

ỨNG DỤNG CỦA MẠNG THẦN KINH NHÂN TẠO (ANN) XỬ LÝ DỮ LIỆU KHÍ TƯỢNG THỦY VĂN

ThS. Lê Xuân Cầu

Trung tâm tư liệu khí tượng thủy văn

Bài báo này sẽ đề cập những vấn đề sau:

1. *Mạng thần kinh nhân tạo làm được những gì?*
2. *Mạng thần kinh nhân tạo là gì? Một số mạng thần kinh nhân tạo thường dùng và các tính chất của chúng*
3. *Ví dụ áp dụng mạng thần kinh nhân tạo trong nghiên cứu thủy văn.*
4. *Kết luận.*

I. GIỚI THIỆU CHUNG

Chúng ta thường xây dựng các mô hình để nghiên cứu hiện tượng vật lý. Các mô hình này được xây dựng bằng các ngôn ngữ riêng. Mỗi một ngôn ngữ là một tập các ký hiệu với sự biểu diễn vật lý xác định. Như vậy, mô hình được hình thành trên cơ sở hệ thống các ký hiệu vật lý. Nghiên cứu kỹ các mô hình đó ta thấy rằng thậm chí khi mô hình cho khả năng tổng quát hoá cao chúng vẫn có những giới hạn nhất định khi đem ứng dụng. Sự sử dụng hệ thống các ký hiệu vật lý tuy có một ý nghĩa đối với sự cảm nhận của con người, song chúng phải thoả mãn các điều kiện ràng buộc nào đó. Trong nhiều trường hợp các điều kiện ràng buộc đó là không thể chấp nhận được. Để thoát ra khỏi các hạn chế của cách tiếp cận ký hiệu đó người ta dùng cách tiếp cận bán ký hiệu. Trong cách tiếp cận bán ký hiệu người ta dùng dấu hiệu để chỉ đến đối tượng. Ta thiết lập các quy tắc về sự tương tác của các dấu hiệu. Chúng ta không nhận thức được sự tương tác đó và cũng không đoán trước được kết quả của sự tương tác đó. Trong lĩnh vực trí tuệ nhân tạo mạng thần kinh nhân tạo (ANN) là một mô hình theo cách tiếp cận bán ký hiệu, còn hệ chuyên gia là một ví dụ về cách tiếp cận xử lý ký hiệu [3].

ANN (Artificial neural network) đã được ứng dụng để thiết lập một quan hệ hàm bất kì giữa các biến đầu ra (outputs) và các biến đầu vào (inputs) của một hệ thống. ANN được dùng để giải các bài toán phân lớp, nhận dạng hay dự báo. Trong thực tế xử lý thông tin ANN đã cho kết quả tốt khi đem ứng dụng trong một số lĩnh vực. Chẳng hạn, ANN được dùng làm mắt nhiều trong điện thoại khi xử lý tín hiệu. Người ta đã dùng ANN trong điều khiển học (tự điều khiển ôtô, robot) và nhận dạng (nhận ra chữ viết tay). ANN đã dùng trong y học để chẩn đoán và chữa bệnh. ANN còn được dùng trong lĩnh vực nghiên cứu tạo tiếng nói hoặc nhận biết tiếng nói [5].

Trong nghiên cứu khí tượng thủy văn, ANN được dùng như là một công cụ gần đúng tổng quát của các biến đầu ra và các biến đầu vào. Ngoài các bài toán phân lớp, tự phân lớp và dự báo ANN còn có thể học một mô hình bất kỳ (mô hình hoá một mô hình). ANN kết hợp với các quá trình tiên xử lý khác nhau giúp ta giải quyết các bài toán phải xử lý những khối lượng thông tin lớn với hàng trăm, hàng nghìn đầu vào, đầu ra. ANN có thể xử lý thông tin từ các máy đo đạc từ xa (từ vệ tinh hay trạm radar), hay từ các trạm đo tự động.

Trong tin học thủy văn đã có một số ứng dụng của ANN, chẳng hạn Minns [3] đã dùng ANN để mô phỏng quan hệ mưa rào-dòng chảy. Solomatine [7] cho ANN học một mô hình thủy động lực trong tối ưu hoá hoạt động của hồ chứa. Van den Boogaard [8] xây dựng một mô hình hỗn hợp bằng cách kết hợp ANN với mô hình số

tri. Câu [1] đã tìm cấu hình của ANN bằng lý thuyết ‘cực tiểu hoá mạo hiểm trung bình theo cấu trúc’ và ứng dụng của nó trong giải các bài toán thiết lập các mối quan hệ.

