

SO SÁNH HIỆU QUẢ GIỮA THUẬT TOÁN HỒI QUY TUYẾN TÍNH VÀ PHI TUYẾN TÍNH KHI DỰ BÁO SÓNG CHẤN ĐỘNG NỔ MÌN TRÊN MỎ THAN NÚI BÉO, QUẢNG NINH

NGUYỄN HOÀNG, BÙI XUÂN NAM
 NGUYỄN TUẤN ANH, NGUYỄN TUẤN THÀNH
Trường Đại học Mỏ-Địa chất
 Email: nguyenhoang@hmg.edu.vn

1. Đặt vấn đề

Nổ mìn là một trong những phương pháp rẻ nhất và hiệu quả nhất trong việc phá vỡ đất đá cứng trên các mỏ lộ thiên [1], [2]. Tuy nhiên, không phải 100 % năng lượng của chất nổ đều tham gia vào việc phá vỡ đất đá. Một số nhà nghiên cứu đã kết luận rằng có tới 80÷85 % năng lượng của chất nổ là vô ích và sinh ra các tác dụng phụ như sóng chấn động (Peak Particle Velocity - PPV), sóng đập không khí, đất đá văng và hậu xung [3]-[9]. Trong số các tác dụng phụ đó thì sóng chấn động là một trong các yếu tố nguy hiểm sinh ra do quá trình nổ mìn trên các mỏ lộ thiên. Nó có thể làm rung động các công trình, nhà cửa, phá hủy các kết cấu công trình, gây mất ổn định tầng và bờ mỏ trên các mỏ lộ thiên,... Do vậy, việc đo đạc và dự báo sóng chấn động sinh ra do nổ mìn một cách chính xác là rất cần thiết để giảm thiểu các tác động tiêu cực của nổ mìn tới môi trường xung quanh.

Trong những năm gần đây, nhiều kỹ thuật và cách tiếp cận mới đã được phát triển để dự báo và giảm thiểu các tác động không mong muốn của sóng chấn động nổ mìn trên mỏ lộ thiên như: các phương pháp máy học, mạng nơ-ron nhân tạo, các thuật toán tiến hóa, hệ thống mạng nơ-ron mờ,... Longjun, Xibing [10] đã sử dụng thuật toán Random Forest (RF) và Support Vector Machine (SVM) để dự báo sóng chấn động nổ mìn với 93 vụ nổ được thu thập phục vụ cho nghiên cứu. Nghiên cứu này đã chỉ ra rằng RF và SVM là những thuật toán máy học phù hợp để dự báo sóng chấn động nổ mìn và SVM đã cung cấp một hiệu suất cao hơn so với RF trong dự báo PPV. Các thuật toán máy học khác như Support Vector Machine (SVM), hồi quy nhiều biến và các công thức thực nghiệm cũng

là cách tiếp cận của Hasanipanah, Faradonbeh [11] để dự báo PPV. Với mục đích này, họ đã thu thập 86 vụ nổ tại mỏ đồng Miduk của Iran để phục vụ nghiên cứu. Các kết quả đã chỉ ra rằng kỹ thuật SVM là một kỹ thuật tiên tiến cho hiệu suất cao hơn rất nhiều so với các kỹ thuật còn lại với sai số bình quân phương RMSE=0,17 và hệ số xác định $R^2=0,95$. Bằng cách tiếp cận khác, Chandar, Sastry [12] đã sử dụng các mô hình hồi quy và hệ thống mạng nơ-ron nhân tạo (ANN) để dự báo PPV với 168 vụ nổ đã được thu thập tại 3 mỏ đá vôi, dolomite và mỏ than. Các kết quả cho thấy mô hình ANN là mô hình tốt nhất được sử dụng trong nghiên cứu của họ với hệ số xác định $R^2=0,878$ cho tổng 3 mỏ. Ngoài ra, còn rất nhiều các nghiên cứu khác đã sử dụng các kỹ thuật tiên tiến và hiện đại để dự báo các tác động không mong muốn sinh ra do nổ mìn trên mỏ lộ thiên [13]-[17].

