



مدیریت تحصیلات تکمیلی

دانشکده علوم

گروه شیمی

پایان نامه جهت اخذ درجه کارشناسی ارشد در رشته شیمی تجزیه

**پیش بینی فعالیت فنولی برخی مشتقات فنول با به کارگیری پارامترهای**

**کوانتومی خالص و روش های QSAR**

اساتید راهنما :

دکتر مسعود نجاتی یزدی

دکتر محمود سنچولی

استاد مشاور:

دکتر علیرضا سامزاده

تهیه و تدوین :

مریم زنگنه

دی ماه ۱۳۹۱

محمد بن عبد الله  
الله أكبر  
الله أكبر

## چکیده

روابط ساختار- فعالیت از جمله روش‌های مهم در کمومتری هستند که بررسی رابطه‌ی بین ساختار و فعالیت ترکیبات، پیش‌بینی فعالیت ترکیبات جدید را براساس اطلاعات مربوط به ساختار شیمیایی آن‌ها امکان‌پذیر می‌سازد. در این تحقیق متغیرهای کوانتومی در حالت خلاً، حلال آب و اکتانول مطابق با روش اونهاگر، متغیرهای پتانسیل الکترواستاتیک پولیتزر، رزونانس مغناطیس هسته برای ۶۷ مولکول از مشتقات فنول محاسبه شد. وابستگی فعالیت ضد میکروبی فنول به این پارامترها با استفاده از رگرسیون خطی چندگانه مورد ارزیابی قرار گرفت. نتایج برای ۶۷ مولکول فنول همبستگی نسبتاً قوی با ضریب خطی با  $R^2_{LOO}=0/43$  نشان می‌دهد. در این مدل متغیر اوربیتال انرژی سهم قابل ملاحظه‌ای در فعالیت بیولوژیکی این سری از مواد دارا می‌باشد. از آنجائیکه ارزیابی اولیه این سری نشان می‌داد که مکانیسم عمل دسته‌ای از آن‌ها متفاوت است. مدل‌سازی برای دو دسته ۵۱ و ۱۶ مولکول به طور جداگانه انجام داده شد. ضریب خطی مدل‌ها به ترتیب  $R^2_{LOO}=0/64$  و  $R^2_{LOO}=0/7$  حاصل گردید. در ادامه تلاش‌هایی برای در یک رژیم قرار دادن تمام مولکول‌ها صورت گرفت و به این منظور سعی شد متغیرهای غیر خطی در مدل با استفاده از شبکه عصبی منظور شود. الگوریتم ژنتیک به همراه شبکه عصبی مدلی ارائه داد که برای کلیه مولکول‌ها مرتبه خطی برابر  $R^2=0/79$  را نشان می‌دهد که مرتبه خطی آن نسبت به مدل خطی اولیه ارتقاء یافته است. مدل ایجاد شده با شبکه عصبی پیشنهاد می‌کند که مکانیسم عمل را متغیربیشترین بار مثبت معین می‌نماید.

واژگان کلیدی: فعالیت فنولی، ارتباط کمی ساختار-فعالیت، مشتقات فنول، پارامترهای کوانتومی

این پایان نامه را ضمن تشکر و سپاس بیکران و در کمال افتخار و اتقان تقدیم می‌نمایم به:  
خدایی که آفرید جهان را، انسان را، عقل را، علم را، معرفت را، عشق را و به کسانی که عشقشان را در وجودم دمید.  
به پدر و مادر مهربان و فداکارم

خدای را بسی شاکرم که از روی کرم، پدر و مادری فداکار نسیم ساخته تا در سایه درخت پر بار وجودشان بیایم و از ریشه آنها شاخ و برگ بگیرم و از سایه وجودشان در راه کسب علم و دانش تلاش نمایم. والدینی که بودنشان تلج افتخاری است بر سرم و نامشان دلیلی است بر بودنم، چرا که این دو وجود، پس از پروردگار، مایه هستی ام بوده‌اند و دستم را گرفتند و راه رفتن را در این وادی زندگی پر از فراز و نشیب آموختند. آموزگاران که برایم زندگی، بودن و انسان بودن را معنا کردند.

به همسر عزیزم  
که سایه مهربانش سایه ساز زندگی‌م می‌باشد، او که اسوه صبر و تحمل بوده و مشکلات مسیر را برایم تسهیل نمود.  
به برادر و خواهر دوست داشتیم  
که همواره در طول تحصیل متحمل زحمت بودند و تکیه‌گاه من در مواجهه با مشکلات، و بودنشان مایه دلگرمی من می‌باشد.

پروردگار بزرگ را شاکرم که با عنایت بی‌کران خویش توفیق پیمودن مسیر پژوهش را به من عطا فرمود. تحقق این دست‌آورد علمی جز در سایه‌ی لطف بی‌انتهای او امکان‌پذیر نمی‌بود.

به مصداق «من لم یسکر المخلوق لم یسکر الخالق» بسی شایسته است از استاد فریخته و فرزانه جناب آقای دکتر محمود سنجلی که با کرامتی چون خورشید، سرزمین دل را روشنی بخشید و گلشن سرای علم و دانش را با راه‌نمایی‌های کارساز و سازنده بارور ساختند؛ تقدیر و تشکر نمایم.

