

Ứng dụng thuật toán XGBoost và tối ưu hóa bầy đàn dự báo tính chất cơ lý của bê tông cốt liệu tái chế

Application of XGBoost algorithm and particle swarm optimization in predicting mechanical properties of recycled aggregate concrete

Lê Bá Anh*

Trường Đại học Giao thông vận tải

*Email liên hệ: baanh.le@utc.edu.vn

Tóm tắt:

Bài báo xây dựng cơ sở dữ liệu về bê tông cốt liệu tái chế được tập hợp từ các nghiên cứu về thực nghiệm đã được công bố trên thế giới. Trong bài báo, nhóm tác giả đã sử dụng giải thuật eXtreme Gradient Boosting (XGB) kết hợp với phương pháp tối ưu hóa bầy đàn để xây dựng mô hình dự báo cường độ chịu nén và độ bền kéo đứt của bê tông sử dụng cốt liệu tái chế. Mô hình đạt được độ chính xác khá cao là 0,93.

Từ khóa: Machine learning; XGB; Bê tông cốt liệu tái chế; Tối ưu hóa bầy đàn.

Abstract:

The article builds a dataset of recycled aggregate concrete gathered from experimental studies published. In the article, the authors used the eXtreme Gradient Boosting (XGB) algorithm combined with the particle swarm optimization method to build a model to predict the ultimate compressive strength and split tensile strength of concrete using recycled aggregate. This model achieves quite high accuracy with an accuracy of 0.93.

Keywords: Machine learning; XGB; Recycled aggregated concrete; Particle swarm optimization.

1. Giới thiệu

Quá trình đô thị hóa nhanh chóng ở các nước đang phát triển cùng các thiên tai bất ngờ là nguyên nhân của nhiều công trình xây dựng bị phá hủy và sinh ra các chất thải xây dựng. Hàng tỷ tấn bê tông bị phá hủy mỗi năm trên khắp thế giới và chỉ có rất ít phần trong đó được tái chế. Điều này đòi hỏi một yêu cầu rất lớn về xử lý các chất thải xây dựng nhằm tránh các tác động kinh tế và môi trường [1]. Do đó, tái chế bê tông là một nhu cầu quan trọng, cũng giúp bảo tồn cốt liệu tự nhiên như sỏi cát từ đó thúc đẩy xã hội phát triển bền vững. Về mặt kinh tế, tái chế cốt liệu giúp giảm chi phí xây dựng vì vật liệu thu được từ các công trình cũ bị phá hủy có thể được tái sử dụng cho công trình mới tại chỗ [2].

Tuy nhiên, ngày nay vật liệu tái chế chủ yếu được sử dụng cho các cấu kiện không quan trọng như vật liệu lấp đầy hoặc lớp nền phụ trong xây dựng đường cao tốc. Một trong những nguyên nhân chủ yếu là do sự thiếu hiểu biết về các tính chất cơ học của bê tông cốt liệu tái chế (Recycled Aggregated Concrete-RAC).

Cốt liệu thu được bằng cách phá dỡ bê tông được gọi là cốt liệu tái chế. Việc sản xuất cốt liệu tái chế từ các mảnh vụn xây dựng làm tăng khả năng nhiễm bẩn từ cốt thép, dầu và các thành phần xây dựng khác. Sự ô nhiễm như vậy có thể gây suy giảm khả năng làm việc của cốt liệu tái chế trong bê tông. Ở nhiều nước châu Âu, việc sử dụng 20% cốt liệu tái chế trong kết cấu bê tông là thông lệ tiêu chuẩn.

Cường độ và mô đun đàn hồi của RAC thấp hơn so với bê tông thông thường [3], [4]. Từ biến, co ngót và biến dạng tối đa của RAC tăng khi tỉ lệ thay thế cốt liệu tái chế tăng. Đặc tính thấm của RAC cao hơn so với bê tông thông thường do khí và chất lỏng có thể đi qua các cốt liệu tái chế. Việc thiết kế thành phần cho RAC cần tối ưu hóa để tăng cường khả năng làm việc của chúng.

