



Tạp chí điện tử
Khoa học và Công nghệ Giao thông
Trang website: <https://jstt.vn/index.php/vn>



Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2025.vn.5.2.99-106>

***Corresponding author:**

Email address:

nghipt@utt.edu.vn

Received: 06/05/2025

Received in Revised Form:
22/06/2025

Accepted: 26/06/2025

Predicting the compressive strength of fly ash concrete using the XGBoost machine learning model

Phung Tang Nghi

University of Transport Technology, 54 Trieu Khuc, Thanh Xuan, Hanoi 100000, Vietnam

Abstract: Fly ash concrete is a type of construction material in which a portion of Portland cement is replaced with fly ash—a byproduct collected from coal combustion in thermal power plants. Using fly ash at an appropriate ratio contributes to a more sustainable concrete mix and offers significant environmental benefits. In this study, the XGBoost machine learning model is applied to predict the compressive strength of fly ash concrete, with the aim of optimizing the mix proportions and ensuring practical performance and quality. The model is trained on a dataset with seven input parameters: cement, sand, coarse aggregate, water, fly ash, admixture, and curing time; while the compressive strength is used as the output target. The model's performance is evaluated using metrics such as mean absolute error (MAE), root mean squared error (RMSE), and the coefficient of determination (R^2). The results show that the XGBoost model is an effective approach for predicting the compressive strength of fly ash concrete, achieving training results of ($R^2 = 0,896$; RMSE = 4,213; MAE = 3,133) and validation results of ($R^2 = 0,868$; RMSE = 4,892; MAE = 3,401).

Keywords: Fly ash concrete; compressive strength; XGBoost; machine learning.



Thông tin bài viết
Dạng bài viết:
Bài báo nghiên cứu

DOI:
<https://doi.org/10.58845/jstt.utt.2025.vn.5.2.99-106>

Tác giả liên hệ:
Địa chỉ Email:
nghipt@utt.edu.vn

Ngày nộp bài: 06/05/2025
Ngày nộp bài sửa: 22/06/2025
Ngày chấp nhận: 26/06/2025

Nghiên cứu dự đoán cường độ chịu nén của bê tông sử dụng tro bay bằng mô hình học máy XGBoost

Phùng Tăng Nghị

Trường Đại học Công nghệ Giao thông vận tải, 54 Triều Khúc, Thanh Xuân, Hà Nội 100000, Việt Nam

Tóm tắt: Bê tông sử dụng tro bay là một loại vật liệu xây dựng được thay thế một phần xi măng Portland bằng tro bay từ phụ phẩm thu được trong quá trình đốt than tại các nhà máy nhiệt điện. Việc sử dụng tro bay với tỷ lệ thích hợp góp phần tạo nên hỗn hợp bê tông bền vững hơn và mang lại những lợi ích đáng kể cho môi trường. Trong nghiên cứu này, mô hình học máy XGBoost được triển khai để dự đoán cường độ chịu nén bê tông sử dụng tro bay nhằm tối ưu tỷ lệ thành phần vật liệu trong cấp phối bê tông, từ đó nâng cao hiệu quả sử dụng và đảm bảo chất lượng bê tông trong điều kiện thực tế. Mô hình được huấn luyện với dữ liệu có 7 tham số đầu vào bao gồm: xi măng, cát, đá, nước, tro bay, phụ gia, thời gian; trong khi cường độ chịu nén được chọn làm mục tiêu đầu ra của mô hình. Việc đánh giá hiệu suất của mô hình được thông qua các chỉ số như sai số tuyệt đối trung bình (MAE), sai số bình phương trung bình (RMSE) và hệ số xác định (R^2). Kết quả cho thấy mô hình học máy XGBoost là phương pháp hiệu quả trong dự đoán cường độ chịu nén bê tông sử dụng tro bay với giá trị huấn luyện là ($R^2 = 0,896$; RMSE = 4,213; MAE = 3,133) và giá trị kiểm chứng là ($R^2 = 0,868$; RMSE=4,892; MAE=3,401).

Từ khóa: Bê tông sử dụng tro bay; cường độ chịu nén; XGBoost; học máy.

