



Article info

Type of article:

Original research paper

DOI:

<https://doi.org/10.58845/jstt.utt.2026.vn.6.1.139-158>

***Corresponding author:**

Email address:

bachpx@utt.edu.vn

Received: 20/01/2026

Received in Revised Form:

05/03/2026

Accepted: 25/03/2026

Analysis of Factors Influencing the Penetration Value of Polyethylene-Modified Bitumen Using a Machine Learning Approach

Bach Xuan Pham*, Thu Trang Nguyen, Bao Khanh Vu

University of Transport Technology, 54 Trieu Khuc Street, Thanh Liet Ward, Hanoi

Abstract: This study presents an advanced machine learning-based approach for analysis of factors affecting the penetration of polyethylene (PE)-modified asphalt binders. A comprehensive experimental dataset comprising 204 samples was synthesized from previous studies, covering seven critical input variables: PE content, melting point, PE type, mixing temperature, mixing time, mixing speed, and the penetration of the base bitumen. The performance of five machine learning algorithms (ExtraTrees, LGBM, Gradient Boosting, Bagging, and AdaBoost) was evaluated through 30 Monte Carlo simulations and 5-fold cross-validation to ensure stability and objectivity. The results indicate that the ExtraTrees model achieved superior performance, with a coefficient of determination $R^2 = 0.982$ and a root mean square error $RMSE = 6.473$ on the testing dataset. Explainable artificial intelligence techniques, utilizing Partial Dependence Plots (PDP), elucidated the non-linear interaction mechanisms of the variables; specifically, the base asphalt penetration (X_7) and PE content (X_1) were identified as the two most decisive factors influencing the penetration of PE-modified asphalt binder. Finally, an intuitive user interface was developed to provide a rapid and accurate prediction tool, facilitating the optimization of the mix design process for modified asphalt binders in practice.

Keywords: Asphalt, polyethylene, penetration, machine learning.



Thông tin bài viết
Dạng bài viết:
Bài báo nghiên cứu

DOI:
<https://doi.org/10.58845/jstt.utt.2026.vn.6.1.139-158>

*Tác giả liên hệ:
Địa chỉ Email:
bachpx@utt.edu.vn

Ngày nộp bài: 20/01/2026
Ngày nộp bài sửa: 05/03/2026
Ngày chấp nhận: 25/03/2026

Phân tích các yếu tố ảnh hưởng đến độ kim lún của nhựa đường biến tính Polyethylene dựa trên tiếp cận học máy

Phạm Xuân Bách*, Nguyễn Thu Trang, Vũ Bảo Khánh
Trường Đại học Công nghệ Giao thông Vận tải, 54 Triều Khúc, Thanh Liệt, Hà Nội

Tóm tắt: Nghiên cứu này trình bày một cách tiếp cận tiên tiến dựa trên học máy để phân tích các yếu tố ảnh hưởng đến độ kim lún của nhựa đường biến tính polyethylene (PE). Một bộ dữ liệu thực nghiệm toàn diện bao gồm 204 mẫu đã được tổng hợp từ các nghiên cứu trước đó, bao quát 7 biến đầu vào quan trọng: hàm lượng PE, điểm nóng chảy, loại PE, nhiệt độ trộn, thời gian trộn, tốc độ trộn và độ kim lún của nhựa đường gốc. Hiệu suất của năm thuật toán học máy (ExtraTrees, LGBM, Gradient Boosting, Bagging và AdaBoost) đã được đánh giá thông qua 30 mô phỏng Monte Carlo và xác thực chéo 5 lớp để đảm bảo tính ổn định và khách quan. Kết quả cho thấy mô hình ExtraTrees đạt hiệu suất vượt trội nhất với hệ số xác định $R^2 = 0.982$ và sai số RMSE = 6.473 trên tập dữ liệu kiểm tra. Kỹ thuật trí tuệ nhân tạo có khả năng giải thích thông qua biểu đồ phụ thuộc từng phần (PDP) đã làm rõ cơ chế tác động phi tuyến tính của các biến; trong đó, độ kim lún nhựa gốc (X7) và hàm lượng PE (X1) là hai yếu tố có ảnh hưởng quyết định nhất đến độ kim lún của nhựa đường biến tính PE. Cuối cùng, một giao diện người dùng trực quan đã được phát triển, cung cấp công cụ dự báo nhanh chóng và chính xác, giúp tối ưu hóa quy trình thiết kế nhựa đường cải tiến trong thực tiễn.

Từ khóa: Nhựa đường, polyethylene, độ kim lún, học máy.

1. Đặt vấn đề

Nhựa đường đóng vai trò thiết yếu trong xây dựng mặt đường bộ nhờ khả năng chống thấm và đặc tính kết dính vượt trội. Tuy nhiên, nhựa đường thông thường chưa biến tính thường bộc lộ những hạn chế đáng kể dưới tác động của tải trọng giao thông và các điều kiện nhiệt độ khắc nghiệt, dẫn đến các hư hỏng phổ biến như nứt và hằn lún vệt bánh xe. Qua đó, nhựa đường biến tính có tích hợp các chất phụ gia khác nhau được phát triển nhằm nâng cao hiệu suất của nhựa đường thông thường. Các chất phụ gia phổ thông cho nhựa đường hiện nay bao gồm các chất phụ gia có tính đàn hồi như

styrene-butadiene-styrene, styrene-isoprene-styrene; các chất phụ gia có tính dẻo như polypropylene, polyvinyl chloride, polyethylene (PE) cùng với các loại vụn cao su [1]. Những chất phụ gia này được lựa chọn để cải thiện độ bền và tuổi thọ của mặt đường, đồng thời mang lại các lợi ích tiềm năng về kinh tế và môi trường.

Trong số các chất phụ gia dẻo, phụ gia PE đã nhận được sự quan tâm đáng kể do cấu trúc phân tử đơn giản và là một trong những loại nhựa được tiêu thụ rộng rãi nhất trên toàn cầu [2]. PE là một polymer mạch dài của ethylene, hoạt động bằng cách tăng độ cứng của chất kết dính, mặc dù

nó thường không cải thiện khả năng hồi phục đàn hồi [3]. Các chất phụ gia PE được phân loại dựa trên mật độ và mức độ nhánh phân tử, và được phân thành nhiều loại riêng biệt: polyethylene tỷ trọng thấp (LDPE), polyethylene tỷ trọng cao (HDPE), polyethylene tỷ trọng thấp mạch thẳng (LLDPE) và sáp polyethylene (PE wax). LDPE đặc trưng bởi các mạch nhánh dài, làm cho nó trở thành một polymer linh hoạt, dẻo dai và thường phân tán đồng đều hơn trong nhựa đường so với các loại khác [3]. Ngược lại, HDPE có cấu trúc mạch thẳng hơn với các phân tử xếp khít nhau, từ đó giúp tăng độ ổn định Marshall của hỗn hợp bê tông nhựa [4]. Ngoài ra, sáp PE với trọng lượng phân tử thấp và dải phân bố trọng lượng phân tử rộng, đôi khi được sử dụng để giảm thiểu hiện tượng tách pha thường gặp trong hỗn hợp nhựa đường dùng phụ gia polyme [5]. Sự khác biệt về cấu trúc phân tử giữa các loại PE không chỉ mang lại những đặc tính kỹ thuật riêng biệt mà còn tạo nền tảng quan trọng để tăng cường đáng kể các đặc tính vật lý và cơ học của nhựa đường thông thường.

Tiềm năng sử dụng PE làm chất phụ gia cho nhựa đường là rất lớn, đặc biệt khi xét đến sự tích tụ khổng lồ của các loại rác thải nhựa. Do đó, việc tái sử dụng nhựa PE ngày càng được quan tâm, để thúc đẩy tính bền vững của môi trường và kinh tế tuần hoàn [6]. Việc kết hợp PE vào nhựa đường làm tăng cường đáng kể các đặc tính vật lý và cơ học của nhựa đường chưa biến tính. Nhiều nghiên cứu đã chỉ ra rằng phụ gia PE làm tăng điểm hóa mềm và độ nhớt, đồng thời giảm độ kim lún của nhựa đường biến tính [5,7–10]. Những cải thiện về mặt lưu biến này dẫn đến một chất kết dính có độ cứng lớn hơn và có khả năng chống biến dạng tốt hơn ở nhiệt độ cao. Trong các chỉ tiêu đánh giá chất lượng, độ kim lún được xem là thông số kỹ thuật then chốt phản ánh trực tiếp sự thay đổi về độ cứng và khả năng chịu tải của nhựa đường sau khi được biến tính bằng PE. Thí nghiệm đo độ kim lún là một trong các thí nghiệm tiêu chuẩn của nhựa đường dùng để xác định độ cứng của nhựa đường. Hiện nay, việc xác định độ kim lún của

nhựa đường biến tính PE thường được thực hiện thông qua phương pháp thực nghiệm. Tuy nhiên, phương pháp thực nghiệm thường đòi hỏi chi phí cao, đồng thời việc chuẩn bị mẫu thí nghiệm nhựa đường biến tính PE đạt yêu cầu về sự đồng đều của hỗn hợp thường yêu cầu nhiều thời gian và công sức chuẩn bị. Vì vậy, việc phát triển một phương pháp khác để xác định độ kim lún của nhựa đường biến tính PE không thông qua phương pháp thực nghiệm là thật sự cần thiết.

Trong vài thập niên trở lại đây, với sự phát triển của trí tuệ nhân tạo, các thuật toán máy học giúp phát triển mô hình để dự báo đặc tính cơ học của nhựa đường biến tính ngày càng trở nên phổ biến hiện nay, và đưa ra các dự đoán có tính chính xác cao [11–13]. Mặc dù vậy, các nghiên cứu hiện nay vẫn còn hạn chế trong việc xây dựng một mô hình dự báo chuyên biệt và tối ưu cho nhựa đường biến tính Polyethylene (PE), đặc biệt là việc tích hợp các mô hình này vào một công cụ ứng dụng thực tiễn để thay thế các thí nghiệm truyền thống vốn tốn kém và mất nhiều thời gian.

Do vậy, mục tiêu chính của nghiên cứu này là trình bày các bước phát triển mô hình học máy giúp dự đoán độ kim lún của nhựa đường biến tính PE thông qua 204 kết quả thí nghiệm thu thập được. Các thuật toán học máy khác nhau được thử nghiệm trong bước phát triển mô hình. Trong số đó, thuật toán cây cực ngẫu nhiên được phát triển bởi Pierre Geurts và các đồng sự, và được giới thiệu lần đầu trong bài báo [14], đã chứng tỏ khả năng dự đoán vượt trội hơn so với các thuật toán khác trong nghiên cứu này, và qua đó được sử dụng để xây dựng mô hình học máy giúp dự đoán độ kim lún của nhựa đường biến tính PE. Cuối cùng, một phần mềm dự báo với giao diện trực quan được phát triển nhằm cung cấp một công cụ hỗ trợ đắc lực cho các kỹ sư trong việc tối ưu hóa quy trình thiết kế và sản xuất nhựa đường biến tính PE thực tiễn.

