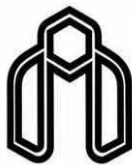


بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه صنعتی شاهرود

دانشکده شیمی

گروه شیمی فیزیک

عنوان :

پیش‌بینی ضریب تراکم‌پذیری همدمای ترکیبات آلی بر اساس توصیف‌کننده‌های

ساختاری مولکولی در محدوده‌ی وسیعی از دما و فشار

دانشجو:

زهرا کاظم نادی ورنوسفادرائی

استاد راهنما:

دکتر زهرا کلانتر

استاد مشاور:

دکتر زینب موسوی

پایان‌نامه جهت اخذ درجه کارشناسی ارشد

بہمن ماہ ۱۳۹۱

اگر تنها ترین تنها شوم، باز خدا هست

او جان شین همه نداشته های من است!

به پاس تعبیر عظیم و انسانیشان از کلمه ی ایثار و از خودگذشتگی

به پاس عاطفه ی سرشار و گرمای امید بخش وجودشان که در این سردترین روزگار ان، بهترین پشتیبان است

به پاس قلب های بزرگشان که فریادس است و سرگردانی و ترس در پناهمشان به شجاعت می گراید

و به پاس محبت های بی دریغشان که هرگز فروکش نمی کند

این مجموعه را به پدر و مادر عزیزم و دو خواهر مهربانم تقدیم می کنم.

## شکر و قدردانی

با نام آنکه خلق کرد، هستی را از رحمت بی انتهای خویش تا انسان در قدم خلقتش به پهنای گسترده ای به بزرگی جهان خیمه  
زند و به او عقل داد تا عظمت خلقت را در یابد. ستایش او را که تجلی وجودش در دو کوهر کرانه زندگی ام، برایم بزرگترین  
نعمت بود، پدر و مادری که هزاران بار چهره های مهربان و دستان کرمان رامی بوسم، آن ها که همواره روشنی بخش راه های  
پرپیچ و خم زندگی ام بودند.

با تمام وجود و در اوج فروتنی از مقام شیخ استاد کرانه سرکار خانم دکتر زهرا کلاستر که در نهایت صبر و بزرگواری تمامی سعی و  
تلاش خود را در جهت اعتلای واقعی ارزش های آموزشی در کالبد هدایت ماور، نمود و نسبت به این جانب مبذول فرمودند،  
کمال قدردانی را می نمایم و با اتقان بیکران از مساعدت های بی دریغ جناب آقای دکتر بهمتی نژاد و سرکار خانم دکتر  
موسوی و سرکار خانم الیاسی و با شکر خالصانه از آقای مصطفی آقا محمدی و آقای احسان عادی و همه کسانی که به نوعی مراد  
به انجام رساندن این مهم یاری نموده اند و در برابر وجود گرامی شان زانوی ادب بر زمین می نهم.

## چکیده

در این تحقیق، مطالعات ارتباط کمی ساختار - خاصیت با استفاده از شبکه‌ی عصبی مصنوعی برای پیش‌بینی ضریب تراکم‌پذیری همدمای ترکیبات آلی مختلف به کار گرفته شده است. سری داده‌ها شامل ضریب تراکم‌پذیری همدمای ۵۶ ترکیب در دما و فشارهای مختلف است که مجموعاً ۴۲۹۷ نقطه داده را تشکیل می‌دهد. برای انتخاب توصیف‌کننده‌های بهینه از دو روش رگرسیون مرحله‌ای (SR) و راهبرد CDFS استفاده شد. سپس مدل‌سازی به وسیله‌ی شبکه‌ی عصبی مصنوعی انجام گرفت و پس از بهینه‌سازی و آموزش شبکه، مدل حاصل برای پیش‌بینی ضریب تراکم‌پذیری همدمای ۷ ترکیب سری تست که در هیچ یک از مراحل مدل‌سازی دخالت نداشته‌اند، به کار گرفته شد. مقایسه‌ی دو مدل حاصل از رگرسیون مرحله‌ای و راهبرد CDFS نشان داد که مدل به دست آمده از CDFS تعداد توصیف‌کننده‌ی کمتر و قدرت پیش‌بینی بالاتری دارد. میانگین مربعات خطای سری تست توسط مدل‌های SR-ANN و CDFS-ANN به ترتیب ۰/۴۴۵۶ و ۰/۳۰۱۲ به دست آمد.

در مرحله‌ی بعد، شبکه‌ی عصبی مصنوعی به همراه روش سهم گروه (GCM) برای پیش‌بینی ضریب تراکم‌پذیری همدمای ترکیبات سری تست به کار گرفته شد. در این مدل‌سازی تعداد گروه‌های تشکیل‌دهنده‌ی هر ترکیب و دو متغیر دما و فشار به عنوان متغیرهای ورودی شبکه‌ی عصبی در نظر گرفته شدند. میانگین مربعات خطای حاصل از این روش ۰/۲۳۰۳ به دست آمد که نشان‌دهنده‌ی برتری روش GCM-ANN نسبت به دو روش SR-ANN و CDFS-ANN می‌باشد.

**کلمات کلیدی:** ارتباط کمی ساختار-ویژگی، شبکه‌ی عصبی مصنوعی، ضریب تراکم‌پذیری همدما،

رگرسیون مرحله‌ای (SR)، راهبرد CDFS، روش سهم گروه (GCM)

نتایج حاصل از این پایان نامه در دو پوستر با عناوین:

**“Nonlinear quantitative structure–property relationship studies of isothermal compressibility of ketons and esters over a wide range of temperatures and pressures”**

❖ در نهمین همایش ملی شیمی دانشگاه پیام نور

**“QSPR study of isothermal compressibility for a diverse set of organic compounds using GA-ANN”**

❖ در پنجمین کنفرانس و کارگاه ریاضی- شیمی ارائه گردید.

## فهرست مطالب

### فصل اول

- ۱-۱- ضریب تراکم‌پذیری همدما و اهمیت آن ..... ۲
- ۲-۱- روش‌های اندازه‌گیری ضریب تراکم‌پذیری همدما ..... ۲
- ۱-۲-۱- اندازه‌گیری  $\kappa_T$  با استفاده از اندازه‌گیری سرعت صوت در دما و فشارهای مختلف ..... ۳
- ۲-۲-۱- اندازه‌گیری  $\kappa_T$  با استفاده از اندازه‌گیری دانسیته در دماها و فشارهای مختلف ..... ۵
- ۳-۱- مروری بر روش‌های پیش‌بینی  $\kappa_T$  ..... ۶
- ۱-۳-۱- پیش‌بینی  $\kappa_T$  با استفاده از معادله حالت ..... ۶
- ۲-۳-۱- پیش‌بینی  $\kappa_T$  با استفاده از شبکه‌ی عصبی مصنوعی ..... ۸

### فصل دوم

- ۱-۲- کمومتریکس ..... ۱۰
- ۲-۲- رابطه‌ی ساختار- خاصیت یا ساختار- فعالیت ..... ۱۱
- ۱-۲-۲- فراهم کردن سری داده‌ها ..... ۱۱
- ۲-۲-۲- رسم ساختار و بهینه‌سازی ..... ۱۲
- ۳-۲-۲- محاسبه‌ی توصیف‌کننده‌ها ..... ۱۲
- ۴-۲-۲- حذف توصیف‌کننده‌های نامناسب ..... ۱۲
- ۵-۲-۲- دسته‌بندی داده‌ها ..... ۱۳
- ۱-۵-۲-۲- دسته‌بندی داده‌ها به روش تحلیل مولفه‌های اساسی (PCA) ..... ۱۳
- ۶-۲-۲- انتخاب بهترین توصیف‌کننده‌ها ..... ۱۴
- ۱-۶-۲-۲- انتخاب بهترین توصیف‌کننده‌ها توسط رگرسیون مرحله‌ای ..... ۱۵
- ۲-۶-۲-۲- انتخاب بهترین توصیف‌کننده‌ها توسط راهبرد CDFS ..... ۱۵
- ۷-۲-۲- ساختن مدل ..... ۱۷
- ۱-۷-۲-۲- رگرسیون خطی چندگانه (MLR) ..... ۱۸
- ۲-۷-۲-۲- شبکه‌ی عصبی مصنوعی ..... ۱۸
- ۲-۷-۲-۲- الف- ساختار شبکه‌های عصبی ..... ۲۰



- ۲۳ ..... ۲-۲-۷-۲-ب- انواع شبکه‌ها از نظر ارتباطات بین نرونی
- ۲۴ ..... ۲-۲-۷-۲-ج- آموزش شبکه عصبی
- ۲۵ ..... ۲-۲-۷-۲-د- معرفی روش‌های مختلف بهینه‌سازی عملکرد شبکه عصبی
- ۲۶ ..... ۲-۲-۸- ارزیابی قدرت پیش‌بینی مدل
- ۲۷ ..... ۲-۳-۳- انتخاب توصیف‌کننده‌ها بر مبنای روش سهم گروه
- ۲۷ ..... ۲-۳-۱- روش سهم گروه
- ۲۸ ..... ۲-۳-۲- شبکه‌های عصبی مصنوعی بر مبنای روش سهم گروه‌ها

### فصل سوم

- ۳۱ ..... ۳-۱-۱- مدل‌سازی ضریب تراکم‌پذیری همدمای سیالات آلی مختلف
- ۳۱ ..... ۳-۱-۱- انتخاب سری داده‌ها
- ۳۱ ..... ۳-۱-۲- رسم و بهینه‌سازی ساختار مولکول‌ها
- ۳۴ ..... ۳-۱-۳- محاسبه‌ی توصیف‌کننده‌ها
- ۳۴ ..... ۳-۱-۴- حذف توصیف‌کننده‌های نامناسب
- ۳۴ ..... ۳-۱-۵- دسته‌بندی داده‌ها
- ۳۵ ..... ۳-۱-۶- انتخاب بهترین توصیف‌کننده‌ها
- ۳۶ ..... ۳-۱-۶-۱- انتخاب بهترین توصیف‌کننده‌ها به روش رگرسیون مرحله‌ای
- ۴۲ ..... ۳-۱-۶-۲- انتخاب بهترین توصیف‌کننده‌ها با استفاده از راهبرد CDFS
- ۴۴ ..... ۳-۱-۷- مدل‌سازی به روش شبکه‌ی عصبی مصنوعی
- ۴۵ ..... ۳-۱-۷-۱- بهینه‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده‌های رگرسیون مرحله‌ای
- ۵۳ ..... ۳-۱-۷-۲- بهینه‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده‌های انتخاب شده توسط راهبرد CDFS
- ۵۹ ..... ۳-۱-۷-۳- مدل‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده‌های حاصل از روش سهم گروه
- ۷۰ ..... ۳-۱-۸- ارزیابی مدل‌های غیر خطی
- ۷۰ ..... ۳-۱-۸-۱- ارزیابی با استفاده از اعتبارسنجی بیرونی
- ۷۸ ..... ۳-۱-۸-۲- ارزیابی با استفاده از آزمون  $\chi^2$  تصادفی
- ۸۰ ..... ۳-۱-۹- بررسی ارتباط توصیف‌کننده‌های وارد شده در مدل با ضریب تراکم‌پذیری همدمای ترکیبات
- ۸۰ ..... ۳-۱-۹-۱- توصیف‌کننده‌ی X0Av
- ۸۱ ..... ۳-۱-۹-۲- توصیف‌کننده‌های 3D-MoRSE

۱۰-۱-۳	بررسی میزان مشارکت توصیف‌کننده‌های منتخب شبکه عصبی	۸۳
۱۱-۱-۳	نتیجه‌گیری	۸۶
۱۲-۱-۳	آینده‌نگری	۸۹
۹۰	پیوست	
۱۱۰	فهرست منابع	

## فهرست جداول

- جدول (۱-۳) - محدوده‌ی دما و فشار و تعداد نقطه داده‌ی استفاده شده در مدل‌سازی برای سیالات آلی مختلف به همراه مرجع مورد استفاده برای هر یک از آنها..... ۳۲
- جدول (۲-۳) - مقایسه‌ی آماره‌های سری ارزیابی مدل‌های به دست آمده از روش رگرسیون مرحله‌ای..... ۳۷
- جدول (۳-۳) - توصیف‌کننده‌های انتخاب شده توسط روش رگرسیون مرحله‌ای..... ۳۸
- جدول (۴-۳) - ماتریس همبستگی توصیف‌کننده‌های انتخاب شده توسط رگرسیون مرحله‌ای..... ۳۸
- جدول (۵-۳) - مقادیر توصیف‌کننده‌های انتخاب شده به روش رگرسیون مرحله‌ای..... ۳۹
- جدول (۶-۳) - توصیف‌کننده‌های ۱۶ مدل QSPR به دست آمده از دسته‌بندی‌های مختلف داده‌ها به همراه پارامترهای آماری به دست آمده برای هر دسته در روش CDFS..... ۴۳
- جدول (۷-۳) - توصیف‌کننده‌های انتخاب شده توسط روش CDFS..... ۴۴
- جدول (۸-۳) - توابع و پارامترهای بهینه‌ی شبکه‌های طراحی شده با توصیف‌کننده‌های حاصل از روش رگرسیون مرحله‌ای..... ۵۱
- جدول (۹-۳) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش رگرسیون مرحله‌ای..... ۵۲
- جدول (۱۰-۳) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS..... ۵۴
- جدول (۱۱-۳) - مشابه جدول (۱۰-۳) حول نقطه‌ی بهینه با گام ۲..... ۵۴
- جدول (۱۲-۳) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونبرگ-مارکوارت و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS..... ۵۵
- جدول (۱۳-۳) - مشابه جدول (۱۲-۳) حول نقطه‌ی بهینه با گام ۲..... ۵۵
- جدول (۱۴-۳) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS..... ۵۶
- جدول (۱۵-۳) - مشابه جدول (۱۴-۳) حول نقطه‌ی بهینه با گام ۲..... ۵۶

جدول (۳-۱۶) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونبرگ- مارکوارت و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS ..... ۵۷

جدول (۳-۱۷) - مشابه جدول (۳-۱۶) حول نقطه‌ی بهینه با گام ۲ ..... ۵۷

جدول (۳-۱۸) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش CDFS ..... ۵۹

جدول (۳-۱۹) - گروه‌های عاملی موجود در مولکول‌های مورد بررسی به عنوان توصیف‌کننده‌های روش سهم گروه ..... ۶۰

جدول (۳-۲۰) - تعداد گروه‌های عاملی موجود در مولکول‌های مختلف سری داده‌ها به عنوان توصیف‌کننده‌های روش سهم گروه ..... ۶۱

جدول (۳-۲۱) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه ..... ۶۵

جدول (۳-۲۲) - مشابه جدول (۳-۲۱) حول نقطه‌ی بهینه با گام ۲ ..... ۶۵

جدول (۳-۲۳) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونبرگ- مارکوارت و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه ..... ۶۶

جدول (۳-۲۴) - مشابه جدول (۳-۲۳) حول نقطه‌ی بهینه با گام ۲ ..... ۶۶

جدول (۳-۲۵) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه ..... ۶۷

جدول (۳-۲۶) - مشابه جدول (۳-۲۵) حول نقطه‌ی بهینه با گام ۲ ..... ۶۷

جدول (۳-۲۷) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونبرگ- مارکوارت و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه ..... ۶۸

جدول (۳-۲۸) - مشابه جدول (۳-۲۷) حول نقطه‌ی بهینه با گام ۲ ..... ۶۸

جدول (۳-۲۹) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش سهم گروه ..... ۷۰

جدول (۳-۳۰) - پارامترهای آماری مربوط به هر یک از شبکه‌های بهینه شده برای سری ارزیابی و سری تست ..... ۷۱

جدول (۳-۳۱) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $y$  تصادفی در مدل SR-ANN . ۷۸

جدول (۳-۳۲) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $y$  تصادفی در مدل CDFS-ANN

۷۹ .....

جدول (۳-۳۳) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $y$  تصادفی در مدل GCM-ANN

۷۹ .....

## فهرست شکل ها

- شکل (۱-۲) - ساختار یک نرون طبیعی..... ۲۰
- شکل (۲-۲) - ساختمان یک نرون محاسباتی..... ۲۱
- شکل (۳-۲) - ساختار یک شبکه‌ی سه لایه..... ۲۲
- شکل (۱-۳) - پراکندگی مولکول‌ها بر روی نمودار PCA..... ۳۵
- شکل (۲-۳) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین..... ۴۷
- شکل (۳-۳) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لونیبرگ-مارکوارت..... ۴۸
- شکل (۴-۳) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین..... ۴۹
- شکل (۵-۳) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونیبرگ-مارکوارت..... ۵۰
- شکل (۶-۳) - نمودار میانگین مربع خطای حاصل از سری ارزیابی بر حسب پارامتر  $\mu$ ..... ۵۱
- شکل (۷-۳) - ساختار شبکه عصبی بهینه شده..... ۵۲
- شکل (۸-۳) - نمودار میانگین مربعات خطای حاصل از سری ارزیابی بر حسب تغییرات پارامتر  $\mu$ ..... ۵۸
- شکل (۹-۳) - نمودار میانگین مربعات خطای حاصل از سری ارزیابی بر حسب تغییرات پارامتر  $\mu$ ..... ۶۹
- شکل (۱۰-۳) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدمما بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل SR-ANN..... ۷۲
- شکل (۱۱-۳) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدمما بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل CDFS-ANN..... ۷۳

- شکل (۳-۱۲) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدم‌ها بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل GCM-ANN ..... ۷۴
- شکل (۳-۱۳) - نمودار باقیمانده‌ها بر حسب مقادیر تجربی ضریب تراکم‌پذیری همدم‌ها برای سری ارزیابی و سری تست. در مدل SR-ANN ..... ۷۵
- شکل (۳-۱۴) - نمودار باقیمانده‌ها بر حسب مقادیر تجربی ضریب تراکم‌پذیری همدم‌ها برای سری ارزیابی و سری تست. در مدل CDFS-ANN ..... ۷۶
- شکل (۳-۱۵) - نمودار باقیمانده‌ها بر حسب مقادیر تجربی ضریب تراکم‌پذیری همدم‌ها برای سری ارزیابی و سری تست. در مدل GCM-ANN ..... ۷۷
- شکل (۳-۱۶) - درصد مشارکت توصیف‌کننده‌ها در مدل SR-ANN ..... ۸۴
- شکل (۳-۱۷) - درصد مشارکت توصیف‌کننده‌ها در مدل CDFS-ANN ..... ۸۴
- شکل (۳-۱۸) - درصد مشارکت توصیف‌کننده‌ها در مدل GCM-ANN ..... ۸۵

# فصل اول

## مقدمه



## ۱-۱- ضریب تراکم‌پذیری همدما<sup>۱</sup> و اهمیت آن

به طور کلی خواص حجمی ترکیبات شیمیایی و وابستگی آن‌ها به دما و فشار از اهمیت خاصی در صنایع شیمیایی برخوردار است. یکی از مهم‌ترین این خواص که در طراحی دستگاه‌های فشار بالا و محاسبات مخازن ضروری است ضریب تراکم‌پذیری همدما،  $\kappa_T$ ، می‌باشد. بنابراین پیش‌بینی دقیق‌تر این کمیت، صحت بیشتر طراحی‌های فشار بالا را در صنایع شیمیایی وابسته به دنبال خواهد داشت. ضریب تراکم‌پذیری همدما که طبق رابطه‌ی (۱-۱) بیان می‌شود، میزان تغییرات حجم مواد با تغییرات فشار را در دمای ثابت مشخص می‌کند.

$$\kappa_T = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial P} \right)_T \quad (1-1)$$

در این رابطه  $\rho$  دانسیته و  $P$  فشار می‌باشد. دیمانسیون ضریب تراکم‌پذیری همدما عکس فشار است و می‌توان آن را بر حسب یکاهای ( $\text{Pa}^{-1}$ ) یا ( $\text{bar}^{-1}$ ) بیان کرد. تغییرات ضریب تراکم‌پذیری همدمای مواد جامد و مایع در برابر دما و فشار، در مقایسه با گازها بسیار کم است.

## ۱-۲- روش‌های اندازه‌گیری ضریب تراکم‌پذیری همدما

برای دستیابی به داده‌های تجربی ضریب تراکم‌پذیری همدمای ترکیبات مختلف بر حسب دما و فشار دو مسیر کلی وجود دارد:

(۱) اندازه‌گیری سرعت صوت در دما و فشارهای مختلف

(۲) اندازه‌گیری دانسیته در دما و فشارهای مختلف

که در ادامه راجع به هر یک از این روش‌ها توضیحاتی ارائه می‌شود.

---

1. Isothermal Compressibility

### ۱-۲-۱-اندازه‌گیری $\kappa_T$ با استفاده از اندازه‌گیری سرعت صوت در دما و فشارهای مختلف

در این روش برای اندازه‌گیری ضریب تراکم‌پذیری همدمای سیالات، ابتدا با اندازه‌گیری سرعت صوت،  $u$ ، در سیال مورد نظر در هر دما و فشار معین، مقدار تراکم‌پذیری آدیاباتیک،  $\kappa_S$ ، طبق رابطه‌ی (۲-۱) به دست می‌آید [۱].

$$u^2 = \frac{1}{\rho \kappa_S} \quad (2-1)$$

که در آن  $\rho$ ، دانسیته و  $S$  آنترופی است و  $\kappa_S$  طبق رابطه‌ی (۳-۱) تعریف می‌شود:

$$\kappa_S = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial P} \right)_S \quad (3-1)$$

بنابراین طبق رابطه‌ی (۲-۱) تنها اندازه‌گیری  $u$  و  $\rho$  در یک دما و فشار معین برای تعیین  $\kappa_S$  یک سیال در همان شرایط ترمودینامیکی کافی است. پس از اندازه‌گیری  $\kappa_S$  می‌توان  $\kappa_T$  را با استفاده از رابطه‌ی (۴-۱) استخراج نمود [۱]:

$$\kappa_T = \kappa_S + \frac{T \alpha_P^2}{\rho C_P} \quad (4-1)$$

که در این رابطه  $C_P$ ، ظرفیت گرمایی سیال در فشار ثابت و  $\alpha_P$ ، ضریب انبساط‌پذیری هم‌فشار است و طبق رابطه‌ی (۵-۱) تعریف می‌شود:

$$\alpha_P = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial T} \right)_P \quad (5-1)$$

مقادیر سرعت صوت یک سیال در دماها و فشارهای مختلف توسط دستگاه اندازه‌گیری سرعت صوت اندازه‌گیری می‌شود. این دستگاه که در فرکانس ۳MHz تنظیم می‌شود سیلندری شامل دو صفحه‌ی پیزوالکتریک، محفظه‌ی قرار گرفتن نمونه، یک تولیدکننده‌ی فرکانس و یک گیرنده‌ی فرکانس است. به وسیله‌ی تولیدکننده‌ی فرکانس و صفحات پیزوالکتریک، امواج صوتی ایجاد شده و بعد از طی کردن طول محفظه‌ی نمونه،  $L$ ، و عبور از آن، توسط آشکارساز دریافت و زمان عبور امواج از نمونه،  $t_s$ ، اندازه‌گیری می‌شود. به این ترتیب سرعت صوت یک ماده از رابطه‌ی (۶-۱) قابل محاسبه خواهد بود:

$$u = \frac{L}{t_s} \quad (6-1)$$

از آن جا که طول محفظه‌ی نمونه تحت تأثیر دما و فشار است، می‌توان آن را به صورت تابعی از دما و فشار طبق رابطه‌ی (۷-۱) بیان نمود [۱].

$$L(P, T) = L_0 \{1 + a(T - T_0)\} \cdot \{1 + b(P - P_0)\} \quad (7-1)$$

که در این رابطه، ضرایب  $a$  و  $b$  با کالیبره کردن دستگاه به دست می‌آیند. مقادیر اندازه‌گیری شده‌ی  $u$  بر حسب دما و فشار در یک رابطه‌ی همبستگی برآزش می‌یابد. در اکثر مراجع علمی این تابع همبستگی به شکل رابطه‌ی (۸-۱) در نظر گرفته می‌شود.

$$\frac{1}{u^2} = (A + BP + CP^2 + DP^3) / (E + FP) \quad (8-1)$$

که در این رابطه ضرایب  $B, C, D, F$ ، ثابت بوده و وابستگی دمایی ضرایب  $A$  و  $E$  مطابق روابط (۹-۱) و (۱۰-۱) بیان می‌شوند [۱].

$$A = A_0 + A_1T + A_2T^2 \quad (9-1)$$

$$E = 1 + E_1T \quad (10-1)$$

صحت ضرایب به دست آمده برای این رابطه با استفاده از پارامترهای آماری تعیین می‌گردد.

مقادیر دانسیته در گستره‌ی دماها و فشارهای مختلف نیز با استفاده از دستگاه مخصوص اندازه‌گیری دانسیته تعیین می‌شود که علاوه بر فشار اتمسفر قادر است مقادیر دانسیته را در فشارها و دماهای متفاوت اندازه‌گیری نماید [۱].

### ۱-۲-۲- اندازه‌گیری $\kappa_T$ با استفاده از اندازه‌گیری دانسیته در دماها و فشارهای مختلف

در این روش ابتدا با استفاده از دستگاه دانسیمتر<sup>۱</sup>، دانسیته‌ی سیال، در دماها و فشارهای متفاوت اندازه‌گیری می‌شود. سپس داده‌های به دست آمده در یک معادله به فرم معادله‌ی تیت<sup>۲</sup> برازش می‌یابد:

$$\rho(T, P) = \frac{\rho_0(T)}{1 - C \ln \left( \frac{B(T) + P}{B(T) + P_0} \right)} \quad (11-1)$$

که در این معادله  $P_0$ ، فشار اتمسفر و برابر ۰/۱ MPa است و وابستگی دمایی کمیت‌های  $\rho_0(T)$  و  $B(T)$  طبق روابط زیر تعریف می‌گردند:

---

1. Densimeter  
2. Tait- equation

$$\rho_0(T) = A_0 + A_1T + A_2T^2 \quad (12-1)$$

$$B(T) = B_0 + B_1T + B_2T^2 \quad (13-1)$$

از طریق برازش مقادیر اندازه‌گیری شده‌ی دانسیته با روابط (۱۱-۱) تا (۱۳-۱) پارامترهای این روابط تعیین می‌گردد. پس از تعیین ضرایب معادله‌ی (۱۱-۱)، می‌توان ضریب تراکم‌پذیری همدم را طبق معادله‌ی (۱۴-۱) برای ترکیب مورد نظر به دست آورد [۳و۲].

$$\kappa_T = \frac{C}{\left(1 - C \ln \left( \frac{B(T) + P}{B(T) + P_0} \right)\right) (B(T) + P)} \quad (14-1)$$

### ۳-۱- مروری بر روش‌های پیش‌بینی $\kappa_T$

#### ۱-۳-۱- پیش‌بینی $\kappa_T$ با استفاده از معادله حالت

همانطور که در رابطه‌ی (۱-۱) ذکر شد ضریب تراکم‌پذیری همدم را می‌توان از مشتق مرتبه‌ی اول معادله‌ی حالت<sup>۱</sup> به دست آورد. معادلاتی که برای توصیف رفتار  $PVT$  مواد نوشته می‌شود، معادله‌ی حالت نام دارد.

در سال ۲۰۰۵ دکتر پارسا فر و همکارانش، معادله‌ی حالت<sup>۲</sup> MLIR را برای پیش‌بینی خواص ترمودینامیکی ترکیبات آلی چگال ارائه دادند. آن‌ها برای ارائه‌ی این معادله، هر ترکیب آلی چگال را به صورت محلول فرضی از گروه‌های تشکیل‌دهنده‌اش در نظر گرفتند و سپس معادله‌ی حالت زیر را برای این سیالات ارائه دادند [۴].

---

1. Equation of State  
2. Modified Linear Isotherm Regularity

$$\left(\frac{Z}{n} - 1\right)V^2 = A_m + B_m\rho^2 \quad (15-1)$$

که در آن  $Z = \frac{P}{\rho RT}$ ، فاکتور تراکم‌پذیری،  $n$  تعداد گروه‌های کربنی،  $A_m$  و  $B_m$  پارامترهای معادله‌ی حالت می‌باشند. آن‌ها سپس با بازآرایی این معادله به شکل معادله‌ی (۱۶-۱) و مشتق‌گیری از آن توانستند معادله‌ی (۱۷-۱) را برای محاسبه‌ی  $\kappa_T$  ارائه کنند:

$$P = n\rho RT + n^3\rho^3RTA_m + n^5\rho^5RTB_m \quad (16-1)$$

$$\kappa_T = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial P} \right)_T = \frac{1}{n\rho RT + 3n^3\rho^3RTA_m + 5n^5\rho^5RTB_m} \quad (17-1)$$

آن‌ها با استفاده از معادله‌ی (۱۷-۱) توانستند  $\kappa_T$  آلکان‌های خطی را با متوسط خطای نسبی ۱/۵٪ پیش‌بینی کنند [۴].

در سال ۲۰۰۵ گوهرشادی و همکارش، معادله حالت GMA<sup>۱</sup> را به فرم زیر برای پیش‌بینی خواص ترمودینامیکی تعدادی از سردکننده‌های هیدروفلوئورو اتر ارائه کردند.

$$P = \left( \frac{B_0RT}{2} - B_1 + TB_2 \ln T \right) \rho^5 + \left( \frac{A_0RT}{2} - A_1 + A_2T \ln T \right) \rho^4 + \frac{\rho RT}{2} \quad (18-1)$$

---

1. Goharshadi–Morsali–Abbaspour

که در این رابطه  $A_0$  تا  $A_2$  و  $B_0$  تا  $B_2$  پارامترهای معادله‌ی حالت هستند. سپس آن‌ها معادله‌ی (۱۹-۱) را با مشتق‌گیری از رابطه‌ی (۱۸-۱) برای محاسبه‌ی ضریب تراکم‌پذیری همدمای این ترکیبات به دست آوردند.

$$\kappa_T = \frac{2}{\rho RT + 4\rho^4(RTA_0 - 2A_1 + 2TA_2 \ln T) + 5\rho^5(B_0RT - 2B_1 + 2B_2T \ln T)} \quad (19-1)$$

آن‌ها توسط این رابطه،  $\kappa_T$  پنتافلورو دی متیل اتر، (۲-دی فلورو متوکسی)-۱ و ۱-تری فلورو اتان و پنتا فلورو اتیل متیل اتر را در دما و فشارهای متفاوت به دست آوردند [۵].

### ۱-۳-۲- پیش‌بینی $\kappa_T$ با استفاده از شبکه‌ی عصبی مصنوعی

در سال‌های اخیر روش‌های قدرتمندتر و در عین حال کارآمدتری برای پیش‌بینی خواص ترمودینامیکی مورد توجه قرار گرفته‌اند. یکی از بهترین این روش‌ها مدل‌سازی QSPR است. نمونه‌ای از پیش‌بینی این خواص با استفاده از شبکه‌ی عصبی مصنوعی در سال ۱۹۹۷ بر روی ۱۰۲ نوع نفت خام خاورمیانه انجام گرفته است که در آن از ۵۲۰ داده تجربی برای طراحی شبکه استفاده شده است. مدل ارائه شده قادر است  $\kappa_T$  انواع نفت خام خاورمیانه را به صورت تابعی از نسبت حلالیت گاز به نفت، وزن مخصوص گاز، وزن مخصوص نفت، درجه حرارت مخزن و فشار مخزن پیش‌بینی نماید [۶].

در سال ۲۰۰۵ نیز دکتر پارسا و همکارانش، توانستند یک مدل شبکه‌ی عصبی موجک بر مبنای سهم گروه‌ها ارائه کنند و  $\kappa_T$  آلکان‌ها را با متوسط خطای نسبی ۱/۳۸٪ پیش‌بینی نمایند [۷].

فصل دوم

كمومترىكس



## ۱-۲- کمومتریکس<sup>۱</sup>

امروزه دستیابی به اطلاعات در سیستم‌های شیمیایی بسیار آسان‌تر از گذشته می‌باشد که این موضوع به استفاده از رایانه در شیمی ارتباط دارد. با استفاده از رایانه، ریاضی و آمار، مجموعه‌ای از یک سری قواعد شیمیایی به نام کمومتریکس شکل می‌گیرد که در عرصه‌های ارزیابی و تفسیر اطلاعات، بهینه کردن و مدل‌سازی فرآیندها و آزمایشات و استخراج حداکثر اطلاعات شیمیایی از داده‌های تجربی، به ما کمک می‌کند. درحقیقت هدف کمومتریکس، بهبود بخشیدن فرآیندهای اندازه‌گیری و استخراج اطلاعات شیمیایی مفیدتر از داده‌های اندازه‌گیری شده فیزیکی و شیمیایی می‌باشد.

طبق یکی از جامع‌ترین تعاریف، کمومتریکس شاخه‌ای از شیمی است که از ریاضی و آمار و منطق برای رسیدن به اهداف زیر استفاده می‌کند:

الف) فرآیندهای تجربی بهینه را طراحی و انتخاب کند.

ب) حداکثر اطلاعات شیمیایی قابل حصول را از تحلیل داده‌های شیمیایی فراهم کند.

ج) بتواند اطلاعات بیشتری درمورد سیستم‌های شیمیایی به دست آورد [۸].

کمومتریکس به عنوان یک شاخه‌ی علمی جوان در دو دهه‌ی 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 نرم‌افزارهای ساده است. بنابراین امروزه کمومتریکس به عنوان ابزاری جدید در همه‌ی شاخه‌های شیمی مورد استفاده قرار می‌گیرد و هر شیمیدان در تحقیقات علمی‌اش به استفاده از این روش‌های آماری و ریاضی نیاز دارد.

---

1. Chemometrics

در سال ۱۹۷۱ سوانت ولد<sup>۱</sup> از دانشگاه اوما در سوئد و کمی بعد بروس کووالسکی<sup>۲</sup> از دانشگاه واشنگتن کلمه‌ی کمومتریکس را در نوشتارهای علمی رواج دادند. در ادامه‌ی رشد کمومتریکس، انجمن بین المللی کمومتریکس در سال ۱۹۷۴ تاسیس و اولین مقاله با عنوان کمومتریکس در سال ۱۹۷۵ به چاپ رسید.

## ۲-۲- رابطه‌ی ساختار- خاصیت یا ساختار- فعالیت

از زمینه‌های مهم کمومتریکس می‌توان به رابطه‌ی کمی ساختار- خاصیت (QSPR)<sup>۳</sup> و رابطه‌ی کمی ساختار- فعالیت (QSAR)<sup>۴</sup> اشاره کرد. QSAR مطالعاتی است که فعالیت‌های بیولوژیکی ترکیبات را به ویژگی‌های ساختاری آن‌ها ارتباط می‌دهد ولی QSPR ارتباط ساختار با خاصیت را مورد بررسی قرار می‌دهد. این مطالعات در واقع سعی در پیدا کردن رابطه‌ای هماهنگ میان فعالیت بیولوژیکی یا خاصیت مواد با ویژگی‌های مولکولی دارد، تا بتواند این قواعد را برای ارزیابی فعالیت یا خاصیت ترکیبات جدید به کار ببرد. برای تعیین این ارتباطات، در هر مطالعه‌ی QSPR یا QSAR باید مرحله‌ی طی شود که در ادامه به چگونگی روند این مطالعات می‌پردازیم.

## ۲-۲-۱- فراهم کردن سری داده‌ها

اولین مرحله، انتخاب یک سری داده مناسب و معتبر می‌باشد. هر چه سری داده‌ها وسیع‌تر و متنوع‌تر باشد، مدل به دست آمده مدلی کلی‌تر و معتبرتر بوده و محدود به یک گروه خاص از ترکیبات نخواهد بود.

- 
1. Svante Wold
  2. Bruce Kowalski
  3. Quantitative Structure- Property Relationship
  4. Quantitative Structure- Activity Relationship

## ۲-۲-۲- رسم ساختار و بهینه‌سازی

در این مرحله از مطالعه، ابتدا ساختار مولکولی هر ترکیب در نرم‌افزار HyperChem7.0 ترسیم می‌شود. سپس با احتساب اتم‌های هیدروژن، ساختار سه بعدی ترکیبات با استفاده از روش نیمه تجربی کوانتومی AM1 بهینه شده و بهینه‌سازی تا زمانی ادامه می‌یابد که جذر میانگین مربعات گرادیان<sup>۱</sup> انرژی به ۰/۰۰۱ کیلو کالری بر مول برسد.

## ۲-۲-۳- محاسبه‌ی توصیف‌کننده‌ها

توصیف‌کننده‌ها مقادیر عددی هستند که ویژگی‌های مختلف ساختاری و الکترونی مولکول‌ها را به طور کمی نشان می‌دهد. در این تحقیق ۱۴۸۱ توصیف‌کننده‌ی ساختاری توسط نرم‌افزار Dragon و همچنین تعدادی توصیف‌کننده‌ی الکترونی، توسط نرم‌افزار HyperChem محاسبه می‌شود.

## ۲-۲-۴- حذف توصیف‌کننده‌های نامناسب

پس از آن که توصیف‌کننده‌ها محاسبه شدند، باید توصیف‌کننده‌هایی که اطلاعات کمتری در مورد تغییرات خواص شیمیایی و یا بیولوژیکی موردنظر دارند، حذف شوند. برای این منظور لازم است که از فرایند کاهش متغیر استفاده گردد، به این ترتیب که توصیف‌کننده‌هایی که برای تمام مولکول‌ها، مقادیر ثابت، تقریباً ثابت و یا صفر داشته باشند، حذف می‌شوند. همچنین توصیف‌کننده‌هایی که کمتر از ۱۰٪ مقادیر غیرصفر و یا بیش از ۹۰٪ مقادیر یکسان دارند، مفید نبوده و کنار گذاشته می‌شوند. در مرحله‌ی بعد باید همبستگی بین توصیف‌کننده‌ها بررسی شود و توصیف‌کننده‌هایی که همبستگی بزرگ‌تر از ۰/۹ دارند حذف شوند. برای این منظور می‌توان از نرم‌افزار SPSS استفاده کرد.

---

1. Root Mean Square Gradient

اما انجام این کار با استفاده از این نرم‌افزار کاری بسیار وقت‌گیر است. بنابراین می‌توان از نرم‌افزار MATLAB استفاده نمود. برای این منظور با استفاده از برنامه‌ای طراحی شده به نام `remcol`، همبستگی توصیف‌کننده‌ها بررسی شده و توصیف‌کننده‌های حذفی مشخص می‌گردند. به این ترتیب تعداد توصیف‌کننده‌ها و به دنبال آن پیچیدگی محاسبات به میزان زیادی کاهش می‌یابد.

## ۲-۲-۵- دسته‌بندی داده‌ها

قبل از ساختن مدل، سری داده‌ها به سه زیر مجموعه تقسیم می‌شوند. اولین زیر مجموعه، سری آموزش نام دارد که رگرسیون مرحله‌ای بر روی آن انجام شده تا مدل‌های مختلفی به دست آیند. دومین زیر مجموعه سری ارزیابی است که مدل‌های به دست آمده از سری آموزش توسط آن ارزیابی شده و به این ترتیب توصیف‌کننده‌های بهینه برای ساخت مدل‌های خطی و غیرخطی تعیین می‌شوند. زیر مجموعه‌ی سوم سری تست یا آزمایش نام دارد که در طول فرایند مدل‌سازی کاربرد نداشته و از آن برای مقایسه‌ی مدل‌های متفاوت استفاده می‌شود.

## ۲-۲-۵-۱- دسته‌بندی داده‌ها به روش تحلیل مولفه‌های اساسی (PCA)<sup>۱</sup>

برای تقسیم‌بندی داده‌ها می‌توان از روش تقسیم تصادفی استفاده کرد و یا برای حصول نتیجه‌ی بهتر، از روش تحلیل مولفه‌های اساسی (PCA) بهره گرفت. در این تحقیق برای دسته‌بندی داده‌ها از این روش استفاده شده است. روش فوق یکی از با ارزش‌ترین نتایج کاربرد جبرخطی است که به وفور در کلیه‌ی اشکال تحلیلی از علوم شبکه‌ی عصبی تا نمودارهای کامپیوتری استفاده شده است، چرا که یک روش آسان و غیرپارامتریک برای استخراج اطلاعات مرتبط از یک مجموعه داده‌ی پیچیده است.

---

1. Principle Component Analysis

اگر در هر مجموعه داده  $n$  متغیر وجود داشته باشد، هر متغیر می‌تواند دارای چند بعد باشد. با توجه به اینکه اغلب درک و شهود فضای چند بعدی دشوار است، روش تحلیل مولفه‌های اساسی ابعاد کلیه‌ی مشاهدات را بر اساس شاخص ترکیبی و دسته‌بندی مشاهدات مشابه کاهش می‌دهد. به طور کلی کاربرد عمده‌ی روش تحلیل مولفه‌های اساسی عبارت است از کاهش تعداد متغیرها و یافتن ساختار ارتباطی بین آن‌ها که در حقیقت همان دسته‌بندی متغیرها است.

برای انجام کار باید توسط نرم‌افزار SPSS و از طریق منوی آنالیز، مولفه‌های اساسی برای مجموعه‌ی توصیف‌کننده‌ها به دست آید. معمولاً دو مولفه‌ی اول مقدار قابل توجهی از پراکندگی داده‌ها را در نظر می‌گیرد. بنابراین نموداری از مولفه‌ی اول بر حسب مولفه‌ی دوم ترسیم می‌شود. شکل این نمودار پراکندگی داده‌ها را نشان می‌دهد. بر اساس این نمودار نقطه داده‌های محیطی در سری آموزش قرار می‌گیرند و سری ارزیابی و تست از داده‌های داخلی به طور تصادفی انتخاب می‌شوند.

## ۲-۲-۶- انتخاب بهترین توصیف‌کننده‌ها

مرحله‌ی بعدی در مطالعات QSPR گزینش بهترین توصیف‌کننده‌های محاسبه شده می‌باشد. با وجود این که در مراحل قبل با به کار بردن روش‌های کاهش متغیر، تعداد توصیف‌کننده‌ها به میزان زیادی کاهش یافته است، اما هنوز تعداد آن‌ها برای ساخت یک مدل مناسب زیاد است. بنابراین می‌توان با به‌کارگیری روش‌هایی، بهترین توصیف‌کننده‌ها را از میان توصیف‌کننده‌های باقی‌مانده، انتخاب نمود. از جمله‌ی این روش‌ها می‌توان به روش رگرسیون مرحله‌ای، الگوریتم ژنتیک و راهبرد CDFS اشاره کرد. در ادامه به اختصار راجع به دو روش رگرسیون مرحله‌ای و راهبرد CDFS که در این تحقیق از آن استفاده شده است توضیحاتی ارائه می‌گردد.

