

ỨNG DỤNG MÔ HÌNH RỪNG CÂY NGẪU NHIÊN ĐỂ DỰ ĐOÁN CƯỜNG ĐỘ CHỊU NÉN CỦA BÊ TÔNG

Trần Văn Quân

Đại học Công nghệ Giao thông vận tải

Nguyễn Quang Hùng

Đại học Thủy lợi

Tóm tắt: Bê tông xi măng là vật liệu được sử dụng rộng rãi nhất hiện nay trong các công trình xây dựng từ cơ sở hạ tầng đến các công trình dân dụng do khả năng chịu nén tốt và giá thành cạnh tranh. Việc tiến hành thí nghiệm xác định cường độ chịu nén của bê tông xi măng đòi hỏi chi phí cao và thời gian thực hiện lâu. Do vậy, việc ứng dụng mô hình rừng cây ngẫu nhiên là một nhánh của trí thông minh nhân tạo vào việc xác định cường độ chịu nén của bê tông xi măng là hết sức có ý nghĩa. Mô hình rừng cây ngẫu nhiên đã được áp dụng vào huấn luyện và kiểm chứng 1030 mẫu cường độ chịu nén của bê tông xi măng. Kết quả dự đoán của mô hình rừng cây ngẫu nhiên cho độ chính xác tương đối cao trong hai trường hợp huấn luyện và kiểm chứng với hệ số tương quan R lần lượt là 0.99 và 0.95. Do vậy, ứng dụng trí thông minh nhân tạo mà cụ thể là mô hình rừng cây ngẫu nhiên vào xác định cường độ chịu nén của bê tông xi măng có sử dụng nhiều loại phụ gia là hoàn toàn khả thi. Mô hình rừng cây ngẫu nhiên còn có thể xác định yếu tố quan trọng nhất ảnh hưởng đến cường độ chịu nén của bê tông xi măng lần lượt là xi măng và độ tuổi của bê tông xi măng, cùng với tro bay là yếu tố ít ảnh hưởng nhất đến cường độ chịu nén của bê tông xi măng.

Từ khóa: Trí tuệ nhân tạo (AI), rừng cây ngẫu nhiên (RF), bê tông, cường độ, dự đoán.

Summary: Cement concrete is the most widely used material today in construction projects from infrastructure to civil construction thanks to its good compressive strength and competitive price. Conducting experiments to determine compressive strength of cement concrete requires high costs and long implementation time. Therefore, the application of the random forest model is a branch of artificial intelligence to determine the compressive strength of cement concrete is very meaningful. Random forest model has been applied to train and test 1030 samples of compressive strength of cement concrete. The predicted results of the random forest model give relatively high accuracy in two training and testing cases with correlation coefficients R respectively 0.99 and 0.95. Therefore, it is feasible to apply artificial intelligence, particularly random forest models, to determine the compressive strength of cement concrete using many additives. The random forest model can also identify the most important factors affecting the compressive strength of cement concrete such as cement and age which is most affecting the compressive strength of cement concrete. The fly ash is the least affected factor on the compressive strength of cement concrete.

Keywords: Artificial Intelligence (AI), Random Forest (RF), Concrete, Compressive Strength, Prediction.

1. ĐẶT VẤN ĐỀ

Bê tông xi măng là vật liệu được sử dụng rộng rãi nhất trong các công trình xây dựng hiện nay từ hạ tầng đến các công trình dân dụng và công nghiệp. Đặc tính ưu việt của bê tông là dễ thi

công, dễ sản xuất giá thành cạnh tranh và đặc biệt là cường độ chịu nén tốt. Do đó hiện nay tại Việt Nam, bê tông xi măng là lựa chọn hàng đầu của các chủ đầu tư trong các công trình xây dựng. Bê tông xi măng là một vật liệu không đồng nhất, là hỗn hợp của xi măng, nước, cốt liệu

