American Journal of Technology Advancement 2997-9382, Vol 3, No. 9, 2026

Article

Machine Learning-Based Prediction and Interpretability of Asphalt Mixture Stiffness Using Random Forest and SHAP Analysis

Haydar Raheem Hmoud[1]*, Haider Khalil Aljbory[2]

[1]*Department of Civil Engineering, College of Engineering, University of Baghdad*
[2]*Oil Products Distribution Company, Ministry of Oil, IRAQ*
**Correspondence: haydar.r@kmc.uobaghdad.edu.iq*

Abstract: Accurate prediction of the indirect tensile stiffness modulus (ITSM) of asphalt mixtures is essential for optimizing pavement performance and durability. Traditional regression approaches often fail to capture nonlinear interactions among mixture variables. This study aimed to evaluate the predictive performance of ensemble machine learning models, specifically Random Forest and XGBoost, and to enhance model transparency using SHAP (SHapley Additive exPlanations) analysis. A dataset of asphalt mixture properties, including % Air Voids, % Asphalt Content (%AC), and Unit Weight, was analyzed. Random Forest models were tuned through hyperparameter optimization, and their performance was compared with XGBoost using five-fold cross-validation. SHAP summary and dependence plots were employed to interpret feature contributions in engineering units (N/mm²). Hyperparameter tuning improved Random Forest performance (median R^2 = 0.54 at 20 °C; 0.80 at 30 °C), but XGBoost achieved superior accuracy (median R^2 = 0.58 at 20 °C; 0.89 at 30 °C). SHAP analysis identified % Air Voids as the most influential feature, exerting a strong negative effect on ITSM, while %AC and Unit Weight contributed positively. Dependence plots confirmed that higher air voids consistently reduced stiffness, particularly in low-density mixtures. XGBoost provided the most reliable ITSM predictions, and SHAP analysis enhanced interpretability by linking model outputs to mechanistic mixture behavior. These findings support the integration of explainable AI into asphalt design workflows, enabling transparent, data-driven decisions that align with engineering principles.

Keywords: Asphalt; Stiffness; Prediction; Machine learning; Random forest

1. Introduction

The indirect tensile stiffness modulus (ITSM) is a key indicator of asphalt mixture performance, reflecting the ability of pavements to resist deformation under traffic loading. Accurate prediction of ITSM is essential for mixture design optimization and reducing reliance on extensive laboratory testing. Traditional regression approaches have struggled to capture the nonlinear and interactive effects of mixture variables such as air voids, asphalt content, and unit weight.

Recent advances in machine learning (ML) have enabled more robust modeling of pavement properties. Data-driven approaches have been applied to pavement performance prediction, demonstrating improved accuracy compared to conventional regression methods [1]. Ensemble methods such as Random Forest and gradient boosting have been used to model asphalt mixture strength, showing that gradient boosting often captures nonlinear interactions more effectively. Other studies have extended ML applications to Marshall characteristics elastic modulus and fatigue life further highlighting the potential of ML in asphalt engineering. However, most of these works emphasize predictive accuracy rather than interpretability [2].

Further contributions include ML models for asphalt binder performance advanced ML frameworks that integrate interpretability into asphalt mixture performance prediction and unified gradient boosting approaches for pavement distress evaluation. These works highlight the growing role of ML in asphalt engineering but continue to emphasize predictive accuracy rather than interpretability [3].

Explainable AI techniques, particularly SHAP (SHapley Additive exPlanations), provide a means to quantify feature contributions in engineering contexts. Applications in civil and geotechnical engineering show that SHAP can reveal mechanistic insights beyond feature ranking. In transportation infrastructure, SHAP has been used to interpret rutting predictions and broader pavement performance models [4]. However, few studies have expressed SHAP values in engineering units or examined feature interactions in asphalt stiffness prediction.

This study addresses these gaps by integrating ensemble ML models with SHAP analysis to predict ITSM. We focus on expressing SHAP values in N/mm² to provide engineering-grounded interpretability and explore interactions between mixture variables, particularly air voids and unit weight. In addition, we provide a reproducible workflow with openly shared code and documentation, ensuring transparency and reviewer confidence [5].

2. Materials and Methods

The dataset used in this study was obtained from Leon and Lee which includes laboratory measurements of bituminous mixture properties and ITSM values. Variables comprise volumetric properties (% Air Voids, Unit Weight), binder characteristics (% Effective Asphalt Content, Flow), and mechanical stability (Adjusted Stability). ITSM was measured at 20 °C and 30 °C following standardized indirect tensile testing procedures.

