

**LỰA CHỌN BIẾN TRONG CHẨN ĐOÁN
ỔN ĐỊNH ĐỘNG HỆ THỐNG ĐIỆN DÙNG MẠNG NEURAL**
FEATURE SELECTION FOR DYNAMIC STABILITY PREDICTION OF
POWER SYSTEM USING NEURAL NETWORK

Nguyễn Ngọc Âu, Quyền Huy Ánh, Phan Việt Thịnh
Trường Đại Học Sư Phạm Kỹ Thuật TP. Hồ Chí Minh

Ngày tòa soạn nhận bài 29/9/2015, ngày phản biện đánh giá 12/10/2015, ngày chấp nhận đăng 20/10/2015

TÓM TẮT

Bài toán chẩn đoán ổn định động hệ thống điện gặp phải vấn đề số biến lớn của tập dữ liệu, nhưng không phải tất cả các biến đều hữu ích. Các biến thừa sẽ gây nhiễu làm giảm tính năng của bộ phân lớp hay chẩn đoán. Lựa chọn biến nhằm mục đích chọn số ít biến đặc trưng cho bộ phân lớp giúp cải thiện độ chính xác nhận dạng. Bài báo đề nghị áp dụng giải thuật Relief để lựa chọn biến và so sánh với phương pháp chọn biến sử dụng hai hàm khoảng cách Fisher và Divergence. Hai bộ nhận dạng được đề xuất để đánh giá độ chính xác nhận dạng là mạng neural GRNN (Generalized Regression Neural Network) và MLPNN (Multilayer Perceptron Neural Network). Kết quả kiểm tra trên sơ đồ IEEE 39-bus cho thấy giải thuật chọn biến Relief với bộ nhận dạng GRNN cho kết quả có số biến nhỏ hơn và độ chính xác kiểm tra cao hơn các phương pháp còn lại. Giải thuật Relief cho số biến giảm đáng kể trong khi độ chính xác nhận dạng được nâng cao hơn so với toàn tập biến.

Từ khóa: chẩn đoán ổn định động, hệ thống điện, mạng neural nhân tạo, lựa chọn biến.

ABSTRACT

Dynamic stability prediction of power system faces with a large number of features, but not all features are useful. The redundant features will cause noise and reduce the performances of classifier. Feature selection aims to select a feature subset for classifier and improve recognition accuracy. This paper suggests the application of Relief algorithm for the feature selection. It is compared with two methods of feature selection that are Fisher discrimination and Divergence. Two recommended models for recognition accuracy are GRNN (Generalized Regression Neural Network) and MLPNN (Multilayer Perceptron Neural Network). Testing results on IEEE 39-bus diagram show that the Relief algorithm with GRNN has yielded the results with smaller number of features and higher accuracy prediction than the other. Relief algorithm has significantly reduced the number of features while the recognition accuracy has been improved over all features.

Keywords: Dynamic stability prediction, power system, neural network, feature selection.

1. GIỚI THIỆU

Hệ thống điện (HTĐ) hiện đại bị áp lực vận hành rất gần với giới hạn biên ổn định, trong khi đó HTĐ luôn phải đối mặt với các kích động bất thường dễ gây ra mất ổn định. Do tính chất phi tuyến của hành vi HTĐ rất cao, các phương pháp phân tích truyền thống tốn nhiều thời gian giải nên gây chậm trễ trong việc ra quyết định. Vì vậy, phát hiện nhanh

mất ổn định HTĐ giúp hệ thống điều khiển ra quyết định kịp thời trở thành yếu tố then chốt đảm bảo vận hành HTĐ ổn định.

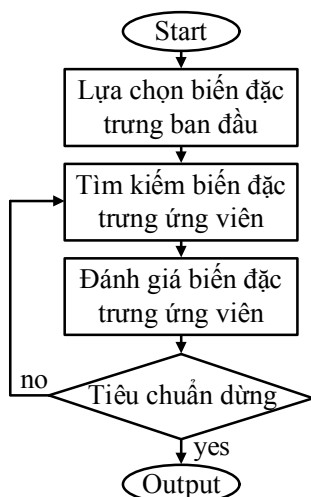
Trong những năm gần đây, Artificial Neural Network (ANN) được áp dụng như là một phương pháp thay thế để giải quyết những vấn đề khó khăn mà những phương pháp phân tích truyền thống không giải quyết được về

