





دانشگاه ولی عصر (عج) رفسنجان

دانشکده علوم پایه

گروه شیمی

پایان نامه کارشناسی ارشد رشته‌ی شیمی گرایش تجزیه

عنوان پایان نامه:

مطالعه ارتباط کمی بین ساختار با فعالیت ضد انگل تری پانوزوماید در یک سری از مشتقات فروکسان – آلکیل نیترات با استفاده از شبکه های عصبی مصنوعی

استاد راهنما:

دکتر زهرا گرکانی نژاد

دانشجو:

نوشین پاریزی

مهر ۱۳۸۸



دانشگاه ولی عصر (عج) رفسنجان

دانشکده علوم پایه

گروه شیمی

پایان نامه کارشناسی ارشد رشته‌ی شیمی گرایش تجزیه

خانم نوشین پاریزی

تحت عنوان:

مطالعه ارتباط کمی بین ساختار با فعالیت ضد انگل تری پانوزوماید در یک سری از

مشتقات فروکسان – آلکیل نیترا با استفاده از شبکه های عصبی مصنوعی



Vali-E-Asr University of Rafsanjan

Faculty of Science

Department of Chemistry

M.Sc.Thesis

**Quantitative structure-anti trypanosomatid-activity relationship study
of some Furoxan-alkylnitrate derivatives by using Artificial Neural
Networks**

Supervisor:

Dr Zahra Garkani-Nejad

By:

Noushin Parizi

(2009/October)

این صفحه بهترین فرصت مکتوبی است که از تمامی معلمان دلسوزه تشکری هر چند ناچیز داشته باشم.

از اولین معلمی که نفستین مرف الفبا را به من آموخت تا به امروز که این دفتر را به یاری معلمی دیگر بستم آنها که اولین بار قلم به دستم دادند و آنها که امروز یاریم کردند.

استاد راهنمای عزیزم سرکار خانم دکتر زهرا گرکانی نژاد

و

کلیه اساتید محترم گروه شیمی

نام و نام خانوادگی: نوشین پاریزی

مقطع تحصیلی: کارشناسی ارشد

رشته و گرایش: شیمی تجزیه

دانشکده: علوم پایه

استاد راهنما: دکتر زهرا گرکانی نژاد

تاریخ دفاع: ۱۳۸۸/۷/۲۷

چکیده

مشتقات آلکیل نیترات - فروکسان و ترکیبات وابسته به آن بازدارنده ضد انگل تری پانوزومائیید بسیار موثری می باشند. خاصیت بازدارندگی این ترکیبات در درمان بیماری های انگلی و عفونی مانند شاگاس (که عامل آن انگل تری پانوزومای کروزای است) و سالک (که عامل آن انگل لیثمانیا است) مورد استفاده قرار می گیرد. هدف از مطالعه حاضر بررسی رابطه کمی ساختار - فعالیت (QSAR) مشتقات آلکیل نیترات - فروکسان و ترکیبات وابسته به آن با استفاده از روش های رگرسیون خطی چند تایی (MLR) و شبکه های عصبی مصنوعی (ANN) و روش حداقل مربعات جزئی (PLS) می باشد.

مقادیر ID₅₀ یک گروه از داده ها شامل ۱۲۰ مشتق آلکیل نیترات - فروکسان و ترکیبات مربوط به آن بکار گرفته شده است. توصیف کننده های مناسبتر (MATS8e، R3u، C029، SIC₁، RDF030m، E3m، Mor11m، Mor28m، Lop) با استفاده از روش رگرسیون خطی چند گانه انتخاب گردید. در برنامه شبکه عصبی از شیوه آموزش پس انتشار خطا در یک شبکه ۹×۹×۱ استفاده شد. خطای استاندارد برای این شبکه ۱/۴۳۶ و ضریب همبستگی بین مقادیر ID₅₀ تجربی و محاسبه شده ۰/۹۹۳ می باشد. برای روش MLR نیز خطای استاندارد برابر با ۷/۵۳ و ضریب همبستگی برابر با ۰/۷۹۲ بدست آمد. تداوم و اعتبار مدل های ANN و MLR با استفاده از تکنیک ارزیابی اعتبار متقاطع مورد بررسی قرار گرفت. نتایج بدست آمده در این پژوهش نشان دهنده برتری مدل شبکه عصبی مصنوعی نسبت به مدل های رگرسیون خطی چندگانه و حداقل مربعات جزئی در پیش بینی مقادیر ID₅₀ مشتقات آلکیل نیترات - فروکسان می باشد.

