

Quantitative Structure-Activity Relationship Study on Thiosemicarbazone Derivatives as Antitubercular agents Using Artificial Neural Network and Multiple Linear Regression

Mehdi Mousavi¹,
Fereidoon Daryaei²,
Omid Ranjbaran³,
Behnam Mohseni⁴,
Saeideh Taheri³,
Abdolreza Hassanzadeh⁵

¹ Associate Professor, Department of Chemistry, Shahid Bahonar University of Kerman, Kerman, Iran

² Assistant Professor, Department of Medicinal Chemistry, Faculty of Pharmacy, Kerman University of Medical Sciences, Kerman, Iran

³ Doctor of Pharmacy, Department of Medicinal Chemistry, Faculty of Pharmacy, Kerman University of Medical Sciences, Kerman, Iran

⁴ PhD in Chemistry, Department of Chemistry, Shahid Bahonar University of Kerman, Kerman, Iran

⁵ Assistant Professor, Pharmaceutics Research Center, Institute of Neuropharmacology, Kerman University of Medical Sciences, Kerman, Iran

(Received September 1, 2019 ; Accepted February 23, 2020)

Abstract

Background and purpose: Nonlinear analysis methods for quantitative structure–activity relationship (QSAR) studies better describe molecular behaviors, than linear analysis. Artificial neural networks are mathematical models and algorithms which imitate the information process and learning of human brain. Some S-alkyl derivatives of thiosemicarbazone are shown to be beneficial in prevention and treatment of mycobacterial infections and this study seeks to find out the relationship between structural features and the anti-tuberculosis activity of these compounds.

Materials and methods: Multiple linear regression and Bayesian regularized artificial neural network (BRANN) for 47 compounds of thiosemicarbazone derivatives were designed using QSAR approaches. Descriptors were selected from a pool of 343 descriptors by stepwise selection and backward elimination. A three layer Bayesian regularized back-propagation feed-forward network was designed, optimized, and evaluated using MATLAB version R2009a.

Results: The best model with 6 descriptors was found using multiple linear regression analysis: $\text{Log MIC} = 2.592 + (0.067 \pm 0.018) \text{PMIX} - (0.066 \pm 0.017) \text{PMIZ} - (1.706 \pm 1.600) \text{Qneg} - (0.235 \pm 0.039) \text{RDF030p} + (0.118 \pm 0.026) \text{RDF 140u} - (0.064 \pm 0.021) \text{RDF060p}$. The best BRANN model was a three-layer network with three nodes in its hidden layer.

Conclusion: The BRANN model has a better predictive power than linear models and may better predict the anti-tuberculosis activity of new compounds with similar backbone of thiosemicarbazone moiety.

Keywords: quantitative structure–activity relationship, multiple linear regression, Bayesian regularized artificial neural network, Thiosemicarbazone derivatives

J Mazandaran Univ Med Sci 2019; 30 (184): 106-118 (Persian).

* Corresponding Author: Abdolreza Hassanzadeh - Pharmaceutics Research Center, Institute of Neuropharmacology, Kerman University of Medical Sciences, Kerman, Iran (E-mail: a_hassanzadeh@kmu.ac.ir)

بررسی رابطه کمی ساختمان با اثر ضد سل مشتقات تیوسمی کاربازونی به روش شبکه عصبی مصنوعی و رگرسیون خطی چندگانه

مهدی موسوی¹
فریدون دریایی²
امید رنجبران³
بهنام محسنی⁴
سعیده طاهری³
عبدالرضا حسن زاده⁵

چکیده

سابقه و هدف: روش‌های مدل سازی غیر خطی برای مطالعات رابطه کمی ساختمان - اثر، راه‌های گویاتری نسبت به روش‌های خطی، برای رفتارهای مولکولی هستند. شبکه‌های عصبی مصنوعی، مدل‌ها و الگوریتم‌های ریاضی هستند که پردازش اطلاعات و یادگیری مغز انسان را تقلید می‌کنند. نشان داده شده است برخی مشتقات S-alkyl تیوسمی کاربازون در پیش‌گیری و درمان عفونت‌های مایکو باکتریایی اثر بخش بوده‌اند. این مطالعه با هدف یافتن ارتباط ساختار با اثر این ترکیبات انجام پذیرفت.

مواد و روش‌ها: در این مطالعه وابستگی کمی کنش و ساختار (QSAR)، از رگرسیون خطی چندگانه و شبکه عصبی مصنوعی با تنظیم بایسین برای 47 ترکیب از مشتقات تیوسمی کاربازون، استفاده گردید. توصیف‌کننده‌ها از روش انتخاب و حذف گام به گام از مجموع 343 توصیف‌کننده، انتخاب شدند. یک شبکه‌ی سه لایه‌ی پیش‌خور پس‌انتشار با تنظیم بایسین به وسیله نرم‌افزار MATLAB نسخه R2009a طراحی، بهینه و ارزیابی شد.

یافته‌ها: پس از رگرسیون خطی چندگانه یک مدل با 6 توصیف‌کننده حاصل شد:
 $\text{Log MIC} = 2/592 + (0/067 \pm 0/018) \text{PMIX} - (0/066 \pm 0/017) \text{PMIZ} - (1/706 \pm 1/600) \text{Qneg} - (0/235 \pm 0/039) \text{RDF030p} + (0/118 \pm 0/026) \text{RDF140u} - (0/064 \pm 0/021) \text{RDF060p}$
سه گره در لایه مخفی بود.

استنتاج: مدل BR-ANN قدرت پیش‌بینی‌کنندگی بیش‌تری نسبت به مدل‌های خطی دارد و احتمالاً بهتر می‌تواند فعالیت ضد سلی ترکیبات جدید با شالوده ساختاری یکسان در میان مشتقات تیوسمی کاربازون را پیش‌بینی کند.

واژه‌های کلیدی: رابطه کمی ساختمان - اثر، رگرسیون خطی چندگانه، شبکه‌ی عصبی مصنوعی با تنظیم بایسین، مشتقات تیوسمی کاربازون

مقدمه

براساس آمار سازمان بهداشت جهانی، سل یکی از ده علت مرگ و میر جمعیت انسانی است و در میان مرگ ناشی از بیماری‌های عفونی، اصلی‌ترین علت مرگ (بالا تر از عفونت ایدز) می‌باشد. در سال 2017،

E-mail: a_hassanzadeh@kmu.ac.ir

مؤلف مسئول: عبدالرضا حسن زاده - کرمان: دانشگاه علوم پزشکی کرمان، دانشکده داروسازی، گروه شیمی دارویی

1. دانشیار، گروه شیمی، دانشگاه شهید باهنر کرمان، کرمان، ایران

2. استادیار، گروه شیمی دارویی، دانشکده داروسازی، دانشگاه علوم پزشکی کرمان، کرمان، ایران

3. دکتر داروساز، گروه شیمی دارویی، دانشکده داروسازی، دانشگاه علوم پزشکی کرمان، کرمان، ایران

