

تعیین عدد تجمع سورفکتانت‌های آنیونی بر اساس تکنیک هدایت سنجی: به کارگیری

تکنیک مدل‌سازی QSAR-ANN برای پیش‌بینی عدد تجمع سورفکتانت‌ها

بهناز عبدوس، سیده مریم سجادی*، احمد باقری

گروه شیمی، دانشکده شیمی، دانشگاه سمنان، سمنان، ایران

تاریخ دریافت: ۰۰/۱۰/۲۴

تاریخ تصحیح: ۰۱/۰۲/۲۰

تاریخ پذیرش: ۰۱/۰۳/۰۸

چکیده

در این مطالعه، از یک روش مدل‌سازی جدید با استفاده از مدل QSAR و شبکه عصبی مصنوعی برای پیش‌بینی عدد تجمع برخی از سورفکتانت‌های آنیونی در محلول آبی در دمای ۲۵ درجه سانتی‌گراد استفاده شده است. عدد تجمع مایسل با استفاده از اندازه‌گیری‌های هدایت الکتریکی و روش اوانس برای سورفکتانت‌های آنیونی در محلول‌های آبی تعیین شد. اما نتایج به دست آمده با استفاده از این روش، نتایج حاصل از روش فلورسانس مطابقت خوبی نداشت از آنجائی که روش فلورسانس روش دقیق‌تری برای محاسبه عدد تجمع مایسل‌ها می‌باشد به همین دلیل از نتایج روش فلورسانس در این مطالعه استفاده شد. به منظور ارتباط ساختار مولکولی این سورفکتانت‌ها با عدد تجمع آن‌ها، مطالعه ارتباط کمی ساختار-خاصیت (QSPR) انجام شد. یک مدل شبکه عصبی مصنوعی (ANN) برای پیش‌بینی عدد تجمع سورفکتانت‌های آنیونی با استفاده از چهار مورد از بیش از ۳۲۰۰ توصیف‌گر مولکولی، محاسبه شده توسط نرم‌افزار Dragon، به عنوان متغیرهای ورودی، توسعه داده شد. اهمیت توصیف‌گرهای انتخابی بر اساس روش ANN محاسبه شدند که ترتیب اهمیت آنها بدین صورت می‌باشد: $MWC04 > MWC05 > X5V > nC$. مجموعه کامل ۲۴ سورفکتانت آنیونی به صورت تصادفی به یک مجموعه آموزشی ۱۶ تایی، یک مجموعه آزمایشی ۴ تایی و یک مجموعه اعتبارسنجی ۴ تایی تقسیم شدند. همچنین از تحلیل رگرسیون خطی چندگانه (MLR) برای ساخت یک مدل خطی با استفاده از توصیف‌گرهای مشابه استفاده شد. ضریب همبستگی (R^2) و ریشه میانگین مربعات خطا (RMSE) مدل‌های ANN و MLR (برای کل مجموعه داده‌ها) به ترتیب ۰/۹۴، ۴/۹۹ و ۰/۸۲، ۸/۳۸ بود. R^2 بالاتر روش ANN نشان داد که رابطه بین توصیف‌گرها و عدد تجمع ترکیبات، غیرخطی است.

کلمات کلیدی: توصیف‌گرهای مولکولی، QSAR، شبکه عصبی مصنوعی، سورفکتانت آنیونی، عدد تجمع.

۱- مقدمه

مواد فعال سطحی (سورفکتانت) به صورت گسترده در ساخت مواد شوینده، آرایشی، دارویی، صنعت نساجی و... کاربرد دارند. این مواد دارای یک گروه آبدوست (سر قطبی) و یک گروه آبگریز (زنجیره هیدروکربنی) می‌باشند این ساختار آمفی‌فیلیکی (دو گانه)، مولکولی با خواص مناسب برای فعالیت سطحی ایجاد می‌کند [۱].

در یک ظرف حاوی مقدار مشخص از آب با افزایش غلظت ماده فعال سطحی به ترتیب مراحل زیر رخ می‌دهند:

۱. در غلظت‌های پایین، مونومرهای مواد فعال سطحی به طور تصادفی در محلول توزیع، که این امر باعث افزایش انرژی آزاد محلول می‌شود. بدین ترتیب برای کاهش انرژی و همچنین به علت خصلت آبدوستی-آبگریزی، این مولکول‌ها در سطح آب قرار می‌گیرند، به طوری که گروه‌های سر با مولکول‌های آب و گروه‌های دم با مولکول‌های هوا در ارتباط هستند. این فرآیند باعث کاهش کشش سطحی آب می‌گردد که علت آن از بین رفتن شبکه مولکول‌های آب می‌باشد که توسط پیوند هیدروژنی ایجاد شده بود.

۲. این فرآیند آنقدر ادامه می‌یابد تا سطح اشباع شود. در این صورت یک شبکه تک لایه‌ای از مولکول‌های مواد فعال سطحی تشکیل می‌شود. پس از اشباع سطح آب، برهم کنش‌های آب دوستی-آب‌گریزی، این بار در داخل توده‌ی آب، باعث تشکیل تجمعاتی از قبیل مایسل، وسیکل، دولایه‌ای، استوانه‌ای، اسفنجی و ... می‌گردند. غلظتی که مایسل در آن تشکیل می‌شود از ویژگی‌های انحصاری هر ماده فعال سطحی بوده و غلظت بحرانی مایسل یا به اختصار CMC نامیده می‌شود [۲]. می‌توان تشکیل مایسل در محلول را حاصل رقابت بین دو نیروی برهمکنش آبگریزی گروه‌های دم و دافعه الکتروستاتیکی گروه‌های سر دانست که اولی در نقش یک نیروی پیش برنده تشکیل تجمع مایسلی و دومی در نقش محدود کننده اندازه مایسل ظاهر می‌شوند. در نقطه‌ی CMC برهمکنش‌های بین اجزاء به کمترین مقدار خود می‌رسد [۳].

۱-۲- عدد تجمع^۱

به طور متوسط هر مایسل حاوی تعداد مشخصی مولکول سورفکتانت است که به آن عدد تجمع (N) می‌گویند. تغییر در دما، غلظت سورفکتانت، قدرت یونی محلول و گروه‌های ساختاری در سورفکتانت ممکن است باعث تغییر در اندازه، شکل و عدد تجمع مایسل شود [۴]. معمولاً هرچقدر CMC سورفکتانت کمتر باشد عدد تجمع آن بالاتر می‌باشد و عملکرد سورفکتانتی مطلوب تر است. تعیین عدد تجمع (N) سورفکتانت‌ها اولاً می‌تواند از دیدگاه مطالعاتی مفید باشد زیرا با تعیین N بسیاری از خصوصیات یک سورفکتانت از قبیل ویژگی‌های ساختاری مولکول، شکل مایسل تشکیل شده (نرمال یا معکوس) و ... را می‌توان پیش بینی کرد و از دیدگاه کاربردی هم حایز اهمیت است زیرا هر چقدر عدد تجمع مایسلی بالاتر باشد نشان دهنده این است که مایسل‌های تشکیل شده قابلیت بالاتری برای انتقال جرم به عنوان حامل دارو (در فرایند رهایش دارو) یا حذف آلاینده‌های هیدروکربنی (در فرایند تصفیه آب یا پساب) دارا می‌باشند.

¹ Aggregation number

تعیین پارامترهای مایسلی از قبیل مقدار CMC، عدد تجمع، خواص ترمودینامیکی و ... سورفکتانت‌ها یا مواد آمی فیلیکی (مانند دارو، مایع یونی فعال سطحی یا ...) همواره یکی از چالش‌های بزرگ محققان در کار با این مواد است. که تعدادی از این مشکلات در بخش ذیل بیان شده است:

الف) مواد سورفکتانتی که طول زنجیره هیدروکربنی کوتاه دارند یا بخش هیدروفوب کوچک دارند مقادیر CMC بالا دارند و دارای عدد تجمع پایین هستند (مانند داروها، مایعات یونی و ...). در نتیجه از لحاظ اقتصادی تعیین پارامترهای مایسلی این مواد فوق العاده هزینه بر می باشد.

ب) پارامترهای مایسلی مواد آمی فیلیکی جدیدی که سنتز می شوند نامشخص است از این رو پیش بینی این خواص در تعیین دقیق این پارامترها می تواند بسیار مفید باشد.

ج) تنها تکنیک معتبر در تعیین عدد تجمع سورفکتانت‌ها و موادی با خصوصیات سورفکتانتی (دارو آمی فیلیک، مایع یونی فعال سطحی و ...) تکنیک فلورسانس با بکار بردن یک پروب مولکولی^۲ می باشد که بعلت گرانی و عدم دسترس بودن دستگاه مورد نظر همواره یکی از چالش‌های تعیین عدد تجمع این مواد می باشد.

