

# پیش‌بینی دمای ذوب مایع‌های یونی بر پایه آنیون بیس (تری‌فلورومتیل سولفونیل) ایمید با رویکرد QSPR

شهربانو رحمان ستایش\*<sup>+</sup>، عبدالحلیم طریک، رحمان زبیدی

دانشکده شیمی، دانشگاه صنعتی شریف، تهران، ایران

**چکیده:** در این پژوهش دمای ذوب مایع‌های یونی بر پایه آنیون بیس (تری‌فلورومتیل سولفونیل) ایمید با استفاده از داده‌های تجربی ۶۴ مایع یونی و رویکرد مطالعه‌ها ارتباط کمی ساختار - ویژگی QSPR مدل سازی شد. ساختار مایع‌های یونی توسط برنامه آوگادرو ترسیم و شکل سه بعدی آن‌ها توسط میدان نیروی مولکولی مرک بهینه سازی شدند. با استفاده از نرم افزارهای درآگون و بدل توصیف کننده‌ها محاسبه شدند. پس از کاهش تعداد توصیف کننده‌ها از الگوریتم کنار-استون برای تقسیم بندی مایع‌های یونی به دو دسته آموزشی و ارزیابی بهره گرفته شد. برای انتخاب بهترین دسته از توصیف کننده‌ها از الگوریتم ژنتیک با تابع برازش فریدمن استفاده شد. سپس با استفاده از روش انتخاب متغیر مرحله‌ای MLR تعداد بهینه متغیرها تعیین شدند و معادله خطی میان متغیرها به دست آورده شد. مدل‌های خطی و غیرخطی با استفاده از روش‌های برازش خطی چندگانه MLR و شبکه عصبی پرسپترون چند-لایه MLP ارایه شدند. مقدارهای ضریب تشخیص  $R^2$  و میانگین انحراف نسبی ARD به ترتیب برابر با ۰/۷۸۷، ۰/۰۴۳ برای مدل خطی و ۰/۷۸۵، ۰/۰۴۲۸ برای مدل غیرخطی به دست آمد. بنابراین هر دو مدل از دقت یکسانی برخوردار بودند. دامنه کاربرد نیز برای مدل‌های ارایه شده تعیین شد. همچنین با استفاده از توصیف کننده‌های به دست آمده، تأثیر ساختار مولکولی در تعیین دمای ذوب بررسی شد.

**واژه‌های کلیدی:** مایع‌های یونی؛ الگوریتم ژنتیک؛ شبکه عصبی مصنوعی؛ دمای ذوب.

**KEYWORDS:** Ionic liquids; Genetic algorithm; Artificial neural network; Melting point.

## مقدمه

هدایت یونی بالا، حلالیت بالا، فشار بخار خیلی کم و اشتعال ناپذیری توجه پژوهشگران زیادی را به خود جلب کرده‌اند و به عنوان حلال‌های ایده‌آل برای شیمی سبز و سنتز پاک به حساب می‌آیند [۳]. ویژگی‌های مشخصه مایع‌های یونی با تغییر آنیون و یا کاتیون آن‌ها تغییر می‌کنند. ساختار یک مایع یونی تأثیر مستقیمی روی ویژگی‌های آن، به ویژه روی ویژگی‌های انتقال دارد [۴،۵]. مایع‌های یونی با وجود ویژگی‌های بسیار خوبی که دارند با توجه به قیمت بالا و اطلاعات کم از ویژگی‌های فیزیکی - شیمیایی آن‌ها

مایع‌های یونی از یک کاتیون آلی غیرمتقارن سنگین شامل نیتروژن یا فسفر مانند آلکیل ایمیدازولیوم، پیرولیدینیوم، آمونیم، فسفونیوم و انواع گوناگونی از آنیون‌های آلی مانند تری‌فلورو استات و یا آنیون‌های معدنی کلرید، برمید، تترافلوروبورات و هگزا فلوروفسفات و ... تشکیل شده‌اند. [۱]. نمک‌های معمول زیر دمای  $80^{\circ}\text{C}$  ذوب نمی‌شوند، ولی بیش تر مایع‌های یونی در دمای اتاق به صورت مایع هستند [۲]. این ترکیب‌ها به دلیل ویژگی‌های فیزیکی - شیمیایی منحصر به فردی که دارند مانند پایداری دمایی بالا،

\*عهده دار مکاتبات

+E-mail: setayesh@sharif.edu

کاربرد گسترده آن‌ها را با محدودیت رو به رو می‌کند. با افزایش استفاده از مایع‌های یونی در فرایندهای شیمیایی و صنعتی، نیاز به ویژگی‌های ترموفیزیکی قابل‌اعتماد برای درک بهتر رفتار ترمودینامیکی و برای توسعه مدل‌های ترمودینامیکی نیز فزونی یافته است [۶]. تخمین زده شده که تا  $10^{18}$  نوع گوناگون مایع یونی می‌تواند وجود داشته باشد [۷]، بنابراین لازم است که درکمان را نسبت به این نوع ترکیب‌ها بیشتر کنیم. دمای ذوب مایع‌های یونی به دلیل تأثیر مستقیم آن با سمیت و ایمنی از ویژگی‌های شیمی - فیزیکی پرکاربرد آن‌ها است. بنابراین تخمین دمای ذوب مایع‌های یونی برای تعیین ایمنی محیط‌زیستی آن‌ها از اهمیت ویژه‌ای برخوردار است. مایع‌های یونی با آنیون بیس (تری‌فلورومتیل‌سولفونیل) امید ویژگی‌های دلخواهی مانند پایداری گرمایی، آماده‌سازی آسان و خالص‌سازی، حلالیت کم در آب، توانایی بالای جذب گازهایی مانند  $H_2S$  و  $CO_2$  نسبت به مایع‌های یونی دیگر و همچنین قابلیت جداسازی و ذخیره کربن را دارا هستند. کاتریتزکی و همکاران [۸] مدلهایی را برای دمای ذوب ترکیب‌های پیریدینیوم برمید ارائه دادند. مدل‌های آن‌ها شامل دو معادله ۲ و ۶ متغیره است که به ترتیب ضریب همبستگی  $0.7125$  و  $0.7883$  دارند. مدل‌های یاد شده اهمیت توصیف‌کننده‌هایی مانند شاخص‌های محتوای اطلاعاتی، شاخص میانگین واکنش‌پذیری نوکولوفیلیک برای اتم N و آنترپوی کل به ازای هر اتم را در تعیین دمای ذوب این ترکیب‌ها نشان دادند. تروها لاک و همکاران [۹] دو مدل برای پیش‌بینی دمای ذوب مایع‌های یونی بر اساس کاتیون‌های ۱ - استخلافی - ۴ - آمینو - ۱،۲،۴ - تری‌ازولیوم و دو آنیون برمید و نیترات با قدرت پیش‌بینی بالایی ارائه کردند. رن و همکاران [۱۰] نیز برای ۲۸۸ نوع مایع یونی با آنیون برمید و کاتیون‌های پیریدینیوم، ایمیدازولیوم، بنزیمیدازولیوم و تریازولیوم مدل‌هایی خطی و غیرخطی ارائه کردند. مدل غیرخطی با ضریب همبستگی  $0.81$  و میانگین انحراف نسبی  $17/75\%$  نتیجه‌های بهتری را از مدل خطی با ضریب همبستگی  $0.712$  و میانگین انحراف نسبی  $24/33\%$  نشان داد. نتیجه‌های گزارش آن‌ها اهمیت ساختار حلقه بنزنی، پیوندهای قابل چرخش، شاخه‌دار شدن، تقارن و برهمکنش‌های الکترونی بین مولکولی در تعیین دمای ذوب این ترکیب‌ها را نشان داد.

