Journalof Farayandno – Vol. 21 – No. 93 (2026): 79-90

Review Article

DOI: 10.22034/farayandno.2026.2070453.2013

This journal is an open access journal licensed under an Attribution-Non Commercial 4.0 International Licenses (CC BY-NC 4.0).

Accurate Prediction of Polymer Specific Heat Capacity Using Advanced Machine Learning Techniques

Zeynab Torabi [1], Elnaz Farzinfar [1], Fatemeh Sepahvand [1], Mohadeseh sarlak [1], Alireza Baghban [2,3*]

[1] Student, Amirkabir University of Technology, Mahshahr Campus, Mahshahr, Iran

[2] Professor, National Iranian South Oilfeilds Company (NISOC), Ahvaz, Iran

[3] National Iranian South Oilfeilds Company (NISOC), Ahvaz, Iran

Received: 28 Dec 2025 **Accepted:** 17 May 2026

Abstract

This study focuses on predicting the specific heat of polymers using experimental data and machine learning algorithms. A comprehensive dataset comprising physical and chemical properties of polymers was utilized as input for various models, including linear regression, Ridge, Lasso, ElasticNet, K-nearest neighbors, artificial neural networks, and decision trees. The performance of these models was evaluated using mean squared error (MSE) and R-squared (R^2) metrics. The results indicate that the Lasso and ElasticNet models provide the best balance between accuracy and generalization to new data. Specifically, the Lasso model achieved an MSE of 728.77 and an R^2 of 0.95, while the ElasticNet model recorded an MSE of 604.34 and an R^2 of 0.95. This study emphasizes that employing regularization techniques and ensemble models can significantly enhance the accuracy and stability of predictions for the thermal properties of polymers, leading to substantial savings in both time and costs associated with experimental testing.

Keyword: Thermal Property, Machine Learning, Intelligent Algorithms, Specific Heat, Polymers.

* Alireza_baghban@alumni.ut.ac.ir

Please Cite This Article Using:

Torabi, Z., Farzinfar, E., Sepahvand, F., sarlak, M., Baghban, A., "Accurate Prediction of Polymer Specific Heat Capacity Using Advanced Machine Learning Techniques", Journal of Farayandno – Vol. 21 – No. 93, pp. 79-90, In Persian, (2026).

مقاله ترویجی

DOI: 10.22034/farayandno.2026.2070453.2013

This journal is an open access journal licensed under an Attribution-Non Commercial 4.0 International Licenses (CC BY-NC 4.0).

پیش‌بینی دقیق ظرفیت گرمایی ویژه پلیمرها با بهره‌گیری از تکنیک‌های نوین یادگیری ماشین

زینب ترابی [1]، الناز فرزین‌فر [1]، فاطمه سپهوند [1]، محدثه سرلک [1]، علیرضا باغبان [3,2]*

[1] دانشجو ، دانشگاه صنعتی امیرکبیر واحد ماهشهر، ماهشهر، ایران

[2] استاد، دانشگاه صنعتی امیرکبیر واحد ماهشهر، ماهشهر، ایران

[3] کارمند، شرکت ملی مناطق نفتخیز جنوب، اهواز، ایران

دریافت: 1404/10/07 پذیرش: 1405/02/27

چکیده

این پژوهش به پیش‌بینی گرمای ویژه پلیمرها با استفاده از داده‌های تجربی و الگوریتم‌های یادگیری ماشین می‌پردازد. مجموعه‌ای از داده‌های آزمایشگاهی شامل ویژگی‌های فیزیکی و شیمیایی پلیمرها به‌عنوان ورودی برای مدل‌های مختلف نظیر رگرسیون خطی، ستیغی، لاسو، الاستیک نت، نزدیک‌ترین همسایگی، شبکه عصبی مصنوعی و درخت تصمیم به‌کار گرفته شد. عملکرد این مدل‌ها با معیارهای میانگین مربع خطا و ضریب رگرسیون ارزیابی گردید. نتایج نشان می‌دهد که مدل‌های لاسو و الاستیک نت بهترین تعادل را بین دقت و قابلیت تعمیم به داده‌های جدید ارائه می‌دهند. به‌طور خاص، مدل لاسو با میانگین مربع خطا معادل 728/77 و ضریب رگرسیون برابر 0/95 و مدل الاستیک نت با میانگین مربع خطا برابر 604/34 و ضریب رگرسیون برابر 0/95 ارزیابی شدند. این مطالعه تأکید می‌کند که استفاده از تکنیک‌های منظم‌سازی می‌تواند دقت و پایداری پیش‌بینی خواص گرمایی پلیمرها را به‌طور قابل توجهی افزایش دهد و در هزینه و زمان آزمایش‌های تجربی صرفه‌جویی نماید.

کلمات کلیدی: خواص گرمایی، یادگیری ماشین، الگوریتم‌های هوشمند، گرمای ویژه، پلیمرها.

* Alireza_baghban@alumni.ut.ac.ir

استناد به مقاله:

ترابی،ز.، فرزین‌فر،الف.، سپهوند،ف.، سرلک،م.، باغبان،ع.، "پیش‌بینی دقیق ظرفیت گرمایی ویژه پلیمرها با بهره‌گیری از تکنیک‌های نوین یادگیری ماشین"، نشریه فرآیندنو، سال بیست و یکم، شماره 93، صص.90-79، (بهار 1405).