Ngày nhận bài: 28/4/2020

Ngày thông qua phản biện: 25/5/2020

Ngày duyệt đăng: 02/6/2020

thô hoặc mịn và cùng với các thành phần phụ gia hóa học như xi lò cao, tro bay, phụ gia hóa dẻo có thể được thêm vào các thành phần bê tông để cải thiện cường độ chịu nén của hỗn hợp bê tông. Thực tế hiện tại, công tác tiến hành xác định cường độ thiết kế, cũng như xác định cấp phối của bê tông xi măng vẫn được tiến hành bằng việc đúc mẫu và xác định cường độ chịu nén của mẫu thông qua các thí nghiệm nén phá hủy. Đối với các cấp phối không sử dụng phụ gia, việc tiến hành thiết kế cấp phối thí nghiệm sẽ đơn giản hơn, tuy nhiên khi sử dụng đến nhiều phụ gia cùng một lúc, ảnh hưởng của các thành phần phụ gia đến cường độ bê tông là bài toán phức tạp do đó cần đúc số lượng mẫu lớn để xác định cường độ cũng như thiết kế cấp phối bê tông xi măng dẫn đến chi phí tăng cao cũng như thời gian dài để xác định cường độ của cấp phối. Do vậy nhằm xác định thành phần cấp phối tối ưu của xi măng, nước, cốt liệu, công cụ số là một giải pháp hữu hiệu có thể ứng dụng.

Sự phát triển của các công cụ số để tìm ra tỷ lệ trộn tối ưu là chủ đề nghiên cứu trong hơn bốn thập kỷ qua [1]. Mục đích của một công cụ số bất kỳ là xác định thành phần vật liệu của mỗi cấp phối cho ra cường độ cấp phối tốt nhất với chi phí hợp lý. Trong những năm gần đây, đã có nhiều mô hình dự báo cường độ nén bê tông được phát triển [2]–[5]. Tuy nhiên các mô hình này áp dụng cho bê tông xi măng với ít thành phần phụ gia và trong giới hạn hàm lượng cấp phối. Miền xác định cường độ của các nghiên cứu với một số lượng nhỏ mẫu thí nghiệm, do vậy khó áp dụng cho nhiều loại bê tông xi măng với nhiều thành phần phụ gia cũng như có sự phân bố lớn các giá trị. Với bê tông xi măng nhiều thành phần phụ gia và có sự phân bố lớn các giá trị cần có một công cụ mô hình số phù hợp để có thể dự đoán và nghiên cứu ảnh hưởng của các thành phần trong cấp phối bê tông xi măng tới cường độ chịu nén.

Thời gian gần đây, mô hình học máy (Machine learning) hay còn gọi là trí thông minh nhân tạo (AI) dựa trên khoa học máy tính đã dần trở nên

phổ biến và được áp dụng trong nhiều lĩnh vực khác nhau. Những ứng dụng rộng rãi của mô hình học máy hay trí thông minh nhân tạo đã được áp dụng trong nhiều lĩnh vực của ngành xây dựng như xác định lực tới hạn của thép với các hình thù khác nhau có nhiều biến số phụ thuộc ảnh hưởng đến lực tới hạn của thép [6], tính chất cơ học của đất [7]. Do đó việc ứng dụng trí thông minh nhân tạo vào xác định cường độ chịu nén của bê tông là hoàn toàn khả thi. Mô hình rừng cây ngẫu nhiên (RF), một nhánh phát triển của trí thông minh nhân tạo đang được áp dụng trong ngành kỹ thuật xây dựng. Do đó, trong bài báo này sẽ sử dụng mô hình rừng cây ngẫu nhiên để dự đoán cường độ chịu nén của bê tông xi măng có sử dụng nhiều loại phụ gia với hàm lượng khác nhau.

2. THIẾT LẬP MÔ HÌNH DỰ BÁO

2.1. Mô hình rừng cây ngẫu nhiên (Random Forest)

Random Forests (RF) chỉ định một họ các phương thức học máy (Machine Learning), bao gồm các thuật toán khác nhau để tạo ra một tập hợp các cây quyết định, như thuật toán Breiman Forest được trình bày bởi Breiman [8] và thường được sử dụng trong tài liệu như một mô hình chuẩn. Bản chất của thuật toán rừng ngẫu nhiên đó là có thể kết hợp được nhiều cây quyết định thay vì chỉ đưa ra lựa chọn dựa trên quyết định của một cây. Từ đó, RF có thể giảm lỗi dự đoán và cải thiện hiệu suất dự đoán. Các bước học tập bao gồm xây dựng một tập hợp các cây quyết định, mỗi nhóm được điều khiển từ một tập hợp con ‘bootstrap, từ tập học ban đầu, tức là sử dụng nguyên tắc đóng bao và sử dụng phương pháp cảm ứng cây gọi là cây ngẫu nhiên. Một thuật toán cảm ứng như vậy, thường dựa trên thuật toán cây phân loại và hồi quy [9].

