



Predicting lightweight concrete strength using stacking model and shap analysis

Bui Gia Linh*, Vu Dinh Phien

University of Transport Technology, Hanoi, Vietnam

Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2024.vn.4.3.48-60>

*Corresponding author:

Email address:

linhbg@utt.edu.vn

Received: 07/06/2024

Revised: 21/07/2024

Accepted: 31/07/2024

Abstract: This study proposes a stacking ensemble machine learning (ML) model for predicting the compressive strength of lightweight aggregate concrete. The model was developed and evaluated using a dataset comprising material composition parameters and experimental compressive strength results. The Stacking Ensemble model achieved high predictive performance with a correlation R of 0.906 and RMSE of 7.782 MPa on the testing dataset. Furthermore, the study employed SHapley Additive exPlanations (SHAP) to analyze the importance of input parameters on compressive strength. The SHAP analysis revealed that natural coarse aggregate content, water-to-binder ratio, natural lightweight aggregate content, and water content were the most influential factors. Two-dimensional SHAP analysis also elucidated the complex relationship between these parameters, indicating that the combination of low CLWA and W enhances the compressive strength. This research not only provides an effective prediction model but also opens up new avenues for optimizing lightweight aggregate concrete compositions based on SHAP analysis, thus contributing to improved quality and efficiency in the construction industry.

Keywords: Machine learning; compressive strength; lightweight concrete; Stacking ensemble; SHAP.



Dự báo cường độ bê tông nhẹ bằng mô hình stacking và shap

Thông tin bài viết

Dạng bài viết:

Bài báo nghiên cứu

DOI:

<https://doi.org/10.58845/jstt.utt.2024.vn.4.3.48-60>

*Tác giả liên hệ:

Địa chỉ Email:

linhbg@utt.edu.vn

Ngày nộp bài: 07/06/2024

Ngày nộp bài sửa: 21/07/2024

Ngày chấp nhận: 31/07/2024

Bùi Gia Linh*, Vũ Đình Phiên

Trường Đại học Công nghệ Giao thông vận tải, Hà Nội, Việt Nam

Tóm tắt: Nghiên cứu sử dụng mô hình Stacking Ensemble nhằm ước tính cường độ nén (CS) của bê tông cốt liệu nhẹ. Mô hình Stacking được đào tạo và đánh giá trên bộ dữ liệu bao gồm các thông số thành phần vật liệu và kết quả thử nghiệm CS thực tế. Mô hình đạt hệ số tương quan R là 0.906 và sai số RMSE là 7.782 MPa trên dữ liệu kiểm chứng. Phương pháp SHAP được sử dụng để phân tích tầm quan trọng của các yếu tố đầu vào. Các yếu tố chính ảnh hưởng đến CS của bê tông bao gồm hàm lượng cốt liệu lớn thông thường, tỷ lệ nước trên chất kết dính, hàm lượng cốt liệu nhẹ cỡ lớn và lượng nước. Đặc biệt, phân tích SHAP hai chiều đã chỉ ra rằng sự kết hợp giữa CLWA và W thấp có thể dẫn đến sự gia tăng CS. Nghiên cứu này không chỉ giới thiệu một mô hình dự đoán hiệu quả mà còn mở ra những hướng nghiên cứu mới về tối ưu hóa thành phần bê tông cốt liệu nhẹ dựa trên phân tích SHAP, từ đó đóng góp vào việc nâng cao chất lượng và hiệu quả trong lĩnh vực xây dựng.

Từ khóa: Học máy; cường độ; bê tông nhẹ; Stacking ensemble; SHAP.

1. Giới thiệu vấn đề

Bê tông truyền thống, dù là vật liệu nền tảng trong xây dựng, đang dần bộc lộ những hạn chế khi đối mặt với những yêu cầu khắt khe của các công trình mới. Trọng lượng lớn của nó không chỉ gây áp lực lên nền móng mà còn làm tăng chi phí thi công. Bê tông cốt liệu nhẹ, với việc thay thế cốt liệu đá bằng các loại nhẹ hơn như đá bọt, xỉ than, đang là một giải pháp hứa hẹn. Ưu điểm vượt trội của bê tông cốt liệu nhẹ là khối lượng thể tích thấp hơn đáng kể so với bê tông truyền thống, giúp giảm tải trọng lên toàn bộ kết cấu và tối ưu hóa chịu lực. Hơn nữa, vật liệu này còn có các tính năng cách nhiệt và âm, góp phần giảm thiểu năng lượng sử dụng, qua đó giúp cải thiện hiệu quả sử dụng của công trình. Sự ra đời của bê tông nhẹ không những đáp ứng nhu cầu của các công trình hiện đại, mà còn mở ra hướng đi mới cho thiết kế tối ưu và phát triển bền vững. Với những ưu điểm về kỹ thuật và tiềm năng ứng dụng lớn, bê tông cốt liệu nhẹ hứa

hẹn sẽ trở thành vật liệu chủ đạo trong tương lai của ngành xây dựng [1].

Bê tông nhẹ, sử dụng cốt liệu nhẹ, là một giải pháp vật liệu đột phá, được tạo thành bằng cách thay thế cốt liệu đá truyền thống bằng các loại cốt liệu có trọng lượng nhẹ (LWA). Nguồn gốc của các LWA này vô cùng đa dạng, từ đất sét nung trương nở, đá phiến nung, cho đến các sản phẩm phụ công nghiệp như xỉ lò cao và tro bay [2]. Sự kết hợp này vừa tạo ra cấu trúc lỗ rỗng đặc trưng, vừa giảm đáng kể khối lượng riêng của bê tông, từ đó mở rộng khả năng ứng dụng trong xây dựng [3]. Ưu điểm nổi bật của loại vật liệu này không chỉ là giảm tải trọng đáng kể cho công trình, giúp tiết kiệm chi phí nền móng và kết cấu [4]. Bê tông nhẹ còn có khả năng hấp thụ, phân tán năng lượng địa chấn [5], giúp cải thiện khả năng chịu động đất của các tòa nhà. Ngoài ra, quá trình sản xuất bê tông nhẹ còn giúp đẩy mạnh phát triển bền vững [6]. Việc sử dụng các sản phẩm phụ công nghiệp làm

cốt liệu không chỉ giúp giảm thiểu chất thải và lượng khí thải nhà kính mà còn tiết kiệm tài nguyên thiên nhiên. Tuy nhiên, cần lưu ý rằng bê tông nhẹ có thể có cường độ chịu nén thấp hơn bê tông thông thường do cốt liệu nhẹ thường có độ rỗng cao [7]. Do đó, cần xem xét kỹ lưỡng việc lựa chọn và thiết kế bê tông nhẹ để đảm bảo công trình đạt yêu cầu về độ bền và an toàn.