4. دکترای شیمی، دانشگاه شهید باهنر کرمان، کرمان، ایران

5. استادیار، مرکز تحقیقات فارماسیوتیکس، پژوهشکده نوروفارماکولوژی، دانشگاه علوم پزشکی کرمان، کرمان، ایران

تاریخ دریافت: 1398/6/10 تاریخ ارجاع جهت اصلاحات: 1398/8/14 تاریخ تصویب: 1398/12/4

به طور متوسط $1/3$ میلیون مرگ در میان جمعیت اچ آی وی منفی و 300 هزار مرگ هم ناشی از سل در بیماران اچ آی وی مثبت گزارش شده است. شیوع سل مقاوم به دارو یک مشکل اساسی برای بهداشت جهانی است. براساس بهترین تخمین، در سال 2017، 558 هزار نفر دچار سل مقاوم به ریفامپیسین (RR-TB)، مؤثرترین داروی خط اول درمان بودند که از میان اینها 82 درصد به چند داروی ضد سل مقاوم بودند (MDR-TB). در این سال به طور گلوبال، $3/5$ درصد از موارد جدید سل و 18 درصد از موارد قدیمی سل مبتلا به MDR-TB و یا RR-TB گزارش شده است (1). مواردی از بیماری سل که به ایزونیازید و ریفامپین مقاومند و علاوه به یکی از فلوئوروکینولونها و حداقل یکی از سه داروی تزریقی خط دوم درمان (آمیکاسین، کاپرئومایسین و کانامایسین) هم مقاومند جزء (Extensively drug-resistant TB) (XDR-TB)، تقسیم‌بندی می‌شوند که طبق اعلان WHO در سال 2017، به میزان $8/5$ درصد از کل موارد MDR-TB گزارش شده است (1). جامعه ما با چالش‌هایی مانند بیماری‌های جدید، مقاومت‌های دارویی و ورود مواد شیمیایی به محیط زیست روبرو است که نیاز فوری به طراحی داروهای جدید با خواص بهبود یافته و با عوارض جانبی کاهش یافته می‌باشد. علاوه بر این، ارزیابی خطر آسیب مواد شیمیایی رها شده در محیط زیست و رویکرد به سمت روش‌های سنتتیک کم خطر برای محیط زیست بیش از پیش محسوس است (2،3). با بروز بیماری‌های جدید و مقاومت‌ها و مشکلات دارویی، نیاز روزافزونی به ارائه‌ی مدل‌های پیش‌گویانه با هدف حل مشکلات سلامت وجود دارد که از مهم‌ترین آنها طراحی داروهای جدید با خواص بهبود یافته و اثرات ناخواسته‌ی کم‌تر است. هزینه‌ی بالا و زمان طولانی کشف و توسعه داروهای موثر، سرمایه‌ی بسیار زیادی از شرکت‌های داروسازی طلب می‌کند. ساخت یک ترکیب جدید به روش‌های سنتی برای تست بیولوژیک، ممکن است چندین ماه طول بکشد.

در این میان موفقیت راهبرد طراحی مولکولی بر پایه‌ی ساختمان، در صنعت داروسازی به اثبات رسیده است (4). روش‌های رایانه‌ای می‌توانند برای کمک و تسریع روند کشف و ساخت داروهای جدید به کار گرفته شوند (5). طراحی مولکولی به کمک رایانه، مجموعه‌ای از شاخه‌ها است که بر پیش‌گویی فعالیت‌ها و خواص ساختارهایی تست نشده و حتی ساخته نشده و مجازی تمرکز دارد که می‌توان به بررسی کمی رابطه‌ی ساختمان با اثر (Quantitative Structure-Activity Relationship: QSAR)، تلاشی در جهت یافتن رابطه بین خواص ساختاری و مولکولی دسته‌ای از ترکیبات مشابه، با عملکرد بیولوژیک آنها، به وسیله‌ی روش‌های آماری، اشاره کرد (6-9).

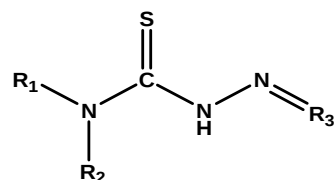
بر خلاف روش‌های کلاسیک آماری همچون آنالیز رگرسیون، شبکه‌های عصبی مصنوعی امکان بررسی روابط پیچیده و غیر خطی را دارند. بر همین اساس این شبکه‌ها به شکل ایده‌آلی برای کاربرد در طراحی دارو براساس روابط QSAR، مناسب هستند. اگرچه روابط غیر خطی بین متغیرها می‌تواند به وسیله مدل‌های متنوع رگرسیونی مدل‌سازی شود ولی طبیعت و ذات این روابط طوری است که چنین مدل‌هایی نمی‌توانند چنین روابطی را به خوبی تجزیه و تحلیل کنند. شبکه‌های عصبی مصنوعی (ANN) به عنوان یک مدل تشخیص روندهای ناپارامتری و غیر خطی می‌تواند روابط غیر خطی پیچیده یا لایه‌های پنهان بین متغیرهای وابسته و مستقل را بیابد و با دقت بهتری نسبت به روش‌های رگرسیونی عمل نماید (10). برای مثال، در مطالعه Chaloulakoua و همکاران در پیش‌بینی تخمین میزان غلظت ازن در فصل تابستان و نیز در مطالعه‌ی نکویی و همکاران در پیش‌بینی فعالیت دارویی برخی از مشتقات آمینواسیدها به عنوان بازدارنده‌های هیستون دی استیلاز جهت درمان سرطان و برخی از بیماری‌ها، برتری مدل شبکه عصبی مصنوعی بر رگرسیون خطی چندگانه نشان داده شده است. به طوری که در مطالعات QSAR، اگر پراکندگی عددی یک متغیر در نمونه‌های مختلف کم باشد، شانس پیدا کردن

مواد و روش ها

این مطالعه به شکل یک مطالعه *in silico* بر اساس اثر ضد سلی یک سری 47 تایی از مشتقات تیوسمی کاربازون صورت پذیرفت.

انتخاب و دسته بندی داده ها

داده های این مطالعه شامل 47 ترکیب از مشتقات تیوسمی کاربازون است. ساختار کلی این ترکیبات در تصویر شماره 1 داده شده است.



تصویر شماره 1: ساختار کلی ترکیبات مشتقات تیوسمی کاربازون

از اثرات ضد سل این ترکیبات که در شرایط برون تن (MIC) بر روی *Mycobacterium tuberculosis* (H37RV) در مطالعات شناخته شده دیگری اندازه گیری شده است، استفاده گردید (20-17). این ترکیبات به دو دسته مرجع (Reference set) و پیش بینی (Prediction set) تقسیم بندی شدند، با توجه به این نکته که دسته پیش بینی باید معرف خوبی از دسته مرجع باشد، دسته مرجع شامل 36 ترکیب و دسته پیش بینی شامل 9 ترکیب در نظر گرفته شد. ساختار مولکول های مورد مطالعه به همراه اثرات ضد سل آن ها در تصویر شماره 2 و جدول شماره 1 نشان داده شده است. در این جدول مولکول هایی که با ستاره (*) نشان داده شده اند در دسته پیش بینی و بقیه در دسته مرجع قرار گرفتند.

