

Dự báo tuổi thọ còn lại của công trình cầu dầm giản đơn bê tông cốt thép dựa trên các mô hình học máy

Predicting the remaining life of reinforced concrete simple girder bridges based on machine learning models

Lê Bá Anh^{1,*}, Hồ Anh Cường¹, Vũ Phương Thảo¹, Nguyễn Trọng Dũng²

¹Trường Đại học Giao thông vận tải

²Trường Đại học Mỏ - Địa chất

*Email liên hệ: baanh.le@utc.edu.vn

Tóm tắt:

Bài báo đã giới thiệu cơ sở dữ liệu về hiện trạng các công trình cầu đồng thời giới thiệu các mô hình học máy phổ biến hiện nay. Hai mô hình điển hình được lựa chọn là eXtreme Gradient Boosting (XGB) và máy vector hỗ trợ (Support Vector Machine – SVM). Nhóm tác giả đã so sánh các mô hình xây dựng được với mô hình trí tuệ nhân tạo đã công bố trước đây để tìm ra mô hình phù hợp nhằm xử lý dữ liệu các công trình cầu dầm giản đơn bê tông cốt thép tại Việt Nam. Mô hình này cũng được đưa ra kiểm chứng bằng cách dự báo tuổi thọ còn lại của một công trình cầu tại Việt Nam.

Từ khóa: Machine learning; XGB; SVM; Tối ưu hóa bầy đàn.

Abstract:

The paper introduced the database on the current state of bridge and introduced popular machine learning models. The two typical models selected are eXtreme Gradient Boosting (XGB) and Support Vector Machine (SVM). The authors compared the built-in models with the previously published artificial intelligence models to find a suitable model to process data of reinforced concrete simple girder bridges in Vietnam. This model is also tested by predicting the remaining life of a bridge in Vietnam.

Keywords: Machine learning; XGB; SVM; Particle Swarm Optimization.

1. Giới thiệu

Công trình cầu đóng một vai trò quan trọng trong kiến trúc của cơ sở hạ tầng Việt Nam, có tổng số 11.261 công trình cầu trên các tuyến đường quốc lộ và địa phương. Theo số liệu thống kê, số lượng các công trình cầu bị hư hỏng chiếm khoảng 9,1% tổng số cầu của quốc gia. Bên cạnh đó, gần 40% số công trình cầu đã trên 50 tuổi, đây là tuổi thọ thiết kế trung bình của một công trình cầu. Để xếp đặt thứ tự ưu tiên trong việc bảo trì và sửa chữa các công trình cầu, việc cần thiết là phải xây dựng một kế hoạch bảo trì dựa trên mô hình hư hỏng của các công trình cầu.

Ở trên thế giới, nhất là ở Hoa Kỳ đã có nhiều mô hình đã được phát triển bằng cách sử dụng hồ

sơ lưu trữ từ National Bridge Inventory (NBI). Theo NBI, trạng thái của kết cấu được đánh giá theo thang điểm từ 0 đến 9, với 9 là điều kiện hoàn hảo và 4 trở xuống cho thấy tình trạng kém. Một công trình cầu bị đánh giá là hư hỏng nếu mặt cầu, kết cấu phần trên hoặc kết cấu phần dưới bị đánh giá là 4 hoặc thấp hơn.