2. Cơ sở dữ liệu

Đối với việc giải các bài toán hồi qui sử dụng mô hình học máy, việc xây dựng nguồn cơ sở dữ liệu đầu vào chính xác, có tính phổ quát cao, và

đáng tin cậy là rất quan trọng. Bộ dữ liệu được thu thập từ các nghiên cứu thực nghiệm được tổng hợp từ 23 bài báo và được tổng hợp trong Bảng 1. Trong tổng số 204 kết quả thí nghiệm được thu thập, có 26 mẫu thí nghiệm có nguồn gốc từ nhựa đường gốc, 4 mẫu nhựa đường sử dụng phụ gia sáp PE, 52 mẫu nhựa đường sử dụng phụ gia LDPE, 2 mẫu nhựa đường sử dụng đồng thời phụ gia LDPE và sáp PE, 9 mẫu nhựa đường sử dụng phụ gia LLDPE, 53 mẫu nhựa đường sử dụng phụ gia HDPE, 13 mẫu nhựa đường sử dụng đồng thời phụ gia LDPE và một loại phụ gia khác PE, 11 mẫu nhựa đường sử dụng đồng thời phụ gia HDPE và một loại phụ gia khác PE, 4 mẫu nhựa đường sử dụng đồng thời phụ gia LLDPE và một loại phụ gia khác PE, 18 mẫu nhựa đường sử dụng phụ gia PE

tái chế, và 12 mẫu nhựa đường sử dụng đồng thời phụ gia PE tái chế và một loại phụ gia khác PE. Mặc dù bộ dữ liệu có sự mất cân bằng về số lượng mẫu giữa các loại PE (LDPE và HDPE chiếm ưu thế), việc giữ lại các nhóm thiểu số như sáp PE hay LLDPE là cần thiết để mô hình nắm bắt được xu hướng biến thiên tổng quát của dòng polymer Polyethylene thay vì chỉ giới hạn ở các loại phổ biến. Nhờ cơ chế xây dựng cây quyết định dựa trên các tập con dữ liệu ngẫu nhiên, các thuật toán sử dụng trong nghiên cứu giúp giảm thiểu sự thống trị của các nhóm đa số lên kết quả dự báo sau này. Điều này đảm bảo cho mô hình duy trì được năng lực khái quát hóa ổn định và tính khách quan ngay cả với bộ dữ liệu thực nghiệm có quy mô vừa và đặc điểm phân phối không đồng đều.

Bảng 1. Tổng hợp dữ liệu từ các nghiên cứu thực nghiệm.

STT	Tài liệu	Số mẫu	STT	Tài liệu	Số mẫu
1	Ho và cộng sự. [5]	20	2	A. A. Yousefi [8]	23
3	Ouyang và cộng sự. [15]	11	4	Sharma và cộng sự. [16]	9
5	Vargas và cộng sự. [17]	7	6	Gama và cộng sự. [18]	2
7	Singh và cộng sự. [19]	3	8	Fang và cộng sự. [20]	5
9	Habib và cộng sự. [9]	15	10	Attaelmanan và cộng sự. [21]	5
11	Casey và cộng sự. [10]	9	12	Jun và cộng sự. [22]	9
13	Fang và cộng sự. [23]	2	14	A.A. Hadidy và T. Yi-qiu [3]	5
15	Nizamuddin và cộng sự. [24]	5	16	Wong và cộng sự. [25]	3
17	Fang và cộng sự. [26]	6	18	As Sk và KSB Prasad [7]	15
19	Yan và cộng sự. [27]	17	20	Fang và cộng sự. [28]	5
21	Costa và cộng sự. [29]	9	22	D.B.Eme và C.Nwaobakata [30]	6
23	H.A.A. Gibreil và C.P. Feng [31]	13			
Tổng số mẫu				204	

Các loại phụ gia PE khác nhau đóng vai trò là một trong các biến đầu vào (X3) để dự đoán biến đầu ra là độ kim lún của nhựa đường biến tính PE (Y) và được mã hóa từ 0 đến 10. Cụ thể, giá trị 0 đại diện cho nhựa đường gốc không trộn phụ gia; các mã từ 1 đến 5 tương ứng với các loại PE đơn lẻ gồm sáp PE (1), LDPE (2), hỗn hợp LDPE và sáp PE (3), LLDPE (4), và HDPE (5). Các giá trị từ 6 đến 8 mô tả các tổ hợp của LDPE, HDPE hoặc LLDPE với các loại phụ gia khác. Cuối cùng, mã 9 và 10 được gán cho nhóm sử dụng PE tái chế đơn lẻ và PE tái chế kết hợp phụ gia. Việc phân loại và

mã hóa này dựa trên đặc điểm cấu trúc phân tử và trạng thái phối trộn thực tế từ bộ dữ liệu 204 mẫu thực nghiệm. Cần lưu ý rằng các giá trị này chỉ đóng vai trò như nhãn định danh cho từng loại PE và không biểu thị quan hệ thứ bậc hay định lượng giữa các loại vật liệu. Việc mã hóa này được sử dụng như một phương pháp biểu diễn biến phân loại trong quá trình xây dựng mô hình. Ngoài ra, mô hình còn xét đến 6 biến đầu vào khác, bao gồm hàm lượng PE (X1), điểm nóng chảy của nhựa PE (X2), nhiệt độ trộn mẫu hỗn hợp nhựa đường gốc và phụ gia PE (X4), thời gian trộn mẫu hỗn hợp

(X5), tốc độ trộn mẫu hỗn hợp (X6), và độ kim lún của nhựa đường gốc (X7). Bảng 2 trình bày chi tiết các giá trị thống kê bao gồm trung bình, độ lệch chuẩn, giá trị nhỏ nhất, lớn nhất, trung vị của từng thông số đầu vào, cũng như của thông số đầu ra

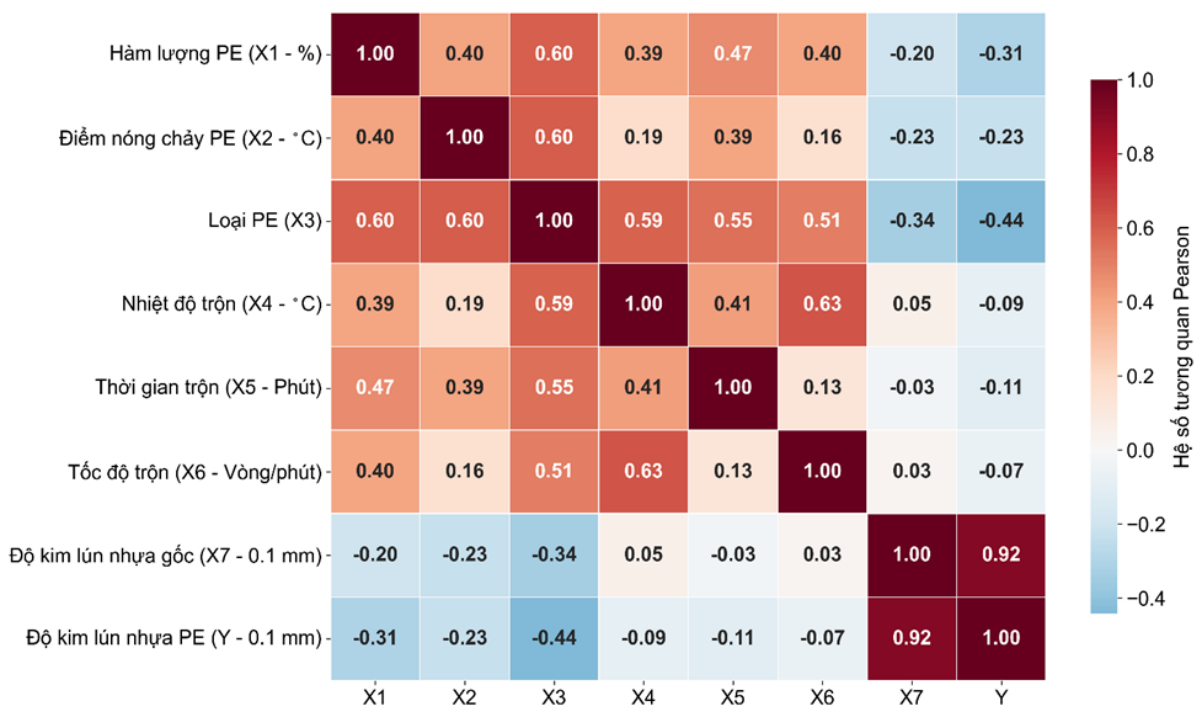
của mô hình. Các biểu đồ hộp ở Hình 2 cũng biểu thị các giá trị thống kê đặc trưng như giá trị trung vị, giá trị trung bình, khoảng tứ phân vị, giá trị lớn nhất và nhỏ nhất không bị xem là giá trị ngoại lai, và giá trị ngoại lai của các biến đầu vào và đầu ra.

Bảng 2. Thống kê mô tả bộ dữ liệu trong nghiên cứu này.

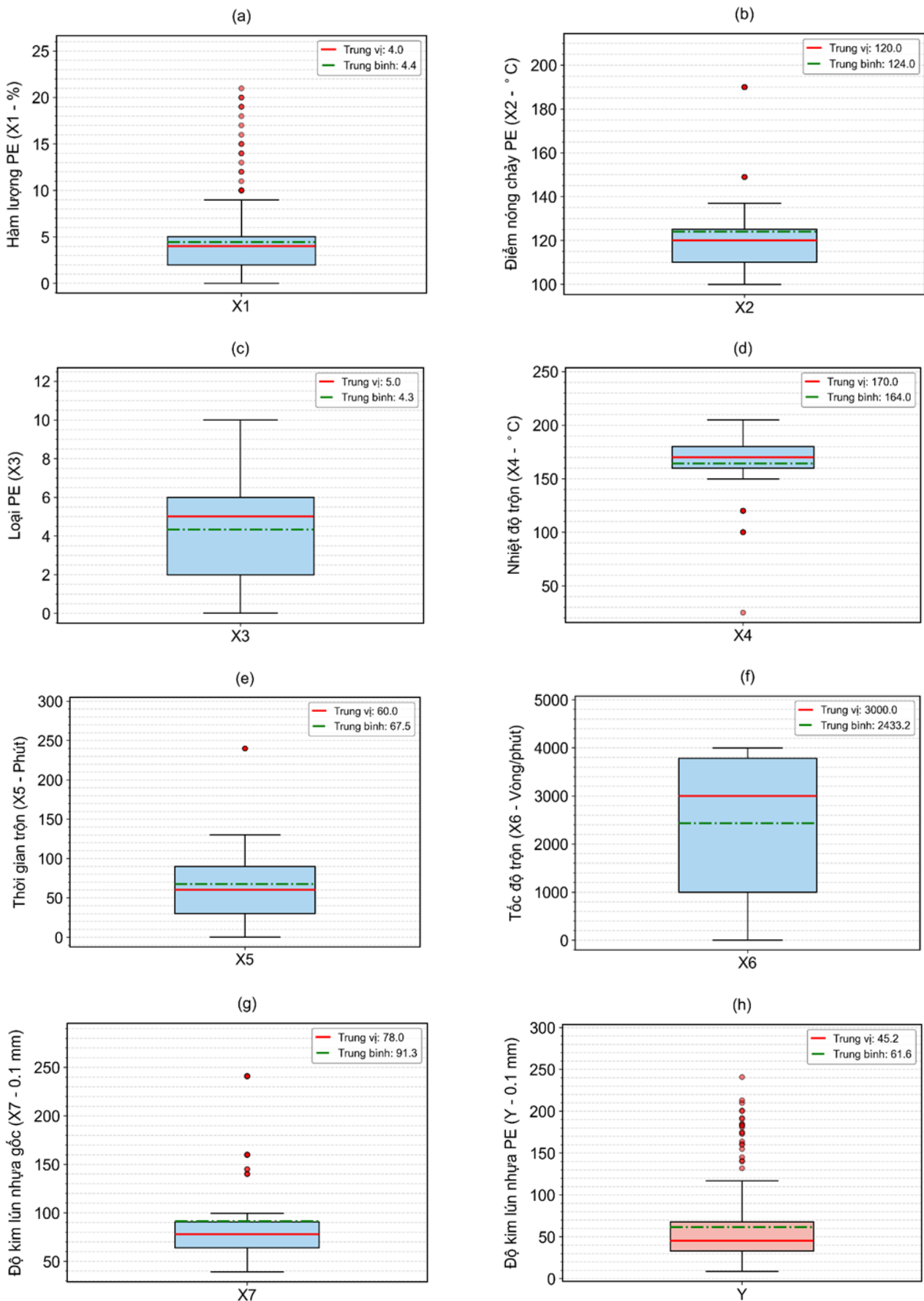
Đặc trưng	Trung bình	Độ lệch chuẩn	Nhỏ nhất	Trung vị	Lớn nhất
Hàm lượng PE (X1 - %)	4.43	4.19	0.00	4.00	21.00
Điểm nóng chảy PE (X2 - °C)	123.98	21.22	100.00	120.00	190.00
Loại PE (X3)	4.33	2.97	0.00	5.00	10.00
Nhiệt độ trộn (X4 - °C)	164.04	24.64	25.00	170.00	205.00
Thời gian trộn (X5 - Phút)	67.53	50.65	0.00	60.00	240.00
Tốc độ trộn (X6 - Vòng/phút)	2433.21	1517.32	0.00	3000.00	4000.00
Độ kim lún nhựa gốc (X7 - 0.1 mm)	91.25	56.51	39.00	78.00	241.00
Độ kim lún nhựa PE (Y - 0.1 mm)	61.61	47.29	8.40	45.23	241.00

