

MẠNG NƠ RON MỜ HỒI QUY VÀ ỨNG DỤNG DỰ BÁO

Nguyễn Minh Thi

Lư Nhật Vinh

Nguyễn Minh Thanh

TÓM TẮT

Bài báo trình bày một mô hình mạng nơ ron mờ hồi quy, là sự mở rộng mạng nơ ron mờ bằng cách thêm các nút hồi tiếp vào lớp thứ hai của mạng nơ ron mờ. Các nút hồi tiếp này đóng vai trò các phần tử nhớ trong mô hình mạng nơ ron mờ hồi quy. Mạng nơ ron mờ hồi quy tỏ ra đạt hiệu quả cao cho những ứng dụng như: dự báo chuỗi thời gian, nhận dạng và điều khiển những hệ phi tuyến.

ABSTRACT

The paper represented a Recurrent Fuzzy Neural Network (RFNN). Temporal relations are embedded in the network by adding feedback connections in the second layer of the fuzzy neural network (FNN). The RFNN expands the basic ability of the FNN to cope with temporal problems. The RFNN is applied in time series prediction, identification, and control of nonlinear systems.

I. GIỚI THIỆU

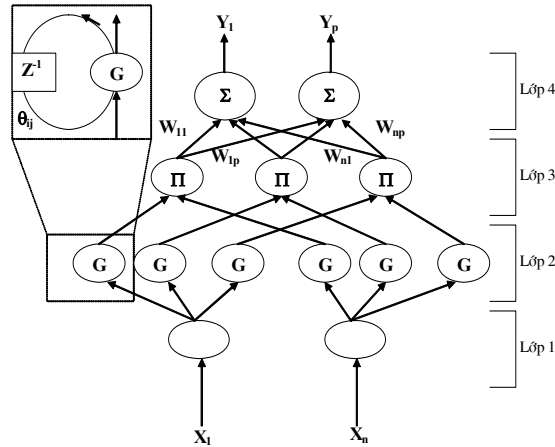
Những năm gần đây mô hình mạng nơ ron đã chứng tỏ khả năng trong việc giải quyết các bài toán bằng cách học từ những mẫu dữ liệu thực tế. Mô hình mạng nơ ron có khả năng xấp xỉ một hàm liên tục với độ chính xác cho trước.

Mô hình logic mờ nhằm vào việc khai thác khả năng xử lý những tri thức như con người. Việc kết hợp mô hình logic mờ và mạng nơ ron mờ ra một hướng mới gọi là mô hình mạng nơ ron mờ nhằm tận dụng những ưu điểm của cả hai mô hình. Nhược điểm chính của mạng nơ ron mờ là chỉ thích hợp với những bài toán có tính chất tĩnh.

Mô hình mạng nơ ron mờ hồi quy là sự mở rộng mạng nơ ron mờ nhằm để giải quyết các bài toán động. Mô hình mạng nơ ron mờ hồi quy thích hợp khi giải quyết các bài toán mà những đầu vào và đầu ra phụ thuộc theo thời gian. Mạng nơ ron mờ hồi quy mở rộng mạng nơ ron mờ bằng cách thêm các nút hồi tiếp vào lớp thứ hai của mạng nơ ron mờ. Các nút hồi tiếp này đóng vai trò các phần tử nhớ trong mô hình mạng nơ ron mờ hồi quy. Việc thêm các nút hồi tiếp này làm cho mạng nơ ron mờ hồi quy có khả năng giải quyết các bài toán động. Do đó, mạng nơ ron mờ hồi quy là một mô hình mở rộng của mạng nơ ron mờ.

II. CẤU TRÚC MẠNG NƠ RON MỜ HỒI QUY

Cấu trúc mô hình mạng nơ ron mờ hồi quy được thể hiện như Hình 1.