Tại Việt Nam, phương pháp phổ biến nhất hiện nay vẫn được sử dụng để xác định sóng chấn động nổ mìn là sử dụng các thiết bị đo chấn động như: Blastmate III, Minimate của Instel (Canada). Tuy nhiên các thiết bị này chỉ ghi lại được sóng chấn động khi các vụ nổ đã được tiến hành mà không có khả năng dự báo trước sóng chấn động sinh ra do quá trình nổ mìn. Các kỹ thuật tiên tiến và hiện đại như các phương pháp máy học, mạng nơ-ron nhân tạo, các thuật toán tiến hóa, hệ thống mạng nơ-ron mờ,... cũng chưa được áp dụng tại Việt Nam.

Trong nghiên cứu này, nhóm tác giả sử dụng các thuật toán hồi quy tuyến tính và hồi quy phi tuyến tính để so sánh và đánh giá khả năng ứng dụng các thuật toán máy học trong dự báo sóng chấn động nổ mìn trên mỏ lộ thiên tại Việt Nam. Linear Regression (LR) và Support Vector Machine

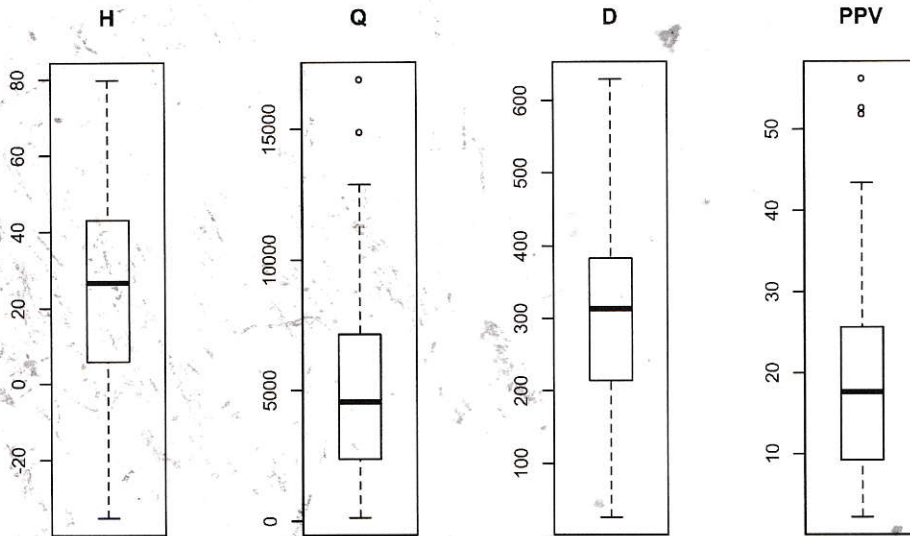
(SVM) là một trong những thuật toán hồi quy tuyến tính và phi tuyến tính được sử dụng cho mục đích này. Mỏ than Núi Béo được xem xét như là một khu vực nghiên cứu điển hình với 108 vụ nổ đã được ghi lại trong nhiều năm.

2. Thu thập dữ liệu

Để thực hiện nghiên cứu này, 108 vụ nổ đã được ghi lại từ các hộ chiếu nổ mìn của mỏ và thiết bị giám sát chấn động Blastmate III. Số liệu sử

dụng cho nghiên cứu bao gồm: sóng chấn động nổ mìn (PPV), độ chênh cao giữa bãi mìn và vị trí đặt máy đo (H), khối lượng thuốc nổ (Q) và khoảng cách đặt máy đo (R).

Các dữ liệu PPV được ghi lại bằng thiết bị Blastmate III của Instatel do Canada sản xuất. Các biến H, Q và R được xem như là các biến đầu vào và PPV là biến đầu ra trong nghiên cứu này. Dữ liệu sử dụng cho nghiên cứu được tóm tắt dưới dạng biểu đồ hộp trong H.1.



H.1. Tóm tắt dữ liệu sóng sử dụng cho nghiên cứu

H.1 tóm tắt các dữ liệu sử dụng trong nghiên cứu này với độ chênh cao giữa bãi mìn và vị trí đặt máy dao động trong khoảng từ -35,16 m đến +79,85 m; Khối lượng thuốc nổ sử dụng cho các vụ nổ dao động từ 131,93 kg đến 16905,69 kg; Khoảng cách đặt máy đo sóng chấn động (Blastmate III) từ 26,18 m đến 629,69m. Các kết quả sóng chấn động thu được nằm trong khoảng từ 2,26 mm/s đến 56,27 mm/s.