## ۲-۶-۱- انتخاب بهترین توصیف‌کننده‌ها توسط رگرسیون مرحله‌ای

در این روش متغیرهای مستقل به صورت مرحله‌ای به ترتیب ضریب همبستگی وارد مدل می‌شوند. در هر مرحله برای سنجش معنادار بودن ضریب همبستگی متغیرهای وارد شده در مدل، آزمون  $F$  انجام می‌گیرد و در صورتی که  $F$  از نظر آماری برای متغیر معنادار باشد، آن متغیر در مدل باقی مانده و در غیر این صورت حذف خواهد شد. به این ترتیب کلیه‌ی مدل‌های ممکن ساخته شده و برای هر یک پارامترهای آماری شامل ضریب همبستگی، مقدار میانگین مربعات خطا و آماره‌ی  $F$  تعیین می‌شود و به این وسیله توصیف‌کننده‌های بهینه به روش رگرسیون مرحله‌ای به دست می‌آیند.

## ۲-۶-۲- انتخاب بهترین توصیف‌کننده‌ها توسط راهبرد CDFS

در مطالعات QSPR ساختن مدل مناسب و معتبر مستلزم انتخاب متغیرهای (توصیف‌کننده) مناسب و اثرگذار است. بنابراین روش‌هایی که برای انتخاب توصیف‌کننده در QSPR به کار می‌روند، حائز اهمیت‌اند. روش معمول برای انتخاب توصیف‌کننده‌ها و ساخت مدل، رگرسیون مرحله‌ای است. اما اشکال بزرگی که در رگرسیون مرحله‌ای وجود دارد، امکان ایجاد مدل‌های شانسی و تصادفی است. این اتفاق بخصوص زمانی که نسبت تعداد داده‌ها به تعداد متغیرها بسیار کم باشد، رخ خواهد داد. گاهی حضور یا عدم حضور بعضی از توصیف‌کننده‌ها در مدل حاصل از یک دسته‌بندی تصادفی، نشان از غلط بودن دسته‌بندی مورد نظر است. بنابراین تجربه نشان داده است که نحوه‌ی دسته‌بندی داده‌ها بر مدل ساخته شده و عملکرد آن اثرگذار است. در ضمن یک مدل تصادفی ممکن است توانایی کمی برای پیش‌بینی داده‌های سری تست داشته باشد [۹۱۰].

برای بررسی اثر دسته‌بندی داده‌ها بر متغیرهای انتخاب شده، روشی به نام CDFS پیشنهاد شده است. راهبرد CDFS در واقع روشی قابل استفاده برای انتخاب توصیف‌کننده‌های اثرگذار و معتبر از

میان تمامی توصیف‌کننده‌های محاسبه شده است. به وسیله این روش می‌توان مدل‌های خطی و غیرخطی را بهبود بخشید و با تعداد متغیر کمتر قدرت پیش‌بینی مدل را افزایش داد [۹۱۰].

برای اجرای این روش، ابتدا تعدادی از مولکول‌ها به عنوان سری تست انتخاب و کنار گذاشته می‌شوند. داده‌های باقی‌مانده، چندین بار به صورت‌های متفاوت به دو سری آموزش و ارزیابی تقسیم‌بندی می‌شوند. همانطور که در بخش ۲-۲-۵ اشاره شد دسته‌بندی داده‌ها می‌تواند به روش تصادفی و یا روش دقیق‌تر تحلیل مولفه‌های اساسی صورت گیرد. هر دسته‌بندی مراحل ساخت و ارزیابی مدل را به طور جداگانه طی می‌کند، به این ترتیب که رگرسیون مرحله‌ای بر روی سری آموزش هر دسته صورت می‌گیرد. سپس مدل‌های به دست آمده توسط سری ارزیابی همان دسته مورد بررسی قرار گرفته تا بهترین مدل برای دسته‌بندی مورد نظر به دست آید. به این ترتیب در انتها به ازای هر دسته‌بندی یک مدل بهینه به دست خواهد آمد. بررسی مدل‌های تولید شده از انواع مختلف دسته‌بندی، نشان می‌دهد که بعضی از توصیف‌کننده‌ها در همه یا اکثر مدل‌ها تکرار شده‌اند. تکرار این توصیف‌کننده‌ها نشان‌دهنده‌ی اثرگذار بودن این متغیرهای مستقل بر متغیر وابسته مورد اندازه‌گیری است. اما تعدادی توصیف‌کننده‌ی دیگر نیز وجود دارد که فقط در یک یا چند مدل به طور تصادفی ایجاد شده‌اند و بنابراین اهمیت چندانی ندارند. توصیف‌کننده‌هایی که بیش از ۵۰٪ در مدل‌های مختلف تکرار شده‌اند به عنوان توصیف‌کننده‌های نهایی انتخاب می‌شوند [۹۱۰].

ویژگی مهم توصیف‌کننده‌های حاصل از CDFS این است که، چون از انواع متفاوت دسته‌بندی مولکول‌ها به دست آمده‌اند، نسبت به نوع و تعداد مولکول‌های موجود در مدل حساس نیستند و به این ترتیب مشکل روش رگرسیون مرحله‌ای، یعنی وابسته بودن توصیف‌کننده‌ها به دسته‌بندی و در نتیجه تصادفی بودن مدل، حل خواهد شد. همین موضوع اهمیت روش CDFS را در مطالعات QSPR اثبات می‌کند.

باید به این نکته اشاره کرد که راهبرد CDFS بسیار شبیه روش ارزیابی تقاطعی است، به طوری که می‌توان آن را نوعی روش ارزیابی تقاطعی متغیر در نظر گرفت. اما تفاوت اساسی بین راهبرد CDFS و ارزیابی تقاطعی این است که در ارزیابی تقاطعی یک مدل ثابت در نظر گرفته می‌شود و برای ارزیابی قدرت پیش‌گویی مدل، به طور تکرارپذیر عمل حذف مرحله‌ای صورت می‌گیرد. اما در تکنیک CDFS همه‌ی مراحل دسته‌بندی داده‌ها، انتخاب توصیف‌کننده‌های بهینه و محاسبه‌ی ضرایب در هر مرحله به طور جداگانه انجام می‌شود [۱۰].

با مقایسه‌ی نتایج حاصل از CDFS با روش رگرسیون مرحله‌ای مرسوم، این نتیجه به دست می‌آید که با استفاده از این روش می‌توان مدلی با تعداد توصیف‌کننده‌های کمتر، که حساسیت کمی نسبت به نحوه‌ی دسته‌بندی داده‌ها دارند، ایجاد نمود. در ضمن نتایج به دست آمده نشان می‌دهند که مدل حاصل از CDFS قدرت پیش‌بینی مدل را به میزان خوبی افزایش می‌دهد [۹ و ۱۰].

## ۷-۲-۲- ساختن مدل

مرحله‌ی اصلی در مطالعات QSPR ساختن مدل می‌باشد. مدل‌های QSPR را می‌توان بر اساس پارامترهایی که به صورت تجربی به دست آمده و یا از طریق تئوری محاسبه شده‌اند، به زبان ریاضی بیان نمود. اگر نتایج این نوع مطالعات همبستگی قابل قبولی را ارائه دهد، علاوه بر شفاف‌سازی نحوه‌ی ارتباط بین خواص مولکول‌ها و ویژگی‌های ساختمانی آن‌ها، به پژوهشگران در پیش‌بینی رفتار مولکول‌های جدید بر اساس رفتار مولکول‌های مشابه کمک می‌کند.

چندین روش متفاوت برای ساختن مدل وجود دارد که به دو دسته‌ی کلی روش‌های خطی و غیرخطی طبقه‌بندی می‌شوند. در مدل خطی با در نظر گرفتن اینکه رابطه‌ی بین توصیف‌کننده‌ها و متغیر وابسته، یک رابطه‌ی خطی است، مدل‌سازی به روش خطی صورت می‌گیرد. زمانی که مدل‌های



خطی نتایج مطلوبی ارائه ندهند و یا روابط غیرخطی بین توصیف‌کننده‌ها و متغیر وابسته وجود داشته باشد، از مدل غیرخطی استفاده می‌شود. از جمله روش‌های خطی می‌توان از رگرسیون خطی چندگانه (MLR)<sup>۱</sup>، آنالیز اجزای اصلی (PCR)<sup>۲</sup> و حداقل مربعات جزئی (PLS)<sup>۳</sup> و از روش‌های غیرخطی، از شبکه‌ی عصبی مصنوعی (ANN)<sup>۴</sup> و الگوریتم ژنتیک (GA)<sup>۵</sup> نام برد.

## ۲-۲-۷-۱- رگرسیون خطی چندگانه (MLR)

رگرسیون خطی چندگانه برای به دست آوردن رابطه‌ی خطی بین یک متغیر وابسته مثل  $y$  و متغیرهای مستقلی مثل  $x_1, x_2, \dots, x_n$  به کار می‌رود. معادله‌ی خطی این رابطه به صورت زیر است:

$$y = b_0 + b_1x_1 + \dots + b_nx_n \quad (۱-۲)$$

که در آن  $n$  تعداد متغیرها،  $b_i$  آامین ضریب رگرسیون و  $x_i$  مقدار عددی آامین متغیر مستقل است.

## ۲-۲-۷-۲- شبکه‌ی عصبی مصنوعی

شبکه‌های عصبی مصنوعی زیر مجموعه‌ای از هوش مصنوعی هستند. این شبکه‌ها الگویی از سیستم‌های عصبی زیستی هستند. انسان‌ها از زمان‌های بسیار دور سعی بر آن داشتند که بیوفیزیولوژی مغز را دریابند، چون همواره مسئله‌ی هوشمندی انسان و قابلیت یادگیری، تعمیم، خلاقیت، انعطاف‌پذیری و پردازش موازی در مغز برای بشر جالب بود و به همین دلایل بکارگیری این

- 
1. Multiple Linear Regression
  2. Principle Component Regression
  3. Partial Least Square
  4. Artificial Neural Network
  5. Genetic Algorithm

قابلیت‌ها در ماشین‌ها بسیار مطلوب می‌نمود. با توجه به این که روش‌های الگوریتمی برای پیاده‌سازی این خصایص در ماشین‌ها مناسب نبودند، در نتیجه از روش‌های مبتنی بر مدل‌های بیولوژیکی استفاده گردید. شبکه‌های عصبی مصنوعی، ساختارهای محاسباتی جدیدی هستند که قادر به پردازش اطلاعات و یادگیری طبق اصول الهام گرفته از سیستم عصبی موجودات زنده می‌باشند و با استفاده از تعداد زیادی مثال آموزش می‌بینند.

مغز به عنوان یک سیستم پردازش اطلاعات با ساختاری موازی، از حدود یک صد میلیارد سلول عصبی به نام نرون تشکیل شده که هر نرون به ده هزار نرون دیگر مرتبط می‌باشند. اکثر نرون‌ها از سه قسمت اساسی تشکیل شده‌اند:

۱- دندریت‌ها گیرنده‌های نرون هستند و سیگنال‌های الکتریکی را به قسمت بدنه‌ی سلول منتقل می‌کنند.

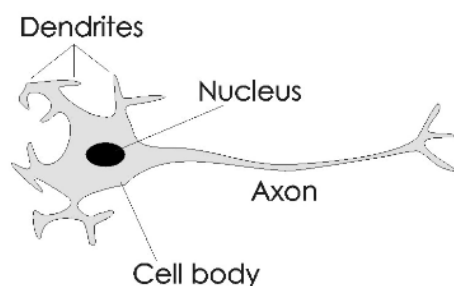
۲- بدنه‌ی سلول که به آن سوما<sup>۱</sup> نیز می‌گویند انجام عملیات لازم بر روی اطلاعات ورودی را بر عهده دارد.

۳- آکسون رشته‌ای طولانی است که اطلاعات خروجی از سوما را دریافت کرده و به خارج نرون می‌فرستد.

محل اتصال آکسون یک سلول به دندریت‌های سلول‌های دیگر را سیناپس<sup>۲</sup> می‌گویند. پیام‌های عصبی تنها به صورت یک طرفه از دندریت به بدنه سلول و از آنجا به آکسون حرکت می‌کنند. وقتی پیام به انتهای آکسون می‌رسد فیبرهای ماهیچه‌ای و یا نرون‌های دیگر که به آن متصل هستند تحریک می‌شوند [۱۱]. شکل (۱-۲) ساختار یک نرون زیستی را نشان می‌دهد.

---

1. Soma  
2. Synapse

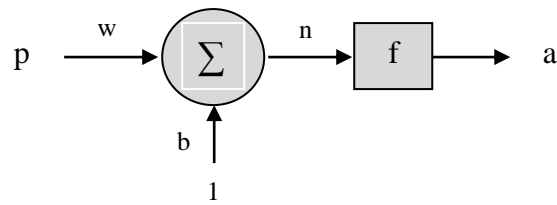


شکل (۱-۲) - ساختار یک نرون طبیعی

شبکه‌های عصبی مصنوعی نیز مشابه مغز از تعداد زیادی نرون تشکیل شده‌اند که از طریق ارتباطاتی به هم متصل شده‌اند. عنصر کلیدی در شبکه‌ی عصبی مصنوعی ساختار جدید سیستم پردازش اطلاعات آن است که از تعداد زیادی عناصر (نرون) با ارتباطات قوی داخلی که هماهنگ با هم برای حل مسائل مخصوص کار می‌کنند، تشکیل شده‌اند. شبکه‌ی عصبی مصنوعی با پردازش روی داده‌های تجربی، قانون نهفته در آن‌ها را به ساختار شبکه منتقل می‌کنند که به این عمل یادگیری می‌گویند. اصولاً توانایی یادگیری، مهم‌ترین ویژگی یک سیستم هوشمند است. سیستمی که بتواند یاد بگیرد منعطف‌تر است و ساده‌تر برنامه‌ریزی می‌شود، بنابراین بهتر می‌تواند در مورد مسائل و معادلات جدید پاسخگو باشد.

## ۲-۲-۷-۲-الف - ساختار شبکه‌های عصبی

یک نرون طبیعی، دنیایی از پیچیدگی‌ها را در خود دارد و به همین دلیل ارائه‌ی یک مدل مصنوعی برای آن، به نحوی که بتواند همه خصوصیات آن را داشته باشد، کار فوق‌العاده مشکلی است. بنابراین در مدلی که ارائه می‌گردد، سعی می‌شود که خصوصیات اصلی یک نرون طبیعی لحاظ شود. شکل (۲-۲) الگوی ساده‌ای از یک نرون محاسباتی را ارائه می‌دهد.



شکل (۲-۲) - ساختمان یک نرون محاسباتی

در شکل (۲-۲) کمیت‌های  $p$  و  $a$  به ترتیب ورودی و خروجی نرون می‌باشند. میزان تأثیر ورودی  $p$  روی خروجی  $a$  به وسیله پارامتر وزن  $w$  تعیین می‌شود. ورودی دیگر یک مقدار ثابت  $1$  است که در جمله‌ی بایاس<sup>۱</sup> ضرب شده و سپس با  $wp$  جمع می‌شود. این حاصل جمع، ورودی خالص  $n$  برای تابع تحریک  $f$  را تشکیل می‌دهد.

$$n = wp + b \quad (۲-۲)$$

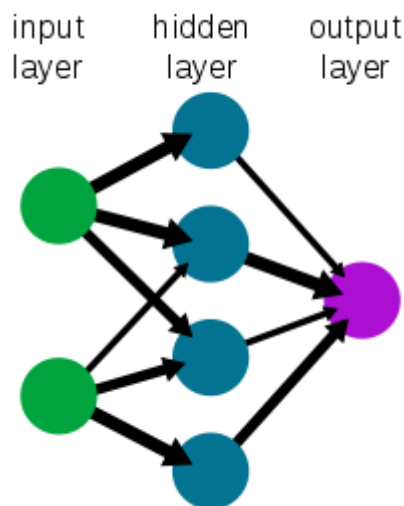
سپس خروجی نرون به کمک معادله‌ی زیر به دست می‌آید:

$$a = f(wp + b) \quad (۳-۲)$$

در اکثر موارد یک سلول عصبی تنها، حتی با ورودی‌های زیاد نمی‌تواند نیاز ما را بر آورده کند. به همین دلیل معمولاً به تعدادی از این سلول‌ها که به صورت موازی عمل می‌کنند، نیاز است. به این سلول‌های موازی یک لایه می‌گویند. در شکل (۳-۲) ساختار یک شبکه‌ی سه لایه مشاهده می‌شود.

---

1. Bias



شکل (۲-۳) - ساختار یک شبکه‌ی سه لایه

در حالت کلی در شبکه‌های عصبی ۳ نوع لایه‌ی نرونی وجود دارد:

۱- **لایه‌ی ورودی:** وظیفه‌ی دریافت اطلاعات خاصی که به شبکه تغذیه شده است، را دارد.

۲- **لایه‌های پنهان:** عملکرد لایه‌های پنهان به وسیله ورودی‌ها و وزن ارتباط‌دهنده‌ی بین آن‌ها و لایه‌های پنهان تعیین می‌شود. به عبارت دیگر، وزن‌های بین واحدهای ورودی و پنهان تعیین می‌کند که چه وقت یک واحد پنهان باید فعال شود.

۳- **لایه‌ی خروجی:** لایه‌ای است که خروجی آن همان خروجی شبکه است. عملکرد واحد خروجی به فعالیت واحد پنهان و وزن ارتباط بین واحد پنهان و خروجی بستگی دارد.

برای محاسبه‌ی خروجی هر نرون در لایه‌ی پنهان و لایه‌ی خروجی، ورودی‌های آن در وزن‌های مربوطه ضرب شده و سپس با هم جمع می‌شوند. آن‌گاه یک تابع انتقال<sup>۱</sup> بر این مجموعه اعمال می‌گردد. تابع انتقال مورد استفاده بر این لایه‌ی پنهان معمولاً تابعی غیرخطی است و چون در روش انتشار برگشتی از مشتق این تابع استفاده می‌شود، بهتر است از توابعی استفاده گردد که مشتق آن‌ها

---

1. Transfer Function

به سادگی قابل محاسبه باشد. نمونه‌هایی از توابع انتقال به کار برده شده در شبکه‌های عصبی مصنوعی عبارتند از:

تابع انتقال pureline، تابع انتقال تانژانت سیگموئید و تابع انتقال لگاریتم سیگموئید.

## ۲-۲-۷-۲-ب - انواع شبکه‌ها از نظر ارتباطات بین نرونی

ارتباطات بین نرونی در یک شبکه‌ی عصبی از اهمیت بسیاری برخوردار می‌باشد و به نوعی قدرت شبکه را تعیین می‌کند. در این نوع تقسیم‌بندی دو نوع شبکه وجود دارد:

۱- شبکه‌های پیش‌خور<sup>۱</sup>

۲- شبکه‌های پس‌خور<sup>۲</sup>

در شبکه‌های پیش‌خور، اتصالات بین نرونی به گونه‌ای است که هر نرون لایه‌ی میانی، می‌تواند ورودی خود را از هر یک از نرون‌های لایه‌ی قبلی خود دریافت کند اما هیچ سیگنالی از نرون به خودش یا به نرون لایه‌ی قبل بر نمی‌گردد. بدین ترتیب در یک شبکه‌ی عصبی، سیگنال‌ها به تدریج از یک لایه‌ی نرونی به لایه‌های بعدی حرکت می‌کنند و در نهایت به لایه‌ی آخر یا خروجی شبکه می‌رسند. در نتیجه جریان اطلاعات در این نوع از شبکه‌ها همیشه در یک مسیر یک طرفه و از ورودی به خروجی می‌باشد.

نوع دیگری از ارتباطات بین نرونی در شبکه‌های عصبی ارتباط بازخوردی است که در شبکه‌های پس‌خور وجود دارد. در این نوع از ارتباطات، خروجی یک نرون به همان نرون یا نرون‌های همان لایه یا به نرون‌های لایه‌ی قبلی اتصال داده می‌شود [۷]. بنابراین این نوع از شبکه‌ها می‌توانند با استفاده از

---

1. Feed forward  
2. Feed Back

حلقه‌های برگشتی، سیگنال‌هایی داشته باشند که در هر دو مسیر ورودی به خروجی و بالعکس حرکت کنند.

## ۲-۲-۷-۲-ج- آموزش شبکه عصبی

شبکه‌های عصبی می‌توانند براساس طراحی خود، سیگنال‌های ورودی را پردازش کنند و به سیگنال‌های خروجی مورد نظر تبدیل نمایند. به طور معمول پس از این که شبکه‌های عصبی طراحی و پیاده‌سازی شدند، باید پارامترهای پیش‌قدر و وزن‌ها ( $w$  و  $b$ ) به ازای مجموعه‌هایی از سیگنال‌های ورودی، به گونه‌ای تنظیم شوند که سیگنال‌های خروجی شبکه، خروجی مطلوب را تشکیل دهند. چنین فرایندی را آموزش شبکه‌ی عصبی می‌نامند. به طور کلی برای آموزش شبکه دو روش وجود دارد:

۱- یادگیری با ناظر<sup>۱</sup>

۲- یادگیری بدون ناظر<sup>۲</sup>

روش یادگیری با ناظر، شامل مراحل است که در آن سیستم از طریق در دست داشتن جواب واقعی آموزش می‌بیند. اما در روش بدون ناظر، شبکه‌ی عصبی باید خودسامان‌ده باشد و بدون در دست داشتن جواب واقعی، بتواند کار آموزش را انجام دهد. در این نوع یادگیری، شبکه می‌آموزد که الگوهای ورودی را به تعداد متناهی از گروه‌ها تقسیم‌بندی کند و به عبارتی عمل خوشه‌بندی را انجام می‌دهد. باید توجه داشت در این حالت فرد طراح است که مقصد و هدف نهایی یعنی چیزی را که باید به آن رسید مشخص می‌کند.

---

1. Supervised learning  
2. Unsupervised learning

رفتار سیستم‌های یادگیر توسط الگوریتم‌های بازگشتی بیان می‌شود که به آنها قوانین یادگیری می‌گویند. انواع مختلفی از قوانین یادگیری وجود دارد که یادگیری عملکردی<sup>۱</sup> یکی از آنهاست. قانون یادگیری در اینجا روندی است که توسط آن ماتریس وزن‌ها و بردارهای پیش‌قدر شبکه‌ی عصبی به نحوی تنظیم می‌شوند که عملکرد شبکه بهینه شود. منظور از بهینه شدن عملکرد شبکه، حداقل شدن خطایی است که بین مقادیر تجربی و پاسخ شبکه وجود دارد. برای بهینه‌سازی عملکرد، شبکه باید پارامترهایش را به گونه‌ای تنظیم کند که شاخص عملکرد<sup>۲</sup> کمترین مقدار را داشته باشد. شاخص عملکرد، معیاری است که به طور کمی، عملکرد شبکه را بهینه می‌کند و رابطه‌ی معکوسی با نحوه‌ی عملکرد شبکه دارد؛ یعنی هرچه مقدار این شاخص کوچک‌تر باشد، عملکرد شبکه بهتر خواهد بود و بالعکس. به شاخص عملکرد اصطلاحاً تابع هدف یا تابع خطا می‌گویند و در بیشتر الگوریتم‌ها، تابع هدف متوسط مربعات خطا<sup>۳</sup> (MSE) می‌باشد [۷].

## ۲-۲-۷-۲-د- معرفی روش‌های مختلف بهینه‌سازی عملکرد شبکه عصبی

الگوریتم‌های بهینه‌سازی تقریباً همگی دوری هستند یعنی؛ در تمام الگوریتم‌های بهینه‌سازی یک مقدار اولیه برای تابع حدس زده می‌شود و بعد این مقدار اولیه طی مراحل طی می‌شود که از قاعده‌ی زیر پیروی می‌کند بهینه می‌شود [۷]:

$$X_{K+1} = X_K + \alpha P_K \quad (۴-۲)$$

که در آن،  $X_K$  تخمین نقطه‌ی مینیمم تابع در مرحله‌ی  $K$  ام است و  $X_{K+1}$  حدس مرحله‌ی بعدی نقطه‌ی مینیمم تابع می‌باشد.  $P_K$  نیز بردار جستجو می‌باشد، بدین معنی که از روی آن، تخمین

---

1. Performance learning  
2. Performance  
3. Mean square error



مرحله‌ی بعد به دست می‌آید.  $\alpha$  یک مقدار اسکالر بین صفر و یک بوده و سرعت یادگیری<sup>۱</sup> نامیده می‌شود. سرعت یادگیری، طول قدم در هر تکرار را در هر مسیر  $P_K$  تعیین می‌کند. انتخاب  $\alpha$  و  $P_K$  روش‌های مختلف مینیمم‌سازی را از هم متمایز می‌کند که چند نمونه از آنها در زیر ذکر شده است:

۱- روش نزول گرادیانی<sup>۲</sup>

۲- الگوریتم پس انتشار خطا<sup>۳</sup>

۳- روش لونیبرگ-مارکوارت<sup>۴</sup>

۴- الگوریتم تنظیم بایزین

## ۲-۲-۸- ارزیابی قدرت پیش‌بینی مدل

برای اطمینان از اینکه مدل به دست آمده، مدل مناسبی است و توانایی پیش‌بینی نمونه‌های مختلفی از یک جمعیت را داراست، باید مدل را ارزیابی کرد. این ارزیابی از طریق پارامترهای آماری صورت می‌گیرد که به وسیله‌ی این شاخص‌های کمی صحت نتایج ارائه شده توسط مدل مورد سنجش قرار می‌گیرند. از جمله‌ی این پارامترهای آماری ضریب همبستگی رگرسیون  $R$  است که هر چه  $R$  یک مدل به یک نزدیک‌تر باشد، آن مدل مناسب‌تر خواهد بود. آماره‌ی  $F$  نیز معیار دیگری برای ارزیابی مدل بوده و مقدار آن بیانگر میزان معنی‌دار بودن کلی مدل است. در این تحقیق میزان  $MSE$  نیز برای مدل‌ها به دست می‌آید و مدلی با کمترین میزان  $MSE$  به عنوان مدل بهینه انتخاب می‌شود.

علاوه بر پارامترهای آماری، به منظور ارزیابی قدرت پیش‌بینی مدل، از شیوه حذف یکتایی (LOO)<sup>۵</sup> و

---

1. Learning rate  
2. Gradient descent algorithm  
3. Error back propagation algorithm  
4. Levenberg-Marquart  
5. Leave One Out

حذف چندتایی (LGO)<sup>۱</sup> استفاده می‌شود. در شیوه LOO هر بار یک داده از سری داده‌ها کنار گذاشته شده و مدل‌سازی با بقیه داده‌ها انجام می‌پذیرد. بدیهی است که در این حالت، به تعداد سری داده‌ها، باید این عمل صورت گیرد، اما در شیوه LGO هر بار یک سری چندتایی از سری داده‌ها کنار گذاشته می‌شود.

علاوه بر آن می‌توان برای بررسی همبستگی‌های شانس مدل‌های غیرخطی، از آزمون  $\chi^2$ -تصادفی استفاده کرد. در این آزمون، مقادیر تجربی به صورت تصادفی در محدوده‌ی همان مقادیر، تغییر داده شده و سپس همبستگی متغیرهای مستقل با متغیرهای وابسته با استفاده از یکی از شاخص‌های آماری که معمولاً  $R^2$  است، مورد بررسی قرار می‌گیرد. اختلاف زیاد بین شاخص آماری به دست آمده از این روش و شاخص آماری به دست آمده از مدل اصلی، نشان‌دهنده‌ی عدم وجود همبستگی شانس می‌باشد. به طور معمول این فرایند چندین بار تکرار می‌گردد.

## ۳-۲- انتخاب توصیف‌کننده‌ها بر مبنای روش سهم گروه<sup>۲</sup>

### ۳-۲-۱- روش سهم گروه

روش سهم گروه روشی ساده، کارآمد و در عین حال معتبر است که به طور گسترده‌ای در پیش‌بینی خواص فیزیکی و ترمودینامیکی مواد به کار می‌رود. این روش در واقع ارتباطی بین خاصیت یک ماده با ساختار شیمیایی آن برقرار می‌کند. مبنای به وجود آمدن روش سهم گروه‌ها این است که اگر چه هزاران ترکیب شیمیایی در دنیای علم و تکنولوژی وجود دارد اما تعداد گروه‌های ساختاری و عاملی سازنده‌ی همه‌ی این ترکیبات، بسیار کمتر از تعداد این ترکیبات می‌باشد. با این ایده، روش سهم گروه‌ها فرض می‌کند که خاصیت فیزیکی یا ترمودینامیکی هر ماده را می‌توان از سهم گروه‌های

---

1. Leave Group Out  
2. Group Contribution Method

موجود در مولکول آن ماده در خاصیت موردنظر، به دست آورد. به این ترتیب یک روش عملی برای پیش‌بینی خواص تعداد زیادی ماده برحسب تعداد کمتری پارامتر به دست می‌آید که این پارامترها، سهم گروه‌های ساختاری را در خاصیت موردنظر مشخص می‌کنند [۷].

لازم به توضیح است که تعریف گروه‌های تشکیل‌دهنده‌ی هر ماده، یک مقوله‌ی اختیاری است. به طور مثال ممکن است یک نفر پیشنهاد کند که باید اتم‌های اصلی شناسایی شوند. در حالی که ما می‌دانیم خواص کربن در الماس نسبت به کربن در گرافیت بسیار متفاوت است. از طرف دیگر، تجربه نشان داده است که اگر چه گروه کربونیل موجود در کتون‌ها نسبت به گروه کربونیل موجود در کربوکسیلیک اسیدها خواص متفاوتی دارد اما حداقل گروه کربونیل موجود در همه‌ی کتون‌ها مشابه می‌باشد. بدیهی است هر چه تمایز بین گروه‌ها بیشتر شود، صحت روش سهم گروه‌ها افزایش می‌یابد تا جایی که سرانجام هر ترکیب فقط شامل گروه‌های مربوط به خودش باشد. اما می‌توان با انتخاب صحیح گروه‌ها به روش سودمندی دست یافت که با وجود تعداد گروه‌های کم، صحت خواص پیش‌بینی شده قابل قبول باشد [۷].

## ۲-۳-۲- شبکه‌های عصبی مصنوعی بر مبنای روش سهم گروه‌ها

در مطالعات QSPR که به طور گسترده‌ای برای مدل‌سازی خواص فیزیکی و شیمیایی مواد به کار می‌رود، خاصیت مورد نظر به توصیف‌کننده‌های مولکولی مرتبط می‌شود. توصیف‌کننده‌های مولکولی مشخصه‌های عددی ماده‌ی خالص هستند که مستقیماً از روی ساختار بهینه شده‌ی مولکول و با استفاده از الگوریتم‌ها و نرم‌افزارهای خاص محاسبه می‌شوند. اما مهم‌ترین عیب این روش، مسیر طولانی و پیچیده‌ی محاسبه‌ی توصیف‌کننده از روی ساختار مولکولی است. در نتیجه مدل‌های حاصل از این مطالعات برای استفاده، چندان ساده نیستند. به همین دلیل در سال‌های اخیر از روش ساده و کارآمد سهم گروه‌ها برای انتخاب توصیف‌کننده‌های مناسب جهت مطالعات QSPR استفاده می‌شود.

در این روش تعداد گروه‌های تشکیل‌دهنده‌ی هر ترکیب به عنوان توصیف‌کننده انتخاب می‌شوند و در صورتی که یکی از گروه‌های تشکیل‌دهنده در ترکیبات مورد مطالعه وجود نداشته باشد توصیف‌کننده‌ی مرتبط با آن گروه، صفر در نظر گرفته می‌شود.

در طی چند سال اخیر از شبکه‌ی عصبی مصنوعی بر مبنای روش سهم گروه‌ها در پیش‌بینی خواص مختلفی از جمله برای پیش‌بینی نقطه‌ی اشتعال آلکان‌ها [۱۲]، دانسیته‌ی مایعات یونی [۱۳]، خصوصیات نقطه‌ی بحرانی تعدادی از هیدروکربن‌ها [۱۴]، هدایت یونی و ویسکوزیته‌ی مایعات یونی [۱۵]، استفاده شده است.

فصل سوم

بخش محاسبات

### ۳-۱- مدل سازی ضریب تراکم پذیری همدمای سیالات آلی مختلف

همانطور که در فصل قبل اشاره شد، شبکه‌ی عصبی مصنوعی یکی از روش‌های کارآمد در پردازش اطلاعات می‌باشد که قدرت یادگیری و محاسباتی بالایی دارد و در علوم مختلف از آن استفاده می‌شود. به همین علت در این پایان‌نامه از روش شبکه‌ی عصبی مصنوعی برای پیش‌بینی ضریب تراکم‌پذیری همدمای ترکیبات آلی مختلف در محدوده‌ی وسیعی از دما و فشار استفاده گردید. آنچه در ادامه می‌آید مراحل طراحی و بهینه‌سازی و همچنین ارزیابی مدل‌های مختلف شبکه‌ی عصبی مصنوعی برای پیش‌بینی  $K_T$  می‌باشد.

#### ۳-۱-۱- انتخاب سری داده‌ها

به منظور پیش‌بینی ضریب تراکم‌پذیری همدمای ۵۶ سیال آلی مختلف، ۴۲۹۷ داده‌ی ضریب تراکم‌پذیری همدما برای این ترکیبات در گستره‌ی وسیعی از دما و فشار جمع‌آوری گردید. نام این ترکیبات به همراه محدوده‌ی دما و فشار، تعداد نقطه داده‌ی مورد استفاده از هر ترکیب برای مدل‌سازی و مرجع مورد استفاده برای هر ترکیب در جدول (۳-۱) آمده است.

#### ۳-۱-۲- رسم و بهینه‌سازی ساختار مولکول‌ها

برای رسم و بهینه‌سازی ساختار مولکول‌ها از نرم‌افزار HyperChem7.0 استفاده گردید. به این ترتیب که ساختار هر ترکیب در این نرم‌افزار ترسیم شد و سپس با احتساب اتم‌های هیدروژن و با استفاده از روش نیمه‌تجربی AM1 ساختار مولکول‌ها بهینه گردید.

جدول (۳-۱) - محدوده‌ی دما و فشار و تعداد نقطه داده‌ی استفاده شده در مدل‌سازی برای سیالات آلی مختلف به همراه مرجع مورد استفاده برای هر یک از آن‌ها

مرجع	تعداد نقطه داده استفاده شده در مدل‌سازی	گستره‌ی فشاری (MPa)	گستره‌ی دمایی (K)	نام ترکیب
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	استیک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	پروپانویک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	بوتانویک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	پنتانویک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	هگزانویک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	هپتانویک اسید
[۱۶]	۷۰	۰/۱ - ۲۵	۲۹۳/۱۵ - ۳۲۳/۱۵	اکتانویک اسید
[۱۷]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۳	پروپیل استات
[۱۷]	۵۴	۰/۱ - ۳۰۰	۲۷۸/۱۵ - ۳۳۸/۱۳	بوتیل استات
[۱۸]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۳	اتیل پروپیونات
[۱۸]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۳	اتیل بوتیرات
[۱۷]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۳	متیل اتیل استات
[۱۹]	۴۲	۰/۱ - ۲۵۰	۲۷۸/۱۵ - ۳۳۸/۱۵	۲-بوتانول
[۲۰]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۵	۲-پنتانول
[۲۱]	۴۲	۰/۱ - ۳۰۰	۲۷۸/۱۵ - ۳۳۸/۱۵	۳-پنتانول
[۲۰]	۵۴	۰/۱ - ۲۰۰	۲۷۸/۱۵ - ۳۳۸/۱۵	۲-هگزانول
[۲۰]	۶۰	۰/۱ - ۳۵۰	۲۷۸/۱۵ - ۳۳۸/۱۵	۴-متیل-۲-پنتانول
[۲۲]	۱۰۰	۰/۱ - ۱۴۰	۲۹۳/۱۵ - ۴۰۳/۱۵	۲-آمینو بوتان
[۳]	۱۰۵	۰/۱ - ۱۴۰	۲۹۳/۱۵ - ۳۵۳/۱۵	پنتیل آمین
[۳]	۱۰۵	۰/۱ - ۱۴۰	۲۹۳/۱۵ - ۳۵۳/۱۵	هگزیل آمین
[۳]	۱۰۵	۰/۱ - ۱۴۰	۲۹۳/۱۵ - ۳۵۳/۱۵	هپتیل آمین
[۲۲]	۱۰۵	۰/۱ - ۱۴۰	۲۹۳/۱۵ - ۴۰۳/۱۵	۲-آمینو اکتان
[۲۳]	۱۲۶	۰/۱ - ۶۰	۲۷۸/۱۵ - ۳۵۸/۱۵	۱-پروپانول
[۲۳]	۱۲۶	۰/۱ - ۶۰	۲۷۸/۱۵ - ۳۵۸/۱۵	۱-بوتانول
[۲۴]	۱۲۶	۰/۱ - ۶۰	۲۷۸/۱۵ - ۳۵۸/۱۵	۱-پنتانول
[۲۵]	۶۶	۰/۱ - ۱۰۰	۲۹۳/۱۵ - ۳۶۳/۱۵	۲-متیل-۱-بوتانول
[۲۶]	۷۷	۰/۱ - ۱۰۰	۳۰۳/۱۵ - ۳۶۳/۱۵	۳-پنتانول
[۲۶]	۷۷	۰/۱ - ۱۰۰	۳۰۳/۱۵ - ۳۶۳/۱۵	۳-متیل-۳-پنتانول
[۲۶]	۷۷	۰/۱ - ۱۰۰	۳۰۳/۱۵ - ۳۶۳/۱۵	۳-اتیل-۳-پنتانول

ادامه‌ی جدول (۱-۳)

نام ترکیب	گستره‌ی دمایی (K)	گستره‌ی فشاری (MPa)	تعداد نقطه داده استفاده شده در مدل‌سازی	مرجع
۱-هگزانول	۲۷۸/۱۵ - ۳۵۸/۱۵	۰/۱ - ۶۰	۱۲۶	[۲۳]
۲-هگزانول	۳۱۳/۱۳ - ۳۶۲/۷۳	۱/۰۱ - ۲۲	۱۳۲	[۲۷]
۱-هپتانول	۲۷۸/۱۵ - ۳۵۸/۱۵	۰/۱ - ۶۰	۱۲۶	[۲۳]
۲-هپتانول	۳۱۳/۱۳ - ۳۶۲/۷۳	۱/۰۱ - ۲۲	۱۳۲	[۲۸]
۱-اکتانول	۲۷۸/۱۵ - ۳۵۸/۱۵	۰/۱ - ۶۰	۱۲۶	[۲۳]
۱-نونانول	۲۷۸/۱۵ - ۳۵۸/۱۵	۰/۱ - ۶۰	۱۲۶	[۲۳]
۱-دکانول	۲۸۸/۱۵ - ۳۵۸/۱۵	۰/۱ - ۶۰	۱۱۲	[۲۳]
۱-دودکانول	۳۲۳/۱۵ - ۳۴۸/۱۵	۰/۱ - ۴۰	۱۹	[۲۹]
۱-تترادکانول	۳۲۳/۱۵ - ۳۴۸/۱۵	۰/۱ - ۴۰	۲۲	[۲۹]
۱-هگزادکانول	۳۴۸/۱۵	۰/۱ - ۴۰	۱۱	[۲۹]
۲-متیل-۲-و۴-پنتان‌دی‌ال	۲۹۳/۱۵ - ۳۱۸/۱۵	۰/۱ - ۱۰۰	۶۶	[۳۰]
<i>n</i> -پروپان	۱۰۰/۰۰ - ۳۶۵/۰۰	۰/۱ - ۱۰۰	۴۰	[۳۱]
<i>n</i> -بوتان	۴۲۵/۰۰ - ۱۴۰/۰۰	۰/۱ - ۶۵	۴۲	[۳۱]
<i>n</i> -پنتان	۲۹۸/۱۵ - ۳۴۸/۱۵	۰/۱ - ۴۰	۱۵	[۳۲]
<i>n</i> -هگزان	۲۹۳/۱۵ - ۳۷۳/۱۵	۰/۱ - ۱۵۰	۸۰	[۳۳]
<i>n</i> -هپتان	۲۹۸/۱۵ - ۳۴۸/۱۵	۰/۱ - ۴۰	۱۵	[۳۲]
<i>n</i> -اکتان	۲۹۸/۰۳ - ۳۴۸/۱۴	۰/۱ - ۴۶۰	۲۸	[۳۱]
<i>n</i> -دودکان	۲۹۳/۱۵ - ۴۳۳/۱۵	۰/۱ - ۱۴۰	۱۱۲	[۳۴]
<i>n</i> -تری‌دکان	۳۱۳/۱۵ - ۳۷۳/۱۵	۱۰ - ۱۰۰	۷۰	[۳۵]
<i>n</i> -پنتادکان	۲۹۳/۱۵ - ۳۸۳/۱۵	۰/۱ - ۱۵۰	۸۴	[۳۶]
<i>n</i> -هپتادکان	۳۱۳/۱۵ - ۳۸۳/۱۵	۰/۱ - ۱۵۰	۷۳	[۳۶]
<i>n</i> -اکتادکان	۳۱۳/۱۵ - ۳۸۳/۱۵	۰/۱ - ۱۵۰	۸۴	[۳۷]
<i>n</i> -نونادکان	۳۱۳/۱۵ - ۳۸۳/۱۵	۰/۱ - ۱۵۰	۸۲	[۳۷]
<i>n</i> -تری‌کوزان	۳۳۳/۱۵ - ۳۹۳/۱۵	۰/۱ - ۱۵۰	۹۷	[۳۸]
<i>n</i> -تتراکوزان	۳۳۳/۱۵ - ۳۹۳/۱۵	۰/۱ - ۱۵۰	۹۲	[۳۸]
<i>n</i> -اکتاکوزان	۳۵۳/۱۵ - ۴۰۳/۱۵	۰/۱ - ۱۵۰	۸۴	[۳۹]
<i>n</i> -هگزاتری‌اکوتان	۳۶۳/۱۵ - ۴۰۳/۱۵	۰/۱ - ۱۵۰	۶۴	[۳۹]



### ۳-۱-۳- محاسبه‌ی توصیف‌کننده‌ها

ساختارهای بهینه شده توسط نرم‌افزار HyperChem به عنوان ورودی به نرم‌افزار Dragon وارد شد و ۱۴۸۱ توصیف‌کننده‌ی ساختاری توسط نرم‌افزار Dragon محاسبه گردید. همچنین ۲۱ توصیف‌کننده‌ی الکترونی، توسط نرم‌افزار HyperChem محاسبه شد. دو متغیر تجربی دما و فشار هم که داده‌های  $K_T$ ، به آن‌ها وابسته اند نیز به آن‌ها اضافه گردید. به این ترتیب ۱۵۰۴ توصیف‌کننده برای هر ترکیب به دست آمد.

