

ỨNG DỤNG ANN TRONG DỰ BÁO ÁP SUẤT NÚT VĨA

Nguyễn Văn Hùng, Đặng Hữu Minh

Đại học Dầu khí Việt Nam

Email: hungnv@pvu.edu.vn

Tóm tắt

Dự báo áp suất nứt vỉa là công việc quan trọng khi lên kế hoạch thiết kế giếng khoan, cho phép kiểm soát, vận hành, kích thích giếng hoạt động hiệu quả. Nếu dự báo không chính xác áp suất nứt vỉa có thể gây ra các vấn đề nghiêm trọng như: mất tuần hoàn dung dịch khoan hoặc xảy ra hiện tượng "kick" dẫn đến phun trào...

Bài báo giới thiệu các ứng dụng của trí tuệ nhân tạo trong lĩnh vực dầu khí, sử dụng phương pháp sử dụng ANN dựa trên các dữ liệu đầu vào gồm: độ sâu, hệ số Poisson, ứng suất địa tĩnh, áp suất lỗ rỗng và dữ liệu đầu ra là áp suất nứt vỉa để xây dựng mô hình dự báo áp suất nứt vỉa cho giếng khoan thuộc bể Nam Côn Sơn. Nhóm tác giả đã so sánh với kết quả dự báo bằng phương pháp truyền thống cho thấy phương pháp sử dụng ANN cho kết quả dự báo áp suất nứt vỉa sát với kết quả đo thực tế nhất.

Từ khóa: Áp suất nứt vỉa, ANN, bể Nam Côn Sơn.

1. Giới thiệu

Áp suất nứt vỉa là thông số quan trọng trong giai đoạn tiền thiết kế hệ thống khoan, khai thác và kích thích vỉa, cũng như để tối ưu công tác khoan. Việc xác định áp suất nứt vỉa bằng phương pháp đo trực tiếp chi phí cao và gặp nhiều khó khăn, vì vậy có thể xác định áp suất nứt vỉa bằng phương pháp gián tiếp. Các thông số như áp suất lỗ rỗng, độ rỗng thành hệ, tỷ trọng đá, tính chất thạch học của đất đá... theo độ sâu được sử dụng làm dữ liệu đầu vào để dự báo áp suất nứt vỉa. Đồng thời, công thức dự báo được sử dụng riêng cho đặc tính vùng dựa trên cơ sở dữ liệu có sẵn được sử dụng.

Năm 1957, Hubbert và Willis lần đầu đưa ra mối tương quan để tính áp suất nứt vỉa dựa vào áp suất lỗ rỗng, hệ số Poisson và ứng suất lớp phủ theo độ sâu [1]. Đến năm 1968, Pennebaker phát triển mối tương quan giữa ứng suất lớp phủ và tuổi thành hệ để xây dựng công thức dự báo [2]. Eaton (1969) tìm ra công thức tính áp suất lỗ rỗng dựa vào ứng suất lỗ rỗng, hệ số Poisson và độ sâu, áp dụng cho vùng Tây Texas và vịnh Mexico [3]. Từ đó, các mối tương quan được xác định dựa trên sự hiệu chỉnh, cải tiến để giúp các kết quả dự báo áp suất nứt vỉa chính xác hơn. Bảng 1 thống kê các phương pháp truyền thống để xác định áp suất nứt vỉa đã được công bố. Mặc dù được

chấp nhận để áp dụng trong việc tính toán, song vẫn còn tồn tại một số nhược điểm dẫn đến kết quả sai số lớn.

Các tài liệu nghiên cứu gần đây cho thấy các phương pháp như sử dụng mạng neuron nhân tạo (ANN), logic mạng mờ và thuật di truyền, được sử dụng rộng rãi trong lĩnh vực dầu khí, kỹ thuật vector hỗ trợ (SVM), mạng chức năng (functional network) và lập luận theo tình huống (case based reasoning)... được sử dụng rộng rãi trong lĩnh vực dầu khí (Hình 1). Nhóm tác giả tổng hợp được 17 ứng dụng của kỹ thuật trí tuệ nhân tạo, bao gồm: (1) phát triển giao diện cho quá trình mô phỏng; (2) minh giải đường log; (3) lựa chọn choòng khoan; (4) chuẩn đoán trong quá trình vận hành bơm; (5) lựa chọn và minh giải mô hình nứt vỉa; (6) phân tích nứt vỉa thủy lực; (7) tối ưu hóa gas-lift; (8) phân loại vỉa dầu khí; (9) các mối tương quan về các tính chất của chất lưu (PVT); (10) phân tích độ rủi ro của dự án thu hồi dầu tăng cường; (11) dự báo chế độ dòng chảy trong đường ống; (12) phân tích sự phá hủy thành hệ; (13) thiết kế và tối ưu hóa nứt vỉa thủy lực; (14) tối ưu hóa khai thác; (15) quản lý thu hồi dầu tăng cường; (16) tối ưu hóa vận hành khoan; (17) tối ưu trong thiết kế giếng sử dụng thuật di truyền [9]. Bảng 2 tổng hợp các tài liệu liên quan tới các ứng dụng của trí tuệ nhân tạo.

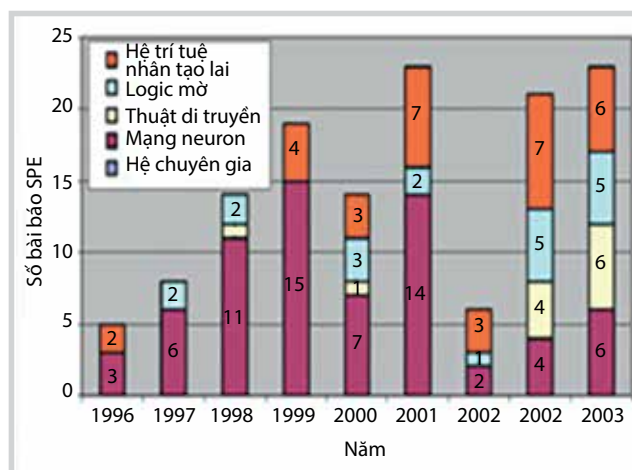
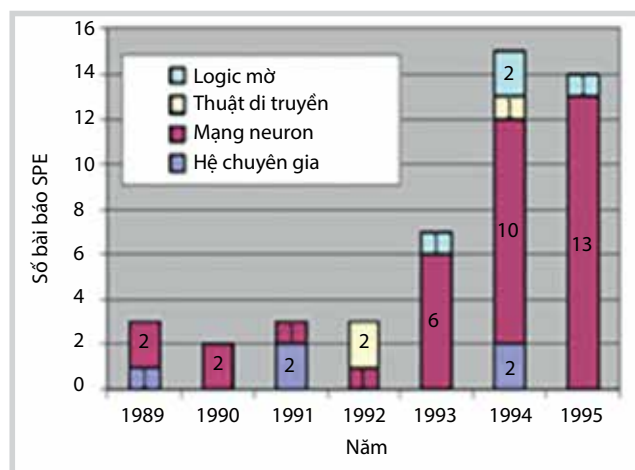
Ngoài các thông số trên, có thể sử dụng trực tiếp các thông số vật lý vỉa như gamma ray, điện trở và mật độ được đo từ kết quả địa vật lý giếng khoan để dự báo áp suất nứt vỉa nhưng có kết hợp hiệu chỉnh mô hình thông qua kết quả đo trực tiếp từ thí nghiệm LOT hay FIT [24].

