



NGHIÊN CỨU ỨNG DỤNG KỸ THUẬT HỌC SÂU VỚI MÔ HÌNH ĐƠN BIẾN BỘ NHỚ DÀI - NGẮN HẠN TRONG HỆ THỐNG QUAN TRẮC VÀ DỰ BÁO NHIỆT ĐỘ THAN TỰ CHÁY MỎ HẦM LÒ

Nguyễn Thế Vinh¹, Nguyễn Lê Thùy Dương^{1,*}, Nguyễn Xuân Đồng¹
Nguyễn Hùng Kiên², Nguyễn Danh Dũng², Trần Hải Dương²

¹Viện nghiên cứu Điện tử, Tin học, Tự động hóa –156A Quán Thánh, P. Quán Thánh, Q. Ba Đình, TP. Hà Nội
²Công ty TNHH MTV PTCN Điện tử, Tự động hóa–156A Quán Thánh, P. Quán Thánh, Q. Ba Đình, TP. Hà Nội

THÔNG TIN BÀI BÁO

CHUYÊN MỤC: Công trình khoa học

Ngày nhận bài: 22/10/2024

Ngày nhận bài sửa: 28/10/2024

Ngày chấp nhận đăng: 02/11/2024

*Tác giả liên hệ:

Email: duongthuynl@gmail.com

TÓM TẮT

Quan trắc trực tuyến và dự báo nhiệt độ vỉa có than tự cháy trong mỏ hầm lò là vấn đề cấp thiết, hiện đang được quan tâm. Bài báo này đề xuất sử dụng phương pháp xây dựng mô hình dự báo nhiệt độ than tự cháy từng giờ trong tám giờ kế tiếp bằng mô hình đơn biến bộ nhớ dài - ngắn hạn (LSTM). Các tham số của mô hình được điều chỉnh thông qua các kiểm thử phù hợp với bài toán đặt ra. Hệ thống quan trắc kết hợp với phương pháp dự báo này góp phần nâng cao hiệu quả sản xuất, an toàn lao động, bảo vệ môi trường, sử dụng hiệu quả tài nguyên than Việt Nam.

Từ khóa: học sâu, hệ thống quan trắc, than tự cháy, bộ nhớ dài - ngắn hạn.

@ Hội Khoa học và Công nghệ Mỏ Việt Nam

1. ĐẶT VẤN ĐỀ

Trong lịch sử khai thác than hầm lò tại Việt Nam đã xảy ra một số vụ cháy mỏ mà nguyên nhân được cho là than có tính tự cháy cao (hay còn gọi là cháy nội sinh). Các vụ cháy nội sinh đã xảy ra tại các Công ty Than Phấn Mễ - Làng Cẩm, Khe Bó, Khe Chuối - Công ty TNHH MTV 91, Tràng Khê II, III - Công ty TNHH MTV Than Hồng Thái, Công ty than Hà Lâm, Khánh Hòa, v.v. Các vụ cháy tiềm ẩn nhiều nguy cơ gây mất an toàn, khó kiểm soát và đã ảnh hưởng nhiều đến các hoạt động sản xuất của mỏ, đặc biệt phải tăng thêm chi phí xử lý, làm tăng giá thành, giảm hiệu quả khai thác [2].

Hiện tượng than tự cháy được giải thích theo các hướng khác nhau dựa trên các nghiên cứu độc lập, riêng biệt với tính chất các vỉa than khác nhau, điều kiện địa chất khác nhau,... Quá trình oxy hóa than

xuất hiện khi oxy phản ứng với than. Phản ứng này chuyển từ dạng phản ứng vật lý sang phản ứng hóa học tùy thuộc điều kiện nhiệt độ. Phản ứng oxy hóa sinh nhiệt, CO, CO₂ và hơi nước. Nghiên cứu quy luật các loại khí sinh ra trong quá trình oxy hóa mẫu than trong phòng thí nghiệm [2],[8], người ta nhận thấy các thông số như H₂, N₂, CH₄, C₂H₂, C₂H₄, C₂H₆, C₃H₆ có thay đổi nhưng ở nhiệt độ cao (>100 °C), trong khi đó hàm lượng khí CO₂ hầu như không thay đổi. Do đó, những thông số này không có ý nghĩa nhiều trong quan trắc trực tuyến và cảnh báo hiện tượng than tự cháy trong mỏ hầm lò. Trong khi đó, nhiệt độ vỉa than, hàm lượng khí CO, hàm lượng khí O₂ là các thông số thay đổi ngay từ giai đoạn ủ nhiệt. Do vậy, chúng được sử dụng là các thông số chính cần quan trắc, cảnh báo, dự báo hiện tượng tự cháy của than trong mỏ hầm lò và đây cũng là những



thông số mà hầu hết các thiết bị/hệ thống quan trắc trực tuyến than tự cháy trên thế giới sử dụng [3],[4],[6],[9].

Các công trình khoa học công bố gần đây cho thấy “học sâu” (*Deep Learning*) có thể là một phương pháp thích hợp cho dữ liệu chuỗi thời gian trong nhiều lĩnh vực. Lim and Zohren [6] đã nghiên cứu mô hình “học sâu” kết hợp mô hình thống kê và các thành phần mạng nơ-ron để cải thiện các phương pháp dự báo chuỗi thời gian. Ikram và cộng sự [4] đã trình bày một nghiên cứu sử dụng mạng nơ-ron tuần hoàn với mô hình đơn biến “bộ nhớ dài - ngắn hạn” (*Long Short Term Memory - LSTM*) để dự đoán nhiệt độ môi trường xung quanh. H.B. Sahu [5] sử dụng mô hình hệ thống suy luận thần kinh mờ thích nghi để dự báo quá trình tự ôxy hóa của than dựa trên kết quả thử nghiệm với 50 mẫu than thu thập từ các mỏ than khác nhau của Ấn Độ. Zhenhua Xie đề xuất mô hình mạng nơ-ron dự đoán nhiệt độ than tự cháy với đầu vào là các giá trị của hàm lượng CO và CO₂.

