



دانشگاه مازندران
دانشکده شیمی

پایان نامه دوره

کارشناسی ارشد رشته شیمی گرایش تجزیه

موضوع

مدلسازی QSPR فاکتور مهاجرت در کروماتوگرافی الکتروسینتیکی
مایسلی و مدل سازی و پیش بینی فاکتور بازداری ترکیبات دارویی
در IAM-LC

استاد راهنما

دکتر محمد حسین فاطمی

استاد مشاور

دکتر محمد رضا حاج محمدی

اساتید داور

دکتر محمد جواد چایچی و دکتر سید ناصر عزیزی

نگارش

هدی شمس الدین

خرداد ماه ۱۳۸۸

تقدیر و تشکر:

سپاس بی کران پروردگار یکتا را که هستی مان بخشید و به طریق علم و دانش رهنمونمان شد و به همنشینی رهروان علم و دانش مفتخرمان نمود و خوشه چینی از علم و معرفت را روزیمان ساخت . اکنون در آستانه راهی نو به پاس نعمات بی حد پروردگار بر خود لازم می دانم سپاسگزار تمام عزیزانی باشم که در برابر سختی ها و نا ملایمات روزگار یاریم نمودند.

سپاس فراوان از از استاد راهنمای عزیزم، آقای دکتر محمد حسین فاطمی که همواره از زحمات و راهنمایی های بی شائبه ایشان بهره مند گردیدم.

از استاد مشاور ارجمندم، آقای دکتر محمد رضا حاج محمدی که در طول انجام این طرح از راهنماییهای ارزشمندشان سود جستم کمال تشکر را دارم.

همچنین از جناب آقای دکتر چایچی و جناب آقای دکتر عزیزی که با سمت استاد مدعو، زحمت مطالعه دقیق پایان نامه را متحمل شدند و همواره از نظریات ارزشمند ایشان استفاده کرده سپاسگزارم.

از نماینده تحصیلات تکمیلی جناب آقای دکتر عمرانی که در جلسه دفاع حضور داشتند، قدردانی و سپاسگزاری می کنم.

تقدیم به دو عشق پاک زندگی، پدر و مادر عزیزم...

آنان که وجودم برایشان همه رنج بود و وجودشان برایم همه مهر. توانشان رفت تا به توانایی برسم و مویشان سپید گشت تا رویم سپید بماند. آنان که فروغ نگاهشان، گرمی کلامشان و روشنی رویشان سرمایه های جاودانی زندگی من است. آنان که راستی قامت در شکستی قامتشان تجلی یافت. در برابر وجود گرامیشان زانوی ادب بر زمین می زنم و با دلی مملو از عشق، محبت و خضوع بر دستشان بوسه می زنم.

تقدیم به خواهر مهربانم که در تمام مراحل تحصیل همواره مشوق و پشتیبان برایم بوده است.

چکیده

در قسمت نخست مدل‌سازی و پیش‌بینی فاکتور مهاجرت برخی از مشتقات بنزن در کروماتوگرافی الکتروسیتیتیکی مایسلی با استفاده از رگرسیون خطی چندتایی و شبکه عصبی مصنوعی انجام گرفت. بدین منظور ابتدا مولکول‌هایی به عنوان سری داده‌ها انتخاب شدند. این مولکول‌ها به سه سری آموزشی، پیش‌بینی و ارزیابی تقسیم شدند. سپس توصیف‌کننده‌های الکترونی، توپولوژی، هندسی و کوانتوم-شیمیایی برای تمام مولکول‌ها مورد محاسبه قرار گرفت. با استفاده از تکنیک‌های انتخاب متغیر مرحله‌ای توصیف‌کننده‌های مهم‌تر انتخاب (توصیف‌کننده‌های وارد شده در مدل عبارتند از: اندیس کایر و هال، ممان اینرسی C، بارسطحی وزن دار شده اتم‌های دهنده پیوند هیدروژنی، بار منفی‌ترین اتم، ممان دو قطبی در راستای محور Y و سپس به کمک روش‌های رگرسیون خطی چندتایی (MLR) و شبکه عصبی مصنوعی (ANN) بین این توصیف‌کننده‌ها و فاکتور مهاجرت مدل‌سازی انجام شد. در مورد شبکه عصبی مصنوعی، ابتدا پارامترهایی نظیر تعداد گره‌های لایه مخفی، ممتوم و سرعت آموزش وزن‌ها و بایاس به روش پس-انتشاربینه شدند. پارامترهای آماری برای مدل‌های بدست آمده قابلیت پیش‌بینی مدل QSPR را نشان می‌دهد که مبین این است که مدل ANN بر مدل MLR برتری چندانی ندارد.

