مقاله مروری

فناوری پلیمر ایران

فصلنامه علمی
سال نهم، شماره ۱، شماره پیاپی ۳۳، بهار ۱۴۰۲
Vol. 9, No. 1, Issue No. 33
Spring 2024, Quarterly
صفحه ۵۸-۴۹

Iran Polymer Technology; Research and Development

سکوی ساخت و تولید هوشمند پلیمرها: مهندسی ژنوم مواد پلیمری

زینب سادات حسینی*
تهران،گروه فرایندهای پلیمریزاسیون، دانشکده مهندسی شیمی، دانشگاه تربیت مدرس

چکیده ...

مواد پلیمری با کارایی عالی، پایه و اساس توسعه فناوری سطح بالا و ساخت و تولید پیشرفته است. اخیراً، مهندسی ژنوم مواد پلیمری(Polymeric Material Genome Engineering) (PMGE) به‌عنوان سکویی اساسی برای ساخت و تولید هوشمند مواد پلیمری مطرح شده است. PMGE رشته‌ای نوظهور است که اصول طرح ژنوم مواد را با علم پلیمر ترکیب می کند تا کشف و توسعه مواد پلیمری جدید را تسریع بخشد. مفهوم PMGE ایجاد پایگاه داده جامعی از خواص پلیمر است که از هر دو روش محاسباتی و تجربی به دست آمده است. می توان از این پایگاه داده برای آموزش مدل‌های یادگیری ماشینی استفاده کرد که قادر است خواص پلیمرهای جدید را پیش بینی کند. به طور کلی، PMGE نشان‌دهنده گامی مهم به سمت تولید هوشمند مواد پلیمری با پتانسیل ایجاد انقلاب در این زمینه همراه با امکان توسعه سریع‌تر و کارآمدتر مواد جدید است. با این حال، توسعه PMGE هنوز در ابتدای راه است و بسیاری از مسائل، باقی مانده که باید مورد توجه قرار گیرد. در این بررسی، مفاهیم بنیادی PMGE و خلاصه‌ای از تحقیقات و دستاوردهای اخیر ارائه می‌شود، سپس مهم‌ترین چالش‌ها به همراه چشم‌انداز آینده ترسیم می‌شود. به‌طورخاص، این مطالعه بر رویکردهای پیش‌بینی خواص، از جمله رویکرد پروکسی و یادگیری ماشین متمرکز است و کاربردهای بالقوه PMGE یعنی کامپوزیت‌های پیشرفته، مواد پلیمری مورد استفاده در سامانه‌های ارتباطی و ساخت مدارهای یکپارچه الکتریکی را مورد بحث قرار خواهد داد.

واژه‌های کلیدی:

پلیمر،
ژنوم مواد،
هوش مصنوعی،
یادگیری ماشین

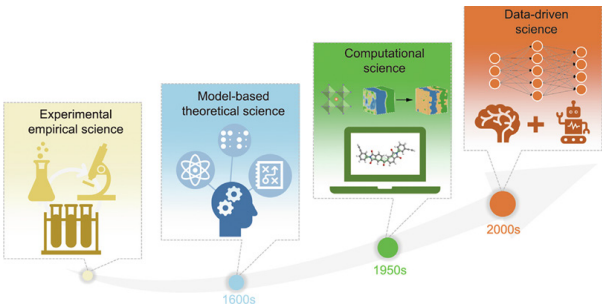
*پست الکترونیکی مسئول مکاتبات:
zeinabsadat_hosseini@modares.ac.ir

۱ مقدمه

در سال ۲۰۱۱، با ظهورمهندسی ژنوم مواد (Material Genome Engineering) (MGE) با هدف سرعت بخشیدن به کشف، طراحی و ساخت مواد، از طریق به‌کارگیری توان هم‌افزای تجربه، مباحث نظری و محاسبات در سامانه‌ای یکپارچه و پرتوان، جوامع علمی و مهندسی به چالش و تحرک کشیده شد. در این رویکرد، مجموعه‌ای وسیع از داده‌های مواد را می‌توان تولید، تجزیه و تحلیل کرد و به اشتراک گذاشت؛ بدین ترتیب محققان می‌توانند ویژگی‌های اصولی مواد را شناسایی و زمان عرضه مواد جدید را به‌طور قابل‌توجهی کوتاه کنند. «مهندسی ژنوم مواد» موجب طراحی و ساخت مواد جدیدی می‌شود که مشکلات را حل می‌کند و رفاه اجتماعی را روز به روز بهبود می‌بخشد، نظیر کشف فلزات که عصر صنعتی شدن از دستاوردهای آن است و قبل از آن غیرقابل تصور بود. به‌طور مشابه، سازندگان رایانه‌های مکانیکی نمی‌توانستند افزایش بیشتر قدرت محاسباتی را که با توسعه مواد نیمه‌رسانا برای ترانزیستورها امکان‌پذیر می‌شود، تصور کنند. کسانی که روی رایانه هدایت آپولو کار می‌کردند، نمی‌دانستند که بیش از نیمی از جمعیت زمین در سال ۲۰۱۸ دستگاه‌هایی در کف دستان خود خواهند داشت که هزار برابر قدرت محاسباتی بیشتری نسبت به رایانه‌هایی که برای هدایت پروازهای فضایی ساخته شده، دارد. بدین ترتیب، به تدریج، کشف مواد و نبوغ مهندسی، مرزهای جدیدی را برای پیشرفت فناوری باز کرد [۲، ۱].

مشابه با تحولات گذشته، تلاش برای طراحی و کشف مواد جدید از طریق تحقیقات علمی، تحولات اجتماعی آینده را دیکته خواهد کرد. حسگرهای زیستی انعطاف‌پذیر می‌توانند در داخل بدن کاشته شوند و زمانی که کارشان انجام شد، به‌طور بی‌ضرر تخریب شوند. موادی که الکتریسیته ساکن و نیروی ترموالکتریک حاصل از فعالیت‌های روزانه را جمع‌آوری می‌کنند، می‌توانند برای تأمین انرژی دستگاه‌های الکترونیکی شخصی به کار روند. چاپگرهای سه‌بعدی می‌توانند کاشتینه‌ی استخوان یا لنزهای تماسی متناسب با بیمار را هنگام مراجعه بیمار به مطب، چاپ کنند. مواد ابررسانای پیشرفته می‌توانند موجب توسعه فناوری‌های اطلاعات کوانتومی برای سامانه‌های ارتباطی و رمزنگاری پیشرفته شود. این پیشرفت‌های بالقوه بر اساس تصورات کنونی ما از امکانات دستکاری دنیایی است که می‌توان آن را به شدت با توسعه مواد جدید تغییر داد[۲،۱]. مواد با کارایی‌عالی، پایه و اساس توسعه فناوری‌های سطح بالا و ساخت و تولید پیشرفته است. تاکنون، علم مواد چهار چارچوب داشته است: چارچوب (Paradigm) علم تجربی آزمایشگاهی، چارچوب علم نظری مبتنی بر مدل، چارچوب علم محاسبات و چارچوب علم داده محور [۴،۳]، همان‌طور که در شکل ۱ نشان داده شده است.

