

ỨNG DỤNG MẠNG NƠON NHÂN TẠO TRONG DỰ BÁO ỚNG BỂ TRÊN MẠNG LƯỚI CẤP NƯỚC

Phạm Thị Minh Lành¹, Hồ Đắc Bảo², Nguyễn Quang Trưởng²

Tóm tắt: Rò rỉ nước trên mạng lưới vẫn đang là vấn đề cần giải quyết của các công ty cấp nước, một trong những nguyên nhân chính của hiện tượng này là do ống vỡ. Các mô hình dự báo nguy cơ ống vỡ cũng không ngừng được cải thiện để dự báo vị trí ống vỡ một cách chính xác và nhanh chóng. Với mục tiêu phát triển một mô hình tốt hơn và có thể ứng dụng vào thực tế, trong nghiên cứu này đã đề xuất sử dụng mô hình mạng nơon nhân tạo để dự báo ống vỡ. Mô hình đề xuất được kiểm chứng bằng mạng lưới cấp nước DMA17 thuộc phường 17 quận Gò Vấp Thành phố Hồ Minh, đánh giá kết quả dự báo bằng chỉ số AUC đã cho thấy mô hình đề xuất đạt chất lượng rất tốt.

Từ khóa: Cấp thoát nước, mạng nơon nhân tạo, thất thoát nước, ống bể vỡ, mô hình dự báo.

1. ĐẶT VẤN ĐỀ

Trong quá trình hoạt động của các mạng lưới cấp nước luôn xảy ra hiện tượng nước bị thất thoát một lượng nhất định. Có rất nhiều nguyên nhân dẫn đến thất thoát nước trong đó chủ yếu là do nguyên nhân kỹ thuật như: rò rỉ trên mạng lưới tuyến ống cấp nước, thi công không đúng kỹ thuật, vỡ ống do đào đường, Trên thực tế cho thấy thất thoát nước phần lớn là do rò rỉ trên đường ống vì đường ống cấp nước được chôn ngầm dưới nền đất nên công tác tìm kiếm rò rỉ rất khó khăn. Đây là vấn đề đặt ra cho tất cả các đơn vị cấp nước trên thế giới cũng như tại Việt Nam. Để phát hiện sớm và loại bỏ nguy cơ rò rỉ nước trên mạng lưới cấp nước cần sử dụng mô hình dự báo nguy cơ ống vỡ.

Mạng nơon đã bắt đầu xuất hiện từ những năm 1940 tuy nhiên mới bắt đầu ứng dụng nhiều trong khoảng 10 năm trở lại đây và ngày một phát triển rộng rãi hơn. Đặc biệt là trong lĩnh vực dự báo và phân loại dữ liệu, mạng nơon nhân tạo đã cho thấy ưu điểm nổi bật khi xử lý các dữ liệu có tính chất phức tạp, khối lượng dữ liệu lớn. Đồng thời kết quả mô hình ứng dụng mạng nơon

nhân tạo hỗ trợ quá trình ra quyết định chính xác và có độ tin cậy cao. Vậy nên trong nghiên cứu này sẽ xem xét việc ứng dụng mạng nơon nhân tạo để xây dựng mô hình dự báo ống vỡ trên mạng lưới cấp nước.

2. TỔNG QUAN VỀ MÔ HÌNH DỰ BÁO ỚNG VỠ

Đã có nhiều nghiên cứu tìm hiểu nguyên nhân vỡ ống trên mạng lưới phân phối để từ đó dự báo khả năng vỡ ống, theo thời gian các phương pháp cũng như nội dung nghiên cứu được mở rộng và xem xét vấn đề toàn diện hơn.

2.1. Các nhân tố ảnh hưởng tới sự kiện ống vỡ

Trong nghiên cứu này sẽ kế thừa kết quả của tác giả (P.T.M. Lành ntk., 2018), để xây dựng mô hình dự báo ống vỡ cần các biến sau: các đặc trưng vật lý của hệ thống (Đường kính ống-D, Chiều dài ống-L, vật liệu ống-Mat), độ tuổi (A) đại diện cho tỉ lệ ăn mòn và đường kính tuyến ống làm việc (RD). Ngoài ra các đại lượng đặc trưng cho đặc điểm hình học của mạng lưới bao gồm vị trí tuyến đường đặt ống (R), số đoạn ống kết nối trên một tuyến ống (N0), và số lần vỡ trước đây (Prior).

