

ارتباط کمی ساختار-فعالیت بر روی یک سری از مشتقات ایمیدازول [۱] و [۲-a] پیریدین کربوکسامید به عنوان عوامل ضد سل

محسن نکوئی نیا^{۱*}، سعید یوسفی نژاد^۲

^۱ گروه شیمی، دانشگاه پیام نور، تهران، ایران

^۲ مرکز تحقیقات علوم بهداشتی، پژوهشکده سلامت، دانشگاه علوم پزشکی شیراز، شیراز، ایران

چکیده: مقاومت بیماری سل به دارو همچنان یکی از مهمترین چالش های پیش رو در درمان این بیماری عفونی است و بنابراین کشف و توسعه داروهای جدید موثر ضد سل همواره مورد توجه محققان است. در این مطالعه، تحلیل ارتباط کمی ساختار-فعالیت (QSAR) بر روی یک سری از مشتقات ایمیدازول [۱] و [۲-a] پیریدین کربوکسامید به عنوان عوامل ضد سل اعمال شد. فعالیت بیولوژیکی ۱۸ ترکیب با روش های رگرسیون خطی چندگانه و شبکه عصبی مصنوعی برآورد شد. چهار توصیف کننده مولکولی (nCl، MATS8m، BELe4 و GATS8e) با استفاده از رگرسیون خطی چندگانه گام به گام انتخاب شدند. بهترین نتایج شبکه عصبی مصنوعی با الگوی ۵-۵-۱ آموزش داده شده با الگوریتم پس انتشار رو به جلو به دست آمد. یک مجموعه آزمون حاوی ۵ ترکیب برای ارزیابی توانایی پیش بینی مدل استفاده شد. نتایج نشان داد که رویکرد شبکه عصبی مصنوعی در مقایسه با رگرسیون خطی چندگانه قدرت پیش بینی بهتری را ارائه می دهد. بر اساس نتایج این مطالعه، الکترون گاتیوی، جرم اتمی و هندسه مولکولی عوامل مهم کنترل کننده فعالیت ضد سل هستند.

m_nekoenia@pnu.ac.ir

نامناسب انجام شود، این روش بی اثر و مقاومت در برابر داروهای ضد سل ایجاد می کند. اگر رژیم های دارویی ضد سل در دوزهای نادرست یا داروهای بی کیفیت توصیه شود، باعث سل مقاوم به چند دارو خواهد شد. علاوه بر این، اگر برای درمان بیماران مبتلا به مشکلات ایمنی یا بیماران مبتلا به HIV استفاده شود، سل مقاوم به داروهای گسترده ایجاد خواهد کرد. داروهای استاندارد ضد سل مانند ایزونیازید، ریفامپیسین، اتامبوتول، دی سیکلوسرین اس و پیرازینامید، ریفابوتین و ریفاپنتین نیز برای جلوگیری از مقاومت و درمان بیماری سل تجویز می شود [۲].

مقاومت این بیماری به دارو همچنان یکی از مهمترین چالش های پیش رو در درمان سل است و بنابراین کشف و توسعه

۱- مقدمه

سل یک بیماری عفونی باکتریایی است که توسط مایکوباکتریوم توبرکلوزیس ایجاد و در سراسر جهان گسترش یافته است. سازمان بهداشت جهانی تخمین زده است که ۸/۸ میلیون مورد جدید سل در سال ۲۰۲۰ وجود داشته است. حدود یک سوم جمعیت جهان به این باکتری مبتلا هستند که ۱۰ درصد از آنها در مقطعی از زندگی خود به این بیماری مبتلا می شوند [۱]. یکی از رایج ترین راهبردها در برابر بیماری عفونی سل، درمان کوتاه با مشاهده مستقیم است. با این حال، اگر درمان به روش های

مجموعه داده ها به طور تصادفی به مجموعه آموزش با ۱۳ ترکیب و مجموعه آزمون شامل ۵ ترکیب تقسیم شدند. از مجموعه آزمون در هیچ یک از مراحل مدل سازی (خطی و غیرخطی) استفاده نشد و صرفاً جهت بررسی قدرت پیش بینی مدل ها مورد استفاده قرار گرفت.

