

# HỌC BẰNG CÁC MÔ HÌNH MẠNG NEURAL

Vũ Thanh Nguyên

## ABSTRACT:

*This paper mentions about neural network models for learning as backpropagation neural network model, radial basis function neural network model, recurrent fuzzy neural network... These models are applied to learn and train on recognition, to analyse and to forecast for investment, finance, optimization, to evaluate for development on agriculture, industry, oriented to develop market, to set up business model...*

## I. TỔNG QUAN VỀ HỌC VÀ MÁY HỌC

Rất khó để định nghĩa một cách chính xác về học. Một trong những nguyên nhân là từ học - learn có ý nghĩa khác nhau trong từng lĩnh vực: tâm lý học, giáo dục, trí tuệ nhân tạo, ... Một định nghĩa rộng nhất, học là một từ dùng để chỉ khả năng một chương trình máy tính để tăng tính thực thi dựa trên những kinh nghiệm đã trải qua. Lợi điểm của các phương pháp học (gọi là máy học - machine learning khi kết hợp với máy tính) là nó phát sinh ra các luật tường minh, có thể được sửa đổi, hoặc được huấn luyện trong một giới hạn nhất định. Các phương pháp máy học hoạt động trên các dữ liệu có đặc tả thông tin.

Máy học là sự tự động của quy trình học và việc học thì tương đương với việc xây dựng những luật dựa trên việc quan sát trạng thái trên cơ sở dữ liệu và những sự chuyển hóa của chúng. Đây là lĩnh vực rộng lớn không chỉ bao gồm việc học từ mẫu, mà còn học tăng cường, học với thầy, ... Các thuật toán học lấy bộ dữ liệu và những thông tin quen thuộc của nó khi nhập và trả về một kết quả câu nói hay một câu ví dụ, một khái niệm để diễn tả những kết quả học. Máy học kiểm tra những ví dụ trước đó và kiểm tra luôn cả những kết quả của chúng khi xuất và học làm cách nào để tái tạo lại những kết quả này và tạo nên những sự tổng quát hóa cho những trường hợp mới.

Nói chung, máy học sử dụng một tập hữu hạn dữ liệu được gọi là tập huấn luyện. Tập này chứa những mẫu dữ liệu mà nó được viết bằng mã theo một cách nào đó để máy

có thể đọc và hiểu được. Tuy nhiên, tập huấn luyện bao giờ cũng hữu hạn do đó không phải toàn bộ dữ liệu sẽ được học một cách chính xác.

Một tiến trình học gồm 2 giai đoạn:

- **Giai đoạn học (learning):** hệ thống phân tích dữ liệu và nhận ra mối quan hệ (có thể là phi tuyến hoặc tuyến tính) giữa các đối tượng dữ liệu.
- **Giai đoạn thử nghiệm (testing):** Mối quan hệ (các luật, lớp...) được tạo ra phải được kiểm nghiệm lại bằng một số hàm tính toán thực thi trên một phần của tập dữ liệu huấn luyện hoặc trên một tập dữ liệu lớn.

Các thuật toán học chia làm 3 loại: học giám sát, học không giám sát và học nửa giám sát.

### ➤ Học có giám sát

Đây là cách học từ những mẫu dữ liệu mà ở đó các kỹ thuật học giúp hệ thống xây dựng cách xác định những lớp dữ liệu. Hệ thống phải tìm một sự mô tả cho từng lớp (đặc tính của mẫu dữ liệu). Thuật toán học có giám sát gồm tập dữ liệu huấn luyện  $M$  cặp:

$$S = \{(x_i, c_j) | i=1, \dots, M; j=1, \dots, C\}$$

Các cặp huấn luyện này được gọi là mẫu, với  $x_i$  là vector  $n$ -chiều còn gọi là *vector đặc trưng*,  $c_j$  là lớp thứ  $j$  đã biết trước.

