 XÂY DỰNG TẬP MẪU

Vì điều kiện vận hành hệ thống điện có dây rộng, mô phỏng off-line được thực hiện để tạo cơ sở dữ liệu ban đầu đại diện cho các thông số chế độ sự cố của hệ thống điện. Sau khi được học off-line, mạng nơ-ron có khả năng lưu trữ thông tin, tự học để cập nhật tri thức, làm giàu tri thức cho mô hình, và khả năng nhận ra các mẫu mới [67]. Xây dựng tập mẫu là công việc quan trọng trong bài toán nhận dạng ứng dụng mạng nơ-ron. Các bước thực hiện xây dựng tập mẫu được trình bày tiếp sau và quy trình như Hình 5.2.

- **Xây dựng sơ đồ thay thế hệ thống điện:** thiết lập sơ đồ hệ thống điện từ hệ thống điện cần nghiên cứu. Các phần tử trong hệ thống điện như đường dây, máy biến áp, máy phát điện,... được mô hình hoá khá chi tiết với sự hỗ trợ các phần mềm chuyên dụng. Trong nghiên cứu này phần PowerWorld được đề nghị sử dụng.
- **Xác định chế độ vận hành:** cần xem xét tất cả chế độ, các mức tải, tham khảo ý kiến chuyên gia để dữ liệu đại diện bao quát các thông số chế độ vận hành của hệ thống điện.
- **Xác định các vị trí sự cố:** xem xét các sự cố khi xảy ra làm xuất hiện các dao động và có thể dẫn đến mất ổn định hệ thống điện. Đây có thể các trường hợp bất thường như cắt điện máy phát, cắt điện đường dây tải lớn, gia tăng tải đột ngột, cắt máy biến áp, và nguy hiểm nhất là ngắn mạch xảy ra. Ngắn mạch cần xem xét tại các bus của các trạm biến áp công suất lớn, các đường dây liên kết giữa các nhà máy điện và giữa các vùng với nhau. Ngắn mạch cần xem xét gồm các dạng như ngắn mạch một pha, hai pha, ba pha,... với vị trí tại tất cả các thanh góp và dọc các đường dây truyền tải. Đối với điểm ngắn mạch trên đường dây được lựa chọn là dọc từ đầu đường dây đến cuối đường dây.
- **Xác định vùng quan sát:** là vùng bao gồm các phần tử hay nhóm các phần tử của hệ thống điện mà ở đó xuất hiện các điểm ngắn mạch gây mất ổn định hệ thống điện. Chú ý các thanh góp chính, các đường dây truyền tải có công suất lớn, các thanh cái tại đầu cực máy phát, ở các trạm biến áp công suất lớn, và chú ý cao mức tải cực đại.
- **Mô phỏng chế độ quá độ:** sử dụng chương trình mô phỏng chuyên dụng để đánh giá ổn định động cho hệ thống điện. Tiến hành mô phỏng và xếp mẫu dữ liệu vào tập ổn định và tập không ổn định theo luật (1.6) như trình bày ở chương 2.
- **Chuẩn hoá dữ liệu:** dữ liệu của các biến có các đơn vị đo khác nhau sẽ làm ảnh hưởng đến kết quả tính toán trong nhận dạng. Dữ liệu thu thập cần chuẩn hoá trước khi đưa vào khâu trích xuất tri thức hay học. Phương pháp chuẩn hoá dữ liệu phổ biến được áp dụng theo biểu thức sau:

$$z_{ij}=[x_{ij}-m(x_{ij})]/\sigma(x_{ij}) \quad (5.1)$$

Trong đó: x_{ij} là dữ liệu ban đầu của biến thứ j ; $m(x_{ij})$ là giá trị trung bình dữ liệu của biến thứ j ; $\sigma(x_{ij})$ là độ lệch chuẩn của dữ liệu của biến thứ j ; z_{ij} là dữ liệu chuẩn hóa của biến thứ j .



Hình 5.2 Quy trình xây dựng tập mẫu

5.3 NGHIÊN CỨU ĐỀ XUẤT QUY TRÌNH CHỌN BIẾN ĐẶC TRƯNG

Số biến đầu vào ban đầu thường rất lớn. Lựa chọn biến là chọn tập biến đặc trưng, thật ra là giảm không gian biến, giúp giảm chi phí đo lường cảm biến, giảm chi phí tính toán và thời gian huấn luyện. Giảm biến phải đảm bảo độ chính xác nhận dạng kỳ vọng, tuy nhiên trong một số trường hợp sẽ cải thiện độ chính xác nhận dạng do loại được biến nhiễu. Vấn đề chọn biến được các tác giả trong [37] trình bày một cách tổng quát với quy trình các bước chọn biến. Tuy nhiên, các công trình công bố trong lĩnh vực chẩn đoán ổn định hệ thống điện chưa trình bày, giải thích một cách có hệ thống về kỹ thuật lọc (Filter) và kỹ thuật bao gói (Wrapper).

Kỹ thuật chọn biến Filter có ưu điểm là đơn giản trong tính toán, tuy nhiên phương pháp này chưa cho biết được độ chính xác nhận dạng. Kỹ thuật Wrapper có ưu điểm là cho biết được độ chính xác nhận dạng, nhưng nhược điểm chính là tốn rất nhiều thời gian tính toán nên không khả thi trong thực hiện với tập biến có số biến ban đầu lớn. Lưu ý rằng độ chính xác nhận dạng là yêu cầu quan trọng xây dựng mô hình nhận dạng. Từ những phân tích trên Nghiên cứu sinh đề nghị quy trình chọn biến kết hợp kỹ thuật Filter và Wrapper (F&W) theo quy trình như Hình 5.3 gồm 5 bước: **Xây dựng tập biến ban đầu, Tìm kiếm tập biến ứng viên, Huấn luyện, Đánh giá, Chọn tập biến.**

- **Bước 1. Xây dựng tập biến ban đầu:** Cơ sở dữ liệu ổn định quá độ hệ thống điện ban đầu là $D = \{x_i, y_i\}_{i=1}^N$, trong đó x_i là vector đặc trưng trạng thái vận hành hệ thống điện với n biến đầu vào, N mẫu, và y_i là biến đầu ra tương ứng. Cơ sở dữ liệu D biểu diễn dưới dạng ma trận như sau:

	x{n biến vào}				y{1 biến ra}
	┌───────────┬───────────┬───────────┬───────────┐				
	x_{11}	x_{12}	...	x_{1n}	y_1
	x_{21}	x_{22}	...	x_{2n}	y_2

	x_{N1}	x_{N2}	...	x_{Nn}	y_N
N mẫu	}				

Biến đầu vào ở chế độ động chứa thông tin chế độ xác lập và chế độ sự cố, kết hợp với ý kiến chuyên gia chọn các biến ban đầu gồm độ lệch điện áp tại các bus ΔV_{bus} , độ lệch công suất tải ΔP_{load} , độ lệch công suất trên các đường dây truyền tải ΔP_{flow} . Biến đầu ra được mã hoá dạng nhị phân, $\{1\}$ là ‘ổn định’, $\{0\}$ là ‘không ổn định’. Điều này làm cho bài toán chẩn đoán ổn định động hệ thống điện là bài toán phân lớp.

- **Bước 2. Tìm kiếm tập biến ứng viên:** Bước này nhằm tìm kiếm các tập biến ứng viên tiềm năng, tức là các tập biến có khả năng cho hệ thống nhận dạng làm việc với độ chính xác cao. Ở bước này, các tiêu chuẩn thống kê để đánh giá chọn tập biến được sử dụng gồm tiêu chuẩn Fisher, Divergence hay SM được trình bày theo biểu thức (4.1), (4.2) và (4.9) của chương 4.
- **Bước 3. Huấn luyện:** Tất cả các tập con biến ứng viên ở bước 2 đều được huấn luyện để tìm tập biến tốt nhất. Đây cũng có thể được xem như là bước tiền xác định bộ phân lớp. Tuy nhiên, trong bước này nghiên cứu sinh đề nghị áp dụng bộ phân lớp K-NN nhờ vào tính đơn giản và tính toán nhanh của nó.

Để đánh giá sau khi học xong mô hình nhận dạng có khả năng nhận biết được các mẫu mới, cũng như không làm mất tính tổng quát của mô hình nghiên cứu. Tập dữ liệu ban đầu được chia thành tập học hay tập huấn luyện và tập kiểm tra. Việc tổng quát hóa của mô hình nhận dạng thể hiện qua việc phân tích độ chính xác tập mẫu kiểm tra hay tập mẫu không tham vào quá trình huấn luyện.

- **Bước 4. Đánh giá:**

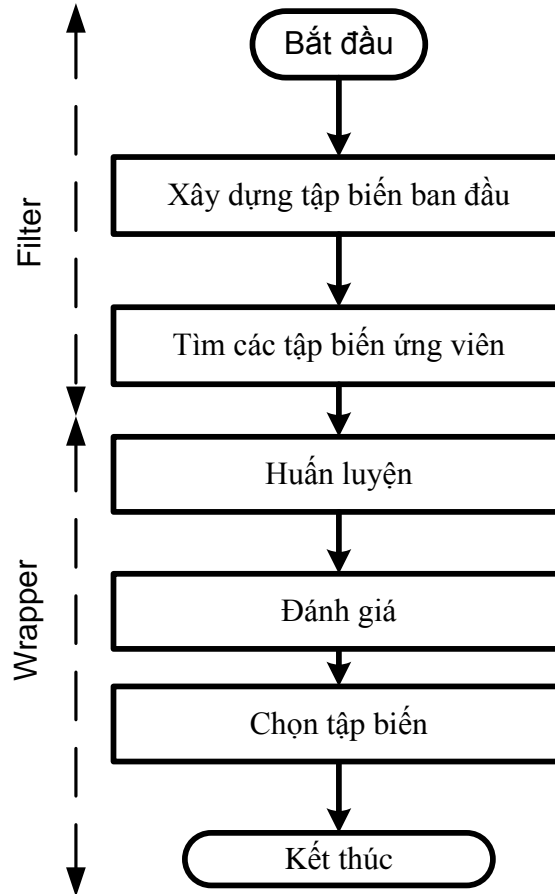
Độ chính xác tổng thể phân lớp được tính theo công thức (5.2).

$$\text{AccRate}(\%) = \frac{n_r}{N} \cdot 100 \quad (5.2)$$

Trong đó: n_r là số mẫu phân lớp đúng. N là tổng số huấn luyện.
Độ sai số phân lớp tổng thể được tính theo công thức (5.3).

$$\text{ErrRate}(\%) = 1 - \text{AccRate} \quad (5.3)$$

- **Bước 5. Chọn tập biến:** Đây là bước ra quyết định chọn tập biến sau cùng. Tập con biến được chọn là tập con biến cho độ chính xác phân lớp cao nhất trong các tập con ứng viên.



Hình 5.3 Quy trình chọn biến đặc trưng F&W

5.4 TRÍCH XUẤT TRI THỨC

Trích xuất tri thức là xây dựng bộ phân lớp hay bộ nhận dạng làm nhiệm vụ học tri thức từ mối quan hệ vào-ra. Sau khi qua khâu chọn biến thì tập biến được chọn là vector z có d biến đặc trưng với $d < n$, d biến là đại diện cho dữ liệu ban đầu với cơ sở dữ liệu mới $D_{new} = \{z_i, y_i\}_{i=1}^N$. Như vậy, đây là bước xây dựng quan hệ đầu vào và đầu ra là $y_{newi} = f_{new}(z_i)$. Trong bước này, các giải thuật học được sử dụng để trích xuất các tri thức về mối quan hệ giữa đầu vào và đầu ra, tri thức được lưu trữ trong cấu trúc của bộ phân lớp.

5.5 ĐÁNH GIÁ

Đây là bước đánh giá về khả năng tổng quát hóa của mô hình sau khi huấn luyện xong. Khả năng tổng quát hóa của mô hình thể hiện qua việc mô hình nhận dạng được các mẫu mới hay mẫu kiểm tra. Độ chính xác tổng thể nhận dạng và lỗi nhận dạng được tính toán theo công thức (5.2) và (5.3). Một yêu cầu quan trọng là bộ phân lớp phải có độ chính xác phân lớp cao. Có nhiều chỉ số để đánh giá hiệu suất phân lớp của bộ phân lớp dựa vào ma trận lẫn lộn (Confusion Matrix) [62]. Bảng 5.1 trình bày ma trận lẫn lộn kích thước 2x2 cho bộ phân lớp có 2 lớp đánh giá độ chính xác nhận dạng. Hai lớp ở đây gồm lớp dương (positive - P) và lớp âm (negative - N). Bảng này cho biết số mẫu chẩn đoán đúng và số mẫu chẩn đoán sai cho mỗi lớp trong 2 lớp.

Bảng 5.1 Ma trận lẫn lộn kích thước 2x2

		Lớp Đúng	
		P	N
Lớp được chẩn đoán	P	TP	FP
	N	FN	TN

Trong đó:

- TP (true positives -TP) là số lượng mẫu dương được bộ phân lớp phân lớp đúng vào lớp dương.
- FP (false positives -FP) là số lượng mẫu âm được bộ phân lớp phân lớp sai vào lớp dương.
- FN (false negatives) là số lượng mẫu âm được bộ phân lớp phân lớp sai vào lớp dương.

- TN (true negatives) là số lượng mẫu âm được bộ phân lớp phân lớp đúng vào lớp âm.

Bảng 5.1 trình bày các chỉ số đánh giá bộ phân lớp dựa vào ma trận lẫn lộn. Ngoài độ chính xác tổng thể như đã trình bày ở công thức (5.2), hai độ đo được dùng để đánh giá chất lượng của thuật toán phân lớp đó là độ chính xác (Precision-Pre) và độ hồi tưởng (Recall-Rec). F-Score là chỉ số đo lường độ chính xác kiểm tra, có xem xét đến cả độ chính xác và độ hồi tưởng. F-Score có thể được xem như là trọng số trung bình của độ chính xác (Pre) và độ hồi tưởng (Rec). F-Score đạt giá trị tốt nhất là 1 và xấu nhất là 0. Chỉ số độ đo F-Score cao thể hiện mô hình phân lớp đạt sự cân bằng giữa độ chính xác (Pre) và độ hồi tưởng (Rec), và khẳng định chất lượng của mô hình phân lớp cao. Các chỉ số độ đo đánh giá chất lượng bộ phân lớp dựa vào ma trận lẫn lộn được trình bày ở bảng 5.2

Bảng 5.2 Các chỉ số độ đo đánh giá bộ phân lớp dựa vào ma trận lẫn lộn

Độ chính xác tổng thể	$AccRate=(TP+TN)/(TP+FN+FP+TN)$
Độ chính xác	$Pre=TP/(TP+FP)$
Độ hồi tưởng	$Rec=TP/(TP+FN)$
F-Score	$F-Score=2 \times Pre \times Rec/(Pre+Rec)$

5.6 BIỂU DIỄN KẾT QUẢ

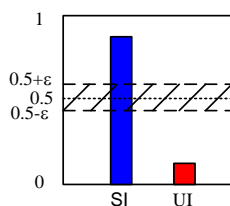
Đây là bước hiển thị kết quả, biên dịch kết quả, và biểu diễn kết quả giúp cho vấn đề trở nên trực quan, dễ hiểu và dễ sử dụng hơn.

Biểu diễn kết quả dạng thanh

Bộ phân lớp sau khi huấn luyện xong được xem là một hàm quyết định, khi nhận tín hiệu đầu vào thì hàm quyết định này tính toán giá trị đầu ra. Giá trị đầu ra là chỉ số ổn định tương đối. Chỉ số ổn định tương đối (relative stable index - SI) và

chỉ số tương đối không ổn định (relative unstable index - UI) đại diện cho trạng thái ổn định và không ổn định của hệ thống điện. Kết quả này có thể biểu diễn một cách trực quan trên một màn hình quan sát. Khi đó SI và UI thể hiện độ ổn định và không ổn định tương đối của hệ thống điện, và có thể biểu diễn như sơ đồ Hình 5.4. Trong đó, chỉ số ổn định tương đối được biểu diễn bằng thanh đứng màu xanh, chỉ số không ổn định tương đối được biểu diễn bằng thanh đứng màu đỏ. Trong bài toán phân lớp có hai lớp thì nhãn lớp của mẫu được gán nhãn số nhị phân. Giá trị đầu ra của các bộ phân lớp được tính toán như luật (4.17), trong đó nhãn ‘1’ là ổn định và nhãn ‘0’ là không ổn định. Giá trị chỉ số ổn định tương đối là từ 0 đến 1. Vì vậy, quan hệ giữa SI và UI theo phương trình (5.4).

$$SI+UI=1 \Leftrightarrow UI=1-SI \quad (5.4)$$

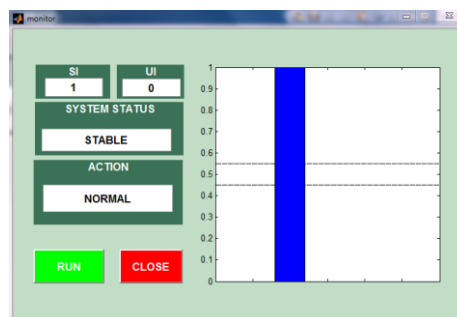


Hình 5.4 Sơ đồ thiết kế màn hình quan sát

Bộ phân lớp được huấn luyện chứa tri thức về trạng thái ổn định hệ thống điện. Hàm quyết định cho kết quả về chỉ số ổn định và cho biết biết điểm ổn định cách biên bao xa. Bài toán phân lớp là hai lớp và gán nhãn nhị phân với nhãn ‘1’ là ổn định và nhãn ‘0’ là không ổn định. Cho nên, biên phân định giữa hai lớp là tại giá trị 0,5. Tuy nhiên, trong thực tế kết quả phân lớp có những điểm nằm ở ngay ranh giới giữa lớp ổn định và lớp không ổn định hoặc vùng lân cận xung quanh giá trị 0,5. Đây là vùng nhận dạng không chắc chắn hay không xác định. Ở Hình 4.4, vùng không chắc chắn là vùng kẻ sọc, vùng này có độ rộng là $(2 \times \varepsilon)$ hay có giá trị từ $(0,5-\varepsilon)$ đến $(0,5+\varepsilon)$, ε có giá trị nhỏ và là giá trị nới lỏng để phân định cho vùng không chắc chắn.

Trong bài báo (III), tác giả đã trình bày một thiết kế màn hình hiển thị trạng thái ổn định động hệ thống điện, sơ đồ thiết kế cụ thể được trình bày ở Hình 5.5. Màn hình giám sát gồm có 5 khâu:

- *Khâu hiển thị chỉ số ổn định SI và UI:* khâu này cho biết chỉ số SI và UI được tính toán trong pha vận hành hay pha nhận dạng.
- *Khâu thể hiện trạng thái hệ thống điện:* Khâu này thể hiện kết quả trạng thái ổn định hệ thống điện được chẩn đoán. Kết quả được chẩn đoán gồm có ba chế độ là: chế độ ‘ổn định’, chế độ ‘không ổn định’, và chế độ ‘cảnh báo’. Chế độ ‘cảnh báo’ là vùng nhận dạng không chắc chắn.
- *Khâu thể hiện hành động hoạt động:* Khâu này chỉ thị cảnh báo điều khiển. Trong trường hợp kết quả nhận dạng là ‘ổn định’ thì chế độ cảnh báo đặt ở chế độ bình thường. Trong trường hợp kết quả nhận dạng là ‘mất ổn định’ thì chế độ cảnh báo đặt ở chế độ ‘điều khiển khẩn cấp’. Trong trường hợp kết quả nhận dạng là ‘cảnh báo’ thì khuyến nghị là xếp vào tình trạng ‘khẩn cấp’ để ra quyết định điều khiển nhằm tránh tình trạng mất ổn định đáng tiếc xảy ra.
- *Khâu cửa sổ màn hình:* Khâu này thể hiện chỉ số ổn định của hệ thống điện. Thanh màu xanh thể hiện cho chỉ số ổn định (SI) và thanh màu đỏ thể hiện cho chỉ số không ổn định (UI).
- *Nút RUN và nút CLOSE:* Màn hình thiết kế nút RUN để khi chọn thì chương trình được đưa vào chế độ hoạt động và chọn nút CLOSE để dừng chương trình.



Hình 5.5 Màn hình biểu diễn trạng thái ổn định hệ thống điện

Hình 5.5 là minh họa cho trường hợp kết quả chẩn đoán hệ thống điện là ổn định. Chỉ số ổn định tương đối $SI = 1$ và chỉ số không ổn định tương đối $UI = 0$.

