HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

NGUYỄN THỊ VÂN ANH

SỬ DỤNG MẠNG NƠ RON TRONG KHAI PHÁ DỮ LIỆU

CHUYÊN NGÀNH : TRUYỀN DỮ LIỆU VÀ MẠNG MÁY TÍNH
MÃ SỐ: 60.48.15

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

NGƯỜI HƯỚNG DẪN KHOA HỌC : ***TS. VŨ VĂN THỎA***

Hà Nội - 2010

MỞ ĐẦU

Trong môi trường cạnh tranh người ta ngày càng cần có nhiều thông tin với tốc độ nhanh để trợ giúp việc ra quyết định và ngày càng nhiều câu hỏi mang tính chất định tính cần phải trả lời dựa trên một khối lượng khổng lồ dữ liệu đã có. Do đó thực tế đã làm phát triển một khuynh hướng kỹ thuật mới làm sao để khai thác tốt cơ sở dữ liệu trong các doanh nghiệp đó là kỹ thuật khai phá dữ liệu (data mining), vận dụng kỹ thuật đó thì dữ liệu giao dịch đóng một vai trò rất quan trọng cho việc hoạch định kế hoạch kinh doanh trên thương trường vào những năm tiếp theo. Kỹ thuật này đã được sử dụng tại nhiều nơi và đã cho kết quả khả quan trong nhiều tổ chức trong và ngoài nước và trên thế giới.

Chương 1: Tổng quan về khai phá dữ liệu

Chương 2: Mạng nơ ron nhân tạo (Artificial neural network)

Chương 3: Ứng dụng mạng nơ ron nhân tạo cho bài toán dự đoán phụ tải điện.

Kết luận và hướng nghiên cứu tiếp theo.

Các từ khóa: Khai phá dữ liệu (datamining), học máy (machine learning), mạng nơ ron (neural network), MLP (Multi-layer Perceptron), SOM (Self-organizer map).

asymmetric fuzzy weight - Decision Support Systems, Vol 24, 1998, 105-126 p.

[12] Rachel Konrad, *Data mining: Digging user info for gold*, ZDNET News, February 7, 2001,

[13] Rekesh Arawal, Ramakrishnan Srikant, *Fast Algorithms for Mining Association*, IBM Almadem Research Center 650 Harry Road, San Jose, CA 95120.

[14] Stuart Russell and Peter Norvig, *Artificial Intelligence - A Modern Approach.* @2003, 1995 by Pearson Education, Inc.

[15] Trần Bách, *Lưới điện và hệ thống điện*. NXB Khoa học và kỹ thuật

[16] The Gartner Group,

[17] Zhe Liao, Jun Wang - Forecasting model of global stock index by stochastic time effective neural network- Expert Systems with Application, Vol.37 (2010), 834-841.

TÀI LIỆU THAM KHẢO

[1] Ben Krose and Patrick van der Smagt, *An Introduction to Neural Networks*, @1996 University of Amsterdam.

[2] David Hand, Heikki Mannila, and Padhraic Smyth, *Principles of Data Mining*, MIT Press, Cambridge, MA, 2001.

[3] Daniel T. Larose, *Discovering Knowledge in Data: An Introduction to Datamining*, NXB. Wiley Interscience.

[4] The Gartner Group,

[5] Joseph. P. Bigus, *Datamining with Neural Network*, @1996 by The McGraw-Hill Companies, Inc.

[6] Lã Văn Út, *Phân tích và điều khiển ổn định Hệ thống điện*, NXB Khoa học và kỹ thuật

[7] Martin T. Hagan, Howard B. Demuth, *Neural Network Design,* copyright@1996 by PWS Publishing Company, USA.

[8] Mehmed Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, ©2003 by John Wiley & Sons.

[9] Mehdi Khashei, Mehdi Bijari - *An Artificial neural network (p, d, q) model for timeseries forecasting* - Expert Systems with Application, Vol. 37(2010) 479-489

[10] M. Becvali, M.Cellura, V. Lo Brano, *A. Marvuglia-Forecasting daily urban electric load using artificial neural networks* - Energy Conversion and Management Vol. 45 (2004) 2879-2900 p.

[11] R. J. Kuo, K. C. Xue - *A decision support system for sales forecasting through fuzzy neural networks with*

Chương 1. Tổng quan về khai phá dữ liệu 1.1 Khái niệm

Theo Gartner Group [4] "*Khai phá dữ liệu là quá trình khám phá các tương quan, mẫu và các xu thế mới có ý nghĩa bằng việc dịch chuyển thông qua lượng lớn các dữ liệu được lưu trữ, và sử dụng các công nghệ nhận dạng mẫu cũng như các công nghệ thống kê, toán*".

Ta có thể phân khai phá dữ liệu thành một trong hai loại sau:

1. Khai phá dữ liệu có tính dự đoán: tức là là sản xuất ra mô hình của hệ thống được mô tả bởi tập dữ liệu được cho.
2. Khai phá dữ liệu có tính mô tả: tức là sản xuất ra thông tin mới, không tầm thường dựa trên tập dữ liệu có sẵn.

1.2 Các nhiệm vụ của khai phá dữ liệu

1.2.1 ***Mô tả:*** Đôi khi, các nhà nghiên cứu và phân tích đơn giản là cố gắng tìm cách mô tả các mẫu và các xu thế nằm trong dữ liệu. Các mô hình khai phá dữ liệu nên là minh bạch ở mức có thể.

1.2.2 ***Ước lượng:*** Ước lượng tương tự như phân loại trừ việc biến mục đích là số chứ không phải là loại.

1.2.3 ***Dự đoán***: Dự đoán giá cả thị trường ba tháng trong tương lai, dự đoán tăng phần trăm trong tai nạn giao thông năm tiếp theo nếu giới hạn tốc độ được tăng lên

1.2.4 ***Phân loại:*** Trong phân loại, có một biến loại mục đích, như là mức thu nhập, có thể được phân đoạn thành ba lớp hoặc ba loại: thu nhập cao, thu nhập giữa và thu nhập thấp

1.2.5 ***Phân cụm*** (Clustering): Phân cụm nhằm vào việc nhóm các bản ghi, hoặc các trường hợp thành các lớp đối tượng tương tự.

1.2.6 ***Luật kết hợp:*** Nhiệm vụ kết hợp cho khai phá dữ liệu là công việc tìm kiếm các thuộc tính "đi cùng nhau", khám phá các luật cho việc xác định mối quan hệ giữa hai hoặc nhiều hơn thuộc tính.

1.3 Quá trình khai phá dữ liệu

1.3.1 *Phát biểu bài toán và đề ra giả thiết*

Trong bước này, một người lên mô hình thường xác định một tập các biến cho sự phụ thuộc không biết trước, và nếu có thể, một dạng chung của sự phụ thuộc này là một giả thiết đầu tiên.