از جناب آقای دکتر معودنجاتی و علیرضا سام زاده بابت تمام زحمات یابی که برای تکمیل رساله داشتم کمال تشکر را دارم.

از جناب آقای دکتر منصور غفاری که الگویی من از نظر سخاوت علمی است زیرا با وجودی که استاد راهنما و مشاورم نبودند و لوزانه مراد حل مشکلات این پایان‌نامه یاری نمودند و همچنین زحمات مطالعه و داوری این رساله را تقبل نمودند صمیمانه تشکر و قدردانی می‌نمایم.

از همکاری صمیمانه آقای بختیار سپری از دانشگاه سنج و آقای اصغر رضایی از دانشگاه قوچان کمال تشکر و قدردانی را دارم.

از کلیه دوستانی که در طول دوران تحصیل در دانشگاه زابل به نحوی اینجانب را مورد لطف و عنایت خویش قرار داده‌اند تشکر می‌نمایم.

## فهرست مطالب

۱	فصل اول: مقدمه
۱-۱	مقدمه
۱-۲-۱	انتخاب سری داده‌ها
۲-۲-۱	تولید توصیف‌گرها
۱-۲-۲-۱	تعریف توصیف‌گر
۲-۲-۲-۱	ویژگی مهم توصیف‌گرها
۳-۲-۲-۱	تقسیم بندی برحسب انواع توصیف‌گرها
۱-۳-۲-۲-۱	توصیف‌گرهای بدون بعد یا توصیف‌گرشمارشی
۲-۳-۲-۲-۱	توصیف‌گرهای یک بعدی یا اثر انگشتی
۳-۳-۲-۲-۱	توصیف‌گرهای دوبعدی یا توپولوژیکی
۴-۳-۲-۲-۱	توصیف‌گرها سه بعدی یا هندسی
۵-۳-۲-۲-۱	توصیف‌گرهای الکترونی
۳-۲-۳-۱	انتخاب متغیر
۳-۱	روش رگرسیون خطی چندگانه
۱-۳-۱	اصول رگرسیون خطی چندگانه
۱-۱-۳-۱	خطای میانگین مربعات (MSE)
۲-۱-۳-۱	پارامتر آماری فیشرف (F)
۲-۳-۱	هم خطی چندگانه
۳-۳-۱	سهم هر توصیف‌گر در مدل پیشنهادی
۴-۱	انتخاب توصیف‌گر مولکولی
۱-۴-۱	روش‌های رگرسیون خطی چندگانه تکرار شونده
۱-۱-۴-۱	روش رگرسیون خطی چندگانه رو به جلو
۲-۱-۴-۱	روش حذف رو به عقب
۳-۱-۴-۱	روش رگرسیون مرحله‌ای
۵-۱	ارزیابی و توان پیش‌بینی مدل
۱-۵-۱	ارزیابی داخلی
۱-۱-۵-۱	ارزیابی تقاطعی گذاشتن یکی در بیرون (LOO)
۲-۱-۵-۱	ارزیابی تقاطعی گذاشتن چند تا در بیرون (LMO)
۲-۱-۵-۱	آزمایش تصادفی کردن-Y
۲-۵-۱	ارزیابی خارج
۱-۲-۵-۱	مجموعه‌های فرد و زوج
۲-۲-۵-۱	مجموعه‌های کالیبراسیون-پیش‌بینی
۲۰	فصل دوم: مروری بر مطالعات انجام شده
۱-۲	مقدمه
۲-۲	مروری بر مطالعات انجام شده
۳-۲	فنونل‌ها
۴-۲	هدف تحقیق
۳۲	فصل سوم: مواد و روش‌ها