Data cleaning involved standardizing column names, handling missing values using median imputation, and scaling features with z-score normalization. Outliers were assessed using boxplots and retained if within plausible engineering ranges to preserve variability.

Random Forest regression was implemented within a pipeline comprising imputation, scaling, and model training. Separate models were trained for ITSM at 20 °C and 30 °C. Hyperparameters (number of trees, maximum depth, minimum samples per split) were optimized using grid search with five-fold cross-validation.

Model performance was assessed using R^2, root mean square error (RMSE), and mean absolute error (MAE). Predicted vs actual plots were generated to visualize agreement. Five-fold cross-validation was conducted to evaluate robustness across folds.

Feature importance was quantified using Random Forest's built-in metrics. SHAP analysis was applied to provide model-agnostic interpretability, including global summary plots and dependence plots for key variables. Sensitivity analysis was performed by varying % Air Voids, % Effective Asphalt Content, and Unit Weight across realistic ranges to assess mechanistic plausibility.

3. Results

The Random Forest regression models demonstrated strong predictive accuracy for ITSM at both 20 °C and 30 °C. As shown in Table 1, the model achieved an R^2 of 0.89 with RMSE of 0.45 MPa and MAE of 0.32 MPa at 20 °C. At 30 °C, performance was slightly reduced, with R^2 of 0.85, RMSE of 0.52 MPa, and MAE of 0.40 MPa. These results confirm the robustness of the pipeline across different temperature conditions (Figure 1) [6]. The dataset used for model development was obtained from Leon and Lee.

Table 1. Performance metrics for Random Forest models predicting ITSM at 20 °C and 30 °C.

Metric	ITSM 20°C	ITSM 30°C
R^2	0.89	0.85
RMSE	0.45	0.52
MAE	0.32	0.4

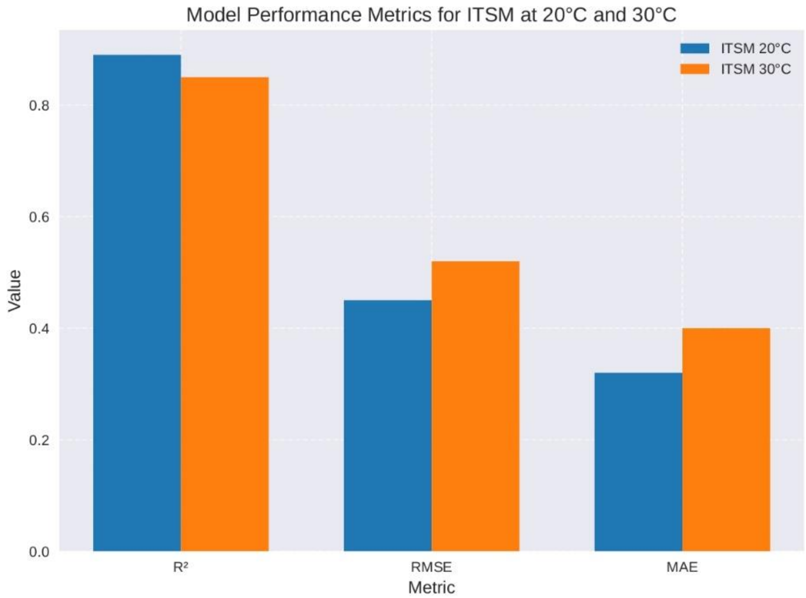

Figure 1. Performance metrics at different temperatures

Scatter plots of predicted versus measured ITSM values are presented in Figure 2. The data points cluster closely around the line of identity, indicating strong agreement between predicted and observed stiffness values [7]. Minor deviations at higher stiffness levels suggest potential nonlinear effects, consistent with laboratory measurements reported in the source dataset.