1.3.2 *Thu thập dữ liệu*

Bước này liên quan tới việc dữ liệu được sưu tập và sinh ra như thế nào. Với hướng tiếp cận này, thì dữ liệu ngẫu nhiên sẽ được sinh, được giả thiết trong phần lớn các ứng dụng khai phá dữ liệu.

1.3.3 *Tiền xử lý dữ liệu*

Trong hướng tiếp cận quan sát, dữ liệu thường được "sưu tập" từ các cơ sở dữ liệu đang tồn tại, các kho dữ liệu, và các trung tâm dữ liệu. Tiền xử lý dữ liệu thường bao gồm ít nhất hai nhiệm vụ sau:

a) Phát hiện (và loại bỏ) các dữ liệu ngoại

như là quá trình học và nội dung học của chúng. Đến nay, có rất nhiều mô hình mạng nơ ron nhân tạo, tuy nhiên do giới hạn của luận văn chúng tôi tìm hiểu sâu về hai mô hình đó là mạng tự tổ chức (SOM) và mạng truyền thẳng đa tầng (MLP) đây là hai mô hình được sử dụng phổ biến và rộng rãi cho các bài toán có dữ liệu lớn, có độ biến thiên cao và ưu điểm của kỹ thuật này chính là khả năng tính xấp xỉ chính xác cho bất kỳ hàm cần dự đoán nào.

Cuối cùng, để ứng dụng quá trình khai phá dữ liệu sử dụng kỹ thuật mạng nơ ron, trong chương 3, chúng tôi giới thiệu bài toán phụ tải điện năng, một trong những bài toán phù hợp với việc sử dụng mạng nơ ron, với đầu ra dự đoán là phụ tải điện ngắn hạn trong vòng 24 giờ tới.

- Với việc thu thập dữ liệu 2 năm 2005 và 2006, chúng tôi đã chuẩn hóa hơn 600 vec tơ làm đầu vào cho mạng nơ ron.
- Với việc sử dụng mạng SOM, chúng tôi đã xây dựng được cấu trúc lưới SOM tối ưu nhất là 8 8, tiếp theo chúng tôi tiến hành phân cụm hiệu quả sử dụng giải thuật K-means với số cụm tốt nhất là 9.
- Để dự đoán phụ tải điện năng trong 24 giờ tới, chúng tôi sử dụng kết quả phân cụm của SOM cùng với một số dữ liệu phụ tải điện quá khứ để tiến hành huấn luyện cho một mạng truyền thẳng đa tầng (MLP) với thiết kế 50 nút tầng ẩn và 24 đầu ra cho các giá trị phụ tải điện của các giờ trong ngày.

Kết quả thu được là hết sức khả quan và có khả năng ứng dụng trong việc dự đoán phụ tải cho bên quản lý hệ thống điện miền Bắc.

và sai số giữa phụ tải dự đoán và phụ tải thực tế là rất nhỏ, trong khi đó đường phụ tải dự đoán ở khoảng từ 1h chiều đến 4h chiều có sai số là khá lớn. Điều này phần lớn xảy ra do dữ liệu phụ tải không được đúng trong khoảng đó. Ngoài ra, trong luận văn này, chúng tôi chỉ dự đoán phụ tải qua dữ liệu phụ tải điện trong quá khứ, nếu muốn được chính xác hơn ta sẽ cần thêm một số điều kiện của thời tiết như nhiệt độ, độ ẩm, … thì đường dự đoán phụ tải điện và phụ tải điện thực tế sẽ có sai số bé hơn.

Kết luận

Luận văn trình bày với cấu trúc 3 chương, với mục đích thể hiện việc khai phá dữ liệu thông qua kỹ thuật mạng nơ ron và minh họa cụ thể qua bài toán phụ tải điện năng. Các kết quả chính của luận văn là:

Ở chương 1, chúng tôi đã nghiên cứu tổng quan về khai phá dữ liệu, phân loại khai phá dữ liệu, đưa ra được các nhiệm vụ của khai phá dữ liệu đó là mô tả, ước lượng, dự đoán, phân loại, phân cụm và cuối cùng là luật kết hợp. Tiếp theo, là phân tích về quá trình khai phá dữ liệu bao gồm 5 bước: Phát biểu bài toán và đề ra giả thiết. Thu thập dữ liệu. Tiền xử lý dữ liệu. Ước lượng mô hình. Diễn giải mô hình và đưa ra kết luận.

Ở chương 2, chúng tôi đi vào nghiên cứu về kỹ thuật mạng nơ ron, là một trong các phương pháp được ứng dụng nhiều và mang lại hiệu quả cao trong các nhiệm vụ khai phá dữ liệu. Bắt đầu với việc giới thiệu về mạng nơ ron sinh học, rồi lên mô hình toán cụ thể cho một nơ ron nhân tạo, chúng tôi tìm hiểu tiếp về kiến trúc của mạng nơ ron nhân tạo gồm mạng truyền thẳng, mạng hồi quy, và khả năng học sửa lỗi của mạng nơ ron

b) Lên tỉ lệ, mã hóa, và lựa chọn các thuộc tính.

Hai lớp nhiệm vụ tiền xử lý này chỉ là các ví dụ mô tả của một phạm vi lớn các hoạt động tiền xử lý trong một quá trình khai phá dữ liệu.

1.3.4 *Ước lượng mô hình*

Sự lựa chọn và thực hiện kỹ thuật khai phá dữ liệu thích hợp là nhiệm vụ chính trong giai đoạn này. Quá trình này không dễ dàng, trong thực hành việc thực thi dựa trên một vài mô hình, và kèm theo là nhiệm vụ chọn được cái tốt. Phần đầu tiên của dữ liệu được gọi là tập học, phần tiếp theo được gọi là tập xác nhận, cũng được gọi là tập kiểm tra. Một mô hình được nhận ra thông qua quá trình khai phá dữ liệu sử dụng các kỹ thuật học quy nạp có thể được ước lượng sử dụng tham số tốc độ lỗi chuẩn như một phép đo việc thực hiện của nó.

1.3.5 *Diễn giải mô hình và đưa ra kết luận*

Trong phần lớn các bài toán, các mô hình khai phá dữ liệu hỗ trợ trong phần ra quyết định. Do vậy, các mô hình thực sự hữu ích thì cần thiết phải diễn giải được bởi vì con người không chắc chắn dựa vào các quyết định của chúng trên các mô hình "khép kín" phức tạp. Để ý rằng các đích của sự chính xác của một mô hình trái ngược với sự chính xác của sự diễn giải của nó. Thông thường, các mô hình đơn giản là diễn giải được nhiều hơn, nhưng chúng cũng kém chính xác hơn. Các phương pháp khai phá dữ liệu hiện đại được mong đợi gặt hái các kết quả chính xác cao sử dụng các mô hình có số chiều cao.