۱-۳- طیف‌سنجی فلورسانس

همان‌طور که بیان شد، رفتار تجمع سورفکتانت‌ها در محلول آبی را می‌توان با عدد تجمع میسل (N) تعیین کرد، که از نتایج طیف‌سنجی فلورسانس با پیرن به عنوان پروب به دست می‌آید. محلول‌ها برای تعیین اعداد تجمع مایسلی به شرح زیر تهیه می‌شوند. از یک محلول مادر ۰٫۱ میلی مولار پیرن در اتانول مطلق، یک حجم مشخص به یک فلاسک حجمی انتقال داده می‌شد. اتانول تبخیر شده و آب مقطر به آن اضافه و محلول به مدت یک شب هم زده می‌شود. از محلول آبی پیرن، محلول‌های استوک سورفکتانت با غلظت سورفکتانت بسیار بالاتر از CMC مربوطه تهیه خواهد شد.

خاموش‌کننده‌ها به طور مشابه در اتانول مطلق حل می‌شوند. از این محلول‌ها، محلول‌های استوک خاموش‌کننده با اضافه کردن حجم مشخصی از محلول خاموش‌کننده اتانولی در یک فلاسک حجمی، تبخیر اتانول، و حل کردن خاموش‌کننده در محلول استوک سورفکتانت/پیرن تهیه می‌شود. غلظت خاموش‌کننده در این محلول‌ها برابر با حداکثر غلظت خاموش‌کننده مخلوط کردن استوک سورفکتانت/پیرن بدون خاموش‌کننده و استوک سورفکتانت/پیرن با خاموش‌کننده در نسبت‌های مشخص، پنج یا شش محلول با غلظت خاموش‌کننده از صفر تا حداکثر غلظت متفاوت تهیه می‌شود.

² Molecular probe

طبق معادله ۱، طیف نشر این محلول‌ها ثابت شد و لگاریتم نسبت شدت I_0/I_Q در یک طول موج خاص در محدوده نشر طیفی در برابر غلظت خاموش‌کننده رسم شد. این نمودار باید یک خط مستقیم از مبدا با شیب برابر با $\frac{1}{[mic]}$ ایجاد کند. ضرب شیب در غلظت مولکول‌های سورفکتانت شرکت‌کننده در تشکیل میسل (یعنی $[S_{tot}] - cmc$) عدد تجمع را مطابق معادله ۲ به دست می‌دهد [۵].

$$\ln\left(\frac{I_0}{I_Q}\right) = \frac{[Q_{mic}]}{[mic]} \quad (1)$$

$$N = \frac{[S_{tot}] - cmc}{[mic]} \quad (2)$$

از این رو با توجه به مشکلات ذکر شده، روش‌های محاسباتی کمومتریکس می‌توانند در پیش‌بینی این پارامترها برای مواد آملی فیلیکی و ترکیبات جدید در این طبقه بسیار مفید باشند. کمومتریکس شاخه‌ای از شیمی بر پایه قوانین ریاضی و آمار است که روش اندازه‌گیری و آزمایش را انتخاب و یا طراحی می‌کند و یا با آنالیز داده‌های شیمیایی بیشترین اطلاعات مفید را استخراج می‌کند [۶]. با توجه به هدف محقق می‌توان از روش‌های کمومتریکس به شکل‌های توصیفی و یا پیشگویانه از آن استفاده کرد. در شکل توصیفی ویژگی شیمیایی مدل‌سازی برای شناسایی ساختار نهان در اطلاعات و رابطه بین نمونه‌ها، انجام می‌شود. در شکل پیشگویانه مدل‌سازی برای شناسایی ساختار، رفتار و خواص نمونه‌های جدید ایجاد می‌شود [۷].

در روش‌های کمومتریکسی QSAR/QSPR، مرحله‌ای با عناوین جمع‌آوری اطلاعات، محاسبه‌ی توصیف‌گرهای مولکولی با توجه به ساختار شیمیایی آن‌ها، گزینش توصیف‌گرها، ایجاد مدل و ارزیابی درونی و بیرونی مدل ایجاد شده وجود دارد. مدل‌های QSAR ارایه شده با نرم‌افزار درآگون امکان محاسبه دقیق خواص فیزیکی، شیمیایی، زیستی و سم‌شناسی ترکیبات آلی و... با استفاده از توصیف‌کننده‌های ساده، الکتروستاتیکی، توپولوژیکی، کوانتومی را فراهم می‌کند تا رابطه بین ساختار و فعالیت مولکول‌ها برقرار شود [۸]. داده‌های حاصل از مطالعه QSAR توسط روش خطی MLR و روش غیر خطی ANN آنالیز می‌شوند. ANN عموماً سیستم‌های ریاضی یادگیری غیرخطی هستند و به مجموعه داده‌های ورودی شناخته شده بدون هیچ فرضی نیاز دارند [۹]. ANN یک نقشه برداری از متغیرهای ورودی و خروجی ایجاد می‌کند که می‌تواند پس از آن برای پیش‌بینی خروجی مجهول به عنوان تابعی از ورودی‌های مناسب استفاده شود [۱۰].

برای نخستین بار در مطالعه حاضر، ترکیب روش QSAR و الگوریتم غیرخطی ANN برای پیش‌بینی عدد تجمع ترکیبات سورفکتانت آنیونی پیشنهاد شده است. توصیف‌گرهای مولکولی با استفاده از نرم‌افزار Dragon محاسبه شدند و عدد تجمع تجربی مولکول‌ها از کار گزارش شده توسط نویسندگان مختلف [۳۰-۲۳] به دست آمد. در مطالعه حاضر، یک رویکرد ANN را برای توسعه یک مدل QSPR برای بهبود R^2 و به دست آوردن پیش‌بینی قابل اعتماد عدد تجمع پیشنهاد شده است. علاوه بر این، سیستم برای مقایسه با روش MLR مورد تجزیه و تحلیل قرار گرفت. در نهایت نشان داده شده است که

مدل غیرخطی ANN می تواند ارتباط رضایت بخشی بین توصیف‌گرهای مولکولی و عدد تجمع ایجاد کند.

۲- نرم افزارهای مورد استفاده

در این کار برای رسم ساختار هر ترکیب (جدول ۲) از نرم افزار GaussView 5.0.8 استفاده شد [۱۱] که بهینه سازی ساختارها بر اساس روش نیمه تجربی PM6 انجام شد. نرم افزار Dragon 5.5-2007 برای محاسبه توصیف‌گرهای مولکولی ساختارهای بهینه شده در مرحله قبل به کار گرفته شد. [۱۲]. تمامی محاسبات در نرم افزار Matlab R2020b انجام شد [۱۳]. ANN-Toolbox که یک جعبه ابزار در نرم افزار Matlab می باشد برای انجام محاسبات مربوط به شبکه عصبی استفاده شد.

۳- تئوری

۳-۱- مدل سازی به روش گام به گام

برای آنالیز در روش رگرسیون خطی چندگانه با روش گام به گام^۳ همه متغیرهای مستقل برای آنالیز وارد می شوند و در طی انجام محاسبات متغیرهای مستقلی که تاثیر محسوسی در ارتباط با متغیرهای وابسته نداشته باشند را از تحلیل حذف می کند و آن‌ها به معادله وارد نمی شوند. در این قسمت قبل از آغاز طراحی مدل، از طریق نرم افزار MATLAB اطلاعات به دو دسته آزمون و آموزش دسته بندی می شوند. فعالیت‌ها و خاصیت‌هایی که به روش تجربی حاصل شده اند بعنوان متغیر مستقل و مقدارهای مرتبط با توصیف‌گرها به عنوان متغیر وابسته در نظر گرفته می شوند. سپس اطلاعات مورد نظر وارد می شوند و نوع روش محاسباتی روش گام به گام انتخاب می شود.

۳-۲- مدل سازی غیر خطی شبکه عصبی مصنوعی

شبکه‌ی عصبی مصنوعی (ANN) برنامه‌های رایانه‌ای هستند که با توپولوژی، ویژگی‌های محاسباتی عناصر و قوانین آموزشی تعریف می شوند که به طور گسترده در مطالعات QSAR/QSPR استفاده می شوند [۱۴، ۱۵]. آن‌ها شامل مجموعه ای از لایه ها هستند که در آن اطلاعات از طریق لایه ها منتقل می شود. لایه اول لایه ورودی نامیده می شود و هر متغیر خارجی به هر نورون خود تغذیه می شود. آخرین لایه، لایه خروجی است که نورون های آن متغیرهای خروجی از شبکه را کنترل می کنند. لایه‌های نورون‌های بین لایه‌های ورودی و خروجی، لایه‌های پنهان نامیده می شوند که هر کدام ممکن است محاسبات را به طور مستقل انجام دهند و نتایج را به لایه دیگری منتقل کنند. دو مرحله اصلی در فرآیند ANN وجود دارد: یادگیری و اعتبار سنجی. در یادگیری یا آموزش، وزن‌ها در پاسخ به اطلاعات ورودی اصلاح می شوند. روش‌های مختلف یادگیری مانند

³ Stepwise

BFGS [۱۴، ۱۶، ۱۷]، انتشار برگشتی [۱۸]، پیش‌خور [۱۹]، انتشار به عقب‌المان [۲۰]، تأخیر زمانی پیش‌خور [۲۱]، هاپفیلد [۲۲] و غیره برای اندازه‌گیری توانایی مدل‌ها برای تولید پاسخ درست وجود دارد.