در این پژوهش، دمای ذوب مایع‌های یونی بر پایه آنیون بیس (تری‌فلورومتیل‌سولفونیل) امید با استفاده از داده‌های تجربی ۶۴ مایع یونی و رویکرد مطالعه‌های ارتباط کمی ساختار - خاصیت

## بخش نظری

### مجموعه داده‌ها

دمای ذوب ۶۴ نوع مایع یونی بر پایه آنیون بیس (تری‌فلورومتیل‌سولفونیل) امید از پایگاه داده IL Thermo استخراج شد [۱۱]. داده‌های جمع‌آوری شده شامل ترکیب‌هایی با پایه‌های کاتیونی ایمیدازولیوم، پیریدینیوم، پیرولیدینیوم و آمونیوم است. نام کاتیون‌های این ترکیب‌ها و همچنین دمای ذوب بر حسب K به دست آمده از روش تجربی، MLR و شبکه عصبی مصنوعی در جدول ۱ آمده است. مقدارهای محاسبه شده توسط برازش خطی بر حسب مقدارهای تجربی دمای ذوب در شکل ۱ نمایش داده شده است.

### محاسبه توصیف‌کننده‌ها

ساختار مولکول‌های مایع‌های یونی توسط برنامه Avogadro ترسیم و شکل سه‌بعدی آن‌ها با استفاده از میدان نیروی مولکولی مرک بهینه سازی شد [۱۲]. سپس از فایل خروجی به دست آمده برای بهینه‌سازی ساختارها توسط روش نیمه تجربی PM6 از نرم‌افزار Gaussian استفاده شد. با استفاده از نرم‌افزارهای Padel [۱۳] و Dragon [۱۴] توصیف‌کننده‌های گوناگون که بالغ بر ۵۰۰۰ بودند، محاسبه شدند. پس از حذف توصیف‌کننده‌ها با مقدارهای ثابت تعداد کل توصیف‌کننده‌ها به ۳۹۵۳ عدد کاهش یافت. برای کاهش اولیه این تعداد توصیف‌کننده‌ها، ماتریس همبستگی متغیرها محاسبه و از متغیرهایی با ضریب همبستگی بالای  $0.95$  متغیری که همبستگی بیشتری با بقیه داشت را به‌عنوان نمونه انتخاب شد. سپس متغیرهایی که مقدارهای واریانس آن‌ها کمتر از  $0.1$  بودند

جدول ۱ - داده‌های دمای ذوب به دست آمده از تجربه، برازش خطی و شبکه عصبی.