1. Đặt vấn đề

Tro bay là sản phẩm phụ trong quá trình đốt than, có các tính chất lý hóa thay đổi đáng kể tùy thuộc vào loại than sử dụng. Các thành phần oxit có trong tro bay như SiO_2 , Al_2O_3 , CaO và Fe_2O_3 chiếm khoảng 90% tổng khối lượng, đóng vai trò quyết định đến tính chất và hiệu quả của tro bay khi sử dụng trong bê tông [1]. Đáng chú ý, hàm lượng silic (SiO_2) và nhôm (Al_2O_3) ảnh hưởng đáng kể đến quá trình hình thành chất kết dính. Trên cơ sở đó, tro bay được sử dụng là vật liệu thay thế một phần xi măng Portland truyền thống, góp phần giảm phát thải khí CO_2 trong quá trình sản xuất đồng thời tận dụng hiệu quả các nguồn phế thải

công nghiệp.

Quá trình thiết kế cấp phối bê tông sử dụng tro bay là một bước quan trọng để xác định tỷ lệ giữa các thành phần vật liệu. Mặc dù các tiêu chuẩn hiện hành cung cấp cơ sở để thiết kế thành phần cấp phối nhưng trong thực tế, việc tối ưu hóa cấp phối để đáp ứng yêu cầu kỹ thuật mất nhiều thời gian và công sức [2]. Trong nghiên cứu thực nghiệm của Joshi (2017), một phần xi măng đã được thay thế bằng tro bay với tỷ lệ từ 10% đến 30%, nhằm đánh giá ảnh hưởng của vật liệu này đến các đặc tính của bê tông sử dụng tro bay như thời gian đông kết, độ bền, khả năng thi công và cường độ chịu nén [3]. Các thử nghiệm về cường

độ chịu nén có tỷ lệ pha trộn tối ưu có thể cải thiện đáng kể cường độ và độ bền, tùy thuộc vào loại tro bay sử dụng. Nghiên cứu của Jatale và cộng sự (2013) đánh giá ảnh hưởng cường độ chịu nén bê tông khi thay thế một phần xi măng bằng tro bay và nhận thấy mức độ phát triển cường độ ở các độ tuổi khác nhau có liên quan đến tỷ lệ nước/xi măng và tỷ lệ tro bay trong hỗn hợp bê tông [4]. Đối với, Wankhede và cộng sự (2014) nghiên cứu tác động của tro bay đối với các đặc tính của bê tông và phát hiện độ sụt của bê tông tăng khi lượng tro bay tăng [5]. Kết quả cho thấy rằng với việc thay thế từ 10% đến 20% xi măng thì cường độ chịu nén tốt hơn sau 28 ngày so với bê tông thông thường, nhưng trong trường hợp thay thế trên 30% xi măng thì cường độ chịu nén tối đa giảm. Các kết quả đạt được cho thấy mức độ hiệu quả của việc sử dụng tro bay phụ thuộc vào sự phối trộn giữa các thành phần vật liệu trong hỗn hợp. Do đó, để tối ưu hóa tỷ lệ cấp phối và giảm thiểu sự phụ thuộc vào các thiết kế thực nghiệm cần xây dựng mô hình tính toán phù hợp để dự báo tỷ lệ cấp phối là rất cần thiết [6].

Trong bối cảnh đó, trí tuệ nhân tạo (AI), đặc biệt là các thuật toán học máy được xem là công cụ hiệu quả để dự báo cường độ chịu nén bê tông tro bay nhờ khả năng học từ dữ liệu [7]. Nhiều nghiên cứu trước đây đã triển khai hiệu quả các mô hình học máy để dự báo một số đặc tính bê tông như độ bền [8], độ sụt [9], độ thấm [10] và cường độ chịu nén [11] với các mô hình học máy đáng chú ý như: cây quyết định tăng cường gradient (XGBoost) [12], CatBoost (CBT) [13], mạng nơron nhân tạo (ANN), máy vector hỗ trợ (SVM), rừng ngẫu nhiên (RF)... Cụ thể, nghiên cứu của Topcu và cộng sự (2008) sử dụng mô hình ANN để dự đoán cường độ chịu nén bê tông chứa tro bay với thời gian 7, 28 và 90 ngày cho độ chính xác cao ($R^2 = 0,96$) [14]. Trong một nghiên cứu khác của Abhilash Gogineni và cộng sự (2024) áp dụng bốn mô hình học máy gồm rừng ngẫu nhiên (RF), máy vector hỗ trợ (SVM), mạng nơron nhân tạo (ANN) và XGBoost để dự đoán cường độ chịu nén bê tông [15]. Kết quả mô hình XGBoost cho hiệu

suất dự đoán vượt trội với $R^2 = 0,997$, cao hơn so với các mô hình còn lại. Sự hiệu quả của các mô hình cho thấy khả năng mô hình hóa tốt các mối quan hệ phi tuyến giữa các tham số đầu vào, đồng thời khai thác thông tin từ dữ liệu thực nghiệm mà không cần giả định trước về dạng hàm toán học.