در قسمت دوم مدل‌سازی و پیش‌بینی فاکتور بازدارندگی ۴۰ ترکیب دارویی در کروماتوگرافی مایع با غشای مصنوعی ثابت در دو فاز متحرک مورفولین پروپان سولفونیک اسید (MOPS) و فسفات بافر سالین (PBS) در $pH = 7.4$ مایسلی با استفاده از رگرسیون خطی چندتایی، روش LFER و شبکه عصبی مصنوعی انجام گرفت. خطای استاندارد برای $\log k_{wPBS}^{IAM}$ در مدل ANN برای سری آموزشی، سری تست ارزیابی داخلی و خارجی به ترتیب برابر ۰/۲۰۵، ۰/۳۲۹ و ۰/۳۸۹ است و در مدل MLR به ترتیب برابر ۰/۲۸۰، ۰/۴۲۶ و ۰/۴۴۸ می‌باشد و همچنین برای $\log k_{wMOPS}^{IAM}$ در مدل ANN برای سری آموزشی، سری تست ارزیابی داخلی و خارجی به ترتیب برابر ۰/۱۴۴، ۰/۵۹۶ و ۰/۵۵۷ است و در مدل MLR به ترتیب برابر ۰/۳۱۸، ۰/۶۱۳ و ۰/۴۵۳ می‌باشد. پارامترهای آماری مبین این است که مدل ANN بر مدل

MLR برتری دارد و برای ارزیابی و بررسی بیشتر مدل مورد نظر تست تصادفی کردن مقدار داده تجربی و ارزیابی تقاطعی مورد بررسی قرار گرفت. در بخش پایانی علاوه بر مدلسازی به روش رگرسیون خطی چندتایی مدلسازی به کمک معادله ارتباط انرژی آزاد خطی (LFER) انجام گرفت. مربع ریشه خطا برای $\log k_{wMOPS}^{IAM}$ و $\log k_{wPBS}^{IAM}$ در مدل MLR به ترتیب برابر ۰/۳۳۲ و ۰/۳۵۱ است در حالیکه در مدل LFER به ترتیب برابر ۰/۳۷۱ و ۰/۵۰۰ می باشد. پارامترهای آماری مدل MLR برتر از مدل LFER است. تست ارزیابی تقاطعی چندتایی و تست تصادفی کردن مقدار داده تجربی بر روی مدل انجام گرفت.

واژه کلیدی :

کروماتوگرافی الکتروسینتیکی مایسلی، رگرسیون خطی چندتایی، شبکه عصبی مصنوعی، ارتباط انرژی آزاد خطی، کروماتوگرافی مایع با غشای مصنوعی ثابت.

فهرست شکل ها

عنوان	صفحه
شکل ۳-۱: اجزای یک نرون بیولوژیکی	۲۴
شکل ۳-۲: نمای ریاضی از یک نرون	۲۷
شکل ۳-۳: چند نمونه از توابع انتقال موجود در ANN	۲۸
شکل ۳-۴: شمایی از یک شبکه عصبی مصنوعی	۲۹
شکل ۳-۵: یک نمونه ساده از شبکه پیش رونده	۳۲
شکل ۳-۶: یک نمونه ساده از یک شبکه پس رونده	۳۳
شکل ۳-۷: نمایی از یک پرسپترون	۳۵
شکل ۴-۱: مکانیسم جداسازی در MEKC (سورفاکتانت آنیونی، حالت شویش معمولی)	۴۷
شکل ۴-۲ (الف): نمودارپیش بینی مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۸۰:۲۰	
SDS: Decanol	۶۵
شکل ۴-۲ (ب): نمودارپیش بینی مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۹۰:۱۰	
SDS: Decanol	۶۵
شکل ۴-۲ (ج): نمودارپیش بینی مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۱۰۰:۰	
SDS: Decanol	۶۶
شکل ۴-۳ (الف): نمودارباقیمانده مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۸۰:۲۰	
SDS: Decanol	۶۶
شکل ۴-۳ (ب): نمودارباقیمانده مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۹۰:۱۰	
SDS: Decanol	۶۷
شکل ۴-۳ (ج): نمودارباقیمانده مقادیر $LogK_m$ توسط MLR بر حسب $LogK_m$ تجربی در نسبت مولی ۱۰۰:۰	