tốc độ tính toán cũng như hiệu suất [1]. Bằng quá trình học cơ sở dữ liệu, mỗi quan hệ phi tuyến vào ra giữa những thông số vận hành HTĐ và tình trạng ổn định có thể tính toán nhanh chóng [2]. Đây là đặc điểm quan trọng, đặc trưng của ANN. Tuy nhiên, ANN làm việc nhanh và hiệu quả thì số biến đầu vào phải nhỏ. Vì vậy, những biến được chọn phải đặc trưng, loại bỏ các biến thừa, gây nhiễu, điều này còn giúp giảm số cảm biến đo lường, giảm thời gian tính toán, và nâng cao độ chính xác nhận dạng.

Trong bài báo này, chúng tôi đề nghị áp dụng giải thuật Relief để chọn biến đầu vào đặc trưng cho ANN trong chẩn đoán ổn định động HTĐ. Giải thuật Relief được so sánh với phương pháp chọn biến sử dụng hai hàm khoảng cách Fisher và Divergence. Kết quả kiểm tra trên sơ đồ IEEE 39-bus với hai bộ nhận dạng GRNN và MLPNN, cùng số biến đặc trưng, giải thuật Relief với bộ nhận dạng GRNN cho kết quả có số biến nhỏ hơn và độ chính xác kiểm tra cao hơn các phương pháp còn lại.

2. LỰA CHỌN BIẾN ĐẶC TRƯNG

Trong các giai đoạn thiết kế mô hình chẩn đoán ổn định động HTĐ dùng ANN, lựa chọn biến đặc trưng là rất quan trọng vì nó ảnh hưởng trực tiếp đến thời gian huấn luyện và độ chính xác chẩn đoán của mô hình. Biến đặc trưng được lựa chọn cần có độ tách biệt dữ liệu giữa lớp ổn định và không ổn định cao sẽ giúp nâng cao độ chính xác nhận dạng.



Hình 1. Quy trình lựa chọn biến đặc trưng

2.1. Quy trình lựa chọn biến đặc trưng

Quy trình được chia thành 4 bước, được trình bày như Hình 1.

2.1.1. Lựa chọn biến đặc trưng ban đầu

Bước này xác định một tập biến đặc trưng đại diện cho cơ sở dữ liệu để huấn luyện ANN. Những biến đặc trưng ban đầu này biểu thị những thông số vận hành HTĐ và bao trùm các trạng thái vận hành của HTĐ. Biến đặc trưng của HTĐ ở chế độ quá độ hay chế độ động là độ thay đổi công suất máy phát, độ thay đổi công suất tải, độ thay đổi công suất trên các đường dây truyền tải, độ sụt điện áp tại các nút,... ngay trong thời điểm xảy ra sự cố.

2.1.2. Tìm kiếm biến đặc trưng ứng viên

Bước này gồm quá trình tìm kiếm những tập con biến ứng viên cho đầu vào bộ phân lớp. Nó chi phối về thời gian tính toán của toàn bộ quá trình lựa chọn biến. Các chiến lược tìm kiếm gồm tìm kiếm tối ưu toàn cục, tìm kiếm tối ưu cục bộ và dựa trên kinh nghiệm. Tìm kiếm tối ưu toàn cục có khả năng tìm ra các tập con biến tốt nhất. Tuy nhiên, chi phí tính toán của nó là rất lớn nên không khả thi trong thực hiện với tập biến có số biến ban đầu lớn. Tìm kiếm tối ưu cục bộ có chi phí tính toán thấp, khả thi trong thực hiện, nên được nhiều công trình đề nghị áp dụng [2,3,4]. Xếp hạng là một phương pháp đơn giản nhất trong tìm kiếm tối ưu cục bộ, thực thi nhanh do chỉ tính toán dựa trên độ đo hay tiêu chuẩn cho đơn biến. Trong phương pháp này, các biến đặc trưng ban đầu được đánh giá tiêu chuẩn cho từng đơn biến. Kết quả được xếp hạng thứ tự quan trọng các biến từ cao đến thấp và chọn ra số biến đặc trưng yêu cầu.

1.1.1. Đánh giá biến đặc trưng ứng viên

Trong bài báo này, biến đặc trưng ứng viên được đánh giá qua giải thuật Relief và phương pháp chọn biến sử dụng hai hàm khoảng cách Fisher và Divergence. Dữ liệu được chuẩn hoá trước khi thực hiện.