Bảng 2. Kết quả phân tích mối tương quan giữa các biến đầu vào

Tương quan	H	Q	D
H	1	0.08	-0.03
Q	0.08	1	-0.03
D	-0.03	-0.03	1

Quá trình sử dụng các thuật toán hồi quy, phân tích tương quan giữa các biến đầu vào là điều cần thiết để loại bỏ những biến có mối tương quan lớn (>0,75 hoặc <-0,75) gây ảnh hưởng tới mức độ chính xác của mô hình dự báo. Một Bảng

phân tích và tính toán các giá trị tương quan giữa các biến đầu vào được xác định trong Bảng 2. Các giá trị trong Bảng 2 cho thấy các biến đầu vào đều có mối tương quan thấp với nhau. Do vậy không cần loại bỏ các biến trong quá trình xây dựng các mô hình dự báo PPV.

3. Phương pháp nghiên cứu

Trong nghiên cứu này, phương pháp máy học với các thuật toán hồi quy tuyến tính (Linear Regression - LR) và hồi quy phi tuyến tính (Support Vector Machine - SVM) được sử dụng để xây dựng các mô hình dự báo sóng chấn động nổ mìn PPV. Bộ dữ liệu bao gồm 108 quan sát được chia thành 2 phần: bộ dữ liệu huấn luyện (92 quan sát) và bộ dữ liệu kiểm tra (16 quan sát).

Các thuật toán LR và SVM sẽ được lập trình trên máy tính điện tử để xây dựng các mô hình dự báo sóng chấn động sử dụng bộ dữ liệu huấn luyện bao gồm 92 quan sát để huấn luyện cho các mô hình dự báo. Bộ dữ liệu kiểm tra bao gồm 16 quan sát sẽ được sử dụng để kiểm tra lại mức độ chính xác của các mô hình dự báo LR và SVM đã xây dựng.

Để đánh giá hiệu quả của quá trình huấn luyện và kiểm tra, một số tiêu chuẩn đánh giá được sử dụng bao gồm: sai số bình phương bình quân phương (RMSE), hệ số xác định (R^2) và sai số tiêu chuẩn (MAE) được xác định theo các công thức (1)-(3):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (1)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

Sai số bình quân phương là một trong những đại lượng cơ bản và thường được sử dụng phổ biến cho việc đánh giá kết quả của mô hình dự báo số trị. Người ta thường hay sử dụng đại lượng sai số bình phương trung bình quân phương (RMSE) biểu thị độ lớn trung bình của sai số. RMSE không chỉ ra độ lệch giữa giá trị dự báo và giá trị thực tế. Giá trị của RMSE nằm trong khoảng $(0, +\infty)$. MAE cũng tương tự như RMSE nhưng các chỉ tiêu đánh giá đơn giản hơn nhiều. RMSE và MAE lý tưởng nhất khi bằng 0. R^2 là hệ số xác định thể hiện mức độ phù hợp của dữ liệu với thuật toán. R^2 nằm trong khoảng $(0,1)$ và lý tưởng nhất khi $R^2=1$ và bằng 0 là tệ nhất.

3.1. Tổng quan về thuật toán Linear Regression

Hồi quy tuyến tính có lẽ là một trong những thuật toán nổi tiếng nhất và được hiểu rõ nhất trong số liệu thống kê và máy học. Các thuật toán hồi quy tuyến tính là những thuật toán học có giám sát trong máy học và là những thuật toán đơn giản. Các thuật toán này nhằm mục đích tìm ra mối quan hệ giữa các đặc tính của một số vấn đề. Cụ thể hơn, từ một tập dữ liệu X, một mô hình Y (phương trình, đồ thị,...) sẽ được xây dựng phù hợp với tập dữ liệu, thể hiện xu hướng thay đổi và mối quan hệ giữa các đặc tính. Khi một mẫu dữ liệu mới được nhập, dựa trên mô hình tích hợp sẵn, có thể dự đoán giá trị của dữ liệu đó [18]-[20].

Nhìn chung, LR là thuật toán hồi quy tuyến tính được thực hiện với giả thiết rằng có một mối quan hệ tuyến tính giữa X và Y. Về mặt toán học, mối quan hệ tuyến tính có thể viết như sau:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (4)$$

Trong đó: y - Các giá trị đầu ra; β - Các hệ số mô hình (các giá trị này là "đã học" trong giai đoạn tập hợp hoặc đào tạo mô hình); β_0 - Hệ số chặn; β_1 - Hệ số cho x_1 (tính năng đầu tiên); β_n - Hệ số cho x_n .