Thuật toán học giám sát tìm kiếm không gian của những giả thuyết có thể, gọi là  $H$ . Đối với một hay nhiều giả thuyết, mà ước lượng tốt nhất hàm không được biết chính xác  $f: x \rightarrow c$ . Đối với công việc phân lớp có thể xem giả thuyết như một tiêu chí phân

lớp. Thuật toán học tìm ra những giả thuyết bằng cách khám phá ra những đặc trưng chung của những ví dụ mẫu thể hiện cho mỗi lớp. Kết quả nhận được thường ở dạng luật (*Nếu ... thì*).

Tùy thuộc vào mức độ của thuật toán học giám sát, người ta có những mô hình học giám sát như sau: *học vẹt*: hệ thống luôn luôn được dạy những luật đúng, rồi có học hội tụ, *học bằng phép loại suy*: hệ thống được dạy phản hồi đúng cho một công việc tương tự, nhưng không xác định, *học dựa trên trường hợp*: trong trường hợp này hệ thống học lưu trữ tất cả các trường hợp, cùng với kết quả đầu ra của chúng, *học dựa trên sự giải thích*: hệ thống sẽ phân tích tập hợp những giải pháp nhằm chỉ ra tại sao mỗi phương pháp là thành công hay không thành công.

### ➤ Học không giám sát

Đây là việc học từ quan sát và khám phá. Hệ thống khai thác dữ liệu được ứng dụng với những đối tượng nhưng không có lớp được định nghĩa trước, mà để nó phải tự hệ thống quan sát những mẫu và nhận ra mẫu. Hệ thống này dẫn đến một tập lớp, mỗi lớp có một tập mẫu được khám phá trong tập dữ liệu. Một thuật toán học giám sát luôn có thể biến đổi thành một thuật toán học không giám sát. Đối với một bài toán mà những mẫu dữ liệu được mô tả bởi  $n$  đặc trưng, người ta có thể chạy thuật toán học giám sát  $n$ -lần, mỗi lần với một đặc trưng khác nhau đóng vai trò thuộc tính lớp, mà chúng ta đang tiên đoán. Kết quả sẽ là  $n$  tiêu chí phân lớp ( $n$  bộ phân lớp), với hy vọng là ít nhất một trong  $n$  bộ phân lớp đó là đúng.

### ➤ Học nửa giám sát

Học nửa giám sát là các thuật toán học tích hợp từ học giám sát và học không giám sát. Việc học nửa giám sát tận dụng những ưu điểm của việc học giám sát và học không giám sát và loại bỏ những khuyết điểm thường gặp trên hai kiểu học này.

## II. CÁC MÔ HÌNH MẠNG NEURAL DÙNG ĐỂ HỌC

### 1. Mô hình mạng neural lan truyền ngược (Backpropagation Neural Network)

Các mô hình sử dụng mạng neuron là một cách tiếp cận khá phổ biến cho vấn đề nêu trên. Điều đó xuất phát từ khả năng ghi nhớ và học của cấu trúc này. Bên cạnh đó, tính ổn định của mạng neuron cũng là một yếu tố quan trọng giúp nó được chọn vì đây là điều kiện quan trọng đặt ra cho bài toán mô hình.

#### Thuật toán lan truyền ngược (BackPropagation Algorithm):

Gọi  $W_{ijk}$  là cung đi từ đỉnh thứ  $k$  của lớp  $l$  vào đỉnh  $j$  của lớp kế tiếp  $l+1$ .

Tín hiệu ra tại mỗi node trên mạng được xác định như sau:

Lớp Input và Output:  $Y_o = Y_l = X$

Lớp ẩn: 
$$Y_H = f(X) = \frac{1}{1 + e^{-X}}$$

▪ Đạo hàm tại lớp ẩn:  $Y' = f(X) \cdot (1 - f(X))$

Sai số của node output thứ  $i$  tại thời điểm

$$e_i = \frac{1}{2} (d_i - y_i)^2$$

xem xét:

Trong đó:  $d_i$ : Giá trị mong muốn nhận được tại node  $i$ ,  $y_i$ : Giá trị thực của hệ thống.