شماره	نام کاتیون	تجربی	برازش خطی	خطای مطلق	شبکه عصبی	خطای مطلق
*۱	tetraethylammonium	۶/۳۷۶	۵۳/۳۷۶	۰.۶۵۸/۰	۹۱/۳۶۲	۶۹۱/۱۳
۲	1-methyl-1-propylpyrrolidinium	۶/۲۸۴	۲۶/۲۸۲	۳۳۵۵/۲	۳۹/۲۷۵	۲۱۱/۹
۳	1-butyl-1-methylpyrrolidinium	۶/۲۶۶	۹۳/۲۷۶	۳۲۸/۱۰	۱۳/۲۷۱	۵۳۴/۴
۴	1-ethyl-3-methylimidazolium	۴/۲۶۳	۲۰/۲۶۳	۱۹۸۷/۰	۵۰/۲۶۱	۸۹۹۳/۱
۵	triethyl(tetradecyl)phosphonium	۱/۳۰۶	۱۰/۳۱۳	۹۹۸۳/۶	۲۸/۳۱۴	۱۷۸/۸
۶	1-hexyl-3-methylimidazolium	۱/۲۷۲	۸۸/۲۵۷	۲۲۴/۱۴	۴۳/۲۵۸	۶۷۶/۱۳
۷	1-ethyl-3-methylpyridinium	۲۷۲	۳۴/۳۰۳	۳۴۵/۳۱	۳۱/۳۰۲	۳۰/۳۰
۸	trimethylsulfonium	۴/۳۲۲	۲۲/۳۲۹	۷۹۱۶/۶	۲۶/۳۲۰	۱۶۴۳/۲
۹	trimethylammonium	۳/۳۵۷	۰.۶/۳۳۵	۲۹۰/۲۲	۷۷/۳۲۵	۵۷۷/۳۱
۱۰	2-butylisoquinolinium	۳۲۱	۷۷/۳۰۹	۲۳۲/۱۱	۵۴/۳۰۹	۴۶۰/۱۱
۱۱	1,3-diethylimidazolium	۶/۲۶۲	۱۳/۲۷۸	۵۲۷/۱۵	۴۰/۲۷۵	۸۰/۱۲
۱۲	ethyl(2-hydroxyethyl)dimethylammonium	۱/۲۷۹	۶۶/۲۸۰	۵۰.۱۷/۱	۵۳/۲۷۸	۶۲۴۲/۰
۱۳	1-butyl-4-methylpyridinium	۴/۲۹۱	۷۵/۲۹۸	۳۴۷۵/۷	۵۹/۳۰۰	۱۹۵/۹
۱۴	triethylsulfonium	۸/۲۶۲	۹۶/۲۵۹	۸۳۸۵/۲	۴۸/۲۵۸	۳۲۰۹/۴
۱۵	1-hexylquinolinium	۲/۳۱۷	۹۵/۳۱۰	۲۴۴۳/۶	۶۹/۳۱۲	۵۰۷۹/۴
۱۶	1-ethyl-1-methylpiperidinium	۷/۳۵۷	۷۲/۳۲۷	۹۸۹/۲۹	۷۷/۳۲۱	۹۳۴/۳۵
۱۷	1-methyl-3-(3-phenylpropyl)-1H-imidazolium	۱/۳۲۱	۹۶/۳۱۶	۱۳۴۲/۴	۰.۲/۳۱۴	۰.۷۵۸/۷
۱۸	1-methyl-3-pentylimidazolium	۱/۲۶۴	۷۷/۲۵۹	۳۳۲۹/۴	۲۵/۲۶۰	۸۵۲۸/۳
۱۹	1-heptyl-3-methylimidazolium	۱/۲۸۰	۲۰/۲۶۱	۸۹۵/۱۸	۹۸/۲۶۱	۱۱۶/۱۸
۲۰	1-decyl-3-methylimidazolium	۱/۲۷۱	۹۳/۲۵۸	۱۶۵/۱۲	۸۱/۲۵۸	۲۹۱/۱۲
۲۱	1-methyl-3-(2-phenylethyl)-1H-imidazolium	۱/۳۱۰	۰.۶/۳۱۰	۰.۳۵۵/۰	۴۲/۳۰۶	۶۷۷۶/۳
*۲۲	1,2-dimethyl-3-propylimidazolium	۲۸۱	۲۷/۲۸۵	۲۷۵۴/۴	۴۳/۲۸۴	۴۲۶/۳
۲۳	1-ethyl-2,3-dimethylimidazolium	۱/۲۹۸	۰.۷/۲۸۷	۰.۲۴/۱۱	۲۸/۲۸۴	۸۱۵/۱۳
۲۴	1-ethyl-1-methylpyrrolidinium	۱/۳۶۳	۴۴/۳۴۶	۶۵۹/۱۶	۶۳/۳۳۲	۴۷۳/۳۰
۲۵	1-octyl-3-methylimidazolium	۸/۲۴۹	۵۲/۲۶۴	۷۱۲/۱۴	۵۰/۲۶۵	۶۹۱/۱۵
۲۶	2,3,5-trimethyl-1-propyl-1H-pyrazolium	۱/۲۹۰	۸۳/۲۸۵	۲۷۱۲/۴	۶۰/۲۷۹	۵۰۴/۱۰
۲۷	methyltripropylammonium	۱/۳۱۹	۹۵/۳۳۶	۸۴۷/۱۷	۶۲/۳۳۰	۵۲۵/۱۱
*۲۸	N,N,N-trimethylpropanaminium	۱/۲۹۲	۶۴/۳۰۶	۵۳۶/۱۴	۴۴/۳۰۱	۳۴۲۳/۹
*۲۹	1-methyl-1-propylpiperidinium	۱/۲۸۵	۰.۵/۳۰۴	۹۴۶/۱۸	۵۴/۲۹۶	۴۳۸/۱۱
۳۰	1-ethylpyridinium	۶/۳۰۳	۵۷/۲۹۵	۰.۳۳۲/۸	۶۹/۲۹۵	۹۰۴۸/۷
*۳۱	1-butylpyridinium	۱/۲۹۹	۱۶/۲۸۵	۹۳۸/۱۳	۰.۳/۲۸۷	۰.۶۹/۱۲
*۳۲	1-pentylpyridinium	۸/۲۷۲	۶۴/۲۸۳	۸۴۵/۱۰	۰.۸/۲۸۶	۲۷۷/۱۳

جدول ۱ - داده‌های دمای ذوب به دست آمده از تجربه، برازش خطی و شبکه عصبی. (ادامه)