ANN là kết quả nghiên cứu của các chuyên gia từ các lĩnh vực khác nhau. Các nhà thần kinh học nghiên cứu các hệ thống mạng thần kinh trong các vùng não của các động vật khác nhau. Các nhà vật lý nghiên cứu sự tương tự giữa xu thế động lực của hệ tương tự não và các hệ động lực phi tuyến đã biết. Các kỹ sư máy tính nghiên cứu việc tạo ra máy tính có khả năng xử lý thông tin tương tự bộ não (xử lý song song). Người nghiên cứu trí tuệ nhân tạo muốn tạo ra các máy với sự thông minh của động vật. Các nhà toán học nghiên cứu các vấn đề toán học của hệ thần kinh nhân tạo. Các nhà tâm lý học nghiên cứu cơ chế xử lý thông tin của con người. Các kỹ sư nghiên cứu sử dụng ANN trong việc giải quyết các bài toán thực tế.

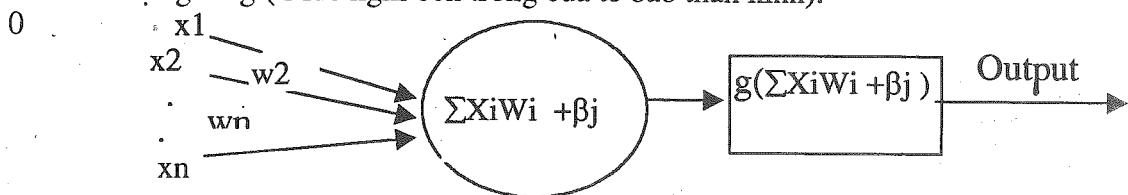
Trong ANN, dữ liệu được biến đổi ở mức biểu diễn số của máy tính. Các số trong máy tính được biểu diễn bởi các bít, bởi vậy các phép toán biến đổi dữ liệu sẽ thực hiện trên các bít riêng biệt. Người làm mô hình không có một ảnh hưởng trực tiếp nào tới các bít mang các thông tin. Như vậy, sau khi biến đổi dữ liệu mà nó mô tả ký hiệu của thế giới tự nhiên thành các dòng bít thì các ký hiệu nguyên thủy không còn có một liên quan nào tới các phép toán trên các bít nữa. Để hiểu và dùng ANN trong nghiên cứu ta cần biết các khái niệm cơ bản của nó.

II. ANN LÀ GÌ ?

Bộ não có thể thực hiện các hoạt động phức tạp và nhận ra được các mẫu rất phức tạp (thậm chí cả khi các mẫu đó có “nhiều”). Bộ não nhận được dòng thông tin từ môi trường bên ngoài. Trên cơ sở dòng thông tin đó, các luật về xu thế và về các đặc trưng của môi trường được hình thành, đó chính là các biểu diễn bên trong bộ não về môi trường bên ngoài (các nét đặc trưng của môi trường). Các biểu diễn bên trong đó cho phép dự đoán trạng thái tương lai của môi trường để sinh vật có thể sinh sống. **Tính chất quan trọng của não là nó có thể học một cái gì đó từ các ví dụ và tổng quát hóa các tính chất đã học được từ các ví dụ.**

ANN là mô hình toán của một hệ tương tự não. Mô hình lấy neuron làm đơn vị xử lý cơ bản. Mỗi neuron được đặc trưng bởi:

- Mức hoạt động(trạng thái phân cực của neuron),
- Giá trị ra (Tốc độ phát hoả của neuron),
- Tập nối các đầu vào của neuron (khớp nối các tế bào thần kinh),
- Giá trị ngưỡng (Mức nghỉ bên trong của tế bào thần kinh).



H.1 Đơn vị xử lý cơ bản (neuron) của ANN

Ở đây: x_1, x_2, \dots, x_n là các đầu vào của neuron thần kinh,

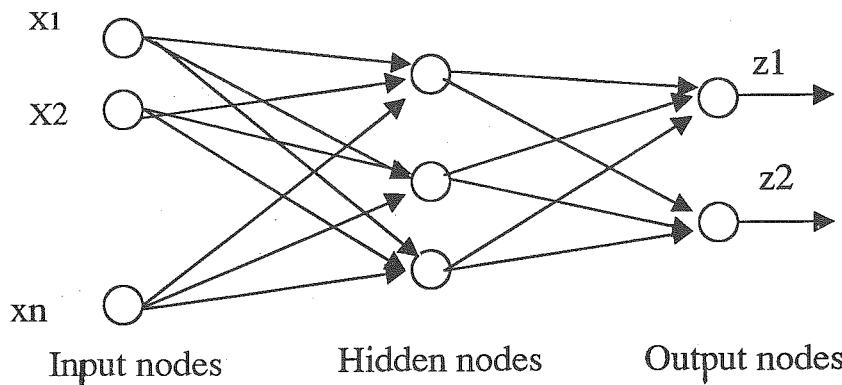
w_1, w_2, \dots, w_n : các trọng số tương ứng với các đầu vào,