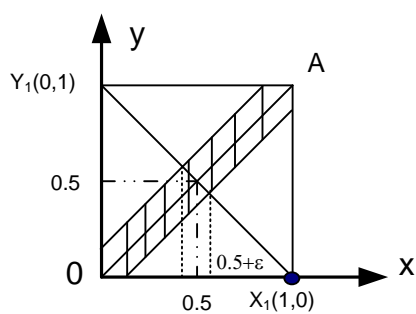
Biểu diễn kết quả dạng điểm trong không gian hai chiều

Với cách biểu diễn màn hình quan sát như vừa trình bày cho biết được chỉ số ổn định tương đối và trình diễn ở dạng thanh. Một ý tưởng được đề xuất để thiết kế màn hình giám sát biểu diễn điểm làm việc trạng thái ổn định động của hệ thống điện. Ý tưởng này chuyển kết quả đầu ra của bộ phân lớp thành điểm làm việc ổn định động của hệ thống điện và biểu diễn trên mặt phẳng không gian hai chiều. Để thực hiện điều này, đặt trục tung là trục không ổn định, y , và trục hoành là trục ổn định, x . Như vậy, phương trình (5.4) được viết lại thành (5.5) hay (5.6).

$$x+y=1 \tag{5.5}$$

$$\Leftrightarrow y= -x+1 \tag{5.6}$$

Phương trình (5.6) cho thấy quỹ đạo điểm làm việc của trạng thái ổn định động hệ thống điện là một đường thẳng. Hình 5.6 là sơ đồ thiết kế màn hình giám trạng thái ổn định động hệ thống điện trong không gian hai chiều. Trong Hình 5.6, tam giác OAX_1 là vùng ổn định, tam giá OAY_1 là vùng không ổn định, và vùng kẻ sọc OA là vùng không xác định. $X_1(1,0)$ là tọa độ của điểm hoàn toàn ổn định với chỉ số ổn định tương đối bằng 1 và chỉ số không ổn định bằng 0. Đây là điểm màu xanh minh họa trên Hình 5.6. $Y_1(0,1)$ là tọa độ của điểm hoàn toàn không ổn định với chỉ số không ổn định tương đối là 1 và chỉ số ổn định tương đối lúc này là 0. Vùng làm việc không chắn chắn của trục ổn định x có độ rộng là $(2 \times \varepsilon)$ hay có độ rộng từ $(0,5-\varepsilon)$ đến $(0,5+\varepsilon)$. Khi đó, vùng làm việc không chắc chắn, vùng kẻ sọc OA , có độ rộng $(2 \times \varepsilon \times \sqrt{2})$, ε có giá trị nhỏ và là giá trị nói lỏng để phân định cho vùng không chắn chắn.



Hình 5.6 Sơ đồ thiết kế màn hình giám sát trong không gian hai chiều

Trong trường hợp bộ phân lớp cho kết quả thuộc miền không chắn chắn thì công cụ tính toán sẽ xếp trường hợp này vào chế độ cảnh báo (warning mode) và theo ý kiến chuyên gia thì việc điều khiển khẩn cấp là cần thiết cho trường hợp này để tránh rủi ro rã lưới.

5.7 NGHIÊN CỨU ĐỀ XUẤT QUY TRÌNH GIẢM KHÔNG GIAN MẪU

Áp dụng kỹ thuật phân lớp trong nhận dạng ổn định quá độ hệ thống điện đối mặt với vấn đề khó khăn về kích thước dữ liệu rất lớn. Một nhu cầu cần thiết là chọn tập dữ liệu tin gọn, đặc trưng mang tính đại diện giúp làm giảm chi phí tính toán, giảm bộ nhớ máy tính, nâng cao độ chính xác nhận dạng.

Có nhiều giải thuật phân cụm dữ liệu, trong đó giải thuật trừ nhóm Chui [68] có nhược điểm là chịu ảnh hưởng rất mạnh vào giá trị bán kính lân cận r_a . Nếu r_a nhỏ thì số lượng nhóm lớn và ngược lại nếu bán kính r_a lớn thì số lượng nhóm quá nhỏ dẫn đến mất thông tin dữ liệu gốc nên không khả thi trong áp dụng đối với không gian dữ liệu có kích thước lớn. Giải thuật tối ưu bày đầy đàn [69] có khả năng tìm kiếm số nhóm tốt, tuy nhiên giải thuật này gặp gánh nặng tính toán nên cũng không phù hợp trong áp dụng cho thuật phân cụm dữ liệu với không gian dữ liệu có kích thước lớn.

Giải thuật phân cụm dữ liệu phổ biến và đơn giản là K-means (KM). Giải thuật K-means có ưu điểm là tính toán nhanh, và xử lý tập dữ liệu có kích thước lớn. Tuy nhiên, giải thuật KM có nhược điểm chính là rất nhạy cảm với tâm được khởi tạo ban đầu và có thể dẫn đến kết quả hội tụ địa phương. Có nhiều phương

pháp để cải thiện chất lượng giải thuật K-means. Những cải thiện này khác nhau ở cách chọn tâm ban đầu. Trong đó, giải thuật trừ nhóm Chui và giải thuật tối ưu bầy đàn có thể áp dụng để khởi tạo tâm cho KM, tuy nhiên qua nghiên cứu phân tích cũng như thực nghiệm thì các giải thuật này không phù hợp áp dụng để khởi tạo tâm cho K-means trong trường hợp dữ liệu có kích thước lớn do gánh nặng tính toán.

Phương pháp khoảng cách Max-Min đơn giản trong tính toán, có ưu điểm là số tâm khởi tạo ban đầu dựa vào khoảng cách Max-Min nên các tâm nhóm được khởi tạo gần đều trong không gian tập dữ liệu, cho nên không làm mất thông tin dữ liệu ban đầu. Do đó, giải thuật Hybrid K-means (HK) được đề xuất áp dụng để khắc phục hạn chế của giải thuật KM. Đây là giải thuật lai được kết hợp bởi phương pháp khoảng cách Max-Min (MMD) và giải thuật KM. Trong đó, MMD là nhằm khởi tạo tâm ban đầu cho KM. Giải thuật HK được đề nghị áp dụng nhằm thực thi giảm kích thước không gian mẫu. Giải thuật KM và HK được trình bày ở Bảng 5.3 và 5.4 tương ứng.

5.7.1 Giải thuật KM

K-means [62], [70] là giải thuật phân cụm mà đại diện các cụm bởi trung tâm của các phần tử trong cụm. Phương pháp này dựa trên độ đo khoảng cách của các đối tượng dữ liệu trong cụm. Tâm đại diện được xem như là trung tâm của cụm. Như vậy, K-means cần khởi tạo một tập k tâm ban đầu, và thông qua đó giải thuật lặp lại các bước gồm gán mỗi đối tượng tới cụm gần tâm, và tính toán lại tâm của mỗi cụm trên cơ sở gán mới cho các mẫu. Quá trình này dừng khi các tâm không đổi.

Thuật toán K-means sẽ tạo ra k cụm dữ liệu $\{C_1, C_2, \dots, C_k\}$ từ một tập dữ liệu chứa N mẫu trong không gian d chiều $x = \{x_{i1}, x_{i2}, \dots, x_{id}\}$, $i=[1,N]$, sao cho hàm tiêu chuẩn đạt giá trị cực tiểu. Tiêu chuẩn độ lệch bình phương, hay hàm mục tiêu D_E đạt giá trị cực tiểu như biểu thức (5.7).

$$D_E = \sum_{i=1}^k \sum_{j=1}^{n_i} \|x_j - m_i\|^2 \quad (5.7)$$

$$m_i = \frac{\sum_{l=1}^{n_i} x_l}{n_i} \quad (5.8)$$

Trong đó: m_i là trọng tâm của cụm C_i , x_i là vector phần tử trong nhóm i , n_i là số lượng các vector phần tử trong nhóm thứ i , $\|\cdot\|$ là khoảng cách Euclide.

Khoảng cách Euclide giữa hai vector x_i và y_i có kích thước l được tính theo biểu thức (5.9).

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^l (x_i - y_i)^2} = \|x_i - y_i\| \quad (5.9)$$

Tham số đầu vào của thuật toán là số cụm k , và tham số đầu ra của thuật toán là các tâm của các cụm dữ liệu.

Bảng 5.3 Giải thuật KM

Giải thuật KM

Đầu vào: $x = \{x_1, x_2, \dots, x_N\}$ tập mẫu ban đầu với số biến là l và số mẫu là N .

Cài đặt k tâm ban đầu.

Đầu ra: $C = \{c_1, c_2, \dots, c_k\} = C = \{c_p\}$ tập tâm đại diện có k tâm ban đầu, $p = [1, k]$.

1. Khởi tạo k tâm ban đầu ngẫu nhiên, $C = \{c_1, c_2, \dots, c_k\}$
 2. Tính toán khoảng cách Euclide từ các mẫu x_i đến các tâm c_k , $d(x, c)$.
 3. Đưa các x_i vào c_k thỏa điều kiện biểu thức (5.7).
 4. Xác định tâm mới dựa vào biểu thức (5.8).
 5. Lặp lại bước 2 đến bước 4, giải thuật dừng khi tâm không đổi.
-

5.7.2 Giải thuật Hybrid K-means (HK)

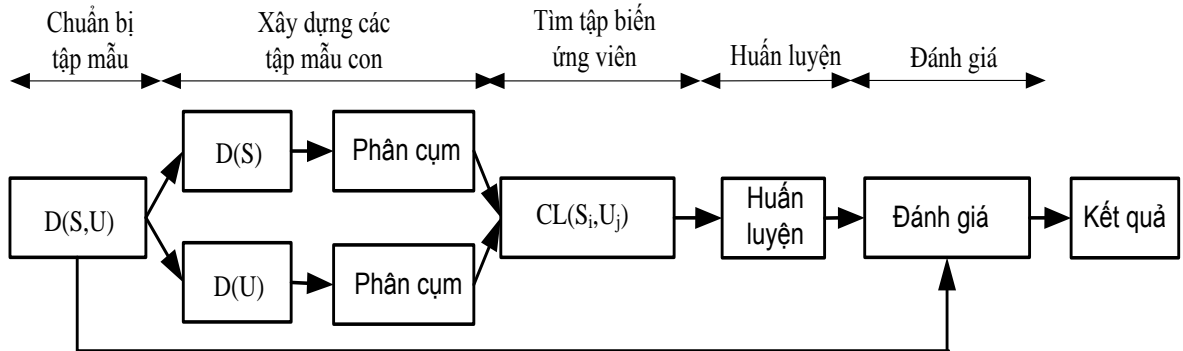
Giải thuật phân cụm KM có nhược điểm là rất nhạy cảm với tâm khởi tạo ban đầu và có thể bị rơi vào cực tiểu địa phương [62], [70]. Giải thuật HK là giải thuật cải tiến của giải thuật KM với chiến lược khởi tạo tâm ban đầu dựa vào phương pháp khoảng cách Max-Min (MMD) [71].

Bảng 5.4 Giải thuật HK

Giải thuật HK	
<i>Đầu vào:</i> $X\{x_1, x_2, \dots, x_N\}$ tập mẫu ban đầu với số biến là l và số mẫu là N .	
<i>Đầu ra:</i> $C\{c_1, c_2, \dots, c_k\} = C\{c_p\}$ tập tâm đại diện có k tâm ban đầu, $p=[1, k]$.	
<i>Khởi tạo tâm ban đầu dựa trên khoảng cách Max-Min</i>	
1. Tính khoảng cách giữa các điểm trong tập dữ liệu và chọn hai điểm có khoảng cách lớn nhất làm hai điểm tâm ban đầu.	
	$d_{ij} = \max \ x_i - x_j\ $ với $1 \leq i, j \leq N; i \neq j; p=2; M=d_{ij}$ (5.10)
	$\theta = M/k$ (5.11)
Chọn hai tâm ban đầu $C\{c_1, c_2\}; c_1=x_i; c_2=x_j$.	
	$X_{\text{now}} = X \setminus C$ (5.12)
Trong đó: x_i, x_j là hai mẫu i và j trong tập mẫu; $\ \cdot\ $ là khoảng cách Euclide. N_{now} là số mẫu của tập mẫu X_{now} tại thời điểm tính toán của giải thuật. θ là khoảng cách giữa hai tâm hay là ngưỡng.	
2. Tính toán khoảng cách nhỏ nhất giữa mỗi mẫu của X_{now} đến mỗi mẫu của $C\{c_1, c_2, \dots, c_p\}$.	
	$d_i = \min \ x_i - c_p\ $ với $i \in [1, N_{\text{now}}], p \in [1, k]$ (5.13)
3. Tìm tâm thứ $p=(p+1)$ với $p \leq k$	
	$dz_i = \max(d_i)$ với $i \in [1, N_{\text{now}}], p \in [1, k]$ (5.14)
Nếu $dz_i > \theta$ thì $C = C \cup x_i; X_{\text{now}} = X \setminus C$	
N_{now} là số mẫu của tập mẫu X_{now} tại thời điểm tính toán của giải thuật.	
4. Lặp lại bước 2 và bước 3, điều kiện dừng: $p > k$.	
<i>Giải thuật KM</i>	
5. Tính toán khoảng cách Euclide từ các mẫu x_i đến các tâm $c_k, d(x, c)$.	
6. Đưa các x_i vào c_k thoả điều kiện công thức (5.7).	
7. Xác định tâm mới dựa vào công thức (5.8).	
8. Lặp lại bước 5 đến bước 7, giải thuật dừng khi tâm không đổi.	

5.7.3 Đề xuất quy trình giảm không gian mẫu

Để thực hiện giảm không gian mẫu, Luận án đề nghị quy trình gồm 5 bước: Chuẩn bị bộ mẫu ban đầu, Xây dựng các tập mẫu con, Tìm các tập con ứng viên, Huấn luyện, và Đánh giá chọn kết quả. Quy trình giảm không gian mẫu được trình bày như Hình 5.7.



Hình 5.7 Quy trình giảm không gian dữ liệu

Bước 1. Chuẩn bị bộ mẫu ban đầu. Đây là bước chuẩn bị bộ mẫu ban đầu cho giai đoạn giảm không gian mẫu. Bộ mẫu đã qua khâu lựa chọn biến đặc trưng là tập mẫu gồm tập mẫu ổn định $D(S)$ và tập mẫu không ổn định $D(U)$. Toàn tập mẫu ký hiệu là $D(S,U)$.

Bước 2. Xây dựng các tập mẫu con. Bước này áp dụng giải thuật phân cụm để rút gọn lần lượt trên từng tập mẫu S và tập mẫu U . Tập S sau khi giảm mẫu là S_R , ký hiệu là $CL(S_R)$. Tập U sau khi giảm mẫu là U_R , ký hiệu là $CL(U_R)$. Ví dụ bộ mẫu ổn định ban đầu $D(S)$ có thể rút gọn lần lượt có số mẫu là $CL(S_R=\{S_1, S_2, \dots, S_i\})$, $i=[1, q]$; Bộ mẫu không ổn định ban đầu $D(U)$ có thể rút gọn lần lượt có số mẫu là $CL(U_R=\{U_1, U_2, \dots, U_j\})$, $j=[1, p]$.

Bước 3. Tìm kiếm các bộ mẫu con ứng viên. Bước này áp dụng phương pháp tìm kiếm các tập mẫu con được đặt tên là ‘*Phương Pháp Tìm Kiếm Tuần Tự-Sequential Search Method (SSM)*’ nhằm xác định số mẫu rút gọn phù hợp S_{Ri} và U_{Rj} , $CL(S_{Ri}, U_{Rj})$ nhờ đánh giá độ chính xác từ bộ phân lớp. Luận án sử dụng kết hợp tuần tự bộ mẫu $(S_R=\{S_1, S_2, \dots, S_i\})$, $i=[1, q]$ và bộ mẫu $(U_R=\{U_1, U_2, \dots, U_j\})$,

$j=[1,p]$) để tạo từng cặp bộ dữ liệu $CL(S_{R_i}, U_{R_j})$. Cụ thể các khả năng hình thành các bộ dữ liệu con là: $\{CL(S_1, U_1), CL(S_1, U_2), \dots, CL(S_1, U_p); CL(S_2, U_1), CL(S_2, U_2), \dots, CL(S_2, U_p); \dots; CL(S_q, U_1), \dots, CL(S_q, U_p)\}$.

Bước 4. Huấn luyện. Các bộ dữ liệu con đã kết hợp ở bước 3, $CL(S_{R_i}, U_{R_j})$, được sử dụng làm các tập học. Tập mẫu nguyên thủy, $D(S, U)$, được sử dụng làm tập kiểm tra. Trong bước này bộ phân lớp K-NN được đề nghị áp dụng nhờ vào tính đơn giản và tính toán nhanh của nó.

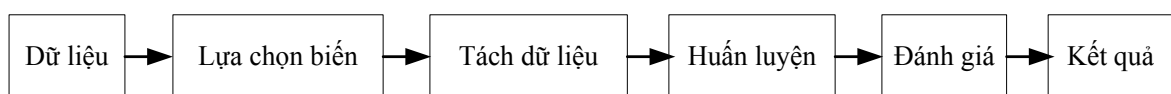
Bước 5. Đánh giá. Độ chính xác phân lớp tổng thể được tính theo công thức (5.2). Đây là bước đánh giá giúp chọn tập mẫu con sau khi rút gọn hay giảm mẫu. Tập mẫu con được chọn là cụm mẫu đạt độ chính xác phân lớp cao nhất trong vùng khảo sát.

5.8 NGHIÊN CỨU ĐỀ XUẤT MÔ HÌNH BỘ PHÂN LỚP CẢI TIẾN NÂNG CAO ĐỘ CHÍNH XÁC NHẬN DẠNG (ADVANCED PARALLEL CLASSIFIER MODEL-APCM)

5.8.1 Xây dựng mô hình phân lớp lai dựa vào năng lượng mẫu

Dữ liệu ổn định động hệ thống điện có kích thước lớn và phức tạp. Điều này dẫn đến bộ phân lớp đơn gặp khó khăn khi học biên giữa các lớp. Trong bài báo đã công bố (II), tác giả đã sử dụng năng lượng mẫu để tách bộ dữ liệu ban đầu thành các nhóm dữ liệu con. Mô hình phân lớp lai (Hybrid Classifier Model-HCM) bây giờ gồm nhiều bộ phân lớp con song song. Mỗi bộ phân lớp con chỉ học biên giữa các lớp của riêng các nhóm dữ liệu con cho nên việc học đơn giản hơn. Vì vậy độ chính xác phân lớp sẽ được cải thiện.

Để xây dựng mô hình HCM, bài báo (II) đã đề nghị quy trình gồm 4 bước: Lựa chọn biến, Tách dữ liệu, Huấn luyện, Đánh giá, Hình 5.8.



Hình 5.8 Quy trình xây dựng mô hình HCM theo năng lượng mẫu

- **Lựa chọn biến:** tác giả đã áp dụng phương pháp xếp hạng biến (FR) với chuẩn Fisher và chuẩn Divergence, và giải thuật SFFS với chuẩn SM như đã

trình bày ở chương 4. Kết quả chọn biến đã trình bày trong các công trình công bố và ở chương 6.

- **Tách dữ liệu:** Mẫu được tách ra dựa vào năng lượng của mẫu. Năng lượng của mẫu được trình bày theo công thức (5.15).

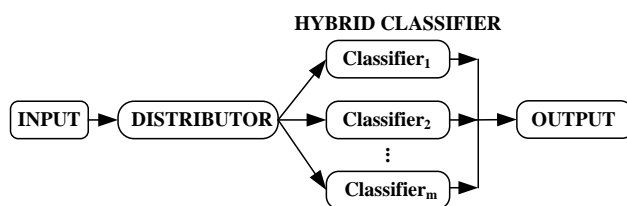
$$E = x_1^2 + x_2^2 + \dots + x_n^2 \quad (5.15)$$

Để tách mẫu, dữ liệu ban đầu được chia thành k tập con hay cụm con. Mỗi cụm con chứa mẫu ổn định và không ổn định, CL(S,U). Các mẫu có mức năng lượng gần nhau sẽ được gom cùng một cụm. Mỗi bộ phân lớp con được xây dựng, huấn luyện và kiểm tra, chỉ trên mỗi một tập dữ liệu con của nó. Như vậy, số bộ phân lớp song song bằng với số cụm con được tách ra. Ví dụ, nếu 2 cụm dữ liệu con được tách ra thì có 2 bộ phân lớp song song trong mô hình HCM được xây dựng.

- **Huấn luyện:** Đây là quá trình trích xuất tri thức hay học quan hệ vào ra.
- **Đánh giá:** Độ chính xác tổng thể phân lớp được tính theo công thức (5.2) như đã trình bày ở phần trên của chương này.