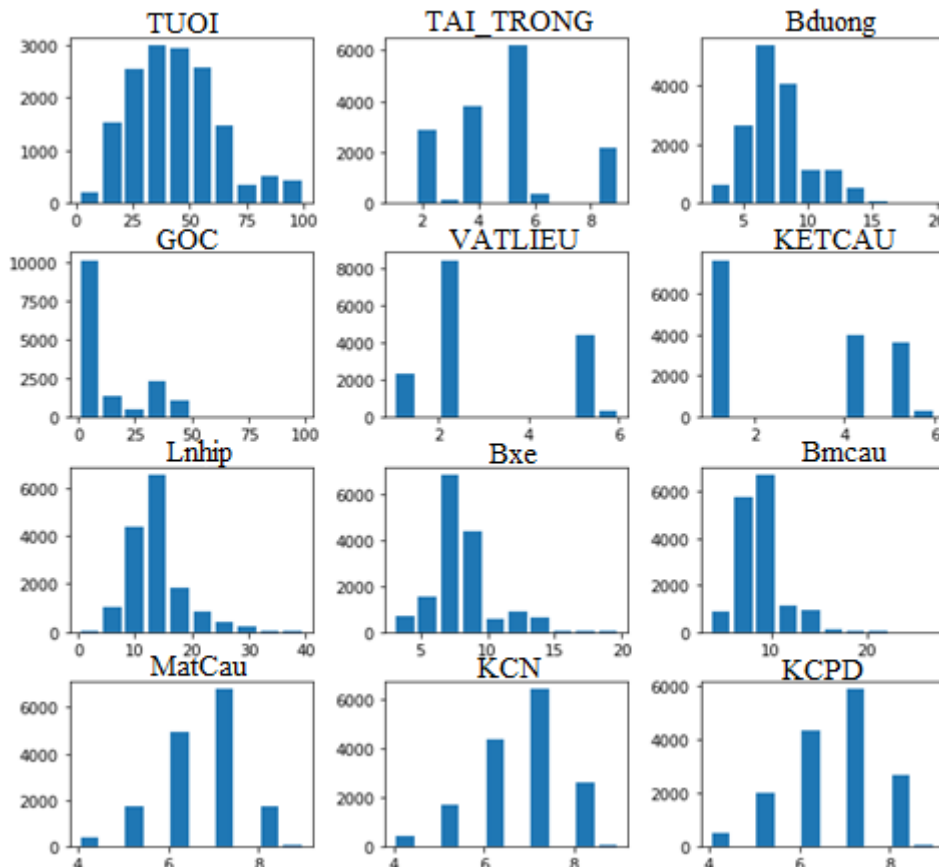
Huang [1] đã phát triển mô hình dựa trên mạng neuron nhân tạo (ANN) để dự đoán sự xuống cấp của bản mặt cầu dựa trên hồ sơ kiểm tra các công trình cầu Wisconsin. Nghiên cứu đã xác định ra 11 yếu tố quan trọng bao gồm: Tuổi, tải trọng thiết kế, lịch sử bảo trì, chiều dài và điều kiện trước đó để dự đoán xếp hạng điều kiện mặt cầu. Mô hình ANN được chứng minh là hoạt

động tốt khi mô hình sự suy giảm của bản mặt cầu theo các biến độc lập. Althaqafi và nhóm tác giả [2] cũng đã chứng minh rằng ANN không chỉ áp dụng tốt cho bản mặt cầu mà cũng có thể sử dụng để áp dụng cho các bộ phận khác của công trình cầu như kết cấu nhịp, kết cấu phần dưới. Bên cạnh ANN, nhiều mô hình học máy khác như máy vector hỗ trợ (SVM), eXtreme Gradient Boosting (XGB) cũng đã được chứng minh cho sự phù hợp với dạng dữ liệu này. Việc áp dụng mô hình trí tuệ nhân tạo để dự báo tuổi thọ còn lại của các công trình cầu tại Việt Nam là một vấn đề đã được đề cập nhiều trong thời gian gần đây. Trong bài báo này, nhóm tác giả sẽ sử dụng các mô hình học máy phổ biến hiện nay như SVM, XGB, so sánh với mô hình ANN dùng để dự báo tuổi thọ còn lại của công trình cầu dầm giản đơn bê tông cốt thép tại Việt Nam.

2. Xây dựng cơ sở dữ liệu

Trong nghiên cứu này, các tác giả đã tiến hành lựa chọn dữ liệu của các công trình cầu dầm giản đơn bê tông cốt thép trong cơ sở dữ liệu NBI. Mỗi công trình cầu bao gồm thông tin:

- (i) Các thông số hình học và thiết kế của công trình cầu như chiều dài nhịp, chiều rộng mặt cầu, loại vật liệu, hồ sơ thiết kế và tải trọng thiết kế;
- (ii) Các điều kiện khai thác như Lưu lượng giao thông hàng ngày trung bình (ADT), độ tuổi và phân loại;
- (iii) Hiện trạng của các công trình cầu. Biểu đồ tần suất của từng tham số đầu vào và tham số đầu ra được biểu diễn trên hình 1.



