



Semnan University



Research Article

Data-Driven Estimation of Rheological Behavior of Asphalt Mixture Using the K-Nearest Neighbors Algorithm

Mojtaba Khodadadi^a , Mohammad Esmaily^b, Hoseein Pahlevany Fard^c,
 Mohammad Hoseein Shirmohammadi^d, Alireza Rashidinejad^e, Mahmoud
 Ghelichkhan^f, Abolfazl Kheirkhah^g

^aAssistant Professor, Faculty of Civil Engineering, K. N. Toosi University of Technology, Tehran, I. R. Iran.

^{b,c,d,e,f,g}Qom Municipality, Qom, I. R. Iran.

PAPER INFO

Paper history:

Received: 2025-11-16

Revised: 2025-12-02

Accepted: 2025-12-04

Keywords:

Machine learning,

Rheological
 characterization,

Data-driven approach,

KNN algorithm,

Asphalt mixture.

ABSTRACT

Dynamic modulus (E^*) and phase angle (ϕ) are key parameters for describing the viscoelastic performance of asphalt mixtures. However, their experimental evaluation involves lengthy testing and costly laboratory procedures. To overcome these limitations, several predictive models have been introduced, among which the Witczak and Hirsch models are the most recognized. In recent years, machine learning (ML) techniques have gained attention in engineering applications due to their strong capabilities in data analysis, optimization, and prediction. This study introduces the K-Nearest Neighbors (KNN) algorithm as an ML-based method to estimate the viscoelastic behavior of asphalt mixtures. The model was trained and validated using an extensive dataset comprising bitumen characteristics, volumetric parameters, and measured values of dynamic modulus and phase angle at various temperatures and loading frequencies, totaling over 5500 data points. The results demonstrate that the proposed ML model provides high prediction accuracy and represents a promising alternative for estimating the viscoelastic properties of asphalt mixtures.

* Corresponding author.

E-mail address: m.khodadadi@kntu.ac.ir

How to cite this article: Khodadadi, M., Esmaily, M., Pahlevany Fard, H., Shirmohammadi, M. H., Rashidinejad, A., Ghelichkhan, M. and Kheirkhah, A. (2025). Data-driven estimation of rheological behavior of asphalt mixture using the k-nearest neighbors algorithm. J. Transport. Infrastruct. Eng., 11(3), 117-136. <https://doi.org/10.22075/JTIE.2025.39733.1746>

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)



مقاله پژوهشی

پیش‌بینی داده‌محور خواص رئولوژیک مخلوط‌های آسفالتی با استفاده از الگوریتم KNN

مجتبی خدادادی^{۱*}، محمد اسماعیلی^۲، حسین پهلوانی‌فرد^۳، محمدحسین شیرمحمدی^۴، علیرضا رشیدی‌نژاد^۵، محمود قلیچ‌خان^۶، ابوالفضل خیرخواه^۷

^۱ استادیار گروه راه و ترابری، دانشکده مهندسی عمران، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران.
^{۲،۳،۴،۵،۶،۷} شهرداری کلان‌شهر قم، قم، ایران.

چکیده

مدول دینامیکی (E^*) و زاویه فاز (ϕ) از پارامترهای کلیدی در توصیف رفتار ویسکوالاستیک مخلوط‌های آسفالتی هستند. با این حال، ارزیابی آزمایشگاهی آن‌ها مستلزم انجام آزمایش‌های طولانی و پرهزینه است. به منظور غلبه بر این محدودیت‌ها، مدل‌های پیش‌بینی تجربی متعددی ارائه شده‌اند که در میان آن‌ها مدل‌های Hirsch و Witczak بیشترین پذیرش را دارند. در سال‌های اخیر، روش‌های یادگیری ماشین (ML) به دلیل توانایی بالای خود در تحلیل داده‌ها، بهینه‌سازی و پیش‌بینی، توجه زیادی در حوزه‌های مختلف مهندسی جلب کرده‌اند. در این پژوهش، الگوریتم KNN به عنوان روشی مبتنی بر یادگیری ماشین برای پیش‌بینی رفتار ویسکوالاستیک مخلوط‌های آسفالتی معرفی شده است. مدل توسعه‌یافته با استفاده از یک پایگاه داده مناسب شامل ویژگی‌های قیر، پارامترهای حجمی و مقادیر مدول دینامیکی و زاویه فاز در دماها و فرکانس‌های مختلف، با بیش از ۵۵۰۰ نقطه داده آموزش و اعتبارسنجی شده است. نتایج نشان داد که مدل پیشنهادی KNN قادر است E^* و ϕ را با دقت زیاد پیش‌بینی کند، به‌طوری که برای مجموعه آزمون مقادیر $R^2=0.86$ ، $Se/Sy=0.33$ برای مدول دینامیکی و مقادیر $R^2=0.78$ ، $Se/Sy=0.45$ برای زاویه فاز به‌دست آمد که بیانگر توان مناسب مدل در بازتولید رفتار رئولوژیک مخلوط‌های آسفالتی در دامنه‌های مختلف دما و فرکانس است.

اطلاعات مقاله

دریافت مقاله: ۱۴۰۴/۰۸/۲۵

بازنگری مقاله: ۱۴۰۴/۰۹/۱۱

پذیرش مقاله: ۱۴۰۴/۰۹/۱۳

واژه‌های کلیدی:

یادگیری ماشین،
خواص رئولوژیک،
رویکرد داده‌محور،
الگوریتم KNN،
مخلوط آسفالتی.

* پست الکترونیک نویسنده مسئول: m.khodadadi@kntu.ac.ir

استناد به این مقاله: خدادادی، م.، اسماعیلی، م.، پهلوانی‌فرد، ح.، شیرمحمدی، م.، ح.، رشیدی‌نژاد، ع.، قلیچ‌خان، م. و خیرخواه، ا. ۱۴۰۴. پیش‌بینی داده‌محور خواص رئولوژیک مخلوط‌های آسفالتی با استفاده از الگوریتم KNN. مهندسی زیرساخت‌های حمل و نقل، ۱۱(۳)، ۱۱۷-۱۳۶.
<https://doi.org/10.22075/JTIE.2025.39733.1746>

۱. مقدمه

مدول دینامیکی E^* و زاویه فاز ϕ ، مهم‌ترین خواص ویسکوالاستیک مخلوط‌های آسفالتی هستند که رفتار رئولوژیک آن‌ها را توصیف می‌کنند. مدول دینامیکی نشان‌دهنده‌ی سختی یک مخلوط آسفالتی است، که بر اساس دمای آزمایش و فرکانس بارگذاری تعریف می‌شود. این ویژگی، به‌عنوان اصلی‌ترین خاصیت مصالح در روش‌های نوین طراحی روسازی بر پایه اصول مکانیکی شناخته می‌شود؛ چرا که تعیین‌کننده نحوه‌ی توزیع تنش‌ها و کرنش‌ها در درون ساختار روسازی است و می‌توان آن را با پدیده‌هایی چون شیاردگی^۱ و ترک‌خوردگی ناشی از خستگی در لایه‌های قیری مرتبط دانست. علاوه بر این، زاویه فاز مشخص‌کننده‌ی تأخیر فازی میان تنش اعمال‌شده و پاسخ کرنش در مواد ویسکوالاستیک است و به‌عنوان شاخصی از مؤلفه‌های ویسکوز (گرانروی) و الاستیک در شرایط خاص آزمایش عمل می‌کند. به‌طور معمول، این خواص دینامیکی ویسکوالاستیک در آزمایشگاه و با استفاده از تجهیزات پیشرفته و توسط تکنسین‌های ماهر اندازه‌گیری می‌شوند. در مواردی که چنین تجهیزاتی در دسترس نباشد، مدول دینامیکی را می‌توان با استفاده از مدل‌های پیش‌بینی، بر پایه ویژگی‌های حجمی مخلوط، دانه‌بندی و خواص قیر، برآورد کرد.

مدل‌های پیش‌بینی متعددی برای تخمین E^* به‌عنوان جایگزینی برای آزمایش‌های آزمایشگاهی توسعه یافته‌اند که پرکاربردترین آن‌ها، مدل‌های پیش‌بینی Witczak (اندری و همکاران، ۱۹۹۹، بری و ویتزک، ۲۰۰۶)، Hirsch (کریستنسن و همکاران، ۲۰۰۳، ژانگ و شن، ۲۰۱۷) و Al-Khateeb (الخطیب و همکاران، ۲۰۰۶) هستند. مطالعات مختلفی توسط پژوهشگران در ارزیابی این مدل‌های پیش‌بینی انجام شده و در مجموع،

نتایج نشان می‌دهند که این مدل‌ها قادر به ارائه برآوردهای قابل‌اطمینان از مدول دینامیکی هستند. با این حال، توجه کمتری به پیش‌بینی زاویه فاز معطوف شده و تنها تعداد محدودی از تحقیقات، مدل‌های پیش‌بینی برای این ویژگی ویسکوالاستیک در مصالح آسفالتی ارائه کرده‌اند (پیکادو سانتوس و همکاران، ۲۰۰۳؛ بیلگیری و همکاران، ۲۰۱۰؛ نایک و بیلگیری، ۲۰۱۴؛ یانگ و یو، ۲۰۱۵؛ ونوداران و بیلگیری، ۲۰۱۵). در سال‌های اخیر، الگوریتم‌های یادگیری ماشین (ML) به‌عنوان ابزارهای مکمل در مدل‌سازی مصالح مهندسی مورد توجه قرار گرفته‌اند. با این حال، در این پژوهش، تمرکز صرفاً بر مرور جامع کاربرد این روش‌ها در حوزه قیر و آسفالت است و از پرداخت تفصیلی به مزایای عمومی آن‌ها پرهیز شده است (خوستوسیلوا و همکاران، ۲۰۲۱). در حوزه بتن و مصالح ساختمانی، چابن و همکاران (۲۰۲۰) مروری تطبیقی بر عملکرد الگوریتم‌های مختلف ML ارائه داده‌اند و نشان داده‌اند که دقت پیش‌بینی این مدل‌ها وابستگی شدیدی به نوع داده و ساختار روابط ورودی-خروجی دارد. خامبرا و شوکلا (۲۰۲۱) نیز کارایی یادگیری ماشین در پیش‌بینی خواص بتن حاوی خاکستر بادی را تأیید نموده‌اند. الگوریتم‌ها و فنون متنوعی از یادگیری ماشین در مسائل مهندسی به‌کار گرفته شده‌اند؛ از جمله، گرادیان بوستینگ^۲، نزدیک‌ترین همسایگان^۳ و رگرسیون بردار پشتیبان^۴ که برای پیش‌بینی مقاومت فشاری بتن مورد استفاده قرار گرفته‌اند (بسکوپلنی و همکاران، ۲۰۲۲). الگوریتم‌های یادگیری ماشین از جمله شبکه‌های عصبی، گرادیان نزولی تصادفی، ماشین بردار پشتیبان (SVM) و درخت تصمیم برای پیش‌بینی عمق یخبندان در روسازی‌های آسفالتی در کشور کره جنوبی به کار گرفته شدند (چوی و همکاران، ۲۰۲۲). باجیک و

2- Gradient Boosting

3- k-Nearest Neighbors

4- Support Vector Regression

1- Rutting

همکاران (۲۰۲۱) و چوی و همکاران (۲۰۲۲) از الگوریتم‌های هوشمند جهت پایش ناهمواری و پیش‌بینی شرایط روسازی بهره گرفتند. همچنین، سلیمانوف و همکاران (۲۰۲۳)، هوانگ و همکاران (۲۰۲۳) و قربانی و همکاران (۲۰۲۳) نشان دادند که این الگوریتم‌ها توان بالایی در برآورد مدول الاستیسیته و رطوبت لایه‌های پایه دارند.