Nhiều công trình thực nghiệm đã tiến hành xác định tương quan giữa thành phần bê tông nhẹ và đặc tính cơ học, đặc biệt là khả năng chịu nén. Kim và cộng sự [8] đã tìm ra rằng thể tích của LWA có quan hệ tỉ lệ nghịch với cường độ nén (CS). Điều này nhấn mạnh vai trò quan trọng của việc lựa chọn và kiểm soát hàm lượng LWA trong thiết kế bê tông. Tuy nhiên, ảnh hưởng của LWA không chỉ giới hạn ở hàm lượng. Loại LWA cũng đóng một vai trò quan trọng đối với CS của bê tông nhẹ [9,10]. Các loại LWA khác nhau có đặc tính lý hóa khác nhau, trực tiếp tác động đến khả năng liên kết với xi măng và hình thành cấu trúc bê tông. Việc lựa chọn loại và tỷ lệ LWA phù hợp, bao gồm tỷ lệ nước - xi, cốt liệu, nước, đều ảnh hưởng đến tính năng của bê tông [11]. Tỷ lệ nước trên xi măng tác động đến quá trình hydrat hóa, trong khi lượng cốt liệu và chất kết dính quyết định độ đặc, độ rỗng và liên kết của bê tông. Các tiêu chuẩn thiết kế đã có hướng dẫn đánh giá ảnh hưởng của các thông số này đến CS của bê tông nhẹ [12]. Tuy nhiên, do sự phức tạp của các đặc tính của cốt liệu cũng như tỷ lệ thành phần, các hướng dẫn này vẫn chưa thống nhất. Vì vậy, cần nghiên cứu các mô hình dự đoán chính xác hơn và cung cấp hướng dẫn thiết kế rõ ràng, toàn diện cho việc sử dụng bê tông nhẹ.

Gần đây, với sự phát triển của công nghệ thông tin, học máy (ML) dần trở thành một công cụ phổ biến và được ứng dụng trong nhiều lĩnh vực. Điểm mạnh của ML nằm ở khả năng phân tích và học hỏi từ dữ liệu, qua đó đem đến các dự đoán và quyết định chính xác. Đặc biệt, ML đã cung cấp những phương án nhanh chóng, hiệu quả để xử lý các vấn đề phức tạp trong kỹ thuật, giúp tăng hiệu quả và năng suất [13-18]. Trong lĩnh vực vật liệu

xây dựng, ML cũng đã chứng minh được tiềm năng to lớn trong các bài toán liên quan bê tông. Các nghiên cứu đã sử dụng ML để dự đoán CS của nhiều loại bê tông, ví dụ bê tông cao su [19], bê tông tái chế [20,21], bê tông có tro, xỉ [22,23], bê tông tự đầm [24], bê tông cường độ cao và siêu cao [25-27]. Những nghiên cứu cho thấy ML không chỉ dự đoán tốt các tính chất được xét tới mà còn giúp tối ưu thành phần và quy trình thiết kế. Sự tích hợp của ML trong ngành vật liệu đã mở ra nhiều hướng đi mới trong ngành xây dựng.

Nhằm mục tiêu dự đoán và tối ưu hóa CS của bê tông nhẹ, bài báo này tập trung khai thác và ứng dụng tiềm năng của ML. Năm mô hình cây được lựa chọn để kiểm chứng, bao gồm Rừng ngẫu nhiên, Tăng cường độ dốc, Tăng cường độ dốc cấp cao, Stacking Ensemble và Voting Ensemble. Quy trình nghiên cứu áp dụng các kỹ thuật xác thực chéo và kỹ thuật SHAP để đảm bảo tính khách quan, độ tin cậy và giải thích mô hình một cách minh bạch. Nghiên cứu được trình bày thành bốn phần, bao gồm thông tin về cơ sở dữ liệu, thuật toán, phân tích kết quả và kết luận. Nghiên cứu này không chỉ dự đoán CS mà còn đánh giá tầm quan trọng của các thông số đầu vào, góp phần tối ưu hóa thành phần và quy trình sản xuất bê tông cốt liệu nhẹ. Kết quả nghiên cứu cung cấp thông tin chuyên sâu, hỗ trợ quá trình phát triển vật liệu xây dựng bền vững và hiệu quả.

2. Cơ sở dữ liệu và thuật toán

2.1. Dữ liệu

Một yếu tố then chốt để phát triển hiệu quả mô hình ML là dữ liệu đầy đủ và tin cậy. Trong bài báo này, bộ dữ liệu được sử dụng được tổng hợp từ kết quả thử nghiệm của 15 công bố uy tín, với 149 mẫu bê tông (20 mẫu bê tông cốt liệu thường và 129 mẫu bê tông cốt liệu nhẹ) và các tỷ lệ hỗn hợp, cốt liệu khác nhau [1,8,10,28-39]. Lưu ý rằng các mẫu bê tông cốt liệu thường được thêm vào dữ liệu như một nhóm đối chứng không chứa cốt liệu nhẹ. Mô hình ML được xây dựng với 9 thông số đầu vào liên quan đến tỷ lệ hoặc hàm lượng hỗn hợp (tính theo kg/m^3) và các đặc tính vật lý. CS của

bê tông nhẹ là biến đầu ra. Bảng 1 thể hiện thống kê về các tham số được sử dụng. Phân tích dữ liệu là nhiệm vụ quan trọng trong khoa học dữ liệu, giúp hiểu rõ dữ liệu và định hướng bài toán. Hình 1 biểu diễn phân bố các biến được sử dụng.

2.2. Mô hình dự đoán

- **Mô hình Rừng ngẫu nhiên**

Rừng ngẫu nhiên (RF) là thuật toán quan trọng trong ML, dùng cho dự báo và phân loại RF xây dựng nhiều cây quyết định từ dữ liệu huấn luyện ngẫu nhiên, giảm overfitting và tăng độ chính xác [40]. RF có khả năng linh hoạt, xử lý dữ liệu lớn, nhiễu, không đồng nhất. Tuy nhiên, RF huấn luyện chậm và đôi khi vẫn xảy ra tình trạng overfitting.

- **Mô hình tăng cường độ dốc**

Tăng cường Độ dốc (GB) là thuật toán ML được sử dụng nhiều cho phân loại và hồi quy [41]. GB kết hợp nhiều mô hình yếu để tạo ra mô hình mạnh, cải thiện liên tục bằng cách xây dựng chuỗi cây quyết định dựa trên phần dư của mô hình trước đó [41]. GB mang lại dự đoán chính xác, phát hiện mẫu phức tạp, xử lý dữ liệu lớn, nhưng dễ overfitting, cần lựa chọn tham số cẩn thận và tốn thời gian huấn luyện.

- **Mô hình tăng cường độ dốc cấp cao**

Tăng cường độ dốc cấp cao (XGBoost) là biến thể mạnh mẽ của GB, được phát triển gần đây nhằm giải quyết nhiều dạng bài toán [42]. XGBoost hoạt động tương tự GB nhưng cải tiến để tăng hiệu suất và tốc độ huấn luyện, sử dụng Gradient Descent để tối ưu hóa hàm mất mát và hỗ trợ nhiều chế độ chính quy hóa. Ưu điểm của XGBoost là xử lý dữ liệu lớn nhanh, tạo mô hình mạnh và ổn định, hỗ trợ nhiều loại biến đầu vào. Tuy nhiên, XGBoost cần nhiều tài nguyên và thời gian huấn luyện, cần điều chỉnh siêu tham số để tối ưu và tránh overfitting.

