

Ứng dụng phương pháp LightGBM dự đoán cường độ liên kết giữa cốt thép bị ăn mòn và lớp bê tông xung quanh

Application of LightGBM for predicting ultimate bond strength between corroded reinforcement and surrounding concrete

Trần Xuân Linh^{a,b*}
Tran Xuan Linh^{a,b*}

^aKhoa Xây dựng, Trường Đại học Duy Tân, Đà Nẵng, Việt Nam

^aFaculty of Civil Engineering, Duy Tan University, Da Nang, 550000, Vietnam

^bViện Nghiên cứu và Phát triển Công nghệ Cao, Trường Đại học Duy Tân, Đà Nẵng, Việt Nam

^bInstitute of Research and Development, Duy Tan University, Da Nang, 550000, Vietnam

(Ngày nhận bài: 16/6/2023, ngày phản biện xong: 27/6/2023, ngày chấp nhận đăng: 01/8/2023)

Tóm tắt

Nghiên cứu này ứng dụng mô hình LightGBM để dự đoán cường độ liên kết của cốt thép bị ăn mòn và lớp bê tông bao quanh. Mô hình được huấn luyện và kiểm nghiệm dựa trên bộ dữ liệu gồm 218 mẫu thí nghiệm sưu tầm từ tài liệu hiện có. Các biến đầu vào để xác định cường độ liên kết giữa cốt thép bị ăn mòn và bê tông xung quanh gồm: cường độ nén của bê tông (X_1); độ dày lớp bê tông bao quanh (X_2); loại cốt thép (X_3); đường kính của cốt thép (X_4); chiều dài đoạn liên kết (X_5) và mức độ ăn mòn (X_6). Kết quả tính toán lấy trung bình sau 30 lần chạy mô hình cho thấy thuật toán LightGBM có thể đạt được hiệu suất dự đoán tốt với sai số bình phương trung bình là 2.580, sai số tuyệt đối trung bình là 1.931, sai số phần trăm tuyệt đối trung bình là 33.379% và hệ số xác định là 0.843. Các chỉ số này cũng tỏ ra vượt trội so với mô hình học máy Rừng ngẫu nhiên. Do đó, mô hình LightGBM có thể hỗ trợ các kỹ sư dự đoán cường độ liên kết của cốt thép bị ăn mòn và lớp bê tông xung quanh.

Từ khóa: Học máy có giám sát; thuật toán LightGBM; trí tuệ nhân tạo; cốt thép bị ăn mòn; cường độ liên kết.

Abstract

This study applies the LightGBM model to predict the ultimate bond strength of the corroded reinforcement and the surrounding concrete layer. The model is trained and tested based on a dataset of 218 samples collected from existing literature. The input variables to determine the bond strength between the corroded reinforcement and the surrounding concrete include: compressive strength (X_1), concrete cover (X_2), steel type (X_3), diameter of steel bar (X_4), bond length (X_5) and corrosion level (X_6). The calculation results after 30 runs of the model show that the LightGBM algorithm can achieve good prediction performance with average root mean square error of 2.580, average mean absolute error of 1.931, average mean absolute percentage error of 33.379% and average coefficient of determination of 0.843. These indicators also proved superior to the Random Forest model. Therefore, the LightGBM model can assist engineers in predicting the bond strength of corroded reinforcement and surrounding concrete.

Keywords: Supervised machine learning; LightGBM algorithm; artificial intelligence; corroded reinforcement; bond strength.

*Tác giả liên hệ: Trần Xuân Linh

Email: tranxuanlinh@duytan.edu.vn

1. Giới thiệu

Cường độ liên kết giữa cốt thép và lớp bê tông xung quanh trong kết cấu bê tông cốt thép (BTCT) là một trong những thông số quan trọng trong tính toán và thiết kế các công trình BTCT. Chỉ số này đặc trưng cho lực tương tác và khả năng duy trì sự kết dính giữa cốt thép và bê tông [1]. Nó có ảnh hưởng trực tiếp đến khả năng chịu tải và tuổi thọ của các công trình BTCT.