Bên cạnh đó, mối tương quan giữa các thông số trong mô hình được biểu thị thông qua ma trận tương quan ở Hình 1. Hệ số tương quan và biểu đồ nhiệt ở Hình 1 biểu thị mức độ tương quan giữa hai thông số. Hai biến có tương quan lớn khi hệ số tương quan lớn, tiến gần tới 1, và được biểu thị bằng màu đỏ đậm của biểu đồ nhiệt. Ngược lại, hai biến có tương quan thấp khi hệ số tương quan nhỏ, tiến gần tới 0, và được biểu thị bằng màu xanh

đương. Phân tích ma trận tương quan chỉ ra mối tương quan lớn giữa độ kim lún của nhựa đường gốc (X7) và độ kim lún của nhựa đường biến tính PE (Y), với hệ số tương quan đạt 0.92. Các biến còn lại (từ X1 đến X6) và đầu ra có hệ số tương quan nhỏ hơn 0.5, cho thấy tương quan trung bình và yếu. Tuy nhiên, tất cả các thông số đầu vào đều được giữ lại và được đánh giá, nhằm tăng tính phổ quát của nghiên cứu.



Hình 1. Ma trận tương quan giữa các biến đầu vào, đầu ra trong bộ dữ liệu độ kim lún.



Hình 2. Phân phối dữ liệu các biến đầu vào, đầu ra được sử dụng trong nghiên cứu

3. Phương pháp nghiên cứu

Nghiên cứu này lựa chọn 5 thuật toán học máy tiêu biểu thuộc nhóm học máy kết hợp bao gồm thuật toán cây cực kỳ ngẫu nhiên, thuật toán tăng cường độ dốc nhẹ, thuật toán tăng cường độ dốc, thuật toán đóng bao và thuật toán tăng cường độ thích nghi. Việc ưu tiên nhóm thuật toán này xuất phát từ khả năng tối ưu hóa việc kết hợp nhiều mô hình dự báo yếu để tạo ra một mô hình mạnh mẽ, có khả năng nắm bắt chính xác các mối quan hệ phi tuyến phức tạp giữa thành phần vật liệu và điều kiện chế tạo nhựa đường. Khác với các phương pháp thống kê tuyến tính truyền thống thường bị hạn chế trong việc mô tả sự biến thiên đa chiều của hệ vật liệu composite, nhóm thuật toán học máy kết hợp được chứng minh là có độ ổn định vượt trội và khả năng giảm thiểu hiện tượng quá khớp hiệu quả, đặc biệt phù hợp để xử lý các điểm gãy và vùng bão hòa đặc tính trong các bộ dữ liệu thực nghiệm có quy mô vừa và nhỏ. Quy trình bao gồm việc huấn luyện, tối ưu hóa và giải thích mô hình để đảm bảo tính chính xác và khả năng ứng dụng thực tiễn.

Thuật toán cây cực kỳ ngẫu nhiên

Thuật toán cây cực kỳ ngẫu nhiên (ExtraTrees) là một thuật toán học máy dựa trên cấu trúc cây quyết định, hoạt động bằng cách xây dựng một số lượng lớn các cây độc lập. Khác với các phương pháp truyền thống, ExtraTrees chọn các ngưỡng chia nhánh một cách hoàn toàn ngẫu nhiên và sử dụng toàn bộ tập dữ liệu huấn luyện để xây dựng cây. Kỹ thuật này giúp giảm thiểu phương sai của mô hình hiệu quả hơn so với Random Forest, đặc biệt hữu ích trong việc xử lý các dữ liệu thực nghiệm có độ nhiễu cao [14].

Thuật toán tăng cường độ dốc nhẹ

Thuật toán tăng cường độ dốc nhẹ (LGBM) là một khung thuật toán tăng cường độ dốc hiệu quả cao, sử dụng kỹ thuật phát triển cây theo chiều sâu thay vì theo cấp độ. LGBM tích hợp hai kỹ thuật cốt lõi là lấy mẫu dựa trên độ dốc một bên và Độ quyền bỏ tính năng, giúp giảm đáng kể thời gian huấn luyện và bộ nhớ tiêu thụ mà vẫn duy trì độ chính xác dự báo vượt trội cho các bài toán hồi quy

phi tuyến [32].

Thuật toán tăng cường độ dốc

Thuật toán tăng cường độ dốc (GB) là một kỹ thuật học máy mạnh mẽ dựa trên việc kết hợp tuần tự các mô hình yếu (thường là cây quyết định) để tạo ra một mô hình dự báo mạnh. Thuật toán hoạt động bằng cách tối ưu hóa hàm mất mát thông qua việc tính toán gradient của các sai số từ mô hình trước đó. Phương pháp này cực kỳ hiệu quả trong việc nắm bắt các mối quan hệ phức tạp giữa các thông số công nghệ và tính chất cơ lý của vật liệu [33].

Thuật toán đóng bao

Thuật toán đóng bao Bagging là phương pháp học máy kết hợp thiết kế để cải thiện độ ổn định và độ chính xác của các thuật toán phân loại và hồi quy. Kỹ thuật này tạo ra nhiều tập dữ liệu con bằng cách lấy mẫu có hoàn lại từ tập dữ liệu gốc, sau đó huấn luyện các mô hình song song và lấy trung bình kết quả dự báo. Phương pháp này giúp giảm hiện tượng quá khớp và tăng tính bền vững cho mô hình dự báo [34].

Thuật toán tăng cường thích nghi

Thuật toán tăng cường thích nghi (AdaBoost) là thuật toán tăng cường hoạt động theo cơ chế lặp, trong đó các mô hình sau tập trung sửa chữa sai sót của các mô hình trước bằng cách tăng trọng số cho các mẫu dữ liệu bị dự báo sai. Quá trình này giúp mô hình thích nghi tốt với các vùng dữ liệu khó dự báo, từ đó tối ưu hóa tổng thể hiệu suất mô hình trên toàn bộ dải dữ liệu thực nghiệm [35].

Mô phỏng Monte Carlo

Mô phỏng Monte Carlo là một kỹ thuật thông kê hiện đại dựa trên việc lặp lại quá trình xáo trộn và phân chia dữ liệu một cách ngẫu nhiên. Phương pháp này cho phép đánh giá hiệu suất của các thuật toán học máy qua nhiều kịch bản khác nhau thay vì chỉ dựa trên một lần phân chia dữ liệu đơn lẻ. Kết quả thu được từ mô phỏng giúp xác định các giá trị kỳ vọng và khoảng tin cậy của mô hình, từ đó đảm bảo tính khách quan, độ ổn định và khả năng tổng quát hóa của kết quả dự báo [36].

Biểu đồ phụ thuộc từng phần

Biểu đồ phụ thuộc từng phần (PDP) được sử dụng để giải thích cách thức các biến đầu vào ảnh hưởng đến giá trị dự báo độ kim lún. Biểu đồ phụ thuộc từng phần là một kỹ thuật giải thích mô hình không phụ thuộc vào thuật toán, được sử dụng để trực quan hóa ảnh hưởng biên trung bình của các biến đầu vào đối với giá trị dự báo. Bằng cách tính toán giá trị trung bình toán học để loại bỏ ảnh hưởng của các biến gây nhiễu, PDP giúp cô lập mối quan hệ trực tiếp giữa đặc trưng đang xét và kết quả đầu ra. Nghiên cứu này triển khai hai dạng biểu đồ chính, PDP một chiều (PDP-1D), dùng để kiểm tra tác động biên của một biến đơn lẻ, thể hiện bản chất của sự phụ thuộc thông qua biểu đồ đường. PDP hai chiều (PDP-2D) dùng để xem xét hiệu ứng tương tác giữa cặp biến cụ thể, được trình bày dưới dạng biểu đồ nhiệt với các dải màu đại diện cho mức độ thay đổi của kết quả dự báo. Việc phân tích PDP giúp xác thực logic của mô hình học máy, đảm bảo các kết quả dự báo phù hợp với các nguyên lý vật lý và kỹ thuật thực nghiệm trong nghiên cứu nhựa đường [37].

Xác thực chéo

Xác thực chéo (CV) là một kỹ thuật thống kê quan trọng nhằm đánh giá khách quan khả năng tổng quát hóa của các mô hình học máy trong bài toán hồi quy. Phương pháp xác thực chéo k-lớp được áp dụng bằng cách chia bộ dữ liệu thành k phần (thông thường k=5 hoặc 10 phần). Trong nghiên cứu này áp dụng kỹ thuật xác thực chéo 5 lớp để đánh giá khả năng tổng quát hóa của mô hình. Quy trình bao gồm k lần lặp lại: tại mỗi lần, một phần dữ liệu được dùng làm tập kiểm tra và k-1 phần còn lại dùng để huấn luyện mô hình. Để đảm bảo tính khách quan, một tập kiểm tra độc lập hoàn toàn được tách riêng khỏi quá trình huấn luyện cho đến bước đánh giá cuối cùng. Hiệu suất chung của mô hình được xác định bằng giá trị trung bình của các chỉ số đo lường qua k vòng lặp, giúp giảm thiểu sai số do việc chia dữ liệu ngẫu nhiên. Kỹ thuật này đóng vai trò then chốt trong việc tinh chỉnh siêu tham số và ngăn chặn hiện tượng quá khớp, từ đó tối ưu hóa độ chính xác dự báo cho

các thuộc tính của nhựa đường cải tiến.

Đánh giá hiệu suất dự báo của mô hình

Hiệu suất của các mô hình được định lượng thông qua các chỉ số thống kê bao gồm hệ số xác định (R^2), sai số tuyệt đối trung bình (MAE), sai số bình phương trung bình (MSE) và căn bậc hai sai số bình phương trung bình (RMSE). Ngoài ra, chỉ số A30 (tỷ lệ dự báo có sai số trong khoảng 30%) cũng được sử dụng để đánh giá độ chính xác của mô hình trong phạm vi sai số kỹ thuật cho phép.