Ngày nhận bài: 2/5/2018. Ngày phản biện đánh giá và sửa chữa: 2/5 - 14/6/2018.

Ngày bài báo được duyệt đăng: 6/3/2019.

Bảng 1. Các phương pháp truyền thống để xác định áp suất nứt vỉa

TT	Thông số đầu vào	Công thức tính	Tài liệu tham khảo
1	Ứng suất địa tĩnh, σ_{ob} Áp suất lỗ rỗng, p_f	$p_{ff} = \frac{\sigma_{ob} + 2p_f}{3}$	[1]
2	Áp suất lỗ rỗng, p_f Hệ số Poisson, μ Ứng suất địa tĩnh, σ_{ob}	$p_{ff} = \frac{\mu}{1 - \mu} (\sigma_{ob} - p_f) + p_f$	[3]
3	Độ sâu, D Áp suất lỗ rỗng, p_f Hệ số F_σ	$p_{ff} = F_\sigma (D - p_f) + p_f$	[4]
4	Mật độ biểu kiến Thời gian truyền sóng địa chấn Ứng suất địa tĩnh, σ_{ob}	Tra đồ thị $\frac{K_b}{\sigma_{ob}}$	[5]
5	Ứng suất địa tĩnh, σ_{ob} Áp suất lỗ rỗng, p_f Hệ số F_σ	$p_{ff} = F_\sigma (\sigma_{ob} - p_f) + p_f$	[6]
6	Áp suất địa tĩnh, σ_{ob} Áp suất lỗ rỗng, p_f Hệ số Poisson, μ Hệ số nén của độ rỗng trong khung đá, α	$p_{ff} = \alpha p_f + \frac{2\mu}{1 - \mu} (\sigma_{ob} - \alpha p_f)$	[7]
7	Gradient ứng suất địa tĩnh, G_{ob} Gradient áp suất lỗ rỗng, G_{fr}	$G_{ff} = G_{fr} + \frac{2\mu}{1 - \mu} (G_{ob} - G_p)$	[8]



Hình 1. Ứng dụng trí tuệ nhân tạo trong lĩnh vực dầu khí [9]

Bảng 2. Ứng dụng của kỹ thuật trí tuệ nhân tạo trong lĩnh vực dầu khí

TT	Phương pháp	Thông số được dự báo	Tài liệu tham khảo
1	Fuzzy logic	Xác định sự rò rỉ khí	[10]
2	Fuzzy logic, GA	Tính chất chất lưu	[11]
3	ANN	Độ rỗng, độ thấm, độ bão hòa	[12]
4	Hybrid System	Độ rỗng, độ thấm của vỉa	[13]
5	ANN	Kẹt cần khoan	[14]
6	SVM	Tính chất của dầu và khí	[15]
7	SVM	Độ thấm vỉa	[16]
8	CSR	Sự cố trong quá trình khoan	[17]
9	ANN	Lựa chọn cấu trúc giàn biển	[18]
10	ANN	Độ ổn định thành giếng khoan	[19]
11	Neuro-fuzzy	Các thông số về độ bền của đất đá	[20]
12	Hybird-SVM	Đặc tính gaslift và tối ưu khai thác	[21]

Nhiều kỹ thuật trí tuệ nhân tạo như mạng logic mờ (fuzzy logic); thuật di truyền (genetic algorithm); huấn luyện có giám sát máy (support vector machines); lập luận theo tình huống (CBR); hệ trí tuệ nhân tạo lai (hybrid systems) được ứng dụng rộng rãi trong công nghiệp dầu khí [25, 26]. Trong bài báo này, nhóm tác giả chỉ tập trung vào bài toán ứng dụng của mạng trí tuệ nhân tạo trong dự báo áp suất nút vỉa.

2. Cơ sở lý thuyết về ANN

2.1. Mô hình ANN

Với mục đích tạo ra mô hình tính toán phỏng theo cách làm việc của neuron trong bộ não con người, vào năm 1943, McCulloch và Pitts [27] đã đề xuất mô hình toán cho một neuron (Hình 2).

Trong mô hình này, một neuron thứ i sẽ nhận các tín hiệu vào x_j với các trọng số tương ứng là w_{ij} , tổng các thông tin vào có trọng số là $\sum_{j=1}^m w_{ij}x_j$. Thông tin đầu ra ở thời điểm $t + 1$ được tính từ các thông tin đầu vào như sau:

$$out(t + 1) = g(\sum w_{ij} x_j(t) - \theta_i) \quad (1)$$

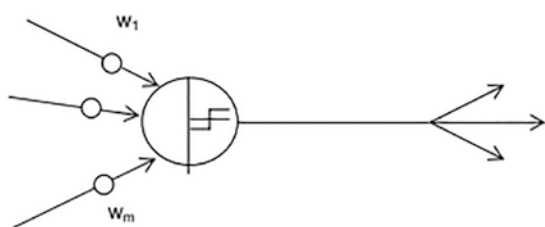
Trong đó, g là hàm kích hoạt (còn gọi là hàm chuyển) có dạng là hàm bước nhảy, đóng vai trò biến đổi từ thông tin đầu vào thành tín hiệu đầu ra:

$$g(f) = \begin{cases} 1 & \text{nếu } f > 0 \\ 0 & \text{nếu } f \leq 0 \end{cases} \quad (2)$$

Như vậy, $out = 1$ (ứng với việc neuron tạo tín hiệu đầu ra) khi tổng các tín hiệu vào lớn hơn ngưỡng θ_i , còn $out = 0$ (neuron không tạo tín hiệu ở đầu ra) khi tổng các tín hiệu vào nhỏ hơn ngưỡng θ_i .