- **Mô hình Stacking Ensemble**

Stacking Ensemble là kỹ thuật ML kết hợp dự báo từ nhiều mô hình để tạo mô hình tổng thể mạnh hơn [43]. Nó sử dụng dự đoán của các mô

hình đơn làm đầu vào cho mô hình meta-model nhằm cải thiện khả năng dự đoán [16]. Quá trình hoạt động gồm huấn luyện mô hình cơ sở trên dữ liệu ban đầu, rồi huấn luyện mô hình meta-model trên dự đoán của các mô hình cơ sở. Các mô hình cơ sở dự đoán trên dữ liệu mới, rồi đưa vào mô hình meta-model để tạo dự đoán cuối cùng.

- **Mô hình Voting Ensemble**

Voting Ensemble là kỹ thuật ML kết hợp dự đoán từ nhiều mô hình cơ sở bằng quy tắc biểu quyết [16]. Các mô hình được huấn luyện trên cùng dữ liệu, sau đó kết hợp bằng biểu quyết để dự đoán cuối cùng. Có hai loại: hard voting (chọn dự đoán phổ biến nhất) và soft voting (trung bình xác suất dự đoán).

2.3. Đánh giá khả năng dự báo mô hình

Để đánh giá mô hình ML, nghiên cứu sử dụng các chỉ số phổ biến như R, RMSE và MAE. Mô hình tốt nhất có R lớn nhất và RMSE, MAE nhỏ nhất. Chi tiết các tiêu chí có thể được tìm hiểu trong tài liệu [44].

3. Kết quả và thảo luận

3.1. Phát triển mô hình ML

Trong phần này, quy trình xây dựng các mô hình dự báo được thực hiện để tạo ra mô hình chính xác và tin cậy. Quá trình gồm hai giai đoạn: huấn luyện (70% dữ liệu) và kiểm chứng (30% dữ liệu). Tỷ lệ 70-30 này được lựa chọn dựa trên khuyến nghị của Sharma và cộng sự [45] và Salcedo-Sanz và cộng sự [46] để đảm bảo hiệu quả phát triển mô hình. Xác thực chéo 10-fold được áp dụng: dữ liệu huấn luyện được chia thành 10 phần, mô hình được huấn luyện 10 lần, mỗi lần dùng 1 phần để xác nhận và 9 phần còn lại để huấn luyện, dữ liệu kiểm chứng giữ nguyên.

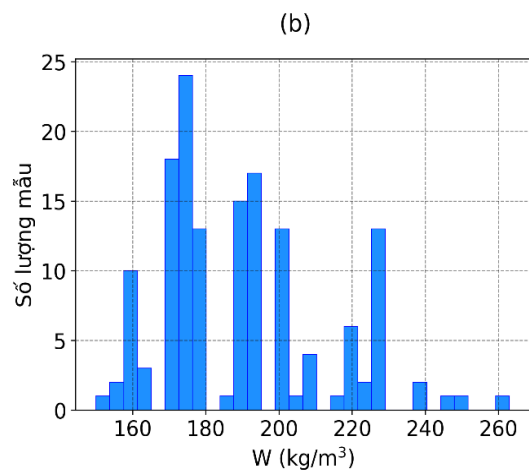
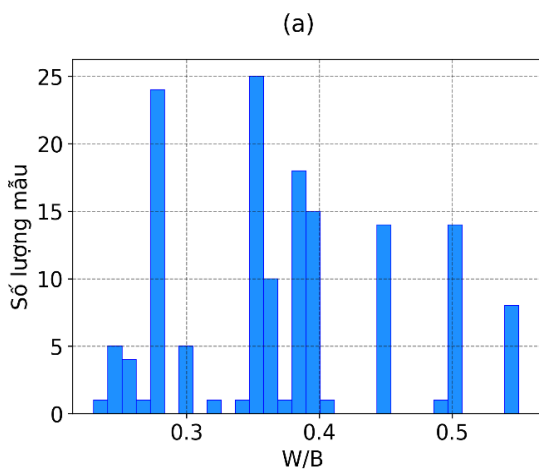
Bảng 2 cho thấy XGBoost và GB đạt hiệu suất tốt nhất trên dữ liệu huấn luyện ($R = 0.999$). Tuy nhiên, hiệu suất giảm đáng kể trên dữ liệu xác thực và kiểm chứng, cho thấy khả năng quá khớp của hai mô hình này. Trong số các mô hình ML đơn lẻ, mô hình Random Forest đạt hiệu suất trên dữ liệu huấn luyện không tốt bằng hai mô hình còn lại,

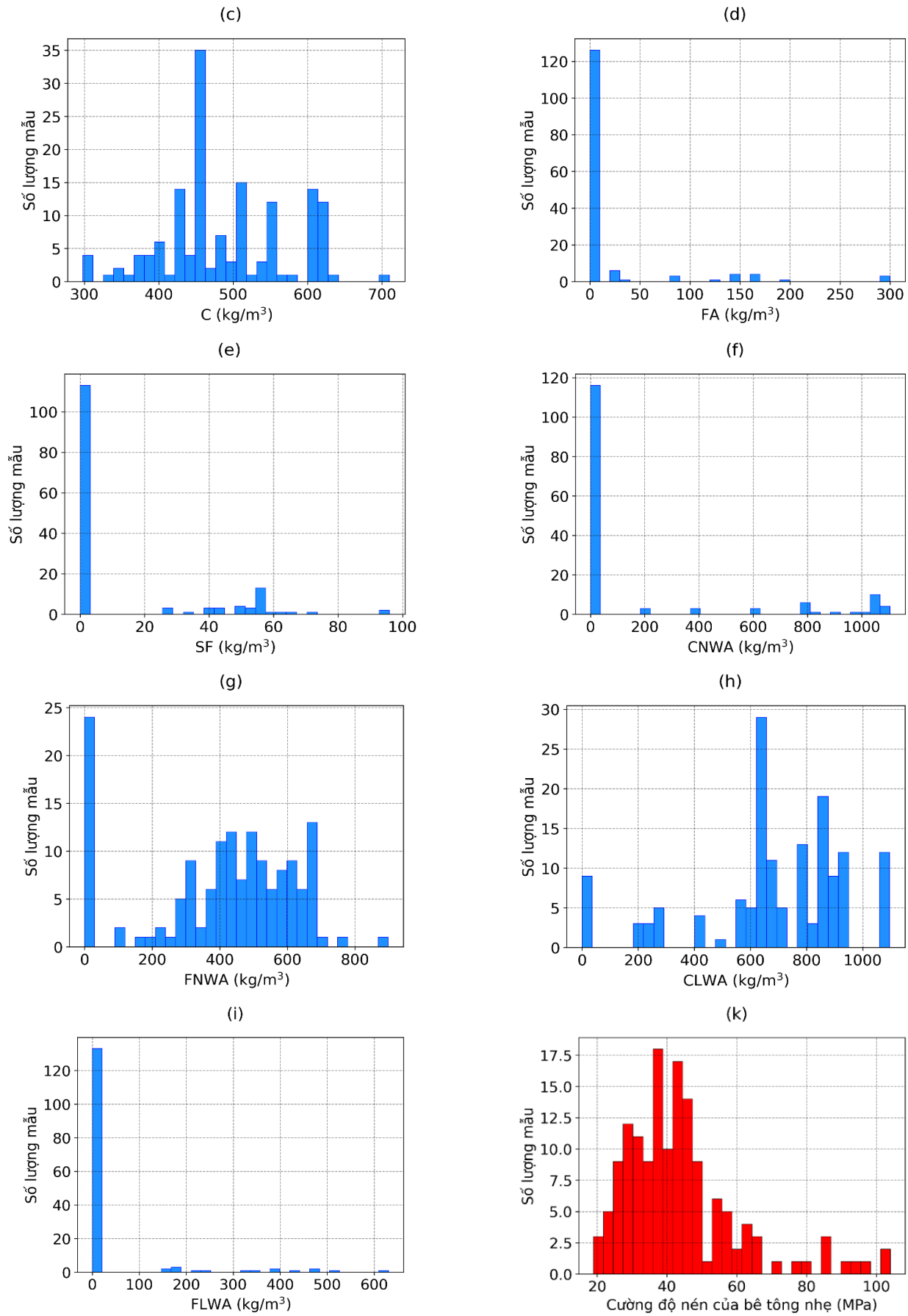
nhưng lại tốt hơn trên dữ liệu kiểm chứng. Điều này là do mô hình GB và XGBoost thường có trọng số thấp hơn và phương sai cao hơn so với Random Forest, nên dễ bị ảnh hưởng bởi các biến động ngẫu nhiên trong tập dữ liệu kiểm chứng. Đối với nhóm mô hình ML kết hợp, mô hình Stacking Ensemble nổi bật là mô hình có khả năng dự đoán tốt nhất trong số 5 mô hình ML khi xét trên tập dữ liệu kiểm chứng, với hệ số tương quan lớn nhất ($R = 0.906$) và sai số nhỏ nhất ($RMSE = 7.782$ MPa). Kết quả đạt được trên tập dữ liệu kiểm chứng là rất quan trọng vì chúng phản ánh hiệu suất dự đoán thực tế của các mô hình ML. Trong mô hình