Khi cung  $W_{ijk}$  được học, giá trị của nó được cập nhật theo công thức sau:

$$w_{jk}(t+1) = w_{jk}(t) + \Delta W_{jk}$$

▪ Trong đó:

$$\begin{aligned} \Delta W_{jk} &= \mu \frac{\partial e_i}{\partial y_i} \cdot \frac{\partial y_i}{\partial x_i} \cdot \frac{\partial x_i}{\partial x_j} \cdot \frac{\partial x_j}{\partial w_{jk}} \\ &= \mu \cdot - (d_i - y_i) \cdot y_i' \cdot \frac{\partial x_i}{\partial x_j} \cdot y_k \\ &= \mu \cdot - (d_i - y_i) \cdot y_i' \cdot \delta_{ij} \cdot y_k \end{aligned}$$

▪ Với:

$$\delta_{ij} = \begin{cases} 1 & , i = j \\ w_{ij} \cdot y_j' & , i = j + 1 \\ \left( \sum_k w_{kj} \cdot y_k' \cdot w_{ik} \right) \cdot y_j' & , i = j + 2 \\ \left( \sum_p w_{pj} \cdot y_p' \cdot \left( \sum_k w_{kp} \cdot y_k' \cdot w_{ik} \right) \right) \cdot y_j' & , i = j + 3 \end{cases}$$

## 2. Mô hình hệ thống mạng neural dạng hàm radial (Radial Basis Function Neural Network).

Hệ thống mạng neural dạng hàm radial sẽ giải quyết vấn đề xấp xỉ một hàm liên tục n biến trên một miền compact [13]. Mô hình này tiến hành lấy đặc trưng cục bộ của hàm, và như vậy sẽ dễ dàng khởi tạo và huấn luyện dữ liệu khi học.

➤ **Cấu trúc mạng:** một mạng RBFNNs gồm có 3 lớp: *lớp đầu vào*, *lớp các hàm Gauss* (số nút là do người sử dụng quy định), *lớp đầu ra*

- Các liên kết từ tầng lớp đầu vào đến tầng các hàm Gauss không có trọng số. Các liên kết ở tầng các hàm Gauss đến tầng lớp đầu ra có trọng số. Mỗi node ở tầng các hàm Gauss có các thông số cần xác định là: trọng tâm, thông số sigma (xác định độ lệch chuẩn của hàm Gauss). Xác định các thông số ở tầng này dùng để phân lớp.
- Các trọng số trên đường liên kết từ tầng các hàm Gauss đến tầng lớp đầu ra được xác định thông qua cách học bình thường: phương pháp học lan truyền ngược, phương pháp học tuyến tính, phương pháp học theo vết cũ.

➤ **Hoạt động của mô hình:** mỗi mẫu dữ liệu nhập sẽ qua k hàm Gauss (giả sử ở tầng này có k node hàm Gauss), hay có thể hiểu là có k lớp, xem mẫu thuộc vào lớp nào qua tính xác suất phân bố chuẩn (là hàm Gauss của các lớp). Các giá trị tính được này được tổ hợp tuyến tính (tính trung bình có trọng số).

## 3. Mô hình mạng neural mờ hồi quy (Recurrent Fuzzy Neural Network)

Mô hình mạng neural này kết hợp từ lý thuyết tập mờ và mô hình mạng neural tận dụng những ưu điểm như có khả năng xấp xỉ một hàm liên tục với độ chính xác cho trước (mạng neural) và khai thác khả năng xử lý những tri thức như con người (lý thuyết tập mờ) [10], [11], [12], [14]. Mô hình này rất thích hợp do yêu cầu đặt ra có những đầu vào và đầu ra phụ thuộc theo thời gian. Mạng neural 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.