انتخاب توصیف کننده ها

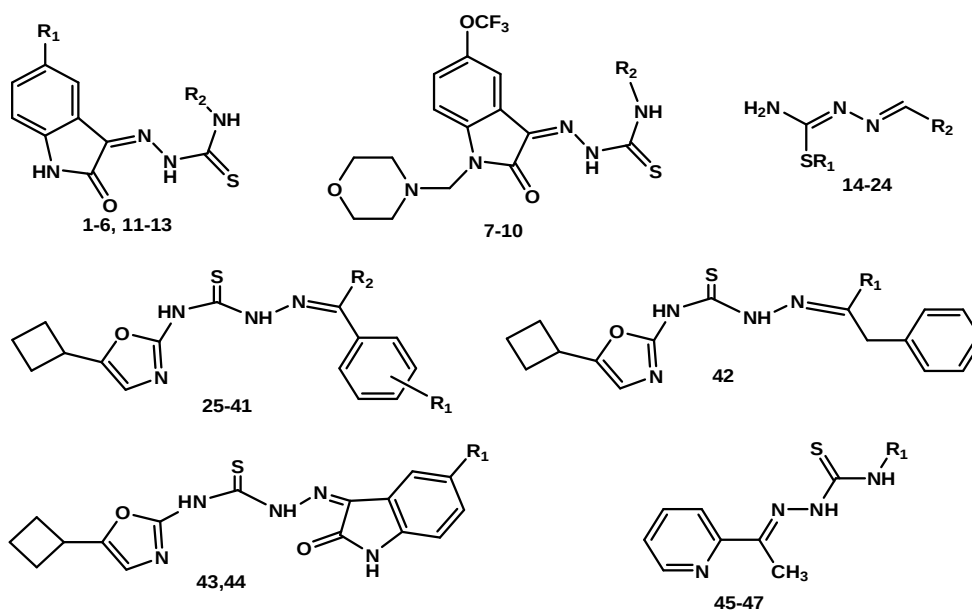
به کمک نرم افزارهای Dragon و HypeChem، تعداد 1947 توصیف کننده برای ترکیبات مورد

یک رابطه ی خطی معتبر افزایش می یابد ولی این مسأله محقق را در انتخاب نمونه هایش کمی محدود می کند (11، 12). به نظر می رسد، زمانی که بین پارامترهای مورد بررسی رابطه غیرخطی آشکاری وجود نداشته باشد در مواردی شبکه عصبی نمی تواند در مقایسه با رگرسیون نتیجه بهتری را ارائه دهد. برای مثال، در مطالعه QSAR منصوریان و همکاران برای ارزیابی فعالیت مشتقات گزانتین به عنوان آنتاگونیست های رسپتور آدنوزین، برتری روش حداقل مربعات جزئی بر رگرسیون تعمیم یافته شبکه عصبی نشان داده شده است (13). شبکه های عصبی مصنوعی، مدل های ریاضی و الگوریتم هایی هستند که برای تقلید از پردازش اطلاعات و کسب دانش مغز انسان طراحی می شوند. این شبکه ها از واحدهای پایه ای به نام نرون های مصنوعی ساخته شده اند. یک نرون یک سری از سیگنال های ورودی را جمع و از طریق یک تابع انتقال آن ها را به سیگنال های خروجی تبدیل می کند. شبکه ای از این نرون ها از طریق تغییر وزن تاثیر هر یک از نرون هایش می تواند روابط را یاد بگیرد. روش های شبکه عصبی مصنوعی متنوع، تاکنون توسعه زیادی پیدا کرده اند که از آن میان روش Back-propagation Network به دلیل توان مدل سازی، به دفعات در مطالعات QSAR به کار گرفته می شود (5).

نتایج مطالعات نشان داده است که بعضی مشتقات تیوسمی کاربازول (تیواستازون) فعالیت خوبی بر ضد *M. tuberculosis* و *M. avium* داشته اند (14). همچنین نشان داده شده است بعضی مشتقات ساخته شده تیوسمی کاربازون ها و آنالوگ های S-alkyl آن ها می توانند در درمان و پیشگیری از عفونت های مایکرو باکتریایی مفید باشند (17-15). این مطالعه با هدف دستیابی به یک رابطه QSAR غیر خطی، انجام پذیرفت تا رابطه ی کمی ساختمان با اثر ضد سل چند مشتق تیوسمی کاربازونی، با ارائه ی مدل به کمک شبکه عصبی مصنوعی بررسی گردد.

برای مطالعات بعدی باقی ماند. این تعداد از توصیف کننده‌ها از دسته‌های مختلف همچون توصیف کننده‌های ساختاری، هندسی، توابع توزیع محوری، گروه‌های عاملی، شارژ و آروماتیسیته انتخاب شدند.

مطالعه، محاسبه و وارد مطالعه شدند. در ارزیابی توصیف کننده‌ها، متغیرهایی که مقادیر ثابت یا صفر داشتند، حذف گردید و از آن‌هایی که همبستگی بالایی ($r > 0/9$) با یکدیگر داشتند یکی را نگه داشته و بقیه حذف شدند. بدین ترتیب تعداد 343 توصیف کننده



تصویر شماره 2: ساختار کلی ترکیبات مورد مطالعه

جدول شماره 1: مشخصات ساختاری ترکیبات و حداقل غلظت بازدارنده گی روی مایکو باکتریوم توپر کلوزیس