Trong thập kỷ qua, các phương pháp học máy (Machine Learning - ML) đã được phát triển để dự đoán cường độ nén và mô đun đàn hồi của RAC và để tối ưu hóa thiết kế tỷ lệ hỗn hợp [5]-[8]. Dữ liệu cần thiết để đào tạo và thử nghiệm mô hình ML có thể được thu thập từ một lượng lớn các công trình nghiên cứu về thực nghiệm đã được công bố. Những nghiên cứu này xem xét các đặc điểm như hàm lượng nước, hàm lượng xi măng và hàm lượng cốt liệu tái chế để dự đoán cường độ nén và mô đun Young của RAC. Ví dụ, Duan và cộng sự [6] đã sử dụng phương pháp mạng neuron nhân tạo (Artificial Neural Network-ANN) để dự đoán các tính chất cơ học của RAC. Phương pháp này được khẳng định là hoạt động tốt hơn so với phương pháp hồi quy phi tuyến tính [7]. Trong nghiên cứu này, nhóm tác giả dự báo tính chất cơ lý của bê tông cốt liệu tái chế bằng mô hình được xây dựng từ giải thuật eXtreme Gradient Boosting (XGB) [9]. Trong những năm gần đây, giải thuật XGB đã chứng tỏ được tính ưu việt của mình khi chiến thắng hầu hết các cuộc thi về máy học, nhất là xử lý các dữ liệu dạng tường minh như dữ liệu của bê tông cốt liệu tái chế. Mô hình được xây dựng từ giải thuật XGB có những ưu điểm như hạn chế được khả năng overfitting nhờ cấu trúc cây của nó. Kết hợp với quá trình tinh chỉnh các thông số nhờ phương pháp tối ưu hóa bầy đàn (Particle Swarm Optimization-PSO), mô hình PSO-XGB có thể đạt được độ chính xác cao.

2. Cơ sở lý thuyết và các nghiên cứu liên quan

2.1. Giải thuật eXtreme Gradient Boosting

Chen và Guestrin [9] đã giới thiệu một thuật toán được phát triển dựa trên mô hình cây quyết định

là giải thuật XGB. Giải thuật này dự đoán kết quả dựa trên các dữ liệu đầu vào như sau:

$$\hat{y}_i = y_i^0 + \eta \sum_{k=1}^M f_k(X_i) \quad (1)$$

Trong đó \hat{y}_i là kết quả dự báo cho mẫu thứ i có vectơ đặc trưng là X_i ; M là số lượng bộ ước tính và mỗi bộ ước tính f_k (với k trong phạm vi từ 1 đến M) tương ứng với một cấu trúc cây độc lập; y_i^0 là giá trị trung bình ban đầu của giá trị đo được trong tập huấn luyện; η là tốc độ học giúp cải thiện mô hình trong khi thêm cây mới và tránh khớp quá mức (overfitting).

Hàm mục tiêu để giảm thiểu bước thứ k được xác định bởi:

$$obj = \gamma T + \sum_{j=1}^T \left[G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] \quad (2)$$

Trong đó T là số lá cây thứ k và ω_j với j từ 1 đến T là trọng lượng của lá; λ và γ là các tham số chuẩn hóa kiểm soát tính đơn giản của cấu trúc cây để tránh khớp quá mức. Các tham số G_j và H_j lần lượt là tổng của các mẫu được liên kết với lá thứ j của độ dốc thứ nhất và thứ hai của hàm mất mát. Cây thứ k được xây dựng bằng cách tách các lá bắt đầu từ một lá. Thay đổi mục tiêu sau khi thêm phần tách là:

$$gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (3)$$

Trong đó G_L và H_L được liên kết với lá bên trái, G_R và H_R được liên kết với lá bên phải sau khi chia tách. Nếu tham số khuếch đại lớn hơn 0, việc tách được chấp nhận. Độ chính xác của mô hình được xây dựng bởi giải thuật XGB phụ thuộc rất nhiều vào các tham số của mô hình huấn luyện. Các tham số quan trọng của mô hình cần được hiệu chỉnh phù hợp lần lượt như sau: Tỉ lệ mẫu con (ζ), suy giảm mất mát tối thiểu (γ), chiều sâu tối đa (d), shrinkage (η), tổng nhỏ nhất của tập con (μ) và tỉ lệ cột theo từng cây con (δ). Để hiệu chỉnh các thông số này, nhóm nghiên cứu đã ứng dụng phương pháp PSO.