شماره	نام کاتیون	تجربی	رگرسیون خطی	خطای مطلق	شبکه عصبی	خطای مطلق
*۳۳	1-propylpyridinium	۶/۳۱۷	۲۹/۲۹۰	۳۰۷/۲۷	۵۳/۲۹۱	۰۶۵/۲۶
۳۴	1-hexylpyridinium	۴/۲۷۶	۹۵/۲۸۳	۵۵۲۸/۷	۸۳/۲۸۶	۴۳۱/۱۰
*۳۵	1-pentyl-1-methylpyrrolidinium	۱/۲۸۱	۲۵/۲۷۶	۸۵۱۹/۴	۹۲/۲۷۰	۱۷۸/۱۰
*۳۶	triethyloctylammonium	۱/۲۸۷	۲۱/۳۱۷	۱۱۴/۳۰	۴۶/۳۱۴	۳۵۶/۲۷
۳۷	N,N-dimethyl-N-propyl-1-butanaminium	۱/۲۹۳	۱۶/۲۸۸	۹۳۵۷/۴	۸۵/۲۸۴	۲۴۷/۸
۳۸	butyltrimethylammonium	۱/۲۸۹	۴۲/۲۵۹	۶۸۴/۲۹	۶۷/۲۶۰	۴۳۱/۲۸
*۳۹	1-hexyl-1-methylpyrrolidinium	۱/۲۷۶	۸۲/۲۷۶	۷۲۰۳/۰	۶۶/۲۷۱	۴۳۵۷/۴
۴۰	N-ethyl-N,N-dimethyl-1-propanaminium	۱/۲۶۳	۸۹/۲۷۹	۷۸۶/۱۶	۷۸/۲۷۹	۶۷۶/۱۶
۴۱	N,N-dimethyl-N-ethyl-N-pentylammonium	۷/۲۶۶	۴۱/۲۷۸	۷۰۹/۱۱	۴۵/۲۷۸	۷۵۳/۱۱
۴۲	1-hexyl-1,4-diazabicyclo[2.2.2]octan-1-ium	۳۰۹	۵۲/۳۲۲	۵۱۸/۱۳	۸۸/۳۱۷	۸۸۴۰/۸
۴۳	1-oleyl-3-methylimidazolium	۳/۲۵۲	۷۲/۲۹۲	۴۱۷/۴۰	۳۶/۲۹۳	۰۶۳/۴۱
۴۴	1-methyl-3-(9E)-9-octadecen-1-yl-1H-Imidazolium	۲/۲۸۹	۲۶/۲۷۸	۹۴۳/۱۰	۹۵/۲۷۷	۲۴۶/۱۱
۴۵	1-methyl-3-(9Z,12Z)-9,12-octadecadien-1-yl-1H-imidazolium	۴/۲۲۶	۵۶/۲۵۵	۱۶۲/۲۹	۱۵/۲۵۲	۷۵۳/۲۵
۴۶	tetrabutylammonium	۱/۳۶۳	۶۳/۳۵۰	۴۶۹/۱۲	۹۱/۳۴۲	۱۸۸/۲۰
۴۷	1,2,3,4,5-pentamethylimidazolium	۱/۳۹۱	۱۸/۳۹۸	۰۸۱۱/۷	۳۳/۳۸۳	۷۶۴/۷
۴۸	N-hexylisoquinolinium	۳/۳۲۸	۳۹/۳۰۵	۹۱۱/۲۲	۸۴/۳۰۶	۴۵۶/۲۱
۴۹	1-carboxy-N,N,N-trimethylmethanaminium	۳۳۰	۴۳/۳۴۴	۴۲۶/۱۴	۲۷/۳۳۵	۲۷۱۵/۵
۵۰	1-butyl-3-methylimidazolium	۳/۲۷۰	۵۶/۲۶۰	۷۹۱۹/۹	۹۹/۲۶۰	۳۵۵۴/۹
*۵۱	triethylhexylammonium	۱/۲۹۳	۵۸/۲۹۹	۴۸۳/۱۶	۸۸/۲۹۴	۷۸۲۹/۱
۵۲	tetrapropylammonium	۱/۳۷۸	۴۶/۳۵۵	۶۴۳/۲۲	۴۱/۳۴۶	۶۹۱/۳۱
۵۳	tributylhexylammonium	۱/۲۹۹	۰۸/۳۱۹	۹۸۱/۱۹	۱۰/۳۱۴	۹۹۹/۱۴
۵۴	3-methyl-1-(naphthalen-2-ylmethyl)-1H-imidazol-3-ium	۶/۳۱۸	۶۶/۳۲۲	۰۶۱۵/۴	۳۱/۳۱۸	۲۹۲۴/۰
۵۵	1,3-dibenzyl-1H-imidazol-3-ium	۹/۳۱۴	۳۵/۳۲۰	۴۵۰۹/۵	۲۲/۳۱۶	۳۲۵۳/۱
*۵۶	1,3-dimethylimidazolium	۱/۲۹۹	۲۵/۳۱۹	۱۴۷/۲۰	۶۱/۳۰۸	۵۰۵۵/۹
۵۷	2-hydroxyethyl-N,N,N-trimethylammonium	۱/۳۰۶	۸۷/۳۲۰	۷۶۶/۱۴	۶۵/۳۱۳	۵۵۰۹/۷
۵۸	4-(3-hydroxypropyl)-4-methylmorpholinium	۲۸۷	۵۳/۲۹۶	۵۲۶۸/۹	۳۳/۲۹۳	۳۳۴/۱۶
*۵۹	1-(5-methylhexyl)-3-methylimidazolium	۳/۲۸۴	۸۵/۲۶۷	۴۵۲/۱۶	۷۰/۲۶۶	۵۹۶/۱۷
*۶۰	1-isobutyl-3-methylimidazolium	۹/۲۵۶	۳۲/۲۷۲	۴۲۵/۱۵	۱۳/۲۶۹	۲۲۹/۱۲
*۶۱	1-(3-methylbutyl)-3-methylimidazolium	۱/۲۷۳	۶۳/۲۷۰	۴۶۶۸/۲	۲۹/۲۶۹	۸۰۵۷/۳
۶۲	1-(4-methylpentyl)-3-methylimidazolium	۳/۲۷۶	۳۴/۲۶۶	۹۵۵۶/۹	۸۳/۲۶۴	۴۷۱/۱۱
*۶۳	1-methyl-3-(1-methylethyl)imidazolium	۱/۲۸۲	۲۵/۲۷۸	۸۴۹۸/۳	۸۹/۲۷۴	۲۱۱۳/۷
۶۴	1-octylquinolinium	۳/۳۲۱	۷۰/۳۰۶	۶۰۲/۱۴	۷۴/۳۰۸	۵۵۹/۱۲

\*نسخه ارزیابی

در پایان نیز هر یک از نسل‌ها توسط تابع تناسب ارزش‌یابی می‌شود. ابزارهای آماری گوناگونی برای اندازه‌گیری تناسب یک مدل بر اساس الگوریتم ژنتیک هنگام فرایند تکامل وجود دارد. استفاده از ابزار فریدمن چندین برتری نسبت به روش معمول کم‌ترین مربع‌ها دارد. فرمول مورد استفاده فریدمن در اینجا به صورت زیر است:

$$LOF = \frac{SSE}{M \left[ 1 - \lambda \left( \frac{c + dp}{M} \right) \right]^2} \quad (1)$$

SSE مجموع مربع‌های خطا، c تعداد پارامترهای مدل، d پارامتر اسموسنس، p تعداد کل متغیرها، M تعداد نمونه در سری آموزشی،  $\lambda$  پارامتری است با مقدار ۰/۹۹ که از صفر شدن مخرج جلوگیری می‌کند. همچنین d از طریق معادله زیر به  $\alpha$  مرتبط است که در آن  $c_{max}$  بیش‌ترین تعداد متغیرهای موردنظر است.

$$d = \alpha \left( \frac{M - c_{max}}{c_{max}} \right) \quad (2)$$

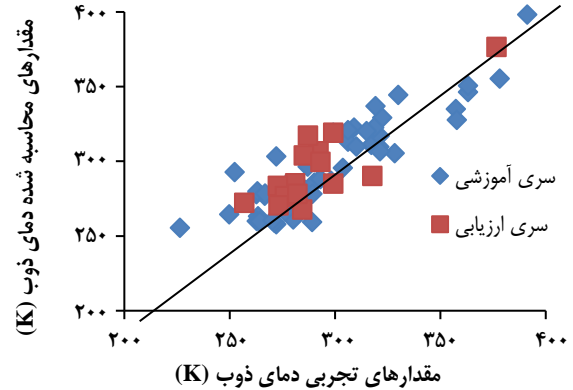
برخلاف روش متداول کم‌ترین مربع‌ها، مقدار LOF را نمی‌توان با افزایش ترم‌های بیشتر در مدل، کاهش داد. ممکن است که جمله نوین SSE را کاهش دهد، ولی از سویی باعث افزایش مقدارهای c و p می‌شود که تمایل به افزایش مقدارهای LOF دارند. بنابراین، افزودن یک جمله ممکن است که SSE را کاهش دهد ولی در عمل باعث افزایش مقدار LOF می‌شود. با محدود کردن تمایل به افزایش جمله‌های بیش‌تر، ابزار LOF در برابر بیش‌تر برآزش بهتر از ابزار SSE عمل می‌کند.