- ۶۷ SDS: Decanol
- ۶۸ شکل ۴-۴: نمودار اثر متوسط هریک از توصیف کننده ها در سه نسبت مولی همزمان
- ۶۹ شکل ۴-۵: منحنی آموزش شبکه
- ۷۰ شکل ۴-۶: پارامترهای بهینه شده شبکه با در نظر گرفتن آموزش اضافی
- ۷۱ شکل ۴-۷: نمودار میله ای نشان دهنده پارامترهای آماری حاصل از مدلسازی فاکتور مهاجرت
- ۷۲ شکل ۴-۸: نمودار پیش بینی مقادیر $LogK_m$ توسط MLR و ANN بر حسب $LogK_m$ تجربی
- ۷۳ شکل ۴-۹: نمودار باقیمانده مقادیر $LogK_m$ توسط MLR و ANN بر حسب $LogK_m$ تجربی
- ۸۰ شکل ۵-۱: نمای یک ستون IAM.PC.DD2
- ۹۳ شکل ۵-۲: نمودار اثر متوسط هریک از توصیف کننده ها در دو فاز متحرک
- ۹۴ شکل ۵-۳(الف): پارامترهای بهینه شده شبکه با در نظر گرفتن آموزش اضافی برای $\log k_{MOPS}^{IAM}$
- ۹۵ شکل ۵-۳(ب): پارامترهای بهینه شده شبکه با در نظر گرفتن آموزش اضافی برای $\log k_{PBS}^{IAM}$
- ۹۸ شکل ۵-۴(الف): نمودار پیش بینی مقادیر $LogK_{MOPS}^{IAM}$ توسط MLR بر حسب $LogK_{MOPS}^{IAM}$ تجربی
- ۹۸ شکل ۵-۴(ب): نمودار پیش بینی مقادیر $LogK_{MOPS}^{IAM}$ توسط ANN بر حسب $LogK_{MOPS}^{IAM}$ تجربی
- ۹۹ شکل ۵-۴(ج): نمودار پیش بینی مقادیر $LogK_{PBS}^{IAM}$ توسط MLR بر حسب $LogK_{PBS}^{IAM}$ تجربی
- ۹۹ شکل ۵-۴(د): نمودار پیش بینی مقادیر $LogK_{PBS}^{IAM}$ توسط ANN بر حسب $LogK_{PBS}^{IAM}$ تجربی
- ۱۰۰ شکل ۵-۵(الف): نمودار باقیمانده مقادیر $LogK_{MOPS}^{IAM}$ توسط MLR بر حسب $LogK_{MOPS}^{IAM}$ تجربی
- ۱۰۰ شکل ۵-۵(ب): نمودار باقیمانده مقادیر $LogK_{MOPS}^{IAM}$ توسط ANN بر حسب $LogK_{MOPS}^{IAM}$ تجربی
- ۱۰۱ شکل ۵-۵(ج): نمودار باقیمانده مقادیر $LogK_{PBS}^{IAM}$ توسط MLR بر حسب $LogK_{PBS}^{IAM}$ تجربی
- ۱۰۱ شکل ۵-۵(د): نمودار باقیمانده مقادیر $LogK_{PBS}^{IAM}$ توسط ANN بر حسب $LogK_{PBS}^{IAM}$ تجربی
- ۱۰۴ شکل ۵-۶: نمودار پیش بینی مقادیر $LogK^{IAM}$ توسط MLR بر حسب $LogK^{IAM}$ تجربی
- ۱۰۵ شکل ۵-۷: نمودار باقیمانده مقادیر $LogK^{IAM}$ توسط MLR بر حسب $LogK_m^{IAM}$ تجربی
- ۱۰۶ شکل ۵-۸: اثر متوسط هر یک از توصیف کننده ها در دو فاز متحرک