در حوزه مخلوط‌های آسفالتی، مطالعات متعددی مستقیماً بر پیش‌بینی خواص رئولوژیک، طراحی آزمایش‌های مارشال و عملکرد مکانیکی تمرکز کرده‌اند. آتاکان و ییلدیز (۲۰۲۳) و تانگا و همکاران (۲۰۲۴) از الگوریتم‌های ML برای برآورد پارامترهای مارشال استفاده نمودند و اوپادیا و همکاران (۲۰۲۴) عملکرد مدل‌های مختلف یادگیری ماشین در پیش‌بینی مقاومت مخلوط‌های آسفالتی اصلاح‌شده را ارزیابی کردند. النبی و همکاران (۲۰۲۴) نیز به مدل‌سازی شاخص IRI به کمک الگوریتم‌های ML پرداختند. گول و همکاران (۲۰۲۲) مدل‌های پیش‌بینی مقاومت و روانی مارشال روسازی‌های آسفالتی را با استفاده از سه تکنیک مختلف محاسبات نرم توسعه دادند. اوانواکوا و همکاران (۲۰۲۲) عملکرد پنج مدل مختلف یادگیری ماشین را با مدل پیش‌بینی تجربی توسعه‌یافته توسط باری و ویتزک (۲۰۰۶) مقایسه کردند و نتیجه گرفتند که مدل‌های پیشنهادی یادگیری ماشین دقت پیش‌بینی بالاتری نسبت به مدل رگرسیون Witczak دارند. در زمینه پیش‌بینی مدول دینامیکی، مارتینز و آنجلون (۲۰۱۰) و لیوا ویلاکورتاس و وارگاس نوردیک (۲۰۱۹) از شبکه‌های عصبی مصنوعی (ANN) استفاده نمودند و بهبود دقت نسبت به مدل‌های تجربی را گزارش کردند. بهنود و محمدی گل‌افشانی (۲۰۲۱) نیز با به‌کارگیری بهینه‌سازی زیست‌جغرافیایی، دقت

پیش‌بینی E^* را ارتقا دادند. همچنین، اُسچ-کاستلبلانکو و همکاران (۲۰۲۳) سه تکنیک مختلف یادگیری ماشین را برای پیش‌بینی خواص رئولوژیک قیرهای اصلاح‌شده با واکس به‌کار گرفتند. رحمان و همکاران (۲۰۲۱) از رگرسیون بردار پشتیبان و روش‌های مجموعه‌ای مبتنی بر درخت تصمیم برای پیش‌بینی شاخص‌های مرتبط با عملکرد مخلوط آسفالتی استفاده کردند. در این بین، الگوریتم‌های یادگیری ماشین نظارت‌شده شامل: شبکه عصبی مصنوعی، سیستم استنتاج فازی عصبی تطبیقی (ANFIS) و برنامه‌نویسی چندبیار^۱ برای پیش‌بینی مقاومت و روانی مارشال مخلوط‌های آسفالتی به‌کار گرفته شدند (گال و همکاران، ۲۰۲۲). بوتلا و همکاران (۲۰۲۲) سه تکنیک مختلف یادگیری ماشین را برای برآورد میزان فعالیت قیر در روسازی‌های آسفالتی بازیافتی به‌کار بردند. مجیدی فرد و همکاران (۲۰۱۹) از روش‌های نوآورانه یادگیری ماشین شامل برنامه‌نویسی بیان ژن^۲ و شبکه عصبی مصنوعی ترکیبی برای پیش‌بینی انرژی شکست نمونه‌های مخلوط آسفالتی استفاده کردند. مطالعاتی نظیر اُسچ کاستلبلانکو و همکاران (۲۰۲۳)، رحمان و همکاران (۲۰۲۱)، بوتلا و همکاران (۲۰۲۲) و مجیدی فرد و همکاران (۲۰۱۹) نشان می‌دهد که کاربرد فراگیر ML در حوزه قیر و آسفالت عمده‌تاً ناظر بر پیش‌بینی تک‌پارامتری است. در خصوص پیش‌بینی همزمان مدول دینامیکی و زاویه فاز، که چالش اصلی این پژوهش محسوب می‌شود، تنها مطالعات محدودی گزارش شده است. شو و هوانگ (۲۰۰۹) از یک روش تفاضلی برای پیش‌بینی مدول دینامیکی و زاویه فاز مخلوط‌های آسفالتی استفاده کردند. در فرآیند پیش‌بینی، ویژگی‌های اصلی مخلوط‌های آسفالتی شامل

1- Multi Expression Programming
2- Gene Expression Programming

اثر ویسکوالاستیک، دانه‌بندی مصالح سنگی و میزان تخلخل هوا در نظر گرفته شدند. روندینلا و همکاران (۲۰۲۳) رویکرد نوآورانه‌ای مبتنی بر یادگیری ماشین برای پیش‌بینی همزمان مدول دینامیکی و زاویه فاز ارائه دادند که بر اساس داده‌های ۹ نوع مخلوط آسفالتی مختلف، با استفاده از آزمایش چهارنقطه‌ای خمشی انجام شد. این پژوهشگران نتیجه گرفتند که روش‌های مبتنی بر یادگیری ماشین عملکرد به مراتب بهتری نسبت به معادلات تجربی به کار رفته دارند. علاوه بر این، روندینلا و همکاران (۲۰۲۴) از روش شبکه عصبی مصنوعی برای پیش‌بینی همزمان مدول دینامیکی و زاویه فاز مخلوط‌های آسفالتی استفاده کردند. در این مطالعه، معادله تجربی Witczak 1-37A که مدل رگرسیونی شناخته شده‌ای است، به عنوان مرجع برای مقایسه عملکرد مدل مورد استفاده قرار گرفت. نتایج این مطالعه نشان داد که پیش‌بینی‌های یادگیری ماشین، دقت چشم‌گیری نشان دادند و در تمامی معیارهای ارزیابی، عملکرد بهتری نسبت به مدل‌های مبتنی بر رگرسیون داشتند. در هر دو پارامتر مدول دینامیکی و زاویه فاز، ضریب همبستگی پیرسون (r) و ضریب تعیین (R^2) بیشتر از ۰/۹۸ به دست آمده بود که این موضوع بیانگر توانمندی و قابلیت اطمینان بالای این ابزار پیش‌بینی است. کمبود مطالعات داده‌محور که بتوانند این دو پارامتر اساسی را به صورت همزمان، در دامنه وسیع دما و فرکانس و با ورودی‌های مهندسی متداول پیش‌بینی کنند، همچنان به چشم می‌خورد و انگیزه اصلی پژوهش حاضر را شکل می‌دهد.

در پژوهش حاضر، الگوریتم نزدیک‌ترین همسایگان (KNN) به عنوان روشی در یادگیری ماشین برای پیش‌بینی همزمان مدول دینامیکی E^* و زاویه فاز ϕ مخلوط‌های آسفالتی پیشنهاد شده است. در مقایسه با

دیگر کاربردهای یادگیری ماشین، مانند شبکه‌های عصبی مصنوعی (ANN)، ماشین‌های بردار پشتیبان (SVM) و درختان رگرسیون با گرادینت تقویتی (GBRT)، الگوریتم KNN در این پژوهش به دلیل ساختار غیرپارامتری ساده، توانایی مناسب در مدل‌سازی روابط غیرخطی، پایداری عملکرد در مجموعه داده‌های با حجم متوسط و کاربرد موفق پیشین در پیش‌بینی خواص مصالح مهندسی انتخاب گردید. بنابراین، هدف مطالعه حاضر، توسعه یک مدل برای پیش‌بینی ویژگی‌های اصلی رئولوژیک مخلوط‌های آسفالتی است، با این نوآوری که با یکپارچه‌سازی یک پایگاه داده شامل: گستره مناسبی از ویژگی‌های قیر، خصوصیات حجمی و شرایط آزمایش، انجام می‌شود. قابل ذکر است که روش KNN قبلاً در زمینه‌های متعددی از جمله شناسایی ترک‌های روسازی و مدل‌سازی داده‌های آزمایش مقاومت مارشال (آکسوی و همکاران، ۲۰۱۲) به کار گرفته شده است. امروزه، الگوریتم KNN به دلیل سادگی توسعه، سهولت درک شهودی، دقت کافی در پیش‌بینی و قابلیت پیاده‌سازی در نرم‌افزارهای صفحه‌گسترده^۱ مورد توجه قرار گرفته است (الدوساری و همکاران، ۲۰۱۹؛ احمد و همکاران، ۲۰۲۳). در مطالعه حاضر نیز سعی می‌شود تا به کمک الگوریتم KNN برای پیش‌بینی خواص ویسکوالاستیک مخلوط‌های آسفالتی، نتایج به دست آمده ارزیابی شده و عملکرد پیش‌بینی آن مورد بحث قرار گیرد. لازم به ذکر است که مدل پیشنهادی جایگزین کامل آزمون‌های آزمایشگاهی نبوده، بلکه با اتکا به تعدادی آزمون‌های پایه در دسترس‌تر، نیاز به انجام آزمون‌های پرهزینه و زمان‌بر دینامیکی در دامنه گسترده دما و فرکانس را

به طور قابل توجهی کاهش می دهد.

۲. الگوریتم KNN

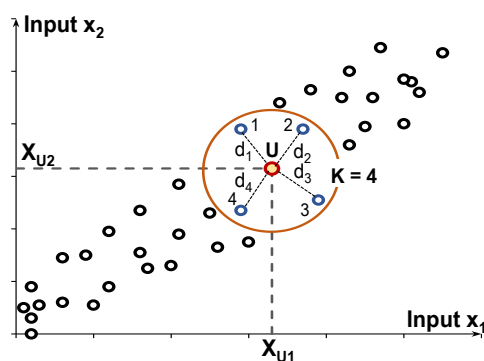
الگوریتم KNN، یک روش یادگیری ماشین تحت نظارت و غیرپارامتری است که برای مسائل رگرسیون و طبقه بندی با استفاده از «شباهت ویژگی» برای پیش بینی مقدار هر نقطه داده جدید استفاده می شود. این الگوریتم یادگیری، دارای هزینه محاسباتی کم و پیاده سازی بسیار ساده می باشد. الگوریتم KNN، بر اساس این فرض شهودی است که اشیاء نزدیک به هم به طور بالقوه مشابه یکدیگر هستند. روش KNN این مزیت را دارد که کاملاً غیرپارامتری است؛ به این معنی که هیچ فرضی در مورد توزیع آماری داده ها انجام نمی شود (ویس و کولیکووسکی، ۱۹۹۱). همچنین، روش KNN برای مدیریت ناپیوستگی ها در بررسی روابط بین متغیرها مناسب است (دار و استاین، ۱۹۹۷). یکی از مزایای مهم این روش، عدم وابستگی به نوع توزیع داده ها است، زیرا از نواحی محلی (همسایگی های نزدیک) برای تحلیل استفاده می کند. همچنین، پیاده سازی روش KNN از نظر محاسباتی کم هزینه است، چرا که تنها نیاز به ساخت یک پایگاه داده دارد. با این حال، یکی از معایب روش KNN زمان بر بودن آن در هنگام پیش بینی است، چرا که برای هر بار پیش بینی، فاصله شباهت باید بین نمونه جدید و تمام نمونه های موجود در پایگاه داده محاسبه شود. همچنین، از آنجا که برای استفاده از این

روش باید کل پایگاه داده در دسترس باشد، روش چندان فشرده یا قابل حملی نیست و ادغام آن در برنامه های بزرگ تر ممکن است دشوار باشد. الگوریتم KNN به یک مجموعه داده آموزشی نیاز دارد که شامل تعدادی نمونه با ورودی ها و خروجی های مشخص باشد. این نمونه های آموزشی به صورت بردارهایی در یک فضای ویژگی چندبعدی تعریف می شوند که دارای n متغیر ورودی (پیش بینی کننده) و m خروجی هستند. به طور کلی، هر نمونه به صورت زیر نمایش داده می شود:

$$x_i = \{x_{i1}, x_{i2}, \dots, x_{in}, y_{i1}, y_{i2}, y_{im}\} \quad (1)$$

که در آن، x_{ij} ها ورودی ها و y_{ij} ها خروجی های مربوط به آن نمونه هستند.