۴-بخش تجربی

۴-۱-مواد

در این تحقیق از سورفکتانت‌های آنیونی سدیم دو دسیل سولفات، لیتیم دو دسیل سولفات و سدیم N-لوریل سارکوسین از شرکت آلمانی Merck با خلوص بیش از ۹۹٪ استفاده شده است. همچنین در این کار جهت محلول‌سازی از آب فوق‌خالص با مقاومت 18.5 MΩ.cm استفاده شده است که برای تهیه آن از یک دستگاه تصفیه آب Direct-Q® 5 UV ساخت کشور فرانسه استفاده گردید.

۴-۲-تجهیزات و دستگاه

همه اندازه‌گیری‌های هدایت سنجی به کمک یک دستگاه هدایت سنج دیجیتال (مدل Consort, C 860، ساخت کشور بلژیک) و یک سل غوطه‌ور با الکتروود پلاتین (با ثابت سل 1cm^{-1}) انجام گرفته است. قبل از هر اندازه‌گیری دستگاه با محلول KCl با غلظت 0.01 مولار کالیبره شد که این محلول دارای هدایت الکتریکی $1413\mu\text{S}\cdot\text{cm}^{-1}$ در دمای 298 K می‌باشد. به منظور ثابت نگه داشتن دما در کل آزمایش‌ها از یک دستگاه سیرکولاتور EYELA NTT-1100 ساخت کشور ژاپن با دقت $\pm 0.1\text{ K}$ استفاده شد. برای برداشتن حجم‌های مورد نیاز از محلول‌های غلیظ تهیه شده، از پیپت‌های میکرولیتری ساخت شرکت آلمانی ایپندورف^۴ در محدوده حجمی ۱۰-۱۰۰ میکرولیتری استفاده شده است.

۴-۳-روش انجام آزمایش

هدایت سنجی ساده‌ترین روش برای مطالعه روی محلول‌های الکترولیت است و همچنان در تعیین مقدار CMC سورفکتانت‌های یونی در حلال قطبی از اهمیت بالایی برخوردار است.

هدایت الکتریکی یک محلول، عکس مقاومت الکتریکی آن است. واحد هدایت ohm^{-1} می‌باشد که در بسیاری از موارد آن را زیمنس (S) می‌نامند. مقدار هدایتی که از طریق دستگاه هدایت سنجی به دست می‌آید، هدایت ویژه است که از طریق رابطه زیر محاسبه می‌شود:

$$k = \frac{l}{A} \frac{1}{R} \quad (3)$$

در این رابطه k هدایت ویژه، R مقاومت محلول، l فاصله بین دو الکتروود و A مساحت سطح الکتروودها است و به نسبت l/A اصطلاحاً ثابت سل می‌گویند. واحد هدایت الکتریکی ویژه (زیمنس بر سانتی متر یا $\text{S}\cdot\text{cm}^{-1}$) است ولی به علت کوچک بودن

⁴ Eppendorf

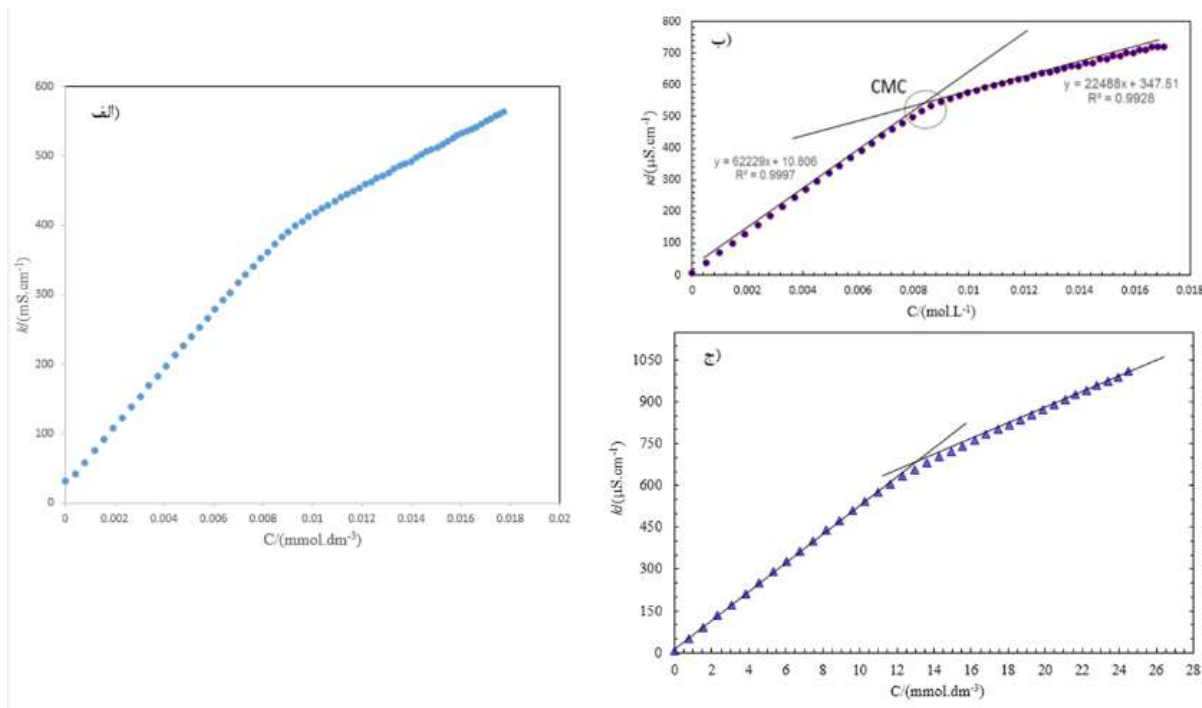
اعداد هدایت الکتریکی آن را بر حسب میکرو (میلی) زیمنس بر سانتیمتر هم بیان می‌کنند.

اندازه‌گیری‌های هدایت سنجی به روش زیر صورت گرفت:

ابتدا یک محلول غلیظ از سورفکتانت مورد مطالعه در حالت خالص (SDS) با غلظت مشخص تهیه شد و به تدریج توسط میکروپیپت ایندورف حجم‌های مشخص از محلول به ۱۰ میلی‌لیتر آب دیونیزه اضافه گردید و هر بار پس از همگن شدن به وسیله مگنت مغناطیسی هدایت الکتریکی اندازه‌گیری شد (تیترا سیون هدایت سنجی). در نهایت، نمودارهای هدایت ویژه بر حسب غلظت سورفکتانت (حالت خالص یا مخلوط) به منظور تخمین CMC رسم شد. لازم به ذکر است همه اندازه‌گیری‌ها در دمای 298K انجام شد.

در نمودار هدایت ویژه بر حسب غلظت سورفکتانت یک تغییر ناگهانی در شیب نمودار اتفاق می‌افتد به غلظتی که این تغییر شیب بوجود می‌آید اصطلاحاً غلظت بحرانی مایسلی یا CMC می‌نامند. در این غلظت خاص اولین تجمع مایسلی در محلول شکل می‌گیرد، تعیین غلظت بحرانی مایسلی (CMC) با استفاده از اندازه‌گیری‌های هدایت الکتریکی روش بسیار مناسب و دقیق برای سورفکتانت‌های یونی در محلول‌های آبی است.

برای تعیین غلظت بحرانی مایسلی سورفکتانت SDS همان‌طور که در شکل ۱ نشان داده شده است به علت این که نمودار هدایت الکتریکی بر حسب غلظت این سورفکتانت نقطه شکست بسیار خوبی دارد تنها با استفاده قطع دادن معادله خط قبل و بعد از نقطه‌ی شکست نمودار میتوان CMC را تعیین کرد. توجه این پدیده به این علت است که قبل از تشکیل مایسل سورفکتانت‌های SDS به صورت مونومرهای آزاد (آنیون) و یون مخالف (کاتیون) هستند که در محلول آبی دارای هدایت الکتریکی بالایی هستند (مخصوصاً یون‌های مخالف) ولی پس از تشکیل مایسل تجمع بزرگی از سورفکتانت‌ها تشکیل می‌شود و علاوه بر این تعدادی از یون‌های مخالف به سطح مایسل اتصال می‌یابند (نه همه آنها) که این موارد (تجمعات بسیار بزرگ و کاهش یونهای مخالف آزاد) سبب کاهش هدایت الکتریکی محلول می‌شود و تغییر شیب نمودار هدایت بر حسب غلظت (در جهت کاهش) می‌شود [۲۳].



شکل ۱) تعیین CMC با استفاده از معادله خط قبل و بعد از شکست نمودار هدایت الکتریکی بر حسب غلظت الف) LDS، ب) SDS و ج) NLS به ترتیب.

بر اساس اندازه‌گیری‌های هدایت سنجی، CMC ترکیبات ذکر شده در جدول ۱ گزارش شده است.

