SoftwareX 31 (2025) 102304

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original Software Publication

A software for predicting Pavement Condition Index (PCI) using machine learning for practical decision-making with an exclusion approach

Huthaifa I. Ashqar [a,*], Amjad Issa [b,c], Sari Masri [a]

[a] Data Science and AI Department, Arab American University, Jenin, Palestine
[b] Civil and Architectural Engineering Department, An-Najah National University, Nablus, Palestine
[c] Construction and Transportation Unit, Scientific Research Centers, An-Najah National University, Nablus, Palestine

ARTICLE INFO

Keywords:
Pavement Condition Index
Machine learning
Random forest
Distress exclusion
Road maintenance
Predictive modelling

ABSTRACT

In Palestine and other resource-constrained settings, determining the Pavement Condition Index (PCI) requires exhaustive visual surveys of up to 19 distress types, which is a process that is both time-consuming and costly to obtain. Despite advances in PCI prediction (2023–2025), existing methods still depend on full-distress assessments, failing to reduce fieldwork burden. We present an open-source machine learning software that classifies pavement into PCI categories (Good, Satisfactory, Fair, Poor, Impassable) by systematically excluding low-utility distresses, reducing inspection effort by up to 40% while achieving an overall accuracy of 82%. The framework integrates features such as pavement age, layer thickness, right-of-way (ROW), average daily traffic (ADT), and heavy-duty vehicle percentage.

Metadata

Nr	Code metadata description	Metadata
C1	Current code version	v1.0
C2	Permanent link to code/ repository used for this code version	https://github.com/HuthaifaAshqar/ PCI_prediction
C3	Permanent link to reproducible capsule	Ht*tps://doi.org/10.5281/ zenodo.15530767
C4	Legal code license	MIT License
C5	Code versioning system used	git
C6	Software code languages, tools and services used	Python 3.10, Jupyter Notebook, pandas, scikit-learn, seaborn, matplotlib, xgboost
C7	Compilation requirements, operating environments and dependencies	See requirements.txt: Python ≥3.10, pandas, scikit-learn, xgboost, matplotlib, seaborn, openpyxl>=3.1.2
C8	If available, link to developer documentation/manual	Not applicable or currently unavailable
C9	Support email for questions	huthaifa.ashqar@aaup.edu

Motivation and significance

Regions such as Palestine consider flexible pavements to be the primary parts of the roadway infrastructure system. Their evaluation is usually based on the Pavement Condition Index (PCI), Present Serviceability Rating (PSR), and International Roughness Index (IRI) scoring. PCI is the most preferred out of the three because of its ease of use and wide acceptance [1]. It, however, relies heavily on field surveys for PCI evaluation, which require identifying up to 19 different types of pavement distress and grading them on severity and extent [2]. This results in a resource-strapped environment being inefficient, labour-intensive, and inconsistent.

Allocated budgets, coupled with a lack of qualified inspectors in Palestinian municipalities, worsen issues around street and road rehabilitation planning. The conventional method demands manual inspection of 100-meter sections to assess for distress patterns such as alligator cracking, shear planes, or rutting for functional regression assessment [3]. Structural and functional interdependencies hinder accurate PCI assessments under resource constraints. Limited mid-range resources lead to delayed decision-making and inefficient rehabilitation prioritization [4,5]. Recent work (2023–2025) using physics-informed neural nets [6], graph CNNs [7] and FCM-XGBoost hybrids [8,9].

To overcome these difficulties, we created a machine learning software that predicts the classes of PCI as Good, Satisfactory, Fair, Poor or Impassable based on historical data on pavement conditions while avoiding certain distress criteria. This method decreases data collection burdens while keeping the accuracy threshold high, which in turn makes it scalable and efficient for municipalities with low inspection capacity.

* Corresponding author.
E-mail address: huthaifa.ashqar@aaup.edu (H.I. Ashqar).

https://doi.org/10.1016/j.softx.2025.102304
Received 29 May 2025; Received in revised form 1 August 2025; Accepted 8 August 2025

We validated the framework using asphalt pavement data from Nablus City, Palestine, from 2020 to 2023, and observed the model's reliability with reduced input data.