فهرست مطالب

عنوان	صفحه
فصل اول: کمومتری	۱
۱-۱ روش های پارامتری	۴
۱-۱-۱ نرم افزارهای مورد استفاده در روش های پارامتری	۵
۱-۱-۱-۱ بسته نرم افزاری Hyperchem	۵
۱-۱-۱-۲ بسته نرم افزاری MOPAC	۵
۱-۱-۱-۳ بسته نرم افزاری SPSS	۵
۱-۱-۱-۴ بسته نرم افزاری Dragon	۶
فصل دوم: شبکه های عصبی مصنوعی	۹
مقدمه	۹
۱-۲ تاریخچه شبکه های عصبی مصنوعی	۱۱
۱-۱-۲ پرسپترون و ماقبل آن	۱۱
۲-۱-۲ بعد از پرسپترون	۱۲
۳-۱-۲ نسل سوم شبکه های عصبی مصنوعی	۱۳
۲-۲ انتظارات از شبکه های عصبی مصنوعی	۱۴
۱-۲-۲ قابلیت یادگیری	۱۴
۲-۲-۲ قابلیت تعمیم (Generalization)	۱۴
۳-۲-۲ پردازش موازی	۱۴

II

- ۱۴..... (Robustness) مقاوم بودن ۴-۲-۲
- ۱۵..... فواید و معایب شبکه های عصبی مصنوعی ۳-۲
- ۱۵..... شبکه های عصبی چند لایه ۴-۲
- ۱۶..... شبکه های بازگشتی ۵-۲
- ۱۷..... شبکه های پیش خور و آموزش ۶-۲
- ۱۸..... ۱-۶-۲ الگوریتم انتشار برگشتی
- ۱۹..... ۲-۶-۲ ساختار شبکه در الگوریتم انتشار برگشتی
- ۲۰..... ۷-۲ آموزش و یادگیری شبکه های عصبی مصنوعی
- ۲۰..... ۱-۷-۲ آموزش با نظارت
- ۲۱..... ۲-۷-۲ آموزش بدون نظارت
- ۲۱..... ۳-۷-۲ آموزش تقویت یافته
- ۲۲..... ۴-۷-۲ آموزش رقابتی
- ۲۲..... ۵-۷-۲ آموزش شبکه لونبرگ-مارکوآرت
- ۲۳..... ۸-۲ سلول عصبی بایاس در شبکه
- ۲۳..... ۹-۲ شبکه های عصبی مصنوعی پس از آموزش
- ۲۴..... ۱۰-۲ زوج های آموزشی و نحوه ارائه آنها به شبکه
- ۲۴..... ۱۱-۲ کاربردهای شبکه های عصبی مصنوعی
- ۲۶..... فصل سوم: معرفی روش کار
- ۲۷..... ۱-۳ انتخاب سری داده ها
- ۲۷..... ۲-۳ محاسبه توصیف کننده ها

III

۲۸.....	۳-۳ تجزیه و تحلیل و ارزیابی آماری توصیف کننده ها
۲۹.....	۴-۳ آنالیز مدل های آماری و انتخاب مدل مناسب
۲۹.....	۱-۴-۳ روش ورود اجباری
۲۹.....	۲-۴-۳ روش انتخاب جلوبرنده
۳۰.....	۳-۴-۳ روش حذف عقب برنده
۳۰.....	۴-۴-۳ روش رگرسیون چند مرحله ای یا گام به گام
۳۰.....	۵-۳ معیارهای مناسب بودن مدل
۳۰.....	۱-۵-۳ ضریب همبستگی
۳۱.....	۲-۵-۳ ضریب تعیین چندگانه
۳۱.....	۶-۳ ارزیابی مدل بدست آمده
۳۲.....	۷-۳ روش اعتبار بخشی متقابل
۳۳.....	فصل چهارم:
۳۳.....	مقدمه
۳۷.....	ارزیابی داده ها
۳۷.....	۱-۴ محاسبه توصیف کننده ها
۴۹.....	۲-۴ انتخاب توصیف کننده ها، مدل سازی و انتخاب بهترین مدل
۵۰.....	۱-۲-۴ شاخص مرکزی شاخه زنی (Lop)
۵۱.....	۲-۲-۴ توصیف کننده های MORSE
۵۱.....	۳-۲-۴ تابع توزیع شعاعی
۵۲.....	۴-۲-۴ توصیف کننده های GETAWAY