Hình 1. Cấu trúc mạng nơ ron mờ hồi quy

Một mạng nơ ron mờ hồi quy gồm 4 lớp như sau:

- Lớp 1: Là lớp nhập gồm N dữ liệu nhập (input).
- Lớp 2: Gọi là lớp các hàm thành viên. Các nút trong lớp này thực hiện việc mờ hoá. Lớp này dùng để tính giá trị hàm thành viên theo hàm phân phối Gauss. Số nút trong lớp 2 là $N \times M$, trong đó M là số luật mờ (số nút của lớp 3).
- Lớp 3: Lớp các luật mờ. Các nút trong lớp này tạo thành cơ sở luật mờ (Fuzzy rule base) gồm M nút. Liên kết giữa lớp 2 và lớp 3 biểu diễn giả thiết của luật mờ. Liên kết giữa lớp 3 và lớp 4 biểu diễn kết luận của luật mờ.
- Lớp 4: Lớp xuất gồm P nút.

Liên kết giữa lớp 3 và lớp 4 được gán trọng số w_{jk} .

Như vậy, số nút của mô hình là: $N + (N \times M) + M + P$

III. HOẠT ĐỘNG CỦA MÔ HÌNH

Chúng ta hãy xem xét hoạt động của các nút trong mỗi lớp. Ký hiệu $u_i^{(k)}$ và $O_i^{(k)}$ tương ứng là input và output của nút thứ i trong lớp k .

- Lớp 1: $O_i^{(1)} = u_i^{(1)} = x_i(t)$, với $i = 1 \div N$
- Lớp 2: $O_{ij}^{(2)} = \exp \left[-\frac{(u_{ij}^{(2)} - m_{ij})^2}{(\sigma_{ij})} \right]$, với $i = 1 \div N, j = 1 \div M$

Trong đó: m_{ij} và σ_{ij} là trọng tâm và độ rộng của hàm thành viên theo phân bố Gauss.

Hơn nữa, input của các nút này là

$$u_{ij}^{(2)}(t) = O_i^{(1)} + \theta_{ij} O_{ij}^{(2)}(t-1), \text{ với } i = 1 \div N, j = 1 \div M$$

Trong đó: θ_{ij} biểu diễn trọng số cho các nút hồi tiếp.

Chúng ta thấy rằng input của các nút trong lớp này có chứa toán hạng $O_{ij}^{(2)}(t-1)$ lưu thông tin trước đó của mô hình. Đây chính là sự khác biệt giữa mạng nơ ron mờ và mạng nơ ron mờ hồi quy.

Như vậy:

$$O_{ij}^{(2)} = \exp \left[-\frac{[O_i^{(1)} + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right] = \exp \left[-\frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right]$$

với $i = 1 \div N, j = 1 \div M$

Mỗi nút trong lớp này có 3 thông số là m_{ij} , σ_{ij} và θ_{ij} .

- Lớp 3: Các nút trong lớp này thực hiện phép toán AND

$$O_j^{(3)} = \prod_{i=1}^N O_{ij}^{(2)} = \prod_{i=1}^N \exp \left[-\frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right], \text{ với } i = 1 \div N, j = 1 \div M$$

- Lớp 4: Các nút trong lớp này thực hiện việc giải mờ.

$$y_k = O_k^{(4)} = \sum_{j=1}^M u_{jk}^{(4)} w_{jk} = \sum_{j=1}^M O_j^{(3)} w_{jk} = \sum_{j=1}^M w_{jk} \prod_{i=1}^N \exp \left[-\frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right]$$

với $i = 1 \div N, j = 1 \div M, k = 1 \div P$

Như vậy, trong mô hình mạng nơ ron mờ hồi quy, các thông số cần phải xác định là m_{ij} , σ_{ij} , θ_{ij} và w_{jk} .

IV. LẬP LUẬN MỜ

Giả sử cho hệ mạng nơ ron mờ hồi quy với nhiều đầu vào (input) và một đầu ra (output). Gọi x_i là biến ngôn ngữ thứ i và α_j là giá trị kích hoạt của luật j , w_j là trọng số của kết nối thứ j .