1.1.2. Tiêu chuẩn dừng

Thông qua thực nghiệm huấn luyện nhiều lần để tìm tập con biến với độ chính xác nhận dạng kỳ vọng. Việc tìm kiếm và đánh giá biến sẽ dừng khi số biến đạt giá trị yêu cầu và độ chính xác nhận dạng cao nhất, đạt giá trị kỳ vọng.

1.1. Các phương pháp xếp hạng

1.1.1. Hàm khoảng cách Fisher

Dựa trên hàm khoảng cách tuyến tính Fisher tìm kiếm độ tách biệt tối ưu tuyến tính giữa hai lớp dữ liệu. Một tập dữ liệu D gồm n mẫu huấn luyện 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 $F(w)$ lớn hơn thì khả năng tách biệt dữ liệu giữa 2 lớp cao hơn [2,3].

$$F(w) = \frac{|m_1 - m_2|^2}{\sigma_1^2 + \sigma_2^2} \quad (1)$$

Trong đó: m_1 là giá trị trung bình của lớp C_1 và σ_2^2 là phương sai của lớp C_2 .

1.1.2. Hàm khoảng cách Divergence

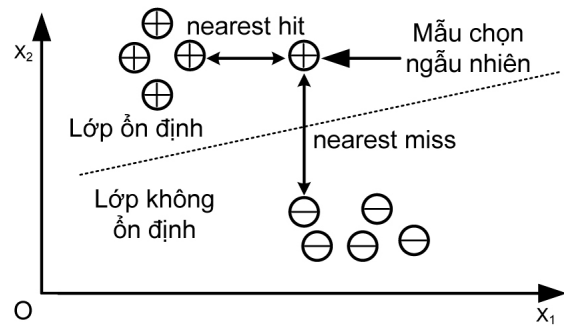
Divergence là tiêu chuẩn đo lường độ phân tán giữa hai lớp dữ liệu. Khoảng cách D giữa 2 lớp được trình bày ở biểu thức (2). Biến có khoảng cách D càng lớn thì dữ liệu giữa 2 lớp càng phân tán [2,4].

$$D = \frac{1}{2} \left(\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} - 2 \right) + \frac{1}{2} (m_1 - m_2)^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \quad (2)$$

Trong đó: σ_1^2, σ_2^2 là phương sai của lớp 1 và lớp 2 của biến; m_1, m_2 là trung bình của lớp 1 và lớp 2 của biến.

1.1.3. Giải thuật Relief

Giải thuật Relief đánh giá thống kê giá trị trọng số của biến dựa trên mẫu. Các biến tốt là biến có các mẫu cùng lớp gần nhau hơn, các mẫu không cùng lớp thì tách biệt hơn [2,5].



Hình 2. Mô tả giải thuật Relief

Ý tưởng của giải thuật Relief được trình bày ở hình 2 [5]. Cho tập học M mẫu gồm S mẫu ổn định và U mẫu không ổn định, chọn ngẫu nhiên x_k ($k = 1, 2, \dots, m$) có m mẫu học từ M mẫu, ($m \leq M$). Relief tính toán khoảng cách nearest hit x_H giữa mẫu gần nhất trong x_k đến mẫu của lớp ổn định S và khoảng cách nearest miss x_M giữa mẫu gần nhất trong x_k đến mẫu của lớp không ổn định U . Sau đó giải thuật cập nhật giá trị trọng số cho tất cả các biến F_i ($i = 1, 2, \dots, N$). Trọng số của các biến phụ thuộc vào giá trị x_k, x_H và x_M . Nếu x_k và x_H của biến F_i có giá trị sai lệch thì hai mẫu cùng lớp là tách biệt, đây là giá trị không mong muốn nên trọng số sẽ giảm. Ngược lại, nếu x_k và x_M của biến F_i có giá trị sai lệch thì hai mẫu khác lớp là tách biệt, đây là giá trị mong muốn nên trọng số sẽ tăng. Toàn bộ quy trình được lặp lại m lần.

Giải thuật Relief cơ bản như sau:

Input: Tập biến $F[N, M]$ có N biến M mẫu

Thiết lập số mẫu ngẫu nhiên $x_k, k=[1, m], m \leq M$

Output: Vec-tơ trọng số $W[i]$ của biến F_i :

$$-1 \leq W[i] \leq 1$$

$$W[i] = 0.0$$

for $k = 1 : m$

for $i = 1 : N$

$$W[i] = W[i] - \text{diff}(i, x_k, x_H) / m + \text{diff}(i, x_k, x_M) / m$$

end

end

Hàm $\text{diff}(i, x_j, x_k)$ tính toán khoảng cách giữa mẫu x_k và x_j của biến F_i , được tính theo công thức (3).