فهرست مطالب

صفحه	عنوان
	فصل اول
	مقدمه
۲	۱-۱ مقدمه
	فصل دوم
	روشهای پارامتری
۷	۱-۲ مقدمه
۸	۲-۲ انتخاب سری داده ها
۹	۳-۲ محاسبه توصیف کننده ها
۹	۱-۳-۲ توصیف کننده های هندسی
۱۰	۲-۳-۲ توصیف کننده های الکترونی
۱۱	۳-۳-۲ توصیف کننده های ساختاری
۱۱	۴-۳-۲ توصیف کننده های توپولوژیکی
۱۲	۱-۴-۳-۲ توصیف کننده های جزء
۱۲	۲-۴-۳-۲ اندیس های توپولوژی
۱۳	۳-۴-۳-۲ توصیف کننده های محیطی
۱۴	۴-۴-۳-۲ توصیف کننده های زیر ساختاری
۱۴	۵-۳-۲ توصیف کننده های فیزیکی - شیمیایی
۱۴	۴-۲ تجزیه و تحلیل و ارزیابی توصیف کننده ها
۱۵	۲-۵ آنالیز مدل های آماری و انتخاب مدل مناسب
۱۶	۲-۶ ارزیابی مدل به دست آمده
۱۷	۲-۷ نرم افزارهای مورد استفاده در این پژوهش

۱۷ Hyperchem بسته نرم افزاری
۱۷ MOPAC بسته نرم افزاری
۱۷ SPSS بسته نرم افزاری
۱۸ Dragon بسته نرم افزاری
۱۸ MATLAB محیط نرم افزاری

شبکه های عصبی مصنوعی

فصل سوم

۲۱ ۱-۳ مقدمه ای بر شبکه های عصبی
۲۱ ۱-۱-۳ شبکه های عصبی
۲۲ ۲-۱-۳ شباهت شبکه عصبی مصنوعی به مغز انسان
۲۳ ۳-۱-۳ روش کار نرون ها
۲۴ ۴-۱-۳ تشابهات و انتظارات
۲۴ ۱-۴-۱-۳ تشابهات
۲۵ ۲-۴-۱-۳ انتظارات
۲۵ ۲-۳ مدلسازی سیستم های عصبی
۲۶ ۱-۲-۳ مدل ریاضی
۲۸ ۲-۲-۳ توابع انتقال
۲۸ ۳-۲-۳ شبکه ها
۳۱ ۴-۲-۳ انواع شبکه ها
۳۲ ۵-۲-۳ معماری شبکه های عصبی
۳۲ ۱-۵-۲-۳ شبکه های پیش رونده
۳۳ ۲-۵-۲-۳ شبکه های پس رونده
۳۴ ۶-۲-۳ لایه های شبکه

۳۴ پرسترونها ۷-۲-۳
۳۵ آموزش شبکه های عصبی ۳-۳
۳۷ انواع شبکه ها ۴-۳
۳۷ شبکه پس انتشار ۱-۴-۳
۳۹ ویژگیهای شبکه های عصبی مصنوعی ۵-۳
۴۰ مزایا و معایب شبکه های عصبی مصنوعی ۱-۵-۳
۴۱ کاربردهای شبکه عصبی مصنوعی ۶-۳