یکی از مراحل مهم در روش QSAR انتخاب مناسب ترین توصیف کننده ها برای پیش بینی بهترین مدل است. از روش رگرسیون خطی چندگانه گام به گام جهت انتخاب توصیف کننده های بکار رفته در مدل استفاده شد. در رگرسیون خطی چندگانه گام به گام، با انتخاب اولین توصیف کننده که بیشترین همبستگی را با متغیر وابسته دارد، شروع می شود. بعد از وارد شدن اولین متغیر، توصیف کننده بعدی که وارد مدل می شود، متغیری است که بیشترین همبستگی را با باقیمانده مدل قبلی دارد. در هر مرحله از مسیر، توصیف کننده ها بر اساس آزمون F جزئی وارد و یا از مدل حذف می شوند [۸].

۳- نتایج و بحث

به منظور ساخت و آزمون مدل، داده ها به دو دسته مجموعه آموزش (۷۵ درصد داده ها) و مجموعه آزمون (۲۵ درصد باقیمانده داده ها) بصورت تصادفی تقسیم شدند. مولکول های شماره (۳، ۴، ۸، ۱۵ و ۱۶ جدول ۲) به عنوان مجموعه آزمون و ۱۳ مولکول دیگر به عنوان مجموعه آموزش تقسیم بندی شدند. از روش رگرسیون چند گانه گام به گام جهت انتخاب بهترین توصیف کننده ها استفاده شد. بر این اساس ۴ توصیف کننده انتخاب و مدل ایجاد شده به صورت زیر بدست آمد:

$$\text{MIC} (\mu\text{M}) = 279.9 + 164.95 \text{ nCl} + 304.6 \text{ MATS8m} - 186.3 \text{ BELe4} + 29.4 \text{ GATS8e}$$

$$N_{\text{train}}=13, N_{\text{test}}=5, R^2_{\text{train}}=0.893, q^2_{\text{LOO}}=0.945, F=115.64, F_{\text{crit}}(95\%)=3.84$$

$$R^2_{\text{test}} < 0.01, \text{RMSE} = 95.82$$

که در آن معادله، MIC حداقل غلظت بازدارنده و nCl، MATS8m، BELe4 و GATS8e چهار توصیف کننده انتخابی است. نتایج پارامترهای آماری از جمله ضریب تعیین مجموعه

داروهای جدید موثر ضد سل بسیار مورد نیاز است. توسعه علوم کامپیوتر کاربرد وسیعی در زمینه کشف دارو پیدا کرده است و طراحی دارو به کمک رایانه به بخشی جدایی ناپذیر از فرآیند کشف دارو در دانشگاه ها و شرکت های داروسازی تبدیل شده است. تبیین روابط کمی ساختار-فعالیت (QSAR) یکی از رویکردهای اصلی طراحی دارو به کمک رایانه است. یک مدل QSAR تلاش می کند تا روابط ثابتی بین تغییرات در مقادیر خواص مولکولی و فعالیت بیولوژیکی برای مجموعه ای از ترکیبات پیدا کند و سپس می تواند برای ارزیابی خواص ترکیبات شیمیایی جدید استفاده شود [۱]. در سال های اخیر مدل سازی QSAR به طور گسترده ای برای شناسایی داروهای جدید ضد سل استفاده شده است [۳-۶]. هدف از انجام این مطالعه، توسعه یک مدل QSAR برای طراحی عوامل ضد سل جدید بر اساس برخی از مشتقات ایمیدازو [۱، ۲-a] پیریدین کربوکسامید با استفاده از شبکه عصبی مصنوعی است.

۲- بخش تجربی

نوزده ترکیب از سری جدید سنتز شده پیریدین-۳-کربوکسامید به عنوان عوامل ضد سل انتخاب شدند [۷]. ساختار مشتقات و همچنین فعالیت ضد سل آنها (ضد سویه H37Rv) در جدول ۱ ارائه شده است.

