

رابطه کمی ساختار-فعالیت تعدادی از مشتق‌های فلاونوئیدی به‌عنوان مهارکننده‌های استیل‌کولین‌استراز بر پایه الگوریتم مونت کارلو

زهرا نظری^۱، شهین احمدی^{۲*} و علی الماسی راد^۳

۱. دانشجوی دکترای عمومی داروسازی، گروه شیمی دارویی، دانشکده داروسازی، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران.

۲. دانشیار گروه شیمی دارویی، دانشکده شیمی دارویی، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران.

۳. استاد گروه شیمی دارویی، دانشکده داروسازی، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران.

چکیده

درسال‌های اخیر، روشی مبتنی بر رویکرد مستقل از بهینه‌سازی مولکول برپایه توصیفگرهای ساختاری مولکولی و توپولوژیکی، به‌عنوان یک رویکرد جذاب در مدل‌سازی QSAR معرفی شده است. یکی از ویژگی‌های کلیدی این رویکرد به استفاده از توصیفگرها برپایه سامانه ورودی-خط ورودی-مولکولی ساده‌شده (SMILES) برای توسعه مدل QSAR اشاره دارد. در این پژوهش، با هدف مطالعه و بررسی QSAR برخی از مشتق‌های فلاونوئیدی به‌عنوان مهارکننده‌های استیل‌کولین‌استراز برپایه الگوریتم مونت کارلو، مجموعه‌ای از ۸۸ مهارکننده استیل‌کولین‌استراز از مقاله‌های معتبر انتخاب شد. ساختار ترکیب‌های با نماد SMILES به‌عنوان فایل‌های ورودی نرم افزار تعریف و به‌طور تصادفی به چهار سری شامل آموزش، آموزش نامرئی، واسنجی و مجموعه اعتبارسنجی تقسیم و مدل‌سازی انجام شد. توصیفگرهای بهینه مورد استفاده در مدل از ترکیب SMILES و نمودار بدون هیدروژن برپایه تابع هدف جدید انتخاب شدند. قدرت پیش‌بینی مدل‌ها با مجموعه‌های اعتبارسنجی، ارزیابی شد. نتیجه‌های به‌دست‌آمده از چهار سری تصادفی نشان داد که مدل‌ها پیشگو و قابل اعتماد از لحاظ آماری، هستند. توصیفگرهای بااهمیت افزاینده و کاهنده فعالیت مهارکنندگی استیل‌کولین‌استراز مشتق‌های فلاونوئید شناسایی شدند. همچنین، pIC_{50} تعداد زیادی از مشتق‌های فلاونوئید طبیعی مهارکننده آنزیم استیل‌کولین‌استراز محاسبه و گزارش شد. با به‌کارگیری مدل به‌دست‌آمده می‌توان pIC_{50} مشتق‌های فلاونوئید را به‌عنوان مهارکننده آنزیم استیل‌کولین‌استراز با روش بهینه‌سازی مونت کارلو و SMILES ترکیب‌ها، پیش‌بینی کرد.

واژه‌های کلیدی: QSAR، فلاونوئیدها، مهارکننده‌های استیل‌کولین‌استراز، SMILES، شاخص آرمانی همبستگی.

بیماری آلزایمر^۱ یا بیماری فراموشی، یک نوع اختلال عملکرد مغزی است که به تدریج توانایی‌های ذهنی بیمار تحلیل می‌رود. آلزایمر رایج‌ترین شکل زوال عقل است. این بیماری علاج‌ناپذیر را نخستین بار روانپزشک آلمانی به نام آلویز آلزایمر در سال ۱۹۰۶ میلادی معرفی کرد. به‌طور کلی، این بیماری در افراد بالای ۶۵ سال بروز پیدا می‌کند؛ هرچند که آلزایمر زودرس (با شیوع کمتر) ممکن است زودتر از این سن نیز رخ دهد [۱]. یک گزارش نشان می‌دهد که حدود ۴۰ تا ۵۰ میلیون نفر در سراسر جهان از آلزایمر رنج می‌برند و این تعداد تا سال ۲۰۵۰ به حدود ۱۳۱/۵ میلیون نفر افزایش خواهد یافت [۲]. علائم این بیماری با از دست‌دادن قدرت حفظ اطلاعات به‌ویژه حافظه موقت در دوران پیری آغاز می‌شود و به تدریج با از دست‌دادن قدرت تشخیص زمان، افسردگی، از دست‌دادن قدرت تکلم، گوشه‌گیری و سرانجام مرگ در اثر ناراحتی‌های تنفسی به پایان می‌رسد [۳].

ارتباط بین علائم بالینی آلزایمر و اختلال در اعصاب کولینرژیک از دیرباز مشخص شده است. طبق فرضیه‌ای کاهش میانجی عصبی استیل کولین به دلیل کاهش تولید آن یا افزایش فعالیت آنزیم تخریب‌کننده آن یعنی استیل کولین استراز (AChE^۲) موجب افت توانایی‌های ذهنی می‌شود [۴]. بنابراین، مهار آنزیم AChE، هدف درمانی برای مدیریت اختلالات مانند میاستنی گراویس، گلوکوم، دمانس و بیماری آلزایمر است [۴].

فلاونوئیدهای غذایی طبیعی با توجه به ویژگی‌های مهارکنندگی AChE، پادامیلوئیدوژن، پاداکسندگی و پادالتهابی آن‌ها به‌عنوان کاندیدهای جایگزین برای درمان آلزایمر توجه چشمگیری را به خود جلب کرده‌اند. شواهد تجربی از این فرضیه حمایت می‌کنند که فلاونوئیدهای ویژه ممکن است تا حدی با تداخل در سایر مسیرهای ایجادکننده آلزایمر مانند تولید و مونتاژ پپتیدهای آمیلوئید-β در توده‌های الیگومری نوروتوکسیک و همچنین، با کاهش تجمع تاو نقش حفاظتی در برابر بیماری آلزایمر ایفا می‌کنند.

شیمی دارویی مدرن بر روی روش‌های طراحی کامپیوتری داروها تمرکز دارد که از نظر زمان و هزینه مناسب هستند [۵ تا ۹]. هر یک از روش‌های شبیه‌سازی کامپیوتری ویژگی‌های بی‌همتای خود را دارند. روش مدل‌سازی رابطه کمی ساختار-فعالیت (QSAR^۳) نقش مهمی در طراحی دارو دارد [۱۰ و ۱۱]. این روش، الگوهای ساختاری مهارکننده‌ها را بهتر شناسایی می‌کند و بینش بهتری در مورد سازوکار اثر دارو ارائه می‌دهد [۱۲ و ۱۳].

یک مدل QSAR توسعه‌یافته در قالب یک معادله ریاضی ارائه شده است که ویژگی‌های شیمیایی مولکول‌های مورد مطالعه را با فعالیت‌های زیستی آن‌ها مرتبط می‌کند [۱۴ و ۱۵]. روشهای متفاوت QSAR، بر پایه توصیفگرهای متفاوتی هستند که محاسبه برخی از آن‌ها آسان، ولی تفسیر آن‌ها دشوار است، مانند توصیفگرهای توپولوژیکی، در حالی که برخی دیگر به راحتی می‌توانند با ساختارهای شیمیایی مرتبط شوند. با این وجود، روش‌شناسی وابسته به ترکیب با در نظر گرفتن محاسبه‌های، آماده‌سازی مولکول، هم‌ترازی و توسعه مدل، زمان‌بر است [۱۶]. روشی مبتنی بر رویکرد مستقل از بهینه‌سازی مولکول که بر پایه ویژگی‌های مولکولی ساختاری و توپولوژیکی است، در سال‌های اخیر به‌عنوان یک رویکرد جذاب

1. Alzheimer
2. Acetylcholinesterase (AChE)
3. Quantitative structure-activity relationship (QSAR)

در مدل‌سازی QSAR معرفی شده است. ویژگی کلیدی این رویکرد به استفاده از توصیفگرها برپایه سامانه ورودی-خط ورودی-مولکولی ساده شده (SMILES^۱) برای توسعه یک مدل QSAR اشاره دارد [۱۷].

در این پژوهش، یک مدل برای ۸۸ مشتق فلاونوئیدی دارای نقش مهارکنندگی استیل‌کولین‌استراز برپایه الگوریتم مونت کارلو ایجاد شد که نسبت به مدل‌های گزارش شده، جامع‌تر بود. تنها SMILES ترکیب‌ها، مستقل از بهینه‌سازی ساختارها، به‌عنوان ورودی مدل به‌کارگرفته شد. این روش برپایه الگوریتم مونت کارلو با یک تابع هدف جدید به نام شاخص آرمانی همبستگی^۲ انجام شد. این تابع جدید افزون بر ضریب تعیین (R^2) مدل، خطای میانگین را نیز در بهینه‌سازی وزن توصیفگرهای مولکولی در نظر می‌گیرد. مدل قابل‌اطمینان و پیشگو با چهار تقسیم تصادفی ارزیابی شد. همچنین، به کمک بهترین مدل به‌دست‌آمده فعالیت مهارکنندگی استیل‌کولین‌استراز تعدادی از مشتقات طبیعی فلاونوئید پیش‌بینی و گزارش شد.

داده‌ها و روش‌ها

این مطالعه به شکل یک مطالعه درون رایانه‌ای^۳ برپایه مهارکنندگی استیل‌کولین‌استراز، یک سری ۸۸ تایی از مشتق‌های فلاونوئید صورت پذیرفت.

سری داده‌ها

در این پژوهش، ۸۸ ترکیب مهارکننده استیل‌کولین‌استراز از مقاله‌های معتبر بین‌المللی انتخاب شدند [۱۸ تا ۲۲]. داده‌های تجربی مربوط به اثر مهارکنندگی این ترکیب‌ها در برابر آنزیم استیل‌کولین‌استراز به‌صورت IC_{50} که میانه غلظت بازدارندگی را نشان می‌دهد به‌کاربرده شد. برای مدل‌سازی رابطه کمی ساختار-فعالیت از تابع $pIC_{50}(-\log IC_{50})$ این ترکیب‌ها که واحد IC_{50} ، مول بر لیتر است به‌عنوان متغیر وابسته استفاده شد، که مقدار آن از ۴/۰۶ تا ۸/۲۳ متغیر است. ساختارهای مولکولی ترکیب‌ها با نرم‌افزار رایگان BIOVIA draw 2016 رسم و به‌صورت نماد SMILES به‌عنوان داده‌های ورودی کورال استفاده شدند.

تفکیک داده‌ها

برای ایجاد مدل QSAR، به‌کمک نرم‌افزار کورال و با روش تصادفی، داده‌ها به چهار سری آموزش، آموزش نامرئی، واسنجی و آزمون تقسیم شدند [۲۳]. به دو سری آموزش و آموزش نامرئی هر کدام در حدود ۳۰ درصد و به دو سری واسنجی و آزمون هر کدام در حدود ۲۰ درصد از داده‌ها اختصاص یافت. این کار ۴ بار به‌صورت تصادفی انجام شد [۲۴]. درصد یکسانی سری‌های آموزش، آموزش نامرئی، واسنجی و آزمون داده‌های هر شکافت با سایر شکافت‌ها در جدول ۱ نشان داده شده است.

1. Simplified molecular-input line-entry system (SMILES)
2. Index of ideality of correlation
3. Insilco

جدول ۱ درصد یکسانی ترکیب‌های سری آموزش، آموزش نامرئی، واسنجی و اعتبارسنجی هر شکافت با سایر شکافت‌ها

سری	شکافت			
	۴	۳	۲	۱
شکافت ۱	۲۸٫۴	۲۸٫۴	۳۰٫۷	۱۰۰
	۳۳٫۳	۳۳٫۳	۳۰٫۲	۱۰۰
	۳۰٫۲	۲۶٫۴	۳۷٫۰	۱۰۰
	۲۸٫۶	۲۹٫۴	۱۱٫۴	۱۰۰
	۱۷٫۶	۲۲٫۹	۴۱٫۲	۱۰۰
شکافت ۲	۲۵٫۰	۲۹٫۵	۱۰۰	
	۲۶٫۴	۳۴٫۰	۱۰۰	
	۲۶٫۴	۳۷٫۷	۱۰۰	
	۲۷٫۸	۲۲٫۹	۱۰۰	
	۱۷٫۶	۱۷٫۱	۱۰۰	
شکافت ۳	۲۸٫۴	۱۰۰		
	۳۳٫۳	۱۰۰		
	۳۰٫۸	۱۰۰		
	۲۲٫۹	۱۰۰		
	۲۲٫۹	۱۰۰		
شکافت ۴	۱۰۰			
	۱۰۰			
	۱۰۰			
	۱۰۰			
	۱۰۰			

انتخاب توصیف‌گرهای مولکولی مناسب

در نرم‌افزار مدل‌سازی کورال، سه گزینه برای انتخاب توصیف‌گرهای بهینه وجود دارد: (۱) توصیف‌گرهای مولکولی بر پایه گراف مولکولی، (۲) توصیف‌گرهای مبتنی بر SMILES و (۳) توصیف‌گرهای هیبریدی که از ترکیب هر دو گراف و SMILES محاسبه می‌شوند [۲۵]. در اینجا، برای ایجاد مدل از توصیف‌گرهای هیبریدی بدون هیدروژن استفاده شد [۲۶]. برای انتخاب توصیف‌گرهای مناسب از بین توصیف‌گرهای هیبریدی یادشده، برپایه جستجوی جامع برای هر مجموعه تصادفی از داده‌ها، چندین مدل متفاوت با گروه‌های متفاوت توصیف‌گرها ایجاد شد و نتیجه‌های مدل‌سازی با این توصیف‌گرها مورد تجزیه و تحلیل قرار گرفت و از بین مدل‌های ایجاد شده، بهترین مجموعه توصیف‌گرها انتخاب شدند. معادله ۱ توصیف‌گرهای بهینه مولکولی را برای مدل PIC50 ترکیب‌های مهارکننده کلاستریل استر ترانسفر^۱ پروتئین، نشان می‌دهد.