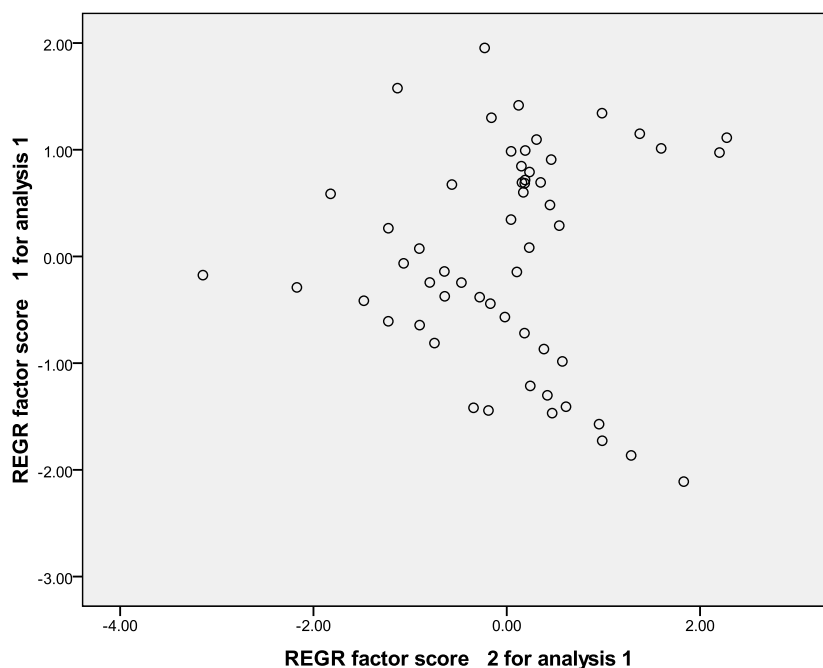
### ۳-۱-۴- حذف توصیف‌کننده‌های نامناسب

با توجه به اینکه تعداد زیاد توصیف‌کننده باعث پیچیدگی محاسبات شده و تعدادی از توصیف‌کننده‌ها حاوی اطلاعات یکسانی هستند، در این مرحله از روش‌های کاهش متغیر استفاده شد. به همین منظور، ابتدا توصیف‌کننده‌هایی که برای تمام مولکول‌ها، مقادیر ثابت و یا تقریباً ثابت داشتند، حذف شدند. به وسیله‌ی نرم‌افزار MATLAB و برنامه‌ی remcol توصیف‌کننده‌هایی که همبستگی بزرگ‌تر از ۰/۹ داشتند، نیز حذف گردید. در پایان این مرحله، ۸۴ توصیف‌کننده از مجموع ۱۵۰۴ توصیف‌کننده باقی ماند.

### ۳-۱-۵- دسته‌بندی داده‌ها

بعد از مشخص شدن توصیف‌کننده‌های مناسب، ابتدا داده‌ها را به سه دسته‌ی آموزش، ارزیابی و تست تقسیم گردید و رگرسیون مرحله‌ای توسط نرم‌افزار SPSS بر روی داده‌های سری آموزش انجام شد. همانطور که در فصل قبل گفته شد برای دسته‌بندی داده‌ها از روش PCA استفاده شد. به این ترتیب که ابتدا توسط نرم‌افزار SPSS از طریق منوی آنالیز مولفه‌های اساسی برای مجموعه‌ی توصیف‌کننده‌ها به دست آمد، سپس با توجه به اینکه دو مولفه‌ی اول مقدار قابل توجهی از پراکندگی

داده‌ها را در نظر می‌گیرد، نموداری از مولفه‌ی اول بر حسب مولفه‌ی دوم ترسیم شد. بر اساس این نمودار نقطه داده‌های محیطی در سری آموزش قرار گرفت و سری ارزیابی و تست از داده‌های داخلی به طور تصادفی انتخاب شدند. شکل (۱-۳) پراکندگی داده‌ها را بر روی نمودار PCA نشان می‌دهد.



شکل (۱-۳) - پراکندگی مولکول‌ها بر روی نمودار PCA

بر اساس شکل (۱-۳) ۶۵۹ داده مربوط به ۸ مولکول در سری ارزیابی، ۶۲۹ داده مربوط به ۷ مولکول در سری تست و ۳۰۰۹ داده باقی‌مانده مربوط به ۴۱ مولکول در دسته‌ی آموزش قرار گرفت.

### ۳-۱-۶- انتخاب بهترین توصیف‌کننده‌ها

حال در این مرحله، ۸۴ توصیف‌کننده‌ی باقی‌مانده یک بار با روش رگرسیون مرحله‌ای و بار دیگر با استفاده از راهبرد CDFS مورد آنالیز قرار گرفت تا بهترین توصیف‌کننده‌ها از میان آن‌ها انتخاب شود.

### ۳-۱-۶-۱- انتخاب بهترین توصیف‌کننده‌ها به روش رگرسیون مرحله‌ای

برای انتخاب توصیف‌کننده‌های مناسب به روش انتخاب مرحله‌ای، از داده‌های سری آموزش توسط نرم‌افزار SPSS رگرسیون مرحله‌ای گرفته شد. برای این منظور از منوی آنالیز در نرم‌افزار SPSS، گزینه‌ی رگرسیون خطی انتخاب گردید و بعد از وارد کردن متغیرهای مستقل و وابسته، از روش مرحله‌ای برای مدل‌سازی استفاده شد. برای انتخاب بهترین توصیف‌کننده‌ها، مدل‌های مختلفی که توسط رگرسیون مرحله‌ای به دست آمده، توسط سری ارزیابی مورد بررسی قرار گرفتند. میانگین مربعات خطا (MSE) و ضریب رگرسیون ( $R^2$ ) محاسبه شده برای سری ارزیابی هر مدل، در جدول (۲-۳) آمده است.

با توجه به پارامترهای آماری به دست آمده در این جدول مدل ۱۱ به عنوان مدل بهینه انتخاب گردید. نام کامل توصیف‌کننده‌های وارد شده در مدل بهینه به همراه ضریب رگرسیون و اثر متوسط آن‌ها در جدول (۳-۳) آمده است. همچنین ماتریس همبستگی بین این توصیف‌کننده‌ها در جدول (۴-۳) ارائه شده که عدم همبستگی بین آن‌ها را نشان می‌دهد.

مقادیر توصیف‌کننده‌های به دست آمده از روش رگرسیون مرحله‌ای برای مولکول‌های مختلف موجود در سری داده‌ها در جدول (۵-۳) ذکر شده است.

جدول (۳-۲) - مقایسه‌ی آماره‌های سری ارزیابی مدل‌های به دست آمده از روش رگرسیون مرحله‌ای

مدل	توصیف‌کننده‌ها	R <sup>2</sup>	MSE
۳	P, T, Mor13m	۰/۷۴۶	۱/۵۵۴۸
۴	P, T, Mor13m, X0Av	۰/۶۹۴	۱/۹۷۳۸
۵	P, T, Mor13m, X0Av, Mor07m	۰/۷۶۶	۱/۵۱۴۳
۶	P, T, Mor13m, X0Av, Mor07m, Mor11m	۰/۷۷۱	۱/۴۲۲۴
۷	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e	۰/۷۸۲	۱/۳۴۸۶
۸	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e	۰/۷۸۹	۱/۳۰۵۱
۹	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m	۰/۷۹۲	۱/۳۱۲۸
۱۰	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex	۰/۷۹۵	۱/۲۹۰۹
۱۱	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex, Hy	۰/۷۹۵	۱/۲۶۷۲
۱۲	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex, Hy, PJI3	۰/۷۸۴	۱/۴۵۸۳
۱۳	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex, Hy, PJI3, MW	۰/۷۷۶	۱/۶۰۵۳
۱۴	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex, Hy, PJI3, MW, Mor23m	۰/۷۶۱	۱/۷۳۹۵
۱۵	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Yindex, Hy, PJI3, MW, Mor23m	۰/۷۶۱	۱/۷۸۱۶
۱۶	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Yindex, Hy, PJI3, MW, Mor23m, E2s	۰/۷۶۲	۱/۵۴۷۷
۱۷	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Yindex, Hy, PJI3, MW, Mor23m, E2s, E2m	۰/۷۶۱	۱/۶۱۴۹
۱۸	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Yindex, Hy, PJI3, MW, Mor23m, E2s, E2m, X1A	۰/۷۵۴	۱/۶۳۷۸
۱۹	P, T, Mor13m, Mor07m, Mor11m, GATS1e, Mor10e, Yindex, Hy, PJI3, MW, Mor23m, E2s, E2m, X1Av	۰/۷۴۰	۱/۷۵۹۲

جدول (۳-۳) - توصیف‌کننده‌های انتخاب شده توسط روش رگرسیون مرحله‌ای

شماره	علامت	اثر متوسط	گروه	نام کامل
۱	P	-۶۴/۷۴۴	experimental	Pressure
۲	T	۴۵/۹۰۰	experimental	Temperature
۳	Mor13m	۲۵/۴۸۸	3D MORSE	signal13/weighted by atomic masses
۴	X0Av	۹/۷۱۸	Topological	average valence connectivity index chi-0
۵	Mor07m	۱۱/۴۹۷	3D MORSE	signal7/weighted by atomic masses
۶	Mor11m	-۱۲/۱۷۴	3D MORSE	signal11/weighted by atomic masses
۷	GATS1e	۱۲/۵۸۴	2D Autocorrelation	lag 1 / weighted by atomic Sanderson electronegativities
۸	Mor10e	-۸/۵۶۴	3D MORSE	signal 10 / weighted by atomic Sanderson electronegativities
۹	Mor26m	۵/۷۷۳	3D MORSE	signal 26/weighted by atomic masses
۱۰	Yindex	۵/۳۸۶	Topological	Balaban Y index
۱۱	Hy	-۳/۲۷۸	empirical	hydrophilic factor 1

جدول (۴-۳) - ماتریس همبستگی توصیف‌کننده‌های انتخاب شده توسط رگرسیون مرحله‌ای

	P	T	Mor13m	X0Av	Mor07m	Mor11m	GATS1e	Mor10e	Mor26m	Yindex	Hy
P	۱										
T	۰/۰۵۴	۱									
Mor13m	-۰/۰۴۵	-۰/۵۰۵	۱								
X0Av	۰/۱۵۷	۰/۰۵۷	-۰/۱۷۷	۱							
Mor07m	۰/۱۴۸	۰/۲۴۸	-۰/۶۴۴	-۰/۳۷۱	۱						
Mor11m	۰/۰۰۸	۰/۳۸۴	-۰/۲۶۷	۰/۲۵۶	-۰/۰۹۴	۱					
GATS1e	۰/۰۰۶	۰/۳۲۸	-۰/۴۱۶	۰/۶۵۶	-۰/۱۶۰	۰/۵۶۴	۱				
Mor10e	-۰/۲۳۳	-۰/۴۲۴	۰/۸۰۵	-۰/۱۵۰	-۰/۷۵۱	-۰/۲۴۱	-۰/۲۵۶	۱			
Mor26m	-۰/۰۳۱	-۰/۲۳۹	۰/۶۷۸	۰/۲۰۸	-۰/۷۹۹	۰/۲۳۳	-۰/۰۷۳	۰/۶۷۹	۱		
Yindex	۰/۰۷۴	-۰/۱۵۰	۰/۲۸۱	۰/۰۷۶	-۰/۲۶۵	-۰/۴۵۱	-۰/۳۵۱	۰/۳۰۲	۰/۲۷۲	۱	
Hy	-۰/۱۵۸	-۰/۱۰۹	۰/۰۶۵	-۰/۱۱۸	-۰/۲۱۴	-۰/۳۸۹	-۰/۱۶۴	۰/۳۸۰	۰/۲۰۰	۰/۵۴۱	۱

جدول (۳-۵) - مقادیر توصیف‌کننده‌های انتخاب شده به روش رگرسیون مرحله‌ای

ردیف	نام مولکول	نام توصیف‌کننده								
		Mor13m	X0Av	Mor07m	Mor11m	GATS1e	Mor10e	Mor26m	Yindex	Hy
۱	استیک اسید	۰/۰۰۹	۰/۵۸۹	۰/۶۳۳	۰/۰۱۴	۱/۸۲۱	-۰/۳۴۳	-۰/۰۷۸	۰/۰۰	۰/۵۸۶
۲	پروپانویک اسید	-۰/۰۴۳	۰/۶۱۳	۰/۶۱۷	۰/۰۰۲	۱/۶۸۴	-۰/۳۳۲	-۰/۱۰۲	۳/۹۲	۰/۳۴۶
۳	بوتانویک اسید	-۰/۰۱۶	۰/۶۲۸	۰/۶۶۳	۰/۰۰۷	۱/۶۲۲	-۰/۲۸۶	-۰/۰۸۸	۲/۶۴	۰/۱۸۴
۴	پنتانویک اسید	-۰/۰۵۶	۰/۶۴۰	۰/۶۶۸	۰/۰۴۵	۱/۵۸۹	-۰/۳۹۴	-۰/۱۱۲	۲/۱۱	۰/۰۶۷
۵	هگزانویک اسید	-۰/۰۹۶	۰/۶۴۸	۰/۷۰۷	۰/۰۶۳	۱/۵۷۱	-۰/۵۴۲	-۰/۱۱۶	۱/۸۱	-۰/۰۲۲
۶	هپتانویک اسید	-۰/۱۳۴	۰/۶۵۵	۰/۸۰۳	۰/۱۰۸	۱/۵۶۱	-۰/۵۵۳	-۰/۱۳۱	۱/۶۲	-۰/۰۹۲
۷	اکتانویک اسید	-۰/۱۶۶	۰/۶۶۰	۰/۸۵۱	۰/۱۴۱	۱/۵۵۷	-۰/۵۲۱	-۰/۱۲۶	۱/۴۹	-۰/۱۴۸
۸	پروپیل استات	-۰/۱۴۷	۰/۶۷۶	۰/۹۱۷	-۰/۰۰۸	۱/۴۴۱	-۱/۰۴۶	-۰/۰۷۸	۲/۱۱	-۰/۶۶۸
۹	بوتیل استات	-۰/۱۶۳	۰/۶۸۰	۰/۹۱۱	۰/۰۰۵	۱/۴۲۶	-۱/۱۱۹	-۰/۰۷۳	۱/۸۱	-۰/۷۱۰
۱۰	اتیل پروپیونات	-۰/۱۴۵	۰/۶۷۶	۰/۸۷۵	-۰/۰۱۶	۱/۴۴۱	-۱/۰۷۴	-۰/۰۶۹	۲/۴۶	-۰/۶۶۸
۱۱	اتیل بوتیرات	-۰/۱۳۸	۰/۶۸۰	۰/۸۸۵	-۰/۰۲۶	۱/۴۲۶	-۱/۰۱۴	-۰/۰۵۸	۲/۱۷	-۰/۷۱۰
۱۲	متیل اتیل استات	-۰/۱۶۵	۰/۶۹۹	۱/۱۵۶	۰/۰۱۰	۱/۴۴۱	-۰/۷۹۱	-۰/۰۹۲	۲/۶۷	-۰/۶۶۸
۱۳	۲-بوتانول	-۰/۰۳۱	۰/۷۲۳	۰/۴۶۶	۰/۰۰۳	۱/۰۰۸	-۰/۶۳۲	-۰/۰۲۳	۳/۹۲	-۰/۷۱۹
۱۴	۲-پنتانول	-۰/۰۴۲	۰/۷۲۰	۰/۵۰۹	۰/۰۱۹	۱/۰۲۴	-۰/۸۲۰	-۰/۰۳۳	۲/۶۴	-۰/۷۶۷
۱۵	۳-پنتانول	-۰/۰۴۰	۰/۷۲۰	۰/۳۹۰	۰/۰۹۷	۱/۰۲۴	-۰/۸۳۸	-۰/۰۳۶	۳/۱۴	-۰/۷۶۷
۱۶	۲-هگزانول	-۰/۰۹۸	۰/۷۱۹	۰/۵۱۴	۰/۱۰۶	۱/۰۴۴	-۰/۷۵۷	-۰/۰۲۹	۲/۱۱	-۰/۸۰۲
۱۷	۴-متیل-۲-پنتانول	-۰/۱۱۸	۰/۷۴۲	۰/۷۰۶	۰/۰۲۹	۱/۰۴۴	-۰/۸۶۲	۰/۰۰۳	۲/۶۷	-۰/۸۰۲
۱۸	۱-پروپانول	-۰/۱۱۸	۰/۷۱۵	۰/۳۵۷	۰/۰۶۴	۲/۱۰۷	-۰/۳۴۳	-۰/۰۳۳	۳/۱۳	۰/۳۷۱
۱۹	۱-بوتانول	-۰/۱۲۳	۰/۷۱۴	۰/۳۶۴	۰/۰۷۸	۲/۰۹۳	-۰/۳۶۲	-۰/۰۲۸	۲/۳۴	۰/۱۶۷

ادامه‌ی جدول (۳-۵)

ردیف	نام مولکول	نام توصیف کننده								
		Mor13m	X0Av	Mor07m	Mor11m	GATS1e	Mor10e	Mor26m	Yindex	Hy
۲۰	۱-پنتانول	-۰/۱۶۸	۰/۷۱۳	۰/۳۴۷	۰/۱۰۶	۲/۰۸۸	-۰/۴۵۰	-۰/۰۳۳	۱/۹۶	۰/۰۳۱
۲۱	۲-متیل-۱-پنتانول	-۰/۲۴۰	۰/۷۴۰	۰/۴۱۲	۰/۰۳۹	۲/۰۸۸	-۰/۲۰۶	-۰/۰۳۰	۳/۱۴	۰/۰۳۱
۲۲	۳-پنتانول	-۰/۲۶۳	۰/۷۴۰	۰/۳۸۹	۰/۰۷۳	۲/۰۸۸	-۰/۷۳۶	۰/۰۰۸	۳/۱۴	۰/۰۳۱
۲۳	۳-متیل-۳-پنتانول	-۰/۳۵۵	۰/۷۶۶	۰/۴۵۴	۰/۰۵۲	۲/۰۸۹	-۱/۰۸۰	-۰/۰۰۵	۴/۱۰	۱/۹۸۴
۲۴	۳-اتیل-۳-پنتانول	-۰/۳۶۲	۰/۷۵۹	۰/۵۵۶	۰/۱۴۰	۲/۰۹۲	-۰/۵۸۱	۰/۰۳۰	۳/۷۵	۱/۷۹۴
۲۵	۱-هگزانول	-۰/۱۹۲	۰/۷۱۲	۰/۳۷۵	۰/۱۳۳	۲/۰۸۹	-۰/۵۰۵	-۰/۰۲۸	۱/۷۳	-۰/۰۶۷
۲۶	۲-هگزانول	-۰/۲۵۸	۰/۷۳۵	۰/۵۹۲	۰/۱۷۸	۲/۰۸۹	-۰/۷۶۵	-۰/۰۲۰	۲/۱۱	-۰/۰۶۷
۲۷	۱-هپتانول	-۰/۲۱۹	۰/۷۱۱	۰/۴۵۰	۰/۱۶۹	۲/۰۹۲	-۰/۵۶۶	-۰/۰۲۶	۱/۵۷	-۰/۱۴۰
۲۸	۲-هپتانول	-۰/۲۷۹	۰/۷۳۲	۰/۵۸۲	۰/۱۷۱	۲/۰۹۲	-۰/۶۸۸	-۰/۰۲۹	۱/۸۱	-۰/۱۴۰
۲۹	۱-اکتانول	-۰/۲۵۵	۰/۷۱۱	۰/۵۲۳	۰/۱۹۹	۲/۰۹۸	-۰/۵۶۳	-۰/۰۳۱	۱/۴۶	-۰/۱۹۸
۳۰	۱-نونانول	-۰/۲۶۷	۰/۷۱۰	۰/۵۵۸	۰/۲۳۱	۲/۱۰۵	-۰/۶۰۳	-۰/۰۳۵	۱/۳۷	-۰/۲۴۴
۳۱	۱-دکانول	-۰/۲۹۵	۰/۷۱۰	۰/۵۶۹	۰/۲۶۷	۲/۱۱۳	-۰/۷۰۲	-۰/۰۴۰	۱/۳۰	-۰/۲۸۳
۳۲	۱-دودکانول	-۰/۳۴۲	۰/۷۱۰	۰/۶۵۶	۰/۳۲۴	۲/۱۲۹	-۰/۷۵۸	-۰/۰۵۸	۱/۱۹	-۰/۳۴۳
۳۳	۱-تترادکانول	-۰/۳۹۸	۰/۷۰۹	۰/۷۶۳	۰/۳۷۱	۲/۱۴۷	-۰/۸۶۵	-۰/۰۸۶	۱/۱۱	-۰/۳۸۸
۳۴	۱-هگزادکانول	-۰/۴۵۴	۰/۷۰۹	۰/۸۱۴	۰/۳۹۸	۲/۱۶۵	-۰/۹۴۶	-۰/۱۲۶	۱/۰۵	-۰/۴۲۳
۳۵	۲-متیل-۲و۴-پنتان دی‌ال	-۰/۳۸۱	۰/۷۱۰	۰/۷۹۲	۰/۰۴۱	۲/۱۲۰	-۰/۷۷۳	-۰/۰۳۷	۳/۰۱	۳/۰۶۶
۳۶	۲-آمینو بوتان	-۰/۲۲۳	۰/۷۷۲	۰/۳۱۲	۰/۱۰۱	۳/۱۲۴	-۰/۴۷۳	-۰/۰۰۵	۳/۹۲	۱/۲۷۹
۳۷	پنتیل آمین	-۰/۱۹۲	۰/۷۳۴	۰/۲۶۷	۰/۱۹۵	۳/۱۰۶	-۰/۲۸۴	-۰/۰۴۹	۱/۹۶	۱/۰۴۸
۳۸	هگزیل آمین	-۰/۲۰۷	۰/۷۳۰	۰/۳۱۵	۰/۲۳۶	۳/۰۹۳	-۰/۲۸۹	-۰/۰۵۷	۱/۷۳	۰/۸۸۰

ادامه‌ی جدول (۳-۵)

ردیف	نام مولکول	نام توصیف کننده								
		Mor13m	X0Av	Mor07m	Mor11m	GATS1e	Mor10e	Mor26m	Yindex	Hy
۳۹	هپتیل آمین	-۰/۲۳۷	۰/۷۲۷	۰/۳۸۳	۰/۲۸۵	۳/۰۸۳	-۰/۴۶۳	-۰/۰۶۵	۱/۵۷	۰/۷۵۲
۴۰	۲-آمینو اکتان	-۰/۳۲۹	۰/۷۴۳	۰/۵۳۶	۰/۲۸۳	۳/۰۷۶	-۰/۶۴۳	-۰/۰۲۸	۱/۶۲	۰/۶۵۱
۴۱	<i>n</i> -پروپان	-۰/۰۳۸	۰/۹۰۲	۰/۱۸۱	۰/۰۲۹	۳/۶۶۷	-۰/۴۴۲	-۰/۰۰۷	۰/۰۰	-۰/۷۹۲
۴۲	<i>n</i> -بوتان	-۰/۰۶۵	۰/۸۵۴	۰/۲۳۱	۰/۰۷۳	۳/۵۰۰	-۰/۶۲۲	-۰/۰۰۵	۳/۱۳	-۰/۸۶۱
۴۳	<i>n</i> -پنتان	-۰/۰۹۷	۰/۸۲۴	۰/۲۴۰	۰/۱۱۷	۳/۴۰۰	-۰/۷۴۷	-۰/۰۱۱	۲/۳۶	-۰/۸۹۸
۴۴	<i>n</i> -هگزان	-۰/۱۳۴	۰/۸۰۵	۰/۲۵۴	۰/۱۷۲	۳/۳۳۳	-۰/۷۰۶	-۰/۰۱۰	۱/۹۶	-۰/۹۲۱
۴۵	<i>n</i> -هپتان	-۰/۱۵۵	۰/۷۹۱	۰/۳۰۲	۰/۲۲۱	۳/۲۸۶	-۰/۶۴۸	-۰/۰۱۶	۱/۷۳	-۰/۹۳۶
۴۶	<i>n</i> -اکتان	-۰/۱۸۱	۰/۷۸۰	۰/۳۶۷	۰/۲۷۳	۳/۲۵۰	-۰/۷۴۵	-۰/۰۲۱	۱/۵۷	-۰/۹۴۶
۴۷	<i>n</i> -دودکان	-۰/۱۸۳	۰/۷۵۶	۰/۲۸۷	۰/۸۸۸	۳/۱۶۷	-۰/۶۹۰	۰/۲۳۵	۱/۲۴	-۰/۹۶۹
۴۸	<i>n</i> -تری دکان	-۰/۱۹۵	۰/۷۵۲	۰/۳۰۶	۱/۰۳۷	۳/۱۵۴	-۰/۶۲۳	۰/۲۴۶	۱/۱۹	-۰/۹۷۲
۴۹	<i>n</i> -پنتادکان	-۰/۳۵۳	۰/۷۴۶	۰/۶۵۹	۰/۵۳۸	۳/۱۳۳	-۱/۱۱۴	-۰/۱۱۶	۱/۱۱	-۰/۹۷۷
۵۰	<i>n</i> -هپتادکان	-۰/۴۰۲	۰/۷۴۲	۰/۷۲۹	۰/۵۷۱	۳/۱۱۸	-۱/۱۳۳	-۰/۱۶۳	۱/۰۵	-۰/۹۸۰
۵۱	<i>n</i> -اکتادکان	-۰/۴۲۶	۰/۷۴۰	۰/۷۷۸	۰/۵۸۱	۳/۱۱۱	-۱/۲۴۸	-۰/۱۸۹	۰/۰۳	-۰/۹۸۲
۵۲	<i>n</i> -نونادکان	-۰/۴۵۳	۰/۷۳۸	۰/۸۲۸	۰/۵۸۴	۳/۱۰۵	-۱/۳۱۳	-۰/۲۱۵	۱/۱۱	-۰/۹۸۳
۵۳	<i>n</i> -تری کوزان	-۰/۵۵۱	۰/۷۳۳	۰/۹۸۹	۰/۵۴۶	۳/۰۸۷	-۱/۴۷۱	-۰/۳۲۵	۰/۹۳	-۰/۹۸۷
۵۴	<i>n</i> -تتراکوزان	-۰/۵۸۰	۰/۷۳۲	۱/۰۳۶	۰/۵۲۸	۳/۰۸۳	-۱/۴۷۹	-۰/۳۵۲	۰/۹۱	-۰/۹۸۷
۵۵	<i>n</i> -اکتاکوزان	-۰/۶۷۷	۰/۷۲۸	۱/۲۰۱	۰/۳۹۴	۳/۰۷۱	-۱/۶۷۴	-۰/۴۴۲	۰/۸۶	-۰/۹۹۰
۵۶	<i>n</i> -هگزاتری آکوتان	-۰/۸۷۸	۰/۷۲۳	۱/۵۳۸	-۰/۰۲۴	۳/۰۵۶	-۱/۹۲۵	-۰/۵۷۰	۰/۷۹	-۰/۹۹۲



### ۳-۱-۶-۲- انتخاب بهترین توصیف‌کننده‌ها با استفاده از راهبرد CDFS

همانطور که اشاره شد در انتخاب توصیف‌کننده به روش رگرسیون مرحله‌ای، امکان ایجاد مدل شانس برای یک دسته‌بندی وجود دارد. البته به دلیل اینکه در این تحقیق تعداد داده‌ها بسیار بیشتر از تعداد توصیف‌کننده‌هاست، احتمال ساخت مدل تصادفی کاهش می‌یابد. اما برای اطمینان بیشتر و مقایسه روش‌ها، راهبرد CDFS به کار گرفته شد. برای انجام این کار ابتدا ۷ مولکولی که به عنوان سری تست انتخاب شده بود کنار گذاشته شدند و داده‌های باقی‌مانده ۱۶ بار به روش‌های متفاوت به دو سری آموزش و ارزیابی تقسیم شدند. سری آموزش مربوط به هر یک از ۱۶ دسته، توسط رگرسیون مرحله‌ای آنالیز گردید و مدل‌های مختلفی به دست آمد سپس مدل‌های به دست آمده برای هر دسته توسط سری ارزیابی همان دسته، مورد بررسی قرار گرفت تا مدل بهینه برای آن دسته‌بندی به دست آید. بنابراین در انتها، ۱۶ مدل متفاوت برای سری داده‌ها ایجاد شد. توصیف‌کننده‌های انتخاب شده در ۱۶ مدل به دست آمده در جدول (۳-۶) ذکر شده است.

جدول (۳-۶) - توصیف‌کننده‌های ۱۶ مدل QSPR به دست آمده از دسته‌بندی‌های مختلف داده‌ها به همراه پارامترهای آماری به دست آمده برای هر دسته در روش CDFS

شماره	توصیف‌کننده‌ها	تعداد توصیف‌کننده
۱	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e, Mor26m, Yindex, Hy	۱۱
۲	P, T, MATS1p, HoF, X0Av, Ds, HATS1m	۷
۳	P, T, Mor13m, X0Av, Mor07m	۵
۴	P, T, MATS1p, HoF	۴
۵	P, T, MATS1p, HoF, G1s, Mor11m, Mor29m	۷
۶	P, T, Mor13m, X0Av, Mor07m	۵
۷	P, T, Mor13m, X0Av, Mor07m, R1m_A, Mor11m, GATS1e, X1A, Mor10e	۱۰
۸	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e	۸
۹	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e, Mor10e	۸
۱۰	P, T, MATS1p, HoF	۴
۱۱	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e	۷
۱۲	P, T, Mor13m, GATS1v, Mor10e, MATS1e, SPAM	۷
۱۳	P, T, Mor13m, X0Av, Mor07m, Mor11m, GATS1e	۷
۱۴	P, T, MATS1p, HoF, X0Av, Ds	۶
۱۵	P, T, Mor13m, X0Av, Mor07m, R1m_A, Mor11m, GATS1e, X1A	۹
۱۶	P, T, MATS1p, HoF, G1s, Mor11m, Mor29m, R3m	۸

همانطور که مشاهده می‌شود ۶ توصیف‌کننده از بین توصیف‌کننده‌های به دست آمده در مدل‌ها، بیش از ۵۰٪ تکرار دارند که نشان‌دهنده‌ی اهمیت این توصیف‌کننده‌هاست. این توصیف‌کننده‌ها در جدول (۳-۷) معرفی شده‌اند.

جدول (۷-۳) - توصیف‌کننده‌های انتخاب شده توسط روش CDFS

شماره	علامت	گروه	اثر متوسط	نام کامل
۱	T	Experimental	۴۹/۴۶۷	Temperature
۲	P	Experimental	-۶۲/۸۰۹	Pressure
۳	X0Av	Topological	۳۳/۶۲۹	average valence connectivity index chi-0
۴	Mor13m	3D MORSE	۳۶/۹۸۶	signal13/weighted by atomic masses
۵	Mor07m	3D MORSE	۱۹/۰۹۰	signal7/weighted by atomic masses
۶	Mor11m	3D MORSE	-۱۰/۹۲۱	signal11/weighted by atomic masses

با توجه به این که توصیف‌کننده‌های تئوری به دست آمده از روش CDFS مشابه ۴ توصیف‌کننده‌ی ابتدایی انتخاب شده به روش رگرسیون مرحله‌ای می‌باشند، بنابراین مقادیر این توصیف‌کننده‌ها نیز مطابق مقادیر ذکر شده در جدول (۷-۳) است.

### ۷-۱-۳- مدل‌سازی به روش شبکه‌ی عصبی مصنوعی

با توجه به پیچیدگی داده‌ها و احتمال وجود روابط غیرخطی بین متغیرهای مستقل و متغیر وابسته، روش غیرخطی شبکه‌ی عصبی مصنوعی به عنوان تکنیکی برتر در مطالعات QSPR برای مدل‌سازی ضریب تراکم‌پذیری همدمما مورد استفاده قرار گرفت. برای این منظور پارامترهای مؤثر بر آموزش موفق شبکه‌ی عصبی، از قبیل تعداد ورودی‌ها، تعداد گره‌های لایه‌ی پنهان، نوع تابع انتقال، نوع تابع آموزش، تعداد دوره‌های آموزش و پارامتر ممنتم،  $\mu$ ، بهینه شدند. در طی فرایند آموزش مقادیر وزن‌های ارتباطی تغییر نموده و معیار برای پایان دادن آموزش و بهینه شدن مقادیر وزن‌ها جهت

پیش‌بینی پارامتر مورد بررسی، مقدار خطای شبکه است که طبق تابع کارایی تعیین شده، محاسبه می‌شود. در این مطالعه، تابع کارایی، میانگین مربعات خطا (MSE) می‌باشد.

در ادامه، چگونگی مدل‌سازی شبکه‌ی عصبی مصنوعی با توصیف‌کننده‌های بهینه‌ی به دست آمده از روش رگرسیون مرحله‌ای و همچنین با توصیف‌کننده‌های حاصل از راهبرد CDFS توضیح داده خواهد شد.

### ۳-۱-۷-۱- بهینه‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده-

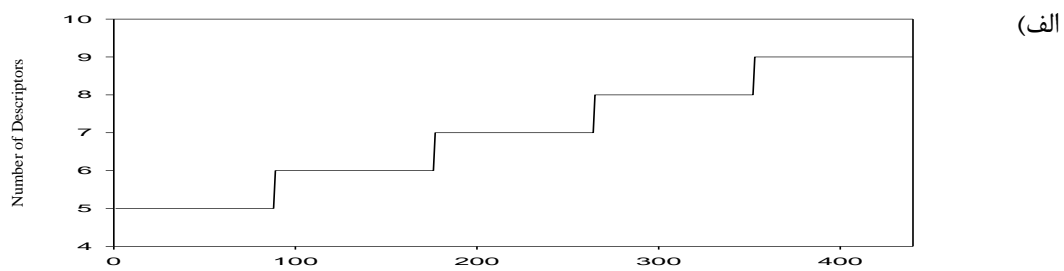
#### های رگرسیون مرحله‌ای

شبکه‌ی عصبی طراحی شده، یک شبکه‌ی سه لایه شامل یک لایه‌ی ورودی، یک لایه‌ی پنهان و یک لایه‌ی خروجی است که تعداد نرون‌های لایه‌ی ورودی برابر تعداد توصیف‌کننده‌ها بوده که باید تعداد توصیف‌کننده‌ها بهینه شود و لایه‌ی خروجی دارای یک نرون است که نشان دهنده‌ی ضریب تراکم‌پذیری همدمای متناظر با هر ورودی می‌باشد. تعداد نرون‌های لایه‌ی پنهان نیز نامعلوم است و باید بهینه شوند.

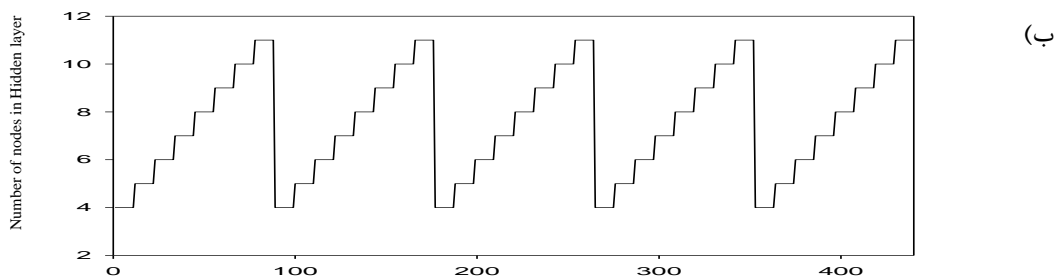
برای بهینه‌سازی تعداد ورودی‌های شبکه، نوع تابع آموزش، نوع تابع انتقال و تعداد گره‌های لایه‌ی پنهان، شبکه‌هایی با ورودی‌هایی از ۲ تا ۱۱ توصیف‌کننده ایجاد شد و با الگوریتم‌های آموزشی تنظیم بایزین و لونیگ-مارکوارت با تعداد متفاوت گره‌های لایه‌ی پنهان از ۲ تا ۱۲ گره، آموزش داده شد. برای انتخاب بهترین تابع انتقال در لایه‌ی پنهان نیز از توابع انتقال تانژانت سیگموئید (tansig) و لگاریتم سیگموئید (logsig) استفاده گردید.

با وجود اینکه در اکثر کارهای انجام شده‌ی قبلی، برای بهینه کردن تعداد ورودی‌ها و تعداد گره‌های لایه‌ی پنهان، ابتدا دور آموزشی ثابت در نظر گرفته شده و پس از بهینه‌سازی این پارامترها، تعداد دور آموزش بهینه می‌شد، اما اگر هر سه پارامتر به طور همزمان بهینه شوند منطقی‌تر است تا

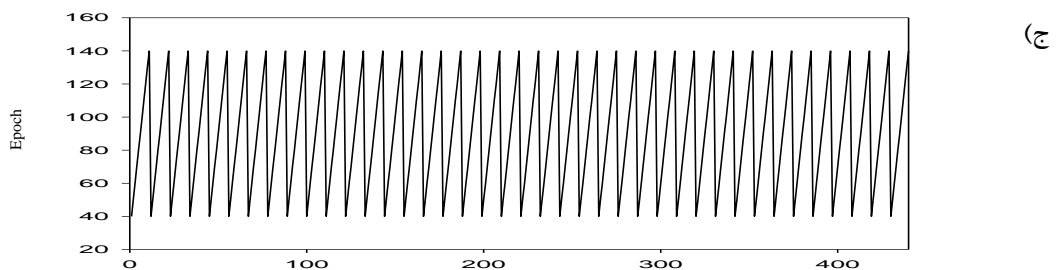
مقادیر بهینه‌ی دقیق‌تری به دست آید. بنابراین در تمامی شبکه‌های طراحی شده در این پایان‌نامه بطور همزمان با تغییر تعداد توصیف‌کننده‌ها و تعداد گره‌های لایه‌ی پنهان، تعداد دورهای آموزش نیز تغییر داده شد. برای این منظور همراه با تغییر تعداد ورودی از ۲ تا ۱۱ و تغییر تعداد گره‌های لایه‌ی پنهان از ۲ تا ۱۲، تعداد دورهای آموزش نیز از ۱۰ تا ۲۰۰ با گام ۲ در نظر گرفته شد و به ازای همه‌ی ترکیب‌های ممکن این سه پارامتر، شبکه‌های طراحی شده بهینه گردید. لازم به تذکر است که در لایه‌ی خروجی از تابع انتقال خطی purelin استفاده گردید. در روند بهینه‌سازی پارامترهای فوق، در هر یک از شبکه‌های طراحی شده به حداقل رساندن میانگین مربعات خطای حاصل از سری ارزیابی به عنوان معیار انتخاب شد. قسمتی از روند تغییرات پارامترهای شبکه در حین بهینه‌سازی همزمان پارامترها به همراه مقادیر MSE به دست آمده برای سری ارزیابی بر حسب یک بردار مرجع در شکل‌های (۳-۲) تا (۳-۵) نشان داده شده است. همچنین مقادیر بهینه به دست آمده برای هر یک از شبکه‌های طراحی شده در جدول (۳-۸) ذکر شده است.