1- مقدمه

پلیمرها به‌عنوان مواد مهندسی با کاربرد وسیع در صنایعی مانند خودروسازی، بسته‌بندی، الکترونیک و پزشکی شناخته می‌شوند؛ این مسئله ناشی از ویژگی‌هایی چون وزن سبک، قابلیت طراحی و پایداری حرارتی و شیمیایی این مواد است [1]. در میان ویژگی‌های متعدد پلیمرها، گرمای ویژه به‌دلیل تأثیر مستقیم در فرآیندهایی مانند انتقال حرارت، شکل‌دهی حرارتی و کنترل حرارتی اهمیت دارد [2]. اگرچه اندازه‌گیری تجربی این خاصیت در دستگاه‌های پیشرفته مانند کالریمتری پویشی تفاضلی دقیق است، این روش معمولاً زمان‌بر، نیازمند هزینه بالا و محدود به شرایط ویژه آزمایشگاهی است [3]. از این رو، ظهور روش‌های داده‌محور، به‌ویژه استفاده از یادگیری ماشین، می‌تواند به‌عنوان روشی مؤثر، سریع و اقتصادی برای پیش‌بینی گرمای ویژه پلیمرها مطرح شود [4].

یادگیری ماشین با تحلیل داده‌های آزمایشگاهی و استخراج روابط پنهان میان ویژگی‌های فیزیکی و شیمیایی پلیمرها، امکان توسعه مدل‌های پیش‌بینی قدرتمند را فراهم می‌سازد. در پژوهش‌های گذشته، الگوریتم‌هایی نظیر شبکه‌های عصبی مصنوعی برای پیش‌بینی رسانایی حرارتی کامپوزیت‌های پلیمری [5]، نزدیک‌ترین همسایگی برای برآورد ویسکوزیته ذاتی پلیمرهای محلول [6]، درخت تصمیم برای تحلیل اثر نانوذرات بر خواص حرارتی پلیمرها [7] و رگرسیون خطی برای تخمین گرمای ویژه پلیمرهای ساده [4] مورد استفاده قرار گرفته‌اند. همچنین در مطالعات جدیدتر، الگوریتم‌های پیشرفته‌تر مانند جنگل تصادفی و گرادیان بوستینگ برای پیش‌بینی گرمای ویژه پلیمرها [8] و مدل‌های شبکه‌های عمیق برای تخمین هم‌زمان چند خاصیت گرمایی [9] به کار گرفته شده‌اند.

با وجود این پیشرفت‌ها، بسیاری از پژوهش‌ها تنها بر یک الگوریتم خاص یا مجموعه‌داده محدود تمرکز داشته‌اند. در تحقیق حاضر، با بهره‌گیری از چند الگوریتم شامل رگرسیون خطی، نزدیک‌ترین همسایگی، درخت تصمیم و شبکه عصبی مصنوعی و مقایسه عملکرد آن‌ها، چارچوبی بهینه برای پیش‌بینی گرمای ویژه پلیمرها ارائه شده است. داده‌های واقعی و متنوع به‌عنوان ورودی به مدل‌ها داده شده و دقت هر الگوریتم بر اساس معیارهایی همچون میانگین مربعات خطا و توان دوم ضریب رگرسیون (R^2) ارزیابی گردید. نتایج نشان دادند که شبکه عصبی بهینه بهترین عملکرد را داشته و می‌تواند به‌عنوان الگویی قابل اعتماد در کاربردهای صنعتی مورد استفاده قرار گیرد.

نوآوری اصلی این پژوهش در استفاده مقایسه‌ای از چند الگوریتم مختلف، بهره‌گیری از داده‌های واقعی و بررسی کاربرد صنعتی مدل پیشنهادی است. یافته‌ها نشان می‌دهند که استفاده از یادگیری ماشین در پیش‌بینی خواص گرمایی پلیمرها نه تنها موجب کاهش هزینه‌ها و صرفه‌جویی در زمان می‌شود، بلکه می‌تواند گامی مؤثر در جهت هوشمندسازی انتخاب و طراحی مواد در مهندسی پلیمر باشد.

نوآوری این پژوهش از چند جهت قابل توجه است:

- برخلاف بسیاری از مطالعات پیشین که تنها از یک الگوریتم یادگیری ماشین بهره برده‌اند، در این تحقیق چندین الگوریتم مختلف به‌صورت مقایسه‌ای مورد استفاده قرار گرفته و عملکرد آن‌ها با یکدیگر سنجیده شده است. این رویکرد امکان انتخاب بهینه‌ترین مدل بر اساس شاخص‌های ارزیابی را فراهم می‌سازد.
- داده‌های متنوع و واقعی از منابع معتبر برای آموزش مدل‌ها استفاده شده‌اند.
- کاربرد عملی مدل پیشنهادی در کاهش هزینه‌های آزمایشگاهی و تسریع فرآیند انتخاب مواد در صنعت بررسی شده است.

نتایج این مطالعه می‌تواند به عنوان گامی مؤثر در جهت هوشمندسازی طراحی مواد و جایگزینی آزمون‌های تجربی پرهزینه در مهندسی پلیمر مورد استفاده قرار گیرد.

2- الگوسازی نظری یا تجربی

2-1- تئوری مدل

در این پژوهش، به‌منظور مدل‌سازی و پیش‌بینی گرمای ویژه پلیمرها، از مجموعه‌ای از الگوریتم‌های یادگیری ماشین شامل رگرسیون خطی، رگرسیون‌های منظم‌شده (ستیغی، لاسو و الاستیک‌نت)، نزدیک‌ترین همسایگی(KNN)، شبکه عصبی مصنوعی، درخت تصمیم و جنگل تصادفی استفاده شده است که در ادامه به اختصار معرفی می‌شوند [10-18].

