Xác Định Tình Trạng Hoạt Động Của Đèn LED Sử Dụng Phương Pháp Giám Sát Tải Không Xâm Nhập Và Học Máy

Nguyễn Ngọc Sơn, Nguyễn Thành Công, Nguyễn Huy Tình, Nguyễn Ngọc An*
Trường Đại học Công Nghệ, Đại học quốc gia Hà Nội
Email: ngocan@vnu.edu.vn

Abstract— **Việc giám sát tình trạng hoạt động của đèn LED chiếu sáng có ý nghĩa quan trọng trong việc nâng cao hiệu quả sử dụng năng lượng và đảm bảo sức khỏe người dùng. Trong nghiên cứu này, chúng tôi áp dụng phương pháp giám sát tình trạng hoạt động của đèn LED sử dụng giám sát tải không xâm nhập (NILM) kết hợp học máy dựa trên tín hiệu dòng điện hiệu dụng để giảm thiểu ảnh hưởng của nhiễu môi trường. Phương pháp giúp xác định các tình huống hoạt động của đèn LED bao gồm bình thường, quá nhiệt (không đủ dòng), quá dòng, bị lỗi và hỏng hoàn toàn. Mô hình được xây dựng trên tập dữ liệu đo được từ các đèn LED và có độ chính xác dự đoán lên đến >96% cho thấy tính khả thi của phương pháp được áp dụng.**

Keywords: NILM, giám sát tải không xâm nhập, dự đoán lỗi đèn LED, phát hiện lỗi đèn LED, giám sát tải thiết bị.

I. GIỚI THIỆU

Hiện nay, đèn LED được sử dụng rộng rãi trong chiếu sáng dân dụng và công nghiệp [1, 2]. So với các hệ thống chiếu sáng truyền thống, đèn LED có nhiều ưu điểm nổi bật vượt trội hơn như tuổi thọ cao [3, 4], hiệu suất chiếu sáng cao [5, 6], mức tiêu thụ năng lượng thấp [7, 8], chỉ số hoàn màu cao [9] và phù hợp với sinh lý con người [10]. Tuy nhiên, chúng cũng gặp phải một số vấn đề như tuổi thọ giảm dần, hiệu suất chiếu sáng bị ảnh hưởng bởi nhiều yếu tố hoạt động. Khi hiệu suất chiếu sáng bị suy giảm, ánh sáng LED sẽ dần trở nên kém chất lượng [11], đôi khi khó phát hiện bằng mắt thường và có thể gây ảnh hưởng đến sức khỏe thị giác người sử dụng. Điều này cũng dẫn đến lãng phí năng lượng điện lớn, đặc biệt khi hệ thống chiếu sáng chiếm khoảng 1/5 tổng nhu cầu tiêu thụ năng lượng điện toàn cầu [12]. Vì vậy, việc giám sát được tình trạng hoạt động của đèn LED đóng vai trò quan trọng trong việc tối ưu hóa sử dụng năng lượng và đảm bảo sức khỏe người dùng.

Hiện có nhiều phương pháp giám sát và dự đoán lỗi đèn LED dựa trên các thông số đo lường khác nhau. Cụ thể, có thể kể đến: (1) đo và phân tích thông số điện đầu ra của nguồn đèn LED, như điện áp dòng điện [13], (2) đo và phân tích chỉ số quang học của đèn LED, như chỉ số nhấp nháy hoặc quang thông [14], (3) đo và phân tích nhiệt độ chip LED [15], (4) đo và phân tích các thông tin tổng hợp từ quang học, nhiệt độ và thông số điện của đèn LED [16]. Tuy nhiên, các phương pháp đo quang học và nhiệt độ thường dễ bị ảnh hưởng bởi môi trường xung quanh[17]. Trong khi đó, phương pháp đo thông số điện đầu ra của nguồn LED thường yêu cầu can thiệp vào phần cứng của đèn LED, làm gián đoạn hệ thống và đôi khi không thuận tiện vì thiết bị đo phải được cài đặt trên đèn LED. Việc này rất bất tiện và trên thị trường cũng không có sẵn nhiều thiết bị đo để đáp ứng nhu cầu này.