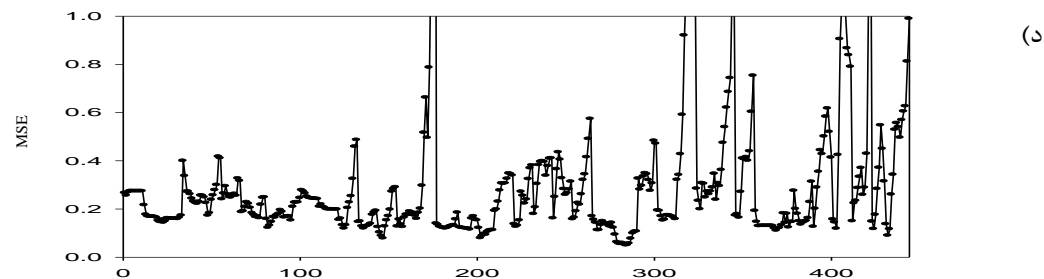
(الف)



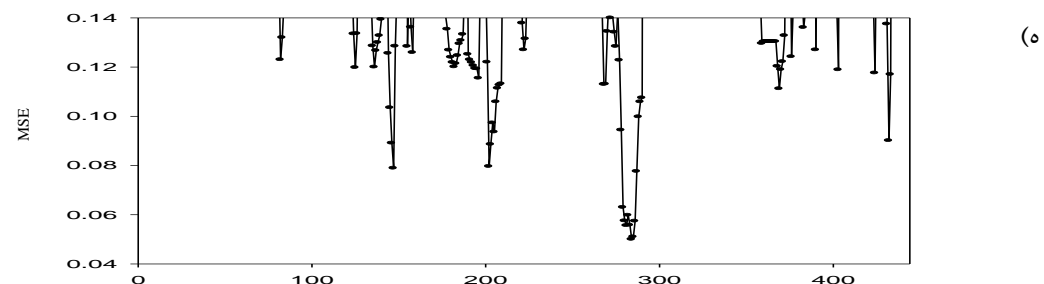
(ب)



(ج)



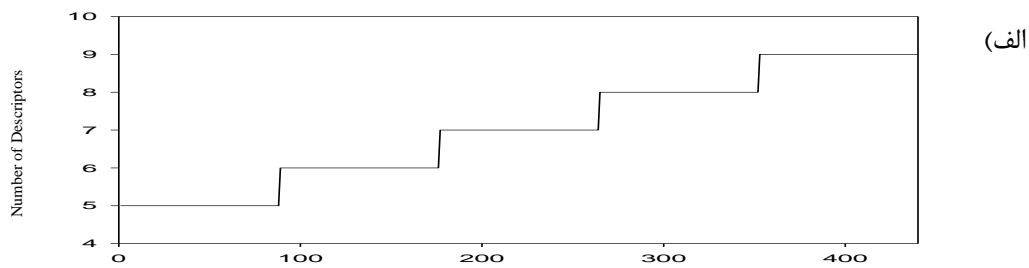
(د)



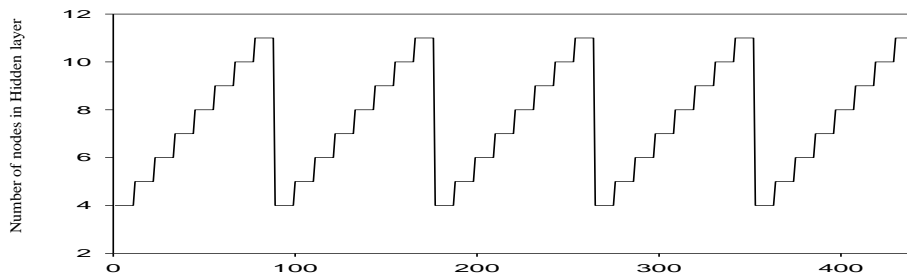
(ه)

بردار مرجع (شمارنده)

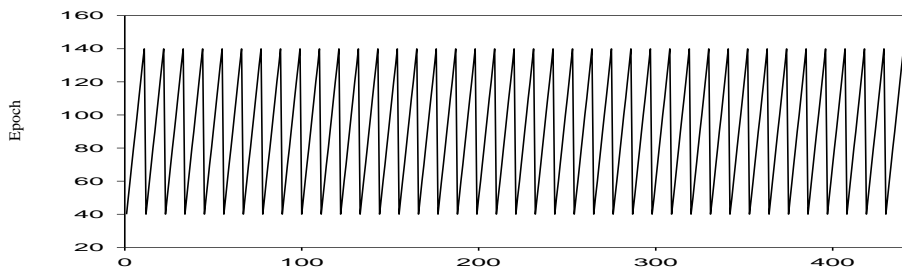
شکل (۳-۲) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی تنظیم بایزین



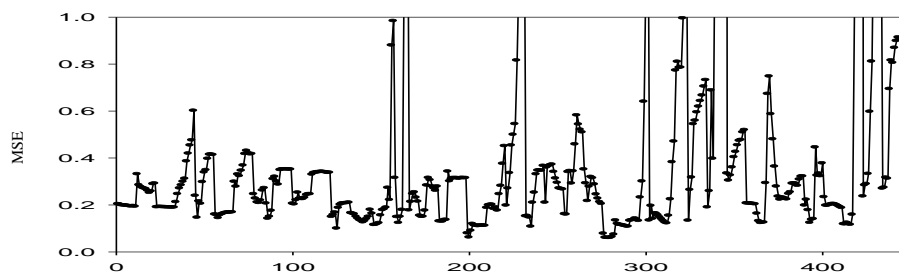
(الف)



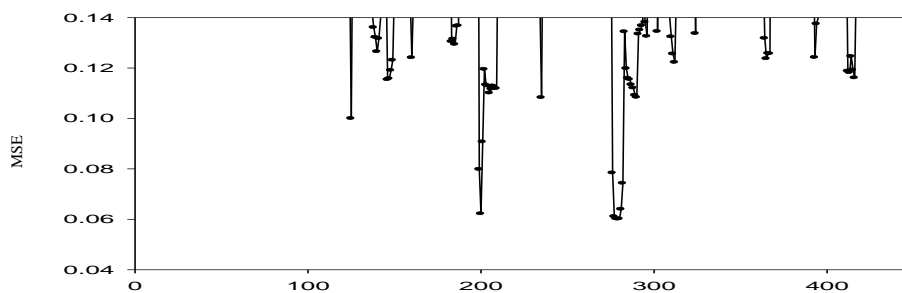
(ب)



(ج)



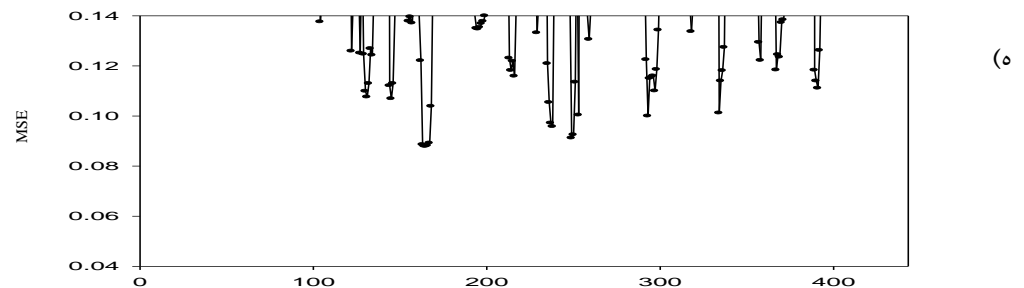
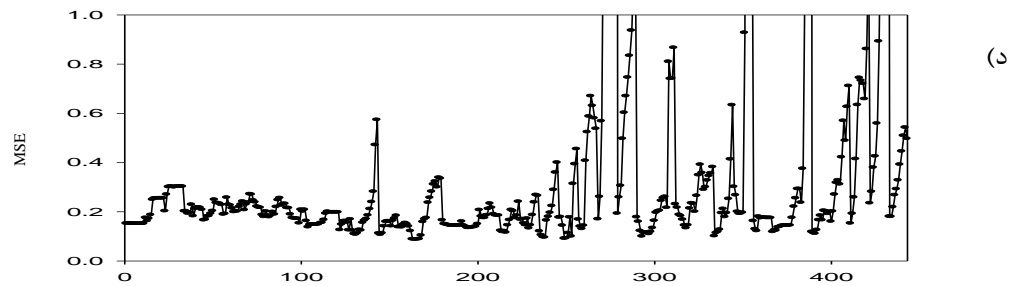
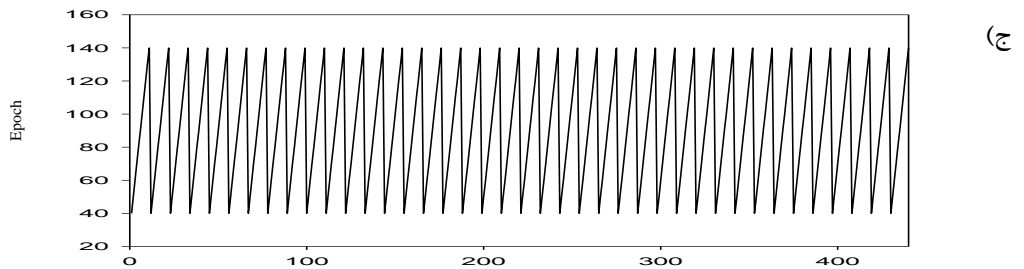
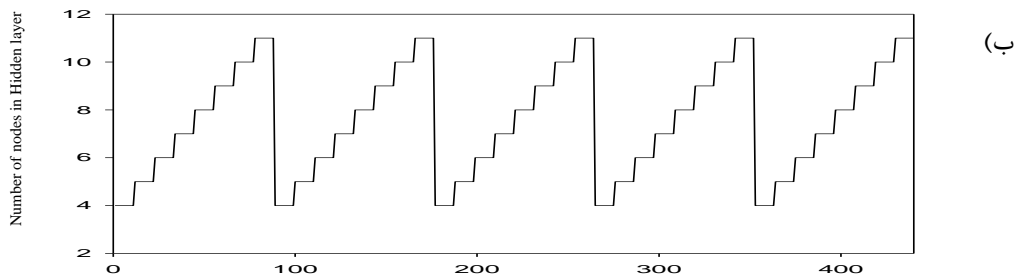
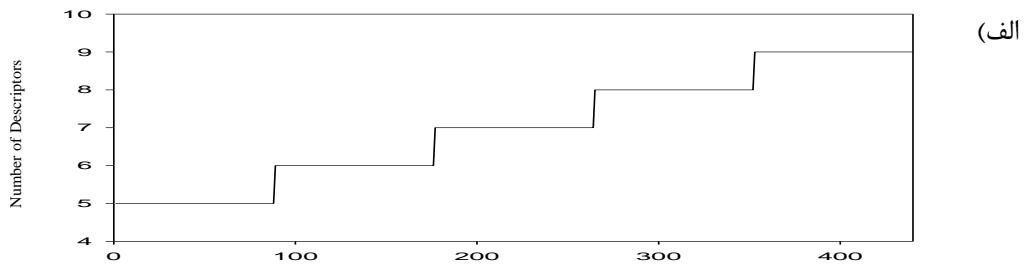
(د)



(ه)

بردار مرجع (شمارنده)

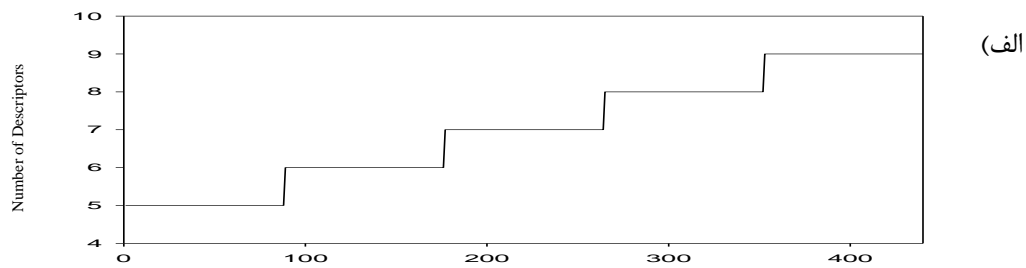
شکل (۳-۳) - نمودارهای (الف) تعداد توصیف‌کننده، (ب) تعداد نرون لایه‌ی مخفی، (ج) دور آموزش، (د) سری MSE ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال لگاریتم سیگموئیدی و الگوریتم آموزشی لوببرگ-مارکوارت



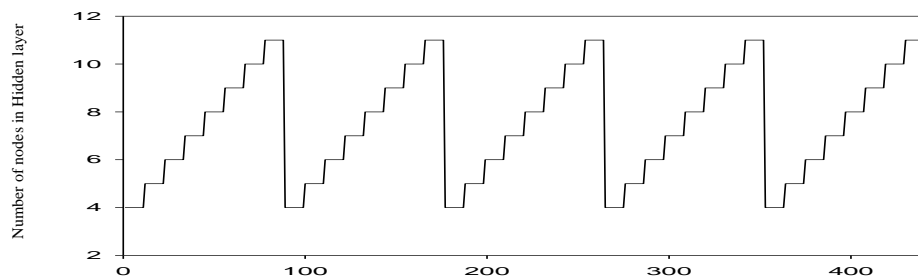
بردار مرجع (شمارنده)

شکل (۳-۴) - نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی تنظیم بایزین

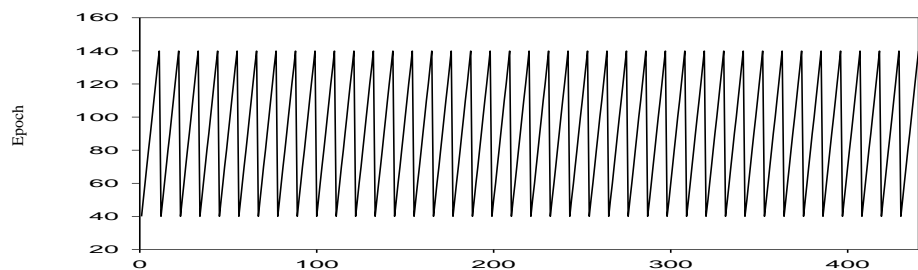




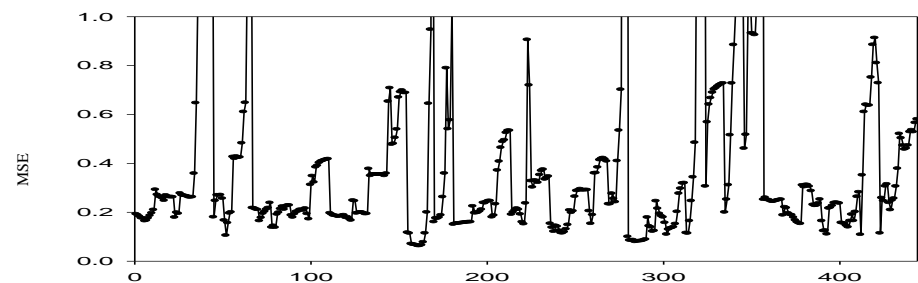
(الف)



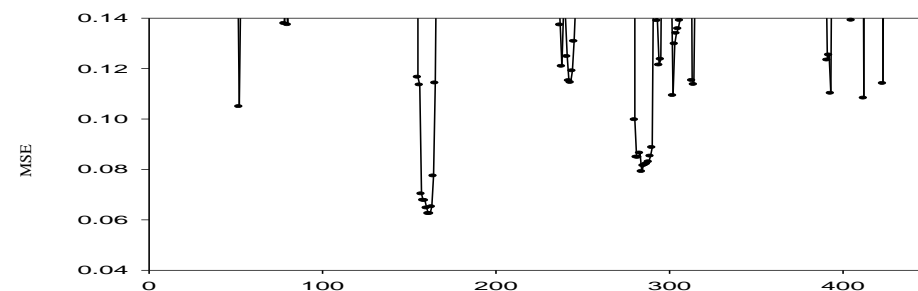
(ب)



(ج)



(د)



(ه)

بردار مرجع (شمارنده)

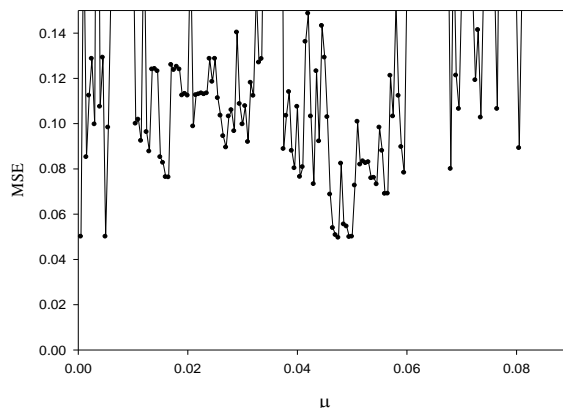
شکل (۳-۵) - نمودارهای (الف) تعداد توصیف کننده، (ب) تعداد نرون لایه مخفی، (ج) دور آموزش، (د) MSE سری ارزیابی و (ه) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن را بهتر نشان می‌دهد، بر حسب بردار مرجع برای تابع انتقال تانژانت سیگموئیدی و الگوریتم آموزشی لونیگ-مارکوارت

جدول (۳-۸) - توابع و پارامترهای بهینه‌ی شبکه‌های طراحی‌شده با توصیف‌کننده‌های حاصل از روش رگرسیون مرحله‌ای

MSE	تعداد دور آموزش	تعداد نرون لایه پنهان	تعداد توصیف‌کننده	تابع آموزش	تابع انتقال
۰/۰۵۰۰	۹۶	۵	۸	تنظیم بایزین	لگاریتم سیگموئید
۰/۰۶۰۱	۵۲	۵	۸	لونبرگ-مارکوارت	لگاریتم سیگموئید
۰/۰۸۷۹	۱۲۸	۱۰	۶	تنظیم بایزین	تانژانت سیگموئید
۰/۰۶۲۵	۸۲	۱۰	۶	لونبرگ-مارکوارت	تانژانت سیگموئید

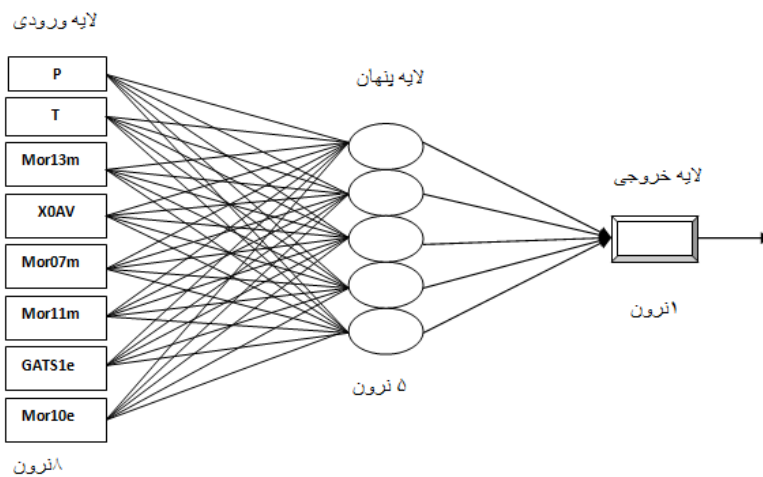
نتایج موجود در این شکل‌ها و جدول (۳-۸) نشان می‌دهد که شبکه‌ای با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئیدی با ۸ توصیف‌کننده در لایه‌ی ورودی، ۵ گره در لایه‌ی پنهان و دور آموزشی ۹۶ کمترین مقدار MSE را برای سری ارزیابی دارد در نتیجه این شبکه به عنوان شبکه‌ی بهینه انتخاب گردید.

جهت بهینه کردن مقدار پارامتر ممنتم، ساختار شبکه‌ی بهینه با مقدار  $\mu$  از ۰/۰۰۰۵ تا ۰/۱ با گام‌های ۰/۰۰۰۵ تغییر داده شد و آنگاه برای هر مورد مقدار میانگین مربع خطا برای سری ارزیابی محاسبه و ثبت گردید که مقدار ۰/۰۴۷۵ به عنوان مقدار بهینه‌ی  $\mu$  انتخاب شد. نمودار مقدار میانگین مربعات خطا برای سری ارزیابی بر حسب  $\mu$  حول نقطه‌ی بهینه در شکل (۳-۶) رسم شده است.



شکل (۳-۶) - نمودار میانگین مربع خطای حاصل از سری ارزیابی بر حسب پارامتر  $\mu$

با توجه به روند بهینه‌سازی ارائه شده، ساختار شبکه‌ی عصبی بهینه شده مطابق شکل (۷-۳) می‌باشد. همچنین مقدار پارامترهای بهینه شده‌ی شبکه‌ی عصبی با استفاده از توصیف‌کننده‌های منتخب حاصل از روش رگرسیون مرحله‌ای در جدول (۹-۳) ارائه شده است.



شکل (۷-۳) - ساختار شبکه عصبی بهینه شده

جدول (۹-۳) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش رگرسیون مرحله‌ای

Train br	تابع آموزش
logsig	تابع انتقال لایه‌ی پنهان
pure line	تابع انتقال لایه‌ی خروجی
۵	تعداد نرون لایه‌ی پنهان
۸	تعداد متغیرهای ورودی
۹۶	تعداد دوره‌های آموزش
۰/۰۴۷۵	پارامتر $\mu$
۰/۰۴۹۵	مقدار MSE سری ارزیابی

### ۳-۱-۷-۲- بهینه‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده-

#### های انتخاب شده توسط راهبرد CDFS

در این بخش نیز از شبکه‌ی عصبی سه لایه، متشکل از یک لایه‌ی ورودی، یک لایه‌ی پنهان و یک لایه‌ی خروجی استفاده شد. برای طراحی یک شبکه‌ی بهینه، مانند قبل باید پارامترهای شبکه بهینه شوند. بر این اساس شبکه توسط دو الگوریتم آموزشی لونبرگ-مارکوارت و تنظیم بایزین، با تعداد متفاوت گره در لایه‌ی پنهان از ۲ تا ۱۰ و همچنین توابع لگاریتم سیگموئید (logsig) و تانژانت سیگموئید (tansig)، به عنوان توابع انتقال لایه‌ی پنهان، و تعداد دور آموزشی از ۱۰ تا ۲۰۰ با گام ۲ آموزش داده شد. در تمامی این شبکه‌ها تعداد متغیر ورودی ثابت و برابر همان ۶ توصیف‌کننده معرفی شده در جدول (۳-۷) است. لازم به تذکر است که برای بهینه‌سازی پارامترهای شبکه‌های عصبی به حداقل رساندن تابع MSE برای سری ارزیابی مدنظر است. با توجه به تعداد زیاد داده‌های به دست آمده از تغییر همزمان تعداد گره‌های لایه پنهان و تعداد دور آموزشی، بخشی از نتایج این بررسی‌ها در جداول (۳-۱۰) تا (۳-۱۷) گزارش شده است.

جدول (۳-۱۰) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۰	۰/۲۶۸۹	۰/۳۴۲۸	۰/۶۰۶۶	۰/۵۶۸۳	۰/۴۴۷۰	۰/۱۸۴۱	۰/۵۵۱۸	۰/۵۸۸۶	۰/۱۸۲۶
	۲۰	۰/۲۷۱۸	۰/۳۱۲۲	۰/۲۶۴۸	۰/۱۹۵۵	۰/۱۹۳۹	۰/۱۳۴۰	۰/۱۶۵۹	۰/۳۱۳۷	۰/۱۱۳۹
	۳۰	۰/۲۵۱۰	۰/۲۷۹۵	۰/۲۲۱۰	۰/۱۷۰۶	۰/۱۸۲۳	۰/۱۲۸۵	۰/۱۴۴۱	۰/۱۹۲۲	۰/۱۵۴۷
	۴۰	۰/۲۲۵۳	۰/۲۷۴۸	۰/۲۰۵۵	۰/۲۳۶۰	۰/۲۱۷۸	۰/۱۴۳۰	۰/۱۲۷۳	۰/۱۶۶۱	۰/۱۵۵۸
	۵۰	۰/۲۱۰۲	۰/۲۵۳۴	۰/۱۷۵۱	۰/۲۳۰۸	۰/۱۶۷۷	۰/۱۴۷۵	۰/۱۰۴۶	۰/۲۳۱۹	۰/۱۵۸۶
	۶۰	۰/۲۰۹۵	۰/۲۴۸۷	۰/۱۷۳۴	۰/۱۳۸۳	۰/۱۵۸۸	۰/۱۴۳۶	۰/۱۰۹۲	۰/۲۳۱۴	۰/۱۴۷۳
	۷۰	۰/۲۰۹۵	۰/۲۴۷۳	۰/۱۵۸۷	۰/۱۳۲۴	۰/۱۶۵۱	۰/۱۴۰۵	۰/۱۲۴۷	۰/۲۱۱۷	۰/۱۴۶۰
	۸۰	۰/۲۰۹۵	۰/۲۴۶۸	۰/۱۵۸۶	۰/۱۳۲۱	۰/۱۷۳۱	۰/۱۳۶۲	۰/۱۳۱۴	۰/۱۹۷۳	۰/۱۶۱۵
	۹۰	۰/۲۰۹۵	۰/۲۴۶۲	۰/۱۵۸۶	۰/۱۳۲۱	۰/۱۷۴۴	۰/۱۳۰۳	۰/۱۳۵۰	۰/۱۷۴۱	۰/۱۹۵۹
	۱۰۰	۰/۲۰۹۵	۰/۲۴۶۱	۰/۱۵۸۶	۰/۱۳۲۳	۰/۱۶۶۶	۰/۱۳۱۲	۰/۱۳۸۲	۰/۱۶۰۶	۰/۱۹۷۸
	۱۱۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۳	۰/۱۷۰۷	۰/۱۳۰۷	۰/۱۴۱۹	۰/۱۵۴۴	۰/۱۸۴۵
	۱۲۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۷۵۴	۰/۱۳۲۲	۰/۱۶۲۸	۰/۱۵۵۲	۰/۱۸۳۶
	۱۳۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۰۹	۰/۱۳۴۰	۰/۱۷۰۴	۰/۱۵۷۵	۰/۱۹۱۲
	۱۴۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۳۳	۰/۱۳۶۱	۰/۱۸۳۲	۰/۱۹۹۱	۰/۲۰۴۴
	۱۵۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۴۴	۰/۱۴۱۱	۰/۱۸۲۰	۰/۲۱۵۶	۰/۲۰۵۰
	۱۶۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۶۱	۰/۱۴۵۴	۰/۱۶۳۵	۰/۲۱۹۷	۰/۲۱۰۵
	۱۷۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۷۲	۰/۱۴۷۵	۰/۱۵۹۳	۰/۲۱۶۳	۰/۲۰۲۱
	۱۸۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۷۷	۰/۱۴۸۳	۰/۱۵۹۱	۰/۲۰۷۹	۰/۲۱۵۸
	۱۹۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۸۳	۰/۱۴۸۴	۰/۱۶۱۲	۰/۲۱۵۰	۰/۲۱۹۶
	۲۰۰	۰/۲۰۹۵	۰/۲۴۶۰	۰/۱۵۸۶	۰/۱۳۲۲	۰/۱۸۸۳	۰/۱۴۷۶	۰/۱۶۲۳	۰/۲۲۲۴	۰/۲۵۶۲

جدول (۳-۱۱) - مشابه جدول (۳-۱۰) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۴۴	۰/۲۱۴۴	۰/۲۶۲۳	۰/۱۹۰۶	۰/۲۳۷۰	۰/۲۱۶۵	۰/۱۴۶۲	۰/۱۱۱۳	۰/۲۰۹۶	۰/۱۵۸۹
	۴۶	۰/۲۱۱۹	۰/۲۵۸۴	۰/۱۷۹۲	۰/۲۳۵۴	۰/۲۰۴۸	۰/۱۴۷۰	۰/۱۰۶۱	۰/۲۲۱۵	۰/۱۵۹۹
	۴۸	۰/۲۱۰۷	۰/۲۵۵۵	۰/۱۷۹۷	۰/۲۳۳۴	۰/۱۹۱۳	۰/۱۴۷۵	۰/۱۰۳۸	۰/۲۲۷۹	۰/۱۵۹۶
	۵۰	۰/۲۱۰۲	۰/۲۵۳۴	۰/۱۷۵۱	۰/۲۳۰۸	۰/۱۶۷۷	۰/۱۴۷۵	۰/۱۰۴۶	۰/۲۳۱۹	۰/۱۵۸۶
	۵۲	۰/۲۰۹۸	۰/۲۵۱۹	۰/۱۷۲۸	۰/۱۹۳۴	۰/۱۷۲۴	۰/۱۴۷۰	۰/۱۰۵۰	۰/۲۳۴۴	۰/۱۵۶۱

جدول (۳-۱۲) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونیبرگ - مارکوارت و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۰	۰/۲۶۹۴	۱/۱۸۲۱	۰/۳۳۴۵	۰/۴۲۲۴	۰/۲۲۶۱	۰/۲۳۶۱	۰/۲۷۱۸	۰/۲۹۹۱	۰/۱۶۶۴
	۲۰	۰/۳۲۴۷	۰/۲۶۸۵	۰/۲۴۵۵	۰/۳۷۱۴	۰/۲۲۷۳	۰/۱۷۸۹	۰/۱۸۲۳	۰/۲۲۸۷	۰/۱۳۹۱
	۳۰	۰/۳۲۸۴	۰/۳۲۳۰	۰/۱۹۸۲	۰/۲۰۲۵	۰/۲۲۵۱	۰/۱۳۱۲	۰/۱۹۲۱	۰/۲۴۹۳	۰/۱۲۷۹
	۴۰	۰/۳۰۵۵	۰/۳۶۸۶	۰/۱۵۹۲	۰/۱۶۸۶	۰/۱۷۷۸	۰/۱۳۳۵	۰/۲۰۹۰	۰/۲۴۵۱	۰/۱۳۲۲
	۵۰	۰/۲۸۱۶	۰/۳۳۰۸	۰/۱۵۵۱	۰/۱۷۱۵	۰/۱۶۴۸	۰/۱۲۹۳	۰/۱۷۵۰	۰/۲۵۲۲	۰/۱۴۲۹
	۶۰	۰/۲۸۱۴	۰/۱۹۵۲	۰/۱۶۸۳	۰/۱۷۴۹	۰/۱۵۵۹	۰/۱۲۷۵	۰/۱۸۳۰	۰/۲۶۹۲	۰/۱۵۱۴
	۷۰	۰/۲۸۱۳	۰/۱۹۵۸	۰/۱۷۰۹	۰/۱۵۶۷	۰/۱۵۷۹	۰/۱۲۴۹	۰/۱۸۹۱	۰/۳۶۴۸	۰/۱۵۸۴
	۸۰	۰/۲۸۱۱	۰/۱۹۵۶	۰/۱۷۲۱	۰/۱۵۵۱	۰/۱۵۹۲	۰/۱۲۱۰	۰/۱۹۶۳	۰/۵۳۰۱	۰/۱۶۵۴
	۹۰	۰/۲۸۱۵	۰/۱۹۵۷	۰/۱۷۲۷	۰/۱۵۴۹	۰/۱۶۰۴	۰/۱۰۸۴	۰/۱۹۹۲	۱/۰۱۳۲	۰/۱۶۲۳
	۱۰۰	۰/۲۸۱۷	۰/۱۹۵۶	۰/۱۷۳۲	۰/۱۵۳۷	۰/۱۶۰۷	۰/۱۰۳۴	۰/۲۰۰۰	۱/۰۰۰۷	۰/۱۷۶۰
	۱۱۰	۰/۲۸۲۵	۰/۱۹۵۵	۰/۱۷۳۳	۰/۱۵۲۸	۰/۱۶۱۲	۰/۱۱۱۳	۰/۱۸۸۸	۱/۰۱۱۸	۰/۱۹۰۰
	۱۲۰	۰/۲۸۳۲	۰/۱۹۵۵	۰/۱۷۳۶	۰/۱۵۲۴	۰/۱۶۱۵	۰/۱۱۸۳	۰/۱۰۹۵	۱/۰۱۱۵	۰/۱۸۳۹
	۱۳۰	۰/۲۸۱۰	۰/۱۹۵۵	۰/۱۷۳۵	۰/۱۵۱۹	۰/۱۶۱۹	۰/۱۴۶۲	۰/۱۳۸۸	۱/۰۱۶۲	۰/۱۸۷۵
	۱۴۰	۰/۲۸۳۷	۰/۱۹۵۶	۰/۱۷۳۶	۰/۱۵۱۴	۰/۱۶۲۳	۰/۱۳۸۴	۰/۱۴۵۷	۱/۰۴۰۲	۰/۱۸۸۷
	۱۵۰	۰/۲۸۳۶	۰/۱۹۸۵	۰/۱۷۳۵	۰/۱۵۰۹	۰/۱۶۲۷	۰/۱۳۹۲	۰/۱۴۹۹	۱/۰۵۲۲	۰/۲۱۷۶
	۱۶۰	۰/۲۸۳۹	۰/۱۹۸۸	۰/۱۷۳۵	۰/۱۵۰۳	۰/۱۶۳۱	۰/۱۴۰۰	۰/۱۵۱۲	۱/۰۹۲۰	۰/۲۷۱۵
	۱۷۰	۰/۲۸۴۴	۰/۲۰۱۵	۰/۱۷۳۴	۰/۱۴۹۸	۰/۱۶۳۵	۰/۱۳۹۸	۰/۱۵۲۷	۱/۱۰۲۲	۰/۲۷۷۱
	۱۸۰	۰/۲۸۴۹	۰/۲۰۶۸	۰/۱۷۳۴	۰/۱۴۹۴	۰/۱۶۴۳	۰/۱۳۹۵	۰/۱۵۲۳	۱/۰۷۲۶	۰/۲۸۰۴
	۱۹۰	۰/۲۸۵۲	۰/۲۰۸۸	۰/۱۷۳۳	۰/۱۴۹۰	۰/۱۶۵۰	۰/۱۳۹۲	۰/۱۵۱۰	۱/۱۳۳۹	۰/۲۸۲۵
	۲۰۰	۰/۲۸۵۵	۰/۲۰۸۴	۰/۱۷۳۳	۰/۱۴۸۶	۰/۱۶۵۷	۰/۱۳۸۹	۰/۱۴۶۷	۱/۱۴۵۸	۰/۲۸۴۳

جدول (۳-۱۳) - مشابه جدول (۳-۱۲) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۹۲	۰/۲۸۱۴	۰/۱۹۵۷	۰/۱۷۲۹	۰/۱۵۶۰	۰/۱۶۰۴	۰/۱۰۵۸	۰/۱۹۹۴	۱/۰۰۵۰	۰/۱۶۲۴
	۹۴	۰/۲۸۱۴	۰/۱۹۵۶	۰/۱۷۲۹	۰/۱۵۳۸	۰/۱۶۰۵	۰/۱۰۴۱	۰/۱۹۹۶	۰/۹۹۵۳	۰/۱۶۶۱
	۹۶	۰/۲۸۱۴	۰/۱۹۵۶	۰/۱۷۳۱	۰/۱۵۳۸	۰/۱۶۰۶	۰/۱۰۳۳	۰/۱۹۹۸	۰/۹۹۲۵	۰/۱۶۲۴
	۹۸	۰/۲۸۱۴	۰/۱۹۵۶	۰/۱۷۳۰	۰/۱۵۳۹	۰/۱۶۰۷	۰/۱۰۳۱	۰/۱۹۹۹	۰/۹۹۵۸	۰/۱۶۹۶
	۱۰۰	۰/۲۸۱۷	۰/۱۹۵۶	۰/۱۷۳۲	۰/۱۵۳۷	۰/۱۶۰۷	۰/۱۰۳۴	۰/۲۰۰۰	۱/۰۰۰۷	۰/۱۷۶۰

جدول (۳-۱۴) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۰	۰/۲۷۴۸	۰/۳۷۵۰	۰/۲۲۵۲	۰/۳۸۴۵	۰/۲۹۱۸	۰/۲۰۶۰	۰/۲۲۵۴	۰/۳۴۰۰	۰/۴۹۷۹
	۲۰	۰/۲۵۹۴	۰/۲۴۹۸	۰/۲۴۷۵	۰/۲۱۱۷	۰/۱۸۷۱	۰/۱۶۴۱	۰/۱۴۷۶	۰/۱۸۰۶	۰/۲۳۷۱
	۳۰	۰/۲۵۶۹	۰/۲۴۶۵	۰/۲۰۴۴	۰/۱۷۴۸	۰/۱۹۷۲	۰/۱۹۷۱	۰/۱۳۱۲	۰/۱۹۰۳	۰/۱۷۴۶
	۴۰	۰/۲۵۰۴	۰/۲۴۰۰	۰/۲۰۴۹	۰/۱۹۶۰	۰/۱۵۷۱	۰/۲۰۹۰	۰/۱۲۷۲	۰/۲۱۱۶	۰/۱۵۰۸
	۵۰	۰/۲۴۶۶	۰/۲۱۳۶	۰/۲۴۷۹	۰/۱۸۶۳	۰/۱۸۱۳	۰/۲۸۸۷	۰/۱۲۶۱	۰/۱۸۴۱	۰/۱۴۷۱
	۶۰	۰/۲۴۴۸	۰/۲۳۷۷	۰/۲۴۷۲	۰/۱۹۸۱	۰/۱۷۹۷	۰/۳۹۰۱	۰/۱۱۱۳	۰/۱۶۲۰	۰/۱۶۹۰
	۷۰	۰/۲۴۳۰	۰/۲۲۸۹	۰/۲۴۶۴	۰/۲۰۵۸	۰/۱۶۶۳	۰/۴۵۳۵	۰/۱۰۶۹	۰/۱۶۱۵	۰/۱۷۱۵
	۸۰	۰/۲۴۲۷	۰/۲۲۶۶	۰/۲۴۳۹	۰/۲۲۰۰	۰/۱۷۲۴	۰/۴۶۵۵	۰/۱۱۲۸	۰/۱۷۸۲	۰/۱۹۱۳
	۹۰	۰/۲۴۲۱	۰/۲۲۶۷	۰/۲۴۲۹	۰/۲۴۱۱	۰/۱۷۰۵	۰/۴۵۱۱	۰/۱۲۸۳	۰/۳۲۲۷	۰/۲۰۰۲
	۱۰۰	۰/۲۴۱۸	۰/۲۲۵۵	۰/۲۴۱۸	۰/۲۴۵۱	۰/۱۶۲۴	۰/۴۳۲۹	۰/۱۵۱۴	۰/۶۳۵۵	۰/۲۱۵۹
	۱۱۰	۰/۲۴۱۷	۰/۲۲۵۳	۰/۲۴۱۱	۰/۲۳۲۷	۰/۱۸۲۶	۰/۴۰۸۴	۰/۱۷۳۱	۱/۰۸۹۳	۰/۲۱۵۵
	۱۲۰	۰/۲۴۱۵	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۸۵۵	۰/۱۸۵۸	۰/۳۹۸۸	۰/۱۷۵۹	۱/۱۵۵۸	۰/۲۱۵۶
	۱۳۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۶۷۸	۰/۱۸۶۰	۰/۳۹۵۳	۰/۱۷۴۳	۱/۵۳۳۸	۰/۲۳۴۹
	۱۴۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۶۳۳	۰/۱۸۴۷	۰/۳۹۵۰	۰/۱۶۸۲	۲/۷۲۷۱	۰/۲۵۱۴
	۱۵۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۶۲۰	۰/۱۸۴۲	۰/۳۹۵۰	۰/۱۵۷۷	۴/۲۶۷۱	۰/۲۶۱۹
	۱۶۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۶۳۶	۰/۱۸۴۰	۰/۳۹۵۰	۰/۱۴۷۴	۵/۱۷۶۳	۰/۲۶۸۷
	۱۷۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۶۹۱	۰/۱۸۳۹	۰/۳۹۵۰	۰/۱۴۱۰	۴/۷۱۳۳	۰/۲۶۷۹
	۱۸۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۷۲۹	۰/۱۸۳۷	۰/۳۹۵۰	۰/۱۴۶۰	۴/۸۱۷۹	۰/۲۴۱۸
	۱۹۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۷۴۵	۰/۱۸۳۶	۰/۳۹۵۰	۰/۱۴۳۸	۴/۹۴۵۰	۰/۲۳۴۲
	۲۰۰	۰/۲۴۱۴	۰/۲۲۵۲	۰/۲۴۰۸	۰/۱۷۰۹	۰/۱۸۳۴	۰/۳۹۵۰	۰/۱۴۴۰	۵/۰۳۸۹	۰/۲۲۴۳

جدول (۳-۱۵) - مشابه جدول (۳-۱۴) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۶۲	۰/۲۴۴۶	۰/۲۳۴۹	۰/۲۴۷۱	۰/۲۰۳۲	۰/۱۷۱۹	۰/۴۰۷۰	۰/۱۰۹۷	۰/۱۶۱۶	۰/۱۶۵۹
	۶۴	۰/۲۴۴۵	۰/۲۳۲۲	۰/۲۴۶۹	۰/۲۱۳۱	۰/۱۶۹۵	۰/۴۲۱۹	۰/۱۰۳۷	۰/۱۶۱۷	۰/۱۶۲۹
	۶۶	۰/۲۴۳۵	۰/۲۳۰۴	۰/۲۴۶۸	۰/۲۰۹۶	۰/۱۶۷۵	۰/۴۳۴۹	۰/۱۰۴۷	۰/۱۶۱۸	۰/۱۶۰۶
	۶۸	۰/۲۴۲۶	۰/۲۳۰۲	۰/۲۴۶۶	۰/۲۰۷۱	۰/۱۶۶۵	۰/۴۴۵۷	۰/۱۰۵۷	۰/۱۶۱۸	۰/۱۴۹۲
	۷۰	۰/۲۴۳۰	۰/۲۲۸۹	۰/۲۴۶۴	۰/۲۰۵۸	۰/۱۶۶۳	۰/۴۵۳۵	۰/۱۰۶۹	۰/۱۶۱۵	۰/۱۷۱۵

جدول (۳-۱۶) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونیگ-مارکوارت و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های CDFS

		تعداد گره در لایه پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۰	۰/۲۹۲۹	۰/۷۹۰۲	۰/۲۷۱۴	۰/۳۱۷۷	۰/۲۴۵۵	۰/۲۵۶۸	۰/۳۱۷۲	۰/۳۸۱۱	۰/۲۵۳۲
	۲۰	۰/۲۹۳۰	۰/۲۶۲۴	۰/۲۲۵۸	۰/۳۲۷۸	۰/۱۵۶۷	۰/۲۸۳۴	۰/۲۷۹۳	۰/۱۷۹۰	۰/۳۵۰۸
	۳۰	۰/۲۹۶۹	۰/۲۲۹۵	۰/۲۵۸۷	۰/۲۲۶۵	۰/۱۴۳۳	۰/۲۸۶۲	۰/۲۷۱۶	۰/۱۶۶۶	۱/۰۱۰۸
	۴۰	۰/۲۴۵۰	۰/۲۲۷۰	۰/۲۸۶۰	۰/۲۲۲۵	۰/۱۴۰۰	۰/۲۴۱۱	۰/۱۶۹۱	۰/۱۴۸۱	۰/۸۸۶۰
	۵۰	۰/۲۲۵۵	۰/۲۲۷۱	۰/۲۸۵۸	۰/۲۲۷۷	۰/۱۳۷۳	۰/۱۹۴۱	۰/۱۱۷۱	۰/۱۵۰۸	۰/۶۹۲۴
	۶۰	۰/۲۱۱۷	۰/۲۲۶۵	۰/۲۸۴۶	۰/۲۵۸۹	۰/۱۳۶۸	۰/۱۹۸۱	۰/۱۵۵۲	۰/۱۵۴۷	۰/۶۹۴۱
	۷۰	۰/۲۰۸۶	۰/۲۲۶۶	۰/۲۸۵۱	۰/۲۷۸۶	۰/۱۳۸۳	۰/۲۰۷۱	۰/۱۶۲۸	۰/۱۶۱۴	۰/۹۰۵۷
	۸۰	۰/۲۰۸۶	۰/۲۲۶۹	۰/۲۸۶۸	۰/۳۰۷۷	۰/۱۴۰۸	۰/۱۸۴۳	۰/۱۶۹۹	۰/۱۶۷۸	۱/۲۱۴۵
	۹۰	۰/۲۰۸۶	۰/۲۲۶۸	۰/۲۸۷۳	۰/۳۱۵۵	۰/۱۴۲۹	۰/۱۷۴۷	۰/۱۷۹۵	۰/۱۷۴۵	۱/۳۱۰۷
	۱۰۰	۰/۲۰۸۵	۰/۲۲۶۸	۰/۲۸۹۰	۰/۲۸۹۹	۰/۱۴۴۳	۰/۲۶۲۹	۰/۱۵۳۰	۰/۱۸۴۶	۱/۰۹۷۲
	۱۱۰	۰/۲۰۸۵	۰/۲۲۶۷	۰/۲۸۷۱	۰/۲۸۲۶	۰/۱۴۵۱	۰/۳۲۸۵	۰/۱۵۶۹	۰/۱۸۸۴	۱/۰۵۷۱
	۱۲۰	۰/۲۰۸۶	۰/۲۲۶۸	۰/۲۸۹۳	۰/۲۸۴۱	۰/۱۴۵۶	۰/۳۶۵۱	۰/۱۴۱۷	۰/۱۹۱۵	۱/۰۴۴۹
	۱۳۰	۰/۲۰۸۵	۰/۲۲۶۷	۰/۲۸۹۳	۰/۲۸۶۱	۰/۱۴۵۹	۰/۳۷۴۱	۰/۱۲۷۷	۰/۱۹۴۸	۱/۰۷۰۱
	۱۴۰	۰/۲۰۸۴	۰/۲۲۶۷	۰/۲۸۹۶	۰/۲۸۷۵	۰/۱۴۶۲	۰/۳۶۵۸	۰/۱۱۸۸	۰/۲۰۰۶	۱/۰۸۲۰
	۱۵۰	۰/۲۰۸۱	۰/۲۲۶۵	۰/۲۸۹۸	۰/۲۸۸۳	۰/۱۴۶۲	۰/۳۵۸۲	۰/۱۰۹۴	۰/۲۱۲۳	۱/۰۸۴۶
	۱۶۰	۰/۲۰۸۱	۰/۲۲۶۵	۰/۲۹۰۰	۰/۲۸۹۲	۰/۱۴۶۴	۰/۳۵۳۷	۰/۱۰۰۲	۰/۲۱۸۴	۱/۰۹۰۲
	۱۷۰	۰/۲۰۸۴	۰/۲۲۶۵	۰/۲۹۰۲	۰/۲۸۹۸	۰/۱۴۶۴	۰/۳۵۰۵	۰/۰۸۹۹	۰/۲۲۲۱	۱/۰۹۳۱
	۱۸۰	۰/۲۰۸۵	۰/۲۲۶۵	۰/۲۹۰۵	۰/۲۹۰۰	۰/۱۴۶۵	۰/۳۴۷۹	۰/۰۶۶۲	۰/۲۲۴۲	۱/۰۹۴۰
	۱۹۰	۰/۲۰۸۴	۰/۲۲۶۶	۰/۲۹۰۳	۰/۲۹۰۳	۰/۱۴۶۵	۰/۳۴۵۵	۰/۰۷۸۹	۰/۲۲۵۸	۱/۰۹۱۸
	۲۰۰	۰/۲۰۸۴	۰/۲۲۶۷	۰/۲۹۰۴	۰/۲۹۰۵	۰/۱۴۶۵	۰/۳۴۳۳	۰/۱۰۲۱	۰/۲۲۶۹	۱/۰۸۹۵

جدول (۳-۱۷) - مشابه جدول (۳-۱۶) حول نقطه‌ی بهینه با گام ۲

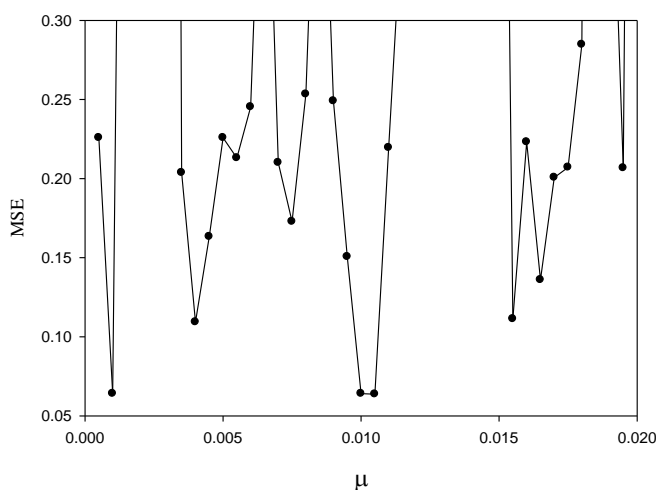
		تعداد گره در لایه پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۷۸	۰/۲۰۸۴	۰/۲۲۶۶	۰/۲۹۰۱	۰/۲۹۰۰	۰/۱۴۶۴	۰/۳۴۸۴	۰/۰۷۰۶	۰/۲۲۳۹	۱/۰۹۴۸
	۱۸۰	۰/۲۰۸۵	۰/۲۲۶۵	۰/۲۹۰۵	۰/۲۹۰۰	۰/۱۴۶۵	۰/۳۴۷۹	۰/۰۶۶۲	۰/۲۲۴۲	۱/۰۹۴۰
	۱۸۲	۰/۲۰۸۴	۰/۲۲۶۶	۰/۲۹۰۳	۰/۲۹۰۱	۰/۱۴۶۴	۰/۳۴۷۴	۰/۰۶۴۱	۰/۲۲۴۶	۱/۰۹۴۷
	۱۸۴	۰/۲۰۸۲	۰/۲۲۶۵	۰/۲۹۰۳	۰/۲۹۰۳	۰/۱۴۶۵	۰/۳۴۶۹	۰/۰۶۴۵	۰/۲۲۴۹	۱/۰۹۵۷



۱۸۶	۰/۲۰۸۴	۰/۲۲۶۳	۰/۲۹۰۳	۰/۲۹۰۲	۰/۱۴۶۵	۰/۳۴۶۴	۰/۰۶۵۹	۰/۲۲۵۲	۱/۰۹۴۴
-----	--------	--------	--------	--------	--------	--------	--------	--------	--------

با توجه به نتایج موجود در این جداول الگوریتم آموزشی لونبرگ- مارکوارت با تابع انتقال تانژانت سیگموئید، تعداد ۸ گره در لایه‌ی پنهان، و دور آموزشی ۱۸۲ دارای کمترین مقدار میانگین مربعات خطا (MSE) برای سری ارزیابی است.