Gần đây, nhiều mô hình học sâu như mạng nơron tích chập (CNN), mạng nơron sâu (DNN) và bộ nhớ ngắn dài hạn (LSTM) và một số mô hình học sâu khác cho thấy tiềm năng trong dự báo. Chẳng hạn, nghiên cứu của Li và cộng sự (2024) sử dụng 471 bộ dữ liệu bê tông chứa tro bay và thử nghiệm sáu mô hình học sâu cho thấy mô hình lai FCNN+CNN có hiệu suất cao nhất với $R^2 = 0,95$, RMSE = 14,18 và MAE = 2,32 [16]. Tương tự, Biswas và cộng sự (2024) áp dụng mô hình DNN để dự báo cường độ chịu nén bê tông sử dụng tro bay (FA-HSC), thu được kết quả với $R^2 = 0,89$ và RMSE = 0,06 [17]. Thực tế, các mô hình học sâu cho thấy độ chính xác cao nhưng cần lượng lớn dữ liệu để huấn luyện và thời gian huấn luyện kéo dài. Trong khi đó các mô hình học máy như XGBoost, ANN, SVM, RF... vẫn đảm bảo hiệu suất tính toán nhờ khả năng huấn luyện nhanh và linh hoạt. Theo đó nghiên cứu này triển khai mô hình XGBoost để dự đoán cường độ chịu nén bê tông sử dụng tro bay, góp phần tối ưu hóa hiệu quả sử dụng tro bay trong bê tông.

2. Dữ liệu nghiên cứu

Dữ liệu sử dụng trong nghiên cứu được thu thập dựa trên kết quả đã được công bố trên Scientific Reports [18], với 233 dữ liệu được lựa chọn sử dụng tro bay gồm 7 tham số đầu vào: xi măng (XM), cát (C), đá dăm (Đ), nước (N), tro bay (TB), phụ gia (PG), thời gian (T) và tham số đầu ra là cường độ chịu nén (F_c). Các tham số được phân tích thống kê để đánh giá sự biến động và xác định các đặc điểm của tập dữ liệu trong Bảng 1. Theo đó, cường độ chịu nén bê tông tro bay dao động từ 8,49 đến 66,42 MPa, với độ lệch chuẩn là 13,3 MPa cho thấy mức độ phân tán lớn về cường độ của các mẫu bê tông. Ngoài ra, Hình 1 cũng trình bày sự phân bố tần suất của các tham số, qua đó

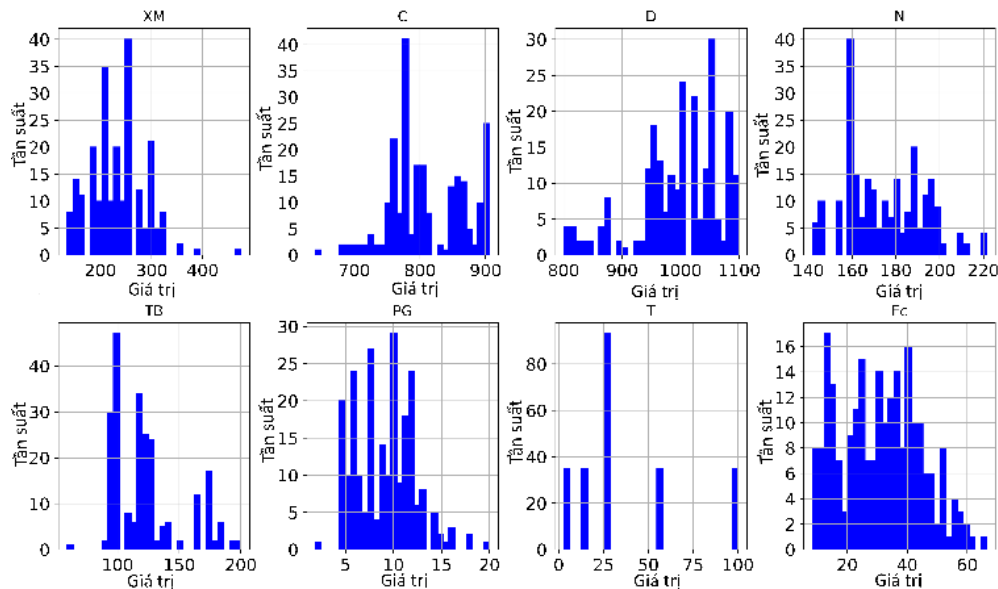
phản ánh rõ đặc điểm phân phối của từng tham số trong tập dữ liệu.