Software description

Software architecture

The software follows a five-layer pipeline (Fig. 1):

- Data ingestion parses ASTM D6433 spreadsheets, which is a standard practice for conducting pavement condition index (PCI) surveys on roads and parking lots, or CSVs into Pandas DataFrames
- Preprocessing and feature engineering normalize distress percentages, log-transform skewed variables, and one-hot encode categorical features
- Model training optimizes Random Forest and XGBoost via stratified 5-fold cross-validation
- Evaluation computes metrics, confusion matrices, and permutation-based feature importances
- Deployment provides a CLI and an optional FastAPI microservice

Software functionalities

Data ingestion

The data_ingestion module accepts either ASTM D6433 field sheets or generic CSVs, checks for the presence of required columns, and outputs a tidy Pandas DataFrame. During the 2020 and 2023 survey campaigns, field data were collected along 98 hundred-metre segments across two arterial and one collector road in Nablus West (Table 1). Each Surveyor documented Active Subsystem's (A-SYS) traffic counting regarding segment width, layer thickness, average daily traffic, and 18 categories of distress (Table 2), labelling each distress severity based on the proportion of slab area affected, per ASTM D6433: Low ($\leq$ 10%),

Table 1
Geometric configurations of the three surveyed roads.

Street Name (St.)	Direction 1 (Right)		Direction 2 (Left)	
	Length	Width Range	Length	Width Range
Tunis St.	1,200	12	1,200	12
Yafa St.	1,200	7.5	1,200	7.5
No. 25 St.	2,500	4.0 – 12.0	2,500	4.0 – 12.0
Total Length (m)	4,900	—	4,900	—