جدول ۱. فرمول مولکولی SDS و LDS و NLS و مقادیر CMC تجربی بدست آمده

فرمول مولکولی	سورفکتانت	نام مختصر	CMC	S ₁	S ₂	β	Agg. No. (con.)	Agg. No. (Flo.)
C ₂₄ H ₃₉ O ₅ Na	Lithium dodecyl sulfate	LDS	10.1 mmol.dm ⁻³	0.040	0.018	0.47	44	60
C ₁₂ H ₂₅ OSO ₃ Na	Sodium dodecyl sulfate	SDS	8.3 mmol.dm ⁻³	0.080	0.029	0.37	40	62
C ₁₅ H ₃₀ NNaO ₂	N-lauryl sarcosine sodium	NLS	13.5 mmol.dm ⁻³	0.050	0.030	0.59	15	17

۴-۴- محاسبه عدد تجمع با استفاده از داده‌های هدایت سنجی با به‌کاربردن مدل اوانس^۵

همانطور که گفته شد، از داده‌های هدایت سنجی، غلظت بحرانی میسل و درجه تفکیک میسل به دست خواهد آمد. درجه تفکیک یون مخالف مایسل، β، را می‌توان از نسبت بین شیب‌های بالای نقطه cmc (S₂)، و زیر نقطه cmc (S₁)، تخمین زد. بنابراین می‌توان عدد تجمع میسل را توسط معادله اوانس [۲۴] تعیین نمود.

$$(1000S_1 - \lambda_C) \left(\frac{N^* - m}{N^{*4/3}} \right)^2 + \lambda_C(1 - \beta) - 1000S_2 = 0 \quad (۴)$$

⁵ evans

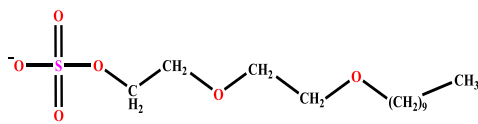
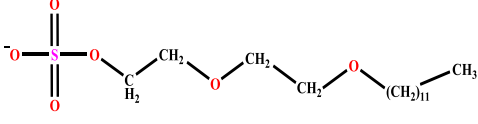
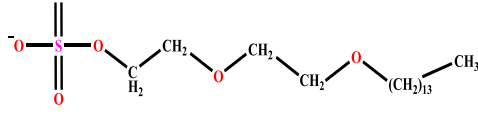
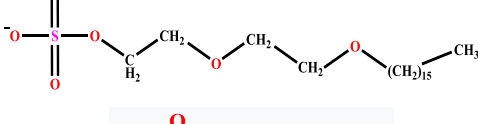
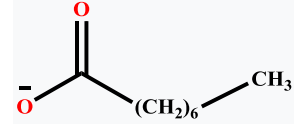
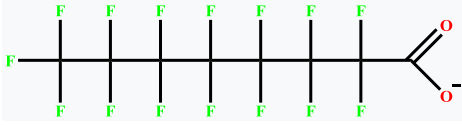
در این رابطه همانطور که گفته شد S_1 و S_2 شیب های پایین و بالا نقطه cmc ، N^* عدد تجمع میسل، m تعداد یون های مخالف تفکیک شده و λC هدایت یونی مولی یون مخالف است. با استفاده از رابطه ۴، عدد تجمع میسل محاسبه شده است که نتایج آن در جدول ۱ گزارش شده است.

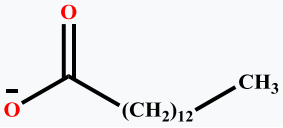
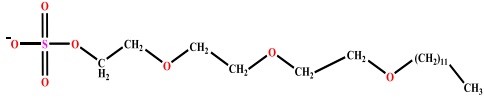
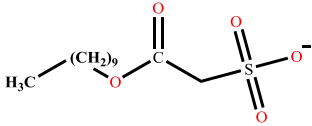
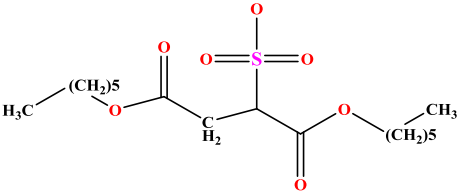
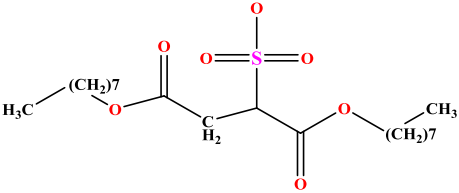
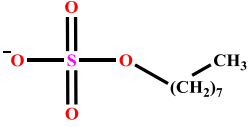
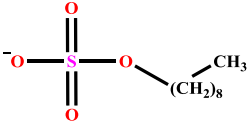
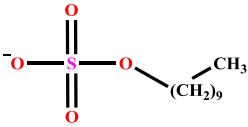
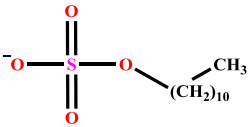
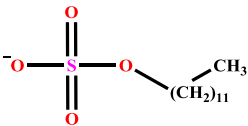
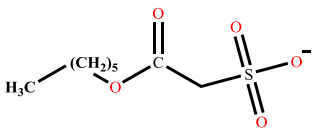
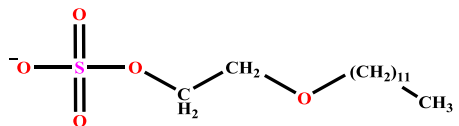
همانطور که مشاهده شد نتایج به دست آمده با استفاده از این روش، از مقدار واقعی گزارش شده با استفاده از روش فلورسانس، انحراف زیادی دارد. از آنجائیکه روش فلورسانس روش دقیق تری برای محاسبه عدد تجمع میسل ها می باشد ولی با مشکلاتی مانند گرانی و عدم در دسترس بودن دستگاه مورد نظر همراه است، از این رو با توجه به مشکلات ذکر شده، روش های محاسباتی کمومتریکس می توانند در پیش بینی این پارامترها برای مواد آملی فیلپکی و ترکیبات جدید در این طبقه بسیار مفید باشند.

۲-۴- مجموعه داده ها

در این مطالعه، ۲۴ نوع سورفکتانت آنیونی به عنوان ورودی توسط Dragon مورد بررسی قرار گرفتند. اسامی این ترکیبات به همراه مقادیر تجربی عدد تجمع سورفکتانت ها در جدول ۲ آمده است. عدد تجمع این مولکول ها از رفرنس های [۲۵-۳۲] گرفته شده است. این مقادیر در محلول های آبی در غلظت های مختلف با استفاده از روشی مبتنی بر تجزیه و تحلیل منحنی های فروپاشی فلورسانس پیرن حل شده با میسل تعیین شد.

جدول ۲. مقادیر تجربی عدد تجمع برای سورفکتانت آنیونی

شماره مولکول.	نماد	ساختار مولکول	عدد تجمع (ANN)	عدد تجمع (تجربی)	مجموعه داده	مرجع.
۱	$C_{10}E_2SO_4Na$		۱۳	۱۱	آموزش	۲۵
۲	$C_{12}E_2SO_4Na$		۲۱	۱۳	آموزش	۲۵
۳	$C_{14}E_2SO_4Na$		۲۰	۱۹	آموزش	۲۵
۴	$C_{16}E_2SO_4Na$		۳۰	۲۶	آموزش	۲۵
۵	C_7CO_2K		۱۵	$1/2 \pm 1/4$ ۴	آموزش	۲۶
۶	PfC_7CO_2K		۱۱	$1/3 \pm 1/8$ ۳	آموزش	۲۶

۷	KMy		۸۱	۸۲	تست	۲۷
۸	SLE3S		۴۵	۴۲	ارزیابی	۲۸
۹	sodium decyl sulfoacetates		۶۷	۷۰	آموزش	۳۰
۱۰	sodium dihexyl sulfosuccinates		۳۴	۳۸	آموزش	۳۰
۱۱	sodium dioctyl sulfosuccinates		۵۷	۵۶	آموزش	۳۰
۱۲	SC ₈ S		۵۰	۳۸	آموزش	۳۱
۱۳	SC ₉ S		۲۴	۳۵	تست	۳۱
۱۴	SC ₁₀ S		۴۷	۴۵	آموزش	۲۹
۱۵	SC ₁₁ S		۵۰	۵۲	ارزیابی	۳۱
۱۶	SC ₁₂ S		۵۳	۵۲	آموزش	۳۱
۱۷	sodium hexyl sulfoacetates		۱۷	۲۰	ارزیابی	۳۰
۱۸	C ₁₂ E ₁ SO ₄ Na		۲۱	۱۹	تست	۳۱

۱۹	Sodium octane sulfonate		۳۰	۲۵/۴	آموزش	۳۱
۲۰	sodium octyl sulfoacetates		۴۱	۴۲	آموزش	۳۰
۲۱	TBADS		۶۱	۵۷	آموزش	۳۲
۲۲	TEADS		۵۹	۶۳	تست	۳۲
۲۳	TMADS		۶۱	۶۴	آموزش	۳۲
۲۴	TPADS		۵۵	۵۷	ارزیابی	۳۲

۳-۴- رسم مولکول، بهینه سازی و محاسبه توصیف‌گرها

ساختار نیمی از مولکول‌ها در نرم افزار GaussView 5.0 رسم و بقیه مولکول‌ها از سایت chemspider استخراج شده اند، و با پسوند chk ذخیره شده اند. این ساختارها با روش semi empirical، بهینه شدند. پس از انجام بهینه سازی مولکول‌ها، ساختار بهینه شده هر یک از این ترکیبات به نرم افزار Dragon منتقل شده و برای هر کدام از این مولکول‌ها ۳۲۲۴ توصیف‌گر محاسبه گردید.