Trong mô hình neuron của McCulloch và Pitts, các trọng số w_{ij} thể hiện ảnh hưởng của khớp nối trong liên kết giữa neuron j (neuron gửi tín hiệu) và neuron i (neuron nhận tín hiệu). Trọng số w_{ij} dương tương ứng với khớp nối kích thích, trọng số âm ứng với khớp nối ức chế và w_{ij} bằng 0 khi không có liên kết giữa 2 neuron. Hàm chuyển g ngoài dạng hàm bước nhảy còn có thể chọn nhiều dạng khác nhau.

Thông qua cách mô hình hóa đơn giản một neuron sinh học như trên, McCulloch và Pitts đã đưa ra một mô



Hình 2. Mô hình ANN [27]

hình ANN có tiềm năng tính toán quan trọng. Nó có thể thực hiện các phép toán logic cơ bản như AND, OR và NOT khi các trọng số và ngưỡng được chọn phù hợp. Sự liên kết giữa các ANN với cách thức khác nhau sẽ tạo nên các loại ANN với tính chất và khả năng làm việc khác nhau.

2.2. Cấu trúc ANN

Nhìn chung, các nghiên cứu đều cho thấy cấu trúc của một ANN gồm 3 phần là lớp dữ liệu đầu vào (input layer), lớp ẩn (hidden layer) và lớp dữ liệu đầu ra (output layer). Mỗi lớp này đều được cấu tạo từ một số lượng neuron. Mỗi neuron là một đơn vị xử lý thông tin của ANN, là yếu tố cơ bản để cấu tạo nên toàn mạng. Các neuron này nhận các dữ liệu vào, xử lý chúng và cho ra một kết quả xử lý để làm dữ liệu đầu vào cho một neuron khác. Thông tin sau khi đi vào lớp đầu vào sẽ đi tới lớp ẩn, lớp này bao gồm các neuron sẽ xử lý thông tin và chuyển sang các lớp tiếp theo, cuối cùng là tới lớp kết quả đầu ra. Mỗi dữ liệu đầu vào sẽ được đánh giá mức độ mạnh yếu hay độ quan trọng thông qua một trọng số w ứng với mỗi neuron. Các trọng số này có thể được thay đổi để đưa ra được giá trị đầu ra phù hợp nhất.

Cấu trúc ANN được xác định bởi 5 yếu tố: (1) số lượng các dữ liệu đầu vào/đầu ra; (2) số lượng các lớp của mạng; (3) số lượng neuron trong mỗi lớp (lớp đầu vào, lớp ẩn, lớp đầu ra) của mạng; (4) các trọng số để liên kết các neuron; (5) cách liên kết giữa các neuron (trong một lớp và giữa các lớp) với nhau.

Đối với ANN, việc quyết định số lượng neuron của lớp ẩn rất phức tạp, phụ thuộc vào nhiều yếu tố: (1) số lượng neuron của lớp dữ liệu đầu vào và đầu ra; (2) lượng dữ liệu của quá trình luyện; (3) hàm sử dụng trong quá trình luyện; (4) thuật toán huấn luyện của mạng neuron.

Để sai số của quá trình xử lý của ANN là thấp nhất, cũng như để quá trình này hoạt động hiệu quả, số lượng neuron của lớp ẩn, cũng như số lớp ẩn cần được lựa chọn phù hợp và tối ưu nhất, theo nguyên tắc: nếu quá ít neuron sẽ dẫn tới sai số cao cho cả hệ thống và yếu tố đầu ra cần dự báo sẽ trở nên rất khó khăn vì số lượng neuron quá ít để có thể thực hiện được. Ngược lại khi quá nhiều neuron thì mạng neuron sẽ hoạt động mất rất nhiều thời gian và trở nên không hiệu quả. Vì vậy, việc quyết định được số lượng lớp ẩn cũng như số lượng neuron trong lớp này vô cùng quan trọng. Thường không có quy tắc nào trong việc lựa chọn các giá trị này, việc xác định được thực hiện bằng phương pháp thử-sai, đến khi nào

giá trị sai số giữa giá trị đầu ra của mạng và giá trị thực nhỏ nhất có thể.

2.3. Thuật toán truyền ngược

Trong số nhiều loại ANN, mạng neuron nhiều lớp ẩn với thuật toán truyền ngược là phương pháp được sử dụng phổ biến nhất hiện nay [28]. Một ANN đặc trưng chứa một lớp đầu vào, một lớp đầu ra và một hoặc nhiều lớp ẩn.

Thuật toán truyền ngược là sự mở rộng của phương pháp bình phương tối thiểu, được phát triển dựa trên cơ sở đào tạo dữ liệu của ANN đa lớp, với mục tiêu làm giảm sai số nhỏ nhất có thể giữa giá trị dự báo và giá trị yêu cầu. Thuật toán truyền ngược sử dụng chuỗi các quy tắc để tính đạo hàm bình phương sai số với biến số là các trọng số và ngưỡng trong các lớp ẩn. Những đạo hàm này được gọi là vector độ nhạy. Một vài thuật toán phi tuyến tính như thuật toán giảm bước (steepest descent algorithm), phương pháp Newton (Newton’s method) và thuật toán gradient liên hợp (conjugate gradient algorithm) được sử dụng để giảm thiểu sai số trong quá trình đào tạo.

Như đã đề cập trước đó, thuật truyền ngược trình bày tổng quát của phương pháp bình phương tối thiểu để giảm hàm mục tiêu, nghĩa là làm hạn chế sai số giữa giá trị dự báo $o(m)$ và giá trị thực y . Hàm mục tiêu được thể hiện như sau:

$$J(m) = E[e^t e] = E[(y - o(m))^T (y - o(m))] \quad (3)$$

Trong đó, m đặc trưng cho trọng số và ngưỡng trong tất cả các neuron.

Với thuật toán giảm bước, giá trị nhỏ nhất của hàm mục tiêu được xác định bằng công thức:

$$w_{ij}^m(k+1) = w_{ij}^m(k) - \zeta \frac{\partial J}{\partial w_{ij}^m} \quad (4)$$

$$\beta_{ij}^m(k+1) = \beta_{ij}^m(k) - \zeta \frac{\partial J}{\partial \beta_{ij}^m} \quad (5)$$

Trong đó: w_{ij}^m là trọng số đặc trưng liên kết neuron thứ i với neuron thứ j ; k là số bước lặp; m là lớp trong mạng neuron; ζ là tốc độ huấn luyện.