3.2. Tổng quan về thuật toán Support Vector Machine (SVM)

Support Vector Machine (SVM) là một phương pháp học máy dựa trên lý thuyết thống kê và được phát triển vào những năm 1990 [21] và tiếp được phát triển bằng thuật toán có hiệu suất cao với một chút điều chỉnh. SVM có khả năng giải quyết các bài toán phân loại và hồi quy. Theo Cortes và Vapnik [21], SVM ban đầu chỉ được sử dụng cho các bài toán phân loại. Đến năm 1996, một phiên bản của SVM cho các bài toán hồi quy được đề xuất bởi Drucker, Burges [22] và chính thức được sử dụng vào năm 1997 với tên gọi Support Vector Regression (SVR) là một nhánh của thuật toán SVM.

Hàm SVM tổng quát được xác định theo công thức sau:

$$f(x) = w \cdot x + b = \sum_{i=1}^k (a_i - a_i^*) K(x, x_i) + b \quad (5)$$

Trong đó: a_i, a_i^* và b thu được thông qua việc giải quyết các vấn đề tối ưu được trình bày trong công thức (6) và (7). Thông thường, một phần nhỏ của a_i, a_i^* không phải là không, và được gọi là vector hỗ trợ.

$$w(a, a^*) = -0,5 \cdot \sum_{i,j=1}^k (a_i - a_i^*) \cdot (a_j - a_j^*) \cdot K(x_i, x_j) + \sum_{i=1}^k y_i (a_i - a_i^*) - \varepsilon \sum_{i=1}^k (a_i + a_i^*) \quad (6)$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^k (a_i - a_i^*) = 0 \\ 0 \leq a_i, a_i^* \leq C, (i=1,2,\dots,k) \end{cases} \quad (7)$$

Trong đó: C là hệ số bất lợi, thể hiện mức độ bất lợi đối với các mẫu lỗi quá mức ε ; $K(x_i, x_j)$ là hàm hạt nhân, giải quyết các vấn đề tính toán có kỹ năng cao.

Một trong các hàm hạt nhân của SVM được sử dụng trong nghiên cứu này là hàm Polynomial và được xác định theo công thức:

$$K(x, y) = [(x \cdot y) + 1]^d; \quad d = (1, 2, \dots) \quad (8)$$

Với: x, y - Lần lượt là các biến đầu vào và đầu ra của hàm; d - Hệ số Poly.

4. Kết quả nghiên cứu và thảo luận

Như đã đề cập trong phần 3, toàn bộ dữ liệu (108 vụ nổ) được chia thành 2 phần. Trong đó 85 % (92 vụ nổ) được sử dụng để xây dựng các mô hình dự báo PPV dựa trên các thuật toán hồi quy tuyến tính LR và thuật toán hồi quy phi tuyến tính SVM. Các thuật toán được lập trình với các tham số của thuật toán thay đổi để tìm ra mô hình dự báo tối ưu dựa trên các biến đầu vào. Trong

ngiên cứu này, các biến H, Q và D được xem như các biến đầu vào để dự báo giá trị đầu ra PPV. Đối với LR, tham số chặn đã được sử dụng để điều chỉnh hiệu suất của mô hình dự báo LR, trong khi D, σ và C là các tham số được sử dụng để điều chỉnh hiệu suất của mô hình dự báo SVM.

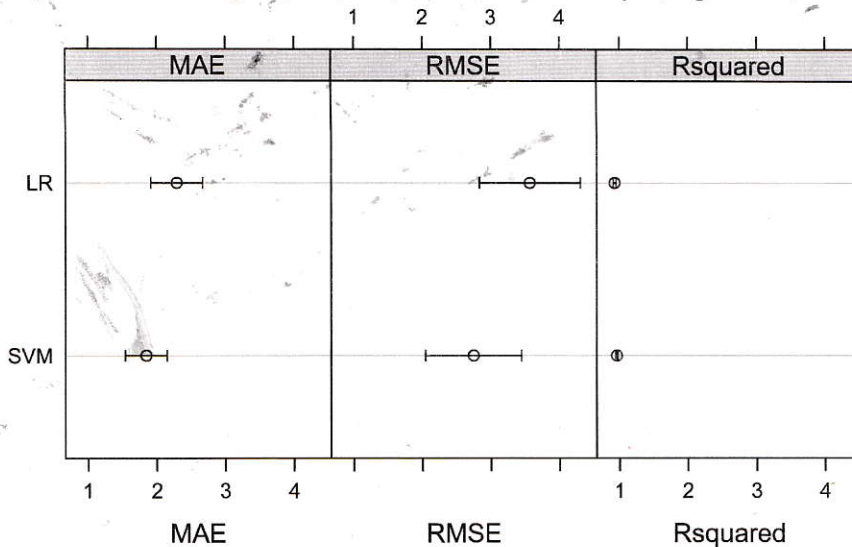
Quá trình thử nghiệm nhóm tác giả đã tìm ra với $D=3, \sigma=0,1$ và $C=1$ thì mô hình SVM đạt giá trị tối ưu. Các mô hình dự báo xây dựng trên bộ dữ liệu huấn luyện được đánh giá thông qua các chỉ số hiệu suất trong công thức (1)-(3) và được thể hiện trong Bảng 3.