2.2. Phương pháp tối ưu hóa bầy đàn (PSO)

Lấy cảm hứng từ hành vi bầy đàn của động vật trong thiên nhiên như chim, ong, cá, ... Kennedy

và cộng sự [10] phát triển phương pháp PSO dựa trên tối ưu hóa ngẫu nhiên. Mỗi hạt di chuyển trong không gian tìm kiếm sẽ liên tục cập nhật thông tin và vị trí của mình cho quần thể, trao đổi này giúp chúng xác định hạt có vị trí tốt nhất. Từ đó, các hạt khác sẽ dựa vào, điều chỉnh hướng bay cũng như tốc độ bay để hướng đến mục tiêu chung của quần thể là lựa chọn được nguồn thức ăn tốt nhất. Vị trí mỗi phân tử qua mỗi bước lặp được cập nhật như sau:

$$x_{t+1}^i = x_t^i + v_{t+1}^i \quad (4)$$

Phương trình dưới đây cập nhật vận tốc của các phân tử:

$$v_{t+1}^i = \omega v_t^i + c_1 r_1 (P_{t(best)}^i - x_t^i) + c_2 r_2 (g_{t(best)} - x_t^i) \quad (5)$$

Ở đây x_t^i và x_{t+1}^i ; v_t^i và v_{t+1}^i lần lượt đại diện cho vị trí và vận tốc của phân tử thứ i tại bước lặp thứ t và $t+1$, tương ứng $P_{t(best)}^i$ là giải pháp tối ưu tốt nhất của hạt thứ i , đây chính là giải pháp tối ưu cục bộ và $g_{t(best)}$ là giải pháp tốt nhất toàn cục của tất cả các phân tử tại thời điểm t ; r_1 và r_2 là các số ngẫu nhiên có giá trị từ 0 đến 1 để xác định

bán kính bay của các hạt sau khi có được hướng bay từ việc chia sẻ thông tin của quần thể với nhau; c_1 và c_2 là các hệ số chia sẻ thông tin của quần thể cục bộ và toàn cục; w là trọng số giúp điều chỉnh bước nhảy của toàn bộ quần thể.

3. Xây dựng cơ sở dữ liệu

Bộ dữ liệu bao gồm 137 mẫu thí nghiệm được thu thập từ 17 nghiên cứu thực nghiệm đã được công bố [11]-[27]. Mỗi mẫu thí nghiệm bao gồm 9 thông số đầu vào: Nước (Water), xi măng(Cement), cát (Sand), cốt liệu thô tự nhiên (NCA), cốt liệu thô tái chế (RCA), chất siêu dẻo (SP), đường kính lớn nhất của cốt liệu tái chế (Dmax_RCA), khối lượng riêng của cốt liệu tái chế (ρ RCA) và lượng nước hấp thụ cốt liệu tái chế (WRCA). Các tham số đầu ra đã được xem xét là cường độ chịu nén (UCS), độ bền kéo đứt (STS). Bảng 1 trình bày chi tiết ký hiệu, vai trò và phân tích thống kê (giá trị tối thiểu, tối đa, trung bình, trung vị và độ lệch chuẩn) của các tham số đầu vào cũng như tham số đầu ra.

Bảng 1. Thống kê xác suất của tập dữ liệu.

