





دانشگاه صنعتی شاهرود

دانشکده: علوم پایه

گروه: شیمی

مطالعه کمی ساختار-فعالیت برخی از ترکیبات سولفونانیلید به عنوان گروهی

جدید از داروهای ضد سرطان و فعالیت ضد HIV برخی از ترکیبات جدید

دانشجو: مهدی محمدرضائی

استاد راهنما:

دکتر منصور عرب چم جنگلی

استاد مشاور:

دکتر زهرا کلانتر کهدمی

پایان نامه کارشناسی ارشد جهت اخذ درجه کارشناسی ارشد

دی ماه ۱۳۸۹

به پاس تعبیر عظیم و انسانی شان از کلمه ایثار و از خودگذشتگی  
به پاس قلب های بزرگشان که فریادرس است و سرگردانی و ترس در پناهشان به شجاعت می گراید  
به پاس عاطفه سرشار و گرمای امید بخش وجودشان که در این سردترین روزگار ان بهترین پشتیبان است  
و به پاس محبت های بی دریغشان که هرگز فروکش نمی کند

این مجموعه را به پدر و مادر عزیزم تقدیم می کنم

خدایا  
سر نوشت مرا خیر بنویس  
سر نوشت مرا خیر بنویس  
تقدیری مبارک  
تا هر چه را که تو دیر می خواهی زود نخواهم،  
و هر چه را تو زود می خواهی دیر نخواهم...م...

به مصداق «من لم یشکر المخلوق لم یشکر الخالق» بسی شایسته است

که با درود فراوان خدمت پدر و مادر بسیار عزیز، دلسوز و فداکارم که پیوسته جرعه نوش جام تعلیم و تربیت، فضیلت و انسانیت آنها بوده‌ام و همواره چراغ وجودشان روشنگر راه من در سختی‌ها و مشکلات بوده است، از ایشان تقدیر و تشکر نمایم.

از زحمات بی‌دریغ اساتید فرهیخته و بزرگوام در دانشگاه صنعتی شاهرود بخصوص استاد ارجمندم جناب آقای دکتر عرب که با راهنمایی‌های برادرانه خود در پیشبرد این پروژه راهگشای اینجانب بودند و استاد گرامیم سرکار خانم دکتر کلانتر که با مشاوره‌های خردمندانه و پرمغزشان یاری‌گر من در این راه بوده و همیشه با روی گشاده پاسخگوی سوالات بنده بودند کمال تشکر و سپاسگزاری را داشته باشم.

همچنین از تمامی همکلاسی‌های عزیزم و بچه‌های دانشکده شیمی و نیز دوستان و هم‌اتاقی‌های گرامییم که لحظاتی سرشار از صفا و صمیمیت را در کنار خود برایم به یادگار گذاشتند و همیشه اینجانب را مورد لطف و محبت خود قرار داده و به من درس صداقت و مهرورزی آموختند بسیار سپاسگزار باشم.

در پایان از زحمات سرکار خانم هنگامه سلیمی که مرا در مراحل مختلف انجام این پروژه یاری نمودند و مهربانانه پاسخگوی پرسش‌های من بودند و تجربیات خود را در اختیار اینجانب قرار دادند کمال تشکر را دارم.

## تعهدنامه

اینجانب مهدی محمدرضائی دانشجوی دوره کارشناسی ارشد رشته شیمی تجزیه دانشکده شیمی دانشگاه صنعتی شاهرود نویسنده پایان‌نامه: مطالعه کمی ساختار- فعالیت برخی از ترکیبات سولفونانیلید به عنوان گروهی جدید از داروهای ضد سرطان و فعالیت ضد HIV برخی از ترکیبات جدید تحت راهنمایی دکتر منصور عرب چمنجنگلی متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تا کنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود» و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجودات زنده (یا بافت‌های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاقی رعایت شده است.

## تاریخ

### امضای دانشجو

#### مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

\* متن این صفحه نیز باید در ابتدای نسخه‌های تکثیر شده پایان‌نامه وجود داشته باشد.