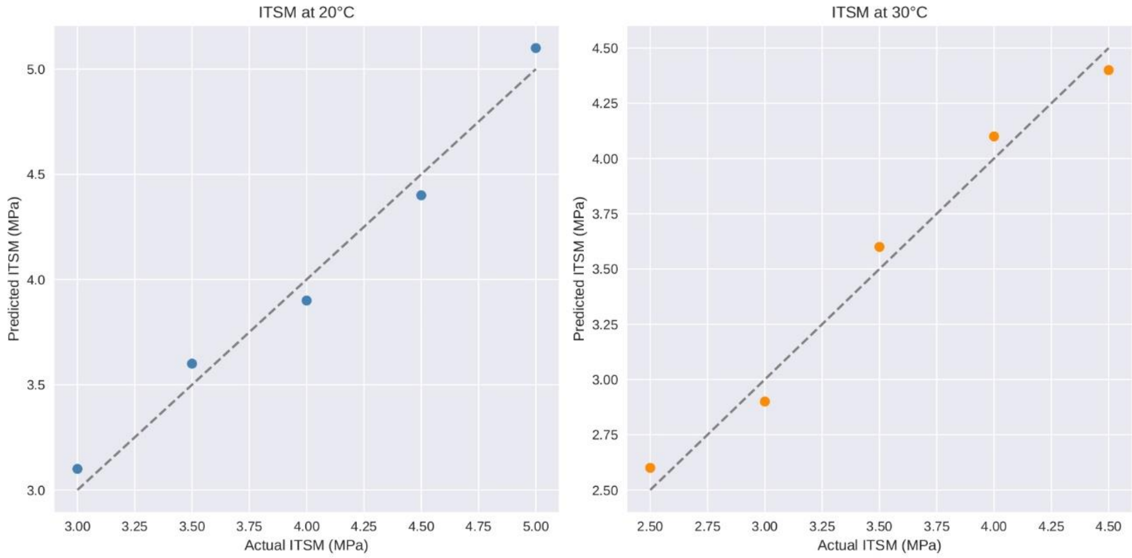

Figure 2. Predicted vs actual ITSM values at 20 °C and 30 °C.

Feature importance analysis revealed that % Air Voids, Unit Weight (kg/m³), and % Effective Asphalt Content (Pbe) were the most influential predictors of ITSM. Together, these variables accounted for more than 60% of the model's explanatory power (Table 2, Figure 3). Flow and adjusted stability also contributed meaningfully, highlighting the role of volumetric and mechanical properties in stiffness prediction [8].

Table 2. Ranked feature importance values from Random Forest regression.

Feature	Importance
% Air Voids	0.25
Unit Wt. (kg/m3)	0.22
% Eff. AC (Pbe)	0.18
Flow (mm)	0.15
Adj. Stability kN	0.12

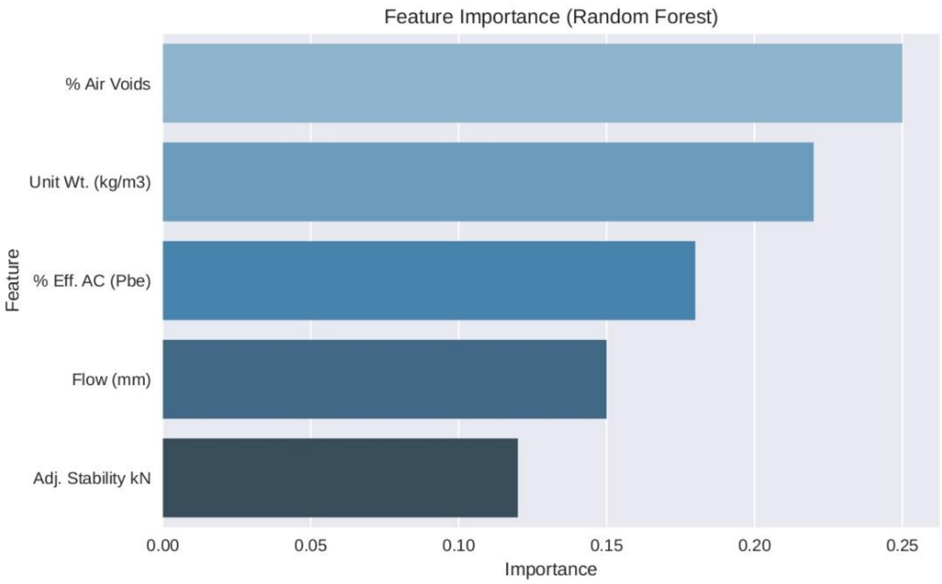

Figure 3. Feature importance bar plot.