Stacking Ensemble, mô hình tổng hợp được huấn luyện để tối ưu hóa hiệu suất dự đoán dựa trên các dự đoán từ các mô hình cơ sở, đảm bảo tính tương thích tốt giữa các mô hình và học cách kết hợp chúng một cách hiệu quả. Trong khi đó, mô hình Voting Ensemble chỉ đơn giản kết hợp dự đoán từ các mô hình cơ sở bằng cách sử dụng cơ chế đơn giản như đa số phiếu hoặc trung bình các dự đoán, có thể làm giảm tính tương thích giữa các mô hình và giảm khả năng tối ưu hóa hiệu suất dự đoán. Do đó, trong nghiên cứu này, mô hình Stacking Ensemble được chọn để trình bày trong các phần tiếp theo.

Bảng 1. Phân tích dữ liệu

	Ký hiệu	Đơn vị	Mean	Std	Min	Q25%	Q50%	Q75%	Max
Biến đầu vào									
Tỉ lệ nước/kết dính	W/B	-	0.37	0.08	0.23	0.30	0.36	0.41	0.55
Nước	W	kg/m ³	189.66	22.51	150	173	190	202	263
Xi măng	C	kg/m ³	488.70	83.80	298	437	460	550	710
Tro bay	FA	kg/m ³	19.35	57.11	0.00	0.00	0.00	0.00	300
Muội silic	SF	kg/m ³	12.72	23.72	0.00	0.00	0.00	0.00	96.00
Cốt liệu lớn thông thường	CNWA	kg/m ³	181.15	366.18	0.00	0.00	0.00	0.00	1105
Cốt liệu mịn thông thường	FNWA	kg/m ³	407.52	221.71	0.00	304	447	583	898
Cốt liệu nhẹ cỡ lớn	CLWA	kg/m ³	684.39	270.48	0.00	626	684	861	1096
Cốt liệu nhẹ cỡ nhỏ	FLWA	kg/m ³	35.35	112.46	0.00	0.00	0.00	0.00	631
Biến đầu ra									
Cường độ nén	CS	MPa	43.64	16.60	19	33	41	48	104





Hình 1. Biểu đồ các tham số

Bảng 2. Tổng hợp kết quả dự đoán các mô hình trên các tập dữ liệu.

Mô hình	Tập dữ liệu	RMSE (MPa)	MAE (MPa)	R
RF	Đào tạo	3.068	2.265	0.980
	Xác nhận	6.651	5.753	0.796
	Kiểm chứng	7.908	5.850	0.901
XGBoost	Đào tạo	0.310	0.061	0.999
	Xác nhận	6.652	5.751	0.798
	Kiểm chứng	8.084	5.450	0.896
GB	Đào tạo	0.310	0.058	0.999
	Xác nhận	8.875	6.673	0.725
	Kiểm chứng	9.955	6.694	0.838
Stacking Ensemble	Đào tạo	2.295	1.707	0.989
	Xác nhận	6.650	5.752	0.829
	Kiểm chứng	7.782	5.781	0.906
Voting Ensemble	Đào tạo	1.077	0.819	0.997
	Xác nhận	6.651	5.751	0.825
	Kiểm chứng	8.218	5.651	0.893

3.2. Mô hình Stacking Ensemble

Kết quả dự đoán CS của bê tông nhẹ bằng mô hình Stacking Ensemble được trình bày. Hình 2a và 2b minh họa phân bố sai số dự đoán trên hai tập dữ liệu. Với tập huấn luyện, phần lớn sai số nằm trong khoảng $[-4, 4]$ MPa, với chỉ 9.6% mẫu có sai số vượt ngoài khoảng này. Với tập kiểm chứng, sai số tập trung chủ yếu trong khoảng $[-20, 20]$ MPa, và chỉ có 4.4% sai số nằm ngoài khoảng trên.

Hình 3a và 3b minh họa mối quan hệ giữa CS thực tế và CS dự đoán bởi mô hình Stacking Ensemble trên tập huấn luyện và tập kiểm chứng tương ứng. Trục hoành biểu diễn CS thực tế của các mẫu bê tông cốt liệu nhẹ, trong khi trục tung biểu diễn CS dự đoán. Đường liền màu đen thể hiện sự tương đồng giữa giá trị dự đoán và giá trị thực tế. Các đường chấm màu đen và đường đứt màu đỏ biểu thị biên độ sai số $\pm 15\%$ và $\pm 30\%$ tương ứng. Mô hình đạt hiệu suất gần như hoàn hảo trên tập huấn luyện, với phần lớn các điểm dữ liệu nằm trên đường $y=x$. Trên tập kiểm chứng, mô hình cũng cho thấy hiệu suất tốt với hầu hết các điểm dữ liệu nằm trong phạm vi sai số $\pm 30\%$.

3.3. Giải nghĩa mô hình ML

SHAP là một phương pháp giải thích mô hình

máy học tiên tiến, cung cấp cái nhìn về cách các đặc trưng đầu vào ảnh hưởng tới dự đoán mô hình. Vận dụng lý thuyết trò chơi và phân phối giá trị Shapley, SHAP định lượng mức độ đóng góp của từng biến tới dự đoán, giải thích cho từng dự đoán cụ thể và cho toàn bộ mô hình. SHAP hỗ trợ phân tích dữ liệu nhằm hiểu rõ cơ chế hoạt động của mô hình, từ đó phát hiện lỗi sai, cải thiện hiệu suất và đặc biệt giải thích kết quả dự đoán một cách tường minh và dễ hiểu [47]. Hình 4 trình bày phân tích ảnh hưởng của các thông số đến CS của bê tông nhẹ bằng phương pháp SHAP. Biểu đồ dạng bee-swarm (hình 4a) và biểu đồ giá trị tuyệt đối (hình 4b) được sử dụng để minh họa kết quả.