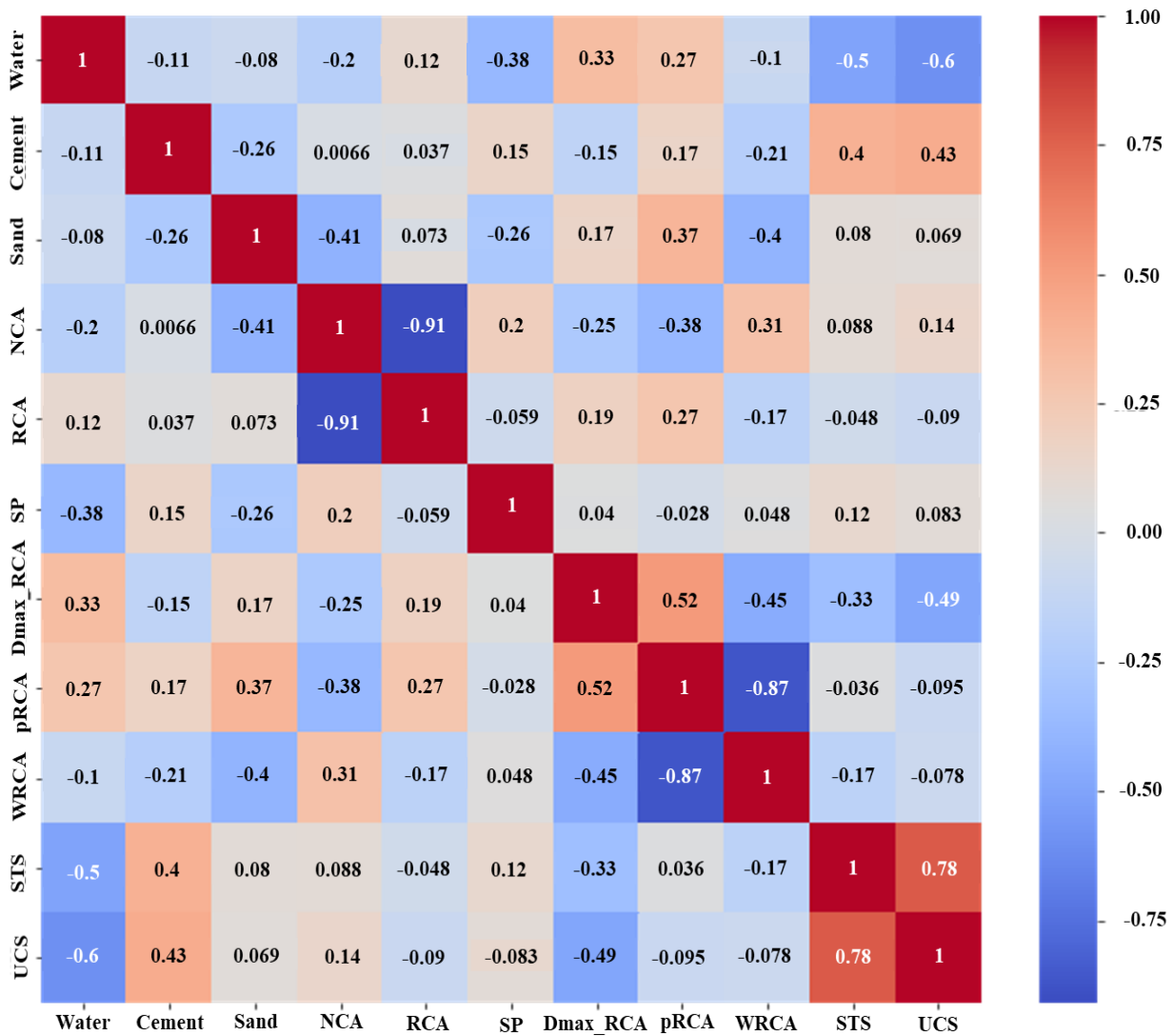
	Số mẫu	Trung bình	Std	min	max
Water	137	181.02	19.26	137	225
Cement	137	354.50	62.63	158	600
Sand	137	738.89	180.84	0	1010
NCA	137	411.57	404.54	0	1168
RCA	137	608.55	333.45	57	1123
SP	137	1.11	1.92	0	7.8
Dmax_RCA	137	18.30	4.06	10	25
ρ RCA	137	2374	151.25	2010	2661
WRCA	137	5.53	2.15	1.9	10.9
STS	137	3.35	0.95	1.4	6.3
UCS	137	49.78	19.02	16.5	108.5

Tập dữ liệu được chia thành hai tập con ngẫu nhiên bao gồm: Tập huấn luyện 70% và tập kiểm chứng 30%. Các dữ liệu đầu vào và đầu ra được chuẩn hóa trong phạm vi [0,1] theo khuyến nghị của Witten và cộng sự [28]. Mối tương quan giữa các thông số của tập dữ liệu được biểu diễn thông qua biểu đồ nhiệt ở hình 1. Các hệ số tương quan nằm trong khoảng [-1;1].

4.Xây dựng mô hình

4.1. Đánh giá độ chính xác

Để đánh giá độ chính xác của mô hình, nhóm nghiên cứu sử dụng ba thông số thống kê lần lượt là R-squared, Mean Absolute Error (MAE) và Root Mean Squared Error (RMSE).



Hình 1. Biểu đồ nhiệt thể hiện mối tương quan giữa các thông số của dữ liệu.

$$R-squared = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{7}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{8}$$

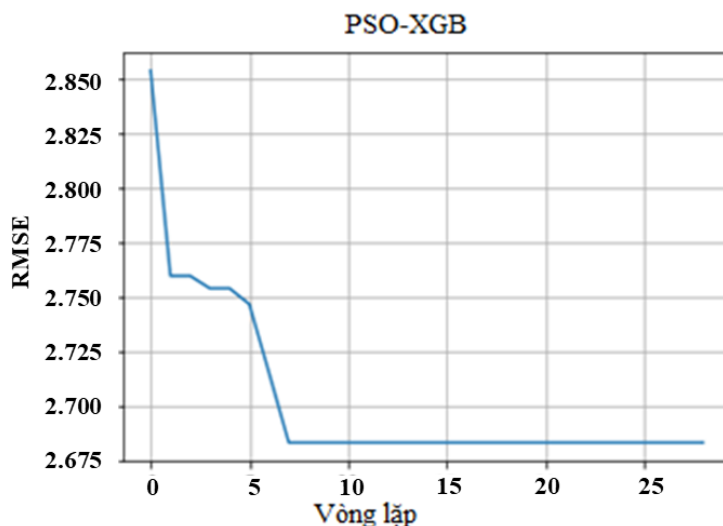
Trong đó $y_i, \hat{y}_i, \bar{y}_i$ lần lượt là giá trị chính xác, giá trị dự báo và giá trị trung bình của các dữ liệu cần đánh giá.