### نتیجه‌ها و بحث

با استفاده از نرم افزار Materials Studio [۱۶] الگوریتم ژنتیک با تابع برآزش فریدمن،  $\alpha$  معادل ۰/۵، احتمال جهش ۰/۴۵، جمعیت ۵۰۰ و تعداد نسل‌ها معادل ۵۰۰۰ پیاده‌سازی شد تا ۶ متغیر برتر برای مدل‌سازی دمای ذوب تعیین شد. مدل خطی به دست آمده در زیر نمایش داده شده است:

$$MT = 183,50 + 94,30(BELe) - 22,00(Mor) + u - 8,40(AATSVe) + 2,00(VPC-5) - 149,42(AVP-4) - 123,35(VE1-D) \quad (3)$$

متغیرهای منتخب عبارت هستند از: کم‌ترین ویژه‌مقدار ماتریس burden



شکل ۱- مقدارهای محاسبه شده برآزش خطی برحسب مقدارهای تجربی دمای ذوب.

نیز حذف شد، پس از انجام این مرحله تعداد توصیف‌کننده‌ها به ۱۲۶۹ عدد کاهش یافت. با توجه به اهمیت تقسیم‌بندی منطقی داده‌ها در مدل‌های QSPR/QSAR، از الگوریتم کنار-استون [۱۵] برای تقسیم بندی مایع‌های یونی به دو دسته آموزشی و ارزیابی بهره گرفته شد. بدین ترتیب از ۶۴ مایع یونی، تعداد ۱۶ مایع یونی به عنوان مجموعه ارزیابی در نظر گرفته شدند و برای مدل‌سازی دمای ذوب از ۴۸ مایع یونی دیگر استفاده شد.

### انتخاب توصیف‌کننده‌ها

#### الگوریتم ژنتیک

برای انتخاب مناسب‌ترین توصیف‌کننده‌ها توسط الگوریتم ژنتیک، تکامل جمعیت شبیه‌سازی شد. جمعیت نسل اول به صورت تصادفی انتخاب شد، هر یک از اعضای این جمعیت توسط کروموزوم‌های دوتایی<sup>(۱)</sup> مشخص شده و زیر مجموعه‌ای از توصیف‌کننده‌ها را نمایش می‌دهند. تعداد ژن‌ها در هر کروموزوم معادل تعداد توصیف‌کننده‌ها است. به هر یک از ژن‌ها عدد ۱ داده می‌شود در صورتی که توصیف‌کننده مربوط به آن در زیرمجموعه موجود باشد و در غیر این صورت عدد صفر به آن داده می‌شود. ادامه جمعیت به نسل‌های بعدی با استفاده از اپراتورهای جهش و تقاطع انجام می‌گیرد. جهش به این صورت است که برخی ژن‌ها به صورت کاملاً تصادفی تغییر می‌کنند. مقدار جهش انتخاب شده در این مطالعه ۰/۴۵ است که با توجه به بزرگی تعداد توصیف‌کننده‌ها و زمان لازم [۱۸] برای همگرایی این عدد انتخاب شد. تقاطع نیز برای ترکیب کردن اطلاعات ژنتیکی والدین برای تولید فرزندان استفاده می‌شود.

(۱) Binary

جدول ۲- ماتریس ضریب‌های همبستگی بین توصیف‌کننده‌ها.

	BELe1	Mor10u	AATS7e	VPC-5	AVP-4	VE1-D
BELe1	۱					
Mor10u	-.۰۴۶۰۶۱	۱				
AATS7e	۰/۵۰۵۱۶۶	-.۰۷۵۱۶	۱			
VPC-5	۰/۱۰۱۳۴۸	-.۰۶۱۶۵	۰/۰۸۶۵۵۸	۱		
AVP-4	-.۰/۱۵۴۹۲	۰/۱۲۹۳۲۲	-.۰/۱۵۶۵۱	۰/۳۸۸۴۶۲	۱	
VE1-D	۰/۳۹۵۱۰۷	-.۰/۳۱۶۹۶	۰/۳۵۳۲۸۸	۰/۱۱۰۵۵۸	-.۰/۱۰۳۶۲	۱

در یافتن پاسخ بهینه می‌شود. این روش در اساس شامل حل کردن معادله زیر می‌باشد:

$$(J^T J + \lambda I) \delta = J^T E$$

که در آن  $J$  ماتریس ژاکوبیان<sup>(۱)</sup> سامانه است،  $\lambda$  فاکتور میرایی<sup>(۸)</sup>،  $\delta$  بردار تغییرهای وزن است که می‌خواهیم به دست آوریم و  $E$  بردار خطا است که شامل خطاهای خروجی برای هر بردار ورودی استفاده شده برای آموزش شبکه است.  $\delta$  تعیین می‌کند که چقدر باید وزن‌های شبکه را تغییر داده شود تا (شاید) به یک پاسخ بهتر برسیم. ماتریس  $J^T$  به عنوان ماتریس هسیان<sup>(۹)</sup> تقریب زده شده شناخته می‌شود.  $\lambda$  در هر تکرار<sup>(۱۰)</sup> تنظیم می‌شود تا فرایند بهینه‌سازی را کنترل کند. اگر کاهش  $E$  سریع باشد، از مقدارهای کم‌تر  $\lambda$  می‌توان استفاده کرد، که الگوریتم را به الگوریتم گاوس - نیوتون<sup>(۱۱)</sup> نزدیک‌تر می‌کند، درحالی‌که اگر یک تکرار میزان ناپیزی مقدارهای باقیمانده را تغییر دهد،  $\lambda$  را می‌توان افزایش داد، که در این صورت به الگوریتم گرادیان نزولی<sup>(۱۲)</sup> شبیه می‌شود. پارامتر بهینه‌سازی برای این نوع شبکه تعداد گره‌ها در لایه پنهان بود که پس از بهینه‌سازی تعداد گره‌ها ۹ در نظر گرفته شد. پیش‌بینی‌های دمای ذوب توسط این مدل در جدول ۱ نمایش داده شده است. همچنین شکل ۲ مقدارهای تجربی دمای ذوب در برابر مقدارهای محاسبه شده آن از طریق این مدل را نشان می‌دهد.