جهت بهینه کردن مقدار  $\mu$ ، شبکه بهینه به دست آمده با مقدار  $\mu$  از ۰/۰۰۰۵ تا ۰/۱ با گام‌های ۰/۰۰۰۵ تغییر داده شد و آنگاه برای هر مورد مقدار میانگین مربعات خطا برای سری ارزیابی محاسبه و ثبت گردید. نمودار MSE بر حسب  $\mu$  حول نقطه‌ی بهینه نیز در شکل (۳-۸) رسم شده است. بر اساس نتایج به دست آمده مقدار بهینه  $\mu$ ، ۰/۰۱۰۵ به دست آمد.



شکل (۳-۸)- نمودار میانگین مربعات خطای حاصل از سری ارزیابی بر حسب تغییرات پارامتر  $\mu$

توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی با استفاده از توصیف‌کننده‌های انتخاب شده به روش CDFS در جدول (۳-۱۸) ارائه شده است.

جدول (۳-۱۸) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش CDFS

Train lm	تابع آموزش
tansig	تابع انتقال لایه‌ی پنهان
pure line	تابع انتقال لایه‌ی خروجی
۸	تعداد نرون لایه‌ی پنهان
۶	تعداد متغیرهای ورودی
۱۸۲	تعداد دوره‌های آموزش
۰/۰۱۰۵	پارامتر $\mu$
۰/۰۶۳۷	مقدار MSE سری ارزیابی

### ۳-۱-۷-۳ - مدل‌سازی شبکه‌ی عصبی مصنوعی با استفاده از توصیف‌کننده‌های

#### حاصل از روش سهم گروه

در مدل‌سازی به روش شبکه‌ی عصبی مصنوعی بر مبنای روش سهم گروه‌ها، ابتدا باید گروه‌های ساختاری مولکول‌ها را مشخص نمود. با توجه به اختیاری بودن تعریف گروه‌های تشکیل‌دهنده‌ی مواد در روش سهم گروه‌ها، در این پایان‌نامه شکستن مولکول‌ها از محل پیوندهای کربن-کربن صورت گرفته است. به این ترتیب ۱۱ گروه عاملی به ترتیبی که در جدول (۳-۱۹) آمده است در مواد مورد بررسی شناسایی شدند. تعداد هر یک از گروه‌های عاملی موجود در مولکول‌های مورد بررسی به عنوان توصیف‌کننده‌های حاصل از روش سهم گروه‌ها انتخاب شدند که مقادیر آن برای تمامی مولکول‌های موجود در سری داده‌ها در جدول (۳-۲۰) آورده شده است. در صورتی که هر گروه عاملی در مولکول مورد نظر وجود نداشته باشد، برای آن عدد صفر در نظر گرفته می‌شود.

جدول (۳-۱۹) - گروه‌های عاملی موجود در مولکول‌های مورد بررسی به عنوان توصیف‌کننده‌های روش سهم گروه

	گروه عاملی	فرمول شیمیایی
۱	متیل	-CH <sub>3</sub>
۲	کربن نوع دوم	-CH <sub>2</sub> -
۳	کربن نوع سوم	>CH-
۴	الکل نوع اول	-CH <sub>2</sub> -OH
۵	الکل نوع دوم	>CH-OH
۶	الکل نوع سوم	>COH-
۷	گروه عاملی کتونی	>C=O
۸	گروه عاملی کربوکسیلیک اسید	-COOH
۹	گروه عاملی استری	-COO-
۱۰	گروه عاملی آمین نوع اول به فرم ۱	-CH <sub>2</sub> -NH <sub>2</sub>
۱۱	گروه عاملی آمین نوع اول به فرم ۲	>CH-NH <sub>2</sub>

جدول (۳-۲۰) - تعداد گروه‌های عاملی موجود در مولکول‌های مختلف سری داده‌ها به عنوان توصیف‌کننده‌های روش سهم گروه

ردیف	نام مولکول	گروه عاملی										
		-CH <sub>3</sub>	-CH <sub>2</sub> -	>CH-	-CH <sub>2</sub> -OH	>CH-OH	>COH-	>C=O	-COOH	-COO-	-CH <sub>2</sub> -NH <sub>2</sub>	>CH-NH <sub>2</sub>
۱	استیک اسید	۱	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰
۲	پروپانویک اسید	۱	۱	۰	۰	۰	۰	۰	۱	۰	۰	۰
۳	بوتانویک اسید	۱	۲	۰	۰	۰	۰	۰	۱	۰	۰	۰
۴	پنتانویک اسید	۱	۳	۰	۰	۰	۰	۰	۱	۰	۰	۰
۵	هگزانویک اسید	۱	۴	۰	۰	۰	۰	۰	۱	۰	۰	۰
۶	هپتانویک اسید	۱	۵	۰	۰	۰	۰	۰	۱	۰	۰	۰
۷	اکتانویک اسید	۱	۶	۰	۰	۰	۰	۰	۱	۰	۰	۰
۸	پروپیل استات	۲	۲	۰	۰	۰	۰	۰	۰	۱	۰	۰
۹	بوتیل استات	۲	۳	۰	۰	۰	۰	۰	۰	۱	۰	۰
۱۰	اتیل پروپیونات	۲	۲	۰	۰	۰	۰	۰	۰	۱	۰	۰
۱۱	اتیل بوتیرات	۲	۳	۰	۰	۰	۰	۰	۰	۱	۰	۰
۱۲	متیل اتیل استات	۳	۰	۱	۰	۰	۰	۰	۰	۱	۰	۰
۱۳	۲-بوتانول	۲	۱	۰	۰	۰	۰	۱	۰	۰	۰	۰
۱۴	۲-پنتانول	۲	۲	۰	۰	۰	۰	۱	۰	۰	۰	۰
۱۵	۳-پنتانول	۲	۲	۰	۰	۰	۰	۱	۰	۰	۰	۰
۱۶	۲-هگزانول	۲	۳	۰	۰	۰	۰	۱	۰	۰	۰	۰
۱۷	۴-متیل-۲-پنتانول	۳	۱	۱	۰	۰	۰	۱	۰	۰	۰	۰
۱۸	۱-پروپانول	۱	۱	۰	۱	۰	۰	۰	۰	۰	۰	۰
۱۹	۱-بوتانول	۱	۲	۰	۱	۰	۰	۰	۰	۰	۰	۰

ادامه‌ی جدول (۳-۲۰)

ردیف	نام مولکول	گروه عاملی										
		-CH <sub>3</sub>	-CH <sub>2</sub> -	>CH-	-CH <sub>2</sub> -OH	>CH-OH	>COH-	>C=O	-COOH	-COO-	-CH <sub>2</sub> -NH <sub>2</sub>	>CH-NH <sub>2</sub>
۲۰	۱-پنتانول	۱	۳	۰	۱	۰	۰	۰	۰	۰	۰	۰
۲۱	۲-متیل-۱-پنتانول	۲	۱	۱	۱	۰	۰	۰	۰	۰	۰	۰
۲۲	۳-پنتانول	۲	۲	۰	۰	۱	۰	۰	۰	۰	۰	۰
۲۳	۳-متیل-۳-پنتانول	۳	۲	۰	۰	۰	۱	۰	۰	۰	۰	۰
۲۴	۳-اتیل-۳-پنتانول	۳	۳	۰	۰	۰	۱	۰	۰	۰	۰	۰
۲۵	۱-هگزانول	۱	۴	۰	۱	۰	۰	۰	۰	۰	۰	۰
۲۶	۲-هگزانول	۲	۳	۰	۰	۱	۰	۰	۰	۰	۰	۰
۲۷	۱-هپتانول	۱	۵	۰	۱	۰	۰	۰	۰	۰	۰	۰
۲۸	۲-هپتانول	۲	۴	۰	۰	۱	۰	۰	۰	۰	۰	۰
۲۹	۱-اکتانول	۱	۶	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۰	۱-نونانول	۱	۷	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۱	۱-دکانول	۱	۸	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۲	۱-دودکانول	۱	۱۰	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۳	۱-تترادکانول	۱	۱۲	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۴	۱-هگزادکانول	۱	۱۴	۰	۱	۰	۰	۰	۰	۰	۰	۰
۳۵	۲-متیل-۲ و ۴-پنتان دی‌ال	۳	۱	۰	۰	۱	۱	۰	۰	۰	۰	۰
۳۶	۲-آمینو بوتان	۲	۱	۰	۰	۰	۰	۰	۰	۰	۰	۱
۳۷	پنتیل آمین	۱	۳	۰	۰	۰	۰	۰	۰	۰	۱	۰
۳۸	هگزیل آمین	۱	۴	۰	۰	۰	۰	۰	۰	۰	۱	۰

ادامه‌ی جدول (۳-۲۰)

ردیف	نام مولکول	گروه عاملی										
		-CH <sub>3</sub>	-CH <sub>2</sub> -	>CH-	-CH <sub>2</sub> -OH	>CH-OH	>COH-	>C=O	-COOH	-COO-	-CH <sub>2</sub> -NH <sub>2</sub>	>CH-NH <sub>2</sub>
۳۹	هپتیل آمین	۱	۵	۰	۰	۰	۰	۰	۰	۰	۱	۰
۴۰	۲-آمینو اکتان	۲	۵	۰	۰	۰	۰	۰	۰	۰	۰	۱
۴۱	<i>n</i> -پروپان	۲	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۲	<i>n</i> -بوتان	۲	۲	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۳	<i>n</i> -پنتان	۲	۳	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۴	<i>n</i> -هگزان	۲	۴	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۵	<i>n</i> -هپتان	۲	۵	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۶	<i>n</i> -اکتان	۲	۶	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۷	<i>n</i> -دودکان	۲	۱۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۸	<i>n</i> -تری‌دکان	۲	۱۱	۰	۰	۰	۰	۰	۰	۰	۰	۰
۴۹	<i>n</i> -پنتادکان	۲	۱۳	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۰	<i>n</i> -هپتادکان	۲	۱۵	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۱	<i>n</i> -اکتادکان	۲	۱۶	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۲	<i>n</i> -نونادکان	۲	۱۷	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۳	<i>n</i> -تری‌کوزان	۲	۲۱	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۴	<i>n</i> -تتراکوزان	۲	۲۲	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۵	<i>n</i> -اکتاکوزان	۲	۲۶	۰	۰	۰	۰	۰	۰	۰	۰	۰
۵۶	<i>n</i> -هگزاتری‌اکوتان	۲	۳۴	۰	۰	۰	۰	۰	۰	۰	۰	۰

به این ترتیب ۱۱ توصیف‌کننده بر مبنای سهم گروه‌ها به علاوه دو متغیر تجربی دما و فشار به عنوان توصیف‌کننده‌های ورودی مدل انتخاب شدند. لایه‌ی خروجی نیز یک نرون دارد که  $K_T$  متناظر با ورودی‌ها را نشان می‌دهد. باز هم باید پارامترهای شبکه‌ی عصبی شامل تعداد نرون‌های لایه‌ی مخفی، تعداد دوره‌های آموزش، نوع تابع انتقال، نوع الگوریتم آموزشی و پارامتر  $\mu$  بهینه گردد. بنابراین تعداد نرون‌های لایه‌ی مخفی از ۲ تا ۱۱ و دور آموزشی از ۱۰ تا ۲۰۰ با گام ۲ در نظر گرفته شدند. برای آموزش شبکه دو الگوریتم آموزشی تنظیم بایزین و لونیبرگ-مارکواریت و برای انتخاب تابع انتقال مناسب دو تابع انتقال تانژانت سیگموئید و لگاریتم سیگموئید مورد استفاده قرار گرفتند، که از ترکیب این دو الگوریتم و دو تابع انتقال، چهار نوع شبکه‌ی عصبی مختلف طراحی گردید. سپس برای هر یک از شبکه‌ها، از روش تغییر همزمان برای تغییر تعداد نرون‌های لایه‌ی مخفی و تعداد دور آموزش جهت آموزش شبکه استفاده گردید و به ازای همه‌ی ترکیب‌های ممکن از این دو پارامتر شبکه‌های طراحی شده بهینه‌سازی شدند. در تمامی این شبکه‌ها مقدار  $\mu$ ، همان مقدار پیش‌فرض ۰/۰۰۰۵ در نظر گرفته شد. در این قسمت نیز مانند بخش گذشته به دلیل تعداد زیاد داده‌های به دست آمده بخشی از نتایج حاصل از این بررسی‌ها برای سری ارزیابی در جداول (۳-۲۱) تا (۳-۲۸) گردآوری شده است.

جدول (۳-۲۱) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه

		تعداد گره در لایه‌ی پنهان									
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
دور آموزشی	۱۰	۱/۴۸۴۲	۱/۳۶۱۱	۱/۳۶۲۹	۰/۸۱۷۳	۰/۹۱۵۳	۱/۱۰۹۶	۱/۵۳۴۱	۰/۸۷۷۳	۰/۷۰۶۴	۰/۸۴۴۷
	۲۰	۱/۰۷۶۶	۰/۹۱۸۵	۰/۵۱۷۶	۶/۰۶۴۰	۰/۲۸۱۵	۰/۴۸۶۴	۰/۸۴۶۰	۰/۴۶۲۷	۱/۵۴۴۹	۰/۴۷۲۴
	۳۰	۰/۲۷۶۸	۰/۵۶۸۹	۰/۲۰۳۴	۱۱/۴۹۲۸	۰/۲۸۸۵	۲/۴۹۱۴	۲/۵۲۱۵	۱/۷۲۹۶	۰/۳۴۳۲	۰/۷۵۵۸
	۴۰	۰/۲۵۷۰	۰/۲۷۴۱	۰/۱۵۹۶	۱۵/۸۳۳۳	۰/۱۷۶۸	۲/۰۸۵۰	۱/۴۶۶۳	۱/۲۲۹۸	۰/۸۹۹۴	۰/۹۲۳۱
	۵۰	۰/۱۵۶۷	۰/۳۳۱۶	۰/۳۳۶۶	۱۵/۴۹۳۷	۰/۱۹۳۲	۱/۳۱۳۹	۱/۴۱۰۹	۱/۰۴۰۵	۰/۹۹۱۴	۱/۳۵۹۵
	۶۰	۰/۱۳۳۲	۰/۴۳۳۲	۰/۱۷۴۱	۱۶/۱۳۱۷	۰/۲۱۷۶	۰/۳۷۷۸	۱/۴۶۱۷	۱/۰۵۰۱	۱/۶۶۱۶	۱/۴۴۹۳
	۷۰	۰/۱۳۵۴	۰/۵۳۹۸	۰/۴۲۳۲	۲۱/۸۴۴۸	۰/۳۱۴۹	۰/۸۳۸۳	۲/۱۹۳۹	۱/۶۵۴۹	۱/۵۷۵۱	۱/۸۱۲۹
	۸۰	۰/۱۳۵۴	۰/۶۰۴۶	۰/۲۶۰۴	۳۱/۱۹۵۲	۰/۲۳۴۵	۱/۱۵۶۸	۳/۶۲۸۷	۱/۳۳۱۵	۰/۸۲۶۵	۲/۲۴۳۶
	۹۰	۰/۱۳۵۴	۰/۵۷۶۱	۰/۱۲۳۷	۴۰/۹۸۸۲	۰/۲۲۸۷	۱/۶۷۱۷	۵/۰۴۸۲	۱/۷۹۰۵	۰/۳۸۱۰	۲/۵۰۹۱
	۱۰۰	۰/۱۳۵۴	۰/۳۸۳۷	۰/۱۱۵۴	۶۰/۱۵۰۰	۰/۲۰۳۳	۲/۰۲۳۱	۴/۵۶۰۹	۱/۹۵۴۱	۰/۲۱۸۴	۲/۶۱۱۳
	۱۱۰	۰/۱۳۵۴	۰/۲۹۱۴	۰/۱۱۰۴	۷۵/۴۳۸۲	۰/۶۲۶۳	۱/۷۸۹۶	۳/۴۸۵۷	۳/۵۴۸۵	۰/۰۸۲۰	۲/۶۱۵۴
	۱۲۰	۰/۱۳۵۴	۰/۲۸۶۵	۰/۱۰۵۹	۸۵/۴۰۱۰	۰/۷۶۹۹	۱/۲۳۸۹	۳/۵۰۷۲	۴/۵۵۹۸	۰/۰۳۳۱	۲/۲۶۳۰
	۱۳۰	۰/۱۳۵۴	۰/۳۲۹۲	۰/۱۰۳۷	۸۸/۶۷۱۷	۰/۷۸۲۸	۰/۷۹۳۶	۲/۶۷۲۷	۵/۵۹۵۶	۰/۰۳۰۲	۱/۳۲۷۸
	۱۴۰	۰/۱۳۵۴	۰/۳۲۴۰	۰/۱۰۴۲	۸۹/۷۱۸۴	۰/۸۲۳۱	۰/۵۵۲۷	۲/۲۱۶۱	۶/۴۰۵۴	۰/۰۲۹۸	۱/۱۴۸۲
	۱۵۰	۰/۱۳۵۴	۰/۲۱۵۱	۰/۱۰۲۲	۸۵/۳۶۱۵	۰/۸۲۲۶	۰/۲۹۹۷	۰/۴۳۰۷	۶/۸۱۷۱	۰/۰۳۲۴	۰/۸۲۴۸
	۱۶۰	۰/۱۳۵۴	۰/۲۰۳۲	۰/۱۰۱۱	۸۰/۵۷۸۱	۰/۸۲۳۶	۰/۲۸۱۴	۲/۴۱۴۶	۶/۹۴۷۲	۰/۰۷۵۰	۰/۷۶۴۱
	۱۷۰	۰/۱۳۵۴	۰/۲۰۲۴	۰/۱۰۰۶	۸۱/۹۲۰۸	۰/۸۴۷۸	۰/۳۰۶۷	۲/۳۶۵۶	۷/۳۲۱۲	۰/۰۸۱۵	۰/۷۰۸۲
	۱۸۰	۰/۱۳۵۴	۰/۲۰۲۵	۰/۰۹۹۵	۸۴/۹۹۸۸	۰/۸۳۶۳	۰/۲۷۸۹	۲/۵۹۹۰	۷/۶۰۰۴	۰/۰۸۸۷	۰/۶۷۵۰
	۱۹۰	۰/۱۳۵۴	۰/۲۰۲۵	۰/۰۹۸۹	۸۴/۹۶۸۸	۰/۶۹۲۱	۰/۲۲۳۴	۳/۰۱۳۲	۷/۶۷۷۴	۰/۰۹۷۳	۰/۷۱۸۵
	۲۰۰	۰/۱۳۵۴	۰/۲۰۲۵	۰/۰۹۷۵	۸۵/۴۸۵۹	۰/۶۳۷۳	۰/۱۹۸۹	۳/۴۶۲۵	۷/۵۷۵۱	۰/۱۱۹۵	۱/۳۱۳۴

جدول (۳-۲۲) - مشابه جدول (۳-۲۱) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۳۲	۰/۱۳۵۴	۰/۳۲۴۸	۰/۱۰۱۷	۸۸/۶۱۷۰	۰/۸۰۳۴	۰/۷۵۲۵	۲/۴۹۸۲	۵/۷۸۴۷	۰/۰۲۹۹
	۱۳۴	۰/۱۳۵۴	۰/۳۱۴۴	۰/۰۹۹۲	۸۹/۱۷۸۹	۰/۸۱۳۶	۰/۷۱۰۶	۲/۳۱۲۱	۵/۹۶۲۱	۰/۰۲۹۷
	۱۳۶	۰/۱۳۵۴	۰/۳۱۸۴	۰/۰۹۹۲	۸۹/۴۳۳۴	۰/۸۲۰۰	۰/۶۶۴۲	۲/۱۶۲۰	۶/۱۲۵۹	۰/۰۲۹۶
	۱۳۸	۰/۱۳۵۴	۰/۳۲۴۶	۰/۱۰۵۲	۸۹/۶۵۳۶	۰/۸۲۲۳	۰/۶۱۱۵	۲/۱۵۸۸	۶/۲۷۴۰	۰/۰۲۹۶
	۱۴۰	۰/۱۳۵۴	۰/۳۲۴۰	۰/۱۰۴۲	۸۹/۷۱۸۴	۰/۸۲۳۱	۰/۵۵۲۷	۲/۲۱۶۱	۶/۴۰۵۴	۰/۰۲۹۸



جدول (۳-۲۳) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونیبرگ -  
 مارکوارت و تابع انتقال لگاریتم سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده  
 از توصیف‌کننده‌های روش سهم گروه

		تعداد گره در لایه‌ی پنهان									
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
دور آموزشی	۱۰	۰/۸۴۹۰	۲/۷۲۶۸	۳/۴۱۴۶	۱/۳۷۲۵	۱/۹۰۹۴	۰/۲۱۰۱	۱/۷۲۱۱	۰/۵۷۵۶	۵/۳۸۴۷	۰/۱۹۵۸
	۲۰	۰/۲۳۹۵	۱/۲۱۴۶	۱/۹۱۳۶	۰/۷۲۶۸	۰/۴۶۷۳	۰/۱۷۲۸	۱/۲۰۴۲	۳/۳۳۷۴	۲۵/۳۳۵۴	۰/۲۴۵۱
	۳۰	۰/۲۶۷۸	۰/۴۶۶۳	۰/۹۷۱۴	۲/۱۲۹۴	۰/۳۵۵۲	۰/۲۰۴۹	۰/۶۴۰۶	۳/۶۹۰۴	۲۹/۴۹۵۸	۰/۲۷۰۰
	۴۰	۰/۴۱۴۵	۱/۰۴۶۸	۰/۸۲۱۴	۰/۸۹۴۱	۰/۱۷۳۱	۰/۱۸۲۴	۰/۶۴۷۸	۱/۸۴۵۵	۳۵/۴۱۷۹	۰/۳۰۸۷
	۵۰	۰/۴۷۱۰	۱/۶۵۴۱	۰/۷۱۲۴	۰/۷۸۴۸	۰/۲۰۰۹	۰/۱۷۶۱	۰/۵۲۴۰	۱/۲۹۳۹	۳۹/۳۶۰۳	۰/۳۲۴۶
	۶۰	۰/۴۶۱۳	۱/۷۶۶۱	۰/۶۱۹۸	۰/۶۹۴۳	۰/۱۸۲۰	۰/۱۷۳۱	۰/۵۴۷۷	۴/۷۰۴۵	۴۲/۱۷۷۵	۰/۳۲۲۵
	۷۰	۰/۴۵۷۲	۱/۷۰۹۵	۰/۵۶۵۴	۰/۷۱۰۵	۰/۱۸۰۵	۰/۱۹۲۸	۰/۵۲۰۶	۴/۵۰۵۴	۴۴/۶۶۶۶	۰/۵۲۳۱
	۸۰	۰/۴۶۱۱	۲/۰۲۵۴	۰/۴۸۱۰	۱/۴۶۶۱	۰/۱۶۴۱	۰/۱۴۸۷	۰/۵۲۷۶	۲/۹۷۳۱	۴۶/۲۷۳۱	۰/۷۷۹۷
	۹۰	۰/۴۵۵۲	۱/۷۹۳۶	۰/۴۴۳۴	۱/۶۷۹۱	۰/۱۶۳۱	۰/۱۵۸۵	۰/۵۴۷۴	۰/۲۴۴۰	۴۲/۸۲۶۶	۰/۸۹۳۶
	۱۰۰	۰/۴۵۷۳	۱/۸۲۳۴	۰/۴۴۲۶	۱/۵۰۸۸	۰/۱۶۷۱	۰/۱۴۸۹	۰/۵۸۸۹	۰/۲۲۳۱	۳۲/۱۷۲۰	۰/۹۹۵۲
	۱۱۰	۰/۴۵۴۵	۰/۵۰۴۸	۰/۴۴۸۲	۱/۲۹۶۶	۰/۱۵۰۸	۰/۲۱۲۱	۰/۵۱۷۹	۰/۴۵۹۲	۲۳/۶۳۸۳	۱/۰۸۷۲
	۱۲۰	۰/۴۵۵۸	۰/۲۳۷۵	۰/۴۴۹۰	۱/۰۸۶۶۰	۰/۱۱۹۱	۰/۱۶۷۶	۰/۴۱۰۶	۰/۶۳۹۴	۱۴/۵۰۵۴	۱/۱۹۱۷
	۱۳۰	۰/۴۵۶۱	۰/۲۴۱۶	۰/۴۴۸۸	۱۴/۷۴۷۷	۰/۱۱۶۱	۰/۱۷۸۳	۰/۴۳۱۱	۱/۰۵۸۱	۳/۲۱۱۰	۱/۷۰۹۶
	۱۴۰	۰/۴۵۵۰	۰/۲۴۲۶	۰/۴۴۸۶	۱۵/۰۵۵۲	۰/۱۱۶۰	۰/۱۸۴۰	۰/۴۳۰۸	۰/۹۷۸۳	۲/۴۹۴۳	۱/۸۷۵۸
	۱۵۰	۰/۴۵۱۲	۰/۱۹۹۴	۰/۴۴۹۳	۱۵/۳۰۰۷	۰/۱۱۵۸	۰/۱۷۹۵	۰/۴۲۲۸	۱/۲۶۰۲	۱/۷۶۴۱	۲/۰۸۹۷
	۱۶۰	۰/۴۵۵۱	۰/۱۹۳۵	۰/۴۴۹۱	۱۵/۴۹۸۳	۰/۱۱۵۶	۰/۱۸۱۹	۰/۴۲۷۴	۰/۹۰۷۹	۱/۱۵۲۹	۲/۴۴۴۹
	۱۷۰	۰/۴۵۳۴	۰/۱۷۳۳	۰/۴۴۸۷	۱۵/۶۶۱۸	۰/۱۱۵۸	۰/۲۳۶۱	۰/۴۵۱۸	۰/۴۹۸۳	۱/۰۰۳۹	۲/۹۰۸۵
	۱۸۰	۰/۴۵۴۶	۰/۱۷۳۸	۰/۴۴۹۰	۱۵/۷۹۹۵	۰/۱۱۵۸	۰/۲۲۷۲	۰/۵۴۵۶	۰/۲۸۲۶	۱/۵۶۸۵	۴/۶۶۹۴
	۱۹۰	۰/۴۵۳۳	۰/۱۷۴۰	۰/۴۴۸۸	۱۵/۹۱۶۳	۰/۱۱۵۷	۰/۲۲۹۷	۰/۶۲۷۱	۰/۲۱۹۲	۰/۸۷۲۷	۴/۳۵۵۳
	۲۰۰	۰/۴۵۴۴	۰/۱۷۴۷	۰/۴۴۸۹	۱۶/۰۱۵۷	۰/۱۱۵۸	۰/۲۳۳۹	۰/۸۴۲۵	۰/۲۱۰۳	۰/۴۹۰۲	۳/۴۰۹۶

جدول (۳-۲۴) - مشابه جدول (۳-۲۳) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۵۶	۰/۴۵۰۷	۰/۱۹۱۶	۰/۴۴۹۰	۱۵/۴۲۴۰	۰/۱۱۵۹	۰/۱۸۳۵	۰/۴۲۸۷	۱/۰۷۳۴	۱/۳۶۳۹
	۱۵۸	۰/۴۵۴۰	۰/۲۱۲۲	۰/۴۴۹۱	۱۵/۴۶۱۹	۰/۱۱۵۸	۰/۱۸۱۵	۰/۴۲۴۲	۰/۹۸۴۴	۱/۲۵۲۱
	۱۶۰	۰/۴۵۵۱	۰/۱۹۳۵	۰/۴۴۹۱	۱۵/۴۹۸۳	۰/۱۱۵۶	۰/۱۸۱۹	۰/۴۲۷۴	۰/۹۰۷۹	۱/۱۵۲۹
	۱۶۲	۰/۴۵۰۹	۰/۱۵۹۶	۰/۴۴۹۱	۱۵/۵۳۳۴	۰/۱۱۵۷	۰/۱۹۴۳	۰/۴۲۹۶	۰/۸۲۸۴	۱/۰۶۶۸
	۱۶۴	۰/۴۵۳۸	۰/۱۷۲۸	۰/۴۴۹۲	۱۵/۵۶۷۳	۰/۱۱۵۸	۰/۲۰۵۷	۰/۴۳۲۴	۰/۷۴۲۸	۰/۹۹۶۹

جدول (۳-۲۵) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی تنظیم بایزین و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه

		تعداد گره در لایه‌ی پنهان									
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
دور آموزشی	۱۰	۰/۷۱۷۹	۱/۳۱۳۴	۱/۳۹۳۲	۰/۷۱۷۵	۱/۱۷۱۷	۱/۱۶۰۲	۱/۳۶۴۷	۰/۵۸۳۲	۱/۰۸۳۲	۰/۵۲۳۹
	۲۰	۰/۳۸۹۹	۰/۵۳۶۲	۰/۵۹۱۲	۲/۰۶۸۰	۰/۶۸۲۷	۰/۳۰۸۷	۰/۵۸۵۵	۰/۳۹۸۰	۰/۶۳۰۶	۰/۲۵۲۸
	۳۰	۰/۴۰۴۵	۰/۱۵۳۷	۴/۸۵۶۴	۰/۳۹۷۲	۰/۲۱۰۳	۰/۱۸۴۹	۰/۴۹۵۳	۰/۲۷۵۴	۰/۲۳۴۹	۰/۵۷۲۹
	۴۰	۰/۴۱۷۰	۰/۱۴۵۴	۸/۴۱۵۱	۰/۳۳۴۰	۰/۱۷۳۰	۰/۱۶۳۱	۰/۲۲۹۴	۰/۱۹۴۵	۰/۲۴۰۴	۲/۵۵۰۱
	۵۰	۰/۴۲۵۶	۰/۱۵۴۰	۹/۲۲۴۴	۰/۳۹۰۵	۰/۱۵۵۲	۰/۱۷۵۶	۰/۲۲۹۰	۰/۳۳۳۷	۰/۱۷۸۶	۳/۶۸۳۸
	۶۰	۰/۴۳۱۰	۰/۱۵۴۱	۱۰/۰۲۹۰	۰/۳۵۸۸	۰/۱۶۴۹	۰/۱۶۹۳	۰/۳۵۸۴	۰/۳۹۳۸	۰/۱۶۰۱	۱/۶۷۷۸
	۷۰	۰/۴۳۴۱	۰/۱۵۴۱	۱۳/۳۶۰۵	۰/۳۶۳۶	۰/۲۲۸۶	۰/۱۹۰۴	۰/۶۸۵۲	۰/۳۱۴۷	۰/۱۶۹۵	۱/۰۴۵۷
	۸۰	۰/۴۳۷۰	۰/۱۵۴۱	۱۵/۱۰۸۸	۰/۳۵۰۷	۰/۲۵۳۰	۰/۲۰۲۶	۱/۵۹۸۳	۰/۲۲۷۱	۰/۲۳۴۵	۰/۸۷۰۶
	۹۰	۰/۴۳۱۴	۰/۱۵۴۱	۳۲/۳۲۲۹	۰/۲۵۰۷	۰/۲۵۸۸	۰/۳۲۴۰	۲/۸۱۰۰	۰/۱۶۵۷	۰/۲۵۲۳	۰/۷۵۱۲
	۱۰۰	۰/۴۲۷۵	۰/۱۵۴۱	۳۴/۴۲۸۴	۰/۲۵۰۹	۰/۲۴۸۶	۰/۲۳۵۳	۳/۹۷۰۷	۰/۱۵۳۹	۰/۲۱۶۳	۰/۶۵۹۴
	۱۱۰	۰/۴۱۸۱	۰/۱۵۴۱	۳۴/۳۵۲۴	۰/۲۳۸۸	۰/۲۳۹۲	۰/۲۲۷۹	۴/۴۵۸۵	۰/۱۹۹۹	۰/۱۶۶۱	۰/۵۷۶۵
	۱۲۰	۰/۴۱۴۰	۰/۱۵۴۱	۳۲/۲۷۳۹	۰/۲۴۳۴	۰/۲۵۴۶	۰/۲۲۴۸	۳/۴۱۰۶	۰/۲۵۰۶	۰/۱۳۱۶	۰/۵۱۵۱
	۱۳۰	۰/۴۱۲۲	۰/۱۵۴۱	۲۹/۱۹۴۵	۰/۲۵۹۶	۰/۲۹۹۹	۰/۲۳۴۴	۲/۴۱۹۲	۰/۲۳۴۱	۰/۱۱۲۱	۰/۴۷۳۶
	۱۴۰	۰/۴۱۱۰	۰/۱۵۴۱	۲۸/۱۳۹۴	۰/۲۶۰۷	۰/۲۷۷۷	۰/۲۵۳۵	۲/۲۲۷۹	۰/۲۳۰۴	۰/۱۲۱۰	۰/۴۳۶۵
	۱۵۰	۰/۴۰۷۸	۰/۱۵۴۱	۲۷/۶۱۸۹	۰/۲۵۹۹	۰/۲۵۳۵	۰/۲۶۸۹	۱/۶۲۵۳	۰/۲۳۹۲	۰/۱۲۴۳	۰/۳۹۱۴
	۱۶۰	۰/۴۰۶۹	۰/۱۵۴۱	۳۶/۶۹۷۹	۰/۲۴۸۶	۰/۲۴۳۳	۰/۲۸۵۵	۰/۹۸۶۲	۰/۲۴۴۴	۰/۱۴۴۷	۰/۳۴۵۸
	۱۷۰	۰/۴۰۶۴	۰/۱۵۴۱	۴۱/۸۹۹۱	۰/۲۳۵۲	۰/۲۵۶۵	۰/۳۱۱۵	۱/۹۳۲۶	۰/۲۴۰۸	۰/۱۰۴۴	۰/۳۳۹۹
	۱۸۰	۰/۴۰۶۴	۰/۱۵۴۱	۳۲/۵۲۱۸	۰/۲۲۳۵	۰/۲۶۲۶	۰/۳۴۶۸	۴/۱۰۲۰	۰/۱۴۴۴	۰/۰۶۹۲	۰/۳۴۴۷
	۱۹۰	۰/۴۰۶۴	۰/۱۵۴۱	۳۰/۱۰۸۳	۰/۲۰۸۱	۰/۲۶۱۱	۰/۳۸۵۳	۸/۸۸۹۶	۰/۱۲۴۹	۰/۰۹۳۶	۰/۳۵۴۸
	۲۰۰	۰/۴۰۶۴	۰/۱۵۴۱	۲۹/۴۸۸۴	۰/۱۶۳۳	۰/۲۵۵۴	۰/۴۲۱۷	۷/۳۹۹۰	۰/۱۱۸۷	۰/۲۲۱۹	۰/۳۸۳۷

جدول (۳-۲۶) - مشابه جدول (۳-۲۵) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
دور آموزشی	۱۸۰	۰/۴۰۶۴	۰/۱۵۴۱	۳۲/۵۲۱۸	۰/۲۲۳۵	۰/۲۶۲۶	۰/۳۴۶۸	۴/۱۰۲۰	۰/۱۴۴۴	۰/۰۶۹۲
	۱۸۲	۰/۴۰۶۴	۰/۱۵۴۱	۳۱/۷۰۶۷	۰/۲۲۱۴	۰/۲۶۲۵	۰/۳۵۴۴	۵/۱۵۲۲	۰/۱۳۷۵	۰/۰۶۶۶
	۱۸۴	۰/۴۰۶۴	۰/۱۵۴۱	۳۱/۰۳۰۷	۰/۲۱۹۴	۰/۲۶۲۴	۰/۳۶۲۲	۶/۶۱۷۱	۰/۱۳۳۰	۰/۰۶۵۱
	۱۸۶	۰/۴۰۶۴	۰/۱۵۴۱	۳۰/۵۴۵۳	۰/۲۱۷۵	۰/۲۶۲۱	۰/۳۶۹۹	۷/۸۱۱۲	۰/۱۳۱۲	۰/۰۶۴۶
	۱۸۸	۰/۴۰۶۴	۰/۱۵۴۱	۳۰/۲۶۳۶	۰/۲۱۵۶	۰/۲۶۱۷	۰/۳۷۷۶	۸/۵۳۵۵	۰/۱۳۱۹	۰/۰۶۶۴

جدول (۳-۲۷) - مقادیر میانگین مربعات خطا مربوط به سری ارزیابی برای شبکه با الگوریتم آموزشی لونیبرگ - مارکوارت و تابع انتقال تانژانت سیگموئید با تعداد گره متفاوت در لایه‌ی پنهان و دور آموزشی مختلف با استفاده از توصیف‌کننده‌های روش سهم گروه

		تعداد گره در لایه‌ی پنهان									
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
دور آموزشی	۱۰	۰/۲۴۱۰	۲/۵۰۱۵	۱/۷۰۳۲	۰/۵۸۲۴	۱/۰۲۱۶	۱/۹۰۳۶	۰/۲۲۷۶	۱/۵۲۱۵	۱/۸۱۸۲	۰/۳۰۷۴
	۲۰	۰/۱۶۴۳	۲/۲۳۳۹	۱/۵۱۶۶	۰/۲۵۲۷	۲/۴۶۱۹	۰/۵۹۸۴	۰/۸۵۰۱	۲/۸۲۳۴	۰/۱۴۷۸	۰/۲۱۷۳
	۳۰	۰/۱۳۹۵	۲/۷۹۹۱	۱/۶۳۸۶	۰/۱۹۲۵	۲/۶۳۸۱	۰/۳۱۵۷	۰/۹۵۶۱	۲/۵۹۱۴	۰/۱۷۴۰	۰/۳۵۸۰
	۴۰	۰/۱۳۵۹	۲/۸۵۳۱	۱/۸۶۷۸	۰/۱۸۱۴	۲/۷۵۹۳	۰/۵۴۳۷	۵/۹۲۶۱	۱/۱۹۷۸	۰/۲۲۱۲	۰/۶۳۴۴
	۵۰	۰/۱۳۵۲	۲/۸۶۰۵	۱/۸۰۷۴	۰/۳۵۳۴	۲/۶۶۱۸	۰/۲۶۵۲	۴۵/۹۹۷۱	۱/۶۸۶۷	۰/۲۶۲۹	۰/۷۶۱۲
	۶۰	۰/۱۳۵۱	۲/۸۶۱۲	۱/۸۲۰۸	۰/۲۲۸۷	۲/۵۳۹۹	۰/۴۵۹۲	۸۸/۵۴۵۸	۱/۲۸۹۱	۰/۲۹۹۰	۰/۸۱۶۹
	۷۰	۰/۱۳۵۰	۲/۸۶۰۰	۱/۹۲۹۵	۰/۲۱۴۹	۲/۰۶۴۴	۰/۵۸۴۷	۹۰/۰۸۳۶	۱/۲۵۷۲	۰/۳۳۱۲	۰/۷۶۵۸
	۸۰	۰/۱۳۴۷	۲/۰۶۶۹	۱/۷۳۰۱	۰/۲۱۲۸	۱/۰۸۸۶	۰/۵۹۵۷	۵۶/۲۱۸۰	۱/۳۲۲۸	۰/۳۵۱۴	۰/۴۷۴۵
	۹۰	۰/۱۳۴۶	۱/۷۳۴۵	۱/۶۶۶۸	۰/۲۱۳۷	۱/۱۱۸۴	۰/۵۶۶۳	۱۲/۲۳۶۶	۱/۳۰۳۹	۰/۳۵۸۳	۰/۲۷۶۹
	۱۰۰	۰/۱۳۴۶	۱/۵۶۲۰	۱/۶۳۰۳	۰/۲۲۲۸	۱/۱۴۱۴	۰/۵۶۷۶	۱۱/۴۹۰۲	۱/۲۶۶۶	۰/۳۷۱۹	۰/۶۵۸۳
	۱۱۰	۰/۱۳۴۶	۱/۵۵۱۲	۱/۶۰۷۵	۰/۱۷۴۶	۱/۱۶۸۸	۰/۵۷۳۴	۱۱/۱۲۲۴	۲/۳۹۶۸	۰/۳۸۴۱	۱/۸۸۳۸
	۱۲۰	۰/۱۳۴۴	۱/۵۵۱۲	۱/۵۹۳۵	۰/۱۷۵۶	۱/۲۰۷۵	۰/۵۸۳۰	۱۱/۷۵۹۶	۲/۴۴۳۸	۰/۴۱۹۳	۱/۹۷۹۹
	۱۳۰	۰/۱۳۴۴	۱/۵۵۰۶	۱/۵۸۵۰	۰/۱۷۹۴	۱/۲۴۰۰	۰/۵۹۷۵	۸/۴۴۷۷	۱/۷۶۸۸	۰/۴۳۷۹	۳/۵۷۰۹
	۱۴۰	۰/۱۳۴۴	۱/۵۵۰۷	۱/۵۸۰۱	۰/۱۸۵۳	۱/۲۵۴۴	۰/۵۷۶۳	۶/۷۴۸۲	۱/۴۳۵۸	۰/۴۳۷۴	۴/۴۸۱۱
	۱۵۰	۰/۱۳۴۴	۱/۵۳۸۳	۱/۵۷۷۵	۰/۱۸۶۸	۱/۲۵۷۴	۰/۵۲۲۲	۶/۸۵۸۱	۱/۲۲۱۶	۰/۴۲۹۰	۴/۶۳۶۱
	۱۶۰	۰/۱۳۴۴	۱/۵۵۵۷	۱/۷۱۷۲	۰/۱۹۲۳	۱/۲۵۷۵	۰/۴۹۷۷	۶/۵۵۸۱	۱/۱۳۳۹	۰/۴۳۶۷	۴/۶۹۰۶
	۱۷۰	۰/۱۳۴۴	۱/۵۵۲۴	۱/۹۴۵۱	۰/۱۹۸۵	۱/۲۵۶۹	۰/۴۸۶۸	۶/۲۰۵۷	۱/۰۴۵۱	۰/۴۴۳۴	۴/۶۸۹۰
	۱۸۰	۰/۱۳۴۴	۱/۵۵۳۹	۱/۹۵۴۳	۰/۲۰۳۷	۱/۲۵۷۵	۰/۴۸۰۹	۶/۰۰۸۴	۰/۹۲۷۳	۰/۴۴۵۷	۴/۲۷۵۹
	۱۹۰	۱/۱۳۴۳	۱/۵۵۳۴	۱/۹۶۰۸	۰/۲۰۹۹	۱/۲۵۷۴	۰/۴۵۴۸	۵/۹۱۵۵	۰/۸۴۷۵	۰/۴۵۲۶	۴/۰۲۹۹
		.									
۲۰۰	۰/۱۳۴۳	۱/۵۵۱۱	۱/۹۶۵۹	۰/۲۱۰۲	۱/۲۵۶۳	۰/۴۱۴۷	۵/۸۷۱۵	۰/۸۳۶۶	۰/۴۵۶۹	۳/۷۸۱۹	

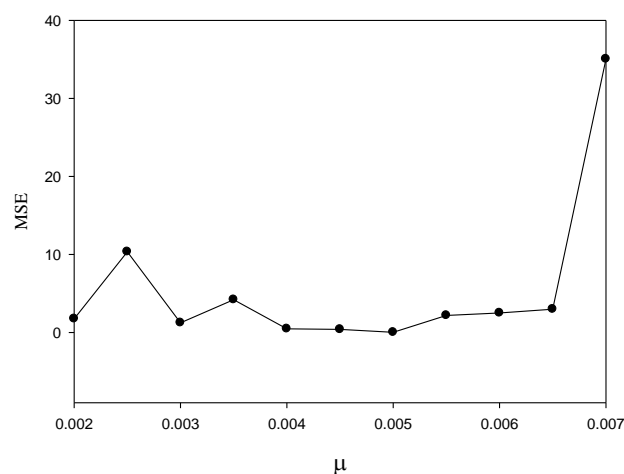
جدول (۳-۲۸) - مشابه جدول (۳-۲۷) حول نقطه‌ی بهینه با گام ۲

		تعداد گره در لایه‌ی پنهان									
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰	
دور آموزشی	۱۸۰	۰/۱۳۴۴	۱/۵۵۳۹	۱/۹۵۴۳	۰/۲۰۳۷	۱/۲۵۷۵	۰/۴۸۰۹	۶/۰۰۸۴	۰/۹۲۷۳	۰/۴۴۵۷	
	۱۸۲	۰/۱۳۴۴	۱/۵۵۳۶	۱/۹۵۵۷	۰/۲۰۴۸	۱/۲۵۶۹	۰/۴۸۰۱	۶/۰۷۸۸	۰/۹۰۶۷	۰/۴۴۵۲	
	۱۸۴	۱/۱۳۴۳	۱/۵۵۲۶	۱/۹۵۷۰	۰/۲۰۵۹	۱/۲۵۷۴	۰/۴۷۹۵	۵/۹۰۶۱	۰/۸۸۸۱	۰/۴۴۲۸	
	.										