۴-۴- انتخاب بهترین توصیف‌گرها با روش گام‌به‌گام

توصیف‌گرها پارامترهای عددی هستند که ویژگی‌های متفاوتی از مولکول‌ها را بیان می‌کنند. به عبارت دیگر، هریک از این توصیف‌گرها، اطلاعات خاصی از مولکول‌ها را در اختیار قرار می‌دهند. تعیین و انتخاب توصیف‌گرها از اهمیت به‌سزایی برخوردار است. در صورتی که توصیف‌گرهای انتخاب شده مناسب نباشند، عمل مدل‌سازی با مشکل رو به رو شده و مدل مناسب به دست نخواهد آمد [۳۶]. بنابراین در انتخاب توصیف‌گرها باید دقت کرد و توصیف‌گرهایی باید انتخاب شوند که در ارتباط با ویژگی‌های خاص مولکول باشند.

بعد از محاسبه ی توصیف‌گرهای مختلف، تعدادی از آن‌ها به عنوان توصیف‌گرهای مناسب که تاثیر بسیاری روی پارامتر مورد بررسی (عدد تجمع) دارند، برای ساخت مدل انتخاب می‌شوند. توصیف‌گرهای مولکولی نتیجه نهایی یک استدلال و روش ریاضی

هستند که اطلاعات شیمیایی را به کد تبدیل می کند و آن ها را به صورت یک نماد نشان می دهد که ارائه دهنده یک مولکول به صورت یک عدد مفید می باشد.

با روش رگرسیون گام به گام توصیف‌گرهایی که با پاسخ تجربی ارتباط بیشتری دارند و تاثیر بیشتری در مدل دارند انتخاب می شوند. این مرحله، شامل یافتن توصیف‌گرهای حاوی اطلاعات مفید است، به طوری که قدرت پیش بینی مدل در سطح قابل قبولی باشد. از بین توصیف‌گرها، ۴ توصیف‌گر به عنوان بهترین توصیف‌گرها برای ۲۴ ترکیب که بیشترین تاثیر را روی عدد تجمع در دمای 25°C داشتند، برای ساخت مدل به کار رفتند که اسامی آنها به همراه نماد، تعریف و دسته در جدول ۳ گزارش شده است.

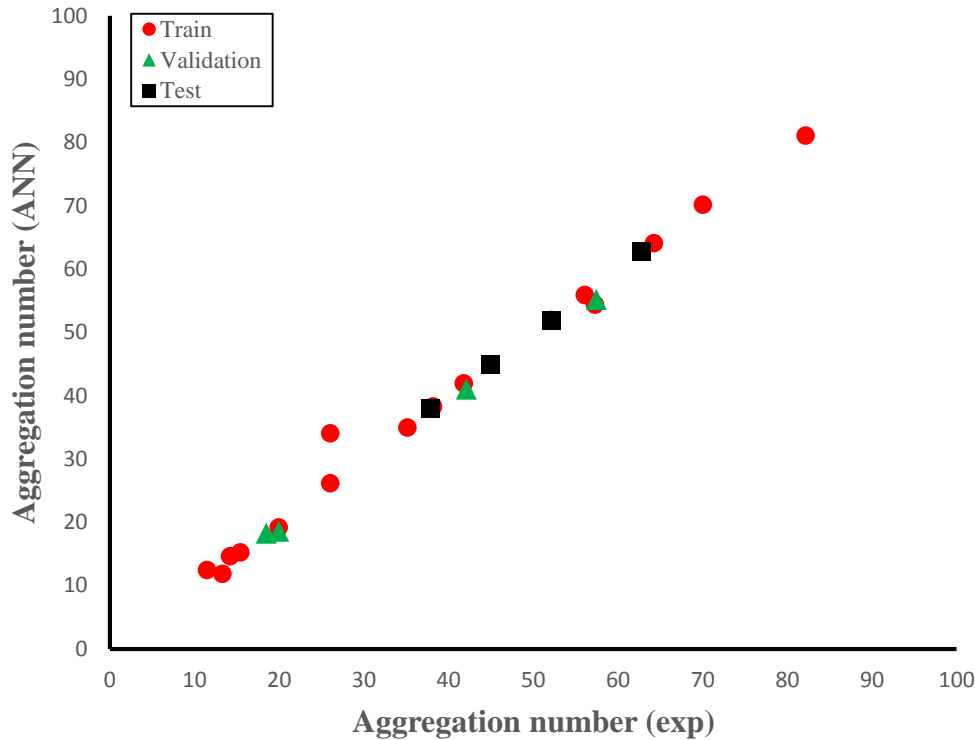
جدول ۳. توصیف‌گرهای ساختاری انتخاب شده برای تحلیل QSAR

ID	Name	Description	Block
۱	nC	number of Carbon atoms	Constitutional indices
۲	MWC04	molecular walk count of order 4	Walk and path counts
۳	X5V	valence connectivity index of order 5	Connectivity indices
۴	MWC05	molecular walk count of order 5	Walk and path counts

همانطور که گفته شد به دلیل جلوگیری از مشکل برازش بیش از حد^۶ تعداد ورودی‌های شبکه عصبی مصنوعی باید حداقل مقدار ممکن باشد به طوری که تعداد وزن‌های شبکه باید کمتر از تعداد نمونه‌ها در مجموعه آموزش باشد که در این صورت وزن ها دقیق محاسبه می شوند. در اینجا ابتدا شبکه عصبی بر روی کل داده ها اجرا شد که نتایج آن در شکل ۲ نمایش داده شده است. سپس به منظور جلوگیری از مشکل برازش بیش از حد تعداد ورودی‌های شبکه از ۱۸۲ توصیف‌گر به ۴ توصیف‌گر کاهش یافت و سپس شبکه عصبی مصنوعی بر روی داده‌ها اجرا گردید. در ابتدا ۲۴ سورفکتانت به صورت تصادفی به سه سری آموزش (۱۶ داده)، ارزیابی (۴ داده) و تست (۴ داده) تقسیم شدند. سپس پارامترهای موثر بر آموزش موفق شبکه عصبی بعد از انتشار از قبیل، تعداد گره‌های لایه پنهان، نوع تابع انتقال، نوع الگوریتم آموزش و تعداد چرخه‌های آموزش در مرحله آموزش شبکه بهینه شدند. به این صورت که شبکه، با تعداد نورون‌های ۱ تا ۵ و با توابع مختلف انتقال و الگوریتم‌های آموزش متفاوت اجرا شد و در نهایت آن نورون و تابع و الگوریتمی که حداقل مقدار MSE را نشان داد و پارامترهای آماری نظیر R^2 مربوط به مجموعه ها مختلف قابل قبول به دست آمد به عنوان شرایط بهینه انتخاب شد که در جدول ۴ گزارش شده است. در حین فرآیند آموزش، مقادیر وزن های ارتباطی تغییر می کند و معیار برای پایان دادن آموزش (بهینه شدن مقادیر وزن ها جهت پیش بینی عدد تجمع)، مقدار خطای شبکه است که طبق تابع کارایی تعیین شده، محاسبه می شود که در این روش، تابع کارایی، ریشه میانگین مربع خطا^۷ در سری ارزیابی می باشد. تعداد بهینه نورون ها در لایه های پنهان بر اساس حداقل $RMSE$ مجموعه آزمون

⁶ overfitting⁷ RMSE

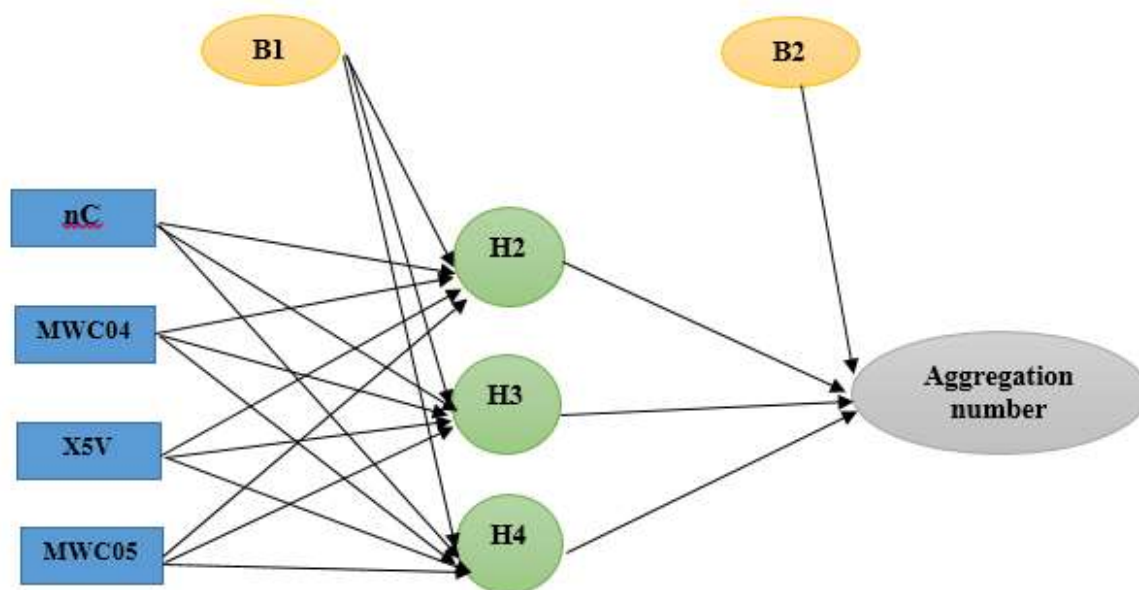
(۴/۸۴) عدد ۳ بود. شکل ۳ شماتیک $(4 \times 3 \times 1)$ ANN را نشان می دهد. بهترین توصیف‌گرها برای ساخت مدل به عنوان ورودی، به شبکه عصبی مصنوعی داده شدند و مقادیر عدد تجمع محاسبه شده توسط شبکه عصبی بهینه به دست آمدند که در جدول ۳ مشاهده می شود. نتایج پارامترهای آماری در بخش ۴-۵ گزارش شده است.