شکل ۱، مثالی را نشان می دهد که تنها دو متغیر ورودی به نام های x_1 و x_2 دارد. هر نقطه داده دارای یک خروجی y است. به عنوان مثال، نقطه ۱ دارای ورودی های x_{11} و x_{21} و خروجی y_1 است. نقطه ۲ نیز دارای ورودی های x_{12} و x_{22} و خروجی y_2 می باشد. برای یک نقطه جدید با نام U که دارای ورودی های x_{U1} و x_{U2} است، در صورتی که $K=4$ باشد، مقدار خروجی تخمینی y_U به صورت میانگین وزنی از فاصله ها d_i و خروجی های y_i چهار نزدیک ترین همسایه محاسبه می شود.



شکل ۱. مثالی از الگوریتم KNN

می‌شود؛ به این صورت که الگوریتم با مقادیر مختلفی از K اجرا شده و مقدار خطا برای هر حالت محاسبه می‌گردد، سپس مقداری از K که کمترین خطا را ایجاد کند به عنوان مقدار بهینه انتخاب می‌شود.

۳. مواد و روش‌ها

۱-۳. مجموعه داده‌ها

به‌منظور پشتیبانی از توسعه الگوریتم یادگیری ماشین (ML)، یک پایگاه داده منحصربه‌فرد با تجمیع اطلاعات از چهار مجموعه داده مختلف گزارش‌شده در ادبیات فنی موضوع تهیه شد. مجموعه داده اول، که با عنوان «مجموعه داده آریزونا» شناخته می‌شود، شامل ۲۱۸۳ نقطه داده بوده و در دانشگاه ایالتی آریزونا توسط Witczak و همکارانش در چارچوب بخش C از پروژه NCHRP 9-19 توسعه یافته است (ویتزک، ۲۰۰۵). این مجموعه شامل داده‌هایی از ۵۷ نوع مخلوط آسفالتی با قیرهای معمولی، اصلاح‌شده با پلیمر و قیر-لاستیکی است. مجموعه داده دوم، موسوم به «مجموعه داده مریلند»، شامل ۲۱۳۲ نقطه داده است که در دانشگاه مریلند توسعه یافته و برای واسنجی مدل پیش‌بینی مدول دینامیکی مخلوط‌های آسفالتی مورد استفاده قرار گرفته است (ویتزک و فونسکا، ۱۹۹۶). این مجموعه شامل نتایج مربوط به ۴۱ نوع مخلوط آسفالتی متراکم ساخته‌شده با قیرهای معمولی است. مجموعه داده سوم، با عنوان «مجموعه داده ویرجینیا»، دارای ۹۸۵ نقطه داده است که توسط فلینش و همکاران (۲۰۰۷) در مؤسسه حمل‌ونقل دانشگاه ویرجینیا تهیه شده است. این مجموعه شامل اطلاعات ۳۱ نوع مخلوط آسفالتی ساخته‌شده با قیر معمولی است. در نهایت، مجموعه داده چهارم، موسوم به «مجموعه داده فلوریدا»، شامل ۲۶۵ نقطه داده بوده و توسط بیرگیسون و همکاران (۲۰۰۴) در دانشگاه فلوریدا و در چارچوب پروژه‌ای برای توسعه قابلیت‌های اندازه‌گیری مدول دینامیکی برای ایالت فلوریدا جمع‌آوری شده است. این پایگاه

مقدار K یک پارامتر تعریف‌شده توسط کاربر است که تعداد نزدیک‌ترین همسایگان را برای انجام پیش‌بینی تعیین می‌کند. این مقدار بر اساس فاصله بین نقطه داده جدید و همسایگان آن انتخاب می‌شود و نقش مهمی در دقت پیش‌بینی ایفا می‌کند. یکی از معیارهای فاصله متداول که برای d_i استفاده می‌شود، فاصله اقلیدسی^۱ است. این فاصله به‌صورت طول یک پاره‌خط بین دو نقطه تعریف می‌شود و می‌توان آن را با استفاده از مختصات دکارتی نقاط به صورت زیر محاسبه کرد:

$$d_i = \sqrt{(x_{U1} - x_{i1})^2 + (x_{U2} - x_{i2})^2} \quad (2)$$

در این رابطه، x_{ij} مختصات ورودی نقطه همسایه و x_{Uj} مختصات ورودی نقطه جدید U هستند. در اینجا، i برابر ۱، ۲، ۳ و ۴ می‌باشد.

این الگوریتم از میانگین وزنی K همسایه‌ی نزدیک استفاده می‌کند، به‌گونه‌ای که وزن هر همسایه معکوس فاصله‌ی آن از نقطه جدید است. بنابراین، همسایگان نزدیک‌تر، نسبت به همسایگان دورتر، تأثیر بیشتری در میانگین دارند. برای مثال، وقتی $K=4$ باشد، مقدار خروجی تخمینی به‌صورت زیر محاسبه می‌شود:

$$y_U = \frac{\sum_{i=1}^4 y_i \times \frac{1}{d_i}}{\sum_{i=1}^4 \frac{1}{d_i}} \quad (3)$$

در این رابطه، d_i فاصله بین نقطه جدید با چهار همسایه نزدیک بوده و y_i خروجی آن‌ها است.

عملکرد این الگوریتم به‌شدت به مقدار پارامتر K وابسته است. مقدار بهینه برای K معمولاً به‌صورت تجربی و با استفاده از روش اعتبارسنجی متقابل^۲ تعیین

1- Euclidean distance
2- Cross-validation

داده شامل اطلاعات مربوط به ۲۷ نوع مخلوط آسفالتی با دانه بندی ها و انواع سنگ دانه های مختلف است.

اگرچه این حجم داده در مقایسه با بسیاری از مطالعات پیشین در حوزه مدل سازی خواص رئولوژیک قابل توجه محسوب می شود، بدیهی است که نمی تواند نماینده تمامی تنوع های بالقوه در مواد تشکیل دهنده آسفالت، دانه بندی ها و شرایط بهره برداری در مقیاس واقعی باشد. از این رو، دامنه ای اعتبار مدل ارائه شده در این پژوهش محدود به بازه ای داده های آزمایشگاهی در دسترس بوده و نتایج آن به عنوان یک ابزار پیش بینی داده محور برای شرایط مشابه، نه به عنوان یک مدل کاملاً عمومی برای کلیه سناریوهای ممکن در مهندسی روسازی، توصیه می شود.

اطلاعات موجود در چهار منبع مذکور گردآوری شده و ویژگی های حجمی اصلی مخلوط ها، خواص قیر و شرایط آزمایش برای تحلیل های این پژوهش در نظر گرفته شده اند. با توجه به اینکه داده های هر چهار منبع،

مطابق استاندارد آزمایشگاهی معتبر به دست آمده اند، نیازی به همگن سازی مجموعه داده ها وجود نداشت. در جدول ۱، نتایج E^* (برحسب MPa) و ϕ چهار منبع آریزونا، مریلند، ویرجینیا و فلوریدا ارائه شده است. این آمار شامل درصد حجمی قیر مؤثر (V_b)، درصد حجمی حفرات هوا (V_a)، فضای خالی بین سنگدانه ها (VMA) برحسب درصد، فضای خالی پر شده با قیر (VFA) برحسب درصد، پارامترهای ویسکوزیته قیر (A و VTS)، دمای آزمایش (T) برحسب درجه سلسیوس و فرکانس بارگذاری (f) برحسب هرتز می باشد. انحراف معیار زیاد در برخی از نتایج مدول دینامیکی ناشی از ناهمگنی ذاتی داده ها در اثر تنوع ترکیبات مخلوط و شرایط گسترده دمایی و فرکانسی است. این پراکندگی به عنوان بازتاب شرایط واقعی عملکرد مواد بوده و با اعمال نرمال سازی داده ها و استفاده از الگوریتم KNN در فرآیند آموزش، اثر منفی آن بر پایداری مدل کنترل شده است.

جدول ۱. نتایج E^* و ϕ برای منابع مختلف بر حسب مشخصات مخلوط آسفالتی

نتایج آزمایش		شرایط آزمایش		مشخصات حجمی				خواص ویسکوزیته قیر		منبع و مقادیر	
φ	E*	F	T	VFA	VMA	V _a	V _b	VTs	A		
48.4	64463	25.0	54.4	79.0	40.7	18.5	22.2	-2.7384	11.3750	حداکثر	آریزونا
2.8	118	0.1	-10.0	54.6	15.5	3.8	10.2	-3.8220	8.3904	حداقل	
23.0	11035	7.1	21.8	66.5	21.0	7.4	14.2	-3.4702	10.4026	میانگین	
11.4	11256	8.8	23.0	6.8	5.9	2.9	3.3	0.3309	0.9060	انحراف معیار	
48.6	42816	25.0	54.4	85.5	22.9	11.3	16.3	-2.8989	11.1165	حداکثر	مربلند
4.0	86	0.1	-17.8	46.5	10.7	2.3	8.5	-3.7237	8.8325	حداقل	
24.6	8339	7.6	19.8	60.3	15.9	6.3	12.0	-3.4395	10.3345	میانگین	
10.3	8416	9.2	24.8	7.3	2.0	1.5	1.9	0.2109	0.5723	انحراف معیار	
38.1	40311	25.0	54.4	69.6	22.8	7.6	15.2	-3.6800	10.9800	حداکثر	ویرجینیا
0.7	125	0.1	-12.2	58.2	15.9	5.1	9.6	-3.6800	10.9800	حداقل	
19.8	9022	7.0	21.0	64.1	18.7	6.7	12.0	-3.6800	10.9800	میانگین	
10.6	9355	8.8	23.5	2.8	1.9	0.5	1.6	0.0000	0.0000	انحراف معیار	
48.9	14398	16.0	40.0	78.5	18.5	4.5	14.8	-3.2640	9.8466	حداکثر	فلوریدا
14.1	317	1.0	10.0	63.8	10.5	3.7	7.3	-3.2640	9.8466	حداقل	
29.7	3845	7.6	24.9	72.4	14.7	4.0	11.1	-3.2640	9.8466	میانگین	
7.1	2937	5.7	13.3	3.3	1.7	0.2	1.7	0.0000	0.0000	انحراف معیار	

۲-۳. ورودی‌های انتخاب‌شده مدل

انتخاب ورودی‌های مدل برای الگوریتم KNN عمدتاً به‌صورت شهودی و بر اساس مرور منابع موجود در ادبیات فنی انجام شده است. با توجه به بررسی کلی منابع، عوامل مهم و کلیدی در تعیین خواص ویسکوالاستیک مخلوط‌های آسفالتی شامل ویسکوزیته قیر، ویژگی‌های حجمی مخلوط و شرایط آزمایش (شامل دما و فرکانس) هستند. بر این اساس، با استفاده از پارامترهای A و VTS، مقدار لگاریتمی ویسکوزیته قیر (برحسب cp) در دمای آزمایش T به‌صورت زیر محاسبه شد:

$$\log(visc) = 10^{[A + VTS \times \log(T_R)]} \quad (۴)$$

$$T_R = \frac{9}{5} \times T + 491.67 \quad (۵)$$

که در آن T_R ، دما بر حسب درجه رانکین (R°) است.