- **رگرسیون خطی**: رابطه خطی بین متغیر هدف و ویژگی‌ها را با استفاده از روش حداقل مربعات خطا مدل‌سازی می‌کند.
- **رگرسیون‌های منظم‌شده:** شامل رگرسیون ستیغی (پنالتی L2)، لاسو (پنالتی L1، منجر به انتخاب ویژگی) و الاستیک‌نت (ترکیب L1 و L2) که برای بهبود عملکرد در حضور ویژگی‌های زیاد یا هم‌خطی به‌کار می‌روند.
- **نزدیکترین همسایگی:** یک الگوریتم غیرپارامتری که پیش‌بینی را بر اساس میانگین خروجی K همسایه نزدیک (با فاصله اقلیدسی) انجام می‌دهد.
- **شبکه عصبی مصنوعی:** ساختاری لایه‌ای با نورون‌های وزندار و توابع فعال‌سازی که با الگوریتم پس‌انتشار خطا آموزش دیده و برای روابط پیچیده غیرخطی مناسب است.
- **درخت تصمیم:** مدل درختی که داده را با قواعد شرطی بر اساس کاهش خطا یا افزایش اطلاعات تقسیم‌بندی می‌کند و تفسیرپذیر است.
- **جنگل تصادفی:** مجموعه‌ای از درختان تصمیم که بر روی نمونه‌های تصادفی از داده و ویژگی‌ها ساخته شده و نتیجه نهایی را با میانگین‌گیری (در رگرسیون) تعیین می‌کند. شایان ذکر است مهمترین ویژگی این الگوریتم دقت بالا و مقاومت در برابر بیش‌برازش می‌باشد.

2-2- داده‌های مورد استفاده در پژوهش

در این پژوهش، برای پیش‌بینی گرمای ویژه پلیمرها از مجموعه داده‌ای شامل ویژگی‌های ساختاری، فیزیکی و شیمیایی نمونه‌های پلیمری استفاده شد. این مجموعه شامل 68 نمونه است که برای هر نمونه، مقادیر متعددی از ویژگی‌های مولکولی ثبت شده‌اند. داده‌های مورد استفاده در مطالعه حاضر از مقاله [4] استخراج شده‌اند و در جدول ضمیمه پیوست شده است.

گرمای ویژه پلیمرها به‌عنوان متغیر خروجی (هدف) در مدل‌های یادگیری ماشین در نظر گرفته شده است. متغیرهای ورودی شامل خصوصیات ساختاری و توصیفگرهای شیمیایی و فیزیکی هستند که به‌طور مستقیم یا غیرمستقیم بر خواص گرمایی پلیمر تأثیرگذارند. برخی از مهم‌ترین این ویژگی‌ها شامل وزن مولکولی، ویژگی‌های ساختاری (از جمله تعداد کل اتم‌ها، پیوندها، زاویه‌ها و زاویه‌های پیچشی)، نوع اتم‌های موجود در ساختار و توصیف‌گرهای زاویه‌ای و

پیچشی خاص برای گروه‌های شیمیایی مختلف مانند C1-O1-C2-O3، H1-C1-C1-H1 و H1-C1-C1-N1 می‌باشند. افزون بر این، ده‌ها توصیفگر دیگر نیز از ساختار مولکولی استخراج شده‌اند.

در مجموع، 188 ویژگی ورودی برای هر نمونه ثبت شده است. این ویژگی‌ها به‌عنوان داده‌های ورودی به مدل‌های یادگیری ماشین داده شده‌اند تا از طریق آن‌ها مقدار گرمای ویژه پیش‌بینی شود. پیش از آموزش مدل‌ها، داده‌ها با روش استانداردسازی (Standard Scaler) مورد پیش‌پردازش قرار گرفتند. سپس، مجموعه‌داده به نسبت 80% برای آموزش و 20% برای آزمون تقسیم شد و داده‌های آموزش جهت ساخت مدل‌های مختلف یادگیری ماشین برای پیش‌بینی گرمای ویژه به‌کار گرفته شد.

2-3- روش‌های مدلسازی

فرایند مدل‌سازی در این پژوهش برای تمامی الگوریتم‌های مورد استفاده به‌صورت مرحله‌ای و ساختارمند انجام شد. در گام نخست، کتابخانه‌های مورد نیاز از جمله NumPy، Pandas، Matplotlib، Seaborn و ابزارهای مربوط به یادگیری ماشین از Scikit-learn فراخوانی شدند تا بستر مناسبی برای پردازش داده‌ها، ساخت مدل‌ها و تحلیل نتایج فراهم گردد. در مرحله بعد، مجموعه داده‌های گردآوری‌شده فراخوانی و بررسی اولیه روی آن‌ها انجام شد. پس از پاک‌سازی و آماده‌سازی داده‌ها، مجموعه به دو بخش آموزش و آزمون تقسیم گردید. داده‌های آموزشی برای ساخت مدل و داده‌های آزمون برای ارزیابی مدل‌های ساخته‌شده مورد استفاده قرار گرفتند. نسبت این تقسیم‌بندی به‌گونه‌ای انتخاب شد که تعادل مناسبی بین آموزش و اعتبارسنجی مدل برقرار گردد. سپس مدل‌های مورد نظر، شامل الگوریتم‌های رگرسیون خطی، رگرسیون منظم‌شده (ریج، لاسو، الاستیک‌نت)، نزدیک‌ترین همسایگی K، شبکه عصبی مصنوعی، جنگل تصادفی و درخت تصمیم، با استفاده از داده‌های آموزشی ساخته شدند. پس از آموزش مدل‌ها، عملکرد آن‌ها بر پایه داده‌های آزمون مورد ارزیابی قرار گرفت تا توانایی تعمیم مدل‌ها به داده‌های جدید بررسی شود. برای تحلیل دقت مدل‌ها، از معیارهای رایج در ارزیابی رگرسیون شامل ضریب رگرسیون (R^2)، میانگین مربعات خطا (MSE) استفاده شد. این معیارها امکان مقایسه کمی عملکرد مدل‌ها و انتخاب مدل بهینه را فراهم نمودند. در نهایت، جهت تفسیر بهتر نتایج، نمودارهایی شامل مقایسه مقادیر پیش‌بینی‌شده و واقعی، توزیع خطاها و سایر نمودارهای تحلیلی با استفاده از ابزارهای بصری‌سازی مانند Matplotlib و Seaborn ترسیم شد. این نمودارها به درک عمیق‌تری از عملکرد مدل‌ها و نقاط قوت و ضعف هر روش کمک کردند.