➤ **Cấu trúc mạng:** một mạng gồm 4 lớp như sau:

**Lớp 1:** Là lớp nhập gồm N dữ liệu nhập (input).

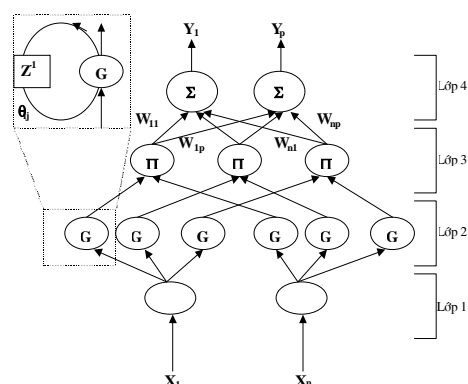
**Lớp 2:** Gọi là lớp các hàm thành viên. Các nút trong lớp này thực hiện việc mờ hóa. Số nút trong lớp 2 là  $N \times M$ , trong đó M là số luật mờ (số node của lớp 3)

**Lớp 3:** Lớp các luật mờ. Các nút trong lớp này tạo thành cơ sở luật mờ 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:  $N + (N \times M) + M + P$



**Cấu trúc mạng RFNN**

➤ **Hoạt động của mô hình**

Ký hiệu  $u_i^{(k)}$  và  $O_i^{(k)}$  tương ứng là đầu vào và đầu ra của nút thứ  $i$  trong lớp  $k$ .

**Lớp 1**

$$O_i^{(1)} = u_i^{(1)} = x_i(t), i = 1 \div N$$

**Lớp 2**

$$O_{ij}^{(2)} = \exp \left[ - \frac{(u_{ij}^{(2)} - m_{ij})^2}{(\sigma_{ij})^2} \right]$$

$i = 1 \div N, j = 1 \div M$

Trong đó  $m_{ij}$  và  $\sigma_{ij}$  tương ứng là trọng tâm và độ rộng của hàm thành viên theo phân bố Gauss. Hơn nữa, đầu vào của các nút này là:

$$u_{ij}^{(2)}(t) = O_i^{(1)} + \theta_{ij} O_{ij}^{(2)}(t-1)$$

$i = 1 \div N, j = 1 \div M$

Trong đó  $\theta_{ij}$  biểu diễn trọng số cho các nút hồi tiếp.

Để dàng thấy rằng đầu vào của các node trong lớp này có chứa toán hạng  $O_{ij}^{(2)}(t-1)$  lưu thông tin trước đó của mô hình. Và đây chính là sự khác biệt giữa mạng neural mờ và mạng neural hồi quy mờ.

Như vậy

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

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

$i = 1 \div N, j = 1 \div M$

Mỗi node trong lớp này có 3 thông số là  $m_{ij}$ ,  $\sigma_{ij}$  và  $\theta_{ij}$ .

**Lớp 3:** Các node trong lớp này thực hiện phép toán AND

$$O_j^{(3)} = \prod_{i=1}^N O_{ij}^{(2)}$$

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

$i = 1 \div N, j = 1 \div M$

**Lớp 4:** Các node 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 này, các thông số cần phải xác định là  $m_{ij}$ ,  $\sigma_{ij}$ ,  $\theta_{ij}$  và  $w_{jk}$ .

➤ **Lập luận mờ**

Giả sử cho hệ mạng neural hồi quy mờ với nhiều đầu vào và một đầu ra. Gọi  $x_i$  là biến ngôn ngữ thứ  $i$  và  $\alpha_j$  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  $i = 1, 2, \dots, n$  là số đầu vào

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

$A_{ij}$  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  $o_{ij}^{(2)}\theta_{ij}$ .

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:

$$y^* = \sum_{j=1}^M w_j \alpha_j \text{ và } \alpha_j = \prod_{i=1}^N \mu A_{ij}(u_{ij})$$

Với M là số luật. Từ mô tả trên, mạng neural hồi quy mờ 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.