شکل ۲) نمودار پراکنندگی داده های توسعه یافته (هدف) برحسب مدل پیش بینی شده ANN (خروجی) با استفاده از ۱۸۲ دیسکریپتور

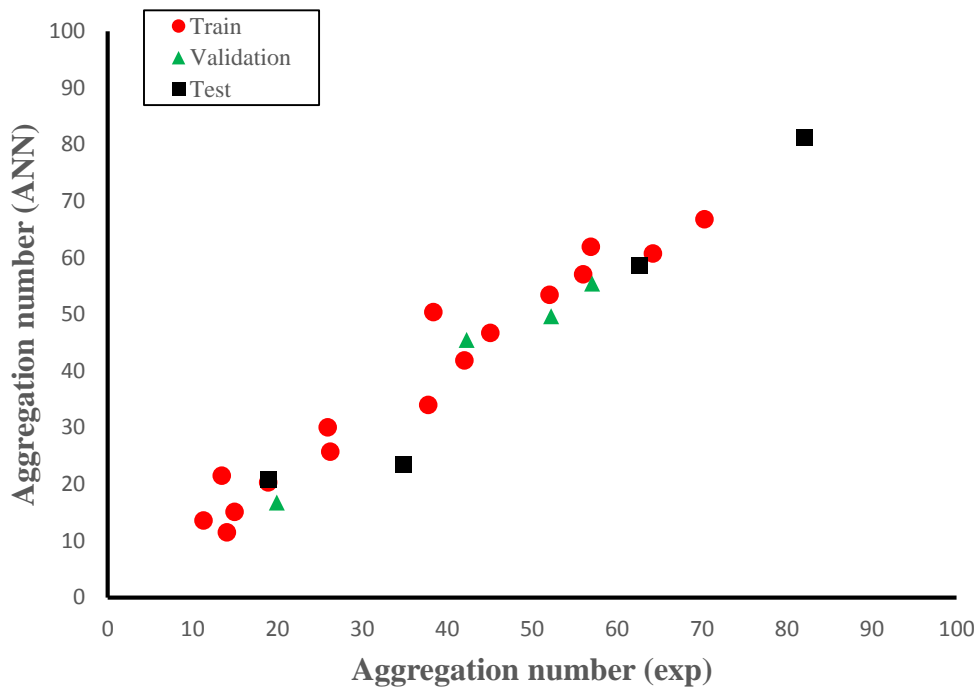
جدول ۴. مشخصات شبکه بهینه در جعبه ابزار matlab

چهار ورودی، یک خروجی و یک لایه پنهان با سه نورون $(4 \times 3 \times 1)$	وضعیت شبکه
مجموعه آموزش: ۷۰ درصد داده‌ها به طور تصادفی انتخاب شده (۱۶ داده)	مجموعه داده
مجموعه آزمون: ۱۵ درصد داده‌ها به طور تصادفی انتخاب شده (۴ داده)	
مجموعه ارزیابی: ۱۵ درصد داده‌ها به طور تصادفی انتخاب شده (۴ داده)	
trainlm	تابع آموزش
logsig	تابع تبدیل لایه پنهان
purelin	تابع تبدیل لایه خروجی
MSE	تابع کارایی
۳	تعداد نورون‌های لایه پنهان
۴	تعداد متغیرهای ورودی (توصیف‌گرها)
۶	تعداد دوره‌های آموزشی

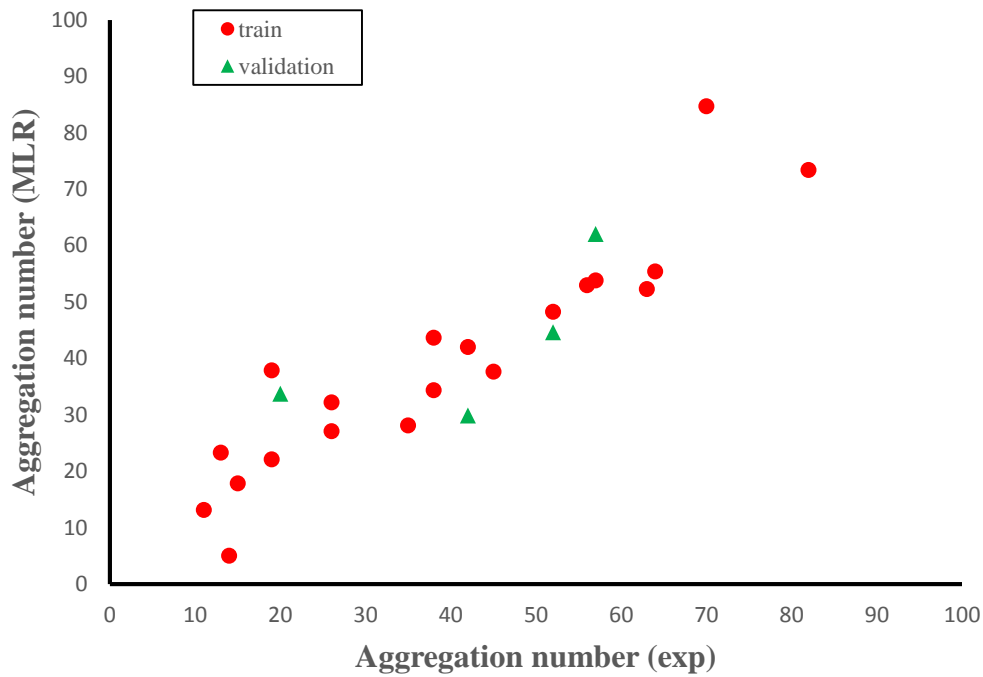


شکل ۳) شماتیک شبکه عصبی مصنوعی با الگوریتم پس انتشار.

به منظور مقایسه نتایج ANN با یک روش خطی، مدل MLR بر روی داده‌ها انجام شد که چون در این روش در مرحله آموزش یک مجموعه داده کافی می‌باشد مجموعه‌های آموزش و تست در شبکه عصبی به عنوان آموزش در MLR استفاده شدند و همان مجموعه ارزیابی در ANN در مدل MLR به عنوان مجموعه ارزیابی استفاده شد. شکل ۴ و ۵ مقایسه‌ای بین نتایج تجربی و پیش‌بینی شده برای برازش کل داده‌ها با استفاده از روش به ترتیب ANN و MLR را نشان می‌دهد. همانطور که واضح است در روش ANN تطابق بیشتری بین نتایج تجربی و پیش‌بینی شده وجود دارد. علاوه بر این ANN با استفاده از کل توصیف‌گرها نیز اجرا گردید که نتایج تطابق زیادی با آنالیز بر مبنای چهار توصیف‌گر نشان داد. نتایج ANN بر اساس ۱۸۲ توصیف‌گر به عنوان ورودی شبکه، در شکل ۲ نمایش داده شده است.



شکل ۴) نمودار پراکندگی داده های توسعه یافته (هدف) برحسب مدل پیش بینی شده ANN (خروجی) با استفاده از ۴ دیسکریپتور انتخابی



شکل ۵) نمودار پراکندگی داده های توسعه یافته (هدف) برحسب مدل پیش بینی شده MLR (خروجی)

۴-۵- نتایج آماری حاصل از مدل سازی عدد تجمع سورفکتانت‌های آنیونی توسط شبکه عصبی مصنوعی و رگرسیون خطی

چند متغیره

ضریب همبستگی (R^2) و همچنین ریشه میانگین مربعات خطا (RMSE) برای داده های مجموعه آموزش، ارزیابی، تست و کل داده ها در جدول ۵ آورده شده اند. قابل ذکر است شبکه عصبی مصنوعی سه بار اجرا شد که در هر بار، مجموعه آموزش، تست و ارزیابی به صورت راندم انتخاب شدند و میانگین R^2 ها و RMSE های مربوط به هر مجموعه به همراه فاصله اطمینان آنها در سطح اطمینان ۹۵٪ در جدول ۵ گزارش شده است. RMSE مقدار پراکندگی مقادیر تجربی حول خط رگرسیون یا به عبارت دیگر میزان خطا را مشخص می کند. برای مقایسه، مدل MLR نیز ۳ بار بر روی داده ها انجام شد که در آن از مجموعه داده های متفاوت آموزش و ارزیابی استفاده شد و نظیر روش ANN میانگین و فاصله اطمینان پارامترهای آماری این سه بار تکرار در جدول ۵ آمده است. نتایج نشان داد که ضریب همبستگی (R^2) بین نتایج تجربی و پیش‌بینی‌شده بر اساس مدل ANN بالاتر از MLR می باشد و بنابراین این روش غیرخطی کارآیی بالایی برای پیش‌بینی عدد تجمع سورفکتانت‌ها دارد.