۱۸۶	۰/۱۳۴۳	۱/۵۵۱۲	۱/۹۵۸۳	۰/۲۰۷۱	۱/۲۵۶۸	۰/۴۷۰۴	۵/۹۹۲۷	۰/۸۷۱۹	۰/۴۵۴۳
۱۸۸	۰/۱۳۴۳	۱/۵۵۴۵	۱/۹۵۹۶	۰/۲۰۸۳	۱/۲۵۶۲	۰/۴۸۱۸	۵/۷۴۹۱	۰/۸۵۸۲	۰/۴۵۷۶

نتایج موجود در این جداول نشان می‌دهد که شبکه‌ی عصبی با تابع آموزشی تنظیم بایزین، تابع انتقال لگاریتم سیگموئید، ۱۰ گره در لایه‌ی پنهان و ۱۳۶ دور آموزش دارای کمترین میزان MSE برای سری ارزیابی است.

در مرحله‌ی بعد برای تعیین مقدار بهینه‌ی  $\mu$ ، با استفاده از شبکه‌ی بهینه، مقدار  $\mu$  از ۰/۰۰۰۵ تا ۰/۱ با گام ۰/۰۰۰۵ تغییر داده شد و بر اساس کمترین مقدار MSE برای هر گام، مقدار بهینه‌ی  $\mu$ ، به دست آمد. نمودار مقدار میانگین مربعات خطا بر حسب  $\mu$  حول نقطه‌ی بهینه نیز در شکل (۳-۹) رسم شده است.



شکل (۳-۹) - نمودار میانگین مربعات خطای حاصل از سری ارزیابی بر حسب تغییرات پارامتر  $\mu$

توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی با استفاده از توصیف‌کننده‌های انتخاب شده به روش سهم گروه در جدول (۳-۲۹) ارائه شده است.

جدول (۳-۲۹) - توابع و مقادیر پارامترهای بهینه شده شبکه‌ی عصبی روش سهم گروه

Train br	تابع آموزش
logsig	تابع انتقال لایه‌ی پنهان
pure line	تابع انتقال لایه‌ی خروجی
۱۰	تعداد نرون لایه‌ی پنهان
۱۳	تعداد متغیرهای ورودی
۱۳۶	تعداد دوره‌های آموزش
۰/۰۰۵	پارامتر $\mu$
۰/۰۲۹۶	مقدار MSE سری ارزیابی

### ۳-۱-۸- ارزیابی مدل‌های غیر خطی

در این مرحله توانایی دو مدل شبکه‌ی عصبی مصنوعی بهینه شده با استفاده از توصیف‌کننده‌های منتخب مولکولی تئوری و یک مدل شبکه‌ی عصبی بهینه شده با استفاده از توصیف‌کننده‌های حاصل از روش سهم گروه‌ها برای پیش‌بینی ضریب تراکم‌پذیری همدمای داده‌های سری تست مورد بررسی قرار گرفت. لازم به ذکر است داده‌های سری تست در هیچ یک از مراحل مدل‌سازی دخالتی نداشته‌اند و همانطور که در قسمت‌های قبل اشاره گردید از ابتدا کنار گذاشته شدند. بنابراین چگونگی پیش‌بینی ضریب تراکم‌پذیری همدمای ۷ مولکول موجود در سری تست بر حسب دما و فشار می‌تواند قدرت پیش‌بینی مدل‌های به دست آمده را به خوبی بیان نماید.

### ۳-۱-۸-۱- ارزیابی با استفاده از اعتبارسنجی بیرونی

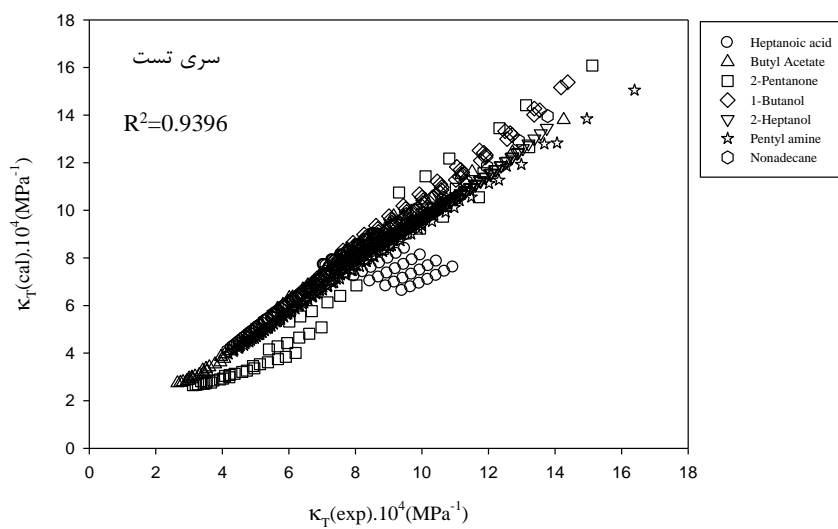
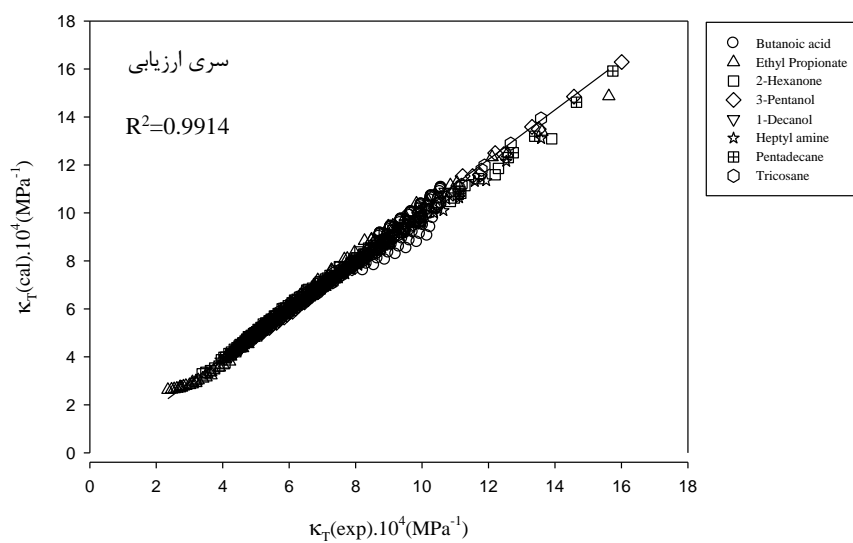
نمودارهای مقادیر پیش‌بینی شده بر حسب مقادیر تجربی برای دو سری ارزیابی و تست در شکل‌های (۳-۱۰) تا (۳-۱۲) رسم شده است. همچنین مقادیر پیش‌بینی شده ضریب تراکم‌پذیری

همدما برای ترکیبات سری تست توسط سه روش مذکور در قسمت پیوست آمده است. همچنین پارامترهای آماری مربوط به هر یک از شبکه‌های بهینه شده برای این دو سری در جدول (۳-۳۰) گردآوری شده است.

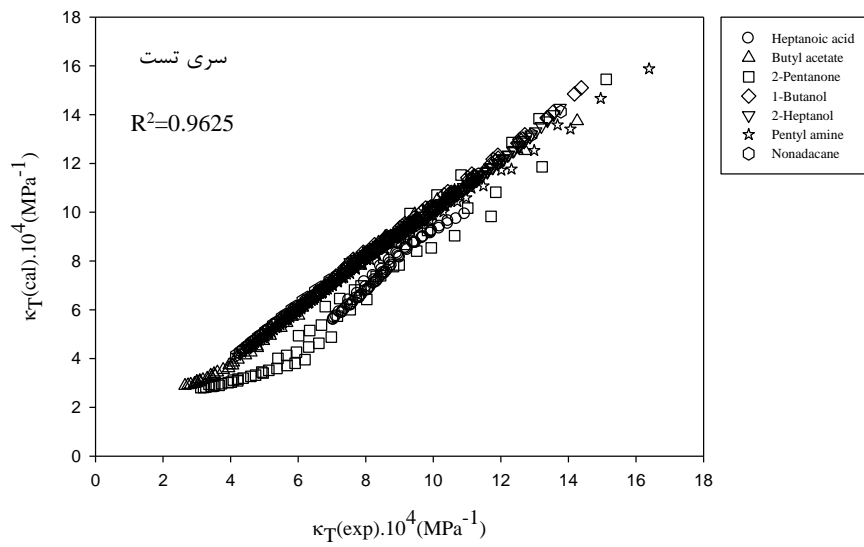
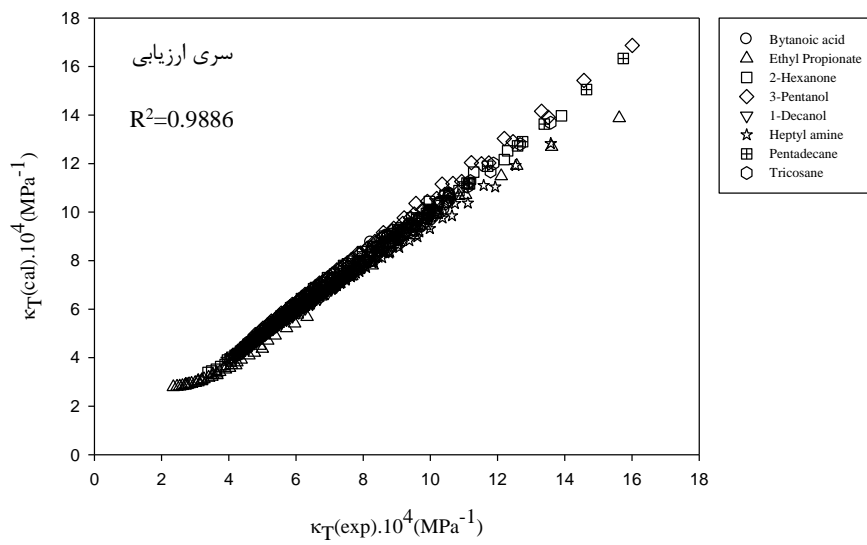
جدول (۳-۳۰) - پارامترهای آماری مربوط به هر یک از شبکه‌های بهینه شده برای سری ارزیابی و سری تست

		SR-ANN	CDFS-ANN	GCM-ANN
سری ارزیابی	R <sup>2</sup>	۰/۹۹۱۴	۰/۹۸۸۶	۰/۹۹۴۸
	MSE	۰/۰۴۹۵	۰/۰۶۳۷	۰/۰۲۹۶
سری تست	R <sup>2</sup>	۰/۹۳۹۶	۰/۹۶۲۵	۰/۹۷۰۵
	MSE	۰/۴۴۵۶	۰/۳۰۱۲	۰/۲۳۰۳

نمودار باقیمانده معیاری برای شایستگی مدل به دست آمده می‌باشد و اگر باقیمانده‌ها به طور یکنواخت حول محور افقی پراکنده باشند، نشان می‌دهد که مدل مناسبی به دست آمده و هیچ خطای سیستماتیکی وجود ندارد. شکل‌های (۳-۱۳) تا (۳-۱۵) نمودارهای باقیمانده را که از تفاضل مقدار تجربی ضریب تراکم‌پذیری همدما از مقدار محاسبه شده توسط مدل‌های مختلف غیرخطی به دست می‌آید، بر حسب مقدار تجربی نشان می‌دهد.

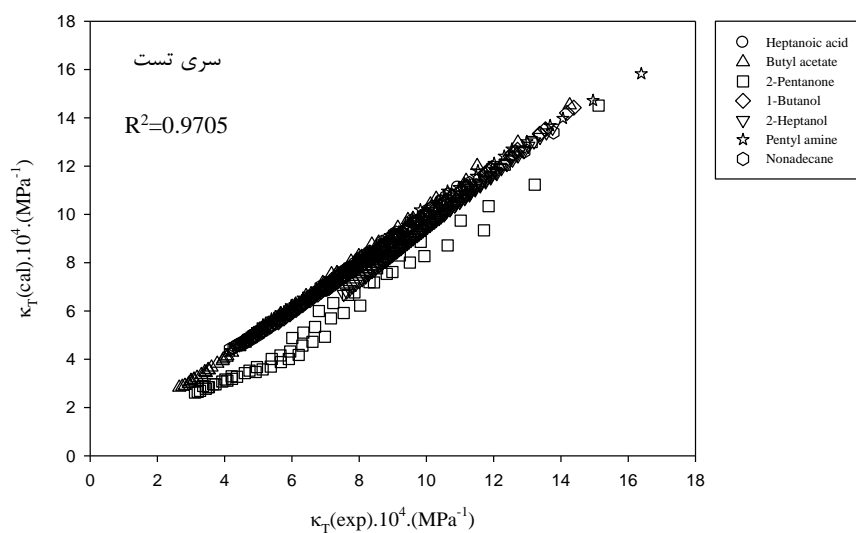
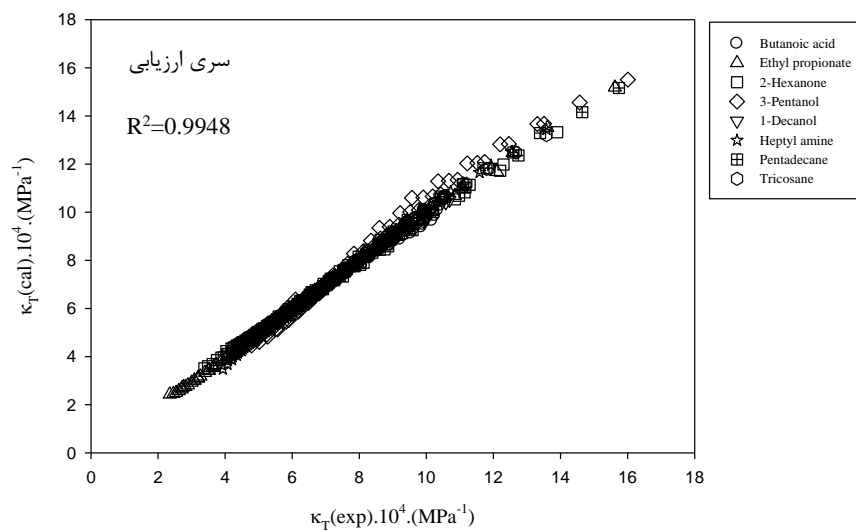


شکل (۱۰-۳) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدمما بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل SR-ANN

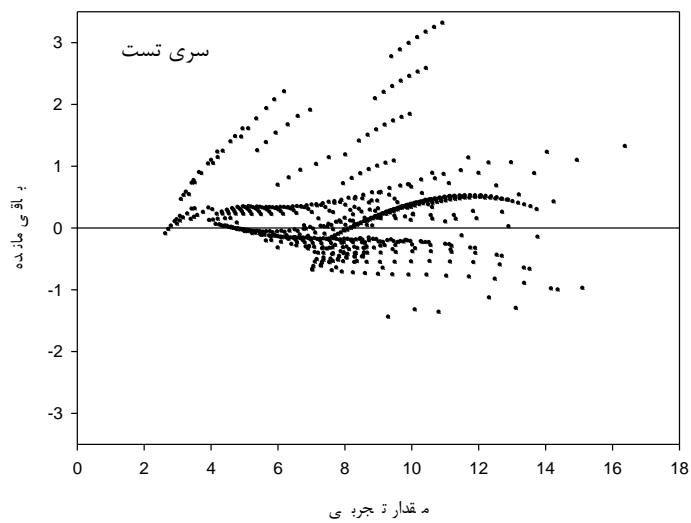
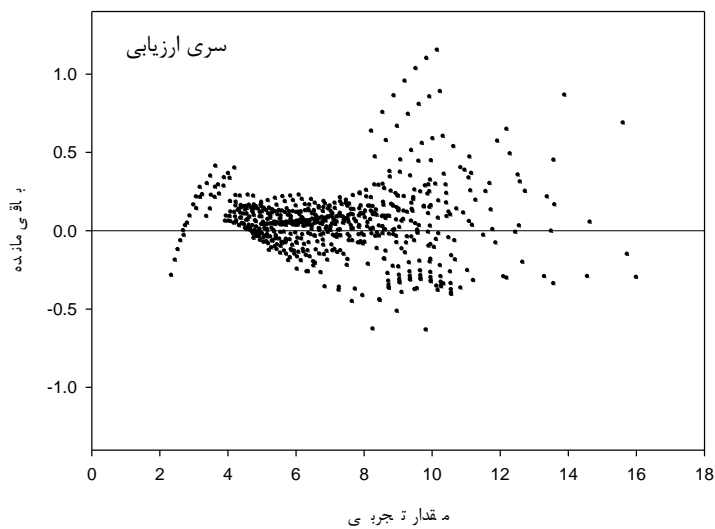


شکل (۱۱-۳) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری هم‌دما بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل CDFS-ANN

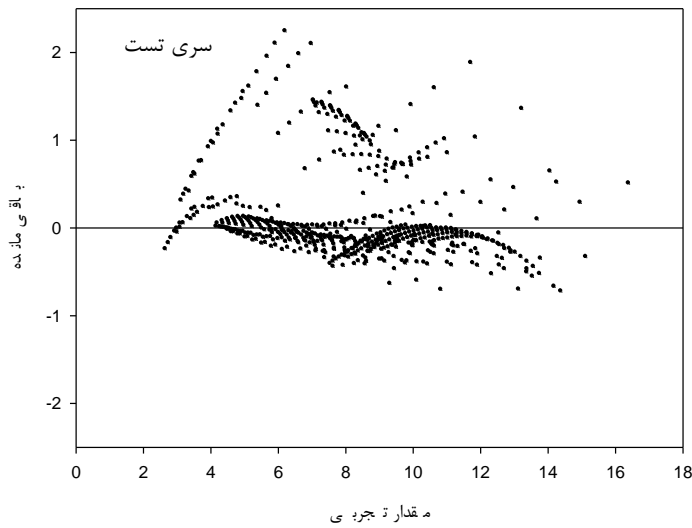
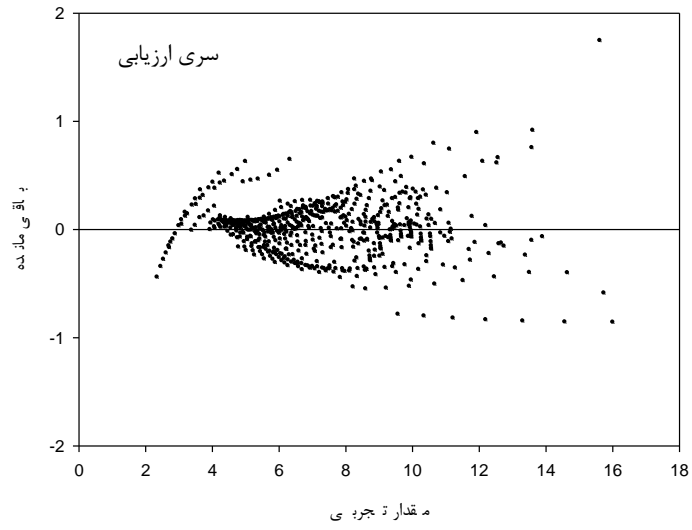




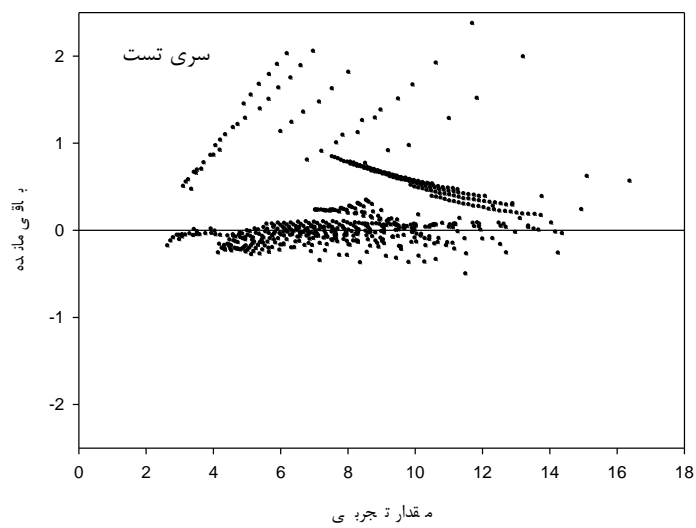
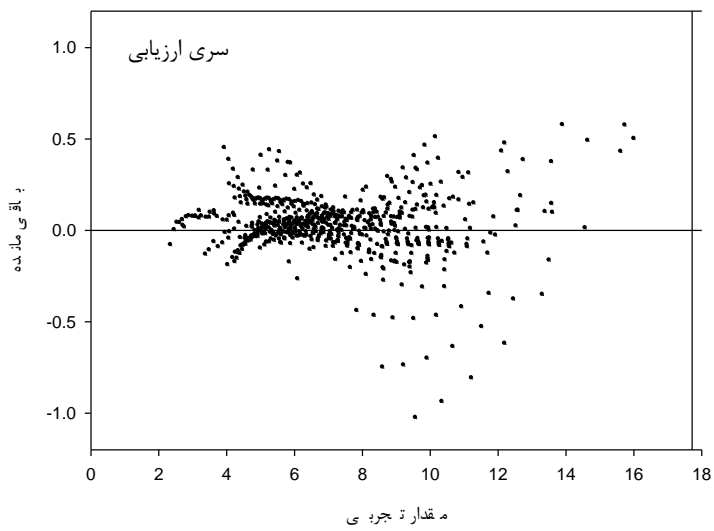
شکل (۱۲-۳) - نمودار مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدمما بر حسب مقادیر تجربی برای سری ارزیابی و سری تست در مدل GCM-ANN



شکل (۳-۱۳) - نمودار باقیمانده‌ها بر حسب مقادیر تجربی ضریب تراکم‌پذیری همدمما برای سری ارزیابی و سری تست در مدل SR-ANN



شکل (۳-۱۴) - نمودار باقیمانده‌ها بر حسب مقادیر تجربی ضریب تراکم‌پذیری همدم‌ها برای سری ارزیابی و سری تست در مدل CDFS-ANN



شکل (۳-۱۵) - نمودار باقیمانده‌ها بر حسب مقادیر تجزی ضریب تراکم‌پذیری هم‌دما برای سری ارزیابی و سری تست در مدل GCM-ANN

### ۳-۱-۸-۲- ارزیابی با استفاده از آزمون $\chi^2$ تصادفی

در این روش مقادیر متغیر وابسته در گستره داده‌های تجربی به صورت تصادفی تغییر داده شدند و شبکه‌های عصبی بهینه شده با استفاده از متغیرهای مستقل اصلی و مقادیر تصادفی از متغیر وابسته توسعه یافت. سپس همبستگی متغیرهای مستقل با متغیرهای وابسته توسط شاخص  $R^2$  مورد بررسی قرار گرفت. در جداول (۳-۳۱) تا (۳-۳۳) مقادیر ضریب رگرسیون برای سری ارزیابی و تست آورده شده است. مقادیر پایین  $R^2$  نشان‌دهنده‌ی عدم وجود همبستگی شانسی در مدل اصلی است.

جدول (۳-۳۱) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $\chi^2$  تصادفی در مدل SR-ANN

شماره‌ی تکرار آزمون	$R^2$ سری ارزیابی	$R^2$ سری تست
۱	۰/۱۶۰۷	۰/۰۰۷۹
۲	۰/۱۹۲۱	۰/۰۱۹۶
۳	۰/۰۴۰۸	۰/۰۵۱۰
۴	۰/۲۱۹۴	۰/۱۶۳۴
۵	۰/۰۰۶۱	۰/۰۱۷۲
۶	۰/۰۰۰۰	۰/۰۰۰۲
۷	۰/۰۱۱۸	۰/۱۷۴۲
۸	۰/۰۰۳۱	۰/۰۱۸۱
۹	۰/۰۶۰۴	۰/۰۰۷۶
۱۰	۰/۰۰۴۵	۰/۰۰۶۸

جدول (۳-۳۲) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $\chi^2$  تصادفی در مدل CDFS-ANN

شماره‌ی تکرار آزمون	$R^2$ سری ارزیابی	$R^2$ سری تست
۱	۰/۰۰۰۸	۰/۰۲۱۶
۲	۰/۲۷۴۳	۰/۳۴۰۰
۳	۰/۰۵۹۵	۰/۰۱۹۵
۴	۰/۱۸۲۰	۰/۰۳۴۵
۵	۰/۲۳۸۶	۰/۰۰۶۳
۶	۰/۱۸۱۶	۰/۱۴۲۲
۷	۰/۰۳۹۸	۰/۱۹۷۱
۸	۰/۲۳۰۸	۰/۰۰۰۲
۹	۰/۰۱۱۵	۰/۱۰۴۰
۱۰	۰/۰۱۲۱	۰/۱۱۷۰

جدول (۳-۳۳) - مقادیر  $R^2$  به دست آمده برای سری ارزیابی و سری تست از آزمون  $\chi^2$  تصادفی در مدل GCM-ANN

شماره‌ی تکرار آزمون	$R^2$ سری ارزیابی	$R^2$ سری تست
۱	۰/۲۲۷۶	۰/۰۰۰۰
۲	۰/۰۰۰۱	۰/۰۱۹۱
۳	۰/۱۱۴۸	۰/۰۱۶۳
۴	۰/۰۹۴۴	۰/۰۱۶۹
۵	۰/۱۸۴۳	۰/۰۹۳۸
۶	۰/۲۳۸۵	۰/۰۷۴۶
۷	۰/۰۵۰۹	۰/۴۳۲۸
۸	۰/۲۲۸۵	۰/۱۵۳۹
۹	۰/۱۳۰۶	۰/۰۷۵۲
۱۰	۰/۰۴۴۲	۰/۱۰۰۸

### ۳-۱-۹- بررسی ارتباط توصیف‌کننده‌های وارد شده در مدل با ضریب تراکم-

#### پذیری همدمای ترکیبات

در این قسمت یک بررسی اجمالی روی اثرات مختلف ۴ توصیف‌کننده‌ی تئوری انتخاب شده با راهبرد CDFS بر ضریب تراکم‌پذیری همدمای ترکیبات مورد مطالعه صورت گرفته است. این توصیف‌کننده‌ها شامل X0Av، Mor13m، Mor07m و Mor11m می‌باشد.

### ۳-۱-۹-۱- توصیف‌کننده‌ی X0Av

X0Av توصیف‌کننده‌ای از گروه اندیس‌های ارتباط مولکولی است. این اندیس‌ها دسته‌ای از توصیف‌کننده‌های توپولوژیکی هستند که اطلاعاتی راجع به ساختمان، اندازه و میزان شاخه‌ای شدن مولکول، نحوه‌ی اتصال اتم‌ها و نوع اتم‌های موجود در یک مولکول را در اختیار قرار می‌دهند. جهت محاسبه‌ی این توصیف‌کننده‌ها ابتدا باید ساختار مولکول را با استفاده از تئوری گراف شیمیایی رسم نمود.

تئوری گراف شیمیایی، شاخه‌ای از علم شیمی-ریاضی می‌باشد که عموماً با شیمی تئوری در ارتباط است. بر اساس این تئوری یک گراف نشان‌دهنده‌ی مجموعه‌ای از عناصر یک گروه و نحوه‌ی ارتباط بین آن‌ها می‌باشد. هر گراف از تعدادی رأس<sup>۱</sup> و یال<sup>۲</sup> تشکیل شده است، که یال‌ها خطوطی هستند که رأس‌ها را به یکدیگر متصل می‌کنند. یک دسته از گراف‌ها، گراف‌های مولکولی هستند، که در این گراف‌ها هر نقطه نشان‌دهنده‌ی یک اتم بوده و خطوط بین نقاط نیز نشان‌دهنده‌ی پیوند شیمیایی بین اتم‌هاست. معمولاً در این گراف‌ها، اتم هیدروژن را نشان نمی‌دهند. این اندیس‌ها اطلاعاتی شامل ساختمان، اندازه و میزان شاخه‌ای شدن مولکول، پیوندها، تعداد اتم‌ها و نوع اتم‌های

---

1. Vertex  
2. Edge

ناجور در مولکول را دارا می‌باشند. این اندیس‌ها اولین بار بوسیله راندیک<sup>۱</sup> در سال ۱۹۷۵ ارائه شد و سپس در سال ۱۹۷۶ توسط هال<sup>۲</sup> اصلاح گردید [۴۰]. این توصیف‌کننده‌ها با رابطه‌ی (۱-۳) قابل محاسبه هستند [۴۱ و ۴۲].

$$\chi^v = \sum_{i=1}^{N_s} \prod_{k=1}^{m+1} \left( \frac{1}{\delta_k^v} \right)^{\frac{1}{2}} \quad (1-3)$$

که در این رابطه  $\delta_k^v$ ، درجه رأس ظرفیت<sup>۳</sup> اتم  $k$ ام در گراف مولکولی است و به شکل رابطه‌ی زیر تعریف می‌شود:

$$\delta_k^v = \frac{(Z_k^v - h_k)}{(Z - Z_k^v - 1)} \quad (2-3)$$

که در آن  $Z$  عدد اتمی اتم  $k$ ،  $Z_k^v$  تعداد الکترون‌های لایه‌ی ظرفیت و  $h_k$  تعداد اتم‌های هیدروژن متصل به اتم  $k$  است. X0Av وارد شده در مدل دارای ضریب تأثیر مثبت و برابر با ۳۳/۶۲۹ بوده و بنابراین با افزایش آن  $K_T$  نیز افزایش می‌یابد.

### ۳-۱-۹-۲- توصیف‌کننده‌های 3D-MoRSE

سه توصیف‌کننده‌ی دیگر انتخاب شده در راهبرد CDFS، جز توصیف‌کننده‌های 3D-MoRSE، به معنای نمایش سه بعدی ساختار مولکول بر اساس تفرق الکترون، هستند. هدف توصیف‌کننده‌های 3D-MoRSE این است که ساختار سه بعدی مولکول را به یک کد ثابت تبدیل کنند. این عمل توسط

---

1. Randic  
2. Hall  
3. Valence vertex degree



معادله‌ی (۳-۳) که به منظور فراهم آوردن منحنی‌های نظری تفرق در مطالعات پراش الکترونی به کار رفته، انجام می‌گیرد [۴۳ و ۴۱].

(۳-۳)

$$Mor(S, A) = I(S, A) = \sum_{i=2}^N \sum_{j=1}^{i-1} A_i A_j \frac{\sin(Sr_{ij})}{Sr_{ij}}$$

در این معادله،  $I$  شدت الکترون پراکنده شده،  $A_i$  و  $A_j$  خاصیت اتمی اتم‌های  $i$  و  $j$ ،  $S$  بیانگر پراکندگی در جهات مختلف،  $N$  تعداد اتم‌ها و  $r_{ij}$  نیز فاصله‌ی بین دو اتم  $i$  و  $j$  را نشان می‌دهد. زاویه پراکندگی به صورت رابطه‌ی زیر تعریف می‌شود:

$$S = 4\pi \cdot \sin\left(\frac{v/2}{\lambda}\right) \quad (۴-۳)$$

در این رابطه،  $v$  زاویه‌ی پراکندگی و  $\lambda$  طول موج پرتو الکترون است [۴۳ و ۴۱]. با توجه به این توضیحات، Mor13m و Mor07m و Mor11m هر سه، توصیف‌کننده‌ی وزن دار شده با جرم اتمی می‌باشند، بنابراین در رابطه‌ی (۳-۳)  $A_i$  و  $A_j$  نشان‌دهنده‌ی جرم اتمی است، با این تفاوت که Mor13m و Mor07m هر دو دارای ضریب تأثیر مثبت و به ترتیب برابر ۳۶/۹۸۶ و ۱۹/۰۹۰ بوده و بنابراین با افزایش آن‌ها ضریب تراکم‌پذیری همدمای نیز افزایش می‌یابد. اما توصیف‌کننده‌ی سوم از این گروه یعنی Mor11m دارای ضریب تأثیر منفی بر ضریب تراکم‌پذیری همدمای و برابر با ۱۰/۹۲۱- است. احتمال داده می‌شود که علت متفاوت بودن ضریب تأثیر Mor11m با دو توصیف‌کننده‌ی دیگر از این گروه مربوط به زاویه‌ی پراکندگی  $S$  در رابطه‌ی (۳-۳) می‌باشد.

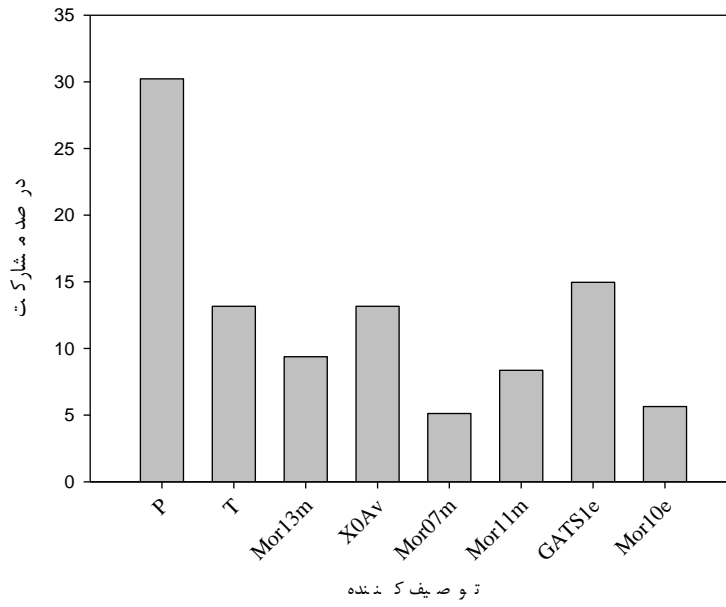
### ۳-۱-۱۰- بررسی میزان مشارکت توصیف‌کننده‌های منتخب شبکه عصبی

میزان مشارکت توصیف‌کننده‌های انتخاب شده به صورت زیر تعیین شد:

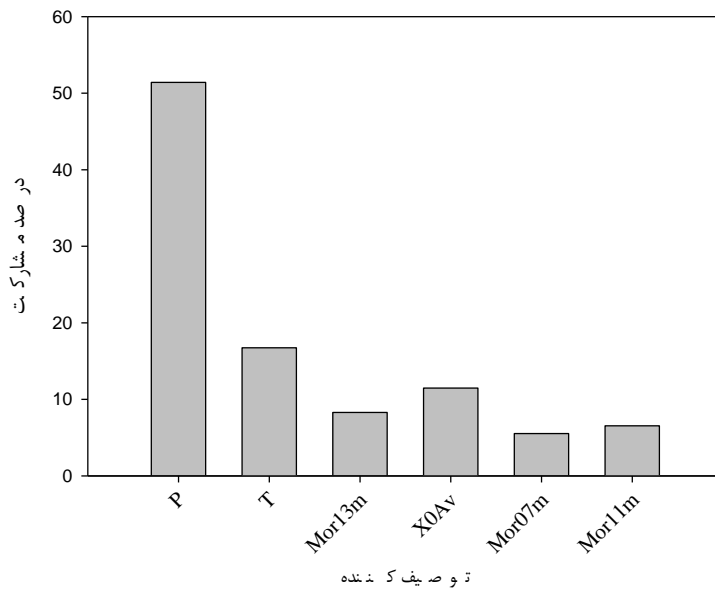
- وزن مربوط به توصیف‌کننده‌ی مورد نظر از شبکه بهینه حذف شد.
- مقدار متغیر وابسته با استفاده از توصیف‌کننده‌های باقی‌مانده برای هر یک از ترکیبات سری ارزیابی پیش‌بینی گردید.
- میانگین خطای مطلق (MAE) حاصل از ترکیبات سری ارزیابی محاسبه شد.
- مراحل بالا برای دیگر توصیف‌کننده‌ها نیز تکرار شد.
- سرانجام درصد مشارکت هر توصیف‌کننده توسط رابطه‌ی (۳-۵) محاسبه گردید:

$$C_i = 100 \frac{\Delta m_i}{\sum_{i=1}^N \Delta m_i} \quad (۳-۵)$$

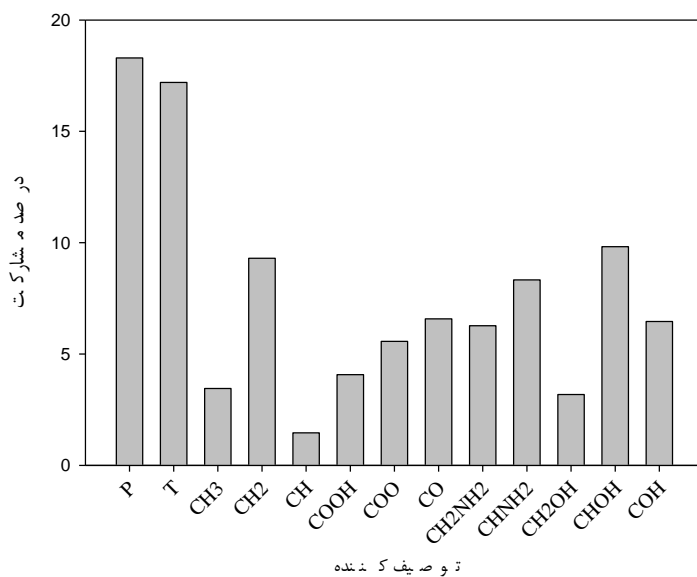
در این رابطه  $C_i$  درصد مشارکت توصیف‌کننده‌ی حذف شده  $i$ ام،  $N$  تعداد توصیف‌کننده‌ها در مدل و  $\Delta m_i$  میانگین خطای مطلق حاصل از سری ارزیابی در غیاب توصیف‌کننده‌ی  $i$ ام را نشان می‌دهد. بنابراین درصد مشارکت توصیف‌کننده‌های منتخب در سه مدل طراحی شده‌ی شبکه عصبی مصنوعی مطابق شکل‌های زیر است:



شکل (۳-۱۶)- درصد مشارکت توصیف کننده‌ها در مدل SR-ANN



شکل (۳-۱۷)- درصد مشارکت توصیف کننده‌ها در مدل CDFS-ANN



شکل (۳-۱۸) - درصد مشارکت توصیف‌کننده‌ها در مدل GCM-ANN

بر اساس شکل‌های به دست آمده، در روش رگرسیون مرحله‌ای و روش CDFS توصیف‌کننده‌ی فشار بیشترین اثر مشارکت را دارا می‌باشد. با این تفاوت که میزان مشارکت فشار در روش CDFS در مقایسه با روش رگرسیون مرحله‌ای به دلیل کاهش تعداد ورودی‌ها افزایش یافته است. این روند در داده‌های تجربی ضریب تراکم‌پذیری هم‌دما نیز مشاهده می‌گردد. یعنی میزان تغییرات  $K_T$  با تغییر فشار در دمای ثابت بیشتر از میزان تغییرات  $K_T$  با تغییرات دما در یک فشار ثابت می‌باشد. اما میزان مشارکت فشار در روش CDFS در مقایسه با روش رگرسیون مرحله‌ای به دلیل کاهش تعداد ورودی‌ها افزایش یافته است.