Kết quả cho thấy sự tương đồng về tầm quan trọng của các thông số giữa hai loại biểu đồ. Các thông số đầu vào có ảnh hưởng lớn nhất đến CS của bê tông nhẹ bao gồm hàm lượng CNWA, W/B, CLWA, W và FNWA. Ngược lại, các biến như C, FA, SF và FLWA thể hiện ảnh hưởng tương đối nhỏ. Phân tích SHAP đã chỉ ra mức độ đóng góp của từng thông số đến CS của bê tông nhẹ. Từ đó, có thể xác định các biến cần được kiểm soát trong quá trình thiết kế để tối ưu hóa CS của loại bê tông này.

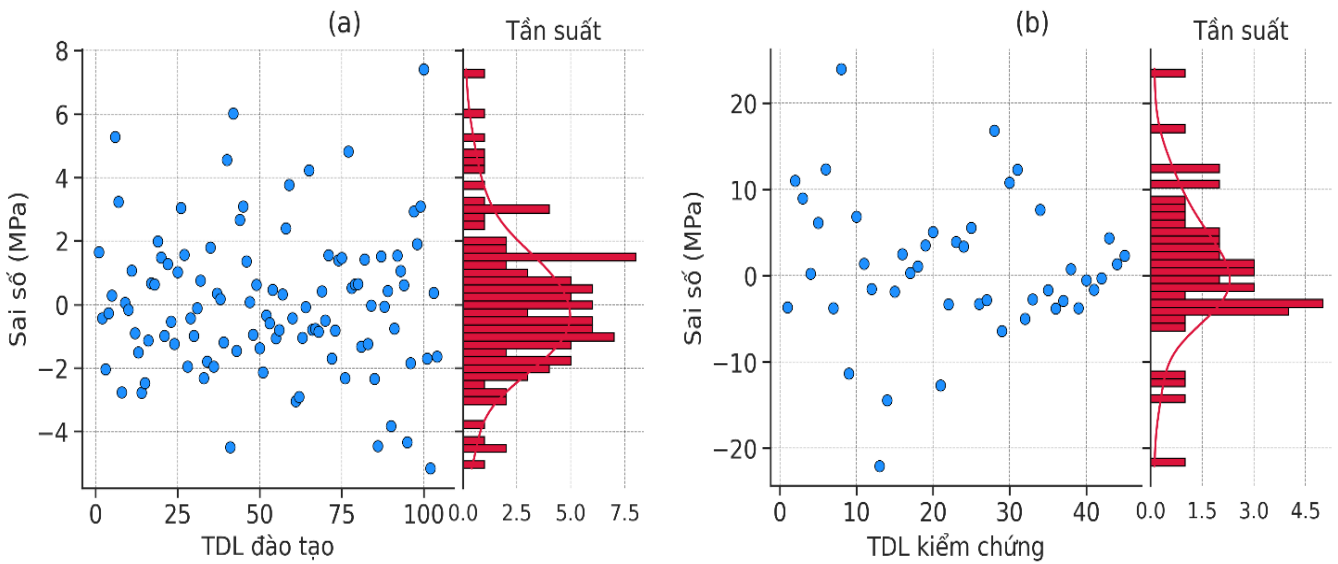
Hình 5 trình bày các mối quan hệ SHAP hai

chiều giữa các cặp biến, tập trung vào 4 biến có ảnh hưởng lớn đến CS của bê tông nhẹ (CNWA, W/B, CLWA, và W), tạo thành 6 biểu đồ.

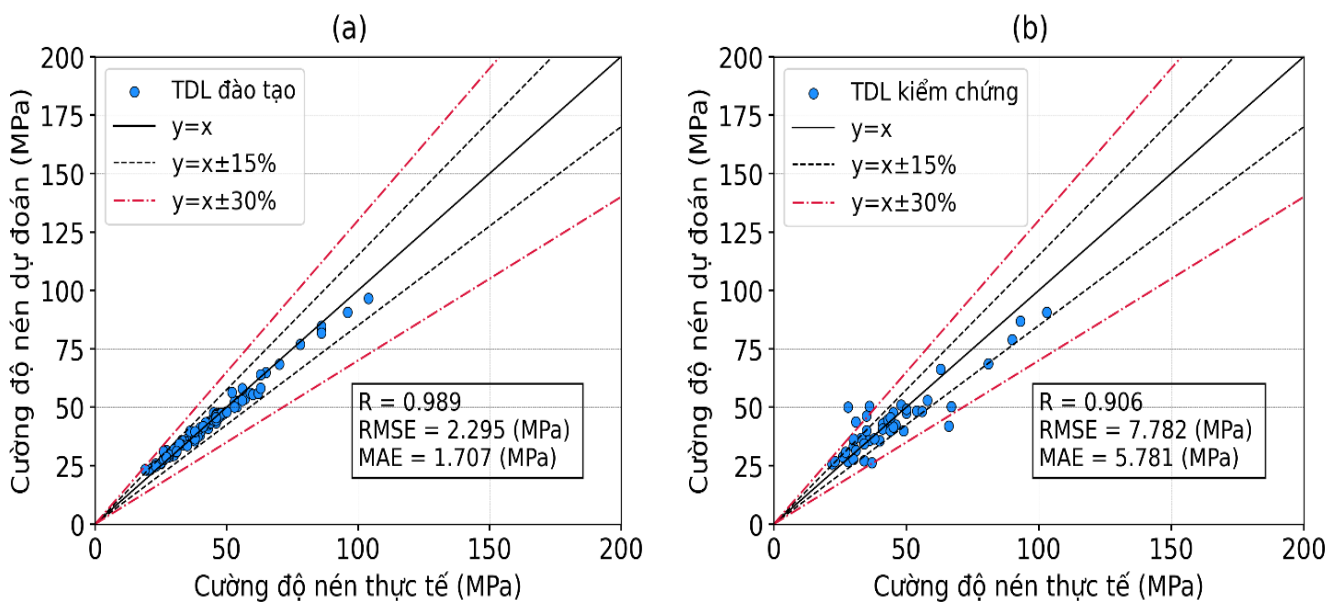
Các hình 5a, 5b, và 5c minh họa mối tương quan giữa CNWA với tỷ lệ W/B, CLWA và W. Kết quả phân tích cho thấy CLWA trong khoảng từ 0 đến 750 kg/m³ không gây biến đổi lớn đến CS và không thể hiện sự phụ thuộc vào W/B, CLWA, và W. Tuy nhiên, khi CLWA và W có giá trị thấp, CS có xu hướng tăng. Các hình 5d và 5e chỉ ra sự giảm dần của CS, đặc biệt khi W/B vượt quá 0.4. Không có mối tương quan rõ ràng giữa W/B với W và CLWA trong việc thay đổi CS. Hình 5f cho thấy

CLWA không ảnh hưởng đáng kể đến CS, và sự thay đổi của W cũng không gây biến động lớn đến CS.