1. Cholesteryl ester transfer protein (CETP)

$$\begin{aligned}
DCW(T, N) = & \sum CW(S_k) + \sum CW(SS_k) \\
& + \sum CW(SSS_k) + CW(BOND) \\
& + CW(NOSP) + CW(HALO) \\
& + CW(HARD) + CW(PAIR) \\
& + CW(Cmax) + CW(Nmax) \\
& + CW(Omax) + CW(Smax) \quad (1)
\end{aligned}$$

در معادله ۱، DCW، مجموع ضرایب همبستگی همه توصیفگرهای مولکولی مدل برای هر ترکیب، S_k ، SS_k و SSS_k ترکیبی از قطعه‌های تک عضوی، دو عضوی و سه عضوی هستند. BOND نشان‌دهنده حضور و یا عدم حضور پیوندهای متفاوت دوگانه (=)، سه‌گانه (#) و فضایی (@، @@) است. NOSP نشان‌دهنده حضور و یا عدم حضور اتم‌های نیتروژن، اکسیژن، گوگرد و فسفر در ساختار ترکیبها است. HALO نشان‌دهنده حضور و یا عدم حضور هالوژن‌ها در ترکیبها است. HARD مربوط به حضور BOND، NOSP و HALO در SMILES ساختارها است. توصیفگرهای $Cmax$ ، $Nmax$ ، $Omax$ و $Smax$ به ترتیب، بیشینه تعداد حلقه‌ها، بیشینه تعداد نیتروژن، بیشینه تعداد اکسیژن، و بیشینه تعداد گوگرد در هر ساختار است. CW، وزن همبستگی یک توصیفگر ویژه را نشان می‌دهد. عامل T، ضریب آستانه است و از آن برای تشخیص توصیفگرهای متفاوت مولکولی به‌دست آمده از SMILES به دو گروه فعال و نادر استفاده می‌شود و N تعداد چرخه‌های بهینه‌سازی مونت کارلو را نشان می‌دهد [۲۷].

وزن‌های همبستگی با روش مونت کارلو و با تابع هدف شاخص آرمانی همبستگی که در مقاله‌های متفاوت به‌عنوان تابع هدف مناسب برای یافتن وزن‌های همبستگی توصیفگرها معرفی شده، به‌دست می‌آیند [۲۸]. در نرم افزار کورال معادله نهایی مدل به‌صورت معادله ۲ است.

$$pIC_{50} = C_0 + C_1 \times DCW(T, N) \quad (2)$$

عامل‌های C_0 و C_1 عددهای ثابت‌اند و برای هر مدل پس از انجام بهینه‌سازی ضرایب و ایجاد رابطه خطی بین متغیر وابسته و DCW مربوط به مولکول‌ها به‌دست می‌آیند. در واقع به‌کمک معادله ۲ می‌توان با داشتن DCW هر ساختار، فعالیت (pIC_{50}) آن را پیش‌بینی کرد.

روش مونت کارلو

معادله ۱ برای محاسبه $DCW(T, N)$ ، مربوط به هر ترکیب، به وزن همبستگی هر توصیفگر مولکولی نیاز دارد. این ضرایب به کمک روش بهینه‌یابی مونت کارلو به‌دست می‌آیند. تابع هدف مورد استفاده در این پژوهش که همان شاخص آرمانی همبستگی است، به‌صورت معادله ۳ تعریف می‌شود [۲۹].

$$TF = (r_T + r_{IT} - |r_T - r_{IT}| \times 0.1) + IIC \times 0.5 \quad (3)$$

که در آن، Γ_T و Γ_{IT} به ترتیب ضریب همبستگی برای سری یادگیری و یادگیری نامرئی است. ضریب آرمانی همبستگی برای سری واسنجی با استفاده از معادله ۴ محاسبه می‌شود.

$$IIC = R_{CAL} \times \frac{\min(-MAE_{CAL}, +MAE_{CAL})}{\max(-MAE_{CAL}, +MAE_{CAL})} \quad (۴)$$

داده‌های مربوط به جمع مقادیر مثبت و منفی Δ_k که مشابه خطای مطلق میانگین است، با معادله‌های ۵ تا ۷ محاسبه می‌شوند.

$$-MAE_{CLB} = -\frac{1}{N} \sum_{y=1}^{N^-} |\Delta_k| \quad (\Delta_k < 0) \quad (۵)$$

$$+MAE_{CLB} = +\frac{1}{N} \sum_{y=1}^{N^+} |\Delta_k| \quad (\Delta_k \geq 0) \quad (۶)$$

$$\Delta_k = Obs_k - Calc_k \quad (۷)$$

که در آن‌ها y تعداد MAE‌های مثبت و منفی، N^- تعداد $\Delta_k < 0$ ، N^+ تعداد $\Delta_k \geq 0$ ، Obs_k و $Calc_k$ مقدارهای آزمایشی و محاسباتی نقطه پایانی (pIC_{50}) است و k از ۱ تا N تغییر می‌کند.

دامنه کاربرد^۱ مدل

فضای فیزیکوشیمیایی، ساختاری یا زیستی، دانش یا اطلاعاتی که مجموعه آموزشی مدل برپایه آن ایجاد و برای پیش‌بینی ترکیب‌های جدید استفاده می‌شود، به‌عنوان دامنه کاربرد مدل شناخته می‌شود. با توجه به اصل سوم سازمان توسعه و همکاری اقتصادی (OECD^۲)، دامنه کاربرد برای اعتبارسنجی مدل QSAR ایجادشده، توصیه می‌شود. در نرم افزار کورال، مدل QSAR مبتنی بر مونت کارلو، پراکندگی ویژگی‌های SMILES در سری‌های آموزش، آموزش نامرئی و واسنجی برای دستیابی به دامنه کاربرد استفاده می‌شود. اگر ترکیبی در گستره دامنه کاربرد قرار نگیرد، به‌عنوان یک داده پرت شناخته می‌شود و پیش‌بینی قابل اعتمادی از آن وجود ندارد. در نرم افزار کورال، اگر نابرابری زیر به‌دست آید، یک ترکیب در گستره دامنه کاربرد مدل قرار می‌گیرد، در غیر این صورت، به‌عنوان یک داده پرت شناخته می‌شود.

$$Defect_{compound} < 2 \times \overline{Defect}_T \quad (۸)$$

که \overline{Defect}_T میانگین نقص آماری برای سری آموزش است. نقص آماری (Defect) را می‌توان به‌عنوان مجموع نقص‌های آماری تمام ویژگی‌های موجود در نماد SMILES توصیف کرد (معادله ۹).

$$Defect_{compound} = \sum_{k=1}^{N_A} Defect_{A_k} \quad (۹)$$

1. Applicability domain
2. Organization for economic co-operation and development

که N_A ، تعداد توصیفگرهای فعال برای یک ترکیب است.

نقص آماری برای یک توصیفگر را می‌توان به صورت معادله ۱۰ تعریف کرد.

$$\text{Defect}_{A_K} = \frac{|P_T(A_K) - P_{CAL}(A_K)|}{N_T(A_K) + N_{CAL}(A_K)} \quad (A_K > 0) \quad (10)$$

$$\text{Defect}_{A_K} = 1 \quad (A_K = 0)$$

عامل‌های $P_T(A_K)$ و $P_{CAL}(A_K)$ ، به ترتیب احتمال یک توصیفگر در سری آموزش و واسنجی هستند.

اعتبارسنجی مدل

اعتبارسنجی آماری مدل‌های QSAR ایجادشده برپایه سه روش ارزیابی می‌شود: (۱) اعتبارسنجی داخلی یا اعتبارسنجی متقابل با تعیین R^2 ، IIC، CCC، Q^2 و F در مجموعه آموزشی، (۲) اعتبارسنجی خارجی با تعیین Q_{F1}^2 ، Q_{F2}^2 ، Q_{F3}^2 ، s ، MAE ، \overline{r}_m^2 و Δr_m^2 با استفاده از ترکیب‌های سری آزمون، و (۳) تصادفی‌سازی داده‌ها یا درهم آمیختگی Y (جدول ۲). رابطه ریاضی این عامل‌های آماری در مقاله‌ها ارائه [۱۳، ۱۶ و ۳۰ تا ۳۲] و در جدول ۲ خلاصه شده‌اند. در جدول ۲، Y_{Obs} نقطه پایانی مشاهده‌شده و Y_{Prd} نقطه پایانی پیش‌بینی شده است. R^2 و R_0^2 مقدارهای ضریب تعیین بین نقاط پایانی مشاهده‌شده و پیش‌بینی شده به ترتیب با و بدون در نظر گرفتن عرض از مبدا هستند و R_T^2 میانگین ضریب تعیین مدل‌های تصادفی‌سازی شده است.

جدول ۲ معادله ریاضی معیارهای متفاوت آماری قابلیت پیش‌بینی برای مدل‌های ایجادشده با کورال

معیار پتانسیل پیش‌بینی	نوع اعتبارسنجی
$R^2 = 1 - \frac{\sum(Y_{obs} - Y_{prd})^2}{\sum(Y_{obs} - \bar{Y})^2}$	داخلی
$Q^2 = 1 - \frac{\sum(Y_{prd} - Y_{obs})^2}{\sum(Y_{obs} - \bar{Y}_{train})^2}$	
$Q_{F1}^2 = 1 - \frac{\sum(Y_{per(test)} - Y_{obs(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2}$	خارجی
$Q_{F2}^2 = 1 - \frac{\sum(Y_{prd(test)} - Y_{obs(test)})^2}{\sum(Y_{obs(test)} - \bar{Y}_{ext})^2}$	
$Q_{F3}^2 = 1 - \frac{\sum(Y_{prd(test)} - Y_{obs(test)})^2}{\frac{n_{ext}}{\sum(Y_{obs(test)} - \bar{Y}_{train})^2}}$	
$R_m^2 = R^2 \times \left(1 - \sqrt{R^2 - R_0^2}\right)$	
$CCC = \frac{2 \sum(X - \bar{X})(Y - \bar{Y})}{\sum(X - \bar{X})^2 + \sum(Y - \bar{Y})^2 + n(\bar{X} - \bar{Y})^2}$	
$MAE = \frac{1}{n} \times \sum Y_{obs} - Y_{prd} $	
$C_{R\beta}^2 = R\sqrt{R^2 - R_T^2}$	

نتیجه‌ها و بحث

مدل‌های QSAR

مدل‌های QSAR برای ترکیب‌های مهارکننده کلاسترین استرترانسفر پروتئین، برپایه تابع هدف شاخص آرمانی همبستگی (HIC) و روش مونت کارلو برای چهار توزیع تصادفی داده‌ها به سری‌های آموزش، آموزش نامرئی، واسنجی و مجموعه اعتبارسنجی محاسبه شده عبارت‌اند از:

مدل به‌دست آمده از شکافت ۱:

$$\text{pIC}_{50} = 0.2862(\pm 0.0909) + 0.1542(\pm 0.0025) \times \text{DCW}(1, 11) \quad (11)$$

$(n_{\text{TRN}} = 28, R_{\text{TRN}}^2 = 0.861; n_{\text{ITRN}} = 28, R_{\text{ITRN}}^2 = 0.823;$
 $n_{\text{CAL}} = 17, R_{\text{CAL}}^2 = 0.850; n_{\text{VAL}} = 15, R_{\text{VAL}}^2 = 0.888)$

مدل به‌دست آمده از شکافت ۲:

$$\text{pIC}_{50} = -0.2090(\pm 0.0746) + 0.0991(\pm 0.0011) \times \text{DCW}(1, 10) \quad (12)$$

$(n_{\text{TRN}} = 29, R_{\text{TRN}}^2 = 0.907; n_{\text{ITRN}} = 30, R_{\text{ITRN}}^2 = 0.867;$
 $n_{\text{CAL}} = 18, R_{\text{CAL}}^2 = 0.721; n_{\text{VAL}} = 11, R_{\text{VAL}}^2 = 0.912)$

مدل به‌دست آمده از شکافت ۳:

$$\text{pIC}_{50} = -1.2488(\pm 0.0910) + 0.1793(\pm 0.0022) \times \text{DCW}(1, 11) \quad (13)$$

$(n_{\text{TRN}} = 28, R_{\text{TRN}}^2 = 0.846; n_{\text{ITRN}} = 27, R_{\text{ITRN}}^2 = 0.844;$
 $n_{\text{CAL}} = 17, R_{\text{CAL}}^2 = 0.880; n_{\text{VAL}} = 16, R_{\text{VAL}}^2 = 0.816)$

مدل به‌دست آمده از شکافت ۴:

$$\text{pIC}_{50} = -2.2701(\pm 0.1051) + 0.1390(\pm 0.0017) \times \text{DCW}(1, 8) \quad (14)$$

$(n_{\text{TRN}} = 29, R_{\text{TRN}}^2 = 0.868; n_{\text{ITRN}} = 29, R_{\text{ITRN}}^2 = 0.868;$
 $n_{\text{CAL}} = 18, R_{\text{CAL}}^2 = 0.820; n_{\text{VAL}} = 12, R_{\text{VAL}}^2 = 0.746)$

ویژگی‌های آماری مدل‌های QSAR ایجاد و محاسبه شده برپایه معادله‌های ۱۱ تا ۱۴ در جدول ۳ نشان داده شده است. این جدول نشان می‌دهد که همه مدل‌های QSAR تولید شده از نقطه نظر آماری مناسب هستند و با الزامات معیارهای اعتبارسنجی متفاوت همخوانی دارند. استحکام مدل‌های QSAR ایجاد شده با مقدارهای عددی R^2 و Q^2 که بیش از ۰٫۷ و ۰٫۶ بودند نشان داده شد [۳۰]. افزون‌براین، مقدار عددی R_m^2 برای مجموعه اعتبارسنجی همه مدل‌های QSAR طراحی شده رضایت‌بخش بود و از معیارهای پیشنهادی روی ۱ و همکارانش پیروی می‌کند [۳۱]. قابل اعتماد بودن مدل‌های QSAR ساخته شده نیز با آزمون تصادفی سازی Y تایید شد [۳۲]. پس از چندین بار تکرار، مدل‌های تصادفی جدید توسعه یافتند که مقدار R^2 همه آن‌ها کمتر از ۰٫۱ بود. این نتیجه‌ها نشان می‌دهد که همبستگی بین pIC_{50} و توصیفگرهای مولکولی برپایه

همبستگی تصادفی نیست. افزون‌براین، برای چهار مدل، CR_p^2 بیشتر از ۰/۷۰ به‌دست آمد، که همبستگی غیرشأنسی مدل‌های توسعه‌یافته را تأیید می‌کند [۳۲].