برای انتخاب ورودی‌های مدل، از ضریب همبستگی پیرسون (r) به‌منظور بررسی اولیه قدرت روابط بین ورودی‌های موجود با مدول دینامیکی و زاویه فاز استفاده شده است. ضریب همبستگی پیرسون بین دو متغیر همواره مقداری بین +۱ و -۱ دارد. مقدار قدر مطلق نزدیک به یک نشان‌دهنده همبستگی کامل و مقدار نزدیک به صفر نشان‌دهنده عدم وجود همبستگی است. علامت مثبت نشان‌دهنده تناسب مستقیم و علامت منفی بیانگر تناسب معکوس بین دو متغیر است.

۳-۳. مجموعه داده‌های آموزش و ارزیابی

در هنگام استفاده از الگوریتم KNN، مجموعه داده باید به دو بخش تقسیم شود: ۱- مجموعه داده آموزشی، که الگوریتم بر اساس آن پیش‌بینی‌های خود را انجام می‌دهد، و ۲- مجموعه داده ارزیابی (اعتبارسنجی)، که برای سنجش عملکرد الگوریتم روی داده‌هایی که قبلاً

دیده نشده‌اند، مورد استفاده قرار می‌گیرد (ایماندوست و بلندرفتار، ۲۰۱۳). بر این اساس، داده‌ها به‌صورت تصادفی^۱ مرتب شده و به دو مجموعه تقسیم شدند. مجموعه نخست، شامل ۷۵ درصد از داده‌ها (شامل ۴۱۷۴ ورودی و خروجی)، به‌عنوان مجموعه آموزشی مورد استفاده قرار گرفت. مجموعه دوم، شامل ۲۵ درصد باقی‌مانده از داده‌ها (۱۳۹۱ ورودی و خروجی)، به‌عنوان مجموعه ارزیابی برای آزمون عملکرد مدل به کار گرفته شد. بنابراین، نتایج بر اساس یک تقسیم‌بندی منفرد از داده‌های موجود به‌دست آمده‌اند. الگوریتم در محیط نرم‌افزار Excel به‌گونه‌ای پیاده‌سازی شد که برای هر مجموعه ورودی مشخص، مقادیر پیش‌بینی‌شده مدول دینامیکی E^* و زاویه فاز ϕ به‌صورت خودکار محاسبه شوند.

همانطور که ذکر گردید، انتخاب مقدار K اهمیت زیادی دارد، زیرا تأثیر مستقیمی بر دقت پیش‌بینی مدل دارد. اگر مقدار K کوچک انتخاب شود، تغییرپذیری نتایج پیش‌بینی‌شده ممکن است زیاد باشد. در حالی که در صورت انتخاب مقادیر بزرگ‌تر، پیش‌بینی‌ها ممکن است بر اساس نقاطی انجام شوند که از نقطه مورد نظر فاصله نسبتاً زیادی دارند (سمان و همکاران، ۲۰۰۶). مقدار بهینه‌ی K معمولاً به‌صورت تجربی و از طریق فرآیند اعتبارسنجی متقابل با مقادیر مختلف K و تعیین حالتی که کمترین خطا را ایجاد می‌کند، تعیین می‌شود. در این پژوهش، برای ارزیابی نسبی عملکرد الگوریتم در مقادیر مختلف K، از معیار ریشه میانگین مربعات خطای نرمال‌شده (NRMSE) استفاده شد که رابطه‌ی آن در معادله (۶) ارائه شده است:

$$NRMSE = NRMSE_{|E^*|} + NRMSE_{\phi} \quad (۶)$$

$$NRMSE_{|E^*|} = \frac{\sqrt{\sum_1^N (|E^*|_i - |E^*|_{pred})^2}}{|E^*|_{mean}} \quad (۷)$$

$$\frac{Se}{Sy} = \frac{\sqrt{\frac{1}{N-1} \times \sum_{i=1}^N (y_i - P_i)^2}}{\sqrt{\frac{1}{N-1} \times \sum_{i=1}^N (y_i - y_m)^2}} \quad (9)$$

که در آن y_i مقدار آزمایشگاهی، P_i مقدار پیش‌بینی شده، N تعداد نمونه‌ها، y_m میانگین مقادیر اندازه‌گیری شده، Se خطای استاندارد پیش‌بینی و Sy انحراف معیار مقادیر آزمایشگاهی را نشان می‌دهد. مقدار کمتر Se/Sy بیانگر عملکرد بهتر مدل و دقت بیشتر پیش‌بینی است. لازم به ذکر است که داده‌های آزمایشگاهی مورد استفاده شامل عدم قطعیت ذاتی اندازه‌گیری مدول دینامیکی و زاویه فاز می‌باشند. با این حال، در این پژوهش، مقادیر ثبت شده به عنوان داده‌های مرجع به کار رفته و عدم قطعیت اندازه‌گیری‌ها به صورت صریح در چارچوب مدل‌سازی وارد نشده است. به منظور طبقه‌بندی و تفسیر کیفی عملکرد مدل‌ها، سینگ و همکاران (۲۰۱۱) یک معیار کیفی ارائه کرده‌اند که در جدول ۲ نشان داده شده است و به عنوان مبنایی برای سنجش کلی دقت و کارایی مدل‌ها استفاده خواهد گردید. محدوده متغیرهای ارائه شده در جدول ۲ بر اساس دامنه‌های متداول گزارش شده در مطالعات پیشین پیش‌بینی مدول دینامیکی و زاویه فاز و مطابق با استانداردهای آزمایشگاهی رایج انتخاب شده‌اند (بری و ویتزک، ۲۰۰۶؛ مارتینز و آنجلون، ۲۰۱۰).

جدول ۲. معیارهای مربوط به پارامترهای آماری

مختلف برازش

Se/Sy	R ²	معیارها
≤ 0.35	≥ 0.90	Excellent
0.36 - 0.55	0.70 - 0.89	Good
0.56 - 0.75	0.40 - 0.69	Fair
0.76 - 0.89	0.20 - 0.39	Poor
≥ 0.90	≤ 0.19	Very Poor

$$NRMSE_{\phi} = \frac{\sqrt{\sum_{i=1}^N (\phi_i - \phi_{pred})^2}}{N \phi_{mean}} \quad (8)$$

که در آن: $NRMSE_{|E^*|}$ ریشه میانگین مربعات خطای نرمال شده برای مقادیر $|E^*|$ ، $NRMSE_{\phi}$ ریشه میانگین مربعات خطای نرمال شده برای مقادیر زاویه فاز ϕ ، N تعداد کل داده‌ها در پایگاه داده، $|E^*|_i$ و ϕ_i مقادیر اندازه‌گیری شده $|E^*|$ و ϕ برای داده‌ی i ام، $|E^*|_{pred}$ و ϕ_{pred} مقادیر پیش‌بینی شده $|E^*|$ و ϕ برای داده‌ی i ام، $|E^*|_{mean}$ و ϕ_{mean} میانگین مقادیر اندازه‌گیری شده $|E^*|$ و ϕ در کل مجموعه داده‌ها می‌باشند. همانطور که در معالات فوق نیز دیده می‌شود، نرمال‌سازی $NRMSE$ بر اساس میانگین مقادیر اندازه‌گیری شده صورت گرفته که موجب امکان‌پذیر نمودن مقایسه نسبی خطا در متغیرهایی با دامنه‌های عددی متفاوت خواهد شد. بر اساس نتایج سه پارامتر فوق، مقدار بهینه K مدل‌ها تعیین خواهد گردید.

۳-۴. ارزیابی نهایی مدل‌ها

برای ارزیابی عملکرد مدل‌های توسعه‌یافته، شش شاخص مختلف برازش^۱ انتخاب و به کار گرفته شدند. این شاخص‌ها عبارت‌اند از: میانگین قدرمطلق خطا (MAE)، میانگین درصدی قدرمطلق خطا (MAPE)، میانگین مربعات خطا (MSE)، ریشه میانگین مربعات خطا (RMSE)، ضریب تعیین (R^2) و نسبت بین خطای استاندارد مقادیر پیش‌بینی شده به مقادیر اندازه‌گیری شده (Se/Sy) که این مقدار، طبق رابطه زیر تعریف می‌شود:

1- Goodness-of-fit

۴. نتایج و بحث

۴-۱. انتخاب ورودی‌های مدل

برای انتخاب ورودی‌های مناسب مدل، مقادیر ضرایب همبستگی پیرسون مورد بررسی قرار گرفت. خروجی این مقادیر بین ورودی‌های موجود، مدول دینامیکی E^* و زاویه فاز ϕ در جدول ۳ ارائه شده‌اند. قابل ذکر است که تحلیل همبستگی پیرسون در این مطالعه صرفاً به‌عنوان ابزار غربالگری اولیه و برای بررسی وابستگی‌های خطی به‌کار رفته و مبنای انحصاری انتخاب متغیرها نبوده است. انتخاب نهایی ورودی‌ها بر اساس ملاحظات فیزیکی مهندسی روسازی انجام شده و وابستگی‌های غیرخطی پیچیده در مرحله آموزش توسط الگوریتم KNN بازنمایی شده‌اند.

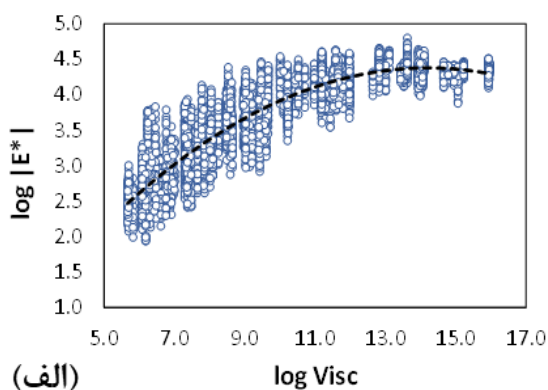
شکل‌های ۲-الف و ۲-ب، به‌ترتیب تغییرات مدول

دینامیکی E^* را در برابر مقادیر لگاریتمی ویسکوزیته قیر و دمای آزمایش نمایش می‌دهند. این دو پارامتر به‌عنوان مهم‌ترین متغیرها با بیشترین همبستگی بر اساس تحلیل ضریب همبستگی پیرسون شناسایی شده‌اند.