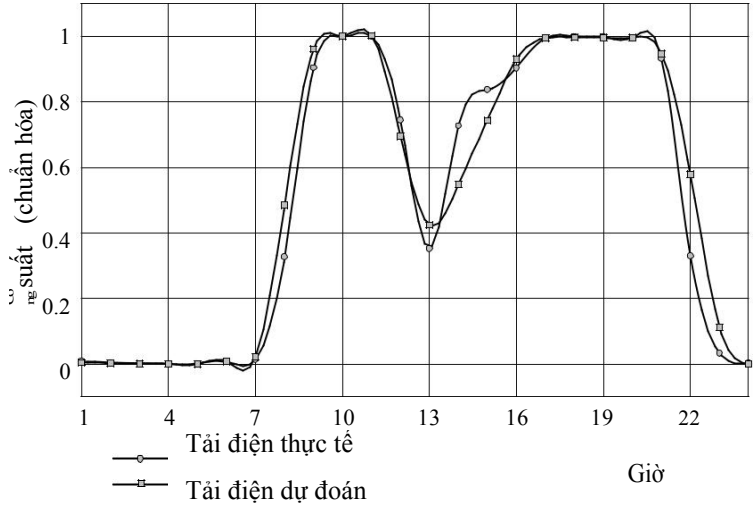
Kết luận chương

Chương 1 nêu ra khái niệm về khai phá dữ liệu, các nhiệm vụ cần thực hiện trong khai phá dữ liệu và cuối cùng là quá trình khai phá dữ liệu. Một trong các phương pháp khai phá được áp dụng nhiều trong các nhiệm vụ khai phá dữ liệu chính là kỹ thuật mạng nơ ron nhân tạo. Chúng ta sẽ tìm hiểu mạng nơ ron nhân tạo với mục đích phân cụm và dự đoán trong chương 2.

Bộ dữ liệu gồm 583 vec tơ được chia thành bộ dữ liệu, bộ dữ liệu huấn luyện bao gồm 555 vec tơ từ ngày 1/1/2005 đến 29/7/2006, bộ dữ liệu kiểm tra bao gồm 28 vectơ từ ngày 30/7/2006 đến ngày 26/8/2006. Ở phần huấn luyện, chúng ta huấn luyện mạng với 5000 bước (epoches).

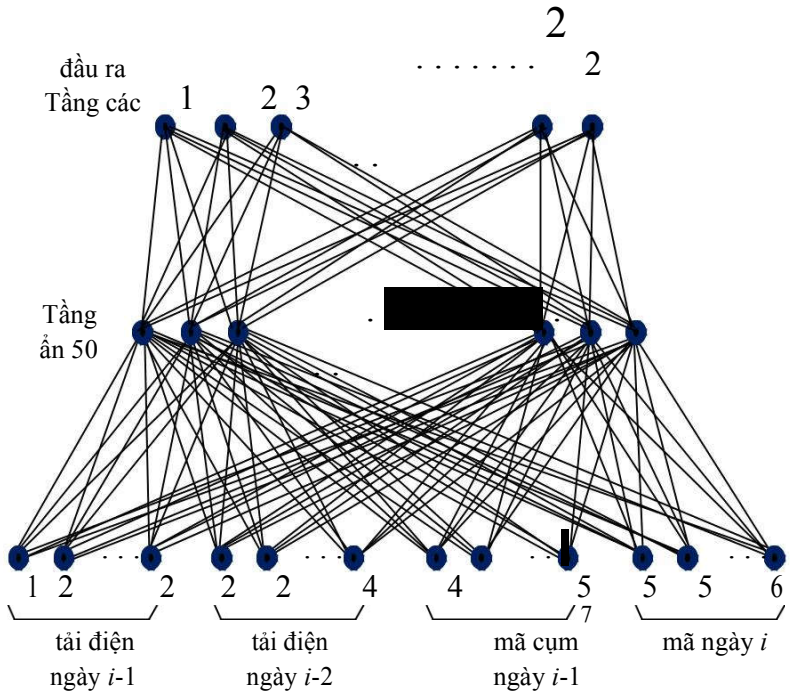
3.5.2 *Đánh giá kết quả*

Với bộ dữ liệu kiểm tra, giả sử ta lấy ngày Thứ ba (15/8/2006) cần được dự đoán, ta có đồ thị so sánh giữa phụ tải điện thực tế và phụ tải dự đoán như sau:



Hình 3.5 Kết quả so sánh giữ a các phụ tải thực tế và phụ tải dự đoán cho Thứ ba (15/8/2006).

Ta thấy kết quả thu được là có thể chấp nhận được, nhất là trong đồ thị các giá trị tại các đỉnh (P_{max}) là khá chính xác, điều này rất quan trọng trong việc dự đoán quá tải lưới điện và có biện pháp dự phòng kịp thời xử lý việc quá tải. Ví dụ trong đồ thị trên, ta thấy rõ đường phụ tải dự đoán thể hiện khá chính xác
ở các phụ tải đỉnh vào khoảng 10h sáng và 5h chiều đến 7h tối



Hình 3.4 Kiến trúc của mạng đa tầng truyền thẳng cho dự đoán phụ tải điện năng.

Hàm hoạt hóa sử dụng trong mạng ở tầng ẩn là hàm sigmoid

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3.14}$$

Đối với các nơ ron ở tầng đầu ra, hàm hoạt hóa sử dụng hàm tuyến tính

$$f(x) = x \ . \tag{3.15}$$

Một biểu thức rút gọn của mô hình được tóm tắt như sau:

$$y_i \quad [\ \ _i, y(\quad), y(\quad), \quad , \quad] \tag{3.16}$$

ở đây, $\mathbf{W}_i$ là trọng số tiếp hợp (hoặc giá trị dốc) của mạng, *CC* là mã cụm của các vec tơ dữ liệu (được gán trong pha SOM), và *DC* là mã định danh ngày được dự đoán.

Chương 2. Mạng nơ ron nhân tạo

2.1 Khái niệm, mô hình và kiến trúc mạng nơ ron nhân tạo.

2.1.1 *Mô hình mạng nơ ron nhân tạo:*

Một nơ ron nhân tạo là một đơn vị xử lý thông tin mà là cơ sở đối với hoạt động của một mạng nơ ron nhân tạo. Một tập các đường kết nối từ các đầu vào x_i khác nhau (tiếp hợp), mỗi đầu vào được phân biệt bởi trọng số hoặc độ dài w_{ki}. Chỉ số đầu tiên chỉ tới nơ ron đang nói đến và chỉ số thứ hai chỉ tới đầu vào của tiếp hợp mà trọng số đó liên quan. Một bộ cộng cho việc tính tổng các tín hiệu đầu vào x_i tính bởi cường độ tiếp hợp tương đối w_{ki}. Việc tính toán mô tả ở đây tạo thành một tổ hợp tuyến tính. Một hàm hoạt hóa *f* để giới hạn biên độ của đầu ra y_k của một nơ ron.