Phương pháp rừng ngẫu nhiên cho phép học song song từ nhiều cây quyết định được xây dựng và huấn luyện ngẫu nhiên với nhiều tập con chứa các mẫu khác nhau. Mỗi cây trong rừng được huấn luyện bởi một tập học con với dữ liệu được phân phối ngẫu nhiên theo nguyên tắc đóng bao và cũng có các tính năng ngẫu nhiên. Các kết quả cuối cùng được đưa ra dưới

dạng giá trị trung bình của mỗi cây quyết định cho bài toán hồi quy hoặc được xác định bằng kết quả đa số cho bài toán phân loại dữ liệu. Với nhiều ưu điểm của rừng ngẫu nhiên (RF), thuật toán này đã được áp dụng rộng rãi trong nhiều ứng dụng khác nhau [10].

2.2. Đánh giá khả năng dự báo của mô hình

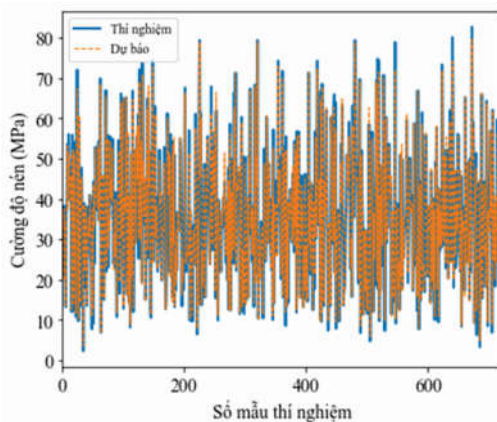
Trong nghiên cứu này, hai tiêu chí được sử dụng là hệ số tương quan (R) (correlation coefficient) và sai số RMSE (Root Mean Square Error) để đánh giá độ chính xác của mô hình rừng cây ngẫu nhiên đã phát triển [6]:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{j=1}^N (p_{0,j} - p_{t,j})^2} \quad (10)$$

$$R = \frac{\sum_{j=1}^N (p_{0,j} - \bar{p}_0)(p_{t,j} - \bar{p}_t)}{\sqrt{\sum_{j=1}^N (p_{0,j} - \bar{p}_0)^2 \sum_{j=1}^N (p_{t,j} - \bar{p}_t)^2}} \quad (11)$$

Trong đó: N là số lượng bộ dữ liệu, p_0 và \bar{p}_0 là giá trị thí nghiệm thực tế và giá trị thí nghiệm thực tế trung bình, p_t và \bar{p}_t là giá trị dự đoán và giá trị dự đoán trung bình, được tính theo mô hình dự báo.

3. DỮ LIỆU THÍ NGHIỆM

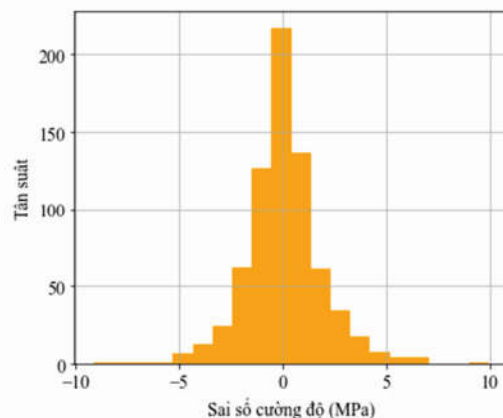


Hình 1: Cường độ nén dự đoán cho phần dữ liệu xây dựng mô hình rừng cây ngẫu nhiên RF