شماره ترکیب	R_1	R_2	حداقل غلظت بازدارندگی رشد ($\mu\text{g/ml}$)	شماره ترکیب	R_1	R_2	حداقل غلظت بازدارندگی رشد ($\mu\text{g/ml}$)
*1	CH_3	C_6H_9	0/911	25	2-OH	H	6/25
2	CH_3	4-F- C_6H_4	1/587	26	2- NO_2	H	0/78
3	CH_3	3-Br- C_6H_4	1/489	27	3- NO_2	H	0/78
*4	CF_3O	4-F- C_6H_4	1/549	28	4- NO_2	H	0/39
5	CF_3O	4-Cl- C_6H_4	1/521	29	4- CH_3	H	3/12
6	CF_3O	4-Br- C_6H_4	1/530	30	4-Cl	H	1/56
7	-	Cycl- C_6H_{11}	5/694	31	4-N(CH_3) $_2$	H	3
8	-	4-F- C_6H_4	3/057	32	4- CH_3O	H	12
9	-	4-Cl- C_6H_4	3/041	33	4-OH, 3- CH_3O	H	6/25
10	-	4-Br- C_6H_4	3/034	34	H	CH_3	3/12
11	CH_3	Cycl- C_6H_{11}	0/795	35	2-OH	CH_3	3/12
12	CH_3	C_6H_5	1/433	*36	4-OH	CH_3	3/12
*13	CH_3	4-Cl- C_6H_4	3/568	37	4- CH_3	CH_3	0/78
14	C_2H_5	4-F- C_6H_4	12/5	38	4- NH_2	CH_3	0/78
15	C_2H_5	4- CF_3 - C_6H_4	50	*39	4- NO_2	CH_3	0/2
16	iso-pr	4-Pyridinyl	12/5	40	H	C_6H_5	0/1
17	iso-pr	4-F- C_6H_4	100	41	4-Br	C_6H_5	0/05
18	iso-pr	4- CF_3 - C_6H_4	50	42	-	C_6H_5 - CH_2	6/25
*19	n- C_4H_9	3-Pyridinyl	25	43	-	H	25
20	n- C_4H_9	4-Pyridinyl	25	*44	-	F	12/5
*21	n- C_4H_9	4-F- C_6H_4	50	*45	H	-	31/3
22	n- C_4H_9	4- CF_3 - C_6H_4	25	46	CH_3	-	7/8
23	CH_2 - C_6H_5	4-Pyridinyl	50	47	C_6H_5	-	15/6
24	CH_2 - C_6H_5	4-F- C_6H_4	10				

محاسبه مدل ها و ارزیابی آن ها

1- مدل سازی خطی

هدف از این مدل سازی یافتن معادله ای خطی است که مقادیر تجربی فعالیت ضد میکروبی را با توصیف کننده های شیمیایی برگرفته از ساختار مولکولی سری ترکیبات مورد مطالعه مرتبط کند. در میان روش های ارائه این نوع مدل ها، رگرسیون خطی چند گانه (MLR) از تعادل مناسبی بین تکرار پذیری، قابلیت تفسیر و کیفیت محاسبات برخوردار است. این روش شکل توسعه یافته ی رگرسیون کلاسیک است که به بیش از یک بعد گسترش داده شده است. در این روش معادلات چند متغیره با شکل کلی زیر تولید می شود.

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

در این معادله y فعالیت بیوشیمیایی و x_i توصیف کننده مولکولی و مقادیر b ضرایب معادله است. بهترین متغیرها به روش انتخاب و حذف گام به گام در نرم افزار SPSS نسخه 11/5 انتخاب شدند. سپس توصیف کننده ها در چهار دسته هندسی، الکترونی، توپولوژیکی و فیزیکی شیمیایی تقسیم شدند و در هر کدام رگرسیون مرحله ای انجام گرفت.

ارزیابی اعتبار مدل خطی

برای ارزیابی مدل خطی از تکنیک ارزیابی متقاطع بهره گیری شد تا اعتبار مدل از نظر استقلال از سری داده ها سنجیده شود. ارزیابی متقاطع که به آن گاهی برآورد چرخشی نیز گفته می شود در آنالیزهای آماری برای تعمیم به یک سری داده ی مستقل به کار گرفته می شود. این روش مشخصاً در سیستم های پیش بینی کننده برای تخمین دقت مدل در عمل کاربرد دارد. طی روش حذف یکتایی، هر بار یک مولکول از سری داده ها کنار گذاشته شد و ساخت مدل با سایر مولکول ها انجام گرفت. سپس اثر ضد سل آن مولکول توسط مدل محاسبه شد. این کار برای تمام مولکول های سری داده ها انجام شد.

معیار ارزیابی مدل Q^2_{LOO} ، از رابطه زیر به دست می آید.

$$Q^2_{LOO} = \frac{PRESS}{SSY} = 1 - \frac{\sum_{i=1}^{47} (y_{exp} - y_{pred})^2}{\sum_{i=1}^{47} (y_{exp} - \bar{y})^2}$$

در این رابطه PRESS مجموع مجذور خطای پیش بینی و SSY مجموع مجذور انحراف از میانگین مقادیر تجربی می باشد. y_{exp} ، y_{pred} و $\bar{y}$ نیز به ترتیب بیان گر مقدار تجربی، مقدار پیش بینی شده و میانگین مقادیر تجربی هستند. در حذف 4 تایی، هر بار 4 مولکول از سری داده ها کنار گذاشته شدند و مدل با سایر مولکول ها ساخته شد. اثرات ضد سل این 4 مولکول با مدل بدست آمده محاسبه گردید. این کار تا زمانی که تمام داده ها برای یک بار از سری خارج شوند، ادامه پیدا کرد. معیار ارزیابی در این روش R^2_{L40} قرار گرفت که محاسبه ی آن مشابه Q^2_{LOO} می باشد. هر چه R^2_{L40} و Q^2_{LOO} مدل به یک نزدیک تر و مجذور مربعات میانگین خطا کمتر باشد، قدرت پیش بینی مدل بیشتر و وابستگی آن به سری داده ها کمتر است.

2- استفاده از شبکه عصبی مصنوعی بایسین برای مدل سازی

از آنجایی که شبکه ی عصبی مصنوعی به گستردگی در مطالعات QSAR مورد استفاده قرار می گیرد و مدل های مناسبی براساس روابط غیر خطی به دست می دهد، در این مطالعه از این روش استفاده شد. برنامه شبکه عصبی مصنوعی در مرکز تحقیقات شیمی ثوری در بخش شیمی دانشگاه باهنر کرمان، در محیط نرم افزار MATLAB نسخه R2009a، نوشته و به کار گرفته شد. برای استفاده از شبکه عصبی مصنوعی ابتدا یک شبکه سه لایه پس انتشار طراحی گردید. توصیف کننده های مدل خطی به عنوان ورودی های این شبکه در نظر گرفته شدند. لایه ی خروجی این شبکه، Log MIC ترکیبات تیوسمی کاربازونی می باشند.

در روند بهینه سازی شبکه، تعداد گره های لایه میانی و تعداد دفعات تکرار شبکه به طور همزمان بهینه شدند. از میان توابع متنوع آموزشی در شبکه های عصبی

مصنوعی که با ویژگی های متفاوتی شبکه را آموزش می دهند، تابع (Bayesian regularization) trainbr برای این شبکه انتخاب شد. این تابع علاوه بر این که از گیر افتادن شبکه در حداقل های موضعی و بروز خطای برازش اضافی جلوگیری می کند، از سرعت قابل قبولی نیز برخوردار می باشد. این تابع همچنین نیاز به دسته بندی های اضافه برای تعیین اعتبار و در نتیجه کوچک شدن دسته ی مرجع را نیز از بین می برد. در روند بهینه سازی شبکه تعداد گره های لایه های مخفی افزایش داده شد تا بهترین نتیجه طی آموزش شبکه بدست بیاید. این کار با اجرای شبکه با تعداد گره های متفاوت در لایه مخفی و مقایسه ی مجذور میانگین مربعات خطای هر دو دسته مرجع و پیش بینی انجام گرفت. در نهایت هدف اصلی شبکه از دیدگاه آماری کاهش مقدار مجذور میانگین مربعات خطاست که با فرمول زیر محاسبه می گردد.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2}$$

در این فرمول RMSE مجذور میانگین مربعات خطا، N تعداد ترکیبات، y_i مقدار تجربی فعالیت زیستی ترکیب i و t_i فعالیت زیستی پیش بینی شده ی ترکیب i، می باشد.