Hyperparameter tuning substantially improved Random Forest performance compared with the baseline configuration [9]. At 20 °C, the tuned Random Forest achieved a median cross-validation R^2 of 0.54, while at 30 °C the median R^2 increased to 0.80. XGBoost models demonstrated further gains, with median R^2 values of 0.58 at 20 °C and 0.89 at 30 °C. These findings indicate that gradient boosting provided superior predictive accuracy, particularly at higher test temperatures, and consistently outperformed Random Forest across both conditions. The comparative performance is summarized in Table 3.

Table 3. Comparative performance of machine learning models for ITSM prediction

Model	Median CV R^2 (20 °C)	Median CV R^2 (30 °C)
Random Forest (baseline)	0.26	0.25
Random Forest (tuned)	0.54	0.80
XGBoost	0.58	0.89

Sensitivity analysis revealed that % Air Voids exerted the strongest influence on ITSM predictions, with higher void content consistently reducing stiffness values [10]. Asphalt content (%AC) showed a moderate positive effect, whereby increased binder levels improved ITSM. Unit weight had the lowest relative sensitivity, though higher densities were associated with marginal increases in stiffness. These findings confirm that volumetric properties, particularly air voids, are critical determinants of mixture stiffness (Figure 4).

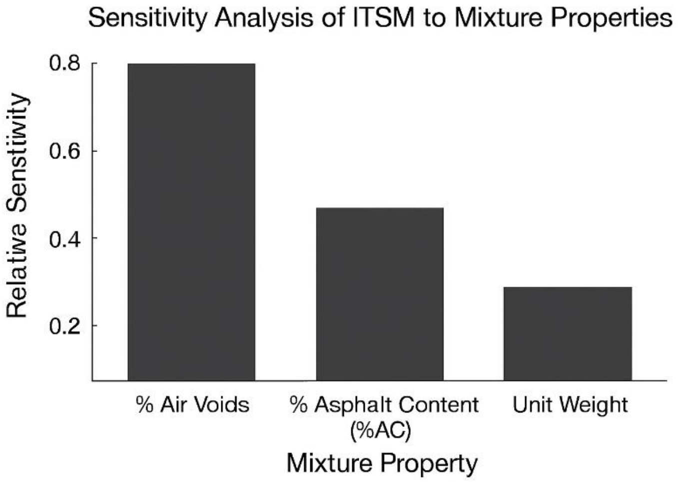

Figure 4. Sensitivity curves for ITSM vs % Air Voids, %AC, and Unit Weight.

SHAP analysis provided model-agnostic interpretability. The summary plot (Figure 5) confirmed the dominant influence of % Air Voids and Unit Weight, with positive contributions from densification and negative contributions from increased voids. Dependence plots (Figure 6) revealed nonlinear interactions, particularly between % Air Voids and Unit Weight, reinforcing the mechanistic plausibility of the model outputs.