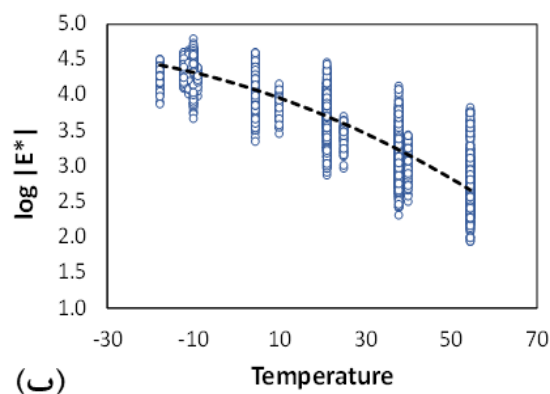
جدول ۳. ضرایب همبستگی پیرسون ویژگی‌های

مخلوط آسفالتی با E^* و ϕ

ویژگی مخلوط آسفالتی	ضریب پیرسون با E^*	ضریب پیرسون با ϕ
Log (Viscosity)	0.85	-0.74
V_b	-0.03	0.10
V_a	-0.05	0.07
VMA	-0.01	0.06
VFA	0.05	-0.01
T	-0.83	0.70
F	0.16	0.02



(الف)

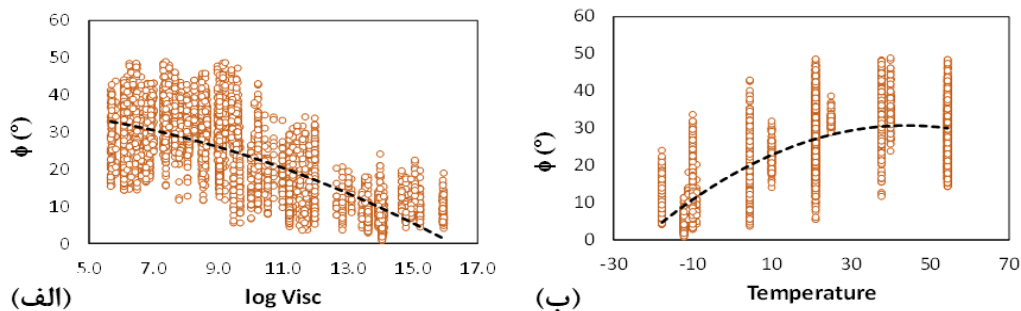


(ب)

شکل ۲. تغییرات زاویه فاز E^* نسبت به الف) ویسکوزیته قیر و ب) دمای آزمایش

به‌گونه‌ای که با افزایش دمای آزمون، مقدار E^* کاهش پیدا می‌کند. شکل‌های ۳-الف و ۳-ب نیز به‌ترتیب تغییرات زاویه فاز ϕ را در برابر مقادیر لگاریتمی ویسکوزیته قیر و دمای آزمایش نشان می‌دهند.

مطابق نتایج شکل ۲، بین ویسکوزیته قیر و مدول دینامیکی یک همبستگی مثبت قوی مشاهده می‌شود؛ به‌طوری که با افزایش ویسکوزیته قیر، مقدار E^* نیز افزایش می‌یابد. در مقابل، بین دمای آزمایش و مدول دینامیکی رابطه‌ای معکوس و قوی وجود دارد،



شکل ۳. تغییرات زاویه فاز ϕ نسبت به الف) ویسکوزیته قیر و ب) دمای آزمایش

بر اساس نتایج شکل ۳، بین دمای آزمون و زاویه فاز یک همبستگی مثبت قوی وجود دارد؛ بدین معنا که با افزایش دما، مقدار ϕ نیز افزایش یافته و رفتار قیر به سمت خاصیت ویسکوز (کاهش الاستیسیته) میل می‌کند. در مقابل، بین ویسکوزیته قیر و زاویه فاز یک همبستگی منفی مشاهده می‌شود. علاوه بر این، یک همبستگی متوسط بین درصد حجمی قیر V_b و درصد هوا V_a با زاویه فاز، و همچنین بین فرکانس آزمایش (f) و مدول دینامیکی E^* وجود دارد. بر اساس این مشاهدات، متغیرهای لگاریتمی ویسکوزیته قیر $(\log Viscosity)$ ، درصد حجمی قیر (V_b) ، درصد هوا (V_a) ، دمای آزمایش (T) و فرکانس آزمایش (f) به عنوان ورودی‌های مدل نخست الگوریتم KNN انتخاب شدند. این مدل با عنوان مدل کاهش یافته^۱ معرفی می‌گردد.

۴-۲. تعیین مقدار بهینه K مدل‌ها

همانطور که در بخش قبل ذکر گردید، بر اساس نتایج سه پارامتر حاصل از معادلات ۶ الی ۸، مقدار بهینه K برای هر کدام از مدل‌ها تعیین خواهد گردید. جدول ۴ تغییرات $NRMSE$ ، $NRMSE_{|E^*|}$ و $NRMSE_{\phi}$ را برای مقادیر مختلف K در مدل کاهش یافته را نشان می‌دهد. همچنین، جدول ۵، تغییرات متناظر برای مدل کامل را ارائه می‌کند.

جدول ۴. تغییرات $NRMSE$ مدل کاهش یافته

مقدار K	$NRMSE_{ E^* }$	$NRMSE_{\phi}$	$NRMSE$
3	0.392	0.223	0.615
5	0.372	0.217	0.589
6	0.378	0.217	0.595
7	0.385	0.218	0.603
9	0.382	0.222	0.604
11	0.382	0.223	0.605

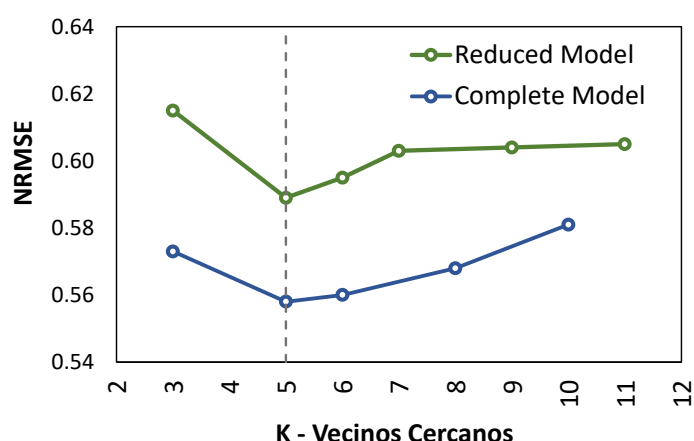
به طور خلاصه، نمودار تغییرات $NRMSE$ برای هر دو مدل یاد شده، در شکل ۴ نمایش داده شده است:

در علم مهندسی روسازی آسفالت، برای تعیین خواص ویسکوالاستیک، معمولاً از مدل‌های هیرچ (کریستنسن و همکاران، ۲۰۰۳؛ ژانگ و شن، ۲۰۱۷) و مدل الخاطب (الخطیب و همکاران، ۲۰۰۶) استفاده می‌شود که در آن‌ها پارامترهای VMA و VFA به عنوان ویژگی‌های کلیدی مؤثر بر رفتار ویسکوالاستیک مخلوط‌های آسفالتی در نظر گرفته می‌شوند. بر این اساس، یک مدل دوم از الگوریتم

۱- Reduced model

جدول ۵. تغییرات NRMSE برای مدل کامل

مقدار K	$NRMSE_{ E^* }$	$NRMSE_{\varphi}$	NRMSE
3	0.377	0.196	0.573
5	0.361	0.197	0.558
6	0.363	0.197	0.560
7	0.368	0.200	0.568
9	0.374	0.207	0.581
11	0.377	0.196	0.573



شکل ۴. NRMSE مدل کاهش یافته و مدل کامل

به‌عنوان مقدار بهینه انتخاب شده و بازه‌ی $K=5$ تا $K=6$ به‌عنوان محدوده‌ی عملی پایدار برای پارامتر K پیشنهاد می‌گردد که تعادل مناسبی بین بایاس و واریانس در فرآیند پیش‌بینی خواص رئولوژیک ایجاد می‌کند. بنابراین، مطابق نتایج این شکل، مقدار بهینه‌ی K برابر ۵ برای هر دو مدل انتخاب می‌گردد، زیرا کمترین مقدار NRMSE را ایجاد کرده است.

۳-۴. ارزیابی عملکرد مدل‌ها

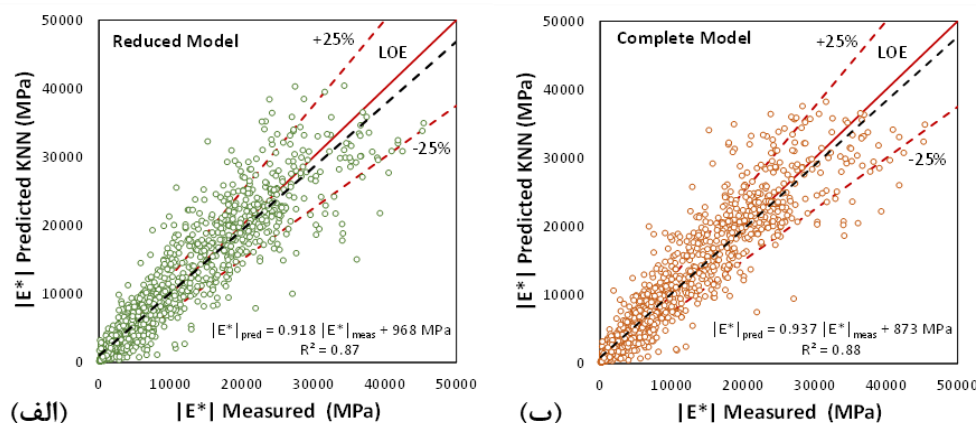
برای ارزیابی عملکرد مدل‌ها، مقادیر اندازه‌گیری شده و پیش‌بینی شده با یکدیگر مقایسه شدند تا میزان پراکندگی داده‌ها در امتداد خط برابری LOE^1 بررسی شود. شکل‌های ۵-الف و ۵-ب مقایسه‌ی مقادیر اندازه‌گیری شده و پیش‌بینی شده‌ی E^* را به ترتیب برای مدل کاهش یافته و مدل کامل نشان می‌دهد.