۳۳	۱-۳-۱- مقدمه .....
۳۴	۲-۳-۲- تولید توصیف‌گرها .....
۳۴	۱-۲-۳- بهینه‌سازی .....
۳۸	۳-۳- پارامترهای الکترونی مولکولی .....
۳۸	۱-۳-۳- انرژی اوربیتال‌ها .....
۳۹	۲-۳-۳- الکترونگاتیویته .....
۳۹	۳-۳-۳- سختی و نرمی .....
۳۹	۴-۳-۳- گشتاور دوقطبی .....
۴۲	۴-۳- سطح مولکول .....
۴۲	۵-۳- خصوصیات کلیدی در سطح مولکول .....
۴۲	۱-۵-۳- پتانسیل الکترواستاتیکی .....
۴۳	۱-۱-۵-۳- پتانسیل الکترواستاتیکی و واکنش‌پذیری .....
۴۴	۲-۱-۵-۳- پتانسیل الکترواستاتیکی و برهمکنش‌های غیر کووالانسی .....
۴۶	۳-۱-۵-۳- محدودیت‌های کاربرد پتانسیل الکترواستاتیکی .....
۴۶	۲-۵-۳- انرژی یونش محلی .....
	۱-۲-۵-۳- انرژی یونش محلی میانگین در سطح مولکول و واکنش‌پذیری کمترین مقدار انرژی یونش محلی
۴۷	میانگین در سطح مولکول .....
۴۸	۳-۲-۵-۳- پتانسیل الکترواستاتیکی و انرژی یونش محلی میانگین .....
۴۹	۶-۳- قطبش‌پذیری .....
۵۰	۷-۳- کاربردهای تابع خصوصیات برهمکنش‌های عمومی .....
۵۱	۲-۷-۳- پیش‌بینی دمای بحرانی .....
۵۱	۳-۷-۳- پیش‌بینی حجم بحرانی .....
۵۱	۴-۷-۳- پیش‌بینی فشار بحرانی .....
۵۲	۵-۷-۳- پیش‌بینی ضریب توزیع .....
۵۲	۱-۵-۷-۳- پیش‌بینی ضریب توزیع آب/اکتانول .....
۵۳	۲-۵-۷-۳- پیش‌بینی زمان بازداری مولکول‌ها در کروماتوگرافی .....
۵۵	۸-۳- محاسبات رزونانس مغناطیس هسته .....
۵۶	۹-۳- ایجاد توصیف‌گرهای کلاسیک با استفاده از نرم‌افزار دراگون .....
۵۷	۱۰-۳- الگوریتم ژنتیک (GA) .....
۵۸	۱-۱۰-۳- مفاهیم الگوریتم ژنتیک باینری .....
۵۹	۲-۱۰-۳- ویژگی‌های الگوریتم ژنتیک .....
۵۹	۱۱-۳- شبکه عصبی مصنوعی .....
۶۰	۱-۱۱-۳- ساختار شبکه عصبی .....
۶۱	۲-۱۱-۳- انواع شبکه‌های عصبی .....
۶۱	۳-۱۱-۳- شبکه‌های Feed-Forward .....
۶۲	۴-۱۱-۳- توابع انتقال .....
۶۳	۵-۱۱-۳- یادگیری شبکه‌های عصبی مصنوعی .....
۶۴	۶-۱۱-۳- روش تحت نظارت .....
۶۴	۷-۱۱-۳- روش بدون نظارت .....

۶۴	..... الگوریتم پس انتشار ۸-۱۱-۳
۶۷	..... فصل چهارم: نتایج و بحث
۶۸	..... ۱-۴ مقدمه
۶۸	..... ۲-۴ مدل سازی
۷۶	..... ۳-۴ سهم تأثیر نسبی هر پارامتر
۷۷	..... ۴-۴ ارزیابی تقاطعی
۷۹	..... ۵-۴ ارزیابی خارجی زوج و فرد
۸۸	..... ۶-۴ آزمایش تصادفی کردن
۸۱	..... ۷-۴ مدل سازی با دسته اول
۹۰	..... ۸-۴ مدل سازی با دسته دوم (۱۶ عدد مولکول)
۹۷	..... ۹-۴ الگوریتم ژنتیک
۱۰۱	..... ۱۰-۴ شبکه عصبی مصنوعی
۱۰۹	..... ۶-۴ پارامترهای کلاسیک
۱۱۸	..... ۱۲-۴ نتیجه گیری
۱۱۸	..... ۱۳-۴ پیشنهادات
۱۰۰	..... فهرست منابع
۱۰۰	..... پیوست



## فهرست جداول

- جدول ۳-۱: جزئیات ساختار، فعالیت بیولوژیکی و آب‌گریزی مولکول‌های مورد استفاده در تحقیق ..... ۳۶
- جدول ۳-۲: پارامترهای کوانتومی محاسبه شده با استفاده از مجموع بارهای اتمی ..... ۴۰
- جدول ۴-۱: مدل ارائه شده بین فعالیت ضد میکروبی فنول و چهار پارامتر کوانتومی مستقل برای ۶۷ مولکول از مشتقات فنول ..... ۷۰
- جدول ۴-۲: تعیین سهم تأثیر نسبی چهار پارامتر انتخاب شده در مدل با استفاده از ۶۷ مولکول ..... ۷۶
- جدول ۴-۳: ارزیابی خارجی زوج و فرد برای ۶۷ مولکول ..... ۸۰
- جدول ۴-۴: آزمایش تصادفی‌سازی-Y برای ۶۷ مولکول ..... ۸۱
- جدول ۴-۵: مدل ارائه شده بین فعالیت ضد میکروبی فنول و هشت پارامتر کوانتومی مستقل برای ۵۱ مولکول از مشتقات فنول ..... ۸۳
- جدول ۴-۶: تعیین سهم تأثیر نسبی هشت پارامتر انتخاب شده در مدل با استفاده از ۵۱ عدد از مشتقات فنول ..... ۸۶
- جدول ۴-۷: ارزیابی خارجی زوج و فرد ۵۱ مولکول از مشتقات فنول ..... ۸۹
- جدول ۴-۸: آزمایش تصادفی‌سازی-Y برای ۵۱ عدد از مشتقات فنول ..... ۸۹
- جدول ۴-۹: مدل ارائه شده بین فعالیت ضد میکروبی فنول و سه پارامتر کوانتومی مستقل برای ۱۶ مولکول از مشتقات فنول ..... ۹۱
- جدول (۴-۱۰): تعیین سهم تأثیر نسبی هشت پارامتر انتخاب شده در مدل با استفاده از ۱۶ مولکول از مشتقات فنول ..... ۹۳
- جدول ۴-۱۱: ارزیابی خارجی زوج و فرد ۱۶ عدد از مشتقات فنول ..... ۹۶
- جدول ۴-۱۲: آزمایش تصادفی‌سازی-Y برای ۱۶ عدد از مشتقات فنول ..... ۹۶
- جدول ۴-۱۳: پارامترهای کوانتومی بهینه شده با استفاده از (GA-MLR) ..... ۹۸
- جدول ۴-۱۴: مقادیر RMSE، ۲۰ شبکه با استفاده از الگوریتم انتشار سریع و تابع Sigmoid ..... ۱۰۱
- جدول ۴-۱۵: پیش‌بینی‌های آماری و عملکرد الگوریتم انتشار سریع برای آموزش، تست و اعتبارسنجی ..... ۱۰۵
- جدول ۴-۱۶: پیش‌بینی‌های آماری و عملکرد الگوریتم انتشار سریع برای همه‌ی داده‌ها ..... ۱۰۵
- جدول ۴-۱۷: مقادیر تجربی و پیش‌بینی ضد میکروبی مشتقات فنول با استفاده از شبکه عصبی ..... ۱۰۶
- جدول ۴-۱۸: تلفیق پارامترهای کلاسیک و کوانتومی ..... ۱۱۰
- جدول ۴-۱۹: ارزیابی خارجی زوج و فرد ۶۷ عدد مولکول ..... ۱۱۷
- جدول ۴-۲۰: آزمایش تصادفی‌سازی-Y برای ۶۷ عدد مولکول ..... ۱۱۷