Bề mặt liên kết giữa cốt thép và bê tông xung quanh thường được coi là lý tưởng trong quá trình thiết kế các kết cấu BTCT. Tuy nhiên, trên thực tế thì bề mặt tiếp xúc này không lý tưởng mà hiện tượng cốt thép bị ăn mòn, rỉ sét là một trong những nguyên nhân phổ biến. Quá trình ăn mòn cốt thép trong các kết cấu BTCT thường xảy ra đối với các công trình xây dựng trong môi trường có độ ẩm cao và các công trình xây dựng trên sông, biển. Sự suy giảm độ pH của dung dịch bên trong các lỗ rỗng của bê tông kèm theo sự xâm nhập của các ion chloride từ môi trường xung quanh đẩy nhanh hiện tượng ăn mòn cốt thép trong các cấu kiện BTCT [2]. Điều này dẫn đến hệ quả là các vết nứt bên trong bê tông hình thành và giảm sự kết dính giữa cốt thép và bê tông. Bên cạnh đó, áp lực lên lớp bê tông bao quanh cùng với sự lan truyền các vết nứt sẽ khiến kết cấu BTCT xuống cấp, hư hại và giảm tuổi thọ Sæther [3].

Hiện tượng ăn mòn cốt thép trong kết cấu BTCT là một quá trình diễn ra phức tạp và bị chi phối bởi nhiều yếu tố như loại cốt thép, loại bê tông, độ dày của lớp bê tông bảo vệ cốt thép, môi trường xung quanh... [4-7]. Một số nghiên cứu [4, 8] đã chỉ ra rằng, cường độ liên kết giữa cốt thép và bê tông xung quanh có thể tăng nhẹ ở mức độ ăn mòn nhất định và giảm khi mức độ ăn mòn này tăng lên. Do đó, việc dự đoán và đánh giá cường độ liên kết giữa cốt thép bị ăn mòn và lớp bê tông xung quanh là một bài toán phức tạp, có ý nghĩa khoa học và thực tiễn cao.

Hiện nay, các mô hình học máy đang ngày càng được ứng dụng nhiều hơn trong các lĩnh vực khác nhau, đặc biệt là trong lĩnh vực xây dựng và vật liệu. Có thể kể ra một số nghiên cứu ứng dụng trí tuệ nhân tạo trong kết cấu BTCT như: dự đoán độ võng của dầm bê tông [9]; dự báo biến dạng từ biến của bê tông xanh [10]; sử dụng mô hình học máy XGBoost để dự đoán biến dạng từ biến của các dầm và sàn BTCT [11]; đánh giá khả năng chọc thủng sàn BTCT [12]; dự đoán vết nứt trong dầm BTCT [13]; dự đoán cường độ chịu cắt của các kết cấu BTCT [14]...

Trong bài báo này, chúng tôi sẽ sử dụng bộ dữ liệu gồm 218 mẫu thí nghiệm được sưu tầm từ các công trình nghiên cứu thử nghiệm trước đó để huấn luyện và kiểm nghiệm mô hình học máy LightGBM trong dự đoán cường độ liên kết giữa cốt thép bị ăn mòn và bê tông xung quanh ở các kết cấu BTCT.

Các phần tiếp theo của bài báo theo trình tự sau: mục 2 trình bày phương pháp nghiên cứu và mô tả tập dữ liệu; mục 3 giới thiệu các kết quả tính toán và thảo luận về hiệu suất của mô hình; ở cuối bài báo nêu các kết luận và hướng nghiên cứu mở rộng.