Trong bài báo này, hệ thống quan trắc trực tuyến hiện tượng than tự cháy trong mỏ hầm lò ứng dụng kỹ thuật “học sâu” với mô hình bộ nhớ LSTM được đề xuất để dự báo về nhiệt độ than biến thiên trong từng giờ và trong tám giờ tiếp theo trong các vỉa than mỏ hầm lò. Đầu tiên, dữ liệu được sắp xếp thứ tự theo thời gian, được xem là chuỗi thời gian. Kế đến là bước biến đổi dữ liệu từ chuỗi thời gian thành biến đầu vào cho mô hình. Sau cùng, mô hình LSTM được huấn luyện và kiểm thử thông qua dữ liệu trong phòng thí nghiệm.

2. DỮ LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Giới thiệu hệ thống quan trắc trực tuyến than tự cháy trong mỏ hầm lò

Mô hình hệ thống hệ thống điều khiển giám sát và thu thập dữ liệu (*Supervisory Control And Data Acquisition - SCADA*) xây dựng để quan trắc than tự cháy gồm có các phần tử: Trung tâm giám sát, cảnh báo (đặt tại phòng điều khiển); Thiết bị giám sát khu vực phòng nổ; Đầu đo nhiệt độ vỉa than an toàn tia lửa; Đầu đo khí CO an toàn tia lửa; Đầu đo khí O₂ an toàn tia lửa (lắp đặt dưới hầm lò). Phần mềm của hệ thống có chức năng: Hiển thị nhiệt độ vỉa than dưới dạng số và đồ thị; Hiển thị hàm lượng khí dưới dạng số và đồ thị; Hiển thị cảnh báo và hướng xử lý lỗi; Cài

đặt ngưỡng cảnh báo; Lưu trữ số liệu; Chức năng phân quyền;... Cấu trúc hệ thống và giao diện chính của phần mềm hệ thống được trình bày trong bài báo [1].

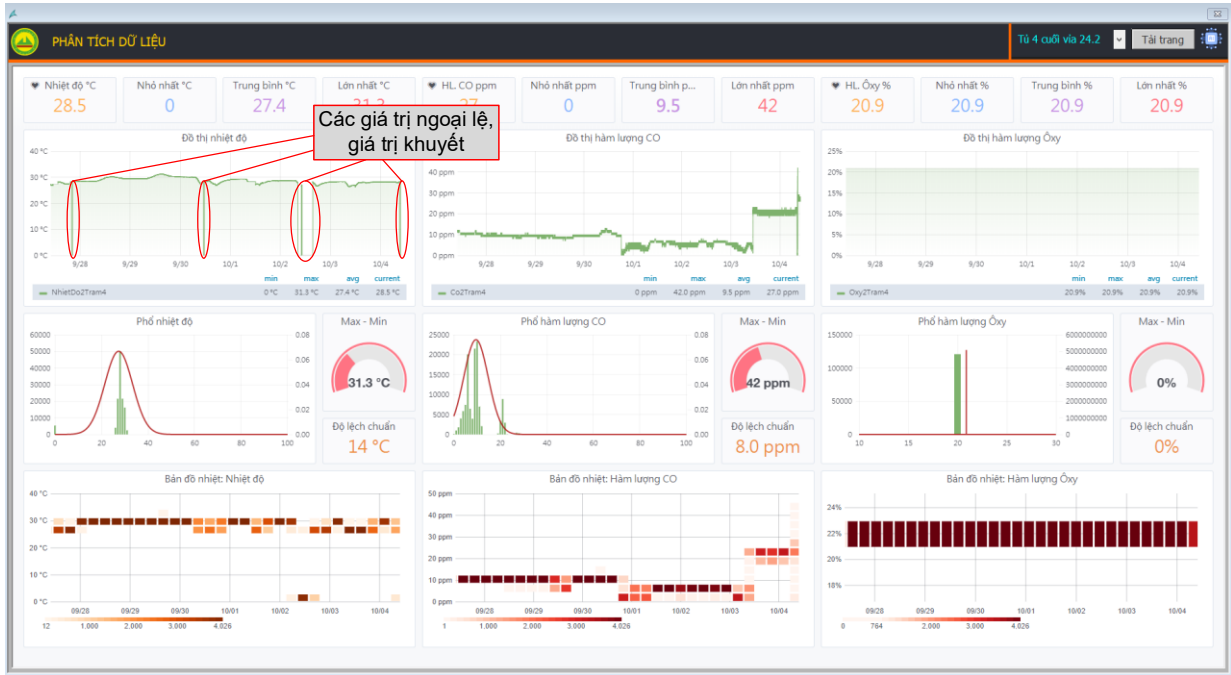
2.2. Dự báo nhiệt độ than tự cháy trong mỏ hầm lò với mô hình LSTM

Với sự phát triển của các thuật toán “học máy” (*Machine Learning*) trong thời gian gần đây đã cung cấp thêm hướng tiếp cận mới với việc xử lý và dự báo chuỗi thời gian đạt được độ chính xác cao. Có thể kể đến mô hình hồi quy tự động ARIMA, SARIMA, hay phương pháp tiếp cận mới hơn nhằm khắc phục những nhược điểm của các mô hình truyền thống là các mạng “học sâu” (*Deep Learning*), điển hình là mạng nơ-ron hồi quy và phiên bản mở rộng của nó là “bộ nhớ dài - ngắn hạn” LSTM. Mạng LSTM được sử dụng nhiều trong các bài toán dự báo chuỗi thời gian với kết quả khá quan trọng nhờ có khả năng ghi nhớ các bước và không bị ảnh hưởng nhiều khi số liệu đầu vào bị thiếu [2], [7].