مدلسازی و پیش بینی فاکتور مهاجرت در MEKC

فصل چهارم

۴۴ کروماتوگرافی الکتروستاتیکی مایسلی ۱-۴
۴۴ مقدمه ۱-۱-۴
۴۵ اصول جداسازی در MEKC ۲-۱-۴
۴۹ بخش تجربی ۲-۴
۴۹ مدل سازی و پیش بینی $\log k'_m$ در سه نسبت مولی متفاوت <i>SDS / Decanol</i> ۱-۲-۴
۴۹ انتخاب سری داده ها ۱-۱-۲-۴
۵۱ محاسبه توصیف کننده ها ۲-۱-۲-۴
۵۲ انتخاب توصیف کننده ها ۳-۱-۲-۴
۵۶ مدل سازی و پیش بینی $\log k'_m$ در سه نسبت مولی متفاوت <i>SDS / Decanol</i> به طور همزمان ۲-۲-۴
۵۶ انتخاب سری داده ها ۱-۲-۲-۴
۵۶ ایجاد و انتخاب توصیف کننده ها ۲-۲-۲-۴
۵۷ ایجاد شبکه های عصبی ۳-۲-۲-۴
۶۲ بحث و نتیجه گیری ۳-۴
۶۲ بررسی مدل پیش بینی فاکتور مهاجرت در سه نسبت مولی <i>SDS / Decanol</i> ۱-۳-۴

۶۲ ۱-۱-۳-۴ تجزیه و تحلیل توصیف کننده ها
۶۳ ۲-۱-۳-۴ ارزیابی و اعتبار مدل
۶۸ ۲-۳-۴ بررسی مدل پیش بینی فاکتور مهاجرت در سه نسبت مولی <i>SDS / Decanol</i> به طور همزمان
۶۸ ۱-۲-۳-۴ تجزیه و تحلیل توصیف کننده ها
۶۹ ۲-۲-۳-۴ بهینه کردن پارامترهای شبکه
۷۱ ۳-۲-۳-۴ ارزیابی کارایی شبکه عصبی مصنوعی

فصل پنجم مدل سازی و پیش بینی فاکتور بازداری در IAM-LC

۷۶ ۱-۵ مقدمه
۷۷ ۱-۱-۵ معادله پارامتر حلالیت
۷۹ ۲-۵ بخش تجربی
۷۹ ۱-۲-۵ مدل سازی و پیش بینی $\log k^{IAM}$ در دو فاز متحرک <i>MOPS</i> و <i>PBS</i>
۷۹ ۱-۱-۲-۵ انتخاب سری داده ها
۸۰ ۲-۱-۲-۵ محاسبه توصیف کننده ها
۸۰ ۳-۱-۲-۵ انتخاب توصیف کننده ها
۸۵ ۴-۱-۲-۵ ایجاد شبکه های عصبی
۸۷ ۲-۲-۵ پیش بینی $\log k^{IAM}$ از طریق مدل سازی <i>MLR</i> و <i>LFER</i> در دو فاز متحرک <i>MOPS</i> و <i>PBS</i>
۸۷ ۱-۲-۲-۵ انتخاب سری داده ها
۸۷ ۲-۲-۲-۵ ایجاد و انتخاب توصیف کننده ها در مدل سازی <i>MLR</i>
۸۸ ۳-۲-۲-۵ ایجاد و انتخاب توصیف کننده ها در مدل سازی <i>LFER</i>
۹۲ ۳-۵ بحث و نتیجه گیری
۹۲ ۱-۳-۵ بررسی مدل سازی و پیش بینی $\log k^{IAM}$ در دو فاز متحرک <i>MOPS</i> و <i>PBS</i>
۹۲ ۱-۱-۳-۵ تجزیه و تحلیل توصیف کننده ها

۹۳ ۲-۱-۳-۵ بهینه کردن پارامترهای شبکه
۹۶ ۳-۱-۳-۵ ارزیابی و اعتبار مدل
۱۰۲ ۲-۳-۵ پیش بینی $\log k^{LAM}$ از طریق مدلسازی MLR و LFER در دو فاز متحرک <i>MOPS</i> و <i>PBS</i>
۱۰۲ ۱-۲-۳-۵ ارزیابی و اعتبار مدل
۱۰۶ ۲-۲-۳-۵ تجزیه و تحلیل توصیف کننده ها
۱۰۷ ۴-۵ نتیجه گیری کلی
۱۰۸ پیشنهادات برای تحقیقات آتی