ارزش دهی شده توسط الکترون‌گاتیوی‌های اتمی ساندرسون (BELe1)<sup>(۱)</sup>، توصیف‌کننده سه بعدی MoRSE سیگنال ۱۰ بدون ارزش‌دهی (Mor10u)<sup>(۲)</sup>، میانگین همبستگی Broto-Moreau در فاصله ۷ ارزش‌دهی شده توسط الکترون‌گاتیوی‌های ساندرسون<sup>(۳)</sup> (AATS7e)، مسیر خوشه‌ای والانس از مرتبه ۵ (VPC-5)<sup>(۴)</sup>، میانگین مسیر والانس از مرتبه ۴<sup>(۵)</sup>، مجموع ضریب‌های آخرین ویژه بردار از ماتریس فاصله‌های توپولوژیکی (VE1-D)<sup>(۶)</sup> [۱۷].  
جدول ۲ ماتریس ضریب‌های همبستگی بین توصیف‌کننده‌های منتخب را نمایش می‌دهد. با توجه به مقدارهای همبستگی می‌توان نتیجه گرفت که توصیف‌کننده‌ها استقلال خطی دارند.

#### مدل‌سازی به روش شبکه عصبی پرسپترون چندلایه

با استفاده از ۶ توصیف‌کننده به دست آمده، از شبکه عصبی پرسپترون چندلایه برای مدل‌سازی غیرخطی استفاده شد. الگوریتم لونیبرگ - مارکوارت برای آموزش شبکه استفاده شد. الگوریتم لونیبرگ - مارکوارت، روشی ساده ولی کارآمد برای تقریب یک تابع است. این الگوریتم به طور مستقیم هدف کمینه کردن میانگین مربع‌های خطای کل نمونه را دارد و نه فقط خطای هر یک از آموزش‌ها به صورت جدا و موجب فراهم سازی قدرت بیش‌تری

(۱) Lowest eigenvalue of the Burden matrix, weighted by atomic Sanderson electronegativities

(۲) 3D MoRSE signal 10 / unweighted

(۳) Average Broto-Moreau autocorrelation - lag 7 / weighted by Sanderson electronegativities

(۴) Valence path cluster / order 5

(۵) Average valence path of order 4

(۶) Coefficient sum of the last eigenvector from topological distance matrix

(۷) Jacobian matrix

(۸) Damping factor

(۹) Hessian

(۱۰) Iteration

(۱۱) Gauss-Newton

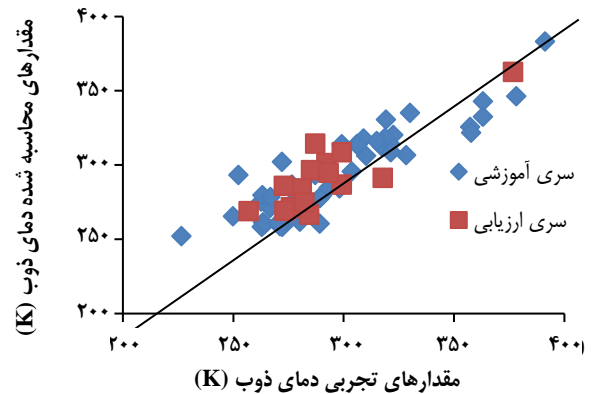
(۱۲) Gradient descent

$x$  نیز ماتریس یا برداری است که می‌خواهیم برای آن مقدارهای قدرت نفوذ را به دست آوریم، که هم شامل مایع‌های یونی مجموعه آموزشی و هم سری ارزیابی می‌باشد. توجه شود که مقدارهای قدرت نفوذ، در صورتی که برای بیش از یک ترکیب در  $x$  از معادله بالا محاسبه شود، برابر عناصر قطری ماتریس به دست آمده است. در شکل ۳ خط عمودی، نشان دهنده قدرت نفوذ  $h^* = \frac{3n}{p}$  است که در اینجا برابر  $0/4375$  است. خط‌های افقی نشان‌دهنده سه برابر انحراف استاندارد داده‌ها است. با توجه به شکل دیده می‌شود که تنها ترکیب ۱۴ در بازه‌ی دامنه کاربرد مدل قرار نگرفته است. هرچند که انحراف استاندارد کمی داشته و پیش‌بینی آن به نسبت خوب است. با استفاده از توصیف‌کننده‌های به دست آمده، مدل‌هایی خطی و غیرخطی برای پیش‌بینی دمای ذوب مایع‌های یونی بر پایه بیس (تری‌فلورومتیل سولفونیل) امید طراحی گردید. مقدارهای دمای ذوب تجربی و همچنین پیش‌بینی شده توسط مدل خطی و غیرخطی نیز در جدول ۱ نمایش داده شده است. با توجه به پارامترهای ارزیابی داخلی و خارجی می‌توان مشاهده کرد که مدل‌های خطی و غیرخطی از لحاظ آماری ارزش‌های یکسانی دارند، هرچند شبکه عصبی مصنوعی در پیش‌بینی سری ارزیابی بهتر از مدل خطی عمل کرده است.

### بررسی توصیف‌کننده‌ها

مقدارهای  $t$  نمایش داده شده در جدول ۴، نشان‌دهنده اهمیت توصیف‌کننده‌ها به ترتیب زیر است:

VPC-5 > AATS7e > AVP-4 > VE1-D > BELe1 > Mor10u  
 VPC-5 بیش‌ترین تأثیر را در تعیین دمای ذوب دارد. این توصیف‌کننده مربوط به گراف مولکولی ۲ - متیل پنتان و ۳-متیل پنتان می‌باشند و دربرگیرنده‌ی موردهایی مانند عدد کتوردیناسیون اتم‌ها، تعداد پیوندهای هیدروژنی و همچنین چگونگی اتصال اتم‌ها به یکدیگر می‌باشد. AATS7e میانگین خود همبستگی فضایی در فاصله ۷ است که خاصیت ارزش‌دهی، الکترونگاتیوی‌های اتمی ساندسون است و به توصیف‌کننده‌های دوبعدی تعلق دارد. در این نوع توصیف‌کننده‌ها اتم‌های موجود در مولکول نقطه‌هایی را در فضای تابع شیمی-فیزیکی مورد نظر نمایش می‌دهند. AVP-4، میانگین مسیر والانس از مرتبه ۴ است. این توصیف‌کننده نیز به‌مانند VPC-5 می‌باشد، با این تفاوت که میانگین‌های مربوط به مسیرها به طول ۴ را در نظر می‌گیرد.