Sử dụng chuỗi quy tắc, đạo hàm của hàm J theo trọng số và ngưỡng được tính toán theo:

$$\frac{\partial J}{\partial w_{ij}^m} = \frac{\partial J}{\partial e_j^m} \frac{\partial e_j^m}{\partial w_{ij}^m} \quad (6)$$

$$\frac{\partial J}{\partial \beta_j^m} = \frac{\partial J}{\partial \varepsilon_j^m} \frac{\partial \varepsilon_j^m}{\partial \beta_j^m} \quad (7)$$

Vi đầu vào của lớp thứ m là một hàm tường minh của trọng số và ngưỡng, nên

$$\varepsilon_j^m = \sum_i w_{ij}^m o_i^{m-1} + \beta_j^m \quad (8)$$

Trong đó, β_j^m là ngưỡng đặc trưng neuron thứ j .

Thành phần thứ 2 phương trình (6) và (7) được tính theo công thức:

$$\frac{\partial e_j^m}{\partial w_{ij}^m} = o_i^{m-1} \quad (9)$$

$$\frac{\partial \varepsilon_j^m}{\partial \beta_j^m} = 1 \quad (10)$$

Độ nhạy của hàm J theo sự thay đổi thành phần thứ j của dữ liệu đầu vào tại lớp m được đặc trưng bởi thông số s_j^m , do đó phương trình (6) và (7) được viết lại như sau:

$$\frac{\partial J}{\partial w_{ij}^m} = s_j^m \frac{\partial e_j^m}{\partial w_{ij}^m} \quad (11)$$

$$\frac{\partial J}{\partial \beta_j^m} = s_j^m \frac{\partial \varepsilon_j^m}{\partial \beta_j^m} \quad (12)$$

Thay ngược lại vào phương trình (4) và (5) ta có:

$$w_{ij}^m(k+1) = w_{ij}^m(k) - \zeta s_j^m o_i^{m-1} \quad (13)$$

$$\beta_j^m(k+1) = \beta_j^m(k) - \zeta s_j^m \quad (14)$$

Dạng ma trận,

$$W^m(k+1) = W^m(k) - \zeta s^m o^{m-1} \quad (15)$$

$$\beta^m(k+1) = \beta^m(k) - \zeta s^m \quad (16)$$

Trong đó

$$s^m = \frac{\partial J}{\partial \varepsilon^m} = \left[\frac{\partial J}{\partial \varepsilon_1^m}, \frac{\partial J}{\partial \varepsilon_2^m}, \dots, \frac{\partial J}{\partial \varepsilon_k^m} \right]^T \quad (17)$$

Vector s^m là vector độ nhạy tại lớp m và được tính từ độ nhạy lớp $m+1$, nó được tính từ công thức của lớp sau, nên gọi là truyền ngược. Ma trận Jacobian biểu diễn phép tính ngược của độ nhạy, được định nghĩa như sau:

$$\frac{\partial \varepsilon^{m+1}}{\partial \varepsilon^m} = \begin{bmatrix} \frac{\partial \varepsilon_1^{m+1}}{\partial \varepsilon_1^m} & \dots & \frac{\partial \varepsilon_1^{m+1}}{\partial \varepsilon_k^m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varepsilon_k^{m+1}}{\partial \varepsilon_1^m} & \dots & \frac{\partial \varepsilon_k^{m+1}}{\partial \varepsilon_k^m} \end{bmatrix} \quad (18)$$

Thành phần cột i , hàng j của ma trận Jacobian được thể hiện trong công thức sau:

$$\begin{aligned} \frac{\partial \varepsilon_i^{m+1}}{\partial \varepsilon_j^m} &= \frac{\partial (\sum_i w_{li}^{m+1} o_i^m + \beta_i^{m+1})}{\partial \varepsilon_j^m} = w_{ji}^{m+1} \frac{\partial o_j^m}{\partial \varepsilon_j^m} \\ &= w_{ji}^{m+1} \frac{\partial f^m(\varepsilon_j^m)}{\partial \varepsilon_j^m} = w_{ji}^{m+1} f^m(\varepsilon_j^m) \end{aligned} \quad (19)$$

Trong đó $f^m(\varepsilon_j^m) = \frac{\partial f^m(\varepsilon_j^m)}{\partial \varepsilon_j^m}$, vì vậy ma trận Jacobian trở thành

$$\frac{\partial \varepsilon^{m+1}}{\partial \varepsilon^m} = W^m f^m(\varepsilon^m) \quad (20)$$

$$f^m(\varepsilon^m) = \begin{bmatrix} f^m(\varepsilon_1^m) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & f^m(\varepsilon_1^m) \end{bmatrix} \quad (21)$$

Sau một chuỗi các phép toán, các vector độ nhạy tại các lớp khác nhau thể hiện qua công thức:

$$\begin{aligned} s^m &= \frac{\partial J}{\partial \varepsilon^m} = \left(\frac{\partial \varepsilon^{m+1}}{\partial \varepsilon^m} \right)^T \frac{\partial J}{\partial \varepsilon^{m+1}} = f^m(\varepsilon^m) (W^{m+1})^T \frac{\partial J}{\partial \varepsilon^{m+1}} \\ &= f^m(\varepsilon^m) (W^{m+1})^T s^{m+1} \end{aligned} \quad (22)$$

Do đó, độ nhạy của sai số có thể được truyền ngược từ lớp cuối cùng đến lớp đầu tiên. Tại lớp thứ l, độ nhạy bằng:

$$\begin{aligned} s_j^l &= \frac{\partial J}{\partial \varepsilon_j^l} = \frac{\partial [(y - o)^T (y - o)]}{\partial \varepsilon_j^l} = \frac{\partial (\sum_{i=1} (y_i - o_i)^2)}{\partial \varepsilon_j^l} \\ &= -2(y_i - o_i) \frac{\partial o_i}{\partial \varepsilon_j^l} = -2(y_i - o_i) \frac{\partial f^l(\varepsilon_j^l)}{\partial \varepsilon_j^l} = -2(y_i - o_i) \\ &= -2(y_i - o_i) f^l(\varepsilon_j^l) \end{aligned} \quad (23)$$

Theo dạng ma trận, phương trình (23) được viết lại thành:

$$s^l = -2 f^l(\varepsilon^l) (y - o) \quad (24)$$

Tóm lại, thuật toán truyền ngược cho mạng neuron đa lớp được tóm tắt gọn như sau: (i) dữ liệu đầu vào được truyền thẳng trong mạng từ lớp đầu tiên đến lớp cuối, (ii) tính toán vector sai số và vector độ nhạy, (iii) vector độ nhạy truyền ngược trong mạng từ lớp cuối đến lớp đầu tiên và (iv) các trọng số và ngưỡng được cập nhật trong suốt quá trình lan truyền bằng quy tắc giảm bước.