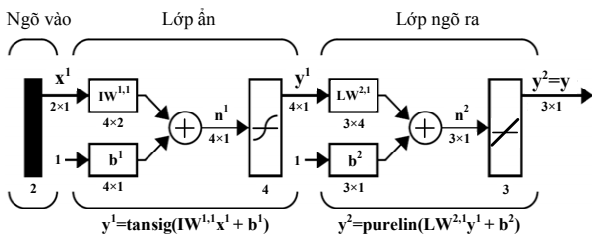
$$\text{diff}(i, x_j, x_k) = \frac{|x_{j i} - x_{k i}|}{\max(F_i) - \min(F_i)} \quad (3)$$

Trong đó: $\max(F_i)$ và $\min(F_i)$ là giá trị mẫu lớn nhất và nhỏ nhất của biến i .

3. MẠNG MLPNN VÀ GRNN

3.1. MLPNN

MLPNN cơ bản như Hình 3.1 gồm có lớp ngõ vào, một hay nhiều lớp ẩn và lớp ngõ ra.



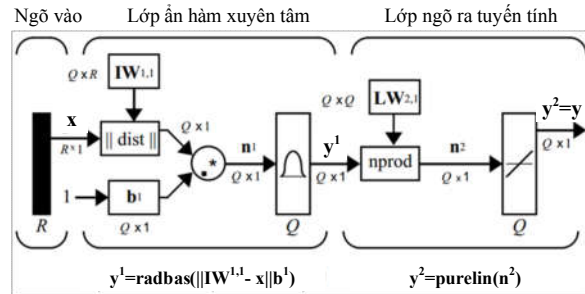
Hình 3. Mạng MLPNN

Số lớp ẩn có thể thay đổi trong quá trình huấn luyện tùy thuộc dữ liệu bài toán. Các nút ngõ ra có giá trị thay đổi phụ thuộc vào lớp ngõ ra mục tiêu. 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 [6]. Có nhiều giải thuật huấn luyện, trong đó thuật toán Levenberg-Marquardt có hiệu suất cao hơn và tốc độ hội tụ nhanh hơn [6]. Thông số mạng ban đầu là ngẫu nhiên, quá trình huấn luyện dừng khi số vòng lặp đạt giá trị tối đa, sai số đạt một mức cho phép hoặc sai số không tăng nữa.

3.2. GRNN

GRNN là biến thể của RBFN (Radial Basis Function Network), là công cụ mạnh trong ứng dụng cho bài toán nhận dạng. So với MLPNN, GRNN yêu cầu nhiều số neural hơn nhưng thời gian huấn luyện ít hơn [7]. GRNN cơ bản như Hình 4 gồm lớp ngõ vào, lớp ẩn hàm xuyên tâm và lớp ngõ ra tuyến tính.

GRNN giống như RBFN nhưng có chút ít thay đổi hơi đặc biệt là ở lớp hàm tổng (summation layer).



Hình 4. Mạng GRNN

Lớp đầu vào kết nối với lớp ẩn, trong đó mỗi neural đại diện cho một mẫu huấn luyện và đầu ra của nó là khoảng cách của biến đầu vào và tâm. Mỗi neural của lớp ẩn kết nối với hai neural của lớp đặc biệt là neural S (S-summation) và neural D (D-summation). Neural S tính toán tổng trọng số đầu ra của lớp ẩn, trong khi neural D tính đầu ra phi trọng số của lớp ẩn.

Trọng số kết nối của neural lớp ẩn và neural S là giá trị ngõ ra mục tiêu tương ứng với mẫu đầu vào. Đối với neural D, trọng số kết nối là đơn vị. Giá trị tính toán của lớp đầu ra tương ứng với vec-tơ mẫu đầu vào là tỉ số của mỗi đầu ra S với mỗi đầu ra D. Vận hành của neural D kể đến thông số spread mà giá trị tối ưu được xác định qua thực nghiệm thử sai [8].

3.3. Luật đầu ra của bộ phân loại

Kết quả tính toán ngõ ra của bộ phân loại có sai số với giá trị nhị phân của đầu ra đã gán nhãn. Ngõ ra được quy định như các trường hợp sau.