نتایج pic_{50} تجربی که از مقالات معتبر استخراج شده و نتایج پیشگویی چهار مدل در جدول ۴ ارائه شده‌اند. افزون‌برآن، مقدارهای DCW مربوط به هر ترکیب که با قراردادن آن‌ها در معادله‌های ۱۱ تا ۱۴، می‌توان pic_{50} را برای هر ترکیب در هر مدل پیشگویی کرد، آمده است. نمایش SMILES مربوط به مشتق‌های فلاونوئید با مشخص کردن سری آموزش (+)، آموزش نامرئی (-)، واسنجی (#) و اعتبارسنجی (*) نیز در این جدول نشان داده شده است. همچنین، دامنه کاربرد مدل برای هر ترکیب در مدل‌های ۱ تا ۴ برپایه معادله مقدار نقص^۱ (معادله ۱۰)، نشان داده شده است (Y نشان‌دهنده این است که ترکیب در دامنه کاربرد مدل وجود دارد و N نشان‌دهنده این است که ترکیب در دامنه کاربرد مدل وجود ندارد). درصد ترکیب‌ها در دامنه کاربرد مدل‌ها به ترتیب ۹۵، ۸۸، ۹۰ و ۸۵ درصد برای مدل‌های ۱ تا ۴ بود. این نشان داد که چهار مدل قادر به پیش‌بینی بیش از ۸۵ درصد از داده‌های جدید هستند.

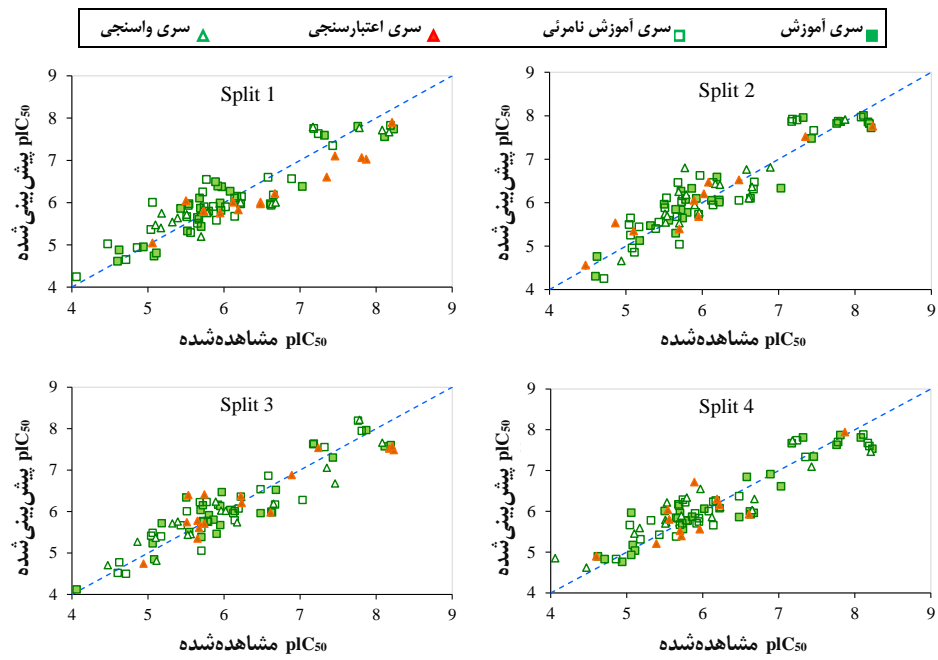
شکل ۱ نمایش گرافیکی مقدارهای داده‌های pic_{50} تجربی در برابر pic_{50} پیش‌بینی شده چهار مدل را نشان می‌دهند. همان‌طور که شکل ۱ نشان می‌دهد، همخوانی خوبی بین داده‌های تجربی و پیش‌بینی شده در مدل‌های پیشنهادی وجود دارد. این نتیجه‌ها تأیید می‌کنند که همه مدل‌های QSAR ساخته‌شده، قابل اعتماد هستند و به خوبی برازش شده‌اند.

جدول ۳ ویژگی‌های آماری برای مدل‌های QSAR مربوط به اثر مهارکنندگی مشتق‌های فلاونوئید برپایه تابع هدف شاخص آرمانی همبستگی برای شکافت‌های متفاوت

مدل*	سری	n	R ²	RMSE	Q ²	Q _{F1} ²	Q _{F2} ²	Q _{F3} ²	CCC	IIC	تصادفی‌سازی Y	cR _p ²	\bar{R}_m^2	ΔR_m^2
۱	آموزش	۲۸	۰/۸۶۱	۰/۳۴۹	۰/۸۰۷				۰/۹۲۵	۰/۶۹۶	۰/۰۱۲	۰/۸۵۵		
	آموزش نامرئی	۲۸	۰/۸۲۳	۰/۴۰۰	۰/۸۰۲				۰/۹۰۳	۰/۵۹۳	۰/۰۳۵	۰/۸۰۵		
	واسنجی	۱۷	۰/۸۵۰	۰/۴۰۶	۰/۷۹۷	۰/۸۵۵	۰/۸۴۸	۰/۸۱۷	۰/۹۱۴	۰/۹۲۱	۰/۰۵۶	۰/۸۲۱	۰/۷۴۵	۰/۱۴۳
	اعتبارسنجی	۱۵	۰/۸۸۸	۰/۴۷۸	۰/۸۴۳				۰/۸۵۱	۰/۴۸۹	۰/۰۶۶		۰/۶۷۹	۰/۱۴۹
۲	آموزش	۲۹	۰/۹۰۷	۰/۳۳۱	۰/۸۹۴				۰/۹۵۱	۰/۸۸۹	۰/۰۲۹	۰/۸۹۲		
	آموزش نامرئی	۳۰	۰/۸۶۷	۰/۴۲۶	۰/۸۴۸				۰/۹۱۵	۰/۶۷۳	۰/۰۱۴	۰/۸۶۰		
	واسنجی	۱۸	۰/۷۲۱	۰/۳۸۰	۰/۶۴۷	۰/۶۸۴	۰/۶۸۳	۰/۸۷۱	۰/۸۴۵	۰/۸۴۸	۰/۰۳۷	۰/۷۰۲	۰/۴۹۸	۰/۲۳۴
	اعتبارسنجی	۱۱	۰/۹۱۲	۰/۳۳۸	۰/۸۵۹				۰/۹۴۴	۰/۶۶۵	۰/۰۹۱		۰/۷۸۷	۰/۰۹۷
۳	آموزش	۲۸	۰/۸۴۶	۰/۳۷۰	۰/۸۲۲				۰/۹۱۶	۰/۷۹۷	۰/۰۱۷	۰/۸۳۷		
	آموزش نامرئی	۲۷	۰/۸۴۴	۰/۳۷۶	۰/۸۲۷				۰/۹۱۹	۰/۸۴۱	۰/۰۳۲	۰/۸۲۸		
	واسنجی	۱۷	۰/۸۸۰	۰/۳۷۷	۰/۸۲۲	۰/۸۷۶	۰/۸۷۶	۰/۸۳۷	۰/۹۱۷	۰/۹۳۷	۰/۰۷۹	۰/۸۳۹	۰/۷۶۹	۰/۱۲۱
	اعتبارسنجی	۱۶	۰/۸۱۶	۰/۴۷۰	۰/۷۶۲				۰/۸۸۲	۰/۷۵۲	۰/۱۲۱		۰/۶۵۵	۰/۱۸۷
۴	آموزش	۲۹	۰/۸۶۸	۰/۳۸۶	۰/۸۵۰				۰/۹۲۹	۰/۸۶۹	۰/۰۶۲	۰/۸۳۶		
	آموزش نامرئی	۲۹	۰/۸۶۸	۰/۳۴۱	۰/۸۴۶				۰/۹۲۴	۰/۵۹۹	۰/۰۴۳	۰/۸۴۶		
	واسنجی	۱۸	۰/۸۲۰	۰/۴۷۶	۰/۷۴۷	۰/۸۰۰	۰/۷۸۴	۰/۷۷۸	۰/۸۶۷	۰/۹۰۵	۰/۰۸۵	۰/۷۷۶	۰/۵۹۳	۰/۲۱۴
	اعتبارسنجی	۱۲	۰/۷۴۶	۰/۴۱۴	۰/۶۰۵				۰/۸۶۳	۰/۷۷۵	۰/۰۹۶		۰/۶۴۶	۰/۰۱۷

* به‌دست‌آمده از شکافت

1. Defect value



شکل ۱ نمایش همبستگی بین pIC₅₀ آزمایشی و محاسباتی برای مدل های ۱ تا ۴

جدول ۴ نمایش SMILES مربوط به مشتق‌های فلاونوئید با مشخص کردن سری آموزش (+)، آموزش نامرئی (-)، واسنجی (#) و اعتبارسنجی (*) و نتیجه‌های آزمایشی و پیش‌گویی هر چهار مدل (قرارگرفتن و قرارنگرفتن در دامنه کاربردی مدل به‌ترتیب با Y و N نشان داده شده است).

شکافت												pIC ₅₀ آزمایشی	شکافت				SMILES	شماره
دامنه کاربرد مدل				پیش‌گویی pIC ₅₀				DCW (T, N)					۴	۳	۲	۱		
۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱			
N	Y	Y	Y	۵,۵۶	۶,۰۳	۵,۷۸	۵,۸۲	۵۶,۳۵	۶۲,۹۲	۶۰,۳۴	۳۵,۸۵	۵,۹۶	*	#	#	+	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4ccc(cc4)c5ccccc5	۱
Y	Y	Y	Y	۶,۲۰	۶,۰۷	۶,۵۹	۵,۸۳	۶۱,۶۸	۶۳,۲۳	۶۸,۵۱	۳۵,۹۳	۶,۱۹	*	-	+	*	COc1cccc1C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۲
N	Y	Y	Y	۶,۲۶	۵,۷۴	۶,۴۳	۶,۱۵	۶۱,۲۸	۶۱,۰۷	۶۶,۹۸	۳۸,۰۲	۶,۱۷	-	#	#	+	COc1cccc(c1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۳
Y	Y	Y	Y	۶,۲۲	۶,۲۳	۶,۸۰	۶,۵۵	۶۱,۱۰	۶۴,۲۶	۷۰,۶۷	۴۰,۶۰	۵,۷۷	-	-	#	-	COc1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۴
Y	Y	Y	Y	۶,۱۰	۶,۰۴	۶,۴۷	۶,۲۷	۵۹,۵۴	۶۳,۰۳	۶۷,۳۲	۳۸,۸۰	۶,۰۸	-	+	*	+	COc1ccc(cc1OC)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۵
Y	Y	N	Y	۶,۵۵	۶,۴۷	۶,۶۳	۶,۳۸	۶۳,۴۶	۶۵,۸۰	۶۸,۹۰	۳۹,۴۷	۵,۹۷	#	+	-	+	COc1cc(cc(OC)c1OC)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۶
N	Y	Y	Y	۵,۹۶	۵,۶۰	۵,۵۳	۵,۷۷	۵۹,۲۲	۶۰,۱۹	۵۷,۸۳	۳۵,۵۵	۵,۷۰	+	-	#	+	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4cccc4Oe5ccccc5	۷
N	Y	Y	Y	۵,۷۹	۶,۰۱	۵,۸۸	۵,۶۷	۵۷,۹۸	۶۲,۸۰	۶۱,۴۱	۳۴,۹۰	۵,۵۱	#	-	-	-	Oc1cccc1C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۸
Y	Y	N	Y	۵,۷۷	۵,۹۲	۵,۷۸	۵,۸۰	۵۷,۸۵	۶۲,۲۱	۶۰,۳۴	۳۵,۷۵	۵,۸۰	+	+	+	+	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4cccc(Oe5ccccc5)c4	۹
Y	Y	Y	Y	۵,۷۵	۵,۶۷	۵,۷۳	۵,۹۹	۵۷,۶۸	۶۰,۶۵	۵۹,۸۸	۳۶,۹۹	۵,۹۵	-	+	-	+	Oc1cccc(c1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۰
Y	Y	Y	Y	۵,۴۰	۵,۷۷	۵,۸۳	۵,۸۴	۵۵,۱۵	۶۱,۲۹	۶۰,۹۲	۳۵,۹۹	۵,۷۲	*	-	+	#	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4ccc(Oe5ccccc5)cc4	۱۱
N	Y	Y	Y	۵,۷۱	۶,۱۷	۶,۱۰	۶,۳۹	۵۷,۴۰	۶۳,۸۳	۶۳,۵۷	۳۹,۵۷	۵,۹۲	-	+	+	+	Oc1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۲
Y	Y	Y	Y	۵,۹۵	۵,۴۷	۵,۹۴	۵,۵۹	۵۹,۱۲	۵۹,۳۰	۶۲,۰۴	۳۴,۳۵	۵,۹۰	-	+	#	-	Cc1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۳
Y	Y	Y	Y	۵,۸۳	۶,۱۳	۵,۶۸	۵,۷۶	۵۸,۲۷	۶۳,۶۲	۵۹,۳۵	۳۵,۴۶	۵,۹۵	-	+	*	*	Clc1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۴
Y	Y	Y	Y	۵,۸۳	۵,۶۱	۵,۶۸	۵,۶۱	۵۸,۲۷	۶۰,۲۱	۵۹,۳۵	۳۴,۴۸	۵,۶۷	-	*	-	+	Brc1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۵
Y	N	Y	Y	۵,۸۷	۵,۸۱	۶,۳۳	۵,۹۱	۵۸,۵۵	۶۱,۵۰	۶۵,۸۹	۳۶,۴۴	۵,۸۶	+	+	+	-	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4ccc(cc4)N5CCCC5	۱۶
N	Y	Y	Y	۵,۶۶	۶,۰۱	۵,۹۴	۶,۱۱	۵۷,۰۵	۶۲,۸۵	۶۲,۰۳	۳۷,۷۴	۶,۱۴	-	+	-	-	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4cccc5ccccc45	۱۷
Y	Y	Y	Y	۵,۷۸	۵,۷۳	۵,۴۷	۵,۵۴	۵۷,۹۱	۶۰,۹۲	۵۷,۲۸	۳۴,۰۷	۵,۳۲	-	#	+	#	O=C1C=C(Oc2cc(CCCCCCN3CCCC3)ccc12)c4ccc(cc4)c5ccccc5	۱۸
N	Y	Y	Y	۵,۶۷	۵,۰۶	۵,۰۴	۵,۲۰	۵۷,۱۳	۵۶,۶۵	۵۲,۹۲	۳۱,۸۵	۵,۷۰	-	-	-	#	CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(CCCCCCN4CCCC4)cc3O2	۱۹
Y	Y	Y	Y	۵,۸۴	۵,۷۸	۵,۸۴	۵,۶۵	۵۸,۳۶	۶۱,۳۳	۶۱,۰۲	۳۴,۷۶	۵,۶۵	+	*	+	-	CCN(CC)CCCCOc1ccc2C(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C	۲۰