3. Kết quả nghiên cứu

3.1. Thông số đầu vào

Kết quả phân tích các mô hình lý thuyết dự báo áp suất nứt vỉa thành hệ cho thấy áp suất nứt vỉa phụ thuộc vào ứng suất địa tĩnh (δ_{ob}), áp suất lỗ rỗng (P_r), hệ số Poisson (μ), tuổi thành hệ, độ sâu và tỷ trọng đất đá. Tùy vào dữ liệu của các giếng khác nhau để chọn thông số đầu vào phù hợp. Cần lưu ý rằng một số thông số đầu vào lại có tương quan với nhau như: hệ số Poisson phụ thuộc vào tỷ trọng riêng, tỷ số thành phần ứng suất hữu hiệu và tuổi thành hệ. Do đó, mối liên hệ giữa khối lượng riêng của thành hệ với gradient áp suất lỗ rỗng và chiều sâu đã gián tiếp bao gồm trong đó và thể hiện được sự ảnh hưởng tới áp suất nứt vỉa. Vì vậy, việc lựa chọn thông số đặc trưng để đưa vào mô hình cần được cân nhắc kỹ. Sau khi tính toán và kiểm nghiệm, nhóm tác giả đề xuất các thông số đầu vào cho mô hình như Bảng 3.

Tổng số 368 điểm dữ liệu được chọn cho một giếng đã được khoan tại bể Nam Côn Sơn nhằm so sánh kết quả tính toán với kết quả đo áp suất nứt vỉa thực tế tại một số vị trí. Các dữ liệu đầu vào gồm: ứng suất địa tĩnh (δ_{ob}), áp suất lỗ rỗng (P_r), hệ số Poisson (μ) và độ sâu. Trong đó, 2/3 dữ liệu được chọn từ tập hợp dữ liệu tổng để thành lập tập dữ liệu đào tạo (training). Trong số 1/3 dữ liệu còn lại, một nửa được chọn để thiết lập tập dữ liệu phê chuẩn (validation) và một nửa là của tập dữ liệu kiểm tra (testing). Những điểm dữ liệu cho mỗi tập hợp được chọn từ tập hợp tổng được phân bố sao cho vẫn giữ được sự đồng nhất giống với tập dữ liệu tổng [29].

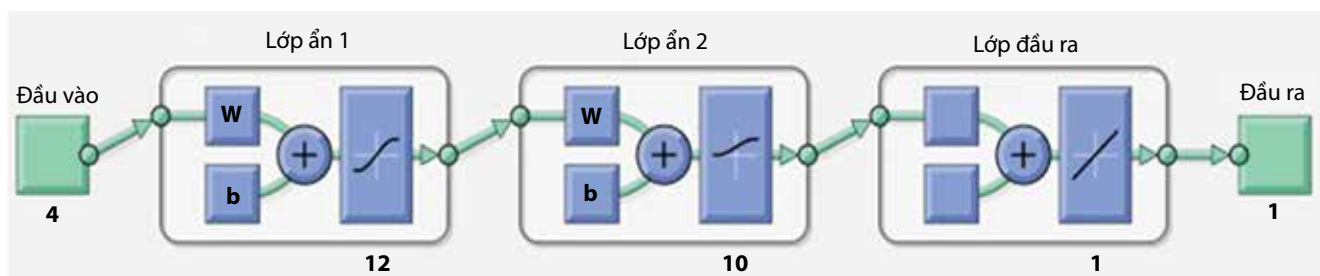
Trong mô hình ANN, nhóm tác giả sử dụng các thông số độ sâu, hệ số Poisson, ứng suất lớp phủ và áp suất lỗ rỗng làm dữ liệu đầu vào để dự báo áp suất nứt vỉa (Bảng 3).

3.2. Xây dựng mô hình

Sau khi dữ liệu được tập hợp, bước tiếp theo trong đào tạo một mạng là tạo ra một đối tượng mạng. Qua quá trình nghiên cứu và thử nghiệm, nhóm tác giả sử dụng mạng lan truyền ngược (feed-forward backpropagation) với 2 lớp ẩn, 1 lớp đầu vào và 1 lớp đầu ra. Tiếp theo là

Bảng 3. Phạm vi của các thông số đầu vào và đầu ra của mạng neuron (368 điểm của một giếng khoan thuộc bể Nam Côn Sơn)

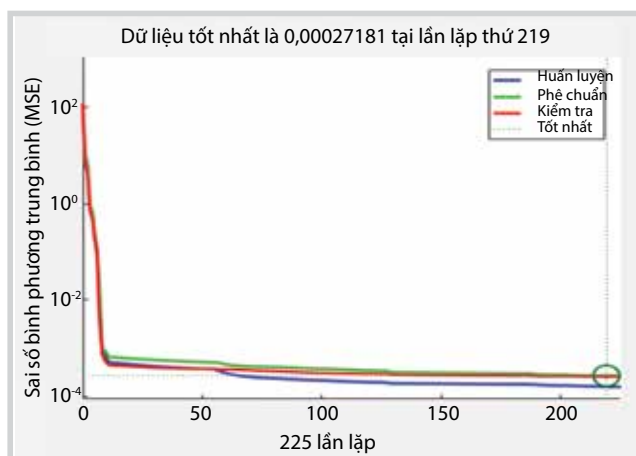
Thông số	Giá trị nhỏ nhất	Giá trị lớn nhất
Độ sâu (m)	225,47	3413,12
Ứng suất địa tĩnh (ppg)	667,101	13872,7
Áp suất lỗ rỗng (ppg)	8,81267	18,8291
Hệ số Poisson	0,188404738	0,414834946
Áp suất nứt vỉa (ppg)	7,90831	18,6784



Hình 3. Sơ đồ cấu trúc mạng neuron

Bảng 4. Trọng số và ngưỡng của lớp ẩn thứ nhất

Neuron	Lớp ẩn 1				Ngưỡng
	Trọng số				
	1	2	3	4	
1	-0,52048	1,5535	-0,01865	-2,2423	2,3779
2	0,61856	-0,01363	-0,94171	-2,0502	-2,2575
3	-1,2951	-0,46325	1,5402	-1,6688	1,262
4	1,3596	-0,21983	1,8014	-0,74996	-1,0919
5	-2,3496	-1,3583	0,08929	-0,45597	0,43587
6	1,7719	0,057823	0,90068	-1,2678	-0,41494
7	-0,98796	0,53053	0,86185	2,1558	0,057315
8	-0,66979	1,8355	-1,8686	-0,23488	-0,66435
9	0,37842	1,5757	-1,098	-1,0741	1,125
10	1,488	0,79276	-0,02608	1,765	1,7931
11	-1,512	-0,40828	-1,0153	1,7788	-2,1895
12	1,6715	1,2146	1,3159	-0,3573	2,6654



Hình 4. Đồ thị biểu diễn hiệu suất của ANN 4-12-10-1

chọn số lượng neuron trong mỗi lớp. Số lượng neuron nhỏ sẽ đào tạo mạng nhanh hơn nhưng không cho kết quả chính xác. Trong khi đó, tăng số lượng neuron có thể tăng thời gian xử lý. Nhóm tác giả sử dụng cấu trúc mạng với 12 neuron trong lớp ẩn thứ nhất và 10 neuron trong lớp ẩn thứ hai (mạng 4-12-10-1). Lựa chọn số lượng neuron không có nguyên tắc rõ ràng nhưng chỉ cần đảm bảo đáp ứng hệ thống sẵn có của người sử dụng, độ chính xác cuối cùng chấp nhận được.