روش سهم گروه در مقایسه با دو روش قبل رفتار متفاوتی دارد. به طوری‌که اولاً میزان مشارکت دما و فشار در این روش تقریباً یکسان بوده و ثانیاً درصد مشارکت این دو توصیف‌کننده‌ی تجربی نسبت به دو روش ذکر شده بسیار کمتر شده است. در روش GCM-ANN با وجود اینکه تعداد گروه‌های ساختاری تشکیل‌دهنده ترکیبات به عنوان توصیف‌کننده‌های ورودی به شبکه‌ی عصبی مصنوعی در نظر گرفته شدند و بنابراین در ماتریس ورودی تعداد زیادی صفر وجود دارد، اما درصد مشارکت

توصیف‌کننده‌های ساختاری مقادیر قابل توجهی دارند. همانطور که از شکل (۳-۱۸) مشاهده می‌شود گروه‌هایی مانند  $-CH_2-$  و  $>CH-NH_2$  و  $>CH-OH$  بعد از دما و فشار درصد مشارکت بیشتری دارند.

### ۳-۱-۱۱- نتیجه‌گیری

از مقایسه‌ی نتایج به دست آمده از سه مدل طراحی شده برای پیش‌بینی ضریب تراکم‌پذیری همدمای سیالات آلی متنوع نکات زیر به دست می‌آید:

- روش سهم‌گروه در مقایسه با دو روش دیگر نتایج بهتری را ارائه می‌دهد. همانطور که در جدول (۳-۳۰) مشاهده می‌گردد، مدل حاصل از این روش دارای میانگین مربعات خطای کمتر و ضریب رگرسیون بالاتری است. با توجه به سادگی روش سهم‌گروه این نتایج از اهمیت بالایی برخوردارند. در واقع این روش پیچیدگی دو روش دیگر را نداشته و نتایج بهتری را نیز ارائه می‌نماید.
- همانطور که در فصل‌های قبل اشاره شد، یکی از اهداف استفاده از روش CDFS رسیدن به قدرت پیش‌بینی بهتر با تعداد توصیف‌کننده کمتر است. نتایج موجود در جدول (۳-۳۰) نشان می‌دهد که ضریب رگرسیون سری تست در روش CDFS در مقایسه با رگرسیون مرحله‌ای افزایش یافته و میزان MSE آن نیز رو به کاهش است. بنابراین با استفاده از روش CDFS توانستیم با تعداد کمتر توصیف‌کننده قدرت پیش‌بینی مدل را افزایش دهیم.
- با بررسی توصیف‌کننده‌های به دست آمده از دو روش رگرسیون مرحله‌ای و راهبرد CDFS در جداول (۳-۳) و (۳-۷)، مشاهده می‌شود که توصیف‌کننده‌های به دست آمده از روش CDFS همان ۶ توصیف‌کننده‌ی اولیه‌ی روش رگرسیون مرحله‌ای است. اما مقایسه‌ی نتایج نهایی به دست آمده از دو مدل نشان داد که اولویت حضور توصیف‌کننده‌ها در مدل غیرخطی شبکه‌ی

عصبی، اثر قابل ملاحظه‌ای بر پیش‌بینی پارامتر مورد بررسی دارد. همانطور که در جدول (۳-۶) دیده می‌شود در همه‌ی مدل‌ها دو توصیف‌کننده‌ی دما و فشار حضور دارند و بنابراین اولویت حضور یکسانی در مدل نهایی خواهند داشت. با وجود اینکه در تمامی این مدل‌ها توسط رگرسیون مرحله‌ای فشار به عنوان اولین توصیف‌کننده انتخاب شده است، اما مشاهده شد که اگر در مدل‌سازی شبکه‌ی عصبی دما در اولویت بالاتری نسبت به فشار قرار گیرد نتایج بهتری به دست خواهد آمد. از طرف دیگر اولویت دو توصیف‌کننده‌ی  $X0Av$  و  $Mor13m$  در روش CDFS متفاوت از رگرسیون مرحله‌ای است. بنابراین نتیجه‌گیری می‌شود که یکی از ویژگی‌های راهبرد CDFS این است که در تعیین اهمیت توصیف‌کننده‌ها در مدل، بهتر و دقیق‌تر عمل کرده ولی روش رگرسیون مرحله‌ای تا حدودی تصادفی رفتار می‌کند.

- همانطور که در فصل‌های قبل بیان شد ویژگی دیگر توصیف‌کننده‌های حاصل از CDFS این است که به نوع دسته‌بندی داده‌ها وابسته نیستند و حساسیت کمی به نحوه‌ی دسته‌بندی داده‌ها دارند. برای اثبات این ویژگی مهم راهبرد CDFS، توصیف‌کننده‌های این روش با چندین دسته‌بندی متفاوت مورد بررسی قرار گرفت و مدل‌های غیرخطی شبکه‌ی عصبی آن به دست آمد. به طور مثال مدل ۱۲ در جدول (۳-۶)، اگر با توصیف‌کننده‌های حاصل از رگرسیون مرحله‌ای که در همان جدول ذکر شده به وسیله‌ی شبکه‌ی عصبی مدل‌سازی غیر خطی شود، ضریب رگرسیون سری تست  $0/8739$  به دست خواهد آمد، اما اگر همین دسته‌بندی را با استفاده از توصیف‌کننده‌های حاصل از CDFS مدل‌سازی کنیم ضریب رگرسیون به دست آمده برای سری تست  $0/9722$  خواهد بود. این میزان تغییرات در ضریب رگرسیون نشان از این دارد که با استفاده از راهبرد CDFS می‌توان توصیف‌کننده‌هایی به دست آورد که برای انواع مختلف دسته‌بندی نتایج مشابهی ارائه می‌کند و این به دلیل وابسته نبودن این توصیف‌کننده‌ها به نحوه‌ی دسته‌بندی داده‌ها است.

- روند تغییرات ضریب تراکم‌پذیری هم‌دما نسبت به تغییرات دما و فشار در تمامی مولکول‌های استفاده شده در سری داده‌ها یکسان است. به این صورت که  $K_T$  در یک دمای ثابت همراه با افزایش فشار کاهش می‌یابد و در یک فشار ثابت با افزایش دما افزایش می‌یابد. اما چهار مولکول پنتانویک اسید، هگزانویک اسید، هپتانویک اسید و اکتانویک اسید روند متفاوتی دارند، به طوریکه با افزایش فشار در یک دمای ثابت میزان  $K_T$  افزایش می‌یابد اما تغییرات آن با دما در فشار ثابت مشابه مولکول‌های دیگر است. نکته‌ی قابل توجه این است که روند متفاوت این چهار سیال از گروه اسیدها، در داده‌های پیش‌بینی شده به روش CDFS و روش سهم‌گروه قابل مشاهده است ولی مدل حاصل از رگرسیون مرحله‌ای تفاوت این مولکول‌ها را تشخیص نمی‌دهد و روند تغییرات  $K_T$  را مانند سایر مولکول‌ها پیش‌بینی می‌کند. این نتایج می‌رساند که شبکه‌ی عصبی مصنوعی طراحی شده با توصیف‌کننده‌های حاصل از روش CDFS و سهم‌گروه در مقایسه با شبکه‌ی عصبی حاصل از رگرسیون مرحله‌ای هوشمندانه‌تر رفتار می‌کند.
- بررسی نمودارهای باقیمانده در شکل‌های (۳-۱۳) تا (۳-۱۵) نشان می‌دهد که گستره‌ی باقیمانده‌های سری تست در مدل حاصل از رگرسیون مرحله‌ای وسیع‌تر است، به این معنا که میزان خطا در این روش نسبت به دو روش دیگر بیشتر است. بنابراین ارزیابی مدل توسط نمودارهای باقیمانده نیز تأیید می‌کند که دو روش CDFS و سهم‌گروه عملکرد بهتری نسبت به روش رگرسیون مرحله‌ای دارند.

### ۳-۱-۱۲- آینده‌نگری

- پژوهشگران می‌توانند نتایج حاصل از این مطالعات را برای پیش‌بینی خواص ترمودینامیکی یا انتقالی دیگر به کار ببرند.
- در این تحقیق توصیف‌کننده‌های روش CDFS از طریق رگرسیون مرحله‌ای بر روی دسته‌بندی‌های مختلف به دست آمده است و توصیف‌کننده‌های حاصل از آن به عنوان ورودی به شبکه‌ی عصبی داده شدند. چون در روش رگرسیون مرحله‌ای تنها برهمکنش‌های خطی بین پارامترها لحاظ می‌شود، می‌توان روش‌هایی همچون الگوریتم ژنتیک و اجتماع مورچگان<sup>۱</sup> را برای انتخاب توصیف‌کننده به روش CDFS به کار گرفت.
- با توجه به توانایی قابل توجه شبکه‌ی عصبی مصنوعی بر مبنای سهم گروه‌ها، می‌توان خواص ترمودینامیکی و انتقالی مختلف را برای مولکول‌های بزرگ به این روش محاسبه نمود.
- می‌توان به جای شبکه‌ی عصبی مصنوعی برای مدل‌سازی غیرخطی از روش‌های دیگر مانند ماشین‌های بردار پشتیبان (SVM)<sup>۲</sup> و حداقل مربعات ماشین‌های بردار پشتیبان (LS-SVM)<sup>۳</sup> استفاده نمود

---

1. Ant Colony Optimization  
2. Support Vector Machines  
3. Least Squares Support Vector Machines



## پیوست

جدول (۱-پ): مقایسه‌ی مقادیر تجربی و مقادیر پیش‌بینی شده‌ی ضریب تراکم‌پذیری همدمما برای ترکیبات سری تست

هپتانویک اسید	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۷/۰۴۱۵	۷/۷۰۷۶	۵/۵۸۵۹	۶/۸۱۲۷
۲	۷/۰۵۰۷	۷/۷۰۲۴	۵/۶۰۳۰	۶/۸۲۱۸
۳	۷/۰۷۸۴	۷/۶۸۷۰	۵/۶۵۴۰	۶/۸۴۹۳
۴	۷/۱۲۴۶	۷/۶۶۱۴	۵/۷۳۸۶	۶/۸۹۵۲
۵	۷/۲۱۷۲	۷/۶۱۰۹	۵/۹۰۶۱	۶/۹۸۷۲
۶	۷/۴۹۶۱	۷/۴۶۴۱	۶/۳۹۳۰	۷/۲۶۵۵
۷	۷/۹۶۵۰	۷/۲۳۳۰	۷/۱۳۹۴	۷/۷۴۰۹
۸	۸/۴۳۹۴	۷/۰۱۶۶	۷/۷۸۶۱	۸/۲۳۷۵
۹	۸/۹۱۹۷	۶/۸۱۲۶	۸/۳۱۸۵	۸/۷۶۳۷
۱۰	۹/۴۰۶۳	۶/۶۱۹۲	۸/۷۲۹۴	۹/۳۲۸۴
۱۱	۷/۳۰۰۵	۷/۹۰۳۰	۵/۸۷۲۶	۷/۰۷۷۶
۱۲	۷/۳۰۹۷	۷/۸۹۷۶	۵/۸۸۹۳	۷/۰۸۷۱
۱۳	۷/۳۳۷۴	۷/۸۸۱۵	۵/۹۳۹۲	۷/۱۱۵۹
۱۴	۷/۳۸۳۶	۷/۸۵۴۷	۶/۰۲۲۱	۷/۱۶۳۸
۱۵	۷/۴۷۶۱	۷/۸۰۱۸	۶/۱۸۶۱	۷/۲۵۹۷
۱۶	۷/۷۵۵۰	۷/۶۴۸۲	۶/۶۶۲۷	۷/۵۴۸۹
۱۷	۸/۲۲۴۰	۷/۴۰۷۰	۷/۳۹۲۳	۸/۰۳۹۱
۱۸	۸/۶۹۸۸	۷/۱۸۱۶	۸/۰۲۲۲	۸/۵۴۶۴
۱۹	۹/۱۷۹۶	۶/۹۶۹۳	۸/۵۳۸۰	۹/۰۷۸۹
۲۰	۹/۶۶۶۸	۶/۷۶۸۵	۸/۹۳۲۳	۹/۶۴۵۵
۲۱	۷/۵۵۶۹	۸/۱۰۶۹	۶/۱۶۸۳	۷/۳۲۷۶
۲۲	۷/۵۶۶۱	۸/۱۰۱۲	۶/۱۸۴۶	۷/۳۳۷۶
۲۳	۷/۵۹۳۸	۸/۰۸۴۲	۶/۲۳۳۳	۷/۳۶۷۶
۲۴	۷/۶۴۰۰	۸/۰۵۶۲	۶/۳۱۴۰	۷/۴۱۷۶
۲۵	۷/۷۳۲۵	۸/۰۰۰۸	۶/۴۷۳۹	۷/۵۱۷۶
۲۶	۸/۰۱۱۴	۷/۸۴۰۰	۶/۹۳۸۴	۷/۸۱۸۱
۲۷	۸/۴۸۰۶	۷/۵۸۸۰	۷/۶۴۸۲	۸/۳۲۴۲
۲۸	۸/۹۵۵۷	۷/۳۵۳۰	۸/۲۵۸۸	۸/۸۴۳۱
۲۹	۹/۴۳۷۰	۷/۱۳۲۲	۸/۷۵۵۶	۹/۳۸۳۱
۳۰	۹/۹۲۵۰	۶/۹۲۳۴	۹/۱۳۱۳	۹/۹۵۲۶

ادامہی جدول (۱-پ)

ہیٹانویک اسید	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۳۱	۷/۸۱۰۷	۸/۳۱۹۵	۶/۴۷۵۱	۷/۵۶۴۴
۳۲	۷/۸۱۹۹	۸/۳۱۳۵	۶/۴۹۰۹	۷/۵۷۴۸
۳۳	۷/۸۴۷۶	۸/۲۹۵۷	۶/۵۳۸۱	۷/۶۰۶۲
۳۴	۷/۸۹۳۸	۸/۲۶۶۳	۶/۶۱۶۵	۷/۶۵۸۳
۳۵	۷/۹۸۶۳	۸/۲۰۸۲	۶/۷۷۱۶	۷/۷۶۲۶
۳۶	۸/۲۶۵۲	۸/۰۳۹۷	۷/۲۲۱۹	۸/۰۷۴۹
۳۷	۸/۷۳۴۷	۷/۷۷۶۳	۷/۹۰۹۰	۸/۵۹۷۷
۳۸	۹/۲۱۰۱	۷/۵۳۱۲	۸/۴۹۷۶	۹/۱۲۹۳
۳۹	۹/۶۹۲۰	۷/۳۰۱۳	۸/۹۷۳۰	۹/۶۷۷۶
۴۰	۱۰/۱۸۰۷	۷/۰۸۴۳	۹/۳۲۸۰	۱۰/۲۵۱۳
۴۱	۸/۰۶۱۹	۸/۵۴۱۴	۶/۷۹۵۱	۷/۷۸۹۹
۴۲	۸/۰۷۱۱	۸/۵۳۵۱	۶/۸۱۰۳	۷/۸۰۰۸
۴۳	۸/۰۹۸۸	۸/۵۱۶۴	۶/۸۵۵۹	۷/۸۳۳۵
۴۴	۸/۱۴۵۰	۸/۴۸۵۵	۶/۹۳۱۵	۷/۸۸۷۹
۴۵	۸/۲۳۷۵	۸/۴۲۴۵	۷/۰۸۱۱	۷/۹۹۶۵
۴۶	۸/۵۱۶۵	۸/۲۴۷۸	۷/۵۱۵۳	۸/۳۲۱۱
۴۷	۸/۹۸۶۲	۷/۹۷۲۳	۸/۱۷۶۶	۸/۸۶۱۳
۴۸	۹/۴۶۲۱	۷/۷۱۶۵	۸/۷۴۰۴	۹/۴۰۶۴
۴۹	۹/۹۴۴۶	۷/۴۷۷۰	۹/۱۹۱۹	۹/۹۶۴۲
۵۰	۱۰/۴۳۴۱	۷/۲۵۱۴	۹/۵۲۳۹	۱۰/۵۴۳۱
۵۱	۸/۳۱۰۵	۸/۷۷۳۰	۷/۱۳۰۵	۸/۰۰۶۱
۵۲	۸/۳۱۹۷	۸/۷۶۶۴	۷/۱۴۵۱	۸/۰۱۷۴
۵۳	۸/۳۴۷۴	۸/۷۴۶۷	۷/۱۸۸۸	۸/۰۵۱۵
۵۴	۸/۳۹۳۶	۸/۷۱۴۲	۷/۲۶۱۲	۸/۱۰۸۱
۵۵	۸/۴۸۶۲	۸/۶۵۰۱	۷/۴۰۴۶	۸/۲۲۱۲
۵۶	۸/۷۶۵۳	۸/۴۶۴۸	۷/۸۲۰۷	۸/۵۵۸۴
۵۷	۹/۲۳۵۳	۸/۱۷۶۳	۸/۴۵۳۱	۹/۱۱۶۷
۵۸	۹/۷۱۱۷	۷/۹۰۹۲	۸/۹۸۹۲	۹/۶۷۶۳
۵۹	۱۰/۱۹۴۸	۷/۶۵۹۷	۹/۴۱۴۲	۱۰/۲۴۴۵
۶۰	۱۰/۶۸۵۱	۷/۴۲۴۹	۹/۷۲۰۷	۱۰/۸۲۹۶
۶۱	۸/۵۵۶۶	۹/۰۱۴۸	۷/۴۸۳۴	۸/۲۱۴۷
۶۲	۸/۵۶۵۸	۹/۰۰۷۹	۷/۴۹۷۳	۸/۲۲۶۵

ادامه‌ی جدول (۱-پ)

هپتانویک اسید	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۶۳	۸/۵۹۳۵	۸/۹۸۷۲	۷/۵۳۸۹	۸/۲۶۲۰
۶۴	۸/۶۳۹۷	۸/۹۵۳۰	۷/۶۰۷۸	۸/۳۲۱۰
۶۵	۸/۷۳۲۳	۸/۸۸۵۹	۷/۷۴۴۳	۸/۴۳۸۶
۶۶	۹/۰۱۱۶	۸/۶۹۱۰	۸/۱۴۰۲	۸/۷۸۸۷
۶۷	۹/۴۸۱۹	۸/۳۸۸۸	۸/۷۴۰۴	۹/۳۶۵۹
۶۸	۹/۹۵۸۸	۸/۱۰۹۷	۹/۲۴۵۹	۹/۹۴۰۷
۶۹	۱۰/۴۴۲۶	۷/۸۴۹۵	۹/۶۴۱۶	۱۰/۵۲۰۲
۷۰	۱۰/۹۳۳۸	۷/۶۰۵۲	۹/۹۲۰۲	۱۱/۱۱۲۴

بوتیل استات	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۸/۸۹	۹/۲۱	۸/۷۶	۹/۰۴
۲	۷/۵۶	۷/۸۸	۷/۶۹	۷/۶۸
۳	۶/۵۷	۶/۸۶	۶/۶۲	۶/۷۰
۴	۵/۸۲	۶/۰۳	۵/۷۲	۵/۹۳
۵	۴/۷۸	۴/۸۱	۴/۵۱	۴/۸۴
۶	۳/۹۶	۳/۸۵	۳/۷۳	۴/۰۰
۷	۳/۴۱	۳/۲۸	۳/۳۰	۳/۴۷
۸	۳/۰۰	۲/۹۴	۳/۰۵	۳/۱۱
۹	۲/۶۵	۲/۷۴	۲/۸۹	۲/۸۳
۱۰	۹/۶۰	۹/۸۲	۹/۶۲	۹/۸۲
۱۱	۸/۰۷	۸/۳۷	۸/۳۵	۸/۲۴
۱۲	۶/۹۵	۷/۲۴	۷/۱۲	۷/۱۳
۱۳	۶/۱۱	۶/۳۴	۶/۰۹	۶/۲۷
۱۴	۴/۹۷	۵/۰۱	۴/۷۳	۵/۰۶
۱۵	۴/۰۸	۳/۹۶	۳/۸۴	۴/۱۳
۱۶	۳/۵۱	۳/۳۴	۳/۳۷	۳/۵۶
۱۷	۳/۰۹	۲/۹۸	۳/۰۹	۳/۱۷
۱۸	۲/۷۴	۲/۷۶	۲/۹۱	۲/۸۷
۱۹	۱۰/۲۹	۱۰/۴۹	۱۰/۴۵	۱۰/۶۶
۲۰	۸/۵۸	۸/۸۹	۸/۹۹	۸/۸۴
۲۱	۷/۳۴	۷/۶۶	۷/۶۱	۷/۵۸
۲۲	۶/۴۲	۶/۶۸	۶/۴۶	۶/۶۲
۲۳	۵/۱۷	۵/۲۳	۴/۹۴	۵/۲۹

ادامہی جدول (۱-پ)

بوتیل استات	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۲۴	۴/۲۲	۴/۰۹	۳/۹۶	۴/۲۸
۲۵	۳/۶۱	۳/۴۱	۳/۴۳	۳/۶۶
۲۶	۳/۱۸	۳/۰۲	۳/۱۲	۳/۲۴
۲۷	۲/۸۲	۲/۷۹	۲/۹۳	۲/۹۱
۲۸	۱۱/۵۱	۱۱/۶۱	۱۱/۷۰	۱۲/۰۱
۲۹	۹/۴۵	۹/۷۷	۹/۹۲	۹/۷۷
۳۰	۷/۹۹	۸/۳۷	۸/۳۱	۸/۲۸
۳۱	۶/۹۲	۷/۲۵	۶/۹۹	۷/۱۷
۳۲	۵/۴۹	۵/۵۹	۵/۲۶	۵/۶۴
۳۳	۴/۴۴	۴/۳۰	۴/۱۴	۴/۵۱
۳۴	۳/۷۸	۳/۵۳	۳/۵۴	۳/۸۲
۳۵	۳/۳۱	۳/۰۹	۳/۱۸	۳/۳۵
۳۶	۲/۹۳	۲/۸۳	۲/۹۷	۲/۹۹
۳۷	۱۲/۷۲	۱۲/۴۳	۱۲/۵۲	۱۲/۹۸
۳۸	۱۰/۱۲	۱۰/۴۲	۱۰/۵۱	۱۰/۴۲
۳۹	۸/۳۸	۸/۸۹	۸/۷۶	۸/۷۵
۴۰	۷/۱۸	۷/۶۷	۷/۳۴	۷/۵۳
۴۱	۵/۶۶	۵/۸۶	۵/۴۶	۵/۸۷
۴۲	۴/۶۰	۴/۴۵	۴/۲۶	۴/۶۵
۴۳	۳/۹۴	۳/۶۲	۳/۶۱	۳/۹۳
۴۴	۳/۴۴	۳/۱۴	۳/۲۳	۳/۴۳
۴۵	۲/۹۹	۲/۸۵	۲/۹۹	۳/۰۵
۴۶	۱۴/۲۶	۱۳/۸۰	۱۳/۷۴	۱۴/۵۲
۴۷	۱۱/۱۸	۱۱/۴۹	۱۱/۳۶	۱۱/۴۰
۴۸	۹/۱۵	۹/۷۶	۹/۳۸	۹/۴۵
۴۹	۷/۷۶	۸/۳۷	۷/۸۱	۸/۰۵
۵۰	۶/۰۱	۶/۳۲	۵/۷۶	۶/۱۸
۵۱	۴/۷۹	۴/۷۱	۴/۴۳	۴/۸۴
۵۲	۴/۰۶	۳/۷۷	۳/۷۲	۴/۰۷
۵۳	۳/۵۳	۳/۲۲	۳/۲۹	۳/۵۵
۵۴	۳/۰۹	۲/۹۰	۳/۰۳	۳/۱۴

ادامہی جدول (۱-پ)

پنٹانوں -۲	مقادیہ تجریہی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۹/۳۱	۱۰/۷۵	۹/۹۴	۹/۱۸
۲	۸/۵۳	۸/۹۹	۸/۱۴	۷/۷۶
۳	۷/۸۶	۷/۷۵	۶/۹۸	۶/۷۷
۴	۶/۸۰	۶/۷۷	۶/۱۳	۶/۰۰
۵	۶/۰۱	۵/۳۳	۴/۹۴	۴/۸۸
۶	۵/۴۰	۴/۱۶	۴/۰۱	۴/۰۱
۷	۴/۹۲	۳/۴۵	۳/۴۴	۳/۴۷
۸	۴/۰۸	۳/۰۴	۳/۱۱	۳/۱۱
۹	۳/۵۳	۲/۸۰	۲/۹۲	۲/۸۴
۱۰	۳/۱۲	۲/۶۵	۲/۸۰	۲/۶۲
۱۱	۱۰/۱۱	۱۱/۴۳	۱۰/۷۱	۹/۹۴
۱۲	۹/۲۱	۹/۴۹	۸/۶۸	۸/۳۰
۱۳	۸/۴۴	۸/۱۳	۷/۴۰	۷/۱۸
۱۴	۷/۲۳	۷/۰۹	۶/۴۶	۶/۳۳
۱۵	۶/۳۴	۵/۵۴	۵/۱۵	۵/۱۰
۱۶	۵/۶۶	۴/۲۹	۴/۱۳	۴/۱۶
۱۷	۵/۱۳	۳/۵۳	۳/۵۲	۳/۵۸
۱۸	۴/۲۲	۳/۰۸	۳/۱۵	۳/۱۹
۱۹	۳/۳۶	۲/۸۲	۲/۹۴	۲/۸۹
۲۰	۳/۲۰	۲/۶۷	۲/۸۲	۲/۶۵
۲۱	۱۰/۸۲	۱۲/۱۷	۱۱/۵۲	۱۰/۷۴
۲۲	۹/۸۳	۱۰/۰۲	۹/۲۵	۸/۸۶
۲۳	۸/۹۹	۸/۵۴	۷/۸۳	۷/۶۱
۲۴	۷/۶۷	۷/۴۲	۶/۸۱	۶/۶۷
۲۵	۶/۶۹	۵/۷۶	۵/۳۷	۵/۳۴
۲۶	۵/۹۶	۴/۴۲	۴/۲۶	۴/۳۲
۲۷	۵/۳۷	۳/۶۱	۳/۵۹	۳/۷۰
۲۸	۴/۳۷	۳/۱۳	۳/۲۰	۳/۲۸
۲۹	۳/۷۳	۲/۸۴	۲/۹۷	۲/۹۶
۳۰	۳/۲۷	۲/۶۸	۲/۸۳	۲/۷۰
۳۱	۱۲/۳۳	۱۳/۴۴	۱۲/۸۵	۱۲/۰۶
۳۲	۱۱/۰۲	۱۰/۹۰	۱۰/۱۷	۹/۷۴

ادامہی جدول (۱-پ)

پنتانون	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۳۳	۹/۹۴	۹/۲۳	۸/۵۳	۸/۲۷
۳۴	۸/۳۱	۷/۹۷	۷/۳۷	۷/۱۹
۳۵	۷/۱۶	۶/۱۳	۵/۷۴	۵/۶۹
۳۶	۶/۳۱	۴/۶۵	۴/۴۷	۴/۵۶
۳۷	۵/۶۷	۳/۷۴	۳/۷۲	۳/۸۸
۳۸	۴/۶۰	۳/۲۱	۳/۲۷	۳/۴۲
۳۹	۳/۹۳	۲/۸۹	۳/۰۱	۳/۰۷
۴۰	۳/۴۴	۲/۷۱	۲/۸۶	۲/۷۸
۴۱	۱۳/۱۳	۱۴/۴۱	۱۳/۸۳	۱۳/۰۰
۴۲	۱۱/۸۵	۱۱/۵۶	۱۰/۸۲	۱۰/۳۴
۴۳	۱۰/۶۳	۹/۷۳	۹/۰۳	۸/۷۱
۴۴	۸/۸۲	۸/۳۷	۷/۷۷	۷/۵۳
۴۵	۷/۵۴	۶/۴۰	۶/۰۰	۵/۹۲
۴۶	۶/۶۱	۴/۸۱	۴/۶۳	۴/۷۲
۴۷	۵/۹۱	۳/۸۴	۳/۸۱	۴/۰۱
۴۸	۴/۷۴	۳/۲۷	۳/۳۲	۳/۵۳
۴۹	۴/۰۲	۲/۹۲	۳/۰۴	۳/۱۶
۵۰	۳/۵۰	۲/۷۳	۲/۸۷	۲/۸۴
۵۱	۱۵/۱۲	۱۶/۰۸	۱۵/۴۵	۱۴/۵۰
۵۲	۱۳/۲۲	۱۲/۶۵	۱۱/۸۶	۱۱/۲۳
۵۳	۱۱/۷۱	۱۰/۵۵	۹/۸۳	۹/۳۴
۵۴	۹/۵۱	۹/۰۳	۸/۴۰	۸/۰۱
۵۵	۸/۰۳	۶/۸۴	۶/۴۳	۶/۲۲
۵۶	۶/۹۸	۵/۰۸	۴/۸۸	۴/۹۳
۵۷	۶/۲۰	۴/۰۰	۳/۹۶	۴/۱۸
۵۸	۴/۹۶	۳/۳۶	۳/۴۱	۳/۶۸
۵۹	۴/۲۱	۲/۹۸	۳/۰۹	۳/۳۰
۶۰	۳/۶۶	۲/۷۶	۲/۹۰	۲/۹۶
۱- بوتانون	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۸/۶۰	۸/۸۳	۸/۹۵	۸/۵۴
۲	۸/۵۳	۸/۷۶	۸/۸۵	۸/۴۶
۳	۸/۲۳	۸/۴۳	۸/۴۷	۸/۱۵

ادامہی جدول (۱-پ)

۱- بوتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۴	۷/۸۹	۸/۰۸	۸/۰۵	۷/۷۹
۵	۷/۵۷	۷/۷۵	۷/۶۹	۷/۴۷
۶	۷/۲۸	۷/۴۵	۷/۳۶	۷/۱۸
۷	۷/۰۱	۷/۱۸	۷/۰۷	۶/۹۲
۸	۶/۷۷	۶/۹۲	۶/۸۱	۶/۶۸
۹	۶/۵۴	۶/۶۹	۶/۵۷	۶/۴۵
۱۰	۶/۳۲	۶/۴۶	۶/۳۵	۶/۲۴
۱۱	۶/۱۲	۶/۲۶	۶/۱۵	۶/۰۴
۱۲	۵/۹۴	۶/۰۷	۵/۹۶	۵/۸۶
۱۳	۵/۷۶	۵/۸۸	۵/۷۹	۵/۶۹
۱۴	۵/۶۰	۵/۷۰	۵/۶۲	۵/۵۳
۱۵	۹/۱۶	۹/۳۸	۹/۵۴	۹/۱۱
۱۶	۹/۰۸	۹/۳۰	۹/۴۴	۹/۰۲
۱۷	۸/۷۴	۸/۹۴	۹/۰۱	۸/۶۷
۱۸	۸/۳۵	۸/۵۳	۸/۵۴	۸/۲۷
۱۹	۸/۰۰	۸/۱۷	۸/۱۴	۷/۹۲
۲۰	۷/۶۸	۷/۸۴	۷/۷۸	۷/۶۰
۲۱	۷/۳۸	۷/۵۴	۷/۴۷	۷/۳۰
۲۲	۷/۱۱	۷/۲۷	۷/۱۸	۷/۰۳
۲۳	۶/۸۶	۷/۰۱	۶/۹۲	۶/۷۸
۲۴	۶/۶۲	۶/۷۷	۶/۶۸	۶/۵۵
۲۵	۶/۴۰	۶/۵۵	۶/۴۶	۶/۳۴
۲۶	۶/۲۰	۶/۳۴	۶/۲۵	۶/۱۴
۲۷	۶/۰۱	۶/۱۴	۶/۰۶	۵/۹۶
۲۸	۵/۸۳	۵/۹۵	۵/۸۸	۵/۷۸
۲۹	۹/۷۷	۹/۹۸	۱۰/۱۷	۹/۷۳
۳۰	۹/۶۸	۹/۸۹	۱۰/۰۵	۹/۶۳
۳۱	۹/۲۹	۹/۴۸	۹/۵۷	۹/۲۳
۳۲	۸/۸۶	۹/۰۳	۹/۰۶	۸/۷۹
۳۳	۸/۴۶	۸/۶۳	۸/۶۱	۸/۳۹
۳۴	۸/۱۰	۸/۲۷	۸/۲۲	۸/۰۴
۳۵	۷/۷۷	۷/۹۵	۷/۸۸	۷/۷۱

ادامہی جدول (۱-پ)

۱- بوتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۳۶	۷/۴۷	۷/۶۴	۷/۵۷	۷/۴۲
۳۷	۷/۱۹	۷/۳۷	۷/۲۸	۷/۱۴
۳۸	۶/۹۴	۷/۱۱	۷/۰۲	۶/۸۹
۳۹	۶/۷۰	۶/۸۶	۶/۷۸	۶/۶۶
۴۰	۶/۴۸	۶/۶۴	۶/۵۶	۶/۴۴
۴۱	۶/۲۷	۶/۴۲	۶/۳۵	۶/۲۳
۴۲	۶/۰۸	۶/۲۲	۶/۱۶	۶/۰۵
۴۳	۱۰/۴۳	۱۰/۶۶	۱۰/۸۴	۱۰/۴۰
۴۴	۱۰/۳۲	۱۰/۵۴	۱۰/۷۱	۱۰/۲۹
۴۵	۹/۸۹	۱۰/۰۹	۱۰/۱۷	۹/۸۳
۴۶	۹/۴۰	۹/۵۸	۹/۶۰	۹/۳۴
۴۷	۸/۹۶	۹/۱۴	۹/۱۱	۸/۹۰
۴۸	۸/۵۵	۸/۷۴	۸/۶۸	۸/۵۰
۴۹	۸/۱۹	۸/۳۸	۸/۳۱	۸/۱۴
۵۰	۷/۸۶	۸/۰۵	۷/۹۷	۷/۸۲
۵۱	۷/۵۵	۷/۷۵	۷/۶۶	۷/۵۲
۵۲	۷/۲۷	۷/۴۶	۷/۳۸	۷/۲۵
۵۳	۷/۰۱	۷/۲۰	۷/۱۲	۶/۹۹
۵۴	۶/۷۶	۶/۹۶	۶/۸۸	۶/۷۵
۵۵	۶/۵۴	۶/۷۲	۶/۶۶	۶/۵۳
۵۶	۶/۳۳	۶/۵۱	۶/۴۵	۶/۳۳
۵۷	۱۱/۱۴	۱۱/۴۰	۱۱/۵۶	۱۱/۱۱
۵۸	۱۱/۰۲	۱۱/۲۷	۱۱/۴۱	۱۰/۹۹
۵۹	۱۰/۵۳	۱۰/۷۵	۱۰/۸۱	۱۰/۴۷
۶۰	۹/۹۸	۱۰/۱۹	۱۰/۱۷	۹/۹۱
۶۱	۹/۴۸	۹/۶۹	۹/۶۳	۹/۴۲
۶۲	۹/۰۳	۹/۲۵	۹/۱۶	۸/۹۹
۶۳	۸/۶۲	۸/۸۵	۸/۷۵	۸/۵۹
۶۴	۸/۲۶	۸/۴۹	۸/۳۸	۸/۲۴
۶۵	۷/۹۲	۸/۱۶	۸/۰۵	۷/۹۱
۶۶	۷/۶۱	۷/۸۵	۷/۷۵	۷/۶۱
۶۷	۷/۳۲	۷/۵۷	۷/۴۷	۷/۳۳



ادامہ ی جدول (۱-پ)

۱- بوتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۶۸	۷/۰۶	۷/۳۰	۷/۲۱	۷/۰۸
۶۹	۶/۸۲	۷/۰۵	۶/۹۷	۶/۸۴
۷۰	۶/۵۹	۶/۸۱	۶/۷۴	۶/۶۲
۷۱	۱۱/۹۱	۱۲/۲۳	۱۲/۳۳	۱۱/۸۷
۷۲	۱۱/۷۷	۱۲/۰۹	۱۲/۱۶	۱۱/۷۳
۷۳	۱۱/۲۱	۱۱/۴۹	۱۱/۴۹	۱۱/۱۵
۷۴	۱۰/۵۸	۱۰/۸۵	۱۰/۷۸	۱۰/۵۲
۷۵	۱۰/۰۳	۱۰/۳۰	۱۰/۱۸	۹/۹۷
۷۶	۹/۵۳	۹/۸۱	۹/۶۶	۹/۴۹
۷۷	۹/۰۸	۹/۳۷	۹/۲۱	۹/۰۶
۷۸	۸/۶۷	۸/۹۷	۸/۸۱	۸/۶۶
۷۹	۳۰۸	۸/۶۰	۸/۴۵	۸/۳۱
۸۰	۷/۹۶	۸/۲۷	۸/۱۳	۷/۹۸
۸۱	۷/۶۵	۷/۹۶	۷/۸۳	۷/۶۸
۸۲	۷/۳۶	۷/۶۷	۷/۵۵	۷/۴۰
۸۳	۷/۱۰	۷/۴۰	۷/۳۰	۷/۱۵
۸۴	۶/۸۵	۷/۱۵	۷/۰۶	۷/۹۱
۸۵	۱۲/۷۱	۱۳/۱۶	۱۳/۱۷	۱۲/۶۷
۸۶	۱۲/۵۶	۱۲/۹۹	۱۲/۹۸	۱۲/۵۲
۸۷	۱۱/۹۲	۱۲/۳۲	۱۲/۲۲	۱۱/۸۵
۸۸	۱۱/۲۱	۱۱/۵۹	۱۱/۴۱	۱۱/۱۴
۸۹	۱۰/۵۹	۱۰/۹۷	۱۰/۷۵	۱۰/۵۲
۹۰	۱۰/۰۴	۱۰/۴۲	۱۰/۱۹	۹/۹۹
۹۱	۹/۵۴	۹/۹۳	۹/۷۰	۹/۵۲
۹۲	۹/۰۹	۹/۴۹	۹/۲۷	۹/۰۹
۹۳	۸/۶۹	۹/۰۹	۸/۸۸	۸/۷۰
۹۴	۸/۳۲	۸/۷۲	۸/۵۳	۸/۳۵
۹۵	۷/۹۸	۸/۳۹	۸/۲۱	۸/۰۳
۹۶	۷/۶۷	۸/۰۷	۷/۹۲	۷/۷۳
۹۷	۷/۳۸	۷/۷۸	۷/۶۴	۷/۴۵
۹۸	۷/۱۲	۷/۵۱	۷/۳۸	۷/۱۹
۹۹	۱۳/۵۴	۱۴/۲۰	۱۴/۰۹	۱۳/۵۳

ادامہی جدول (۱-پ)

۱- بوتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱۰۰	۱۳/۳۷	۱۴/۰۱	۱۳/۸۷	۱۳/۳۴
۱۰۱	۱۲/۶۵	۱۳/۲۴	۱۳/۰۰	۱۲/۵۸
۱۰۲	۱۱/۸۶	۱۲/۴۱	۱۲/۱۰	۱۱/۷۸
۱۰۳	۱۱/۱۶	۱۱/۷۰	۱۱/۳۶	۱۱/۰۹
۱۰۴	۱۰/۵۵	۱۱/۰۸	۱۰/۷۳	۱۰/۵۰
۱۰۵	۱۰/۰۰	۱۰/۵۴	۱۰/۲۰	۹/۹۷
۱۰۶	۹/۵۱	۱۰/۰۶	۹/۷۳	۹/۵۱
۱۰۷	۹/۰۷	۹/۶۱	۹/۳۱	۹/۰۹
۱۰۸	۸/۶۷	۹/۲۲	۸/۹۴	۸/۷۰
۱۰۹	۸/۳۰	۸/۸۵	۸/۶۰	۸/۳۵
۱۱۰	۷/۹۷	۸/۵۱	۸/۲۹	۸/۰۳
۱۱۱	۷/۶۶	۸/۱۹	۷/۹۹	۷/۷۳
۱۱۲	۷/۳۸	۷/۸۹	۷/۷۲	۷/۴۵
۱۱۳	۱۴/۳۸	۱۵/۳۸	۱۵/۱۰	۱۴/۴۲
۱۱۴	۱۴/۱۸	۱۵/۱۶	۱۴/۸۵	۱۴/۲۰
۱۱۵	۱۳/۳۸	۱۴/۲۷	۱۳/۸۵	۱۳/۳۳
۱۱۶	۱۲/۵۰	۱۳/۳۲	۱۲/۸۳	۱۲/۴۲
۱۱۷	۱۱/۷۳	۱۲/۵۱	۱۲/۰۰	۱۱/۶۵
۱۱۸	۱۱/۰۵	۱۱/۸۲	۱۱/۳۱	۱۰/۹۹
۱۱۹	۱۰/۴۶	۱۱/۲۱	۱۰/۷۴	۱۰/۴۱
۱۲۰	۹/۹۲	۱۰/۶۷	۱۰/۲۱	۹/۹۰
۱۲۱	۹/۴۴	۱۰/۱۹	۹/۷۷	۹/۴۴
۱۲۲	۹/۰۱	۹/۷۵	۹/۳۷	۹/۰۳
۱۲۳	۸/۶۲	۹/۳۵	۹/۰۰	۸/۶۵
۱۲۴	۸/۲۶	۸/۹۸	۸/۶۷	۸/۳۰
۱۲۵	۷/۹۳	۸/۶۳	۸/۳۶	۷/۹۸
۱۲۶	۷/۶۳	۸/۳۱	۸/۰۷	۷/۶۹

۲-ہپتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۹/۷۹	۹/۴۵	۹/۷۶	۹/۲۱
۲	۹/۶۶	۹/۳۴	۹/۶۴	۹/۰۷
۳	۹/۵۳	۹/۲۳	۹/۵۲	۸/۹۳
۴	۹/۴۱	۹/۱۳	۹/۴۱	۸/۸۰

ادامہی جدول (۱-پ)