2.2 Mô hình dự báo ống vỡ

Một số nghiên cứu trên thế giới đã đề xuất các mô hình thường dùng trong xác suất thống kê để dự báo sự kiện ống vỡ như sau:

¹ Khoa Kỹ thuật hạ tầng đô thị, ĐH Kiến trúc Tp. Hồ Chí Minh

² Khoa Kỹ thuật Xây dựng, ĐH Bách Khoa Tp. Hồ Chí Minh

Mô hình số mũ: Nhận định ban đầu cho rằng ống vỡ chủ yếu là do thành ống bị ăn mòn và vì vậy tuổi ống được sử dụng để đánh giá. Shamir & nnk. (1979) đã đưa ra mô hình số mũ để mô tả hiện tượng này. Nhược điểm của phương pháp này là khi xem xét số điểm vỡ trên một đoạn ống cấp nước thì các ống truyền dẫn có chiều dài lớn sẽ có số lần vỡ nhiều và các đoạn ống nhánh phân phối có chiều dài nhỏ và tỉ lệ vỡ ống cao nhưng lại đưa ra kết quả số lần vỡ ít hơn nhiều. Như vậy, cần phải xem xét tới các đặc điểm hình học của mạng lưới cấp nước nhiều hơn, như số điểm đầu nối trên một tuyến ống hay tuyến đường mà ống làm việc (vì môi trường làm việc khác nhau thì độ ăn mòn cũng khác nhau).

Mô hình hồi quy Bayesian: G. Kabir (2016) đã xây dựng mô hình hồi quy Bayesian trên cơ sở lý thuyết tính xác suất của sự kiện ống vỡ. Kết quả cho thấy ảnh hưởng của các tính chất vật lý ống dẫn quan trọng hơn tác động của môi trường đặt ống nhưng xác suất dự báo phụ thuộc nhiều vào xác suất tiên định, nghĩa là những tuyến ống xuất hiện vỡ nhiều lần thì xác suất xuất xảy ra vỡ trên những tuyến có đặc điểm tương tự sẽ cao hơn. Với các biến không dẫn đến sự cố vỡ ống trong lịch sử thì mô hình dự báo xác suất không chính xác.

Mô hình hồi quy logistic tổng quát: Mô hình hồi quy tuyến tính tương đối đơn giản, dễ sử dụng tuy nhiên kết quả chỉ đạt độ chính xác đến 70% (T. Wengström, 1993). Khả năng vỡ ống trong mô hình được đánh giá qua tỉ lệ vỡ ống trong 1 năm/1km. Để tăng độ chính xác của kết quả dự báo, các nghiên cứu gần đây sử dụng mô hình hồi quy logistic tổng quát với biến số ở dạng nhị phân 0|1 biểu thị khả năng ống không vỡ|vỡ.

Mô hình mạng trí tuệ nhân tạo: Một số nghiên cứu như H. Al-barqawi & nnk. (2008) tiếp cận theo phương pháp mạng nơron nhân tạo (Artificial Neural Network- ANN) để xác định hiệu suất làm việc của hệ thống cũng như tỉ lệ vỡ ống trong thời gian làm việc, ngoài ra còn đề xuất đưa phân tích phân tầng (AHP) vào mô hình ANN để tăng hiệu suất mô hình (D. Achim & nnk., 2007). Dữ liệu khảo sát được phân loại thành các cấp độ khác nhau theo các tiêu chí nhất định, sau đó đánh giá

trọng số của mỗi yếu tố và phân tích tổng hợp để đưa ra giá trị đóng góp trung bình của mỗi yếu tố vào sự kiện ống vỡ. Một trong những nhược điểm của nghiên cứu này là thời gian chạy mô hình ANN lớn và chi phí khảo sát số liệu đầu vào khá cao, nên mô hình mới chỉ đáp ứng cho nghiên cứu học thuật.

Kế thừa và phát triển các nghiên cứu trước đây kết hợp với mục tiêu xây dựng một mô hình dự báo có chất lượng tốt, độ tin cậy cao phù hợp với điều kiện Việt Nam, tác giả đề xuất như sau:

- Đánh giá trực tiếp khả năng xuất hiện điểm vỡ trên ống không qua tỉ lệ vỡ trên 1km như nghiên cứu mô hình số mũ của tác giả (U. Shamir nnk., 1979).

- Tương tự như nghiên cứu về mô hình hồi quy logistic, tác giả tiếp tục sử dụng số nhị phân để đánh giá hiện tượng ống vỡ|không vỡ. Bên cạnh đó, để cải thiện kết quả của mô hình Bayesian, nghiên cứu sẽ sử dụng ANN để tăng tính ngẫu nhiên của mô hình.

- Nhược điểm về hiệu suất của mô hình ANN mà nghiên cứu trước đây đã đưa ra thì với sự phát triển công nghệ thông tin ngày nay hoàn toàn có thể khắc phục được vậy nên trong nghiên cứu này tiếp tục sử dụng ANN để xây dựng mô hình.