## چکیده:

تحقیقات اخیر نشان داده‌اند که استروژن نقش بسزایی در بیماری سرطان سینه دارد، بنابراین پژوهش‌هایی در رابطه با داروهایی که بتوانند به نحوی فعالیت این هورمون را در بدن محدود نمایند انجام شده است. این پژوهش‌ها به دو دسته تقسیم می‌شوند، اول پژوهش بر روی ترکیباتی که به طور مستقیم بر فعالیت استروژن در بدن اثر می‌گذارند و دوم تحقیق بر روی داروهایی که از طریق ایجاد محدودیت در تولید استروژن عمل می‌نمایند. یکی از این دسته ترکیبات که می‌توانند تولید استروژن را محدود نمایند، مشتقات سولفونانیلید هستند. در بررسی اول رابطه کمی ساختار-فعالیت ۲۱ آنالوگ سولفونانیلید با استفاده از رگرسیون خطی چندگانه و شبکه عصبی مصنوعی مورد مطالعه قرار گرفت. از میان تعداد زیاد توصیف‌کننده محاسبه شده فقط ۲۰ توصیف‌کننده توسط رگرسیون خطی مرحله‌ای انتخاب گردید که فقط ۶ تا از آنها به عنوان ورودی شبکه استفاده شدند. سری داده‌ها به طور تصادفی به دو سری آموزش و تست که به ترتیب شامل ۱۶ و ۵ ترکیب بودند تقسیم شد. مدل‌های رگرسیون خطی و شبکه عصبی با استفاده از روش ارزیابی تقاطعی ساخته شدند. توانایی پیش‌بینی مدل‌ها با استفاده از سری تست و روش رد مرحله‌ای تک‌تک مورد ارزیابی قرار گرفت و مقادیر MSE به ترتیب برابر با ۰/۰۲۲۷ و ۰/۰۱۳۲ برای MLR و ۰/۰۰۹۸ و ۰/۰۱۶۱ برای ANN به دست آمد. نتایج به دست آمده نشان دهنده قدرت بالای مدل شبکه عصبی مصنوعی در پیش‌بینی توان بازدارندگی ترکیبات مورد بررسی می‌باشند.

در بررسی دوم رابطه ساختار-فعالیت مشتقات ۵-اکسوپیرولیدین-۳-کربوکسامید مورد پژوهش قرار گرفت. این ترکیبات از طریق دخالت در برهمکنش میان CCR5 و HIV مانع از گسترش بیماری می‌شوند. مجموعه داده‌های مورد بررسی شامل ۱۰۴ ترکیب از مشتقات ۵-اکسوپیرولیدین-۳-کربوکسامید می‌باشد. در این تحقیق سری داده‌ها به طور تصادفی به سه سری آموزش، ارزیابی و تست تقسیم شد که هر یک به ترتیب شامل ۶۲، ۲۱ و ۲۱ ترکیب می‌باشند. مدل شبکه عصبی با استفاده از ۱۹ توصیف‌کننده با قدرت بالایی مقادیر  $pic_{50}$  مربوط به هر یک از ترکیبات را پیش‌بینی نمود. مقادیر mse برای سری تست و رد مرحله‌ای تک‌تک به ترتیب ۰/۱۲۱۰ و ۰/۱۴۰۵ می‌باشد.

**واژه‌های کلیدی:** مطالعه ساختار-فعالیت، شبکه عصبی مصنوعی، رگرسیون خطی چندگانه، سولفونانیلید.

نتایج به دست آمده از این پژوهش طی دو پوستر تحت عناوین

1. Quantitative Structure-Activity Relationships of some of Sulfonanilide derivatives as anti-breast cancer agents
2. Non-linear QSAR study of anti-HIV-1 activity of some recently synthesized pipredine-4-carboxamide derivatives

در هفدهمین سمینار شیمی تجزیه ایران در دانشگاه کاشان ارائه گردید.