Bên cạnh đó, có một số nghiên cứu chọn hướng giám sát không xâm nhập, sử dụng học máy [18, 14]. H. Jiang và cộng sự đã sử dụng SVM để phân loại các loại lỗi đèn LED với độ chính xác (Accuracy) đạt 65.4% trên tập test [18]. Tuy nhiên, phương pháp còn hạn chế do dữ liệu bị mất cân bằng khi lấy qua các thông số như cường độ sáng trung bình, mức độ duy trì lumen và chỉ số tạo màu bị ảnh hưởng bởi môi trường. Trong khi đó, Y. Shang và cộng sự đã sử dụng SVM để giám sát các đèn LED FSL với độ chính xác (Accuracy) 100%, các đèn LED OSRAM với độ chính xác (Accuracy) 89.3% [14] trong điều kiện lý tưởng. Tuy nhiên, phương pháp sẽ không còn hiệu quả khi gặp nhiễu ánh sáng hoặc hệ thống chiếu sáng bị thay đổi.

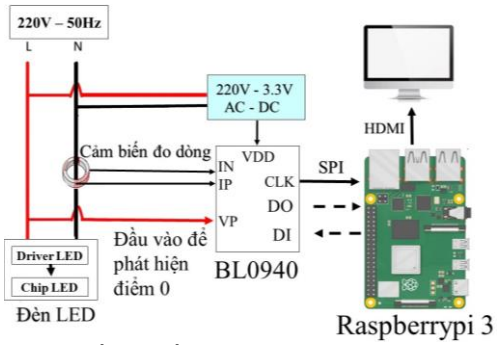
Trong nghiên cứu này, chúng tôi áp dụng phương pháp giám sát tải không xâm nhập (Non-intrusive load monitoring - NILM) kết hợp học máy để xác định tình trạng hoạt động của đèn LED dựa trên giá trị dòng điện hiệu dụng đo được từ nguồn đầu vào của đèn LED nhằm giảm bớt ảnh hưởng của nhiễu môi trường. Các mô hình được xây dựng và đối sánh bao gồm SVM, Random Forest và XgBoost với độ chính xác dự đoán > 96% và ít chịu ảnh hưởng của nhiễu ánh sáng tự nhiên hay các hạn chế của các phương pháp đã nêu ở trên. Các mô hình thu được sau quá trình training được sử dụng để phân loại và đánh giá tập dữ liệu kiểm tra.

Phần tiếp theo của bài báo được tổ chức như sau: trong phần II, chúng tôi miêu tả mô hình thiết bị đo đạc, cách thu thập dữ liệu, mô tả dữ liệu thu thập được. Trong phần III, chúng tôi mô tả cách xây dựng mô hình và các tham số đánh giá. Phần IV cung cấp các kết quả đánh giá mô hình. Cuối cùng, chúng tôi kết luận bài báo trong phần V.