3. ỨNG DỤNG MÔ HÌNH ANN DỰ BÁO ỐNG VỠ

3.1 Mạng nơron nhân tạo (Artificial Neural Network - ANN)

Xuất phát từ ý tưởng mô phỏng bộ não của con người, mạng nơron nhân tạo (ANN) bao gồm nhiều nơron kết nối với nhau để xử lý thông tin qua các thuật toán. Một nơron có rất nhiều sợi nhánh để tiếp nhận nhiều dữ liệu cùng lúc, tùy thuộc vào từng bài toán mà các tín hiệu có thể được trao đổi qua lại giữa các nơron. Giả sử nơron i có n tín hiệu đầu vào, mỗi tín hiệu x_i được gán một trọng số W_i trở thành $x_i W_i$ là một đại lượng vô hướng. Tổng Σ_{ii} (Hình 1) là đối số duy nhất của hàm kích hoạt f_{ij} . Tín hiệu đầu ra y_i là đại lượng vô hướng. Hàm kích hoạt có thể là hàm tuyến tính, hàm ngưỡng, hàm Sigmoid, hàm Sigmoid ngược... Hàm Sigmoid thuận lợi cho thuật toán lan truyền ngược vì dễ lấy đạo

hàm, giúp giảm đáng kể tính toán trong quá trình huấn luyện, thích hợp cho các bài toán có đầu ra mong muốn rơi vào khoảng $[0,1]$ nên được lựa chọn trong nghiên cứu này.

Mạng nơron nhân tạo thường có hai loại chính là mạng truyền thẳng và mạng hồi quy. Tín hiệu truyền theo một chiều từ trái qua phải được gọi là mạng truyền thẳng, mạng có cấu trúc cũng như thuật toán đơn giản nên đạt hiệu suất cao. Trong mạng này không có liên kết mở rộng từ đơn vị đầu ra tới đầu vào nhưng mạng nơron hồi quy lại có khả năng này. Trong bài toán dự báo ồng vỡ, các thuộc tính của ồng là độc lập, không có mối liên kết thông tin giữa các ồng vậy nên mạng nơron truyền thẳng sẽ phù hợp cho bài toán này.

3.1.1. Mạng nơron nhân tạo truyền thẳng (Feedforward Artificial Neural Network)

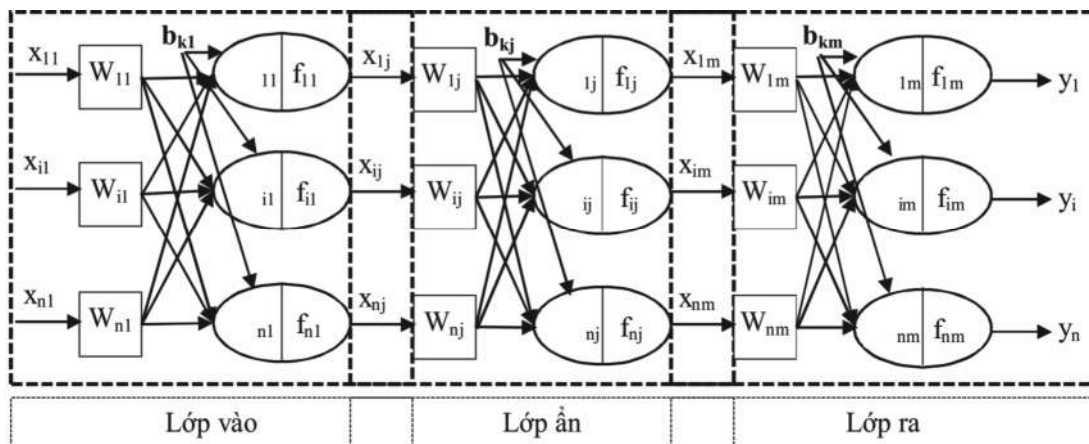
Mạng nơron thường có ba lớp là lớp vào, lớp ẩn và lớp ra như Hình 1, trong đó:

- Lớp vào làm nhiệm vụ nhận tín hiệu, với mỗi nơron là một tín hiệu đầu vào. Thông tin ở lớp vào x_{il} có thể là một hoặc nhiều yếu tố x .
- Lớp ẩn có nhiệm vụ xử lý thông tin đầu vào và đưa kết quả xử lý tới lớp ra. Lớp này có thể gồm một hoặc nhiều nơron nhân tạo và một mạng nơron sẽ có một hoặc nhiều lớp ẩn.
- Lớp ra cũng có thể có một hoặc nhiều nơron,

tín hiệu lớp vào sau khi được xử lý có thể đưa ra một hoặc nhiều tín hiệu ở lớp ra.

Chức năng của một mạng nơron được quyết định bởi các nhân tố chính: số lớp, quá trình học hay huấn luyện mạng và thuật toán huấn luyện mạng. Một mạng nơron thông thường thì có một lớp vào và một lớp ra, một hoặc nhiều lớp ẩn. Tuy nhiên, theo tác giả Jeff Heaton (2008) thì rất hiếm khi sử dụng nhiều hơn một lớp ẩn. Bên cạnh đó, tác giả cũng cho rằng việc lựa chọn số lớp ẩn cũng không quan trọng bằng số nơron trong lớp ẩn vì đây mới là giá trị quyết định cấu trúc của mạng nơron. Vậy nên trong nghiên cứu này sẽ sử dụng mạng nơron có một lớp ẩn.