β_j : Giá trị ngưỡng của neuron,

$g()$ là hàm chuyển đổi (thường là các hàm nén : hàm lôgíc).

Rosenblatt [4] đã chứng minh một kết quả quan trọng của mô hình nhân tạo neuron (perceptron). Định lý hội tụ của perceptron như sau: Nếu một hàm được tính bởi Perceptron thì Perceptron có thể được dạy để học hàm đó.

Mạng thần kinh nhân tạo đa lớp gồm ít nhất 3 lớp các nút (nút là các neuron giả lập). Lớp đầu vào là lớp thu động cho phép đưa vào ANN các ví dụ để học. Tiếp đến là lớp ẩn (lớp mà các nút của nó không có liên hệ trực tiếp với các lớp bên ngoài). Lớp ẩn tính các tổ hợp phi tuyến của các đầu vào. Lớp đầu ra tính các tổ hợp phi tuyến của đầu ra của các nút ẩn [2].



H.2. Mạng thần kinh nhân tạo đa lớp
(ANN với 1 lớp ẩn)

Ở đây: x_1, x_2, \dots, x_n là các đầu vào của ANN,
 y_1, y_2, \dots , là các đầu ra của ANN.

Mạng thần kinh nhân tạo BPN (Error Backpropagation Network) (H.2) thường hay được dùng hơn cả [6]. Các tín hiệu đầu vào ANN sẽ được phân bố giữa các nút ẩn. Sau đó các nút ẩn sẽ biến đổi chúng thành các tín hiệu ra, những tín hiệu này sẽ được truyền cho các đầu ra của ANN.

N_{inp} các đầu vào của ANN được ký hiệu: $x_i, i=1, \dots, N_{inp}$
 N_{hid} các đầu ra y_j của lớp ẩn sẽ là:

$$j_1, \dots, N_{hid}$$

N_{out} các giá trị ra của các đầu ra sẽ là:

$$Z_{ki} = g\left(b_{0k} + \sum_{j=1}^{N_{hid}} b_{jk} y_j\right), \quad k=1, \dots, N_{out}$$

Hàm chuyển đổi $g(u)$ giới hạn giữa 0 và +1 là hàm lôgic sau:

$$g(u) = \frac{1}{1+e^{-u}}$$

Như vậy, ANN sẽ có tất cả $(N_{inp}+1) \cdot N_{hid} + (N_{hid}+1) \cdot N_{out}$ tham số (Các trọng số a và b). Các giá trị của các tham số a và b sẽ nhận được bởi quá trình dạy ANN sao cho sai số nhận được từ ANN nhỏ nhất.

Gọi các giá trị quan trắc được của đầu ra OBS_i và các giá trị đầu ra của ANN là MOD_i khi đó sai số quân bình bình phương MSE của ANN sẽ là:

$$MSE = \frac{1}{N_{\text{examp}} \cdot N_{\text{out}}} \sum_{i=1}^{N_{\text{examp}} \cdot N_{\text{out}}} (OBS_i - MOD_i)^2$$

ở đây : N_{examp} là số các ví dụ để ANN học.

Bởi vì ta biết MSE có thể biểu diễn bằng giải tích qua các biến đầu vào và các tham số , thêm vào đó MSE là hàm khả vi cho nên ta có thể tìm giá trị của các trọng số a và b sao cho MSE đạt giá trị cực tiểu. Phương pháp gradien (độ dốc lớn nhất) thường được dùng để tìm các trọng số của ANN.

Ngày nay ngoài mạng thần kinh BPN có dạng như trên người ta còn thường hay dùng các mạng thần kinh như RBFN (Radial Basic Function Net) và SOM (Self-organization Map), LVQ (Learning Vector Quatizer) hoặc RNN (Reccurent Neural Network) để xử lý thông tin.