اولین چارچوب نیاز به آزمون و خطا دارد که منجر به چرخه‌های تحقیقاتی طولانی برای کشف مواد می‌شود. در چارچوب دوم، قوانین علمی با جمع‌بندی های تجربی و ساخت مدل‌های فیزیکی کشف می‌شوند. در چارچوب سوم، حالات میکروسکوپی اتم‌ها یا مولکول‌ها توسط رایانه‌ها شبیه‌سازی می‌شود تا خواص ماکروسکوپی مواد به دست آید. با پیشرفت علم اطلاعات و هوش‌مصنوعی (Artificial Intelligence) (AI)، چارچوب چهارم در اوایل دهه ۲۰۰۰ پدیدار شد. تحقیقات مواد اکنون وارد عصر مبتنی بر داده (چهارمین چارچوب) شده است. این چارچوب رویکردی پژوهشی است که از الگوریتم‌ها برای تجزیه و تحلیل حجم زیادی از داده‌ها و یافتن قوانین بین آن‌ها استفاده می‌کند. برخلاف چارچوب دوم و سوم، چارچوب چهارم می‌تواند داده‌های ناشناخته را بر اساس داده‌های تجربی موجود استنباط و پیش‌بینی کند . ترکیب این چهار چارچوب باعث ظهور انواع مواد پیشرفته شده است. چارچوب چهارم بر اساس اهداف منتج از داده‌ها موجب تسریع تحقیقات مواد و کاهش هزینه از طریق سنتز مجازی، پیش‌بینی خواص و غربالگری می‌شود و در حال تبدیل شدن به چارچوبی انقلابی است [۵]–[۷]. علم کلان‌داده‌ها یکی از مبانی بین‌رشته‌ای گروه‌هایی نظیر بیوانفورماتیک، شیمی انفورماتیک و مواد انفورماتیک است. به‌عنوان دستاورد برجسته‌ای در بیوانفورماتیک، آلفافولد۲ (AlphaFold2) تا حدی در پیش‌بینی توالی‌ها و ساختارهای سه‌بعدی پروتئین‌ها از متخصصان دیگر پیشی گرفته است. در شیمی‌انفورماتیک، استفاده از هوش‌مصنوعی برای کشف داروهای جدید، روشی کارآمد و شناخته شده است. برخلاف بیوانفورماتیک و شیمی انفورماتیک، که در حال حاضر



شکل ۱ توسعه چهار چارچوب پژوهش در مواد: چارچوب های علم تجربی آزمایشگاهی، چارچوب علم نظری مبتنی بر مدل، چارچوب علم محاسبات و چارچوب علم داده محور.

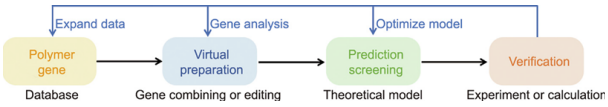
به خوبی تثبیت شده‌اند، مواد انفورماتیک هنوز رشته‌ای به سرعت در حال رشد است. به عنوان پیشتاز، مهندسی ژنوم مواد (MGE) در حال تبدیل شدن به سکویی اساسی برای ساخت و تولید مواد هوشمند است. با توسعه MGE، طراحی سفارشی و تهیه بهینه مواد، مزیت‌های بالقوه‌ای را ایجاد کرده است [۸].

۲ توسعه مهندسی ژنوم مواد پلیمری

ژنوم شیوه‌نامه‌ای ارثی برای تولید، پیشبرد و حفظ موجود زنده است. به بیانی دیگر در هر سلول تعداد زیادی خط مشی‌های کدگذاری شده وجود دارد. این دستورالعمل‌ها، برای هدایت کلیه فعالیت‌های سلول و تولید مواد ضروری لازم هستند. به هر مجموعه کامل از این خط مشی‌ها، ژنوم می‌گویند [۹]. با استفاده از این تعریف از ژنوم، در مهندسی ژنوم مواد پلیمری(PMGE) با استفاده از محاسبات نظری و پایگاه‌های داده به پیش‌بینی و غربالگری و راستی‌آزمایی برای دستیابی به کدگذاری خاص برای هر ماده و طراحی منطقی، ساخت هوشمند و تسریع در طراحی و توسعه مواد پلیمری پرداخته می‌شود (شکل ۲) [۱۱، ۱۰].

۱ تعریف «ژن» پلیمر و طراحی «پلیمرهای مجازی». بر اساس قوانین خاصی از تجزیه و تحلیل داده‌های شیمیایی موجود و تجربه کارشناسان، عوامل مرتبط با خواص مواد، مانند مواد شیمیایی گروه‌ها و عناصری که پلیمرها را تشکیل می‌دهند، به‌اصطلاح به‌عنوان ژن پلیمرها تعریف می‌شوند. سپس، مجموعه‌ای از «پلیمرهای مجازی» را می‌توان با ترکیب یا ویرایش ژن پلیمری طراحی کرد (به‌عنوان مثال، تنظیم ترکیب زنجیره‌ای پلیمرها) [۸]..

۲ پیش‌بینی و غربالگری ویژگی‌های پلیمر با کارایی بالا. رابطه کمی ساختار–خواص(Quantitative Structure Property Relationship) (QSPR) پلیمرها بر اساس داده‌های تجربی یا شبیه‌سازی برای پیش‌بینی خواص «پلیمرهای مجازی طراحی شده» به دست می‌آید. در مرحله بعد، غربالگری برای به‌دست آوردن پلیمرهای جدید امیدوارکننده با توجه به الزامات



شکل ۲مفهوم و مراحل PMGE. بر اساس پایگاه داده، (۱) تعریف ژن های پلیمری و پلیمرهای مجازی طراحی شده (۲) پیش‌بینی بازده بالا و غربالگری خواص پلیمر از طریق محاسبات نظری یا آزمایش‌هایی با توان بالا (۳) بهینه‌سازی و (۴) تأیید نتایج غربال‌شده از طریق آزمایش‌ها.

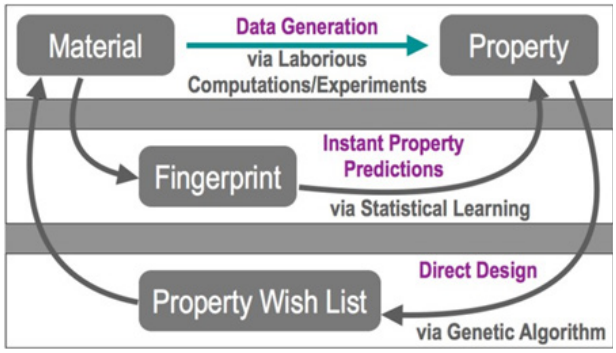
عملکردی مورد نظر انجام می‌شود [۸].

۳ و ۴ بهینه‌سازی و تأیید. مواد پلیمری غربال‌شده برای تأیید قابلیت اطمینان نتایج غربالگری و بهینه‌سازی مدل پیش‌بینی، سنتز می‌شوند. علاوه بر این، محاسبات نظری با دقت بالا می‌تواند برای تأیید نتایج غربالگری استفاده شود [۸].