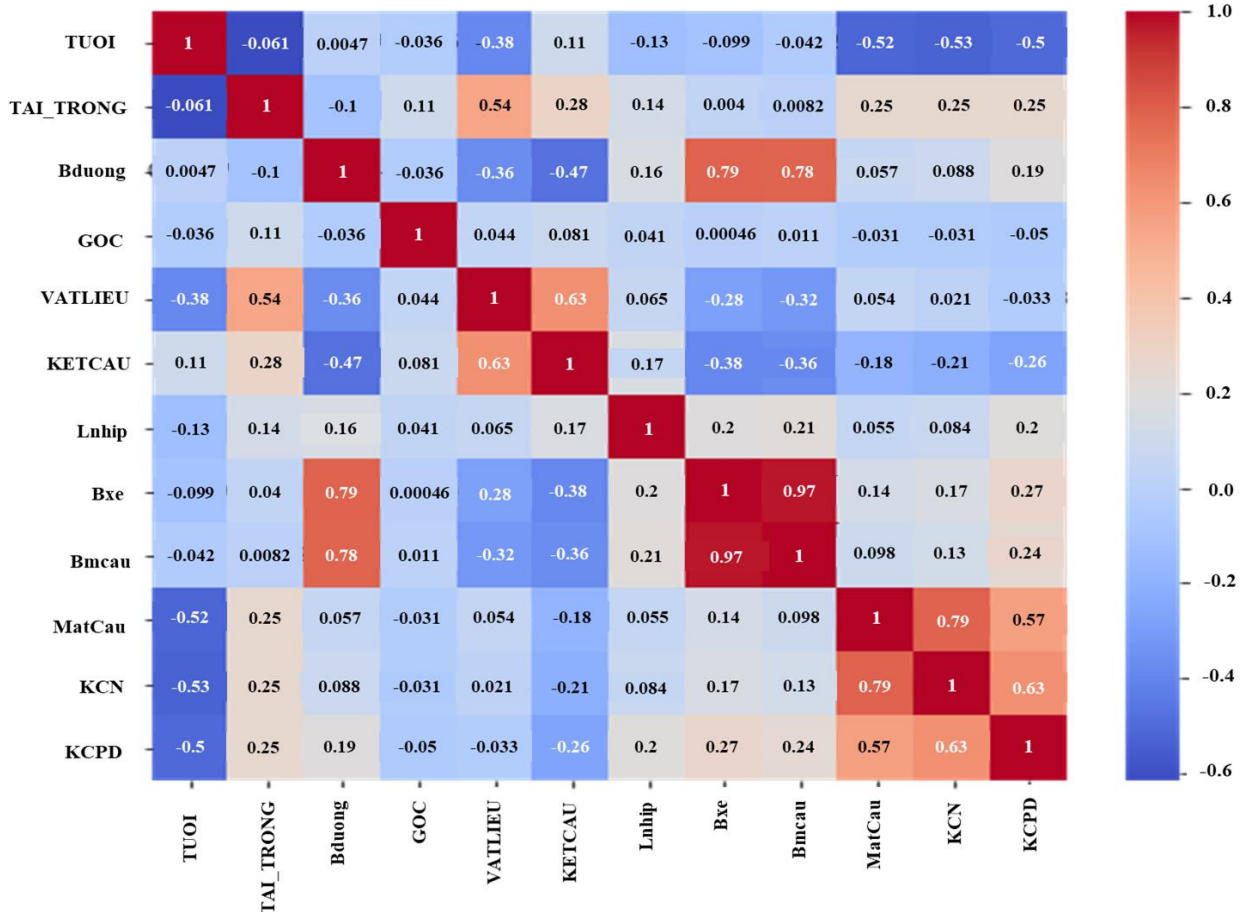
Hình 1. Biểu đồ tần suất của từng tham số Tuổi thọ, Tải trọng, Bề rộng mặt đường, góc nghiêng, vật liệu, loại kết cấu, chiều dài nhịp, bề rộng làn xe chạy, bề rộng toàn cầu, trạng thái của mặt cầu, trạng thái của kết cấu nhịp, trạng thái của kết cấu phần dưới.

Trong bộ dữ liệu thu thập được bao gồm 15.492 công trình cầu có tuổi thọ từ 0 đến 100 tuổi, vật liệu

thiết kế là bê tông cốt thép thường hoặc bê tông cốt thép dự ứng lực; kết cấu nhịp dầm giản đơn; số làn

xe chạy từ 1 đến 6 lần; lưu lượng xe ngày đêm nằm trong khoảng từ 0 đến 11.913 xe; tải trọng thiết kế là các tải trọng tiêu chuẩn; bề rộng phần xe chạy trên cầu từ 3 đến 19,8 m và với bề rộng toàn cầu từ 3,6 đến 27 m. Trong tập dữ liệu này, có thể loại bỏ những công trình cầu có giá trị nhỏ hơn 3 do số lượng những công trình cầu quá ít để có thể đưa vào

dữ liệu. Mỗi tương quan giữa các yếu tố của dữ liệu được thể hiện trong hình 2. Theo như biểu đồ nhiệt thấy rõ các yếu tố như tuổi thọ, lưu lượng xe, tải trọng thiết kế và khai thác, số lượng dầm, chiều dài nhịp,... ảnh hưởng nhiều đến trạng thái hiện tại của công trình trong khi góc nghiêng của cầu ít ảnh hưởng đến trạng thái hiện tại.