ANN có một số tính chất quan trọng sau:

1. ANN có thể gần đúng của một quan hệ hàm bất kỳ,
2. ANN có thể học từ các dữ liệu,
3. ANN có khả năng tổng quát hoá những cái đã học cho những dữ liệu chưa biết.

Các tính chất trên được suy ra từ định lý về “ Sự tồn tại mạng thần kinh ánh xạ Kolmogorov ” : Một hàm liên tục bất kỳ $f(x_1, x_2, \dots, x_n)$ xác định trên khoảng I^n ($I = [0,1]$) có thể được biểu diễn dưới dạng:

$$f(x) = \sum_{j=1}^{2n+1} \chi_j \left(\sum_{i=1}^n \Psi_{ij}(x_i) \right)$$

Ở đây: χ_j , Ψ_{ij} là các hàm liên tục một biến. Ψ_{ij} là hàm đơn điệu ,không phụ thuộc vào hàm f .

Một tính chất ưu việt cơ bản của mạng thần kinh nhân tạo là nó có một thuật toán để học đơn giản nhưng rất mạnh cho phép ANN có thể phù hợp với môi trường bên ngoài.

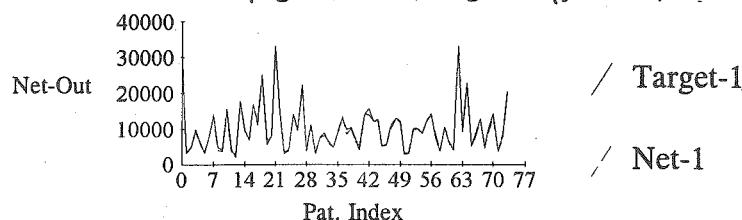
III. VÍ DỤ ÁP DỤNG MẠNG THẦN KINH NHÂN TẠO TRONG NGHIÊN CỨU THỦY VĂN

Lưu lượng nước Q của dòng chảy không ổn định trong sông là một hàm phụ thuộc nhiều yếu tố. Nó phụ thuộc vào mực nước H, độ dốc mặt nước I, cường suất mực nước dH/dt , sức cản thủy lực (hệ số sêzi), v...v.. Quan hệ này, phi tuyến và rất phức tạp. Từ công thức Jone, ta thấy rằng: $Q = f(H, dH/dt)$. Khả năng gần đúng tổng quan của ANN được thể hiện khi tiến hành xác định các vòng lũ cho sông Hồng, trạm Sơn Tây, năm 1971. ANN được áp dụng ở đây có 2 đầu vào là các biến H và dH/dt , 4 nút ẩn, 1 đầu ra (biến lưu lượng Q). Trong mùa lũ tại trạm Sơn Tây có 93 lần đo lưu lượng. Tính toán được tiến hành theo hai cách:

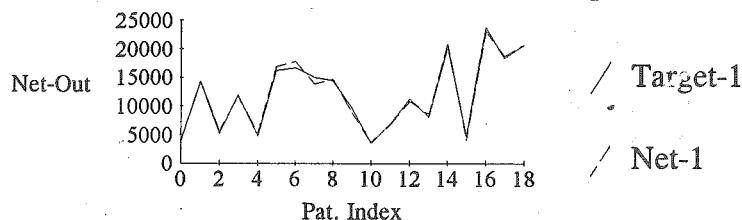
- 1.Từ tập số liệu quan trắc có được (93 ví dụ) hai tập con sẽ được tạo ra một cách ngẫu nhiên. Tập thứ nhất gồm 74 ví dụ (vectô các giá trị đầu vào và đầu ra) dùng để dạy ANN. Tập thứ hai gồm 19 ví dụ dùng để kiểm tra độc lập kết quả tính (các ví dụ này không tham gia trong việc tìm các trọng số của ANN).

2.Tập số liệu quan trắc (93 ví dụ) dùng để dạy ANN. Tập số liệu trích lũ bằng phương pháp thủ công sẽ được dùng để so sánh với kết quả trích lũ bằng ANN.
Kết quả tính toán cho cách thứ nhất được chỉ ra ở hình 3 và hình 4:

H.3. Lưu lượng thực đo (dùng để dạy ANN)