یکی از انواع راهبردهای پیش‌بینی شامل یافتن ویژگی‌های کلیدی است که می‌تواند خواص مواد را از طریق داده‌کاوی ارزیابی کند. ویژگی کلیدی قابل‌محاسبه‌ای به‌عنوان نماینده استخراج می‌شود. خواص ماکروسکوپی که به‌دست‌آوردن دقیق آن‌ها از محاسبات نظری دشوار است به متغیرهای نماینده قابل محاسبه تبدیل می‌شوند. سپس، مواد پلیمری را می‌توان با مقایسه متغیرهای نماینده مربوط غربال کرد. به‌عنوان مثال، شارما و همکاران [۱۲]. از شکاف انرژی ساختارهای پلیمری که به‌راحتی می‌توان با استفاده از نظریه تابع چگالی (Density Fourier Transform) (DFT) محاسبه کرد، استفاده کرد تا ولتاژ شکست و تلفات دی الکتریک را نشان دهد. آن ها با استفاده از ثابت دی‌الکتریک و شکاف انرژی به‌عنوان معیار غربالگری، مجموعه‌ای از دی الکتریک‌های پلیمری تمام آلی را به دست آوردند. راهبرد متغیر نماینده گاهی اوقات تجربی است، برای مثال، ژو و همکاران [۱۳]. داده‌های تجربی و محاسباتی موجود بیش از ۴۰۰ پلیمر را از پایگاه داده پلی اینفو (PolyInfo) تجزیه‌وتحلیل کرد. آن‌ها دریافتند که دمای تجزیه ۵٪ (T_{d5}) پلیمرها به انرژی تفکیک پیوند (Bond Dissociation Energy) (BDE) ضعیف‌ترین پیوند در ساختار پلیمر بستگی دارد، جایی که مقدار همبستگی پیرسون نزدیک به ۰/۷ است. بنابراین، BDE را می‌توان ویژگی‌ای کلیدی برای ارزیابی پایداری حرارتی مواد پلیمری در نظر گرفت. سپس، آن‌ها از ژنوم مواد پلیمری برای از بین بردن تضاد بین پایداری حرارتی بالا و انرژی پخت کم رزین‌ها استفاده کردند. شکاف انرژی محاسبه شده توسط DFT به‌عنوان نماینده قابل پردازش در نظر گرفته شد. با استفاده از مدل‌های پیش‌بینی ویژگی‌های کلیدی پیشنهادی، غربالگری دو مرحله‌ای برای به دست آوردن ساختار بهینه پلی (سیلان آریل استیلن) (Poly (Silane Arylacetylene)) (PSA) انجام شد. سپس، ساختار PSA حاوی ۲،۷–دی اتینیل نفتالین غربال شد. تأیید تجربی نشان داد که رزین PSA جدید دمای تجزیه حرارتی ۶۵۵ درجه سانتی‌گراد و آنتالپی پخت $J \cdot g^{-1}$ ۲۴۱/۹ را نشان می‌دهد که بیانگر خواص جامع عالی است. علاوه بر این، نسبت مدول توده به مدول برشی (K/G) می‌تواند برای نشان دادن چقرمگی پلیمرها استفاده شود. پس از محاسبه و غربالگری نماینده K/G برای چقرمگی و نماینده BDE برای پایداری

حرارتی، گائو و همکاران [۱۴]. پلی‌آمید استیلن اختتام‌یافته جدیدی (Acetylene Terminated Polyimide) (ATPI) را معرفی کرد که می‌تواند برای افزایش چقرمگی رزین‌های PSA به روش اختلاط استفاده شود. برای رزین کوپلیمر شده ATPI و PSA، چقرمگی به‌طور قابل‌توجهی بهبود می‌یابد، در عین حال مقاومت حرارتی حفظ می‌شود. همان‌طور که در بالا توضیح داده شد، طراحی و غربال کردن مواد پلیمری با استفاده از مدل پیش‌بینی متغیرهای نماینده، مؤثر و قابل‌اعتماد است. کلید پیش‌بینی نماینده در استخراج روابط اساسی بین ویژگی‌های هدف و عوامل فیزیکی میکروسکوپی یا ماکروسکوپی نهفته است. یادگیری ماشین (Machine Learning) (ML) می‌تواند قوانین اساسی را از داده‌ها استخراج کند و داده‌های ناشناخته را پیش‌بینی، استنتاج یا طبقه‌بندی کند. این راهبردی دیگر برای دستیابی به پیش‌بینی با توان بالا و غربالگری در PMGE است [۱۶، ۱۵]. مدل ML آموزش‌دیده بر روی داده‌های تجربی قابل‌اعتماد می‌تواند به‌طور مستقیم ویژگی مواد را پیش‌بینی کند. به‌عنوان مثال، بر اساس پایگاه‌های داده مانند پلیمر ژنوم (Polymer Genome) و پاپ کم (PubChem)(شکل ۳).

زانگ و همکاران [۱۶] از روشی چندلایه برای ایجاد مدل‌های پیش‌بینی ML برای QSPR بین خواص هدف (یعنی دمای تجزیه حرارتی و گرانروی) و ساختارهای پلیمری استفاده کرد (شکل ۴). با ترکیب ژن، آن‌ها ۳۶۸ رزین انتخابی برای غربالگری به دست آوردند. با استفاده از دو مدل ML، خواص رزین‌های انتخابی پیش‌بینی شد و با توان عملیاتی بالا غربال شد. سپس مجموعه‌ای از رزین‌ها با قابلیت پردازش بهینه و مقاومت حرارتی بالا به‌دست آمد. راستی‌آزمایی تجربی نشان داد که رزین غربال‌شده (PSNP-MV) دارای خواص جامع عالی در پردازش و مقاومت در برابر حرارت است. هنگامی که داده‌های تجربی محدود یا با کیفیت پایین هستند، از محاسبات نظری یا

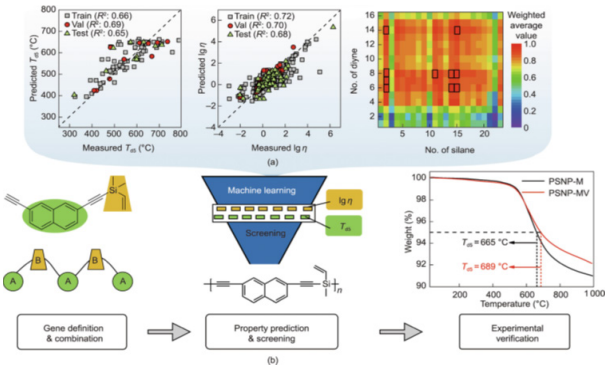


شکل ۳ سه مرحله تعیین ویژگی مواد از طریق پایگاه‌های داده: مرحله تولید داده، مرحله پیش‌بینی ویژگی فوری و مرحله طراحی مستقیم.

داده‌های شبیه‌سازی می‌توان برای آموزش مستقیم مدل‌های ML استفاده کرد. مدل‌های به‌دست‌آمده می‌توانند پیش‌بینی‌های قابل اعتمادی را ارائه دهند.

در صورتی که برخی از داده‌ها از محاسبات، شبیه‌سازی یا پایگاه داده، قابلیت اطمینان پایینی داشته باشند، بهبود کیفیت داده‌ها با مدلی جایگزین با چند عامل تثبیت‌کننده راهبردی مؤثر است. انحراف بین داده‌های با قابلیت اطمینان کم (مانند داده‌های شبیه‌سازی) و داده‌های با قابلیت اطمینان بالا (مانند تجربی) را می‌توان آموزش داد و به مدلی مبتنی بر ML اجازه داد تا تفاوت‌های آن‌ها را ارزیابی کند و سپس کیفیت داده‌ها را بهبود بخشد. با استفاده از راهبردهای امیدوارکننده فوق، مشکل کمبود داده‌های تجربی حل می‌شود و می‌توان به طراحی و غربالگری پلیمرها پرداخت. محاسبات نظری همچنین می‌تواند برای تخمین خواص پلیمر و غربالگری خواص هدف مورد استفاده قرار گیرد، اما گاهی اوقات می‌تواند زمان‌بر باشد. مدل‌های ML می‌توانند بر محدودیت‌های محاسبات نظری، به‌ویژه هزینه‌های محاسباتی زمان‌بر غلبه کنند. هنگامی‌که تعداد ژن‌های پلیمر افزایش می‌یابد، فضای ساختار پلیمر به صورت تصاعدی افزایش می‌یابد و محاسبه خواص پلیمر غیرعملی است. مدل ML می‌تواند در مدت زمان کوتاهی به پیش‌بینی این خواص پلیمری دست یابد. به‌طورکلی، استفاده از مدل‌های ML مزایای دقت پیش‌بینی بالا، چرخه توسعه کوتاه و کاربرد وسیع را به همراه دارد. این مزایا به خوبی با الزامات طراحی مواد و غربالگری در PMGE مطابقت دارد [۱۷].