2-4- نتایج و بحث

برای پیش‌بینی گرمای ویژه پلیمرها، چندین مدل یادگیری ماشین مورد ارزیابی قرار گرفتند و نمودار ضریب رگرسیون آن‌ها برای داده‌های آموزش و تست در شکل 1 و شکل 2 و جدول 1 ارائه شده است. براساس داده‌های ارائه شده در جدول 1 می‌توان به تفاوت عملکرد مدل‌های مختلف یادگیری ماشین براساس دو معیار میانگین مربعات خطا و ضریب R^2 برای داده‌های آموزش و ازمون را مشاهده کرد.

مدل رگراسیون خطی ساده، با وجود داشتن R^2 بسیار بالا (0/99) و MSE پایین در داده‌های آموزش، در داده‌های آزمون عملکرد ضعیفی از خود نشان داد و MSE بسیار بالایی ثبت کرد. این امر به وضوح نشان‌دهنده بیش برازش شدید و عدم توانایی مدل در تعمیم به داده‌های جدید است. مدل‌های منظم شده، شامل ستیغی، لاسو و الاسایک

نت با اعمال تکنیک‌های منظم‌سازی، توانستند بیش برازش به وجود آمده در مدل کلاسیک خطی را به طور چشمگیری کاهش دهند و عملکرد تعمیم‌یافته بهتری بر روی داده‌های آزمون ارائه کنند. در حالی که رگراسیون ستیغی با R^2 برابر 0/89 نیز عملکرد قابل قبول و منطقی داشت، مدل‌های لاسو و الاستیک نت با ثبت R^2 برابر 0/95 و MSE پایین روی داده‌های آزمون، بهترین تعادل بین دقت مدل و قابلیت تعمیم را به نمایش گذاشتند. این نتایج تاییدی بر این اصل خواهد بود که استفاده از منظم‌سازی در مدل‌های خطی می‌تواند دقت و پایداری پیش‌بینی‌ها را به طور قابل توجهی افزایش دهد.

جدول 1- نتایج ارزیابی مدل های مختلف یادگیری ماشین

مدل	میانگین خطای مربعی در آموزش	میانگین خطای مربعی در آزمون	ضریب تعیین (R^2) در آموزش	ضریب تعیین (R^2) در آزمون
رگرسیون خطی	0/38	1841989545/89	0/99	125624/10
رگرسیون ریج	41/42	1565/11	0/99	0/89
رگرسیون لاسو	88/08	728/77	0/99	0/95
رگرسیون الاستیک نت	167/58	604/34	0/98	0/95
نزدیک‌ترین همسایگی	2247/68	2500/37	0/78	0/82
شبکه عصبی (ScikitLearn)	2/05	751/06	0/99	0/94
شبکه عصبی (Keras)	116/47	1905/15	0/98	0/87
درخت تصمیم	401/44	2548/29	0/96	0/82
جنگل تصادفی	187/43	1402/85	0/98	0/90

شبکه‌های عصبی مصنوعی نیز نتایج قابل توجهی ارائه کردند. نسخه پیاده‌سازی شده با کتابخانه Scikit-Learn با R^2 برابر 0/94 روی داده‌های آزمون، توانست دقت بالایی همراه با تعمیم‌پذیری مناسب داشته باشد، در حالی که نسخه ارائه شده با کتابخانه Keras با R^2 پایین‌تر (0/87) و MSE بالاتر، نشانه‌هایی از بیش برازش را نشان داد. این اختلاف می‌تواند ناشی از تفاوت در معماری شبکه، تنظیمات ابرپارامتر یا حجم داده‌های آموزشی باشد. در مدل‌های مبتنی بر درخت تصمیم، تفاوت عملکرد مشخصی بین درخت تصمیم و جنگل تصادفی مشهود است. مدل درخت تصمیم به دلیل تمایل بیش از حد به تطابق پیدا کردن با داده‌های آموزش، R^2 پایین‌تری روی داده‌های آزمون (0/82) داشت و بنابراین قابلیت تعمیم محدودی ارائه کرد. در مقابل، جنگل تصادفی با بهره‌گیری از تکنیک اجماع مدل‌ها[1] توانست بیش برازش را کاهش داده و R^2 برابر 0/90 روی داده‌های آزمون ثبت کند، که نشان‌دهنده بهبود قابل توجهی در تعمیم پذیری است. مدل نزدیکترین همسایگی نیز با وجود نداشتن بیش برازش شدید، عملکرد متوسطی داشت.