IV

۵۲.....	۵-۲-۴ توصیف کننده های WHIM
۵۳.....	۶-۲-۴ شاخص های خودهمبستگی دو بعدی
۵۳.....	۷-۲-۴ اجزاء اتم در مرکز.....
۵۵.....	۳-۴ ایجاد شبکه های عصبی (ANN)
۵۸.....	۴-۴ ارزیابی اعتبار مدل ها.....
۶۰.....	۵-۴ روش حداقل مربعات جزئی (PLS)
۶۲.....	۶-۴ تصادفی کردن پارامتر وابسته Y
۶۴.....	۷-۴ بحث و نتیجه گیری
۶۶.....	منابع و مأخذ

فصل اول

کمومتری

مقدمه

پیشرفت دانش دستگاهی در مطالعات شیمی و توانایی اتصال دستگاههای مختلف اندازه‌گیری به کامپیوتر، ما را با حجم زیادی از اطلاعات مواجه می‌سازد و به اصطلاح با انفجار داده‌ها^۱ مواجه هستیم. بنابراین برای شیمی‌دانان استفاده همزمان از روش‌های اندازه‌گیری در شیمی تجزیه، کامپیوتر و روش‌های آماری ضروری و مفید بنظر می‌رسد.

تحولات همزمان روش‌های اندازه‌گیری و علوم کامپیوتر باعث ایجاد نظم و قوانین جدیدی در شیمی گردید که به عنوان کمومتری^۲ مطرح می‌گردد. این عنوان اولین بار توسط ولد^۳ که دانشمندی سوئدی بود در سال ۱۹۷۰ میلادی مطرح گردید و در سال ۱۹۷۴ با همکاری یک شیمی‌دان آمریکایی بنام کووالسکی^۴ انجمن بین‌المللی کمومتری تاسیس گشت [۱].

تعاریف زیادی برای کمومتری ذکر شده که مهمترین آنها بصورت زیر است [۲]:

1 - Data Explosion
2 - Chemometrics
3- Svante Wold
4- Kowalski

« کمومتری یکسری قواعد شیمیایی است که با استفاده از علم ریاضی، آمار، رایانه و منطق صوری روش‌های تجربی بهینه‌ایی را طراحی یا انتخاب می‌کند که این روش‌ها را می‌توان جهت کسب بیشترین اطلاعات مفید از تجزیه و تحلیل سیستم‌های شیمیایی بکار برد.»
 به عبارت دیگر، کمومتری ابزاری است که به کمک آن شیمیدانان می‌توانند بطور موثری در راهی که از اندازه‌گیری به اطلاعات و در ادامه به دانستن ختم می‌شود حرکت کنند.

برخی از کاربردهای کمومتری:

- ۱- بهینه‌سازی ساختار ملکولی [۳]
- ۲- تخمین پارامترهای سینتیکی [۴]
- ۳- بهینه‌سازی پارامترهای کنترل کیفی [۴]
- ۴- بررسی رفتار کروماتوگرافی غیر خطی
- ۵- پردازش سیگنال [۵و۶]
- ۶- فعالیت دارو [۷]
- ۷- کالیبراسیون غیر خطی و مدلسازی [۷]

برخی از روش‌های آماری بکار رفته در کمومتری:

- ۱- رگرسیون خطی چند گانه^۱ (MLR)
- ۲- آنالیز مولفه‌های اصلی^۲ (PCA)
- ۳- حداقل مربعات جزئی^۳ (PLS)
- ۴- الگوریتم ژنتیک^۴ (GA)
- ۵- شبکه‌های عصبی مصنوعی^۵ (ANN)

کمومتری به عنوان یک شاخه علمی جوان در سه دهه اخیر به سرعت توسعه پیدا کرده است. این رشد سریع مدیون پیشرفت سریع دستگاه‌های هوشمند و خودکار آزمایشگاهی و هم‌چنین امکان استفاده از کامپیوترهای قدرتمند و نرم افزارهای ساده است. بنابراین کمومتری به عنوان یک وسیله در همه قسمت‌های شیمی و بیشتر در زمینه شیمی تجزیه مورد استفاده قرار گرفته است. امروزه یک

1- Multiple linear Regression
 2- Principle component Analysis
 3- Partial Least squares
 4- Genetic algorithm
 5- Artificial Neural Network

شیمیدان تجزیه به صورت فزاینده‌ایی به استفاده از روش‌های آماری و ریاضی در کارهای روزمره‌اش نیاز پیدا می‌کند.