فهرست جدول ها

صفحه	عنوان
۱۲	جدول ۱-۲ نام برخی از توصیف کننده های جزء.....
۱۳	جدول ۲-۲ نام برخی از توصیف کننده های توپولوژی
۵۰	جدول ۱-۴: مقادیر تجربی $\log k'_m$ در سه نسبت مولی متفاوت <i>SDS / Decanol</i>
۵۱	جدول ۲-۴: نام برخی از توصیف کننده های محاسبه شده
	جدول ۳-۴ (الف): مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log k'_m$ در نسبت مولی ۸۰:۲۰
۵۲ <i>SDS / Decanol</i>
	جدول ۳-۴ (ب): مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log k'_m$ در نسبت مولی ۹۰:۱۰ .
۵۲ <i>SDS / Decanol</i>
	جدول ۳-۴ (ج): مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log k'_m$ در نسبت مولی ۱۰۰:۰
۵۳ <i>SDS / Decanol</i>
۵۳	جدول ۴-۴: ماتریس همبستگی بین توصیف کننده های انتخاب شده
۵۳	جدول ۴-۵: مقادیر توصیف کننده ها برای مدلسازی فاکتور مهاجرت
۵۵	جدول ۴-۶: مقادیر تجربی و محاسبه شده با استفاده از مدلسازی MLR
۵۷	جدول ۴-۷: مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log K'_m$
۵۸	جدول ۴-۸: مقادیر تجربی و محاسبه شده $\log K'_m$ با استفاده از مدل های ANN و MLR و سری آموزشی ، پیش بینی ، ارزیابی ..
۶۴	جدول ۴-۹ (الف): پارامترهای آماری حاصل از مدلسازی فاکتور مهاجرت در نسبت مولی ۸۰:۲۰ <i>SDS:Decanol</i>
۶۴	جدول ۴-۹ (ب): پارامترهای آماری حاصل از مدلسازی فاکتور مهاجرت در نسبت مولی ۹۰:۱۰ <i>SDS:Decanol</i>
۶۴	جدول ۴-۹ (ج): پارامترهای آماری حاصل از مدلسازی فاکتور مهاجرت در نسبت مولی ۱۰۰:۰ <i>SDS: Decanol</i>

- جدول ۴-۱۰: پارامترهای آماری حاصل از مدلسازی فاکتور مهاجرت ۷۱
- جدول ۵-۱: مقادیر تجربی $\log k^{IAM}$ در دو فاز متحرک *MOPS* و *PBS* ۸۲
- جدول ۵-۲ (الف): مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log k_{wMOPS}^{IAM}$ ۸۳
- جدول ۵-۲ (ب): مشخصات توصیف کننده های انتخاب شده جهت مدلسازی $\log k_{wPBS}^{IAM}$ ۸۳
- جدول ۵-۳: ماتریس همبستگی بین توصیف کننده های انتخاب شده ۸۳
- جدول ۵-۴: مقادیر توصیف کننده ها برای مدلسازی فاکتور بازداری ۸۴
- جدول ۵-۵: مقادیر تجربی و محاسبه شده $\log k_w^{IAM}$ با استفاده از مدل های *ANN* و *MLR* ۸۶
- جدول ۵-۶: ماتریس همبستگی بین توصیف کننده ها در مدل *LFER* ۸۹
- جدول ۵-۷: مقادیر عددی توصیف کننده ها در مدل *LFER* ۹۰
- جدول ۵-۸: مقادیر محاسبه شده $\log k^{IAM}$ به دو روش *MLR* و *LFER* در دو فاز متحرک *MOPS* و *PBS* ۹۱
- جدول ۵-۹: مقادیر بهینه پارامترهای شبکه ۹۵
- جدول ۵-۱۰ (الف): پارامترهای آماری حاصل از مدلسازی $\log k_{wMOPS}^{IAM}$ ۹۶
- جدول ۵-۱۰ (ب): پارامترهای آماری حاصل از مدلسازی $\log k_{wPBS}^{IAM}$ ۹۶
- جدول ۵-۱۱: مقادیر R^2 حاصل از چندین تست تصادفی کردن مقادیر تجربی برای $\log k_{wMOPS}^{IAM}$ و $\log k_{wPBS}^{IAM}$ ۹۷
- جدول ۵-۱۲: پارامترهای آماری مدل *MLR* و *LFER* جهت پیش بینی $\log k^{IAM}$ ۱۰۲
- جدول ۵-۱۳: مقادیر R^2 حاصل از ۳۰ تست تصادفی کردن مقادیر تجربی برای $\log k_{wMOPS}^{IAM}$ و $\log k_{wPBS}^{IAM}$ ۱۰۳