RMSE là thước đo tiêu chuẩn được áp dụng rộng rãi để định lượng sai lệch của các mô hình hồi quy. Do cơ chế bình phương phần dư, chỉ số này có độ nhạy cao đối với các sai số biên độ lớn và các giá trị ngoại lai trong tập dữ liệu dự báo. Nhằm cung cấp cái nhìn toàn diện về đặc điểm phân phối và sự biến động của sai số, RMSE thường được phân tích song song với sai số tuyệt đối trung bình MAE. Bên cạnh đó, sai số phần trăm tuyệt đối trung bình MAPE đóng vai trò là một thước đo độc lập về thang đo, giúp đánh giá độ chính xác của hệ thống dự báo dưới dạng tỷ lệ phần trăm so với giá trị thực nghiệm. Một mô hình được coi là đạt hiệu suất tối ưu và sở hữu khả năng dự báo vượt trội khi đồng thời đạt được giá trị A30 và R^2 cao nhất (tiến gần đến 1) và các chỉ số sai số (RMSE, MAE, MAPE) ở mức thấp nhất. Các công thức định nghĩa cho các thước đo thống kê này được trình bày như sau:

$$R^2 = 1 - \frac{\sum_1^N (P_i - A_i)^2}{\sum_1^N (\bar{A} - A_i)^2} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_1^N (P_i - A_i)^2} \tag{2}$$

$$MAE = \frac{1}{N} \sum_1^N |P_i - A_i| \tag{3}$$

$$MAPE = \frac{1}{N} \sum_1^N \frac{|P_i - A_i|}{A_i} \times 100\% \tag{4}$$

$$A30 = \frac{1}{N} \sum_1^N \begin{cases} 1, & \text{if } \left(\frac{|P_i - A_i|}{A_i} \right) \leq 0.3 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

Trong đó, P_i và A_i đại diện cho giá trị thực tế và giá trị dự đoán cho mẫu thứ i , \bar{A} giá trị trung bình

của tất cả các giá trị thực tế, và N là tổng số mẫu.

4. Kết quả và thảo luận

4.1. Xây dựng mô hình học máy

Quá trình xây dựng và đánh giá mô hình được thực hiện dựa trên quy trình chạy 30 mô phỏng Monte Carlo nhằm đảm bảo tính ổn định và

khách quan của các thuật toán dự báo. Hiệu suất của 5 thuật toán học máy, bao gồm ExtraTrees, LGBM, Gradient Boosting, Bagging và AdaBoost, được so sánh chi tiết dựa trên các chỉ số R² và RMSE trung bình qua các tập huấn luyện, tập kiểm tra và tập xác thực chéo 5 lần.

Bảng 3. Hiệu suất dự báo các mô hình trên tập huấn luyện và tập kiểm tra theo tiêu chí R² và RMSE

	Thuật toán	R ² Trung bình	RMSE Trung bình
Tập huấn luyện	ExtraTrees	0.990	4.588
	LGBM	0.931	12.180
	GradientBoosting	0.984	5.760
	Bagging	0.977	6.933
	AdaBoost	0.929	12.287
Tập kiểm tra	ExtraTrees	0.982	6.473
	LGBM	0.948	11.164
	GradientBoosting	0.981	6.713
	Bagging	0.973	7.992
	AdaBoost	0.936	12.277
Xác thực chéo 5 lần	ExtraTrees	0.927	11.895
	LGBM	0.831	18.411
	GradientBoosting	0.931	11.726
	Bagging	0.908	13.586
	AdaBoost	0.869	16.322

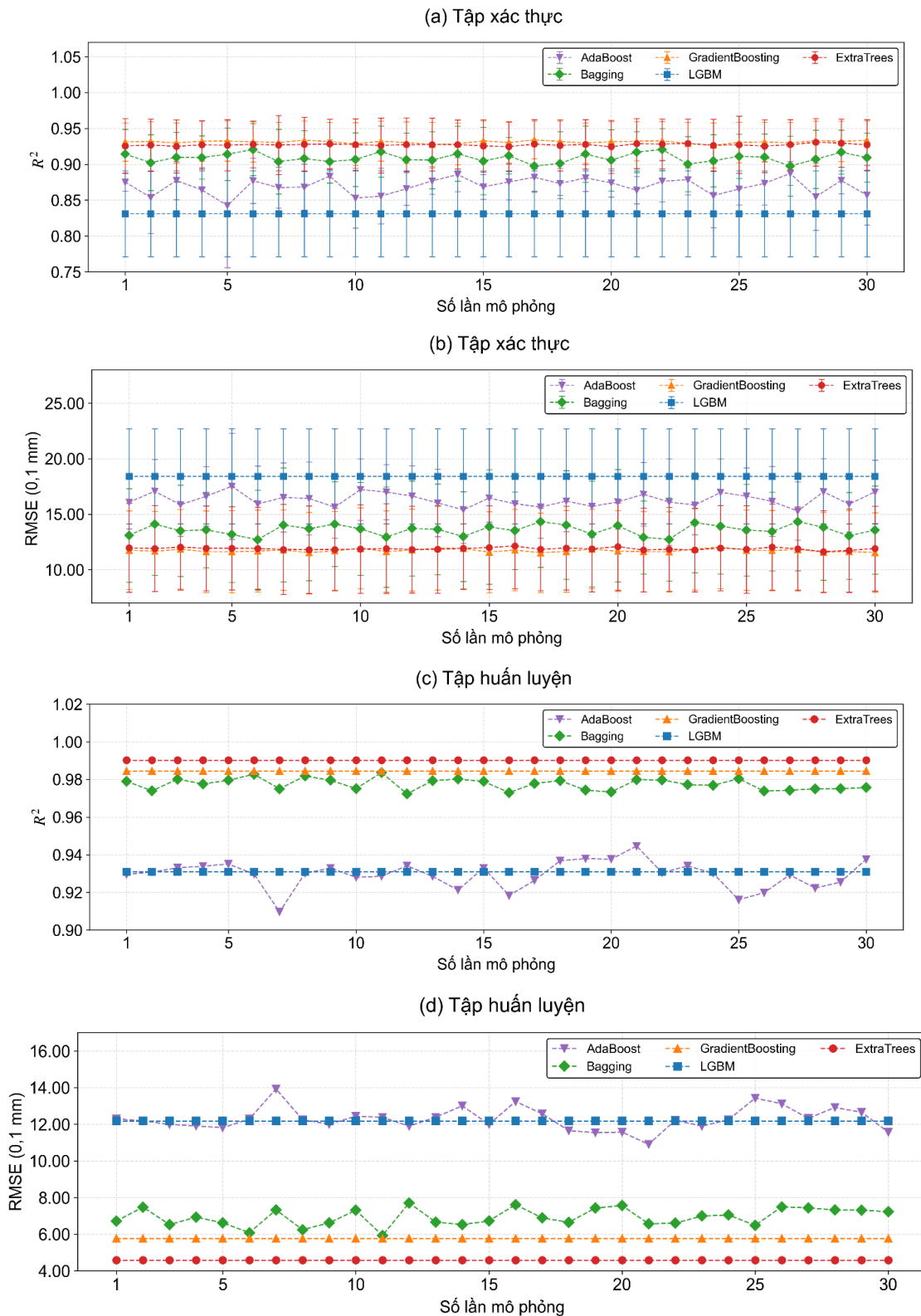
Kết quả thu được từ Hình 3 cung cấp trực quan về độ ổn định của các mô hình trong 30 lần mô phỏng. Tại tập huấn luyện (Hình 3c,d), các thuật toán ExtraTrees và Gradient Boosting thể hiện sự vượt trội với giá trị R² gần như tuyệt đối lần lượt là 0.990 và 0.984, đồng thời duy trì đường RMSE thấp và phẳng, minh chứng cho khả năng học các đặc trưng phức tạp từ dữ liệu nhựa đường cải tiến rất hiệu quả. Đối với tập xác thực chéo (Hình 3a,b), chỉ số phản ánh trung thực nhất năng lực dự báo thực tế của ExtraTrees và Gradient Boosting tiếp tục đứng đầu với dải sai số hẹp, khẳng định tính ổn định của các thuật toán này khi dữ liệu thực nghiệm biến động.

Phân tích định lượng từ Bảng 3 cho thấy mô hình ExtraTrees đạt hiệu suất cao nhất trên tập huấn luyện với R² = 0.990 và sai số thấp nhất với RMSE = 4.588. Đối với dữ liệu chưa được học tại tập kiểm tra, mô hình ExtraTrees vẫn đạt hiệu suất

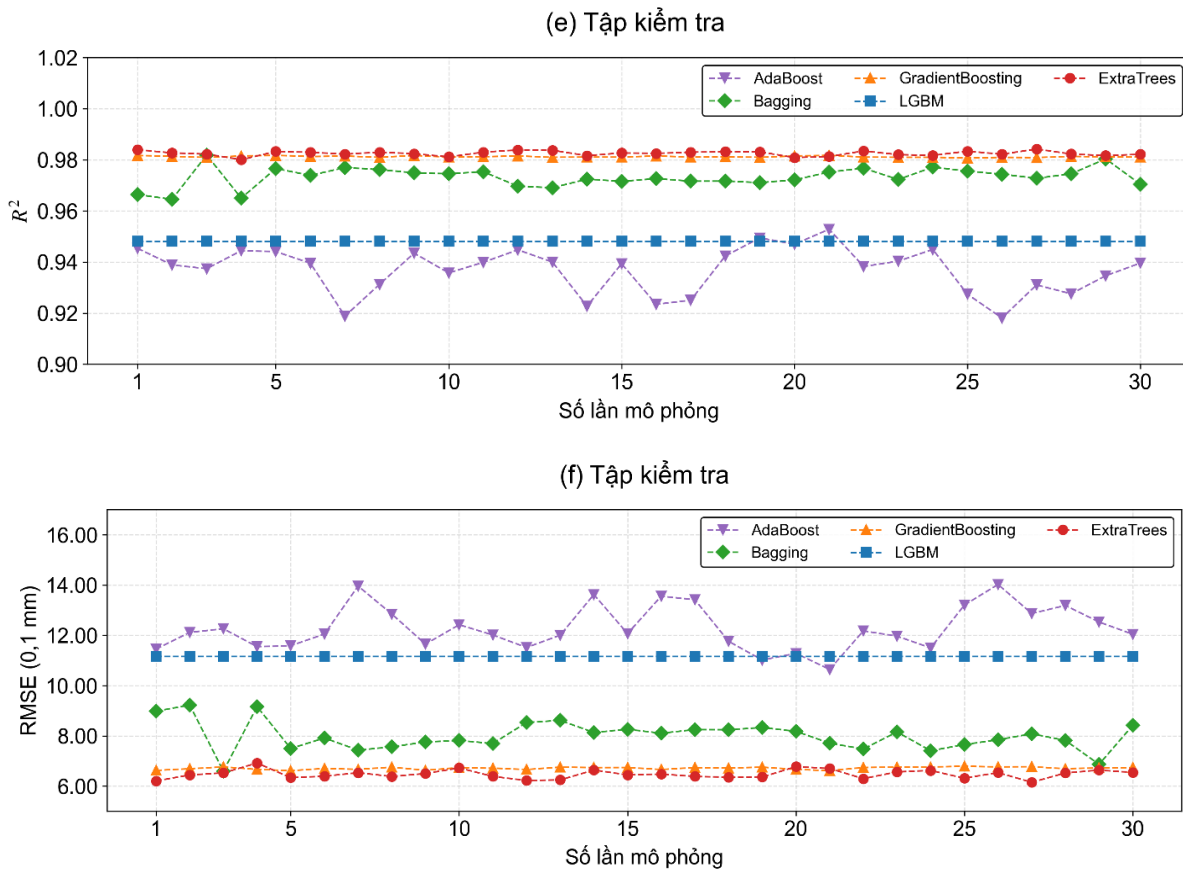
dự báo cao với R² = 0.982 và RMSE = 6.473, vượt trội hơn hẳn so với các mô hình khác như Bagging (R² = 0.973) hay AdaBoost (R² = 0.936). Đáng chú ý, trong quy trình xác thực chéo 5 lần, Gradient Boosting và ExtraTrees là hai thuật toán duy nhất duy trì được hệ số xác định trên 0.92, với giá trị trung bình trên tập xác thực lần lượt là 0.931 và 0.927. Ngược lại, thuật toán LGBM cho khả năng dự báo thấp nhất trên tập xác thực với hệ số R² thấp nhất (0.831) và sai số RMSE cao nhất (18.411). Điều này cho thấy với bộ dữ liệu đặc thù về nhựa đường có quy mô vừa và nhỏ, mô hình LGBM có thể chưa phát huy tối đa hiệu suất so với các thuật toán khác dựa trên cây ngẫu nhiên hoặc tăng cường truyền thống. Tổng hợp các kết quả so sánh, thuật toán ExtraTrees đã thể hiện các chỉ số dự báo vượt trội, ổn định và có sự cân bằng tốt nhất giữa khả năng học tập và tính tổng quát hóa trên mọi tập dữ liệu. Mặc dù Gradient Boosting có