Một luật suy diễn mờ được biểu diễn như sau:

R^j : Nếu u_{1j} là A_{1j} , u_{2j} là A_{2j} , .., u_{nj} là A_{nj} , thì $y = w_j$

Trong đó:

$i = 1, 2, \dots, n$.

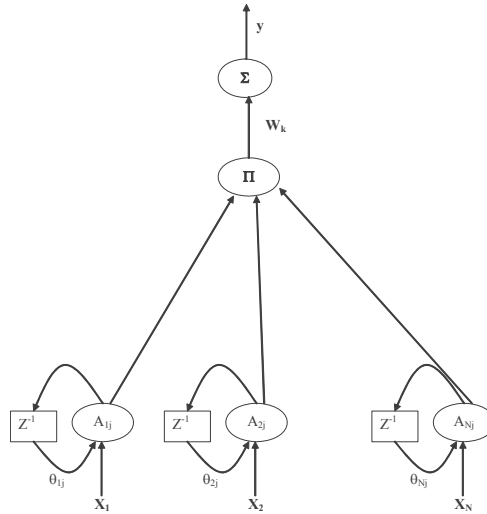
n là số đầu vào (input)

$u_{ij} = x_i + \theta_{ij} * o_{ij}^{(2)}(t-1)$

A_{ij} là các tập mờ

w_j là trọng số kết nối

Đầu vào của mỗi hàm thành viên là đầu vào x_i của mạng cộng với số hạng $o_{ij}^{(2)}\theta_{ij}$. Sơ đồ kết nối dựa trên luật suy diễn mờ như Hình 2.



Hình 2. Cấu trúc kết nối trên cơ sở luật mờ thứ j

Hệ thống mờ với những thành phần nhớ (Feedback unit) có thể được xem như là một hệ suy luận mờ động và giá trị suy luận được tính bởi

$$y^* = \sum_{j=1}^M w_j \alpha_j$$

với M là số luật và

$$\alpha_j = \prod_{i=1}^N \mu A_{ij}(u_{ij})$$

Mạng nơ ron mờ hồi quy là một hệ suy luận mờ có các phần tử nhớ. Sau khi huấn luyện xong, các thông số trong mạng cùng với các phần tử nhớ đã xác định tri thức.

V. GIẢI THUẬT HỌC CHO MẠNG NƠ RON MỜ HỒI QUY

Mạng nơ ron mờ hồi quy sử dụng giải thuật học lan truyền ngược (Back Propagation) để xác định các thông số trong mô hình. Để đơn giản trong việc trình bày, chúng ta xem xét trường hợp mô hình chỉ có một đầu ra.

Mục đích của việc học là cực tiểu hóa hàm lỗi :

$$E = \frac{1}{2} \sum_t (y^{(d)}(t) - y(t))^2 = \frac{1}{2} \sum_t (y^{(d)}(t) - O^{(4)}(t))^2$$

Trong đó: $y^{(d)}(t)$ là output mong muốn và $y(t) = O^{(4)}(t)$ là output hiện tại của mô hình tại mẫu dữ liệu thứ t .

Như chúng ta đã biết, trong giải thuật lan truyền ngược, các thông số sẽ được cập nhật như sau:

$$W(t+1) = W(t) + \Delta W(t) = W(t) + \eta \left(-\frac{\partial E(t)}{\partial W} \right)$$

Trong đó: W là vector thông số của mô hình và η là tốc độ học.