در واقع (Bayesian regularized Artificial Neural Network) BR-ANN برای کاهش این مقدار

شرط خطایی را شامل می شود که وزن ها را از طریق اصلاح مقادیر بزرگ تنظیم می کند. هر ارتباطی بین نرون ها از طریق اعدادی حقیقی که وزن نامیده می شوند برقرار می شود. وزن مقداری مثبت یا منفی است که در سیگنال نرون پردازشگر ضرب شده و به نرون بعدی منتقل می شود. نرون های لایه ورودی مقادیر خود را از متغیرهای مستقل دریافت می کنند در حالی که لایه میانی مقادیر خود را از نرون های قبلی خود دریافت می کند که حاصل تاثیر وزن ها است. در نهایت لایه ی خروجی مقادیر را از نرون های قبلی گرفته و به صورت

خروجی ارائه می دهد. از آنجایی که شبکه ای با وزن های بزرگ معمولاً مدلی با منحنی های بزرگ را ارائه می دهند، ترجیح با مقادیر کوچک برای وزن های شبکه است. این روش آموزش از مقدار MSW که در واقع میانگین حداقل وزن های شبکه است به عنوان تابع تنظیم بهره می برد. در این شبکه در لایه های ورودی و خروجی از تابع خطی و در لایه ی میانی یک تابع هایپربولیک تانژانت به عنوان توابع انتقال شبکه استفاده شد. داده های ورودی و اهداف به ترتیب از توصیف کننده های مستقل برگزیده ی مدل خطی و مقادیر تجربی Log MIC تشکیل شدند. هر دو سری این داده ها پیش از ورود به شبکه عصبی نرمال شدند.

BR-ANN با تابع آموزشی trainbr تابع آموزشی تنظیم بایسین می باشد، اجرا شد. در این تابع μ پارامتر تنظیم Marquardt است که روش تنظیم بایسین برای توقف آموزش از آن استفاده می کند. آموزش شبکه وقتی مقدار μ از 10^{40} بیش تر شد، متوقف گردید.

3- روش تجزیه و تحلیل اطلاعات

بررسی پارامترهای آماری R^2 ، SE، F توسط نرم افزار MATLAB نسخه ی R2009a انجام شد.

یافته ها

مدل سازی رگرسیون خطی چندگانه

در مرحله ی نخست مدل سازی ترکیبات تیوسمی کاربازون جهت مطالعه بررسی رابطه کمی ساختمان با اثر ضد سل آن ها، مدل سازی خطی به روش رگرسیون خطی چندگانه با انتخاب توصیف کننده ها به روش گام به گام و حذف متغیر ها انجام گرفت و اعتبار آن ها ارزیابی شد، که بهترین مدل با معادله زیر به دست آمد.

$$\begin{aligned} \text{Log MIC} = & 2/592 + (0/067 \pm 0/018) \text{PMIX} - \\ & (0/066 \pm 0/017) \text{PMIZ} - (1/706 \pm 1/600) \text{Qneg} - \\ & (0/235 \pm 0/039) \text{RDF030p} + (0/118 \pm 0/026) \\ & \text{RDF 140u} - (0/064 \pm 0/021) \text{RDF060p} \end{aligned}$$

شاخص‌های آماری مربوط به تعیین اعتبار این مدل خطی بهینه در جدول شماره 2 نشان داده شده است.

در جدول شماره 3 همبستگی توصیف کننده ها با یکدیگر و با مقادیر تجربی Log MIC نشان داده شده است. توصیف کننده ها همبستگی خوبی با یکدیگر ندارند و این به این معنی است که آن ها نماینده مستقلی از ساختار هستند و هر کدام وجه منحصر به فردی از ساختار را توصیف می کنند. هر چه این همبستگی کمتر باشد، مدل قابل اعتمادتری به دست می آید.

جدول شماره 2: مشخصات آماری مدل بهینه ی رگرسیون خطی چندگانه

$R^2_{\text{Reference}}$	$R^2_{\text{Prediction}}$	Q^2_{LOO}	R^2_{L40}	RMSE _{Reference}	RMSE _{Prediction}	RMSE _{L40}
0/8923	0/6785	0/6820	0/8242	0/2358	0/4261	0/2842

جدول شماره 3: ضرایب همبستگی بین توصیف کننده ها و مقدار تجربی فعالیت ضد سل

Log MIC	RDF140u	RDF060p	RDF030p	Qoeg	PMIZ	PMIX
					1	PMIX
					0/016	PMIZ
				1	0/160	Qoeg
			0/618	0/098	0/314	RDF030p
		1	0/657	0/626	0/340	RDF060p
	1	0/334	0/643	0/449	0/007	RDF140u
1	0/253	0/737	0/791	0/491	0/401	Log MIC

شبکه عصبی مصنوعی با تنظیم بایسین (BR-ANN)

با توجه به رفتارهای غیر خطی مولکول ها در عملکردهای زیستی، انتظار می رود که روش های مدل سازی غیر خطی برای پیش بینی ویژگی های زیستی آن ها مدل مطلوب تری ارائه دهند. بنابراین شبکه عصبی مصنوعی پس انتشار با تنظیم بایسین برای مدل سازی غیر خطی 47 مولکول طراحی شد که حاصل بهینه سازی، یک شبکه سه لایه بود که مشخصات آن در جدول شماره 4 و ویژگی های آماری آن در جدول شماره 5 آورده شده است.

تعداد بهینه گره های لایه ی میانی با مقایسه ی تغییرات RMSE که در نمودار شماره 1 نمایش داده شده است، 3 گره انتخاب شد. همان طور که در نمودار نمایان است RMSE بهینه در هر دو سری مرجع و

پیش بینی با 3 گره در لایه مخفی حاصل شده است و با افزایش گره ها مقدار RMSE با وجود کاهش در سری مرجع، در سری پیش بینی افزایش می یابد.

