

Nghiên cứu ứng dụng trí tuệ nhân tạo trong dự báo cường độ vật liệu xây dựng giao thông ở Việt Nam

■ PGS. TS. ĐÀO VĂN ĐÔNG

Viện Chiến lược và Phát triển Giao thông vận tải

■ TS. PHẠM THÁI BÌNH; TS. TRỊNH HOÀNG SƠN

Trường Đại học Công nghệ Giao thông vận tải

TÓM TẮT: Một trong những yêu cầu kỹ thuật quan trọng của vật liệu xây dựng công trình giao thông là cường độ. Vật liệu trước khi sử dụng cho các công trình giao thông cần phải được kiểm tra đánh giá và xác định cường độ. Các tiêu chuẩn hiện hành thường quy định việc xác định hay dự báo cường độ của vật liệu thông qua các phương pháp thí nghiệm phá hủy, không phá hủy hoặc thông qua các quan hệ lý thuyết được xác lập bằng thực nghiệm. Những phương pháp này thường tốn kinh phí và thời gian nhưng lại có độ chính xác không cao do số lượng thí nghiệm bị hạn chế hoặc khó xác lập được đầy đủ các yếu tố ảnh hưởng đến cường độ của vật liệu. Trí tuệ nhân tạo (AI) là một giải pháp mới với trọng tâm là các thuật toán học máy (ML) có thể thiết lập được các mô hình dự báo các tính chất của vật liệu xây dựng, trong đó có cường độ, với độ chính xác cao, nhanh chóng và không tốn kinh phí. Bài báo giới thiệu những phương pháp học máy đang được sử dụng có hiệu quả cùng nguyên lý sử dụng những phương pháp đó để dự báo cường độ vật liệu xây dựng. Các kiến nghị về giải pháp ứng dụng hiệu quả phương pháp ML cũng được đề cập để trí tuệ nhân tạo trở thành phương pháp dự báo cường độ vật liệu xây dựng có hiệu quả ở Việt Nam.

TỪ KHÓA: Cường độ vật liệu, trí tuệ nhân tạo, học máy.

ABSTRACT: One of the important technical requirements of construction materials is strength. Construction materials before being used need to be tested and evaluated for strength. Current standards generally prescribe the determination or prediction of the strength of materials through destructive, non-destructive testing methods or through experimentally established theoretical relationships. These methods are often costly and time consuming, but often have low accuracy due to the limited number of experiments or the difficulty of fully establishing the factors affecting the strength of the material. Artificial intelligence (AI) is a new solution with a focus on machine learning (ML)

algorithms that can build predictive models of the properties of building materials, including strength, with high accuracy, fast and cost saving. This paper introduces the machine learning methods being used effectively and the principles of using those methods to predict the strength of construction materials. Recommendations on solutions for effective application of the ML methods are also mentioned so that artificial intelligence becomes an effective method of predicting the strength of construction materials in Vietnam.

KEYWORDS: Strength of material, artificial intelligence, machine learning.

1. ĐẶT VẤN ĐỀ

Vật liệu xây dựng giữ vai trò quan trọng cấu thành các kết cấu công trình giao thông. Chất lượng vật liệu xây dựng có vai trò quyết định chính đến chất lượng công trình xây dựng. Trong số các vật liệu xây dựng giao thông thông dụng, các loại vật liệu gồm bê tông xi măng, bê tông asphalt và thép là những vật liệu được sử dụng phổ biến nhất hiện nay khi xây dựng kết cấu hạ tầng giao thông. Cường độ của vật liệu như cường độ chịu nén, cường độ chịu kéo, cường độ chịu kéo khi uốn, cường độ chịu cắt, là những tính chất quan trọng phản ánh chất lượng của vật liệu xây dựng. Cường độ của vật liệu phụ thuộc vào các yếu tố đầu vào chế tạo nên vật liệu như thành phần, cấu trúc, tính chất của nguyên liệu chế tạo cũng như tuổi của chính vật liệu. Quan hệ giữa các yếu tố đầu vào và cường độ của vật liệu thường khá phức tạp và tuân theo các quy luật phi tuyến. Vì vậy, việc dự báo cường độ của vật liệu là khá phức tạp và chỉ mang tính gần đúng.