Đặt $e(t) = y^{(d)}(t) - y(t)$, ta có:

$$\frac{\partial E(t)}{\partial W} = -e(t) \frac{\partial y(t)}{\partial W} = -e(t) \frac{\partial O^{(4)}(t)}{\partial W}$$

Do đó, các thông số m_{ij} , σ_{ij} , θ_{ij} và w_{jk} sẽ được cập nhật tương ứng như sau:

$$w_{jk}(t+1) = w_{jk}(t) - \eta^w \frac{\partial E}{\partial w_{jk}}$$

$$m_{ij}(t+1) = m_{ij}(t) - \eta^m \frac{\partial E}{\partial m_{ij}}$$

$$\sigma_{ij}(t+1) = \sigma_{ij}(t) - \eta^\sigma \frac{\partial E}{\partial \sigma_{ij}}$$

$$\theta_{ij}(t+1) = \theta_{ij}(t) - \eta^\theta \frac{\partial E}{\partial \theta_{ij}}$$

Trong đó:

$$\frac{\partial E}{\partial w_{jk}} = -e(t) O_j^{(3)}$$

$$\frac{\partial E}{\partial m_{ij}} = -e(t) \sum_{j=1}^M w_{jk} \frac{\partial O_j^{(3)}}{\partial m_{ij}} = -e(t) \sum_{j=1}^M w_{jk} O_j^{(3)} \frac{2[x_i(t) + O_j^{(2)}(t-1)\theta_{ij} - m_{ij}]}{(\sigma_{ij})^2}$$

$$\frac{\partial E}{\partial \sigma_{ij}} = -e(t) \sum_{j=1}^M w_{jk} \frac{\partial O_j^{(3)}}{\partial \sigma_{ij}} = -e(t) \sum_{j=1}^M w_{jk} O_j^{(3)} \frac{2[x_i(t) + O_j^{(2)}(t-1)\theta_{ij} - m_{ij}]}{(\sigma_{ij})^3}$$

$$\frac{\partial E}{\partial \theta_{ij}} = -e(t) \sum_{j=1}^M w_{jk} \frac{\partial O_j^{(3)}}{\partial \theta_{ij}} = -e(t) \sum_{j=1}^M w_{jk} \frac{-2[x_i(t) + O_j^{(2)}(t-1)\theta_{ij} - m_{ij}] O_j^{(2)}(t-1)}{(\sigma_{ij})^2}$$

Giải thuật học

Học có giám sát với tập mẫu $\{X_s, Y_s\}$

Mỗi khi đưa một mẫu $X_s = (x_1, x_2, \dots, x_n)$ vào mạng, ta thực hiện các công việc sau:

- Lan truyền mẫu X_s qua mạng để tính giá trị đầu ra (output) theo:

$$y_k = \sum_{j=1}^M w_{jk} \prod_{i=1}^N \exp \left[- \frac{[x_i(t) + \theta_{ij} O_{ij}^{(2)}(t-1) - m_{ij}]^2}{(\sigma_{ij})^2} \right]$$

- Tính lỗi tại mẫu học dựa trên sai lệch:

$$e(t) = y^{(d)}(t) - y(t)$$

- Cập nhật các thông số m_{ij} , σ_{ij} , θ_{ij} và w_{jk} theo:

$$w_{jk}(t+1) = w_{jk}(t) - \eta^w \frac{\partial E}{\partial w_{jk}}$$

$$m_{ij}(t+1) = m_{ij}(t) - \eta^m \frac{\partial E}{\partial m_{ij}}$$

$$\sigma_{ij}(t+1) = \sigma_{ij}(t) - \eta^\sigma \frac{\partial E}{\partial \sigma_{ij}}$$

$$\theta_{ij}(t+1) = \theta_{ij}(t) - \eta^\theta \frac{\partial E}{\partial \theta_{ij}}$$

VI. ỨNG DỤNG MẠNG NƠ RON MỜ HỒI QUY VÀO DỰ BÁO

Chúng ta tiến hành cài đặt thử nghiệm mô hình mạng nơ ron mờ hồi quy trên để dự báo một số mặt hàng ảnh hưởng đến nền kinh tế Việt Nam và ảnh hưởng của chúng có tác động rất lớn đến các hoạt động kinh tế khác.