Số lượng nơron lớp ẩn phụ thuộc vào số lượng biến vào và biến ra, thuật toán huấn luyện mạng, hàm lỗi. Nếu số nơron lớp ẩn quá ít sẽ làm mạng nơron không nhận dạng được tín hiệu đầy đủ trong tập dữ liệu dẫn đến hiện tượng thiếu ăn khớp. Ngược lại số nơron lớp ẩn nhiều quá sẽ làm tăng thời gian huấn luyện mạng hoặc không đáp ứng tốt một đầu vào mới, nghĩa là thông tin trong dữ liệu mẫu không đủ đặc trưng để huấn luyện mạng. Trong nghiên cứu này sẽ lựa chọn ANN có hiệu quả dự báo tốt nhất bằng cách thay đổi số nơron ẩn và sử dụng kỹ thuật kiểm định bằng cách chia tập huấn luyện thành ba phần là kiểm định, huấn luyện và kiểm tra.



Hình 1. Mô hình mạng nơron nhân tạo truyền thẳng ba lớp

3.1.2. Quá trình huấn luyện mạng nơron

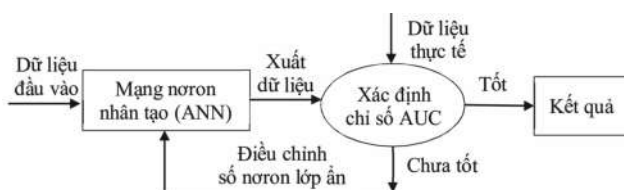
Quá trình học hay huấn luyện mạng là quá trình cơ bản, quan trọng nhất của mạng nơron và

được phân thành ba loại: học có giám sát, học không có giám sát, học tăng cường. Trong lĩnh vực dự báo khả năng vỡ ồng, các dữ liệu vào và

dữ liệu đích đều có sẵn nên phương pháp học có giám sát là phù hợp nhất. Các tín hiệu vào được gán trọng số sau đó so sánh và điều chỉnh các giá trị này phù hợp với các tập đích mong muốn. Hệ số bias (b_k) được đưa vào để điều chỉnh các giá trị trọng số sau mỗi lần lặp. Sai số của ANN được xác định bằng tổng bình phương trung bình (Mean Squared Error – MSE). Quá trình huấn luyện dừng lại khi số lần học đủ lớn hoặc khi sai số giữa tập ra và tập đích đạt giá trị đủ nhỏ. Do dữ liệu ống vỡ lớn nên phân tách dữ liệu phức tạp, cần sử dụng các thuật toán tối ưu để tìm nghiệm phù hợp. So sánh các thuật toán tối ưu thường được sử dụng cho mạng nơron, nghiên cứu lựa chọn thuật toán Levenberg-Marquardt vì thuật toán này cho kết quả hội tụ nhanh và giảm đáng kể khối lượng tính toán.

3.2 Mô hình ANN dự báo ống vỡ

Dựa trên lý thuyết về mạng nơron nhân tạo kết hợp với đặc điểm của hiện tượng ống vỡ trên mạng lưới cấp nước và tính năng của công cụ MATLAB, nghiên cứu đề xuất mô hình ANN dự báo khả năng ống vỡ trên mạng lưới cấp nước như Hình 2.



Hình 2. Mô hình ANN dự báo ống vỡ

Các bước thực hiện mô hình đề xuất và thuộc tính của nó như sau:

Bước 1. Xây dựng dữ liệu đầu vào bằng phương pháp thống kê và xử lý số liệu đưa vào ANN. Ở đây dữ liệu thống kê được phân chia ngẫu nhiên làm hai bộ dữ liệu với tỉ lệ 7:3 trong đó 70% dữ liệu huấn luyện chia thành 3 nhóm là 70% luyện mạng, 15% kiểm định và 15% kiểm tra. Bộ thứ hai có 30% dữ liệu để dự báo khả năng vỡ ống.

Bước 2. Tạo ANN trong MATLAB như Hình 3

- ANN có cấu trúc 3 lớp gồm: lớp vào (có 8

yếu tố), lớp ẩn (có 1 lớp và số nơron trong lớp ẩn (số nơron ẩn) sẽ thay đổi) và lớp ra (1 yếu tố ống vỡ| không vỡ).

- Thuật toán huấn luyện (Training) là Levenberg - Marquardt.

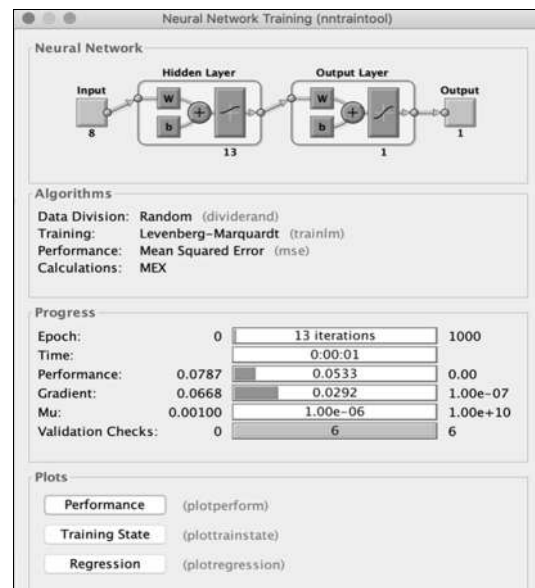
- Hàm lỗi làm hàm xác định độ lệch trung bình bình phương (Mean Squared Error-MSE).