Các tiêu chuẩn quốc gia và quốc tế hiện nay đều quy định việc dự báo đánh giá hay xác định cường độ của vật liệu xây dựng bằng hai nhóm phương pháp chính, gồm nhóm phương pháp không phá hủy và nhóm phương pháp phá hủy. Nhóm phương pháp không phá hủy như siêu âm, phương pháp độ cứng bề mặt, phương pháp nhỏ bật, phương pháp điện - từ... thường cho kết quả gần đúng do chỉ đánh giá cường độ vật liệu gián tiếp thông qua các

thông số về cấu trúc, mật độ hay độ cứng của vật liệu [1]. Nhóm phương pháp phá hủy, do xác định trực tiếp khả năng chịu lực của vật liệu trên các mẫu thử khi gia tải cho đến khi phá hoại mẫu, nên cho kết quả xác định cường độ chính xác hơn so với nhóm phương pháp không phá hủy. Tuy nhiên, nhóm phương pháp phá hủy thường giới hạn số lượng mẫu thí nghiệm, cần thời gian chờ đủ tuổi của vật liệu khi thí nghiệm và khó xác định được quan hệ giữa các yếu tố đầu vào và cường độ của vật liệu. Vì vậy, một số nghiên cứu đã được tiến hành với mục đích dự báo được giá trị cường độ của vật liệu xây dựng thông qua các phương trình hồi quy đơn giản dựa trên hai hoặc ba tham số ảnh hưởng. Tuy nhiên, với cách tiếp cận này, việc xem xét hai hoặc ba tham số đầu vào chưa phản ánh hết bản chất của bài toán dự báo cường độ bởi vì cường độ của vật liệu phụ thuộc vào rất nhiều các tham số khác nhau. Vì vậy, việc sử dụng cách tiếp cận thông qua các phương trình hồi quy đơn giản với nhiều tham số đầu vào là khó khăn và phức tạp.

Trong những năm gần đây, các kỹ thuật trí tuệ nhân tạo (Artificial Intelligence - AI) và điển hình là các mô hình học máy (Machine Learning - ML) đã và đang được phát triển và ứng dụng rộng rãi trong việc giải quyết rất nhiều bài toán thực tế, trong đó có bài toán dự báo cường độ của vật liệu xây dựng [2, 3, 4]. Theo đó, khi cung cấp được dữ liệu đầu vào kèm các thông số ảnh hưởng, các chương trình học máy sẽ được huấn luyện và có khả năng phân tích, học hỏi và dự báo để đưa ra được các kết quả đầu ra. Cách tiếp cận này được đánh giá là ưu việt so với các cách tiếp cận dự báo truyền thống, cụ thể là: (1) có độ chính xác cao hơn trong dự báo, (2) có thể giải quyết bài toán phức tạp xem xét được nhiều tham số với cơ sở dữ liệu lớn [5]. Tuy nhiên, các kỹ thuật trí tuệ nhân tạo cũng rất nhạy cảm với các sai số từ việc sử dụng dữ liệu và lựa chọn mô hình phù hợp và trong một số bài toán lớn đòi hỏi nguồn lực về máy móc thiết bị và thời gian lớn [5].

Với mục tiêu nghiên cứu và định hướng về việc áp dụng các kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu xây dựng, bài báo này trình bày một số nội dung liên quan đến nguyên lý áp dụng các kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu xây dựng, nhất là vật liệu xây dựng giao thông; tổng hợp và phân tích một số mô hình phổ biến được áp dụng trong và ngoài nước trong việc dự báo cường độ của vật liệu. Trên cơ sở đó đưa ra các kết luận và kiến nghị cho việc áp dụng các kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu nói riêng và định hướng bài toán dự báo tính chất của vật liệu xây dựng nói chung.