kết quả xác thực chéo cao hơn một chút về R^2 , nhưng ExtraTrees lại cho thấy ưu thế về độ chính xác trên tập kiểm tra và khả năng giảm thiểu phương sai sai số. Do đó, ExtraTrees được lựa

chọn làm mô hình đại diện trong nghiên cứu này để tiến hành các phân tích sâu hơn về tầm quan trọng của các biến đầu vào và dự báo thực tế cho giá trị độ kim lún của nhựa đường cải tiến PE.



Hình 3. Hiệu suất dự báo các mô hình học máy sau 30 mô phỏng Monte Carlo trên tập xác thực chéo, tập huấn luyện và tập kiểm tra



Hình 3. (tiếp)

4.2. Kết quả dự đoán đại diện

Các chỉ số đánh giá thống kê của mô hình trên tập huấn luyện và tập kiểm tra được trình bày trên Bảng 4. Các thông số thống kê trên tập huấn luyện cho thấy tính chính xác trong khả năng dự đoán của mô hình với giá trị R^2 đạt 0.99. So sánh

các chỉ số đánh giá thống kê của mô hình trên tập huấn luyện và trên tập kiểm tra không cho thấy sự chênh lệch quá lớn giữa các thông số. Điều này cho thấy sự ổn định của mô hình khi làm việc với các tập dữ liệu mới, qua đó khẳng định tính khái quát của mô hình.

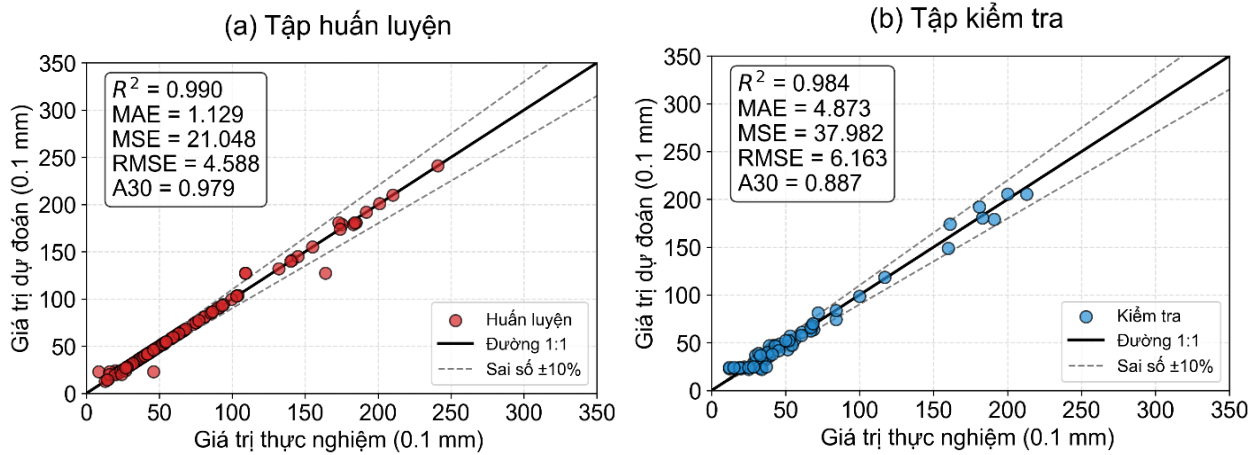
Bảng 4. Các chỉ số đánh giá thống kê của mô hình

Tiêu chí đánh giá	R^2	MAE	MSE	RMSE	A30
Tập huấn luyện	0.990	1.129	21.048	4.588	0.979
Tập kiểm tra	0.984	4.873	37.982	6.163	0.887

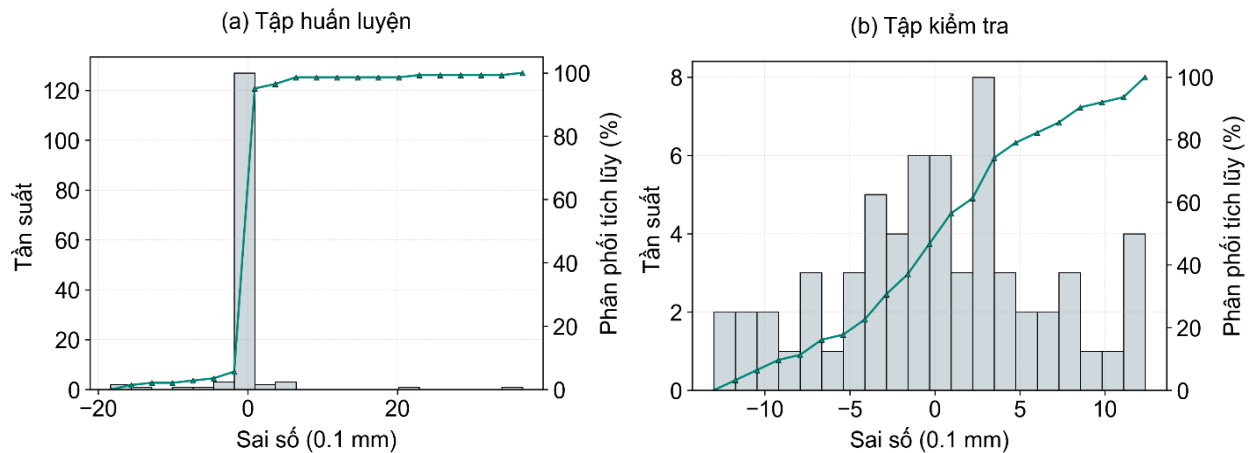
Kết quả hồi quy của mô hình trên tập huấn luyện và tập kiểm tra được biểu thị ở Hình 4a và 4b. Trên mỗi đồ thị, đường hồi quy cùng với đường giới hạn trên và đường giới hạn dưới (tương ứng với sai số 10% tính từ đường hồi quy) được thể hiện để chỉ ra sự tương quan giữa kết quả dự đoán và kết quả từ thực nghiệm. Đa số các điểm dữ liệu đều nằm trong khoảng 10% giới hạn trên và giới hạn dưới tính từ đường hồi quy, qua đó cho thấy sự tương quan gần giữa kết quả dự đoán và kết quả thực nghiệm, và khẳng định khả năng dự

đoán của mô hình với độ chính xác cao.

Ngoài ra, các giá trị sai số giữa giá trị dự đoán và giá trị thực nghiệm của tập huấn luyện và tập kiểm tra cũng được biểu thị ở Hình 5a, 5b. Mức sai số trên tập huấn luyện được biểu thị ở Hình 5a là rất nhỏ, với hơn 90% kết quả sai số xấp xỉ 0. Mức sai số trên tập kiểm tra đối với dữ liệu kim lún nằm trong khoảng [-1,5 ; 1,5] 0.1 mm, với hơn 80% kết quả sai số tập trung trong khoảng [-5; 5] 0.1 mm. Qua đó cho thấy khả năng dự đoán chính xác của mô hình.



Hình 4. Kết quả hồi quy của mô hình trên 2 tập dữ liệu: (a) Tập huấn luyện, (b) Tập kiểm tra.



Hình 5. Kết quả sai số mô phỏng của mô hình trên 2 tập dữ liệu: (a) Tập huấn luyện, (b) Tập kiểm tra.

Phân tích sâu hơn về đặc điểm phân phối sai số cho thấy mô hình ExtraTrees thể hiện năng lực dự báo ổn định và chính xác nhất ở dải độ kim lún thấp dưới 100 (0.1 mm), tương ứng với các loại nhựa đường có độ cứng cao. Điều này có thể giải thích do mật độ dữ liệu thực nghiệm tập trung cao ở vùng giá trị này, giúp mô hình học được các đặc trưng vật liệu tốt hơn. Ngược lại, tại dải độ kim lún

cao trên 150 (0.1 mm) đại diện cho nhựa đường mềm, sai số có xu hướng dao động lớn hơn do tính không đồng nhất của dữ liệu biên. Nhận xét này chỉ ra rằng mô hình đặc biệt hiệu quả trong việc thiết kế cấp phối các loại nhựa đường biến tính PE yêu cầu độ cứng lớn để kháng hàn lún.

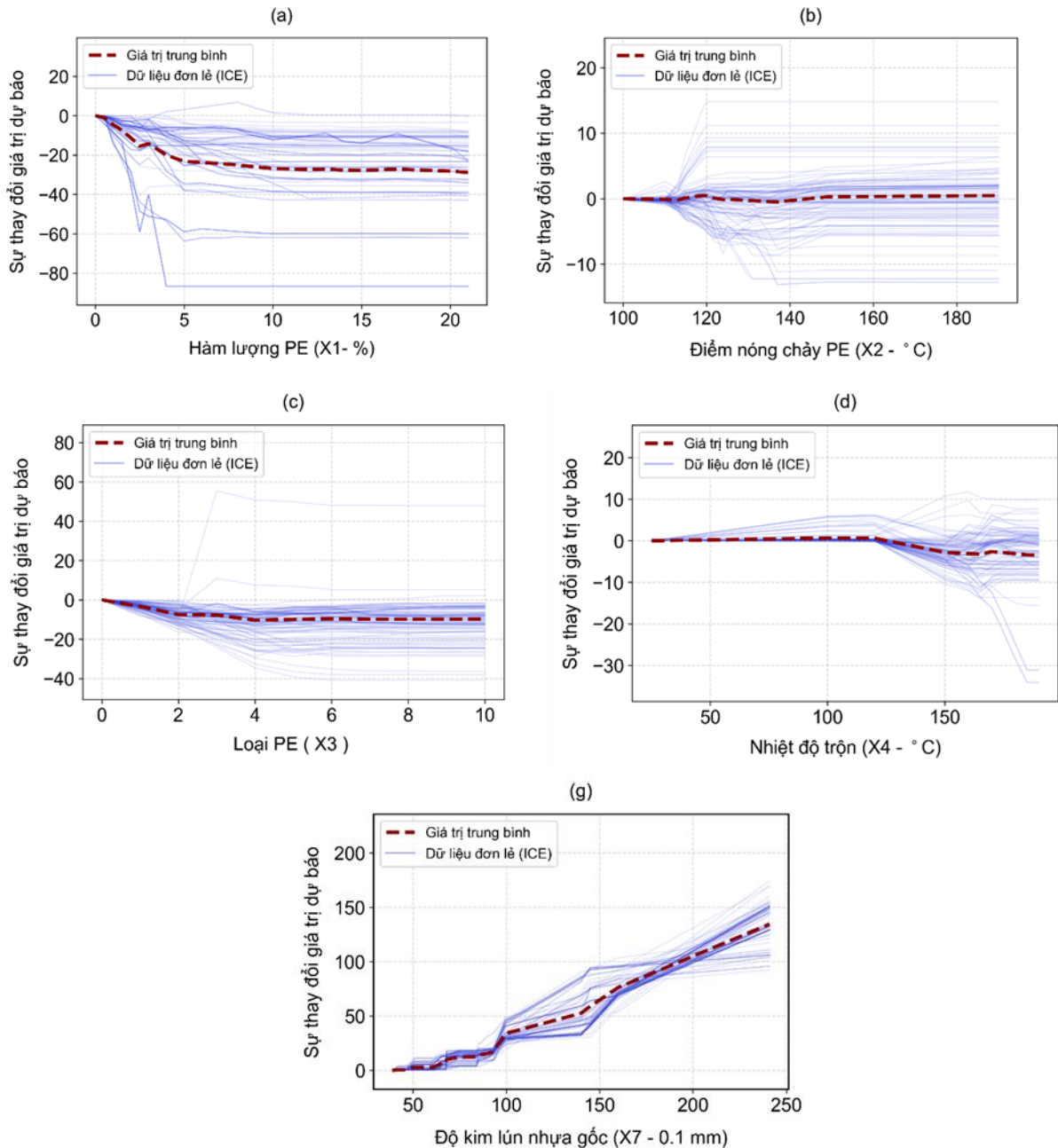
4.3. Phân tích ảnh hưởng của các yếu tố đầu vào đến hiệu suất dự báo của mô hình