فصل اول

مقدمه

۱- مقدمه:

برای فهم پدیده های شیمیایی، حفظ محیط زیست، کشف و گسترش مواد جدید و زمینه های دیگر شیمی هنوز توانایی حل مسائل به طور کامل وجود ندارد و برای انجام بعضی از مسائل نیاز به سیستم های بسیار پیچیده می باشد که انجام آنها هزینه های بسیار زیاد و مطالعات گسترده را در بر دارد. برای رفع بخشی از این مشکل می توان از روشهای محاسباتی کموتریکس استفاده کرد. این روش، روشی کارآمد جهت استخراج حداکثر اطلاعات مفید از سیستم های شیمیایی می باشد. امروزه کموتریکس به عنوان شاخه ای از علم شیمی است که از روشهای ریاضی و آمار بر پایه منطق برای رسیدن به اهداف زیر استفاده می کند:

(۱) طراحی و انتخاب شرایط بهینه آزمایشات

(۲) فراهم کردن حداکثر اطلاعات ممکن به وسیله تجزیه و تحلیل داده های شیمیایی

(۳) کمک به آگاهی و درک بهتر سیستم های شیمیایی

(۴) ارائه مدل‌های مناسب جهت پیش بینی فرایندهای شیمیایی

کموتریکس اولین بار توسط دانشمند سوئدی به نام ولد در سال ۱۹۷۲ به کار گرفته شد [۱] و توسط کووالکسی توسعه داده شد [۲]. چندین تعریف برای کموتریکس بیان شده که غالباً در منابع تجزیه ای به کار می روند به عنوان مثال کموتریکس می تواند به عنوان کاربرد روشهای ریاضی، آمار، نمودار و نمادین برای افزایش اطلاعات

شیمیایی که می توانند از داده ها استخراج شوند، تعریف شود یا کمومتریکس شامل روشهایی است که در آن استفاده از انفورماتیک، ریاضی و آمار منجر به بهینه کردن اکتسابی، آنالیز و تفسیر داده های شیمیایی می شود. در جای دیگر کمومتریکس، گرفتن اطلاعات از داده های شیمیایی چند متغیره با استفاده از آمار و ریاضی معرفی می شود.

امروزه کمومتریکس در شاخه های مختلف شیمی مورد استفاده قرار گرفته است که برخی از آنها عبارتند از: کنترل فرایندها، پردازش علائم، تجزیه و تحلیل و شناخت الگوها، بهینه کردن شرایط [۳]. یکی از زمینه های مهم کاربرد کمومتریکس در مطالعاتی است که خواص مولکولها را به ویژگی های ساختاری آنها نسبت می دهد. هنگامی که ارتباط بین رفتار و ساختار مواد پیدا شود، خواص و فعالیت هر تعدادی ترکیب، حتی ترکیباتی که تا کنون سنتز نشده اند می توانند به سادگی با کامپیوتر پیش بینی شوند. بنابراین می توان ایده آل ترین ترکیبات را برای سنتز و آزمایش در آزمایشگاه انتخاب کرد [۴]. این نوع از مطالعات به بررسی کمی ارتباط ساختمان با فعالیت، QSAR و همچنین بررسی کمی ارتباط ساختمان با ویژگی، QSPR، معروف می باشند. هدف از اینگونه مطالعات این است که خواص بیولوژیکی، فیزیکی و شیمیایی یک ماده بر اساس ترکیب شیمیایی اش پیش بینی شود [۵].