Hình 2. Biểu đồ nhiệt thể hiện mối tương quan giữa các hệ số của dữ liệu.

Theo kết quả nghiên cứu của Witten và cộng sự [3], các dữ liệu đã được chuẩn hóa nằm trong khoảng [0,1] để giảm thiểu các sai số trong quá trình học tập của mô hình. Tập dữ liệu sau khi chuẩn hóa được phân chia ngẫu nhiên thành tập huấn luyện và tập kiểm chứng với tỉ lệ là 80% và 20% (tương ứng với 12.393 mẫu và 3.099 mẫu). Quá trình này đảm bảo giai đoạn đào tạo các mô hình AI có thể được thực hiện với khả năng khái quát hóa cao nhất.

Để đánh giá độ chính xác của mô hình, có thể sử dụng hệ số độ chính xác:

$$Acc = \frac{Số\ dự\ báo\ chính\ xác}{Số\ lần\ dự\ báo} \quad (1)$$

Giá trị đo độ chính xác này nằm trong khoảng từ 0 đến 1 với các giá trị càng tiến tới 1 thì độ chính xác càng cao.

3. Các phương pháp học máy

Hiện nay, có rất nhiều phương pháp học máy đã được nghiên cứu và phát triển phù hợp với tập dữ liệu đã được giới thiệu ở trên. Điển hình trong số đó là: ANN, SVM và mô hình cây quyết định. Các mô hình này được lựa chọn để áp dụng cho tập dữ liệu này với các nguyên nhân sau: ANN, SVM và XGB thuộc về ba mảng khác nhau của các phương pháp học máy: Mạng trí tuệ nhân tạo, mô hình kernel và mô hình cây. Trong nhiều mô hình học máy, XGB đã được chứng minh là mô hình tốt nhất hiện nay

dành cho các dữ liệu dạng bảng khi mô hình này chiến thắng hầu hết các cuộc thi về mô hình học máy đối với dạng dữ liệu này. Trong bài báo này tác giả áp dụng hai mô hình SVM và XGB kết hợp với kết quả của mô hình ANN đã được công bố để từ đó lựa chọn ra mô hình tối ưu.

3.1. Giải thuật eXtreme Gradient Boosting

Giải thuật XGB là giải thuật được giới thiệu bởi Chen và Guestrin [4] là một thuật toán được phát triển căn cứ vào mô hình cây quyết định. Giải thuật này dự đoán kết quả dựa trên các dữ liệu đầu vào như sau:

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

Trong đó y_i là kết quả dự báo là kết quả dự đoán 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:

$$obj = \gamma T + \sum_{j=1}^T \left[G_j \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] \quad (3)$$

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

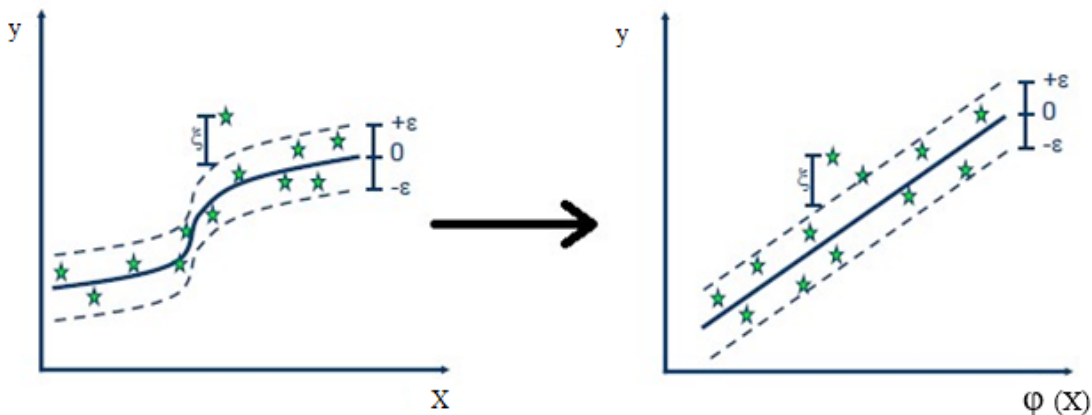
hóa kiểm soát tính đơn giản của cấu trúc cây để tránh khớp quá mức. Các tham số G_j và H_j lần lượt là tổng của các mẫu được liên kết với lá thứ j của độ dốc thứ nhất và thứ hai của hàm mất mát. Cây thứ k được xây dựng bằng cách tách các lá bắt đầu từ một lá. Thay đổi mục tiêu sau khi thêm phân tách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