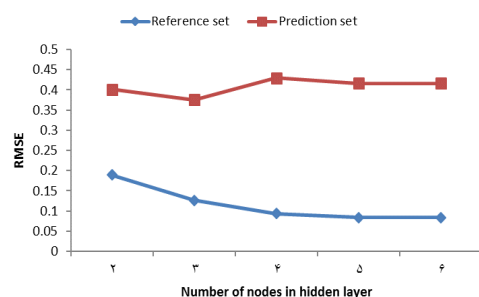
شرایط توقف شبکه و ویژگی های آن تا زمان توقف در نمودارهای شماره 2 و 3 نشان داده شده است. در نمودار شماره 3، مقدار Gradient معرف محدوده ای است که اختلاف مقادیر پیش بینی با مقادیر تجربی در آن قرار می گیرد. متغیر mu در واقع شاخص تنظیم کننده Marquardt می باشد که تعیین کننده زمان توقف آموزش شبکه است. متغیر val fail معرف دفعات شکست اعتبار سنجی شبکه در هر بار تکرار آن است. مقادیر شاخص های آماری مدل شبکه ی عصبی مصنوعی در مقایسه با مدل خطی در جدول شماره 6 ارائه شده است. مقادیر Log MIC تجربی و پیش بینی شده برای هر دو گروه مرجع و پیش بینی توسط مدل خطی و شبکه عصبی در جدول شماره 7 نشان داده شده است.

جدول شماره 4: مشخصات شبکه عصبی سه لایه بهینه شده

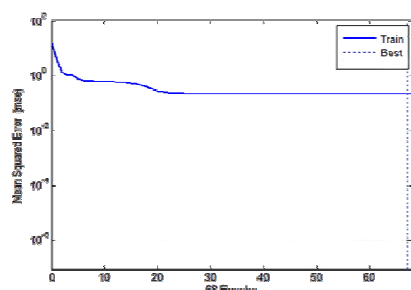
3	تعداد گره های لایه ی میانی
1	تعداد گره های لایه ی خروجی
67	تعداد دفعات تکرار
MSW	تابع تنظیم
Trainbr	تابع آموزشی
Tansig	تابع انتقال

جدول شماره 5: مشخصات آماری مربوط به مدل شبکه ی عصبی برای سری های مرجع و پیش بینی شونده

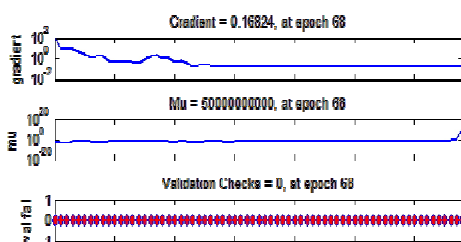
RMSE	R^2	Type of set
0/127	0/969	دسته مرجع
0/375	0/921	دسته پیش بینی



نمودار شماره 1: تغییرات مجذور میانگین مربعات خطا در دو دسته مرجع و پیش بینی به ازای تعداد گره های لایه مخفی



نمودار شماره 2: تغییرات میانگین مربعات خطا در مقابل تعداد دفعات تکرار



نمودار شماره 3: تغییرات شاخص‌های شبکه در مقابل تعداد دفعات تکرار

بحث

در مطالعات طراحی دارو بر پایه‌ی کامپیوتر، تحلیل آماری روابط بین ویژگی‌های ساختاری و اثرات درمانی با استفاده از روش‌های متنوع، درک روابط ساختار با اثرات زیستی را تسهیل می‌کند. توسعه‌ی این روش‌ها همواره در جهت سرعت بخشی و بهبود کیفیت این مطالعات صورت می‌گیرد.

در این مطالعه توصیف‌کننده‌های مختلفی برای پیش‌بینی فعالیت ضد سل 47 مولکول از مشتقات تیوسمی کاربازون به روش‌های خطی چندگانه و شبکه عصبی مصنوعی محاسبه شد و مدل‌های مطلوب از مجموعه‌ای از توصیف‌کننده‌های الکترونی و هندسی به‌ست آمدند.

توصیف‌کننده‌های هندسی PMIZ و PMIX (Principal moments of inertia) بیانگر ممان اینرسی اصلی مولکول بوده و نشانگر چگونگی توزیع جرم اتم‌های مولکول در جهات مختلف می‌باشد. این توصیف‌کننده‌ها مقادیر کمی فیزیکی مرتبط با دینامیک

جدول شماره 6: مقادیر شاخص‌های آماری داده‌های گروه مرجع و

پیش‌بینی مربوط به مدل شبکه‌ی عصبی در مقابل مدل خطی

BR-ANN (دسته پیش‌بینی)	MLR (دسته پیش‌بینی)	BR-ANN (دسته مرجع)	MLR (دسته مرجع)	
0/921	0/8785	0/969	0/8923	R ²
0/675	0/4261	0/127	0/2358	RMSE

جدول شماره 7: مقادیر تجربی و پیش‌بینی شده Log MIC توسط

مدل‌های خطی و شبکه عصبی

شماره ترکیب	لگاریتم حداقل غلظت بازدارندگی (Exp)	لگاریتم حداقل غلظت بازدارندگی (BR-ANN)	لگاریتم حداقل غلظت بازدارندگی (MLR)
*1	-0/04	0/12	0/33
2	0/20	0/27	0/11
3	0/17	0/24	0/62
*4	0/54	0/33	0/46
5	0/18	0/18	0/29
6	0/18	0/15	0/08
7	0/76	0/72	0/66
8	1/50	1/28	0/83
9	0/48	0/57	0/50
10	0/48	0/46	0/20
11	-0/10	-0/11	0/07
12	2/00	1/21	1/62
*13	0/55	0/43	0/36
14	1/10	1/49	1/46
15	1/70	1/75	1/65
16	1/10	0/25	1/44
17	2/00	1/66	1/50
18	1/70	1/59	1/69
*19	1/40	1/49	1/56
20	1/40	1/28	1/39
*21	0/49	1/29	0/27
22	1/40	1/48	1/60
23	1/70	1/44	1/51
24	-0/70	-0/46	0/37
25	0/80	0/80	1/00
26	-0/11	-0/11	-0/05
27	-0/11	-0/08	0/14
28	-0/41	-0/43	-0/25
29	0/49	0/18	0/12
30	0/19	0/08	0/13
31	0/49	0/59	0/53
32	1/10	0/58	0/69
33	0/80	0/88	1/15
34	0/49	0/42	0/57
35	0/49	0/52	0/29
*36	1/40	1/11	0/78
37	-0/11	-0/06	-0/22
38	-0/11	-0/06	-0/16
*39	0/16	0/28	0/27
40	-1/00	-0/80	-0/53
41	-1/30	-1/38	-1/20
42	0/80	0/65	-0/39
43	1/70	1/47	1/41
*44	1/10	1/00	0/73
*45	0/19	0/54	0/36
46	0/89	0/95	0/65
47	1/19	1/23	0/91

*: دسته پیش‌بین

چرخشی یک مولکول هستند. ممان اینرسی حول هر محور، از طریق فرمول زیر محاسبه می‌شود به طوری که تعداد اتم‌ها و m_i ، r_i به ترتیب جرم اتمی و فاصله عمودی از محور مورد نظر اتم i از مولکول هستند.

$$I = \sum_{i=1}^A m_i r_i^2$$

برای هر یک از ابعاد X , Y , Z ممان اینرسی به صورت زیر تعریف می‌شود.

$$I_{XX} = \sum_{i=1}^A m_i \cdot (y_i^2 + z_i^2)$$

$$I_{YY} = \sum_{i=1}^A m_i \cdot (x_i^2 + z_i^2)$$

$$I_{ZZ} = \sum_{i=1}^A m_i \cdot (x_i^2 + y_i^2)$$

مقادیر بین محوری، محصولات ممان اینرسی هستند و به صورت زیر تعریف می‌گردد.

$$I_{XY} = \sum_{i=1}^A m_i \cdot x_i \cdot y_i$$

$$I_{XZ} = \sum_{i=1}^A m_i \cdot x_i \cdot z_i$$

$$I_{YZ} = \sum_{i=1}^A m_i \cdot y_i \cdot z_i$$

با توجه به نتایج حاصل در مدل با افزایش ممان اینرسی حول محور X مقدار Log MIC افزایش و حول محور Z باعث کاهش این مقدار می‌شود.