2. Phương pháp nghiên cứu

2.1. Mô hình LightGBM

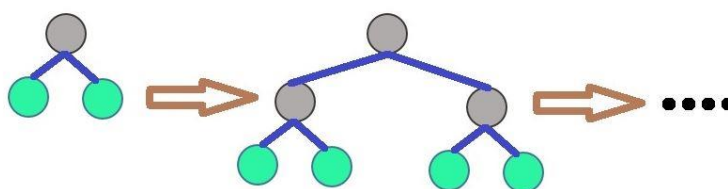
LightGBM là một thuật toán được phát triển bởi Microsoft Research Asia [15] dựa trên phương pháp cây quyết định tăng cường (*Gradient Boosting Decision Tree* (GBDT)). Mô hình này có một số ưu điểm như tốc độ huấn luyện và hiệu quả tính toán cao, có thể sử dụng được với các bài toán dự đoán có số lượng dữ liệu lớn với độ chính xác cao. Nhược điểm của LightGBM là tương đối nhạy cảm với hiện tượng khớp quá mức (*over-fitting*) đặc biệt là đối với các bộ dữ liệu có số lượng ít [16].

Thuật toán GBDT sử dụng phương pháp sắp xếp trước để chọn và phân tách các chỉ số. Ưu

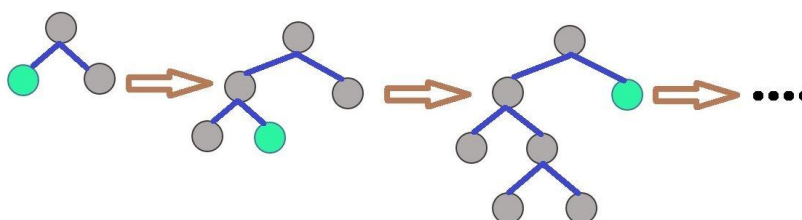
điểm của phương pháp này là có thể xác định chính xác điểm chia tách. Tuy nhiên, nó đòi hỏi nhiều thời gian và bộ nhớ trong quá trình tính toán. Thuật toán LightGBM là phương pháp dựa trên biểu đồ và chiến lược tăng trưởng theo chiều lá của cây (*Trees Leaf-wise Growth Strategy*). Phương pháp này chỉ định giới hạn độ sâu tối đa nhằm tăng tốc độ huấn luyện mô hình và giảm mức tiêu thụ bộ nhớ [16].

Trong các thuật toán cây quyết định phát triển cây theo cấp độ (độ sâu) (*Level-wise Tree Growth*) (Hình 1) do quá trình tìm kiếm và

phân tách nhiều lá có mức độ thu thập thông tin thấp là không cần thiết, nên làm tăng mức tiêu thụ bộ nhớ bổ sung và giảm hiệu quả dự đoán. Do đó, mô hình LightGBM phát triển cây theo chiều lá (*Trees Leaf-wise Growth Strategy*) (Hình 2) [17]. Chiến lược tăng trưởng theo chiều lá tỏ ra hiệu quả hơn vì nó chỉ tách lá có mức thu thập thông tin lớn nhất trên cùng một lớp. Trong phương pháp này, thông số giới hạn độ sâu tối đa của cây được áp dụng vào quá trình sinh trưởng của cây nhằm hạn chế hiện tượng khớp quá mức (*over-fitting*) [16].



Hình 1. Chiến lược tăng trưởng theo độ sâu của cây (*Level-wise Tree Growth*)



Hình 2. Chiến lược tăng trưởng theo chiều lá của cây (*Trees Leaf-wise Growth Strategy*)

Trong bài báo này chúng tôi sử dụng thuật toán LightGBM với các thông số cơ bản sau: phương pháp 'boosting': 'gbdt', đối tượng 'objective': 'regression' và số lá tối đa trên một cây 'num_leaves': 10.