Trong nghiên cứu này, chúng tôi đề xuất ứng dụng kỹ thuật “học sâu” kết hợp với hệ thống SCADA truyền thống để xây dựng một hệ thống SCADA thế hệ mới nhằm dự đoán trước những diễn biến của than tự cháy với mục đích hỗ trợ mỏ giải quyết vấn đề trước khi trạng thái tự cháy trong các vỉa than trở nên nghiêm trọng. Phương pháp tiếp cận là dựa trên kỹ thuật “học sâu” với mô hình LSTM đơn biến để dự báo nhiệt độ than tự cháy mỏ hầm lò trong 8 giờ tiếp theo nhằm hỗ trợ các mỏ chủ động lập kế hoạch phòng ngừa và đưa ra các giải pháp khai thác an toàn, hiệu quả. Với nghiên cứu này, đầu ra của mô hình LSTM là chuỗi giá trị nhiệt độ vỉa than tự cháy được dự báo từng giờ trong 8 giờ tiếp theo, đầu vào là chuỗi giá trị nhiệt độ trong vòng một tuần trước đó. Chuỗi giá trị của tập dữ liệu đầu vào gồm có 120.960 giá trị được hệ thống lấy mẫu với chu kỳ 5 giây/lần.

2.2.1. Tiền xử lý chuỗi dữ liệu

Trước khi áp dụng mô hình dự báo nhiệt độ than tự cháy, chuỗi dữ liệu đầu vào cần được xử lý qua thuật toán làm sạch với hai bước: (a) Xử lý các giá trị bị khuyết NaN (*Not a number*) và (b) Xử lý các giá trị ngoại lệ với quy tắc 3 σ cho phân phối chuẩn. Hình 1 minh họa dữ liệu thô có giá trị bị khuyết, giá trị ngoại lệ.



Hình 1. Minh họa dữ liệu thô

a) Xử lý các giá trị bị khuyết

Trong quá trình hoạt động của hệ thống, có những trường hợp mất điện, bảo trì, sửa chữa,... khi đó dữ liệu bị khuyết, tức là NaN (*Not a Number*). Có nhiều giải pháp để xử lý giá trị NaN. Cách thứ nhất là bỏ dữ liệu bị khuyết ra khỏi quá trình xây dựng mô hình. Cách làm này đơn giản nhưng có những hạn chế nhất định. Nếu có quá nhiều dữ liệu bị khuyết thì sẽ không còn thông tin gì cho việc xây dựng hoặc sử dụng mô hình. Cách thứ hai được lựa chọn vì giúp ta có thể giải quyết vấn đề dữ liệu bị khuyết là thay thế các giá trị bị khuyết bởi một giá trị trung bình của tập dữ liệu rồi dùng giá trị đó để xây dựng, sử dụng mô hình.

b) Xử lý các giá trị ngoại lệ

Trong chuỗi dữ liệu nhiệt độ vỉa than có những giá trị ngoại lệ có thể xuất phát từ nhiều nguyên nhân: Thiết bị đo nhiệt độ vỉa than có thể bị lỗi phần cứng; Nhiễu điện từ hoặc môi trường có thể làm sai lệch giá trị nhiệt độ;.... Tiêu chuẩn hóa z-score là một phương pháp phổ biến để xác định các giá trị ngoại lệ trong một tập dữ liệu. Việc xác định và loại bỏ các giá trị ngoại lệ là rất quan trọng để đảm bảo tính chính

xác và tin cậy của dữ liệu nhiệt độ vỉa than. Sử dụng phương pháp tiêu chuẩn hóa z-score có thể giúp phát hiện và xử lý các giá trị bất thường này một cách hiệu quả trước khi được sử dụng trong mô hình dự báo. Trong phân phối chuẩn, giả sử μ là giá trị trung bình và σ là độ lệch chuẩn. Quy tắc 3 σ cho phân phối chuẩn nói rằng: 68% các điểm dữ liệu nằm trong khoảng $\mu \pm \sigma$; 95% các điểm dữ liệu nằm trong khoảng $\mu \pm 2\sigma$; 99.7% các điểm dữ liệu nằm trong khoảng $\mu \pm 3\sigma$. Với một giá trị dữ liệu x thì điểm chuẩn z-score của nó được tính bởi công thức (1) [10]:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

trong đó:

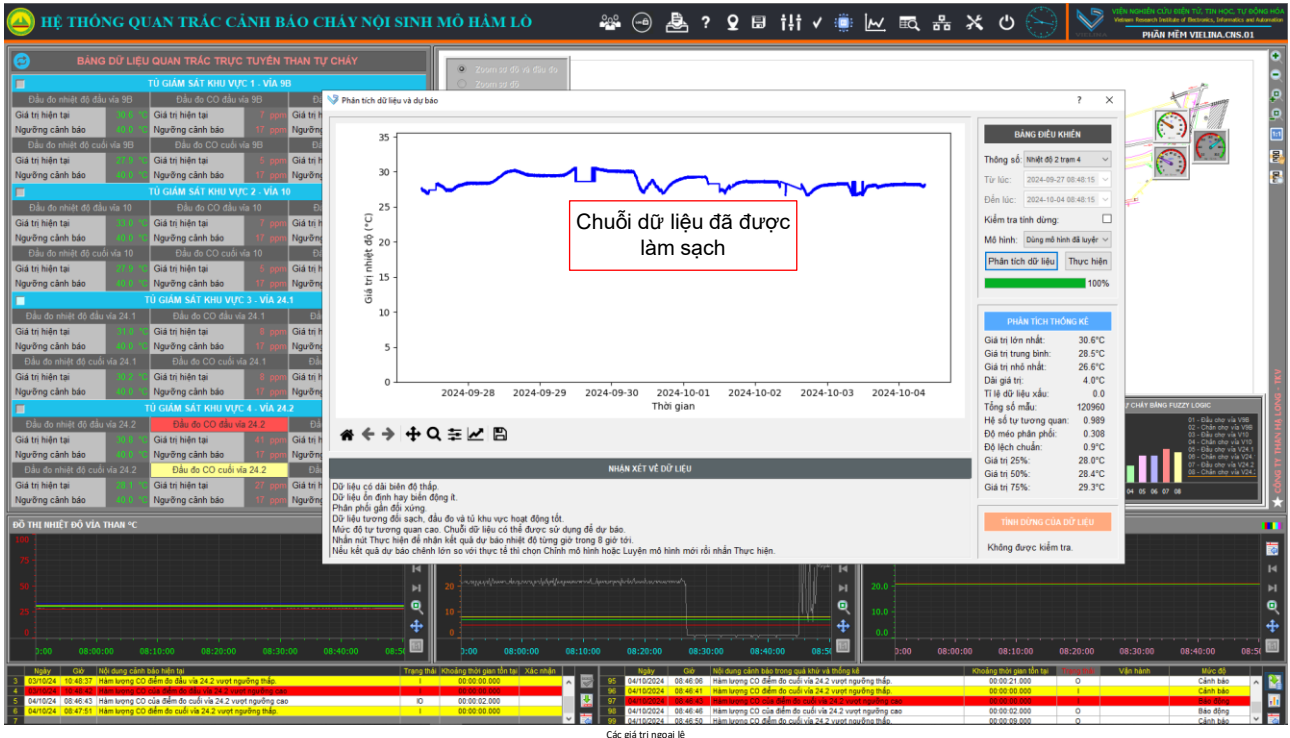
z là điểm chuẩn z-score của một giá trị dữ liệu x thuộc tập dữ liệu;

x là giá trị của điểm dữ liệu cần tính z-score;

μ là giá trị trung bình của tập dữ liệu;

σ là độ lệch chuẩn của tập dữ liệu.

Những điểm có điểm chuẩn z-score nằm ngoài đoạn $[-3, 3]$ có thể được coi là các điểm ngoại lệ, từ đó thay thế các giá trị ngoại lệ bằng giá trị trung bình. Hình 2 minh họa dữ liệu đã được xử lý.



Hình 2. Minh họa kết quả dữ liệu đã được xử lý

2.2.2 Xây dựng mô hình dự báo

Bước tiền xử lý chuỗi dữ liệu trình bày phần trên được tự động thực hiện bằng phần mềm của hệ thống. Sau khi được làm sạch, thu được mỗi chuỗi có độ dài 120.960 mẫu ứng với 7 ngày x 24 giờ. Từ đó, đề xuất sử dụng mô hình LSTM với cấu trúc: 200 đơn vị ở lớp đầu vào, 200 đơn vị ở lớp ẩn và 8 đơn vị ở lớp đầu ra. Hàm kích hoạt ReLU và tối ưu hóa

Adam giúp mô hình hội tụ nhanh và hiệu quả. Mô hình dự đoán 8 bước thời gian tiếp theo dựa trên 24 bước thời gian trước đó. Việc chọn cửa sổ trượt 24 bước có nghĩa là mô hình sẽ sử dụng 24 giá trị nhiệt độ trước đó để dự báo 8 giá trị nhiệt độ tiếp theo. Các siêu tham số (hyperparameters) của mô hình được trình bày chi tiết trong Bảng 1.

Bảng 1. Các siêu tham số của mô hình dự báo

| TT | Tham số | Giá trị | Giải thích |
|----|------------------------|---------|---|
| 1 | n_steps_in | 24 | Số bước đầu vào cho mô hình là 24 (ứng với 24 giờ). |
| 2 | n_steps_out | 8 | Số bước đầu ra mà mô hình dự đoán (dự đoán 8 giờ tiếp). |
| 3 | n_features | 1 | Số đặc trưng đầu vào (một loại dữ liệu, cụ thể là nhiệt độ). |
| 4 | LSTM units (1st layer) | 200 | Lớp này có vai trò quan trọng trong việc xử lý và học các đặc trưng từ dữ liệu chuỗi thời gian. |
| 5 | LSTM units (2nd layer) | 200 | Có nhiệm vụ xử lý thêm và học sâu hơn các đặc trưng từ dữ liệu chuỗi thời gian. Thông tin từ chuỗi thời gian trước đó (do lớp LSTM đầu tiên cung cấp) được xử lý thêm giúp mô hình hiểu rõ hơn về quan hệ giữa các bước thời gian trong chuỗi, nhờ đó cải thiện độ chính xác của dự đoán. Đầu ra của lớp LSTM thứ hai (còn được gọi là hidden state cuối cùng) sẽ được chuyển đến lớp Dense. Lớp Dense sẽ sử dụng thông tin này để thực hiện các dự đoán cuối cùng cho n_steps_out = 8. |
| 6 | Dense units | 8 | Lớp Dense có nhiệm vụ tạo ra đầu ra cuối cùng của mô hình. Mỗi đơn vị trong lớp Dense sẽ dự đoán một giá trị của chuỗi đầu ra. |