همانطور که در جداول و شکل فوق دیده می‌شود، به‌منظور تعیین مقدار بهینه پارامتر کلیدی K در الگوریتم KNN، تحلیل حساسیت عملکرد مدل نسبت به تغییرات این پارامتر انجام شد. بدین منظور، مقادیر NRMSE برای چندین مقدار مختلف K شامل $K=3,5,6,7,9,11$ محاسبه و نتایج در جدول ۴ و شکل ۴ ارائه گردیده است. نتایج نشان می‌دهد که حداقل مقدار NRMSE برای هر دو ساختار مدل (کاهش یافته و کامل) در $K=5$ به دست آمده و در ناحیه‌ی پیرامونی این مقدار، به‌ویژه در $K=6$ ، تغییرات خطا بسیار محدود بوده و عملکرد مدل پایدار باقی می‌ماند. افزایش بیشتر مقدار K به تدریج منجر به افزایش خطا می‌شود که می‌تواند ناشی از ورود همسایه‌های دورتر و کم‌ارتباط و در نتیجه افزایش بایاس مدل باشد. در حالی که مقادیر کمتر K منجر به نوسانات بیشتر پیش‌بینی و افزایش واریانس مدل می‌شوند. بر این اساس، مقدار $K=5$

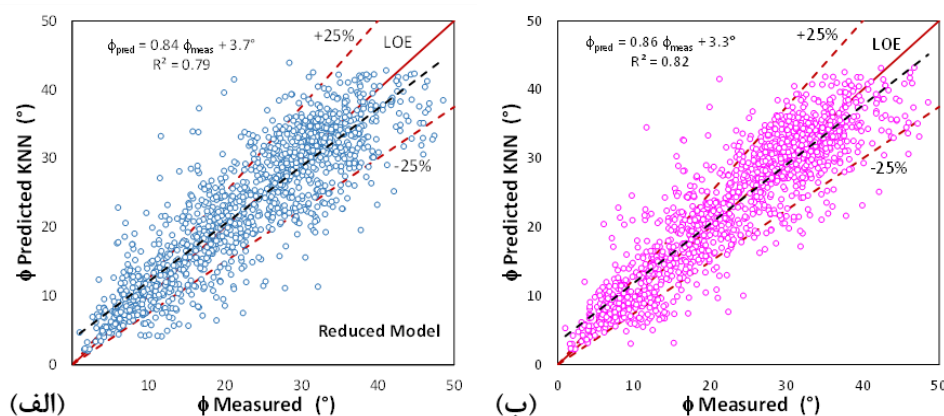
1- Line of Equality

در این مدل‌ها ارائه می‌کند.

همچنین، شکل ۶ مقایسه‌ی مشابهی را برای زاویه فاز ϕ



شکل ۵. مقادیر پیش‌بینی شده و اندازه‌گیری شده E^* مدل: (الف) کاهش یافته و (ب) کامل



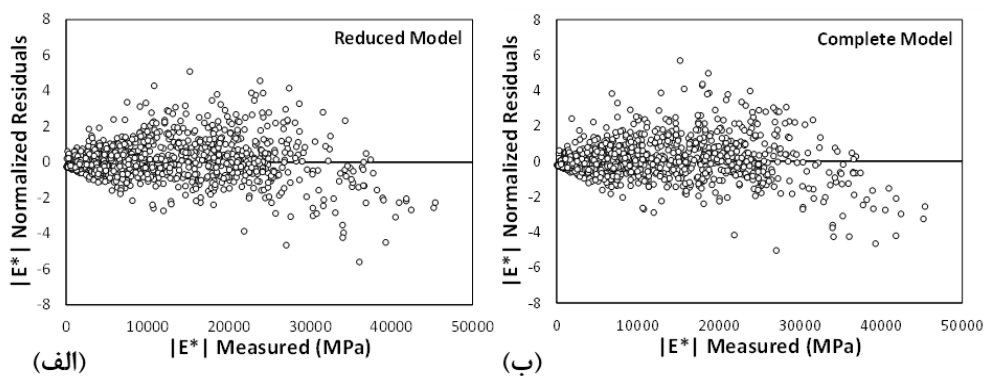
شکل ۶. مقادیر پیش‌بینی شده و اندازه‌گیری شده ϕ مدل: (الف) کاهش یافته و (ب) کامل

علاوه بر این، شکل ۷-الف و ۷-ب به ترتیب، باقیمانده‌های نرمال‌شده^۱ را برای مقادیر E^* در مدل کاهش‌یافته و مدل کامل نشان می‌دهند. همچنین، شکل ۸-الف و ۸-ب به ترتیب باقیمانده‌های نرمال‌شده‌ی مربوط به ϕ را برای مدل‌های کاهش‌یافته و مدل کامل ارائه می‌دهند. قابل ذکر است که داده‌های این پژوهش به‌صورت مجموعه‌ای از مقادیر مستقل آزمایشگاهی استفاده شده و تحلیل خودهمبستگی باقیمانده‌ها به‌طور صریح انجام نشده است. از این‌رو نتایج آماری بر مبنای فرض استقلال نسبی نمونه‌ها تفسیر می‌شوند.

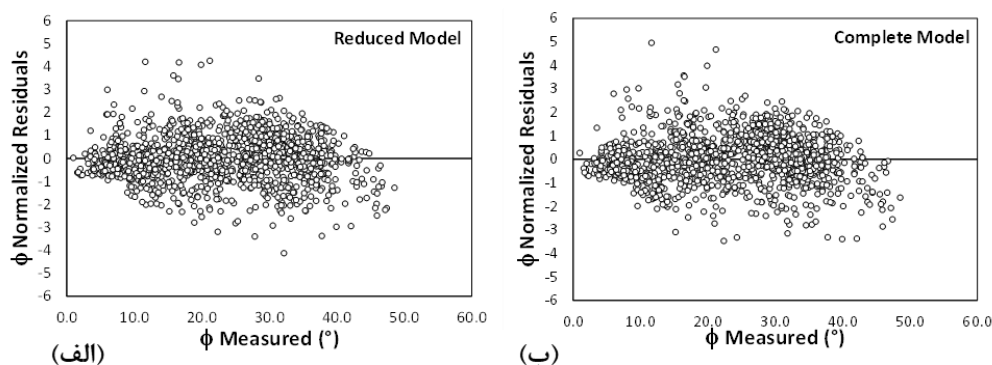
برای هر دو مدل، نقاط داده‌ی E^* در امتداد خط برابری LOE در یک ناحیه‌ی نسبتاً باریک قرار گرفته و در دو سوی این خط بدون وجود انحراف قابل‌توجهی (bias) توزیع شده‌اند. نقاط داده‌ی مربوط به زاویه فاز ϕ نیز در امتداد خط برابری قرار گرفته و به‌صورت فشرده در اطراف آن تجمع یافته‌اند.

در شکل‌های ۵ و ۶، دو خط اضافی نشان‌دهنده‌ی محدوده‌ی پراکندگی $\pm 25\%$ حول خط برابری LOE افزوده شده‌اند. طبق نتایج شکل ۵، برای هر دو مدل، بیش از ۸۶٪ از مقادیر پیش‌بینی‌شده‌ی E^* در این محدوده قرار دارند، در حالی‌که نتایج شکل ۶ نشان می‌دهند که برای مقادیر ϕ ، بیش از ۷۰٪ داده‌ها دارای انحرافی کمتر از $\pm 25\%$ می‌باشند.

1- Normalized Residuals



شکل ۷. مقادیر باقیمانده‌های نرمال شده E^* برای: الف) مدل کاهش یافته و ب) مدل کامل



شکل ۸. مقادیر باقیمانده‌های نرمال شده ϕ برای: الف) مدل کاهش یافته و ب) مدل کامل

در ادامه، رابطه‌ی بین مقادیر پیش‌بینی شده و اندازه‌گیری شده از طریق شیب و عرض از مبدأ خط رگرسیون مورد بررسی قرار می‌گیرد.

مطابق شکل ۵-الف، برای مدل کاهش یافته مقادیر E^* عرض از مبدأ کم (۹۶۸ مگاپاسکال) و شیبی نزدیک به یک (۰/۹۱۸) حاصل شده است. برای مقادیر زاویه فاز ϕ نیز در شکل ۶-الف عرض از مبدأ کم (۳/۷ درجه) و شیبی نزدیک به یک (۰/۸۴) کسب شده است. همچنین، برای مدل کامل، مقادیر E^* ، نتایج شکل ۵-ب عرض از مبدأ کم (۸۷۳ مگاپاسکال) و شیبی نزدیک به یک (۰/۹۳۷) را نشان می‌دهد. در نهایت، برای مقادیر ϕ ، شکل ۶-ب عرض از مبدأ کم (۳/۳ درجه) و شیب برابر با ۰/۸۶ در مدل کامل را نمایش می‌دهد. از این‌رو، هر دو مدل کاهش یافته و کامل، در مقایسه با نتایج مشابه گزارش شده در مطالعات پیشین (ترن و هال، ۲۰۰۵؛ ابو عبدو و همکاران، ۲۰۰۹)

مطابق نتایج شکل‌های ۷ و ۸، در هر چهار حالت، انحراف قابل توجهی (bias) در نتایج خواص ویسکوالاستیک مشاهده نمی‌گردد. این شواهد بیانگر این است که مقادیر پیش‌بینی شده، تطابق مناسبی با مقادیر اندازه‌گیری شده دارند. به منظور ارزیابی رسمی نرمال بودن توزیع باقیمانده‌های مدل‌ها، علاوه بر بررسی‌های گرافیکی ارائه شده در شکل‌های ۷ و ۸، آزمون آماری شاپیرو-ویلک نیز روی مجموعه خطاهای پیش‌بینی اجرا شد. نتایج این آزمون در سطح معناداری ۰/۰۵ نشان داد که فرض صفر مبنی بر نرمال بودن توزیع باقیمانده‌ها رد نمی‌شود. از این‌رو، می‌توان نتیجه گرفت که توزیع خطاها از الگوی تقریباً نرمال تبعیت کرده و استفاده از شاخص‌های آماری مبتنی بر مفروضه نرمال بودن شامل ضریب تعیین R^2 و معیار RMSE برای ارزیابی عملکرد مدل در این مطالعه معتبر و قابل اتکا است.

و مطابق دسته‌بندی جدول ۲ در بخش قبل، می‌توانند در سطح بسیار خوب^۱ ارزیابی شوند.

با وجود آنکه ویژگی‌های حجمی VMA و VFA در تحلیل پیرسون همبستگی ضعیفی با E^* و ϕ نشان دادند، افزودن این پارامترها به مدل کامل موجب بهبود نتایج پیش‌بینی شد؛ به‌طوری که شیب‌ها به مقدار ۱ نزدیک‌تر و مقادیر عرض از مبدأ کوچکتر شدند.

در انتها، جدول ۶ مقایسه‌ی عملکرد پیش‌بینی دو مدل KNN را بر اساس معیارهای خطا (MAE، MSE، MAPE)، ضریب تعیین R^2 ، نسبت خطای استاندارد مقادیر پیش‌بینی‌شده به انحراف معیار مقادیر اندازه‌گیری‌شده (Se/Sy) و ارزیابی کلی عملکرد طبق معیارهای معرفی‌شده جدول ۴ نشان می‌دهد. از آنجا که مدل KNN یک روش ناپارامتریک و فاقد تابع درست‌نمایی و تعداد پارامتر مشخص است، به‌کارگیری معیارهای اطلاعاتی نظیر AIC و BIC در این چارچوب امکان‌پذیر و رایج نیست. از این‌رو، مقایسه مدل‌ها بر مبنای شاخص‌های متداول دقت پیش‌بینی انجام شده است.