ادامه جدول ۴

شکافت												pIC ₅₀ آزمایشی	شکافت				SMILES	شماره
دامنه کاربرد مدل				پیشگویی pIC ₅₀				DCW (T, N)										
۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱		۴	۳	۲	۱		
Y	Y	Y	N	۵,۸۰	۵,۴۶	۵,۵۹	۵,۲۹	۵۸,۰۶	۵۹,۲۴	۵۸,۴۸	۳۲,۴۶	۵,۵۶	*	#	+	-	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۲۱
Y	Y	Y	Y	۵,۷۳	۵,۴۴	۵,۵۴	۵,۲۳	۵۷,۵۶	۵۹,۱۳	۵۷,۹۹	۳۲,۶۹	۵,۵۲	#	#	#	+	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۲۲
Y	Y	Y	Y	۵,۱۸	۴,۸۵	۴,۹۶	۴,۷۴	۵۳,۵۷	۵۵,۳۲	۵۲,۰۷	۲۸,۸۵	۵,۰۸	+	+	-	+	<chem>CN(C)C(C)C4CCCC4C1=CC(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۲۳
Y	Y	Y	Y	۵,۷۰	۵,۷۴	۵,۷۴	۵,۷۲	۵۷,۳۶	۶۱,۱۰	۶۰,۰۲	۳۵,۲۳	۵,۵۱	#	*	#	#	<chem>CCN(CC)CCCCCOc1ccc2C(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۲۴
N	N	N	Y	۵,۶۶	۵,۴۲	۵,۴۹	۵,۲۷	۵۷,۰۷	۵۹,۰۱	۵۷,۴۹	۳۲,۹۳	۵,۰۴	-	-	-	-	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۲۵
N	Y	Y	Y	۵,۵۹	۵,۴۰	۵,۴۴	۵,۴۰	۵۶,۵۷	۵۸,۸۹	۵۶,۹۹	۳۳,۱۶	۵,۱۷	#	-	-	#	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۲۶
N	N	Y	Y	۵,۰۴	۴,۸۲	۴,۸۶	۴,۸۱	۵۲,۵۷	۵۵,۰۹	۵۱,۰۷	۲۹,۳۲	۵,۱۱	+	#	-	+	<chem>CN(C)C(C)C4CCCC4C1=CC(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۲۷
Y	Y	Y	N	۵,۵۶	۵,۷۱	۵,۶۵	۵,۷۹	۵۶,۳۶	۶۰,۸۶	۵۹,۰۲	۳۵,۷۰	۵,۷۴	+	*	+	*	<chem>CCN(CC)CCCCCOc1ccc2C(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۲۸
Y	Y	Y	Y	۵,۵۲	۵,۳۹	۵,۳۹	۵,۴۴	۵۶,۰۷	۵۸,۷۸	۵۶,۴۹	۳۳,۴۰	۵,۷۰	*	+	*	+	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۲۹
Y	N	Y	Y	۵,۴۵	۵,۳۷	۵,۳۴	۵,۴۷	۵۵,۵۷	۵۸,۶۶	۵۵,۹۹	۳۳,۶۳	۵,۱۰	#	#	*	#	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۳۰
Y	N	Y	Y	۴,۹۰	۴,۷۸	۴,۷۶	۴,۸۸	۵۱,۵۷	۵۴,۸۶	۵۰,۰۷	۲۹,۷۹	۴,۶۲	+	-	+	+	<chem>CN(C)C(C)C4CCCC4C1=CC(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۳۱
Y	N	Y	Y	۵,۴۳	۵,۶۷	۵,۵۵	۵,۸۷	۵۵,۳۶	۶۰,۶۳	۵۸,۰۲	۳۶,۱۶	۵,۴۳	-	-	-	+	<chem>CCN(CC)CCCCCOc1ccc2C(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۳۲
Y	Y	Y	Y	۵,۳۸	۵,۳۵	۵,۳۰	۵,۵۱	۵۵,۰۷	۵۸,۵۴	۵۵,۴۹	۳۳,۸۶	۵,۶۵	-	*	+	-	<chem>CN(C)c1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCCC4)cc3O2</chem>	۳۳
Y	Y	Y	Y	۴,۷۶	۴,۷۴	۴,۶۶	۴,۹۵	۵۰,۵۷	۶۲,۵۴	۴۹,۰۷	۳۰,۲۶	۴,۹۴	+	*	#	+	<chem>CN(C)C(C)C4CCCC4C1=CC(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۳۴
Y	Y	Y	Y	۴,۶۲	۴,۷۱	۴,۵۶	۵,۰۳	۴۹,۵۷	۵۴,۳۹	۴۸,۰۷	۳۰,۷۳	۴,۴۷	#	#	*	-	<chem>CN(C)C(C)C4CCCC4C1=CC(=O)C=C(Oc2c1)c3ccc(cc3)N(C)C</chem>	۳۵
Y	Y	Y	Y	۷,۳۴	۶,۶۸	۷,۶۶	۷,۱۰	۶۹,۱۶	۶۷,۱۶	۷۹,۳۳	۴۴,۱۹	۷,۴۶	+	#	-	*	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(O)cc3O2</chem>	۳۶
Y	Y	Y	Y	۷,۰۹	۷,۳۱	۷,۴۸	۷,۳۵	۶۷,۳۷	۷۱,۲۰	۷۷,۴۸	۴۵,۷۸	۷,۴۳	#	+	+	-	<chem>COc1cc(O)c(O)c1C2=CC(=O)c3ccc(O)cc3O2</chem>	۳۷
Y	Y	Y	Y	۷,۳۳	۷,۰۶	۷,۵۲	۶,۶۱	۶۹,۰۸	۶۹,۵۹	۷۷,۸۹	۴۰,۹۶	۷,۳۵	-	#	*	*	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(O)cc3O2</chem>	۳۸
Y	Y	Y	Y	۷,۸۸	۷,۵۷	۸,۰۱	۷,۵۶	۷۳,۰۳	۷۲,۹۳	۸۲,۸۱	۴۷,۱۵	۸,۱۱	-	+	+	+	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCN4CCC(C)CC4)cc3O2</chem>	۳۹
Y	Y	Y	Y	۷,۸۱	۷,۶۶	۷,۹۷	۷,۷۲	۷۲,۵	۷۳,۵۱	۸۲,۴۴	۴۸,۱۸	۸,۰۸	+	#	-	#	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCN4CCN(CCO)CC4)cc3O2</chem>	۴۰

ادامه جدول ۴

شکافت												pIC ₅₀ آزمایشی	شکافت				SMILES	شماره
دامنه کاربرد مدل				پیشگویی pIC ₅₀				DCW (T, N)					۴	۳	۲	۱		
۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱							
Y	Y	Y	Y	۷,۹۴	۷,۹۷	۷,۹۱	۷,۰۳	۷۳,۴۵	۷۵,۴۸	۸۱,۸۸	۴۳,۷۰	۷,۸۷	*	+	#	*	<chem>COc1ccc(cc1OC)C2=CC(=O)c3ccc(OCCN4CCC(C)CC4)cc3O2</chem>	۴۱
Y	Y	N	Y	۷,۷۰	۸,۲۲	۷,۸۷	۷,۷۷	۷۱,۷۴	۷۷,۰۹	۸۱,۴۷	۴۸,۵۲	۷,۷۸	-	#	-	#	<chem>COc1cc(cc(OC)c1OC)C2=CC(=O)c3ccc(OCCN4CCC(C)CC4)cc3O2</chem>	۴۲
Y	Y	N	Y	۷,۸۷	۷,۹۵	۷,۸۶	۷,۰۶	۷۲,۹۵	۷۵,۳۶	۸۱,۳۸	۴۳,۹۳	۷,۸۱	+	-	+	*	<chem>COc1ccc(cc1OC)C2=CC(=O)c3ccc(OCCCN4CCC(C)CC4)cc3O2</chem>	۴۳
Y	Y	Y	Y	۷,۶۳	۸,۲۰	۷,۸۲	۷,۸۱	۷۱,۳۴	۷۶,۹۷	۸۰,۹۷	۴۸,۷۵	۷,۷۶	+	-	+	+	<chem>COc1cc(cc(OC)c1OC)C2=CC(=O)c3ccc(OCCCN4CCC(C)CC4)cc3O2</chem>	۴۴
Y	Y	Y	Y	۷,۸۱	۷,۵۶	۷,۹۶	۷,۶۰	۷۲,۵۳	۷۲,۸۲	۸۲,۳۱	۴۷,۳۹	۷,۳۲	+	-	+	+	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCC(C)CC4)cc3O2</chem>	۴۵
Y	Y	Y	Y	۷,۷۴	۷,۶۴	۷,۹۲	۷,۷۶	۷۲,۰۰	۷۳,۳۹	۸۱,۹۴	۴۸,۴۱	۷,۱۸	#	+	-	-	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCN4CCN(CCO)CC4)cc3O2</chem>	۴۶
Y	Y	Y	Y	۷,۷۴	۷,۵۴	۷,۹۱	۷,۶۳	۷۲,۰۳	۷۲,۰۰	۸۱,۸۱	۴۷,۶۲	۷,۲۴	-	*	-	-	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCC(C)CC4)cc3O2</chem>	۴۷
Y	Y	Y	Y	۷,۶۷	۷,۶۳	۷,۸۷	۷,۷۹	۷۱,۵۱	۷۳,۲۷	۸۱,۴۴	۴۸,۶۴	۷,۱۷	+	-	-	#	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCN(CCO)CC4)cc3O2</chem>	۴۸
Y	Y	Y	Y	۷,۶۷	۷,۵۲	۷,۸۶	۷,۶۷	۷۱,۵۳	۷۲,۵۸	۸۱,۳۱	۴۷,۸۶	۸,۱۷	-	*	+	#	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCC(C)CC4)cc3O2</chem>	۴۹
Y	Y	Y	Y	۷,۶۰	۷,۶۱	۷,۸۲	۷,۸۳	۷۱,۰۱	۷۳,۱۶	۸۰,۹۴	۴۸,۸۸	۸,۱۹	-	+	+	-	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCN(CCO)CC4)cc3O2</chem>	۵۰
Y	Y	Y	Y	۷,۵۳	۷,۴۸	۷,۷۶	۷,۷۴	۷۰,۵۳	۷۲/۳۵	۸۰,۳۱	۴۸,۳۲	۸,۲۳	+	*	*	+	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCC(C)CC4)cc3O2</chem>	۵۱
Y	Y	Y	Y	۷,۴۶	۷,۵۷	۷,۷۲	۷,۹۰	۷۰,۰۱	۷۲,۹۲	۷۹,۹۵	۴۹,۳۵	۸,۲۱	#	*	+	*	<chem>COc1ccc(cc1)C2=CC(=O)c3ccc(OCCCCCN4CCN(CCO)CC4)cc3O2</chem>	۵۲
Y	Y	Y	Y	۵,۲۱	۵,۷۶	۵,۴۰	۵,۶۴	۵۳,۸۲	۶۱,۲۱	۵۶,۵۵	۳۴,۷۱	۵,۳۹	*	#	-	#	<chem>COc1cc2OC(=CC(=O)c2cc1OC)c3ccc(cc3)[O]4CCN(Cc5ccccc5)CC4</chem>	۵۳
Y	Y	Y	Y	۴,۸۳	۵,۲۷	۵,۵۳	۴,۹۴	۵۱,۰۹	۵۸,۰۴	۵۷,۸۶	۳۰,۱۵	۴,۸۶	-	#	*	-	<chem>COc1cc2OC(=CC(=O)c2cc1OC)c3ccc(OCC4CCN(Cc5ccccc5)CC4)cc3</chem>	۵۴
Y	Y	Y	Y	۵,۳۱	۵,۷۲	۵,۱۲	۵,۷۵	۵۴,۵۵	۶۰,۹۵	۵۳,۷۶	۳۵,۴۲	۵,۱۸	-	+	+	#	<chem>COc1cc2OC(=CC(=O)c2cc1OC)c3ccc(cc3)[O]4CCN(Cc5ccccc5)CC4</chem>	۵۵
Y	Y	Y	Y	۴,۹۳	۵,۲۳	۵,۲۵	۵,۰۵	۵۱,۸۲	۵۷,۷۹	۵۵,۰۶	۳۰,۸۶	۵,۰۶	+	+	-	*	<chem>COc1cc2OC(=CC(=O)c2cc1OC)c3ccc(OCC4CCN(Cc5ccccc5)CC4)c3</chem>	۵۶
Y	Y	N	Y	۶,۶۱	۶,۲۸	۶,۳۳	۶,۳۸	۶۳,۹۰	۶۴,۵۷	۶۵,۹۳	۳۹,۵۲	۷,۰۳	+	-	+	+	<chem>COc1cc2OC=C(C(=O)c2cc1OC)c3ccc(cc3)[O]4CCN(Cc5ccccc5)CC4</chem>	۵۷
Y	Y	N	Y	۶,۲۳	۵,۷۹	۶,۴۶	۵,۶۸	۶۱,۱۷	۶۱,۴۱	۶۷,۳	۳۴,۹۶	۶,۱۴	-	-	-	-	<chem>COc1cc2OC=C(C(=O)c2cc1OC)c3ccc(OCC4CCN(Cc5ccccc5)CC4)cc3</chem>	۵۸
Y	Y	Y	Y	۶,۷۱	۶,۲۴	۶,۰۵	۶,۴۹	۶۴,۶۴	۶۴,۳۲	۶۳,۱۳	۴۰,۲۳	۵,۸۹	*	#	*	+	<chem>COc1cc2OC=C(C(=O)c2cc1OC)c3ccc(cc3)[O]4CCN(Cc5ccccc5)CC4</chem>	۵۹
N	Y	N	Y	۶,۳۴	۵,۷۵	۶,۱۸	۵,۷۹	۶۱,۹۱	۶۱,۱۵	۶۴,۴۳	۳۵,۶۷	۵,۷۹	#	-	#	#	<chem>COc1cc2OC=C(C(=O)c2cc1OC)c3ccc(OCC4CCN(Cc5ccccc5)CC4)c3</chem>	۶۰