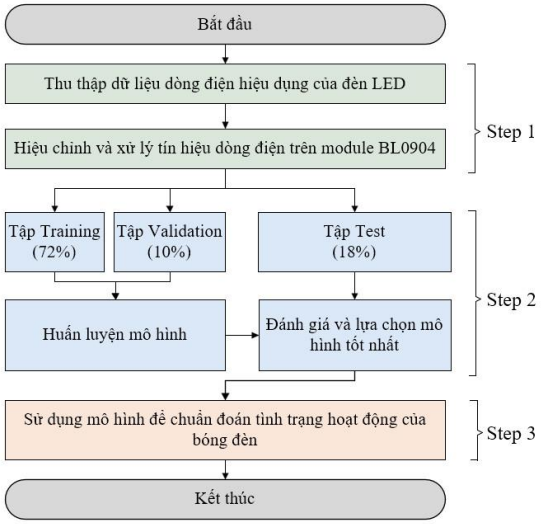
ISBN 978-604-80-8932-0

II. PHƯƠNG PHÁP GIÁM SÁT VÀ THIẾT LẬP HỆ THỐNG ĐO

A. *Hệ thống giám sát không xâm nhập và đo đạc*



Hình 1: Sơ đồ hệ thống giám sát không xâm nhập



Hình 2: Sơ đồ nguyên lý của hệ thống

Hệ thống giám sát không xâm nhập và đo đạc sử dụng trong nghiên cứu này được mô tả trong Hình 1. Trong đó, cảm biến dòng SCT013-100A và IC/module đo đạc BL0940 được sử dụng để thu thập giá trị dòng điện hiệu dụng về bộ xử lý Raspberry Pi 3 để theo dõi và lưu trữ dữ liệu. So với các cảm biến và IC cùng chức năng, cảm biến dòng SCT013-100A có ưu điểm có độ chính xác cao, tần số làm việc lên đến 1KHz, dễ lắp đặt. Trong khi đó, IC BL0940 có tần số lấy mẫu cao, độ chính xác cao, không cần hiệu chỉnh, và khả năng xử lý nhiễu tốt với tốc độ truyền dữ liệu lên đến 900 KHz. Hệ thống giám sát được thiết lập bằng cách kẹp cảm biến đo dòng vào dây cấp nguồn cho đèn LED, mà không cần phải tác động vào thiết kế trước đó của thiết bị. Các linh kiện đo đạc và chuyển đổi nguồn có giá thành rẻ hơn so với các sản phẩm có sẵn trên thị trường. Với kích thước nhỏ gọn, hệ đo có thể lắp đặt đơn giản ở nhiều vị trí, phù hợp để ứng dụng rộng rãi trong các hộ gia đình.

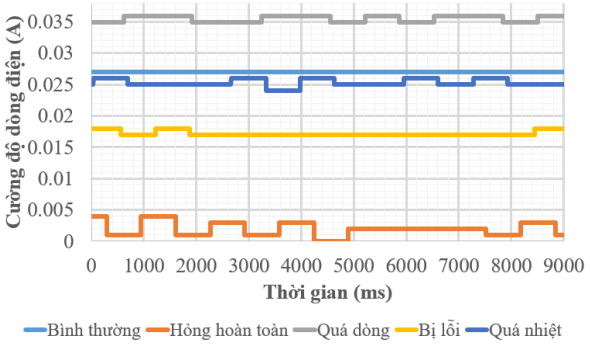
Quá trình đo đạc, xử lý tín hiệu dòng điện, chuẩn bị dữ liệu xây dựng mô hình được minh họa trong Hình 2. Đầu tiên, cường độ dòng điện xoay chiều được đo bởi cảm biến dòng. Sau đó, tín hiệu được hiệu chỉnh, xử lý trên phần cứng module BL0940. Các bước hiệu chỉnh bao gồm: (1) Tín hiệu cảm biến dòng điện xoay chiều 220V-50Hz thu thập được khuếch đại bởi bộ PGA để tăng cường độ tín hiệu, (2) Tín hiệu khuếch đại được lấy mẫu ADC ở tần số cao, (3) Dữ liệu sau khi siêu rời rạc hóa được lọc bởi bộ lọc SINC3 để loại bỏ thành phần tần số cao, (4) Loại bỏ các thành phần DC của tín hiệu, (5) Dữ liệu sau khi lọc thông cao được nhân với chính nó để tính giá trị bình phương, (6) Giá trị đầu ra được lọc bởi bộ lọc thông thấp để loại bỏ các thành phần tần số cao, (7) Giá trị sau khi lọc thông thấp được cộng với giá trị hiệu chuẩn I_RMSOS, sau đó được lấy căn bậc hai để tính toán RMS của tín hiệu, (8) Sau đó, giá trị RMS được tính trung bình theo một số mẫu để cải thiện độ chính xác. Cuối cùng, hệ nhúng Raspberry Pi thu dữ liệu từ module BL0940 thông qua giao thức SPI.