جدول ۶. ارزیابی عملکرد الگوریتم KNN برای مجموعه داده‌های آموزش

معیارهای خطا	E^*		ϕ	
	مدل کاهش‌یافته	مدل کامل	مدل کاهش-یافته	مدل کامل
MAE	2138	2066	3.7	3.3
MAPE	0.37	0.35	0.23	0.19
MSE	12241410	11547556	4.1	21
RMSE	3499	3398	5.5	4.5
R^2	0.87	0.88	0.75	0.82
Se	3404	3333	5.3	4.3
Sy	9722	9722	10.8	10.8
Se/Sy	0.35	0.34	0.49	0.40
ارزیابی	Good Excellent	Good Excellent	Good Good	Good Good

1- Very good

مقایسه‌ی شاخص‌های خطا برای پیش‌بینی‌های $|E^*|$ و ϕ نشان‌دهنده‌ی بهبود جزئی در مدل کامل نسبت به مدل کاهش‌یافته است. همین الگو در مورد ضریب تعیین R^2 و نسبت Se/Sy نیز مشاهده می‌شود. ارزیابی کلی مدل‌ها برای پیش‌بینی $|E^*|$ ، با مقادیر R^2 بزرگ‌تر از ۰/۸۷ و مقادیر Se/Sy کوچک‌تر از ۰/۳۵، بین خوب تا عالی^۲ متغیر است. برای زاویه فاز (ϕ) نیز ارزیابی در سطح خوب قرار دارد، با R^2 بزرگ‌تر از ۰/۷۹ و Se/Sy کمتر از ۰/۴۴ برای هر دو مدل. در مجموع، در این مطالعه، با مقایسه‌ی دو مدل کاهش‌یافته و کامل الگوریتم KNN توسعه‌یافته، مدل کامل که شامل VMA و VFA به‌عنوان ورودی‌های تکمیلی است، در تمامی شش شاخص عملکردی، کمی بهتر از مدل کاهش‌یافته عمل می‌کند. علاوه بر این، معیارهای عملکرد مدل به‌صورت جداگانه برای مجموعه داده‌های آزمون نیز محاسبه شد که نتایج آن در جدول ۷ ارائه شده است.

جدول ۷. ارزیابی عملکرد الگوریتم KNN برای مجموعه داده‌های آزمون

معیارهای خطا	E^*		ϕ	
	مدل کاهش‌یافته	مدل کامل	مدل کاهش-یافته	مدل کامل
MAE	2360	2225	4.1	3.7
MAPE	0.40	0.38	0.23	0.21
MSE	14304124	12705624	30	25
RMSE	3780	3560	5.5	5.0
R^2	0.84	0.86	0.75	0.78
Se	3355	3180	5.3	4.9
Sy	9722	9722	10.8	10.8
Se/Sy	0.34	0.33	0.49	0.45
ارزیابی	Good Excellent	Good Excellent	Good Good	Good Good

مقایسه معیارهای عملکرد مدل برای مجموعه‌های آموزش و آزمون، یعنی نتایج جدول ۶ و ۷ نتایج نشان

2- Excellent

Driven به منظور توسعه مدل مبتنی بر یادگیری مستقیم از داده‌های آزمایشگاهی موجود و بدون فرض روابط تحلیلی از پیش تعریف‌شده است و به معنای اتکا به مجموعه‌های میدانی بسیار گسترده نیست. نتایج زیر از بررسی‌های صورت‌گرفته در این پژوهش حاصل شده است:

- مقدار بهینه‌ی K برابر ۵، به‌عنوان پارامتر نهایی انتخاب گردید، زیرا کمترین مقدار NRMSE را فراهم کرد. عملکرد هر دو مدل، بر اساس شاخص‌های خطا و معیارهای برازش آماری، در بازه‌ی خوب تا عالی ارزیابی شد؛ با این تفاوت که مدل کامل اندکی عملکرد بهتری از مدل کاهش‌یافته نشان داد. برای خاصیت E^* ، مقادیر R^2 برابر ۰/۸۸ و Se/Sy برابر ۰/۳۴ به‌دست آمد. همچنین، برای Φ نیز این مقادیر به‌ترتیب ۰/۸۲ و ۰/۴۰ به‌دست آمد.
- اگرچه تحلیل پیرسون نشان داد که همبستگی بین پارامترهای حجمی VMA و VFA و ویژگی‌های ویسکوالاستیک مخلوط‌های آسفالتی ضعیف است. اما افزودن این پارامترها، دقت پیش‌بینی مقادیر E^* و Φ را بهبود بخشید. این پارامترها به‌صورت معمول و عادی در فرآیند طرح اختلاط مخلوط آسفالتی محاسبه می‌شوند و تعیین آن‌ها نیازمند زمان یا هزینه‌ی اضافی نیست و از این‌رو کآمد خواهند بود. این نتیجه با شواهد تجربی موجود مبنی بر اهمیت قابل توجه این پارامترها در رفتار مکانیکی مخلوط‌های آسفالتی همخوانی دارد.
- هر دو مدل کاهش‌یافته و کامل الگوریتم KNN توسعه‌یافته قادرند به‌طور هم‌زمان، تخمین‌های دقیقی از مقادیر E^* و Φ را ارائه دهند. در نتیجه، این دست‌آورد نیاز به انجام آزمایش‌های آزمایشگاهی پرهزینه و زمان‌بر را کاهش می‌دهد. با این حال، باید توجه داشت که عمومیت‌بخشی

می‌دهد که اختلاف محسوسی میان عملکرد دو مجموعه وجود ندارد و نشانه‌ای از بیش‌برازش قابل توجه مشاهده نمی‌شود. همچنین باید تأکید شود که هر دو مدل یادگیری ماشین، به‌گونه‌ای آموزش داده شده‌اند که به‌صورت هم‌زمان قادر به پیش‌بینی مدول دینامیکی E^* و زاویه فاز Φ باشند. بدین ترتیب، این مدل‌ها قادرند بینش جامع‌تر و دقیق‌تری از ویژگی‌های ویسکوالاستیک بنیادی مخلوط‌های آسفالتی را ارائه دهند.

۵. نتیجه‌گیری

در این مطالعه سعی شد تا یک رویکرد جامع و دقیق برای توسعه‌ی یک الگوریتم نوآورانه‌ی یادگیری ماشین تحت عنوان K نزدیک‌ترین همسایه (KNN) ارائه گردد که هدف آن پیش‌بینی هم‌زمان مدول دینامیکی E^* و زاویه فاز Φ در مخلوط‌های آسفالتی است. مهم‌ترین مزیت الگوریتم KNN نسبت به روش‌های متداول پیش‌بینی در این است که برای برقراری روابط بین عوامل مؤثر، نیازی به فرضیات آماری، ساده‌سازی‌ها یا ملاحظات تجربی ندارد. برای توسعه‌ی مدل‌ها، یک پایگاه داده منحصربه‌فرد از ترکیب چهار مجموعه داده‌ی معتبر موجود در منابع علمی گردآوری شد که شامل ویژگی‌های حجمی اصلی مخلوط‌ها، خصوصیات قیر و شرایط آزمایش‌ها می‌باشد. متغیرهای ورودی مورد نیاز برای مدل‌های پیش‌بینی، در مرحله‌ی طراحی اولیه‌ی مخلوط قابل اندازه‌گیری یا استخراج هستند. مدل نخست، با عنوان مدل کاهش‌یافته از ورودی‌های $T, V_a, V_b, \log(\text{Visc})$ و f استفاده کرد. همچنین، مدل دوم، با عنوان مدل کامل علاوه بر این متغیرها، VMA و VFA را نیز به‌عنوان ورودی‌های تکمیلی در نظر گرفت. هر دو مدل در محیط Excel پیاده‌سازی شدند و کل پایگاه داده به دو بخش تقسیم شد که ۷۵٪ آن‌ها برای آموزش و ۲۵٪ برای ارزیابی به‌کار گرفته شدند. قابل به ذکر است که در این پژوهش، اصطلاح Data-

پایگاه داده و هزینه‌ی محاسباتی نسبی مدل‌های پیاده‌سازی شده همچنان از چالش‌های موجود است.

روش تحقیق ارائه شده می‌تواند در مطالعات آینده برای تحلیل پایگاه‌های داده‌ی بزرگ‌تر و متنوع‌تر به کار گرفته شود تا ابزاری قدرتمند و کارآمد برای پیش‌بینی دقیق خواص مخلوط‌های آسفالتی و به‌کارگیری آن‌ها در روش‌های طراحی روسازی مکانیکی-تجربی فراهم گردد.

۷. مطالعات آتی

فرایند اعتبارسنجی مدل در این پژوهش بر اساس یک تقسیم‌بندی تصادفی منفرد با نسبت ۷۵٪ داده‌های آموزشی و ۲۵٪ داده‌های آزمون انجام شده است. این رویکرد با هدف ایجاد توازن میان حجم داده‌های آموزشی کافی و مجموعه‌ی آزمون نماینده برای ارزیابی عملکرد مدل، و با توجه به ماهیت داده‌های آزمایشگاهی و محدودیت در دسترسی به پایگاه‌های داده‌ی مستقل اتخاذ گردیده است. لازم به تأکید است که در این مطالعه از روش‌های پیشرفته‌تر اعتبارسنجی

نظیر k-fold cross-validation یا مجموعه داده‌ی مستقل منتشرنشده استفاده نشده و بدین ترتیب نتایج ارائه‌شده بیانگر ارزیابی اولیه عملکرد مدل در دامنه‌ی داده‌های موجود می‌باشند. توسعه‌ی چارچوب اعتبارسنجی با بهره‌گیری از این روش‌های قوی‌تر در مطالعات آتی به‌عنوان یکی از مسیرهای مهم پژوهشی پیشنهاد می‌شود. لازم به تأکید است که نتایج این پژوهش صرفاً شامل پیش‌بینی مقادیر میانگین پاسخ و ارزیابی دقت مدل بر اساس شاخص‌های آماری کلاسیک بوده و تحلیل عدم قطعیت یا برآورد فواصل اطمینان برای مقادیر پیش‌بینی‌شده در چارچوب این مطالعه انجام نشده است. از آنجا که در کاربردهای مهندسی روسازی، آگاهی از دامنه‌ی تغییرات احتمالی نتایج پیش‌بینی در اتخاذ تصمیمات محافظه‌کارانه از اهمیت ویژه‌ای برخوردار است، توسعه‌ی چارچوب‌های احتمالاتی نظیر روش بوت‌استرپ، شبیه‌سازی مونت کارلو یا به‌کارگیری مدل‌های یادگیری ماشین مبتنی بر توزیع احتمال به‌عنوان مسیر مهم پژوهش‌های آتی پیشنهاد می‌شود.