2.2. Bộ dữ liệu thực nghiệm

Trong nghiên cứu này, chúng tôi sử dụng một bộ dữ liệu thực nghiệm gồm 218 mẫu được sưu tầm từ các công trình nghiên cứu trước đó [5, 8, 18-23]. Các yếu tố ảnh hưởng đến cường độ liên kết giữa cốt thép bị ăn mòn và bê tông xung quanh được sử dụng làm biến đầu vào của mô hình LightGBM gồm: cường độ nén bê tông (X_1); độ dày lớp bê tông bao quanh (X_2); loại

cốt thép (X_3); đường kính của cốt thép (X_4); chiều dài đoạn liên kết (X_5) và mức độ ăn mòn (X_6) [4]. Các chỉ số thống kê của các biến đầu vào và biến đầu ra được trình bày chi tiết trong Bảng 1. Loại cốt thép (X_3) được mã hóa nhị phân nhằm mục đích thuận tiện cho quá trình huấn luyện và kiểm nghiệm mô hình: 0 thể hiện cốt thép tròn trơn và 1 cho cốt thép có gờ. Ngoài ra, cường độ nén (X_1) của các mẫu bê tông đa số được đo từ các mẫu bê tông hình lập phương $150 \times 150 \times 150$ mm. Đối với các trường hợp các mẫu bê tông có hình dạng khác, cường độ nén được quy đổi tương đương [4].

Bảng 1. Mô tả thống kê các biến đầu vào và đầu ra

| Biến số | Ký hiệu | Min | Trung bình | ĐLC | Độ xiên | Max |
|--------------------------|----------------|-------|------------|-------|---------|--------|
| Cường độ nén (MPa) | X ₁ | 22.13 | 33.65 | 11.88 | 0.70 | 52.10 |
| Lớp bê tông bảo vệ (mm) | X ₂ | 15.00 | 60.66 | 41.13 | 1.22 | 147.50 |
| Loại cốt thép | X ₃ | -- | -- | -- | -- | -- |
| Đường kính cốt thép (mm) | X ₄ | 12.00 | 16.19 | 4.56 | 1.01 | 25.00 |
| Chiều dài liên kết (mm) | X ₅ | 36.60 | 92.25 | 71.69 | 3.52 | 500.00 |
| Độ ăn mòn (%) | X ₆ | 0.00 | 4.20 | 8.38 | 5.84 | 80.00 |
| Cường độ liên kết (MPa) | Y | 1.30 | 10.94 | 6.96 | 0.82 | 31.70 |

(ĐLC: độ lệch chuẩn)

3. Kết quả tính toán và thảo luận

Với mục đích giảm tác động tiêu cực của các biến số có giá trị lớn, dữ liệu sẽ được chuẩn hóa bằng phép biến đổi dữ liệu điểm Z (*Z-score data transformation*) (công thức 1). Để đánh giá hiệu suất tính toán của mô hình, nghiên cứu này sử dụng các thông số gồm sai số toàn phương trung bình (*RMSE*) (công thức 2), sai số tuyệt đối trung bình (*MAE*) (công thức 3), sai số phần trăm tuyệt đối trung bình (*MAPE*) (công thức 4) và hệ số xác định (*R²*) (công thức 5). Việc sử dụng nhiều tiêu chí đánh giá khác nhau sẽ làm rõ hơn những ưu và nhược điểm của mô hình ở các khía cạnh khác nhau.

$$X_N = \frac{X_O - m_X}{s_X} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_{A,i} - Y_{P,i})^2} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N \frac{|Y_{A,i} - Y_{P,i}|}{Y_{A,i}} \quad (3)$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \frac{|Y_{A,i} - Y_{P,i}|}{Y_{A,i}} \quad (4)$$