Thêm vào một độ nghiêng áp dụng bên ngoài định nghĩa là b_k. Độ nghiêng này có hiệu ứng làm tăng thêm hoặc thấp xuống đầu vào mạng của hàm hoạt hóa, phụ thuộc vào liệu nó là âm hay dương. Nơ ron là một mô hình trìu tượng của một nơ ron tự nhiên, và các khả năng xử lý của nó được hình thức hóa sử dụng các chú thích sau. Đầu tiên, có một vài đầu vào: __. _ _ _. ··· . __. Mỗi đầu vào __ được nhân lên bởi trọng số __. , ở đây *k* là chỉ số của một nơ ron được cho trong một mạng nơ ron. Tổng trọng số của các tích _ ___, với _ _ _. ··· . __ thường được xem như là "*net*" trong mạng nơ ron nhân tạo:

_ __ _ _ _ __ _ _ _ _ _ ... _ ___ _ _ _ _ (2.1)

Sử dụng ___ _ _ và đầu vào __ _ _ thì

$$net_k = x_0 w_{k0} + x_1 w_{k1} + x_2 w_{k2} + ... + x_m w_{km}$$
$$= \sum_{i=1}^{m} x_i w_{ki} \quad (2.2)$$

Do đó, $netk$ có thể tính bằng tích hai vec tơ m chiều:

$$net_k = \mathbf{X} \cdot \mathbf{W} \quad (2.3)$$

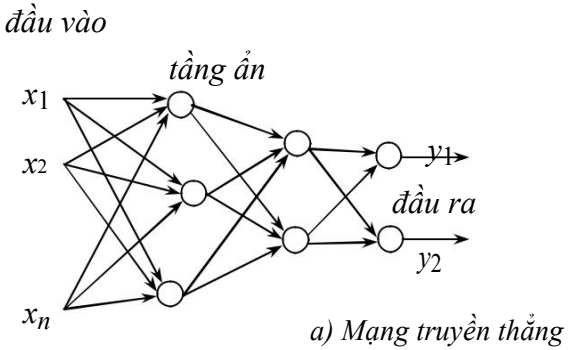
Cuối cùng, một nơ ron nhân tạo tính toán đầu ra y_k bằng một hàm có đối số là $netk$:

$$y_k = f(net_k) \quad (2.4)$$

Hàm f được gọi là hàm hoạt hóa. Có rất nhiều dạng hàm hoạt hóa.

2.1.2 *Kiến trúc mạng nơ ron nhân tạo*

Kiến trúc của một mạng nơ ron nhân tạo được định nghĩa bởi các đặc điểm của môt nút và các đặc điểm của sự kết nối của nút đó trong một mạng. Thông thường, kiến trúc mạng được xác định bởi số đầu vào của mạng, số đầu ra của mạng, toàn bộ số nút cơ sở mà thường là bằng các phần tử xử lý cho toàn bộ mạng, sự tổ chức và tương kết của chúng. Các mạng nơ ron nói chung được phân thành hai loại: mạng truyền thẳng và mạng hồi quy.



a) Mạng truyền thẳng

3.5 Thủ tục học sử dụng cho mạng nơ ron truyền thẳng đa tầng (MLP):

Giải thuật học cho bài toán này để huấn luyện mạng đa tầng truyền thẳng (MLP) là lan truyền ngược có thích ứng với hệ số quán tính

$$w_{ji}(t+1) = w_{ji}(t) \quad \eta \frac{\partial E}{\partial w_{ji}}(t) + \alpha \Delta w_{ji}(t \quad 1) \quad (3.13)$$

ở đây, tham số quán tính α giảm sự ảnh hưởng của bước trước lên bước hiện thời. Tốc độ học ban đầu được dùng là η=0.01, và hệ số quán tính là α=0.5.

3.5.1 *Kiến trúc của mạng MLP cho bài toán.*

Mô hình mạng MLP được thực hiện bảo một mạng hai tầng kết nối đầy đủ (đầu vào, tầng ẩn và tầng đầu ra) với các đặc điểm cấu trúc như sau:

- ✓ Giá trị phụ tải 24 giờ của ngày trước ngày dự đoán (i 1).
- ✓ Giá trị phụ tải 24h trước ngày dự đoán hai ngày (i 2).
- ✓ Vec tơ 9 thành phần chỉ ngày trước ngày dự đoán có định danh cụm mà đường cong phụ tải của nó thuộc về (ví dụ 100000000 cụm số 1, 010000000 cụm số 2, v.v…).
- ✓ Vec tơ 7 thành phần chỉ định ngày trước ngày dự đoán (1000000 là Chủ Nhật, 0000001 là thứ Bảy).

Đầu ra được thiết lập là 24, thể hiện sự dự đoán phụ tải 24 giờ cho ngày cần dự đoán.

10 10	0.773768
12 12	0.727853
14 14	0.764637

Bảng 3.2 Giá trị của $I_{\frac{*}{S}}$ cho 5 mô hình mạng SOM.

3.4 Mức thứ hai của sự trìu tượng: Sự phân cụm hiệu quả

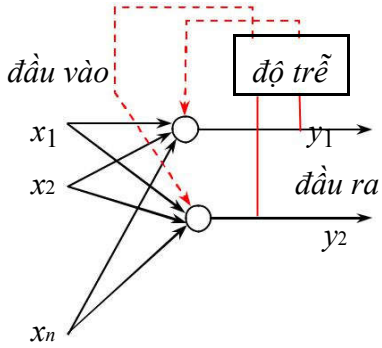
Trên các vec tơ nguyên mẫu của mạng SOM được huấn luyện, giải thuật *k*-mean phân cụm được thực hiện, thông thường bằng việc tối thiểu hóa hàm đích hoặc hàm lỗi:

$$E = \sum_{k=1}^{C} \sum_{x \in Q_k} \| \mathbf{x} - \mathbf{c}_k \|^2 \qquad (3.11)$$

ở đây C là số các cụm được đặt vào; Q_k chỉ cụm thứ k; $\mathbf{c}_k$ là trung tâm của cụm k, thu được bởi công thức:

$$\mathbf{c}_k = \frac{1}{N_k} \sum_{x \in Q_k} \mathbf{x}_i \qquad (3.12)$$

Giải thuật này được áp dụng lặp với tất cả các giá trị C trong vùng từ 2 tới $C_{\max}$. Giá trị của $C_{\max}$ được chọn dựa vào *U*-matrix của SOM được huấn luyện, và $C_{\max}$ ở đây ta chọn là 10. Sau quá trình phân cụm, và tính toán hiệu quả, các kết quả cho thấy việc chọn số cụm bằng 9 là cho kết quả phân cụm dữ liệu phụ tải hợp lý nhất[10]. Tiếp theo, 9 cụm được gán mã định danh đó sẽ được dùng làm đầu vào cho mạng đa tầng truyền thẳng MLP.



b) Mạng hồi quy

Hình 2.1 Kiến trúc mạng truyền thẳng và mạng hồi quy.