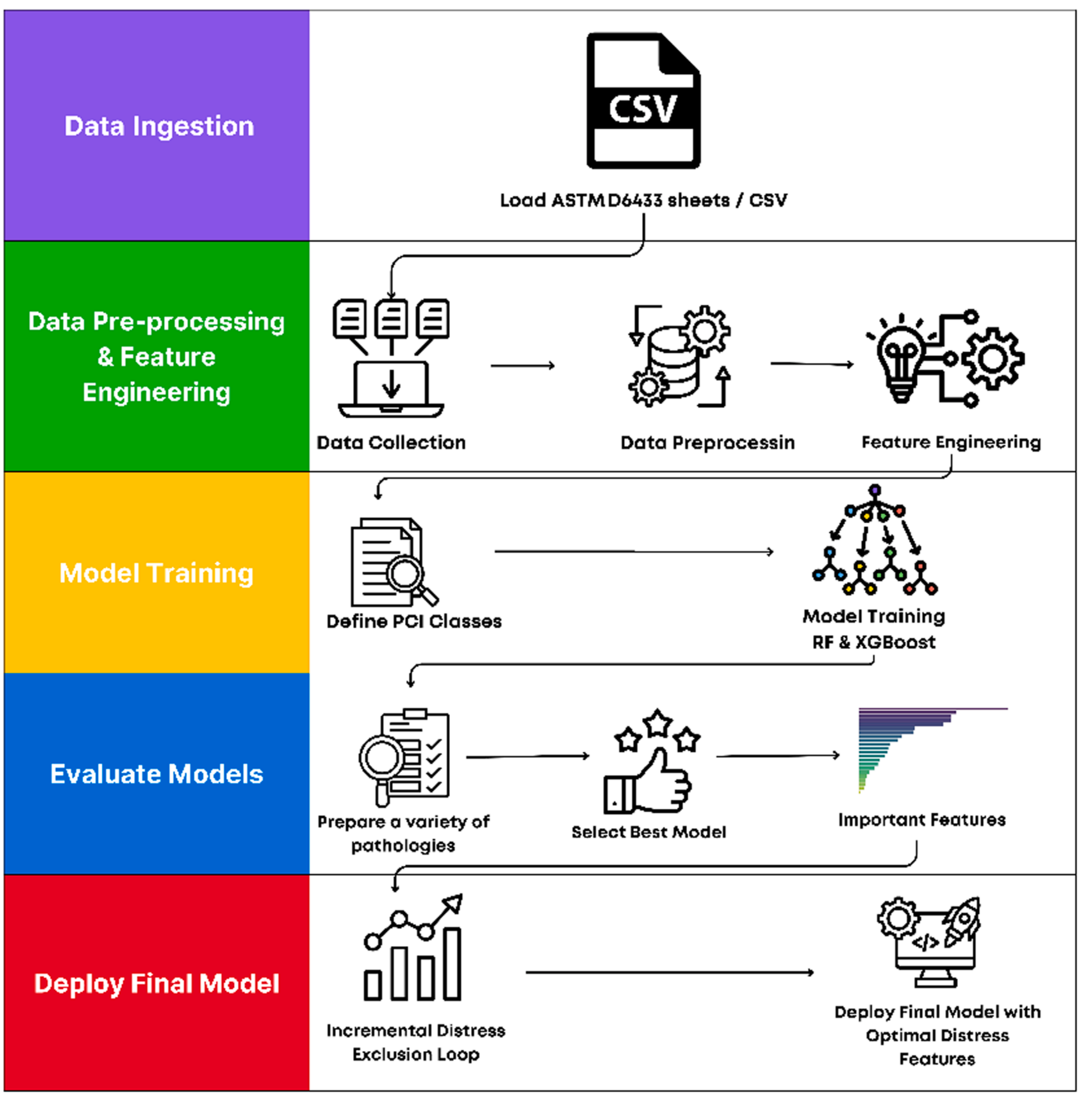

Fig. 1. Five-layer workflow for distress-exclusion-based PCI classification. Direction 1 (Right) and Direction 2 (Left) denote opposite survey passes and are analytically equivalent.

Table 2
Counts of segments exhibiting each distress (after cleaning).

Possible Types of Flexible Pavement Distresses				
(A) Cracking	(B) Patching and Potholes	(C) Surface Deformation	(D) Surface Defects	Group (E) Miscellaneous Distresses
10. Longitudinal and Transverse Cracks	11. Patching	6. Depressions	11. Polished Aggregate	9. Lane/ Shoulder Drop Off
1. Alligator Cracking	4. Bumps and Sags	5. Corrugation	2. Bleeding	14. Rail-Road Crossing*
7. Edge Cracking	13. Potholes	15. Rutting	19. Weathering and Raveling	
8. Joint Reflection Cracking		16. Shoving		
3. Block Cracking		18. Swelling		
17. Slippage Cracking				

* Not applicable, due to the lack of railroads in Nablus City

Medium (10–25%), and High (> 25%). To reduce typographical errors, the drop-down validation incorporated within the raw sheets was digitized prior to ingestion.

The numeric identifiers (1–19) correspond to ASTM D6433 distress codes and our field data template, ensuring consistent classification across surveyors. Twelve rows (10.9%) with > 5% missing data were removed. We also identified the 1–5 PCI rating scale rubric as shown in Table 3, which provides a structured and interpretable framework for classifying pavement conditions into actionable categories, ranging from "Good" to "Impassable." Aligned with ASTM D6433 standards, each class corresponds to a specific PCI range and guides maintenance decisions ensuring that model outputs are not only accurate but also practically meaningful for infrastructure planning and resource allocation. The selection of the 1–5 PCI classification scale in this study is directly aligned with established pavement evaluation standards, particularly ASTM D6433, which defines threshold ranges for PCI-based condition assessment. By mapping the continuous PCI scores into five discrete categories including Good, Satisfactory, Fair, Poor, and Impassable, the scale offers both interpretability and alignment with widely accepted industry practices. This categorical structure is commonly used by transportation agencies and asset management systems to prioritize maintenance actions and allocate resources. Furthermore, this approach simplifies the decision-making process while preserving the fidelity of the original PCI assessment framework. The software also allows for other scales to be implemented, if the user is interested to changing them.

Table 3
Rubric for the 1–5 PCI classification scale used in the study, aligning PCI ranges with condition categories and corresponding maintenance recommendations to support practical decision-making.

Rating	PCI Range	Condition Class	General Description	Recommended Action
5	PCI > 86	Good	Pavement is in excellent condition with minimal to no visible distresses. Surface is smooth, safe, and aesthetically acceptable.	Preventive maintenance (e.g., crack sealing, surface treatments) to extend life.
4	70 < PCI ≤ 86	Satisfactory	Pavement shows minor to moderate signs of aging or wear, with few distresses of low severity. Functionality is unaffected.	Light maintenance; monitor regularly for emerging issues.
3	56 < PCI ≤ 70	Fair	Noticeable surface distresses are present. Structural integrity is still sound, but serviceability is declining.	Minor rehabilitation or targeted repair planning is advised.
2	40 < PCI ≤ 56	Poor	Pavement has significant distress, often including moderate to severe cracking, rutting, or potholes. Ride quality is diminished.	Prioritize for rehabilitation to prevent further degradation.
1	PCI ≤ 40	Impassable	Pavement is severely damaged and fails to meet minimum service standards. May pose safety risks.	Immediate and major rehabilitation or full reconstruction is required.