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

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 thì việc tách được chấp nhận.

3.2. Máy vector hỗ trợ

Thuật toán SVM [5] thuộc nhóm phương pháp học máy kernel. Phương pháp này đã được áp dụng thành công cho nhiều vấn đề thực tế trong cả hai lĩnh vực hồi quy và phân loại. SVM dạng chuẩn nhận dữ liệu đầu vào và phân loại chúng thành hai lớp khác nhau. Để làm như vậy, tập dữ liệu ban đầu được chuyển đổi thành một không gian đa chiều mới, nơi tồn tại siêu phẳng hỗ trợ. Việc chuyển đổi như vậy được thực hiện với sự trợ giúp của hàm kernel tính toán khoảng cách giữa mỗi cặp điểm dữ liệu trong không gian được chuyển đổi. Siêu phẳng được tối ưu hóa bằng cách tối đa hóa khoảng cách của nó đến điểm dữ liệu được chuyển đổi gần nhất.



Hình 3. SVM phi tuyến.

Đối với SVM phi tuyến tính (hình 3), các hàm kernel như hàm đa thức (5) và hàm Cơ sở xuyên

tâm Gaussian (6) biến đổi dữ liệu thành không gian đặc trưng nhiều chiều hơn để có thể thực hiện phân tách tuyến tính (7) và (8).

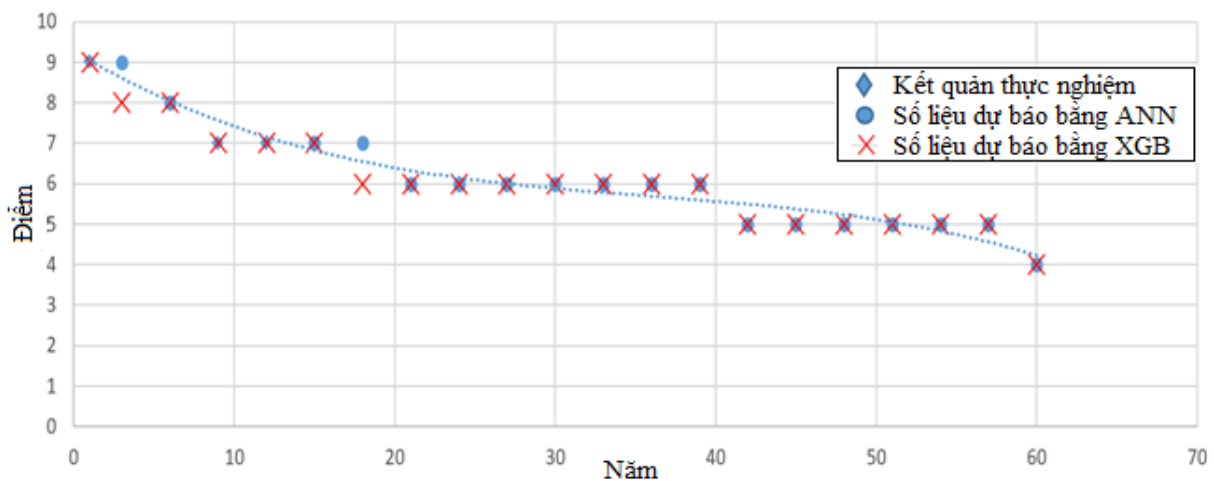
Có thể thấy rằng, mô hình được xây dựng bằng phương pháp XGB đạt được độ chính xác cao nhất so với các mô hình còn lại.