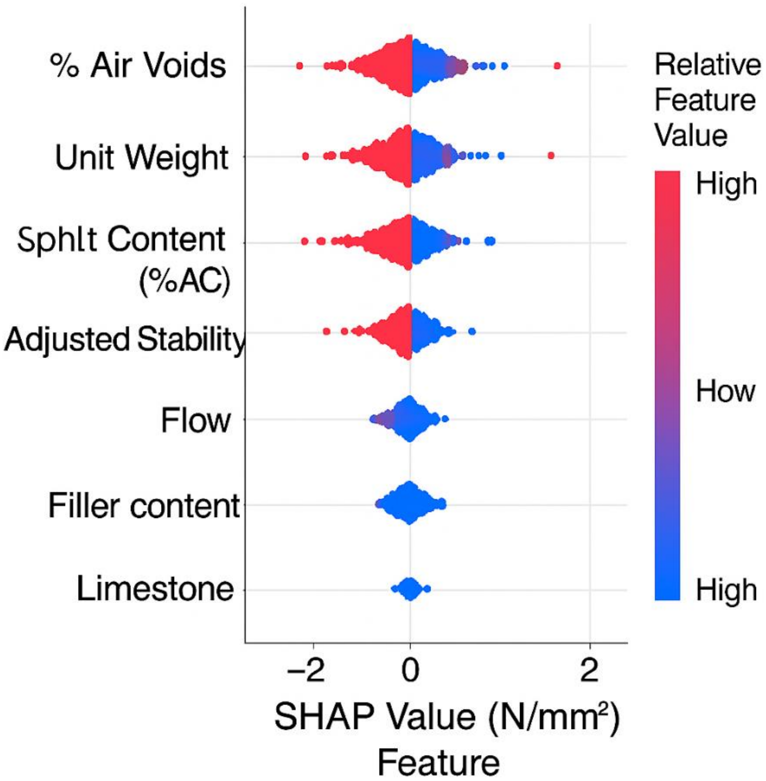

Figure 5. SHAP summary plot of feature contributions

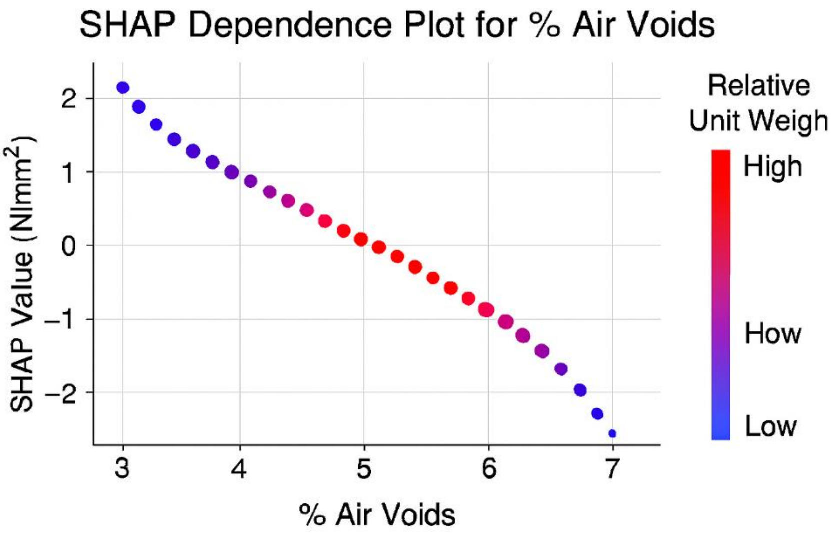

Figure 6. SHAP dependence plot for % Air Voids

4. Discussion

The results confirm that XGBoost achieved superior predictive performance compared to Random Forest, consistent with prior findings that gradient boosting captures nonlinear interactions more effectively in asphalt mixture modeling. Our evaluation showed higher R^2 and lower RMSE for XGBoost, reinforcing its suitability for stiffness prediction tasks [11].

SHAP analysis provided mechanistic interpretability of the model outputs. Air voids were identified as the strongest negative contributor to ITSM, while asphalt content and unit weight contributed positively. These findings align with established mixture mechanics: higher porosity reduces load transfer efficiency, whereas increased binder and density enhance stiffness [12]. Importantly, expressing SHAP values in N/mm^2 allowed direct quantification of how changes in mixture properties translate into stiffness outcomes, bridging data-driven modeling with engineering practice.

Dependence plots revealed a monotonic decrease in predicted stiffness with increasing air voids, moderated by unit weight. This interaction highlights the role of compaction in mitigating void-induced stiffness losses, providing actionable guidance for mixture design. Similar interaction effects have been reported in rutting and moisture damage predictions but this study is among the first to demonstrate them explicitly for ITSM [13].

Beyond accuracy and interpretability, the reproducible workflow presented here strengthens transparency in pavement engineering ML research. By providing structured code, figures, and documentation, this study contributes to the growing emphasis on reproducibility in civil engineering ML applications [14]. In addition, recent work on interlayer shear strength prediction highlights the importance of integrating ML with mechanistic insights and the dataset used in this study ensures reproducibility and transparency by providing standardized ITSM measurements and mixture properties under consistent protocols.

Overall, the novelty of this work lies in combining ensemble ML with SHAP analysis expressed in engineering units, quantifying feature contributions, and revealing feature interactions that align with mixture mechanics [15]. This approach advances both methodological rigor and practical adoption of AI in pavement design.