۳ مدل‌های الگوریتمی در فناوری‌های ژنوم مواد

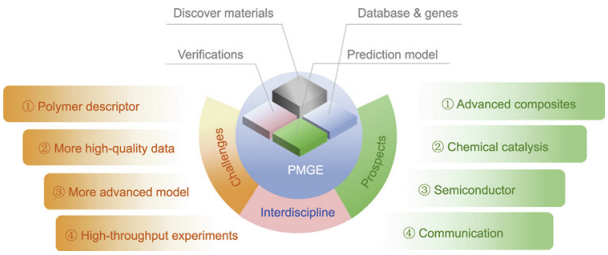


شکل ۴ پیش‌بینی ML و غربالگری رزین با کارایی بالا. (الف) مدل‌های ML از دما و گرانروی تجزیه حرارتی و گراف حرارتی خواص برای ۳۶۸ رزین انتخابی(ب) طراحی رزین جدید PSNP-MV با خواص عالی به کمک رویکرد ژنوم مواد تقویت‌شده با ML

به‌دلیل حجم زیاد وظایف محاسباتی و پردازش خودکار نتایج محاسباتی، استفاده کامل از داده‌های عظیم و پیچیده و آشکارسازی ارتباطات بالقوه بین عوامل مربوط و عملکرد مواد بسیار چالش‌برانگیز است. ابزارهای محاسباتی، از جمله نرم‌افزارهای مختلف و مدل‌های الگوریتم، می‌توانند برای حل چالش‌های موجود در پردازش مقادیر زیادی داده به کار گرفته شوند.
رگرسیون، طبقه‌بندی یا خوشه‌بندی معمولاً در پردازش داده‌ها به کار می‌رود. رگرسیون عمدتاً با اعداد پیوسته یا واقعی سروکار دارند، در حالی‌که در طبقه‌بندی یا خوشه‌بندی با وارد کردن داده‌ها با یا بدون برچسب منجر به نتیجه‌ای مجزا می‌شوند. هدف هر دو رگرسیون و طبقه‌بندی، کشف رابطه بین نقاط داده از طریق یک مدل پیش‌بینی و دستیابی به نتیجه پیش‌بینی‌شده قابل اعتماد است. برای این منظور، برخی از الگوریتم‌ها، مانند ماشین‌های بردار پشتیبان (Support Vector Machines) (SVM)، جنگل‌های تصادفی (Random Forests)، شبکه‌های عصبی بیزی (Bayesian Neural Networks) و غیره، می‌توانند هم برای وظایف رگرسیون و هم برای طبقه‌بندی در مورد داده‌های پلیمری و مواد زیست پزشکی استفاده شوند (جدول ۱).

۴ کاربردهای مهندسی

همانطور که در شکل ۵ نشان داده شده است، PMGE می‌تواند توسعه مواد پلیمری را در کاربردهای مهندسی مختلف سرعت بخشد، به ویژه زمانی که دو یا چند ویژگی با یکدیگر در تضاد باشند.
• کامپوزیت‌های با ماتریس رزین پیشرفته. علاوه بر رزین‌های پلیمری، PMGE برای طراحی ساختاری و بهبود خواص الیاف پلیمری با استحکام بالا و مدول بالا استفاده می‌شود. PMGE همچنین می‌تواند برای تنظیم پیوند فصل مشترک بین رزین و الیاف و بهینه‌سازی فرایندپذیری مواد کامپوزیتی استفاده شود. علاوه بر این، مدل‌های ML را می‌توان بر اساس داده‌های شبیه‌سازی اجزای محدود و آزمایش‌های کامپوزیت‌ها، آموزش داد. با یک مدل ML آموزش‌دیده، می‌توان عملکرد مواد کامپوزیت را به‌سرعت پیش‌بینی کرد و امکان تحقق طراحی منطقی کامپوزیت‌های پیشرفته را فراهم کرد[۸،۳۹].
•مهندسی شیمی و کاتالیزور. طراحی منطقی و غربال کاتالیزورهای مختلف، از جمله مواد کاتالیزوری متخلخل و کاتالیزورهای پلیمری‌شدن، می‌تواند از طریق راهبرد طراحی مواد (ML بهبودیافته) تسریع شود. ریزساختار، عملکرد ماکروسکوپی و کارایی صنعتی پلی‌اولفین، نوع کاتالیزورها را تعیین می‌کنند. بنابراین، طراحی ساختاری کاتالیزورها کلید پیشرفت صنعت پلی‌الفین است. به‌عنوان مثال، طراحی منطقی مکان‌های فعال کاتالیزور زیگلر ناتا و پیش‌بینی انتخاب‌پذیری پیکربندی کاتالیزور متالوسن برای پلیمری‌شدن پروپن هنوز چالش برانگیز است. رویکرد ML مبتنی بر داده می‌تواند راهبردی امیدوارکننده برای کشف و طراحی کاتالیزور پلیمری ارائه دهد [۴۰].
• مواد نیمه‌هادی آلی پلیمری. چنین سامانه‌های پلیمری به تحرک الکترون بالا، بازده تابندگی بالا، ویژگی‌های چرخش بالا، رسانایی بالا و غیره نیاز دارند. به‌دست‌آوردن مواد پلیمری با خواص چندگانه با استفاده از روش سنتی آزمون و خطا دشوار است. طراحی پلیمرهای کونژوگه با استفاده از PMGE می‌تواند تحقیقات روی مواد نیمه‌رسانای آلی پلیمری با خواص بهبودیافته را تسریع کند [۸،۳۹].
• ارتباطات. پلیمرهای مورد استفاده در زمینه فناوری ارتباطات با بسامد بالا به‌طور همزمان به خواص مکانیکی افزایش‌یافته، مقاومت در برابر حرارت و خواص الکترومغناطیسی نیاز دارند. به عنوان مثال، مواد پلیمری مورد استفاده در تجهیزات ارتباطی نسل ششم (G_6) باید دارای ثابت دی‌الکتریک نسبتاً کم و تلفات دی‌الکتریک کم باشند. علاوه بر این، پلیمرهای با کارایی بالا که در بسته‌بندی تراشه‌ها استفاده می‌شوند باید دارای مقاومت حرارتی بالا، ضریب انبساط حرارتی پایین، سختی بالا، چقرمگی بالا، عایق الکتریکی بالا و ثابت دی‌الکتریک پایین باشند. بنابراین، کاربردهای مهندسی فوق نیازمند کشف مواد پلیمری پیشرفته با خواص بهبودیافته است و PMGE بدون شک بهترین انتخاب است. از طریق پیش‌بینی و غربالگری با توان بالا، PMGE می‌تواند به هدف کشف مواد پلیمری با کارایی عالی برسد [۳۹، ۸].