هدف از این بررسی ها پیدا کردن معادلاتی است، که این معادلات توابی از یک یا تعداد بیشتری توصیف کننده بوده و دقیقاً خواص یا فعالیت تجربی را پیش بینی می کنند. به مجموعه ابزارها و روشهایی که به این منظور مورد استفاده قرار می گیرند روش های پارامتری گویند. در روش های پارامتری سعی می شود بین یک سری توصیف کننده های مولکولی با فعالیت یا خاصیت مورد نظر ارتباط منطقی برقرار نمایند. توصیف کننده های مولکولی که به این منظور استفاده می شوند، مقادیر عددی می باشند که جنبه های مختلف ساختاری مولکول را به طور کمی نشان می دهند.

مقالات متعددی در مورد استفاده از روشهای پارامتری در مطالعات QSPR / QSAR منتشر شده اند. به عنوان مثال به موارد زیر اشاره می شود:

۱- مدلسازی QSPR پیش بینی نقطه ابری سور فاکتانت های غیر یونی در سال ۲۰۰۶ [۶].

۲- پیش بینی رفتار بازدارندگی یک سری از الکل های اشباع با استفاده از اندیس های توپولوژی توسط جانکس و همکارانش [۷].

۳- کاربرد QSAR در مطالعات علف کشها [۸].

۴- پیش بینی اندیس بازداري یک سری از ترکیبات سمی در کروماتوگرافی گازی توسط گرگانی نژاد و همکارانش [۹].

۵- استفاده از QSPR برای پیش بینی نیمه عمر تخریب نوری هیدروکربنهای آروماتیک چند حلقه ای کلردار در اتمسفر [۱۰].

۶- مطالعات QSPR برای تخمین ثابت های اسیدی بعضی مشتقات اسیدهای آروماتیک با استفاده از رگرسیون خطی چندتابی [۱۱].

همانطور که می دانیم استفاده از شبکه های عصبی مصنوعی^۱، ANN، در شیمی در چند دهه اخیر رشد کرده است. شبکه های عصبی مصنوعی، منطق فازی^۲ و الگوریتم ژنتیک^۳ همگی جزء دسته ای از علوم کامپیوتر به نام هوش مصنوعی قرار می گیرند. این روشها تفاوت زیادی با سایر روشهای محاسباتی و همچنین با یکدیگر دارند. این سیستم ها قواعد را از مثالهایی که به آنها داده می شود یاد می گیرند بصورتی که هیچ کامپیوتر معمولی نمی تواند چنین کاری را انجام دهد می توانند به سرعت مسائل محاسباتی مشکل را حل کنند. آنها ساختار درونی خود را به صورت تجربی به نحوی تنظیم می کنند که بهترین پاسخ ممکن را برای داده های ورودی ایجاد کنند. استفاده از شبکه های عصبی مصنوعی در مطالعات QSPR / QSAR خصوصاً هنگامی که ارتباط بین توصیف کننده و فعالیت یا خواص مورد نظر خطی نبوده و یا اینکه بین آنها برهمکنش هایی وجود داشته باشد باعث بهبود مدل حاصله می شود.

از کاربردهای شبکه های عصبی مصنوعی می توان به موارد زیر اشاره کرد:

۱- مدل سازی بازداري کروماتوگرافی یونی [۱۲-۱۳].

۲- پیش بینی ثابت دی الکتریک [۱۴].

۳- مدل سازی حلالیت دارو در مخلوط های آب- حلال [۱۵].

۴- پیش بینی و جداسازی ترکیبات کایرال [۱۶].

۵- طراحی سرامیک های دی الکتریک [۱۷].

۶- آنالیز کمی داده های چند متغیره [۱۸].

¹ - Artificial neural network

² - Fuzzy logic

³ - Genetic algorithm

هدف از این پروژه مدل سازی و پیش بینی فاکتور مهاجرت یک سری از ترکیبات در کروماتوگرافی الکتروسینتیکی مایسلی با استفاده از گرسپون خطی چند گانه و شبکه عصبی مصنوعی و نیز مدل سازی زمان بازداری ترکیبات دارویی در کروماتوگرافی مایع با غشای مصنوعی ثابت می باشد.



فصل دوم
روشهای پارامتری