- Hàm kích hoạt nơron ẩn dạng logsig (sigmoid).

- Khởi tạo trọng số và hệ số bias ban đầu ngẫu nhiên nhỏ.

Bước 3. Xuất kết quả dự báo và so sánh với số liệu thực tế. Đánh giá chất lượng kết quả.

Bước 4. Thay đổi số nơron ẩn để cải thiện kết quả.



Hình 3. Các thông số cài đặt trong mô hình ANN trong MATLAB

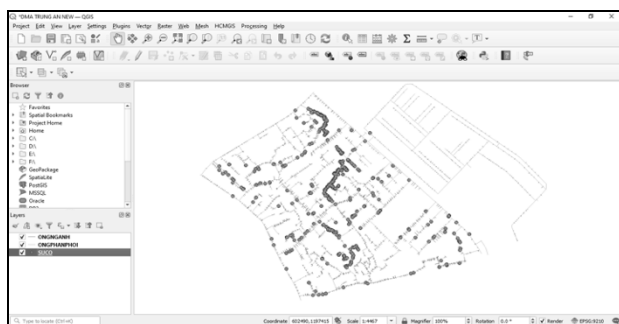
3.4 Kiểm định mô hình bằng tiêu chuẩn AUC

Các kết quả mô hình ANN dự báo sai sẽ rơi vào hai trường hợp, một là ống hoạt động hoàn toàn bình thường nhưng mô hình lại dự báo ống vỡ, hai là ống bị vỡ nhưng mô hình lại dự báo ống không vỡ. Sai lầm loại 1 có ảnh hưởng tới chi phí kiểm tra đường ống vỡ nhưng sai lầm loại 2 gây ra hậu quả lớn hơn vì không những nước bị thất thoát mà còn có nguy cơ chất ô nhiễm đi vào trong đường ống. Như vậy, khi kết quả dự báo mắc hai loại sai lầm này thì cần có một tiêu chuẩn để đánh

giá chất lượng mô hình. Theo Hanley JA nnk. (1982), ROC (Receiver Operating Characteristic) là đường cong nối các điểm có hoành độ là tỉ lệ dự báo đúng xác suất ống không vỡ (False Positive Rate) và tung độ là tỉ lệ dự báo đúng xác suất ống vỡ (True Positive Rate). Bằng phương pháp tích phân người ta xác định giá trị diện tích dưới đường ROC gọi là chỉ số AUC (Area Under the Curve), AUC cho biết giá trị xác suất lấy ngẫu nhiên một kết quả dự báo ống vỡ được xếp hạng cao hơn kết quả dự báo ống không vỡ bao nhiêu. Chỉ số AUC càng cao thì mô hình dự báo càng chính xác trong việc phân loại các lớp. Theo Hanley JA nnk. (1982) thì hiệu suất của mô hình dự báo được đánh giá theo tiêu chuẩn AUC như Bảng 1.

Bảng 1. Tiêu chuẩn AUC

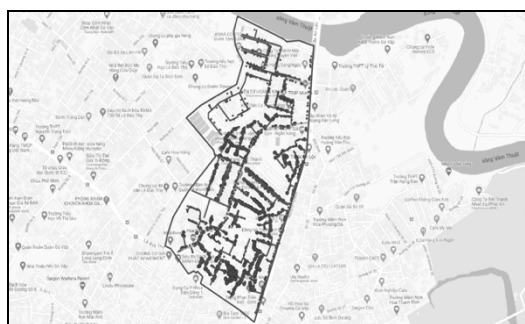
AUC	Ý nghĩa diễn giải
> 0,9	Rất tốt
0,8 - 0,9	Tốt
0,7 - 0,8	Trung bình
0,6 - 0,7	Không tốt
0,5 - 0,6	Kết quả không sử dụng được



4. KIỂM CHỨNG MÔ HÌNH

4.1. Thu thập số liệu mạng lưới cấp nước DMA17 Trung An quận Gò Vấp

Công ty cổ phần cấp nước Trung An thuộc Tổng công ty cấp nước Sài Gòn quản lý cấp nước trên địa bàn quận Gò Vấp (trừ phường 1), quận 12 và huyện Hóc Môn. Mạng lưới được chia thành các khu vực cấp nước (DMA-District Meter Area), trong đó có DMA17 (Hình 4) thuộc phường 17 quận Gò Vấp cũng là một trong những khu vực quản lý của công ty. Thu thập dữ liệu từ hồ sơ quản lý của công ty và tham vấn các nhân viên phụ trách kỹ thuật cũng như công nhân vận hành, sửa chữa hệ thống để hiểu rõ hơn về số liệu. Mạng lưới cấp nước DMA17 bao gồm các ống nhánh đường kính 25mm, ống chính đường kính từ 50mm đến 250mm và 1683 đồng hồ khách hàng. Theo số liệu thống kê từ năm 2014 đến tháng 7/2020 mạng lưới liên tục được mở rộng, tính đến thời điểm khảo sát thì tổng chiều dài ống nhánh là 3416,68m. Mạng lưới ống nhánh có ba loại vật liệu là uPVC, PE và HDPE với tỉ lệ chiều dài lần lượt chiếm 36,4%, 63% và 0,6% so với tổng chiều dài mạng lưới ống nhánh.