## فهرست مطالب

### فصل اول

- ۱-۱-۱- کمومتریکس ..... ۲
- ۱-۲-۱- ارتباط کمی ساختار - فعالیت ..... ۳
- ۱-۲-۱- جمع‌آوری سری داده‌ها ..... ۴
- ۱-۲-۲-۱- رسم ساختارهای مولکولی و بهینه‌سازی آنها ..... ۴
- ۱-۲-۳-۱- استخراج توصیف‌کننده‌ها ..... ۵
- ۱-۲-۴-۱- بدست آوردن بهترین توصیف‌کننده‌ها برای ورود به مدل ..... ۸
- ۱-۲-۵-۱- مدل‌سازی ..... ۱۰
- ۱-۲-۶-۱- مدل‌سازی به روش MLR ..... ۱۰
- ۱-۳-۱- شبکه‌های عصبی ..... ۱۴
- ۱-۳-۱-۱- سیستم عصبی زیستی ..... ۱۴
- ۱-۳-۲-۱- مدل ریاضی نرون ..... ۱۶
- ۱-۳-۳-۱- معماری شبکه‌های عصبی ..... ۱۸
- ۱-۳-۴-۱- آموزش شبکه عصبی ..... ۲۲
- ۱-۳-۵-۱- شبکه‌های پس انتشار ..... ۲۴
- ۱-۳-۶-۱- ارزیابی مدل ..... ۳۵

### فصل دوم

- ۱-۲-۱- سرطان ..... ۳۸



- ۳۹ ..... ۱-۱-۲- انواع سرطان‌ها.
- ۴۱ ..... ۲-۱-۲- سرطان سینه
- ۴۶ ..... ۳-۱-۲- استروژن
- ۴۹ ..... HIV-۲-۲
- ۴۹ ..... HIV-1-۱-۲-۲- چیست؟
- ۵۲ ..... ۲-۲-۲- ایدز چیست؟
- ۵۲ ..... ۳-۲-۲- راه‌های انتقال
- ۵۳ ..... ۴-۲-۲- سیر بیماری و علائم آن
- ۵۴ ..... ۵-۲-۲- تشخیص بیماری
- ۵۵ ..... ۶-۲-۲- درمان
- ۵۵ ..... ۳-۲-۲- پارامترهای معرفی خواص دارویی ترکیبات
- ۵۶ ..... ۱-۳-۲- توان دارویی
- ۵۷ ..... ۲-۳-۲- دوز متوسط کشنده
- ۵۸ ..... ۳-۳-۲- بازدارندگی
- ۵۹ ..... ۴-۲-۲- پیشینه کارهای انجام شده در QSAR ترکیبات دارویی
- ۵۹ ..... ۱-۴-۲- سرطان
- ۶۲ ..... HIV-1-۲-۴-۲
- فصل سوم
- ۶۷ ..... ۱-۳-۱- سری داده‌ها

۶۹	۲-۳- نرم افزارهای مورد استفاده
۶۹	۱-۲-۳ Hyperchem
۷۰	۲-۲-۳ Dragon
۷۱	۳-۲-۳ PASW Statistics
۷۲	۴-۲-۳ MATLAB
۷۳	۳-۳- انتخاب توصیف کننده های مهم یا کاهش متغیرهای مستقل
۷۶	۴-۳- انتخاب مدل مناسب
۷۷	۵-۳- فرآیند مدل سازی QSAR
۷۸	۱-۵-۳- مدل رگرسیون خطی چندگانه (MLR)
۷۹	۲-۵-۳- مدل سازی با استفاده از شبکه عصبی مصنوعی (ANN)
۸۷	۳-۵-۳- معماری شبکه عصبی مصنوعی بهینه شده
۸۸	۶-۳- ارزیابی و مقایسه مدل ها
۸۸	۱-۶-۳- ارزیابی مدل های به دست آمده با استفاده از سری تست
۹۴	۲-۶-۳- نتیجه گیری نهایی
فصل چهارم	
۹۶	۱-۴- مقدمه
۹۶	۲-۴- نرم افزارهای مورد استفاده
۹۷	۳-۴- محاسبات و مدل سازی
۹۷	۱-۳-۴- سری داده ها