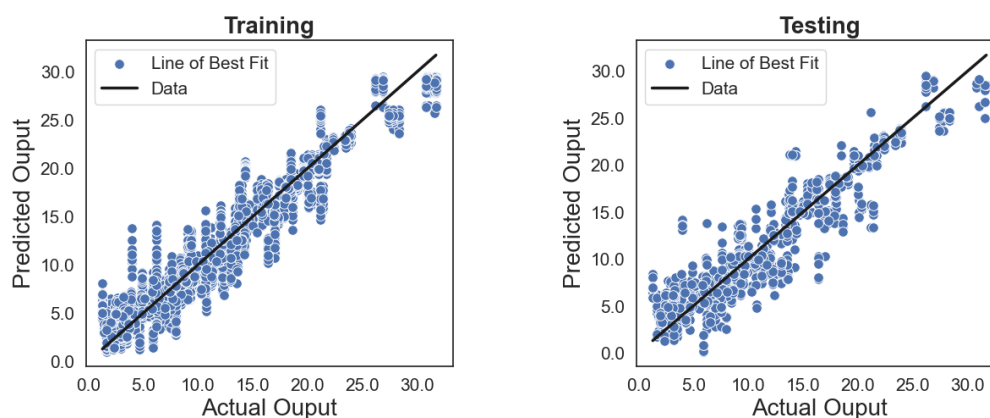
$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_{A,i} - Y_{P,i})^2}{\sum_{i=1}^N (Y_{A,i} - \bar{Y})^2} \quad (5)$$

trong đó X_N và X_O lần lượt là các biến đã chuẩn hóa và biến ban đầu; m_X và s_X thể hiện giá trị trung bình và độ lệch chuẩn của biến ban đầu; $Y_{A,i}$ và $Y_{P,i}$ là kết quả đầu ra thực tế và dự đoán bằng mô hình của mẫu dữ liệu thứ i^{th} ; \bar{Y} là giá

trị trung bình của các biến đầu ra thực tế; N là số lượng mẫu dữ liệu.

Mô hình được huấn luyện và kiểm nghiệm dựa trên ngôn ngữ *Python* và thư viện *scikit-learn*. Tập dữ liệu được chia ngẫu nhiên thành hai tập con gồm: tập thử nghiệm 15% và tập huấn luyện chiếm 85% tổng số mẫu dữ liệu. Mô hình được chạy 30 lần. Kết quả dự đoán cường độ liên kết giữa cốt thép bị ăn mòn và bê tông xung quanh của mô hình *LightGBM* trong cả giai đoạn huấn luyện (*training*) và thử nghiệm (*testing*) sử dụng dữ liệu của 30 lần chạy thể hiện trên Hình 3. Ngoài ra, kết quả tính toán của mô hình *LightGBM* cũng được so sánh với mô hình Rừng ngẫu nhiên (*Random Forest* (RF)) [24] ở bốn chỉ số khác nhau nhằm mục đích đánh giá toàn diện hơn hiệu suất tính toán của *LightGBM* (Bảng 2). So sánh cho thấy mô hình *LightGBM* tỏ ra vượt trội so với mô hình RF ở tất cả các chỉ số trong cả giai đoạn huấn luyện và thử nghiệm. Cụ thể, các chỉ số trung bình sau 30 lần chạy của mô hình *LightGBM* ở giai đoạn kiểm nghiệm lần lượt là: $RMSE = 2.580$, $MAE = 1.931$, $MAPE = 33.379\%$ và $R^2 = 0.843$.

Điều này cho thấy mô hình *LightGBM* có thể được sử dụng như là một công cụ có độ tin cậy cao để dự đoán cường độ liên kết giữa cốt thép bị ăn mòn và lớp bê tông bảo vệ xung quanh trong các kết cấu BTCT.



Hình 3. Minh họa của mô hình dự đoán trong giai đoạn huấn luyện (training) và thử nghiệm (testing).