یکی از زمینه‌های مهم کاربرد کمومتری در مطالعاتی است که خواص ملکول‌ها را به ویژگی‌های ساختاری آنها نسبت می‌دهد. از نظر شیمیدانان فعالیت و خواص یک ترکیب ناشی از ویژگی‌های ساختاری آن است.

وقتی خصوصیات ساختاری گونه‌ها و فعالیت آنها توسط اعداد و ارقام بیان شود می‌توان رابطه ریاضی یا کمی بین ساختار و فعالیت گونه‌ها ایجاد کرد که به اصطلاح به آن QSAR می‌گویند. این رابطه می‌تواند برای پیش‌بینی پاسخ بیولوژیکی یا شیمیایی دیگر ساختارها مورد استفاده قرار گیرد.

در واقع هدف QSAR، ایجاد رابطه‌ایی منطقی بین کمیت‌ها و یا خواص ترکیبات (فعالیت) و ساختار شیمیایی آنها می‌باشد و این قانون برای ملکول‌های جدید مورد استفاده قرار می‌گیرد [۸].

نتایج این مطالعات علاوه بر توجیه و تفسیر ارتباط بین خواص ملکول‌ها و خصوصیات ساختاری آنها به محققین در پیش‌بینی رفتار ملکول‌های جدید بر اساس رفتار ملکول‌های مشابه کمک می‌کند. به مجموعه ابزارها و روش‌هایی که به این منظور مورد استفاده قرار می‌گیرد روش‌های پارامتری گویند. در روش‌های پارامتری سعی می‌شود بین یک سری توصیف‌کننده‌های ملکولی با فعالیت یا خاصیت موردنظر ارتباط منطقی برقرار نمایند. توصیف‌کننده‌های ملکولی که به این منظور استفاده می‌شوند، مقادیر عددی می‌باشند که جنبه‌های مختلف ساختار ملکولی را به طور کمی نشان می‌دهند.

تاثیر استخلاف‌های مجزا (گروه‌ها) به فعالیت‌های بیولوژیکی ترکیبات مختلف بعنوان بنیان اساسی تمامی مدل‌های QSAR به شمار می‌رود. به همین علت فعالیت‌های بیولوژیکی باید بر حسب واحدهای مولی تعریف گردند.

مضافاً مطالعات کمی ساختار - فعالیت، بین فعالیت‌های مختلف بیولوژیکی و خصوصیات فیزیکی‌شیمیایی (آنالیز هانش^۱، روش اکسترانترمودینامیک^۲) و یا وجود یا عدم وجود خصوصیات ساختاری خاص، هماهنگی ایجاد می‌کنند (آنالیز فری - ویلسون^۳) [۹-۱۲]. مفهوم آنالیز هانش جهت ایجاد ارتباط بین فعالیت‌های مختلف بیولوژیکی مانند آزمایشگاهی و غیر آزمایشگاهی با یکدیگر بکار می‌رود [۱۲]. در آنالیز هانش، ضرایب توزیع یا انکسار مولی و یا خصوصیات چربی دوستی، الکترونی، قطبی و استری استخلاف‌ها با فعالیت‌های بیولوژیکی مرتبط می‌باشند [۹ و ۱۲]. با توجه به زمینه پارامتر اندازه‌گیری شده، روابط مذکور می‌تواند در رابطه با یک یا چندین خصوصیت فیزیکی شیمیایی

1- Hansch

2- Extrathermodynamic

3- Free-Wilson

بصورت خطی یا غیر خطی باشد. در آنالیز فری- ویلسون، متغیرهای شاخص که دارای مقادیر صفر یا ۱ می‌باشند، وجود یا عدم وجود خصوصیات ساختاری را کدگذاری می‌کنند [۱۱]. ترکیب پارامترهای هانش و فری - ویلسون در یک روش ترکیبی منجر به مدل‌هایی می‌گردد که قابلیت کاربرد هر دو روش را گسترش می‌دهند.

رگرسیون خطی یا غیر خطی چندگانه بعنوان یک ابزار آماری جهت استخراج مدل‌های بهتر، بررسی میزان اهمیت مدل‌های مذکور و اهمیت هر متغیر مستقل در معادله رگرسیون بکار می‌رود. سایر روش‌های آماری مانند آنالیز متمایز کننده، آنالیز ترکیب اصلی (PCA) یا آنالیز حداقل مربعات جزئی (PLS) بعنوان جانشینی برای آنالیز رگرسیون به شمار می‌روند [۱۱]. روش‌های جدیدتر تشابه ملکولها را ضمن در نظر گرفتن خصوصیات مختلف فیزیکوشیمیایی و سایر خصوصیاتشان، با فعالیت‌های بیولوژیکی آنها مورد مقایسه قرار می‌دهند.