Các tham số trong tập dữ liệu có xu hướng phân bố tập trung tại những khoảng giá trị nhất định thay vì rải đều. Với tham số XM, C, Đ và N, Fc

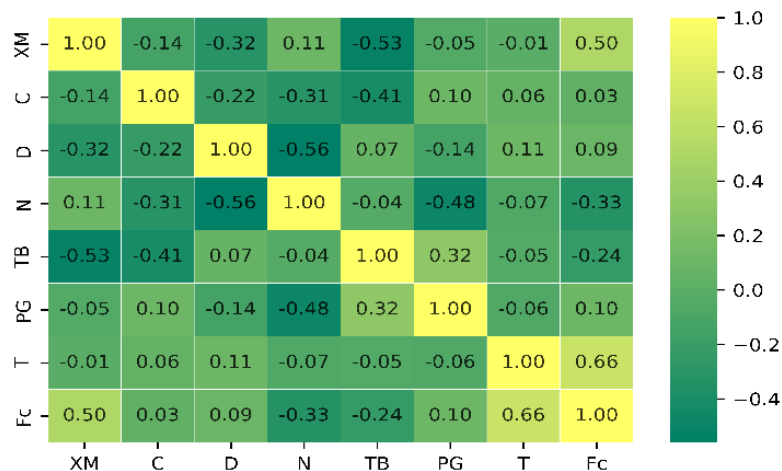
đều có khoảng giá trị giao động lớn trong thành phần cấp phối. Các tham số PG và TB có khoảng dao động hẹp, cho thấy sự kiểm soát chặt chẽ trong sử dụng phụ gia và tro bay. Đặc biệt, tham số T có cụm giá trị phổ biến là 28 ngày.

Bảng 1. Bảng thống kê giá trị

STT	Tham số	Viết tắt	Đơn vị	Giá trị nhỏ nhất	Giá trị lớn nhất	Độ lệch chuẩn (std)
1	Xi măng	XM	kg/m ³	134,7	475	52,075
2	Cát	C	kg/m ³	641	905,9	57,697
3	Đá	Đ	kg/m ³	801	1098	71,067
4	Nước	N	kg/m ³	142	221,4	17,619
5	Tro bay	TB	kg/m ³	59	200,1	27,98
6	Phụ gia	PG	kg/m ³	1,9	20	3,034
7	Thời gian	T	Ngày	3	100	30,657
8	Cường độ chịu nén	Fc	Mpa	8,49	66,42	13,3



Hình 1. Phân bố dữ liệu của tham số



Hình 2. Giá trị tương quan của tham số

Phân tích giá trị tương quan giúp làm rõ mối quan hệ giữa các tham số, trong đó hệ số tương quan Pearson được tính toán và hiển thị cho từng

cặp tham số. Kết quả trong Hình 2 cho thấy, tham số thời gian và cường độ chịu nén có giá trị tương quan cao nhất là 0,66 trong khi các tham số còn lại

có tương quan thấp hơn, đảm bảo tính độc lập giữa các tham số.

3. Phương pháp nghiên cứu

3.1. Tăng cường độ dốc cực đại (XGBoost)

XGBoost được Chen và cộng sự phát triển vào năm 2015 và là một trong những thuật toán tiên tiến trong nhóm boosting, nổi bật nhờ khả năng xử lý hiệu quả dữ liệu lớn và tính ổn định cao trong dự đoán [19]. Mô hình này sử dụng thuật toán tăng cường theo độ dốc để xây dựng từng cây học yếu sao cho mỗi cây mới tối thiểu hóa phần lỗi còn lại của mô hình hiện tại. Hàm mục tiêu trong XGBoost được thiết kế tối ưu, gồm hai phần: (1) hàm mất mát đo lường sai số dự đoán và (2) hàm điều chuẩn (regularization) giúp kiểm soát độ phức tạp của mô hình nhằm giảm thiểu nguy cơ quá khớp [20]. Nhờ cơ chế này, XGBoost không chỉ cải thiện tốt độ chính xác của mô hình mà còn tăng cường khả năng khái quát hóa.