5.8.2 Mô hình HCM đề nghị

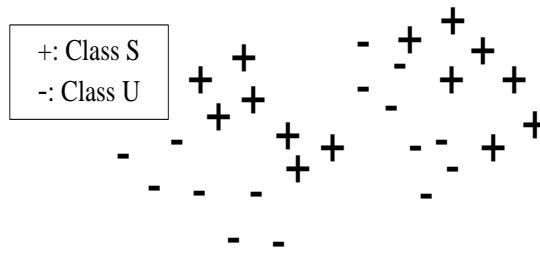
Mô hình HCM đề nghị như Hình 5.9 gồm: Đầu vào, Bộ phân phối, Bộ phân lớp lại, và Đầu ra.



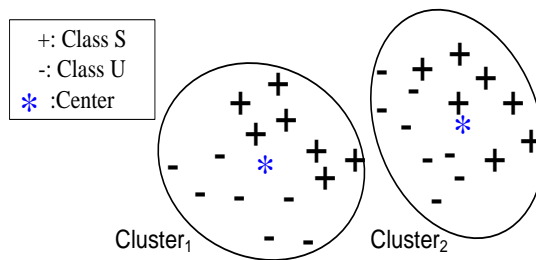
Hình 5.9 Mô hình HCM đề nghị

- **Đầu vào:** Đầu vào có vec-tơ với biến được lựa chọn ở bước chọn biến.
- **Bộ phân phối:** Một hình đã được thiết kế thì năng lượng của một mẫu của một cụm con nằm trong khoảng giá trị từ α_p đến α_q , $p=[1,1]$, $q=[2,h]$. Dựa theo luật (5.16) Bộ phân phối sẽ gửi mẫu kiểm tra hay mẫu mới đến đúng Bộ phân lớp con của nó.

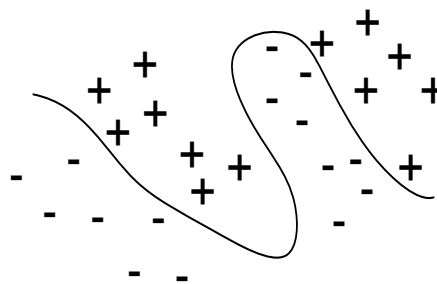
đơn kết nối song song nhau, được đề nghị tên gọi rút gọn cho tiện lợi là ‘Mạng nơ-ron song song cải tiến’ dựa trên kỹ thuật phân cụm dữ liệu. Kỹ thuật phân cụm dữ liệu được áp dụng nhằm tách dữ liệu thành các tập con. Giải thuật HK, đã trình bày ở trên, được áp dụng để tách dữ liệu huấn luyện thành các tập con. Như vậy, mỗi bộ phân lớp con chỉ học biên giữa các lớp của tập dữ liệu con riêng cho nên việc học trở nên đơn giản hơn, và nâng cao độ chính xác phân lớp. Minh họa cụ thể cho trường hợp này như Hình 5.10.



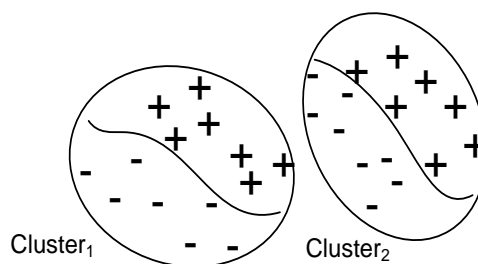
Hình 5.10a



Hình 5.10b



Hình 5.10c



Hình 5.10d

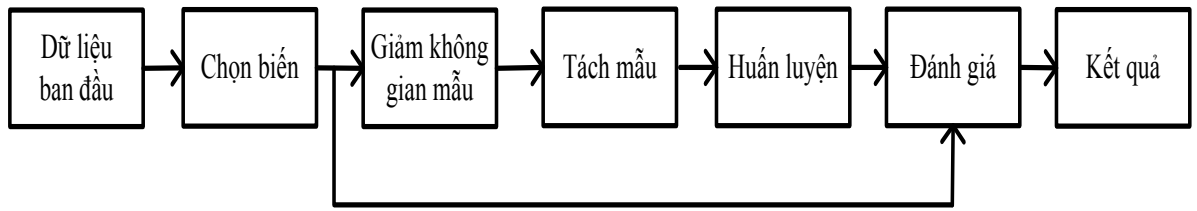
Hình 5.10 Minh họa ý tưởng xây dựng mô hình APCM

Trong đó:

- Hình 5.10a thể hiện tập dữ liệu ban đầu là toàn bộ tập dữ liệu, ký hiệu (-) cho lớp U và (+) cho lớp S.
- Hình 5.10b tách bộ dữ liệu ban đầu thành hai tập con, (*) ký hiệu cho tâm của các tập con.
- Hình 5.10c thể hiện ranh giới của hai lớp của toàn bộ tập dữ liệu, là đường cong gấp khúc cho nên bộ nhận dạng sẽ khó khăn trong việc học toàn bộ biên phân cách này.
- Hình 5.10d so sánh với hình 5.10c thì mỗi tập dữ liệu con có ranh giới với độ gấp khúc giảm đi rất nhiều cho nên mỗi bộ nhận dạng con sẽ học với biên phân cách đơn giản hơn, do vậy kết quả là độ chính xác nhận dạng sẽ nâng cao.

5.8.3.1 Quy trình xây dựng mô hình APCM

Các khâu xây dựng mô hình phân lớp cải tiến cụ thể gồm 5 bước như sau: **Chọn biến, Giảm không gian mẫu, Tách mẫu, Đánh giá**, Hình 5.11. Trong đó khâu chọn biến và giảm không gian biến đã được trình bày như phần trên.



Hình 5.11 Quy trình xây dựng mô hình mạng nơ-ron cải tiến

- **Khâu tách mẫu:**

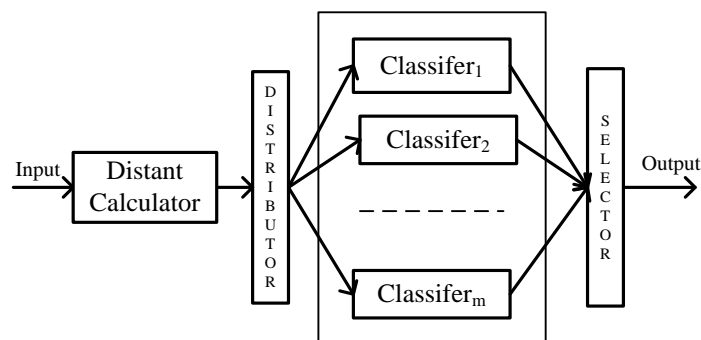
Khâu tách mẫu nhận dữ liệu từ khâu giảm không gian mẫu, áp dụng giải thuật HK để tách dữ liệu đầu vào thành các tập con. Mỗi tập con có một tâm đại diện sau khi thực hiện giải thuật tách nhóm. Mỗi Bộ phân lớp con học với dữ liệu của từng tập con riêng biệt, và như vậy mỗi Bộ phân lớp con có thông số khác nhau do được huấn luyện độc lập. Dữ liệu kiểm tra là dữ liệu ban đầu, dữ liệu chưa giảm không gian mẫu.

- **Khâu đánh giá:**

Đây là khâu đánh giá độ chính xác nhận dạng để chọn số bộ phân lớp. Số bộ phân lớp được chọn ứng với độ chính xác nhận dạng cao nhất trong vùng khảo sát.

5.8.3.2 Vận hành mô hình

Vấn đề cần giải quyết tiếp là mô hình sau khi được thiết kế xong thì làm việc với mẫu mới hay mẫu kiểm tra như thế nào. Luận án đề nghị mô hình APCM như Hình 5.12, gồm 4 bộ phận: **Bộ tính toán khoảng cách**, **Bộ phân phối**, **Bộ nhận dạng**, **Bộ chọn ngõ ra**.



Hình 5.12 Mô hình APCM đề nghị

▪ **Bộ tính toán khoảng cách (Distant Calculator):**

Trong quá trình thiết kế mô hình thì tâm của các cụm con, thực thi ở khâu tách mẫu, được lưu trữ là $CC_{Rk}(c_{Rk})$, $k=[1,m]$. Mẫu mới bất kỳ x_i đưa vào mô hình trước tiên phải qua bộ tính khoảng cách. Bộ tính toán khoảng cách tính khoảng cách Euclide từ x_i đến các tâm $CC_{Rk}(c_{Rk})$. Giá trị khoảng cách nhỏ nhất đến tâm của tập con nào thì mẫu x_i sẽ được gán nhãn thuộc về tập con hay bộ phân lớp con đó tương ứng, theo luật (5.17).

$$\left\{ \begin{array}{l} d(X_i, CC_{Rk}(c_{Rk})) \rightarrow \min \\ \text{Nếu } \{d(X_i, (CC_{Rk}(c_{Rk})))\} = \min \text{ thì } X_i \in \text{Cluster}_k \text{ and labelled } X_i \text{ is 'k', } k=[1,m] \end{array} \right. \quad (5.17)$$

▪ **Bộ phân phối mẫu (Distributor):**

Bộ phân phối có nhiệm vụ phân phối mẫu kiểm tra đến bộ phân lớp của nó theo luật (5.18).

$$\text{Nếu labelled } X_i \text{ is 'k', } k=[1,m] \text{ thì } X_i \in \text{Classifier}_m \quad (5.18)$$

▪ **Bộ nhận dạng (Classifier):**

Bộ nhận dạng hay Bộ phân lớp gồm các Bộ phân lớp con kết nối song song. Các bộ phân lớp con nhận tín hiệu từ Bộ phân phối mẫu, tính toán và gửi kết quả đến Bộ lựa chọn.

▪ **Bộ chọn ngõ ra (Selector):**

Nhãn lớp của mẫu được gán nhãn số nhị phân. Có hai trường hợp gán nhãn số nhị phân cho ngõ ra.

Trường hợp 1: Đối với bài toán phân lớp có hai lớp thì lớp ổn định được gán nhãn là 1 và lớp không ổn định được gán nhãn là 0. Đối với mạng nơ-ron tính toán phân lớp dựa vào ngưỡng làm biên quyết định và có sai số trong tính toán [26], [37]. Giá trị đầu ra của các bộ phân lớp được tính toán như luật (5.19).

$$\left\{ \begin{array}{l} \text{Nếu } y \leq 0.5 \text{ thì } y=0; \text{ Nếu } y > 0.5 \text{ thì } y=1 \\ \text{Nếu } y=1 \text{ thì 'Ổn định'; Nếu } y=0 \text{ thì 'Mất ổn định'} \end{array} \right. \quad (5.19)$$

Trường hợp 2: Trường hợp bài toán phân lớp có hai lớp có thể gán nhãn nhị phân là $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ cho lớp ổn định, và $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ cho lớp không ổn định. Sau đây trình bày cho

trường hợp tổng quát với bài toán phân lớp gồm có M lớp thì ma trận ngõ ra như sau:

$$\begin{array}{l} \text{Lớp 1} \\ \text{Lớp 2} \\ \vdots \\ \text{Lớp M} \end{array} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{l} F_1(x) \\ F_2(x) \\ \vdots \\ F_M(x) \end{array}$$

Ngõ ra tuân theo luật số lớn [72], kết quả nhận dạng được xếp vào lớp thứ k như sau:

$$F_k(x) > F_j(x) \quad \forall j \neq k \quad (5.20)$$

Trong đó $F_k(x)$ và $F_j(x)$ là kết quả tính toán từ hàm ánh xạ mạng nơ-ron.

$$F(x) = \begin{bmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_M(x) \end{bmatrix} \quad (5.21)$$

5.9 TÓM TẮT CHƯƠNG 5

Trong chương này tác giả trình bày các nghiên cứu các nội dung cụ thể như sau:

- Quy trình xây dựng tập mẫu.
 - Quy trình chọn biến.
 - Quy trình giảm không gian mẫu.
 - Quy trình xây dựng mô hình bộ phân lớp song song cải tiến.
- ✓ Về vấn đề xây dựng tập mẫu: quy trình lấy mẫu đã đề xuất cho bài toán nhận dạng ổn định động hệ thống điện, và được áp dụng vào xây dựng tập mẫu với kết quả áp dụng trình bày ở chương 6 và các công trình công bố (I-VI).
 - ✓ Vấn đề chọn biến: Nghiên cứu đề xuất quy trình chọn biến một cách có hệ thống cho bài toán nhận dạng ổn định động hệ thống điện. Kết quả nghiên cứu chọn biến đã được công bố trong các bài báo khoa học đã công bố và ở chương 6.
 - ✓ Vấn đề giảm không gian mẫu: Tác giả đã nghiên cứu và đề xuất quy trình giảm không gian mẫu cho bài toán nhận dạng ổn định động hệ thống điện. Thực hiện giảm không gian mẫu, kết quả là mỗi nhóm con chỉ cần có một tâm

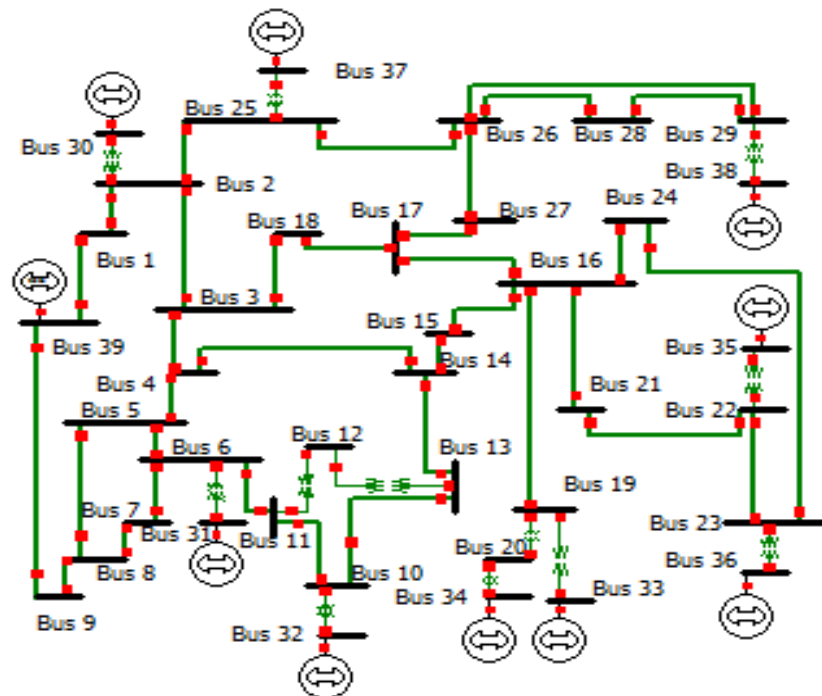
đại diện. Như vậy, việc cập nhật mẫu mới khi cần thiết thật ra là cập nhật lại tâm của nhóm đó tương ứng dựa vào khoảng cách hình học Euclide. Điều này có ý nghĩa rất quan trọng là làm cho mô hình không tăng dung lượng bộ nhớ lưu trữ mẫu mới, linh hoạt trong việc cập nhật làm giàu tri thức mới. Vấn đề giảm không gian dữ liệu được công bố trong bài báo (I).

- ✓ Trong khâu biểu diễn kết quả phân lớp tác giả đã giới thiệu một thiết kế công cụ trực quan hiển thị trạng thái ổn định động hệ thống điện. Trạng thái ổn định biểu diễn bằng thanh màu xanh thể hiện chỉ số ổn định tương đối hệ thống điện. Trạng thái không ổn định biểu diễn bằng thanh màu đỏ thể hiện chỉ số không ổn định tương đối hệ thống điện. Màn hình hiển thị có thể được sử dụng như công cụ huấn luyện cho các vận hành viên quan sát và đề ra ra chiến lược điều khiển trong các trường hợp sự cố khẩn cấp. Vấn đề này được công bố trong bài báo (III).
- ✓ Vấn đề xây dựng mô hình mạng nơ-ron song song cải tiến cho bài toán nhận dạng ổn định động hệ thống điện: Luận án đã nghiên cứu phát hiện và đề xuất được phương pháp để phân chia tập dữ liệu thành các tập con làm cơ sở xây dựng mô hình mạng nơ-ron song song cải tiến. Phương pháp tách tập dữ liệu thành các tập con dựa vào năng lượng mẫu được công bố ở bài báo (II). Phương pháp đề xuất giải quyết được mục tiêu quan trọng cho mô hình nhận dạng ổn định động HTĐ là nâng cao độ chính xác phân lớp. Ý tưởng tách tập mẫu thành các tập con gần nhau vô cùng nhóm dựa vào giải thuật phân cụm dữ liệu được trình bày ở chương 6.

Chương 6. ỨNG DỤNG MÔ HÌNH APCM CHẨN ĐOÁN NHANH ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN IEEE 39-BUS

6.1 GIỚI THIỆU SƠ ĐỒ HỆ THỐNG ĐIỆN CHUẨN IEEE 39-BUS

Hệ thống điện IEEE 10-máy 39-bus gồm có 39 bus, trong đó 10 bus là bus máy phát, 12 máy biến áp, 10 máy phát, 34 đường dây truyền tải và 19 tải. 10 máy phát được kết nối từ bus 30 đến bus 39, trong đó bus 39 được chọn là bus chuẩn, 9 bus được là bus PV, 29 bus còn lại được là bus PQ, có 2 cấp điện áp khác nhau là 345kV và 20kV. Sơ đồ hệ thống điện IEEE 39 bus được trình bày ở Hình 6.1.



Hình 6.1 Sơ đồ hệ thống điện IEEE 39 bus

Mô hình máy phát điện của hệ thống điện IEEE 10-máy 39-bus sử dụng mô hình GENPWTwoAxis. Điều khiển kích từ sử dụng mô hình IEEET1. Điều khiển tần số sử dụng mô hình TGVO1. Mỗi đường dây trên hệ thống có chiều dài khác nhau với điện trở, điện kháng và dung dẫn khác nhau. Điện trở, điện kháng và dung dẫn được cho trong đơn vị tương đối trên tổng chiều dài của các đường dây truyền tải dựa vào điện áp 345kV. Tải được mô tả bởi trở kháng cố định. Máy biến áp gồm có điện trở R_T và điện kháng X_T , các giá trị này là giá trị tương đương của

cuộn dây sơ cấp và thứ cấp của máy biến áp, và được cho trong đơn vị tương đối dựa trên điện áp 20kV cho cuộn sơ cấp và 345kV cho cuộn thứ cấp. Tất cả các giá trị trên hệ thống điện được cho trên cùng hệ thống cơ bản 100MVA.

Đây là sơ đồ đơn tuyến nổi tiếng được nhiều công trình giới thiệu và áp dụng [16], [26], [27], [31]. Tất cả mô hình các phần tử trong sơ đồ được lấy từ thư viện phần mềm PowerWorld. Thông số các phần tử của HTĐ IEEE 39-bus được trình bày ở phần phụ lục.

6.2 XÂY DỰNG TẬP MẪU VÀ TẬP BIẾN BAN ĐẦU

Thực hiện mô phỏng off-line để thu thập dữ liệu cho đánh giá ổn định động HTĐ với các mức tải từ (20,30,...,120)% tải cơ bản, thời gian cắt ngắn mạch cài đặt là 50ms [12]. Các sự cố được xem xét là ngắn mạch ba pha, một pha chạm đất, hai pha tại tất cả các thanh góp và dọc các đường dây truyền tải với mỗi khoảng cách 5% chiều dài đường dây. Ở mỗi mẫu tải, phần mềm Powerworld 18 tính toán mẫu công suất phát tương ứng nhờ công cụ tính toán phân bố công suất tối ưu (OPF).

Biến đầu vào ở chế độ động hay quá độ chứa thông tin chế độ xác lập và chế độ sự cố, kết hợp với ý kiến chuyên gia chọn các biến ban đầu gồm độ lệch điện áp tại các bus $\{\Delta V_{bus}\}$, độ lệch công suất tải $\{\Delta P_{load}\}$, độ lệch công suất trên các đường dây truyền tải $\{\Delta P_{flow}\}$.

Biến đầu ra được gán nhãn nhị phân, $\{1\}$ là lớp ổn định và $\{0\}$ là lớp không ổn định. Biến đầu vào và biến đầu ra là x và y . Tổng số biến đầu vào là $104(39+19+46)$ biến. Trong đó có 39 biến độ lệch điện áp tại các bus, 19 biến độ lệch công suất tác dụng của tải, và 46 biến độ lệch công suất tác dụng trên nhánh. Số biến đầu ra 1 biến đầu ra $y\{1,0\}$.

Quan sát kết quả mô phỏng, trạng thái của hệ thống điện là ổn định khi độ lệch góc rô-to của bất kỳ hai máy phát không quá 180° , không ổn định khi độ lệch góc rô-to của bất kỳ hai máy phát vượt quá 180° . Từ kết quả mô phỏng, có 3551 mẫu với 2649 mẫu ổn định (S) và 902 mẫu không ổn định (U), ký hiệu tập dữ là $D(S,U)=D(2649,902)$. Tập dữ liệu được chuẩn hoá trước khi đưa vào huấn luyện theo biểu thức (5.1).