Mạng truyền thẳng nếu các lan truyền xử lý từ phía đầu vào tới phía đầu ra là theo một hướng thống nhất, không có bất cứ lặp hoặc ngược lại.

Mạng hồi quy được dùng trong các trường hợp khi có thông tin hiện thời đưa vào mạng đó, nhưng chuỗi đầu vào là rất quan trọng, và chúng ta cần mạng nơ ron đó lưu trữ một bản ghi của các đầu vào trước tiên và khuếch đại chúng với dữ liệu hiện thời đó để sinh ra câu trả lời.

Mặc dù có rất nhiều mô hình mạng nơ ron được đề xuất trong cả hai loại trên, mạng nơ ron truyền thẳng đa tầng với giải thuật học lan truyền ngược là mô hình được sử dụng rộng rãi nhất trong các ứng dụng thực tế.

2.2 Quá trình học và nội dung học

Nhiệm vụ chính của mạng nơ ron là phải học một mô hình của thế giới mà ở đó nó được nhúng vào và duy trì mô hình đó đủ phù hợp với thế giới thực để thu được một số kết quả xác định của ứng dụng liên quan. Quá trình học được dựa trên các

mẫu dữ liệu từ thế giới thực, và ở đây nằm trên sự khác nhau cơ bản giữa thiết kế của một mạng nơ ron và một hệ xử lý thông tin cổ điển. Để mô tả một trong các luật học, ta xét trường hợp đơn giản của một nơ ron k, được chỉ trong Hình 2.1 tạo thành chỉ một nút tính toán trong mạng nơ ron đó. Nơ ron k được thực hiện bằng vec tơ đầu vào $X(n)$, ở đây n chỉ thời gian rời rạc, hay chính xác hơn là bước thời gian của quá trình lặp liên quan đến việc điều chỉnh các trọng số đầu vào w_{ki}. Mỗi mẫu dữ liệu cho việc huấn luyện mạng nơ ron nhân tạo bao gồm vec tơ đầu vào $X(n)$ và đầu ra liên quan $d(n)$.

Xử lý vec tơ đầu vào $X(n)$, một nơ ron *k* sản xuất đầu ra được gọi là $y_k(n)$

$$y_k = f\left(\sum_{i=1}^{m} x_i w_{ki}\right) \qquad (2.5)$$

Nó thể hiện đầu ra duy nhất của mạng đơn giản này, và nó được so sánh với một phản hồi mong muốn hoặc đích ra $d_k(n)$ đã được cho. Một lỗi được sinh ra tại đầu ra được định nghĩa:

$$e_k(n) = d_k(n) \quad y_k(n) \qquad (2.6)$$

Tín hiệu lỗi được sinh ra khởi động một cơ chế điều khiển giải thuật học, mục đích của nó là phải áp dụng một chuỗi các chiều chỉnh sửa lỗi tới các trọng số đầu vào của một nơ ron. Các trọng số sửa lỗi được thiết kế để làm tín hiệu đầu ra $y_k(n)$ tiến sát tới đích mong muốn V từng bước từng bước một. Mục đích này thu được bằng việc tối thiểu hóa hàm tính giá $E(n)$, nó là giá trị tức thời của năng lượng lỗi, trong ví dụ này nó là $e_k(n)$. Quá trình học dựa trên việc tối thiểu hàm tính giá được chỉ tới

Kích cỡ sơ đồ	Độ dài huấn luyện thô	Độ dài huấn luyện tinh
66	200	20000
88	200	20000
1010	200	20000
1212	200	20000
1414	200	20000

Bảng 3.1 Độ dài huấn luyện của các pha thô và tinh cho 5 mạng SOM được kiểm tra.

Độ dài *T* của các pha "huấn luyện thô" và "tinh chỉnh" là như nhau cho mỗi SOM được sử dụng. Các độ dài huấn luyện *T* của hai pha được tổng kết trong

Bảng 3.1. Với mỗi SOM được huấn luyện, chúng ta sẽ tính toán chỉ số $I_s(t)$, tổng của các chỉ số chất lượng trước đây:

$$I_s(t) = \alpha M(t) + \beta E(t) \qquad (3.9)$$

ở đây, *t* là bước lặp (với một khoảng là tăng thêm 50). Mạng SOM được huấn luyện tốt nhất là cái được huấn luyện với số khoảng t_* cho mỗi I_s giả sử là một giá trị nhỏ nhất, là:

$$I_s^* = I_s(t^*) = \min\{I_s(t)\} \qquad (3.10)$$

Giá trị của I_s^* nhận được cho mỗi SOM được tổng kết trong Bảng 3.4. Trong luận văn này, mạng SOM tốt nhất được chọn là cái gồm 64 nơ ron (sơ đồ 8 8 nơ ron).

Kích cỡ sơ đồ	I_s^*
6 6	0.737839
8 8	0.670167

lân cận $\sigma(t)$ ban đầu nhận giá trị là 5 giảm đơn điệu xuống 1 trong quá trình làm thô (giá trị của bán kính lân cận được giữ bằng 1 trong suốt quá trình làm tinh).

3.3.2 *Chọn mô hình mạng SOM tốt nhất*

Chất lượng của mạng SOM thường được đánh giá dựa trên sự phân giải và khả năng duy trì hình học tập dữ liệu. Do đó, để đánh giá mạng SOM, tại mỗi bước lặp ta tính thêm vào vào các phép đo sau:

a. Phép đo độ méo trung bình (ADM) và được định nghĩa là

$$ADM = \frac{1}{N}\sum_{i=1}^{N}\sum_{j}^{M} h_{b_i j} \| \mathbf{x}_i - \mathbf{w}_j \|^2 \qquad (3.7)$$

ở đây N là số các vec tơ mẫu được dùng cho huấn luyện sơ đồ; M là số các đơn vị sơ đồ và $\|\|_2$ là khoảng cách Euclidean.

b. Phép đo độ phân giải (trung bình lỗi lượng tử):

$$QE = \frac{1}{N}\sum_{i=1}^{N} \| \mathbf{x}_i - \mathbf{w}_{b_i} \|^2 \qquad (3.8)$$

Năm sơ đồ tự tổ chức (6 6; 8 8; 10 10; 12 12; 14 14), được mô tả bởi mạng lục giác, kiểm tra và phân cụm của dữ liệu phụ tải được dẫn ra từ một trong số chúng, cái nào được chọn là do sự quan sát của hai chỉ số chất lượng là *ADM*, và *QE*. Ở bài toán này, SOMs được sử dụng khởi đầu với các giá trị ngẫu nhiên. Đối với mỗi thành phần xi, các giá trị của các vec tơ trọng số được phân bố đồng đều trong miền $[\min(xi)\max(xi)]$.