Bảng 3. Các chỉ số hiệu suất của các mô hình dự báo trên tập dữ liệu huấn luyện

Phương pháp	MAE					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LR	0,858	1,574	2,191	2,287	2,863	4,640
SVM	0,837	1,441	1,681	1,844	2,032	3,977
Phương pháp	RMSE					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LR	1,100	2,548	3,097	3,562	4,009	8,802
SVM	0,928	1,710	2,424	2,742	2,688	8,018
Phương pháp	R^2					
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LR	0,765	0,903	0,943	0,925	0,966	0,997
SVM	0,856	0,946	0,970	0,957	0,978	0,993

Từ Bảng 3, có thể thấy cả 2 thuật toán LR và SVM hoạt động khá tốt khi dự báo PPV tại mỏ than Núi Béo. Tuy nhiên, nhìn vào các giá trị trung bình (Mean) của MAE, RMSE và R^2 trên Bảng 3 có thể thấy rằng mô hình dự báo SVM cung cấp một hiệu

suất cao hơn mô hình dự báo LR với $MAE=1,844$, $RMSE=2,742$ và $R^2=0,957$. Các kết quả so sánh hiệu suất của các mô hình được minh họa rõ hơn trong H.2 với mức độ tin cậy 95 % của các mô hình dự báo đã xây dựng trên bộ dữ liệu huấn luyện.



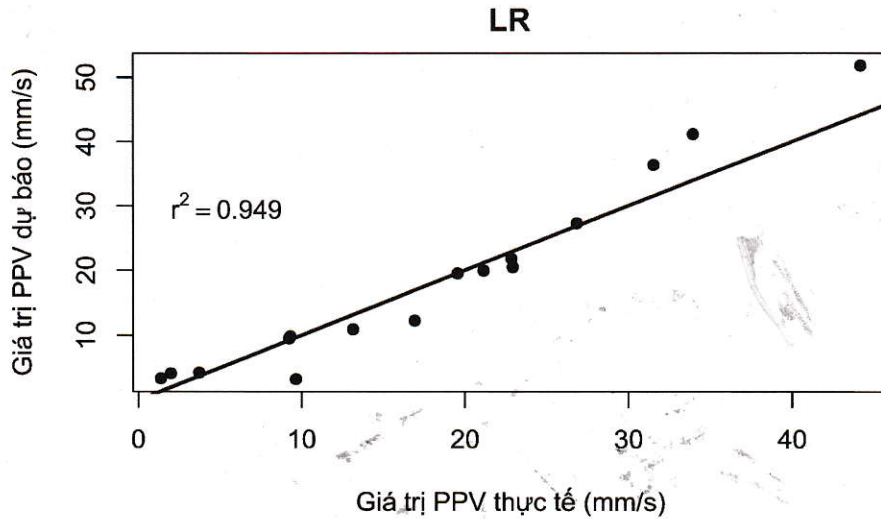
H.2. Mức độ tin cậy 95 % của các mô hình dự báo PPV cho mỏ than Núi Béo

Trong thực tế, các mô hình dự báo sau khi xây dựng dựa trên bộ dữ liệu huấn luyện cần phải kiểm tra lại bằng các giá trị đo thực tế và điều chỉnh mô hình dự báo (nếu cần). Do vậy, 16 quan sát còn lại (tương đương 15 % còn lại của dữ liệu) được sử dụng như là những dữ liệu chưa từng thấy cho dự báo PPV bằng mô hình đã xây dựng. Hiệu suất của các mô hình dự báo được đánh giá thông qua cả