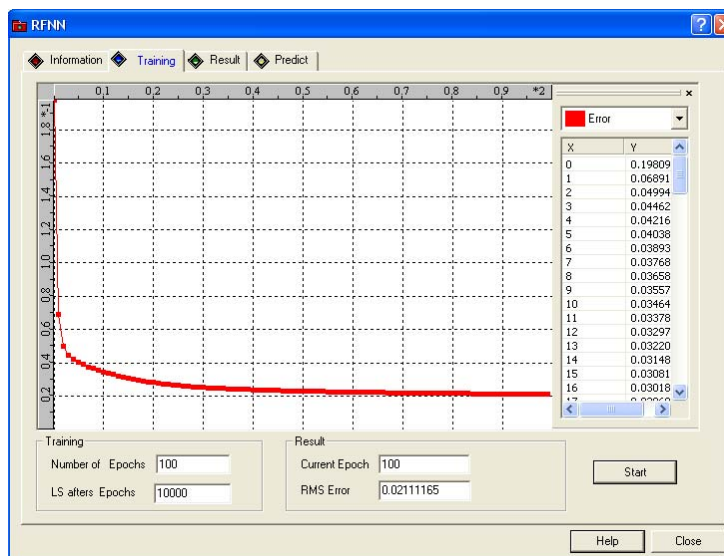
Ngoài ra, để kiểm tra và so sánh khả năng dự báo của mô hình so với các phương pháp đang được ứng dụng rộng rãi trong kinh tế, chương trình được cài đặt thêm phương pháp dự báo chuỗi ARIMA, là mô hình phân tích dự báo kinh tế cổ điển đang được sử dụng rộng rãi trong các ngành dự báo của kinh tế ở Việt Nam và trên thế giới, nhằm đối chiếu và so sánh phương pháp dự báo của mô hình với phương pháp dự báo chuỗi ARIMA.

➤ Dữ liệu thực nghiệm

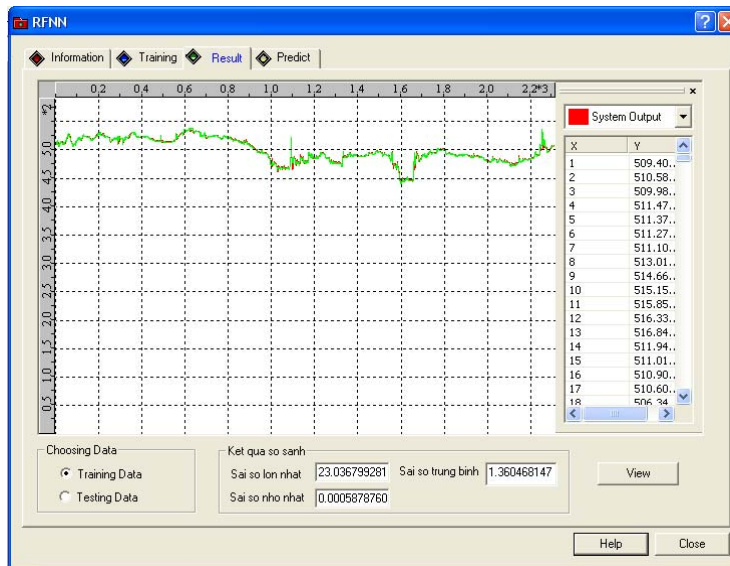
Dữ liệu thực nghiệm là giá cả một số mặt hàng như: vàng, đôla, gạo, cà phê, xi măng được cung cấp từ Viện Kinh Tế TP. Hồ Chí Minh. Dữ liệu về giá cả các mặt hàng này được cập nhật hàng ngày, và có thể sử dụng dữ liệu từ nhiều năm trước, từ năm 2002 đến 2007, tức bao gồm khoảng hơn 2000 số liệu cho mỗi bộ dữ liệu của mỗi loại mặt hàng thử nghiệm.