## فهرست اشکال

- شکل ۳-۱: مشتق مولکول فنول..... ۳۵
- شکل ۳-۲: گراف پتانسیل الکترواستاتیکی در سطح تئوری HF/6-31G\*Opt بر روی سطح مولکول ۲-اتیل فنول. گستره پتانسیل الکترواستاتیکی برحسب کیلو کالری بر مول. قرمز بیشتر از ۳۶/۱۰۳۹، زرد بین ۳۶/۱۰۳۹ و ۱۴/۳۹۶۲، سبز بین ۱۴/۳۹۶۲ و ۷/۳۱۱۶، آبی منفیتر از ۷/۳۱۱۶-..... ۵۵
- شکل ۳-۳: گراف انرژی محلی میانگین در سطح تئوری HF/6-31G\*Opt بر روی سطح مولکول ۲-اتیل فنول. انرژی‌های یونش برحسب الکترون ولت بر مول: قرمز بیشتر از ۱۸/۸۴۴۴، زرد بین ۱۸/۸۴۴۴ و ۱۶/۴۰۶۹، سبز بین ۱۶/۴۰۶۹ و ۱۳/۹۶۹۴، آبی کوچکتر از ۱۳/۹۶۹۴..... ۵۵
- شکل ۳-۴: پیکربندی شبکه عصبی چند لایه ..... ۶۱
- شکل (۳-۵): شبکه Feed-Forward شامل ۲ لایه ورودی، ۱ لایه مخفی شامل دو عصب و ۱ لایه خروجی ..... ۶۲
- شکل (۳-۶): انواع مختلف توابع انتقال (a) حد تابع انتقال سخت (b) تابع انتقال خطی (c) تابع انتقال Sigmoid - Logistic- Sigmoid تابع انتقال (d) Tangential ..... ۶۳
- شکل ۴-۱: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله مدل با استفاده از ۴ پارامتر (توصیف‌گر) کوانتومی برای ۶۷ مولکول از مشتقات فنول..... ۷۵
- شکل ۴-۲: تفاوت فعالیت ضد میکروبی فنول تجربی و فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی برای ۶۷ مولکول از مشتقات فنول..... ۷۶
- شکل ۴-۳: تعیین اثر نسبی و سهم ۴ پارامتر انتخاب شده در مدل با استفاده از ۶۷ مولکول..... ۷۷
- شکل ۴-۴: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله ارزیابی تقاطعی (LOO) برای ۶۷ مولکول از مشتقات فنول..... ۷۸
- شکل ۴-۵: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله ارزیابی تقاطعی (LMO) برای ۶۷ مولکول از مشتقات فنول..... ۷۹
- شکل ۴-۶: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله مدل با استفاده از ۸ پارامتر (توصیف‌گر) کوانتومی برای ۵۱ مولکول از مشتقات فنول..... ۸۵
- شکل ۴-۷: تفاوت فعالیت ضد میکروبی فنول تجربی و فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی برای ۵۱ مولکول از مشتقات فنول..... ۸۵
- شکل ۴-۸: تعیین اثر نسبی و سهم ۸ توصیف‌گر انتخاب شده در مدل با استفاده از ۵۱ عدد از مشتقات فنول... ۸۶
- شکل ۴-۹: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله (LOO) با استفاده از ۸ پارامتر (توصیف‌گر) کوانتومی برای ۵۱ مولکول از مشتقات فنول..... ۸۷
- شکل ۴-۱۰: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله (LOO) با استفاده از ۸ پارامتر (توصیف‌گر) کوانتومی برای ۵۰ عدد از مشتقات فنول..... ۸۸

شکل ۴-۱۱: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله (LMO) با استفاده از ۸ پارامتر (توصیف‌گر) کوانتومی برای ۵۱ مولکول از مشتقات فنول..... ۸۸

شکل ۴-۱۲: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به‌وسیله (LMO) با استفاده از ۸ پارامتر (توصیف‌گر) کوانتومی برای ۵۰ عدد از مشتقات فنول..... ۸۹