| TT | Tham số | Giá trị | Giải thích |
|----|-----------|---------|---|
| | | | Với 8 đơn vị, lớp Dense sẽ tạo ra một chuỗi đầu ra gồm 8 bước thời gian ứng với kết quả dự báo 8 giờ tiếp theo. |
| 7 | Epochs | 200 | Mô hình sẽ trải qua toàn bộ tập dữ liệu huấn luyện 200 lần (200 là giá trị được lựa chọn sau nhiều lần thử nghiệm với tiêu chí về thời gian huấn luyện, giá trị của hàm mất mát, tính tổng quát hóa của mô hình). |
| 8 | Optimizer | Adam | Thuật toán tối ưu Adam thường hội tụ nhanh và hiệu quả trên nhiều loại bài toán khác nhau. Khả năng điều chỉnh tốc độ học tự động giúp giảm thiểu nguy cơ học quá mức hoặc học không đủ. |

Nội dung chính của bảng là mô tả các siêu tham số của mô hình dự báo dựa trên kiến trúc LSTM sử dụng các thư viện TensorFlow và Keras. Tham số "n_steps_in" xác định số bước đầu vào của mô hình, ở đây là 24, tương ứng với việc mô hình sẽ nhận 24 giá trị liên tiếp từ chuỗi thời gian (nhiệt độ) để dự báo giá trị tương lai. Việc chọn số bước đầu vào lớn giúp mô hình có thêm thông tin về xu hướng quá khứ để cải thiện dự báo. Tham số tiếp theo là "n_steps_out", đây là số bước mà mô hình sẽ dự báo trong tương lai, ở đây là 8, nghĩa là mô hình sẽ dự đoán các giá trị cho 8 thời điểm tiếp theo (tương ứng với dự báo 8 giờ tới). Số bước đầu ra phù hợp giúp cân bằng giữa độ phức tạp của mô hình và tính chính xác của dự báo. Vì mô hình chỉ sử dụng một đặc trưng duy nhất từ dữ liệu, cụ thể là nhiệt độ nên tham số "n_features" được thiết lập là 1, dạng đơn biến (univariate). LSTM units (1st layer) là số lượng đơn vị LSTM trong lớp đầu tiên. Với 200 đơn vị, mô hình có khả năng học các mẫu từ chuỗi thời gian một cách chi tiết. Số lượng lớn đơn vị giúp mô hình nắm bắt nhiều thông tin phức tạp từ dữ liệu. Lớp LSTM thứ hai cũng có 200 đơn vị, tiếp tục học các đặc trưng sâu hơn từ lớp LSTM đầu tiên. Lớp này giúp mô hình hiểu rõ hơn về quan hệ giữa các bước thời gian và cải thiện độ chính xác của dự báo. Lớp Dense có 8 đơn vị, đóng vai trò đầu ra cuối cùng, dự đoán 8 giá trị cho chuỗi đầu ra (tương ứng với n_steps_out = 8). Lớp này tổng hợp thông tin từ các lớp LSTM trước đó để đưa ra dự báo. Mô hình được huấn luyện qua 200 epoch, tức là dữ liệu huấn luyện sẽ được lặp lại qua 200 lần. Con số này đảm bảo mô hình có đủ thời gian để học nhưng không quá lớn để tránh overfitting (học quá mức, làm giảm khả năng tổng quát hóa). Thuật toán Adam được sử dụng để tối ưu hóa quá trình huấn luyện. Adam là một trong những thuật toán tối ưu hóa hiệu quả nhất hiện nay, kết hợp giữa tốc độ học và

tính thích nghi, giúp mô hình học nhanh chóng mà vẫn giữ được độ chính xác cao.

Mô đun phần mềm dự báo nhiệt độ than tự cháy sử dụng mô hình LSTM được xây dựng với nhiều tính năng như: tải, hiển thị và phân tích dữ liệu nhiệt độ quá khứ; hiển thị kết quả dự báo nhiệt độ dưới dạng đồ thị và số; cung cấp các tùy chọn cho người dùng khi thực hiện dự báo;... Chức năng cung cấp tùy chọn cho người dùng khi thực hiện dự báo thể hiện tính linh hoạt của phần mềm, mỗi tùy chọn mang lại một ưu điểm khác nhau, tùy thuộc vào nhu cầu sử dụng. Tùy chọn thứ nhất là "Sử dụng mô hình đã huấn luyện". Đây là lựa chọn sử dụng mô hình LSTM đã được huấn luyện trước đó với dữ liệu quá khứ. Mô hình này đã qua quá trình tối ưu, phù hợp cho các tình huống mà người dùng không cần điều chỉnh thêm mô hình mà chỉ cần sử dụng ngay lập tức. Ưu điểm của cách này là tiết kiệm thời gian, vì không cần phải huấn luyện lại mô hình và dự báo có thể được thực hiện nhanh chóng (khoảng 30 giây). Tuy nhiên, nếu dữ liệu thực tế mới thay đổi so với dữ liệu đã được mô hình huấn luyện trước đó, độ chính xác của dự báo có thể không cao. Tùy chọn thứ hai là "Chỉnh mô hình". Tùy chọn này cho phép người dùng tinh chỉnh mô hình dự báo hiện tại dựa trên các thông số hoặc dữ liệu mới. Thời gian hiệu chỉnh mô hình khoảng 400 giây (với máy tính công nghiệp cấu hình core i7, bộ nhớ Ram 16GB). Việc hiệu chỉnh giúp mô hình thích ứng với các điều kiện thực tế hơn, đặc biệt khi có thêm dữ liệu mới hoặc khi hệ thống cần điều chỉnh các thông số của mô hình để đạt được độ chính xác cao hơn. Đây là lựa chọn phù hợp khi mô hình đã có độ chính xác tương đối, nhưng cần điều chỉnh thêm để phù hợp với các biến động hiện tại hoặc cải thiện tính dự báo. Tùy chọn cuối cùng là "Huấn luyện mô hình mới". Với tùy chọn này, mô hình sẽ được xây dựng từ đầu dựa trên dữ liệu mới nhất. Cách này