Q_{neg} هم یک توصیف کننده الکترونی است که بیان گر مقدار کمی بار خالص منفی اتمی می‌باشد. توصیف کننده‌های مربوط به دانسیته‌ی موضعی الکترون، توسط شیمی بر پایه‌ی کامپیوتر محاسبه می‌شوند و می‌توان آن‌ها را از توصیف کننده‌های شیمی کوانتومی به شمار آورد.

بارهای الکتریکی در مولکول نیروی محرکه واکنش‌های الکترواستاتیک هستند و دانسیته‌ی موضعی الکترون نقش عمده‌ای در واکنش‌های شیمیایی، ویژگی‌های فیزیکوشیمیایی و تمایل اتصال به گیرنده ایفا می‌کنند.

Q_{neg} حداکثر بار منفی یک مولکول است که به صورت $Q_{\text{neg}} = \max_a (q_a^-)$ محاسبه می‌شود و q^- بار منفی خالص اتمی است. با توجه به ضریب منفی این توصیف کننده در مدل، افزایش Q_{neg} ، کاهش Log MIC را به دنبال دارد.

توصیف کننده‌های RDF030p ، RDF140u و RDF060p جزء دسته‌ی تابع توزیع شعاعی و از توصیف کننده‌های هندسی هستند. تابع توزیع محوری با توجه به چینش سه بعدی اتم‌ها منحصر به فرد است و در برابر جابه‌ایی و چرخش کل مولکول تغییرناپذیر است. به علاوه این تابع می‌تواند برای نمایش اطلاعات خاص در یک ساختار معین سه بعدی به انواع خاص اتم‌ها یا فواصل معین محدود شود، تا به طور مثال ممانعت فضایی یا ویژگی‌های ساختار-اثر یک مولکول را توصیف کند. شکل اصلی کد تابع توزیع شعاعی به طریق زیر محاسبه می‌شود. f فاکتور مقیاس، w خصوصیات ویژه‌ی اتمی i و j ، r_{ij} فاصله‌ی بین اتم i و j و A تعداد اتم‌هاست. β پارامتر هموارسازی است، که بیانگر توزیع احتمال فاصله بین اتمی یک اتم خاص است. β می‌تواند معرف حرکات اتم‌ها باشد.

$$g(R) = f \cdot \sum_{i=1}^{A-1} \sum_{j=i+1}^A w_i \cdot w_j \cdot e^{-\beta \cdot (R-r_{ij})^2}$$

نتایج مقادیر پیش‌بینی درد و مدل و اعتبار سنجی آن‌ها، نشان می‌دهد که مدل حاصل از شبکه در مقایسه با مدل خطی بهبود قابل ملاحظه‌ای یافته و مقادیر مجذور میانگین مربعات خطا برای هر دو گروه مرجع و پیش‌بینی کاهش پیدا کرده است (جدول شماره 6). مقایسه همبستگی مقادیر پیش‌بینی شده در دو مدل با مقادیر تجربی در هر دو گروه مرجع و پیش‌بینی، که در مقایسه نتایج به‌دست آمده است و در نمودارهای 4، 5، 6 و 7 مشخص است، نشانگر برتری مدل شبکه عصبی بر مدل خطی در این مطالعه است. به علاوه کاهش مقادیر باقیمانده که اختلاف مقادیر پیش‌بینی با مقادیر تجربی

با توجه به نتایج مطلوب در پیش‌بینی اثرات ضد سل این دسته از ترکیبات، می‌توان انتظار داشت که با ارائه‌ی مقادیر محاسبه شده برای ترکیباتی با ساختارهای مشابه به شبکه‌ی آموزش دیده‌ی حاصل، اثرات زیستی آن‌ها با اعتبار قابل قبولی پیش‌بینی شود. بدین ترتیب به کمک یافته‌های این مطالعه می‌توان، مشتقات تیوسمی کاربازونی را طوری منطقی طراحی کرد که فعالیت ضد باکتریایی آن‌ها افزوده گردد. این مطالعه نشان داد که توصیف‌گرهای الکترونی و هندسی بیش‌ترین تأثیر را در فعالیت ضد سلی مشتقات تیوسمی کاربازونی ایفا می‌کنند. هر چند در این نوع از مطالعات فرض ثابت دانستن شرایط محیطی و نادیده گرفتن برهم‌کنش‌های ابرالکترونی مولکول‌های زیستی با مولکول‌های مورد نظر در شرایط درون تنی اجتناب‌ناپذیر است، اما با توجه به نقاط قوت این نوع تخمین‌ها و همچنین افزایش چشمگیر در سرعت غربالگری تعداد بسیار زیادی از ترکیبات، می‌توان از این ابزار به خوبی در معرفی ترکیبات رهبر برای مراحل بعدی مطالعات طراحی دارو، سود جست.

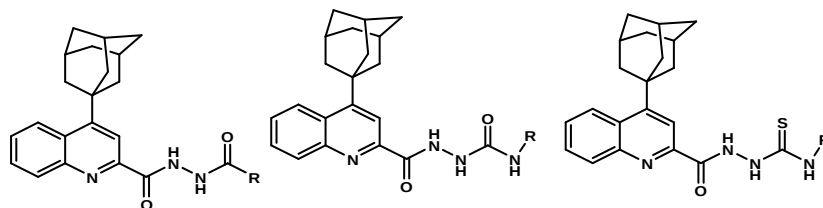
سپاسگزاری

مطالعه حاضر بخشی از طرح تحقیقاتی مصوب دانشگاه علوم پزشکی و خدمات بهداشتی درمانی کرمان به شماره 91/173 می‌باشد که از دانشگاه علوم پزشکی کرمان به عنوان تأمین‌کننده‌ی منابع مالی مطالعه حاضر تشکر و قدردانی می‌شود.

هستند، در دو گروه مرجع و پیش‌بینی، با مقایسه‌ی نمودارهای 8، 9، 10 و 11، به خوبی برتری مدل شبکه عصبی مصنوعی بر مدل خطی را نشان می‌دهد.