شکل ۴-۱۳: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با..... ۹۲

استفاده از ۳ پارامتر (توصیف‌گر) کوانتومی برای ۱۶ مولکول از مشتقات فنول..... ۹۲

شکل ۴-۱۴: تفاوت فعالیت ضد میکروبی فنول تجربی و فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی برای ۱۵ مولکول از مشتقات فنول ۹۲ شکل ۴-۱۵: تعیین اثر نسبی و سهم ۳ توصیف‌گر انتخاب شده در مدل با استفاده از ۱۶ مولکول از مشتقات فنول..... ۹۳

شکل ۴-۱۶: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از ارزیابی تقاطعی (LOO) برای ۱۶ مولکول از مشتقات فنول..... ۹۴

شکل ۴-۱۷: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از ارزیابی تقاطعی (LOO) برای ۱۵ مولکول از مشتقات فنول..... ۹۴

شکل ۴-۱۸: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از ارزیابی تقاطعی (LMO) برای ۱۶ مولکول از مشتقات فنول..... ۹۵

شکل ۴-۱۹: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از ارزیابی تقاطعی (LMO) برای ۱۵ مولکول از مشتقات فنول..... ۹۵

شکل ۴-۲۰: عملکرد شبکه در نرون‌های مختلف مخفی با استفاده از الگوریتم انتشار سریع (QP)..... ۱۰۲

شکل ۴-۲۱: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از شبکه عصبی برای داده‌های آموزش..... ۱۰۲

شکل ۴-۲۲: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از شبکه عصبی برای داده‌های تست..... ۱۰۳

شکل ۴-۲۳: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از شبکه عصبی برای داده‌های اعتبارسنجی..... ۱۰۳

شکل ۴-۲۴: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی با استفاده از شبکه عصبی برای تمام داده‌ها..... ۱۰۴

شکل ۴-۲۵: تفاوت فعالیت ضد میکروبی فنول تجربی و فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب ۱۰۴ فعالیت ضد میکروبی فنول تجربی برای ۶۷ مولکول از مشتقات فنول..... ۱۰۴

شکل ۴-۲۶: درصد تأثیر توصیف‌گر انتخاب شده در مدل ساخته شده با استفاده از شبکه عصبی مصنوعی..... ۱۰۶

شکل ۴-۲۷: یک شبکه پرسپترون چند لایه رو به جلو (MLP) شامل ۹ تا ورودی، یک لایه مخفی با ۳ تا نرون و یک خروجی ..... ۱۰۸

شکل ۴-۲۸: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی توسط مدل با استفاده ۱۱ پارامتر (توصیف‌گر) کوانتومی و کلاسیک ..... ۱۱۴

شکل ۴-۲۹: تفاوت فعالیت ضد میکروبی فنول تجربی و فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی برای ۶۷ مولکول از مشتقات فنول با استفاده از مدل پارامترهای کلاسیک و کوانتومی ..... ۱۱۴

شکل ۴-۳۰: تعیین اثر نسبی و سهم ۱۱ توصیف‌گر انتخاب شده در مدل با استفاده از ۶۷ مولکول از مشتقات فنول ..... ۱۱۵

شکل ۴-۳۱: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به وسیله ارزیابی تقاطعی (LOO) برای ۶۷ مولکول از مشتقات فنول ..... ۱۱۵

شکل ۴-۳۲: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به وسیله ۱۱۶ ارزیابی تقاطعی (LOO) بدون مولکول شماره ۴۰ ..... ۱۱۶

شکل ۴-۳۴: فعالیت ضد میکروبی فنول پیش‌بینی شده برحسب فعالیت ضد میکروبی فنول تجربی به وسیله ارزیابی تقاطعی (LMO) بدون مولکول شماره ۴۰ ..... ۱۱۷

# فصل اول

## مقدمه

## ۱-۱- مقدمه

امروزه بسیاری از دانشمندان شیمی به نحوی از روش‌های آماری و ریاضی برای کسب داده‌های شیمیایی استفاده می‌نمایند که به آن کمومتری<sup>۱</sup> اطلاق می‌شود. یکی از مهمترین زمینه‌های کاربرد کمومتری مطالعه‌ی ارتباط بین خواص مولکول‌ها با ویژگی‌های ساختاری آن‌ها می‌باشد. این نوع از مطالعات که تحت عنوان ارتباط کمی ساختار-فعالیت، QSAR<sup>۲</sup> و یا ارتباط کمی ساختار-خاصیت، QSPR<sup>۳</sup> معروف می‌باشند به نحوه‌ی ارتباط موجود بین خصوصیات مولکول‌ها و ویژگی ساختاری آن‌ها می‌پردازند. بررسی رابطه‌ی بین ساختار و فعالیت ترکیبات، پیش‌بینی فعالیت ترکیبات جدید را براساس اطلاعات مربوط به ساختار شیمیایی آن‌ها امکان‌پذیر می‌سازد. مدل‌های QSAR/QSPR روابط ریاضی بین ساختار شیمیایی با خواص فیزیکی، شیمیایی و فعالیت بیولوژیکی هستند.