شکل ۵ چالش‌ها و چشم اندازهای PMGE مسائل مربوط به پایگاه‌های اطلاعاتی و ژن‌های پلیمری، مدل‌های پیش‌بینی ویژگی و راستی‌آزمایی‌ها

۵ چالش‌ها وچشم‌اندازهای PMGE

مهندسی ژنوم مواد پلیمری هنوز در مراحل اولیه خود است و مسائل زیادی وجود دارد که باید مورد توجه قرار گیرند که در ادامه به برخی از آن‌ها اشاره شده است.

جدول ۱ معرفی برخی از الگوریتم‌های رایج.

	وظیفه	الگوریتم	ایده الگوریتم	ویژگی	کاربرد
۱	رگرسیون	شبکه عصبی انتشار برگشتی[1] (BPNN)	یک شبکه پیشران چند لایه آموزش داده شده بر اساس الگوریتم انتشار برگشت خطا، که به طور مداوم وزن‌ها و آستانه‌های شبکه را از طریق انتشار برگشتی تنظیم می‌کند تا مجموع مربعات خطاهای شبکه را به حداقل برساند.	دارای قابلیت‌های نقشه‌برداری غیرخطی، قابلیت‌های خودآموزی و خودسازگاری قوی، قابلیت تعمیم بالا و تحمل خطا است. اما سرعت همگرایی آهسته دارد و دارای مشکلات کمینه سازی محلی است. وابستگی به نمونه بالاست.	[۱۸]-[۲۰]
۲	رگرسیون و طبقه‌بندی	شبکه عصبی تابع شعاعی[2] (RBFNN)	RBF به‌عنوان تابع فعال‌سازی لایه نورون‌های پنهان استفاده می‌شود و لایه خروجی ترکیبی خطی از خروجی لایه نورون های پنهان است.	ساختار ساده است، آموزش مختصر است، سرعت همگرایی یادگیری سریع است، می‌تواند هر تابع غیرخطی را تقریب بزند و بر مشکل کمینه سازی محلی غلبه کند. توانایی تقریب، توانایی طبقه بندی و سرعت یادگیری بهتر از BPNN است. اما به نورون های بیشتری نیاز دارد	[۲۱]-[۲۳]
۳	طبقه‌بندی	شبکه عصبی احتمالی[3] (PNN)	این شاخه ای از RBFNN است که تخمین تابع چگالی و نظریه تصمیم بیزی را بر اساس شبکه RBF ترکیب می‌کند.	تحمل خطا خوب است، نتیجه طبقه‌بندی به انتخاب تابع شعاعی حساس نیست، تعداد نورون‌ها در هر لایه ثابت است و در هنگام تغییر نمونه نیازی به آموزش مجدد نیست. اما هر نمونه باید محاسبه و ذخیره شود.	[۲۱، ۲۴]
۴	رگرسیون و طبقه‌بندی	ماشین بردار پشتیبانی و رگرسیون بردار پشتیبان[4] (SVM&SVR)	حداکثر خط جدایی یا ابر صفحه را برای طبقه‌بندی نمونه ایجاد کند و تعادلی بین دقت یادگیری مدل و توانایی یادگیری برای به دست آوردن بهترین توانایی ارتقا را پیدا می‌کند	می‌تواند مسائل غیرخطی را با دقت بالا و توانایی تعمیم خوب حل کند. اجرای نمونه‌های آموزشی در مقیاس بزرگ دشوار است و حل مسائل با چند طبقه‌بندی دشوار است.	[۲۰، ۲۵]-[۲۷]
	رگرسیون و طبقه‌بندی	جنگل تصادفی	چندین درخت تصمیم گیری را ایجاد می‌کند. درختان مختلف نتایج پیش‌بینی متفاوتی برای یک نمونه دارند. با ترکیب این نتایج، نتیجه نهایی میانگین تمام نتایج درخت تصمیم گیری است.	برای مجموعه داده های نامتعادل، خطاها را می‌توان متعادل کرد، داده‌های نمونه‌های بزرگ را با دقت خوب می‌توان پردازش کرد. اما در مواجهه با نمونه‌های کوچک، نتایج طبقه‌بندی خوبی ممکن است به دست نیاید. در برخی از مشکلات رگرسیون بیزی را به راحتی می‌توان استفاده کرد.	[۲۸]-[۳۱]
	خوشه‌بندی	K- خوشه‌بندی (K-means Clustering)	چندین درخت تصمیم گیری را به‌صورت تصادفی ایجاد می‌کند. درختان مختلف نتایج پیش‌بینی متفاوتی برای یک نمونه دارد، از فاصله به‌عنوان شاخص ارزیابی شباهت برای خوشه‌بندی نمونه‌های با شباهت زیاد در یک خوشه استفاده می‌کند.	الگوریتم ساده و سریع با قابلیت تفسیر قوی، اثر خوشه‌بندی خوب، مناسب برای تعیین عوامل دشوار، حساس به نویز و نقاط غیرعادی، اثر خوشه‌بندی ضعیف بر روی داده‌های به شدت نامتعادل دارد و پردازش حجم نمونه‌های بزرگ به زمان طولانی نیاز دارد.	[۳۲]-[۳۵]
	رگرسیون و طبقه‌بندی	شبکه عصبی پیچشی[5] (CNN)	نمایش چند لایه هدف با استفاده از پیچش و ساختار شبکه چندلایه. انتظار می‌رود که اطلاعات انتزاعی موجود در داده‌ها را بتوان از طریق ویژگی‌های سطح بالای چندلایه بیان کرد تا استحکام ویژگی بهتری به دست آید.	اتصال محلی و اشتراک وزن تا حد زیادی تعداد عوامل را کاهش می دهد، عدم نیاز به انتخاب دستی ویژگی‌ها، بدون فرایند پیش‌پردازش پیچیده در هنگام پردازش داده‌های تصویری، اما هنگام تنظیم عوامل، حجم نمونه بزرگ مورد نیاز است	

[1]Back Propagation Neural Network, [2] Radial Basis Function Neural Network, [3] Probabilistic Neural Network, [4] Support Vector Machine & Support Vector Regression, [5] Convolutional Neural Network

۵–۱ تعریف ژن و توضیح ساختار مولکولی

ساختارهای زنجیره‌ای منحصربه‌فرد و ساختارهای چندمقیاسی پیچیده پلیمرها، چالش‌هایی را برای تعریف ژن و توصیف ساختاری پلیمرها ایجاد می‌کند. توسعه روش‌های پیشرفته‌تری برای توصیف ویژگی‌های ساختاری پلیمرها ضروری است. برای تعریف ژن پلیمری، برای متعادل‌کردن انعطاف‌پذیری طراحی ساختار با قابلیت دسترسی سنتز در آزمایش‌ها، ژن‌های پلیمری را می‌توان با توجه به مسیرهای ترکیبی سامانه‌های پلیمری هدف تعریف کرد. روش‌های موجود مانند بیگ اسمایلز (BigSMILES)، نمایش نمودار، اثر انگشت مولکولی و غیره را می‌تواند بررسی کند. علاوه بر این، رویکردهای جدیدی از انفورماتیک یا ریاضیات را می‌توان به آن معرفی کرد. در BigSMILES، قطعات پلیمری با فهرستی از واحدهای تکرارشونده نشان داده می‌شوند که BigSMILES را به انتخابی عالی برای نمایه‌سازی شناسه‌ها در سامانه پایگاه داده‌های پلیمری تبدیل کرده است. علاوه بر این، ژن‌های پلیمری باید به‌طور منظم تجزیه‌وتحلیل، طبقه‌بندی و برچسب‌گذاری شوند. قوانین داده‌کاوی با تجربه کارشناسان باید برای بهبود دقت و منطقی بودن تعریف ژن پلیمری و توصیف ساختار ترکیب شوند [۱۰،۴۱].