4.2. Mô hình PSO-XGB

Sau khi xây dựng mô hình XGB phụ thuộc vào các tham số: Tỷ lệ mẫu con (ζ), suy giảm mất mát

tối thiểu (γ), chiều sâu tối đa (d), shrinkage (η), tổng nhỏ nhất của tập con (μ) và tỉ lệ cột theo từng cây con (δ). Nhóm nghiên cứu đã sử dụng phương pháp PSO để tiến hành hiệu chỉnh nhằm lựa chọn bộ tham số tối ưu. Bộ tham số giúp mô

hình XGB đạt được độ chính xác R-squared lớn nhất là $\zeta = 0,577$; $\gamma = 3,745$; $d = 8$; $\eta = 0,54$; $\mu = 3,218$; $\delta = 0,894$. Hình 2 thể hiện quá trình tối ưu hóa của mô hình PSO-XGB với biểu đồ của độ chính xác phụ thuộc vào số vòng lặp.



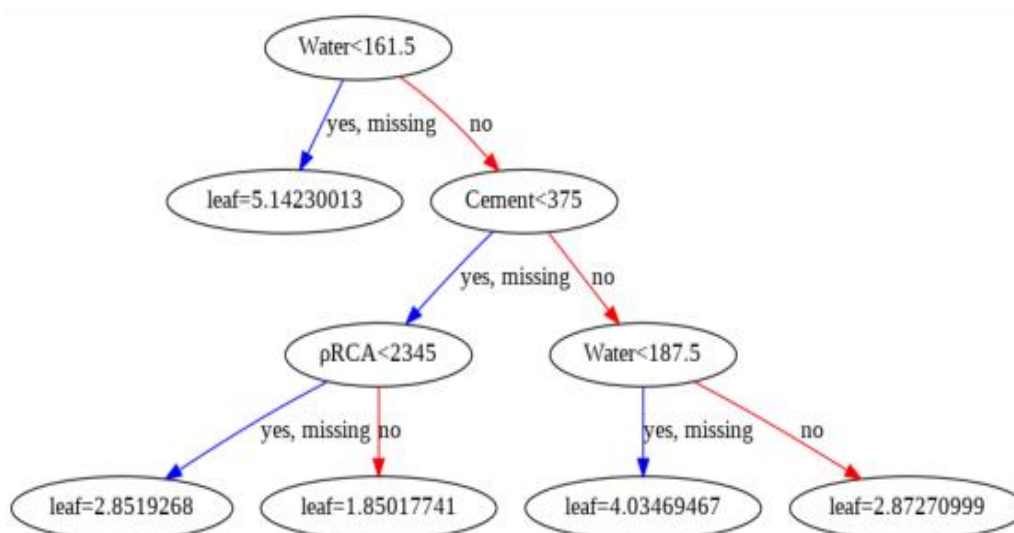
Hình 2. Quá trình tối ưu hóa của mô hình PSO-XGB.

5. Kết quả và nhận xét

Kết quả thể hiện độ chính xác của mô hình được trình bày ở bảng 2 thông qua ba thông số R-squared, MAE và RMSE của tập dữ liệu huấn luyện và tập dữ liệu kiểm chứng. Đồng thời Hình 4 và Hình 5 cũng thể hiện sự tương quan giữa giá trị dự báo và giá trị chính xác. Mô hình XGB thu được được thể hiện ở hình 3.

Bảng 2. Kết quả độ chính xác của mô hình.

Kết quả	Tập dữ liệu	R-squared	MAE	RMSE
UCS	Train	0.97	1.92	3.24
	Test	0.93	3.82	5.42
STS	Train	0.90	0.25	0.28
	Test	0.87	0.27	0.35



Hình 3. Mô hình XGB.