2. NGUYÊN LÝ CƠ BẢN ÁP DỤNG CÁC MÔ HÌNH HỌC MÁY TRONG DỰ BÁO CƯỜNG ĐỘ VẬT LIỆU XÂY DỰNG

Bài toán dự báo cường độ của vật liệu xây dựng được coi là bài toán phân tích hồi quy. Trong đó, có hai loại biến số cần quan tâm trong việc xây dựng cơ sở dữ liệu: (1) biến cường độ được coi là biến phụ thuộc (dependent variable) và (2) các tham số ảnh hưởng được coi là các biến độc lập (independent variables). Để áp dụng hiệu quả các mô hình

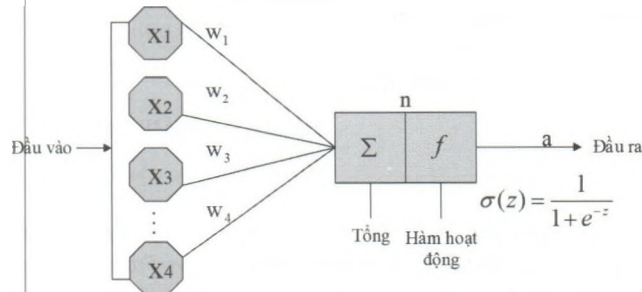
học máy trong dự báo cường độ vật liệu xây dựng, việc xác định và lựa chọn hợp lý các tham số ảnh hưởng đóng vai trò quan trọng, ảnh hưởng tới độ chính xác của mô hình dự báo cường độ. Đối với mỗi loại vật liệu xây dựng các tham số ảnh hưởng này có thể là khác nhau. Đối với bê tông xi măng là loại và hàm lượng của các vật liệu chế tạo, phương pháp dưỡng hộ, tuổi của bê tông [2]. Đối với bê tông asphalt là loại và hàm lượng của các vật liệu chế tạo [3]. Đối với thép là thành phần và hàm lượng các nguyên tố hợp kim cũng như công nghệ chế tạo sản phẩm thép [4]. Nhìn chung, việc lựa chọn các tham số ảnh hưởng phụ thuộc vào từng loại vật liệu và chuyên môn của người tiến hành phân tích dự báo.

Việc xây dựng được một cơ sở dữ liệu đủ lớn để áp dụng các kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu là quan trọng. Về mặt nguyên lý, dữ liệu càng lớn thì các thuật toán sẽ được đào tạo và học càng tốt và cho kết quả dự báo có độ tin cậy càng cao. Tuy nhiên, không có một căn cứ nào để xác định số lượng mẫu như thế nào là đủ lớn. Việc lựa chọn độ lớn của tập dữ liệu phụ thuộc vào nguồn lực của nhóm nghiên cứu có thể tạo ra hoặc được kế thừa. Các dữ liệu về các tham số ảnh hưởng và giá trị cường độ của vật liệu thường được xác định trực tiếp tại phòng thí nghiệm theo các tiêu chuẩn nhất định. Vì vậy, để tạo ra tập dữ liệu đủ lớn cần có các nguồn lực về trang thiết bị, nhân lực và kinh phí để tiến hành thí nghiệm. Một số nghiên cứu sử dụng tập dữ liệu chỉ dưới 100 dữ liệu [9], nhưng có nghiên cứu sử dụng nhiều hơn và có thể lên tới hàng nghìn dữ liệu [4]. Tuy vậy, dữ liệu dù lớn hay nhỏ vẫn phải đảm bảo tiêu chí về chất lượng, độ tin cậy của dữ liệu và tính đại diện của tập dữ liệu dùng để phân tích và mô hình hóa [5].