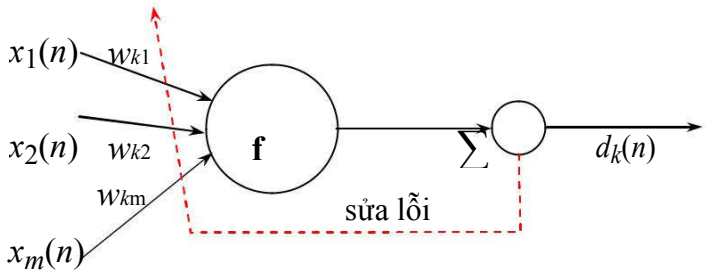
việc học sửa chữa lỗi. Đặc biệt, tối thiểu hóa hàm $E(n)$ dẫn tới một luật học thường được gọi là luật delta hoặc luật Widrow-Hoff. Gọi ■■ |■| là giá trị của hệ số trọng số cho nơ ron *k* hoạt động bởi đầu vào ■ | ■▌ tại bước thời gian *n*. Theo luật delta, sự điều chỉnh Δw n được định nghĩa bởi:

$$\Delta w_{kj}(n) = \eta \cdot e_k(n) \cdot x_j(n) \qquad (2.7)$$

ở đây, η là hằng số dương quyết định tốc độ học. Do đó, luật delta được phát biểu như sau: Sự điều chỉnh được làm tới một hệ số trọng số của một kết nối nơ ron đầu vào là tỉ lệ với tích của tín hiệu lỗi và giá trị đầu vào của kết nối hiện tại. Sau khi $\Delta wkj(n)$ đã được tính toán, giá trị cập nhật của trọng số tiếp hợp được quyết định bởi:

$$w_{kj}(n+1) = w_{kj}(n) + \blacktriangle w_{kj}(n) \qquad (2.8)$$

$wkj(n)$ và $wkj(n+1)$ có thể được xem là giá trị cũ và giá trị mới của trọng số tiếp hợp wkj.



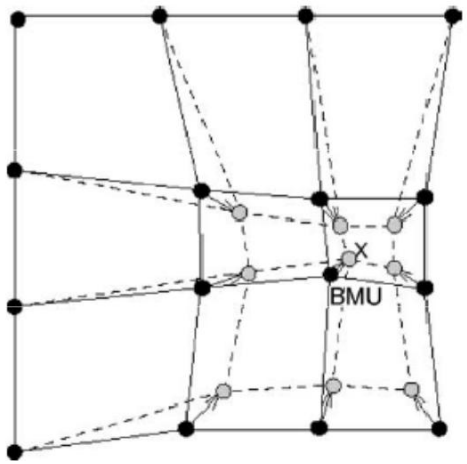
Hình 2.2 Học sửa lỗi thực hiện thông qua các điều chỉnh các trọng số.

Học sửa lỗi được áp dụng trong kiến trúc mạng nơ ron phức tạp hơn nhiều. Quá trình sửa chữa trọng số sẽ tiếp tục với các mẫu huấn luyện mới hoặc dùng cùng các mẫu dữ liệu trong các vòng lặp tiếp theo. Khi để kết thúc quá trình lặp được định nghĩa bởi một tham số đặc biệt hoặc tập các tham số gọi là tiêu chuẩn dừng. .

2.3 Một số kiểu mạng nơ ron nhân tạo

2.3.1 *Mạng tự tổ chức (Kohonen feature maps)*

"*Kohonen feature maps*" là mạng nơ ron truyền thẳng sử dụng giải thuật học không giám sát, thông qua một quá trình tự tổ chức, cấu hình các đơn vị đầu ra thành một sơ đồ hình học hoặc không gian. Các mạng Kohonen thể hiện dưới dạng của "sơ đồ tự tổ chức" (SOM - Self-organized maps), bản thân nó thể hiện một lớp đặc biệt của các mạng nơ ron. Mục đích của SOM là biến đổi một tín hiệu đầu vào có số chiều cao, phức tạp thành một sơ đồ rời rạc có số chiều thấp đơn giản hơn. Do vậy, các SOM thường thích hợp cho việc phân tích theo cụm, ở đây, các mẫu ẩn phía dưới trong các bản ghi và các trường được tìm kiếm. Các mạng SOM cấu trúc các nút đầu ra thành các cụm các nút, ở đây các nút trong sự xấp xỉ gần sát hơn tương đương với nhau hơn là các nút khác ở xa hơn.



Hình 3.3 Cập nhật đơn vị phù hợp nhất và các lân cận của nó về phía mẫu đầu vào đánh dấu là x. Các đường đen đậm và đứt quãng thể hiện trước và sau khi cập nhật.

Hàm lân cận được sử dụng là hàm Gaussian:

$$h_{i,j}(t) = e^{-\frac{\|r_i - r_j\|^2}{2\sigma^2(t)}} \quad (3.5)$$

ở đây r_i và r_j là các vị trí của các đơn vị i và j trên lưới SOM; $\sigma(t)$ trung bình được gọi là bán kính lân cận. Hàm lân cận nhận giá trị lớn nhất cho đơn vị thắng cuộc và giảm đơn điệu với việc tăng khoảng cách trên lưới sơ đồ $\|r_i - r_j\|$.

Tốc độ học được giảm theo luật luỹ thừa:

$$\eta(t) = \eta_0(0.001/\eta_0)^{t/T} \quad (3.6)$$

với T là độ dài huấn luyện, và η_0 là giá trị ban đầu ta cho là 0.01 và giảm xuống 0.001 trong suốt quá trình huấn luyện. Bán kính

3.3 Mức đầu tiên của sự trìu tượng: huấn luyện của SOM

3.3.1 *Chuẩn bị các vec tơ đầu vào*

Một sơ đồ tự tổ chức SOM sử dụng một tập các vec tơ nguyên mẫu mô tả tập dữ liệu và thực hiện "phép chiếu duy trì hình học" của các nguyên mẫu đó từ một không gian đầu vào d chiều vào một lưới thấp chiều. Mỗi nơ ron của một SOM được thể hiện bởi một vec tơ trọng số d chiều (vec tơ nguyên mẫu) $\mathbf{w}=[w_1,\ldots,w_d]$, ở đây d là số chiều của các vec tơ đầu vào. Nơ ron có vec tơ trọng số sát nhất với vec tơ đầu vào $\mathbf{x}$ được gọi là đơn vị phù hợp nhất (BMU). Gọi $\mathbf{w}_j$ là vec tơ trọng số của một đơn vị sơ đồ và $\mathbf{w}_{bi}$ là vec tơ trọng số của BMU, ta có chỉ số b_i là:

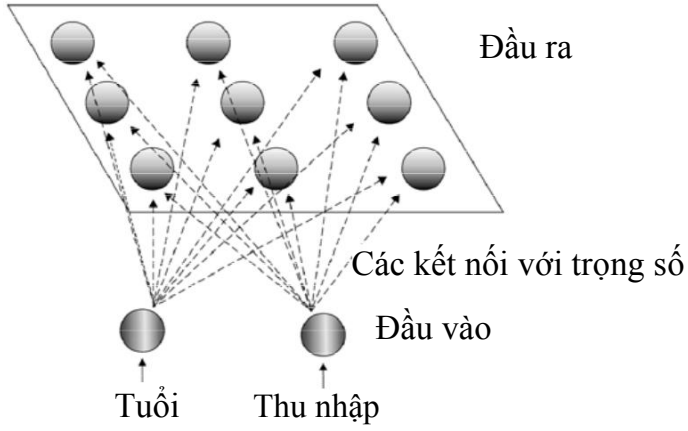
$$\|\mathbf{x}(t)-\mathbf{w}_{b_i}(t)\| = \min_j \|\mathbf{x}(t)-\mathbf{w}_j(t)\| \qquad (3.3)$$

ở đây $\min_j$ chỉ giá trị nhỏ nhất của j đạt được tại thời điểm t, $\|\cdot\|$ là phép đo khoảng cách, thường là khoảng cách Euclidean,

Sau sự thể hiện của các vec tơ dữ liệu $\mathbf{x}$, các vec tơ trọng số của các nơ ron được cập nhật theo luật học sau:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \eta(t)h_{b_i,j}(t)[\mathbf{x}(t) - \mathbf{w}_j(t)] \qquad (3.4)$$

ở đây, t là thời gian lặp; $\mathbf{x}(t)$ là một vec tơ đầu vào lấy ngẫu nhiên từ dữ liệu đầu vào tại thời điểm t; $h_{b_i,j}(t)$ là một phần tử lân cận trung tâm trên BMU; $\eta(t)$ là tốc độ học.



Hình 2.3 Kiến trúc thông thường của SOM.

Các sơ đồ tự tổ chức là dựa trên việc học cạnh tranh, ở đâu các nút đầu ra cạnh tranh giữa chúng với nhau để trở thành nút thắng cuộc, chỉ nút đó là được hoạt hóa bởi một sự quan sát đầu vào đầu đặc biệt. Kiến trúc SOM thông thường như

Hình 2.3.

Các mạng nơ ron cạnh tranh nằm trong lớp các mạng hồi quy, và chúng dựa trên các giải thuật học không giám sát, như là giải thuật cạnh tranh. Trong việc học cạnh tranh, các nơ ron đầu ra của mạng nơ ron tính toán trong số chính chính tiến tới tích cực (được "cháy").

Giải thuật mạng Kohonen:

Với mỗi vec tơ đầu vào *x*, do:

Cạnh tranh. Đối với mỗi nút đầu vào *j*, tính toán giá trị $D(w_j, x_n)$ của hàm tính giá. Ví dụ, với khoảng cách Euclidean, $D(w_j, x_n) = \sqrt{\sum (w_{ij} - x_{ni})^2}$. Tìm nút *j* chiến thắng mà làm $D(w_j, x_n)$ là nhỏ nhất trên tất cả các nút đầu ra.

Hợp tác. Định danh tất cả các nút đầu ra *j* phía trong lân cận của *j* định nghĩa bởi kích thước lân cận *R*. Với các nút này, làm như sau với tất cả các trường bản ghi đầu vào:

- Thích ứng. Điều chỉnh các trọng số:

$$w_{ij,\text{new}} = w_{ij,\text{current}} - \eta(x_{ni} - w_{ij,\text{current}})$$

Điều chỉnh tốc độ học và kích thước lân cận, khi cần. Dừng lại khi gặp điều kiện kết thúc.

2.3.2 *Mạng nơ ron truyền thẳng đa tầng (Multilayer Perceptrons).*

Mạng truyền thẳng đa tầng là một trong các lớp phổ biến và quan trọng nhất trong các ứng dụng của mạng nơ ron. Mạng nơ ron này thông thường bao gồm một tập các đầu vào mà hình thành nên tầng đầu vào của mạng, một hoặc nhiều hơn các tầng ẩn của các nút tính toán, và cuối cùng một tầng đầu ra của các nút tính toán. Việc xử lý là theo một hướng thẳng dựa trên cơ sở từng lớp. MLP đã được ứng dụng thành công để giải quyết một số các bài toán khó và đa dạng bằng việc huấn luyện mạng dưới

Phụ tải cao điểm là nhân tố quyết định việc huy động nguồn điện trong khi đó phụ tải thấp điểm lại giữ vai trò quan trọng trong việc quyết định phối hợp và điều chỉnh các nguồn điện nhằm đảm bảo vận hành kinh tế. Dự báo chính xác phụ tải cao điểm có hiệu quả lớn về kinh tế vào mùa lũ khi cần khai thác tối đa nguồn thủy điện, trong khi nước các hồ chứa đang xả thì các nguồn điện khác huy động chỉ nhằm mục đích an toàn, bù điện áp và phủ đỉnh.

3.2 Chuẩn hóa các vec tơ dữ liệu phụ tải

Dữ liệu phụ tải được dùng trong luận văn là từ 1/1/2005 đến 26/8/2006: tập dữ liệu bao gồm 583 vec tơ, mỗi vec tơ gồm 24 thành phần (đã loại bỏ một số vec tơ không phù hợp). Trước khi chia tách dữ liệu phụ tải vào các cụm khác nhau sử dụng SOM, chúng được chuẩn hóa sử dụng kỹ thuật mô tả trong phương trình sau:

$$x' = \frac{x - \bar{x}}{\text{Std}_x} \tag{3.1}$$

$$x_{\text{norm}} = \frac{1}{1 + e^{-x'}} \tag{3.2}$$

ở đây:

x là giá trị của thành phần chung của các vec tơ dữ liệu bước khi chuẩn hóa;

$\bar{x}$ là giá trị trung bình trên thành phần;

$\text{Std}_x = \sqrt{\frac{1}{N}\sum N(x_i - \bar{x})^2}$ (*N* số dữ liệu trong một thành phần) là độ lệch chuẩn của cùng thành phần đó;

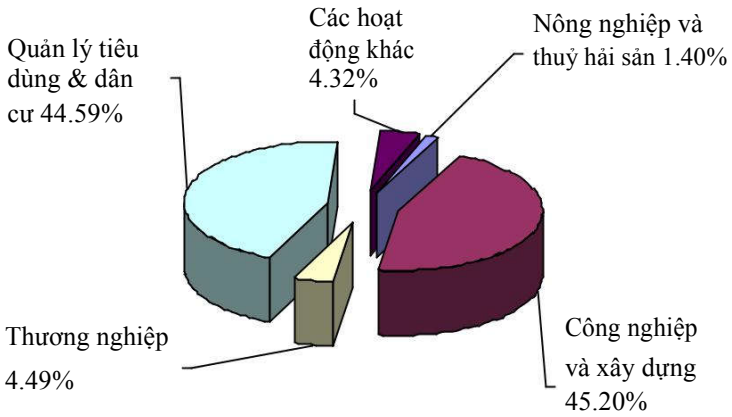
x_{norm} là giá trị thành phần sau chuẩn hóa.