مدل‌های QSAR/QSPR ابزارهای مطلوبی هستند که شیمیدانان را قادر می‌سازند تا عملکردهای شیمیایی جدید و حتی سنتز نشده را پیش‌بینی کنند. توصیف‌گرهای مولکولی نقش اساسی در مطالعات QSAR/QSPR دارند و پیدا کردن توصیف‌گرهای مولکولی جدید با همبستگی بیشتر برای فعالیت/خاصیت، پیش‌برنده تحقیقات QSAR/QSPR است. صدها توصیف‌گر مولکولی در دسترس هستند که می‌توانند جنبه‌های مختلفی از متغیرهای ساختاری را شرح دهند و برای ارتباط با انواع مختلف خواص فیزیکی، شیمیایی و فعالیت بیولوژیکی به کار روند. از بین این توصیف‌گرها، توصیف‌گرهای الکترونی به دست آمده از محاسبات شیمی کوانتوم عمومیت عمده‌ای دارند و یک رقابتی بین پیچیدگی محاسبات و صحت روش‌های محاسباتی شیمی کوانتوم (مانند

---

<sup>۱</sup> Chemometrics

<sup>۲</sup> Quantitative structure-activity relationship

<sup>۳</sup> Quantitative structure-property relationship

روش‌های نیمه تجربی و *ab initio* وجود دارد. پیشرفت‌های اخیر در سخت‌افزار محاسباتی و گسترش الگوریتم‌های کارآمد به توسعه محاسبات شیمی کوانتوم مولکولی کمک کرده است. ارتباط کمی ساختار- فعالیت براساس این فرضیه استوار است که ساختار مولکولی باید متغیر مسئولی برای خواص بیولوژیکی و شیمی فیزیک آن‌ها باشد. ارتباط کمی ساختار- فعالیت می‌تواند فعالیت بیولوژیکی ماده جدید یا آزمایش نشده را از روی ساختار مولکولی ترکیبات مشابه که تاکنون تشخیص داده شده‌اند، به دست آورد.

### ۱-۲-۲-۱- مراحل اصلی مدل‌سازی QSAR/QSPR

این مراحل عبارتند از:

(الف) انتخاب سری داده‌ها

(ب) تولید توصیف‌گرها

(ج) انتخاب متغیر

(د) مدل‌سازی QSAR

(ه) ارزیابی و پیش‌بینی مدل‌ها

#### ۱-۲-۱- انتخاب سری داده‌ها<sup>۱</sup>

داده‌های شیمیایی از آزمایش تجربی و یا مقالات تهیه می‌شوند که باید صحت آن‌ها به دقت ارزیابی شود چون صحت مدل جدا از صحت داده‌های ورودی نیست. داده‌ها باید از یک منبع گرفته شوند.

#### ۱-۲-۲-۱- تولید توصیف‌گرها<sup>۲</sup>

تعداد زیادی از نرم‌افزارها، مجموعه‌های گسترده‌ای از توصیف‌گرهای تئوری مختلف از گراف‌های دو بعدی تا مختصات سه بعدی X,Y,Z را محاسبه می‌کنند که بطور مثال برخی از نرم‌افزارهای رایج در این زمینه دراگون (DRAGON, Surface analysis suite) می‌باشد.

<sup>۱</sup>Data set

<sup>۲</sup>Descriptor generation

### ۱-۲-۲-۱- تعریف توصیف‌گر

توصیف‌گرهای مولکولی<sup>۱</sup> پارامترهای عددی<sup>۲</sup> هستند که ویژگی‌های متفاوتی از مولکول را بیان می‌کنند. توصیف‌گرهای مولکولی می‌توانند خواص شیمی-فیزیکی تجربی مولکول یا اندیس‌های تئوری محاسبه شده به وسیله فرمول‌های ریاضی و الگوریتم‌های محاسباتی باشند. به عبارت دیگر، هر یک از این توصیف‌گرها، اطلاعات خاصی از مولکول را در اختیار می‌گذارند. تعیین و انتخاب توصیف‌گرها از اهمیت به‌سزایی برخوردار است. در صورتی که توصیف‌گرهای انتخاب شده مناسب نباشند، عمل مدل‌سازی با مشکل مواجه شده و مدل مناسب به دست نخواهد آمد. بنابراین در انتخاب توصیف‌گرها باید دقت کرد و توصیف‌گرهایی باید انتخاب شوند که در ارتباط با ویژگی خاص مولکول باشند.

بیشتر از ۳۰۰۰ توصیف‌گر از تئوری‌ها و رویکردهای مختلف تعریف شده به‌دست آمده‌اند و به وسیله نرم‌افزارهای خاصی قابل محاسبه می‌باشند. هر توصیف‌گر مولکولی اطلاعات خاصی از کل اطلاعات شیمیایی را در برمی‌گیرد در نتیجه تعداد توصیف‌گرها به‌طور پیوسته با افزایش نیاز به کسب اطلاعات عمیق‌تر در مورد سیستم‌های بیولوژیکی و شیمیایی، افزایش می‌یابد. یک نکته مورد توجه در استفاده از توصیف‌گرها در مسائل مدل‌سازی، اطلاعاتی است که در برمی‌گیرند. برخی افراد گمان می‌کنند که بهتر است از اکثر توصیف‌گرها برای فرآیند مدل‌سازی استفاده کنند، این نادرست است بلکه بهترین توصیف‌گرها آن‌هایی هستند که محتوای اطلاعات آن‌ها با محتوای اطلاعات پاسخ قابل مقایسه باشد.