### III. KẾT LUẬN

Các phương pháp học trên mô hình neural đã được tác giả và các cộng sự nghiên cứu và áp dụng rất nhiều trong thực tiễn (có kết hợp với các thuật toán khác như GA, hệ lai...) như xây dựng hệ thống quản trị tài chính, dự báo chứng khoán [6], xây dựng mô hình hỗ trợ ra quyết định marketing [8], phân tích dự báo kinh tế [1],[2],[9], truy vấn bản tự động [7], phân tích dự báo các sản phẩm chiến lược nền kinh tế quốc dân [9], nhận dạng chữ tiếng Việt... Hiện nay tác giả đang xây dựng và thử nghiệm trên mô hình mạng neural mờ hồi quy và mạng neural dạng hàm radial hệ thống đánh giá chất lượng giảng dạy và học cho các trường phổ thông trung học, cao đẳng và đại học trên địa bàn Tp. HCM.

### IV. TÀI LIỆU THAM KHẢO

1. TS. Vũ Thanh Nguyên; Báo cáo nghiệm thu đề tài giải quyết một số vấn đề phân tích dự báo kinh tế ứng dụng trong ngành công nghiệp tại thành phố Hồ Chí Minh của Sở Khoa Học và Công Nghệ Tp. HCM 10/2003.
2. ThS. Nguyễn Quốc Tòng; Báo cáo về các phương pháp định tính và định lượng được ứng dụng trong các công tác phân tích dự báo của Viện Kinh Tế thành phố - 2000.
3. Duc Truong Pham, Liu Xing; Neural Networks for Identification, Prediction and Control. Springer – Verlag London Limited 1998.
4. T.T. Chow, Z.Lin and C.L. Song. Applying Neural Network and Genetic Algorithm In System Optimization. 7th International IBPSA Conference, 2001.
5. Trương Hải Bằng. Ứng dụng của hệ thống lai trong quản trị tài chính. Luận văn Thạc Sĩ CNTT - ĐHKHTN , 07/2001.
6. Nguyễn Phước Đại. Nghiên cứu một số thuật toán máy học và ứng dụng trong truy vấn văn bản tự động. Luận văn Thạc Sĩ CNTT - ĐHKHTN, 06/2003.
7. Võ Nguyễn Hương Giang. Sử dụng hệ lai xây dựng mô hình hỗ trợ ra quyết định marketing. Luận văn Thạc Sĩ CNTT - ĐHKHTN, 07/2004.
8. Lư Nhật Vinh. Kết hợp mạng neural và logic mờ để giải quyết bài toán kinh tế. Luận văn Thạc Sĩ CNTT - ĐHKHTN, 07/2001.
9. Dương Ngọc Hiếu. Sử dụng hệ lai (Fuzzy Logic+Neural Network, Genetic + Neural Network) trong công tác phân tích dự báo các sản phẩm chiến lược của nền kinh tế quốc dân. Luận văn Thạc Sĩ CNTT - ĐHBK, 07/2005.
10. C-H. Lee, C.-C Teng; Identification and Control of Dynamic Systems Using Recurrent Fuzzy Neural Networks; p.349 – p.366. IEEE Transactions on Fuzzy System 08/2000.
11. F-J.Lin, R.-J. Wai; Hybrid Control Using Recurrent Fuzzy Neural Network for Linear Induction Motor Servo Drive; p.102 – p.115; IEEE Transactions on Fuzzy System 02/2001.
12. P.Tino, C.Schittenkopf, G.Dorffner. Financial Volatility Trading Using Recurrent Neural Network; p. 865 – p.874; IEEE Transactions on Neural Networks 07/2001.
13. S. Seshagiri, H. K. Khalil; Output Feedback Control of Nonlinear Systems Using RBF Neural Networks; p. 69 – p.79; IEEE Transactions on Neural Networks 01/2000.
14. F.-J.Lin, R.-J Wai, C.-M.Hong; Hybrid Supervisory Control Using Recurrent Fuzzy Neural Network for Tracking Periodic Inputs; p.68 – p.90; IEEE Transactions on Neural Networks 01/2001.