- Trường hợp 1: Nếu mã hóa ngõ ra $\{1\}$ là ‘ổn định’, và $\{0\}$ là ‘không ổn định’ thì:
 Nếu $y > 0.5 \Rightarrow y = 1$ ổn định
 Nếu $y \leq 0.5 \Rightarrow y = 0$ không ổn định (4)
- Trường hợp 2: Ma trận ngõ ra M lớp như sau:

$$F_1(X)$$

$$F_1(X) = F_2(X) \quad (5)$$

.....

$$F_M(X)$$

Ngõ ra tuân theo luật số lớn [9], kết quả nhận dạng được xếp vào lớp thứ k nếu:

$$F_k(x) > F_j(x) \forall j \neq k \quad (6)$$

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

3.4. Huấn luyện và đánh giá mô hình nhận dạng

Quá trình huấn luyện và kiểm tra được thực hiện k lần. Phần trăm độ chính xác nhận dạng của mô hình trong huấn luyện hoặc kiểm tra được tính trung bình trong k lần thực hiện. Hiệu suất của bộ phân loại được đánh giá theo tỷ lệ phần trăm huấn luyện đúng hoặc kiểm tra đúng và được xác định theo phương trình (7).

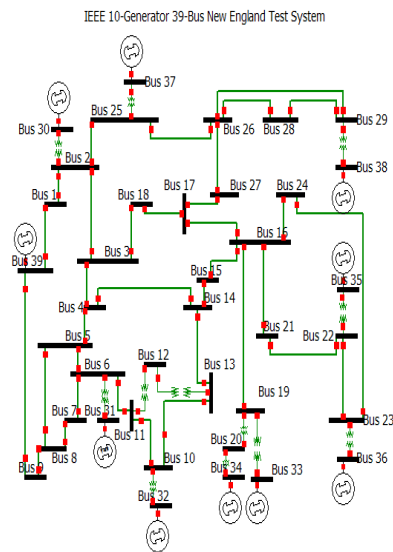
$$\text{Classification_Rate\%} = \frac{R}{S} \cdot 100 \quad (7)$$

Trong đó: R là số mẫu đúng, S là tổng số mẫu. Giá trị kỳ vọng độ chính xác nhận dạng của mô hình được đề nghị trong [10] phải đạt hơn 90%.

4. ÁP DỤNG ĐÁNH GIÁ ỔN ĐỊNH ĐỘNG HTĐ IEEE 10 MÁY 39-BUS

4.1. Sơ đồ IEEE 10 máy 39-bus

HTĐ IEEE 39-bus New England là một HTĐ tương đương của các hệ thống con của vùng đông bắc Hoa Kỳ và Canada [11]. Hệ thống gồm có 39 bus, trong đó 10 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 31 được coi là bus Slack, 9 bus được gọi là bus PV, 29 bus còn lại được gọi là bus PQ, có 2 cấp điện áp là 345kV và 20kV. Hệ thống được cho như ở Hình 5.



Hình 5. HTĐ IEEE 39-bus New England

4.2. Tạo cơ sở dữ liệu

Mô phỏng để tạo cơ sở dữ liệu với sự trợ giúp của phần mềm PowerWorld được thực hiện trên hệ thống IEEE 10 máy 39-bus, xét sự cố ngắn mạch 3 pha cân bằng tại các bus và dọc các đường dây truyền tải ở những vị trí 25%, 50% và 75% đường dây, với các mức tải từ 20% đến 120% tải cơ bản. Thuộc tính động của HTĐ trong quá trình ngắn mạch dựa trên quan sát mối quan hệ giữa các góc công suất của các máy phát điện. HTĐ ổn định nếu bất kỳ góc công suất tương đối của máy phát thứ i δ_i so với máy phát thứ j δ_j không vượt quá 180° . Ngược lại, nếu góc công suất tương đối vượt quá 180° thì HTĐ mất ổn định. Các mẫu dữ liệu thu thập tương ứng được xếp vào lớp ổn định và không ổn định.

Kết quả mô phỏng thu được tập mẫu gồm 1819 mẫu, trong đó 1436 mẫu ổn định, 383 mẫu không ổn định. Toàn bộ dữ liệu được chuẩn hoá trước khi huấn luyện.