cho phép tạo ra một mô hình hoàn toàn mới, có khả năng phản ánh chính xác tình trạng hiện tại của hệ thống. Ưu điểm chính là tính cập nhật, giúp mô hình phản ánh được dữ liệu mới nhất. Nhược điểm là tốn nhiều thời gian và tài nguyên máy tính do quá trình huấn luyện bắt đầu từ đầu. Tuy nhiên, trong một số trường hợp, điều này là cần thiết để đảm bảo rằng mô hình dự báo có thể thích ứng với những thay đổi mới và đưa ra kết quả chính xác hơn. Phương pháp này thích hợp cho các tình huống mà mô hình cũ không còn hiệu quả hoặc không còn phản ánh đúng các xu hướng hiện tại. Trên cơ sở lý thuyết cũng như thực tế, mô hình dựa trên LSTM hay bất kỳ mô hình học máy nào khác, đều phụ thuộc vào dữ liệu huấn luyện. Nếu dữ liệu mới xuất hiện các mẫu hoặc xu hướng khác biệt với những gì mà mô hình cũ đã được huấn luyện thì mô hình cũ không nắm bắt được những sự thay đổi này, khi đó sẽ phải huấn luyện lại.

Việc cung cấp các tùy chọn giúp người dùng có thể linh hoạt trong việc sử dụng phần mềm dựa trên mục tiêu và điều kiện cụ thể. Tùy chọn “Sử dụng mô hình đã huấn luyện” phù hợp cho các trường hợp cần dự báo nhanh chóng mà không cần huấn luyện lại mô hình. “Chỉnh mô hình” cho phép tinh chỉnh thêm khi có dữ liệu mới hoặc cần cải thiện độ chính xác. “Huấn luyện mô hình mới” là lựa chọn cho các tình huống yêu cầu mô hình hoàn toàn mới. Mặc dù tạo mô hình mới nhưng không đòi hỏi người sử dụng có kiến thức chuyên sâu về công nghệ, các thao tác chỉ gồm lựa chọn, nhấn chuột và chờ đợi phần mềm thực hiện. Cách tiếp cận này rất hiệu quả trong việc đảm bảo tính linh hoạt và cập nhật của các hệ thống dự báo dựa trên trí tuệ nhân tạo.

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Tiến hành thử nghiệm

Mục tiêu của thử nghiệm là kiểm tra khả năng của mô hình LSTM đơn biến trong dự báo nhiệt độ của than tự cháy với các điều kiện giả lập, mô phỏng hiện tượng tự cháy mỏ hầm lò trong 8 giờ tiếp theo. Thử nghiệm này nhằm đánh giá tiềm năng ứng dụng của mô hình trong thực tế.

Các thiết bị và hệ thống được chuẩn bị gồm có: Hệ thống SCADA với nhiệm vụ thu thập và hiển thị dữ liệu thời gian thực về nhiệt độ của các vỉa than trong mỏ hầm lò; Mô hình học sâu LSTM đơn biến được sử dụng để phân tích và dự báo chuỗi dữ liệu nhiệt độ của than tự cháy; Tập dữ liệu đầu vào gồm