Công thức tổng quát của hàm mục tiêu tại vòng lặp thứ t được viết như sau:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (1)$$

Trong đó, l là hàm mất mát (ví dụ: sai số bình phương hoặc hàm mất mát logistic), y_i là giá trị thực tế, $\hat{y}_i^{(t-1)}$ là giá trị dự đoán ở vòng lặp trước, $f_t(x_i)$ là cây quyết định mới tại vòng lặp t , và $\Omega(f_t)$ là thành phần điều chỉnh độ phức tạp của cây.

3.2. Hiệu suất của mô hình

Hiệu suất của mô hình được đánh giá bằng các chỉ số: hệ số xác định (R^2), sai số bình phương trung bình (RMSE) và sai số tuyệt đối trung bình (MAE). Cụ thể, RMSE và MAE đo sai số giữa giá trị dự đoán và giá trị thực tế, tuy nhiên khác nhau về cách đánh giá độ lớn của sai số [21]. Về nguyên tắc, giá trị RMSE và MAE càng nhỏ thì mô hình càng chính xác. Ngoài ra, hệ số xác định R^2 đánh giá mức độ phù hợp của mô hình, giá trị R^2 nằm trong khoảng từ 0 đến 1, trong đó giá trị càng gần 1 cho thấy mô hình giải thích tốt dữ liệu thực tế, còn giá trị gần 0 cho thấy mô hình giải thích kém.

Các chỉ số này được tính bằng các công thức sau:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_0 - y_p)^2} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_0 - y_p| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_0 - y_p)^2}{\sum_{i=1}^N (y_0 - \bar{y}_i)^2} \quad (4)$$

Trong đó, y_0 và y_p là giá trị đo được thực tế và giá trị của mô hình tương ứng, $\bar{y}_i$ là giá trị trung bình và N là tổng số các dữ liệu đầu vào.

3.3. Lựa chọn đặc trưng

Lựa chọn đặc trưng dựa trên mức độ quan trọng của các tham số đầu vào và là một trong những kỹ thuật phổ biến trong học máy [22]. Phân tích tầm quan trọng không chỉ giúp hiểu rõ hơn về vai trò của từng tham số (chẳng hạn như hàm lượng xi măng, nước, tro bay, phụ gia...), mà còn hỗ trợ loại bỏ các tham số không cần thiết, từ đó giúp giảm chiều dữ liệu, rút ngắn thời gian huấn luyện, và tránh hiện tượng quá khớp. Ngoài ra, lựa chọn đặc trưng hợp lý giúp tăng cường hiệu suất dự đoán vì mô hình tập trung vào các tham số thực sự có ý nghĩa thống kê và kỹ thuật.

3.4. Trình tự thực hiện

Để xây dựng một mô hình học máy có độ chính xác và khả năng khái quát tốt, cần có quy trình thực hiện được triển khai một cách khoa học, hợp lý và tuần tự. Quy trình này thường bao gồm các bước chính như sau:

- Thu thập và tổng hợp dữ liệu đầu vào: Dữ liệu được thu thập từ các thí nghiệm liên bao gồm các tham số đầu vào: xi măng (XM), nước (N), cát (C), đá (Đ), tro bay (TB), phụ gia (PG), và thời gian dưỡng hộ (T), cùng với tham số đầu ra là cường độ chịu nén (F_c).

- Xử lý và phân chia dữ liệu: Dữ liệu được phân tích và chuẩn hóa về khoảng giá trị [0-1]. Sau đó được chia ngẫu nhiên với 70% huấn luyện và 30% kiểm chứng.

- Huấn luyện mô hình: Mô hình XGBoost được sử dụng để huấn luyện và kiểm chứng. Quá trình huấn luyện được thực hiện kết hợp với phương pháp GridSearchCV để tối ưu hóa các

siêu tham số như số lượng cây ($n_estimators = 500$), tốc độ học ($learning_rate = 0,01$), độ sâu cây ($max_depth = 5$).