### ۱-۲-۲-۱- ویژگی مهم توصیف‌گرها

توصیف‌گرهای تعریف شده باید ویژگی‌های زیر را داشته باشند:

الف) تفسیر ساختاری داشته باشند.

ب) حداقل با یک خاصیت همبستگی مطلوبی داشته باشد.

<sup>۱</sup> Molecular descriptor

<sup>۲</sup> Numerical indices



ج) باید بین ساختارهای مشابه مثل ایزومرها تمایز قائل شود.

د) باید برای یک ساختار موضعی به کار رود.

و) باید مستقل باشد.

ز) ساده باشد.

#### ۱-۲-۲-۳- تقسیم بندی بر حسب انواع توصیف‌گرها

توصیف‌گرهای مختلفی در محاسبات مورد استفاده قرار می‌گیرند که عبارتند از:

الف) توصیف‌گر بدون بعد یا شمارشی

ب) توصیف‌گر یک بعدی یا اثر انگشتی

ج) توصیف‌گر دو بعدی یا توپولوژیکی

د) توصیف‌گر سه بعدی یا هندسی

ه) توصیف‌گر الکترونی

#### ۱-۲-۲-۳-۱- توصیف‌گرهای بدون بعد یا توصیف‌گر شمارشی<sup>۱</sup>

توصیف‌گرهای به دست آمده از فرمول شیمیایی توصیف‌گر بدون بعد نامیده می‌شوند. این توصیف‌گرهای مولکولی نیازی به اطلاعات مربوط به ساختار مولکولی و اتصالات اتم‌ها ندارند. تعداد اتم‌ها و پیوندها، همین‌طور مجموع یا میانگین خواص اتمی از این نوع دسته از توصیف‌گرها به شمار می‌روند. این توصیف‌گرها به آسانی محاسبه می‌شوند، ذاتاً تفسیرپذیرند، به ساختار مولکولی بهینه نیازی ندارند و مستقل از کنفورماسیون‌اند. مانند عدد اتمی، وزن مولکولی (MW)، تعداد انواع اتم‌ها و به طور کلی توصیف‌گرهای اساسی<sup>۲</sup> و هر تابعی از خواص اتمی جزء این گروه می‌باشند. خواص اتمی رایج عبارتند از جرم اتمی، بار اتمی، شعاع واندروالسی و شعاع کووالانسی، قطبش‌پذیری اتمی هستند.

<sup>۱</sup> Dimensionless descriptor or count descriptor

<sup>۲</sup> Constitutional descriptor

### ۱-۲-۲-۳-۲- توصیف‌گرهای یک بعدی یا اثر انگشتی<sup>۱</sup>

توصیف‌گرهای مولکولی که می‌توانند از اطلاعات زیر ساختاری مولکول محاسبه شوند مثل تعداد گروه‌های عاملی، اجزاء زیر ساختاری متعلق به توصیف‌گرهای یک بعدی یا 1D است. این توصیف‌گرها که به‌عنوان اثر انگشتی معرفی می‌شوند، بردارهایی دوبعدی هستند که در حضور زیر ساختار با یک و در غیاب آن با صفر مشخص می‌شود.

### ۱-۲-۲-۳-۳- توصیف‌گرهای دوبعدی یا توپولوژیکی<sup>۲</sup>

یک توصیف‌گر دوبعدی (2D) چگونگی اتصال اتم‌ها در یک مولکول را برحسب ماهیت پیوندهای شیمیایی معین می‌کند. این دسته از توصیف‌گرها اطلاعاتی راجع به ساختمان، اندازه و میزان شاخه‌ای شدن مولکول، نحوه اتصال اتم‌ها و نوع اتم‌های موجود در یک مولکول را در اختیار ما قرار می‌دهند. جهت محاسبه این توصیف‌گرها ابتدا باید ساختار مولکول را به صورت یک گراف شیمیایی رسم نمود.

تئوری گراف شیمیایی<sup>۳</sup>، شاخه‌ای از علم شیمی-ریاضی می‌باشد که عموماً با شیمی تئوری در ارتباط است. بر اساس این تئوری یک گراف نشان‌دهنده مجموعه‌ای از عناصر یک گروه و نحوه ارتباط بین آن‌ها می‌باشد. هر گراف از تعدادی رأس و لبه تشکیل شده است. لبه‌ها خطوطی می‌باشند که رأس‌ها را به هم متصل کرده‌اند. یک دسته از گراف‌ها، گراف‌های ملکولی هستند. در این گراف‌ها هر نقطه نشان‌دهنده یک اتم بوده و خطوط بین نقاط نیز نشان‌دهنده پیوند شیمیایی بین اتم‌هاست. در این نوع از گراف‌ها منظور از مسیر، فاصله‌ای است که در طول آن هیچ نقطه‌ای دو بار تکرار نشود. معمولاً در گراف‌ها، اتم هیدروژن را نشان نمی‌دهند. این اندیس‌ها اطلاعاتی شامل ساختمان، اندازه و میزان شاخه‌ای شدن مولکول، پیوندها، تعداد اتم‌ها

<sup>۱</sup> One-dimensional descriptors or fingerprints

<sup>۲</sup> Two-dimensional descriptors or topological descriptors

<sup>۳</sup> Chemical graph theory

و نوع اتم های ناجور در مولکول را دارا می باشند. این اندیس ها اولین بار به وسیله راندیک در سال ۱۹۷۵ ارائه شده و سپس در سال ۱۹۷۶ توسط هال و کایر اصلاح شد.