6.3 CHỌN BIẾN ĐẶC TRƯNG

Sau khi thực hiện xây dựng bộ mẫu và biến xong, áp dụng quy trình chọn biến đặc trưng F&W để thực thi chọn biến như Hình 5.3.

Bước 1. Tập biến ban đầu:

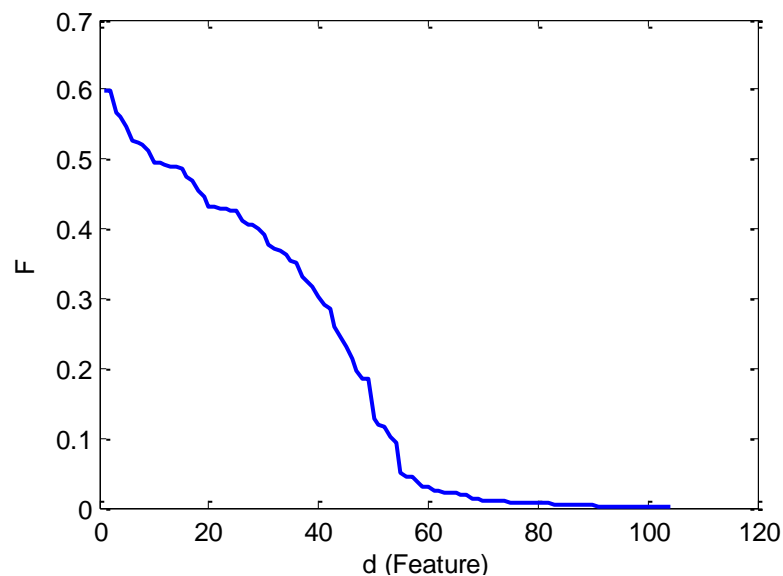
Biến đầu vào và biến đầu ra là x và y . Tổng số biến đầu vào là 104 biến. Số biến đầu ra là 1 biến, $y \in \{1,0\}$. Tập mẫu $D(S,U)=D(2649,902)$.

Bước 2. Tìm kiếm tập biến ứng viên:

- **Giải thuật áp dụng chọn biến:**

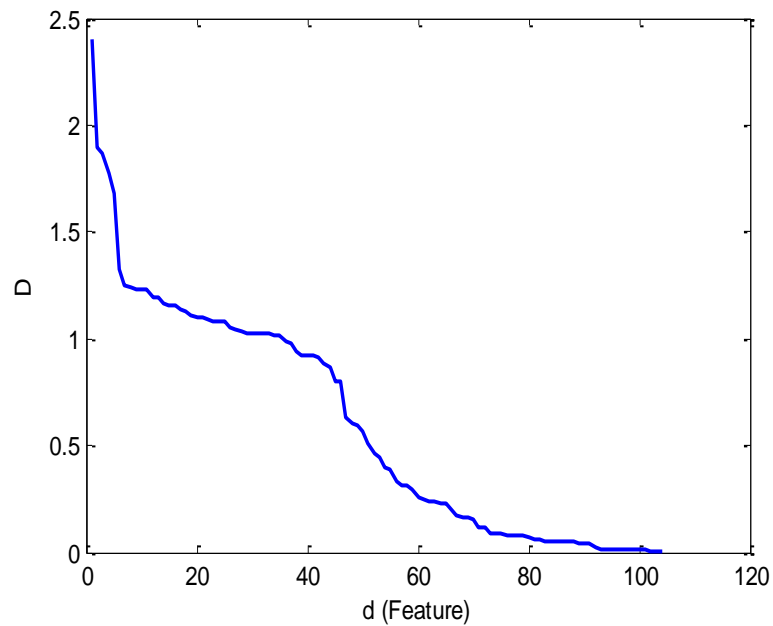
Để thực hiện chọn biến hai giải thuật được giới thiệu áp dụng là giải thuật xếp FR và giải thuật chọn biến SFFS theo quy trình chọn biến Hình 5.3.

- **Tiêu chuẩn áp dụng chọn biến:**

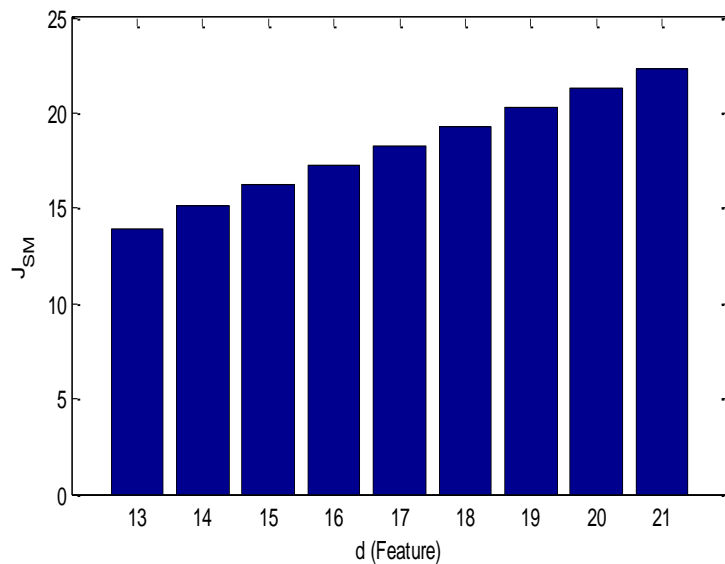


Hình 6.2 Xếp hạng biến theo chuẩn Fisher

Giải thuật xếp hạng biến (FR) với chuẩn chọn biến Fisher và Divergence theo biểu thức (4.1) và (4.2) tương ứng. Giải thuật Sequential Forward Floating Selection (SFFS) với chuẩn chọn biến SM theo biểu thức (4.9). Kết quả tính toán giá trị khoảng cách (F) và xếp hạng biến theo chuẩn Fisher được trình bày ở Hình 6.2. Kết quả tính toán giá trị khoảng cách (D) và xếp hạng biến theo chuẩn Divergence được trình bày ở Hình 6.3. Kết quả tính toán giá trị khoảng cách (J_{SM}) theo chuẩn SM bằng giải thuật SFFS trình bày ở Hình 6.4.



Hình 6.3 Xếp hạng biến theo chuẩn Divergence



Hình 6.4 Giá trị khoảng cách tính toán biến bằng giải thuật SFFS theo chuẩn SM

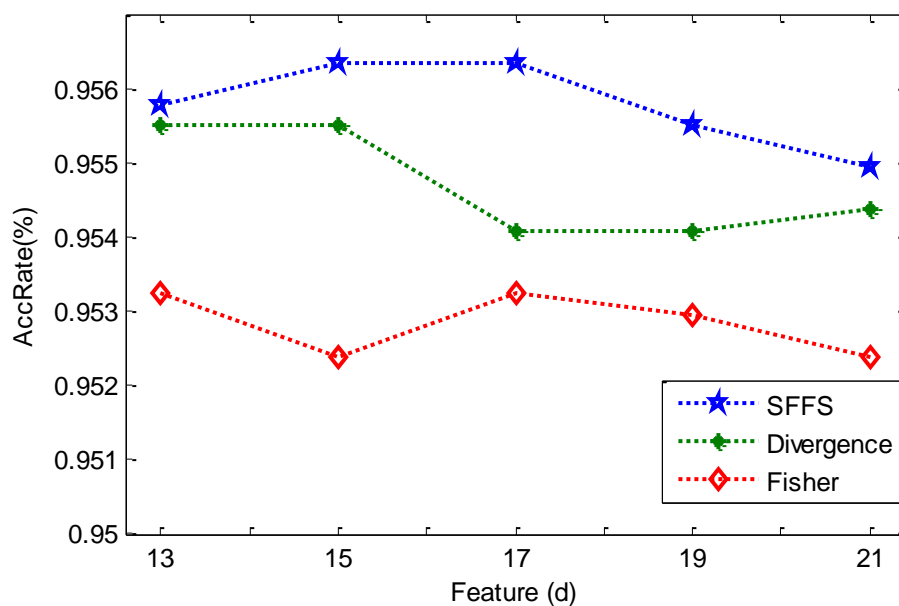
▪ **Bước 3. Huấn luyện:**

Bộ phân lớp K-NN (1-NNC, K=1) được đề nghị áp dụng để học tri thức từ cơ sở dữ liệu ổn định động HTĐ. Bộ phân lớp 1-NNC còn có vai trò quan trọng là đánh giá độ chính xác kiểm tra để chọn tập biến nhờ vào đặc điểm tính toán nhanh

và đơn giản của nó. Tập mẫu học và tập mẫu kiểm tra được xây dựng từ tập mẫu ban đầu, $D(S,U)=D(2649,902)$. Để đánh giá khách quan và thể hiện khả năng tổng quát hóa của mô hình nghiên cứu, tập mẫu ban đầu được chia ngẫu nhiên thành 10 tập con (kfold=10). Mỗi tập con học có 2834 mẫu S và 812 mẫu U, mỗi tập kiểm tra có 265 mẫu S và 90 mẫu U. Kết quả đánh giá độ chính xác được tính trung bình của 10 lần thực hiện.

• **Bước 4. Đánh giá:**

Kết quả tính toán đánh giá độ chính xác kiểm tra cho giải thuật chọn biến SFFS theo chuẩn SM, và giải thuật xếp hạng FR theo chuẩn Fisher và Divergence được trình bày ở đồ thị Hình 6.5. Độ chính xác phân lớp kiểm tra được tính theo biểu thức (5.2).



Hình 6.5 Đánh giá chọn tập biến

▪ **Bước 5. Kết quả chọn tập biến đặc trưng:**

Từ kết quả chọn biến Hình 6.5 độ chính xác phân lớp của tập biến được chọn bởi giải thuật SFFS theo chuẩn SM cao hơn giải thuật FR theo tiêu chuẩn Fisher và Divergence. Như vậy, Có 15 biến được chọn từ giải thuật SFFS. Ngoài ra, tập biến ban đầu, 104 biến, cũng được đánh giá độ chính xác kiểm tra. Bảng 6.1 trình

bày kết quả đánh giá kiểm tra tại 15 biến được chọn và 104 biến ban đầu, 1-NNC, kfold=10. Tập biến được chọn, 15 biến, được chọn trình bày ở Bảng 6.2.

Bảng 6.1 Độ chính xác kiểm tra phân lớp 1-NN, d=15 và d=104, kfold=10

Bộ phân lớp	d(feature)	AccRate(%)
1-NN	104	95,60
	15	95,63

Bảng 6.2 Các biến được chọn

STT	Biến		Ký hiệu
1	$\text{del}V_{\text{Bus}2}$	Độ lệch điện áp Bus 2	X1
2	$\text{del}V_{\text{Bus}20}$	Độ lệch điện áp Bus 20	X2
3	$\text{del}V_{\text{Bus}25}$	Độ lệch điện áp Bus 25	X3
4	$\text{del}V_{\text{Bus}30}$	Độ lệch điện áp Bus 30	X4
5	$\text{del}V_{\text{Bus}37}$	Độ lệch điện áp Bus 37	X5
6	$\text{del}P_{\text{Load}15}$	Độ lệch công suất tác dụng của tải 15	X6
7	$\text{del}P_{\text{Load}23}$	Độ lệch công suất tác dụng của tải 23	X7
8	$\text{del}P_{\text{Load}31}$	Độ lệch công suất tác dụng của tải 31	X8
9	$\text{del}P_{\text{Flow}2-30}$	Độ lệch công suất tác dụng của nhánh 2-30	X9
10	$\text{del}P_{\text{Flow}10-32}$	Độ lệch công suất tác dụng của tải 10-32	X10

11	$delP_{Flow16-21}$	Độ lệch công suất tác dụng của tải 16-21	X11
12	$delP_{Flow19-33}$	Độ lệch công suất tác dụng của tải 19-33	X12
13	$delP_{Flow20-34}$	Độ lệch công suất tác dụng của tải 20-34	X13
14	$delP_{Flow26-29}$	Độ lệch công suất tác dụng của tải 26-29	X14
15	$delP_{Flow28-29}$	Độ lệch công suất tác dụng của tải 28-29	X15

6.4 GIẢM KHÔNG GIAN MẪU

Áp dụng kỹ thuật phân lớp trong nhận dạng ổn định quá độ hệ thống điện đối mặt với vấn đề khó khăn về kích thước dữ liệu lớn. Tuy nhiên, thu thập và lưu trữ tất cả mọi mẫu là không khả thi. Vấn đề cần thiết đặt ra là chọn những mẫu đại diện cho không gian mẫu nhằm giảm gánh nặng thu thập dữ liệu, cũng như tiết kiệm tài nguyên bộ nhớ máy tính, giảm chi phí.

Sau khi thực hiện xây dựng bộ mẫu và qua khâu chọn biến thì bước kế tiếp là giảm không gian mẫu.

Bước 1. Chuẩn bị tập mẫu ban đầu:

Tập mẫu bây giờ gồm 15 biến như Bảng 6.2, tập mẫu kiểm tra là tập mẫu nguyên thủy $D(S,U)=D(2649,902)$. Một điều chú ý ở đây là tập mẫu kiểm tra là tập mẫu gốc, không tham gia vào quá trình huấn luyện, việc này để kiểm tra mô hình có khả năng nhận biết được mẫu mới hay không. Tập mẫu huấn luyện là tập mẫu thứ cấp được tìm kiếm ở các bước tiếp sau.

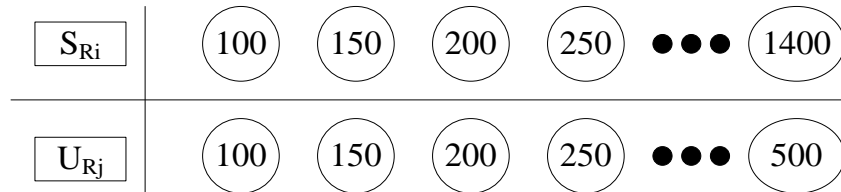
Bước 2. Xây dựng các tập mẫu con:

▪ **Giải thuật áp dụng:**

Hai giải thuật được áp dụng để rút gọn dữ liệu là giải thuật K-means (KM) và giải thuật K-means lai (HK) và so sánh kết quả.

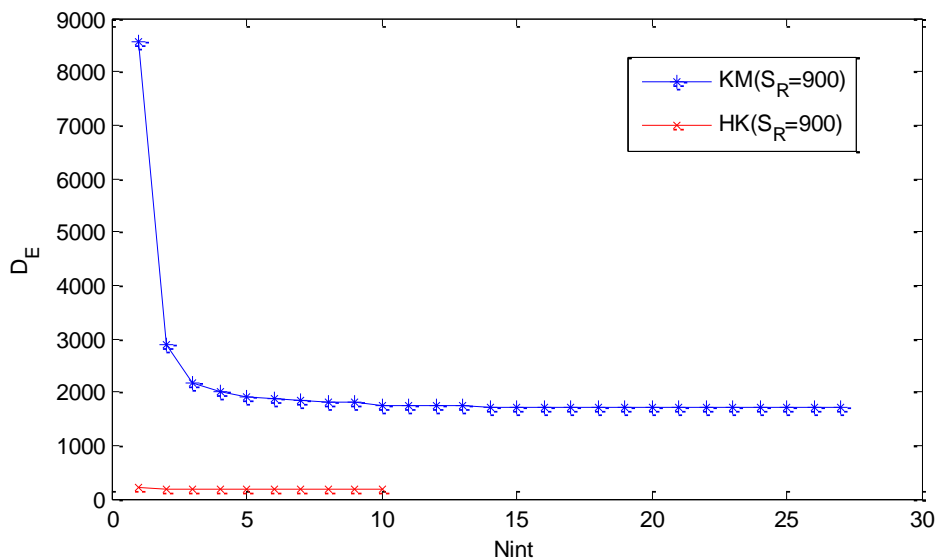
▪ **Thực hiện:**

Thực thi rút gọn dữ liệu trên tập mẫu S và tập mẫu U để hình thành các tập mẫu con. Số mẫu lớp S giảm từ 2649 mẫu xuống với dãy từ 100 đến 1400 mẫu, $CL(S_{R_i}=\{100, 150, \dots, 1400\}, i=[1,27])$. Số mẫu lớp U giảm từ 902 mẫu xuống dãy từ 100 đến 500 mẫu, $CL(U_{R_j}=\{100, 150, \dots, 500\}, j=[1,9])$, Hình 6.6.

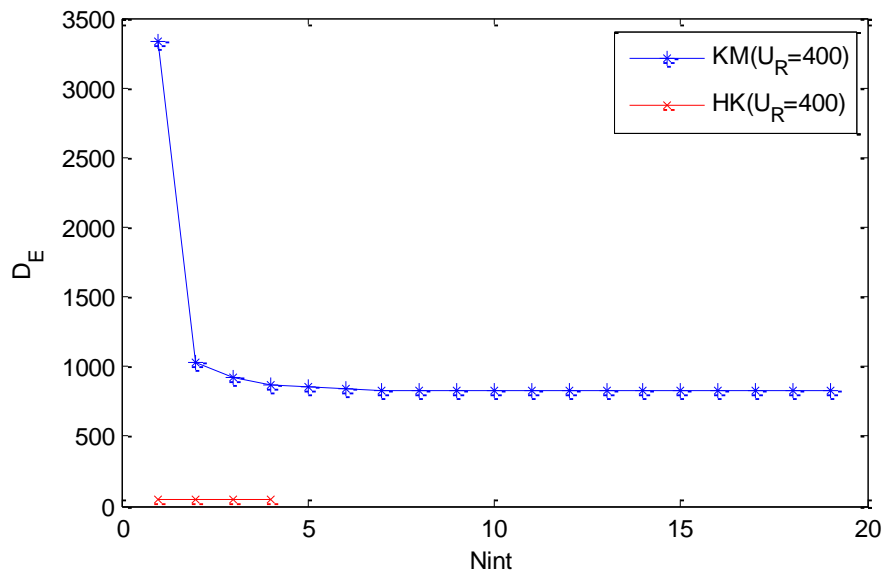


Hình 6.6 Các nhóm mẫu của tập S và tập U được rút gọn

Hình 6.7 và 6.8 trình bày đặc tính hội tụ của giải thuật KM và HK khi thực thi phân cụm giảm không gian mẫu. Hình 6.7 thể hiện quan hệ giữa số lần lặp (N_{int}) và giá trị mục tiêu (D_E) cho trường hợp giảm số mẫu D(S=2649) xuống $CL(S_R=900)$. Hình 6.8 thể hiện quan hệ giữa số lần lặp (N_{int}) và giá trị mục tiêu cho trường hợp giảm mẫu D(U=902) xuống $CL(U_R=400)$.

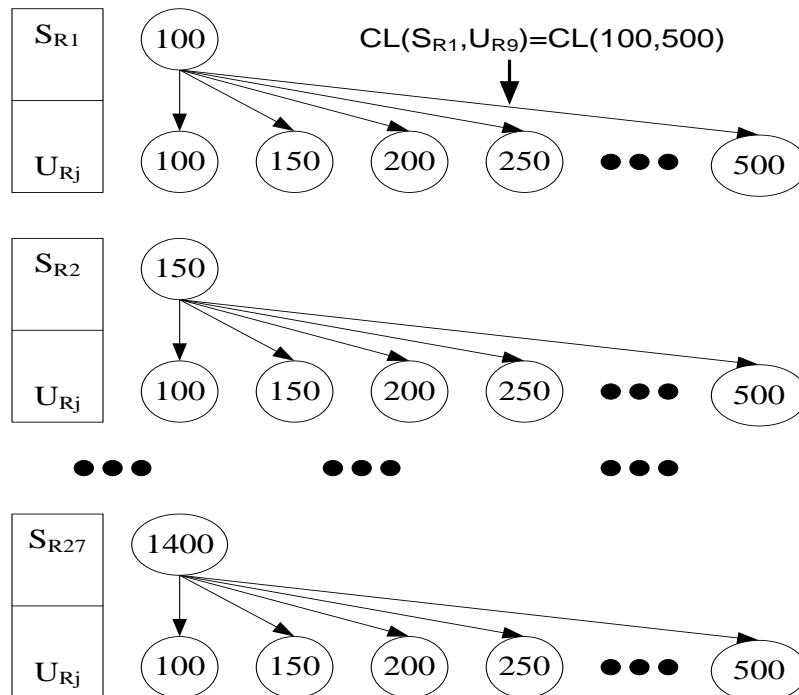


Hình 6.7 Đặc tuyến hội tụ của giải thuật HK và KM tại S_R=900.