- ۹۸ ..... ۲-۳-۴- رسم و بهینه‌سازی ساختار مولکولها
- ۱۰۲ ..... ۳-۳-۴- محاسبه توصیف‌کننده‌های مولکولی
- ۱۰۲ ..... ۴-۳-۴- انتخاب توصیف‌کننده‌های مهم
- ۱۰۵ ..... ۴-۴- ایجاد مدل QSAR
- ۱۰۵ ..... ۱-۴-۴- مدل‌سازی با استفاده از شبکه عصبی مصنوعی (ANN)
- ۱۱۴ ..... ۲-۴-۴- معماری شبکه عصبی بهینه شده
- ۱۱۵ ..... ۳-۴-۴- ارزیابی مدل شبکه عصبی مصنوعی
- ۱۲۲ ..... ۴-۴-۴- مدل رگرسیون خطی چندگانه
- فصل پنجم
- ۱۲۴ ..... ۱-۵- بررسی توصیف‌کننده‌های انتخاب شده
- ۱۲۴ ..... ۱-۱-۵- توصیف‌کننده‌های گروه *2D Autocorrelation* [۴۴]
- ۱۲۶ ..... ۲-۱-۵- توصیف‌کننده‌های *3D-MorSE* [۴ و ۴۵]
- ۱۲۷ ..... ۳-۱-۵- توصیف‌کننده‌های *BCUT* [۴۶]
- ۱۲۸ ..... ۴-۱-۵- توصیف‌کننده‌های *Constitutional* [۴۷]
- ۱۲۹ ..... ۵-۱-۵- توصیف‌کننده‌های گروه *GETAWAY* [۴۸]
- ۱۳۰ ..... ۶-۱-۵- توصیف‌کننده‌های *RDF* [۴۹]
- ۱۳۱ ..... ۷-۱-۵- توصیف‌کننده‌های گروه *Topological*
- ۱۳۲ ..... ۸-۱-۵- توصیف‌کننده‌های گروه *WHIM* [۵۱]
- ۱۳۳ ..... ۹-۱-۵- توصیف‌کننده‌ی *Hy*

- ۵-۲- بررسی میزان مشارکت توصیف‌کننده‌ها در مدل شبکه عصبی ..... ۱۳۳
- ۵-۲-۱- بررسی میزان مشارکت توصیف‌کننده‌ها در مدل شبکه عصبی مشتقات سولفونانیلید..... ۱۳۴
- ۵-۲-۲- بررسی میزان مشارکت توصیف‌کننده‌ها در مدل شبکه عصبی مشتقات ۵-اکسوپیرولیدین-۳- ..... ۱۳۵
- ۵-۳- بحث و نتیجه‌گیری در مورد خواص بازدارندگی ترکیبات سولفونانیلید..... ۱۳۶
- ۵-۴- بحث و نتیجه‌گیری در مورد خواص بازدارندگی ترکیبات ۵-اکسوپیرولیدین- ..... ۱۳۷
- ۵-۵- آینده‌نگری..... ۱۳۷

## فهرست اشکال

- شکل ۱-۱-۱- شمای سلول عصبی زیستی ..... ۱۵
- شکل ۱-۲-۱- مدل ریاضی نرون ..... ۱۶
- شکل ۱-۳-۱- برخی از توابع انتقال مورد استفاده در شبکه‌های عصبی ..... ۱۸
- شکل ۱-۴-۱- نمایش یک نرون تک‌لایه ..... ۱۹
- شکل ۱-۵-۱- نمایش یک شبکه تک‌لایه با R ورودی و S نرون ..... ۱۹
- شکل ۱-۶-۱- نمایش یک شبکه تک‌لایه با R ورودی و S نرون ..... ۲۰
- شکل ۱-۷-۱- نمایش یک شبکه چند لایه ..... ۲۱
- شکل ۱-۸-۱- شمای کلی عملکرد شبکه عصبی مورد بحث ..... ۲۳
- شکل ۱-۹-۱- شبکه دو لایه tansig/purelin ..... ۲۵
- شکل ۱-۱۰-۱- شماتیک کلی از یک ماتریس ژاکوبین ..... ۲۸
- شکل ۱-۲-۱- تبدیل تستوسترون به استرودیول بوسیله آروماتازها ..... ۴۸
- شکل ۱-۲-۲- تبدیل آندروستندیون به استرون بوسیله آروماتازها ..... ۴۸
- شکل ۱-۳-۲- مکانیسم عمل آروماتازها در تبدیل آندروژن‌ها به استروژن‌ها ..... ۴۸
- شکل ۱-۴-۲- تصویر شماتیک از ویروس ایدز ..... ۵۰
- شکل ۱-۳-۱- ساختار اصلی ترکیبات مورد بررسی ..... ۶۷
- شکل ۱-۲-۳- نمودار مقادیر  $R^2$  بر حسب تعداد توصیف‌کننده‌ها ..... ۷۶
- شکل ۱-۳-۳- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش تنظیم‌بازین به همراه تابع تبدیل تانژانت هایپربولیک ..... ۸۳
- شکل ۱-۴-۳- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش تنظیم‌بازین به همراه تابع تبدیل لگاریتم زیگموئید ..... ۸۴
- شکل ۱-۵-۳- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش الگوریتم لونبرگ-مارکورت به همراه تابع تبدیل تانژانت هایپربولیک ..... ۸۵