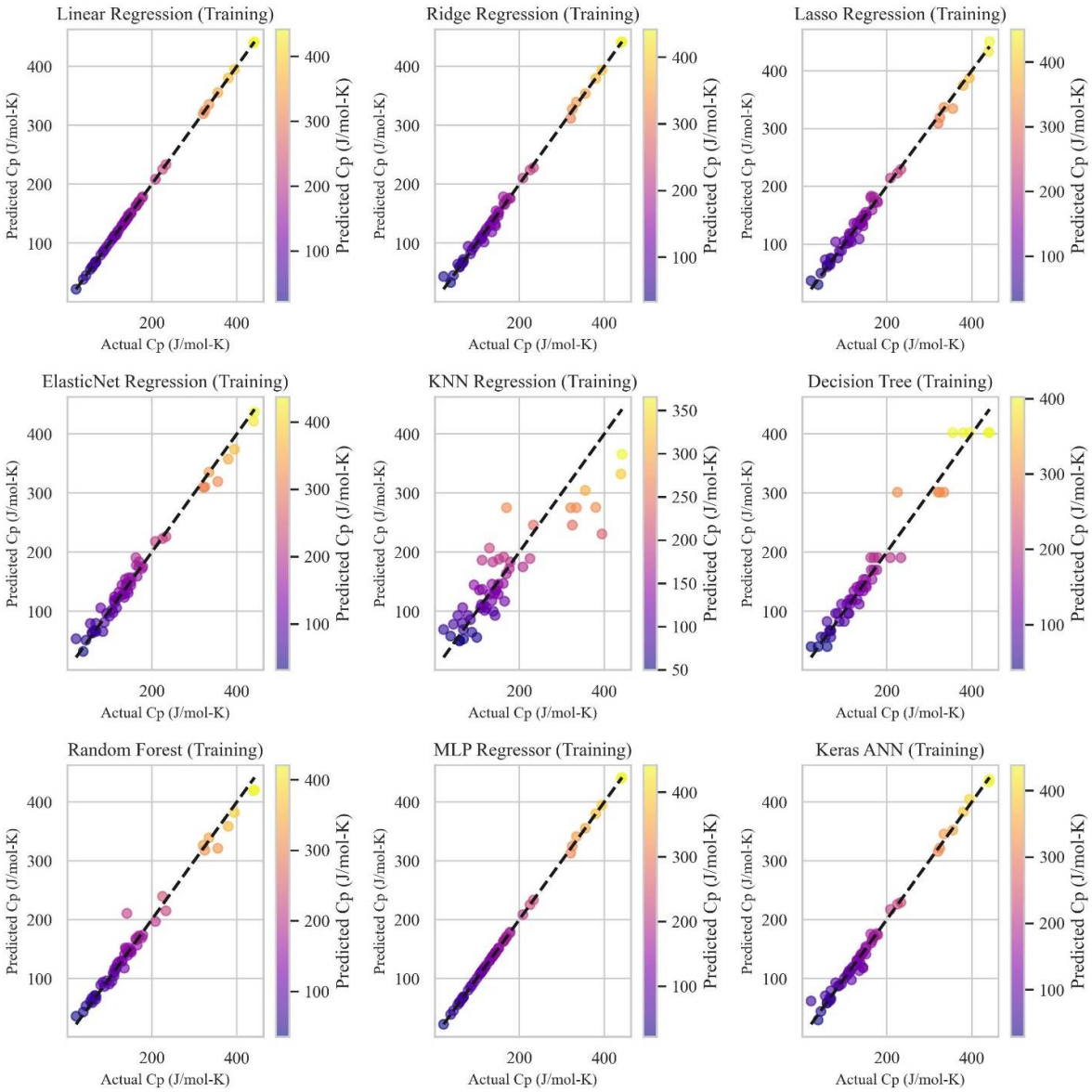
[1] ensemble

در این مدل R^2 برابر 0/82 و MSE نسبتاً بالا روی داده‌های آزمون نشان می‌دهد که دقت آن کمتر از مدل‌های خطی منظم شده و شبکه‌های عصبی بوده و احتمالاً نیازمند بهینه‌سازی پارامتر k و معیار فاصله است.
جدول 2 مقادیر آزمایشگاهی و پیشبینی شده توسط مدل های مختلف را برای 4 پلیمر Poly(styrene)، Nylon612، Poly(n-butyl methacrylate) و Nylon69 نشان می‌دهد.