Bảng 2. So sánh kết quả tính toán của mô hình LightGBM và mô hình RF

| Giai đoạn | Thông số | LightGBM | | RF | |
|------------|----------------|------------|---------------|------------|---------------|
| | | Trung bình | Độ lệch chuẩn | Trung bình | Độ lệch chuẩn |
| Huấn luyện | RMSE (MPa) | 1.988 | 0.142 | 2.576 | 0.055 |
| | MAE (MPa) | 1.515 | 0.110 | 2.022 | 0.046 |
| | MAPE (%) | 25.113 | 2.280 | 34.950 | 1.562 |
| | R ² | 0.918 | 0.013 | 0.862 | 0.007 |
| Thử nghiệm | RMSE (MPa) | 2.580 | 0.294 | 3.152 | 0.461 |
| | MAE (MPa) | 1.931 | 0.232 | 2.471 | 0.363 |
| | MAPE (%) | 33.379 | 9.103 | 42.413 | 11.098 |
| | R ² | 0.843 | 0.051 | 0.780 | 0.062 |

4. Kết luận

Trong nghiên cứu này chúng tôi ứng dụng mô hình học máy LightGBM để giải bài toán dự đoán cường độ liên kết giữa cốt thép bị ăn mòn và lớp bê tông bao quanh trong kết cấu BTCT. Một bộ dữ liệu gồm 218 mẫu thí nghiệm được sưu tầm từ các công trình nghiên cứu thực nghiệm trước đó đã được sử dụng để huấn luyện và kiểm nghiệm mô hình. Các biến đầu vào độc lập của mô hình LightGBM gồm: cường độ nén, độ dày lớp bê tông bảo vệ, loại cốt thép, đường kính cốt thép, chiều dài đoạn liên kết và mức độ ăn mòn. Kết quả tính toán và so sánh sau 30 lần chạy ngẫu nhiên cho thấy mô hình LightGBM có khả năng dự đoán với độ chính xác cao ở cả hai giai đoạn huấn luyện và thử nghiệm thể hiện ở tất cả các thông số RMSE, MAE, MAPE và R².

Ngoài ra, việc phát triển mô hình LightGBM kết hợp với các công cụ tối ưu hóa các tham số mô hình để áp dụng vào những bài toán khác trong lĩnh vực xây dựng cũng là hướng mở rộng có tiềm năng của nghiên cứu này.

Tài liệu tham khảo

- [1] Nepal J., Chen H. P., & Alani A. M. (2013). Analytical modelling of bond strength degradation due to reinforcement corrosion. *Key Engineering Materials* (569), 1060-1067. DOI: 10.4028/www.scientific.net/KEM.569-570.1060
- [2] Hoang N.-D., Chen C.-T., & Liao K.-W. (2017). Prediction of chloride diffusion in cement mortar using Multi-Gene Genetic Programming and Multivariate Adaptive Regression Splines. *Measurement* (112), 141-149. DOI: <https://doi.org/10.1016/j.measurement.2017.08.031>
- [3] Sæther I. (2011). Bond deterioration of corroded steel bars in concrete. *Structure and Infrastructure Engineering* 7(6), 415-429. DOI: 10.1080/15732470802674836