Đối với giá trị dự báo cường độ chịu nén UCS, độ chính xác của mô hình là R-squared = 0.93 với

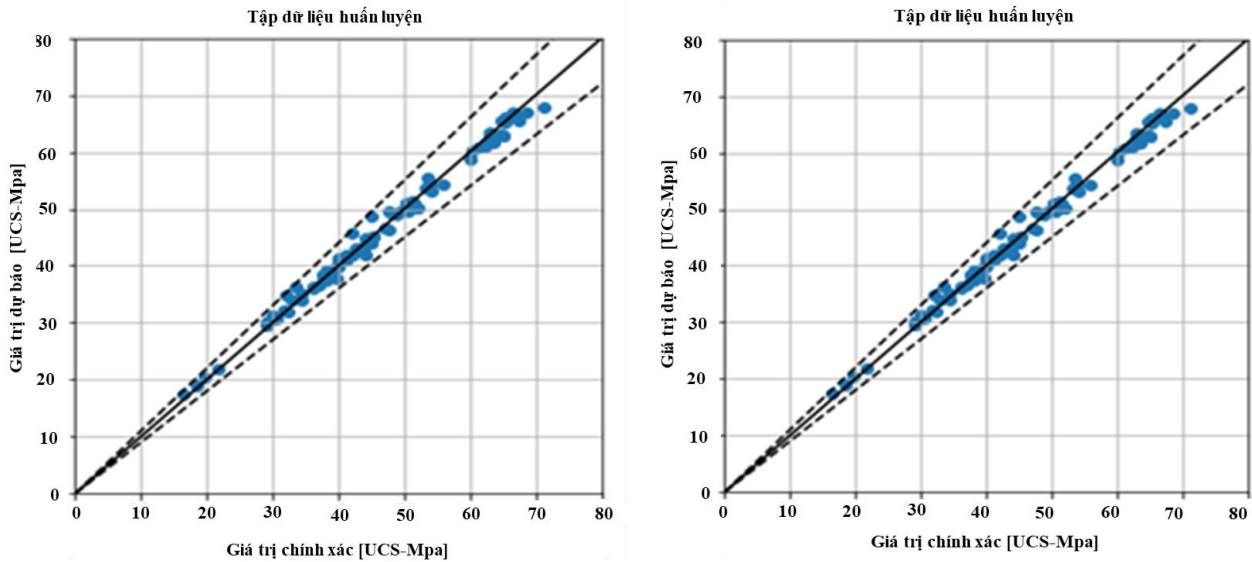
sai số MAE là 3.82 MPa và RMSE là 5.42 MPa. Các thông số tương ứng đối với độ bền kéo đứt

STS là $R\text{-squared} = 0.87$, $MAE = 0.27$ MPa và $RMSE = 0.35$ MPa. Với các kết quả thu được, có thể thấy rằng mô hình PSO-XGB là một công cụ dự báo tốt.

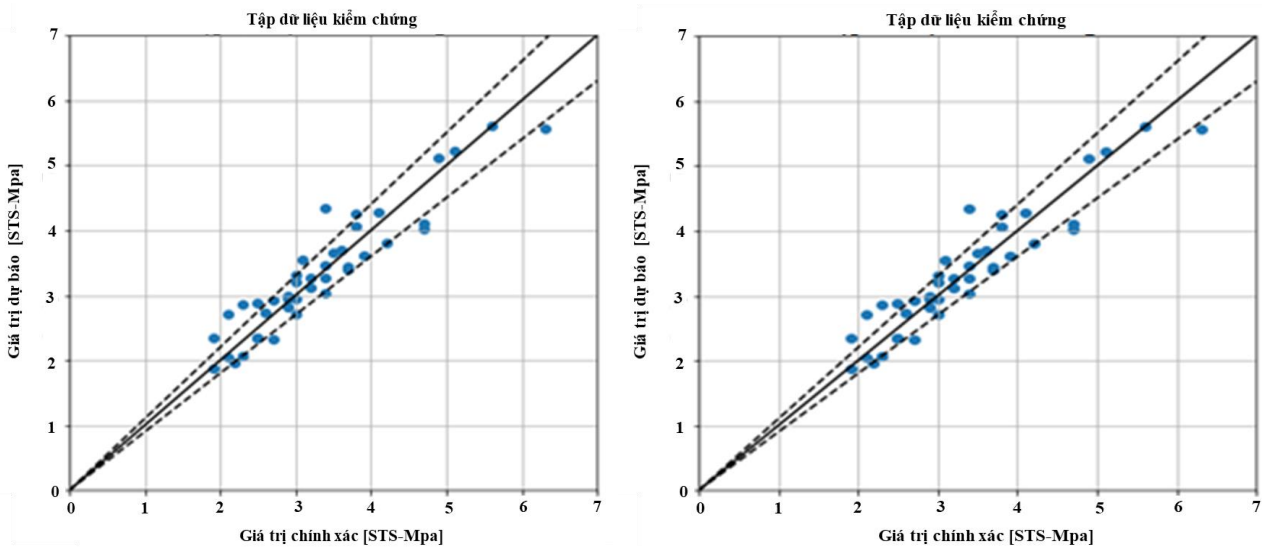
6. Kết luận

Bài báo xây dựng cơ sở dữ liệu về bê tông cốt liệu tái chế được tập hợp từ các nghiên cứu về