B. *Đèn LED và các dạng tình trạng hoạt động*

Đèn LED có cấu tạo chính bao gồm: nguồn LED, chip LED, tản nhiệt và vỏ. Các cuộc kiểm tra lỗi trên đèn LED cho thấy đèn có thể bị ảnh hưởng bởi tuổi thọ bởi điều kiện hoạt động, nhiệt độ môi trường. Ta có thể phân loại tình trạng hoạt động của đèn LED theo các nhóm chính gồm: hoạt động bình thường, ảnh hưởng bởi nhiệt độ cao, dòng điện không đủ để hoạt động cho hiệu suất ánh sáng đảm bảo, dòng điện vượt quá dòng định mức và hỏng hoàn toàn. Trong nghiên cứu này, chúng tôi sử dụng loại đèn LED Model: A55N4/5W.H, Rạng Đông, có thông số kỹ thuật của nhà sản xuất như trong Bảng 1.

BẢNG 1. THÔNG SỐ KỸ THUẬT CỦA ĐÈN LED A55N4/5W.H, RẠNG ĐÔNG

Thông số	**Giá trị**
Công suất	5W
Điện áp hoạt động	150-220V AC
Quang thông	475lm (6500K)
Hiệu suất ánh sáng	95lm/W (6500K)
Nhiệt độ làm việc	10 - 40°C
Tuổi thọ	2000 giờ (L70)

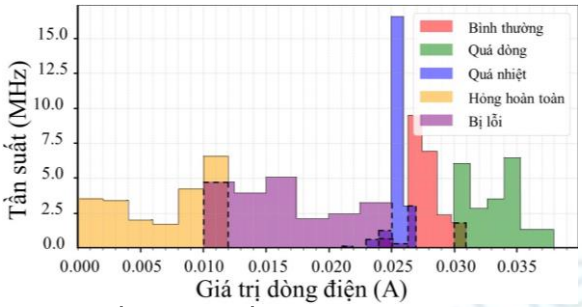


Hình 3: Biểu đồ các mức cường độ dòng điện hiệu dụng của đèn LED tại các tình trạng.

C. Cách thức thu thập dữ liệu

BẢNG 2. PHÂN BỐ DỮ LIỆU GIỮA CÁC NHÃN TRONG TẬP DỮ LIỆU ĐO ĐƯỢC

Nhãn	Số lượng	Phần trăm
Bình thường	2400	20%
Quá nhiệt	2400	20%
Bị lỗi	2400	20%
Quá dòng	2400	20%
Hỏng hoàn toàn	2400	20%



Hình 4: Đồ thị phân bố dữ liệu của các nhãn trong tập dữ liệu đo được

Để xây dựng mô hình, chúng tôi khảo sát hoạt động của 300 bộ đèn LED A55N4/5W.H, trong đó có 110 bộ hỏng hoàn toàn, 100 bộ bị lỗi, 30 bộ hoạt động bình thường, 30 bộ bị ảnh hưởng quá nhiệt giảm dòng (trường hợp này sử dụng bộ đèn hoạt động bình thường trong môi trường nhiệt độ khắc nghiệt), 60 bộ bị quá dòng. Đo đạc trên các thiết bị này, chúng tôi thu thập được 12000 dữ liệu về dòng điện hiệu dụng từ 300 thiết bị đèn LED với tổng thời lượng là 30 giờ. Mỗi đoạn dữ liệu có chiều dài là 9 giây, và được ghi nhận mỗi phút thông qua bộ đo năng lượng điện. Số lượng dữ liệu của các nhãn được lấy đồng đều nhau như mô tả trong Bảng 2, nhằm giảm ảnh hưởng của mất cân bằng dữ liệu [19].

III. CHUẨN BỊ DỮ LIỆU VÀ MÔ HÌNH

Dữ liệu thu thập sau đó sẽ được chia một cách ngẫu nhiên thành hai tập dữ liệu bao gồm tập training 72%, tập test 18% và tập validation 10% (Bước 2), bằng cách sử dụng API của thư viện scikit-learn [20]. Phân bố của từng tập dữ liệu được thể hiện ở Bảng 2, 3, cho thấy tập dữ liệu có tính cân bằng giữa số lượng các nhãn.