Hình 6.8 Đặc tuyến hội tụ của giải thuật HK và KM tại $U_R=400$.

Bước 3. Tìm kiếm các tập con ứng viên.



Hình 6.9 Các khả năng kết hợp hình thành tập mẫu rút gọn

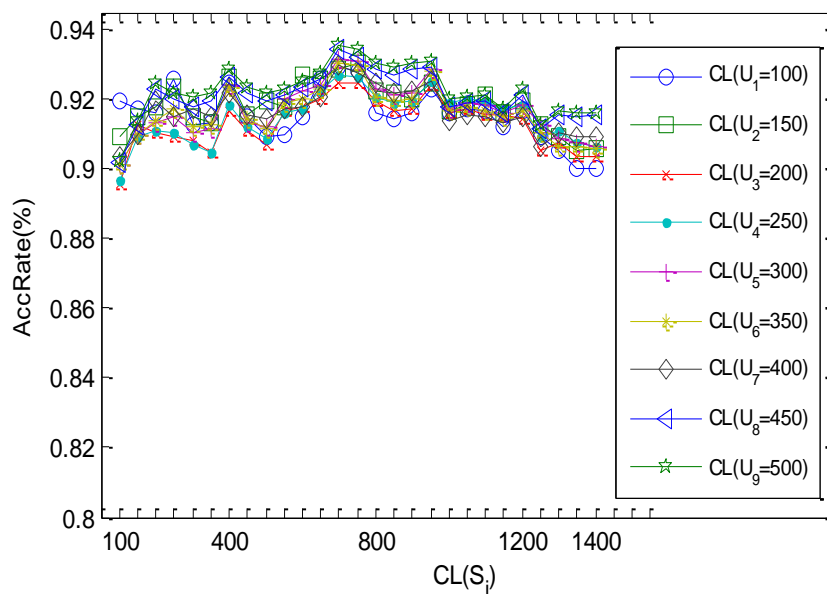
Áp dụng phương pháp tìm kiếm tuần tự SSM (trình bày ở mục 5.7.3 chương 5) để xây dựng các tập mẫu con hay các tập mẫu đã rút gọn của tập S và tập U kết hợp, và kết quả hình thành các tập con $CL\{(S_{Ri}, U_{Rj}), i=[1,27], j=[1,9]\}$. Các tập

mẫu con này là tập mẫu thứ cấp, đóng vai trò tập mẫu huấn luyện. Tập kiểm tra là tập mẫu nguyên thủy $D(2649,902)$. Hình 6.9 trình bày và thể hiện các khả năng kết hợp hình thành tập mẫu rút gọn $CL\{(S_{Ri}, U_{Rj})\}$.

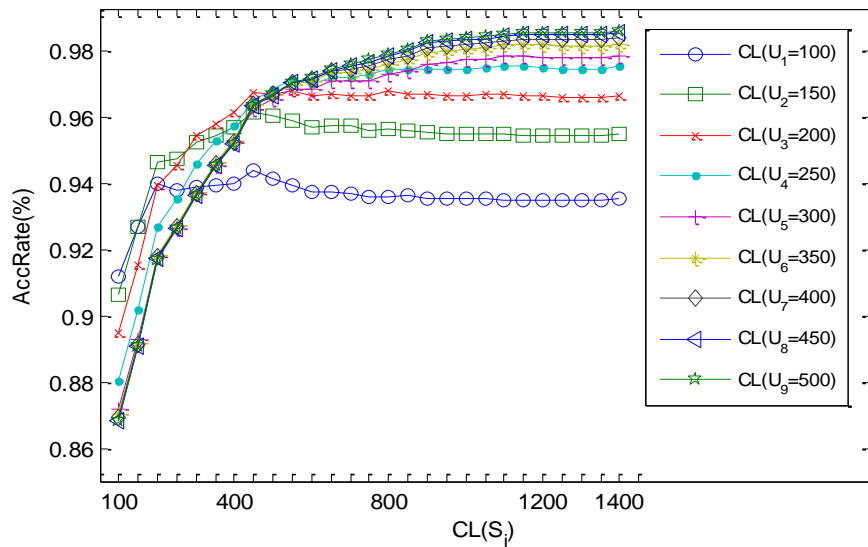
Bước 4. Huấn luyện.

❖ Bộ phân lớp 1-NN (1-NNC)

Các bộ dữ liệu con đã kết hợp ở bước 3, $CL(S_{Ri}, U_{Rj})$, được sử dụng làm các tập học. Tập mẫu nguyên thủy, $D(S,U)$, được sử dụng làm tập kiểm tra. Hình 6.10 và Hình 6.11 tương ứng cho kết quả đánh giá độ chính xác kiểm tra với số mẫu rút gọn bởi giải thuật KM và HK. Từ Hình 6.10 và Hình 6.11 kết quả đánh giá độ chính xác kiểm tra chọn tập mẫu con có độ chính xác cao trong vùng khảo sát, kết quả được trình bày ở Bảng 6.3.



Hình 6.10 Đánh giá độ chính xác phân lớp, rút dữ liệu với giải thuật KM, 1-NNC



Hình 6.11 Đánh giá độ chính xác phân lớp, rút dữ liệu với giải thuật HK, 1-NNC

Bảng 6.3 Độ chính xác kiểm tra phân lớp 1-NN với không gian mẫu giảm

Giải thuật	Cụm dữ liệu	AccRate(%)
KM	CL(700,500)	93,6
HK	CL(900,400)	98,0

Bảng 6.3, Với giải thuật KM thì có độ chính xác phân lớp cao nhất trong vùng khảo sát đạt 93,6% tại cụm dữ liệu $CL(S,U)=CL(700,500)$, với HK có độ chính xác phân lớp đạt 98% tại cụm dữ liệu $CL(S,U)=CL(900,400)$. Với kết quả này, tiến hành huấn luyện với bộ phân lớp MLPC, GRNNC, và SVMC. Đây là các bộ phân lớp được các công trình áp dụng trong thời gian gần đây [27], [29], [31], [37], [51], [73]. Các công cụ bộ phân lớp được sự hỗ trợ của phần mềm Matlab.

❖ **Bộ phân lớp mạng nơ-ron MLP, GRNN và SVM:**

- **Bộ phân lớp MLP:**

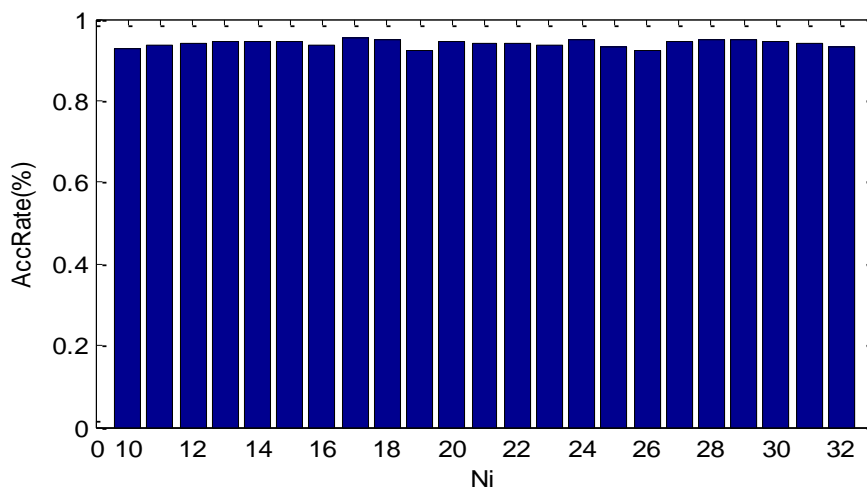
Mạng nơ-ron perceptron (MLP) gồm 3 lớp là lớp vào, lớp ẩn và lớp ra. Trong đó, Biến đầu vào là 15 biến, lớp nơ-ron ẩn dùng hàm kích hoạt sigmoid và lớp ra dùng hàm tuyến tính purelin, số biến đầu ra là 1.

Thông số được cài đặt cho toàn bộ các lần huấn luyện:

- Sai số đích = $1e-5$.
- Số chu trình huấn luyện = $1e3$.
- Giải thuật học và cập nhật trọng số Levenberg-Marquardt.
- Các thông số khác của mạng nơ-ron chọn theo giá trị mặc định.

Giải thuật học Levenberg-Marquardt được áp dụng vì đây là giải thuật được nhiều công trình áp dụng và khuyến dùng nhờ tính toán nhanh và cho độ chính xác cao [27], [59], [73]. Số nơ-ron ẩn được chọn theo thực nghiệm, trong [74] đề nghị số nơ-ron ẩn có giá trị từ 2/3 đến 2 lần tổng số biến đầu vào và đầu ra.

Hình 6.12 trình bày kết quả thực nghiệm tìm số nơ-ron ẩn, thể hiện quan hệ độ chính xác phân lớp với số nơ-ron ẩn (N_i) có giá trị từ 10 nơ-ron đến 32 nơ-ron.



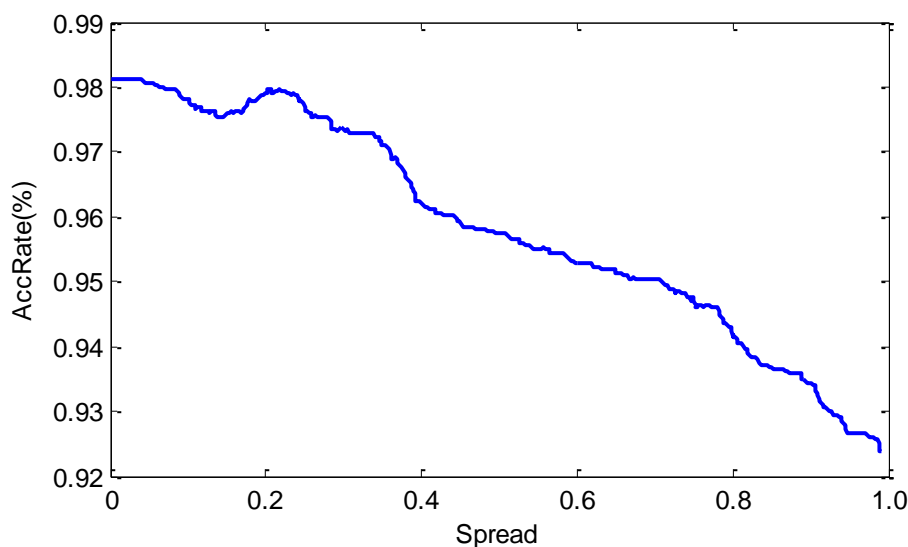
Hình 6.12 Thực nghiệm tìm số nơ-ron ẩn

- **Bộ phân lớp GRNN:**

GRNN là biến thể của RBFN (Radial Basis Function Network), là công cụ mạnh trong ứng dụng cho bài toán nhận dạng, thời gian huấn luyện nhanh. GRNN cơ bản gồm lớp ngõ vào, lớp ẩn hàm xuyên tâm, lớp tổng và lớp ngõ ra tuyến tính.

GRNN có cấu trúc giống như RBFN ngoại trừ lớp tổng (summation layer). Thông số độ rộng spread có giá trị tốt nhất được xác định qua thực nghiệm [31].

Hình 6.13 trình bày kết quả thực nghiệm tìm hệ số Spread, thể hiện quan hệ độ chính xác phân lớp với hệ số Spread có giá trị từ 0.01 đến 1.



Hình 6.13 Thực nghiệm tìm hệ số Spread

- **Bộ phân lớp SVM:**

SVM là giải thuật nổi tiếng trong máy học cho bài toán phân lớp. Cho tập dữ liệu huấn luyện $T = \{x_i, y_i\}_{i=1}^N$. Trong đó, x_i là vectơ dữ liệu đầu vào kích thước n biến và N mẫu, và $y_i \in \{+1, -1\}$ là nhãn lớp của mẫu x_i . SVM yêu cầu tìm lời giải tối ưu bài toán (3.26) như đã trình bày ở chương 3. Hàm nhân $\Phi(\cdot)$ là hàm ánh xạ dữ liệu trong không gian đầu vào vào không gian thuộc tính mà ở đó dữ liệu tách biệt tuyến tính. Hàm nhân với hàm bán kính cơ sở thường được khuyến dùng cho các áp dụng xây dựng mô hình bộ phân lớp SVM. Đặt K_s (KernelScale) là hệ số của hàm cơ sở bán kính. Thuật toán SVM được hỗ trợ trong phần mềm Matlab, ở đó việc xây dựng mô hình bộ phân lớp SVM qua tìm kiếm tập thông số (C, K_s) tốt nhất sao cho bộ phân lớp đạt kết quả nhận dạng trên tập kiểm tra cao nhất. Các cặp thông số (C, K_s) thường được tìm kiếm qua thực nghiệm thử sai, và cặp thông số được chọn là cặp cho kết quả kiểm tra với độ chính xác cao nhất trong vùng tìm kiếm. Việc tìm kiếm kết hợp tất cả giá trị các cặp thông số (C, K_s) để cho kết quả

tối ưu là công việc với khối lượng tính toán rất lớn. Cho nên, trong thực tế nghiên cứu thì phương pháp kết hợp tuần tự các giá trị được đề xuất cho $C=\{\exp(-5),\exp(-4),\dots,\exp(5)\}$ và $Ks=\{\exp(-5),\exp(-4),\dots,\exp(5)\}$ để xây dựng mô hình cho bộ phân lớp SVM.

Kết quả độ chính xác phân lớp được trình bày ở Bảng 6.4 với tập dữ liệu huấn luyện là $CL(S,U)=CL(900,400)$. Độ chính xác (Pre), độ hồi tưởng (Rec) và hệ số F-score được trình bày ở Bảng 6.5.

Bảng 6.4 Độ chính xác phân lớp với MLPC, GRNNC và SVMC

MLPC			GRNNC			SVMC		
N_i	AccRate(%)		Sp	AccRate(%)		(C,Ks)	AccRate(%)	
	Huấn luyện	Kiểm tra		Huấn luyện	Kiểm tra		Huấn luyện	Kiểm tra
17	92.7	95.2	0.051	100	98.1	(0.4493,3.004)	96.5	98.0

Bảng 6.5 Độ chính xác, độ hồi tưởng và hệ số F-Score

	MLPC	GRNNC	SVMC
Pre	0.9702	0.9849	0.9947
Rec	0.9658	0.9898	0.9788
F-Score	0.9679	0.9873	0.9867

6.5 ÁP DỤNG MÔ HÌNH CẢI TIẾN

Bước 1. Chọn biến và mẫu ban đầu

Tập mẫu được chọn 15 biến như Bảng 6.2, tập mẫu ban đầu là $D(S,U)=D(2649,902)$.

Bước 2. Giảm không gian mẫu

Tập mẫu đã rút gọn tìm được là $CL(S,U)=CL(900,400)$ bằng giải thuật HK như ở Bảng 6.3.

Bước 3. Tách mẫu

Áp dụng giải thuật HK tách nhóm dữ liệu $CL(S,U)=CL(900,400)$ ra số nhóm từ 2 nhóm đến 10 nhóm. Số mẫu của từng tập con tách ra được trình bày ở Bảng 6.6. Các nhóm con này đóng vai trò làm dữ liệu học. Dữ liệu kiểm tra là bộ mẫu nguyên thủy $D(S,U)=D(2649,902)$.

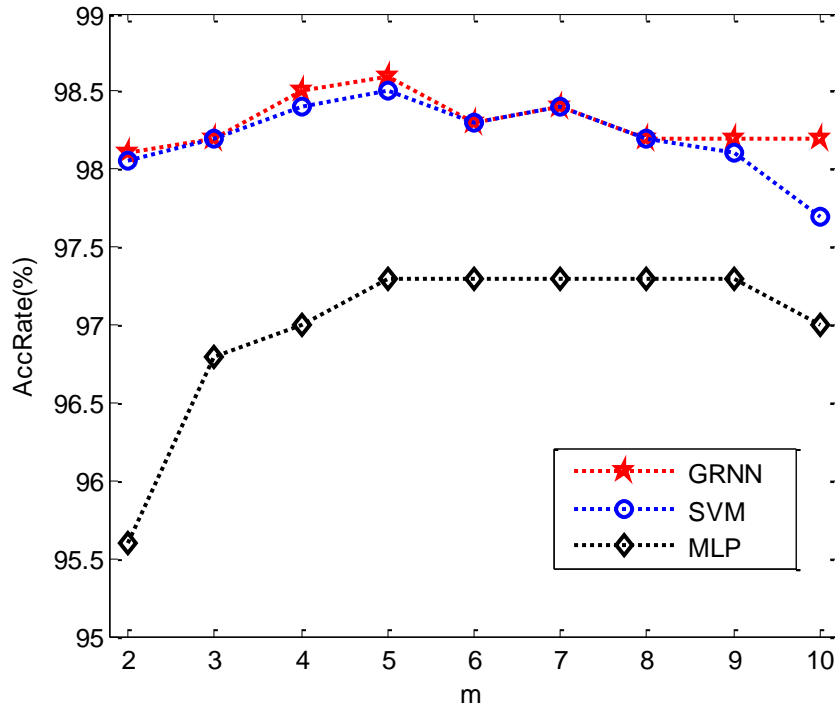
Bước 4. Huấn luyện

Bảng 6.6 Các nhóm mẫu được tách ra từ $CL(900,400)$

Bộ mẫu đã rút gọn	Số nhóm tách ra			
	2	3	4	5
$CL(900,400) \rightarrow$	$CL_1(634,154)$	$CL_1(580,119)$	$CL_1(238,97)$	$CL_1(173,25)$
\downarrow	$CL_2(266,254)$	$CL_2(242,199)$	$CL_2(205,149)$	$CL_2(184,145)$
\downarrow		$CL_3(78,82)$	$CL_3(66,68)$	$CL_3(69,68)$
\downarrow			$CL_4(391,86)$	$CL_4(387,85)$
\downarrow				$CL_5(87,77)$
Số nhóm tách ra				
6	7	8	9	10
$CL_1(177,25)$	$CL_1(154,14)$	$CL_1(155,14)$	$CL_1(151,14)$	$CL_1(116,23)$
$CL_2(119,112)$	$CL_2(119,112)$	$CL_2(87,94)$	$CL_2(82,79)$	$CL_2(86,84)$
$CL_3(74,65)$	$CL_3(61,62)$	$CL_3(61,62)$	$CL_3(63,62)$	$CL_3(67,64)$
$CL_4(182,76)$	$CL_4(184,76)$	$CL_4(60,38)$	$CL_4(50,36)$	$CL_4(50,36)$
$CL_5(85,77)$	$CL_5(76,51)$	$CL_5(76,51)$	$CL_5(76,47)$	$CL_5(4,8)$
$CL_6(263,45)$	$CL_6(266,44)$	$CL_6(282,43)$	$CL_6(192,13)$	$CL_6(193,15)$
	$CL_7(40,41)$	$CL_7(40,41)$	$CL_7(38,44)$	$CL_7(53,70)$
		$CL_8(139,57)$	$CL_8(74,63)$	$CL_8(67,54)$
			$CL_9(174,42)$	$CL_9(175,42)$
				$CL_{10}(89,4)$

Tiến hành huấn luyện với dữ liệu của từng tập con các bộ phân lớp MLP, GRNN, và SVM. Độ chính xác phân lớp kiểm tra được trình bày ở Hình 6.14. Kết quả đánh giá độ chính xác phân lớp kiểm tra đạt cao nhất tại số nhóm được tách ra là 5 nhóm. Ký hiệu là $DP=\{(CL1 \cup CL2 \cup CL3 \cup CL4 \cup CL5)\}$. Như vậy, mô hình sau khi huấn luyện hoàn thành sẽ có 5 bộ phân lớp con song song. Mỗi bộ phân

lớp được huấn luyện với từng bộ dữ liệu con riêng, kết quả là mỗi bộ phân lớp con có thông số cấu trúc hoàn toàn khác nhau.



Hình 6.14 Đồ thị độ chính xác nhận dạng với số Bộ phân lớp song song

Kết quả đánh giá độ chính xác phân lớp trình bày ở Bảng 6.7 cho mô hình APCM đề nghị với bộ phân lớp MLP (APCM&MLP), Bảng 6.8 cho mô hình APCM đề nghị với bộ phân lớp GRNN (APCM&GRNN), và Bảng 6.9 cho mô hình đề nghị với bộ phân lớp SVM (APCM&SVM). Độ chính xác (Pre), độ hồi tưởng (Rec) và hệ số F-score được trình bày ở Bảng 6.10 cho mô hình APCM gồm 5 bộ phân lớp song song.