thực nghiệm đã được công bố trên thế giới. Đồng thời, với sự trợ giúp của giải thuật XGB kết hợp với PSO để xây dựng được mô hình dự báo cường độ chịu nén và độ bền kéo đứt của bê tông sử dụng cốt liệu tái chế. Mô hình này đạt được độ chính xác khá cao với độ chính xác $R\text{-squared}$ là 0,93 đối với UCS và 0.87 với STS.



Hình 4. Kết quả dự đoán và kết quả chính xác của UCS của tập dữ liệu huấn luyện và tập dữ liệu kiểm chứng (đường nét liền thể hiện giá trị trùng nhau, đường nét đứt thể hiện sai số $\pm 10\%$).



Hình 5. Kết quả dự đoán và kết quả chính xác của STS của tập dữ liệu huấn luyện và tập dữ liệu kiểm chứng (đường nét liền thể hiện giá trị trùng nhau, đường nét đứt thể hiện sai số $\pm 10\%$).

Lời cảm ơn

Nghiên cứu này được tài trợ bởi trường Đại học Giao thông vận tải, trong đề tài mã số T2022-CT-005.

Tài liệu tham khảo

[1] T. C. Hansen; “Recycled aggregates and recycled aggregate concrete second state-of-the-art report developments 1945–1985”. Mater. Struct. 1986; 19(3):201–246. DOI:10.1007/BF02472036.

- [2] H. Zhang, Y. Zhao, T. Meng, and S. P. Shah; “Surface Treatment on Recycled Coarse Aggregates with Nanomaterials”. *J. Mater. Civ. Eng.* 2016; 28(2). DOI:10.1061/(ASCE)MT.1943-5533.0001368.
- [3] J. Xiao, W. Li, Y. Fan, and X. Huang; “An overview of study on recycled aggregate concrete in China (1996–2011)”. *Constr. Build. Mater.* 2012; 31:364–383. DOI:10.1016/j.conbuildmat.2011.12.074.
- [4] M. Etxeberria, E. Vázquez, A. Marí, and M. Barra; “Influence of amount of recycled coarse aggregates and production process on properties of recycled aggregate concrete”. *Cem. Concr. Res.* 2007; 37(5):735–742. DOI:10.1016/j.cemconres.2007.02.002.
- [5] Y. Yang et al.; “Concrete Mix Design for Completely Recycled Fine Aggregate by Modified Packing Density Method”. *Materials.* 2020; 13(16). DOI:10.3390/ma13163535.
- [6] Z.H. Duan, S. C. Kou, and C. S. Poon; “Prediction of compressive strength of recycled aggregate concrete using artificial neural networks”. *Constr. Build. Mater.* 2013; 40:1200–1206. DOI:10.1016/j.conbuildmat.2012.04.063.
- [7] N. Deshpande, S. Londhe, and S. Kulkarni; “Modeling compressive strength of recycled aggregate concrete by Artificial Neural Network, Model Tree and Non-linear Regression”. *Int. J. Sustain. Built Environ.* 2014; 3(2):187–198. DOI:10.1016/j.ijbsbe.2014.12.002.
- [8] A. Gholampour, A. H. Gandomi, and T. Ozbakkaloglu; “New formulations for mechanical properties of recycled aggregate concrete using gene expression programming”. *Constr. Build. Mater.* 2017; 130:122–145. DOI:10.1016/j.conbuildmat.2016.10.114.
- [9] T. Chen, C. Guestrin; “Xgboost: A scalable tree boosting system”; in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 13 - 17 August 2016; San Francisco, California, USA. 2016; pp.785-794.
- [10] J. Kennedy and R. Eberhart; “Particle swarm optimization”; in *Proc. ICNN'95 - International Conference on Neural Networks*, 27 November – 01 December 1995; Perth, WA, Australia. IEEE; 1995; 4:1942-1948.
- [11] J. M. V. Gómez-Soberón; “Porosity of recycled concrete with substitution of recycled concrete aggregate: An experimental study”. *Cement and Concrete Research.* 2002; 32(8):1301-1311. DOI:10.1016/S0008-8846(02)00795-0.
- [12] M. Etxeberria, A. R. Marí, and E. Vázquez, “Recycled aggregate concrete as structural material”. *Materials and Structures.* 2007; 40(5):529-541. DOI:10.1617/s11527-006-9161-5.
- [13] L. Evangelista and J. de Brito; “Mechanical behaviour of concrete made with fine recycled concrete aggregates”. *Cement and Concrete Composites.* 2007; 29(5):397-401. DOI:10.1016/j.cemconcomp.2006.12.004.
- [14] S. C. Kou, C. S. Poon, and D. Chan; “Influence of Fly Ash as Cement Replacement on the Properties of Recycled Aggregate Concrete”. *Journal of Materials in Civil Engineering.* 2007; 19(9):709-717. DOI:10.1061/(ASCE)0899-1561(2007)19:9(709).
- [15] S. C. Kou, C. S. Poon, and D. Chan; “Influence of fly ash as a cement addition on the hardened properties of recycled aggregate concrete”. *Materials and Structures.* 2008; 41(7):1191-1201. DOI:10.1617/s11527-007-9317-y.
- [16] K. -H. Yang, H. -S. Chung, and A. F. Ashour; “Influence of Type and Replacement Level of Recycled Aggregates on Concrete Properties”. *Aci Materials Journal.* 2008; 105(3):289-296.
- [17] R. Kumutha and K. Vijai; “Strength of concrete incorporating aggregates recycled from demolition waste”. *ARPN Journal of Engineering and Applied Sciences.* 2010; 5(5):65-72.
- [18] C. J. Zega and A. A. D. Maio; “Recycled Concretes Made with Waste Ready-Mix Concrete as Coarse Aggregate”. *Journal of Materials in Civil Engineering.* 2011; 23(3):281-286. DOI:10.1061/(ASCE)MT.1943-5533.0000165.