Bảng 5. Bảng xếp hạng độ quan trọng của các biến

Xếp hạng	Đặc trưng	Nhỏ nhất	Lớn nhất	Hiệu số
1	Độ kim lún nhựa gốc (X7 - 0.1 mm)	36.36	174.16	137.80
2	Hàm lượng PE (X1 - %)	52.77	77.76	24.99
3	Thời gian trộn (X5 - Phút)	58.21	62.94	4.73
4	Loại PE (X3)	61.47	66.09	4.62
5	Điểm nóng chảy PE (X2 - °C)	59.76	63.88	4.12
6	Tốc độ trộn (X6 - Vòng/phút)	60.17	62.86	2.69
7	Nhiệt độ trộn (X4 - °C)	59.97	62.19	2.22

Để làm rõ cơ chế tác động của các yếu tố thành phần và quy trình chế tạo đến đặc tính của nhựa đường cải tiến, phân tích PDP dựa trên mô hình ExtraTrees được thực hiện và đưa ra kết quả định lượng tại Bảng 5. Kết quả trong bảng này cho thấy độ kim lún nhựa gốc (X7) là yếu tố có ảnh hưởng lớn nhất đến giá trị dự báo với hiệu số tác động lên tới 137.80. Điều này khẳng định rằng tính

chất của nhựa đường nền đóng vai trò cốt lõi trong việc hình thành cấu trúc và độ cứng của sản phẩm cuối cùng. Kết quả này tương đồng với nghiên cứu của Tang và cộng sự [38], khi chỉ ra rằng nhựa đường gốc quyết định khung độ nhớt ban đầu và khả năng trương nở của polymer, từ đó trực tiếp chi phối các chỉ tiêu cơ lý của sản phẩm nhựa đường cải tiến.



Hình 6. Biểu đồ phụ thuộc một phần (PDP-1D) mô tả ảnh hưởng của các yếu tố X1 - X7 đến giá trị dự báo độ kim lún Y.

Phân tích biểu đồ PDP-1D (Hình 6) cho thấy mối quan hệ phi tuyến tính rõ rệt giữa các biến đầu vào và độ kim lún dự báo. Đối với độ kim lún nhựa

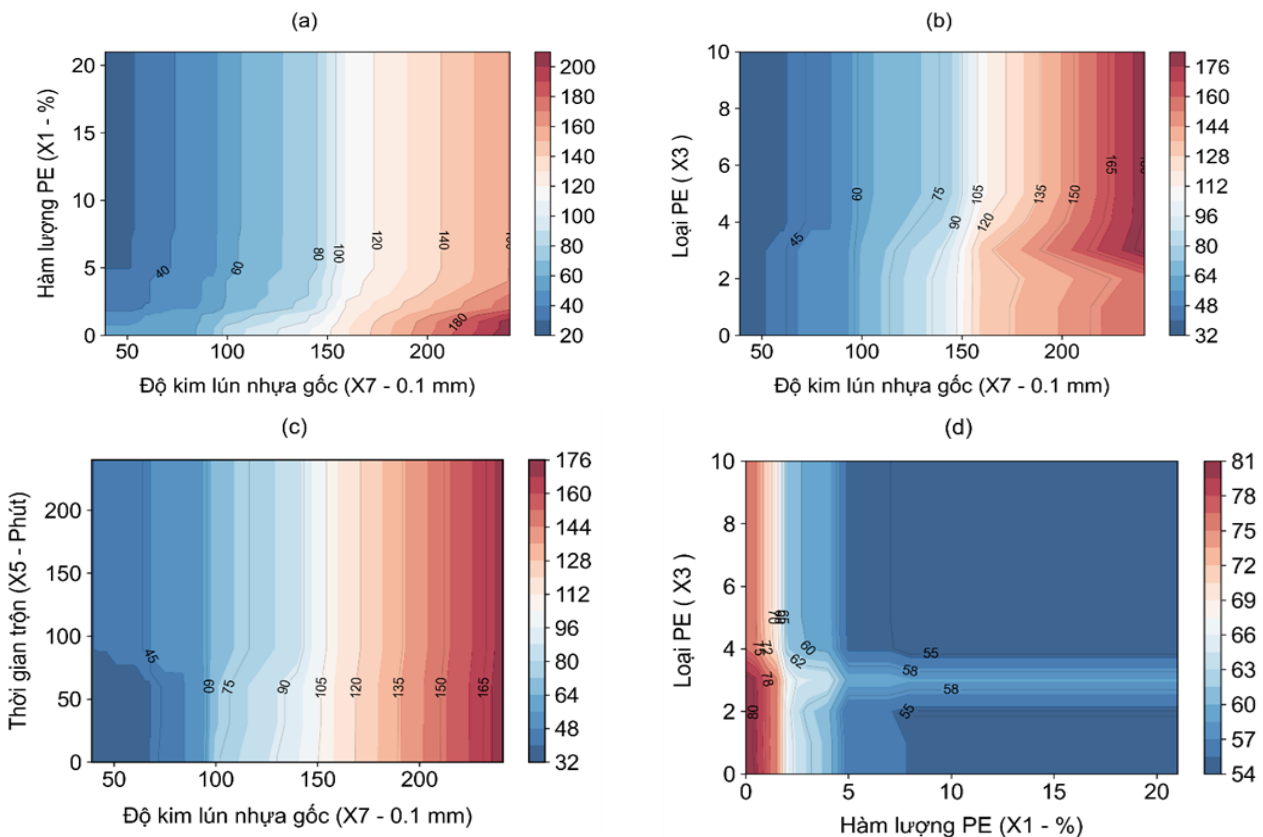
gốc (X7) (Hình 6g), giá trị dự báo tăng mạnh và gần như tuyến tính khi X7 vượt ngưỡng 100 (0.1 mm), phản ánh sự duy trì đặc tính từ nhựa nền sang

nhựa cải tiến. Ngược lại, hàm lượng PE (X1) (Hình 6a) cho thấy xu hướng làm giảm độ kim lún nhanh chóng trong khoảng từ 0% đến 5%. Khi hàm lượng PE vượt quá ngưỡng này, sự thay đổi của độ kim lún dần trở nên ổn định, cho thấy hiện tượng bão hòa của polymer trong hỗn hợp. Xu hướng này phù hợp với các quan sát thực nghiệm của Hýzl và cộng sự [39], khi chứng minh rằng hiệu quả của phụ gia polymer thường đạt đến trạng thái bão hòa sau một ngưỡng hàm lượng nhất định, nơi mà mạng lưới polymer đã được hình thành liên tục trong nền nhựa. Các yếu tố kỹ thuật như nhiệt độ trộn (X4) (Hình 6d) và thời gian trộn (X5) (Hình 6e) cũng thể hiện ngưỡng tối ưu. Sự gia tăng các giá trị này giúp cải thiện độ phân tán của PE. Tuy nhiên, sự gia tăng quá cao của các biến này có thể dẫn đến hiện tượng lão hóa nhiệt hoặc phân hủy chuỗi polymer, làm giảm hiệu quả cải thiện tính chất vật liệu [40,41].

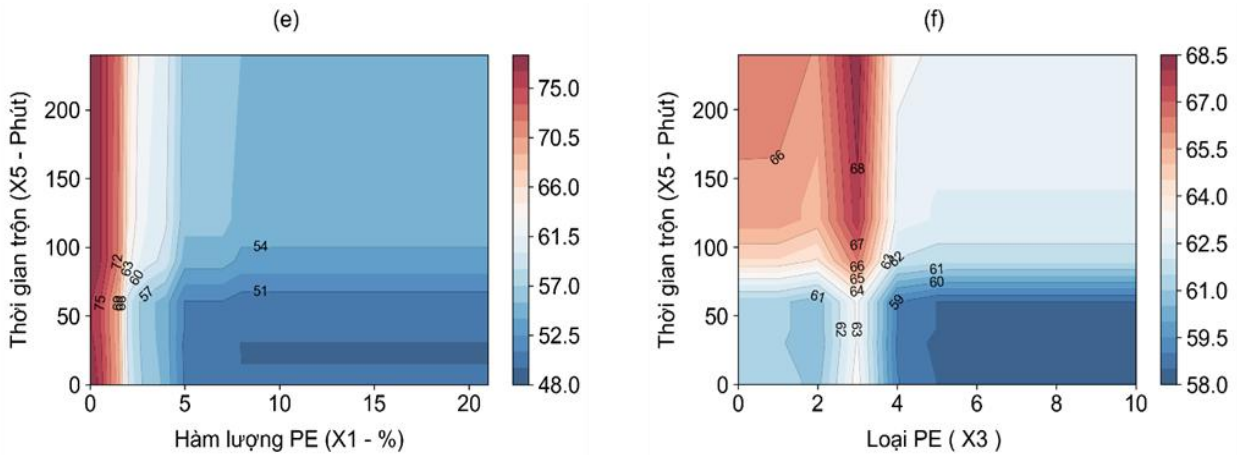
Tại Hình 6b, điểm nóng chảy PE (X2) cho thấy tác động không đáng kể đến giá trị dự báo, với đường trung bình gần như đi ngang trong toàn

bộ dải khảo sát. Đối với loại PE (X3) tại Hình 6c, giá trị độ kim lún giảm nhẹ khi thay đổi mã loại từ 0 đến 4 và sau đó duy trì trạng thái ổn định, phản ánh sự khác biệt về khả năng tương thích của các dòng polymer khác nhau đối với nhựa đường. Cuối cùng, tốc độ trộn (Hình 6f) thể hiện vai trò quan trọng trong giai đoạn phân tán ban đầu khi làm thay đổi rõ rệt giá trị dự báo trong khoảng 0 - 1000 vòng/phút, trước khi đạt đến ngưỡng bão hòa và không gây ra sự biến động lớn ở các tốc độ cao hơn.