شکل ۳-۶- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش الگوریتم لونبرگ مارکورت به همراه تابع تبدیل لگاریتم زیگموئید ..... ۸۶

شکل ۳-۷- تصویر شماتیک ساختار هندسی شبکه عصبی مصنوعی به دست آمده پس از بهینه‌سازی ..... ۸۷

شکل ۳-۸- نمودار مقادیر پیش‌بینی شده ( $\log(1/C50)$  توسط الف) روش MLR و ب) شبکه عصبی مصنوعی بهینه بر حسب مقادیر تجربی برای سری تست ..... ۸۹

شکل ۳-۹- نمودار مقادیر پیش‌بینی شده ( $\log(1/C50)$  با استفاده از الف) مدل رگرسیون خطی چندگانه و ب) شبکه عصبی مصنوعی بهینه بر حسب مقادیر واقعی به روش رد مرحله‌ای تک‌تک ..... ۹۱

شکل ۳-۱۰- نمودار مقادیر خطای مطلق در ارزیابی مدل الف) رگرسیون خطی چندگانه و ب) شبکه عصبی مصنوعی بهینه به روش رد مرحله‌ای تک‌تک در مقابل مقادیر تجربی ( $\log(1/C50)$  ..... ۹۱

شکل ۴-۱- ساختار پایه مشتقات ۵-اکسوپیرولیدین-۳-کربوکسامید ..... ۹۸

شکل ۴-۲- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش تنظیم‌بازین به همراه تابع تبدیل تانژانت هایپربولیک ..... ۱۰۸

شکل ۴-۳- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش تنظیم‌بازین به همراه تابع تبدیل لگاریتم زیگموئید ..... ۱۰۹

شکل ۴-۴- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش الگوریتم لونبرگ-مارکورت به همراه تابع تبدیل تانژانت هایپربولیک ..... ۱۱۰

شکل ۴-۵- نمودارهای الف) تعداد توصیف‌کننده‌ها، ب) تعداد گره‌ها، ج) تعداد دوره‌های آموزش و د) مقادیر MSE برای تابع آموزش الگوریتم لونبرگ مارکورت به همراه تابع تبدیل لگاریتم زیگموئید ..... ۱۱۱

شکل ۴-۶- تصویر شماتیک ساختار هندسی شبکه عصبی مصنوعی به دست آمده پس از بهینه‌سازی ..... ۱۱۴

شکل ۴-۷- نمودار مقادیر پیش‌بینی شده  $pIC_{50}$  با استفاده از شبکه عصبی مصنوعی بر حسب مقادیر واقعی در سری ارزیابی ..... ۱۱۶

شکل ۴-۸- نمودار مقادیر پیش‌بینی شده  $pIC_{50}$  با استفاده از شبکه عصبی مصنوعی بر حسب مقادیر واقعی در سری تست ..... ۱۱۷

- شکل ۴-۹- نمودار مقادیر پیش‌بینی شده  $pIC_{50}$  با استفاده از مدل شبکه عصبی مصنوعی بهینه شده برحسب مقادیر واقعی به روش رد مرحله‌ای تک‌تک ..... ۱۲۰
- شکل ۴-۱۰- نمودار مقادیر خطای مطلق در ارزیابی مدل شبکه عصبی مصنوعی به روش LOO در مقابل مقادیر تجربی  $pIC_{50}$  ..... ۱۲۰
- شکل ۵-۱- نمودار درصد مشارکت توصیف‌کننده‌های به کار گرفته شده در مدل شبکه عصبی برای ترکیبات سولفونامید ..... ۱۳۵
- شکل ۵-۲- نمودار درصد مشارکت توصیف‌کننده‌های به کار گرفته شده در مدل شبکه عصبی برای ترکیبات ۵-اکسوپیرولیدین-۳-کربوکسامید ..... ۱۳۵