Hình 4. Dữ liệu GIS của mạng lưới cấp nước DMA17, Trung An, quận Gò Vấp.

4.2. Xây dựng mô hình ANN dự báo ống vỡ sử dụng phần mềm MATLAB

Bảng 2. Mô tả ảnh hưởng của các yếu tố tới sự kiện ống vỡ

TT	Loại biến	Tên biến	Giá trị	Ảnh hưởng tới biến ống vỡ (F=1)
1	Đặc trưng vật lí	Tuổi ống (A)	3 tháng đến 12 năm	Số điểm vỡ (N0) chủ yếu xuất hiện trên các ống vật liệu (Mat) số 1-uPVC và số 3-PE hoạt động trong thời gian (A) từ 5 đến 7 năm và từ 11 đến 13 năm.
2		Số điểm đầu nối (N0)	2 đến 100 điểm	
3		Vật liệu (Mat)	uPVC, HDPE, PE	

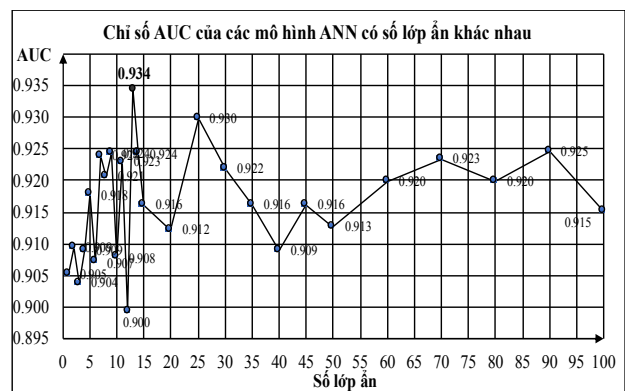
TT	Loại biến	Tên biến	Giá trị	Ảnh hưởng tới biến ống vỡ (F=1)
4		Đường kính (D)	25mm	Số liệu ống vỡ quan sát là trên mạng lưới ống nhánh nên đường kính chỉ có một loại là 25mm và đoạn ống vỡ có chiều dài ngắn (< 8m).
5		Chiều dài (L)	0,1 đến 12m	
6	Đặc trưng môi trường làm việc	Tuyến đường (R)	Mã tuyến được kí hiệu từ 2 đến 129	Ống quan sát là ống nối từ ống chính tới nhà dân có áp lực làm việc và tải trọng tác dụng bên ngoài không lớn nên nguyên nhân ống vỡ do tải trọng không thể hiện rõ trong số liệu này và các điểm vỡ nằm trên các tuyến đường khác nhau.
7	Lịch sử vỡ	Prior	0 đến 1 lần	Lịch sử vỡ ống Prior cho thấy số ống vỡ lần đầu nhiều hơn ống vỡ lần thứ hai và tập trung chủ yếu trên các tuyến ống đường kính (RD) là 100mm.
8	Đặc trưng tải trọng	Đường kính tuyến (RD)	50, 100, 150, 200 (mm)	

Dữ liệu thống kê đường ống trên mạng lưới cấp nước DMA17 được chia theo tỉ lệ 7:3 tương đương với 1818:779. Bảng 2 mô tả dữ liệu mẫu cho mô hình ANN với 4 tập dữ liệu dạng ma trận kích thước như sau: tập huấn luyện [8 x 1818]; tập đích [1 x 1818]; tập kiểm tra [8 x 779]; tập so sánh [1 x 779] dùng để kiểm chứng kết quả dự báo.

4.3. Kết quả mô hình ANN

Khi thay đổi số lớp ẩn thì kết quả dự báo cũng thay đổi theo, bằng phương pháp thử sai tác giả bắt đầu từ ANN có 1 nơron ẩn và tăng dần đến 100 nơron, sau mỗi lần mô phỏng cho một kết quả mới. So sánh kết quả mô hình sau mỗi lần thay đổi số nơron ẩn với dữ liệu thực tế và xác định giá trị AUC theo lý thuyết đã trình bày ta có biểu đồ như Hình 5. Chỉ số AUC đạt giá trị nhỏ nhất là 0,900 và lớn nhất là 0,934 khi mạng có 13 nơron ẩn, so với tiêu chuẩn thì mô hình dự báo đạt mức tốt và rất tốt.