تمام ساختارها در نرم افزار کم درآرسم شدند. ترکیبات مورد بررسی به نرم افزارهای پرمک^۳ منتقل و ساختار هندسی مولکول ها ابتدا توسط روش میدان نیرو MM+ بهینه شد. سپس جهت بهینه سازی دقیق تر ساختار هندسی از الگوریتم نیمه تجربی AM1 استفاده شد. مولکول های بهینه شده به نرم افزار دراگون^۴ منتقل و توصیف کننده ها محاسبه و مورد استفاده قرار گرفت. توصیف کننده ها ثابت یا تقریباً ثابت با انحراف معیار نزدیک به صفر و همچنین توصیف کننده هایی با ضریب همبستگی بیشتر از ۰/۹۰ حذف شدند. برای هر مولکول ۶۷۱ توصیف کننده های باقی ماند. توصیف کننده های باقیمانده مشتقات پیریدین-۳-کربوکسامید در ماتریس با ابعاد (۱۸×۶۷۱) ذخیره شد.

Artificial neural network (ANN) ^۱
ChemDraw Ultra Version 16.0.1 ^۲
Hyperchem Release 8.0.8 ^۳
Dragon Version 5.5 ^۴

جدول ۱. ساختار شیمیایی مشتقات ایمیدازو [۱-۲a] پیریدین-۳-کربوکسامید و مقادیر MIC تجربی

| MIC (μM) | ساختار | شماره مولکول | MIC (μM) | ساختار | شماره مولکول |
|----------|--------|--------------|----------|--------|--------------|
| ۰/۸ | | ۱۰ | ۰/۲۲ | | ۱ |
| ۴۷/۴۱ | | ۱۱ | ۱۱۷/۹۱ | | ۲ |
| ۰/۴۲ | | ۱۲ | ۵۶/۰۶ | | ۳ |
| ۰/۴ | | ۱۳ | ۱۰۶/۸۸ | | ۴ |
| ۰/۱ | | ۱۴ | ۲۰۰/۱۳ | | ۵ |
| ۰/۱۹ | | ۱۵ | ۲۰۰ | | ۶ |
| ۱۶۱/۶۶ | | ۱۶ | ۱۶۷/۵۸ | | ۷ |
| ۲۰۹/۲۹ | | ۱۷ | ۰/۴۴ | | ۸ |
| ۲۰۰/۱۳ | | ۱۸ | ۰/۴۲ | | ۹ |

به دست آمده از آزمون اعتبارسنجی متقاطع (q^2 LOO) در مدل SW-MLR برابر با ۰/۹۴۵ شد که نشان دهنده پایایی مدل پیشنهادی است.

آموزش (R^2_{train}) نشان می دهد که ارتباط خطی نسبتا مناسبی بین توصیف کننده های انتخابی و حداقل غلظت بازدارنده مولکول ها در سری آموزش وجود دارد. همچنین، پارامتر آماری

جدول ۲. مقادیر توصیف کننده های مولکولی، MIC تجربی و پیش بینی شده با مدل های ANN و SW-MLR