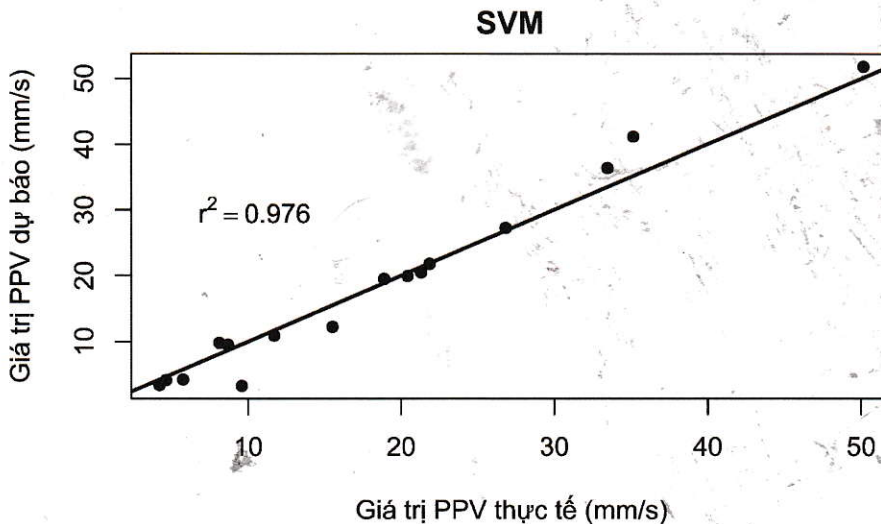
bộ dữ liệu huấn luyện và bộ dữ liệu kiểm tra. Các kết quả trong Bảng 4 cho thấy hiệu suất của các mô hình dự báo LR và SVM trên bộ dữ liệu thử nghiệm khá tốt. Một lần nữa, mô hình dự báo SVM lại cho kết quả dự báo tốt hơn mô hình LR với $MAE=1,809$, $RMSE=2,591$ và $R^2=0,976$ trên bộ dữ liệu kiểm tra. H. 3 và H.4 minh họa mức độ phù hợp của dữ liệu đối với các mô hình LR và SVM.

Bảng 4. Hiệu suất của các mô hình dự báo trên tập dữ liệu huấn luyện và tập kiểm tra

Phương pháp	Dữ liệu huấn luyện			Dữ liệu kiểm tra		
	MAE	RMSE	R ²	MAE	RMSE	R ²
LR	2,287	3,562	0,925	2,724	3,707	0,949
SVM	1,844	2,742	0,957	1,809	2,591	0,976



H.3. Mối quan hệ giữa các giá trị thực tế và giá trị dự báo PPV bằng thuật toán hồi quy tuyến tính LR



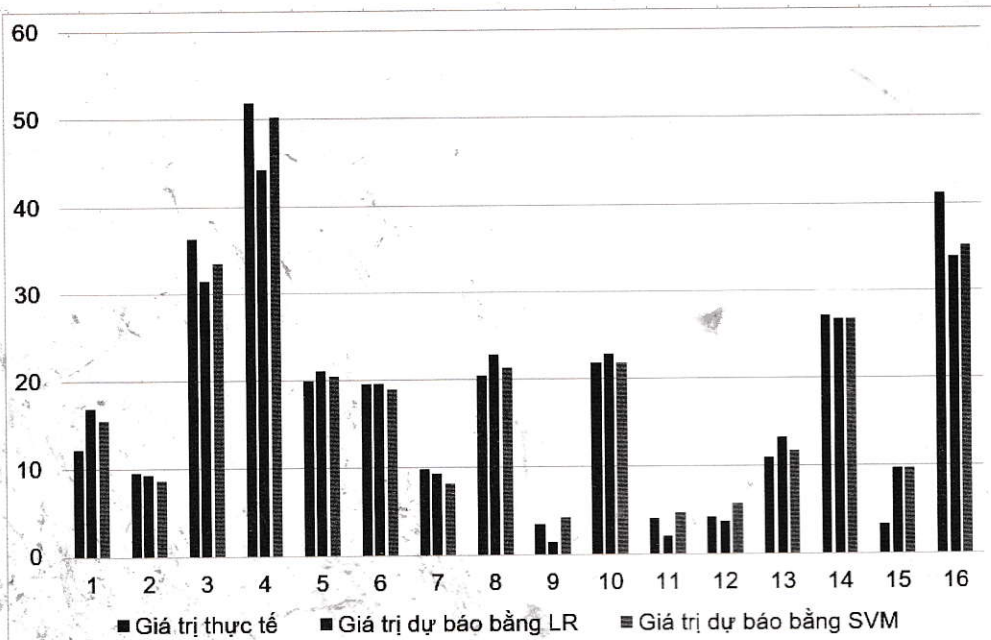
H.4. Mối quan hệ giữa các giá trị thực tế và giá trị dự báo PPV bằng thuật toán hồi quy phi tuyến tính SVM