Feature engineering

Binary flags remain boolean, while numerical variables undergo min-max scaling. Categorical variables that describe road attributes are converted with one-hot encoding. All missing values below 5% are imputed with the median for numeric values and mode for categorical values; rows exceeding that threshold are removed. Categorical text like surface type and surveyed year are one-hot encoded; lane count is one-hot encoded to avoid an artificial ordinal scale. Distress areas are transformed from segments to the percentage of segment area and log-transformed to reduce right skew.

Model training

The `trainer.py` module is responsible for training both Random Forest and XGBoost models. The hyperparameters are set through YAML files, command line interface flags, or directly, and they are fine-tuned with 5-fold cross-validation. Class imbalance is managed using class_weight for Random Forest or scale_pos_weight for XGBoost. The grid search optimizes the number of trees set to {100, 200, 400}, and maximum depth set to {None, 10, 20}, while learning rate η is set to {0.05 or 0.1} for XGBoost. Stratification is utilized over the five PCI condition classes to maintain the “Impassable” cases which are rare.

Random Forest and XGBoost were chosen for their proven effectiveness in handling complex, non-linear relationships and high-dimensional datasets, which are typical in pavement condition modelling. RF offers robustness against overfitting and provides interpretable feature importance, making it well-suited for identifying the most influential distress indicators. XGBoost, on the other hand, is known for its superior predictive accuracy and efficiency, especially with imbalanced datasets, due to its regularization techniques and gradient boosting framework. Both models support multi-class classification, making them ideal for predicting PCI categories while maintaining computational efficiency and interpretability, which are key considerations for practical deployment in infrastructure management systems.

Distress-exclusion engine

A loop that iteratively excludes distress features whose No-Crack ratio, which is defined as the fraction of segments without that distress, exceeds the set threshold. Features that rarely occur are removed without degrading model performance, and the model is retrained at each step with metrics logged.

Evaluation and visualisation

The package prints accuracy, precision, recall and F1-score, generates CSV reports and saves Matplotlib figures such as the feature-importance bar-plot (Fig. 2).

Illustrative examples

According to Table 4, both models' classification results showed that the Random Forest and XGBoost models are effective in predicting pavement conditions, achieving an overall accuracy of 82%. Even though both models perform the same on average, they differ in precision, recall, and F1-score for different categories of pavement conditions which is indicative of the advantages and disadvantages of each algorithm with respect to specific classification problems Table 5.

The models achieve above 90% on precision and recall for well-maintained pavements, performing exceptionally. PCI class thresholds (Good $\geq$ 86, Satisfactory 70–86, Fair 56–70, Poor 40–56, Impassable $<$ 40) follow ASTM D6433. Random Forest yielded a precision of 0.95 and recall of 0.93, with XGBoost reporting 0.93 and 0.94 respectively. Both achieve high F1-scores (0.94 and 0.93), indicating accurate classification with minimal errors. In the Impassable category both models perform well, Random Forest and XGBoost having equal F1-scores of 0.93. XGBoost's higher recall (0.97 vs. 0.94) more effectively captures critical failures, indicating superior performance in those categories. The Random Forest and XGBoost algorithms are appropriate for distinguishing between well-maintained and critically deteriorated roads. However, Better performance in the middle categories might require more training data or improved feature selection. With these frameworks, agencies are able to automate pavement assessment, optimize maintenance workflows, and save costs on manual inspections. For context, recent physics-informed neural-network work reported 80 % accuracy [6] and an FCM–XGBoost approach achieved 83 % [8], placing our 82 % squarely within the current state-of-the-art.

Alligator Cracking, Rut Depth, and Longitudinal Cracking have the most influence and impact the models' outcomes the most. Rut depth and longitudinal cracking reveal problems stemming from structural deterioration, while alligator cracking demonstrates deformation due to water. Each of these features strongly indicates structural degradation, which is why they are so important. This gives road authorities the ability to focus on only the most relevant distresses, saving up to 40% inspection time while maintaining accuracy. This reduces reliance on lengthy strategic documents and facilitates proactive maintenance. Within the CLI, vendors can integrate the tool into asset-management systems, and researchers can apply MIT-licensed packages to benchmark tests, for example, comparing advanced approaches such as LiDAR-enhanced distress detection.