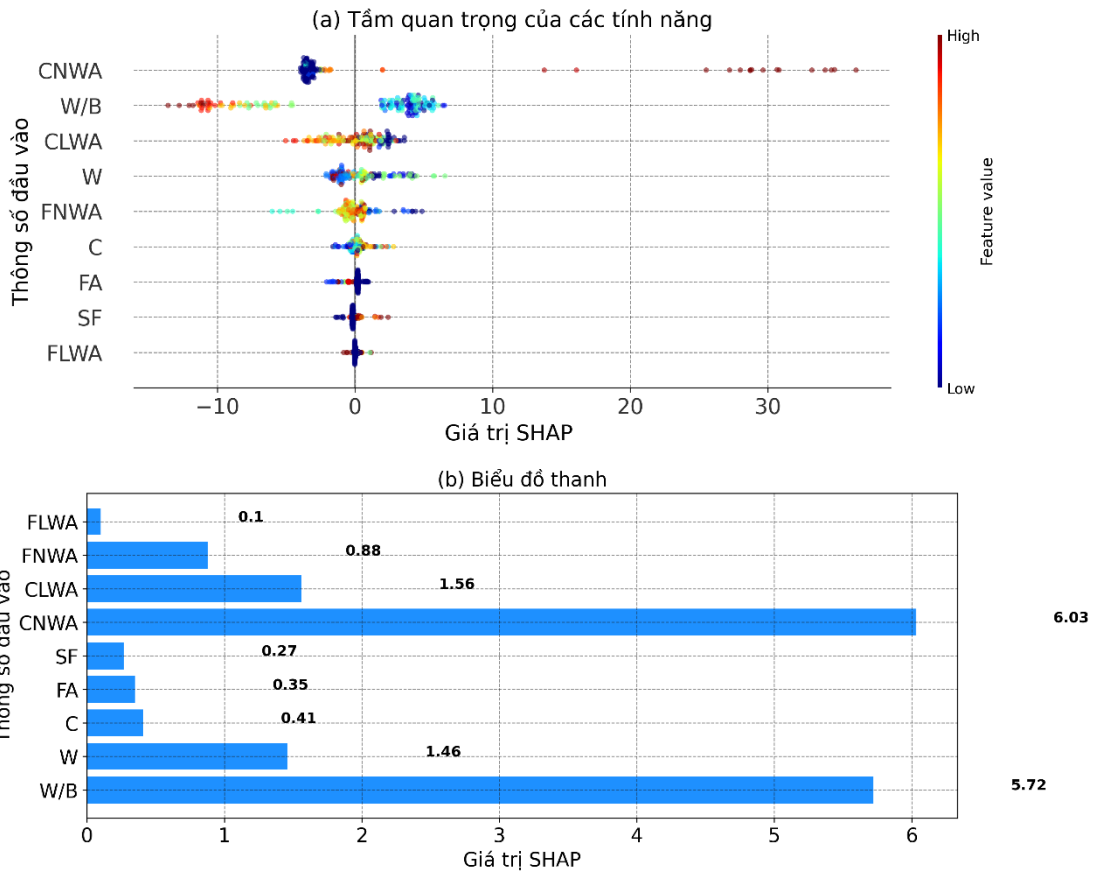
Kết quả phân tích SHAP hai chiều cung cấp góc nhìn chi tiết hơn về mối quan hệ phức tạp giữa các thông số đầu vào và CS của bê tông cốt liệu nhẹ. Mặc dù các biến riêng lẻ như CLWA và W có thể không gây ảnh hưởng lớn khi thay đổi trong một phạm vi nhất định, nhưng sự kết hợp của chúng, đặc biệt là khi có giá trị thấp, có thể làm tăng CS. Điều này nhấn mạnh tầm quan trọng của việc xem xét sự tương tác giữa các biến trong quá trình tối ưu hóa thiết kế bê tông.



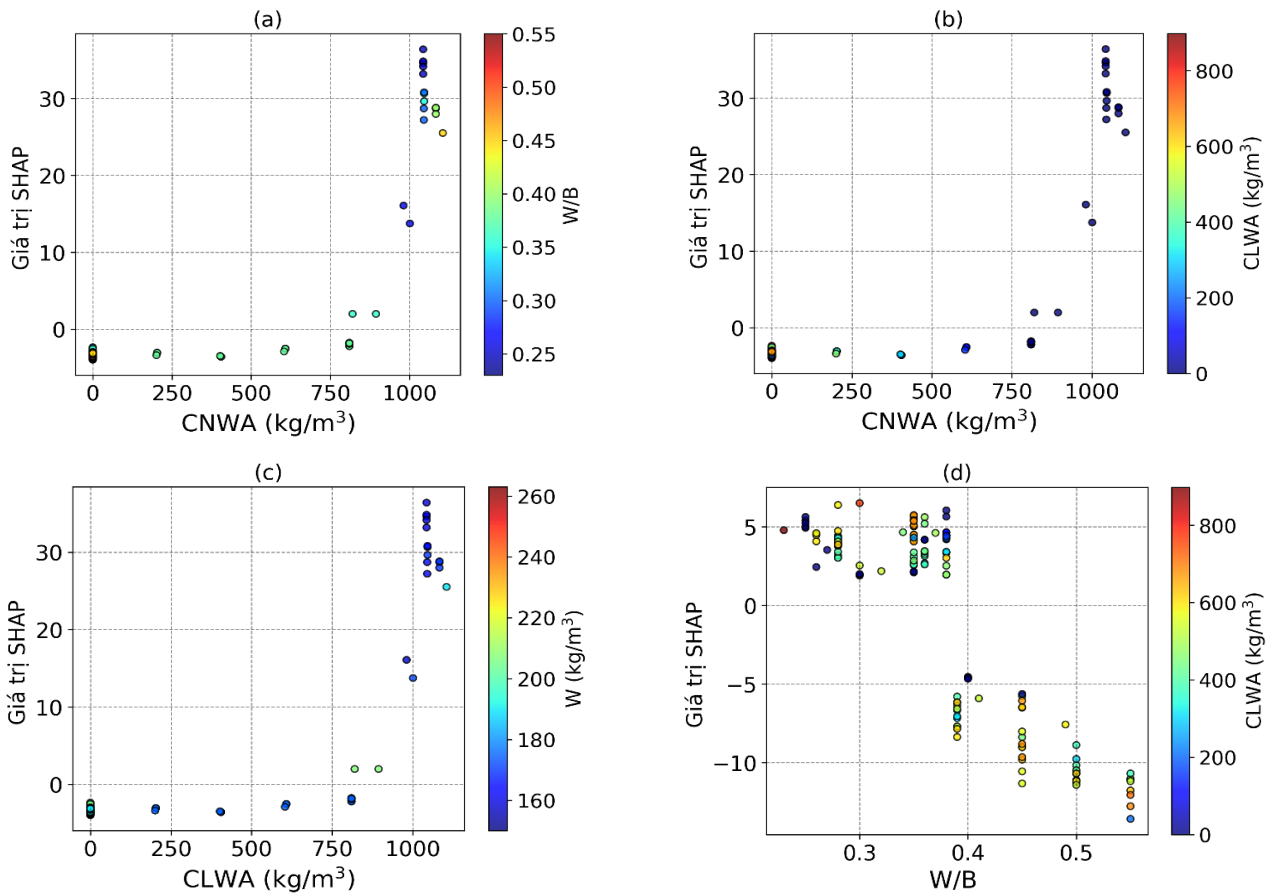
Hình 2. Biểu đồ sai số của mô hình (a) tập đào tạo, (b) tập kiểm chứng