شکل ۲- مقدارهای محاسبه شده از شبکه عصبی بر حسب مقدارهای تجربی دمای ذوب.

پارامترهای مورد استفاده برای ارزیابی مدل که شامل ضریب تشخیص  $R^2$ ، ضریب تشخیص تعدیل شده  $R_{adj}^2$ ، میانگین انحراف نسبی ARD است به صورت زیر تعریف می‌شوند:

$$R^r = 1 - \frac{\sum_{i=1}^n (y_i^{exp} - y_i^{cal})^r}{\sum_{i=1}^n (y_i^{exp} - \bar{y})^r} \quad (4)$$

$$R_{adj}^r = 1 - \left[ \frac{(1 - R^r)(n-1)}{n-k-1} \right] \quad (5)$$

$$ARD = \frac{\sum_{i=1}^n \left| \frac{y_i^{exp} - y_i^{cal}}{y_i^{exp}} \right|}{n} \quad (6)$$

که در آن  $y_i^{exp}$  مقدار تجربی دمای ذوب،  $y_i^{cal}$  مقدار محاسبه شده،  $\bar{y}$  مقدار متوسط دمای ذوب تجربی،  $n$  تعداد نقطه‌های در نمونه داده‌ها،  $k$  تعداد متغیرها در مدل است. پارامترهای ارزیابی درونی و بیرونی برای مدل شبکه عصبی پرسپترون چندلایه نیز در جدول ۳ مورد بررسی قرار گرفته‌اند.

### دامنه کاربرد مدل

شکل ۳ که به نمودار ویلیامز معروف است، انحراف استاندارد بر حسب قدرت نفوذ دیده می‌شود. قدرت نفوذ  $h$ ، توسط معادله زیر محاسبه می‌شود:

$$h = x^T (X^T X)^{-1} x^T \quad (7)$$

که در آن  $X$  ماتریسی شامل  $n$  مایع یونی (ردیف‌ها) و  $k$  توصیف‌کننده مربوط به مایع یونی (ستون) از مجموعه آموزشی است.

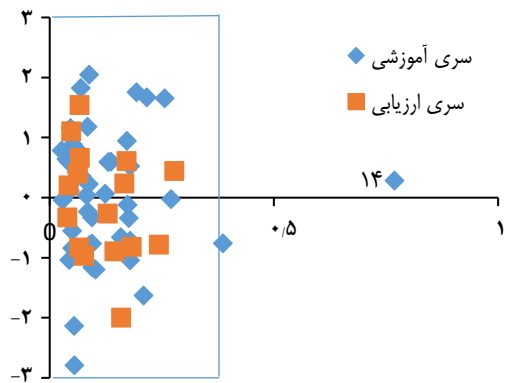
شاخص‌های VED، به‌عنوان توصیف‌کننده، فاصله‌های بین اتم‌ها در گراف مولکولی را در خود جای می‌دهد. BELE1 کم‌ترین ویژه‌مقدار ماتریس Burden، ارزش‌دهی شده توسط الکترون‌گاتیوی اتمی ساندرسون است که از ویژه‌مقدارهای مثبت و منفی ماتریس مجاورت به دست می‌آید. این توصیف‌کننده‌ها، توزیع‌های سطحی بارهای مثبت، بارهای منفی، دهنده‌های پیوند هیدروژنی، پذیرنده‌های پیوند هیدروژنی، منطقه‌هایی با قطبیت بالا، مناطق با قطبیت پایین را مشخص می‌کنند. به‌طور کلی این توصیف‌کننده اطلاعات مربوط به فاصله‌ها، الکترون‌گاتیوی‌ها و انواع اتم را در خود ذخیره می‌کنند. Mor10u توصیف‌کننده سه‌بعدی MorSE است، این دسته از توصیف‌کننده‌ها بر اساس ایده به‌دست آوردن اطلاعات از مختصات سه‌بعدی اتم‌ها و تبدیل‌های استفاده شده در مطالعه‌های پراش الکترونی می‌باشد. این توصیف‌کننده اهمیت ساختار سه‌بعدی مولکول را نشان می‌دهد.

### نتیجه‌گیری

در این مطالعه توانایی رویکرد مطالعه‌های ارتباط کمی ساختار - خاصیت QSPR در پیش‌بینی ویژگی‌های شیمی - فیزیکی مانند دمای ذوب مایع‌های یونی بر پایه آنیون بیس (تری‌فلورومتیل‌سولفونیل) امید با رویکرد QSPR بررسی شد. در مدل‌سازی، داده‌های مربوط به دمای ذوب ۶۴ نوع مایع یونی استخراج و بیش از ۵۰۰۰ نوع توصیف‌کننده گوناگون محاسبه شدند. با استفاده از الگوریتم ژنتیک و برازش خطی چند گانه مرحله‌ای تعداد توصیف‌کننده‌های بهینه برابر با ۶ به‌دست آمد که به عنوان توصیف‌کننده‌های برتر انتخاب شدند. و دو مدل خطی و غیرخطی تشکیل شدند. مقدارهای ضریب تشخیص  $R^2$  و میانگین انحراف نسبی ARD به ترتیب برابر با ۰/۷۸۷، ۰/۴۳ برای مدل خطی و ۰/۷۸۵، ۰/۴۲۸ برای مدل غیرخطی به‌دست آمد. با توجه به پارامترهای ارزیابی درونی و بیرونی دیده شد که مدل‌های خطی و غیرخطی از لحاظ آماری ارزش‌های یکسانی دارند، هرچند شبکه عصبی مصنوعی در پیش‌بینی سری ارزیابی بهتر از مدل خطی عمل کرده است. بررسی توصیف‌کننده‌ها اهمیت عدد کوردیناسیون اتم‌ها، تعداد پیوندهای هیدروژنی، الکترون‌گاتیوی اتم‌های موجود، توزیع بارها و ساختار سه‌بعدی را نشان داد. همه ترکیب‌ها به غیر از ترکیب شماره ۱۴ در بازه‌ی دامنه کاربرد

جدول ۳ - بررسی کیفیت مدل‌ها بر روی سری‌های آموزشی و ارزیابی و کل داده‌ها.