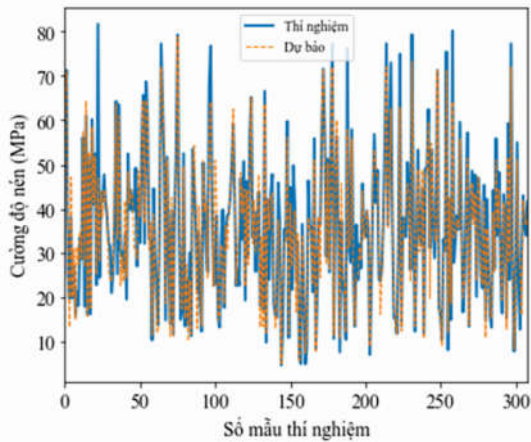
Trong nghiên cứu này, cơ sở dữ liệu về cường độ bê tông xi măng được khai thác từ nghiên cứu của Yeh [10]. Bộ cơ sở dữ liệu bao gồm 1030 dữ liệu về cường độ chịu nén của bê tông xi măng với việc sử dụng xi măng và các loại phụ gia. Các dữ liệu đầu vào của mô hình rừng cây ngẫu nhiên bao gồm 8 dữ liệu đầu vào bao gồm xi măng, xi lò cao, tro bay, nước, phụ gia hóa dẻo, cốt liệu thô, cốt liệu mịn, tuổi bê tông. Dữ liệu đầu ra của mô hình rừng cây ngẫu nhiên lựa chọn là cường độ chịu nén của bê tông xi măng. Bộ dữ liệu được sử dụng trong công việc này được chia ngẫu nhiên thành hai bộ dữ liệu con bằng cách sử dụng phân phối thống nhất, trong đó 70% dữ liệu được sử dụng để xây dựng các mô hình RF và 30% dữ liệu còn lại được dùng trong việc kiểm chứng mô hình đã được xây dựng.

4. KẾT QUẢ

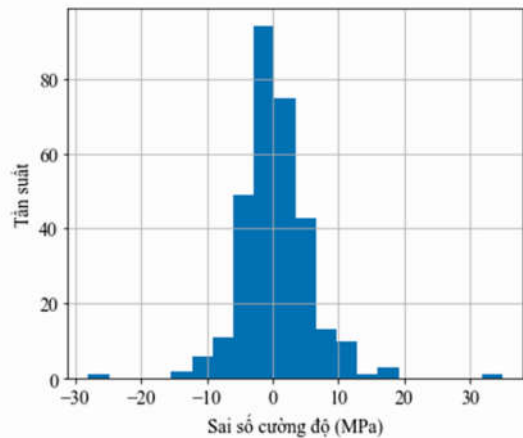
Hình 1 và hình 2 lần lượt là kết quả của cường độ nén dự đoán cho phần dữ liệu xây dựng mô hình và tần suất sai số tương ứng giữa giá trị dự đoán của mô hình và giá trị thực. Hình 3 và hình 4 lần lượt là kết quả của cường độ nén dự đoán cho phần kiểm chứng và tần suất sai số tương ứng giữa giá trị dự đoán của mô hình RF cho kiểm chứng và giá trị thực.



Hình 2: Tần suất sai số giữa cường độ nén dự đoán bởi RF và giá trị thực tế cho phần dữ liệu xây dựng mô hình



Hình 3: Cường độ nén dự đoán cho phần kiểm chứng bởi mô hình rừng cây ngẫu nhiên RF



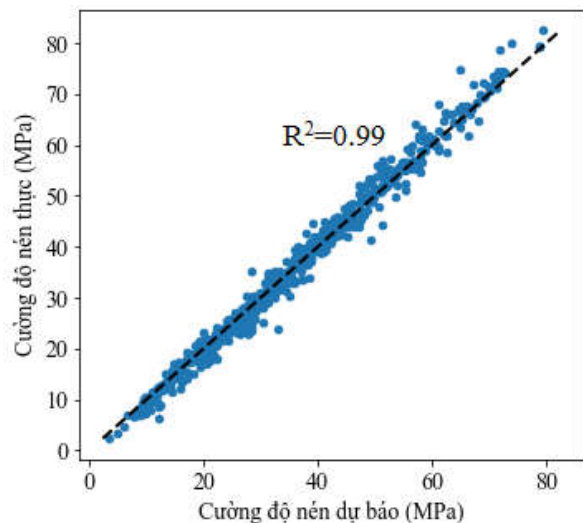
Hình 4: Tần suất sai số giữa cường độ nén dự đoán bởi RF và giá trị thực tế cho phần kiểm chứng