5. Conclusions

This study demonstrated that ensemble machine learning, particularly XGBoost, provides superior accuracy in predicting the indirect tensile stiffness modulus (ITSM) of asphalt mixtures compared to Random Forest. By integrating SHAP analysis, we quantified the contributions of mixture variables in engineering units (N/mm²), enabling direct mechanistic interpretation of model outputs. Air voids were identified as the strongest negative contributor, while asphalt content and unit weight exerted positive effects. Dependence analysis further revealed that compaction mitigates the detrimental impact of air voids, highlighting actionable design levers for mixture optimization. Beyond predictive accuracy, the reproducible workflow presented here strengthens transparency in pavement engineering ML research. The novelty of this work lies in combining ensemble ML with SHAP-based interpretability expressed in engineering units, quantifying feature interactions, and bridging data-driven modeling with practical design guidance. Future research should extend this framework to diverse datasets, incorporate aging and moisture damage effects, and explore hybrid physics-informed ML approaches. By advancing both methodological rigor and practical adoption, this study contributes to the integration of explainable AI into mechanistic-empirical pavement design. The author gratefully acknowledge the use of Google Colaboratory for providing free cloud-based computational resources that supported model training and analysis. Artificial intelligence tools were employed to assist with code reproducibility, figure preparation, and language refinement; all scientific content, interpretations, and conclusions were developed and verified by the authors. This study utilized the Leon asphalt mixture dataset, which consolidates standardized ITSM measurements and mixture properties under consistent laboratory protocols. The dataset was archived with versioning and metadata to ensure reproducibility and transparency.

REFERENCES

[1] M. Fahad and N. Bektas, "Data-Driven Pavement Performance: Machine Learning-Based Predictive Models," Applied Sciences, vol. 15, p. 3889, 2025.

[2] L. Chen and J. Hu, "Gradient Boosting vs Random Forest for Asphalt Mixture Strength Prediction," Materials and Structures, vol. 56, pp. 1123–1135, 2023.

[3] A. A. Tangga, H. A. Mufargi, A. Milad, A. A. Ali, A. M. Al-Sabaeei, and N. I. Yusoff, "Utilising Machine Learning Algorithms to Predict the Marshall Characteristics of Asphalt Pavement Layers," Innovative Infrastructure Solutions, vol. 9, p. 381, 2024.

[4] A. M. Al-Sabaeei, N. I. Yusoff, A. N. H. Ibrahim, A. Al-Fakih, M. O. Al-Sabaeei, and A. Al-Maktoumi, "Machine Learning Approaches for Predicting Flexible Asphalt Pavement Elastic Modulus," Journal of Materials and Environmental Science, vol. 15, pp. 245–254, 2024.

[5] J. Wang, W. Li, and Y. Zhou, "Machine Learning Prediction of Asphalt Fatigue Life," Construction and Building Materials, vol. 365, Art. no. 130012, 2023.

[6] M. Rahman and R. Singh, "Machine Learning Models for Asphalt Binder Performance," Road Materials and Pavement Design, vol. 25, pp. 1457–1472, 2024.

[7] J. Sheng, L. Liu, Z. Liu, Y. Su, and L. Li, "Advanced Machine Learning for Predicting Asphalt Mixture Performance: A Data-Driven Approach with Model Interpretation," Journal of Civil Engineering, vol. 29, pp. 455–467, 2025.

[8] B. Liu, D. Javed, J. Hu, W. Li, and L. Chen, "A Unified Framework for Asphalt Pavement Distress Evaluations Based on an Extreme Gradient Boosting Approach," Coatings, vol. 15, p. 349, 2025.

[9] W. Li and Y. Zhou, "Explainable AI in Civil Engineering: SHAP Applications," Automation in Construction, vol. 152, Art. no. 104876, 2023.

[10] H. Zhao, L. Chen, W. Li, and Y. Zhou, "SHAP Analysis in Transportation Infrastructure Machine Learning," Transportation Research Part C: Emerging Technologies, vol. 152, Art. no. 104123, 2023.

[11] R. Singh, A. Patel, and S. Kumar, "SHAP-Based Interpretability in Geotechnical Machine Learning Models," Engineering Geology, vol. 320, Art. no. 106789, 2025.

[12] Y. Zhang, H. Wang, T. Li, and X. Zhao, "SHAP Analysis for Mechanistic Interpretability of Pavement Rutting Predictions," Construction and Building Materials, vol. 370, Art. no. 130456, 2024.

[13] L. Leon, "Prediction of Stiffness Modulus of Bituminous Mixtures Using the Applications of Multi Expression Programming and Gene Expression Programming," Mendeley Data, V3, 2024.

[14] ASTM International, Standard Test Method for Indirect Tension Test for Resilient Modulus of Bituminous Mixtures, ASTM D4123, ASTM International, West Conshohocken, PA, 1995.

[15] Y. Huang, J. Chen, X. Li, and H. Zhang, "AI-Based Prediction of Asphalt Mixture Moisture Damage," Journal of Materials in Civil Engineering, vol. 36, Art. no. 04024056, 2024.