Hình 3 biểu diễn cấu trúc mạng mô phỏng. Những hàm hoạt động cho những lớp ẩn là hàm tag-sigmoid và

log-sigmoid, hàm pure linear được sử dụng cho lớp đầu ra. Những hàm đó được biểu diễn như sau:

$$tansig(x) = \frac{2}{1 + e^{-2x}} - 1; \text{logsig}(x) = \frac{1}{1 + e^{-x}}; \text{purelin}(x) = x$$

3.3. Phân tích hiệu suất mạng neuron sau khi huấn luyện

Sau khi mạng được đào tạo và đạt tới 225 phép lặp có thể thu được đồ thị biểu diễn hiệu suất của quá trình đào tạo như Hình 4. Hình 4 biểu diễn tổng sai số bình phương cho 3 tập dữ liệu với số lần lặp cho mạng có cấu trúc 4-12-10-1 sử dụng thuật toán đào tạo lan truyền ngược LMA (Levenberg-Marquardt Algorithm). Ứng xử của đường cong thể hiện quá trình đào tạo thành công và 3 tập dữ liệu được chọn phù hợp từ tập dữ liệu ban đầu. Hiệu suất tốt nhất là $MSE = 2,7181 \times 10^{-4}$ tại lần lặp thứ 219.

Các thông số về trọng số và ngưỡng của các lớp trong mô hình được trình bày trong Bảng 4 - 6.

Hình 5 trình bày dữ liệu đào tạo, xác nhận, kiểm tra và dữ liệu tổng. Đường nét đứt thể hiện kết quả sau khi đào tạo trùng với dữ liệu mục tiêu. Đường nét liền trình bày sự hồi quy tuyến tính tốt nhất giữa kết quả đào tạo và mục tiêu. Giá trị R biểu thị một mối quan hệ giữa kết quả và mục tiêu. Nếu R = 1 thì đó là một mối quan hệ tuyến tính

Bảng 5. Trọng số và ngưỡng của lớp ẩn thứ hai

		Lớp ẩn 2												Ngưỡng
		Trọng số												
		1	2	3	4	5	6	7	8	9	10	11	12	
-1,4609	-0,83637	0,32721	-0,34836	0,467	-0,61183	1,3799	0,75256	1,4912	1,3204	-1,0869	3,4815			
-1,017	-0,3032	-1,139	-0,45539	-1,3098	0,92884	0,96758	-0,10303	0,90322	-1,8692	0,072192	2,675			
-0,48565	0,10705	-0,32639	1,9532	0,79354	-1,1854	0,40528	-0,17407	-1,3132	-1,6964	-0,60623	1,7451			
0,3212	-0,91785	1,3249	-1,6033	-0,35895	-1,5348	-1,1905	-0,99824	0,14685	-1,1742	0,9503	-1,1224			
-1,1074	-0,79508	0,13514	1,5733	0,56099	0,97494	-1,0279	0,82365	-0,08593	-0,46225	1,6648	0,35091			
0,25754	0,70328	0,01723	1,1042	-0,10609	-0,02995	1,1658	-1,4554	1,1411	0,94204	-1,1189	0,34509			
0,71036	-1,3651	-0,71857	0,83441	0,81094	-0,06848	-1,391	-1,3585	0,47523	0,88189	1,262	1,2354			
-0,84301	1,3998	0,018484	-1,497	0,76883	1,1825	-0,11162	0,067826	1,1283	-1,5051	-1,0708	-1,8199			
-1,5087	0,77821	0,30796	-0,78059	1,36	0,40713	-1,0926	-1,4463	-0,24013	-0,25396	-1,2874	-2,775			
-0,79359	-0,03767	0,85067	-0,75599	1,7767	0,45024	1,9821	1,4187	-0,2011	0,061264	-0,5548	-3,3202			

Bảng 6. Trọng số và ngưỡng của lớp đầu ra

		Lớp đầu ra										Ngưỡng
		Trọng số										
		1	2	3	4	5	6	7	8	9	10	
0,50363	0,63168	-0,13513	-0,51143	0,49316	0,52492	0,19885	-0,53623	0,15883	-0,75647	-1,1629		