BẢNG 3. PHÂN BỐ CỦA TỪNG TẬP DỮ LIỆU

Tập dữ liệu / Nhãn	Tập training	Tập test	Tập validation
Bị lỗi	1718	445	237
Bình thường	1734	419	247
Quá dòng	1709	454	237
Quá nhiệt	1766	411	2 23
Hỏng hoàn toàn	1722	431	247
Tổng cộng	8649	2160	1191
Tỷ lệ (%)	72	18	10

Các thuật toán học máy được sử dụng bao gồm SVM [21], Random Forest [22], và XgBoost [23]. Các chỉ số đánh giá độ hiệu quả của mô hình bao gồm độ chính xác (accuracy), precision, recall, macro F1-score [24], và AUC [25]. Các mô hình được xây dựng sử dụng thư viện scikit-learn. Về các tham số, mô hình SVM được huấn luyện với kernel là Radial Basis Function (RBF). Đối với Random Forest, số lượng cây quyết định (n_estimators) được thiết lập là 100, phương pháp đánh giá chất lượng của cây quyết định (criterion) được đặt là "entropy". Với XgBoost, tham số objective được đặt là "binary:logistic" và tree_methods được sử dụng là "hist". Tất cả các tham số còn lại ở các mô hình được thiết lập theo mặc định.

IV. KẾT QUẢ VÀ THẢO LUẬN

Mô hình sau huấn luyện được đánh giá trên tập test với các chỉ số được mô tả trong Bảng 4. Kết quả cho thấy tất cả các mô hình đều có khả năng dự đoán tốt trên tập Test với độ chính xác và chỉ số macro F1-score trong khoảng từ 91% đến trên 96%. Nguyên nhân giúp các mô hình có khả năng dự đoán tốt có thể do các nhóm dữ liệu có sự phân tách tương đối rõ như được minh họa trong Hình 4. Trong đó, mô hình Random Forest có chỉ số đánh giá tốt nhất, với tất cả chỉ số đánh giá của mô hình này đều tốt hơn so với hai mô hình còn lại.

BẢNG 4. KẾT QUẢ ĐÁNH GIÁ MÔ HÌNH

Mô hình / Chỉ số	SVM	Random Forest	XgBoost
Accuracy	0.915	0.962	0.944
Precision	0.91	0.962	0.944
Recall	0.91	0.962	0.944
Marco F1-Score	0.91	0.961	0.944
Marco AUC	0.994	0.997	0.997

Ngoài ra, tình trạng overfitting của các mô hình cũng được đánh giá thông qua việc so sánh độ lỗi dự đoán (1 - Accuracy) của mô hình trên ba tập dữ liệu training, test, và validation. Kết quả kiểm tra được tổng hợp trong Bảng 5. Nếu mô hình có độ lỗi trên tập test và validation lớn hơn vượt trội so với độ lỗi trên tập training thì mô hình có khả năng bị overfitting [26]. Bảng 5 cho thấy sự chênh lệch không lớn giữa test error và training error của các mô hình và sự ổn định tương đối giữa validation error và test error. Do đó, có thể thấy các mô hình khớp với dữ liệu và cho kết quả dự đoán hợp lý.

BẢNG 5: ĐỘ LỖI DỰ ĐOÁN CỦA CÁC MÔ HÌNH

Mô hình	Training error	Test error	Validation error
SVM	0.091	0.085	0.104
RF	0.005	0.038	0.049
xgBoost	0.005	0.056	0.061

ISBN 978-604-80-8932-0

BẢNG 6. SO SÁNH HIỆU NĂNG MÔ HÌNH VỚI CÁC NGHIÊN CỨU ĐÃ CÓ
(Env. effect: ảnh hưởng từ môi trường)

		Precision	Accuracy	Env. effect
This study	SVM	91.7%	91.3%	Low
	RF	96.2%	96.2%	
	XgBoost	94.4%	94.4%	
Jiang et al (SVM) [18]		82%	65.4%	High
Shang et al [14]	SVM - Quadric (OSRAM LED)	-	89.3%	High
	SVM - Quadric (FSL LED)	-	100%	High