روشهای بررسی کمی رابطه ساختار با فعالیت (QSAR) به منظور برقراری ارتباط بین فعالیت بیولوژیکی و خواص فیزیکوشیمیایی ترکیبات دارویی مورد استفاده قرار می‌گیرد. وابستگی لیگاند‌های آنزیمی به محل‌های پیوندشان از لحاظ نیروهای برهم‌کنشی مختلف بین ملکول‌های غیر ترکیبی قابل توجه می‌باشد. در برخی موارد حتی تفسیر مکانیسمی از برهم‌کنش یک لیگاند باهدف بیولوژیکی خود با استفاده از معادلات QSAR امکان‌پذیر می‌باشد. ارزیابی اعتبار متقاطع به معنی مقایسه سیستم‌های کاملاً مرتبط به هم بیولوژیکی، اعتبار مدل‌های بدست آمده را تقویت می‌کند. در فرآیند طراحی دارو و بهینه‌سازی نهایی، QSAR امکان یافتن و اصلاح فرضیه‌های کاری را بطور منطقی ممکن می‌سازد.

۱-۱ روش‌های پارامتری

همانطور که پیش‌تر مطرح شد کمومتری علم استفاده از ریاضی، آمار و کامپیوتر در شیمی می‌باشد و یکی از پرکاربردترین زمینه‌های کاربرد آن QSAR یعنی مطالعات کمی ارتباط ساختار - فعالیت می‌باشد. به مجموعه ابزارها و روش‌هایی که به این منظور مورد استفاده قرار می‌گیرند روش‌های پارامتری گویند. روشهای پارامتری مجموعه تکنیک‌هایی است که برای مدل‌سازی و پیش‌بینی فعالیت‌های بیولوژیکی مورد استفاده قرار می‌گیرند. با استفاده از این روش‌ها می‌توان ارتباطی منطقی بین فعالیت یا خاصیت موردنظر با تعدادی از توصیف‌کننده‌های ملکولی برقرار نمود. توصیف‌کننده‌های ملکولی متغیرهای مستقلی هستند که ویژگی‌های ملکول را بطور کمی بیان می‌کنند. معمولاً از روش رگرسیون خطی چند گانه (MLR) جهت ایجاد این روابط استفاده می‌شود. در رگرسیون خطی چندگانه، ارتباط بین توصیف‌کننده‌ها بعنوان متغیر مستقل و فعالیت بیولوژیکی، به صورت یک متغیر وابسته در قالب یک معادله ریاضی طرح‌ریزی می‌شود. قابل ذکر است که متغیرهای مستقل بر اساس

اصول علمی که انتظار می‌رود با متغیر وابسته در ارتباط باشند، انتخاب می‌شوند. روش‌های آماری در مورد ارتباط بین متغیرهای وابسته و مستقل کاربرد وسیعی دارند و نرم‌افزارهای آماری متفاوتی جهت تسریع و سهولت انجام عملیات محاسباتی مورد استفاده قرار می‌گیرند.

۱-۱-۱ نرم‌افزارهای مورد استفاده در روش‌های پارامتری

در این بخش مختصراً به شرح بسته نرم‌افزارهایی که در این پروژه مورد استفاده قرار گرفته‌اند می‌پردازیم:

۱-۱-۱-۱ بسته نرم‌افزاری Hyperchem

با این نرم‌افزار ابتدا شکل ملکول بطور تقریبی بر روی صفحه نمایشگر کامپیوتر رسم می‌گردد. سپس توسط روش‌های کوانتوم - مکانیک، ساختار ملکول بهینه می‌شود. با استفاده از این نرم‌افزار می‌توان اطلاعات فراوانی مانند زوایای پیوندی، طول پیوندها، زوایای پیچش و... را به دست آورد. این بسته نرم‌افزاری، قابلیت ترسیم ساختار کلاسترها و ترکیبات پیچیده را دارد و در کتابخانه آن ساختار اسیدهای آمینه و نوکلئیک اسیدها موجود است.