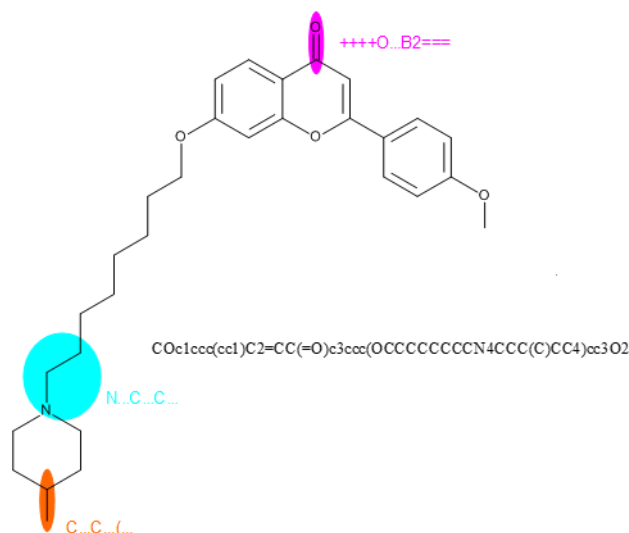
ادامه جدول ۴

شکافت												pIC ₅₀ آزمایشی	شکافت				SMILES	شماره
دامنه کاربرد مدل				پیشگویی pIC ₅₀				DCW (T, N)					۴	۳	۲	۱		
۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱			
Y	Y	Y	Y	۴,۸۵	۴,۱۲	۳,۹۵	۴,۲۵	۵,۱۲۳	۵,۰۶	۴,۱۹۱	۲۵,۶۶	۴,۰۶	#	+	-	-	<chem>Cc1c(O)ccc2C(=O)C=C(Oc12)c3ccccc3</chem>	۶۱
N	Y	Y	Y	۶,۱۴	۶,۰۴	۶,۲۵	۵,۸۷	۶,۰۴۷	۶۲,۹۸	۶۵,۱۱	۳۶,۱۷	۵,۶۹	+	+	#	+	<chem>Cc1c(OCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۶۲
Y	Y	N	Y	۶,۰۷	۶,۰۲	۶,۲۰	۵,۹۰	۵۹,۹۷	۶۲,۸۷	۶۴,۶۱	۳۶,۴۱	۶,۰۲	+	#	*	-	<chem>Cc1c(OCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۶۳
Y	Y	Y	Y	۶,۰۰	۶,۰۰	۶,۱۵	۵,۹۴	۵۹,۴۷	۶۲,۷۵	۶۴,۱۱	۳۶,۶۴	۶,۶۲	+	+	#	-	<chem>Cc1c(OCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۶۴
Y	Y	Y	Y	۵,۹۳	۵,۹۸	۶,۱۰	۵,۹۸	۵۸,۹۷	۶۲,۶۳	۶۳,۶۲	۳۶,۸۸	۶,۶۱	*	*	#	+	<chem>Cc1c(OCCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۶۵
Y	Y	Y	Y	۵,۸۶	۵,۹۶	۶,۰۵	۶,۰۱	۵۸,۴۷	۶۲,۵۲	۶۳,۱۲	۳۷,۱۱	۶,۱۲	#	-	+	*	<chem>Cc1c(OCCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۶۶
Y	Y	Y	Y	۶,۱۷	۶,۲۳	۶,۴۶	۶,۱۱	۶۰,۷۲	۶۴,۲۳	۶۷,۲۴	۳۷,۷۴	۵,۶۸	-	-	-	+	<chem>CC1CCCN1CCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4</chem>	۶۷
Y	Y	Y	Y	۶,۱۰	۶,۲۱	۶,۴۱	۶,۱۵	۶۰,۲۲	۶۴,۱۱	۶۶,۷۴	۳۷,۹۸	۶,۲۳	#	*	#	-	<chem>CC1CCCN1CCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4</chem>	۶۸
Y	Y	Y	Y	۶,۰۳	۶,۱۹	۶,۳۶	۶,۱۸	۵۹,۷۳	۶۳,۹۹	۶۶,۲۴	۳۸,۲۱	۶,۶۵	#	#	#	-	<chem>CC1CCCN1CCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4</chem>	۶۹
Y	Y	Y	Y	۵,۹۶	۶,۱۷	۶,۳۱	۶,۲۲	۵۹,۲۳	۶۳,۸۸	۶۵,۷۴	۳۸,۴۴	۶,۶۷	+	-	-	*	<chem>CC1CCCN1CCCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4</chem>	۷۰
Y	Y	Y	Y	۵,۸۹	۶,۱۶	۶,۲۶	۶,۲۵	۵۸,۷۳	۶۳,۷۶	۶۵,۲۴	۳۸,۶۸	۵,۷۲	#	+	+	-	<chem>CC1CCCN1CCCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4</chem>	۷۱
Y	Y	Y	Y	۶,۱۹	۶,۴۲	۶,۱۶	۵,۹۰	۶۱,۵۶	۶۵,۴۵	۶۴,۰۰	۳۶,۳۹	۵,۷۴	-	*	-	-	<chem>Cc1c(OCCN2CCC(O)CC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۲
Y	Y	Y	Y	۶,۲۲	۶,۴۰	۶,۱۱	۵,۹۴	۶۱,۰۶	۶۵,۳۳	۶۳,۷۰	۳۶,۶۲	۵,۵۳	#	*	-	-	<chem>Cc1c(OCCN2CCC(O)CC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۳
Y	Y	Y	Y	۶,۱۵	۶,۳۸	۶,۰۶	۵,۹۷	۶۰,۵۶	۶۵,۲۱	۶۳,۲۰	۳۶,۸۶	۶,۲۲	*	*	+	+	<chem>Cc1c(OCCCN2CCC(O)CC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۴
Y	Y	Y	Y	۶,۰۸	۶,۳۶	۶,۰۱	۶,۰۱	۶۰,۰۶	۶۵,۰۹	۶۲,۷۰	۳۷,۰۹	۶,۲۲	-	-	+	#	<chem>Cc1c(OCCCN2CCC(O)CC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۵
Y	Y	Y	Y	۶,۰۱	۶,۳۴	۵,۹۶	۶,۰۵	۵۹,۵۶	۶۴,۹۸	۶۲,۲۰	۳۷,۳۳	۵,۵۰	+	+	+	*	<chem>Cc1c(OCCCCN2CCC(O)CC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۶
Y	Y	Y	Y	۵,۹۳	۵,۹۸	۶,۱۰	۵,۹۸	۵۸,۹۷	۶۲,۶۳	۶۳,۶۲	۳۶,۸۸	۶,۶۳	-	-	+	#	<chem>Cc1c(OCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۷
Y	Y	Y	Y	۵,۸۶	۵,۹۶	۶,۰۵	۶,۰۱	۵۸,۴۷	۶۲,۵۲	۶۳,۱۲	۳۷,۱۱	۶,۴۸	+	+	-	*	<chem>Cc1c(OCCCCN2CCCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۷۸
Y	Y	Y	Y	۵,۷۸	۵,۷۸	۶,۰۷	۵,۷۶	۵۲,۸۹	۶۱,۳۲	۶۳,۳۲	۳۵,۴۶	۵,۷۸	-	+	+	#	<chem>CN1CCN(CCCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4)CC1</chem>	۷۹
Y	Y	Y	Y	۵,۷۱	۵,۷۶	۶,۰۲	۵,۷۹	۵,۳۹	۶۱,۲۰	۶۲,۸۲	۳۵,۶۹	۵,۷۳	+	+	-	*	<chem>CN1CCN(CCCCOe2ccc3C(=O)C=C(Oc3e2C)c4ccccc4)CC1</chem>	۸۰

ادامه جدول ۴

شکافت												pIC ₅₀ آزمایشی	شکافت				SMILES	شماره
دامنه کاربرد مدل				پیشگویی pIC ₅₀				DCW (T, N)					۴	۳	۲	۱		
۴	۳	۲	۱	۴	۳	۲	۱	۴	۳	۲	۱							
Y	Y	Y	Y	۶,۹۱	۶,۸۸	۶,۸۱	۶,۵۷	۶۶,۰۶	۶۸,۴۷	۷۰,۷۹	۴۰,۷۰	۶,۸۹	+	*	#	-	<chem>CCN(CC)CCCCOc1ccc2C(=O)C=C(Oc2c1C)c3ccccc3</chem>	۸۱
Y	N	Y	Y	۶,۸۴	۶,۸۷	۶,۷۶	۶,۶۰	۶۵,۵۶	۶۸,۳۵	۷۰,۲۹	۴۰,۹۴	۶,۵۸	+	-	#	-	<chem>CCN(CC)CCCCCOc1ccc2C(=O)C=C(Oc2c1C)c3ccccc3</chem>	۸۲
Y	N	Y	Y	۶,۳۷	۶,۵۴	۶,۵۲	۵,۹۷	۶۲,۱۷	۶۶,۲۵	۶۷,۸۷	۳۶,۸۶	۶,۴۸	-	-	*	*	<chem>CN(C)CCCCOc1ccc2C(=O)C=C(Oc2c1C)c3ccccc3</chem>	۸۳
Y	N	Y	Y	۶,۳۰	۶,۵۲	۶,۴۷	۶,۰۱	۶۱,۶۷	۶۶,۱۴	۶۷,۳۷	۳۷,۱۰	۶,۶۸	#	+	-	#	<chem>CN(C)CCCCCOc1ccc2C(=O)C=C(Oc2c1C)c3ccccc3</chem>	۸۴
Y	Y	Y	Y	۶,۰۳	۵,۵۱	۵,۶۹	۵,۹۷	۵۹,۷۴	۵۹,۵۷	۵۹,۵۰	۳۶,۸۶	۵,۵۴	*	+	#	+	<chem>Cc1c(OCCCCN2CCOCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۸۵
Y	Y	Y	Y	۵,۹۷	۵,۴۹	۵,۶۴	۶,۰۱	۵۹,۲۵	۵۹,۴۶	۵۹,۰۰	۳۷,۱۰	۵,۰۶	+	-	-	-	<chem>Cc1c(OCCCCN2CCOCC2)ccc3C(=O)C=C(Oc13)c4ccccc4</chem>	۸۶
Y	Y	N	N	۴,۹۰	۴,۵۲	۴,۳۰	۴,۶۲	۵۱,۵۸	۵۳,۱۸	۴۵,۴۸	۲۸,۰۶	۴,۶۰	*	-	+	+	<chem>COC(=O)[C@@H]1CCCN1CCCCOc2ccc3C(=O)C=C(Oc3c2C)c4ccccc4</chem>	۸۷
N	Y	N	N	۴,۸۳	۴,۵۰	۴,۲۵	۴,۶۵	۵۱,۰۸	۵۳,۰۶	۴۴,۹۸	۲۸,۲۹	۴,۷۱	+	-	-	-	<chem>COC(=O)[C@@H]1CCCN1CCCCOc2ccc3C(=O)C=C(Oc3c2C)c4ccccc4</chem>	۸۸

تفسیر مکانیکی، یکی دیگر از نتیجه‌های مهم مدل‌سازی QSAR است. به این معنی که توصیفگر افزایشده و کاهشده فعالیت گزارش شود و برپایه این نتیجه‌ها بتوان ترکیب‌هایی با فعالیت بالا پیشنهاد کرد. در روش بهینه‌سازی مونت کارلو به کمک نرم افزار کورال، با توجه به علامت و مقدارهای وزن همبستگی توصیفگرها می‌توان توصیفگرهای افزایشده یا کاهشده فعالیت را تشخیص داد. برای تشخیص عوامل افزایشده و کاهشده، سه اجرای مستقل مونت کارلو روی هر شکافت انجام می‌شود و ضرایب مربوط به توصیفگرها دارای سه حالت متفاوت خواهند بود. اگر ضریب یک توصیفگر در هر سه اجرا مثبت باشد، در این صورت به‌عنوان توصیفگر افزایشده معرفی می‌شود. در صورتی که ضریب توصیفگر منفی باشد، توصیفگر کاهشده خواهد بود و در صورتی که این ضریب در برخی اجراها مثبت و در برخی از اجراها منفی شود، این توصیفگر تعریف نشده است. از توصیفگرهای افزایشده، می‌توان به حضور همزمان اکسیژن و پیوند دوگانه (==O.B2+++), حضور دست‌کم یک حلقه (1.....), توالی یک نیتروژن آلیفاتیک و دو کربن آلیفاتیک (N...C...C...), و توالی دو کربن آلیفاتیک همراه با شاخه‌دارشدن ((...C...C...)) اشاره کرد. یک توصیفگر کاهشده مشترک در هر چهار شکافت هم وجود دارد که توالی کربن آلیفاتیک، شاخه‌دارشدن و پیوند دوگانه ((...C...=...)) است.