سری	پارامتر	دمای ذوب	
		خطی	شبکه عصبی
آموزشی	R	۸۹۷/۰	۸۹۲/۰
	$R^2$	۸۰۴/۰	۷۹۶/۰
	$R^2_{adj}$	۸۰۰/۰	۷۹۲/۰
	Std. Error	۸۷۸/۱۵	۲۰۴/۱۶
	ARD	۰۴۳/۰	۰۴۴/۰
ارزیابی	R	۸۴۴/۰	۸۵۴/۰
	$R^2$	۷۱۳/۰	۷۲۹/۰
	$R^2_{adj}$	۶۹۲/۰	۷۰۹/۰
	Std. Error	۷۲۹/۱۴	۳۱۳/۱۴
	ARD	۰۴۱/۰	۰۳۹/۰
کل	R	۸۸۷/۰	۸۸۶/۰
	$R^2$	۷۸۷/۰	۷۸۵/۰
	$R^2_{adj}$	۷۸۳/۰	۷۸۱/۰
	Std. Error	۵۸۳/۱۵	۶۵۰/۱۵
	ARD	۰۴۳/۰	۰۴۲۸/۰



شکل ۳- قلمرو کاربرد مدل ارایه شده برای دمای ذوب مایع‌های یونی.



جدول ۴- آماره‌های مربوط به برازش خطی.

Descriptor	Type	Coefficient	Std. Error	t-Value	p-Value	F-Value
VPC-5	topological	۰.۶۳۹۷/۲۰	۴۸۶۲۱/۲	۰.۶۹۷۱/۸	.	۱۲۰/۶۵
AATS7e	topological	۴۴۱۲۱/۸-	۲۰.۶۴۳/۱	۹۹۶۸۶/۶-	.	۹۵۶/۴۸
AVP-4	topological	۴۱۸/۱۴۹-	۲۶۴۵۱/۲۶	۶۸۸۹۷/۵-	.	۳۶۴/۳۲
BELe1	BCUT	۳۷۷۴۶/۹۴	۱۱۷۲/۲۴	۹۱۳۲۸/۳	۰.۰۳۴/۰	۳۱۴/۱۵
VE1-D	topological	۳۵۵/۱۲۳-	۷۴۵۷۳/۲۵	۷۹۱۲۸/۴-	۰.۰۰۰۲/۰	۹۵۶/۲۲
Mor10u	3D-MoRSE	۰.۹۶۹/۳۲-	۱۱۲۰۲/۷	۱۰.۶۹۸/۳-	۰.۰۳۴۳/۰	۶۵۳/۹

### قردانی

نویسندگان از همکاری دانشکده شیمی دانشگاه صنعتی شریف در این کار پژوهشی تشکر می‌نمایند.

مدل قرار گرفته و از پیش‌بینی‌های خوبی برخوردار بودند. اطلاعات ساختاری ارائه شده در این کار می‌تواند برای طراحی منطقی مجموعه نوین مایع‌های یونی مورد استفاده قرار گیرد.

تاریخ دریافت: ۱۳۹۷/۴/۸ ؛ تاریخ پذیرش: ۱۳۹۷/۸/۲۸

### مراجع

- [1] Earle M.J., Seddon K.R., [Ionic Liquids. Green Solvents for the Future](#), *Pure Appl. Chem.*, **72**: 1391-1398 (2000).
- [2] Keskin S., Kayrak-Talay D., Akman U., Hortaçsu Ö., [A Review of Ionic Liquids Towards Supercritical Fluid Applications](#), *J. Supercrit. Fluids.*, **43**: 150–180 (2007).
- [3] Gardas R.L., Freire M.G., Carvalho P.J., Marrucho I.M., Fonseca I.M.A., Ferreira A.G.M., Coutinho J.A.P., [A Review of Ionic Liquids Towards Supercritical Fluid Applications](#), *J. Chem. Eng. Data.*, **52**: 80–88 (2007).
- [4] Seki S., Kobayashi T., Kobayashi Y., Takei K., Miyashiro H., Hayamizu K., Tsuzuki S., Mitsugi T., Umabayashi Y., [Effects of Cation and Anion on Physical Properties of Room-Temperature Ionic Liquids](#), *J. Mol. Liq.*, **152**: 9–13(2010).
- [5] Zhang S., Sun N., He X., Lu X., Zhang X., [Physical Properties of Ionic Liquids: Database and Evaluation](#), *J. Phys. Chem. Ref. Data*, **35**: 1475–1517 (2006).
- [6] Wang X., Chi Y., Mu T., [A Review on the Transport Properties of Ionic Liquids](#), *J. Mol. Liq.*, **193**: 262–266 (2014).
- [7] Plechkova N.V., Seddon K.R., [Applications of Ionic Liquids in the Chemical Industry](#), *Chem. Soc. Rev.*, **37**: 123–150 (2008).
- [8] Katritzky A. R., Lomaka A., Petruskhin R., Jain R., Karelson M., Visser A. E., Rogers R. D., [QSPR Correlation of the Melting Point for Pyridinium Bromides, Potential Ionic Liquids](#), *J. Chem. Inf. Comput. Sci.*, **42**: 71-74 (2002).

- [9] Trohalaki S., Pachter R., Drake G. W., Hawkins T., [Quantitative Structure-Property Relationships for Melting Points and Densities of Ionic Liquids](#), *Energy Fuels*, **19**: 279-284 (2005).
- [10] Ren Y., Qin J., Liu H., Yao X., Liu M., [QSPR Study on the Melting Points of a Diverse Set of Potential Ionic Liquids by Projection Pursuit Regression](#), *QSAR Comb. Sci.*, **28**: 1237-1244 (2009).
- [11] <https://ilthermo.boulder.nist.gov> (accessed 2017)
- [12] Hanwell M. D., Curtis D. E., Lonie D. C., Vandermeersch T., Zurek E., Geoffrey R., [Hutchison Avogadro: an Advanced Semantic Chemical Editor, Visualization, and Analysis Platform.](#), *J. Chem.*, **4**: 17-23 (2012).
- [13] Yap C. W., [PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints](#), *J. Comput. Chem.*, **32**: 1466-1474(2011).
- [14] Mauri A., Consonni V., Pavan M., Todeschini, R., [Dragon Software: an Easy Approach to Molecular Descriptor Calculations](#), *Match*, **56**: 237-248 (2006.).
- [15] Kennard R.W., Stone L.A., [Computer-Aided Design of Experiments](#), *Technomet.*, **11**: 137-148 (1969).
- [16] <http://accelrys.com/products/collaborative-science/biovia-materials-studio/>
- [17] Todeschini R., Consonni V., ["Handbook of Molecular Descriptors"](#), WILEY-VCH Verlag GmbH. (2000).
- [18] Haupt R.L., Haupt E., ["Practical Genetic Algorithms"](#), 2 nd ed., John Wiley & Sons, Inc. (2004).