۱-۱-۱-۲ بسته نرم‌افزاری MOPAC

MOPAC یک نرم‌افزار اوربیتال ملکولی نیمه تجربی برای مطالعه ساختارهای ملکولی و واکنش‌های شیمیایی است. این نرم‌افزار با استفاده از خروجی Hyperchem که شامل زوایای پیچشی و زوایای پیوندی و طول پیوندها است محاسبات را انجام می‌دهد. از نتایج حاصل از چنین محاسباتی می‌توان اطلاعاتی از قبیل گرمای تشکیل، پتانسیل یونیزاسیون، انرژی الکترونی، انرژی دافعه هسته - هسته، گشتاور دو قطبی و بسیاری از فاکتورهای دیگر را در مورد ملکول‌ها بدست آورد. با وجود اینکه این برنامه بر اساس مفاهیم تئوری کوانتومی و مفاهیم ترمودینامیک عمل کرده از محاسبات پیشرفته ریاضی هم کمک می‌گیرد. کاربران نیازی به دانستن مفاهیم عمیق فوق ندارند. داده‌های ورودی و خروجی این برنامه تا حد ممکن ساده هستند تا جایی که می‌توان به راحتی توجه خود را به سایر جنبه‌های شیمیایی مسئله معطوف کرد.

۱-۱-۱-۳ بسته نرم‌افزاری SPSS

این بسته نرم‌افزاری، توانایی تحلیل کلی اطلاعات را داراست و عمومی‌ترین نرم‌افزار آماری است. امروزه آمار، مجموعه‌ای از مفاهیم و روش‌هایی است که در هر زمینه پژوهش، برای گردآوری اطلاعات مربوط به آن و انجام نتیجه‌گیری‌ها در شرایطی که عدم قطعیت وجود دارد به کار می‌رود. SPSS دارای توانایی بسیار بالایی است.

در این پژوهش از این نرم‌افزار برای بدست آوردن بهترین مدل، حذف متغیرهای با همبستگی بالا و ارزیابی و اعتبار مدل استفاده شده است.

۴-۱-۱-۱ بسته نرم‌افزاری Dragon

دراگون نرم‌افزاری است که با استفاده از یک فایل ورودی، تعداد زیادی از توصیف کننده‌های ملکولی را که محاسبه آنها مشکل است، محاسبه می‌کند. فایل ورودی نرم‌افزار دراگون می‌تواند فایل (*.hin) مربوط به نرم‌افزار Hyperchem باشد. این نرم‌افزار توسط گروه کمومتریکس دانشگاه میلانو طراحی شده است و قابلیت محاسبه ۱۴۹۷ توصیف کننده برای هر ملکول را دارا است. مقاله‌های فراوانی در مورد استفاده از روشهای پارامتری در مطالعات QSAR تاکنون منتشر شده‌اند که به برخی از آنها در زیر اشاره می‌گردد:

- پیش‌بینی اندیس بازدارنده کوواتس^۱ در دسته بزرگی از ترپن‌ها توسط همتی‌نژاد و همکارانش [۱۳].

- پیش‌بینی زمان بازدارنده باقی مانده آفت‌کش‌ها در بافت‌های حیوانی توسط کنوز^۲ و همکارانش [۱۴].

- پیش‌بینی ضرایب نسبت خون/هوا در ترکیبات آلی فرآر توسط گل محمدی [۱۵].

- پیش‌بینی فعالیت بازدارنده‌های فورین توسط وراچارت چیوان^۳ و همکارانش [۱۶].

- پیش‌بینی فعالیت بازدارندگی هپاراناز^۴ توسط جلالی هروی و همکارانش [۱۷].

- پیش‌بینی بازدارندگی مشتقات بنزیمیدازول وابسته به فلز توسط گرکانی‌نژاد و همکارانش [۱۸].

- مدل‌سازی تحرک الکتروفوریتی^۵، زمان بازدارنده و فاکتور پاسخ نسبی ترکیبات مختلف توسط گرکانی‌نژاد و همکارانش [۱۹-۲۰].

در پیشرفت‌های سه دهه اخیر از شبکه‌های عصبی مصنوعی، منطق فازی و الگوریتم ژنتیکی در کمومتری استفاده شده است. این روشها جزء دسته‌ای از علوم کامپیوتر به نام هوش مصنوعی، AI^۶، قرار می‌گیرند که تفاوت اساسی با سایر روش‌های محاسباتی و هم‌چنین با یکدیگر دارند. در سیستم‌های ریاضی متداول، دانسته‌ها بصورت قواعد روشن و از پیش تعیین شده‌ای در سیستم وجود

1- Kovats

2- E.Konoz

3- Apilak Warachart cheewan

4 - Heparanase

5 - Electrophoretic mobility

6- Artificial Intelligent

دارد، در حالی که این سیستم‌ها قواعد را توسط آموزش از مثال‌هایی که به آنها داده می‌شوند فرامی‌گیرند.