Impact

The newly introduced instruments allow road agencies to customize ASTM D6433 surveys by focusing only on the subset of distresses with a significant impact on PCI prediction. Insightful testing along three urban corridors in Nablus demonstrated that removing certain low-utility distress categories results in inspection time savings of approximately 40%, or about three-person days per kilometre. Such labour savings reduce the duration of the data-to-decision cycle and enable the restricted maintenance budget to be allocated towards the actual maintenance work instead of data collection. Aside from the operational features, this pipeline has an unobstructed MIT license and is open-sourced, which means, any interested parties can contribute to setting a reproducible standard for other researchers conducting multimodal PCI forecasting, such as using LiDAR point clouds or crack detection using vision-transformers. Its command-line interface, alongside the optional FastAPI micro-service, allows commercial asset-management companies to integrate the decision-making logic into their platform, empowering low- and middle-income cities with automatic pavement advisory services.

Limitations and future work

This study is constrained by a geographically limited dataset from Nablus City, which may affect model generalizability. However, the dataset and the results of the study were validated by transportation engineers from the city. Further, the model can be trained and fine-tuned (i.e., calibrated) on any new dataset from a different geographical location. While the model demonstrates strong internal validity through stratified 5-fold cross-validation on the Nablus dataset, we recognize the importance of evaluating its performance on external datasets to ensure broader applicability. Future work will focus on benchmarking the framework across diverse geographic locations, pavement types, and environmental conditions. Plans are underway to collaborate with transportation agencies in other cities to obtain annotated PCI datasets, which will allow us to test the model's generalizability and adjust distress exclusion thresholds accordingly. This external validation will further strengthen the utility of the proposed tool and support its integration into pavement management systems beyond the study area.

Conclusions

This paper proposes a ML software model for predicting the PCI as an alternative to visual inspection approaches employed in developing countries like Palestine. Typically, calculating PCI involves determining the type, severity, and extent of 19 different pavement distresses which is both technically challenging and subjective due to the expertise required and the inherent labour-intensive nature of the process. We constructed a model using data from roads in western Nablus City. The

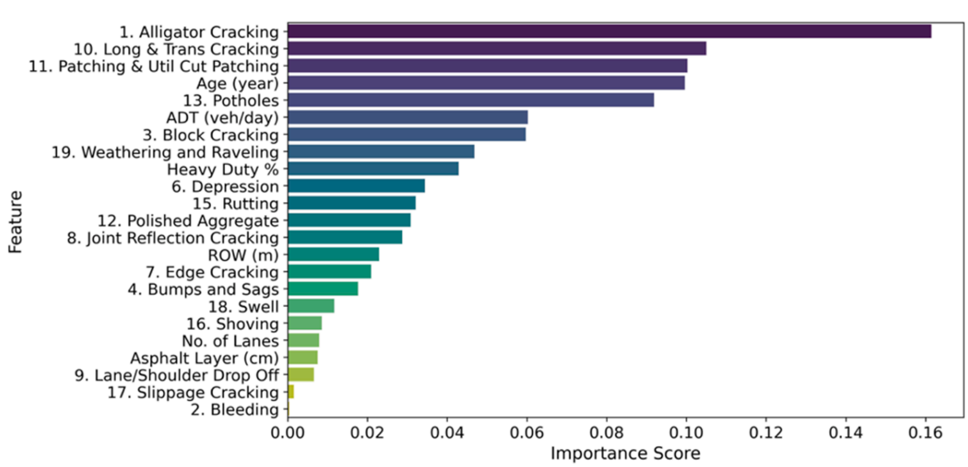

Fig. 2. Feature importance of distresses for PCI prediction.

Table 4
Pseudocode for algorithm used for PCI classification and distress exclusion.

Step 1: Load dataset with PCI values, age (year), pavement layer thickness (m), ROW (m), ADT (veh/day), heavy duty percentage (%), number of lanes, and distress features
Step 2: For each row in dataset:
If PCI $>$ 86, classify as 'Good'
Else if 70 $<$ PCI $\leq$ 86, classify as 'Satisfactory'
Else if 56 $<$ PCI $\leq$ 70, classify as 'Fair'
Else if 40 $<$ PCI $\leq$ 56, classify as 'Poor'
Else classify as 'Impassable'
Step 3: For each distress feature:
If the *No-Crack ratio* for this feature $>$ 95% (i.e., absent in $>$95% of segments):
Drop feature from dataset
Step 4: Set 'Class' column as target variable (y)
Set remaining distress features as input variables (X)
Step 5: Train Random Forest model with 5-fold cross-validation and evaluate performance
Step 6: Obtain feature importance scores from trained model
Step 7: Train XGBoost model with 5-fold cross-validation and evaluate performance
Step 8: If XGBoost outperforms RF:
Select XGBoost as best model
Else:
Select RF as best model
Step 9: Set exclusion threshold to 95%
Initialize threshold based on the *No-Crack ratio
While threshold $\geq$ 0%:
Remove any distress features whose *No-Crack ratio* exceeds the current threshold
Retrain best-performing model
Evaluate model performance
Reduce threshold by 5%
If all distress features are removed:
Stop iteration
Step 10: Select the model with the best performance after incremental exclusion and deploy model for PCI prediction with optimized feature set