Có thể thấy, trong trường hợp này thuật toán hồi quy phi tuyến tính hoạt động tốt hơn nhiều so với thuật toán hồi quy tuyến tính và SVM là thuật toán hồi quy phi tuyến tính điển hình có thể áp dụng để dự báo sóng chấn động nổ mìn cho mỏ than Núi Béo. Điều này cho thấy giữa các biến đầu vào không có mối quan hệ tuyến tính và

chúng là những biến độc lập, phù hợp với các kết quả trong Bảng 2. Việc dự báo sóng chấn động trong thực tế dựa trên những biến độc lập (không có mối quan hệ tuyến tính) nếu không sử dụng các thuật toán hồi quy phi tuyến tính (như SVM) là rất khó và cho độ chính xác không cao. H.5 minh họa mức độ chính xác của các mô hình

dự báo so với các giá trị đo thực tế trên bộ dữ liệu kiểm tra. Các kết quả cho thấy SVM cung

cấp một hiệu năng ưu việt hơn so với LR và rất gần với giá trị đo thực tế.



H.5. Mức độ chính xác của các mô hình LR và SVM trên tập dữ liệu kiểm tra

5. Kết luận

Trí thông minh nhân tạo là một phương pháp hiện đại với độ chính xác cao thông qua các thuật toán máy học để dự báo sóng chấn động nổ mìn trên mô lộ thiên dựa trên các số liệu được thống kê một cách chính xác. Nghiên cứu này đã chứng minh hiệu quả của các thuật toán hồi quy phi tuyến tính trong dự báo sóng chấn động nổ mìn bằng thuật toán SVM là xuất sắc với hiệu suất của mô hình đạt các giá trị MAE=1,809, RMSE=2,591 và R²=0,976. Các kết quả dự báo sóng chấn động nổ mìn cho mỏ than Núi Béo bằng thuật toán hồi quy phi tuyến tính SVM cho phép các kỹ sư và nhà quản lý mỏ dự đoán được trước các giá trị của sóng chấn động sinh ra và kiểm soát được các tác động không mong muốn tới môi trường xung quanh với sai số bình quân của các kết quả dự báo chỉ bằng 1,81 %. Trong nghiên cứu này, LR là một phương pháp đơn giản để dự báo PPV sinh ra do nổ mìn. Tuy nhiên mức độ chính xác không cao và cần được tiếp tục nghiên cứu để cải thiện mức độ chính xác của mô hình khi dự báo PPV. SVM là phương pháp cho hiệu quả dự báo PPV cao hơn, tuy nhiên mức độ phức tạp khi xây dựng mô hình SVM là lớn hơn. Đây là một cách tiếp cận mới sử dụng kỹ thuật SVM cho dự báo sóng chấn động nổ mìn trên mô lộ thiên ở Việt Nam và có thể ứng dụng trong thực tế tùy thuộc vào điều kiện cụ thể của từng mỏ. □

TÀI LIỆU THAM KHẢO

- Bhandari, S., Engineering rock blasting operations. A. A. Balkema. 388, 1997: p. 388.
- Hustrulid, W.A., Blasting principles for open pit mining: general design concepts. 1999: Balkema.
- Hasanipanah, M. và cộng sự. Feasibility of indirect determination of blast induced ground vibration based on support vector machine. Measurement, 2015. 75: p. 289-297.
- Hasanipanah, M., và cộng sự. Prediction of air-overpressure caused by mine blasting using a new hybrid PSO-SVR model. Engineering with Computers, 2017. 33(1): p. 23-31.
- Khandelwal, M. và T. Singh, Prediction of blast induced air overpressure in opencast mine. Noise & Vibration Worldwide, 2005. 36(2): p. 7-16.
- Khandelwal, M. và T. Singh, Prediction of blast-induced ground vibration using artificial neural network. International Journal of Rock Mechanics and Mining Sciences, 2009. 46(7): p. 1214-1222.
- Monjezi, M., H.A. Khoshalan, và A.Y. Varjani, Prediction of flyrock and backbreak in open pit blasting operation: a neuro-genetic approach. Arabian Journal of Geosciences, 2012. 5(3): p. 441-448.
- Monjezi, M., M. Hasanipanah, và M. Khandelwal, Evaluation and prediction of blast-induced ground vibration at Shur River Dam, Iran,

by artificial neural network. *Neural Computing and Applications*, 2013. 22(7-8): p. 1637-1643.