| MIC (μM) | توصیف کننده | | | | شماره مولکول | | | |
|-----------------------|-------------|--------|-------|-----|--------------|--------|-------|--------|
| | ANN | SW-MLR | تجربی | nCl | | MATS8m | BELe4 | GATS8e |
| ۰/۲۱۸ | ۰/۲۰۵ | ۰/۲۲ | ۰ | ۰ | -۰/۱۳۹ | ۱/۳۳۵ | ۰/۲۸۳ | ۱ |
| ۱۱۷/۸۷۳ | ۱۰۶/۸۵۸ | ۱۱۷/۹۱ | ۰ | ۰ | ۰/۱۴۵ | ۱/۱۸۷ | ۰/۱۳۴ | ۲ |
| ۵۵/۷۸۶ | ۳۵/۰۰۰ | ۵۶/۰۶ | ۰ | ۰ | ۰/۰۹۷ | ۱/۵۰۶ | ۰/۲۰۹ | ۳† |
| ۱۰۶/۸۷۴ | ۵۵/۹۱۳ | ۱۰۶/۸۸ | ۰ | ۰ | ۰/۰۵۲ | ۱/۳۳۳ | ۰/۲۹ | ۴† |
| ۲۰۰/۳۶۳ | ۱۸۷/۹۵۲ | ۲۰۰/۱۳ | ۱ | ۱ | ۰/۰۹ | ۱/۳۳۱ | ۰/۶۲۹ | ۵ |
| ۲۰۰/۰۵۸ | ۲۰۶/۱۴۱ | ۲۰۰ | ۱ | ۱ | -۰/۰۱۳ | ۱/۳۳۱ | ۰/۴۵ | ۶ |
| ۱۶۷/۶۹۵ | ۱۶۹/۲۳۵ | ۱۶۷/۵۸ | ۱ | ۱ | -۰/۰۸۱ | ۱/۴۴ | ۰/۵۹ | ۷ |
| ۰/۴۳۱ | ۶۶/۹۲۲ | ۰/۴۴ | ۰ | ۰ | ۰/۰۴۳ | ۱/۲۶۳ | ۰/۳۱۴ | ۸† |
| ۰/۳۹۳ | ۶/۶۹۱ | ۰/۴۲ | ۰ | ۰ | ۰/۰۰۴ | ۱/۵۳۴ | ۰/۳۸۷ | ۹ |
| ۰/۷۶۳ | ۱۹/۷۱۹ | ۰/۸ | ۰ | ۰ | -۰/۰۷ | ۱/۳۹۹ | ۰/۷۴۱ | ۱۰ |
| ۴۷/۶۰۷ | ۳۷/۲۰۵ | ۴۷/۴۱ | ۰ | ۰ | -۰/۰۲۶ | ۱/۳۳۱ | ۰/۴۹۹ | ۱۱ |
| ۰/۸۴۲ | ۲۱/۹۰۴ | ۰/۴۲ | ۰ | ۰ | -۰/۰۹۱ | ۱/۳۶۱ | ۰/۷۹۲ | ۱۲ |
| ۰/۴۲۹ | -۲۲/۵۱۵ | ۰/۴ | ۰ | ۰ | -۰/۱۵۴ | ۱/۵۴۸ | ۱/۱۱۹ | ۱۳ |
| ۰/۲۲۰ | -۲/۳۸۷ | ۰/۱ | ۰ | ۰ | -۰/۱۶۷ | ۱/۴۳۸ | ۱/۲۱ | ۱۴ |
| ۰/۱۹۸ | ۱۹۶/۲۸۵ | ۰/۱۹ | ۱ | ۱ | -۰/۰۴۷ | ۱/۴۳۸ | ۱/۱۴۵ | ۱۵† |
| ۱۶۱/۶۵۵ | ۱۶۰/۷۹۶ | ۱۶۱/۶۶ | ۱ | ۱ | -۰/۱۲۷ | ۱/۵۱ | ۱/۲۲۳ | ۱۶† |
| ۲۰۹/۳۹۶ | ۲۰۹/۲۴۳ | ۲۰۹/۲۹ | ۱ | ۱ | -۰/۰۲۴ | ۱/۳۰۳ | ۰/۴۹۲ | ۱۷ |
| ۲۰۰/۱۲۹ | ۲۰۴/۵۶۰ | ۲۰۰/۱۳ | ۱ | ۱ | -۰/۰۵۲ | ۱/۳۹ | ۱/۱۷۴ | ۱۸ |

†مجموعه آزمون

جدول ۳. ماتریس همبستگی و عامل تورم واریانس برای چهار توصیف کننده انتخابی

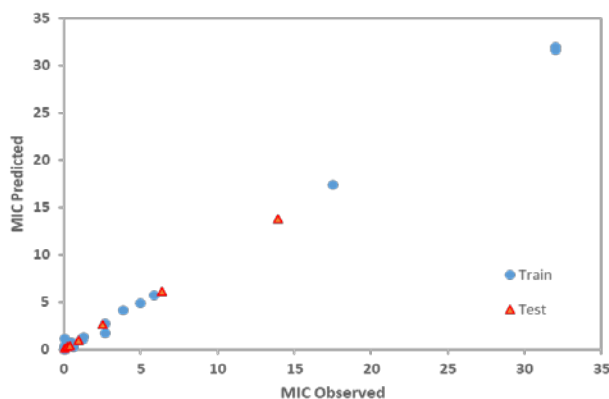
| | nCl | MATS8m | BELe4 | GATS8e | VIF |
|--------|--------|--------|-------|--------|------|
| nCl | ۱/۰۰ | | | | ۱/۰۶ |
| MATS8m | ۰/۰۵۴ | ۱/۰۰ | | | ۲/۰۵ |
| BELe4 | -۰/۱۷۱ | -۰/۵۸۴ | ۱/۰۰ | | ۱/۷۱ |
| GATS8e | ۰/۰۴۱ | -۰/۶۶۶ | ۰/۵۵۳ | ۱/۰۰ | ۱/۹۹ |

است، به این معنی که مدل قابل قبول بود و مشکل چند خطی وجود ندارد.