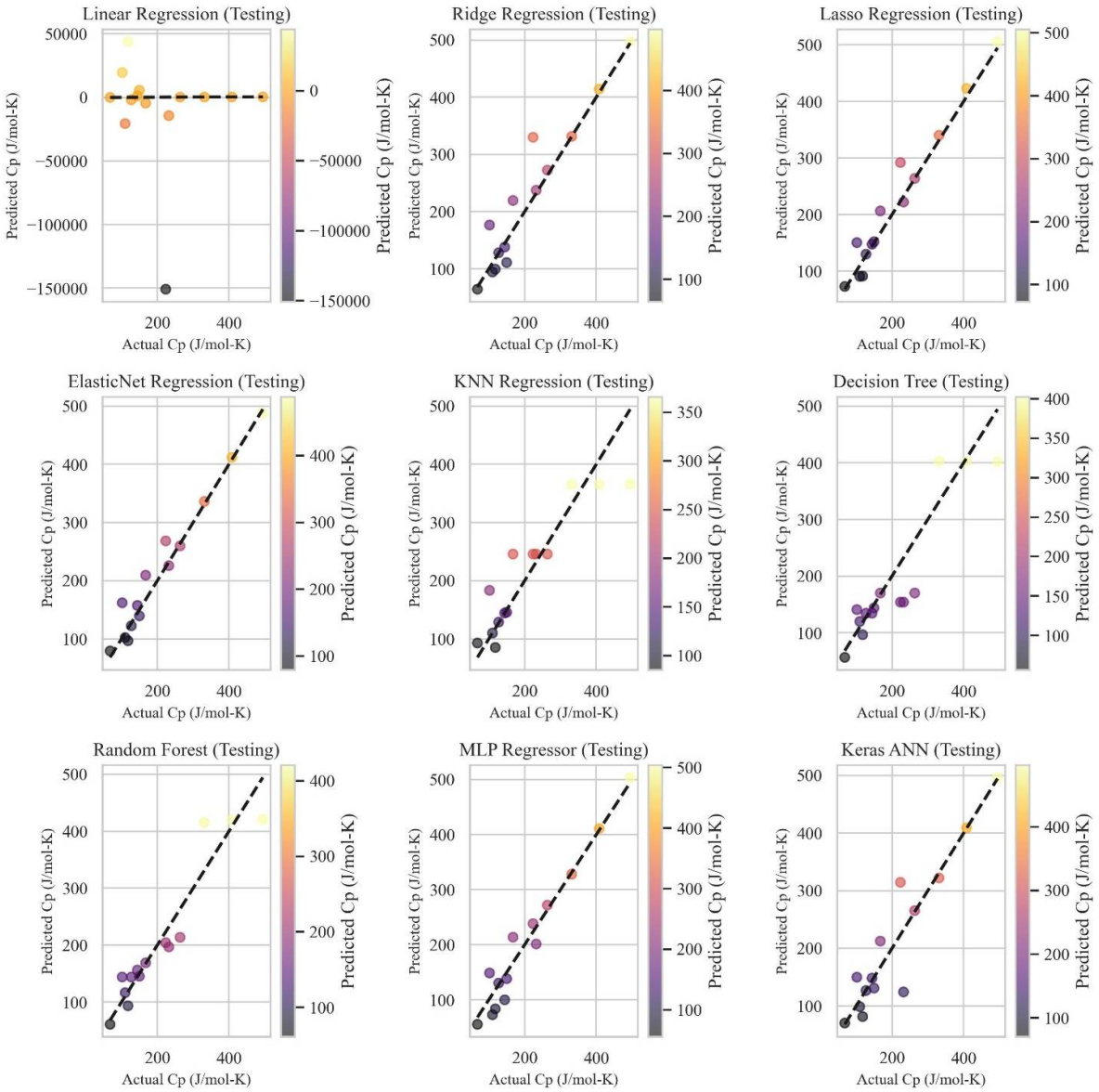
جدول 2- مقدار واقعی و پیشبینی شده Cp برای مدل های مختلف و 4 نمونه پلیمر

مقادیر Cp	پلیمر			
	Poly (n-butyl methacrylate)	Nylon612	Poly(styrene)	Nylon69
مقدار واقعی	263/41	494/48	127/38	408/08
رگرسیون خطی	275/8103	490/8397	-1823/19	408/5947
رگرسیون ریج	272/3824	496/9648	128/0616	414/1959
رگرسیون لاسو	263/8892	505/3966	130/0508	422/702
رگرسیون الاستیک نت	259/6889	487/5666	122/4495	411/7867
نزدیک‌ترین همسایگی	245/6633	365/74	128/78	365/74
درخت تصمیم	169/9167	402/2132	134/1873	402/2132
جنگل تصادفی	213/5897	420/7009	143/8953	419/8384
شبکه عصبی (ScikitLearn)	271/8971	503/8072	130/3287	410/825
شبکه عصبی(Keras)	265/4179	496/4916	126/8244	409/0018

با توجه به نتایج به دست آمده، مدل‌های لاسو و الاستیک نت بهترین ترکیب از دقت و توان تعمیم به داده‌های جدید را ارائه کرده‌اند و می‌توانند گزینه‌های مناسب و مطمئنی برای پیش‌بینی محسوب شوند. در مقابل، رگراسیون خطی ساده و درخت تصمیم بدون استفاده از تکنیک‌های منظم‌سازی، به دلیل بیش برازش شدید یا تعمیم پذیری ضعیف، گزینه‌های مناسبی برای استفاده بر روی این داده‌ها نیستند. این تحلیل نشان می‌دهد که اعمال تکنیک‌های منظم‌سازی و انتخاب مدل‌های تجمیعی یا شبکه‌های عصبی، نقش حیاتی در افزایش دقت و پایداری مدل‌های پیش‌بینی دارد.