Kết quả dự đoán của các mô hình trong nghiên cứu này được so sánh với hiệu quả của các mô hình được xây dựng một số nghiên cứu đi trước trong Bảng 6. H. Jiang và cộng sự [18] đã sử dụng bộ dữ liệu gồm các thông số: cường độ trung bình của chiếu sáng, mức độ duy trì lumen, nhiệt độ màu (CCT) và chỉ số hoàn màu (RA) nhằm phân loại 04 trạng thái lỗi của đèn LED bao gồm: lỗi chip, lỗi bao phủ đèn LED, quá nhiệt và quá tải điện. Do các thông số này chịu ảnh hưởng mạnh từ nhiễu môi trường nên hiệu quả dự đoán của mô hình chưa cao. Trong khi đó, Y. Shang và cộng sự [14] đã sử dụng bộ dữ liệu có thông số là đặc tính thời gian - tần số từ tín hiệu ánh sáng phát ra của đèn LED nhằm nhận diện các trạng thái: không có lỗi, lỗi loại I, lỗi loại II với đèn của hãng OSRAM và không có lỗi, lỗi loại III đối với hãng FSL với kết quả độ chính xác cao trong điều kiện thí nghiệm. Tuy nhiên, phương pháp này yêu cầu dữ liệu đo đạc tín hiệu ánh sáng có độ chính xác cao, thiết bị đo đạc phức tạp, theo đó chịu ảnh hưởng mạnh từ nhiễu môi trường và khó triển khai trong thực tế. Trong khi đó, nghiên cứu này xây dựng các mô hình học máy dựa trên dữ liệu từ dòng điện hiệu dụng của đầu vào đèn LED có hiệu quả dự đoán tốt (>90%), ít chịu ảnh hưởng của các nhiễu môi trường, ví dụ như ánh sáng tự nhiên, và không đòi hỏi khắt khe về thiết bị đo đạc khi áp dụng thực tế.

V. KẾT LUẬN

Trong nghiên cứu này, chúng tôi đã áp dụng một phương pháp giám sát tải không xâm nhập (NILM) sử dụng học máy để theo dõi và dự đoán tình trạng hoạt động của đèn LED. Các mô hình được huấn luyện cho khả năng dự đoán tốt với độ chính xác đạt đến 96.2% với mô hình Random Forest. Tuy nghiên cứu mới chỉ được thực hiện trên loại đèn LED A55N4/5W.H nhưng nghiên cứu này cũng có thể được áp dụng trên các loại đèn LED cũng như các hệ thống chiếu sáng khác nhau.

Phương pháp của chúng tôi có khả năng dự đoán tốt và ít chịu ảnh hưởng bởi điều kiện môi trường, không đòi hỏi cao về phần cứng. Bên cạnh đó, phương pháp thể hiện ưu thế rõ rệt so với các phương pháp giám sát tải có xâm nhập do không cần can thiệp vào hệ thống phần cứng của thiết bị chiếu sáng, giảm chi phí cài đặt.

Với kích thước nhỏ gọn, và dễ lắp đặt, hệ thống giám sát không xâm nhập này có tiềm năng được ứng dụng rộng rãi không chỉ trong các hệ thống chiếu sáng lớn mà còn trong chiếu sáng ở hộ gia đình.

TÀI LIỆU THAM KHẢO

[1] P. Mottier, LED for lighting applications, no. 134, John Wiley & Sons, 2010.

[2] H. H. Lê, "Lessons learned from "Piloting Led lights in fishing instead of traditional light sources to reduce green house gas emission" project," 2016.

[3] S. M. Rahman, J. Kim, G. Lerondel, Y. Bouzidi, K. Nomenyo và L. Clerget, "Missing research focus in end-of-life management of light-emitting diode (LED) lamps," 2017.

[4] Đ. Q. Thạch, N. B. Thi, N. Q. Hải và C. N. Văn, "Hệ thống đèn tín hiệu đường sắt sử dụng công nghệ LED," Bản B của Tạp chí Khoa học và Công nghệ Việt Nam, tập 64, 2022.

[5] M. M. Aman, G. B. Jasmon, H. Mokhlis và A. H. A. Bakar, "Analysis of the performance of domestic lighting lamps," Energy policy, no. 52, p. 482–500, 2013.

[6] D. Feezell và S. Nakamura, "Invention, development, and status of the blue light-emitting diode, the enabler of solid-state lighting," Comptes Rendus Physique, no. 19, p. 113–133, 2018.