مقادیر توصیف کننده های انتخاب شده، مقادیر MIC تجربی و پیش بینی شده توسط مدل SW-MLR در جدول ۲ آورده شده است. ضریب تعیین مجموعه آزمون (R^2_{test}) از ۰/۰۱ کمتر و ریشه میانگین مربعات خطا (RMSE) مجموعه آزمون به برابر ۹۵/۸۲ بدست آمد. اگرچه این مدل نتایج خوبی برای ارزیابی مقاطع مجموعه آموزش را نشان داد اما در پیشگویی مجموعه آموزش و به ویژه مجموعه آزمون ناموفق بود و از این رو مدل

همخطی چندگانه^۱ بین توصیف کننده های انتخاب شده از طریق محاسبه ضرایب همبستگی جفتی و فاکتور تورم متغیر^۲ ارزیابی شد (جدول ۳). ماتریس همبستگی نشان می دهد که MATS8m با BELe4 ($p < ۰/۰۵$ ، $r = -۰/۵۸۴$) و GATS8e با BELe4 ($p < ۰/۰۵$ ، $r = -۰/۶۶۶$) و GATS8e با MATS8m ($p < ۰/۰۵$ ، $r = ۰/۵۵۳$) همبستگی ضعیفی داشته و بین سایر متغیرها، هیچ همبستگی مشاهده نمی شود. ستون آخر جدول ۳ نشان می دهد که مقادیر VIF چهار متغیر مستقل کمتر از ۵

^۱ multi-collinearity
^۲ variation inflation factors



شکل ۱. مقادیر MIC واقعی و پیش بینی شده از مدل پیشنهادی شبکه عصبی

بالاتری در مجموعه آزمون برخوردار است. تفسیر توصیف کننده های بکار رفته در مدل، می تواند درک بهتری از مکانیسم مهار کنندگی ترکیبات مورد بررسی را نشان دهد. اولین توصیف کننده وارد شده به مدل تعداد اتم های کلر (nCl) است. علامت مثبت ضریب رگرسیون بیانگر اثر مثبت این متغیر بر حداقل غلظت بازدارندگی است و با افزایش تعداد اتم کلر، MIC افزایش می یابد. در واقع حضور اتم کلردر این دسته از مولکول ها سبب کاهش در فعالیت آنها در برابر مایکوباکتریوم توبرکلوزیس می گردد. دومین توصیف کننده وارد شده به مدل MATS8m است که یک توصیف کننده همبستگی دویعدی وزن شده توسط جرم اتمی در طول معینی (lag-8) از مولکول است. MATS8m برهمکنش جفت اتم ها در فاصله توپولوژیکی برابر با ۸ را نشان می دهد [۹]. این توصیف کننده ویژگی هایی همانند شکل هندسی مولکول و سایز را نشان می دهد. انتخاب این توصیف کننده با ضریب مثبت نشان دهنده تاثیر مستقیم سایز مولکول بر حداقل غلظت بازدارندگی مشتقات پیریدین-کربوکسامید است. سومین توصیف کننده وارد شده به مدل BELe4 است که کمترین مقدار ویژه شماره ۴ از ماتریس بوردن است که توسط الکتروننگاتیوهای ساندرسون وزن دار شده است. این توصیف کننده به عنوان یک توصیف کننده BCUT^۶ شناخته می شوند که شامل مشارکت همه اتم ها است و بنابراین توپولوژی کل مولکول را منعکس می کنند [۱۰]. توصیف کننده های BCTU برای رمزگذاری خواص اتمی مربوط به برهمکنش های بین مولکولی طراحی شده اند. حضور این توصیف کننده در مدل با

غیر خطی شبکه عصبی مصنوعی برای ساخت یک مدل مناسب برای مجموعه آموزش و آزمون استفاده شد.