جدول ۵. پارامترهای آماری مدل ANN و MLR.

Set of data	R^2		RMSE	
	ANN	MLR	ANN	MLR
Total	0.94±0.04	0.82±0.002	4.99±1.76	8.38±0.10
Training	0.93±0.06	0.84±0.03	5.22±2.54	7.97±0.26
Test	0.92±0.10	-	4.84±3.44	-
Validation	0.93±0.08	0.51±0.28	3.76±2.67	10.20±1.46

با توجه به نتایج به دست آمده، مدل سازی با روش شبکه عصبی مصنوعی روشی مناسب برای پیش بینی عدد تجمع سورفکتانت ها می باشد.

۴-۶- اثر متغیرهای ورودی

وزن ها مجموعه ای از مقادیر عددی هستند که بین توصیف‌گرهای ورودی و نورون های لایه پنهان و همچنین بین نورون ها و لایه خروجی قرار دارند. ماتریس وزن عصبی را می توان برای ارزیابی اهمیت نسبی هر متغیر ورودی بر روی متغیر خروجی

استفاده کرد. اهمیت نسبی هر توصیف‌گر با استفاده از روش عددی، روش گارسون، به شرح زیر تعیین شد [۳۵].

$$Q_{md} = \frac{\sum_{n=1}^h |w_{mn}v_{nd}| / \sum_{t=1}^N |w_{tn}|}{\sum_{m=1}^N \sum_{n=1}^h |w_{mn}v_{nd}| / \sum_{t=1}^N |w_{tn}|} \quad (5)$$

جایی که وزن بین m امین ورودی و n امین واحد پنهان با w_{tn} نشان داده می شود و v_{nd} نماینده وزن بین n امین واحد پنهان و خروجی d است. در مدل ANN نهایی، درصد تأثیر متغیرهای ورودی بر روی عدد تجمع با ترکیب وزن‌های اتصال پنهان ورودی و پنهان خروجی محاسبه شد. همانطور که نتایج گزارش شده در جدول ۶ نشان می دهد، اهمیت توصیف‌گرهای ورودی به ترتیب زیر افزایش می یابد:

nC > X5V > MWC05 > MWC04

جدول ۶: ماتریس وزن موثر برای مدل ANN

nC	Input descriptors			Hidden neurons	Hidden to out
	MWC04	X5V	MWC05		
-۰,۸۰۰۱	-۰,۸۱۴۰	-۰,۹۲۶۴	-۰,۸۳۴۶	H1	-۰,۲۵۰۷
۲,۱۰۱۶	۰,۳۳۵۹	-۱,۲۹۶۴	۱,۰۰۳۵	H2	۱,۱۶۱۵
-۱,۴۵۱۱	-۰,۲۷۹۵	-۱,۰۰۱۱	۰,۰۵۹۳	H3	۰,۱۵۴۷
۳۹,۹	۱۳,۱	۲۹,۶	۱۷,۴	Relative importance (%)	

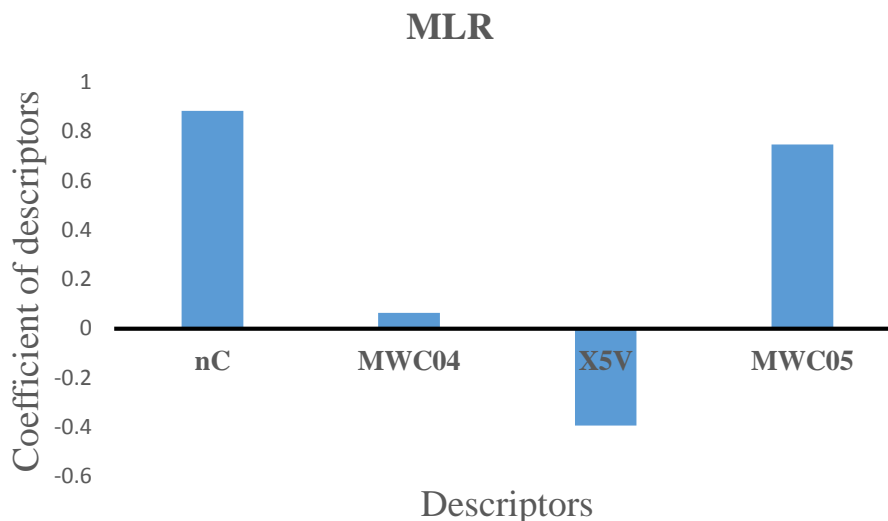
در ادامه توضیح مختصری در مورد این توصیف‌گرها داده می‌شود.

nC: توصیف‌گر nC نشان دهنده تعداد اتم‌های کربن است. همبستگی مثبتی با شاخص عدد تجمع نشان می‌دهد. بنابراین، با افزایش تعداد اتم‌های کربن در ترکیبات (افزایش آب‌گریزی)، افزایش قابل توجهی در مقدار شاخص عدد تجمع وجود خواهد داشت. (همانطور که مشاهده می‌شود در سدیم اوکتیل سولفات (SC₈S) مقدار (nC=8) و سدیم دو دسیل سولفات (SC₁₂S) مقدار (nC=12)، دارای مقادیر عدد تجمع به ترتیب ۳۸ و ۵۲ می‌باشند. هم‌چنین در ترکیبات شماره ۱ تا ۴ نیز مشاهده می‌شود که با افزایش طول زنجیره هیدروکربنی، عدد تجمع نیز افزایش می‌یابد). در نتیجه ترکیباتی که تعداد اتم‌های کربن بیشتری دارند (اسکلت مولکولی بزرگ‌تر و اندازه) مقادیر بالاتری از عدد تجمع را نشان می‌دهند. از سوی دیگر، در ترکیباتی که در آنها تعداد کربن هیبرید شده sp³ کمتری وجود دارد، مقدار عدد تجمع پایین می‌ماند [۳۲].

MWC04: تعداد مسیر مولکولی مرتبه ۴ است. در واقع توصیف‌گرهایی که تعداد سیر برگشتی را در طول‌های مختلف نشان می‌دهند. در شکل ۶ ضرایب توصیف‌گرها بر اساس مدل MLR بر حسب نام توصیف‌گر رسم شده است که بر اساس آن، **MWC04** همبستگی مثبتی با عدد تجمع نشان می‌دهد. بنابراین با افزایش این شاخص، طول زنجیره هیدروکربنی نیز افزایش پیدا می‌کند. با افزایش طول زنجیره هیدروکربنی، مهاجرت مولکول‌های سورفکتانت روی سطح زیادتر می‌شود، در نتیجه مولکول‌های بیش‌تری روی سطح قرار می‌گیرند و برهمکنش بین مولکول‌های آب که در سطح قرار دارند کاهش پیدا می‌کند بنابراین کشش سطحی کم می‌شود یعنی سطح سریع‌تر اشباع می‌شود و در نتیجه عدد تجمع افزایش پیدا می‌کند. این توصیف‌گر برای مولکول‌های حلقوی و غیر حلقوی بدون هیچ مشکلی تعریف شده است. این توصیف‌گر از نظر مفهومی بسیار ساده است، به راحتی توسط یک روش جمع‌بندی نوع مورگان محاسبه می‌شود، و در آلکان‌های غیرحلقه‌ای بسیار متمایز هستند و معیار خوبی برای پیچیدگی ساختاری است. هم‌چنین **MWC05** نیز تعداد مسیر مولکولی مرتبه ۵ است [۳۳].

X5V: شاخص اتصال ظرفیت است که یک شاخص توپولوژیکی برای پیچیدگی مولکولی است که حضور هترواتم‌ها و پیوندهای دوگانه و سه‌گانه را در مولکول نشان می‌دهد [۳۴]. با توجه به شکل ۶، این توصیف‌گر همبستگی منفی با شاخص عدد تجمع

دارد. از آنجائی که این شاخص تحت تاثیر پیوندهای دوگانه و سه گانه قرار دارد بنابراین هر چه این پیوندها در ساختار مولکول بیشتر باشد، به دلیل افزایش پدیده رزونانس و پخش بار در سطح مولکول، عدد تجمع کاهش پیدا می کند.