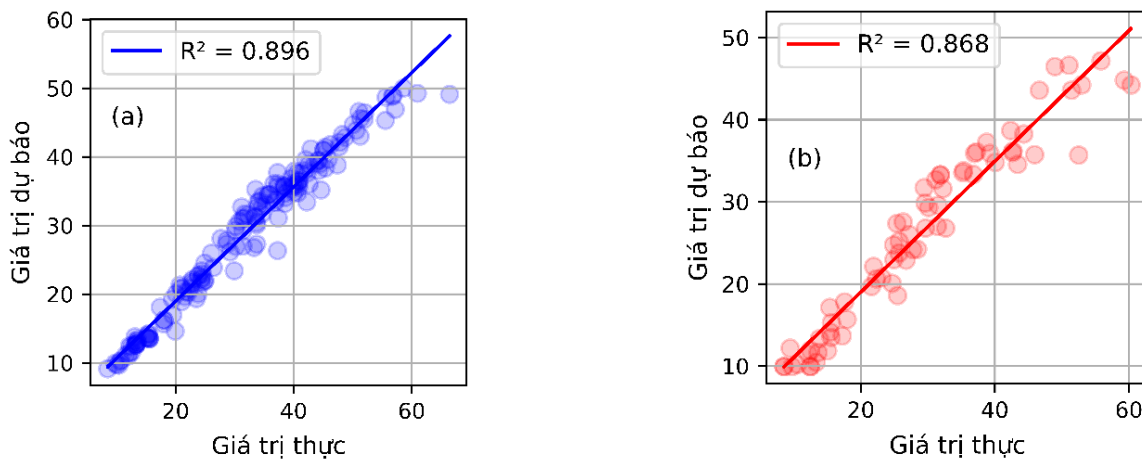
- Phân tích và lựa chọn đặc trưng: Các đặc trưng được phân tích để xác định tham số nào có ảnh hưởng lớn đến kết quả dự báo.

4. Kết quả và thảo luận

Kết quả của mô hình XGBoost được đánh giá dựa trên hệ số xác định (R^2). Hình 3a và Hình 3b lần lượt thể hiện mức độ phân tán dữ liệu huấn luyện và kiểm chứng. Dễ dàng nhận thấy các điểm dữ liệu nằm gần sát với đường hồi quy tuyến tính, cho thấy hiệu suất tốt giữa giá trị dự đoán và thực tế. Cụ thể, hệ số xác định $R^2 = 0,896$ đối với dữ liệu

huấn luyện và 0,868 đối với dữ liệu kiểm chứng, Sự chênh lệch nhỏ giữa hai giá trị này cho thấy mô hình có khả năng khái quát tốt.

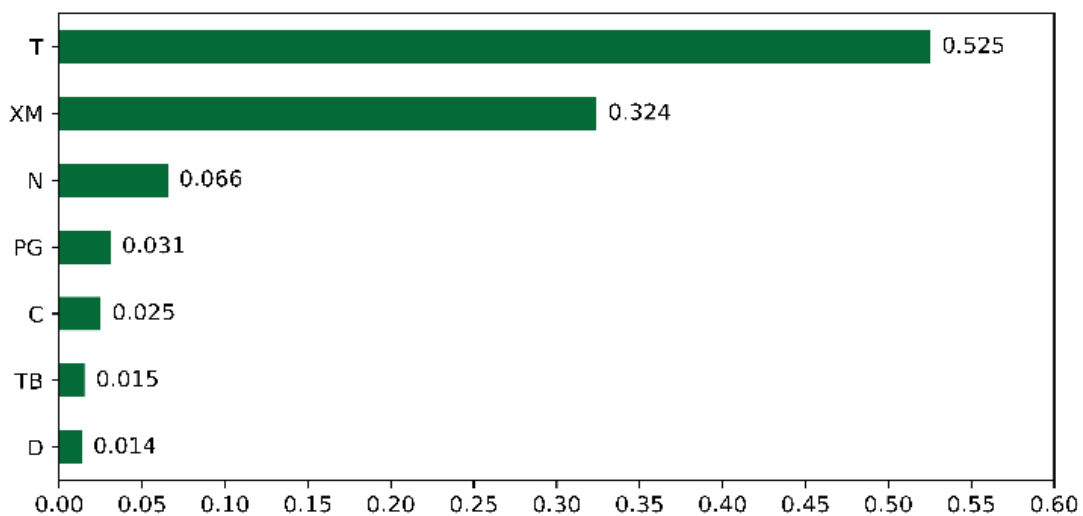
Các chỉ số thống kê như RMSE và MAE được trình bày ở Bảng 2 với giá trị RMSE lần lượt là 4,213 đối với tập huấn luyện và 4,892 với tập kiểm chứng; trong khi đó, giá trị MAE tương ứng là 3,133 và 3,401. Những sai số này đều nằm trong giới hạn, phản ánh sự chênh lệch giữa giá trị huấn luyện và kiểm chứng là không đáng kể. Các kết quả đạt được, cho thấy việc sử dụng mô hình XGBoost vào quá trình thiết kế và dự báo cường độ chịu nén không chỉ khả thi mà còn có tiềm năng áp dụng rộng rãi trong thực tế.