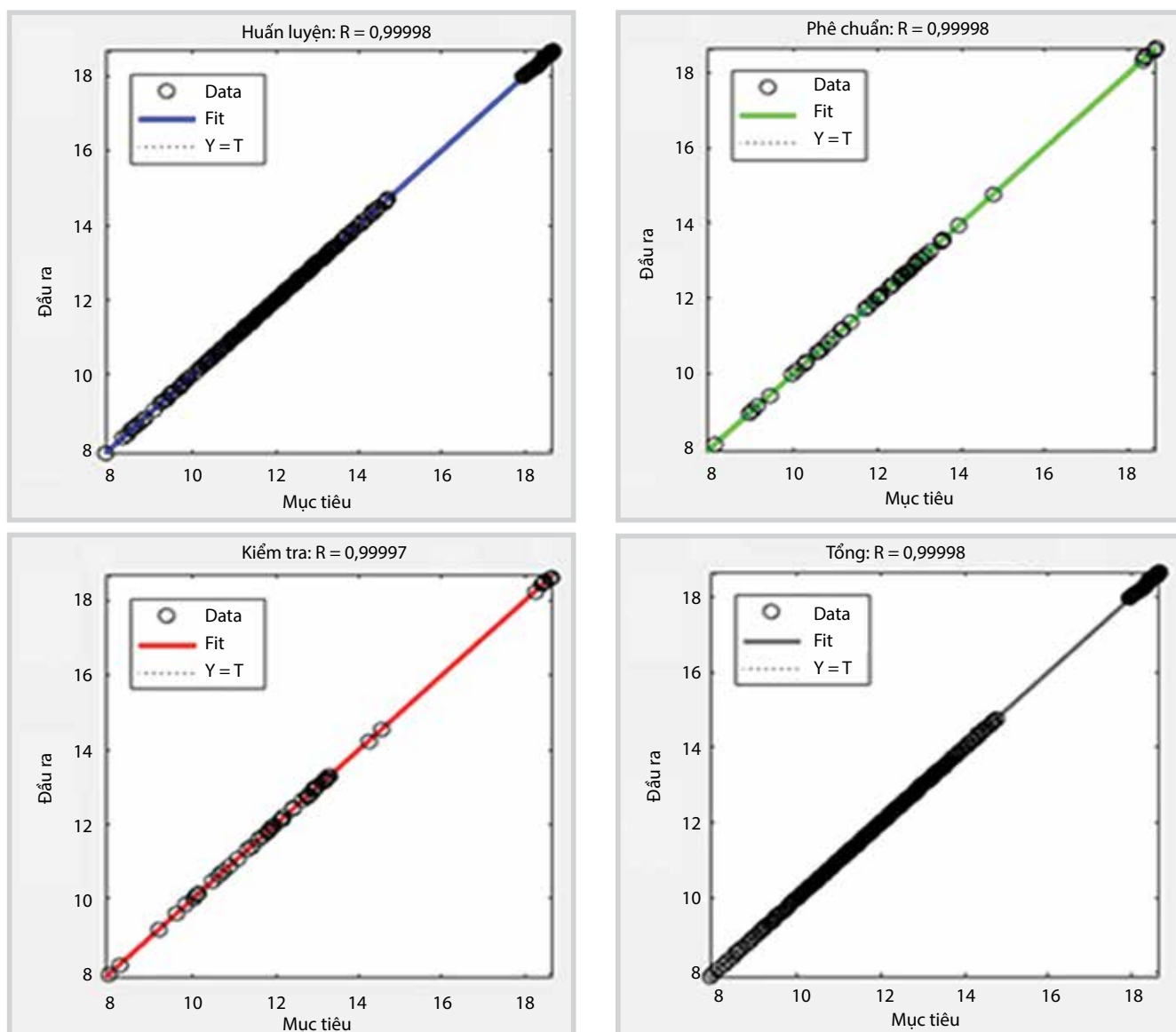
chính xác giữa kết quả đào tạo và mục tiêu [29].

Kết quả của dữ liệu được đào tạo bởi mô hình ANN trong nghiên cứu này chỉ ra đây là một mối quan hệ chính xác: kết quả đào tạo, kết quả xác nhận, kiểm tra và kết quả tổng cộng đều cho các giá trị R lớn hơn 0,9995.

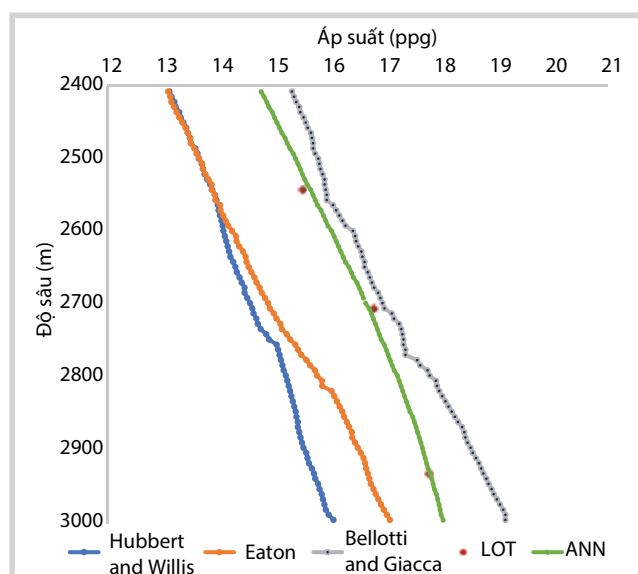
3.4. Sử dụng ANN để dự báo áp suất nứt vỉa, so sánh với các phương pháp truyền thống

Sau khi có mô hình dự báo với độ tin cậy cao, nhóm tác giả tiến hành dự báo áp suất nứt vỉa cho một giếng khoan thuộc bể Nam Côn Sơn. Kết quả dự báo được trình bày trong Hình 6, trên hình cũng thể hiện kết quả dự báo bằng phương pháp truyền thống thường và phương pháp xác định áp suất nứt vỉa trực tiếp thông qua thí nghiệm "Leak off test", LOT tại 3 độ sâu của giếng: 2542,59m, 2705,42m và 2932,09m (Bảng 1).

Thông qua kết quả so sánh trong Hình 6 có thể thấy rằng mô hình lý thuyết Hubbert và Willis, Eaton nằm khá xa kết quả đo thực nghiệm LOT, trong khi kết quả của mô hình Bellotti và Glacca nằm sát hơn nhưng vẫn còn khác biệt. Ngược lại, đường cong dự báo áp suất nứt vỉa bằng phương pháp ANN qua các điểm áp suất nứt vỉa của thí nghiệm LOT. Vì vậy, phương pháp ANN có kết quả dự báo chính xác nhất so với các phương pháp dự báo truyền thống (Hubbert and Willis, Eaton, Bellotti and Giacca).



Hình 5. Đồ thị hồi quy của ANN



Hình 6. Áp suất nứt via theo các phương pháp dự báo, xác định khác nhau

4. Kết luận

Nhóm tác giả đã thực hiện dự báo áp suất nứt vỉa cho một giếng khoan thuộc bể Nam Côn Sơn, với số điểm dữ liệu là 368. Từ kết quả nghiên cứu có thể thấy phương pháp sử dụng ANN có ưu điểm sau:

- Sử dụng nhiều điểm dữ liệu khác nhau, phù hợp với vùng nghiên cứu rộng, dữ liệu đa dạng;
- Thông số đầu vào cho mô hình dễ dàng có được;
- Đơn giản và dễ sử dụng trong quá trình trước khi khoan;
- Độ chính xác cao hơn so với các phương pháp dự báo truyền thống.

Kết quả dự báo áp suất nứt vỉa cho giếng khoan ở bể Nam Côn Sơn đủ độ tin cậy để sử dụng cho quá trình

thiết kế giếng ở khu vực lân cận và cần kiểm nghiệm thêm trước khi nhân rộng tính ứng dụng của phương pháp này.