- [4] Güneysisi E. M., Mermerdaş K., & Gültekin A. (2016). Evaluation and modeling of ultimate bond strength of corroded reinforcement in reinforced concrete elements. *Materials and Structures* 49(8), 3195-3215. DOI: 10.1617/s11527-015-0713-4
- [5] Yalciner H., Eren O., & Sensoy S. (2012). An experimental study on the bond strength between reinforcement bars and concrete as a function of concrete cover, strength and corrosion level. *Cement and Concrete Research* 42(5), 643-655. DOI: <https://doi.org/10.1016/j.cemconres.2012.01.003>
- [6] Güneysisi E., Gesoğlu M., Karaboğa F., & Mermerdaş K. (2013). Corrosion behavior of reinforcing steel embedded in chloride contaminated concretes with and without metakaolin. *Composites Part B: Engineering* 45(1), 1288-1295. DOI: <https://doi.org/10.1016/j.compositesb.2012.09.085>
- [7] Molina F. J., Alonso C., & Andrade C. (1993). Cover cracking as a function of rebar corrosion: Part 2—Numerical model. *Materials and Structures* 26(9), 532-548. DOI: 10.1007/bf02472864
- [8] Almusallam A. A., Al-Gahtani A. S., Aziz A. R., & Rasheeduzzafar (1996). Effect of reinforcement corrosion on bond strength. *Construction and Building Materials* 10(2), 123-129. DOI: [https://doi.org/10.1016/0950-0618\(95\)00077-1](https://doi.org/10.1016/0950-0618(95)00077-1)
- [9] Al-Zwainy F. M. S., Zaki R. I. K., Al-saadi A. M., & Ibraheem H. F. (2018). Validity of artificial neural modeling to estimate time-dependent deflection of reinforced concrete beams. *Cogent Engineering* 5(1), 1-15. DOI: 10.1080/23311916.2018.1477485
- [10] Sadowski L., Nikoo M., Shariq M., Joker E., & Czarniecki S. (2019). The Nature-Inspired Metaheuristic Method for Predicting the Creep Strain of Green Concrete Containing Ground Granulated Blast Furnace Slag. *Materials* 12(2), 293.
- [11] Nguyen H., Nguyen N.-M., Cao M.-T., Hoang N.-D., & Tran X.-L. (2021). Prediction of long-term deflections of reinforced-concrete members using a novel swarm optimized extreme gradient boosting machine. *Engineering with Computers* (38), 1-13.
- [12] Hoang N.-D. (2019). Estimating Punching Shear Capacity of Steel Fibre Reinforced Concrete Slabs Using Sequential Piecewise Multiple Linear Regression and Artificial Neural Network. *Measurement* (137), 58-70. DOI: <https://doi.org/10.1016/j.measurement.2019.01.035>
- [13] Ghadimi S., & Kourehli S. S. (2017). Multiple crack identification in Euler beams using extreme learning machine. *KSCE Journal of Civil Engineering* 21(1), 389-396. DOI: 10.1007/s12205-016-1078-0
- [14] Kurtoglu A. E., Gulsan M. E., Abdi H. A., Kamil M. A., & Cevik A. (2017). Fiber reinforced concrete corbels: Modeling shear strength via symbolic regression. *Computers and Concrete* 20(1), 065-075
- [15] G. Ke, Qi M. et al (2017). LightGBM: A highly efficient gradient boosting decision tree [C]. *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*
- [16] Retrieved from <https://lightgbm.readthedocs.io/en/latest/Features.html>
- [17] Shi H. (2007). *Best-first decision tree learning*. The University of Waikato, Hamilton, New Zealand.
- [18] Auyeung Y., Balaguru P., & Chung L. (2000). Bond Behavior of Corroded Reinforcement Bars. *Materials Journal* 97(2), 214-220. DOI: 10.14359/826
- [19] Shima H. (2002). Local bond stress-slip relationship of corroded steel bars embedded in concrete. *Proceeding of the third international symposium on bond in concrete, Budapest, pp 153–158*.
- [20] Zhao Y., & Jin W. (2002). Test study on bond behavior of corroded steel bars and concrete. *J Zhejiang Univ (Engineering Science Edition)* 36(4), 352–356.
- [21] Fang C., Lundgren K., Chen L., & Zhu C. (2004). Corrosion influence on bond in reinforced concrete. *Cement and Concrete Research* 34(11), 2159-2167. DOI: <https://doi.org/10.1016/j.cemconres.2004.04.006>
- [22] Horrigmoe G., Sæther I., Antonsen R., & Arntsen B. (2007). Laboratory investigations of steel bar corrosion in concrete. *Background document SB3.10. Sustainable bridges: assessment for future traffic demands and longer lives. A project co-funded by the European Commission within the Sixth Framework Programme 2007*.
- [23] Chung L., Jay Kim J.-H., & Yi S.-T. (2008). Bond strength prediction for reinforced concrete members with highly corroded reinforcing bars. *Cement and Concrete Composites* 30(7), 603-611. DOI: <https://doi.org/10.1016/j.cemconcomp.2008.03.006>
- [24] Breiman L. (2001). Random Forests. *Machine Learning* 45(1), 5-32. DOI: 10.1023/a:1010933404324.