Hình 3. Giá trị R^2 : (a) Giá trị huấn luyện; (b) Giá trị kiểm chứng

Bảng 2. Bảng kết quả sai số của mô hình XGBoost

STT	Tham số	Giá trị huấn luyện	Giá trị kiểm chứng
1	RMSE	4,213	4,892
2	MAE	3,133	3,401



Hình 4. Lựa chọn đặc trưng (%)

Phân tích lựa chọn đặc trưng cho thấy tác động của từng tham số đầu vào đến kết quả dự đoán cường độ nén của bê tông sử dụng tro bay có sự khác nhau trong Hình 4. Kết quả cho thấy rằng tham số thời gian (T) có ảnh hưởng lớn nhất chiếm 52,5%, tiếp theo là xi măng (XM) chiếm 32,4%. Mặt khác, các tham số còn lại đưa ra ít hơn đáng kể so với dự đoán, với nước (N) chiếm 6,6%, phụ gia (PG) chiếm 3,1%, cát (C) chiếm 2,5%, tro bay (TB) chiếm 1,5% và đá (Đ) chiếm 1,4%. Sự khác biệt này không phủ nhận vai trò của tro bay trong hỗn hợp bê tông, mà phản ánh đặc điểm riêng của tham số đầu vào với mục tiêu dự đoán. Trong giai đoạn đầu, phản ứng của tro bay diễn ra chậm hơn so với phản ứng thủy hóa của xi măng. Do đó, khi mô hình dự đoán cường độ chịu nén tại các mốc thời gian ngắn tác động trực tiếp của tro bay chưa thực sự thể hiện rõ rệt.

5. Kết luận

Kết quả nghiên cứu mô hình XGBoost có khả năng dự báo hiệu quả cường độ chịu nén của bê tông sử dụng tro bay, với hệ số xác định R^2 lần lượt đạt 0,896 và 0,868 trên các tập dữ liệu huấn luyện và kiểm chứng. Nghiên cứu có ý nghĩa đối với các kỹ sư và nhà thiết kế đưa ra quyết định nhanh chóng, tiết kiệm thời gian và giảm chi phí thử nghiệm. Tuy nhiên, hạn chế về quy mô dữ liệu và tính đa dạng của các tham số đầu vào cho thấy cần có các nghiên cứu tiếp theo với tập dữ liệu phong phú hơn và kiểm định mô hình trong nhiều điều kiện thực tế khác nhau để nâng cao độ tin cậy trong thực tiễn xây dựng.