شکل ۲ مشتق فلاونوئید مهارکننده آنزیم استیل کولین استراز (ترکیب شماره ۵۱) با بالاترین اثر مهارکنندگی در بین ترکیب‌ها موجود در مدل همراه با برخی عامل‌های ساختاری افزایشده فعالیت

در مورد توصیفگر توالی یک نیتروژن آلیفاتیک و دو کربن آلیفاتیک (N...C...C...), به‌عنوان توصیفگر افزایشده، می‌توان به مثال‌هایی اشاره کرد. ترکیب شماره ۸۱ (pIC50=۶/۸۹) و ۸۳ (pIC50=۶/۴۸) از نظر ساختاری خیلی شبیه به هم هستند، با این تفاوت که ترکیب شماره ۸۱ دارای این توصیفگر و ترکیب ۸۳ بدون این توصیفگر است.

در مورد توصیفگر افزایشنده ((...C...C...C))، با بررسی داده‌ها، به صورت کلی می‌توان گفت با افزایش تعداد n که همان افزایش تعداد کربن‌های آلیفاتیک است، فعالیت افزایش پیدا کرده است. در چندین چارچوب کلی متفاوت از داده‌ها، به صورت کلی این روند افزایشی مشاهده می‌شود. برای مثال، ترکیب شماره ۲۸ ($pIC_{50}=5.74$) در مقایسه با ترکیب شماره ۲۴ ($pIC_{50}=5.51$)، ترکیب شماره ۲۹ ($pIC_{50}=5.70$) در مقایسه با ترکیب شماره ۲۵ ($pIC_{50}=5.04$)، ترکیب شماره ۴۹ ($pIC_{50}=8.17$) در مقایسه با ترکیب شماره ۴۵ ($pIC_{50}=7.32$)، ترکیب شماره ۵۰ ($pIC_{50}=8.19$) در مقایسه با ترکیب شماره ۴۸ ($pIC_{50}=7.17$)، ترکیب شماره ۶۴ ($pIC_{50}=6.62$) در مقایسه با ترکیب شماره ۶۲ ($pIC_{50}=5.69$)، ترکیب شماره ۶۹ ($pIC_{50}=6.64$) در مقایسه با ترکیب شماره ۶۷ ($pIC_{50}=5.97$)، ترکیب شماره ۷۴ ($pIC_{50}=6.22$) در مقایسه با ترکیب شماره ۷۲ ($pIC_{50}=5.74$)، ترکیب شماره ۸۴ ($pIC_{50}=6.68$) در مقایسه با ترکیب شماره ۸۳ ($pIC_{50}=6.48$)، فعالیت افزایش پیدا کرده است.

به عنوان نتیجه مدل سازی، فعالیت یک سری از فلاونوئیدها بر پایه بهترین مدل (مدل مربوط به شکافت ۳) که بالاترین عامل‌های آماری برای سری واسنجی را داشت، pIC_{50} مربوط به تعدادی از فلاونوئیدهای طبیعی پیشگویی و گزارش شدند. نام فلاونوئیدها و نتیجه‌های این پیشگویی در جدول ۵ گزارش شده است. همان‌طور که در جدول مشاهده می‌شود تعدادی از فلاونوئیدهای طبیعی گزارش شده، فعالیتی بیش از فعال‌ترین ترکیب‌های موجود در مدل دارند. ترکیب‌های شماره ۳۵، ۵۴، ۶۰ و ۶۲ تا ۶۶ دارای pIC_{50} بالای ۸ هستند. این ترکیب‌ها به عنوان ترکیب‌های فعال طبیعی برای مهار آنزیم استیل کولین استراز می‌توانند استفاده شوند. بر پایه جستجوی جامعی که در مقالات معتبر انجام شده است، فعالیت تعدادی از این ترکیب‌ها به صورت تجربی در آزمایشگاه اندازه‌گیری شده است که قابل مقایسه با فعالیت‌های به دست آمده از راه مدل است. مقدارهای تجربی pIC_{50} این ترکیب‌ها در جدول ۵ همراه با ذکر مرجع مربوط آمده است.

جدول ۵ نام، SMILES فلاونوئیدهای طبیعی و pIC₅₀ پیشگویی شده آنها بر پایه بهترین مدل (مدل ۳)

شماره	نام فلاونوئید	SMILES	شکافت ^۳	pIC ₅₀ گزارش شده	مرجع
۱	آرتلاستین	<chem>CC(=CCC1=C(Oe2c(CC=C(C)C)c(O)c(CC=C(C)C)c(O)c2C1=O)c3ccc(O)cc3O)C</chem>	۷,۷۷	-	-
۲	آرتلاستین	<chem>CC(=CCc1c(O)c(CC=C(C)C)c2OC3=C(C(Oc4cc(O)ccc34)C=C(C)C)C(=O)c2c1O)C</chem>	۵,۳۷	-	-
۳	آرتلاستوکرومن	<chem>CC(=CCc1c2OC(C)C=Cc2c(O)c3C(=O)C4=C(Oc13)c5ccc(O)cc5OC4C=C(C)C)C</chem>	۷,۲۴	-	-
۴	آرتوکارپین	<chem>CC(=CCc1c(O)c2OC(=CC(=O)c2e1O)c3ccc(O)cc3O)C</chem>	۴,۵۲	-	-
۵	آرتلاستوکارپین	<chem>CC(=CCc1c(O)c(CC=C(C)C)c2OC3=C(CC(Oc4cc(O)ccc34)C(C)(C)O)C(=O)c2c1O)C</chem>	۵,۱۷	-	-
۶	کارپلاتوفوران	<chem>CC(=CCc1c2OC(Cc2c(O)c3C(=O)C4=C(Oc13)c5ccc(O)cc5OC(C4)C(C)(C)O)C(C)(C)O)C</chem>	۶,۱۱	-	-
۷	آرتلاستوفوران	<chem>CC(=CCC1=C(Oc2c3CC(Oc3c(CC=C(C)C)c(O)c2C1=O)C(C)(C)O)c4ccc(O)cc4O)C</chem>	۷,۵۹	-	-
۸	پریمولین	<chem>Oc1cccc2OC(=CC(=O)c12)c3cccc3</chem>	۳,۵۱	-	-
۹	کریسین	<chem>Oc1cc(O)c2C(=O)C=C(Oc2c1)c3cccc3</chem>	۴,۵۴	۴,۵۰	[۳۳]
۱۰	تکتوکریزین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3cccc3</chem>	۴,۶۱	۴,۴۷	[۳۴]
۱۱	پرایتین	<chem>Oc1ccc(O)c2C(=O)C=C(Oc12)c3cccc3</chem>	۵,۱۸	-	-
۱۲	آپیژنین	<chem>Oc1ccc(cc1)C2=CC(=O)c3c(O)cc(O)cc3O2</chem>	۷,۳۵	۵,۱۱	[۳۵]
۱۳	آکاستین	<chem>COc1ccc(cc1)C2=CC(=O)c3c(O)cc(O)cc3O2</chem>	۷,۴۲	-	-
۱۴	گنکوانین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3ccc(O)cc3</chem>	۵,۳۸	-	-
۱۵	اکیوتیدینین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3cccc3O</chem>	۵,۰۶	۴,۹۸	[۳۶]
۱۶	بایکلین	<chem>Oc1cc2OC(=CC(=O)c2c(O)c1O)c3cccc3</chem>	۴,۷۳	۴,۹۶	[۳۷]
۱۷	اوروکسیلین	<chem>COc1c(O)cc2OC(=CC(=O)c2c1O)c3cccc3</chem>	۴,۴۷	-	-
۱۸	نگلتین	<chem>COc1cc2OC(=CC(=O)c2c(O)c1O)c3cccc3</chem>	۴,۸۰	-	-
۱۹	نوروگونین	<chem>Oc1cc(O)c2C(=O)C=C(Oc2c1O)c3cccc3</chem>	۵,۴۹	-	-
۲۰	ووگونین	<chem>COc1c(O)cc(O)c2C(=O)C=C(Oc12)c3cccc3</chem>	۵,۵۲	۵,۲۰	[۳۶]
۲۱	جرالدون	<chem>COc1ccc(cc1O)C2=CC(=O)c3ccc(O)cc3O2</chem>	۶,۸۲	-	-
۲۲	تیتونین	<chem>COc1ccc2C(=O)C=C(Oc2c1)c3ccc(OC)c(O)c3</chem>	۵,۶۰	-	-
۲۳	لوتولین	<chem>Oc1cc(O)c2C(=O)C=C(Oc2c1)c3ccc(O)c(O)c3</chem>	۵,۸۷	۴,۷۶	[۳۷]
۲۴	۶-هیدروکسی لوتولین	<chem>Oc1ccc(cc1O)C2=CC(=O)c3c(O)c(O)c(O)cc3O2</chem>	۷,۴۵	-	-
۲۵	کرایزوریول	<chem>COc1ccc(cc1O)C2=CC(=O)c3c(O)cc(O)cc3O2</chem>	۷,۳۷	-	-
۲۶	دیوسمتین (کریزوریول)	<chem>COc1ccc(cc1O)C2=CC(=O)c3c(O)cc(O)cc3O2</chem>	۷,۳۷	-	-
۲۷	پیلوین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3ccc(OC)c(O)c3</chem>	۶,۰۹	-	-
۲۸	ولوتین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3ccc(O)c(OC)c3</chem>	۶,۰۹	-	-
۲۹	نورارتوکارپین	<chem>Oc1ccc(C2=CC(=O)c3c(O)cc(O)cc3O2)c(O)c1</chem>	۶,۹۲	-	-
۳۰	آرتوکارپین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3ccc(O)cc3O</chem>	۵,۶۵	-	-
۳۱	اسکوتالارین	<chem>Oc1ccc(cc1)C2=CC(=O)c3c(O)c(O)c(O)cc3O2</chem>	۷,۹۵	-	-
۳۲	هیسپیدولین	<chem>COc1c(O)cc2OC(=CC(=O)c2c1O)c3ccc(O)cc3</chem>	۵,۰۷	-	-
۳۳	سوربیفولین	<chem>COc1cc2OC(=CC(=O)c2c(O)c1O)c3ccc(O)cc3</chem>	۵,۳۹	-	-
۳۴	پکتولیناریژنین	<chem>COc1ccc(cc1)C2=CC(=O)c3c(O)c(OC)c(O)cc3O2</chem>	۸,۱۶	-	-
۳۵	سیرسیماریتین	<chem>COc1cc2OC(=CC(=O)c2c(O)c1OC)c3ccc(O)cc3</chem>	۵,۱۰	-	-

ادامه جدول ۵

شماره	نام فلاونوئید	SMILES	شکافت ^۳	پIC ₅₀ گزارش شده	مرجع
۳۶	میکانین	<chem>COc1ccc(cc1)C2=C(O)C(=O)c3c(O)c(OC)c(OC)cc3O2</chem>	۷,۶۶	-	-
۳۷	ایزوسکوتالارین	<chem>Oc1ccc(cc1)C2=CC(=O)c3c(O)cc(O)c(O)c3O2</chem>	۷,۹۰	-	-
۳۸	زاپوتینین	<chem>COc1ccc2OC(=CC(=O)c2c1O)c3c(OC)cccc3OC</chem>	۵,۴۰	-	-
۳۹	زاپوتین	<chem>COc1ccc2OC(=CC(=O)c2c1OC)c3c(OC)cccc3OC</chem>	۵,۱۲	-	-
۴۰	سروسیلین	<chem>COc1cc(OC)cc(c1)C2=CC(=O)c3c(OC)c(OC)ccc3O2</chem>	۷,۹۶	-	-
۴۱	التین	<chem>COc1c(O)c2C(=O)C=C(Oc2c(OC)c1OC)c3cccc3</chem>	۵,۳۵	-	-
۴۲	تریستین	<chem>Oc1cc(O)c2C(=O)C=C(Oc2c1)c3cc(O)c(O)c(O)c3</chem>	۶,۰۱	-	-
۴۳	تریسین	<chem>COc1cc(cc(OC)c1O)C2=CC(=O)c3c(O)cc(O)cc3O2</chem>	۷,۱۷	-	-
۴۴	کوریمبوسین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1)c3cc(OC)c(OC)c(OC)c3</chem>	۶,۵۲	-	-
۴۵	نپتین	<chem>COc1c(O)cc2OC(=CC(=O)c2c1O)c3ccc(O)c(O)c3</chem>	۵,۱۷	-	-
۴۶	پدالتین	<chem>COc1cc2OC(=CC(=O)c2c(O)c1O)c3ccc(O)c(O)c3</chem>	۵,۴۹	-	-
۴۷	نودی فلورین	<chem>COc1cc(ccc1O)C2=CC(=O)c3c(O)c(O)c(O)cc3O2</chem>	۷,۵۱	-	-
۴۸	جاسیوسیدین	<chem>COc1cc(ccc1O)C2=CC(=O)c3c(O)c(OC)c(O)cc3O2</chem>	۷,۶۶	-	-
۴۹	سیرسیلیول	<chem>COc1cc2OC(=CC(=O)c2c(O)c1OC)c3ccc(O)c(O)c3</chem>	۵,۲۰	-	-
۵۰	اوپاتیلین	<chem>COc1ccc(cc1OC)C2=CC(=O)c3c(O)c(OC)c(O)cc3O2</chem>	۷,۹۰	-	-
۵۱	سیرسیلینول	<chem>COc1cc(ccc1O)C2=CC(=O)c3c(O)c(OC)c(OC)cc3O2</chem>	۷,۸۱	-	-
۵۲	اوپاتورین	<chem>COc1ccc(cc1O)C2=CC(=O)c3c(O)c(OC)c(OC)cc3O2</chem>	۷,۸۱	-	-
۵۳	سینستین	<chem>COc1ccc(cc1OC)C2=CC(=O)c3c(OC)c(OC)c(OC)cc3O2</chem>	۸,۱۹	-	-
۵۴	هیپولاتین	<chem>Oc1ccc(cc1O)C2=CC(=O)c3c(O)cc(O)c(O)c3O2</chem>	۷,۴۱	-	-
۵۵	اونوپوردین	<chem>COc1c(O)cc(O)c2C(=O)C=C(Oc12)c3ccc(O)c(O)c3</chem>	۶,۲۲	-	-
۵۶	وایتین	<chem>COc1cc(O)c2C(=O)C=C(Oc2c1OC)c3cccc(O)c3OC</chem>	۵,۷۰	-	-
۵۷	نوادنسن	<chem>COc1ccc(cc1)C2=CC(=O)c3c(O)c(OC)c(O)c(OC)c3O2</chem>	۸,۴۱	-	-
۵۸	زانتومیکرول	<chem>COc1c(O)c2C(=O)C=C(Oc2c(OC)c1OC)c3ccc(O)cc3</chem>	۵,۴۹	-	-
۵۹	تنگرتین	<chem>COc1ccc(cc1)C2=CC(=O)c3c(OC)c(OC)c(OC)c(OC)c3O2</chem>	۸,۷۰	-	-
۶۰	سرپیلین	<chem>COc1ccc(C2=CC(=O)c3c(O)cc(OC)c(OC)c3O2)c(OC)c1OC</chem>	۷,۸۷	-	-
۶۱	هیمنوکسین	<chem>COc1ccc(cc1OC)C2=CC(=O)c3c(O)c(OC)c(O)c(OC)c3O2</chem>	۸,۴۲	-	-
۶۲	گاردنید دی	<chem>COc1ccc(cc1O)C2=CC(=O)c3c(O)c(OC)c(OC)c(OC)c3O2</chem>	۸,۳۳	-	-
۶۳	نوبلتین	<chem>COc1ccc(cc1OC)C2=CC(=O)c3c(OC)c(OC)c(OC)c(OC)c3O2</chem>	۸,۷۱	-	-
۶۴	اسکاپوسین	<chem>COc1cc(cc(O)c1OC)C2=CC(=O)c3c(O)c(OC)c(O)c(OC)c3O2</chem>	۸,۳۴	-	-