Tài liệu tham khảo

1. M.King Hubbert, David G.Willis. *Mechanics of hydraulic fracturing*. Society of Petroleum Engineers. 1957; 210: p. 153 - 163.
2. E.S.Pennebaker. *An engineering interpretation of seismic data*. Fall Meeting of the Society of Petroleum Engineers of AIME, Houston, Texas. 29 September - 2 October.
3. Ben A.Eaton. *Fracture gradient prediction and its application in oilfield operations*, Journal of Petroleum Technology. 1969; 21(10): p. 1353 - 1360.
4. W.R.Matthews, John Kelly. *How to predict formation pressure and fracture gradient from electric and sonic logs*. Oil and Gas Journal. 1967: p. 39 - 43.
5. L.A.MacPherson, L.N.Berry. *Prediction of fracture gradients from log derived moduli*. The Log Analyst. 1972.
6. Stan A.Christman. *Offshore fracture gradients*. Journal of Petroleum Technology. 1973; p. 910 - 914.
7. R.A.Anderson, D.S.Ingram, A.M.Zanier. *Determining fracture pressure gradients from well logs*. Journal of Petroleum Technology. 1973.
8. P.Bellotti, D.Giacca. *Seismic data can detect overpressures in deep drilling*. Oil and Gas Journal. 1978.
9. E.M.Shokir. *Neuron network determines shaly-sand hydrocarbon saturation*. Oil & Gas Journal. 2001.
10. Henrique V.da Silva, Celso K.Morooka, Ivan R.Guilherme, Tiago C.da Fonseca, José R.P.Mendes. *Leak detection in petroleum pipelines using a fuzzy system*. Journal of Petroleum Science and Engineering. 2005; 49(3 - 4): p. 223 - 238.
11. S.J.Cuddy, P.W.J.Glover. *The application of fuzzy logic and genetic algorithms to reservoir characterization and modeling*. Soft Computing for Reservoir Characterization and Modeling. 2002.
12. Alpana Bhatt, Hans B.Helle. *Committee neuron networks for porosity and permeability prediction from well logs*. Geophysical Prospecting. 2002: p.645 - 660.
13. Fatai Adesina Anifowose, Abdulazeez Abdurraheem. *Prediction of porosity and permeability of oil and gas reservoirs using hybrid computational intelligence models*. North Africa Technical Conference and Exhibition, Cairo, Egypt. 14 - 17 February, 2010.
14. S.R.Shadizadeh, F.Karimi, M.Zoveidavianpoor. *Drilling stuck pipe prediction in Iranian oil fields: An artificial neuron network approach*. Iranian Journal of Chemical Engineering. 2010; 7(4): p. 29 - 41.
15. Fatai Adesina Anifowose, AbdIAzeem Oiyafemi Ewenla, Safriyu Ijiyemi Eludiora. *Prediction of oil and gas reservoir properties using support vector machines*. International Petroleum Technology Conference, Bangkok, Thailand. 15 - 17 November, 2011.
16. R.Gholami, A.R.Shahraki, M.Jamali Paghaleh. *Prediction of hydrocarbon reservoirs permeability using support vector machine*. Mathematical Problems in Engineering. 2012.
17. HZ Raja, F Sormo, ML Vinther. *Case-based reasoning: predicting real-time drilling problems and improving drilling performance*. SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain. 25 - 28 September, 2011.
18. Md. Alhaz Uddin, Mohammed Jameel, Hashim Abdul Razak. *Application of artificial neuron network in fixed offshore structures*. Indian Journal of Geo-Marine Sciences. 2012.
19. R.Keshavarzi, R.Jahanbakhshi, M.Rashidi. *Predicting formation fracture gradient. In oil and gas wells: A neuron network approach*. 45th U.S. Rock Mechanics/ Geomechanics Symposium, San Francisco, California. 26 - 29 June, 2011.
20. M.Heidarian, H.Jalalifar, A.Rafati. *Prediction of rock strength parameters for an Iranian oil field using neurofuzzy method*. Journal of AI and Data Mining. 2016; 4(2): p. 229 - 234.
21. T.O.Odedele, H.D.Ibrahim. *Predicting oil well gas lift performance and production optimization using hybrid particle swarm optimization and fuzzy support vector machines*. World Congress on Engineering. 2016.
22. Schlumberger Oilfield Glossary. Oilfield glossary. <http://www.glossary.oilfield.slb.com/>.
23. Adam Bourgoyne Jr, Keith Miliheim, Martin Chenevert, KS Young Jr. *Applied drilling engineering*. Society of Petroleum Engineers. 1986.
24. Jincai Zhang, Shang-Xian Yin. *Fracture gradient prediction: an overview and an improved method*. Petroleum Science. 2017; 14(4): p. 720 - 730.

25. Opeyemi Bello, Javier Holzmann, Tanveer Yaqoob, Catalin Teodoriu. *Application of artificial intelligence methods in drilling system design and operations: a review of the state of the art*. Society of Petroleum Engineers. 2015; 5(2): p. 121 - 139.
26. Agnar Aamodt, Enric Plaza. *Case-Based reasoning: Foundational issues, methodological variations, and system approaches*. Artificial Intelligence Communications. 1994; 7(1): p. 39 - 52.
27. Warren S. McCulloch, Walter Pitts. *A logical calculus of the ideas immanent in nervous activity*. Bulletin of Mathematical Biophysics. 1943; 5(4): p.115 - 133.
28. Adel Malalah, Ibrahim Sami N a s h a w i . *Estimating the fracture gradient coefficient using neuron networks for a field in the Middle East*. Journal of Petroleum Science and Engineering. 2005; 49(3 - 4): p. 193 - 211.
29. M.H.Beale, M.T.Hagan, H.B.Demuth. *Neuron network toolbox user's guide*. The MathWorks. 2015.

APPLICATION OF ARTIFICIAL NEURAL NETWORK IN FRACTURE PRESSURE PREDICTION

Nguyen Van Hung, Dang Huu Minh

Petrovietnam University

Email: hungnv@pvu.edu.vn

Summary

Prediction of formation fracture pressure is an essential task in designing safer drilling operations and economical well planning, allowing effective control, operation and stimulation. Errors in fracture pressure prediction can lead to several serious problems such as lost circulation and kick, and even blowout.

This paper presents an overview on application of artificial intelligent in the petroleum industry. Then an artificial neural network model will be used with depth, overburden stress, Poisson's ration, and pore pressure as the input data and fracture pressure as the output data of one well in Nam Con Son basin for predicting formation fracture pressure. The results obtained from the model are compared with those obtained from conventional method. The comparison shows that the ANN method is promising and under some circumstances it is superior to the available techniques.

Key words: Formation fracture pressure, artificial neural network, Nam Con Son basin.