9. Verma, A. và T. Singh, Intelligent systems for ground vibration measurement: a comparative study. *Engineering with Computers*, 2011. 27(3): p. 225-233.

10. Longjun, D., và cộng sự Comparisons of random forest and support vector machine for predicting blasting vibration characteristic parameters. *Procedia Engineering*, 2011. 26: p. 1772-1781.

11. Hasanipanah, M., và cộng sự Forecasting blast-induced ground vibration developing a CART model. *Engineering with Computers*, 2017. 33(2): p. 307-316.

12. Chandar, K.R., V. Sastry, và C. Hegde, A Critical Comparison of Regression Models and Artificial Neural Networks to Predict Ground Vibrations. *Geotechnical and Geological Engineering*, 2017. 35(2): p. 573-583.

13. Faradonbeh, R.S. và M. Monjezi, Prediction and minimization of blast-induced ground vibration using two robust meta-heuristic algorithms. *Engineering with Computers*, 2017: p. 1-17.

14. Faradonbeh, R.S., và cộng sự Prediction of ground vibration due to quarry blasting based on gene expression programming: a new model for peak particle velocity prediction. *International Journal of Environmental Science and Technology*, 2016. 13(6): p. 1453-1464.

15. Taheri, K., và cộng sự A hybrid artificial bee colony algorithm-artificial neural network for forecasting the blast-produced ground vibration. *Engineering with Computers*, 2017. 33(3): p. 689-700.

16. Sheykhi, H., và cộng sự Forecasting ground vibration due to rock blasting: a hybrid intelligent approach using support vector regression and fuzzy C-means clustering. *Engineering with Computers*, 2017: p. 1-9.

17. Hasanipanah, M., và cộng sự Estimation of blast-induced ground vibration through a soft computing framework. *Engineering with Computers*, 2017: p. 1-9.

18. Weisberg, S., *Applied linear regression*. Vol. 528. 2005: John Wiley & Sons.

19. Seber, G.A. và A.J. Lee, *Linear regression analysis*. Vol. 936. 2012: John Wiley & Sons.

20. Montgomery, D.C., E.A. Peck, và G.G. Vining, *Introduction to linear regression analysis*. Vol. 821. 2012: John Wiley & Sons.

21. Cortes, C. và V. Vapnik, Support-vector networks. *Machine learning*, 1995. 20(3): p. 273-297.

22. Drucker, H., và cộng sự. Support vector regression machines. in *Advances in neural information processing systems*. 1997.

Ngày nhận bài: 15/10/2017

Ngày gửi phản biện: 16/11/2017

Ngày nhận phản biện: 20/02/2018

Ngày chấp nhận đăng bài: 10/04/2018

Từ khóa: *sóng chấn động; thuật toán hồi quy; Linear Regression; Support Vector Machine; máy học*

SUMMARY

Ground vibration (PPV) is one of the undesirable effects induced by blasting operations in open-pit mine, and it can cause damage to surrounding structures. Therefore, predicting ground vibration is essentially necessary to reduce the environmental effects of blasting. In this paper, the authors used two regression algorithms to predict PPV in Nui Beo open-pit coal mine, including Linear Regression (LR) and Support Vector Machine (SVM).

LR represented for linear regression algorithms, whereas SVM represented for non-linear regression algorithms. To employ this study, 108 blasting events were recorded with four parameters, including ground vibration (PPV), elevation between blast sites and monitor device (H), the capacity of explosive (Q), and the distance between blast sites and monitor device (D).

The indicators for evaluating the performance of the predictive models were used including Root Mean Square Error (RMSE), Coefficient of determination (R^2), and Mean Absolute Error (MAE). The results indicated that LR and SVM are suitable for predicting PPV in Nui Beo open-pit coal mine and SVM technique provides higher performance than LR technique with MAE=1.801, RMSE=2.591, and $R^2=0.976$. In addition, the results are also the basis of development of other predictive models in blasting operations of open-pit mine in Vietnam.