Biểu đồ Hình 5 cũng cho thấy thay đổi số lớp ẩn có ảnh hưởng tới kết quả mô hình tuy nhiên biên độ dao động của AUC trong khoảng từ 5 đến 30 nơron ẩn là nhiều nhất. Khi tiếp tục tăng số lớp ẩn thì biên độ này giảm dần điều này có nghĩa là chất lượng kết quả không thể cải thiện thêm nữa. Mô hình ANN 13 nơron ẩn dự báo ống vỡ thực hiện tất cả 13 lần lặp và giá trị lỗi (MSE) qua các lần được trình bày trong Hình 6.



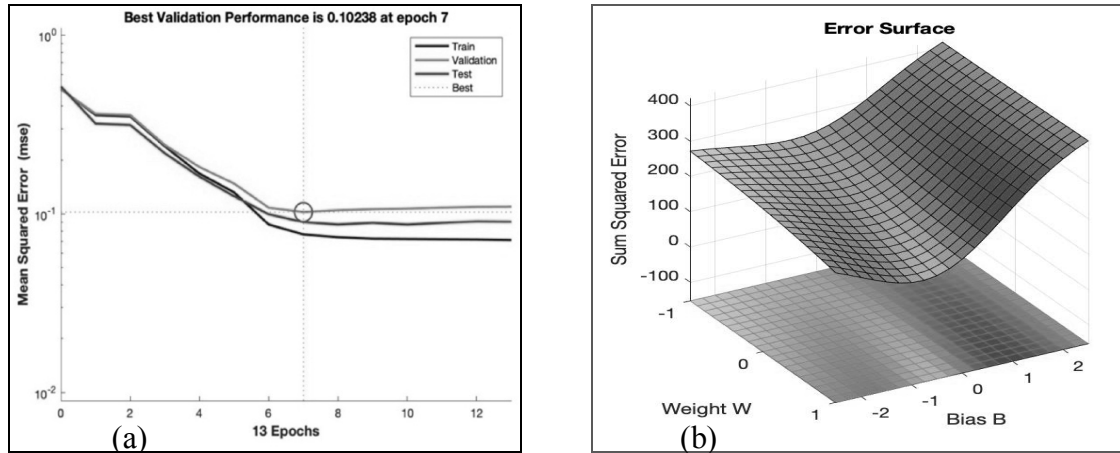
Hình 5. Biểu đồ chỉ số AUC của các mô hình ANN có số nơron lớp ẩn khác nhau

Tập kiểm định (Validation) là đường trên cùng trong Hình 6 (a) được mô hình đánh giá là đạt hiệu suất tối ưu ở bước lặp (epoch) thứ 7 khi đó giá trị $MSE = 0,10238$ (Best Validation Performance). Mô tả tương quan giữa tổng độ lệch (Sum Squared Error), trọng số Weight W và số Bias B ta có Hình 6 (b) – Bề mặt lỗi (Error Surface), biểu đồ cũng cho thấy khi số lần lặp tăng (lúc này chỉ số Bias giảm) thì các thông số của mô hình ANN có xu hướng bão hoà và không tìm thêm được điểm cực tiểu.

Xuất 779 kết quả mô hình và so sánh với dữ liệu thực tế có số ống vỡ là 292 và không vỡ là 487 ống, kết quả dự báo 205 ống vỡ và 574 ống không vỡ. Như vậy mô hình đã dự báo sai 109

trường hợp trong đó dự báo 98 ống không vỡ nhưng thực tế bị vỡ (sai $98/292=34\%$) và dự báo 11 ống vỡ nhưng thực tế là không vỡ (sai $11/487=2\%$). Sử dụng tiêu chuẩn AUC đã trình bày ở trên để kiểm tra chất lượng mô hình ANN

13 nơron ẩn. Kết quả tính toán và biểu diễn đường cong ROC (Hình 7) đã cho thấy $AUC = 0,934$, nghĩa là mô hình đạt mức rất tốt và so với tất cả các trường hợp mô phỏng thì chỉ số AUC của mô hình này tốt nhất.



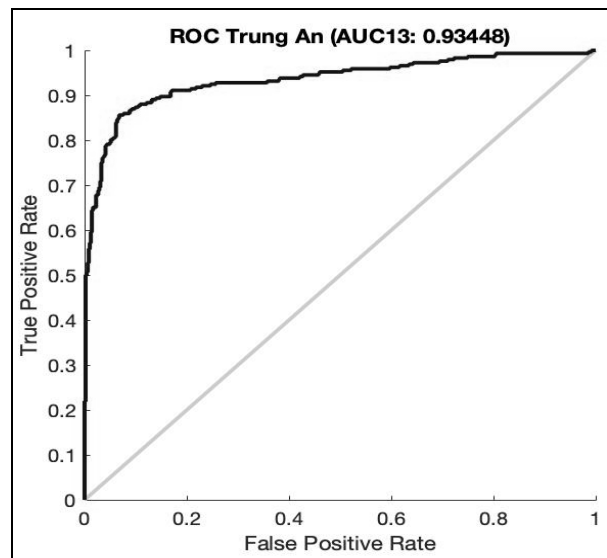
Hình 6. Giá trị MSE và bề mặt lỗi của mô hình ANN 13 nơron ẩn