## فهرست جداول

جدول ۱-۳- ترکیبات مورد بررسی در تحقیق	۶۸
جدول ۲-۳- توصیف کننده های محاسبه شده توسط نرم افزار Dragon	۷۰
جدول ۳-۳- کل توصیف کننده های انتخاب شده به همراه کلاس آنها	۷۵
جدول ۴-۳- توصیف کننده های مورد استفاده در مدل سازی شبکه عصبی مصنوعی به همراه کلاس مربوطه	۷۷
جدول ۵-۳- ماتریس همبستگی برای توصیف کننده های انتخاب شده	۷۷
جدول ۶-۳- مقادیر MSE و $R^2$ بدست آمده برای ارزیابی تقاطعی سری آموزش از مدل های ۲ تا ۶	۷۹
جدول ۷-۳- پارامترهای شبکه های بهینه بدست آمده بر اساس مقادیر میانگین مربعات خطای استاندارد (MSE)	۸۷
جدول ۸-۳- نتایج حاصل از مدل رگرسیون خطی چندگانه و شبکه عصبی برای سری تست	۸۸
جدول ۹-۳- نتایج حاصل از ارزیابی مدل رگرسیون خطی چندگانه و شبکه عصبی مصنوعی با استفاده از روش رد مرحله ای تک تک	۹۰
جدول ۱۰-۳- پارامترهای آماری محاسبه شده برای مدل رگرسیون خطی چندگانه و شبکه عصبی مصنوعی	۹۲
جدول ۱۱-۳- مقادیر $R^2$ برای سری های آموزش و تست پس از چندین آزمون Y- تصادفی	۹۳
جدول ۱-۴- جزئیات ساختاری ترکیبات ۱ تا ۳۰	۹۹
جدول ۲-۴- جزئیات ساختاری ترکیبات ۳۰ الی ۶۳	۱۰۰
جدول ۳-۴- جزئیات ترکیبات ۶۴ الی ۱۰۴	۱۰۱
جدول ۴-۴- توصیف کننده های انتخاب شده توسط روش رگرسیون خطی مرحله ای به همراه کلاس مربوطه	۱۰۴
جدول ۵-۴- پارامترهای شبکه بهینه بدست آمده بر اساس مقادیر میانگین مربعات خطای استاندارد (MSE)	۱۱۲



- جدول ۴-۶- توصیف‌کننده‌های مورد استفاده در مدل‌سازی شبکه عصبی مصنوعی بهینه به همراه کلاس مربوطه ..... ۱۱۲
- جدول ۴-۷- ماتریس همبستگی برای توصیف‌کننده‌های انتخاب شده ..... ۱۱۳
- جدول ۴-۸- مقادیر مشاهده شده و پیش‌بینی شده  $pIC_{50}$  برای ترکیبات سری ارزیابی با استفاده از شبکه عصبی بهینه ..... ۱۱۶
- جدول ۴-۹- مقادیر مشاهده شده و پیش‌بینی شده  $pIC_{50}$  برای ترکیبات سری تست با استفاده از شبکه عصبی بهینه ..... ۱۱۷
- جدول ۴-۱۰- نتایج حاصل از ارزیابی مدل شبکه عصبی مصنوعی با استفاده از روش رد مرحله‌ای تک‌تک ..... ۱۱۸
- جدول ۴-۱۱- مقادیر  $R^2$  برای سری‌های آموزش، ارزیابی و تست پس از چندین آزمون  $Y$ -تصادفی ..... ۱۲۱
- جدول ۴-۱۲- پارامترهای آماری محاسبه شده برای مدل شبکه عصبی مصنوعی ..... ۱۲۲

# فصل اول

مقدمه

## ۱-۱- کمومتریکس

کمومتریکس شاخه‌ای از شیمی است که ابزارهای ریاضی و آمار را برای سامان بخشیدن به داده‌های شیمیایی و به‌دست آوردن اطلاعات بیشتر از آنها به کار می‌گیرد. از قابلیت‌های کمومتریکس می‌توان به نگرش چند متغیره به مسئله، ساخت و ارزیابی مدل‌هایی با قابلیت پیش‌بینی، مقایسه نتایج به‌دست آمده از روش‌های مختلف و تعریف و استفاده از شاخص‌هایی که قادرند کیفیت اطلاعات استخراج شده و مدل‌های به‌دست آمده را بسنجند، اشاره کرد [۲،۱]. پیشرفت کامپیوترها و منابع نرم افزاری در سال‌های اخیر باعث رشد روزافزون استفاده از این قابلیت‌ها در شاخه‌های مختلف شیمی، و به ویژه شیمی تجزیه شده است.