Ngoài ra, sự tương tác giữa các cặp biến quan trọng nhất được làm rõ thông qua biểu đồ PDP-2D (Hình 7). Về mặt trực quan, các biểu đồ nhiệt trong Hình 7 minh họa giá trị độ kim lún dự báo thông qua thang đo màu sắc, trong đó màu xanh đại diện cho giá trị độ kim lún thấp (vật liệu cứng hơn), màu đỏ đại diện giá trị độ kim lún cao (vật liệu mềm hơn). Sự phân bố của các dải màu này cho phép xác định nhanh chóng các vùng tương tác tối ưu giữa hai biến đầu vào để đạt được hiệu suất mong muốn.



Hình 7. Biểu đồ phụ thuộc từng phần (PDP-2D) mô tả ảnh hưởng của các cặp biến đầu vào đến dự báo giá trị đầu ra



Hình 7. (tiếp)

Kết quả tại Hình 7a chỉ ra rằng sự kết hợp giữa nhựa đường gốc có độ kim lún thấp và hàm lượng PE cao tạo ra sản phẩm có độ cứng lớn nhất (độ kim lún thấp nhất). Điều này chứng minh hiệu ứng cộng hưởng mạnh giữa pha nền nhựa cứng và pha gia cường polymer. Đặc biệt, sự tương tác giữa loại PE và độ kim lún nhựa gốc (Hình 7b) khẳng định rằng các loại PE với cấu trúc mạch và độ kết tinh khác nhau sẽ có mức độ tương thích khác biệt với các phân đoạn hóa học trong nhựa nền [38,39]. Sự tương thích này ảnh hưởng trực tiếp đến khả năng phân tán và hình thái học của nhựa đường, từ đó quyết định hiệu quả cải thiện độ kim lún.

Bên cạnh các yếu tố về thành phần, sự tương tác giữa điều kiện chế tạo và đặc tính phụ gia cũng đóng vai trò quan trọng trong việc hình thành tính chất của nhựa đường cải tiến PE. Tại Hình 7c, biểu đồ tương tác giữa thời gian trộn (X5) và độ kim lún nhựa gốc (X7) cho thấy giá trị độ kim lún dự báo phụ thuộc chặt chẽ vào độ cứng ban đầu của chất kết dính. Khi sử dụng nhựa đường gốc mềm hơn (X7 cao), việc gia tăng thời gian trộn đóng vai trò quan trọng trong việc đảm bảo polymer phân tán đồng nhất để đạt được độ cứng mong muốn. Ngược lại, đối với nhựa gốc cứng, giá trị dự báo đạt trạng thái ổn định khá nhanh, cho thấy sự bão hòa tương tác giữa PE và nhựa nền diễn ra sớm hơn, giúp tối ưu hóa thời gian chế tạo mẫu mà vẫn đảm bảo chỉ tiêu kỹ thuật. Kết quả này cũng tương đồng với các quan sát về tối ưu hóa quy trình của Khan và cộng sự [41].

Mối quan hệ tương tác giữa loại PE (X3) và hàm lượng PE (X1) tại Hình 7d phản ánh sự khác biệt về hiệu quả cải thiện giữa các dòng phụ gia. Kết quả chỉ ra rằng mỗi loại PE với cấu trúc mạch nhánh và độ kết tinh khác nhau sẽ có một ngưỡng hàm lượng tối ưu riêng biệt để làm giảm độ kim lún. Một số loại PE cho thấy hiệu quả làm cứng vật liệu rõ rệt ngay ở hàm lượng thấp, trong khi các loại PE khác yêu cầu nồng độ cao hơn để hình thành mạng lưới polymer liên tục trong nền nhựa đường. Kết quả này phù hợp với nghiên cứu [38,39] về sự tương thích của polymer trong nhựa đường.

Sự tương tác giữa thời gian trộn và hàm lượng PE (Hình 7e) làm nổi bật vai trò của động học quá trình hòa trộn. Để đạt được độ kim lún thấp nhất (vùng màu xanh đậm), việc kết hợp giữa hàm lượng PE cao và thời gian trộn dài là cần thiết để đảm bảo sự trương nở hoàn toàn của polymer. Tuy nhiên, mô hình cũng chỉ ra vùng bão hòa nơi mà việc tiếp tục gia tăng thời gian trộn không còn mang lại sự thay đổi đáng kể về độ cứng, giúp tránh hiện tượng lão hóa nhiệt không cần thiết cho vật liệu. Kết quả này tương đồng với nghiên cứu [41].

Cuối cùng, Hình 7f minh họa mối quan hệ giữa thời gian trộn (X5) và loại PE (X3). Biểu đồ này khẳng định mức độ tương thích khác nhau của các loại PE đối với quy trình chế tạo mẫu. Một số loại PE có cấu trúc mạch phức tạp đòi hỏi thời gian trộn dài hơn để phá vỡ các liên kết nội tại và đạt được độ phân tán tối ưu. Những phát hiện này thông qua mô hình ExtraTrees cho phép các nhà thực nghiệm thiết kế quy trình chế tạo riêng biệt

cho từng loại phụ gia, đảm bảo tính kinh tế và kỹ thuật trong sản xuất nhựa đường cải tiến PE.

Mặc dù độ kim lún của nhựa đường gốc (X7) được xác định là biến quan trọng nhất, đóng vai trò nền tảng định hình đặc tính của sản phẩm cuối cùng, giá trị của mô hình nằm ở khả năng nắm bắt chính xác tác động của quá trình biến tính. Việc đưa X7 vào mô hình đóng vai trò như một biến chuẩn hóa, cho phép thuật toán phân tách ảnh hưởng của bản chất nhựa nền khỏi ảnh hưởng của phụ gia. Kết quả phân tích PDP cho thấy, dù có trọng số thấp hơn X7, các biến như hàm lượng PE (X1) và loại PE (X3) vẫn tạo ra những thay đổi rõ rệt và phi tuyến tính lên độ kim lún. Điều này minh chứng rằng mô hình không chỉ dựa vào đặc tính vật liệu gốc mà thực sự đã học được cơ chế tương tác giữa PE và nhựa đường, cho phép dự báo hiệu quả sự cải thiện đặc tính vật liệu khi thay đổi các thông số biến tính.

Như vậy, khả năng của mô hình ExtraTrees trong việc nhận diện và mô hình hóa các mối tương tác đa biến phức tạp không chỉ tái xác thực các quy

luật vật lý dưới góc độ khoa học vật liệu, mà còn khẳng định ưu thế vượt trội của mô hình học máy với khả năng giải thích (XAI). Kết quả này cung cấp một công cụ phân tích tin cậy, giúp chuyển đổi từ các quy trình thiết kế nhựa đường biến tính dựa trên kinh nghiệm sang cách tiếp cận tối ưu hóa dựa trên dữ liệu.

5. Giao diện người dùng

Để tối ưu hóa quy trình thiết kế nhựa đường cải tiến mà không đòi hỏi các phép tính phức tạp, mô hình ExtraTrees tối ưu đã được tích hợp vào một nền tảng phần mềm độc lập. Một phần mềm với giao diện thân thiện với người dùng được phát triển để tăng tính ứng dụng của nghiên cứu vào thực tế sử dụng (Hình 8). Phần mềm được xây dựng sử dụng ngôn ngữ lập trình Python và tích hợp mô hình được phát triển từ nghiên cứu với các biến đầu vào và đầu ra của bộ dữ liệu. Người dùng có thể điều chỉnh các thông số đầu vào, bằng cách kéo các thanh tương tác đến giá trị mong muốn, sau đó thực hiện lệnh tính toán để phần mềm cho ra giá trị độ kim lún dự báo.



Hình 8. Giao diện người dùng dự báo độ kim lún nhựa đường PE

Ngoài ra, phần mềm cũng có thể được sử dụng để tối ưu hóa các thông số đầu vào. Kết quả dự đoán của phần mềm có thể được sử dụng như giá trị tham khảo ban đầu, để từ đó tinh chỉnh các thông số đầu vào một cách phù hợp, nhằm có được giá trị kim lún mong muốn. Với thiết kế đơn giản và thân thiện với người dùng, phần mềm là một công cụ hữu hiệu phục vụ trong các việc thiết kế và nghiên cứu nhựa đường biến tính có sử dụng phụ gia PE. Mặc dù giao diện phần mềm hiện tại dựa trên mô hình có độ chính xác dự báo cao, nghiên cứu xác định việc kiểm chứng bằng các mẫu thí nghiệm đúc thực tế tại phòng thí nghiệm là bước đi quan trọng tiếp theo. Trong giai đoạn tới, nghiên cứu sẽ thực hiện các thí nghiệm độc lập trên các nguồn nhựa đường và phụ gia PE cụ thể tại thị trường nội địa. Kết quả này sẽ đóng vai trò là bộ dữ liệu xác nhận chéo cuối cùng để tinh chỉnh các trọng số của mô hình, giúp công cụ trở nên hoàn thiện, tin cậy và đáp ứng tốt các yêu cầu đặc thù trong thực tiễn thiết kế cấp phối nhựa đường biến tính tại Việt Nam.

6. Kết luận

Thông qua các kết quả nghiên cứu và thảo luận, bài báo rút ra những kết luận chính sau đây. Trước hết, nghiên cứu đã chứng minh các thuật toán học máy, đặc biệt là ExtraTrees, có khả năng dự báo chính xác độ kim lún của nhựa đường cải tiến PE với độ tin cậy rất cao. Mô hình này đạt được sự cân bằng tối ưu giữa khả năng học tập dữ liệu thực nghiệm và tính tổng quát hóa trên các tập dữ liệu chưa biết, thể hiện qua hệ số xác định R^2 đạt 0.982 trên tập kiểm tra. Việc ứng dụng kỹ thuật trí tuệ nhân tạo có khả năng giải thích thông qua biểu đồ PDP đã làm rõ cơ chế tác động phi tuyến tính của các biến đầu vào. Trong đó, độ kim lún của nhựa đường gốc (X7) và hàm lượng phụ gia PE (X1) là hai được xác định là hai biến quan trọng nhất ảnh hưởng đến hiệu suất dự báo đầu ra. Phân tích chuyên sâu cho thấy hàm lượng PE thể hiện tác động làm cứng nhựa đường rõ rệt nhất trong khoảng từ 0% đến 5%, sau đó hiệu quả dần bão hòa do sự hình thành mạng lưới polymer liên tục trong nền nhựa đường. Đồng thời, sự tương tác

phức tạp giữa các loại PE khác nhau và điều kiện chế tạo như nhiệt độ hay thời gian trộn cũng đóng vai trò then chốt trong việc hình thành tính chất của nhựa đường cải tiến. Việc tối ưu hóa các thông số kỹ thuật này thông qua mô hình dự báo không chỉ giúp ngăn ngừa hiện tượng lão hóa nhiệt mà còn đảm bảo tính kinh tế trong quá trình sản xuất. Cuối cùng, giao diện người dùng trực quan được phát triển trong nghiên cứu này đóng vai trò là một công cụ hỗ trợ đắc lực cho kỹ sư vật liệu, cho phép dự báo nhanh kết quả thực nghiệm và tối ưu hóa thiết kế cấp phối nhựa đường cải tiến PE trong thực tiễn. Hơn nữa, hướng nghiên cứu này sở hữu tiềm năng mở rộng rất lớn khi có thể áp dụng quy trình phân tích và các thuật toán học máy tương tự để dự báo các chỉ tiêu cơ lý khác cho đa dạng các loại nhựa đường biến tính polymer khác.

Việc tích hợp dữ liệu thực nghiệm quy mô lớn kết hợp với các kỹ thuật giải thích mô hình tiên tiến trong tương lai sẽ cho phép xây dựng các hệ thống hỗ trợ quyết định toàn diện, giúp tối ưu hóa vật liệu và thúc đẩy quá trình chuyển đổi số trong lĩnh vực xây dựng hạ tầng giao thông.