Bảng 6.7 Kết quả huấn luyện trung bình gồm 5 bộ MLPC song song

Số nhóm tách ra	MLPC						
	Số nơ-ron ẩn					AccRate(%)	
						Huấn luyện	Kiểm tra
5	18	28	15	19	11	95,4	97,3

Bảng 6.8 Kết quả huấn luyện gồm 5 bộ GRNNC song song

Số nhóm tách ra	GRNNC						
	Hệ số Spread					AccRate(%)	
						Huấn luyện	Kiểm tra
5	0,01	0,216	0,178	0,01	0,01	99,3	98,7

Bảng 6.9 Kết quả huấn luyện gồm 5 bộ SVMC song song

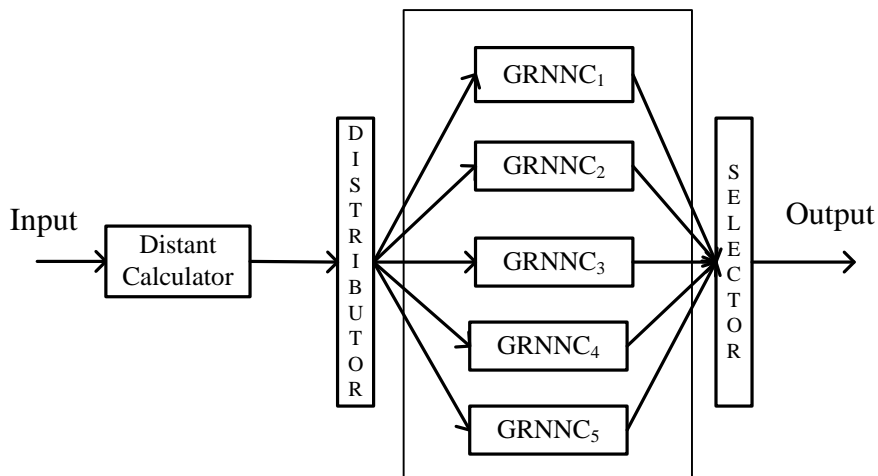
Số nhóm tách ra	SVMC						
	(C,Ks)					AccRate (%)	
						Huấn luyện	Kiểm tra
5	(1.6487, 0.3697)	(2.0138, 0.3679)	(0.6065, 0.2231)	(4.4817, 0.4066)	(0.8187, 0.2466)	97.3	98.5

Bảng 6.10 Độ chính xác, độ hồi tưởng và hệ số F-Score của mô hình APCM

5MLPC					
Số bộ phân lớp	1	2	3	4	5
Pre	0.9964	0.9501	1.000	0.9939	0.9937
Rec	0.9691	0.9675	0.9775	0.9727	0.9754
F-Score	0.9826	0.9587	0.9886	0.9833	0.9845
5GRNNC					
Pre	0.9859	0.9760	0.9908	0.9845	1.000
Rec	0.9964	0.9645	0.9953	0.9986	1.000
F-Score	0.9911	0.9702	0.9931	0.9915	1.000
5SVMC					
Pre	1.000	0.9820	0.9908	0.9979	1.000
Rec	0.9793	0.9686	1.000	0.9821	0.9876
F-Score	0.9895	0.9753	0.9954	0.9899	0.9938

Kết quả xây dựng mô hình:

- Bảng 6.7, 6.8, và 6.9 cho thấy bộ phân lớp áp dụng mạng nơ-ron GRNN cho kết quả nhận dạng cao hơn bộ phân lớp áp dụng mạng nơ-ron MLP là 1,4% và bộ phân lớp áp dụng SVM là 0,1%. Như vậy, Bộ phân lớp áp dụng mạng GRNN được chọn áp dụng cho mô hình. Mô hình APCM&GRNN gồm 5 bộ phân lớp con làm việc song song, có sơ đồ được trình bày như Hình 6.15.
- Kết quả số liệu của quá trình thiết kế mô hình được trình bày Bảng 6.11



Hình 6.15 Mô hình APCM gồm 5 bộ phân lớp song song GRNNC

Bảng 6.11 Tóm tắt kết quả xử lý dữ liệu

	Số biến	Mẫu
Dữ liệu ban đầu	104	D(2694,902)
Lựa chọn biến	15	D(2694,902)
Giảm mẫu	15	CL(900,400)
Tách mẫu	15	$CL_1(173,25) \cup CL_2(184,145) \cup CL_3(69,68) \cup CL_4(387,85) \cup CL_5(87,77)$

6.6 NHẬN XÉT

❖ Giảm không gian biến:

- Hình 6.2 và Hình 6.3 thể hiện độ quan trọng của từng biến được xếp từ cao xuống thấp bằng giải thuật xếp hạng FR, chuẩn Fisher và Divergence.
- Hình 6.4 trình bày độ quan trọng của từng tập con biến với giải thuật SFFS, chuẩn SM. Kết quả cho thấy số biến trong tập con biến tăng đồng biến với độ tách biệt nhóm J_{SM} .
- Kết quả chọn biến Hình 6.14 có độ chính xác nhận dạng cho tập biến được chọn bởi giải thuật SFFS cao hơn giải thuật xếp hạng FR. Điều này chứng tỏ điểm mạnh của SFFS ở khả năng mở rộng không gian tìm kiếm so với FR. Mặc dù giải thuật FR đơn giản hơn nhưng việc kết hợp từng biến tốt đơn lẻ chưa đảm bảo hình thành một tập biến tốt so với SFFS.
- Có 15 biến được chọn từ giải thuật SFFS ở Bảng 6.2 hay số biến giảm đi 6,9 lần so với 104 biến ban đầu nhưng độ chính xác kiểm tra vẫn không suy giảm.

Như vậy, đề xuất áp dụng giải thuật chọn biến đã phát huy được hiệu quả và kết quả giảm đi một số lượng lớn biến không cần thiết, gây nhiễu. Điều này còn có ý nghĩa rất lớn trong việc giảm chi phí thiết bị đo lường cảm biến, đơn giản hóa sơ đồ đấu nối, và giúp hệ thống xử lý nhanh hơn.

❖ Giảm không gian mẫu:

- Hình 6.7, áp dụng giải thuật KM để giảm không gian mẫu với $D(2649)$ xuống $CL(S_R=900)$ thì số lần lặp là 27 lần, trong khi giải thuật đề nghị áp dụng HK chỉ lặp có 10 lần. Kết quả tương tự như ở Hình 6.8, áp dụng giải thuật KM để giảm không gian mẫu với $D(U_R=902)$ xuống $CL(U_R=400)$ số lần lặp là 19 lần, trong khi giải thuật đề nghị áp dụng HK chỉ lặp có 4 lần. Kết quả cho thấy giải thuật đề nghị áp dụng HK có số lần lặp ít, giá trị hàm mục tiêu cải thiện đáng kể so với giải thuật KM.
- Hình 6.10 và 6.11, giải thuật KM, bộ phân lớp 1-NNC, cho độ chính xác phân lớp cao nhất trong vùng khảo sát đạt 93,6% tại $CL(700,500)$ hay tổng

số mẫu đã giảm đến 66,2%. Trong khi, với giải thuật HK tại CL(300,150) hay tổng số mẫu giảm đến 87,3% với độ chính xác phân lớp kiểm tra đạt đến 95,2%.

- Kết quả số mẫu sau rút gọn được chọn ở Bảng 6.3 với giải thuật áp dụng HK, số mẫu từ D(2649,902) giảm xuống CL(900,400) hay không gian mẫu ổn định giảm còn 33,87% và không gian mẫu không ổn định giảm còn 44,3% so với số mẫu ban đầu, trong khi kết quả kiểm tra đạt 98%.
- So với giải thuật KM, giải thuật giảm mẫu HK đề nghị áp dụng cải thiện độ chính xác đáng kể, số lần lặp ít hơn khi tìm tâm hay hội tụ nhanh hơn. Điều này khẳng định giải thuật HK đã khắc phục được nhược điểm hội tụ địa phương của giải thuật KM.
- Bảng 6.4, với bộ mẫu rút gọn CL(900,400) làm dữ liệu học, độ chính xác kiểm tra phân lớp của GRNNC cho kết quả cao hơn MLPC và SVMC tương ứng là 2,9% và 0,1%.
- Bảng 6.5 giá trị F-Score của các bộ phân lớp cao, trong đó thấp nhất trong ba bộ phân lớp là MLP nhưng F-Score cũng đạt là 0.9679. Giá trị F-Score của GRNN cao nhất và đạt đến 0.9873, cho thấy giá trị độ chính xác và độ hội tụ đạt sự cân bằng tốt, và điều này thể hiện chất lượng phân lớp cao.

Như vậy, giải thuật đề nghị HK khẳng định khả năng giảm không gian mẫu một cách hiệu quả với độ chính xác cao hơn giải thuật KM. Thực thi giảm không gian mẫu dẫn đến kết quả là mỗi nhóm dữ liệu có một tâm đại diện riêng. Điều này ý nghĩa quan trọng là làm cho tập dữ liệu học trở nên tin gọn, đặc trưng mang tính đại diện, giảm chi phí tính toán, tiết kiệm bộ nhớ máy tính. Giúp hệ thống nhận dạng xử lý nhanh, linh hoạt trong học mẫu mới, và tự động hóa làm mới tri thức hệ thống.

❖ **Xây dựng mô hình mạng nơ-ron song song cải tiến APCM:**

- Hình 6.14 trình bày kết quả thực nghiệm độ chính xác phân lớp với số mạng nơ-ron song song từ 2 đến 10 tương ứng với số nhóm dữ liệu được tách ra từ 2 đến 10 nhóm và số mẫu của mỗi nhóm con như ở Bảng 6.5. Kết quả độ chính xác phân lớp với mô hình cải tiến APCM có độ chính xác cao

nhất cho cả ba bộ phân lớp tại số nhóm bằng 5 với độ chính xác cho 5 bộ phân lớp song song MLP, GRNN, và SVM tương ứng là 97,3% (Bảng 6.7), 98,7% (Bảng 6.8), và 98,5% (Bảng 6.9). Trong khi với mô hình đơn MLP, GRNN, và SVM có độ chính xác tương ứng ở Bảng 6.4 là 95,2%, 98,1%, 98%. Kết quả cho thấy độ chính xác phân lớp của mô hình cải tiến tăng so với độ chính xác phân lớp đơn cho cả ba bộ phân lớp. Điều này cho thấy kết quả có tính thống nhất cho cả ba bộ phân lớp. Các công trình đã nghiên cứu chẩn đoán ổn định hệ thống điện đã công bố chẳng hạn các bài báo [32], [37], [48] có độ chính xác từ 94% đến 97%, cho nên kết quả đánh giá đánh độ chính xác phân lớp trong nghiên cứu hoàn toàn chấp nhận được và cho thấy mô hình đề xuất đã nâng cao được độ chính xác phân lớp.

- Mô hình song song gồm có 5 bộ phân lớp con:
 - Bảng 6.4 và Bảng 6.6, với mô hình APCM&MLP thì độ chính xác kiểm tra tăng 2,1% so với mô hình đơn. Bảng 6.4 và Bảng 6.7, với mô hình APCM&GRNN thì độ chính xác kiểm tra tăng 0,6% so với mô hình đơn. Bảng 6.4 và 6.8, với mô hình APCM&SVM thì độ chính xác tăng 0,5% so với mô hình đơn. Điều này khẳng định mô hình đề xuất đã cải thiện được độ chính xác phân lớp.
 - Bảng 6.10 giá trị F-Score của các Bộ phân lớp cao, giá trị 0.9587 là giá trị thấp nhất của một bộ phân lớp trong năm bộ phân lớp MLPC. Các giá trị F-Score của một bộ phân lớp trong năm bộ phân lớp GRNNC đạt giá trị cao nhất bằng 1. Kết quả đánh giá cho thấy giá trị độ chính xác và độ hồi tưởng đạt sự cân bằng cao, điều này thể hiện chất lượng phân lớp cao.

Như vậy, kết quả nghiên cứu cho thấy mô hình đề xuất đã đạt được mục tiêu quan trọng là nâng cao độ chính xác nhận dạng.

6.7 TÓM TẮT CHƯƠNG 6

Chương 6 trình bày các kết quả nghiên cứu đã được áp dụng và kiểm tra trên sơ đồ chuẩn IEEE 39-bus. Các kết quả cụ thể như sau:

- ✓ Xây dựng tập mẫu: áp dụng quy trình xây dựng tập mẫu đã trình bày ở chương 4 vào xây dựng tập mẫu cho sơ đồ IEEE 39-bus. Bộ dữ liệu có 3551 mẫu với 2649 mẫu ổn định và 902 mẫu không ổn định, $D(S,U)=D(2649,902)$.
- ✓ Giảm không gian biến: Hai phương pháp chọn biến đã được áp dụng là phương pháp xếp hạng biến với chuẩn Fisher và chuẩn Divergence, phương pháp tìm kiếm tiến lùi SFFS với chuẩn SM. Kết quả nghiên cứu cho thấy giải thuật SFFS cho kết quả tốt hơn phương pháp xếp hạng FR. Điều này có thể giải thích là phương pháp SFFS đã mở rộng được không gian tìm kiếm so với phương pháp xếp hạng nên đã truy tìm được tập biến cho độ chính xác cao hơn.
- ✓ Giảm không gian mẫu: Giải thuật K-means là giải thuật phân cụm dữ liệu tính toán nhanh đơn giản, tuy nhiên K-means có nhược điểm là dễ rơi vào hội tụ địa phương. Tác giả đã đề nghị áp dụng giải thuật K-means lai (Hybrid K-means -HK) nhằm tránh nhược điểm của K-means. Kết quả nghiên cứu cho thấy giải thuật HK khắc phục được điểm yếu của giải thuật K-means, giúp giảm không gian mẫu hiệu quả và cho độ chính xác cao. Vấn đề giảm không gian dữ liệu đã được công bố trong bài báo (I).
- ✓ Cải tiến mô hình mạng nơ-ron song song: áp dụng quy trình xây dựng mô hình mạng nơ-ron song song cải tiến đề xuất, kết quả nghiên cứu cho thấy mô hình đề xuất giải quyết được mục tiêu nâng cao độ chính xác phân lớp. Mỗi một mô hình con có một tâm đại diện được lưu trữ. Điều này giúp giảm không gian bộ nhớ rất lớn, làm cho mô hình tự nhận biết mẫu mới nhanh, và có tính tự thích nghi với mẫu mới cao, việc cập nhật mẫu mới và làm tươi bộ nhớ đơn giản nhanh chóng.

Chương 7. KẾT LUẬN

7.1 KẾT QUẢ ĐẠT ĐƯỢC

Hệ thống điện hiện đại đối mặt với áp lực vận hành cận biên giới hạn ổn định do tốc độ phát triển của nguồn phát, đường dây không theo kịp tốc độ phát triển của phụ tải. Điều này đưa đến hệ thống điện vận hành càng dễ mất ổn định. Trong những năm gần đây trên thế giới đã chứng kiến rất nhiều sự cố tan rã hệ thống điện nghiêm trọng gắn liền với sự mất ổn định của hệ thống điện và đã gánh lấy những thiệt hại kinh tế vô cùng to lớn.

Xử lý quá trình quá độ trong hệ thống điện có một giá trị rất quan trọng để nâng cao độ tin cậy, tính ổn định chế độ vận hành hệ thống điện và đòi hỏi khắc khe về thời gian. Để đánh giá tính ổn định quá độ của hệ thống điện trong những dao động lớn do sự cố gây ra, các phương pháp truyền thống tỏ ra kém hiệu quả. Chẩn đoán nhanh mất ổn định động hệ thống điện thì cơ hội để điều khiển đưa hệ thống điện trở về trạng thái ổn định khả thi hơn. Vì vậy, một nhu cầu là cần hệ thống chẩn đoán nhanh ổn định động hệ thống điện với độ chính xác cao. Từ mục đích và nhiệm vụ nghiên cứu đặt ra, luận án đã hoàn thành những nội dung khoa học và thực tiễn như sau:

❖ Về khoa học:

1. Đề xuất quy trình xây dựng tập biến cho bài toán nhận dạng ổn định động hệ thống điện. Giới thiệu và áp dụng thành công hai giải thuật chọn biến, SFFS và FR, vào giảm không gian biến một cách hiệu quả. Kết quả kiểm tra trên sơ đồ IEEE 39-bus, giải thuật SFFS cho kết quả chọn biến tốt hơn phương pháp FR nhờ vào mở rộng không gian tìm kiếm của nó.
2. Áp dụng thành công giải thuật giảm không gian mẫu K-means lai (Hybrid K-means – HK), cải tiến từ giải thuật phân cụm dữ liệu nổi tiếng K-means, khai phá dữ liệu ổn định động hệ thống điện. Kết quả kiểm tra trên tập mẫu của sơ đồ IEEE 39-bus cho thấy giải thuật HK đã khắc phục được nhược điểm K-means, giúp nâng cao chất lượng phân cụm dữ liệu, và giảm không gian mẫu một cách hiệu quả. Việc giảm không gian mẫu làm cho mỗi nhóm mẫu con có

một tâm đại diện. Điều này có ý nghĩa rất quan trọng là làm cho mô hình không tăng dung lượng bộ nhớ lưu trữ mẫu mới, linh hoạt trong việc cập nhật làm giàu tri thức mới, và giúp mô hình có khả năng bao phủ dữ liệu, thích nghi với mẫu mới rất cao.

3. Đề xuất quy trình xây dựng hệ thống nhận dạng thông minh chẩn đoán nhanh ổn định động hệ thống điện dựa trên cơ sở mạng nơ-ron, khai phá dữ liệu. Luận án đã phát triển thành công mô hình mạng nơ-ron song song cải tiến cho bài toán nhận dạng ổn định động hệ thống điện. Mô hình đề xuất đạt được mục tiêu quan trọng là nâng cao độ chính xác phân lớp. Với cách tiếp cận trình bày trong luận án, các khâu thực hiện đã được quy trình hóa, điều này giúp mô hình nhận dạng hoàn toàn có thể mở rộng cho các trường hợp nhiều lớn khác nhau.
4. Bài toán nhận dạng trong đề tài có thể tổng hợp cho bài toán điều khiển ổn định hệ thống điện chẳng hạn như điều khiển sa thải phụ tải.

❖ **Về thực tiễn:**

5. Xây dựng cách tiếp cận nhận dạng nhanh trạng thái ổn định động hệ thống điện khi có sự cố ngắn mạch nghiêm trọng xảy ra chỉ dựa trên các biến đầu vào là độ thay đổi công suất tác dụng tải, độ thay đổi công suất tác dụng trên các nhánh, và sụt điện áp tại các nút. Kết quả kiểm tra cho độ chính xác nhận dạng cao về trạng thái ổn định hệ thống điện mà không cần giải hệ phương trình vi phân như phương pháp truyền thống.
6. Hệ thống nhận dạng thông minh được xây dựng có thể được sử dụng như công cụ trợ giúp các điều độ viên đề ra chiến lược điều khiển trong những tình huống khẩn cấp, và huấn luyện các điều độ viên hệ thống điện xử lý các tình huống dựa trên các kịch bản sự cố.
7. Việc giảm không gian biến có ý nghĩa rất lớn trong việc giảm chi phí thiết bị đo lường cảm biến, đơn giản hóa sơ đồ đấu nối, và giúp hệ thống tăng tốc độ xử lý. Việc giảm không gian mẫu có ý nghĩa quan trọng trong việc tiết kiệm bộ nhớ lưu trữ mẫu, giảm chi phí thu thập dữ liệu, và giúp mô hình dễ dàng cập nhật dữ liệu mới cũng như tính thích nghi với mẫu mới rất cao.

7.2 HƯỚNG PHÁT TRIỂN

Từ kết quả nghiên cứu trong luận án, các kiến nghị cho nghiên cứu cần tiếp tục triển khai trong tương lai như sau:

- Nghiên cứu phối hợp với khâu điều khiển khẩn cấp để giữ ổn định hệ thống điện trong trường hợp mất ổn định, chẳng hạn như điều khiển sa thải phụ tải.
- Nghiên cứu mở rộng áp dụng phương pháp nghiên cứu trong đề tài cho nhận dạng ổn định điện áp, ổn định tần số.
- Nghiên cứu mở rộng áp dụng mô hình đề xuất vào lưới điện Việt Nam.