Các chuẩn hóa thường được thực hiện để giảm độ biến thiên của các thành phần vec tơ.

Trong pha tiền xử lý, dữ liệu ban đầu được chuẩn hóa để giảm sự biến thiên của chúng và sau đó nhóm lại thành các cụm liên quan tới các kiểu phụ tải khác nhau. Việc phân nhóm được thực hiện bởi một kỹ thuật có hiệu quả cao đó là mạng tự tổ chức (SOM). Nhờ có ứng dụng thủ tục phân cụm này, mỗi vec tơ dữ liệu phụ tải được đánh nhãn với một mã định danh liên quan tới cụm mà nó thuộc về và được sử dụng. Sự phân cụm thực sự được thực hiện bằng việc thực thi giải thuật k-means trên các vectơ của SOM được huấn luyện trước đó với dữ liệu phụ tải. Sau đó, chúng ta mô tả một ứng của của một MLP huấn luyện với dữ liệu phụ tải đối với một vấn đề dự đoán phụ tải trong thời gian 24h.

3.1.2 *Phụ tải hệ thống điện miền Bắc*

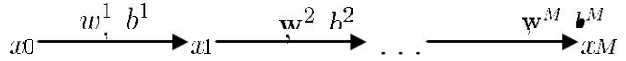
Các thành phần cấu thành phụ tải được thể hiện ở biểu đồ sau:



Hình 3.2 Các thành phần cấu thành phụ tải.

dạng có giám sát với một thuật giải được nhiều người biết đến là giải thuật lỗi lan truyền ngược (Back-propagation).

Giả sử ta có biểu đồ của một MLP như sau:



ở đây, x_m là tập các đầu vào tại tầng *m*, $m = 0, \ldots, M$, và $\mathbf{w}_m$ là một ma trận $n_m \times n_{m-1}$ đối với tất cả $m = 1, \ldots, M$. Có $M + 1$ tầng nơ ron, và M tầng các trọng số tiếp hợp. Chúng ta sẽ thay đổi các trọng số $\mathbf{w}$ và độ dốc b để đầu ra thực sự x_M trở nên sát hơn với đầu ra mong muốn d.

Giải thuật lan truyền ngược bao gồm các bước sau:

1. **Truyền thẳng.** Vec tơ đầu vào x_0 được biến đổi tới vec tơ đầu ra x_M bởi tính giá trị công thức:

$$x_i^m = f(y_i^m) = f\left(\sum_{j=1}^{n_{m-1}} \mathbf{w}_{ij}^m x_j^{m-1} + \mathbf{b}_i^m\right) \qquad (2.9)$$

   với $i = 1$ đến M.

2. **Tính toán lỗi**. Việc tính toán lỗi chính là tìm hệ số nhạy cảm thể hiện sự khác nhau giữa đầu ra mong muốn và đầu ra thực sự x_M được tính toán

$$s_i^M = f'(y_i^M)(d_i - x_i^M) \qquad (2.10)$$

3. **Truyền ngược.** Tín hiệu lỗi tại các đơn vị đầu ra được truyền ngược lại qua toàn bộ mạng, bằng việc tính giá trị

$$s_j^{m-1} = f'(y_j^{m-1}) \sum_{i=1}^{n_m} s_i^m \mathbf{w}_{ij}^m \qquad (2.11)$$

   với *m* từ *M* đến 1.

4. **Cập nhật trọng số.** Các trọng số tiếp hợp và các độ dốc được cập nhật sử dụng các kết quả của truyền thẳng và lan truyền ngược:

$$\Delta \mathbf{w}_{ij}^{m} = \eta s_i^m x_j^{m\ 1} \qquad (2.12)$$

(2.13)

được tính từ $m = 1$ đến M.

Kết luận chương

Chương 2 trình bày về khái niệm, mô hình và kiến trúc của mạng nơ ron nhân tạo trong đó giới thiệu hai kiểu mạng nơ ron nhân tạo là mạng nơ ron tự tổ chức và mạng nơ ron truyền thẳng đa tầng được dùng rộng rãi cho các bài toán phân cụm và dự đoán có dữ liệu đầu vào lớn và nhiễu.

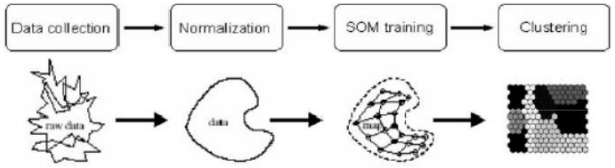
Chương 3. ỨNG DỤNG MẠNG NƠ RON CHO BÀI TOÁN DỰ ĐOÁN TẢI ĐIỆN NĂNG

3.1 Giới thiệu bài toán phụ tải điện năng

Rất nhiều nghiên cứu đã chỉ ra độ nhạy cảm do sự quá tải tiêu thụ điện năng đặc biệt tập trung vào sự dự đoán giới hạn 24 giờ tới. Một sự tiếp cận khá phổ biến trong các vấn đề dự đoán phụ tải ngắn hạn được tạo thành bởi các giải thuật dựa trên các mạng nơ ron nhân tạo. Phần lớn các mô hình cho dự đoán phụ tải ngắn hạn sử dụng kiến trúc mạng nơ ron được biết đến như các "perceptron" đa tầng (MLPs). Đặc biệt, với vấn đề dự đoán phụ tải ngắn hạn, rất nhiều các ứng dụng chỉ ra rằng mạng nơ ron nhân tạo có khả năng học các thuộc tính của phụ tải điện, và hơn nữa yêu cầu phân tích sâu và kỹ để được khám phá.

3.1.1 *Mô tả công việc*

Trong luận văn này, chúng tôi sử dụng mạng nơ ron nhân tạo dựa trên mô hình dự báo sử dụng dữ liệu phụ tải điện thu thập tại khu vực miền Bắc - Việt Nam. Các giá trị phụ tải được lấy từ tổng điện năng tiêu thụ của khu vực này. Lượng điện tiêu thụ được tính bởi lượng điện sinh hoạt người dân địa phương và cả lượng điện tiêu thụ của các nhà dùng điện công nghiệp.



Hình 3.1 Sơ đồ các bước phân cụm trên SOM.