- [19] G. Fathifazl, A. G. Razaqpur, O. B. Isgor, A. Abbas, B. Fournier, and S. Foo; “Creep and drying shrinkage characteristics of concrete produced with coarse recycled concrete aggregate”. *Cement and Concrete Composites*. 2011; 33(10):1026-1037. DOI:10.1016/j.cemconcomp.2011.08.004.
- [20] M. C. Rao, S. Bhattacharyya, and S. Barai, “Influence of field recycled coarse aggregate on properties of concrete”. *Materials and Structures*. 2011; 44(1):205-220. DOI:10.1617/s11527-010-9620-x.
- [21] A. Abd Elhakam, A. E. Mohamed, and E. Awad; “Influence of self-healing, mixing method and adding silica fume on mechanical properties of recycled aggregates concrete”. *Construction and Building Materials*. 2012; 35:421-427. DOI:10.1016/j.conbuildmat.2012.04.013.
- [22] P. Pereira, L. Evangelista, and J. De Brito; “The effect of superplasticizers on the mechanical performance of concrete made with fine recycled concrete aggregates”. *Cement and concrete composites*. 2012; 34(9):1044-1052. DOI:10.1016/j.cemconcomp.2012.06.009.
- [23] L. Butler, J. S. West, and S. L. Tighe; “Effect of recycled concrete coarse aggregate from multiple sources on the hardened properties of concrete with equivalent compressive strength”. *Construction and Building Materials*. 2013; 47:1292-1301. DOI:10.1016/j.conbuildmat.2013.05.074.
- [24] C. Thomas, J. Setién, J. A. Polanco, P. Alaejos, and M. Sánchez de Juan; “Durability of recycled aggregate concrete”. *Construction and Building Materials*. 2013; 40:1054-1065. DOI:10.1016/j.conbuildmat.2012.11.106.
- [25] K. H. Younis and K. Pilakoutas; “Strength prediction model and methods for improving recycled aggregate concrete”. *Construction and Building Materials*. 2013; 49:688-701. DOI:10.1016/j.conbuildmat.2013.09.003.
- [26] D. Pedro, J. de Brito, and L. Evangelista; “Performance of concrete made with aggregates recycled from precasting industry waste: influence of the crushing process”. *Materials and Structures*. 2015; 48(12):3965-3978. DOI:10.1617/s11527-014-0456-7.
- [27] C. Thomas, I. Sosa, J. Setién, J. A. Polanco, and A. I. Cimentada; “Evaluation of the fatigue behavior of recycled aggregate concrete”. *Journal of Cleaner Production*. 2014; 65:397-405. DOI:10.1016/j.jclepro.2013.09.036.
- [28] I. H. Witten, E. Frank; “Data mining: practical machine learning tools and techniques with Java implementations”. *ACM Sigmod Record*. 2002; 31:76–77.