5. KẾT LUẬN VÀ KIẾN NGHỊ

Ứng dụng ANN trong dự báo khả năng ống vỡ đã được thực hiện trong các nghiên cứu trước đây và tiếp tục được cải thiện trong nghiên cứu này để đánh giá khả năng vỡ ống trên mạng lưới cấp nước tại Việt Nam. Nghiên cứu đã đạt được kết quả sau:

- Trong dự báo ống vỡ thì ANN ở dạng truyền thẳng và học có giám sát, cấu trúc mạng chỉ cần một lớp ẩn và số nơron trong lớp ẩn sẽ thay đổi cho đến khi kết quả dự báo đạt chỉ số AUC tốt nhất. Kết quả ứng dụng thực tế cho thấy nếu số nơron ẩn quá nhỏ thì kết quả dự báo có chỉ số AUC không cao, nếu số lượng nơron ẩn lớn thì kết quả không được cải thiện thêm nữa, vậy nên trong nghiên cứu này tác giả đề xuất số lượng nơron ẩn dao động từ 5-30 nơron.

- Hiệu quả đạt được khi áp dụng trên mạng lưới cấp nước DMA17 Trung An, quận Gò Vấp, thành phố Hồ Chí Minh ở mức rất tốt ($AUC = 0,934$). Qua đó cho thấy mô hình ANN trong dự báo ống vỡ là hoàn toàn phù hợp và có độ chính xác cao.



Hình 7. Đường cong ROC và AUC của mô hình ANN 13 nơron ẩn

Nghiên cứu có thể được cải thiện trong tương lai về việc đánh giá đặc điểm hình học của điểm vỡ cũng như nguyên nhân ống vỡ (thông thường những ống vỡ do yếu tố bên ngoài tác động thì điểm vỡ không bị ảnh hưởng bởi 8 yếu tố đã xét trong bài toán). Bên cạnh đó, cũng có thể đánh giá cụ thể hơn về những điểm vỡ tại vị trí phụ tùng đầu nối.

TÀI LIỆU THAM KHẢO

- Phạm Thị Minh Lành, Vũ Thị Vân Anh, Phạm Hà Hải (2018) “*Đề xuất mô hình dự báo ống vỡ trên mạng lưới cấp nước*”, Tạp chí Khoa Học Kỹ Thuật Thủy lợi và Môi trường, số 60.
- U. Shamir and Charles D.D Howard (1979) “*An analytic approach to scheduling pipe replacement*” Journal American water Work. Assoc., vol. 71, no. 5, tr. 248.
- G. Kabir, S. Tesfamariam, and R. Sadiq (2016) “*Bayesian Model Averaging for the Prediction of Water Main Failure for Small to Large Canadian Municipalities*”, Canada Journal Civil Eng., vol. 43, no. 3, tr. 233–240.
- T. Wengström (1993) “*Comparative analysis of Pipe Break Rates*”, Chalmers Univ. Technol.
- H. Al-barqawi and T. Zayed (2008) “*Infrastructure Management: Integrated AHP / ANN Model to Evaluate Municipal Water Mains Performance*”, Journal of Infrastructure Systems, tr. 305–318.
- D.Achim; F. Ghotb; and K. J. McManus (2007), “*Prediction of water pipe asset life using neural networks*,” J. Infrastruct. Syst., vol. 13, no. 1, tr. 26–30.
- Barbara G.Tabachnick and Linda S. Fidell (2007), *Using Multivariate Statistics*. United States of America: Pearson Education. Inc.
- Jeff Heaton (2008), *Introduction to Neural Network with Java, Second edition*, United States of America: Heaton Research Inc, tr. 159.
- Demuth, Howard B. and Beale, Mark H. and De Jess, Orlando and Hagan, Martin T. (2014), *Neural Network Design, Second edition*, Martin Hagan, Oklahoma State University, Stillwater, OK, United States.
- Hanley JA, McNeil BJ., (1982) *The meaning and use of the area under a receiver operating characteristic (ROC) curve*, Radiology, vol. 143(1):29-36.
- I. H. Witten, E. Frank, and M. a Hall, (2011), *Data Mining: Practical Machine Learning Tools and Techniques*, Elsevier Inc.

Abstract:

APPLICATION OF ARTIFICIAL NEURAL NETWORKS TO PREDICT PIPE FAILURE IN WATER SUPPLY NETWORKS

Water leakage on the water supply network is still a problem to be solved by water supply companies. One of the main causes of this phenomenon is due to pipe failure. Prediction models of water pipe failure are also constantly being improved to accurately and quickly predict the position of pipe leakage. With the aim of developing a better prediction, this study proposed application of the neural network model to predict pipe failure on water supply network. The proposed model is verified by the water supply network in ward 17 (DMA17), Go Vap district, Ho Minh city. The quality of the proposed model was evaluated by using the AUC standard. The results shows that the proposed model is very good.

Keywords: Artificial neural networks, water leakage, pipe failure, predictive models, water supply networks.

Ngày nhận bài: 21/12/2020

Ngày chấp nhận đăng: 31/12/2020