Một trong những ưu điểm của việc ứng dụng các mô hình trí tuệ nhân tạo là việc kiểm chứng mô hình được thực hiện trên tập dữ liệu kiểm chứng (testing dataset) tách biệt và độc lập với dữ liệu được sử dụng trong việc đào tạo và xây dựng mô hình (training dataset). Vì vậy, việc xác minh hay kiểm chứng độ chính xác của mô hình là khách quan và đảm bảo độ tin cậy. Tuy nhiên, cần phải xác định được tỷ lệ phù hợp trong việc phân chia số lượng mẫu cho các tập dữ liệu đào tạo và tập dữ liệu kiểm chứng. Một số nghiên cứu đã được tiến hành đánh giá hiệu quả của các mô hình trí tuệ nhân tạo sử dụng các tỷ lệ phân chia mẫu khác nhau và kiến nghị rằng tỷ lệ 70/30 hoặc 60/40 hoặc 50/50 cho tập đào tạo/tập kiểm chứng thường cho kết quả dự báo chính xác hơn so với các tỷ lệ khác [5]. Trong đó, tỷ lệ 70/30 được sử dụng phổ biến trong các nghiên cứu áp dụng mô hình trí tuệ nhân tạo trong dự báo tính chất của vật liệu [6].

Khi xây dựng được cơ sở dữ liệu phù hợp, việc áp dụng các kỹ thuật trí tuệ nhân tạo để dự báo cường độ vật liệu xây dựng dựa trên việc phân tích dữ liệu thông qua các hàm, thuật toán của các mô hình học máy để tìm ra quy luật giữa các biến đầu vào và các biến đầu ra, từ đó điều chỉnh và tối ưu hóa các trọng số cho các tham số ảnh hưởng để đạt được quy luật phù hợp nhất với biến đầu ra tương ứng với kết quả độ chính xác của mô hình là cao nhất [5]. Điển hình là đối với thuật toán mạng thần kinh nhân tạo (ANN), sử dụng hàm kích hoạt (activation function) là hàm sigmoid

để phân tích mối liên hệ giữa các biến số đầu vào và biến số đầu ra để đưa ra kết quả dự báo (Hình 2.1).



Hình 2.1: Mô hình neural nhân tạo

Để đánh giá độ chính xác của các mô hình học máy, thông thường các nghiên cứu sử dụng các chỉ số thống kê kê định lượng như các chỉ số hệ số xác định (R^2), giá trị sai số bình phương trung bình gốc (RMSE) và giá trị sai số tuyệt đối trung bình. Giá trị của các chỉ số này được tính theo công thức sau:

$$R^2 = \frac{\sum_{j=1}^N (p_{0,j} - \bar{p}_0)(p_{i,j} - \bar{p}_i)}{\sqrt{\sum_{j=1}^N (p_{0,j} - \bar{p}_0)^2 \sum_{j=1}^N (p_{i,j} - \bar{p}_i)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (p_{0,j} - p_{i,j})^2} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{j=1}^N (p_{0,j} - p_{i,j}) \quad (3)$$

Trong đó: N - Số lượng mẫu trong tập dữ liệu; p_0 và \bar{p}_0 -

Giá trị thực và giá trị trung bình của các mẫu thí nghiệm; p_i và \bar{p}_i - Giá trị dự báo và giá trị dự báo trung bình. Giá trị R^2 càng cao và càng gần 1 thì mô hình càng chính xác. Ngược lại, giá trị RMSE và MAE càng nhỏ và càng gần 0 thì mô hình càng chính xác (Bảng 3.1).