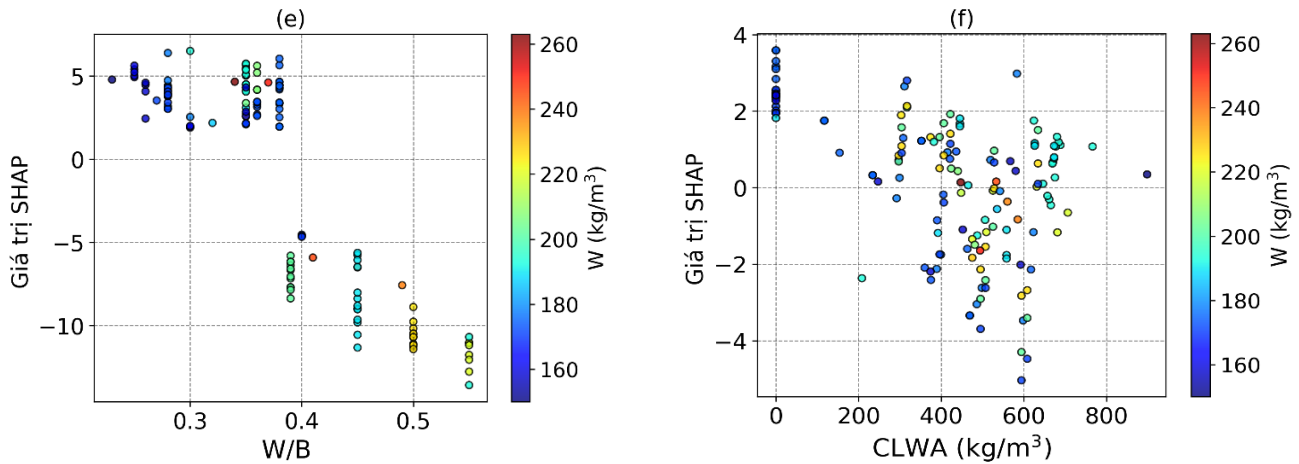


Hình 3. Biểu đồ hồi quy của mô hình (a) tập đào tạo, (b) tập kiểm chứng.



Hình 4. Phân tích SHAP





Hình 5. Phân tích SHAP hai chiều giữa 4 biến quan trọng nhất đối với CS của bê tông nhẹ.

4. Kết luận

Nghiên cứu này đã giới thiệu một mô hình dự đoán CS của bê tông cốt liệu nhẹ dựa trên kỹ thuật ML Stacking Ensemble, kết hợp với phương pháp phân tích SHAP để làm rõ vai trò và mối liên hệ giữa các thông số đầu vào. Mô hình Stacking Ensemble đã thể hiện khả năng dự đoán chính xác và độ tin cậy cao, vượt trội hơn so với các mô hình đơn lẻ như RF, GB hay XGBoost. Phân tích SHAP không chỉ xác định được những thông số quan trọng nhất ảnh hưởng đến CS mà còn làm sáng tỏ các tương tác phức tạp giữa chúng, cung cấp thông tin giá trị cho việc tối ưu hóa thành phần bê tông cốt liệu nhẹ.

Tuy nhiên, nghiên cứu còn hạn chế về kích thước và tính đa dạng của tập dữ liệu, độ phức tạp của mô hình Stacking Ensemble. Tương lai, cần mở rộng tập dữ liệu, nghiên cứu kỹ thuật ML khác và tích hợp SHAP vào thiết kế bê tông để cải thiện độ chính xác và tạo ra bê tông tốt hơn.

Tài liệu tham khảo

- [1] J.A. Rossignolo, M.V. Agnesini, J.A. Morais. (2003). Properties of high-performance LWAC for precast structures with Brazilian lightweight aggregates. *Cement and Concrete Composites*, 25(1), 77-82.
- [2] A. Elsharief, M.D. Cohen, J. Olek. (2005). Influence of lightweight aggregate on the microstructure and durability of mortar. *Cement and Concrete Research*, 35(7), 1368-1376.
- [3] Y. Ke, S. Ortola, A.L. Beaucour, H. Dumontet. (2010). Identification of microstructural characteristics in lightweight aggregate concretes by micromechanical modelling including the interfacial transition zone (ITZ). *Cement and Concrete Research*, 40(11), 1590-1600.
- [4] L.K. Minapu, M. Ratnam, U. Rangaraju. (2014). Experimental study on light weight aggregate concrete with pumice stone, silica fume and fly ash as a partial replacement of coarse aggregate. *International Journal of Innovative Research in Science, Engineering and Technology*, 3(12), 18130-18138.
- [5] T. Wu, X. Yang, H. Wei, X. Liu. (2019). Mechanical properties and microstructure of lightweight aggregate concrete with and without fibers. *Construction and Building Materials*, 199, 526-539.
- [6] S.K. Patel, R.K. Majhi, H.P. Satpathy, A.N. Nayak. (2019). Durability and microstructural properties of lightweight concrete manufactured with fly ash cenosphere and sintered fly ash aggregate. *Construction and Building Materials*, 226, 579-590.
- [7] B.M. Herki. (2020). Lightweight concrete using local natural lightweight aggregate. *Journal of Critical Reviews*, 7(4), 490-497.
- [8] Y.J. Kim, Y.W. Choi, M. Lachemi. (2010). Characteristics of self-consolidating concrete using two types of lightweight coarse aggregates. *Construction and Building*