در مطالعه‌ای که توسط شهلائی و همکاران روی ترکیبات مهارکننده رسپتورهای کموکینی (eceptors Chemokine) - به عنوان ترکیبات با فعالیت بالقوه ضد التهاب - صورت گرفت، هم‌نشان داده شد که روش‌های غیر خطی آنالیز به صورت بهتری فعالیت بیولوژیکی ترکیبات مورد مطالعه را پیش‌بینی می‌کند (21). در مورد استخراج رابطه‌ی ساختار - فعالیت ضد سلی ترکیبات مورد مطالعه، مستندی در مقالات پیدا نشد اما در مطالعه‌ای که توسط Patel و همکاران در پیدا کردن رابطه‌ی ساختار - فعالیت ضد سلی مشتقات هیدرازیدی، سمی کاربازیدی و تیوسمی کاربازیدی از کینولین‌های دارای استخلاف آدامانتان در موقعیت چهار (تصویر شماره 4) صورت گرفته، مشخص شده است که گذاشتن استخلافات بزرگ و استخلافات الکترون کشنده روی موقعیت پارای حلقه‌ی فنیل در موقعیت R، می‌تواند فعالیت ضد سلی این ترکیبات را افزایش دهد (22).

در این مطالعه توصیف‌کننده‌های مختلفی برای هر مولکول محاسبه، و از بین آن‌ها 6 توصیف‌کننده‌ای که بیش‌ترین ارتباط را با فعالیت ضد سلی ترکیب‌ها داشتند، انتخاب شدند. سپس مدل‌سازی و پیش‌بینی این توصیف‌کننده‌ها با روش‌های MLR و ANN انجام شد. نتایج حاصل از مدل‌سازی با شبکه‌ی عصبی به خوبی بهبود در شاخص‌های آماری و توان پیش‌بینی مدل نسبت به مدل خطی را نمایش می‌دهد.



شکل شماره ۴: برخی مشتقات هیدرازید، سمی کاربازید و تیوسمی کاربازید از کینولین 4-آدامانتان-1-یل با اثر ضد سل

References

1. World Health Organization (WHO). Global Tuberculosis Report 2018. Geneva: World Health Organization; 2018. Available at: <https://www.who.int/medicinedocs/en/m/abstract/Js23553en>. Accessed May 2, 2019.
2. Martin YC. Quantitative drug design: a critical introduction. New York: Marcel Dekker; 1978.
3. Sanz F, Martín M, Pérez J, Turmo J, Mitjana A, Moreno V, et al. Quantitative approaches to drug design. Amsterdam: Elsevier; 1983.
4. Saliner AG. Molecular quantum similarity in QSAR: application in computer-aided drug design. PhD thesis, Universitat de Girona, Girona, 2004.
5. Terfloth L, Gasteiger J. Neural networks and genetic algorithms in drug design. *Drug Discov Today* 2001; 6 (suppl 2): 102-108.
6. Thomas G. Medicinal Chemistry: an introduction. 2nd ed. Chichester: John Wiley & Sons, Ltd; 2008.
7. Hansch C, Hoekman D, Gao H. Comparative QSAR: toward a deeper understanding of chemobiological interactions. *Chem Rev* 1996; 96(3): 1045-1076.
8. Schmidli H. Multivariate prediction for QSAR. *Chemometr Intell Lab Syst* 1997; 37(1): 125-136.
9. Hansch C, Kurup A, Garg R, Gao H. Chem-bioinformatics and QSAR: a review of QSAR lacking positive hydrophobic terms. *Chem Rev* 2001; 101(3): 619-672.
10. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; 346: 1075-9.
11. Chaloulakou A, Saisana M, Spyrellis N. Comparative assessment of neural networks and regression models for forecasting summertime ozone in Athens. *Sci Total Environ* 2003; 313: 1-13.
12. Nekoei M, Mohammad Hosseini M, Rahimi M, Alavi-Gharahbagh A. Linear and nonlinear quantitative structure-activity relationship for prediction of drug activity of some amino acid derivatives. *J Appl Res Chem (JARC)* 2012; 6(4): 61-53 (Persian).
13. Mansourian M, Saghaie L, Fassihi A, Madadkar-Sobhani A, Mahnam K. Linear and nonlinear QSAR modeling of 1,3,8-substituted-9-deazaxanthines as potential selective A2BAR antagonists. *Med Chem Res* 2013; 22: 4549-4567.
14. Bermudez LE, Reynolds R, Kolonoski P, Aralar P, Inderlied CB, Young LS. Thiosemicarbazole (thiacetazone-like) compound with activity against *Mycobacterium avium* in mice. *Antimicrob Agents Chemother* 2003; 47(8): 2685-2687.
15. Karali N, Gürsoy A, Kandemirli F, Shvets N, Kaynak FB, Ozbey S, et al. Synthesis and structure-antituberculosis activity relationship of 1H-indole-2,3-dione derivatives. *Bioorg Med Chem* 2007; 15(17): 5888-5904.
16. De Logu A, Saddi M, Onnis V, Sanna C, Congiu C, Borgna R, et al. In vitro antimycobacterial activity of newly synthesised S-alkylisothiosemicarbazone derivatives and synergistic interactions in combination with rifamycins against *Mycobacterium avium*. *Int J Antimicrob Agents* 2005; 26(1): 28-32.
17. Cocco MT, Congiu C, Onnis V, Pellerano ML, De Logu A. Synthesis and antimycobacterial activity of new S-alkylisothiosemicarbazone derivatives. *Bioorg Med Chem* 2002; 10(3): 501-506.

18. da S Maia PI, Pavan FR, Leite CQF, Lemos SS, de Sousa GF, Batista AA, et al. Vanadium complexes with thiosemicarbazones: Synthesis, characterization, crystal structures and anti-*Mycobacterium tuberculosis* activity. *Polyhedron* 2009; 28(2): 398-406.
19. Güzel Ö, Karali N, Salman A. Synthesis and antituberculosis activity of 5-methyl/trifluoromethoxy-1 H-indole-2, 3-dione 3-thiosemicarbazone derivatives. *Bioorg Med Chem* 2008; 16(19): 8976-8987.
20. Sriram D, Yogeeswari P, Thirumurugan R, Pavana RK. Discovery of new antitubercular oxazolyl thiosemicarbazones. *J Med Chem*. 2006; 49(12): 3448-3450.
21. Shahlai M, Fassihi A, Saghaie L. Application of PC-ANN and PC-LS-SVM in QSAR of CCR1 antagonist compounds: a comparative study. *Eur J Med Chem* 2010; 45(4): 1572-1582.
22. Patel SR, Gangwal R, Sangamwar AT, Jain R. Synthesis, biological evaluation and 3D-QSAR study of hydrazide, semicarbazide and thiosemicarbazide derivatives of 4-(adamantan-1-yl) quinoline as anti-tuberculosis agents. *Eur J Med Chem* 2014; 85: 255-267.