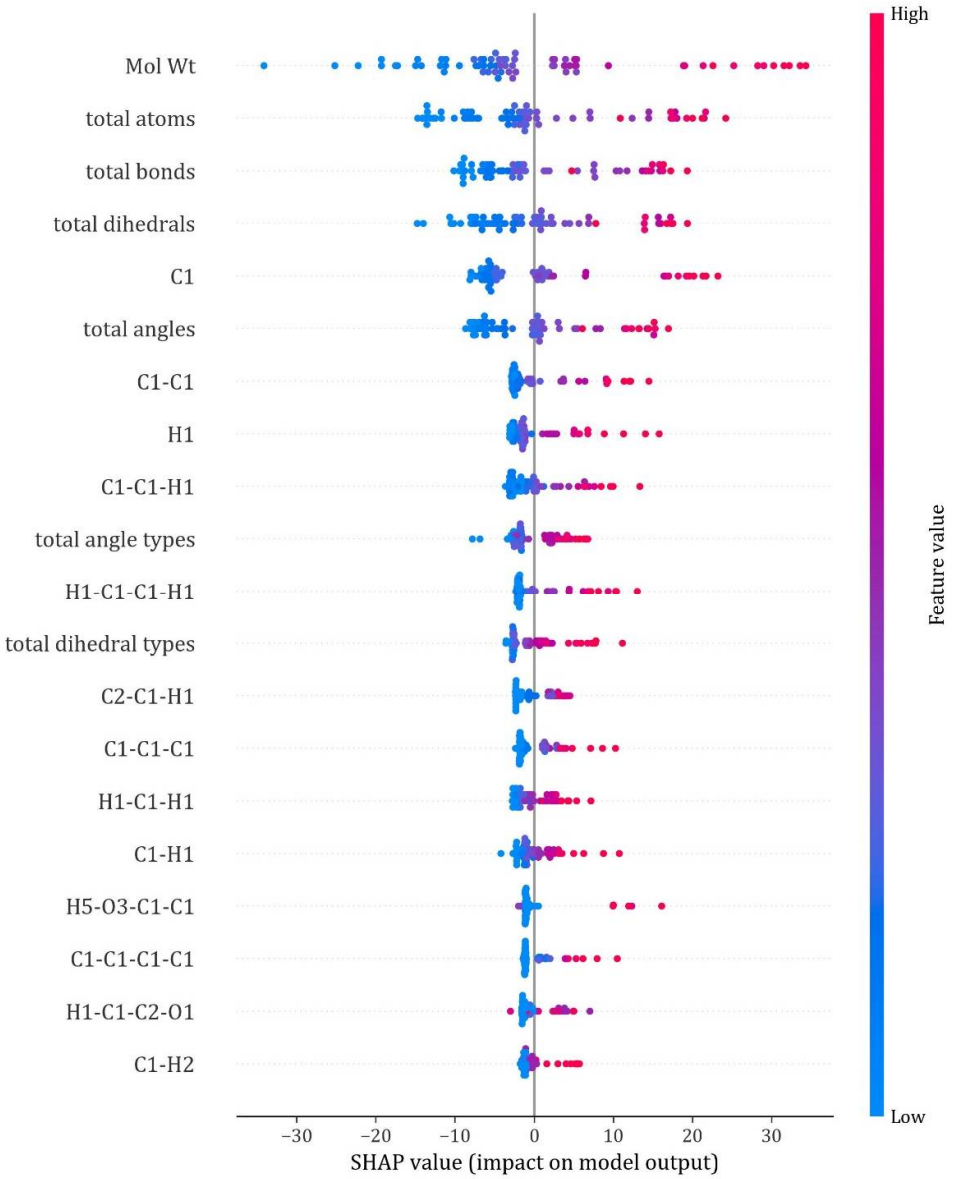
شکل 1- نمودارهای ضریب رگرسیون مدل‌های ارائه شده برای داده‌های آموزش



شکل 2- نمودارهای ضریب رگرسیون مدل‌های ارائه شده برای داده‌های تست

شکل 3 مربوط به نمودار SHAP بوده و ابزاری برای تفسیر خروجی مدل‌های یادگیری ماشین است که نشان می‌دهد هر ویژگی ورودی چقدر و در چه جهتی روی پیش‌بینی نهایی تأثیر می‌گذارد. این نمودار با استفاده از تئوری بازی‌ها و مقدار شپلی، سهم هر ویژگی را در تفاوت بین پیش‌بینی مدل و مقدار میانگین پیش‌بینی‌ها محاسبه می‌کند. مقادیر مثبت SHAP یعنی آن ویژگی باعث افزایش خروجی شده و مقادیر منفی یعنی باعث کاهش آن شده است. در نمودار خلاصه SHAP، ویژگی‌ها به ترتیب اهمیت از بالا به پایین مرتب می‌شوند و رنگ نقاط نشان‌دهنده مقدار واقعی آن ویژگی (بالا یا پایین) است. در شکل 3، جرم مولکولی بیشترین تاثیر را داشته اما همانطور که مشخص است اکثر ویژگی‌ها تأثیر کمی روی خروجی مدل دارند. مقادیر پایین ویژگی‌ها (آبی) خروجی را کاهش و مقادیر بالا (قرمز)

خروجی را افزایش می‌دهند. با اینکه دامنه تأثیر برخی نقاط پرت تا ±30 می‌رسد، به طور کلی مدل رفتار خطی و قابل پیش‌بینی با قدرت تفکیک پایین بین ویژگی‌ها دارد.



شکل 3- نمودارهای شاپ جهت تعیین پارامترهای تاثیرگذار بر روی مدل

3- نتیجه گیری و پیشنهادات

در این پژوهش، به منظور پیش‌بینی گرمای ویژه پلیمرها از مجموعه‌ای از الگوریتم‌های یادگیری ماشین شامل مدل‌های رگرسیون خطی ساده و منظم‌شده (ریج، لاسو، الاستیک‌نت)، مدل‌های غیرخطی مانند نزدیک‌ترین همسایگی K، درخت تصمیم و جنگل تصادفی و همچنین شبکه‌های عصبی مصنوعی با دو پیاده‌سازی (Scikit-learn و Keras) استفاده شد. هدف اصلی، بررسی دقت این مدل‌ها در یادگیری روابط پیچیده بین توصیفگرهای ساختاری و شیمیایی پلیمرها و گرمای ویژه آن‌ها بود. نتایج نشان داد که مدل‌های ساده مانند رگرسیون خطی، با وجود عملکرد