- Materials*, 24(1), 11-16.
- [9] J.A. Bogas, A. Gomes, M.F.C. Pereira. (2012). Self-compacting lightweight concrete produced with expanded clay aggregate. *Construction and Building Materials*, 35, 1013-1022.
- [10] N.U. Kockal, T. Ozturan. (2011). Strength and elastic properties of structural lightweight concretes. *Materials & Design*, 32(4), 2396-2403.
- [11] J. Kanadasan, H.A. Razak. (2014). Mix design for self-compacting palm oil clinker concrete based on particle packing. *Materials & Design (1980-2015)*, 56, 9-19.
- [12] D.J. Akers, R.D. Gruber, B.W. Ramme, M.J. Boyle, J.G. Grygar, S.K. Rowe, T.W. Bremner, E.S. Kluckowski, S.R. Sheetz, R.G. Burg. (2003). Guide for structural lightweight-aggregate concrete, ACI 213R-03. *American Concrete Institute (ACI), Michigan (2003)*. http://civilwares.free.fr/ACI/MCP04/213r_03.PDF
- [13] Q.H. Nguyen, H.-B. Ly, T.-A. Nguyen, V.-H. Phan, L.K. Nguyen, V.Q. Tran. (2021). Investigation of ANN architecture for predicting shear strength of fiber reinforcement bars concrete beams. *Plos One*, 16(4), e0247391.
- [14] T.-A. Nguyen, H.-B. Ly, H.-V.T. Mai, V.Q. Tran. (2021). On the training algorithms for artificial neural network in predicting the shear strength of deep beams. *Complexity*, 2021, 5548988, 1-18.
- [15] H.-A. Le, D.-A. Le, T.-T. Le, H.-P. Le, T.-H. Le, H.-G.T. Hoang, T.-A. Nguyen. (2022). An Extreme Gradient Boosting approach to estimate the shear strength of FRP reinforced concrete beams. *Structures*, 45, 1307-1321.
- [16] T.-A. Nguyen, S.H. Trinh, M.H. Nguyen, H.-B. Ly. (2023). Novel ensemble approach to predict the ultimate axial load of CFST columns with different cross-sections. *Structures*, 47, 1-14.
- [17] T.-A. Nguyen, H.-B. Ly, V.Q. Tran. (2022). Predicting shear strength of slender beams without reinforcement using hybrid gradient boosting trees and optimization algorithms. *Frontiers of Structural and Civil Engineering*, 16, 1267–1286.
- [18] T.-A. Nguyen, K.N. Le, H.-B. Ly. (2023). Universal boosting ML approaches to predict the ultimate load capacity of CFST columns. *The Structural Design of Tall and Special Buildings*, 33(2), e2071.
- [19] H.-B. Ly, T.-A. Nguyen, H.-V.T. Mai, V.Q. Tran. (2021). Development of deep neural network model to predict the compressive strength of rubber concrete. *Construction and Building Materials*, 301, 124081.
- [20] F. Khademi, S.M. Jamal, N. Deshpande, S. Londhe. (2016). Predicting strength of recycled aggregate concrete using artificial neural network, adaptive neuro-fuzzy inference system and multiple linear regression. *International Journal of Sustainable Built Environment*, 5(2), 355-369.
- [21] Z.-H. Duan, S.-C. Kou, C.S. Poon. (2013). Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Construction and Building Materials*, 40, 1200-1206.
- [22] S. Sahoo, T.R. Mahapatra. (2018). ANN Modeling to study strength loss of Fly Ash Concrete against Long term Sulphate Attack. *Materials Today: Proceedings*, 5(11), 24595–24604.
- [23] I.B. Topcu, M. Sarıdemir. (2008). Prediction of compressive strength of concrete containing fly ash using artificial neural networks and fuzzy logic. *Computational Materials Science*, 41(3), 305-311.
- [24] E.M. Golafshani, G. Pazouki. (2018). Predicting the compressive strength of self-compacting concrete containing fly ash using a hybrid artificial intelligence method. *Computers and Concrete*, 22(4), 419-437.
- [25] O. Abuodeh, J.A. Abdalla, R.A. Hawileh. (2019). Prediction of compressive strength of ultra-high performance concrete using SFS and ANN. *2019 8th International Conference on*

- Modeling Simulation and Applied Optimization (ICMSAO), IEEE, 2019*, pp 1-5.
- [26] J. Abellán García, J. Fernández Gómez, N. Torres Castellanos. (2020). Properties prediction of environmentally friendly ultra-high-performance concrete using artificial neural networks. *European Journal of Environmental and Civil Engineering*, 2319-2343.
- [27] I.-C. Yeh. (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, 28(12), 1797-1808.
- [28] U.J. Alengaram, H. Mahmud, M.Z. Jumaat. (2011). Enhancement and prediction of modulus of elasticity of palm kernel shell concrete. *Materials & Design*, 32(4), 2143-2148.
- [29] M. Aslam, P. Shafiq, M.Z. Jumaat, M. Lachemi. (2016). Benefits of using blended waste coarse lightweight aggregates in structural lightweight aggregate concrete. *Journal of Cleaner Production*, 119, 108-117.
- [30] J.A. Bogas, J. de Brito, J.M. Figueiredo. (2015). Mechanical characterization of concrete produced with recycled lightweight expanded clay aggregate concrete. *Journal of Cleaner Production*, 89, 187-195.
- [31] L.H. Nguyen, A.-L. Beaucour, S. Ortola, A. Noumowé. (2014). Influence of the volume fraction and the nature of fine lightweight aggregates on the thermal and mechanical properties of structural concrete. *Construction and Building Materials*, 51, 121-132.
- [32] T.H. Wee, M.S. Chin, M.A. Mansur. (1996). Stress-Strain Relationship of High-Strength Concrete in Compression. *Journal of Materials in Civil Engineering*, 8(2), 70-76.
- [33] J.A. Bogas, A. Gomes. (2013). A simple mix design method for structural lightweight aggregate concrete. *Materials and Structures*, 46, 1919-1932.
- [34] J.M. Chi, R. Huang, C.-C. Yang, J.J. Chang. (2003). Effect of aggregate properties on the strength and stiffness of lightweight concrete. *Cement and Concrete Composites*, 25(2), 197-205.
- [35] Y.W. Choi, Y.J. Kim, H.C. Shin, H.Y. Moon. (2006). An experimental research on the fluidity and mechanical properties of high-strength lightweight self-compacting concrete. *Cement and Concrete Research*, 36(9), 1595-1602.
- [36] M. Gesoğlu, T. Özturan, E. Güneyisi. (2007). Effects of fly ash properties on characteristics of cold-bonded fly ash lightweight aggregates. *Construction and Building Materials*, 21(9), 1869-1878.
- [37] E. Güneyisi, M. Gesoğlu, E. Booya. (2012). Fresh properties of self-compacting cold bonded fly ash lightweight aggregate concrete with different mineral admixtures. *Materials and Structures*, 45, 1849-1859.
- [38] O. Kayali. (2008). Fly ash lightweight aggregates in high performance concrete. *Construction and Building Materials*, 22(12), 2393-2399.
- [39] C.-C. Yang, R. Huang. (1998). Approximate strength of lightweight aggregate using micromechanics method. *Advanced Cement Based Materials*, 7(3-4), 133-138.
- [40] L. Breiman. (2001). Random Forest. *Machine Learning*, 45, 5-32.
- [41] J.H. Friedman. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189-1232.
- [42] T. Chen, C. Guestrin. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 2016, pp 785-794.
- [43] Q. Li, Z. Song. (2023). Prediction of compressive strength of rice husk ash concrete based on stacking ensemble learning model. *Journal of Cleaner Production*, 382, 135279.
- [44] H.-B. Ly, E. Monteiro, T.-T. Le, V.M. Le, M. Dal, G. Regnier, B.T. Pham. (2019). Prediction and sensitivity analysis of bubble dissolution time in 3D selective laser sintering using ensemble decision trees. *Materials*, 12(9),

- 1544.
- [45] A. Sharma, M.K. Goyal. (2015). Bayesian network model for monthly rainfall forecast. *In Proceedings of the 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN);, IEEE: Kolkata, India, 2015*, pp 241-246.
- [46] S. Salcedo-Sanz, R.C. Deo, L. Carro-Calvo, B. Saavedra-Moreno. (2016). Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theoretical and Applied Climatology*, 125, 13-25.
- [47] S.M. Lundberg, S.-I. Lee. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems* 30, 4768-4777.