همچنین، ویژگی‌های چند مقیاسی باید در توصیف ساختار پلیمرها در نظر گرفته شود. به عنوان مثال، اطلاعات زنجیره و حالت توده‌ای (به‌عنوان مثال، ساختار بلوری و ساختار اتصال عرضی شده و پخت‌شده) را می‌توان از محاسبات نظری، شبیه‌سازی‌ها و آزمایش‌ها به‌دست آورد. اخیراً، هو و همکاران [۴۲]، از توصیفگرهای چگالی اتصال عرضی برای پیش‌بینی عملکرد رزین‌های اپوکسی پخت شده استفاده کرد. توزیع و پراکندگی پلیمری بر ساختارهای پلیمرها تأثیر می‌گذارد و باعث ایجاد تغییراتی در خواص پلیمر می‌شود. داده‌های پراکنش را می‌توان شناسایی، برچسب‌گذاری کرد و در پایگاه داده پلیمری گنجاند و سپس به‌عنوان یکی از ورودی‌ها استفاده کرد.

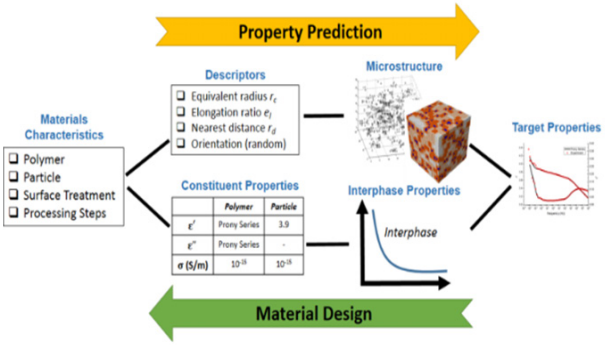
۵–۲ مدل پیش‌بینی نماینده برای ویژگی‌های کلیدی

یافتن یا ایجاد ویژگی‌های کلیدی بیشتری که نمایانگر ویژگی‌های پلیمر هستند، مانند مقاومت در برابر حلال، مقاومت در برابر سایش، مقاومت در برابر ضربه و خاصیت اتصال فصل مشترک ضروری است. علاوه بر این، برای دستیابی به پیش‌بینی سریع خواص پلیمر و غربالگری چندمرحله‌ای، باید روش‌های محاسبه سریع‌تر نماینده‌ها مانند روش اتصال مولکولی و روش مشارکت گروهی ایجاد شود [۸].

۵–۳ ML پیش‌بینی خواص پلیمر

فناوری‌های ژنوم مواد، الگوی سنتی تحقیق و توسعه مواد را تغییر داده است. با این حال، برای پاسخگویی به نیازها و چالش‌های رو به رشد در تحقیق و توسعه مواد پلیمری، تلاش‌های زیادی لازم است. ترسیم رابطه بین اجزا، ساختارها و خواص مواد پلیمری پیچیده و چالش‌برانگیز است. علاوه بر این، توانایی تعمیم مدل‌های پیش‌بینی خواص پلیمر قوی نیست و رابطه ساختار–ویژگی چند مقیاسی را نمی‌توان دقیقاً توصیف کرد. همه این مسائل کاربردهای پیش‌بینی ML را در PMGE محدود می‌کنند. این چالش عمدتاً ناشی از دشواری در دستیابی به داده‌های کافی از مواد پلیمری است که عموماً مجموعه داده‌ها نمونه کوچکی هستند. در این راستا، یا توسعه ابزارهای تجربی با توان عملیاتی بالا یا افزایش داده‌ها بسیار مطلوب است. برای رسیدگی به این چالش‌ها، داده های پایگاه‌های باز را می‌توان از طریق پردازش زبان مورد بهره‌برداری و استخراج قرار داد. با باز بودن و اشتراک‌گذاری سکوی PMGE، محققان بیشتری به‌طور فعال داده‌ها را وارد می‌کنند. علاوه بر این، محققان باید به استفاده از داده‌های تجربی با کیفیت پایین توجه کنند. به‌طور خاص، تمام داده‌ها باید استاندارد شوند تا کیفیت داده‌ها بهبود یابد. همچنین، ابزارهای مرتبط برای استخراج ویژگی‌های تصویر را می‌توان در تحلیل تصویر از مواد پلیمری برای کمک به استخراج اطلاعات تصویر استفاده کرد (شکل ۶) [۲،۴۳].

الگوریتم‌های پیشرفته می‌توانند برای توسعه راهبردهای ML برای حل مشکل مقادیر کم داده، مانند یادگیری انتقال، یادگیری



شکل ۶ نمونه‌هایی از استفاده از داده‌های پایگاه‌های باز برای پیش‌بینی و طراحی. فرایند بالا: پیش‌بینی خواص مواد از انتخاب مواد تا خواص دی‌الکتریک کامپوزیت شبیه‌سازی‌شده. فرایند پایین: طراحی از عملکرد مواد هدف تا اجزای ضروری و شیمی. با استفاده از الگوریتم بهینه‌سازی و فرا مدل‌سازی، عملکرد هدف با ریزساختار و فاز میانی نماینده ترسیم می‌شود که می‌تواند اجزای لازم و پردازش را با استفاده از مدل‌های داده‌کاوی پشتیبانی کند.

تحت نظارت و یادگیری فعال مورد استفاده قرار گیرند. معرفی الگوریتم و استفاده از اطلاعات میکرو یا نانو ساختاری با توصیف‌کننده‌های ساختار مولکولی برای ایجاد مدل‌های پیش‌بینی ML با معانی فیزیکی امیدوارکننده است. به‌عنوان مثال، سازوکار وابستگی بسامد برای ویژگی دی‌الکتریک پلیمری را می‌توان با توضیحی ساختاری برای آموزش مدل ML معرفی کرد. این می‌تواند در ایجاد روابط چند مقیاسی دقیق ساختار و خواص سودمند باشد [۴۴].

۵-۴ آزمایش های با توان بالا

سامانه آزمایشی برای سنتز پلیمرهای با کارایی بالا با هدف غربالگری سریع پلیمرها، گسترش پایگاه‌های داده و بهینه‌سازی مدل‌های پیش‌بینی باید ایجاد شود. تحقیقات بین‌رشته‌ای رویکردی موثر برای این موضوع است که شامل روش‌های علمی و فناوری مختلف مانند اطلاعات، کنترل سامانه و فناوری میکروسیال است. همچنین، دسترسی سنتزی پلیمرها و خواص پردازش مناسب برای تولید در مقیاس بزرگ باید در PMGE در نظر گرفته شود. فراتر از پیش‌بینی و غربالگری، لازم است نیازهای کاربردهای مهندسی را از هم جدا کرد و سپس راهبرد طراحی معکوس PMGE ایجاد کرد. طراحی معکوس ساختارهای پلیمری اهمیت PMGE را غنی می‌کند و طراحی منطقی و ساخت هوشمند پلیمرها را تحقق می‌بخشد [۲۸].