اندیس هوزویا<sup>۱</sup>، شمار قدم های مولکولی<sup>۲</sup>، توصیف گرهای شکل کایر<sup>۳</sup>، اندیس واینر<sup>۴</sup> و اندیس هارای<sup>۵</sup> مثال هایی از این نوع توصیف گرها هستند.

#### ۱-۲-۲-۳-۴- توصیف گرها سه بعدی یا هندسی<sup>۶</sup>

توصیف گرهای هندسی نسبت به نوع توپولوژی از اهمیت بیشتری برخوردار می باشند. یکی از خصوصیات بارز این توصیف گرها، ارتباط آن ها با ساختار سه بعدی مولکول ها می باشد. بنابراین قبل از محاسبه این توصیف گر ها باید پایدارترین ساختار فضایی مولکول تعیین شود که بهینه کردن مولکول توسط روش های مکانیک کوانتومی انجام می شود.

#### ۱-۲-۲-۳-۵- توصیف گرهای الکترونی

توصیف گرهای الکترونی بیانگر خواص الکترونی مولکول ها می باشند و اطلاعات مفیدی راجع به برهم کنش های الکترونی بین مولکولی و درون مولکولی ارائه می دهند. این نوع از برهم کنش ها خواص فیزیکی و شیمیایی مولکول ها را به شدت تحت تأثیر قرار می دهند. اهمیت این برهم کنش ها به ویژه هنگامی که مولکول های عمل کننده قطبی باشند، بیشتر می شود. با توجه به این که مولکول ها از الکترون، پروتون و نوترون تشکیل شده اند، بنابراین بسیاری از رفتارهای مولکول را می توان با توجه به تأثیرات متقابل الکتریکی درک و توجیه نمود. اندازه و قدرت نیروهای حاصل از برهم کنش های الکترونی، بار جزئی اتم ها، نحوه توزیع بار در مولکول، فاصله بین بارهای مخالف، گشتاور دو قطبی و قطبش پذیری مولکول ها را می توان در این مورد نام برد.

<sup>۱</sup> Hosoya index

<sup>۲</sup> Molecular walk counts

<sup>۳</sup> Kier shap descriptor

<sup>۴</sup> Wiener index

<sup>۵</sup> Haray indices

<sup>۶</sup> Three-dimentional descriptors or geometrical descriptors

بنابراین دانستن مقادیر فوق برای یک مولکول می تواند در پیش بینی و توجیه رفتار آن در محیط مفید باشد.

هر نوع توصیف‌گری که بتواند به نحوی یکی از عوامل مؤثر در برهم کنش های الکترونی را بیان کند، در گروه توصیف‌گرهای الکترونی قرار می گیرد. برای محاسبه‌ی توصیف‌گرهای الکترونی و هندسی بایستی ساختار مولکول‌ها را بهینه کرده و سپس توصیف‌گرهای مزبور را محاسبه نمود. برای بهینه کردن ساختار مولکول‌ها عموماً از روش های اوربیتال مولکولی استفاده می شود. این روش ها به طور کلی به دو گروه روش های آغازین و نیمه تجربی تقسیم می شوند. در روش های آغازین که امروزه به روش های *ab initio* نیز معروفند، جهت محاسبه پارامترهای الکترونی ساختاری از حل دقیق معادله شرودینگر استفاده می کنند. روشن است که برای حل دقیق معادله شرودینگر ناگزیر از حل تعداد زیادی انتگرال پیچیده و دشوار هستیم. به طور مثال برای حل تابع موج یک مولکول متوسط باید بیش از  $10^6$  انتگرال و معادله درجه بالا را حل نمود. البته بسیاری از انتگرال های فوق مقادیری برابر یک یا نزدیک به صفر دارند. مشخص است که چنانچه روشی قادر به استفاده از تقریب های مناسب جهت حذف این انتگرال ها باشد بسیار مطلوب بوده و باعث صرفه جویی در زمان لازم برای محاسبات خواهد شد. به علاوه حذف این انتگرال ها باعث می شوند که بتوان مولکول های بزرگتر را بررسی کرد.

تئوری تقریب اوربیتال مولکولی اغلب از داده های تجربی مربوط به اتم ها و سیستم های مولکولی جهت تخمین کمیت های قرار دادی که پارامتر محاسباتی محسوب می شوند، استفاده می کند. از این رو به این تکنیک ها، روش های نیمه تجربی گفته می شود. در محاسبات نیمه تجربی عموماً از یک هامیلتونی ساده‌تر از هامیلتونی مولکولی استفاده می‌شود و ضمناً در این محاسبات از داده‌های تجربی و یا از پارامترهایی که می‌توانند برای هماهنگی با داده‌های تجربی و یا پارامترهای که می‌توانند برای هماهنگی با داده‌های تجربی تنظیم شوند، کمک گرفته می‌شود.