➤ Kết quả thử nghiệm

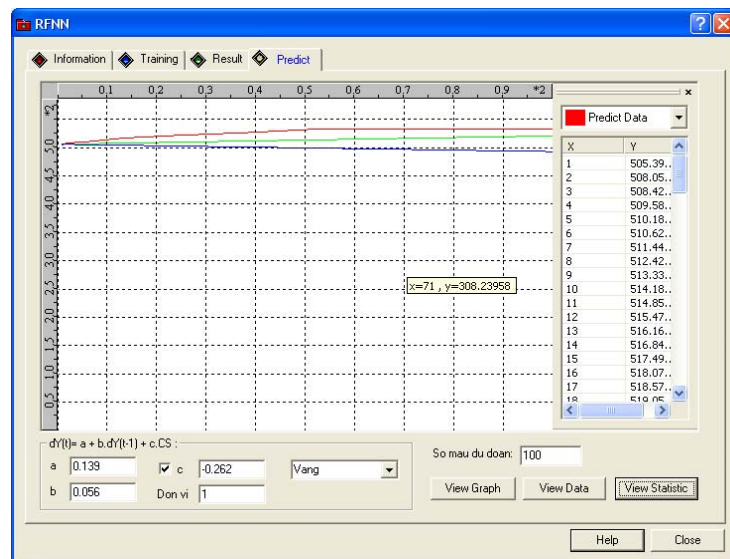
Kết quả thử nghiệm (có đối chiếu với mô hình ARIMA)



Hình 3. Kết quả huấn luyện



Hình 4. Kết quả kiểm tra



Hình 5. Kết quả dự báo

VII. KẾT LUẬN

Mạng nơ ron mờ hồi quy là một mạng kết nối nhiều tầng hồi quy cho việc thực hiện suy luận mờ sử dụng những luật mờ động. Mạng gồm có bốn lớp, trong đó có hai lớp ẩn và một lớp hồi tiếp. Quan hệ thời gian được nhúng trong mạng bằng cách xây dựng thêm các nút hồi tiếp vào một mạng nơ ron mờ và các nút hồi tiếp này đóng vai trò là các phần tử nhớ. Mạng nơ ron mờ hồi quy với khả năng lưu trữ thông tin tạm thời cho phép mở rộng phạm vi ứng dụng để giải quyết những vấn đề thời gian.

Mạng nơ ron mờ hồi quy có những điểm mạnh sau:

- Trong các ứng dụng, mạng nơ ron mờ hồi quy có cấu trúc nhỏ hơn và cũng ít các tham số lập hơn mạng nơ ron mờ.
- Mạng nơ ron mờ hồi quy có khả năng xử lý các hệ động và lưu trữ thông tin tạm thời.

- Mạng nơ ron mờ hồi quy rất thành công trong việc xấp xỉ một hàm liên tục với độ chính xác mong muốn.
- Mạng nơ ron mờ hồi quy chỉ là một mạng tĩnh như mạng nơ ron mờ khi tham số $\theta=0$. Nên có thể khẳng định rằng mạng nơ ron mờ hồi quy là một mô hình mở rộng của mạng nơ ron mờ.

Mô hình mạng nơ ron mờ hồi quy là một mô hình mới và tỏ ra đạt hiệu quả cao cho những ứng dụng như: dự báo chuỗi thời gian, nhận dạng và điều khiển những hệ phi tuyến.

TÀI LIỆU THAM KHẢO

- [1]. Nguyễn Hoàng Phương, Bùi Công Cường, Nguyễn Doãn Phước, Phan Xuân Minh, Chu Văn Hỷ, *Hệ mờ và ứng dụng*, Nhà xuất bản khoa học và kỹ thuật, 1998.
- [2]. Ching-Hung Lee and Ching-Cheng Teng, *Identification and Control of Dynamic Systems Using Recurrent Fuzzy Neural Networks*, IEEE Trans. Fuzzy Systems, Vol. 8, pp. 349 – 366.
- [3]. Chin-Teng Lin & C.S. Geogre Lee, *Neural Fuzzy Systems*, Prentice-Hall International, Inc.
- [4]. Tom M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.