خوب روی داده‌های آموزشی، در تعمیم‌پذیری به داده‌های آزمون دچار ضعف شده و بیش‌برازش قابل‌توجهی داشتند. مدل ریج توانست تعادل خوبی بین دقت و جلوگیری از بیش‌برازش برقرار کند و عملکردی بهتر از لاسو و الاستیک‌نت داشت. مدل نزدیک‌ترین همسایگی K نیز با حفظ دقت در هر دو مجموعه داده، نشان داد که در پیش‌بینی‌های غیرخطی می‌تواند موفق عمل کند. از میان همه‌ی مدل‌ها، شبکه عصبی پیاده‌سازی‌شده با Scikit-learn بهترین عملکرد را داشت؛ به‌طوری که ضریب رگرسیون نزدیک به 1 در داده‌های آموزش و آزمون نشان‌دهنده توانایی بالا در مدل‌سازی رفتار پیچیده پلیمرها بود. در مقابل، مدل شبکه عصبی مبتنی بر Keras به دلیل تنظیمات غیربهینه و احتمالاً ساختار نامناسب شبکه، دچار بیش‌برازش شدید شده و در پیش‌بینی داده‌های آزمون ناتوان ظاهر شد. در نهایت، مدل جنگل تصادفی نیز با ایجاد تعادل بین دقت بالا و جلوگیری از بیش‌برازش، عملکرد بسیار خوبی ارائه داد و از پایداری مناسبی در برابر نویز برخوردار بود. با توجه به اهمیت دقت و قابلیت تعمیم در کاربردهای صنعتی و تحقیقاتی، توصیه می‌شود در مطالعات آتی از شبکه‌های عصبی با بهینه‌سازی دقیق‌تر ساختار و ابرپارامترها استفاده شود و همچنین روش‌های منظم‌سازی پیشرفته برای کاهش بیش‌برازش در نظر گرفته شود. علاوه بر این، گسترش مجموعه داده‌ها و بهبود کیفیت داده‌های ورودی می‌تواند موجب افزایش دقت پیش‌بینی‌ها گردد. در مجموع، این پژوهش نشان داد که بهره‌گیری از مدل‌های یادگیری ماشین می‌تواند ابزاری مؤثر و جایگزین مناسبی برای روش‌های پرهزینه و زمان‌بر آزمایشگاهی در پیش‌بینی خواص گرمایی پلیمرها باشد.

4- منابع

[1] A. Patil, A. Patel, R. Purohit, An overview of polymeric materials for automotive applications, *Materials Today: Proceedings*, 4 (2017) 3807-3815.
[2] X. Wei, Z. Wang, Z. Tian, T. Luo, Thermal transport in polymers: a review, *Journal of Heat Tr*ansfer ,143 (2021) 072101.
[3] E. Ghanbari, S.J. Picken, J.H. van Esch, Analysis of differential scanning calorimetry (DSC): determining the transition temperatures, and enthalpy and heat capacity changes in multicomponent systems by analytical model fitting: E .Ghanbari et al, *Journal of Thermal Analysis and Calorimetry*, 148 (2023) 12393-12409.
[4] R. Bhowmik, S. Sihn, R. Pachter, J.P. Vernon, Prediction of the specific heat of polymers from experimental data and machine learning methods, *Polymer*, 220 (2021) 123558.
[5] R.S. Bhoopal, P. Sharma, S. Kumar, A. Pandey, R. Beniwal, R. Singh, Prediction of effective thermal conductivity of polymer composites using an artificial neural network approach, *Special Topics & Reviews in Porous Media: An International Journal* , 3 (2012).
[6] R. Mortazavi, S. Mortazavi, A. Troncoso, Wrapper-based feature selection using regression trees to predict intrinsic viscosity of polymer, *Engineering with Computers*, 38 (2022) 2553-2565.
[7] C. Joo, H. Park, H. Kwon, J. Lim, E. Shin, H. Cho, J. Kim, Machine learning approach to predict physical properties of polypropylene composites: application of MLR, DNN, and random forest to industrial data, *Polymers*, 14 (2022) 3500.
[8] K. Ishikiriyama, Machine learning prediction of heat capacity of polymers as a function of temperature, *Polymer*, (2025) 129171.
[9] M.-X. Zhu, H.-G. Song, Q.-C. Yu, J.-M. Chen, H.-Y. Zhang, Machine-learning-driven discovery of polymers molecular structures with high thermal conductivity, *International Journal of Heat and Mass Transfer*, 162 (2020) 120381.
[10] G. Chen, R. Hassan, A. Baghban, Pioneering accurate hydrogen generation from ammonia using cutting-edge modeling techniques, *Journal of Industrial and Engineering Chemistry* ,(2026).

[11] M. Gheytanzadeh, F. Rajabhasani, A. Baghban, S. Habibzadeh, O. Abida, A. Esmaeili, M.T. Munir, Estimating hydrogen absorption energy on different metal hydrides using Gaussian process regression approach, *Scientific Reports*, 12 (2022) 21902.
[12] R. Hassan, A. Baghban, Beyond the surface: Quasi-SMILES machine learning approaches for precise estimation of organic sorption, *Materials Today Communications*, (2025) 114126.
[13] R. Hassan, A. Baghban, Predicting CO2 adsorption in KOH-activated biochar using advanced machine learning techniques, *Scientific reports*, 15 (2025) 24410.
[14] R. Hassan, A. Baghban, AI-based algorithms for estimating hydrochar properties in terms of biomass ultimate analysis, *Results in Engineering*, 26 (2025) 105141.
[15] R. Soleimani, R. Hassan, A. Baghban, Towards Cleaner Water by Leveraging AI for Optimizing Coagulation Processes in Microplastic Removal, *Journal of Environmental Chemical Engineering*, (2025) 120725.
[16] R. Soleimani, R. Hassan, Z. Behtouei, A. Baghban, Data-driven prediction of thermal conductivity of soil using advanced computational schemes, *International Communications in Heat and Mass Transfer*, 169 (2025) 109780.
[17] R. Soleimani, M.M. Kouchi, Z. Behtouei, Z. Ghasemi, A. Baghban, Exploring the solubility potential of anti-cancer and supportive agents in supercritical CO2 through advanced computational intelligence techniques, *Journal of CO2 Utilization*, 102 (2025) 103227.
[18] Q. Xie, A. Baghban, Data-Driven strategies for predicting metal ion adsorption Behaviour in clay mineral frameworks, *Minerals Engineering*, 246 (2026) 110377.