4. Áp dụng mô hình vào việc dự báo tuổi thọ còn lại của công trình cầu dầm giản đơn bê tông cốt thép tại Việt Nam

Để kiểm chứng mô hình mạng xây dựng, cần đánh giá công trình cầu Đồng Niên nằm trên Quốc lộ 5, km 51+ 138. Kết cấu nhịp bao gồm cả hai dạng bê tông cốt thép thường và có dạng mặt cắt ngang chữ T. Theo số liệu thu thập được, cầu được kiểm định lần thứ 1 năm 2008, công trình cầu ở trạng thái tốt, không có bất kỳ dấu hiệu hư hỏng. Đánh giá điểm theo thang NBI là 9.

Năm 2018, kiểm tra hiện trạng cầu thấy xuất hiện hiện tượng như sau: Hiện trạng cầu xuất hiện nhiều vết nứt trên thân trụ T6, T7, gối cầu bị nghiêng lệch. Đánh giá điểm theo thang NBI là 7 điểm. Năm 2020, kiểm định cầu lần thứ 2. Các vết nứt xuất hiện trên trụ, dầm chủ không phát triển so với kết quả theo dõi từ năm 2018. Đánh giá điểm theo thang NBI là 7 điểm.

Với dữ liệu đầu vào chỉ có 3 điểm, việc đánh giá mức độ suy giảm và tuổi thọ còn lại của công trình cầu là tương đối khó khăn. Tuy nhiên khi kết hợp với dữ liệu của các công trình cầu khác và dự báo bằng các phương pháp học máy, chúng ta có thể đưa ra đường cong suy giảm trạng thái của kết cấu được trình bày trong hình 5.



Hình 5. Đường cong suy giảm của kết cấu nhịp cầu Đồng Niên.

Dựa vào kết quả khảo sát của 3 lần kiểm định và dự báo của đường cong suy giảm, có thể dự báo rằng tuổi thọ công trình cầu còn lại khoảng 50 năm. Kết quả trên hình 5 xuất hiện một số sai khác trong kết quả dự báo của mô hình ANN và mô hình XGB. Điều này là do cả hai mô hình đều có sai số dẫn đến các sự sai lệch trong quá trình dự báo.

5. Kết luận

Bài báo đã sử dụng các phương pháp học máy kết hợp với tập dữ liệu các công trình cầu dầm giản đơn bê tông cốt thép thế giới để xây dựng mô hình dự báo trạng thái của kết cấu cầu. Kết quả thu được là mô hình quan hệ giữa hiện trạng của công trình cầu với các yếu tố như chiều dài cầu, tuổi thọ, bề rộng cầu... Mô hình có thể xây dựng được đường cong suy giảm của kết cấu và từ đó có thể xác định tuổi thọ

còn lại của một công trình cầu. Nhóm tác giả cũng đã kiểm chứng mô hình với một công trình cầu tại Việt Nam.

Lời cảm ơn

Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo trong đề tài mã số “B2021-GHA-03”: Nghiên cứu xây dựng mô hình dự báo mức độ an toàn và tuổi thọ của công trình hạ tầng ven biển và hải đảo bằng công nghệ trí tuệ nhân tạo (AI) và dữ liệu giám sát (condition monitoring information).

Tài liệu tham khảo

[1] Y. –H. Huang; “Artificial neural network model of bridge deterioration”. Journal of Performance

- of Constructed Facilities. 2010; 24(6):597-602. DOI: 10.1061/(ASCE)CF.1943-5509.0000124.
- [2] E. Althaqafi, E. Chou; “Developing Bridge Deterioration Models Using an Artificial Neural Network.” *Infrastructures*. 2022; 7. DOI:10.3390/infrastructures7080101.
- [3] I. H. Witten, E. Frank; “Data mining: Practical machine learning tools and techniques with Java implementations”. *ACM Sigmod Record*. 2002; 31(1):76–77.
- [4] T. Chen, C. Guestrin; “Xgboost: A scalable tree boosting system”; in *Proc. the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 13-17 August 2016; San Francisco, California, USA. 2016; pp.785-794.
- [5] C. Cortes, V. Vapnik; “Support-vector networks”. *Machine Learning*. 1995; 20:273–297. DOI:10.1007/BF00994018.
- [6] L. B. Anh, N. V. Ân; “Ứng dụng mạng neuron nhân tạo trong việc dự báo tuổi thọ còn lại của công trình cầu”; Hội nghị Khoa học Cán bộ trẻ lần thứ XVI tháng 05/2022; Viện Khoa học công nghệ xây dựng (IBST), Bộ Xây dựng; Hà Nội, Việt Nam. 2022; tr. 526-531.