TÀI LIỆU THAM KHẢO

- [1] Y. V. Makarov, V. I. Reshetov, V. a. Stroeve, and N. I. Voropai, “Blackout prevention in the United States, Europe, and Russia,” *Proc. IEEE*, vol. 93, no. 11, pp. 1942–1954, 2005.
- [2] B. Liscouski and W. Elliot, “U.S.-Canada Power System Outage Task Force,” *System*, vol. 40, no. April, p. 238, 2004.
- [3] J. A. Laghari, H. Mokhlis, A. H. A. Bakar, and H. Mohamad, “Application of computational intelligence techniques for load shedding in power systems : A review,” *ENERGY Convers. Manag.*, vol. 75, no. August 2003, pp. 130–140, 2013.
- [4] Z. Dong and P. Zhang, *Emerging Techniques in Power System Analysis*. Springer, 2010.
- [5] A. M. A. Haidar, A. Mohamed, A. Hussain, and N. Jaalam, “Artificial Intelligence application to Malaysian electrical powersystem,” *Expert Syst. Appl.*, vol. 37, no. 7, pp. 5023–5031, 2010.
- [6] P. Lakra and M. Kirar, “Load Shedding techniques for System With Cogeneration : a Review,” vol. 4, no. 3, 2015.
- [7] Lã Văn Út, *Phân tích & điều khiển ổn định hệ thống điện*. NXB Đại Học Quốc Gia TP.HCM, 2011.
- [8] M. Pavella, D. Ernst, and D. Ruiz-Vega, “Transient Stability of Power Systems A Unified Approach to Assessment and Control,” *Kluwer Acad. Publ.*, pp. 1–254, 2000.
- [9] Nguyễn Hoàng Việt and Phan Thị Thanh Bình, *Ngăn Mạch & Ổn Định Trong Hệ Thống Điện*. NXB Đại Học Quốc Gia TP.HCM, 2011.

- [10] P. Kundur, N. J. Balu, and M. G. Lauby, *Power system stability and control*. McGraw-Hill, Inc, 1994.
- [11] P. Kundur, J. Paserba, and V. Ajjarapu, “Definition and classification of power system stability,” *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, 2004.
- [12] J. D. Glover, M. S. Sarma, and T. Overbye, *Power System Analysis and Design*, Fifth Edit. Global Engineering: Christopher M. Shortt Acquisitions, 2012.
- [13] T. a. Lipo, *Analysis of Synchronous Machines*, Second Edi, 545p, CRC Press, 2012.
- [14] C. Cecati and H. Latafat, “Time domain approach compared with direct method of Lyapunov for transient stability analysis of controlled power system,” *Int. Symp. Power Electron. Power Electron. Electr. Drives, Autom. Motion*, pp. 695–699, 2012.
- [15] B. Bonvini, S. Massucco, a. Morini, and T. Siewierski, “A comparative analysis of power system transient stability assessment by direct and hybrid methods,” *Proc. 8th Mediterr. Electrotech. Conf. Ind. Appl. Power Syst. Comput. Sci. Telecommun. (MELECON 96)*, vol. 3, pp. 1575–1579, 1996.
- [16] M. A. Pai, *Energy Function Analysis For Power System Stability*. Kluwer Academic Publishers, 1989.
- [17] J. I. F. Icatons, S. Application, and Y. Xue, “Extended Equal Area Criterion,” vol. 4, no. 1, pp. 44–52, 1989.
- [18] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, “Support vector machine-based algorithm for post-fault transient

- stability status prediction using synchronized measurements,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, 2011.
- [19] M. Giroletti, M. Farina, and R. Scattolini, “Electrical Power and Energy Systems Short Communication A hybrid frequency / power based method for industrial load shedding,” *Int. J. Electr. Power Energy Syst.*, vol. 35, no. 1, pp. 194–200, 2012.
- [20] T. Shekari, F. Aminifar, and M. Sanaye-Pasand, “An analytical adaptive load shedding scheme against severe combinational disturbances,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4135–4143, 2015.
- [21] M. Moazzami and a. Khodabakhshian, “A new optimal adaptive under frequency load shedding Using Artificial Neural Networks,” *2010 18th Iran. Conf. Electr. Eng.*, pp. 824–829, 2010.
- [22] H. Shah and K. Verma, “PMU-ANN based Approach for Real Time Voltage Stability Monitoring,” *6th Int. Conf. Power Syst.*, no. New Delhi, India, IEEE, 2016.
- [23] O. Shariati, A. A. M. Zin, A. Khairuddin, M. Pesaran H.A, and M. Aghamohammadi, “An Integrated Method for under Frequency Load Shedding Based on Hybrid Intelligent System-Part I: UFLS Design,” *Power and Energy Engineering Conference (APPEEC), 2012 Asia-Pacific*, 2012.
- [24] O. Shariati, A. A. M. Zin, A. Khairuddin, M. Pesaran H.A, and M. Aghamohammadi, “An Integrated Method for under Frequency Load Shedding Based on Hybrid Intelligent System-Part II: UFLS Design,” in *Power and Energy Engineering Conference (APPEEC), 2012 Asia-Pacific*, 2012.

- [25] R. Zhang, Y. Xu, Z. Y. Dong, K. Meng, and Z. Xu, “Intelligent Systems for Power System Dynamic Security Assessment: Review and Classification,” *4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, 2011, pp. 134–139.
- [26] S. Kalyani and K. S. Swarup, “Pattern analysis and classification for security evaluation in power networks,” *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, pp. 547–560, 2013.
- [27] A. Karami and S. Z. Esmaili, “Transient stability assessment of power systems described with detailed models using neural networks,” *Int. J. Electr. Power Energy Syst.*, vol. 45, no. 1, pp. 279–292, 2013.
- [28] A. Sharifian and S. Sharifian, “A new power system transient stability assessment method based on Type-2 fuzzy neural network estimation,” *Int. J. Electr. Power Energy Syst.*, vol. 64, pp. 71–87, 2015.
- [29] Y. Xu *et al.*, “Assessing Short-Term Voltage Stability of Electric Power Systems by a Hierarchical Intelligent System,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 27, no. 8, pp. 1686–1696, 2016.
- [30] M. K. Paramathma, D. Devaraj, and S. R. B., “Artificial Neural Network based Static Security Assessment Module using PMU Measurements for Smart Grid Application,” *IEEE*, pp. 1–5, 2016.
- [31] A. M. a. Haidar, M. W. Mustafa, F. a. F. Ibrahim, and I. a. Ahmed, “Transient stability evaluation of electrical power system using generalized regression neural networks,” *Appl. Soft Comput.*, vol. 11, no. 4, pp. 3558–3570, 2011.
- [32] A. Y. Abdelaziz and M. A. El-Dessouki, “Transient Stability Assessment using Decision Trees and Fuzzy Logic Techniques,” *Int. J.*

Intell. Syst. Appl., vol. 5, no. 10, pp. 1–10, 2013.

- [33] A. Karami, “Power system transient stability margin estimation using neural networks,” *Int. J. Electr. Power Energy Syst.*, vol. 33, no. 4, pp. 983–991, 2011.
- [34] P. K. Olulope, K. a. Folly, S. P. Chowdhury, and S. Chowdhury, “Prediction of critical clearing time using artificial neural network,” *2011 IEEE Symp. Comput. Intell. Appl. Smart Grid*, no. Ci, pp. 1–5, 2011.
- [35] I. B. Sulistiawati, M. Abdillah, and A. Soeprijanto, “Neural network based transient stability model to analyze the security of Java-Bali 500 kV power system,” *Proc. 2011 Int. Conf. Electr. Eng. Informatics*, no. July, pp. 1–6, 2011.
- [36] V. L. Paucar and F. C. Fernandes, “A methodology based on neural networks for the determination of the critical clearing time of power systems transient stability,” *Proceedings. Int. Conf. Power Syst. Technol.*, vol. 4, no. Lm, pp. 2669–2673, 2002.
- [37] R. Zhang, S. Member, Y. Xu, and Z. Y. Dong, “Feature Selection For Intelligent Stability Assessment of Power Systems,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–7, 2012.
- [38] N. Amjady and S. F. Majedi, “Transient stability prediction by a hybrid intelligent system,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1275–1283, 2007.
- [39] A. Y. Abdelaziz, M. R. Irving, A. M. El-Arabaty, and M. M. Mansour, “Out-of-step prediction based on artificial neural networks,” *Electr. Power Syst. Res.*, vol. 34, no. 2, pp. 135–142, 1995.

- [40] A. M. El-Arabaty, H. a. Talaat, M. M. Mansour, and A. Y. Abd-Elaziz, “Out-of-step detection based on pattern recognition,” *Int. J. Electr. Power Energy Syst.*, vol. 16, no. 4, pp. 269–275, 1994.
- [41] Q.H.Anh, “The applycation of pattern recognition for fast analysis of the dynamic stability of electrical power system,” *Perganon*, pp. 1–13, 1994.
- [42] Y. Xu, Z. Y. Dong, L. Guan, R. Zhang, and F. Luo, “Preventive Dynamic Security Control of Power Systems Based on Pattern Discovery Technique,” vol. 27, no. 3, pp. 1236–1244, 2012.
- [43] S. A. Siddiqui, K. Verma, K. R. Niazi, and M. Fozdar, “Preventive and Emergency Control of Power System for Transient Stability Enhancement,” *JEET*, vol. 10, pp. 742–750, 2015.
- [44] B. Sun, M. Liu, L. Zhu, N. Liu, X. Qiu, and Z. Zhuang, “Emergency Control Strategy Based on Multi-agent Theory under Blackout,” *Energy Power Eng.*, vol. 5, no. 4, pp. 717–721, 2013.
- [45] N. I. Voropai, D. N. Efimov, P. V Etingov, and D. a Panasetsky, “Smart Emergency Control in Electric Power Systems,” pp. 1658–1664, 2011.
- [46] Y. V Makarov, N. I. Voropai, and D. N. Efimov, “Complex Emergency Control System Against Blackouts in Russia,” in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, PA, USA*, 2008, pp. 1–8.
- [47] N. I. A. Wahab, A. Mohamed, and A. Hussain, “Feature selection and extraction methods for power systems transient stability assessment employing computational intelligence techniques,” *Neural Process.*

- Lett.*, vol. 35, no. 1, pp. 81–102, 2012.
- [48] K. S. Swarup, “Artificial neural network using pattern recognition for security assessment and analysis,” *Neurocomputing*, vol. 71, no. 4–6, pp. 983–998, 2008.
- [49] Y. Xu, Z. Y. Dong, K. Meng, R. Zhang, and K. P. Wong, “Real-time transient stability assessment model using extreme learning machine,” *IET Gener. Transm. Distrib.*, vol. 5, no. 3, p. 314, 2011.
- [50] K. R. Niazi, C. M. Arora, and S. L. Surana, “Power system security evaluation using ANN: feature selection using divergence,” *Proc. Int. Jt. Conf. Neural Networks, 2003.*, vol. 3, pp. 2094–2099, 2003.
- [51] S. Zarrabian, R. Belkacemi, and A. A. Babalola, “Intelligent mitigation of blackout in real-time microgrids: Neural network approach,” *Power Energy Conf. Illinois (PECI), 2016 IEEE*, 2016.
- [52] Y. Zhou, J. Wu, L. Hao, L. Ji, and Z. Yu, “Transient Stability Prediction of Power Systems Using Post-disturbance Rotor Angle Trajectory Cluster Features,” *Electr. Power Components Syst.*, no. September, 2016.
- [53] W. D. Oliveira, J. P. A. Vieira, U. H. Bezerra, D. A. Martins, and G. Rodrigues, “Power system security assessment for multiple contingencies using multiway decision tree,” *Electr. Power Syst. Res.*, vol. 148, pp. 264–272, 2017.
- [54] I. S. Saeh, M. W. Mustafa, Y. S. Mohammed, and M. Almakhtar, “Static Security classification and Evaluation classifier design in electric power grid with presence of PV power plants using C-4.5,” *Renew. Sustain. Energy Rev.*, vol. 56, pp. 283–290, 2016.

- [55] P. H. Đ. Dục, *Mạng Noron & Ứng Dụng Trong Điều Khiển Tự Động*. Nhà Xuất Bản Khoa Học và Kỹ Thuật, 2009.
- [56] N. Đ. Thúc and H. Đ. Hải, *Tri Tuệ Nhân Tạo - Mạng Noron-Phuong Pháp & Ứng Dụng*. Nhà Xuất Bản Giáo Dục, 2000.
- [57] S. Haykin, *Neural Networks and Learning Machines*, Third Edit. Pearson Prentice Hall, 2009.
- [58] D. Graupe, *Principles of Artificial Neural Networks*. World Scientific, 2007.
- [59] M. H. Beale, M. T. Hagan, and H. B. Demuth, “Neural Network Toolbox™ User’s Guide R2014a,” 2014.
- [60] D. F. Specht, “A general regression neural network,” *Neural Networks, IEEE Trans.*, vol. 2, no. 6, pp. 568–576, 1991.
- [61] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Fourth Edi. Elsevier Inc, 2009.
- [62] A. R. Webb and K. D. Copsey, *Statistical Pattern Recognition*, Third Edit. A John Wiley & Sons, Ltd., Publication, 2011.
- [63] B. E. Boser, T. B. Laboratories, I. M. Guyon, T. B. Laboratories, and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifiers,” *Proc. 5th Annu. Work. Comput. Learn. Theory*, p. page 144-152, 1992.
- [64] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Proc. Mach. Learn.*, vol. 20, pp. 273–297, 1995.
- [65] H. Byun and S. Lee, “Applications of Support Vector Machines for Pattern Recognition: A Survey,” *Springer-Verlag Berlin Heidelb. 2002*, pp. 213–236, 2002.

- [66] M. Cheriet, N. Kharm, C.-L. Liu, and C. Y. Suen, *Character Recognition Systems - A Guide for Students and Practitioners*. A John Wiley Inc., 2007.
- [67] K. Y. Lee and M. A. El-Sharkawi, *Modern Heuristic Optimization Techniques*. A John Wiley & Sons. Inc. Publication, 2008.
- [68] S. L. Chiu, “Fuzzy model Identification base on cluster estimation,” *J. Intell. Fuzzy Syst.*, vol. 2, pp. 267–278, 1994.
- [69] J. Kennedy and R. Eberhart, “Particle Swarm Optimization,” *IEEE Int. Conf. , Perth, WA, Aust.*, no. ISBN: 0-7803-2768-3, pp. 1942–1948, 1995.
- [70] I. H. Witten, E. Frank, and M. a. Hall, *Data Mining: Practical Machine Learning Tools and Techniques, Third Edition*, vol. 54, no. 2. Elsevier Inc, 2011.
- [71] A. K. Menahem Friedman, *Introduction To Pattern Recognition : statical, structural, neural, and fuzzy logic approaches*. Imperial College Press, 1999.
- [72] S. Haykin, *Neural Networks-A Comprehensive Foundation*, Second Edi. Tom Robbins, 1999.
- [73] S. Kalyani and K. S. Swarup, “Electrical Power and Energy Systems Pattern analysis and classification for security evaluation in power networks,” *Int. J. Electr. POWER ENERGY Syst.*, vol. 44, no. 1, pp. 547–560, 2013.
- [74] K. G. Sheela and S. N. Deepa, “Review on methods to fix number of hidden neurons in neural networks,” *Math. Probl. Eng. Hindawi Publ. Corp.*, p. 11 p, 2013.

DANH MỤC CÁC CÔNG TRÌNH CÔNG BỐ

Các bài báo đã công bố

- I. **Ngoc Au Nguyen**, Trong Nghia Le, Huy Anh Quyen, Thi Thanh Binh Phan, "*Data reduction for dynamic stability classification in power system*", IETE Journal of Research, DOI:10.1080/03772063.2017.1417752, ISSN: 0377-2063 (Print) 0974-780X (Online), Jan. 2018. (SCIE)
- II. **Ngoc Au Nguyen**, Trong Nghia Le, Huy Anh Quyen, Binh Phan Thi Thanh, Thai Binh Nguyen "*Hybrid Classifier Model for Dynamic Stability Prediction in Power System*", IEEE International Conference On System Science And Engineering (ICSSE 2017), p. 67-70, 2017 IEEE.
- III. **Nguyen Ngoc Au**, Quyen Huy Anh, Phan Thi Thanh Binh, "*Dynamic Stability Recognition Of Power System Using Generalized Regression Neural Networks*" Journal of Technical Education Science, N^o 40A(10/2017), ISSN 1859-1272, 2017.
- IV. **Nguyen Ngoc Au**, Quyen Huy Anh, Phan Thi Thanh Binh, "*Feature subset selection in dynamic stability assessment power system using artificial Neural networks*" Science & Technology Development Journal, ISSN 1859-0128, Vol.18, p. 15-24, No.K3 – 2015.
- V. **Nguyen Ngoc Au**, Quyen Huy Anh, Phan Thi Thanh Binh, "*Dynamic stability assessment of power system using Multilayer feedforward Neural networks with Reduced feature selection*", The 2nd International Conference On Green Technology And Sustainable Development, 2014 (GTSD'14).
- VI. **Nguyen Ngoc Au**, Quyen Huy Anh, Phan Viet Thinh, "*Feature Selection For Dynamic Stability Prediction Of Power System Using Neural Network*", Journal of Technical Education Science, N^o 34(2015), ISSN 1859-1272, 2015.
- VII. Quyen Huy Anh, **Nguyen Ngoc Au**, Nguyen Vu Phuong Thao, "*Design dual input power system stabilizer for multi-machine system based on focused-time-delay Neural netwok*", Journal of Technical Education Science, N0 25(2013), ISSN 1859-1272.

Các bài báo liên quan

- VIII. **Nguyễn Ngọc Âu**, Lê Trọng nghĩa, Quỳên Huy Ánh, Phan Thị Thanh Bình, "*Sa thải phụ tải dựa trên nhận dạng nhanh ổn định động hệ thống điện*",

- ISSN 1859-1531 Tạp Chí Khoa Học Và Công Nghệ, Đại Học Đà Nẵng-Số 11(129).2017 QUYỀN 2, p. 6-11, 2017.
- IX. Trong Nghĩa Le, **Ngoc Au Nguyen**, Huy Anh Quyen, “*Emergency control of load shedding based on coordination of artificial Neural network and analytic hierarchy process algorithm*”, IEEE International Conference On System Science And Engineering (ICSSE 2017), p. 67-70, 2017 IEEE.
- X. L.T. Nghĩa, T.T. Giang, **N.N. Au**, Q.H. Anh, D.N. An, “*Emergency Control of Load Shedding Based on Fuzzy-AHP Algorithm*”, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 6 Issue 09, p. 185-191, September – 2017.
- XI. Phan Thị Thanh Bình, Nguyễn Thụy Mai Khanh, **Nguyễn Ngọc Âu**. “*Phân tích tình ổn định điện áp khi có máy phát điện gió DFIG*”, Tạp chí Phát triển KH&CN, Tập 19, Số K5-2016, trang 5-12, 2016.
- XII. Trong Nghĩa Le, Huy Anh Quyen, **Ngoc Au Nguyen**, “*Application of fuzzy-analytic hierarchy process algorithm and fuzzy load profile for load shedding in power systems*”, Electrical Power and Energy Systems 77 (2016) 178–184, 2016. (SCIE)

Chủ nhiệm đề tài cấp trường trọng điểm:

1. “*Hệ thống nhận dạng đánh giá ổn định động hệ thống điện*”, 2016-2017, T2017-68TĐ.
2. “*Ứng dụng mạng thần kinh nhân tạo chẩn đoán ổn định động hệ thống điện nhiều máy*”, 2015-2016, T2016-48TĐ.
3. “*Lựa chọn mẫu trong đánh giá thông minh ổn định động hệ thống điện*”, 21014-2015, T2015-34TĐ.
4. “*Đánh giá ổn định hệ thống điện nhiều máy phát*”, 2012-2013, T2013-02TĐ.