از میان روش‌های فوق، استفاده از شبکه‌های عصبی مصنوعی در شیمی، نقش برجسته‌ای پیدا کرده و کاربرد آن در شیمی به سرعت رو به افزایش است.

از میان شاخه‌های مختلف شیمی، شیمی تجزیه یکی از بخش‌هایی است که بیشترین استفاده را از شبکه‌های عصبی مصنوعی (ANN) کرده است. جلالی هروی و همکارانش استفاده از ANN را در شبیه‌سازی و جستجوی طیف‌های جرمی گزارش کرده‌اند [۲۱]. پیش‌بینی میزان جابجایی $^1\text{C NMR}$ [۲۲]، شناسایی الگوها از داده‌های کروماتوگرافی [۲۳]، پیش‌بینی ضریب پاسخ آشکارسازهای شعله - یونشی [۲۰] و پیش‌بینی ضرایب انتخابگری الکترودهای یون گزین [۲۲] از دیگر موارد گزارش شده از کاربردهای ANN است.

با استفاده از شبکه‌های عصبی مصنوعی می‌توان خواص و فعالیت مولکول‌ها را پیش‌بینی یا اینکه از آن در آنالیز مخلوط‌های پیچیده استفاده نمود [۲۴]، نوع ترکیب و زیر گروه‌های آن را از روی طیف آن شناسایی کرد [۲۴] و توالی پروتئین‌ها را در DNA^2 مشخص نمود [۲۵]. از شبکه‌های عصبی مصنوعی به خوبی می‌توان در شناخت الگوها و طبقه‌بندی داده‌ها استفاده نمود [۲۶]. همچنین از این شبکه‌ها در مطالعات QSAR، هنگامی که ارتباط بین توصیف‌کننده‌ها و کمیت مورد مدل‌سازی پیچیده و غیر خطی بوده و یا اینکه پردازش سریع حجم وسیعی از اطلاعات مورد نیاز باشد می‌توان به خوبی استفاده کرد.

آنالیز کمی داده‌های چند متغیره [۲۷] از دیگر کاربردهای شبکه‌های عصبی مصنوعی می‌باشد. پیش‌بینی آنتالپی آلکان‌ها [۲۲]، پیش‌بینی ثابت دی‌الکتریک [۲۲]، پیش‌بینی تحرک الکتروفورتیک سولفونامیدها در الکتروفورز موئینه [۱۹] و پیش‌بینی جداسازی ترکیبات کایرال توسط جولین - ارتیز^۳ و همکاران [۲۸] از جمله مقالات منتشر شده راجع به استفاده از شبکه‌های عصبی مصنوعی در مطالعات QSAR می‌باشد.

یکی از مهم‌ترین ویژگی‌های شبکه عصبی مصنوعی در این است که شبکه‌های فوق، سیستمی برنامه‌ریزی شده با قواعد از پیش تعیین شده نمی‌باشد و می‌توانند طی فرآیندی که آموزش نامیده می‌شود ساختار درونی خود را به صورت تجربی به نحوی تنظیم کنند که بهترین پاسخ ممکن را برای داده‌های ورودی ایجاد کنند. استفاده از شبکه‌های عصبی مصنوعی به جای روش رگرسیون خطی چند

1- Nuclear Magnetic Resonance
2- Deoxyribonucleic Acid
3- Julin-Ortiz

تایی در مطالعات QSAR، به خصوص هنگامی که ارتباط بین توصیف کننده و فعالیت موردنظر خطی نبوده و یا اینکه بین آنها برهم کنشی‌هایی وجود داشته باشد باعث بهبود مدل حاصله خواهد شد که این پروژه این مطلب را به وضوح نشان داده است و در آن سعی شده تا با استفاده از سه روش شبکه‌های عصبی مصنوعی و رگرسیون خطی چندگانه و روش حداقل مربعات جزئی فعالیت یکسری از مشتقات فروکسان - آلکیل نیترات مدل‌سازی شود. در فصل‌های بعدی روش‌های QSAR و تکنیک شبکه‌های عصبی مصنوعی بطور کامل بحث شده است.