۸. مراجع

- Abu Abdo, A., Bayomy, F., Nielsen, R., Weaver, T. and Jung, S. J. 2009. "Prediction of the dynamic modulus of Superpave mixes". Proc. 8th Int. Conf. Bearing Capacity Roads, Railways, Airfields, 305–314. DOI: 10.1201/9780203865286.ch33
- Ahmed, M. S., N'diaye, M., Attouch, M. K., et al. 2023. "K-nearest neighbors prediction and classification for spatial data". J. Spat. Econom., 4: 12. <https://doi.org/10.1007/s43071-023-00041-2>
- Aksoy, A., Iskender, E. and Kahraman, H. T. 2012. "Application of the intuitive KNN estimator for prediction of the Marshall test (ASTM D1559) results for asphalt mixtures". Constr. Build. Mater. 34: 561–569. <https://doi.org/10.1016/j.conbuildmat.2012.02.091>
- Al-Dosary, N. M. N., Al-Hamed, S. A. and Aboukarima, A. M. 2019. "K-nearest neighbors method for prediction of fuel consumption in tractor–chisel plow systems". Eng. Agric., 39(6): 729–736. <https://doi.org/10.1590/1809-4430-Eng.Agric.v39n6p729-736/2019>
- Al-Khateeb, G., Shenoy, A., Gibson, N. and Harman, T. 2006. "A new simplistic model for dynamic modulus predictions of asphalt paving mixtures". J. Assoc. Asphalt Paving Technol., 75: 1254–1293. <https://doi.org/10.1016/j.cscm.2022.e01580>
- Alnaqbi, A., Zeiada, W. and Al-Khateeb, G. G. 2024. "Machine learning modeling of pavement performance and IRI prediction in flexible pavement". Innov. Infrastruct. Solut., 9: 385. <https://doi.org/10.1007/s41062-024-01688-y>

- Andrei, D., Mirza, W. and Witzczak, M. W. 1999. "Development of a revised predictive model for the dynamic (complex) modulus of asphalt mixtures". NCHRP Rep. 1-37A. <https://doi.org/10.4236/ojce.2020.103018>
- Atakan, M. and Yıldız, K. 2023. "Prediction of Marshall design parameters of asphalt mixtures via machine learning algorithms based on literature data". Road Mater. Pavement Design, 25(3): 454–473. <https://doi.org/10.1080/14680629.2023.2213774>
- Bajic, M., Pour, S. M., Skar, A., Pettinari, M., Levenberg, E. and Alstrøm, T. S. 2021. "Road roughness estimation using machine learning". arXiv preprint arXiv:2107.01199. <https://doi.org/10.48550/arXiv.2107.01199>
- Bari, J. and Witzczak, M. W. 2006. "Development of a new revised version of the Witzczak E* predictive model for hot mix asphalt mixtures". J. Assoc. Asphalt Paving Technol., 75: 381–424. <https://doi.org/10.1016/j.conbuildmat.2014.11.011>
- Behnood, A. and Mohammadi Golafshani, E. 2021. "Predicting the dynamic modulus of asphalt mixture using machine learning techniques". Constr. Build. Mater., 266 (Part A): 120983. <https://doi.org/10.1016/j.conbuildmat.2020.120983>
- Beskopylny, A. N., Stelmakh, S. A., Shcherban, E. M., Mailyan, L. R., Meskhi, B., Razveeva, I., Chernilnik, A. and Beskopylny, N. 2022. "Concrete strength prediction using machine learning methods". Appl. Sci., 12: 10864. <https://doi.org/10.3390/app122110864>
- Biligiri, K. P., Kaloush, K. and Uzan, J. 2010. "Evaluation of asphalt mixtures' viscoelastic properties using phase angle relationships". Int. J. Pavement Eng., 11(2): 143–152. <https://doi.org/10.1080/10298430903033354>
- Birgisson, B., Roque, R., Kim, J. and Pham, L. V. 2004. "The use of complex modulus to characterize the performance of asphalt mixtures and pavements in Florida". University of Florida. https://doi.org/10.1080/14680629.2018.1441065?urlappend=%3Futm_source%3Dresearchgate.net%26medium%3Darticle
- Botella, R., Lo Presti, D., Vasconcelos, K., et al. 2022. "Machine learning techniques to estimate the degree of binder activity of reclaimed asphalt pavement". Mater. Struct., 55: 112. <https://doi.org/10.1617/s11527-022-01933-9>
- Chaabene, W. B., Flah, M. and Nehdi, M. L. 2020. "Machine learning prediction of mechanical properties of concrete". Constr. Build. Mater., 260: 119889. <https://doi.org/10.1016/j.conbuildmat.2020.119889>
- Christensen, D. W., Pellinen, T. and Bonaquist, R. F. 2003. "Hirsch model for estimating the modulus of asphalt concrete". J. Assoc. Asphalt Paving Technol., 72: 97–121. [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0002099](https://doi.org/10.1061/(ASCE)MT.1943-5533.0002099)
- Choi, H. J., Kim, S., Kim, Y. and Won, J. 2022. "Predicting frost depth of soils in South Korea using machine learning". Sustain., 14: 9767. <https://doi.org/10.3390/su14159767>
- Dhar, V. and Stein, R. 1997. "Intelligent decision support methods". Prentice-Hall, Upper Saddle River, N. J. https://doi.org/10.1007/978-1-4615-1147-2_7
- Flintsch, G., Loulizi, A., Diefenderfer, S. D., Galal, K. A. and Diefenderfer, B. K. 2007. "Asphalt materials characterization in support of M-E PDG implementation". Rep. VTRC 07-CR10, Virginia Tech. <https://doi.org/10.3141/2057-14>
- Ghorbani, B., Yaghoubi, E., Wasantha, P. L. P., van Staden, R., Guerrieri, M. and Fragomeni, S. 2023. "Machine learning-based prediction of resilient modulus for blends of tire-derived aggregates and demolition wastes". Road Mater. Pavement Design, 25(4): 694–715. <https://doi.org/10.1080/14680629.2023.2222176>
- Gul, M. A., Islam, M. K., Awan, H. H., et al. 2022. "Prediction of Marshall stability and Marshall flow using ML". Symmetry, 14: 2324. <https://doi.org/10.3390/sym14112324>
- Huang, Y., Molavi Nojumi, M., Ansari, S., Hashemian, L. and Bayat, A. 2023. "Evaluating the use of ML for moisture content prediction". Road Mater. Pavement Design, 24(12): 2910–2928. <https://doi.org/10.1080/14680629.2023.2182135>
- Justo-Silva, R., Ferreira, A. and Flintsch, G. 2021. "Review on ML techniques for pavement performance". Sustain., 13: 5248. <https://doi.org/10.3390/su13095248>
- Khambra, G. and Shukla, P. 2021. "Novel ML applications on fly ash-based concrete". Mater. Today Proc., 80: 6245. <https://doi.org/10.1016/j.matpr.2021.07.262>
- Leiva-Villacorta, F. and Vargas-Nordbeck, A. 2019. "Neural network-based model to estimate dynamic modulus E*". J. Soft Comput. Civ. Eng., 3(2): 1–15. <https://doi.org/10.1016/j.jrmge.2019.03.005>
- Majidifard, H., Jahangiri, B., Buttlar, W. G. and Alavi, A. H. 2019. "New ML-based prediction models for fracture energy". Measurement, 135: 438–451. <https://doi.org/10.1016/j.measurement.2018.11.081>

- Martínez, F. and Angelone, S. 2010. "The estimation of the dynamic modulus using ANN". 11th Int. Conf. Asphalt Pavements, Vol. 1, 354–363. <https://doi.org/10.1016/j.conbuildmat.2020.119912>
- Naik, A. K. and Biligiri, K. P. 2014. "Predictive models to estimate phase angle". J. Mater. Civ. Eng., 27(8). [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0001197](https://doi.org/10.1061/(ASCE)MT.1943-5533.0001197)
- Picado-Santos, L., Capita, S. D. and Pais, J. C. 2003. "Stiffness modulus and phase angle prediction models for high modulus bituminous mixtures". Int. J. Pavements, 2(3). <https://doi.org/10.1016/j.conbuildmat.2018.09.160>
- Rahman, S., Bhasin, A. and Smit, A. 2021. "Exploring ML to predict asphalt mixture performance". Constr. Build. Mater., 295: 123585. <https://doi.org/10.1016/j.conbuildmat.2021.123585>
- Rondinella, F., Daneluz, F., Baldo, N. and Hofko, B. 2023. "Improved predictions of asphalt concretes' dynamic modulus". Constr. Build. Mater., 400: 132709. <https://doi.org/10.1016/j.conbuildmat.2023.132709>
- Rondinella, F., Daneluz, F., Hofko, B. and Baldo, N. 2024. "ML approach for simultaneous prediction of E* and phase angle". Adv. Res. Technol. Inf. Innov. Sustain., 1935. https://doi.org/10.1007/978-3-031-48858-0_40
- Seman, G., Shoop, S., McGrath, S. and Rollings, R. 2006. "Soil strength prediction with KNN". Proc. 59th Can. Geotech. Conf. <https://doi.org/10.1177/03611981241245679>
- Shu, X. and Huang, B. 2009. "Predicting dynamic modulus with differential method". Road Mater. Pavement Design, 10(2): 337–359. [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0001048](https://doi.org/10.1061/(ASCE)MT.1943-5533.0001048)
- Singh, D., Zaman, M. and Commuri, S. 2011. "Evaluation of predictive models for estimating dynamic modulus". Transp. Res. Record, 2210: 57–72. https://doi.org/10.3141/2210-07?urlappend=%3Futm_source%3Dresearchgate.net%26medium%3Darticle
- Suleymanov, A., Tuktarova, I., Belan, L., et al. 2023. "Spatial prediction of soil properties using random forest, k-nearest neighbors and cubist approaches in the foothills of the Ural Mountains, Russia". Model. Earth Syst. Environ., 9: 3461–3471. <https://doi.org/10.1007/s40808-023-01723-4>
- Tangga, A. A., Al Mufargi, H., Milad, A., et al. 2024. "Utilising machine learning algorithms to predict the Marshall characteristics of asphalt pavement layers". Innov. Infrastruct. Solut. 9: 381. <https://doi.org/10.1007/s41062-024-01698-w>
- Tran, N. and Hall, K. 2005. "Evaluating the predictive equation for dynamic moduli". J. Assoc. Asphalt Paving Technol., 74: 1–17. <https://www.tandfonline.com/action/showCitFormats?doi=10.1080/10298436.2020.1841191>
- Upadhyay, A., Thakur, M. S. and Sihag, P. 2024. "Predicting Marshall stability of carbon fiber–reinforced asphalt concrete". Int. J. Pavement Res. Technol., 17: 102–122. <https://doi.org/10.1007/s42947-022-00223-5>
- Useche-Castelblanco, J. S., Reyes-Ortiz, O. J. and Alvarez, A. E. 2023. "Application of ML models for wax-modified asphalt binders". Constr. Build. Mater., 395: 132352. <https://doi.org/10.1016/j.conbuildmat.2023.132352>
- Uwanuakwa, I. D., Busari, A., Ali, S. I. A., et al. 2022. "Comparing ML models with Witczak NCHRP 1-40D model". Arab. J. Sci. Eng., 47: 13579–13591. <https://doi.org/10.1007/s13369-022-06935-x>
- Venudharan, V. and Biligiri, K. P. 2015. "Estimation of phase angles using resilient modulus test". Constr. Build. Mater., 82: 274–286. <https://doi.org/10.1016/j.conbuildmat.2015.02.061>
- Weiss, S. M. and Kulikowski, C. A. 1991. "Computer systems that learn". Morgan Kaufmann Publ., San Mateo, CA. [https://doi.org/10.1016/S0957-4174\(02\)00026-X](https://doi.org/10.1016/S0957-4174(02)00026-X)
- Witczak, M. W. 2005. "NCHRP Proj. 9-19, Superpave support and performance models management". Final Rep. <https://doi.org/10.1051/mateconf/201712007003>
- Witczak, M. W. and Fonseca, O. A. 1996. "Revised predictive model for dynamic (complex) modulus". Transp. Res. Record, 1540(1): 15–23. <https://doi.org/10.1177/0361198196154000103>
- Yang, X. and You, Z. 2015. "New predictive equations for dynamic modulus and phase angle using nonlinear least-squares regression". J. Mater. Civ. Eng., 27(3). [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0001070](https://doi.org/10.1061/(ASCE)MT.1943-5533.0001070)
- Zhang, C. and Shen, S. 2017. "Modification of the Hirsch dynamic modulus prediction model". J. Mater. Civ. Eng., 29(12). [https://doi.org/10.1061/\(ASCE\)MT.1943-5533.0002099](https://doi.org/10.1061/(ASCE)MT.1943-5533.0002099)