PHỤ LỤC

Bảng PL1. Công suất máy phát

Bus	MaxMvar	MinMVar	MinMW	MaxMW	Unit No.
30	800	-500	0.00	350.00	Gen10
31	800	-500	0.00	1150.00	Gen2
32	800	-500	0.00	750.00	Gen3
33	800	-500	0.00	732.00	Gen4
34	800	-300	0.00	608.00	Gen5
35	800	-500	0.00	750.00	Gen6
36	800	-500	0.00	660.00	Gen7
37	800	-500	0.00	640.00	Gen8
38	800	-500	0.00	930.00	Gen9
39	1500	-1000	0.00	1100.00	Gen1

Bảng PL2. Mô hình máy phát điện đồng bộ GENPWTwoAxis

Unit No.	Bus	H	R_a	x'_d	x'_q	x_d	x_q	T'_{do}	T'_{qo}
1	39	500.0	0	0.006	0.008	0.02	0.019	7.0	0.7
2	31	30.3	0	0.0697	0.170	0.295	0.282	6.56	1.5
3	32	35.8	0	0.0531	0.0876	0.2495	0.237	5.7	1.5
4	33	28.6	0	0.0436	0.166	0.262	0.258	5.69	1.5
5	34	26.0	0	0.132	0.166	0.67	0.62	5.4	0.44
6	35	34.8	0	0.05	0.0814	0.254	0.241	7.3	0.4
7	36	26.4	0	0.049	0.186	0.295	0.292	5.66	1.5
8	37	24.3	0	0.057	0.0911	0.290	0.280	6.7	0.41
9	38	34.5	0	0.057	0.0587	0.2106	0.205	4.79	1.96
10	30	42.0	0	0.031	0.008	0.1	0.069	10.2	0.0167

Bảng PL3. Thông số thiết bị điều khiển kích từ IEEE1

Bus	T_r	K_a	T_a	V_{rmax}	V_{rmin}	K_e	T_e	K_f	T_f	E_1	$S_E(E_1)$	E_2	$S_E(E_2)$
39	0	5	0.06	1	-1	-0.0485	0.25	0.04	1	0.75	0.08	1.1	0.26
38	0	6.2	0.05	1	-1	-0.633	0.405	0.057	0.5	0.75	0.66	1.1	0.88
37	0	5	0.06	1	-1	-0.0198	0.5	0.08	1	0.75	0.13	1.1	0.34
36	0	5	0.06	1	-1	-0.525	0.5	0.08	1	0.75	0.08	1.1	0.314
35	0	40	0.02	10	-10	1	0.785	0.03	1	0.75	0.67	1.1	0.91
34	0	5	0.02	1	-1	-0.0419	0.471	0.0754	1.246	0.75	0.064	1.1	0.251
33	0	40	0.02	6.5	-6.5	1	0.73	0.03	1	0.75	0.53	1.1	0.74
32	0	5	0.02	1	-1	-0.047	0.528	0.0845	1.26	0.75	0.072	1.1	0.282
31	0	40	0.02	10.5	-10.5	1	1.4	0.03	1	0.75	0.62	1.1	0.85
30	0	40	0.02	10	-10	1	0.785	0.03	1	0.75	0.67	1.1	0.91

Bảng PL4. Thông số thiết bị điều chỉnh tần số TGOV1

Bus	R	T_1	V_{max}	V_{min}	T_2	T_3	D_t
30	0.05	0.4	1	0	1.5	5	0
31	0.05	0.4	1	-0.05	1.5	5	0
32	0.05	0.4	1	-0.05	1.5	5	0
33	0.05	0.4	1	-0.05	1.5	5	0
34	0.05	0.4	1	-0.05	1.5	5	0
35	0.05	0.4	1	-0.05	1.5	5	0
36	0.05	0.4	1	-0.05	1.5	5	0
37	0.05	0.4	1	-0.05	1.5	5	0
38	0.05	0.4	1.2	-0.05	1.5	5	0
39	0.1	0.8	1.5	-0.05	1.5	5	0

Bảng PL5. Thông số máy biến áp

From Bus	To Bus	R_T	X_T	Tap
12	11	0.0016	0.0435	1.0060
12	13	0.0016	0.0435	1.0060
6	31	0.0000	0.0250	0.85714
10	32	0.0000	0.0200	1.0700
19	33	0.0007	0.0142	1.0700
20	34	0.0009	0.0180	1.0090
22	35	0.0000	0.0143	1.0250
23	36	0.0005	0.0272	1.0000
25	37	0.0006	0.0232	1.0250
2	30	0.0000	0.0181	1.0250
29	38	0.0008	0.0156	1.0250
19	20	0.0007	0.0138	1.0600

Bảng PL6. Thông số trở kháng đường dây

From Bus	To Bus	Branch Device Type	R	X	B
1	2	Line	0.0035	0.0411	0.6987
1	39	Line	0.0010	0.0250	0.7500
2	3	Line	0.0013	0.0151	0.2572
2	25	Line	0.0070	0.0086	0.1460
2	30	Transformer	0.0000	0.0181	0.0000
3	18	Line	0.0011	0.0133	0.2138
3	4	Line	0.0013	0.0213	0.2214
4	14	Line	0.0008	0.0129	0.1382
4	5	Line	0.0008	0.0128	0.1342
5	8	Line	0.0008	0.0112	0.1476
5	6	Line	0.0002	0.0026	0.0434
6	11	Line	0.0007	0.0082	0.1389
6	7	Line	0.0006	0.0092	0.1130
7	8	Line	0.0004	0.0046	0.0780
8	9	Line	0.0023	0.0363	0.3804
9	39	Line	0.0010	0.0250	1.2000
10	32	Transformer	0.0000	0.0200	0.0000
10	13	Line	0.0004	0.0043	0.0729
10	11	Line	0.0004	0.0043	0.0729
12	13	Transformer	0.0016	0.0435	0.0000
12	11	Transformer	0.0016	0.0435	0.0000
13	14	Line	0.0009	0.0101	0.1723
14	15	Line	0.0018	0.0217	0.3660
15	16	Line	0.0009	0.0094	0.1710
16	24	Line	0.0003	0.0059	0.0680
16	21	Line	0.0008	0.0135	0.2548
16	19	Line	0.0016	0.0195	0.3040
16	17	Line	0.0007	0.0089	0.1342
17	27	Line	0.0013	0.0173	0.3216
17	18	Line	0.0007	0.0082	0.1319
19	33	Transformer	0.0007	0.0142	0.0000
19	20	Transformer	0.0007	0.0138	0.0000
20	34	Transformer	0.0009	0.0180	0.0000
21	22	Line	0.0008	0.0140	0.2565
22	35	Transformer	0.0000	0.0143	0.0000
22	23	Line	0.0006	0.0096	0.1846

23	36	Transformer	0.0005	0.0272	0.0000
23	24	Line	0.0022	0.0350	0.3610
25	37	Transformer	0.0006	0.0232	0.0000
25	26	Line	0.0032	0.0323	0.5130
26	29	Line	0.0057	0.0625	1.0290
26	28	Line	0.0043	0.0474	0.7802
26	27	Line	0.0014	0.0147	0.2396
28	29	Line	0.0014	0.0151	0.2490
29	38	Transformer	0.0008	0.0156	0.0000
31	6	Transformer	0.0000	0.0250	0.0000

Bảng PL7. Thông số trở kháng thứ tự không đường dây

From Bus	To Bus	Branch Device Type	R	X	C
1	2	Line	0.008750	0.102750	0.698700
1	39	Line	0.002500	0.062500	0.750000
2	3	Line	0.003250	0.037750	0.257200
2	25	Line	0.017500	0.021500	0.146000
2	30	Transformer	0.000000	0.018100	0.000000
3	18	Line	0.002750	0.033250	0.213800
3	4	Line	0.003250	0.053250	0.221400
4	14	Line	0.002000	0.032250	0.138200
4	5	Line	0.002000	0.032000	0.134200
5	8	Line	0.002000	0.028000	0.147600
5	6	Line	0.000500	0.006500	0.043400
6	11	Line	0.001750	0.020500	0.138900
6	7	Line	0.001500	0.023000	0.113000
7	8	Line	0.001000	0.011500	0.078000
8	9	Line	0.005750	0.090750	0.380400
9	39	Line	0.002500	0.062500	1.200000
10	32	Transformer	0.000000	0.020000	0.000000

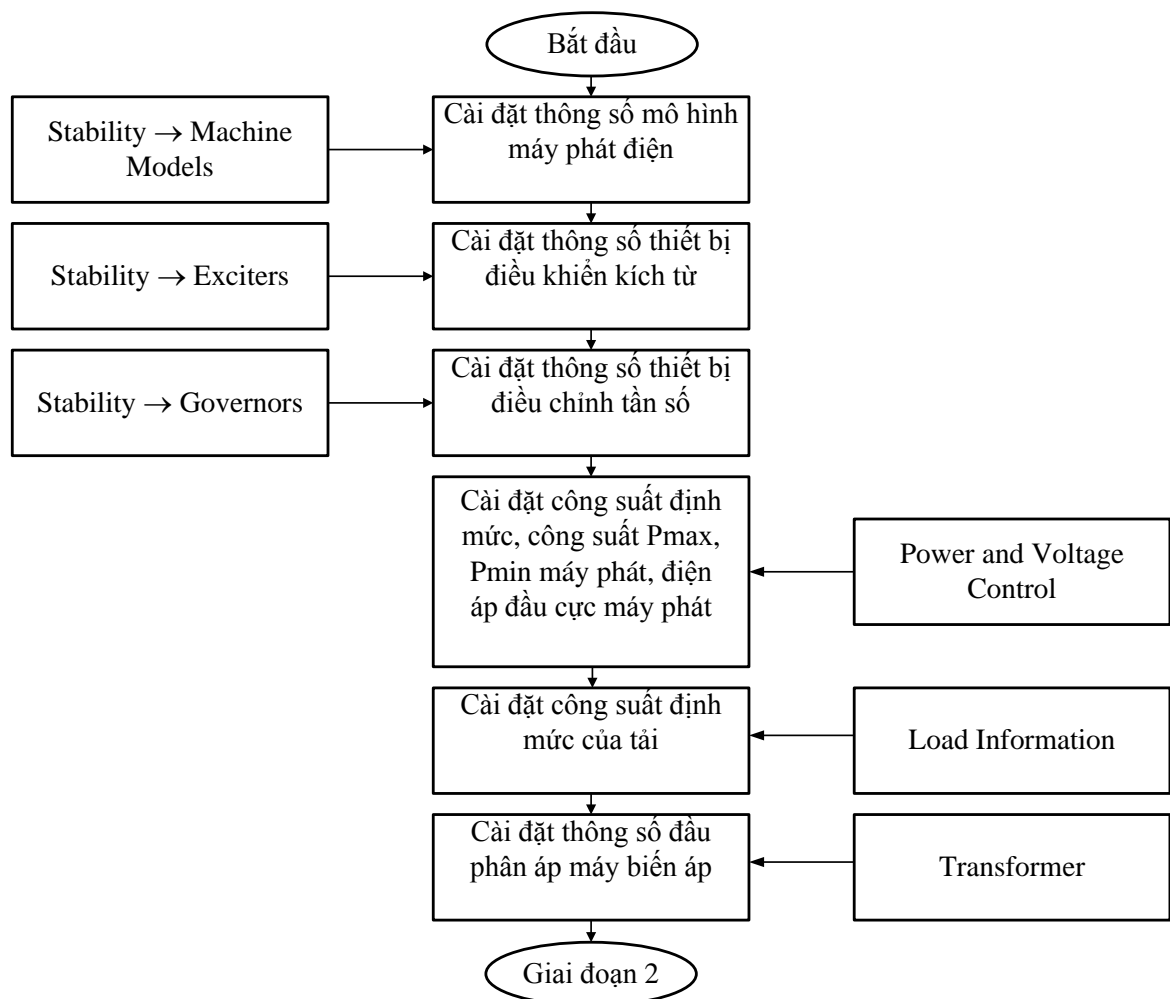
10	13	Line	0.001000	0.010750	0.072900
10	11	Line	0.001000	0.010750	0.072900
12	13	Transformer	0.001600	0.043500	0.000000
12	11	Transformer	0.001600	0.043500	0.000000
13	14	Line	0.002250	0.025250	0.172300
14	15	Line	0.004500	0.054250	0.366000
15	16	Line	0.002250	0.023500	0.171000
16	24	Line	0.000750	0.014750	0.068000
16	21	Line	0.002000	0.033750	0.254800
16	19	Line	0.004000	0.048750	0.304000
16	17	Line	0.001750	0.022250	0.134200
17	27	Line	0.003250	0.043250	0.321600
17	18	Line	0.001750	0.020500	0.131900
19	33	Transformer	0.0007	0.0142	0.0000
19	20	Transformer	0.0007	0.0138	0.0000
20	34	Transformer	0.0009	0.0180	0.0000
21	22	Line	0.002000	0.035000	0.256500
22	35	Transformer	0.0000	0.0143	0.0000
22	23	Line	0.001500	0.024000	0.184600
23	36	Transformer	0.0005	0.0272	0.0000
23	24	Line	0.005500	0.087500	0.361000
25	37	Transformer	0.0006	0.0232	0.0000
25	26	Line	0.008000	0.080750	0.513000
26	29	Line	0.014250	0.156250	1.029000
26	28	Line	0.010750	0.118500	0.780200
26	27	Line	0.003500	0.036750	0.239600
28	29	Line	0.003500	0.037750	0.249000
29	38	Transformer	0.0008	0.0156	0.0000
31	6	Transformer	0.0000	0.0250	0.0000

PL8. Các khâu thực thi chương trình PowerWorld để lấy mẫu

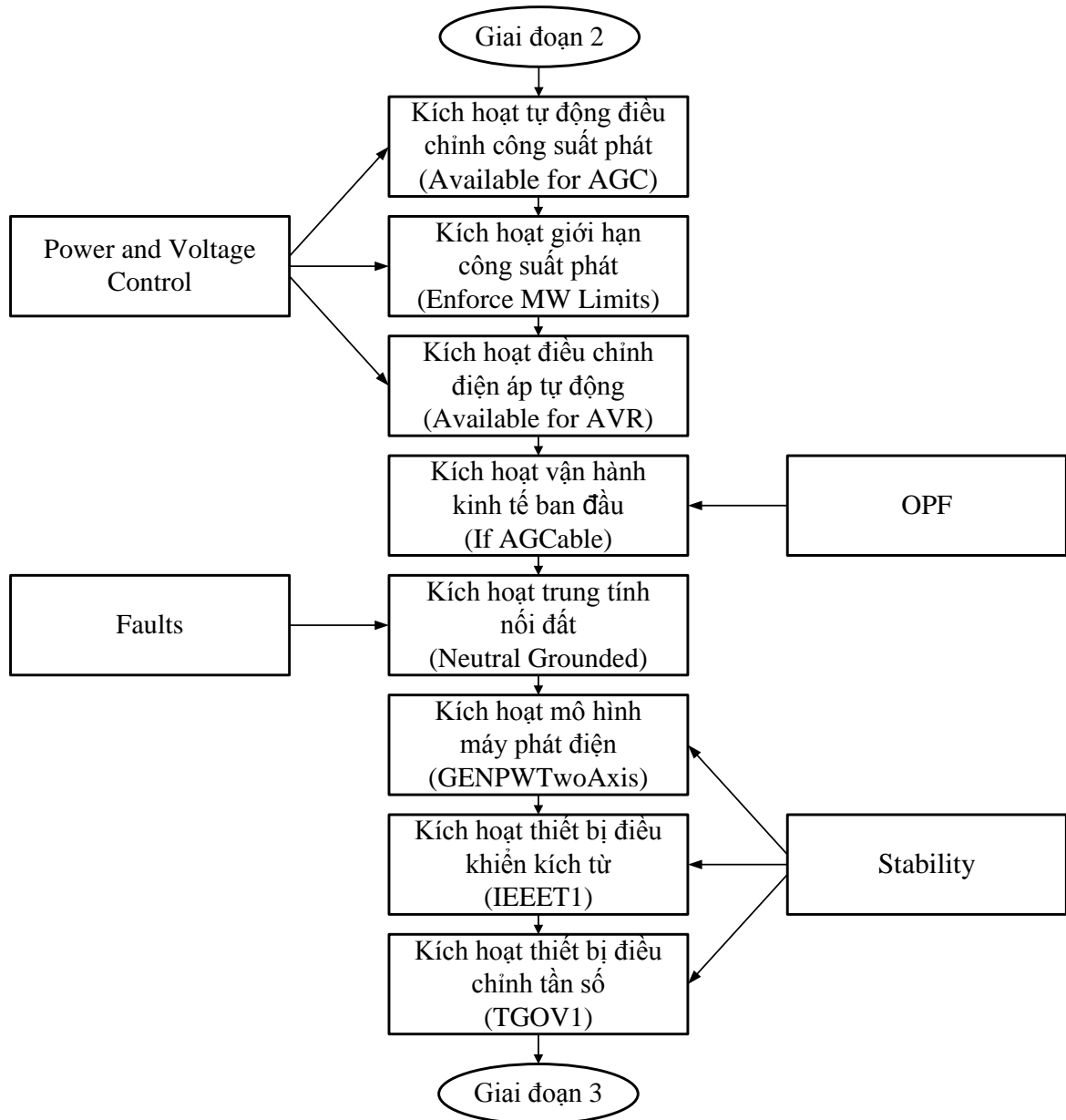
Các giai đoạn mô phỏng phần mềm PowerWorld để thu thập mẫu được đúc kết thành 4 giai đoạn gồm:

- Giai đoạn 1: Thực hiện cài đặt các thông số chuẩn của các mô hình trong hệ thống điện.
- Giai đoạn 2: Thực hiện kích hoạt các thông số trong mô hình hệ thống điện.
- Giai đoạn 3: Chạy phân bố công suất tối ưu OPF.
- Giai đoạn 4: Mô phỏng ổn định quá độ, đánh giá ổn định/không ổn định và lấy mẫu.

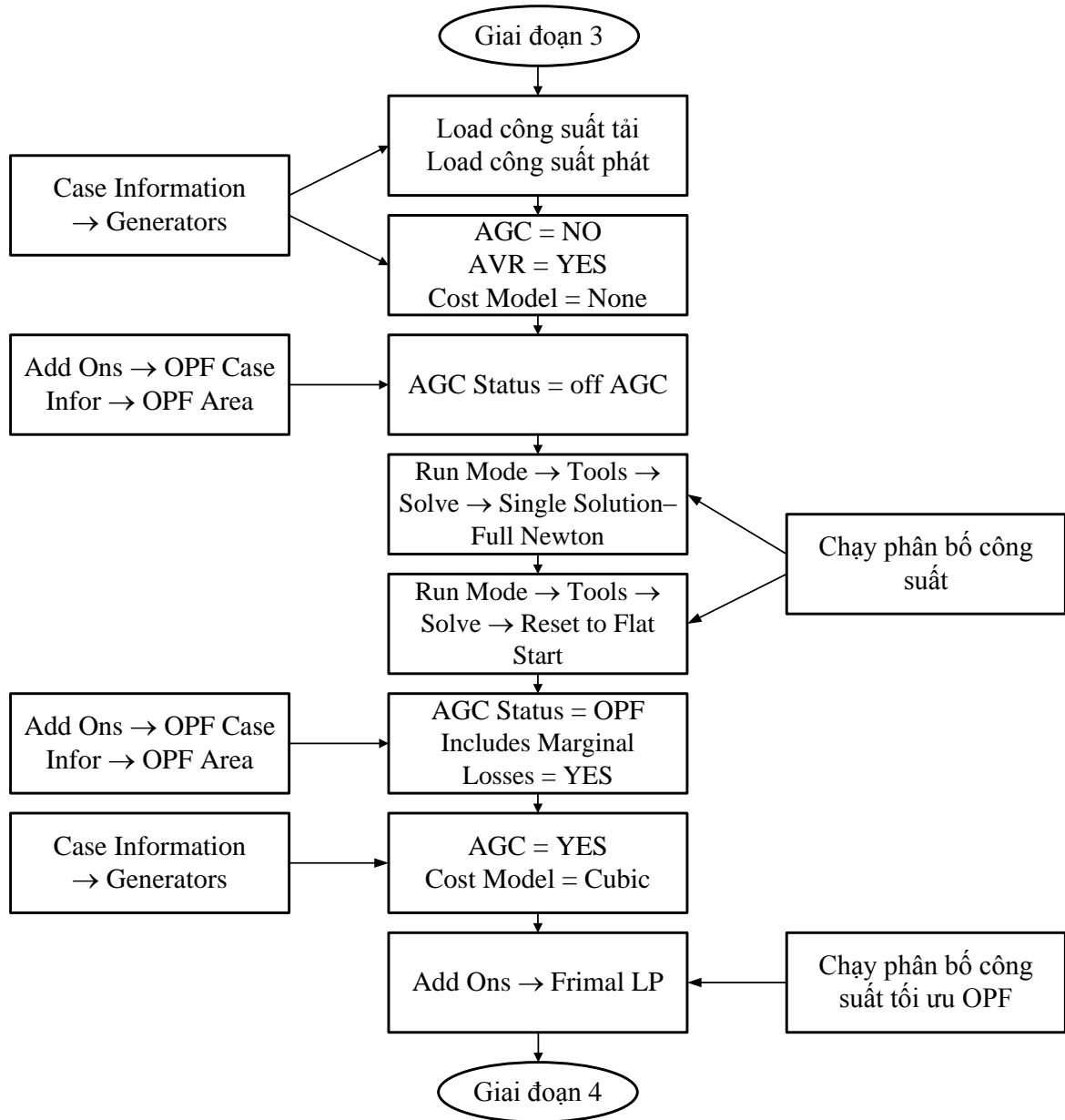
Giai đoạn 1: Thực hiện cài đặt các thông số cho các mô hình trong hệ thống điện.



- *Giai đoạn 2: Thực hiện kích hoạt các thông số trong mô hình hệ thống điện.*



▪ *Giai đoạn 3: Chạy phân bố công suất tối ưu OPF.*



- *Giai đoạn 4: Mô phỏng ổn định quá độ, đánh giá ổn định/không ổn định và lấy mẫu.*