4.3. Biến đầu vào và biến đầu ra

Mỗi mẫu dữ liệu trong tập mẫu học được biểu diễn dưới dạng vec-tơ bao gồm những biến đầu vào là độ lệch công suất phát, độ lệch công suất tải, độ sụt điện áp tại các bus, độ lệch công suất phân bố trên các đường dây truyền tải, $x = \{\Delta P_{gen}, \Delta Q_{gen}, \Delta V_{bus}, \Delta P_{load}, \Delta Q_{load}, \Delta P_{flow}, \Delta Q_{flow}\}$. Biến đầu ra y được gán nhãn cho lớp 1 hay lớp ‘ổn định’, cho lớp 2 hay

lớp ‘không ổn định’. Do sai số trong tính toán, ngõ ra của bộ nhận dạng sẽ được bầu chọn theo luật số lớn như công thức (6).

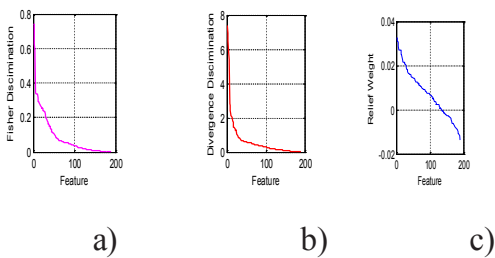
Tổng số biến đầu vào là 189 biến và 2 biến đầu ra.

4.4. Phân chia dữ liệu

Tập dữ liệu được chia ngẫu nhiên thành 6 tập con, kích thước bằng nhau. Mỗi tập gồm 303 mẫu (239 mẫu ổn định, 64 mẫu không ổn định). Như vậy, mỗi tập huấn luyện có 1516 mẫu (1197 mẫu ổn định, 319 mẫu không ổn định) và tập kiểm tra có 303 mẫu (239 mẫu ổn định, 64 mẫu không ổn định).

4.5. Kết quả tính toán xếp hạng biến

Kết quả tính toán và xếp hạng biến của các hàm Fisher, Divergence, Relief được trình bày ở Hình 6.



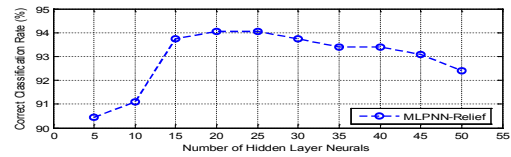
a) Fisher b) Divergence c) Relief

Hình 6. Kết quả xếp hạng biến

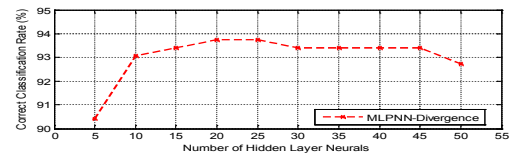
4.6. Kết quả nhận dạng chính xác

MLPNN được chọn có 3 lớp là lớp đầu vào, lớp ẩn và lớp ra. Giải thuật học và cập nhật trọng số được chọn là thuật toán Levenberg-Marquardt. Hàm kích hoạt lớp ẩn và lớp ngõ ra là hàm tansig và purelin.

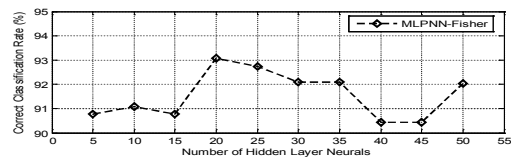
Việc huấn luyện mạng là công việc khó khăn, trong đó việc chọn số neural trong lớp ẩn cho MLPNN và chọn hệ số spread cho GRNN là rất quan trọng, vì nó ảnh hưởng rất lớn đến độ chính xác. Bằng phương pháp thực nghiệm thử sai huấn luyện nhiều lần với số neural ẩn và giá trị hệ số spread khác nhau để tìm kết quả tốt nhất. Kết quả thực thi được trình bày ở hình 7, 8, 9, 10, 11 và 12.