Table 5
Classification results of RF and XGBoost across PCI classes. Highest value for each class and metric is highlighted in grey.

Algorithm	Class	Precision	Recall	F1-Score	Accuracy
Random Forest	Good	0.95	0.93	0.94	0.82
	Satisfactory	0.90	0.70	0.79	
	Fair	0.63	0.85	0.72	
	Poor	0.72	0.62	0.66	
	Impassable	0.92	0.94	0.93	
	Average	**0.83**	**0.82**	**0.82**	
XGBoost	Good	0.93	0.94	0.93	0.82
	Satisfactory	0.99	0.65	0.78	
	Fair	0.63	0.85	0.72	
	Poor	0.71	0.62	0.67	
	Impassable	0.90	0.97	0.93	
	Average	**0.84**	**0.82**	0.82	

model utilized parameters including but not limited to road age, pavement thickness, ROW, ADT, heavy vehicle percentage, number of lanes, and major distresses. ML Random Forest and XGBoost models yielded an overall accuracy of 82% for well-maintained and significantly deteriorated roads. This software enables transportation agencies to streamline pavement monitoring with an overall classification accuracy of 82% and precision/recall above 90% for well-maintained roads. By excluding infrequent distresses, inspection time is reduced by up to 40%, allowing resources to be reallocated from data collection to maintenance execution.

CRediT authorship contribution statement

Huthaifa I. Ashqar: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Amjad Issa:** Writing – original draft, Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Sari Masri:** Writing – review & editing, Visualization, Validation, Software, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. S. for Testing and Materials. Standard test method for measurement of fatigue crack growth rates: designation: E 647-08. ASTM International; 2008.
[2] Ahmed NS. Machine learning models for pavement structural condition prediction: a comparative study of random forest (RF) and eXtreme Gradient Boosting (XGBoost). Open J Civil Eng 2024;14(04):570–86. https://doi.org/10.4236/ojce.2024.144031.
[3] Vyas V, Singh AP, Srivastava A. Prediction of asphalt pavement condition using FWD deflection basin parameters and artificial neural networks. Road Mater Pavem Des 2021;22(12):2748–66. https://doi.org/10.1080/14680629.2020.1797855.
[4] Galehouse L, Moulthrop JS, Hicks RG. Principles of pavement preservation: definitions, benefits, issues, and barriers. Tr News 2003;(228).
[5] Nyirandayisabye R, Li H, Dong Q, Hakuzweyezu T, Nkinahamira F. Automatic pavement damage predictions using various machine learning algorithms: Evaluation and comparison. Result Eng 2022;16:100657. https://doi.org/10.1016/j.rineng.2022.100657.
[6] Kargah-Ostadi N, Vasylevskyi K, Ablets A, Drach A. Physics-informed neural networks to advance pavement engineering and management. Road Mater Pavem Des 2024;25(11):2382–403. https://doi.org/10.1080/14680629.2024.2315073.
[7] Boonsiripant S, Athan C, Jedwanna K, Lertworawanich P, Sawangsuriya A. Comparative analysis of deep neural networks and graph convolutional networks for road surface condition prediction. Sustainability 2024;16(22):9805. https://doi.org/10.3390/su16229805.
[8] Lin L, Li S, Wang K, Guo B, Yang H, Zhong W, Liao P, Wang P. A new FCM–XGBoost system for predicting Pavement Condition Index. Expert Syst Applic 2024;249:123696. https://doi.org/10.1016/j.eswa.2024.123696.
[9] Alnaqbi A, Al-Khateeb GG, Zeiada W. Hybrid machine learning applications in pavement engineering: predicting spalling with PSO-GBM. Discov Civ Eng 2025;2:100. https://doi.org/10.1007/s44290-025-00254-4.