Tài liệu tham khảo

- [1]. A. Harmaji, M.C. Kirana, R. Jafari. (2024). Machine learning to predict workability and compressive strength of low-and high-calcium fly ash-based geopolymers. *Crystals*, 14(10), 830. <https://doi.org/10.3390/cryst14100830>
- [2]. B. Chen, L. Wang, Z. Feng, Y. Liu, X. Wu, Y. Qin, L. Xia. (2023). Optimization of high-performance concrete mix ratio design using machine learning. *Engineering Applications of Artificial Intelligence*, 122, 106047. <https://doi.org/10.1016/j.engappai.2023.106047>
- [3]. R. Joshi. (2017). Effect on compressive strength of concrete by partial replacement of cement with fly ash. *International Research Journal of Engineering and Technology*, 4(2), 315-318.
- [4]. A. Jatale, K. Tiwari, S. Khandelwal. (2013). Effects on compressive strength when cement is partially replaced by fly-ash. *IOSR Journal of Mechanical and Civil Engineering*, 5(4), 34-43.
- [5]. P.R. Wankhede, V.A. Fulari. (2014). Effect of fly ash on properties of concrete. *International Journal of Emerging Technology and Advanced Engineering*, 4(7), 284-289.
- [6]. I.-C. Yeh. (2006). Analysis of strength of concrete using design of experiments and neural networks. *Journal of Materials in Civil Engineering*, 18(4), 597-604. [https://doi.org/10.1061/\(ASCE\)0899-1561\(2006\)18:4\(597\)](https://doi.org/10.1061/(ASCE)0899-1561(2006)18:4(597))
- [7]. O.Y. Al-Jarrah, P.D. Yoo, S. Muhaidat, G.K. Karagiannidis, K. Taha. (2015). Efficient machine learning for big data: A review. *Big Data Research*, 2(3), 87-93. <https://doi.org/10.1016/j.bdr.2015.04.001>
- [8]. R. Parichatprecha, P. Nimityongskul. (2009). Analysis of durability of high performance concrete using artificial neural networks. *Construction and Building Materials*, 23(2), 910-917. <https://doi.org/10.1016/j.conbuildmat.2008.04.015>
- [9]. C.-L. Hwang, M.-F. Hung. (2005). Durability design and performance of self-consolidating lightweight concrete. *Construction and Building Materials*, 19(8), 619-626. <https://doi.org/10.1016/j.conbuildmat.2005.01.003>
- [10]. X. Hu, B. Li, Y. Mo, O. Alsawi. (2021). Progress in artificial intelligence-based prediction of concrete performance. *Journal of Advanced Concrete Technology*, 19(8), 924-936. <https://doi.org/10.3151/jact.19.924>
- [11]. S.P. Rajput. (2022). Applying artificial intelligence to predict green concrete

- compressive strength. *Artificial Intelligence for Renewable Energy Systems, Elsevier*, pp 131-149. <https://doi.org/10.1016/B978-0-323-90396-7.00003-1>
- [12]. I.B. Mustapha, Z. Abdulkareem, M. Abdulkareem, A. Ganiyu. (2024). Predictive modeling of physical and mechanical properties of pervious concrete using XGBoost. *Neural Computing and Applications*, 36(16), 9245-9261. <https://doi.org/10.1007/s00521-024-09553-w>
- [13]. A.N. Beskopylny, S.A. Stel'makh, E.M. Shcherban', L.R. Mailyan, B. Meskhi, I. Razveeva, A. Chernil'nik, N. Beskopylny. (2022). Concrete strength prediction using machine learning methods CatBoost, k-nearest neighbors, support vector regression. *Applied Sciences*, 12(21), 10864. <https://doi.org/10.3390/app122110864>
- [14]. İ.B. Topçu, 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. <https://doi.org/10.1016/j.commatsci.2007.04.009>
- [15]. A. Gogineni, I.K. Panday, P. Kumar, R.K. Paswan. (2024). Predicting compressive strength of concrete with fly ash and admixture using XGBoost: a comparative study of machine learning algorithms. *Asian Journal of Civil Engineering*, 25(1), 685-698. <https://doi.org/10.1007/s42107-023-00804-0>
- [16]. H. Li, H. Chung, Z. Li, W. Li. (2024). Compressive Strength Prediction of Fly Ash-Based Concrete Using Single and Hybrid Machine Learning Models. *Buildings*, 14(10), 3299. <https://doi.org/10.3390/buildings14103299>
- [17]. R. Biswas, M. Kumar, D.R. Kumar, P. Samui, Pradeep T, M.K. Rajak, D.J. Armaghani, S. Singh. (2024). Application of novel deep neural network on prediction of compressive strength of fly ash based concrete. *Nondestructive Testing and Evaluation*, 1-31. <https://doi.org/10.1080/10589759.2024.2426703>
- [18]. P. Li, Y. Zhang, J. Gu, S. Duan. (2024). Prediction of compressive strength of concrete based on improved artificial bee colony-multilayer perceptron algorithm. *Scientific Reports*, 14(1), 6414. <https://doi.org/10.1038/s41598-024-57131-w>
- [19]. T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, T. Zhou. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1-4.
- [20]. M. Niazkar, A. Menapace, B. Brentan, R. Piraei, D. Jimenez, P. Dhawan, M. Righetti. (2024). Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). *Environmental Modelling & Software*, 174, 105971. <https://doi.org/10.1016/j.envsoft.2024.105971>
- [21]. C.J. Willmott, K. Matsuura. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82.
- [22]. A. Altmann, L. Toloşi, O. Sander, T. Lengauer. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10), 1340-1347. <https://doi.org/10.1093/bioinformatics/btq134>