۵-۵ ایجاد و مدیریت پایگاه داده

به‌عنوان پایه‌ای مهم در تحقیق و توسعه مواد مبتنی بر داده، داده‌های مواد پلیمری موجود در منابع مختلف تا حدی مانع از توسعه سریع سامانه استاندارد داده و ایجاد واردات داده می‌شود. الگو برای مقابله با انواع مختلف داده‌ها و قالب‌ها و همچنین مدیریت منظم و ذخیره‌سازی داده‌ها پیش‌نیاز است. همچنین می‌توان از هوش مصنوعی برای جمع‌آوری و طبقه‌بندی خودکار آخرین داده‌ها استفاده کرد. علاوه بر این، انتخاب عوامل آزمایشی تأثیر زیادی بر عملکرد و ساختار مواد دارند. شرایط تجربی مربوط به داده‌ها باید تکمیل و در پایگاه داده به‌روزرسانی شود تا نتایج شبیه‌سازی تأیید شود [۳۹].

۶ نتیجه‌گیری

PMGE فناوری نسل بعدی مواد، این قابلیت را دارد که هزینه تحقیقات مواد را کاهش دهد، محدودیت‌های عملکردی را متعادل کند و حتی پیشرفت‌هایی در مواد پلیمری ایجاد کند. PMGE می‌تواند روش‌های سنتی طراحی پلیمر را متحول کند و پیشرفت تحقیقات در علم مواد را ارتقا دهد. اصطلاح «ژنوم پلیمر» به رویکردی انفورماتیک برای پیش‌بینی و طراحی خواص پلیمر با استفاده از یادگیری ماشین اشاره دارد. PMGE بستری را فراهم می‌کند که در آن می‌توان پلیمرها را برای پیش‌بینی خواص آن‌ها جستجو و تجزیه‌وتحلیل کرد. این به ویژه برای طراحی پلیمرهای جدید با ویژگی‌های خاص یا برای درک این که چگونه تغییرات در ساختار پلیمر می‌تواند بر خواص کلی آن‌ها تأثیر بگذارد، مفید است. سکوی ژنوم پلیمر به کاربران اجازه می‌دهد تا ساختار شیمیایی پلیمر را با استفاده از رشته SMILES وارد کنند، که نمادی متنی است که ساختار گونه‌های شیمیایی را توصیف می‌کند. سپس این سکو از مدل‌های شبکه‌های عصبی عمیق چندوظیفه‌ای برای پیش‌بینی خواص مختلف پلیمر استفاده می‌کند. با این حال، بسیاری از مسائل باقی مانده است که باید مورد توجه قرار گیرند، زیرا PMGE هنوز در مراحل اولیه است. اشتراک بین‌رشته‌ای اطلاعات، ریاضیات، مهندسی کنترل و غیره می‌تواند مشکلاتی مانند پیش‌بینی ویژگی و تأیید تجربی را حل کند. در آینده، پیش‌بینی و غربالگری و همچنین طراحی معکوس ممکن است تحقق یابد. PMGE را می‌توان به‌عنوان سکویی عمومی و پایدار برای طراحی و کاربردهای پلیمری تصور کرد. محققان می‌توانند از PMGE برای طراحی منطقی پردازش، ترکیب و کارایی مواد پلیمری جدید در زمینه‌های کامپوزیت‌های پیشرفته، نیمه‌هادی‌ها، سامانه‌های ارتباطی و موارد دیگر استفاده کنند. با به کارگیری مهندسی ژنوم مواد، می‌توان نسل جدیدی از مواد پیشرفته برای ایجاد پیشرفت‌های خارق‌العاده‌ای که قادر به جرقه زدن نوآوری‌های جدید انقلابی باشد، ایجاد کرد.

مراجع

1. Pablo J.J.de, Jones B., Kovacs C.L., Ozolins V., and Ramirez A.P., The Materials Genome Initiative, the Interplay of Experiment, Theory and Computation, Curr. Opin. *Solid State Mater. Sci*., 18, 99–117, **2014**.

2. Pablo J.J.de, Jackson N.E., Webb M.A, Zhao J.C, New Frontiers for the Materials Genome Initiative, *NPJ Comput. Mater.*, 5, 41-64, **2019**.

3. Yuan W.L., He L., Tao G.H., and Shreeve J.M., Materials-Genome Approach to Energetic Materials," *Accounts Mater*. Res., 2, 692–696, **2021**.

4. Xie J., Su Y., Zhang D., and Feng Q., "A Vision of Materials Genome Engineering in China," *Engineering*, 10, 10–12, **2022**.

5. Rizkin B. A.and Hartman R. L., Supervised Machine Learning for Prediction of Zirconocene-Catalyzed α-Olefin Polymerization," *Chem. Eng. Sci*., 210, 115224-115236, **2019**.

6. Xu P., Chen H., Li M., and Lu W., New Opportunity: Machine Learning for Polymer Materials Design and Discovery," Adv. T*heory Simulations*, 5, 2100565-2100575, **2022**.

7. Agrawal A.and Choudhary A., Perspective: Materials Informatics and Big Data: Realization of the 'Fourth Paradigm' of Science in Materials Science, *APL Mater*., 4, 053208-053218, **2016**.

8. Gao L., Wang L., Lin J., and Du L., An Intelligent Manufacturing Platform of Polymers: Polymeric Material Genome Engineering, *Engineering*, 27, 31–36, **2023**.

9. Kannan S., Subbaram K., and Faiyazuddin M., Artificial Intelligence in Vaccine Development: Significance and Challenges Ahead, in a Handbook of Artificial Intelligence in Drug Delivery, Elsevier, Maldives, 467–486, **2023**.

10. Mannodi-Kanakkithodi A., Chandrasekaran A., Kim C., Huan T.D., Pilania G., Botu V., Ramprasad R., Scoping the Polymer Genome: A Roadmap for Rational Polymer Dielectrics Design and Beyond," *Mater. Today*, 21, 785–796, **2018**.

11. Zhao H., Li X., Zhang Y., Schadler L.S., Chen W., and Brinson L.C., Perspective: NanoMine: A Material Genome Approach for Polymer Nanocomposites Analysis and Design, *APL Mater*., 4, 053204-053214, **2016**.

12. Sharma V., Wang C., Lorenzini R.G., Ma R., Zhu Q., Sinkovits D.W., Pilania G., Oganov A. R., Kumar S., Sotzing G.A., Boggs S. A. and Ramprasad R., Rational Design of All Organic Polymer Dielectrics, *Nat. Commun*., 5, 4845-4853, **2014**.

13. Zhu J., Chu M., Chen Z., Wang L., Lin J., and Du L., Rational Design of Heat-Resistant Polymers with Low Curing Energies by a Materials Genome Approach," *Chem. Mater.*, 32, 4527–4535, **2020**.

14. Gao G., Zhang S., Wang L., Lin J., Qi H., Zhu J., Du L., and Chu M., Developing Highly Tough, Heat-Resistant Blend Thermosets Based on Silicon-Containing Arylacetylene: A Material Genome Approach, *ACS Appl. Mater. Interfaces*, 12, 27587–27597, Jun. **2020**.

15. Mannodi-Kanakkithodi A., Pilania G., Huan T. D., Lookman T., and Ramprasad R., Machine Learning Strategy for Accelerated Design of Polymer Dielectrics, *Sci. Rep.*, 6, 20952-20962, **2016**.

16. Zhang S., Du S., Wang L., Lin J., Du L., Xu X., Gao L., Design of Silicon-Containing Arylacetylene Resins Aided by Machine Learning Enhanced Materials Genome Approach, *Chem. Eng*. J., 448, 137643-137652, **2022**.