از یک شبکه عصبی پرسپترون چند لایه با الگوریتم آموزشی پس انتشار برای پیش بینی مقادیر سری آزمون پس از مرحله آموزش استفاده شد. الگوریتم پس انتشار یک روش یادگیری نظارت شده است که می تواند به دو مرحله تقسیم شود: انتشار و به روز رسانی وزنها. دو مرحله تا زمانی که عملکرد شبکه به اندازه کافی خوب باشد تکرار می شوند. در الگوریتم های پس انتشار، مقادیر خروجی با پاسخ صحیح مقایسه می شوند تا مقدار تابع خطای از پیش تعریف شده محاسبه شود. با استفاده از این اطلاعات، الگوریتم وزن هر اتصال را تنظیم می کند تا مقدار تابع خطا به حداقل کاهش یابد.

از نرم افزار متلب^۲ برای پیاده سازی شبکه های عصبی دو لایه و با الگوی ۵-۵-۱ استفاده شد. مدل به ترتیب از یک تابع سیگنویید^۳ و یک تابع خطی بعنوان توابع فعال سازی لایه های پنهان و خروجی استفاده کرد. با ۱۳ مولکول سری آموزش و بر اساس روش بهینه یابی الگوریتم لونبرگ-مارکوارت^۴ مقادیر بایاس و وزن ها بهینه شد. توصیف کننده های مولکولی انتخاب شده در مدل های خطی SW-MLR به عنوان ورودی شبکه استفاده شد. قبل از یادگیری شبکه، مقادیر ماتریس ورودی و بردار خروجی بین -۱ و ۱ نرمال سازی شدند. تعداد بهینه گره ها^۵ در لایه پنهان در این شبکه ۵ عدد انتخاب شد. مقادیر توصیف کننده های انتخاب شده، مقادیر MIC تجربی و پیش بینی شده توسط مدل ANN در جدول ۲ آورده شده است. همچنین نمودار مقادیر MIC پیشگویی شده و مقادیر تجربی برای هر دو مجموعه آموزش و آزمون در شکل ۱ نشان داده شده است. نتایج نشان می دهد که تطابق بسیار خوبی بین مقادیر تجربی و پیش بینی وجود دارد. ریشه میانگین مربعات خطا (RMSE) برای مجموعه آزمون برابر ۰/۱۲۲ بدست آمد که نشان دهنده قدرت پیشگویی بسیار خوب شبکه عصبی است. مقایسه نتایج بدست آمده با نتایج رگرسیون خطی چندگانه گام به گام نشان می دهد که شبکه عصبی از قدرت پیشگویی به مراتب

^۱ predefined error function

^۲ Matlab R2018

^۳ sigmoid

^۴ Levenberg-Marquardt optimization

^۵ nodes

3.R.C. Khunt. V.M. Khedkar. R.S. Chawda. N.A. Chauhan. A.R. Parikh, E.C. Coutinho, *Bioorganic Med. Chem. Lett.* 22, 666 (2012).

4. A. Nayyar. A. Malde. R. Jain, E. Coutinho, *Bioorg. Med. Chem.*, 14, 847 (2006).

5. L. Friggeri. F. Ballante. R. Ragno. I. Musmuca. D. De Vita. F. Manetti. M. Biava. L. Scipione. R. Di Santo. R. Costi. M. Feroci, S. Tortorella, *J. Chem. Inf. Model.* 53, 1463 (2013).

6.M. N. Gomes. R.C. Braga. E.M. Grzelak. B. J. Neves. E. Muratov. R. Ma. L. L. Klein. S. Cho. G. R. Oliveira. S. G. Franzblau, C.H. Andrade, *Eur. J. Med. Chem.*, 137, 126 (2017).

7. O.K. Onajole. S. Lun. Y.J. Yun. D.Y. Langue. M. Jaskula-Dybka. A. Flores. E. Frazier. A.C. Scurry. A. Zavala. K. R. Arreola. B. Pierzchalski. A. J.-L. Ayitou, W. R. Bishai, *Chem. Biol. Drug Des.* 96, 1362 (2020).

8. R.N. Forthofer. E.S. Lee, M. Hernandez, *Biostatistics: A Guide to Design, Analysis and Discovery*, Elsevier Science 2006.

9. X. Wang. X. Meng. F. Li. J. Ding. C. Ji, H. Wu, *Chemosphere*, 226, 159 (2019).

10. R.V.C. Guido. G.H.G. Trossini. M.S. Castilho. G. Oliva. E. I. Ferreira, A.D. Andricopulo, *J. Enzyme. Inhib. Med. Chem.* 23, 964 (2008).