Kết quả cho thấy mô hình rừng cây ngẫu nhiên dự đoán cường độ nén cho khoảng 700 mẫu bê tông với độ chính xác tương đối cao cho phần huấn luyện. Sự sai lệch giá trị dự đoán là rất nhỏ ± 5 MPa, và sai số xấp xỉ 0 MPa với khoảng hơn 500 mẫu, đây là một kết quả rất tốt cho phần xây dựng mô hình rừng cây ngẫu nhiên với sai số RMSE = 1.88 MPa. Do vậy, mô hình RF sau khi được xây dựng thành công sẽ được kiểm chứng bởi khoảng hơn 300 mẫu cường độ nén còn lại. Kết quả trong hình 3, 4 cho thấy mô hình RF sau khi đã được huấn luyện cho kết quả tương đối tốt với khoảng 300 mẫu cường độ nén còn lại. Sai số trong mô hình kiểm chứng tập chung chủ yếu trong khoảng ± 10 MPa với sai số RMSE = 5.45

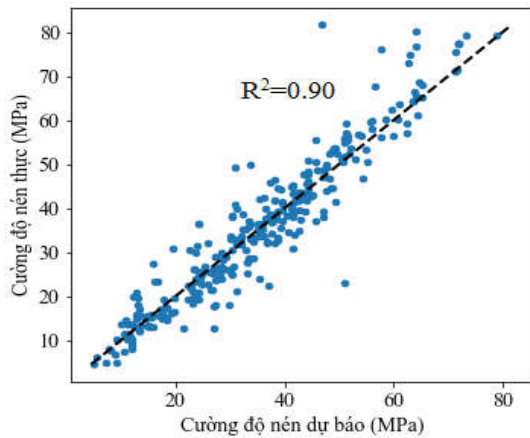
Mô hình hồi quy cho 2 phần xây dựng mô hình và kiểm chứng được thể hiện ở hình 6 và hình 7. Từ hình trên ta thấy khả năng dự báo của mô hình là tương đối cao khá sát với cường độ chịu nén thực tế, tuy nhiên vẫn có sự sai số với các cường độ chịu nén lớn. Giá trị tương quan R thu được cho phần xây dựng mô hình là $R=0.99$ còn phần kiểm chứng là $R=0.95$. Điều này cho thấy việc áp dụng mô hình rừng cây ngẫu nhiên RF cho việc dự báo cường độ chịu nén của bê tông là rất khả thi.

Tuy nhiên, hệ số tương quan R dành cho phần kiểm chứng chưa thực sự cao như phần dữ liệu dùng để xây dựng mô hình.

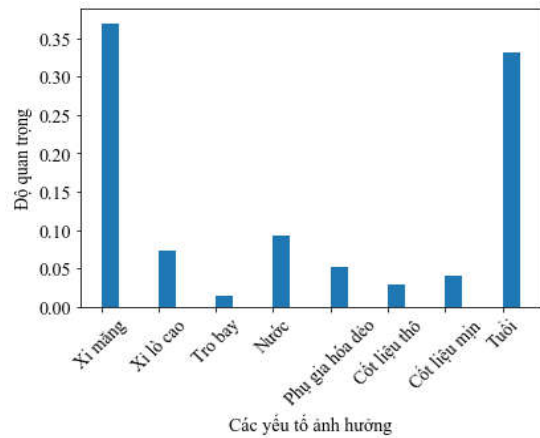
Đặc biệt bằng vào mô hình rừng cây ngẫu nhiên RF, vai trò quan trọng của từng yếu tố đến cường độ chịu nén của bê tông cũng được biểu diễn trong hình 7. Mức độ ảnh hưởng của các tham số tới cường độ của bê tông có thể được xếp lần lượt như hàm lượng xi măng, tuổi bê tông, lượng nước....