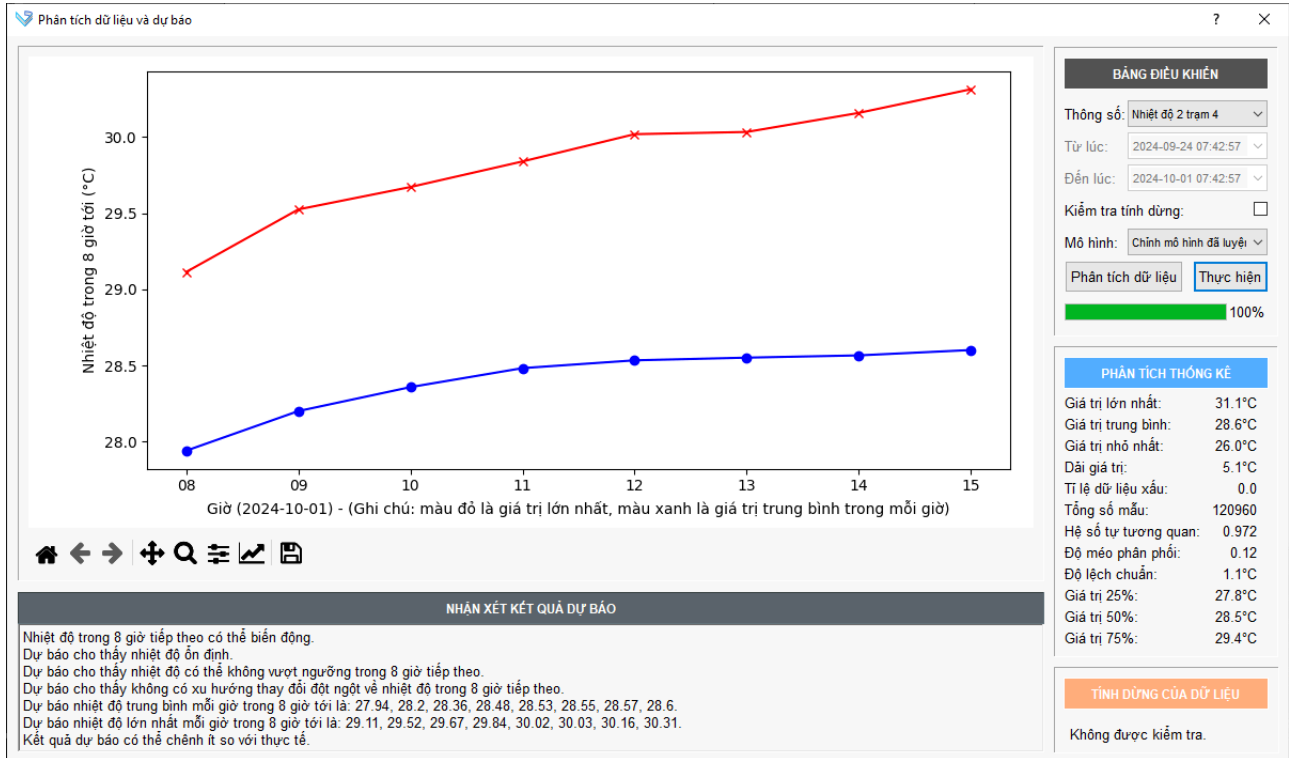
120.960 giá trị nhiệt độ giả lập của các vỉa than, thu thập với chu kỳ 5 giây/lần trong một tuần trước đó.

Thử nghiệm được tiến hành từng bước. Trước tiên, dữ liệu được tiền xử lý để đảm bảo tính chính xác và liên tục. Các giá trị được kiểm tra nhằm loại bỏ sự bất thường và cấu trúc thành chuỗi thời gian để đưa vào mô hình LSTM. Tiếp theo, mô hình LSTM đơn biến sẽ được huấn luyện trên tập dữ liệu để học các xu hướng và quy luật thời gian của hiện tượng tự cháy than. Cuối cùng, sau khi huấn luyện, mô hình thực hiện dự báo nhiệt độ của vỉa than trong 8 giờ tiếp theo, hiển thị và tự động nhận xét kết quả dự báo (mức độ biến động của nhiệt độ, khả năng có/không có cảnh báo nhiệt độ vượt ngưỡng,...) trên màn hình máy tính.

3.2. Các kết quả thử nghiệm và thảo luận

Đầu ra của mô hình là chuỗi giá trị nhiệt độ dự báo từng giờ trong 8 giờ tiếp theo, được trình bày trong Hình 3. Nội dung chính của hình thể hiện giao diện của mô đun phần mềm phân tích và dự báo nhiệt độ vỉa than. Phần chính giữa của giao diện là biểu đồ dự báo nhiệt độ theo giờ trong 8 giờ tới, đường bên trên thể hiện giá trị nhiệt độ lớn nhất trong mỗi giờ, đường bên dưới biểu thị giá trị trung bình nhiệt độ trong mỗi giờ. Các thông báo nhận xét kết quả dự báo (nhiệt độ có xu hướng ổn định, có khả năng vượt ngưỡng cảnh báo hay không,...) được trình bày phía bên dưới của giao diện, từ đó giúp người dùng nắm bắt và ứng phó kịp thời.

Các đặc trưng thống kê của sai số dự báo, được tính từ tập giá trị dự báo và tập giá trị nhiệt độ thực lấy mẫu trong phòng thí nghiệm của các lần thử nghiệm được trình bày trong Bảng 2. Nội dung của bảng trình bày về các lần thử nghiệm dự báo khác nhau (6 lần), với các đặc trưng sai số được đo lường và ghi nhận theo các thông số thống kê. Mỗi cột đại diện cho một lần dự báo cụ thể, với các thông số như trung bình, độ lệch chuẩn, giá trị nhỏ nhất, phần trăm phân vị và giá trị lớn nhất. Thông số “mean” thể hiện giá trị trung bình của sai số dự báo ở mỗi lần. Thông số “std” biểu thị mức độ phân tán của sai số. Các thông số “25%”, “75%” thể hiện các giá trị ở mốc 25% và 75% của phân bố sai số, giúp nhận định khoảng giữa của dữ liệu sai số, như 25% của lần 1 là -3.56, nghĩa là 25% giá trị sai số nằm dưới mức này. Thông số “min”, “max” là các giá trị sai số giữa giá trị nhiệt độ dự báo và giá trị nhiệt độ thực trong mỗi lần thử nghiệm.