شکل ۶ ضرایب توصیف‌گرها بر اساس مدل MLR بر حسب نام توصیف‌گر

۵- نتیجه گیری

در مطالعه حاضر از مدل QSPR برای ارزیابی رابطه بین ساختار و فعالیت ۲۴ سورفکتانت آنیونی استفاده شد. نرم افزار Dragon برای محاسبه توصیف‌گرهای مولکولی استفاده شد که در آن ۴ توصیف‌گر با استفاده از روش انتخاب ویژگی توضیح داده شده انتخاب و سپس با مدل غیرخطی ANN و روش خطی MLR مورد ارزیابی قرار گرفتند. با استفاده از اندازه‌گیری‌های هدایت الکتریکی و روش اوانس، عدد تجمع مایسل برای سورفکتانت‌های آنیونی در محلول‌های آبی تعیین شد که مشاهده شد نتایج به دست آمده با استفاده از این روش، از مقدار واقعی گزارش شده با استفاده از روش فلورسانس، انحراف زیادی دارد. از آنجائیکه روش فلورسانس روش دقیق تری برای محاسبه عدد تجمع مایسل‌ها می باشد ولی با مشکلاتی مانند گرانی و عدم در دسترس بودن دستگاه مورد نظر همراه است، از این رو با توجه به مشکلات ذکر شده، روش‌های محاسباتی کمومتریکس می توانند در پیش بینی این پارامترها برای مواد آمفی فیلیکی و ترکیبات جدید در این طبقه بسیار مفید باشند. ضریب همبستگی (R^2) و ریشه میانگین مربعات خطا (RMSE) مدل‌های ANN و MLR (برای کل مجموعه داده) به ترتیب ۰/۹۴ ، ۴/۹۹ و ۰/۸۲ ، ۸/۳۸ به دست آمد. نتایج این مطالعه از کارایی ANN در تشخیص همبستگی ساختار-فعالیت با قدرت پیش‌بینی بالا پشتیبانی می‌کند.

۶- تقدیر و تشکر

نویسندگان از حمایت های دانشگاه سمنان برای انجام پروژه تشکر و قدردانی می نماید.

۷- مراجع

- [1] A. Bagheri, Comparison of the interaction between propranolol hydrochloride (PPL) with anionic surfactant and cationic surface active ionic liquid in micellar phase, *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, **615** (2021) 126183.
- [2] T. Gaudin, P. Rotureau, I. Pezron, G. Fayet, New QSPR Models to Predict the Critical Micelle Concentration of Sugar-Based Surfactants, *Industrial & Engineering Chemistry Research*, **55** (2016) 11716.
- [3] Ž. Medoš, S. Friesen, R. Buchner, M. Bešter-Rogač, Interplay between aggregation number, micelle charge and hydration of cationic surfactants, *Physical Chemistry Chemical Physics*, **22** (2020) 9998.
- [4] Lindman, B. and H. Wennerström, *Micelles*. 1980, Springer Berlin / Heidelberg, 1.
- [5] J. van Stam, S. Depaemelaere, F.C. De Schryver, Micellar Aggregation Numbers - A Fluorescence Study, *Journal of Chemical Education*, **75** (1998) 93.
- [6] Massart, D. L.; Vandeginste, B. G.; Buydens, L.; Lewi, P.; Smeyers-Verbeke, J. *Handbook of chemometrics and qualimetrics: Part A*; Elsevier Science Inc., 1997.
- [7] Wold, S. *Chemometrics and Intelligent Laboratory Systems*, **30** (1995) 109.
- [8] Michielan, L.; Moro, S. *Journal of chemical information and modeling*, **50** (2010) 961.
- [9] Bertolaccini L, Solli P, Pardolesi A, Pasini A (2017) An overview of the use of artificial neural networks in lung cancer research. *Journal of thoracic disease* **9** (4):924
- [10] Willis MJ, Montague GA, Di Massimo C, Tham MT, Morris AJ (1992) Artificial neural networks in process estimation and control. *Automatica* **28** (6):1181.
- [11] W. LI; H. XIE, Y. HUANG, L. SONG, Y. SHAO, K. QIU, Application of Gaussian 09/GaussView 5.0 in Analytical Chemistry Teaching. *Journal of Kunming Medical University*, **12** (2016) 134.
- [12] K.W. Ely Setiawan, Mudasir Mudasir, QSAR modeling for predicting the antifungal activities of gemini imidazolium surfactants against *Candida albicans* using GA-MLR methods, 2021.
- [13] J. Ghasemi, S. Ahmadi, Combination of Genetic Algorithm and Partial Least Squares for Cloud Point Prediction of Nonionic Surfactants from Molecular Structures, *Annali di Chimica*, **97** (2007) 69.
- [14] Ramalingam V, Palaniappan B, Panchanatham N, Palanivel S (2006) Measuring advertisement effectiveness—a neural network approach. *Expert systems with applications* **31** (1):159.
- [15] Zupan J, Gasteiger J, *Neural networks in chemistry and drug design*. John Wiley & Sons, Inc, (1999).
- [16] Fletcher R, A new approach to variable metric algorithms. *The computer journal* **13** (3) (1970) 317.

- [17] Goldfarb D, A family of variable-metric methods derived by variational means. *Mathematics of computation* **24** (109) (1970) 23.
- [18] Goh AT Back-propagation neural networks for modeling complex systems. *Artificial intelligence in engineering*, **9** (3) (1995) 143.
- [19] Iqbal A, Aftab S, A Feed-Forward and Pattern Recognition ANN Model for Network Intrusion Detection. *International Journal of Computer Network & Information Security* **11** (4) (2019).
- [20] Nidhyananthan SS, Shenbagalakshmi V, Assessment of dysarthric speech using Elman back propagation network (recurrent network) for speech recognition. *International Journal of Speech Technology* **19** (3), (2016) 577.
- [21] Ahmed S, IP Traffic forecasting using focused time delay feed forward neural network. *Bahria University Journal of Information & Communication Technologies (BUJICT)*, **2** (1) (2009).
- [22] Fermo IR, Cavali TS, Bonfim-Rocha L, Srutkoske CL, Flores FC, Andrade CM, Development of a low-cost digital image processing system for oranges selection using hopfield networks. *Food and Bioproducts Processing* **125** (2021) 181.
- [23] M. Arezoo, B. Ahmad, *J. Of Applied Chemistry*, **61** (1400) 175, in Persian.
- [24] B.W. Barry, R. Wilson, Micellar molecular weights and hydration of ethoxylated anionic and cationic surfactants, *Colloid and Polymer Science*, **256** (1978) 44.
- [25] C. La Mesa, Micelle aggregation numbers from colligative properties, *Colloid and Polymer Science*, **268** (1990) 959.
- [26] Anachkov SE, Danov KD, Basheva ES, Kralchevsky PA, Ananthapadmanabhan KP, Determination of the aggregation number and charge of ionic surfactant micelles from the stepwise thinning of foam films. *Advances in Colloid and Interface Science* **183-184**, (2012)55.
- [27] M. Aoudia, B. Al-Haddabi, Z. Al-Harhi, A. Al-Rubkhi, Sodium Lauryl Ether Sulfate Micellization and Water Solubility Enhancement Towards Naphthalene and Pyrene: Effect of the Degree of Ethoxylation, *Journal of Surfactants and Detergents*, **13** (2010) 103.
- [28] N.V. Lebedeva, A. Shahine, B.L. Bales, Aggregation Number-Based Degrees of Counterion Dissociation in Sodium n-Alkyl Sulfate Micelles, *The Journal of Physical Chemistry B*, **109** (2005) 19806.
- [29] D.J. Jobe, V.C. Reinsborough, Micellar properties of sodium alkyl sulfoacetates and sodium dialkyl sulfosuccinates in water, *Canadian Journal of Chemistry*, **62** (1984) 280.
- [30] H.V. Tartar, A.L.M. Lelong, Micellar Molecular Weights of Some Paraffin Chain Salts by Light Scattering, *The Journal of Physical Chemistry*, **59** (1955) 1185.
- [31] M. Benrraou, B.L. Bales, R. Zana, Effect of the Nature of the Counterion on the Properties of Anionic Surfactants. 1. Cmc, Ionization Degree at the Cmc and Aggregation Number of Micelles of Sodium, Cesium, Tetramethylammonium, Tetraethylammonium, Tetrapropylammonium, and Tetrabutylammonium Dodecyl Sulfates, *The Journal of Physical Chemistry B*, **107** (2003) 13432.

- [32] P. De, K. Roy, Greener chemicals for the future: QSAR modelling of the PBT index using ETA descriptors, *SAR and QSAR in Environmental Research*, **29** (2018) 319.
- [33] G. Rücker, C. Rücker, On Topological Indices, Boiling Points, and Cycloalkanes, *Journal of Chemical Information and Computer Sciences*, **39** (1999) 788.
- [34] T. Hancock, R. Put, D. Coomans, Y. Heyden, Y. Everingham, A performance comparison of modern statistical techniques for molecular descriptor selection and retention prediction in chromatographic QSRR studies, *Chemometrics and Intelligent Laboratory Systems*, **76** (2005) 185.
- [35] Fathollahi M, Sajady H, QSPR modeling of decomposition temperature of energetic cocrystals using artificial neural network. *Journal of Thermal Analysis and Calorimetry*, **133** (3) (2018) 1663.
- [36] N. Fakhari, Z. Kalantar, N. godarzi, *Linear and non-linear QSPR study of n-alkanes viscosity over a wide range of temperature and pressure*, 2012.