11. M.H. Fatemi, Z.G. Chahi, *SAR and QSAR in Environmental Research* 23, 155 (2012).

ضریب منفی نشان دهنده تاثیر معکوس این توصیف کننده بر حداقل غلظت بازدارندگی مشتقات پیریدین-کربوکسامید است. و در انتها، چهارمین توصیف کننده GATS8e می باشد که از دسته توصیف کننده همبستگی دوبرعی است. توصیفگر توپولوژیکی GATS8e اطلاعات ساختاری مولکولی دوبرعی و همچنین الکترونگاتیوی ساندرسون را رمزگذاری می کند که مستقیماً به واکنش پذیری شیمیایی یک مولکول مرتبط است [۱۱]. ضریب مثبت این توصیف کننده در مدل بیانگر تاثیر مستقیم این توصیف کننده بر حداقل غلظت بازدارندگی این مشتقات است.

۴- نتیجه گیری

تحلیل QSAR بر روی مشتقات امیدازول [۱ و ۲-a] پیریدین کربوکسامید با استفاده از رگرسیون خطی چندگانه و شبکه عصبی مصنوعی انجام شد. بیش از ۶۰۰ توصیف کننده نظری برای هر مولکول محاسبه شد. بهترین مجموعه از توصیف کننده های محاسبه شده با روش گام به گام انتخاب شد. برای مدلسازی QSAR از رگرسیون خطی چندگانه و شبکه عصبی مصنوعی به عنوان سیستم خطی و غیرخطی استفاده شد. برای هر دو مدل، مجموعه آموزش کیفیت آماری خوبی را نشان دادند اما نتایج پیش بینی های مجموعه آزمون مدل غیرخطی به مراتب از مدل خطی برتری دارد. بر اساس نتایج مدل، الکترونگاتیوی، جرم اتمی و هندسی مولکول عوامل مهمی در کنترل فعالیت ضد باکتری سل هستند.

مراجع

1. P.T V. Nguyen. T. Van Dat. S. Mizukami. D.L. H. Nguyen. F. Mosaddeque. S.N. Kim. D.H.B. Nguyen. O.T. Dinh. T.L.Vo. G.L.T. Nguyen. C. Quoc Duong. S. Mizuta. D.N.H. Tam. M.P. Truong. N. T. Huy, K. Hirayama, *Malar. J.* 20, 264 (2021).

2. E. Yuanita. Sudirman. N.K.T. Dharmayani. M. Ulfa, J. Syahri, *J. Clin. Tuberc. Other Mycobact. Dis.* 21, 100203 (2020).

Quantitative structure–activity relationship on a series of imidazole [1, 2-a] pyridinecarboxamide derivatives as anti-tuberculosis agents

Mohsen Nekoeinia^{1,*}, Saeed Yousefinejad²

¹Department of chemistry, Payame Noor Unvierstiy, Tehran, Iran

² Research Center for Health Sciences, Institute of Health, Shiraz University of Medical Sciences, Shiraz, Iran

Abstract: Tuberculosis drug resistance is still one of the most important challenges in the treatment of this infectious disease, and therefore the discovery and development of new effective anti-tuberculosis drugs are always of interest to researchers. In this study, Quantitative structure – activity relationship (QSAR) analysis was applied on a series of imidazole[1,2-a] pyridinecarboxamide derivatives as anti-tuberculosis agents. The biological activity of the 18 derivatives were estimated by multiple linear regression and artificial neural network approaches. The four molecular descriptors (nCl, MATS8m, BELe4 and GATS8e) were selected by using stepwise multiple linear regression. The best results of artificial neural network were obtained with a 5-5-1 architecture trained with the feed forward backpropagation algorithm. An external test set containing 5 compounds for evaluating the model's predictive ability was used. The results showed that the artificial neural network approach provides better predictive power compared with multiple linear regression. According to the results of this study, electronegativity, atomic masses and molecular geometry have been found to be important factors controlling the anti-tuberculosis activity.

Keywords: QSAR, Artificial neural networks, Imidazole [1,2-a] pyridinecarboxamide derivatives, Anti-tuberculosis