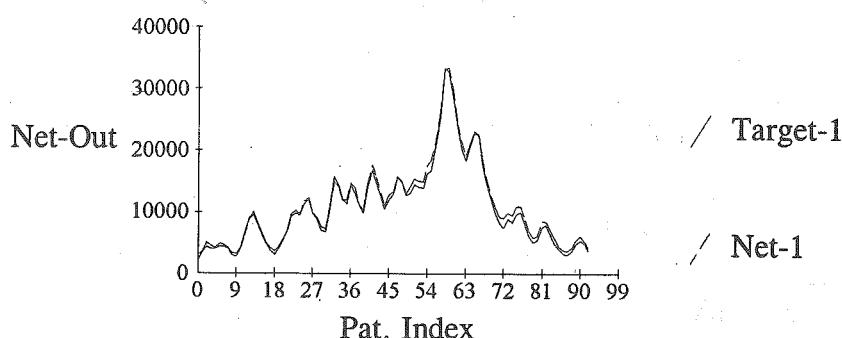
.4. Lưu lượng thực đo(dùng để kiểm tra kết quả của ANN)



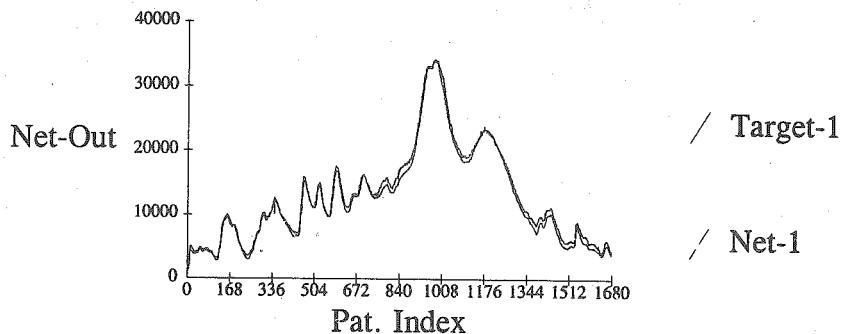
Các giá trị Target là giá trị lưu lượng nước thực đo. Các giá trị Net là các giá trị lưu lượng nước tính toán từ ANN. Các kết quả của việc dạy và kiểm tra ANN cho thấy các số liệu thực đo và tính toán sai số nhỏ. Tính tổng quan (kết quả áp dụng ANN cho tập số liệu hoàn toàn mới) của ANN rất tốt. ANN nhận được có thể dùng để trích lũ.

Kết quả tính toán cho cách thứ hai được thể hiện ở hình 5 và hình 6:

H.5. Lưu lượng thực đo (dùng để dạy ANN)



H.6. Lưu lượng trích lũ (Dùng để so sánh kết quả ANN)



Kết quả cho thấy đường trích lũ bằng thủ công hầu như trùng với đường trích lũ bằng ANN.

IV. KẾT LUẬN

2. ANN là công cụ xử lý thông tin mới, nó cho phép giải quyết những bài toán rất phức tạp (phân lớp, tự phân lớp, nhận dạng và gần đúng các biến đầu ra, đầu vào) với hàng nghìn đầu vào, đầu ra. ANN có thể xử lý những khối lượng thông tin rất lớn và tốc độ xử lý nhanh (xử lý song song).
3. ANN rất linh hoạt khi dùng để giải các bài toán khác nhau, đầu tư ít song hiệu quả cao.
4. Trong bài này tác giả chỉ giới thiệu những khái niệm cơ bản của một ANN thông dụng mà không đề cập tới các vấn đề liên quan tới cách dạy và học của ANN cũng như việc tìm cấu trúc tối ưu của ANN để sự thể hiện của mô hình là tốt hơn cả.

TÀI LIỆU THAM KHẢO

1. Cau, L. X. (1998). Establishment of relationships through theory of Structural Risk Minimisation. Individual Study. IHE, Delft, The Netherlands.
2. Laurene Fausett. (1994). Fundamentals of Neural Networks. Prentice Hall International, Inc.
3. Minns, A. W. (1998). Artificial Neural Networks as Subsymbolic Process Descriptors, PhD, thesis, Balkema, Rotterdam.
4. Rosenblatt, F. Principle of Neurodynamics. Spartan, New York, 1962 .
5. D. E. Rumelhart, B. Widrow, M.A. Lens. The basis ideas in Neural Networks. Communication of the ACM march 1994/Vol.37, No.3 .
6. M. Smith. (1993) Neural networks for statistical modelling. Van Nostrand Reyngold, N. Y.
7. D. P. Solomatine, L. A. Avila Torres. (1996) Neural network approximation of a hydrodynamic model in optimizing reservoir operation. Hydroinformatics'96, Muller (ed.).
8. H.F.P Van den Boogaard & A. C. H. Kruisbrink. Hybrid modelling by integrating neural networks and numerical models . Hydroinformatics'96, Muller (ed.), 1996.