17. Song X., Lv L., Sun W., and Zhang J., A Radial Basis Function-Based Multi-Fidelity Surrogate Model: Exploring Correlation Between High-Fidelity and Low-Fidelity Models, Struct. *Multidiscip. Optim*., 60, 965–981, **2019**.

18. Huang Y., Zhang T., Zhu Z., Gong W., and Xia X., PM2.5 Concentration Estimation with 1-km Resolution at High Coverage over Urban Agglomerations in China Using the BPNN-KED Approach and Potential Application, *Atmos. Res.*, 258, 105628-105642, **2021**.

19. Chen Y., Mao Y., Pan X., Jin W., and Qiu T., "Verification and Comparison of Three Prediction Models of Ischemic Stroke in Young Adults Based on the Back Propagation Neural Networks, *Medicine (Baltimore)*., 100, e25081-e25086, **2021**.

20. Du B., Lund P. D., Wang J., Kolhe M., and Hu E., Comparative Study of Modelling the Thermal Efficiency of a Novel Straight Through Evacuated Tube Collector with MLR, SVR, BP and RBF Methods, *Sustain. Energy Technol. Assessments*, 44, 101029-101039, **2021**.

21. Han H.-G., Ma M.-L., and Qiao J.-F., Accelerated Gradient Algorithm for RBF Neural Network, *Neurocomputing*, 441, 237–247, **2021**.

22. Fang Q., "Estimation of Navigation Mark Floating Based on Fractional-Order Gradient Descent with Momentum for RBF Neural Network," *Math. Probl. Eng*., 2021, 1–10, **2021**.

23. Wu H., Zhao Y.-P., and Tan H.-J., Novel Radial Basis Function Network Based on Dynamic Time Warping and Kalman Filter for Real-Time Monitoring of Supersonic Inlet Flow Patterns, *J. Aerosp. Eng*., 34, 04021041- 04021053, **2021**.

24. Duriagina Z. A., Tkachenko R.O., Trostianchyn A.M. , Lemishka I.A. , Kovalchuk A.M., Kulyk V.V. , Kovbasyuk T.M.., Determination of the Best Microstructure and Titanium Alloy Powders Properties Using Neural Network," *J. Achiev. Mater. Manuf. Eng*., 1, 25–31, **2018**.
25. Jueyendah S., Lezgy-Nazargah M., Eskandari-Naddaf H., and Emamian S. A., "Predicting the Mechanical Properties of Cement Mortar Using the Support Vector Machine Approach," *Constr. Build. Mater*., 291, 123396-123408, **2021**.
26. Zhu Q., Wang Y., and Luo Y., "Improvement of Multi-Layer Soil Moisture Prediction Using Support Vector Machines and Ensemble Kalman Filter Coupled with Remote Sensing Soil Moisture Datasets over an Agriculture Dominant Basin in China," *Hydrol. Process*., 35, e14154-e14176, **2021**.
27. Sun Y. T., Bai H. Y., Li M. Z., and Wang W. H., Machine Learning Approach for Prediction and Understanding of Glass-Forming Ability, *J. Phys. Chem. Lett*., 8, 3434–3439, **2017**.
28. Tkachenko R., Duriagina Z., Lemishka I., Izonin I., and Trostianchyn A., Development of Machine Learning Method of Titanium Alloy Properties Identification in Additive Technologies, *Eastern-European J. Enterp. Technol*., 3, 23–31, **2018**.
29. Wolfensberger D., Gabella M., Boscacci M., Germann U., and Berne A., RainForest: a Random Forest Algorithm for Quantitative Precipitation Estimation over Switzerland, *Atmos. Meas. Tech.*, 14, 3169–3193, **2021**.
30. Speiser J. L., "A Random Forest Method with Feature Selection for Developing Medical Prediction Models with Clustered and Longitudinal Data," *J. Biomed. Inform*., 117, 103763-103774, **2021**.
31. Li M., Xu Y., Men J., Yan C., Tang H., Zhang T., Li H., "Hybrid Variable Selection Strategy Coupled with Random Forest (RF) for Quantitative Analysis of Methanol in Methanol-Gasoline via Raman Spectroscopy, Spectrochim. *Acta Part A Mol. Biomol. Spectrosc*., 251, 119430-119438, **2021**.
32. Fan Y., Bai J., Lei X., Lin W., Hu Q., Wu G., Guo J., Tan G.,PPMCK: Privacy-Preserving Multi-Party Computing for K-means Clustering, *J. Parallel Distrib. Comput*., 154, 54–63, **2021**.
33. Rim B., Lee S., Lee A., Gil H.-W., and Hong M., Semantic Cardiac Segmentation in Chest CT Images Using K-Means Clustering and the Mathematical Morphology Method, *Sensors*, 21, 2675-2694, **2021**.
34. Liu B., Zhang T., Li Y., Liu Z., and Zhang Z., "Kernel Probabilistic K-Means Clustering, *Sensors*, 21, 1892-1908, **2021**.
35. Xiong P., Liu H., Tian Y., Chen Z., Wang B., and Yang H., "Helicopter Maritime Search Area Planning Based on a Minimum Bounding Rectangle and K-means Clustering, Chinese *J. Aeronaut*., 34, 554–562, **2021**.
36. Liz H., Sánchez-Montañés M., Tagarro A., Domínguez-Rodríguez S., Dagan R., and Camacho D., Ensembles of Convolutional Neural Network Models for Pediatric Pneumonia Diagnosis, F*utur. Gener. Comput. Sys*t.,122, 220–233, **2021**.
37. Li M. and Ruan Z., A Novel Decoding Method for Motor Imagery Tasks with 4D Data Representation and 3D Convolutional Neural Networks, *J. Neural Eng*., 18, 046029-046050, **2021**.
38. Jia S.and Hu P., "ChrNet: A Re-trainable Chromosome-Based 1D Convolutional Neural Network for Predicting Immune Cell Types, *Genomics*, 113, 2023–2031, **2021**.
39. Qiu Y., Wu Z., Wang J., Zhang C., and Zhang H., Introduction of Materials Genome Technology and Its Applications in the Field of Biomedical Materials, *Materials (Basel).*, 16, 1906-1925, **2023**.
40. Boyd P. G., Lee Y., and Smit B., Computational Development of the Nanoporous Materials Genome, *Nat. Rev. Mater*., 2, 17037-17052, **2017**.
41. Lin T.S., Coley C. W., Mochigase H., Beech H. K., Wang W., Woods E., Craig S. L., Johnson J. A., Kalow J. A., Jensen K. F., Olsen B. D.,"BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules, *ACS Cent. Sci*., 5, 1523–1531, **2019**.
42. Hu Y., Zhao W., Wang L., Lin J., and Du L., Machine-Learning-Assisted Design of Highly Tough Thermosetting Polymers, *ACS Appl. Mater. Interfaces*, 14, 55004–55016, **2022**.
43. Shetty P.and Ramprasad R., Machine-Guided Polymer Knowledge Extraction Using Natural Language Processing: The Example of Named Entity Normalization," *J. Chem. Inf. Model.*, 61, 5377–5385, **2021**.
44. Wu S.F.and Hsieh Y.-T., The Assessment on the Lifetime Performance Index of Products with Gompertz Distribution Based on the Progressive Type I Interval Censored Sample, *J. Comput. Appl. Math*., 351, 66–76, **2019**.