Tài liệu tham khảo

- [1] A. Behnood, M.M. Gharehveran. (2019). Morphology, rheology, and physical properties of polymer-modified asphalt binders, *European Polymer Journal*, 112, 766–791. <https://doi.org/10.1016/j.eurpolymj.2018.10.049>
- [2] E. Masad, K.L. Roja, A. Rehman, A. Abdala. (2020). A review of asphalt modification using plastics: a focus on polyethylene. *Texas A&M University: Qatar, Doha*. <https://doi.org/10.13140/RG.2.2.36633.77920>
- [3] A.I. Al-Hadidy, T. Yi-qiu. (2009). Effect of polyethylene on life of flexible pavements. *Construction and Building Materials*, 23(3), 1456–1464. <https://doi.org/10.1016/j.conbuildmat.2008.07.004>
- [4] A. Issa, A.A. Sheikah, R. Nazzal, A. Jaber. (2022). Study the effect of adding high-density polyethylene on the asphalt mixture.

- Proceedings of International Structural Engineering and Construction*, 9(1). [http://www.doi.org/10.14455/ISEC.2022.9\(1\).M AT-01](http://www.doi.org/10.14455/ISEC.2022.9(1).M AT-01)
- [5] S. Ho, R. Church, K. Klassen, B. Law, D. MacLeod, L. Zanzotto. (2006). Study of recycled polyethylene materials as asphalt modifiers, *Canadian Journal of Civil Engineering*, 33(8), 968–981. <https://doi.org/10.1139/106-044>
- [6] D. Wang, A. Baliello, L. Poulikakos, K. Vasconcelos, M.R. Kakar, G. Giancontieri, E. Pasquini, L. Porot, M. Tušar, C. Riccardi, M. Paseto, D.L. Presti, A.C. Falchetto. (2022). Rheological properties of asphalt binder modified with waste polyethylene: An interlaboratory research from the RILEM TC WMR. *Resources, Conservation and Recycling*, 186, 106564. <https://doi.org/10.1016/j.resconrec.2022.106564>
- [7] A.S. Sk, K.S.B. Prasad. (2012). Utilization of waste plastic as a strength modifier in surface course of flexible and rigid pavements. *International Journal of Engineering Research Application*, 2(4), 1185–1191.
- [8] A.A. Yousefi. (2003). Polyethylene dispersions in bitumen: The effects of the polymer structural parameters. *Journal of Applied Polymer Science*, 90(12), 3183–3190. <https://doi.org/10.1002/app.12942>
- [9] N.Z. Habib, I. Kamaruddin, M. Napiah, I.M. Tan. (2011). Rheological properties of polyethylene and polypropylene modified bitumen, *International Journal of Civil and Environmental Engineering*, 3(2), 96–100.
- [10] D. Casey, C. McNally, A. Gibney, M.D. Gilchrist. (2008). Development of a recycled polymer modified binder for use in stone mastic asphalt. *Resources, Conservation and Recycling*, 52(10), 1167–1174. <https://doi.org/10.1016/j.resconrec.2008.06.002>
- [11] H.-G.T. Hoang, T.-A. Nguyen, H.-L. Nguyen, H.-B. Ly. (2022). Neural network approach for GO-modified asphalt properties estimation. *Case Studies in Construction Materials*, 17, e01617. <https://doi.org/10.1016/j.cscm.2022.e01617>
- [12] H.-G.T. Hoang, T.-A. Nguyen, H.-L. Nguyen, H.-B. Ly. (2025). An Effective Extreme Gradient Boosting Approach to Predict the Physical Properties of Graphene Oxide Modified Asphalt. *International Journal of Pavement Research and Technology*. <https://doi.org/10.1007/s42947-025-00636-y>
- [13] H.-G.T. Hoang, H.-V.T. Mai, H.L. Nguyen, H.-B. Ly. (2024). Application of extreme gradient boosting in predicting the viscoelastic characteristics of graphene oxide modified asphalt at medium and high temperatures, *Frontiers of Structural and Civil Engineering*, 18, 899–917. <https://doi.org/10.1007/s11709-024-1025-y>
- [14] P. Geurts, D. Ernst, L. Wehenkel. (2006). Extremely randomized trees. *Machine Learning*, 63, 3–42. <https://doi.org/10.1007/s10994-006-6226-1>
- [15] C. Ouyang, S. Wang, Y. Zhang, Y. Zhang. (2006). Low-density polyethylene/silica compound modified asphalts with high-temperature storage stability. *Journal of Applied Polymer Science*, 101(1), 472–479. <https://doi.org/10.1002/app.23029>
- [16] S. Sharma, S. Sharma, N. Upadhyay. (2019). Composition Based Physicochemical Analysis of Modified Bitumen by High-density polyethylene (HDPE) and Low-density polyethylene (LDPE)., *Oriental Journal of Chemistry*, 35(3), 1167-1173. <https://doi.org/10.13005/ojc/350336>
- [17] M.A. Vargas, M.A. Vargas, A. Sánchez-Sólis, O. Manero. (2013). Asphalt/polyethylene blends: Rheological properties, microstructure and viscosity modeling. *Construction and Building Materials*, 45, 243–250. <https://doi.org/10.1016/j.conbuildmat.2013.03.064>

- [18] D.A. Gama, Y. Yan, J.K.G. Rodrigues, R. Roque. (2018). Optimizing the use of reactive terpolymer, polyphosphoric acid and high-density polyethylene to achieve asphalt binders with superior performance. *Construction and Building Materials*, 169, 522–529. <https://doi.org/10.1016/j.conbuildmat.2018.02.206>
- [19] B. Singh, L. Kumar, M. Gupta, G.S. Chauhan. (2013). Polymer-modified bitumen of recycled LDPE and maleated bitumen. *Journal of Applied Polymer Science*, 127, 67–78. <https://doi.org/10.1002/app.36810>
- [20] C. Fang, T. Li, Z. Zhang, X. Wang. (2008). Combined modification of asphalt by waste PE and rubber. *Polymer Composites*, 29(10) (2008), 1183–1187. <https://doi.org/10.1002/pc.20424>
- [21] M. Attaelmanan, C.P. Feng, A.I. Al-Hadidy. (2011). Laboratory evaluation of HMA with high density polyethylene as a modifier, *Construction and Building Materials*, 25(5), 2764–2770. <https://doi.org/10.1016/j.conbuildmat.2010.12.037>
- [22] L. Jun, Z. Yuxia, Z. Yuzhen. (2008). The research of GMA-g-LDPE modified Qinhuangdao bitumen. *Construction and Building Materials*, 22(6), 1067–1073. <https://doi.org/10.1016/j.conbuildmat.2007.03.007>
- [23] C. Fang, R. Yu, Y. Li, M. Zhang, J. Hu, M. Zhang. (2013). Preparation and characterization of an asphalt-modifying agent with waste packaging polyethylene and organic montmorillonite. *Polymer Testing*, 32(5), 953–960. <https://doi.org/10.1016/j.polymertesting.2013.04.006>
- [24] S. Nizamuddin, M. Jamal, R. Gravina, F. Giustozzi. (2020). Recycled plastic as bitumen modifier: The role of recycled linear low-density polyethylene in the modification of physical, chemical and rheological properties of bitumen. *Journal of Cleaner Production*, 266, 121988. <https://doi.org/10.1016/j.jclepro.2020.121988>
- [25] S.F. Wong, A.A. Htwe, S.H. Oh, T.Y. Leo, J. Cheng, B.K. Tay. (2017). Utilization of waste plastics in stone mastic asphalt for infrastructural applications. *Materials Science Forum*, 905, 55–59. <https://doi.org/10.4028/www.scientific.net/MSF.902.55>
- [26] C. Fang, C. Wu, J. Hu, R. Yu, Z. Zhang, L. Nie, S. Zhou, X. Mi. (2014). Pavement properties of asphalt modified with packaging-waste polyethylene. *Journal of Vinyl Additive Technology*, 20(1), 31–35. <https://doi.org/10.1002/vnl.21328>
- [27] K. Yan, H. Xu, L. You. (2015). Rheological properties of asphalts modified by waste tire rubber and reclaimed low density polyethylene, *Construction and Building Materials*, 83, 143–149. <https://doi.org/10.1016/j.conbuildmat.2015.02.092>
- [28] C. Fang, M. Zhang, R. Yu, X. Liu. (2015). Effect of preparation temperature on the aging properties of waste polyethylene modified asphalt. *Journal of Materials Science & Technology*, 31(3), 320–324. <https://doi.org/10.1016/j.jmst.2014.04.019>
- [29] L. Costa, S. Fernandes, H. Silva, J. Oliveira. (2017). Study of the interaction between asphalt and recycled plastics in new polymer modified binders (PMB). *Ciência & Tecnologia Dos Materiais*, 29(1), e192–e197. <https://doi.org/10.1016/j.ctmat.2016.04.005>
- [30] D.B. Eme, C. Nwaobakata. (2019). Effect of low density polyethylene as bitumen modifier on some properties of hot mix asphalt. *Nigerian Journal of Technology*, 38(1), 1–7. <http://dx.doi.org/10.4314/njt.v38i1.1>
- [31] H.A.A. Gibreil, C.P. Feng. (2017). Effects of high-density polyethylene and crumb rubber powder as modifiers on properties of hot mix asphalt. *Construction and Building Materials*, 142, 101–108

- <https://doi.org/10.1016/j.conbuildmat.2017.03.062>.
- [32] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html> (accessed January 16, 2026).
- [33] J.H. Friedman. (2001). Greedy function approximation: a gradient boosting machine, *Annals of Statistics*, 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- [34] L. Breiman. (1996). Bagging predictors. *Machine Learning*, 24, 123–140. <https://doi.org/10.1007/BF00058655>
- [35] R.E. Schapire. (2013). Explaining AdaBoost. *Empirical Inference*, Springer Berlin Heidelberg, Berlin, Heidelberg, 37–52. https://doi.org/10.1007/978-3-642-41136-6_5.
- [36] C.P. Robert, G. Casella. (2000). Monte Carlo Statistical Methods. *Technometrics*, 42(4). <https://doi.org/10.2307/1270959>
- [37] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*.
- [38] N. Tang, W. Huang, G. Hao. (2021). Effect of aging on morphology, rheology, and chemical properties of highly polymer modified asphalt binders. *Construction and Building Materials*, 281, 122595. <https://doi.org/10.1016/j.conbuildmat.2021.122595>
- [39] P. Hýzl, O. Dašek, M. Varaus, D. Stehlík, P. Coufalík, J. Dašková, I. Krčmová, P. Nekulova. (2016). The effect of compaction degree and binder content on performance properties of asphalt mixtures. *The Baltic Journal of Road and Bridge Engineering*, 11(3), 222–232. <https://doi.org/10.3846/bjrbe.2016.26>
- [40] A.M. Zaltuom. (2018). A review study of the effect of air voids on asphalt pavement life. *Proceedings of First Conference for Engineering Sciences and Technology*, 618–625. <https://doi.org/10.21467/proceedings.4.29>
- [41] M.A. Khan, M.S. Khan, B. Nasir, A. Khan, M.M.S. Sabri, M. Ahmad, W. Qamar, R.A. Gonzalez-Lezcano. (2024). Performance optimization of asphalt pavements using binder film thickness as a criterion in innovative mix design compared to Marshall and Superpave methods, *Frontiers in Materials*, 11, 1488310. <https://doi.org/10.3389/fmats.2024.1488310>