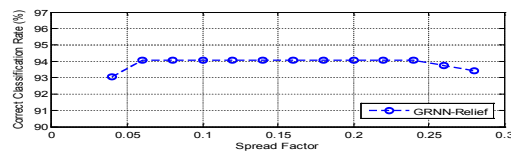
Hình 7. Kết quả huấn luyện với số neural ẩn thay đổi của MLPNN-Relief



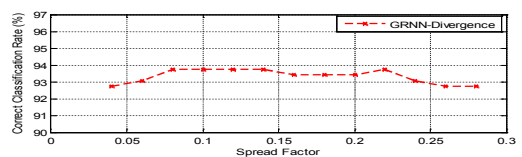
Hình 8. Kết quả huấn luyện với số neural ẩn thay đổi của MLPNN-Divergence



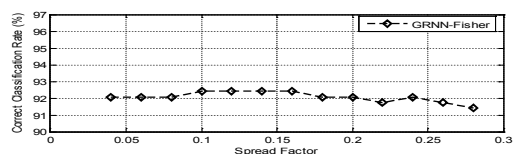
Hình 9. Kết quả huấn luyện với số neural ẩn thay đổi của MLPNN-Fisher



Hình 10. Kết quả huấn luyện với hệ số spread thay đổi của GRNN-Relief

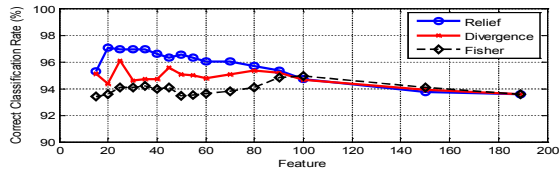


Hình 11. Kết quả huấn luyện với hệ số spread thay đổi của GRNN-Divergence

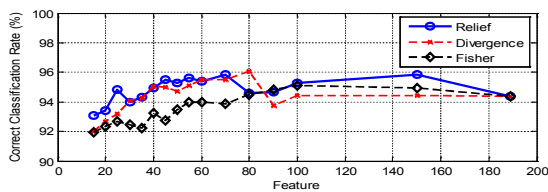


Hình 12. Kết quả huấn luyện với hệ số spread thay đổi của GRNN-Fisher

Hình 13 và hình 14 trình bày kết quả so sánh độ chính xác nhận dạng của các phương pháp Relief, Divergence và Fisher sử dụng mạng GRNN và MLPNN.



Hình 13. So sánh độ chính xác nhận dạng cho các phương pháp sử dụng GRNN



Hình 14. So sánh độ chính xác nhận dạng cho các phương pháp sử dụng MLPNN

Kết quả so sánh độ chính xác nhận dạng của hai mạng GRNN và MLPNN tại số biến là 15 với các phương pháp Relief, Divergence, Fisher trình bày ở Bảng 1. So sánh độ chính xác nhận dạng của GRNN-Relief tại 15 biến và 189 biến trình bày ở Bảng 2. So sánh thời gian huấn luyện MLPNN-Relief và GRNN-Relief được trình bày ở Bảng 3.

Bảng 1. Độ chính xác nhận dạng kiểm tra của GRNN và MLPNN

Mạng	Số biến	Relief	Divergence	Fisher
GRNN	15	95.27%	95.10%	93.40%
MLPNN	15	93.07%	92.02%	91.97%

Bảng 2. So sánh độ chính xác nhận dạng của GRNN-Relief tại 15 biến và 189 biến

Số biến	15	189
GRNN-Relief	95.27%	93.56%

Bảng 3. So sánh thời gian huấn luyện MLPNN-Relief và GRNN-Relief

Số biến	MLPNN-Relief	GRNN-Relief
189	121.73s	1.49s
15	5.6s	0.84s

5. BÀN LUẬN

Hình 7 đến 9 và Hình 10 đến 12 cho độ chính xác nhận dạng cao nhất tại tập biến có số biến là 15, số neural ẩn của MLPNN là 20 và hệ số spread của GRNN là 0.1. Kết quả thấy giải thuật chọn biến Relief với bộ nhận dạng GRNN cho kết quả số biến nhỏ hơn, độ chính xác kiểm tra cao hơn các phương pháp còn lại.

Kết quả từ Hình 13, 14 và Bảng 1 cho thấy:

- Đối với GRNN: phương pháp Relief, Divergence và Fisher, khi số biến đặc trưng là 15 biến, độ chính xác nhận dạng của GRNN lần lượt là 95.27%, 95.10% và 93.40%.
- Đối với MLPNN: phương pháp Relief, Divergence và Fisher, khi số biến đặc trưng là 15 biến, độ chính xác nhận dạng của MLPNN lần lượt là 93.07%, 92.02% và 91.97%.
- Như vậy, GRNN-Relief cho độ chính xác nhận dạng kiểm tra cao hơn GRNN-Divergence và GRNN-Fisher là 0.17% và 1.87%. MLPNN-Relief cho độ chính xác nhận dạng kiểm tra cao hơn MLPNN-Divergence và MLPNN-Fisher là 1.05% và 1.1%.