مقایسه مدل با مدل‌های گزارش شده

گزارش‌های بسیاری در مورد مدلسازی QSAR برخی از مشتقات فلاونوئیدی دارای نقش مهارکنندگی استیل‌کولین‌استراز وجود دارد. در سال ۲۰۱۴، گوپال و همکارانش یک مدل QSAR سه بعدی بر ۲۴ مشتق فلاونوئید با توصیفگرهای استخراج شده از نرم افزار VLifeQSAR و به کمک روش حداقل مربع‌های جزئی ارائه کردند. حدود ۱۹ ترکیب در سری آموزش و ۵ ترکیب در سری آزمون به صورت تصادفی قرار گرفتند (جدول ۶). در این مدل، نتایج آماری مدل مقایسه R² برای سری آموزش و آزمون به ترتیب، ۰/۸۲۲۷ و ۰/۷۸۹۳ گزارش شد [۳۸]. در سال ۲۰۱۷،

داس و همکارانش یک مدل QSAR بر ۲۱ مشتق فلاونوئیدی را به روش وایزش^۱ خطی گام‌به‌گام برازش دادند. توصیفگرهای مولکولی به کمک نرم‌افزار ChemSketch استخراج شد. در این مطالعه، به علت کوچک بودن تعداد داده‌ها از تقسیم داده‌ها به دو سری آموزش و آزمون چشم‌پوشی و تنها به اعتبارسنجی درونی اکتفا شد. مقدار R^2 برای مدل ۰/۷۳۵ به دست آمد [۳۹]. در سال ۲۰۱۸ موتوکوماران و همکارانش، تحلیل تصویر چندمتغیره-رابطه کمی ساختار-فعالیت (MIA-QSAR) بر روی ۵۹ مشتق فلاونوئید را با روش حداقل مربع‌های جزئی انجام دادند. داده‌ها به روش تصادفی به دو سری آموزش و آزمون به ترتیب با تعداد ۳۹ و ۱۳ ترکیب تقسیم شد و مقدار R^2 برای سری آموزش و آزمون به ترتیب ۰/۹۵۵ و ۰/۹۴۸ به دست آمد. در این روش با وجود مقدار R^2 به نسبت بالا، تفسیر روشنی از متغیرهای افزایشنده و کاهشنده ارائه نشد [۴۰] در سال ۲۰۱۹ کارماکار و همکارانش ۳۰ ترکیب فلاونوئید را سنتز کردند و یک مدل QSAR بر روی آن‌ها بدون انجام اعتبارسنجی خارجی ارائه دادند. آن‌ها از توصیفگرهای نرم‌افزار دارگون و PaDEL در ایجاد مدل استفاده کردند و از روش وایزش تقریب عملکرد ژنتیکی (GFA^2)، برای مدلسازی بهره گرفتند و مقدار R^2 ، ۰/۶۸۳ را برای مدل گزارش کردند. آن‌ها حضور یک گروه هیدروکسیل در موقعیت ۳ حلقه فلاون را به عنوان توصیفگر افزایشنده معرفی کردند [۴۱].

در سال ۲۰۲۳ سوفی و همکارانش یک مدل QSAR بر روی ۴۵ مشتق از فلاونوئیدها برازش دادند. آن‌ها داده‌ها را به روش تصادفی به ترتیب به دو سری آموزش و آزمون ۳۶ و ۹ تایی تقسیم کردند. توصیفگرها به کمک کم آفیس و کم اسکتچ استخراج و به روش وایزش خطی گام‌به‌گام، برازش انجام شد. برپایه نتایج مدلسازی مقادیر R^2 برای سری آموزش و آزمون به ترتیب، ۰/۸۰۱ و ۰/۸۲۴ به دست آمد. در این مطالعه، از جمله عوامل اصلی حاکم بر فعالیت پادکولین استراز، انکسار مولی گزارش شد. آن‌ها حدود ۲۳ مشتق جدید هم برپایه توصیفگرهای مهم، گزارش کردند [۴۲]. در سال ۲۰۲۳، نور و همکارانش یک مدل QSAR، بر روی ۳۱ مشتق از فلاونوئید با نقش مهارکنندگی استیل کولین استراز برازش دادند. آن‌ها ساختار ترکیب‌ها را به روش DFT، بهینه و داده‌ها را به روش تصادفی به دو سری آموزش (۲۵ ترکیب) و آزمون (۶ ترکیب) تقسیم کردند. آن‌ها از ویژگی‌های الکترونیکی، توصیفگرهای توپولوژیکی و لیپینسکی به کمک وایزش خطی چندگانه برای ایجاد مدل بهره بردند و مقدار R^2 برای سری آموزش و آزمون به ترتیب ۰/۷۸۹ و ۰/۷۶۰ گزارش کردند. آن‌ها برپایه نتایج مدلسازی ۹ مشتق جدید از فلاونوئید طراحی و گزارش کردند [۴۳]. با توجه به مرور کاملی که از مدل‌های QSAR در این زمینه بیان شد، یکی از نقص‌های کلی در مدل‌ها تعداد بسیار کم ترکیب‌ها و گاهی عدم اعتبارسنجی خارجی مدل‌ها است. افزون‌بر آن، همان‌طور که اشاره شد، برخی از مدل‌ها مانند مدل MIA-QSAR قابل تفسیر نبودند. برای مقایسه نتیجه‌های مدلسازی حاضر با مدل‌های گزارش‌شده توسط سایر پژوهشگران برای مشتقات فلاونوئیدی به عنوان مهارکننده‌های استیل کولین استراز، با یک بررسی جامع، شش مدل در جدول ۶ گزارش شده است. همان‌گونه که در جدول خلاصه شده، مطالعه حاضر نسبت به مطالعه‌های دیگر، بیشترین تعداد داده‌ها در سری آموزش و آزمون را دارند. توصیفگرهای به کاررفته در این مدل تنها برپایه فرمت SMILES ترکیب‌ها، بدون نیاز به بهینه‌سازی سه بعدی ساختار که روشی وقت‌گیر است، ایجاد شده است. در مدل‌های شماره ۲ و ۴،

1. Regression
2. Genetic function approximation (GFA)

اعتبارسنجی خارجی برای مدلسازی انجام نشده است و فقط سری آموزش وجود دارد. در مورد مدل شماره ۳، مدلسازی برپایه تصویر دو بعدی مشتقات فلاونوئیدی انجام شده است. در این مورد نمی‌توان تفسیر روشنی از مدل برپایه سری‌های عاملی یا سایر توصیفگرهای مولکولی ارائه داد. در مورد مدل شماره ۱، مدل وایزشی مورد استفاده، از نوع روش حداقل مربع‌های جزئی است. بنابراین، این مدل هم، تفسیر روشنی از توصیفگرها ارائه نمی‌دهد. مدل ارائه شده حاضر دارای مقادیر ضریب تعیین (R^2) به نسبت بالایی در مقایسه با سایر روش‌های ارائه شده است. یکی دیگر از نتیجه‌های این مطالعه پیش‌بینی و گزارش PIC_{50} بیش از ۶۰ مشتق فلاونوئیدی طبیعی به عنوان مهارکننده‌های استیل کولین‌استراز است که تنها فعالیت تعداد محدودی از این ترکیب‌ها در آزمایشگاه اندازه‌گیری و در مقاله‌ها گزارش شده است.

جدول ۶ مقایسه مدل‌های QSAR گزارش شده برای مدلسازی مشتقات فلاونوئیدی به عنوان مهارکننده‌های استیل کولین‌استراز

شماره	تعداد کل ترکیب‌ها	تعداد ترکیب‌ها در سری آموزش	تعداد ترکیب‌ها در سری آزمون	نوع توصیفگر و یا نرم‌افزار مورد استفاده	مدل وایزشی	R^2 train	R^2 test	مرجع
۱	۲۴	۱۹	۵	VLifeQSAR	حداقل مربع‌های جزئی	۰٫۸۲۳	۰٫۷۸۹	[۳۸]
۲	۲۱	۲۱	-	ACD ChemSketch	وایزش خطی چندگانه گام به گام	۰٫۷۳۵	-	[۳۹]
۳	۵۲	۳۹	۱۳	2D image Descriptors	حداقل مربع‌های جزئی	۰٫۹۵۵	۰٫۹۴۸	[۴۰]
۴	۳۰	۳۰	-	Dragon and PaDEL	وایزش خطی چندگانه	۰٫۶۸۳	-	[۴۱]
۵	۴۵	۳۶	۹	Chemoffice and ChemsSketch	وایزش خطی چندگانه گام به گام	۰٫۸۰۱	۰٫۸۲۴	[۴۲]
۶	۳۱	۲۵	۶	Electronic Properties, Lipinski and Topological Descriptors	وایزش خطی چندگانه گام به گام	۰٫۷۸۹	۰٫۷۶۰	[۴۳]
۷	۸۸	۷۲	۱۶	CORAL	وایزش خطی	۰٫۸۴۶	۰٫۸۱۶	مطالعه حاضر

نتیجه‌گیری

روش‌های محاسباتی به عنوان ابزاری قدرتمند در طراحی و پیش‌بینی فعالیت داروها علیه بیماری‌های گوناگون، جایگاه ویژه‌ای یافته‌اند و توجه روزافزونی را به خود جلب کرده‌اند. در این پژوهش، مدل‌های QSAR برای ۸۸ مشتق فلاونوئیدی به عنوان مهارکننده‌های استیل کولین‌استراز، با روش مونت کارلو و برپایه شاخص آرمانی همبستگی (IIC) توسعه داده شدند. فرایند ساخت این مدل‌ها با به کارگیری نرم‌افزار CORAL و توصیف‌گر هیبریدی بهینه‌ای انجام گرفت که ترکیبی از SMILES و HFGs بود. مدل‌های پیشنهادی با شاخص‌های اعتبارسنجی داخلی و خارجی و همچنین آزمون تصادفی‌سازی Y ارزیابی شدند. نتایج حاکی از ضریب تعیین رضایت‌بخش برای داده‌های آزمون بود. توصیف‌گرهای مؤثر در افزایش یا کاهش فعالیت مهارکنندگی استیل کولین‌استراز ترکیب‌های مورد مطالعه، با تحلیل وزن همبستگی و ویژگی‌های متفاوت مولکولی در طی چندین مرحله اجرای روش مونت کارلو شناسایی شدند. افزون‌براین، فعالیت مهارکنندگی استیل کولین‌استراز بیش از ۶۰ مشتق فلاونوئیدی طبیعی

به‌عنوان مهارکننده‌های استیل‌کولین‌استراز پیش‌بینی و مقدار pIC50 آن‌ها گزارش شد. برپایه نتایج به‌دست‌آمده از مدل‌سازی، امکان طراحی و شناسایی ترکیب‌های جدید با فعالیت مهارکنندگی بالاتر برای مقابله با بیماری آلزایمر فراهم شده است.