Hình 5: Kết quả hồi quy mạng ANN cho phần xây dựng mô hình



Hình 6: Kết quả hồi quy mô hình rừng cây ngẫu nhiên RF cho phần kiểm chứng



Hình 7: Các yếu tố ảnh hưởng tới cường độ chịu nén của bê tông phân tích bằng mô hình rừng cây ngẫu nhiên RF

5. KẾT LUẬN VÀ KIẾN NGHỊ

Bài báo đã giới thiệu mô hình rừng cây ngẫu nhiên RF là một trong những mô hình của trí thông minh nhân tạo (AI) trong việc xác định cường độ chịu nén của bê tông xi măng có sử dụng đến nhiều loại phụ gia cùng với các độ tuổi khác nhau của bê tông. Dữ liệu dùng cho mô phỏng được tập hợp từ kết quả thí nghiệm đủ độ tin cậy. Kết quả cho thấy rõ mô hình rừng cây ngẫu nhiên khả thi trong việc xác

định cường độ chịu nén của bê tông với hệ số tương quan cho mô hình là 0.99 và kiểm chứng là 0.95. Sai số của mô hình và áp dụng cho kiểm chứng là rất nhỏ chủ yếu tập chung vào khoảng sai số từ ± 5 MPa. Ngoài ra, mô hình rừng cây ngẫu nhiên có thể chỉ ra tầm quan trọng của từng thành phần cấp phối, với xi măng và tuổi của bê tông có tầm ảnh hưởng nhất đến cường độ chịu nén của bê tông, ngược lại hàm lượng tro sử dụng ảnh hưởng nhỏ nhất đến cường độ chịu nén của bê tông.

TÀI LIỆU THAM KHẢO

- [1] P. Chopra, R. K. Sharma, and M. Kumar, 'Prediction of Compressive Strength of Concrete Using Artificial Neural Network and Genetic Programming', *Advances in Materials Science and Engineering*, vol. 2016, pp. 1–10, 2016, doi: 10.1155/2016/7648467.
- [2] G. Cusatis, A. Mencarelli, D. Pelessone, and J. Baylot, 'Lattice Discrete Particle Model (LDPM) for failure behavior of concrete. II: Calibration and validation', *Cement and Concrete Composites*, vol. 33, no. 9, pp. 891–905, Oct. 2011, doi: 10.1016/j.cemconcomp.2011.02.010.
- [3] G. Cusatis, D. Pelessone, and A. Mencarelli, 'Lattice Discrete Particle Model (LDPM) for failure behavior of concrete. I: Theory', *Cement and Concrete Composites*, vol. 33, no. 9, pp. 881–890, Oct. 2011, doi: 10.1016/j.cemconcomp.2011.02.011.
- [4] K. Miled, O. Limam, and K. Sab, 'A probabilistic mechanical model for prediction of aggregates' size distribution effect on concrete compressive strength', *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 12, pp. 3366–3378, Jun. 2012, doi: 10.1016/j.physa.2012.01.051.

- [5] A. Naija, K. Miled, and O. Limam, ‘A discrete micromechanical model for predicting HSC compressive strength based on a yield design approach’, *Construction and Building Materials*, vol. 175, pp. 714–725, Jun. 2018, doi: 10.1016/j.conbuildmat.2018.04.197.
- [6] H.-B. Ly *et al.*, ‘Hybrid Artificial Intelligence Approaches for Predicting Critical Buckling Load of Structural Members under Compression Considering the Influence of Initial Geometric Imperfections’, *Applied Sciences*, vol. 9, no. 11, p. 2258, Jan. 2019, doi: 10.3390/app9112258.
- [7] S. Kiran, B. Lal, and S. S. Tripathy, ‘Shear Strength Prediction of Soil based on Probabilistic Neural Network’, *Indian Journal of Science and Technology*, vol. 9, no. 41, Nov. 2016, doi: 10.17485/ijst/2016/v9i41/99188.
- [8] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Taylor & Francis, 1984.
- [9] L. Breiman, *Classification and Regression Trees*. Routledge, 2017.
- [10] I.-C. Yeh, ‘Modeling of strength of high-performance concrete using artificial neural networks’, *Cement and Concrete Research*, vol. 28, no. 12, pp. 1797–1808, Dec. 1998, doi: 10.1016/S0008-8846(98)00165-3.