Bảng 2, GRNN-Relief với 15 biến thì độ chính xác kiểm tra là 95.27%, trong khi 189 biến độ chính xác kiểm tra là 93.56%, hay số biến giảm đến 12.6 lần, độ chính xác nâng cao 1.71%, điều này cho thấy áp dụng giải thuật Relief loại được các biến thừa, gây nhiễu, nâng cao độ chính xác.

Bảng 3, GRNN-Relief có thời gian huấn luyện nhanh hơn đáng kể so với MLPNN-Relief, cụ thể tại số biến 189 và 15 biến thì thời gian huấn luyện giảm 81.6 và 6.6 lần. Kết quả này thể hiện rõ sự khác biệt trong cơ chế học của GRNN và MLPNN. MLPNN với giải thuật lan truyền ngược phải tính sai số qua nhiều vòng lặp mới kết thúc quá trình huấn luyện, trong khi GRNN chỉ cần tính tâm của

hàm bán kính là cho ra kết quả ngõ ra. Vì vậy, GRNN có thời gian huấn luyện nhanh hơn rất nhiều so với MLPNN.

6. KẾT LUẬN

Bài báo đề nghị áp dụng giải thuật Relief để lựa chọn biến và so sánh với phương pháp chọn biến sử dụng hai hàm khoảng cách Fisher và Divergence trong chẩn đoán nhanh ổn định động HTĐ.

Kết quả kiểm tra với hai bộ nhận dạng GRNN và MLPNN với cùng số biến đặc trưng, GRNN-Relief cho độ chính xác cao hơn các phương pháp chọn biến còn lại. GRNN-Relief với số biến là 15 biến giảm đến 12.6 lần, độ chính xác nhận dạng nâng cao 1.71%, thời gian huấn luyện giảm 81.6 lần, điều này cho thấy giải thuật Relief chọn được các biến đặc trưng nhất, loại biến nhiễu giúp nâng cao độ chính xác, rút ngắn thời gian huấn luyện.

TÀI LIỆU THAM KHẢO

- [1] M. R. Aghamohammadi, F. Mahdavizadeh, R. Bagheri, *Power System Dynamic Security Classification Using Kohonen Neural Networks*, 2009 IEEE, pp.1-7.
- [2] Rui Zhang, Yan Xu, Zhao Yang Dong and David J Hill, *Feature Selection For Intelligent Stability Assessment of Power Systems*, 2012 IEEE, pp.1-7.
- [3] C. A. Jensen and M. A. El-Sharkawi, *Power system security assessment using neural networks: feature selection using Fisher discrimination*, IEEE Trans. Power Systems, vol. 16, Nov. 2001, pp.757-763.
- [4] K. Niazi, C. Arora, and S. Surana, *Power system security evaluation using ann: feature selection using divergence*, Electric Power Systems Research, vol. 69, Feb. 2004, pp.161-167.
- [5] Taylor, *Computational methods of feature selection*, Edited by Huan Liu, Horoshi Motoda, 2008, pp.169-176.
- [6] I. S. Isa, Z. Saad, S. Omar, M. K. Osman, K. A. Ahmad, *Suitable MLP Network Activation Functions for Breast Cancer and Thyroid Disease Detection*, 2010 IEEE, pp.39-44.
- [7] Matlab R2011b, *Neural Network Toolbox User's Guide*.
- [8] Behnam Ababaei, Teymour Sohrabi and Farhad Mirzaei, *Assessment of radial basis and generalized regression neural networks in daily reservoir inflow simulation*, Elixir Comp. Sci. & Engg. 42, 2012, pp.6074-6077.
- [9] Simon S. Haykin, *Neural Networks A Comprehensive Foundation*, 2nd Edition, 1999, pp.84-86.
- [10] Quyen Huy Anh, *The application of pattern recognition methods for fast analysis of the dynamic stability of electrical power system*, Electrical technology, No. 2, pp.1-13, 1994, Pergamon.
- [11] Hussain Hassan Al Marhoon, *A Practical Method for Power Systems Transient Stability and Security*, B.S. University of New Orleans, 2008, pp.56-60.
- [12] Prabha Kundur, *Power System Stability and Control*, New York: McGraw Hill, 1994, pp.17-39.