از نظر شیمی‌دانان فعالیت‌ها و خواص یک ترکیب ناشی از ویژگی‌های ساختاری آن است. هرگاه مطالعات به صورت ارتباط بین ساختار مولکولی و خواص مشاهده شده مولکول انجام گیرد، به آن ارتباط کمی ساختار - ویژگی (QSPR)<sup>۱</sup> می‌گویند مانند یافتن رابطه‌ای بین خواص نظیر نقطه‌ی جوش و ساختار مولکولی و یا پیش‌بینی فاکتور بازدارنده ترکیبات آلی روی ستون‌های کروماتوگرافی. اما وقتی خواصی از نوع بیولوژیکی (مانند فعالیت دارویی) توصیف می‌شود، اشاره به ارتباط کمی ساختار - فعالیت (QSAR)<sup>۲</sup> خواهد داشت [۳]. در مطالعات QSAR سعی بر این است تا رابطه‌ی هماهنگ میان فعالیت‌های شیمیایی و فیزیکی با ویژگی‌های مولکولی پیدا شود، به گونه‌ای که بتوان این قواعد را برای ارزیابی فعالیت ترکیبات جدید به کار برد. در واقع نتایج این نوع مطالعات علاوه بر شفاف‌سازی نحوه‌ی ارتباط بین خواص مولکول‌ها و ویژگی‌های ساختمانی آنها، به پژوهشگران در پیش‌بینی رفتار مولکول‌های جدید بر اساس رفتار مولکول‌های مشابه کمک می‌کند [۴].

از جمله روش‌هایی که به منظور مطالعه ارتباط خطی ساختار فعالیت مورد استفاده قرار می‌گیرند

<sup>۱</sup>Quantitative Structure-Property Relationship (QSPR)

<sup>۲</sup>Quantitative Structure-Activity Relationship (QSAR)

می‌توان به روش‌های رگرسیون خطی چندگانه<sup>۱</sup>، رگرسیون اجزای اصلی<sup>۲</sup> و روش حداقل مربعات جزئی<sup>۳</sup> اشاره کرد. روش‌های دیگری مانند شبکه‌های عصبی مصنوعی<sup>۴</sup> نیز ارتباط غیرخطی میان ساختار و ویژگی‌های ترکیبات را مورد مطالعه قرار می‌دهند. در تحقیق حاضر مطالعه ارتباط کمی ساختار-فعالیت ترکیبات موردنظر با استفاده از روش‌های خطی و غیرخطی انجام گرفته و نتایج با هم مقایسه شده‌اند.

## ۱-۲- ارتباط کمی ساختار - فعالیت

کورن هانش<sup>۵</sup> برای اولین بار در سال ۱۹۶۴ مفهوم رابطه‌ی کمی ساختار - فعالیت یا همان QSAR را مطرح کرد. او نشان داد که توصیف‌کننده‌های کمی<sup>۶</sup> مولکولی در ترکیبات دارویی با ویژگی‌های بیولوژیک این ترکیبات ارتباط نزدیکی دارند. او و دانشمندان پس از وی، از توصیف‌کننده‌های ساده‌ای برای توصیف ساختار ترکیبات استفاده کردند که توانایی توصیف ویژگی‌های الکترونی، هندسی، فضایی و هیدروفوبیک ترکیبات دارویی را داشتند. هر مطالعه QSAR شامل مراحل زیر می‌باشد:

- جمع‌آوری سری داده‌ها
- رسم ساختارهای مولکولی و بهینه‌سازی آنها
- استخراج توصیف‌کننده‌ها
- بدست آوردن بهترین توصیف‌کننده‌ها برای ورود به مدل
- مدل‌سازی
- ارزیابی مدل

<sup>۱</sup>Multiple Linear Regression(MLR)

<sup>۲</sup>Principal Components Regression (PCR)

<sup>۳</sup>Partial Least Square (PLS)

<sup>۴</sup>Artificial Neural Network (ANN)

<sup>۵</sup> Corren Hansch

<sup>۶</sup> Quantitative Descriptors