[7] M. M. Omran, V. Aziz và K. Youssef, "Energy Efficient Lighting System," in 2019 21st International Middle East Power Systems Conference (MEPCON), 2019.

[8] N. Đ. Nhật, Đ. T. Tiến, N. D. Q. Trâm và T. V. Đàn, "Hiệu quả sử dụng đèn LED trong khai thác thủy sản trên tàu lưới vây xa bờ tại huyện GIO LINH, tỉnh Quảng Trị," Hue University Journal of Science: Agriculture and Rural Development, tập 129, p. 43–49, 2020.

[9] Z. Guo, H. Lu, T. Shih, Y. Lin, Y. Lu và Z. Chen, "Spectral Optimization of Candle-Like White Light-Emitting Diodes With High Color Rendering Index and Luminous Efficacy," Journal of Display Technology, no. 12, pp. 1393-1397, 2016.

[10] C.-Y. Chen, M.-D. Ke, J.-H. Wu và P.-J. Wu, "The Investigation on Physiological Influences and the Working Efficiencies of Several Lighting in Market," in 2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2013.

[11] N. Narendran, J. D. Bullough, N. Maliyagoda và A. Bierman, "Lighting Research Center," Transportation Research, 2020.

[12] S. Uddin, H. Shareef, A. Mohamed, M. A. Hannan và K. Mohamed, "LEDs as energy efficient lighting systems: A

detailed review," in 2011 IEEE Student Conference on Research and Development, 2011.

[13] C. Ding và T. Zhang, "Research on health monitoring of LED lighting systems," in 2016 Prognostics and System Health Management Conference (PHM-Chengdu), 2016.

[14] Y. Shang, F. Sun, Q. Fang, B. Chen và J. Xie, "A novel fault diagnosis strategy for LED lamps via light output time-frequency characteristics analysis and machine learning," Heliyon, no. 9, 2023.

[15] A. Freddi, G. Ippoliti, M. Marcantonio, D. Marchei, A. Monteriu và M. Pirro, "A Fault Diagnosis and prognosis LED lighting system for increasing reliability in energy efficient buildings," in IET Conference on Control and Automation 2013: Uniting Problems and Solutions, 2013.

[16] V. Gupta, B. Basak và B. Roy, "A Fault-Detecting and Motion-Sensing Wireless Light Controller for LED Lighting System," in 2020 IEEE Calcutta Conference (CALCON), 2020.

[17] B. T. Lê, T. T. Đặng và others, "An overview on cooling led lamps," Journal of Technical Education Science, p. 22–26, 2013.

[18] H. Jiang, Q. Ma, F. Yang và M. Shen, "LED device fault diagnosis base on neural network and SVM model analysis ," in 2017 14th China International Forum on Solid State Lighting: International Forum on Wide Bandgap Semiconductors China (SSLChina: IFWS), 2017.

[19] F. Thabtah, S. Hammoud, F. Kamalov và A. Gonsalves, "Data imbalance in classification: Experimental evaluation," Information Sciences, no. 513, p. 429–441, 2020.

[20] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot và E. Duchesnay, "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, no. 12, p. 2825–2830, 2011.

[21] X. Li và Y. Guo, "Active Learning with Multi-Label SVM Classification.," in IjCAI, 2013.

[22] V. Svetnik, A. Liaw, C. Tong, J. C. Culberson, R. P. Sheridan and B. P. Feuston, "Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling," Journal of Chemical Information and Computer Sciences, vol. 43, p. 1947–1958, November 2003.

[23] L. Yi, N. Liu, J. Liu, L. Fan, J. Liu và H. Song, "Non-intrusive load identification method based on XGBoost and GRU by stacking ensemble learning," International Journal of Advanced Mechatronic Systems, no. 9, p. 109–119, 2021.

[24] R. Yacouby và D. Axman, "Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models," in Proceedings of the first workshop on evaluation and comparison of NLP systems, 2020.

[25] J. Huang và C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," IEEE Transactions on knowledge and Data Engineering, no. 17, p. 299–310, 2005.

[26] Ian Goodfellow, Yoshua Bengio, & Aaron Courville (2016). Deep Learning. MIT Press.

ISBN 978-604-80-8932-0