فصل دوم

شبکه های عصبی مصنوعی

مقدمه

پایه‌سازی ویژگی‌های شگفت‌انگیز مغز در یک سیستم مصنوعی (سیستم دینامیکی ساخته بشر) همیشه وسوسه‌انگیز و مطلوب بوده است. محققینی که طی سالها در این زمینه فعالیت کرده‌اند بسیاری، لیکن نتیجه این تلاشها صرفنظر از یافته‌های ارزشمند، باور هرچه بیشتر این اصل بوده است که مغز بشر دست نیافتنی است. با تاکید به این نکته که گذشته از متافیزیک، دور از دسترس بودن ایده‌ال هوش طبیعی را می‌توان با عدم کفایت دانش موجود بشر از فیزیولوژی عصبی پذیرفت. باید اذعان داشت که عالی بودن هدف و کافی نبودن دانش موجود، خود سبب انگیزش پژوهش‌های بیشتر و بیشتر در این زمینه بوده و خواهد بود. هم‌چنان که امروز شاهد بروز چنین فعالیت‌هایی در قالب شبکه‌های عصبی مصنوعی هستیم.

طی چند دهه اخیر، تلاش‌های بسیار جدی جهت طراحی مدارات الکترونیکی که قادر باشند شبکه‌های عصبی زیستی را همانندسازی کنند صورت گرفته است. شبکه‌های مدل شده که با نام الگوهای شبکه‌های عصبی شناخته شده‌اند گسترش یافته و مدل‌سازی شده‌اند. برخی از این نمونه‌ها به گونه‌ایی بسیار نزدیک عملکرد شبکه عصبی زیستی را همانندسازی کرده‌اند و برخی دیگر تفاوت بسیاری دارند.

طی صد سال اخیر تلاش‌های بسیار جدی برای مدل کردن یک نرون طبیعی صورت گرفته و پیشرفت قابل ملاحظه‌ای در این راستا انجام شده است. برای مدل سازی یک شبکه عصبی مصنوعی^۱ (ANN) می‌توان از یک مدل ریاضی که خصوصیات یک سیستم بیولوژیک را توصیف کند استفاده کرد. سپس با استفاده از یک رایانه می‌توان این مدل را به سرعت و به نحو مناسبی شبیه‌سازی کرد. جهت ساده کردن و بهبود عملکرد این مدل تغییرات لازم صورت می‌گیرد.

شبکه‌های عصبی مصنوعی در واقع مثلثی هستند که سه ضلع مفهومی دارد:

۱- سیستم تجزیه و تحلیل داده‌ها

۲- نرون یا سلول عصبی

۳- شبکه یا قانون کار گروهی نرونها

در یک تعریف کلاسیک، هایکین^۲ می‌گوید: شبکه عصبی عبارت است از مجموعه‌ای عظیم از پردازشگرهای موازی که استعداد ذاتی برای ذخیره اطلاعات تجربی و بکارگیری آن دارند و یک شبکه حداقل در دو مورد شبیه مغز است:

۱- مرحله‌ای موسوم به یادگیری دارد.

۲- وزنه‌های سیناپسی جهت ذخیره دانش بکار می‌روند.

بنابه تعریفی دیگر شبکه‌های عصبی مصنوعی (ANN)، ساختاری (شبکه‌ای) است که متشکل از تعدادی واحد (نرون‌های مصنوعی) که در داخل شبکه به هم وصل شده‌اند، است. هر واحد دارای یک مشخصه ورودی / خروجی (I/O) می‌باشد و محاسبه یا عملی جزئی را اجرا می‌کند. خروجی هر واحد، با توجه به مشخصه (I/O) آن، اتصالات درونیش به سایر واحدها و احتمالاً ورودی‌های خارجی، تعیین می‌گردد. از آنجا که آموزش دستی شبکه امکان‌پذیر است، از این رو شبکه معمولاً کارکردی کلی از یک حالت یا حالت‌های بیشتری از آموزش را به دست می‌آورد. شبکه عصبی مصنوعی (ANN) متشکل از یک شبکه نیست، بلکه خانواده‌هایی متشکل از شبکه‌های گوناگون می‌باشد. عمل یا عملکرد کلی شبکه‌های عصبی مصنوعی توسط توپولوژی شبکه، خصوصیات نرون منفرد^۳، روند یادگیری^۴ و داده‌های آموزش معین می‌شود. به منظور کاربردی شدن، یک شبکه عصبی مصنوعی (ANN) می‌بایستی ابزارهایی برای ارتباط با دنیای خارج داشته باشد [۲۹ و ۳۰].

1- Artificial Neural Networks (ANN)

2- Haykin

3- Individual neuron

4- Learning