مراجع

- [1] Migliaccio R, Cacciamani F. The temporal lobe in typical and atypical Alzheimer disease. *Handb Clin Neurol*. 2022;187:449-66. doi: **10.1016/B978-0-12-823493-8.00004-3**
- [2] Prince M, Wimo A, Guerchet M, Ali GC, Wu YT, Prina M. World Alzheimer Report 2015-The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International (ADI); 2015.
- [3] Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimers Dement*. 2013;9(1):63-75.e2. doi: **10.1016/j.jalz.2012.11.007**
- [4] Khan H, Amin S, Kamal MA, Patel S. Flavonoids as acetylcholinesterase inhibitors: Current therapeutic standing and future prospects. *Biomedicine & Pharmacotherapy*. 2018;101:860-70. doi: **10.1016/j.biopha.2018.03.007**
- [5] Soleymani N, Ahmadi S, Shiri F, Almasirad A. QSAR and molecular docking studies of isatin and indole derivatives as SARS 3CLpro inhibitors. *BMC Chemistry*. 2023;17(1):32. doi: **10.1186/s13065-023-00947-w**
- [6] Lorenz S, Amsel AK, Puhlmann N, Reich M, Olsson O, Kümmerer K. Toward application and implementation of in silico tools and workflows within benign by design approaches. *ACS Sustainable Chemistry & Engineering*. 2021;9(37):12461-75. doi: **10.1021/acssuschemeng.1c03070**
- [1] Zhao P, Peng Y, Xu X, Wang Z, Wu Z, Li W, et al. In silico prediction of mitochondrial toxicity of chemicals using machine learning methods. *Journal of Applied Toxicology*. 2021;41(10):1518-26. doi:**10.1002/jat.4141**
- [8] Azimi A, Ahmadi S, Kumar A, Qomi M, Almasirad A. SMILES-based QSAR and molecular docking study of oseltamivir derivatives as influenza inhibitors. *Polycyclic aromatic compounds*. 2022;43(4):3257-3277. doi: **10.1080/10406638.2022.2067194**
- [9] Nikkar M, Sayyadikordabadi R, Alizadeh O, Ghasemi G. Lipophilicity and biological activity study of several caspofungin antifungal drugs using QSAR and Monte Carlo methods. *Iranian Journal of Chemistry and Chemical Engineering*. 2022;41(12):4132-4150. doi: **10.30492/IJCCE.2022.535892.4880**
- [10] Ahmadi S, Ghanbari H, Lotfi S, Azimi N. Predictive QSAR modeling for the antioxidant activity of natural compounds derivatives based on Monte Carlo method. *Molecular Diversity*. 2021;25(1):87-97. doi: **10.1007/s11030-019-10026-9**
- [11] Shahin A, Roohallah K, Maryam M. Prediction of anti-cancer activity of 1,8-naphthyridin derivatives by using of genetic algorithm-stepwise multiple linear regression. 2018;28(3):181-194. doi: **10.29252/iau.28.3.181**
- [12] Ahmadi S, Moradi Z, Kumar A, Almasirad A. SMILES-based QSAR and molecular docking study of xanthone derivatives as α -glucosidase inhibitors. *Journal of Receptors and Signal Transduction*. 2022;42(4):361-372. doi: **10.1080/10799893.2021.1957932**
- [13] Hamzehali H, Lotfi S, Ahmadi S, Kumar P. Quantitative structure–activity relationship modeling for predication of inhibition potencies of imatinib derivatives using SMILES attributes. *Scientific Reports*. 2022;12(1):21708. doi: **10.1038/s41598-022-26279-8**

- [14] Ahmadi S, Abdolmaleki A, Jebeli Javan M. In silico study of natural antioxidants. *Vitam Horm.* 2023;121:1-43. **doi: 10.1016/bs.vh.2022.09.001**
- [15] Ghiasi T, Ahmadi S, Ahmadi E, Talei Babil Olyai M, Khodadadi Z. The index of ideality of correlation: QSAR studies of hepatitis C virus NS3/4A protease inhibitors using SMILES descriptors. *SAR and QSAR in Environmental Research.* 2021;32(6):495-520. **doi: 10.1080/1062936X.2021.1925344**
- [16] Ghasedi N, Ahmadi S, Ketabi S, Almasirad A. DFT based QSAR study on quinolone-triazole derivatives as antibacterial agents. *Journal of Receptors and Signal Transduction.* 2022;42(4):418-428. **doi: 10.1080/10799893.2021.1988971**
- [17] Ahmadi S, Mardinia F, Azimi N, Qomi M, Balali E. Prediction of chalcone derivative cytotoxicity activity against MCF-7 human breast cancer cell by Monte Carlo method. *Journal of Molecular Structure.* 2019;1181:305-311. **doi: 10.1016/j.molstruc.2018.12.089**
- [18] Luo W, Chen Y, Wang T, Hong C, Chang LP, Chang CC, et al. Design, synthesis and evaluation of novel 7-aminoalkyl-substituted flavonoid derivatives with improved cholinesterase inhibitory activities. *Bioorganic & medicinal chemistry.* 2016;24(4):672-680. **doi: 10.1016/j.bmc.2015.12.031**
- [19] Luo W, Su Y-B, Hong C, Tian RG, Su LP, Wang YQ, et al. Design, synthesis and evaluation of novel 4-dimethylamine flavonoid derivatives as potential multi-functional anti-Alzheimer agents. *Bioorganic & medicinal chemistry.* 2013;21(23):7275-7282. **doi: 10.1016/j.ejmech.2016.06.022**
- [20] Singh M, Kaur M, Singh N, Silakari O. Exploration of multi-target potential of chromen-4-one based compounds in Alzheimer's disease: Design, synthesis and biological evaluations. *Bioorganic & Medicinal Chemistry.* 2017;25(24):6273-6285. **doi: 10.1016/j.bmc.2017.09.012**
- [21] Shen Y, Zhang J, Sheng R, Dong X, He Q, Yang B, et al. Synthesis and biological evaluation of novel flavonoid derivatives as dual binding acetylcholinesterase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry.* 2009;24(2):372-380. **doi: 10.1080/14756360802187885**
- [22] Li RS, Wang XB, Hu XJ, Kong LY. Design, synthesis and evaluation of flavonoid derivatives as potential multifunctional acetylcholinesterase inhibitors against Alzheimer's disease. *Bioorganic & Medicinal Chemistry Letters.* 2013;23(9):2636-2641. **doi: 10.1016/j.bmcl.2013.02.095**
- [23] Ghaedi A. Predicting the cytotoxicity of ionic liquids using QSAR model based on SMILES optimal descriptors. *Journal of Molecular Liquids.* 2015;208:269-279. **doi: 10.1016/j.molliq.2015.04.049**
- [24] Benfenati E, Toropov AA, Toropova AP, Manganaro A, Gonella Diaza R. CORAL software: QSAR for anticancer agents. *Chemical biology & drug design.* 2011;77(6):471-476. DOI: 10.1111/j.1747-0285.2011.01117.x
- [25] Das NR, Sharma T, Mallick A, Toropova AP, Toropov AA, Achary PGR. Computational approach in designing and development of novel inhibitors of AKR1C1. In: Swarnkar T, Patnaik S, Mitra P, Misra S, Mishra M, editors. *Ambient Intelligence in Health Care.* Singapore: Springer; 2023. p. 325-337. **doi: 10.1007/978-981-19-6068-0_32**
- [26] Toropova AP, Toropov AA. CORAL: QSAR models for carcinogenicity of organic compounds for male and female rats. *Computational biology and chemistry.* 2018;72:26-32. **doi: 10.1016/j.compbiolchem.2017.12.012**

- [27] Singh R, Kumar P, Devi M, Lal S, Kumar A, Sindhu J, et al. Monte Carlo based QSGFEAR: prediction of Gibb's free energy of activation at different temperatures using SMILES based descriptors. *New Journal of Chemistry*. 2022;46(39):19062-19072. **doi: 10.1039/D2NJ03515D**
- [28] Ahmadi S, Lotfi S, Kumar P. Quantitative structure–toxicity relationship models for predication of toxicity of ionic liquids toward leukemia rat cell line IPC-81 based on index of ideality of correlation. *Toxicology Mechanisms and Methods*. 2022;32(4):302-312. **doi: 10.1080/15376516.2021.2000686**
- [29] Ahmadi S. Mathematical modeling of cytotoxicity of metal oxide nanoparticles using the index of ideality correlation criteria. *Chemosphere*. 2020;242:125192. **doi: 10.1016/j.chemosphere.2019.125192**
- [30] Golbraikh A, Tropsha A. Beware of q²!. *Journal of Molecular Graphics and Modelling*. 2002;20(4):269-276. **doi: 10.1016/S1093-3263(01)00123-1**
- [31] Roy PP, Roy K. QSAR studies of CYP2D6 inhibitor aryloxypropanolamines using 2D and 3D descriptors. *Chemical Biology & Drug Design*. 2009;73(4):442-455. **doi: 10.1111/j.1747-0285.2009.00791.x**
- [32] Rücker C, Rücker G, Meringer M. γ -Randomization and its variants in QSPR/QSAR. *Journal of chemical information and modeling*. 2007;47(6):2345-2357. **doi: 10.1021/ci700157b**
- [33] Taslimi P, Caglayan C, Gulcin İ. The impact of some natural phenolic compounds on carbonic anhydrase, acetylcholinesterase, butyrylcholinesterase, and α -glycosidase enzymes: An antidiabetic, anticholinergic, and antiepileptic study. *Journal of Biochemical and Molecular Toxicology*. 2017;31(12):21995. **doi: 10.1002/jbt.21995**
- [34] Cuong NM, Khanh PN, Nhung LTH, Ha NX, Huong TT, Bauerova K, et al. Acetylcholinesterase inhibitory activities of some flavonoids from the root bark of *Pinus krempfii* Lecomte: in vitro and in silico study. *Journal of Biomolecular Structure and Dynamics*. 2024;42(9):4888-4901. **doi: 10.1080/07391102.2023.2223664**
- [35] Balkis A, Tran K, Lee YZ, Balkis KN, Ng K. Screening flavonoids for inhibition of acetylcholinesterase identified baicalein as the most potent inhibitor. *Journal of Agricultural Science*. 2015;7(9):26-35. **doi:10.5539/jas.v7n9p26**
- [36] Goel N, Garg A, Nagendra C, Reddy AM, Biswas R, Natarajan R, et al. In-vitro and in-silico cholinesterase inhibitory activity of bioactive molecules isolated from the leaves of *Andrographis nallamalayana* JL Ellis and roots of *Andrographis beddomei* CB Clarke. *Journal of Molecular Structure*. 2024;1301:137406. **doi: 10.1016/j.molstruc.2023.137406**
- [37] Xie Y, Yang W, Chen X, Xiao J. Inhibition of flavonoids on acetylcholine esterase: Binding and structure–activity relationship. *Food & Function*. 2014;5(10):2582-2589. **doi: 10.1039/c4fo00287c**
- [38] Goyal M, Grover S, Dhanjal JK, Goyal S, Tyagi C, Grover A. Molecular modelling studies on flavonoid derivatives as dual site inhibitors of human acetyl cholinesterase using 3D-QSAR, pharmacophore and high throughput screening approaches. *Medicinal Chemistry Research*. 2014;23:2122-2132. **doi: 10.1007/s00044-013-0810-2**
- [39] Das S, Laskar MA, Sarker SD, Choudhury MD, Choudhury PR, Mitra A, et al. Prediction of anti-Alzheimer's activity of flavonoids targeting acetylcholinesterase in silico. *Phytochemical Analysis*. 2017;28(4):324-331. **doi: 10.1002/pca.2679**
- [40] Muthukumaran P, Rajiniraja M. MIA-QSAR based model for bioactivity prediction of flavonoid derivatives as acetylcholinesterase inhibitors. *Journal of Theoretical Biology*. 2018;459:103-110. **doi: 10.1016/j.jtbi.2018.09.030**

- [41] Karmakar A, Ambure P, Mallick T, Das S, Roy K, Begum NA. Exploration of synthetic antioxidant flavonoid analogs as acetylcholinesterase inhibitors: An approach towards finding their quantitative structure–activity relationship. *Medicinal Chemistry Research*. 2019;28:723-741. doi: **10.1007/s00044-019-02330-8**
- [42] Soufi H, Moussaoui M, Baammi S, Baassi M, Salah M, Daoud R, et al. Multi-combined QSAR, molecular docking, molecular dynamics simulation, and ADMET of flavonoid derivatives as potent cholinesterase inhibitors. *Journal of Biomolecular Structure and Dynamics*. 2024;42(12):6027-6041. doi: **10.1080/07391102.2023.2238314**
- [43] Nour H, Hashmi MA, Belaidi S, Errougui A, El Kouali M, Talbi M, et al. Design of acetylcholinesterase inhibitors as promising anti-Alzheimer's agents based on QSAR, molecular docking, and molecular dynamics studies of liquiritigenin derivatives. *ChemistrySelect*. 2023;8(32):e202301466. doi: **10.1002/slct.202301466**

Quantitative structure-activity relationship of some flavonoid derivatives as acetylcholinesterase inhibitors based on Monte Carlo algorithm

Z. Nazari¹, S. Ahmadi^{2,*}, A. Almasirad³

1. M.D. Student of Pharmacy, Department of Medicinal Chemistry, Faculty of Pharmacy, Tehran Medical Science, Islamic Azad University, Tehran, Iran.
2. Associate Prof. of Department of Pharmaceutical Chemistry, Faculty of Pharmaceutical Chemistry, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.
3. Professor of Department of Medicinal Chemistry, Faculty of Pharmacy, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.

Abstract: A method based on an independent approach of molecule optimization based on structural and topological molecular features has recently been introduced as an attractive approach in the QSAR modeling. A key feature of this approach refers to the use of descriptors based on the Simplified Molecular Input Line Input System (SMILES) to develop the QSAR model. In this research, by studying the QSAR of some flavonoid derivatives as acetylcholinesterase inhibitors based on the Monte Carlo algorithm, a set of 88 acetylcholinesterase inhibitors was selected from peer reviewed papers. The structures of the compounds in SMILES format were defined as software input files and randomly divided into four series including training, invisible training, calibration and validation sets and modeling was done. The optimal descriptors used in the model were selected from the combination of SMILES and hydrogen suppressor graph based on a new objective function. Finally, the predictive power of models was evaluated by validation sets. The results of four random series showed that predictive and statistically reliable models were obtained. The most important descriptors that increase and decrease the acetylcholinesterase inhibitory activity of flavonoid derivatives were identified. Finally, acetylcholinesterase inhibitory (pIC_{50}) of a large number of natural flavonoid derivatives were calculated and reported. Using the obtained model, it is possible to predict pIC_{50} of flavonoid derivatives as acetylcholinesterase enzyme inhibitors using Monte Carlo and SMILES optimization methods.

Keywords: QSAR, Flavonoids, Acetylcholinesterase inhibitor, SMILES, Index of ideality of correlation.

* Corresponding author Email: s.ahmadi@iautmu.ac.ir & ahmadi.chemometrics@gmail.com