3. CÁC MÔ HÌNH HỌC MÁY ĐƯỢC SỬ DỤNG PHỔ BIẾN TRONG DỰ BÁO CƯỜNG ĐỘ VẬT LIỆU XÂY DỰNG

Hiện nay, các kỹ thuật trí tuệ nhân tạo đã được phát triển rộng rãi, có nhiều các mô hình, thuật toán học máy đã được tạo ra. Các thuật toán này được phát triển và viết trên các ngôn ngữ lập trình khác nhau. Các ứng dụng như Python, Matlab, Weka và các ứng dụng khác đã tích hợp các thuật toán, mô hình trong các hộp/gói công cụ của các ứng dụng này. Việc áp dụng chúng trong việc giải quyết bài toán dự báo cường độ của vật liệu xây dựng không còn là công việc quá khó khăn. Tuy nhiên, bài toán dự báo cường độ vật liệu xây dựng sử dụng trí tuệ nhân tạo là bài toán liên ngành, đòi hỏi người/nhóm phân tích dự báo cần phải có kiến thức chuyên môn về vật liệu xây dựng để thiết kế bài toán sao cho phù hợp và lựa chọn được dữ liệu đảm bảo độ chính xác và tin cậy, vừa phải có kiến thức chuyên môn về công nghệ thông tin để lựa chọn được mô hình một cách phù hợp và lựa chọn đúng, chính xác các tham số của thuật toán khi áp dụng vào mô hình dự báo cường độ của vật liệu xây dựng.

Bảng 3.1 thống kê một số nghiên cứu gần đây sử dụng ML để dự báo các tính chất cường độ điển hình của 3 vật liệu sử dụng phổ biến trong xây dựng giao thông, gồm bê tông xi măng (BTXM), bê tông asphalt (BTAP) và thép.

Bảng 3.1. Một số nghiên cứu dự báo cường độ của vật liệu sử dụng trí tuệ nhân tạo

TT	Loại cường độ để dự báo	Loại vật liệu	Tài liệu tham khảo	Mô hình sử dụng	Số lượng mẫu	Các tham số đầu vào được sử dụng	Kết quả (R^2 , RMSE, MAE)
1	Cường độ chịu nén	BTXM	[7]	Random Forest (RF)	1030	Thành phần cấp phối và tuổi của bê tông	(RF) $R^2 = 0,98$; RMSE = 1,88 MPa
2	Cường độ kéo gián tiếp	BTXM	[8]	Artificial neural networks (ANN) và Fuzzy logic systems (FLS)	210	Thành phần cấp phối và tuổi của bê tông	(ANN) $R^2 = 0,998$; RMSE = 0,1406 MPa (FLS) $R^2 = 0,995$; RMSE = 0,2955 MPa
3	Cường độ chịu nén	BTXM trong môi trường biển	[6]	Support Vector Machines (SVM)	116	Hàm lượng tro bay, xỉ lò cao; nhiệt độ; độ ẩm, ion (Mg^{2+} , SO_4^{2-} , Cl); tuổi ngày	(SVM) $R^2 = 0.96$
4	Thông số Marshall (độ bền, độ dẻo, độ rỗng)	BTAP	[9]	ANN và Least Square-Support Vector Machines (LSSVM)	90	Thành phần cấp phối; tỷ lệ bitum; loại phụ gia	(ANN) $R^2 = 0,876$ (LSSVM) $R^2 = 0,922$

5	Cường độ chịu kéo	Thép	[4]	K-Nearest Neighbors (KNN), RF	5540	Thành phần hóa học và phương pháp luyện thép	(KNN) R ² = 0,742; RMSE = 14,29 MPa; MAE = 10,10 MPa (RF) R ² = 0,775; RMSE = 12,75MPa; MAE = 9,14MPa
---	-------------------	------	-----	-------------------------------	------	----------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------

4. KẾT LUẬN VÀ KIẾN NGHỊ

Các kỹ thuật trí tuệ nhân tạo đã và đang được xem như là cách tiếp cận tiên tiến, có độ chính xác cao được sử dụng trong các bài toán dự báo, trong đó có dự báo cường độ của vật liệu xây dựng. Thông qua việc tổng hợp và phân tích các nghiên cứu dự báo cường độ của vật liệu sử dụng các kỹ thuật trí tuệ nhân tạo khác nhau, bài báo đưa ra các kết luận như sau:

- Việc sử dụng kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu xây dựng là một cách tiếp cận tiên tiến, có độ chính xác cao, có thể dùng để thay thế các cách tiếp cận dự báo truyền thống (phương pháp không phá hủy), tiết kiệm được chi phí và thời gian so với cách tiếp cận thí nghiệm trực tiếp (phương pháp phá hủy). Tuy nhiên, các nghiên cứu cần phải xây dựng được ngân hàng cơ sở dữ liệu đủ lớn, đủ tin cậy và phù hợp để áp dụng hiệu quả các kỹ thuật trí tuệ nhân tạo trong dự báo cường độ của vật liệu xây dựng.

- Khi áp dụng các kỹ thuật trí tuệ nhân tạo, việc lựa chọn và xây dựng được cơ sở dữ liệu phù hợp và chính xác có vai trò quan trọng, ảnh hưởng nhiều tới độ chính xác của mô hình dự báo.

- Việc lựa chọn các kỹ thuật trí tuệ nhân tạo phù hợp với mỗi bộ cơ sở dữ liệu cũng ảnh hưởng tới kết quả của mô hình dự báo.

- Việc sử dụng các mô hình, thuật toán trí tuệ nhân tạo trong việc phát triển các phần mềm, thiết bị dự báo tự động hoặc các hệ thống thông minh là rất có tiềm năng và triển vọng trong lĩnh vực xây dựng nói chung và trong việc dự báo tính chất về cường độ của vật liệu xây dựng giao thông nói riêng.

Tài liệu tham khảo

[1]. V.M. Malhotra and N.J. Carino (2004), *Handbook of Nondestructive testing of concrete*, 2nd edition, CRC Press.

[2]. Wassin Ben Chaabene, Majdi Flash, Moncef L. Nehdi (2020), *Machine learning prediction of mechanical properties of concrete: Critical review*, *Construction and Building Materials*.

[3]. Hamed Majidifard, Behnam Jahangiri, William G. Buttler, Amir H. Alavi (2018), *New machine learning-based prediction models for fracture energy of asphalt mixtures*, *Measurement*.

[4]. Estela Ruiz, Diego Ferreño, Miguel Cuartas, Ana López, Valentín Arroyo and Federico Gutiérrez-Solana (2020), *Machine learning algorithms for the prediction of the strength of steel rods: an example of data-driven manufacturing in steelmaking*, *International Journal of Computer Integrated Manufacturing*.

[5]. Yayin Xu, Ying Zhou, Przemyslaw Sekula, Lieyun Ding (2021), *Machine learning in construction: From shallow to deep learning*, *Developments in the Built Environment*.

[6]. Ling, H., et al. (2019), *Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment*, *Construction and Building Materials*, 206, pp.355-363.

[7]. Trần Văn Quân, Nguyễn Quang Hùng (2020), *Ứng dụng mô hình rừng cây ngẫu nhiên để dự đoán cường độ chịu nén của bê tông*, *Tạp chí KH&CN Thủy lợi*, số 60.

[8]. I.B. Topçu, M. Saridemir (2008), *Prediction of mechanical properties of recycled aggregate concretes containing silica fume using artificial neural networks and fuzzy logic*, *Comput. Mater. Sci*.

[9]. Sunil Khuntia, Aditya Kumar Das, Monika Mohanty and Mahabir Panda (2014), *Prediction of Marshall Parameters of Modified Bituminous Mixtures Using Artificial Intelligence Techniques*, *International Journal of Transportation Science and Technology*.

Ngày nhận bài: 14/12/2021

Ngày chấp nhận đăng: 10/01/2022

**Người phản biện: TS. Nguyễn Ngọc Lân
TS. Bùi Thị Quỳnh Anh**