Hình 3. Kết quả dự báo nhiệt độ vỉa than 8 giờ tiếp theo, từ 8h đến 15h

Bảng 2. Các đặc trưng thống kê của sai số dự báo

| Thông số | Lần 1 | Lần 2 | Lần 3 | Lần 4 | Lần 5 | Lần 6 |
|----------|--------------|-------|-------|-------|-------------|-------|
| mean | -2.59 | 0.36 | -0.04 | 0.08 | 0.32 | -0.11 |
| std | 1.92 | 2.09 | 1.68 | 1.65 | 1.45 | 1.29 |
| min | -4.50 | -1.06 | -4.45 | -4.27 | -2.63 | -3.01 |
| 25% | -3.56 | -0.83 | -1.09 | -0.93 | -0.65 | -0.98 |
| 75% | -1.63 | 1.08 | 0.03 | -0.02 | 0.43 | 0.81 |
| max | -0.66 | 2.77 | 1.06 | 1.14 | 3.20 | 2.36 |

Số liệu thử nghiệm cho thấy sai số dự báo dao động từ -4,50 °C đến 3,20 °C. Kết quả cho thấy sai số dự báo chấp nhận được, nằm trong giới hạn có thể áp dụng được trong thực tế, đảm bảo độ tin cậy khi sử dụng để quan trắc và cảnh báo sớm than tự cháy trong mỏ hầm lò, từ đó mở ra triển vọng áp dụng thực tế.

4. KẾT LUẬN

➤ Hệ thống quan trắc kết hợp với phương pháp dự báo nhiệt độ than tự cháy bằng kỹ thuật “học sâu” với mô hình mạng đơn biến “bộ nhớ dài - ngắn hạn” LSTM đóng vai trò quan trọng trong an toàn khai thác than hầm lò từ đó nâng cao hiệu quả sản xuất.

➤ Bằng cách áp dụng những tiến bộ khoa học công nghệ, hệ thống giúp các mỏ có thể chủ động dự báo và ứng phó với biến động của nhiệt độ than tự cháy, từ đó giảm thiểu nguy cơ tai nạn và tăng hiệu quả trong quá trình khai thác.

➤ Kết quả thử nghiệm đã cho thấy tính khả thi, hiệu quả của việc áp dụng những thành tựu của cuộc Cách mạng công nghiệp lần thứ tư với mô hình LSTM trong hệ thống quan trắc than tự cháy. Chất lượng dự báo sẽ tiếp tục được cải tiến để đạt được độ chính xác cao hơn trong giai đoạn áp dụng thực tế tại các mỏ than hầm lò □

**TÀI LIỆU THAM KHẢO**

- [1]. Nguyễn Xuân Đồng và nnk (2024), Nghiên cứu thiết kế xây dựng hệ thống quan trắc trực tuyến và ứng dụng kỹ thuật lô-gic mờ trong cảnh báo hiện tượng than tự cháy mỏ hầm lò, *Tạp chí Công nghiệp Mỏ*; Số 6-24.
- [2]. Viện Khoa học Công nghệ Mỏ (2014), Nghiên cứu đánh giá tính tự cháy của than và đề xuất các giải pháp kỹ thuật công nghệ phòng ngừa tự cháy ở các mỏ than hầm lò Việt Nam, *Báo cáo tổng kết đề tài*.
- [3]. Automate Technology Co., Ltd. (2013), Cảm biến đo nhiệt độ đa điểm cho vỉa than, *Xuzhou China Mining*.
- [4]. Ikram, B. A. O., Abdelhakim (2019), Deep Learning architecture for temperature forecasting in an IoT LoRa based system, *In Proceedings of the 2nd International Conference on Networking, Information Systems & Security*.
- [5]. H.B. Sahu (2013), Forecasting Spontaneous Heating Susceptibility of Indian Coals Using Neuro Fuzzy System, <https://link.springer.com/article/10.1007/s10706-013-9618-6>.
- [6]. Lim, B., & Zohren, S. (2020), Time Series Forecasting With Deep Learning: A Survey https://link.springer.com/chapter/10.1007/978-3-030-59338-4_19.
- [7]. Manjula C. Belavagi, Girija Attigeri (2022), Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison, *IEEE Access Volume 10*.
- [8]. Ran Vijay Kumar Singh (2019), Spontaneous Heating and Fire in Coal Mines, *The 9th Asia-Oceania Symposium on Fire Science and Technology*.
- [9]. Yu Zheng (2019), Monitoring the spontaneous combustion of coal stack based on ZigBee and LabView, *Journal of Physics Conference Series*.
- [10]. <https://towardsai.net/p/machine-learning/machine-learning-standardization-z-score-normalization-with-mathematics>

**RESEARCH ON APPLICATION OF DEEP LEARNING TECHNIQUE
WITH LONG SHORT TERM MEMORY MODEL IN THE MONITORING
AND FORECASTING SYSTEM OF COAL SPONTANEOUS COMBUSTION
TEMPERATURE IN UNDERGROUND MINES**

Vinh The Nguyen¹, Duong Thuy Le Nguyen^{1,*}, Dong Xuan Nguyen¹
Kien Hung Nguyen², Dung Danh Nguyen², Duong Hai Tran²

¹Vietnam Reseach Institute of Electronics, Informatics and Automation

²Electronics and Automation Technology Development Company Limited

ARTICLE INFOR

TYPE: Research Article

Received: 22/9/2024

Revised: 28/10/2024

Accepted: 02/11/2024

^{1,*} Corresponding author:

Email: duongthuynl@gmail.com

ABSTRACT

Online monitoring and forecasting of coal seam temperature in underground mines with spontaneous combustion is an urgent issue that is currently receiving attention. In this paper, a method of building a model to forecast the coal spontaneously combustion temperature every hour for the next eight hours using a univariate Long Short Term Memory (LSTM) model is proposed. The parameters of the model are adjusted through tests suitable for the given problem. The monitoring system combined with this forecasting method contributes to improving production efficiency, labor safety, environmental protection and effective use of Vietnam coal resources.

Keywords: deep learning, online monitoring system, coal spontaneous combustion, long short term memory.

@ Vietnam Mining Science and Technology Association