۲-ہپتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۵	۹/۲۹	۹/۰۴	۹/۳۱	۸/۶۷
۶	۹/۱۷	۸/۹۴	۹/۲۱	۸/۵۴
۷	۹/۰۵	۸/۸۴	۹/۱۱	۸/۴۱
۸	۸/۹۴	۸/۷۵	۹/۰۱	۸/۲۸
۹	۸/۸۲	۸/۶۷	۸/۹۲	۸/۱۵
۱۰	۸/۷۱	۸/۵۸	۸/۸۳	۸/۰۳
۱۱	۸/۶۱	۸/۵۰	۸/۷۵	۷/۹۱
۱۲	۸/۵۰	۸/۴۲	۸/۶۶	۷/۷۹
۱۳	۸/۴۰	۸/۳۴	۸/۵۸	۷/۶۷
۱۴	۸/۲۹	۸/۲۷	۸/۵۰	۷/۵۶
۱۵	۸/۱۹	۸/۱۹	۸/۴۳	۷/۴۴
۱۶	۸/۰۹	۸/۱۲	۸/۳۵	۷/۳۳
۱۷	۸/۰۰	۸/۰۵	۸/۲۸	۷/۲۲
۱۸	۷/۹۰	۷/۹۸	۸/۲۱	۷/۱۱
۱۹	۷/۸۱	۷/۹۱	۸/۱۴	۷/۰۱
۲۰	۷/۷۲	۷/۸۵	۸/۰۸	۶/۹۰
۲۱	۷/۶۳	۷/۷۹	۸/۰۱	۶/۸۰
۲۲	۷/۵۴	۷/۷۲	۷/۹۵	۶/۷۰
۲۳	۱۰/۵۲	۱۰/۰۸	۱۰/۵۰	۹/۹۹
۲۴	۱۰/۳۸	۹/۹۵	۱۰/۳۶	۹/۸۴
۲۵	۱۰/۲۴	۹/۸۳	۱۰/۲۳	۹/۷۰
۲۶	۱۰/۱۱	۹/۷۱	۱۰/۱۰	۹/۵۵
۲۷	۹/۹۸	۹/۶۰	۹/۹۸	۹/۴۱
۲۸	۹/۸۵	۹/۴۹	۹/۸۶	۹/۲۷
۲۹	۹/۷۳	۹/۳۸	۹/۷۴	۹/۱۳
۳۰	۹/۶۰	۹/۲۸	۹/۷۴	۹/۰۰
۳۱	۹/۴۸	۹/۱۸	۹/۵۲	۸/۸۶
۳۲	۹/۳۶	۹/۰۹	۹/۴۲	۸/۷۳
۳۳	۹/۲۵	۸/۹۹	۹/۳۳	۸/۶۰
۳۴	۹/۱۳	۸/۹۰	۹/۲۳	۸/۴۷
۳۵	۹/۰۱	۸/۸۱	۹/۱۴	۸/۳۴
۳۶	۸/۹۱	۸/۷۳	۹/۰۵	۸/۲۲

ادامہی جدول (۱-پ)

۲-ہپتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۳۷	۸/۸۰	۸/۶۴	۸/۹۷	۸/۱۱
۳۸	۸/۶۹	۸/۵۶	۸/۸۸	۷/۹۹
۳۹	۸/۵۹	۸/۴۹	۸/۸۰	۷/۸۷
۴۰	۸/۴۸	۸/۴۱	۸/۷۲	۷/۷۵
۴۱	۸/۳۸	۸/۳۳	۸/۶۴	۷/۶۴
۴۲	۸/۲۸	۸/۲۷	۸/۵۷	۷/۵۳
۴۳	۸/۱۹	۸/۱۹	۸/۵۰	۷/۴۲
۴۴	۸/۰۹	۸/۱۲	۸/۴۲	۷/۳۱
۴۵	۱۱/۲۸	۱۰/۷۷	۱۱/۳۱	۱۰/۸۱
۴۶	۱۱/۱۴	۱۰/۶۳	۱۱/۱۵	۱۰/۶۶
۴۷	۱۱/۹۹	۱۰/۴۹	۱۰/۵۰	۱۰/۵۱
۴۸	۱۰/۸۴	۱۰/۳۶	۱۰/۸۴	۱۰/۳۵
۴۹	۱۰/۷۰	۱۰/۲۳	۱۰/۷۰	۱۰/۲۰
۵۰	۱۰/۵۶	۱۰/۱۰	۱۰/۵۶	۱۰/۰۵
۵۱	۱۰/۴۳	۹/۹۸	۱۰/۴۳	۹/۹۰
۵۲	۱۰/۲۹	۹/۸۷	۱۰/۳۱	۹/۷۶
۵۳	۱۰/۱۶	۹/۷۵	۱۰/۱۹	۹/۶۲
۵۴	۱۰/۰۳	۹/۶۵	۱۰/۰۷	۹/۴۸
۵۵	۹/۹۰	۹/۵۳	۹/۹۵	۹/۳۳
۵۶	۹/۷۸	۹/۴۳	۹/۸۵	۹/۲۰
۵۷	۹/۶۶	۹/۳۴	۹/۷۴	۹/۰۷
۵۸	۹/۵۴	۹/۲۴	۹/۶۴	۸/۹۴
۵۹	۹/۴۳	۹/۱۵	۹/۵۴	۸/۸۱
۶۰	۹/۳۱	۹/۰۶	۹/۴۵	۸/۶۸
۶۱	۹/۲۰	۸/۹۷	۹/۳۶	۸/۵۶
۶۲	۹/۰۹	۸/۸۸	۹/۲۷	۸/۴۳
۶۳	۸/۹۸	۸/۸۰	۹/۱۸	۸/۳۱
۶۴	۸/۸۷	۸/۷۲	۹/۰۸	۸/۱۹
۶۵	۸/۷۷	۸/۶۳	۹/۰۱	۸/۰۷
۶۶	۸/۶۷	۸/۵۶	۸/۹۴	۷/۹۶
۶۷	۱۲/۰۹	۱۱/۵۷	۱۲/۲۱	۱۱/۷۰
۶۸	۱۱/۹۱	۱۱/۳۹	۱۲/۰۱	۱۱/۵۲

ادامہی جدول (۱-پ)

۲-ہپتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۶۹	۱۱/۷۶	۱۱/۲۴	۱۱/۸۳	۱۱/۳۶
۷۰	۱۱/۶۰	۱۱/۰۸	۱۱/۶۵	۱۱/۱۹
۷۱	۱۱/۴۵	۱۰/۹۳	۱۱/۴۹	۱۱/۰۳
۷۲	۱۱/۳۰	۱۰/۷۸	۱۱/۳۳	۱۰/۸۷
۷۳	۱۱/۱۶	۱۰/۶۵	۱۱/۱۸	۱۰/۷۲
۷۴	۱۱/۰۱	۱۰/۵۱	۱۱/۰۴	۱۰/۵۶
۷۵	۱۰/۸۷	۱۰/۳۹	۱۰/۸۹	۱۰/۴۱
۷۶	۱۰/۷۳	۱۰/۲۶	۱۰/۷۶	۱۰/۲۶
۷۷	۱۰/۶۰	۱۰/۱۴	۱۰/۶۳	۱۰/۱۲
۷۸	۱۰/۴۶	۱۰/۰۲	۱۰/۵۱	۹/۹۷
۷۹	۱۰/۳۳	۹/۹۱	۱۰/۳۹	۹/۸۳
۸۰	۱۰/۲۰	۹/۸۰	۱۰/۲۷	۹/۶۹
۸۱	۱۰/۰۸	۹/۷۰	۱۰/۱۶	۹/۵۵
۸۲	۹/۹۶	۹/۵۹	۱۰/۰۵	۹/۴۱
۸۳	۹/۸۴	۹/۴۹	۹/۹۵	۹/۲۸
۸۴	۹/۷۲	۹/۴۰	۹/۸۵	۹/۱۵
۸۵	۹/۶۰	۹/۳۱	۹/۷۵	۹/۰۲
۸۶	۹/۴۸	۹/۲۱	۹/۶۵	۸/۸۹
۸۷	۹/۳۷	۹/۱۳	۹/۵۷	۸/۷۷
۸۸	۹/۲۶	۹/۰۴	۹/۴۷	۸/۶۵
۸۹	۱۲/۹۰	۱۲/۴۵	۱۳/۱۸	۱۲/۶۱
۹۰	۱۲/۷۲	۱۲/۲۵	۱۲/۹۶	۱۲/۴۳
۹۱	۱۲/۵۶	۱۲/۰۷	۱۲/۷۵	۱۲/۲۶
۹۲	۱۲/۳۹	۱۱/۸۸	۱۲/۵۴	۱۲/۰۸
۹۳	۱۲/۲۳	۱۱/۷۱	۱۲/۳۶	۱۱/۹۱
۹۴	۱۲/۰۷	۱۱/۵۵	۱۲/۱۷	۱۱/۷۴
۹۵	۱۱/۹۱	۱۱/۳۹	۱۱/۹۹	۱۱/۵۷
۹۶	۱۱/۷۶	۱۱/۲۴	۱۱/۸۳	۱۱/۴۱
۹۷	۱۱/۶۰	۱۱/۰۸	۱۱/۶۶	۱۱/۲۴
۹۸	۱۱/۴۵	۱۰/۹۴	۱۱/۵۲	۱۱/۰۹
۹۹	۱۱/۳۱	۱۰/۸۱	۱۱/۳۷	۱۰/۹۳
۱۰۰	۱۱/۱۷	۱۰/۶۷	۱۱/۲۳	۱۰/۷۸

ادامہی جدول (۱-پ)

۲-ہپتانول	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱۰۱	۱۱/۰۳	۱۰/۵۴	۱۱/۰۹	۱۰/۶۳
۱۰۲	۱۰/۸۹	۱۰/۴۲	۱۰/۹۵	۱۰/۴۸
۱۰۳	۱۰/۷۵	۱۰/۳۰	۱۰/۸۳	۱۰/۳۳
۱۰۴	۱۰/۶۲	۱۰/۱۸	۱۰/۷۱	۱۰/۱۸
۱۰۵	۱۰/۴۹	۱۰/۰۷	۱۰/۵۸	۱۰/۰۴
۱۰۶	۱۰/۳۶	۹/۹۶	۱۰/۴۸	۹/۹۰
۱۰۷	۱۰/۲۴	۹/۸۶	۱۰/۳۷	۹/۷۷
۱۰۸	۱۰/۱۱	۹/۷۵	۱۰/۲۶	۹/۶۳
۱۰۹	۹/۹۹	۹/۶۵	۱۰/۱۶	۹/۵۰
۱۱۰	۹/۸۸	۹/۵۶	۱۰/۰۶	۹/۳۷
۱۱۱	۱۳/۷۵	۱۳/۴۵	۱۴/۲۷	۱۳/۵۸
۱۱۲	۱۳/۵۷	۱۳/۲۲	۱۴/۰۱	۱۳/۳۹
۱۱۳	۱۳/۳۸	۱۳/۰۱	۱۳/۷۶	۱۳/۲۰
۱۱۴	۱۳/۲۰	۱۲/۸۰	۱۳/۵۳	۱۳/۰۱
۱۱۵	۱۳/۰۳	۱۲/۶۰	۱۳/۳۱	۱۲/۸۳
۱۱۶	۱۲/۸۶	۱۲/۴۱	۱۳/۱۰	۱۲/۶۵
۱۱۷	۱۲/۶۹	۱۲/۲۲	۱۲/۸۹	۱۲/۴۸
۱۱۸	۱۲/۵۲	۱۲/۰۵	۱۲/۷۰	۱۲/۳۰
۱۱۹	۱۲/۳۶	۱۱/۸۷	۱۲/۵۱	۱۲/۱۲
۱۲۰	۱۲/۲۰	۱۱/۷۱	۱۲/۳۴	۱۱/۹۶
۱۲۱	۱۲/۰۵	۱۱/۵۵	۱۲/۱۶	۱۱/۷۹
۱۲۲	۱۱/۹۰	۱۱/۴۰	۱۲/۰۰	۱۱/۶۳
۱۲۳	۱۱/۷۴	۱۱/۲۵	۱۱/۸۴	۱۱/۴۷
۱۲۴	۱۱/۶۰	۱۱/۱۱	۱۱/۶۹	۱۱/۳۱
۱۲۵	۱۱/۴۵	۱۰/۹۷	۱۱/۵۵	۱۱/۱۵
۱۲۶	۱۱/۳۱	۱۰/۸۴	۱۱/۴۱	۱۲/۰۰
۱۲۷	۱۱/۱۷	۱۰/۷۱	۱۱/۲۸	۱۰/۸۴
۱۲۸	۱۱/۰۳	۱۰/۵۹	۱۱/۱۵	۱۰/۶۹
۱۲۹	۱۰/۹۰	۱۰/۴۷	۱۱/۰۲	۱۰/۵۵
۱۳۰	۱۰/۷۷	۱۰/۳۵	۱۰/۹۰	۱۰/۴۰
۱۳۱	۱۰/۶۴	۱۰/۲۴	۱۰/۷۹	۱۰/۲۶
۱۳۲	۱۰/۵۱	۱۰/۱۳	۱۰/۶۷	۱۰/۱۲

ادامہی جدول (۱-پ)

پنتیل آمین	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۹/۸۱۷۶	۹/۶۳۳۸	۱۰/۱۹۶۵	۱۰/۱۹۰۲
۲	۸/۹۲۷۴	۸/۶۵۸۴	۹/۱۰۴۶	۹/۱۲۴۸
۳	۸/۱۸۵۰	۷/۹۰۲۶	۸/۲۸۶۵	۸/۲۵۶۶
۴	۷/۵۶۱۸	۷/۲۹۳۰	۷/۶۳۶۲	۷/۵۵۹۶
۵	۷/۰۳۰۹	۶/۷۸۱۱	۷/۰۹۲۷	۶/۹۹۹۲
۶	۶/۵۷۳۰	۶/۳۳۸۹	۶/۶۲۳۶	۶/۵۴۵۸
۷	۶/۱۷۳۸	۵/۹۴۹۴	۶/۲۱۰۸	۶/۱۷۴۴
۸	۵/۸۲۲۶	۵/۶۰۲۴	۵/۸۴۳۷	۵/۸۶۴۷
۹	۵/۵۱۱۱	۵/۲۹۰۹	۵/۵۱۵۳	۵/۶۰۰۸
۱۰	۵/۲۳۲۸	۵/۰۱۰۳	۵/۲۲۰۶	۵/۳۷۰۷
۱۱	۴/۹۸۲۷	۴/۷۵۷۰	۴/۹۵۵۸	۵/۱۶۴۸
۱۲	۴/۷۵۶۶	۴/۵۲۸۲	۴/۷۱۷۷	۴/۹۷۴۷
۱۳	۴/۵۵۱۱	۴/۳۲۱۴	۴/۵۰۳۶	۴/۷۹۰۴
۱۴	۴/۳۶۳۶	۴/۱۳۴۵	۴/۳۱۱۲	۴/۵۹۵۲
۱۵	۴/۱۹۱۷	۳/۹۶۵۸	۴/۱۳۸۲	۴/۳۵۵۳
۱۶	۱۰/۶۲۶۹	۱۰/۲۹۰۴	۱۰/۹۲۴۰	۱۰/۹۶۵۸
۱۷	۹/۵۹۱۹	۹/۱۸۴۰	۹/۶۸۳۲	۹/۷۸۵۵
۱۸	۸/۷۴۱۱	۸/۳۴۰۰	۸/۷۶۹۶	۸/۸۱۷۸
۱۹	۸/۰۳۵۱	۷/۶۶۸۳	۸/۰۵۳۲	۸/۰۳۶۶
۲۰	۷/۴۳۹۶	۷/۱۰۹۸	۷/۴۵۹۶	۷/۴۰۵۹
۲۱	۶/۹۳۰۱	۶/۶۳۰۵	۶/۹۴۹۸	۶/۸۹۴۴
۲۲	۶/۴۸۹۱	۶/۲۱۰۳	۶/۵۰۲۵	۶/۴۷۵۲
۲۳	۶/۱۰۳۵	۵/۸۳۶۸	۶/۱۰۵۲	۶/۱۲۶۵
۲۴	۵/۷۶۳۳	۵/۵۰۲۱	۵/۷۵۰۰	۵/۸۳۰۷
۲۵	۵/۴۶۰۹	۵/۲۰۰۸	۵/۴۳۱۴	۵/۵۷۴۵
۲۶	۵/۱۹۰۲	۴/۹۲۸۹	۵/۱۴۵۰	۵/۳۴۷۴
۲۷	۴/۹۴۶۳	۴/۶۸۳۳	۴/۸۸۷۶	۵/۱۴۰۸
۲۸	۴/۷۲۵۶	۴/۴۶۱۳	۴/۶۵۶۱	۴/۹۴۵۳
۲۹	۴/۵۲۴۷	۴/۲۶۰۷	۴/۴۴۷۹	۴/۷۴۷۰
۳۰	۴/۳۴۱۰	۴/۰۷۹۵	۴/۲۶۰۸	۴/۵۱۸۸
۳۱	۱۱/۵۳۰۸	۱۱/۰۲۶۰	۱۱/۷۲۲۶	۱۱/۸۰۴۷
۳۲	۱۰/۳۲۲۸	۹/۷۶۵۳	۱۰/۳۰۸۶	۱۰/۵۰۳۱

ادامہ ہی جدول (۱-پ)

پنٹیل آمین	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۳۳	۹/۳۴۵۱	۸/۸۱۹۰	۹/۲۸۷۰	۹/۴۲۹۴
۳۴	۸/۵۴۴۱	۸/۰۷۶۲	۸/۴۹۷۵	۸/۵۵۷۸
۳۵	۷/۸۷۵۴	۷/۴۶۵۳	۷/۸۴۹۷	۷/۸۵۱۴
۳۶	۷/۳۰۸۳	۶/۹۴۵۰	۷/۲۹۶۵	۷/۲۷۶۸
۳۷	۶/۸۲۱۱	۶/۴۹۱۱	۶/۸۱۲۷	۶/۸۰۵۷
۳۸	۶/۳۹۷۸	۶/۰۸۸۹	۶/۳۸۳۶	۶/۴۱۴۴
۳۹	۶/۰۲۶۵	۵/۷۲۹۱	۶/۰۰۰۳	۶/۰۸۳۷
۴۰	۵/۶۹۷۹	۵/۴۰۵۵	۵/۶۵۶۴	۵/۷۹۸۸
۴۱	۵/۴۰۵۱	۵/۱۱۳۷	۵/۳۴۷۳	۵/۵۴۸۴
۴۲	۵/۱۴۲۵	۴/۸۵۰۰	۵/۰۶۹۴	۵/۳۲۳۱
۴۳	۴/۹۰۵۴	۴/۶۱۱۶	۴/۸۱۹۳	۵/۱۱۴۰
۴۴	۴/۶۹۰۴	۴/۳۹۶۳	۴/۵۹۴۵	۴/۹۰۹۱
۴۵	۴/۴۹۴۴	۴/۲۰۱۷	۴/۳۹۲۴	۴/۶۸۶۶
۴۶	۱۲/۵۴۲۸	۱۱/۸۵۳۳	۱۲/۶۰۳۱	۱۲/۷۰۸۸
۴۷	۱۱/۱۲۷۲	۱۰/۴۱۰۴	۱۰/۹۸۶۸	۱۱/۲۷۹۱
۴۸	۱۰/۰۰۰۸	۹/۳۴۴۶	۹/۸۴۱۸	۱۰/۰۹۲۸
۴۹	۹/۰۹۰۷	۸/۵۲۰۳	۸/۹۷۱۸	۹/۱۲۴۸
۵۰	۸/۳۳۹۳	۷/۸۵۰۳	۸/۲۶۴۳	۸/۳۳۶۸
۵۱	۷/۷۰۸۱	۷/۲۸۴۴	۷/۶۶۴۹	۷/۶۹۴۱
۵۲	۷/۱۷۰۰	۶/۷۹۳۵	۷/۱۴۲۵	۷/۱۶۶۶
۵۳	۶/۷۰۵۵	۶/۳۶۰۰	۶/۶۷۹۹	۶/۷۲۸۸
۵۴	۶/۳۰۰۴	۵/۹۷۳۱	۶/۲۶۶۹	۶/۳۶۰۰
۵۵	۵/۹۴۳۸	۵/۶۲۵۶	۵/۸۹۶۴	۶/۰۴۳۸
۵۶	۵/۶۲۷۴	۵/۳۱۲۲	۵/۵۶۳۴	۵/۷۶۷۵
۵۷	۵/۳۴۴۶	۵/۰۲۹۱	۵/۲۶۳۸	۵/۵۲۱۴
۵۸	۵/۰۹۰۴	۴/۷۷۳۳	۴/۹۹۴۲	۵/۲۹۶۴
۵۹	۴/۸۶۰۴	۴/۵۴۲۱	۴/۷۵۱۶	۵/۰۸۱۹
۶۰	۴/۶۵۱۵	۴/۳۳۳۱	۴/۵۳۳۵	۴/۸۶۰۰
۶۱	۱۳/۶۷۸۱	۱۲/۷۸۷۵	۱۳/۵۷۸۵	۱۳/۶۷۹۸
۶۲	۱۲/۰۱۲۵	۱۱/۱۲۸۷	۱۱/۷۲۴۳	۱۲/۱۱۵۵
۶۳	۱۰/۷۱۲۱	۹/۹۲۳۲	۱۰/۴۳۷۵	۱۰/۸۰۹۹
۶۴	۹/۶۷۶۸	۹/۰۰۴۹	۹/۴۷۵۹	۹/۷۳۹۱



ادامہی جدول (۱-پ)

پنتیل آمین	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۶۵	۸/۸۳۲۲	۸/۲۶۷۸	۸/۷۰۵۰	۸/۸۶۳۷
۶۶	۸/۱۲۹۶	۷/۶۵۱۰	۸/۰۵۶۱	۸/۱۴۷۶
۶۷	۷/۵۳۵۵	۷/۱۱۹۳	۷/۴۹۲۷	۷/۵۵۸۹
۶۸	۷/۰۲۶۳	۶/۶۵۱۸	۶/۹۹۴۹	۷/۰۷۰۶
۶۹	۶/۵۸۴۷	۶/۲۳۵۵	۶/۵۵۰۷	۶/۶۶۰۱
۷۰	۶/۱۹۸۱	۵/۸۶۲۰	۶/۱۵۲۲	۶/۳۰۹۶
۷۱	۵/۸۵۶۵	۵/۵۲۵۵	۵/۷۹۴۰	۶/۰۰۵۱
۷۲	۵/۵۵۲۵	۵/۲۲۱۷	۵/۴۷۱۵	۵/۷۳۵۹
۷۳	۵/۲۸۰۱	۴/۹۴۷۰	۵/۱۸۱۲	۵/۴۹۲۷
۷۴	۵/۰۳۴۶	۴/۶۹۸۸	۴/۹۱۹۹	۵/۲۶۵۸
۷۵	۴/۸۱۲۰	۴/۴۷۴۴	۴/۶۸۴۸	۵/۰۴۰۱
۷۶	۱۴/۹۵۴۱	۱۳/۸۴۶۵	۱۴/۶۶۴۶	۱۴/۷۱۹۶
۷۷	۱۲/۹۸۶۶	۱۱/۹۳۱۵	۱۲/۵۲۹۴	۱۳/۰۱۴۱
۷۸	۱۱/۴۸۲۲	۱۰/۵۶۱۷	۱۱/۰۷۸۴	۱۱/۵۸۲۷
۷۹	۱۰/۳۰۳۷	۹/۵۳۴۷	۱۰/۰۱۴۴	۱۰/۴۰۲۶
۸۰	۹/۳۵۴۴	۸/۷۲۱۲	۹/۱۷۳۱	۹/۴۳۳۶
۸۱	۸/۵۷۲۶	۸/۰۴۷۴	۸/۴۷۱۰	۸/۶۳۸۵
۸۲	۷/۹۱۷۳	۷/۴۷۰۷	۷/۸۶۴۳	۷/۹۸۳۸
۸۳	۷/۳۵۹۵	۶/۹۶۵۹	۷/۳۲۹۴	۷/۴۴۰۶
۸۴	۶/۸۷۸۹	۶/۵۱۷۷	۶/۸۵۲۵	۶/۹۸۴۸
۸۵	۶/۴۶۰۱	۶/۱۱۶۲	۶/۴۲۴۷	۶/۵۹۶۹
۸۶	۶/۰۹۲۰	۵/۷۵۴۸	۶/۰۳۹۹	۶/۲۶۱۵
۸۷	۵/۷۶۵۶	۵/۴۲۸۶	۵/۶۹۳۴	۵/۹۶۶۸
۸۸	۵/۴۷۴۲	۵/۱۳۳۸	۵/۳۸۱۲	۵/۷۰۳۱
۸۹	۵/۲۱۲۳	۴/۸۶۷۳	۵/۱۰۰۱	۵/۴۶۱۱
۹۰	۴/۹۷۵۷	۴/۶۲۶۴	۴/۸۴۷۰	۵/۲۲۸۰
۹۱	۱۶/۳۹۰۱	۱۵/۰۵۱۳	۱۵/۸۸۰۲	۱۵/۸۳۰۱
۹۲	۱۰/۰۵۷۵	۱۲/۸۳۱۹	۱۳/۴۱۱۶	۱۳/۹۷۷۰
۹۳	۱۲/۳۱۴۳	۱۱/۲۶۸۸	۱۱/۷۶۹۴	۱۲/۴۱۳۱
۹۴	۱۰/۹۷۲۱	۱۰/۱۱۵۳	۱۰/۵۸۹۱	۱۱/۱۱۷۲
۹۵	۹/۹۰۵۵	۹/۲۱۴۴	۹/۶۷۰۲	۱۰/۰۴۸۵
۹۶	۹/۰۳۶۵	۸/۴۷۶۵	۸/۹۱۰۷	۹/۱۶۸۶

ادامہی جدول (۱-پ)

پنتیل آمین	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۹۷	۸/۳۱۴۴	۷/۸۴۹۸	۸/۲۵۸۲	۸/۴۴۲۷
۹۸	۷/۷۰۴۴	۷/۳۰۴۲	۷/۶۸۴۳	۷/۸۴۰۲
۹۹	۷/۱۸۱۹	۶/۸۲۱۳	۷/۱۷۳۰	۷/۳۳۵۱
۱۰۰	۶/۷۲۹۲	۶/۳۸۹۵	۶/۷۱۴۵	۶/۹۰۶۴
۱۰۱	۶/۳۳۳۰	۶/۰۰۱۳	۶/۳۰۱۸	۶/۵۳۷۲
۱۰۲	۵/۹۸۳۱	۵/۶۵۱۱	۵/۹۳۰۰	۶/۲۱۴۵
۱۰۳	۵/۶۷۱۹	۵/۳۳۴۶	۵/۵۹۴۸	۵/۹۲۸۱
۱۰۴	۵/۳۹۳۱	۵/۰۴۸۵	۵/۲۹۲۷	۵/۶۶۸۶
۱۰۵	۵/۱۴۱۹	۴/۷۸۹۸	۵/۰۲۰۶	۵/۴۲۴۷

نونادکان	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۱	۸/۸۳۸	۸/۸۲۸	۹/۰۴۳	۸/۹۶۵
۲	۸/۰۳۷	۸/۰۱۷	۸/۱۷۳	۸/۲۰۸
۳	۷/۳۹۱	۷/۳۹۲	۷/۵۰۱	۷/۵۵۹
۴	۶/۸۵۲	۶/۸۸۶	۶/۹۵۷	۷/۰۰۷
۵	۹/۹۸۹	۹/۸۹۲	۱۰/۱۶۴	۹/۹۹۴
۶	۸/۹۸۸	۸/۸۵۶	۹/۰۹۸	۹/۰۸۲
۷	۸/۱۷۸	۸/۰۸۹	۸/۲۸۳	۸/۳۰۳
۸	۷/۵۲۸	۷/۴۸۹	۷/۶۳۹	۷/۶۴۲
۹	۶/۹۸۴	۶/۹۹۶	۷/۱۰۵	۷/۰۸۱
۱۰	۶/۵۱۹	۶/۵۶۸	۶/۶۴۳	۶/۶۰۰
۱۱	۶/۱۲۳	۶/۱۹۱	۶/۲۳۸	۶/۱۹۲
۱۲	۵/۷۷۶	۵/۸۵۰	۵/۸۷۶	۵/۸۴۲
۱۳	۵/۴۷۱	۵/۵۴۱	۵/۵۵۳	۵/۵۴۴
۱۴	۵/۱۹۹	۵/۲۵۸	۵/۲۶۳	۵/۲۸۹
۱۵	۴/۹۶۱	۴/۹۹۹	۵/۰۰۱	۵/۰۷۱
۱۶	۴/۷۴۱	۴/۷۶۲	۴/۷۶۵	۴/۸۸۵
۱۷	۴/۵۴۴	۴/۵۴۴	۴/۵۵۱	۴/۷۲۴
۱۸	۴/۳۶۵	۴/۳۴۵	۴/۳۵۹	۴/۵۸۶
۱۹	۱۰/۶۳۴	۱۰/۵۱۸	۱۰/۸۴۲	۱۰/۵۷۴
۲۰	۹/۵۰۲	۹/۳۳۶	۹/۶۰۷	۹/۵۷۴
۲۱	۸/۶۱۱	۸/۴۷۹	۸/۷۰۷	۸/۷۲۳
۲۲	۷/۸۹۰	۷/۸۲۲	۸/۰۰۷	۸/۰۰۰

ادامہی جدول (۱-پ)

نوناڈکان	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۲۳	۷/۲۹۵	۷/۲۸۹	۷/۴۳۲	۷/۳۸۷
۲۴	۶/۷۹۱	۶/۸۳۲	۶/۹۳۶	۶/۸۶۲
۲۵	۶/۳۶۴	۶/۴۳۲	۶/۵۰۳	۶/۴۱۶
۲۶	۵/۹۹۳	۶/۰۷۱	۶/۱۱۷	۶/۰۳۳
۲۷	۵/۶۶۹	۵/۷۴۳	۵/۷۷۲	۵/۷۰۸
۲۸	۵/۳۷۹	۵/۴۴۴	۵/۴۶۱	۵/۴۲۹
۲۹	۵/۱۱۹	۵/۱۷۰	۵/۱۸۱	۵/۱۹۱
۳۰	۴/۸۸۷	۴/۹۱۸	۴/۹۲۹	۴/۹۸۷
۳۱	۴/۶۷۷	۴/۶۸۷	۴/۶۹۹	۴/۸۱۱
۳۲	۴/۴۹۱	۴/۴۷۶	۴/۴۹۳	۴/۶۶۱
۳۳	۴/۳۱۵	۴/۲۸۳	۴/۳۰۸	۴/۵۳۱
۳۴	۴/۱۵۷	۴/۱۰۸	۴/۱۴۰	۴/۴۱۹
۳۵	۱۱/۳۳۴	۱۱/۲۲۱	۱۱/۵۴۸	۱۱/۲۰۲
۳۶	۱۰/۰۵۹	۹/۸۶۵	۱۰/۱۵۱	۱۰/۱۰۷
۳۷	۹/۰۶۳	۸/۹۰۲	۹/۱۵۵	۹/۱۷۶
۳۸	۸/۲۶۳	۸/۱۷۸	۸/۳۹۲	۸/۳۸۶
۳۹	۷/۶۱۹	۷/۶۰۱	۷/۷۷۴	۷/۷۱۸
۴۰	۷/۰۷۳	۷/۱۱۲	۷/۲۴۴	۷/۱۴۵
۴۱	۶/۶۱۰	۶/۶۸۶	۶/۷۸۲	۶/۶۵۸
۴۲	۶/۲۱۵	۶/۳۰۳	۶/۳۷۰	۶/۲۴۰
۴۳	۵/۸۶۳	۵/۹۵۷	۶/۰۰۲	۵/۸۸۵
۴۴	۵/۵۶۱	۵/۶۴۰	۵/۶۷۱	۵/۵۸۲
۴۵	۵/۲۸۴	۵/۳۵۰	۵/۳۷۲	۵/۳۲۲
۴۶	۵/۰۴۰	۵/۰۸۴	۵/۱۰۲	۵/۰۹۹
۴۷	۴/۸۱۹	۴/۸۳۹	۴/۸۵۷	۴/۹۰۷
۴۸	۴/۶۱۸	۴/۶۱۵	۴/۶۳۶	۴/۷۴۳
۴۹	۴/۴۳۶	۴/۴۱۰	۴/۴۳۷	۴/۶۰۱
۵۰	۴/۲۶۸	۴/۲۲۴	۴/۲۵۷	۴/۴۷۹
۵۱	۱۲/۹۱۱	۱۲/۹۲۳	۱۳/۱۷۵	۱۲/۶۰۹
۵۲	۱۱/۲۷۸	۱۱/۱۰۶	۱۱/۳۶۲	۱۱/۲۹۷
۵۳	۱۰/۰۴۹	۹/۸۶۶	۱۰/۱۲۹	۱۰/۱۸۸
۵۴	۹/۰۹۳	۸/۹۷۳	۹/۲۲۲	۹/۲۵۰

ادامہی جدول (۱-پ)

نوناڈکان	مقادیر تجربی	MLR-ANN	CDFS-ANN	GCM-ANN
۵۵	۸/۳۱۴	۸/۲۸۶	۸/۵۰۵	۸/۴۵۶
۵۶	۷/۶۶۸	۷/۷۲۱	۷/۹۰۲	۷/۷۷۶
۵۷	۷/۱۳۴	۷/۲۳۷	۷/۳۸۰	۷/۱۹۸
۵۸	۶/۶۷۳	۶/۸۰۷	۶/۹۱۵	۶/۷۰۲
۵۹	۶/۲۷۸	۶/۴۲۰	۶/۵۰۱	۶/۲۸۱
۶۰	۵/۹۲۹	۶/۰۶۷	۶/۱۲۶	۵/۹۲۰
۶۱	۵/۶۲۰	۵/۷۴۴	۵/۷۸۸	۵/۶۱۲
۶۲	۵/۳۴۹	۵/۴۴۶	۵/۴۸۱	۵/۳۴۷
۶۳	۵/۱۰۲	۵/۱۷۱	۵/۲۰۲	۵/۱۱۹
۶۴	۴/۸۷۹	۴/۹۲۰	۴/۹۴۹	۴/۹۲۴
۶۵	۴/۶۷۶	۴/۶۸۹	۴/۷۲۲	۴/۷۵۶
۶۶	۴/۴۹۶	۴/۴۷۸	۴/۵۱۵	۴/۶۱۱
۶۷	۱۳/۷۷۶	۱۳/۹۶۰	۱۴/۱۲۳	۱۳/۳۹۲
۶۸	۱۱/۹۵۷	۱۱/۸۴۱	۱۲/۰۴۳	۱۱/۹۵۷
۶۹	۱۰/۵۸۹	۱۰/۴۲۰	۱۰/۶۶۱	۱۰/۷۴۸
۷۰	۹/۵۳۷	۹/۴۱۸	۹/۶۶۷	۹/۷۲۷
۷۱	۸/۶۷۹	۸/۶۶۴	۸/۸۹۶	۸/۸۶۳
۷۲	۷/۹۹۲	۸/۰۵۲۸	۸/۲۵۲۰	۸/۱۲۴
۷۳	۷/۴۱۴	۷/۵۳۶۴	۷/۶۹۸۸	۷/۴۹۵
۷۴	۶/۹۲۳	۷/۰۸۰۰	۷/۲۰۷۰	۶/۹۵۶
۷۵	۶/۴۹۴	۶/۶۷۱۴	۶/۷۶۸۲	۶/۴۹۸
۷۶	۶/۱۲۴	۶/۲۹۹۱	۶/۳۷۱۸	۶/۱۰۶
۷۷	۵/۸۰۱	۵/۹۵۷۴	۶/۰۱۲۷	۵/۷۷۰
۷۸	۵/۵۱۱	۵/۶۴۳۱	۵/۶۸۶۷	۵/۴۸۳
۷۹	۵/۲۴۸	۵/۳۵۲۳	۵/۳۸۹۴	۵/۲۳۵
۸۰	۵/۰۱۵	۵/۰۸۶۳	۵/۱۲۰۹	۵/۰۲۳
۸۱	۴/۸۰۶	۴/۸۴۱۹	۴/۸۷۷۵	۴/۸۴۰
۸۲	۴/۶۱۰	۴/۶۱۷۹	۴/۶۵۶۹	۴/۶۸۲

## فهرست منابع

- [1]. J. L. Daridon, A. Lagrabette, B. Lagourette, "Speed of sound, density, and Compressibilities of heavy synthetic cuts from ultrasonic measurements under pressure", **J. Chem. Thermodyn**, 30 (1998), 607-623
- [2]. S. Matsuo, T. Makita, "Volumetric Properties of 1-Alkanols at Temperatures in the Rang 298-348 K and Pressures up to 40 MPa", **Int. J. Thermophys**, 10 (1989), 885-897
- [3]. Y. Miyake, A. Baylaucq, F. Plantier, "High-pressure (up to 140 MPa) density and derivative properties of some (pentyl-, hexyl-, and heptyl-) amines between (293.15 and 353.15) K", **J. Chem. Thermodyn**, 40 (2008), 836-845
- [4]. G. A. Parsafar, Z. Kalantar, "Extension of linear isotherm regularity to long chain primary, secondary and tertiary alcohols, ketones and 1-carboxylic acids by group contribution method", **Fluid Phase Equilibr**, 234 (2005) 11-21
- [5]. E. K. Goharshadi, F. Moosavi, "Prediction of thermodynamic properties of some hydrofluoroether refrigerants using a new equation of state", **Fluid. Phase. Equilibr**, 238 (2005) 112-119
- [6]. R. Gharbi, "Estimating the Isothermal Compressibility Coefficient of Undersaturated Middle East Crudes Using Neural Networks", **Energy. Fuel**, 11 (1997), 372-378
- [7]. کلانتر ز، (۱۳۸۵)، پایان نامه دکتری: "پیش بینی خواص ترمودینامیکی و انتقالی سیالات آلی چگال با استفاده از روش سهم گروه‌ها و شبکه‌ی عصبی موجک"، دانشکده‌ی شیمی، دانشگاه صنعتی اصفهان
- [8]. D. L. Massart, B. G. Vanderginste, L. M. C. Buydens, "Handbook of Chemometrics and Qualimetrics Part A", **Elsevier**, Amesterdom, 1997
- [9]. B. Hemmateenejad, K. Javadnia, M. Elyasi, "Quantitative structure-retention relationship for the Kovats retention indices of a large set of terpenes: A combined data splitting-feature selection strategy", **Anal. Chim. Acta**, 592 (2007), 72-81

- [10]. B. Hemmateenejad, K. Javadnia, M. Elyasi, R. Miri, "Linear and nonlinear quantitative structure–property relationship models for solubility of some anthraquinone, anthrone and xanthone derivatives in supercritical carbon dioxide", **Anal. Chim. Acta**, 610 (2008), 25-34
- [۱۱]. حسن زاده سورشجانی، ح، پیش‌بینی ویسکوزیته مخلوط چندتایی سیالات با استفاده از شبکه عصبی موجک، پایان‌نامه کارشناسی ارشد، دانشکده شیمی، دانشگاه صنعتی اصفهان، ۱۳۸۸
- [12]. Y. Pan, J. Jiang, Z. Wang, "Quantitative structure–property relationship studies for predicting flash points of alkanes using group bond contribution method with back-propagation neural network", **J. Hazard. Mater**, 147 (2007) 424–430
- [13]. J. O. Valderrama, A. Reategui, R. E. Rojas, "Density of Ionic Liquids Using Group Contribution and Artificial Neural Networks", **Ind. Eng. Chem. Res.**, 2009
- [14]. F. Gharagheizi, A. Eslamimanesh, A. H. Mohammadi, "Determination of Critical Properties and Acentric Factors of Pure Compounds Using the Artificial Neural Network Group Contribution Algorithm", **J. Chem. Eng. Data**, 56, (2011), 2460–2476
- [15]. H. Matsuda, H. Yamamoto, K. Kurihara, K. Tochigi, "Computer-aided reverse design for ionic liquids by QSPR using descriptors of group contribution type for ionic conductivities and viscosities", **Fluid. Phase. Equilib**, 261 (2007) 434–443
- [16]. W. T. Vong, F. N. Tsai, "Densities, Molar Volumes, Thermal Expansion Coefficients, and Isothermal Compressibilities of Organic Acids from 293.15 to 323.15 K and Pressures up to 25 MPa", **J. Chem. Eng. Data**, 42 (1997), 1116-1120
- [17]. R. Malhotra, L. A. Woolf, "PVT Property Measurements for the Liquids Propyl Acetate, Butyl Acetate, and 1-Methylethyl Acetate from (278 to 338) K and (0.1 to 380) MPa", **J. Chem. Eng. Data**, 41 (1996), 1366-1370
- [18]. R. Malhotra, L. A. Woolf, "PVT Property Measurement for the Liquids Ethyl Propionate and Ethyl Butyrate from (278 to 338) K and (0.1 to 380) MPa", **J. Chem. Eng. Data**, 41 (1996), 1371-1374

- [19]. R. Malhotra, L. A. Woolf, “Thermodynamic Properties of Butan-2-one at Temperatures from 278 to 338 K and pressures from 0.1 to 280 MPa; Predictions for Higher Ketones”, **J. Chem. Thermodyn**, 24 (1992), 1207-1217
- [20]. R. Malhotra, L. A. Woolf, “Volumetric measurements of liquid Pentan-2-one, hexan-2-one, 4-methylpentan-2-one at Temperatures from 278.15 K to 338.15 K and pressures in the range from 0.1 MPa to 386 MPa”, **J. Chem. Thermodyn**, 28 (1996), 1411-1421
- [21]. R. Malhotra, L. A. Woolf, W. E. Price, “Thermodynamic properties of pentan-3-one at temperatures from 278 K to 338 K and pressures from 0.1 MPa to 280 MPa”, **J. Chem. Thermodyn**, 25 (1993), 361-366
- [22]. M. Yoshimura, A. Baylaucq, J. P. Bazile, “Volumetric Properties of 2-Alkylamines (2-Aminobutane and 2-Aminooctane) at Pressures up to 140 MPa and Temperatures between (293.15 and 403.15) K”, **J. Chem. Eng. Data**, 54 (2009), 1702-1709
- [23]. M. J. Dávila, R. Alcalde, M. Atilhan, S. Aparicio, “*PVT* measurements and derived properties of liquid 1-Alkanols”, **J. Chem. Thermodyn**, 47 (2012) 241–259
- [24]. C. Boned, A. Baylaucq, J. P. Bazile, “Liquid density of 1-pentanol at pressures up to 140 MPa and from 293.15 to 403.15 K”, **Fluid. Phase. Equilib**, 270 (2008) 69–74
- [25]. M. Dzida, “Study of the Effects of Temperature and Pressure on the Thermodynamic and Acoustic Properties of 2-Methyl-1-butanol at Temperatures from 293K to 318K and Pressures up to 100MPa” **Int. J. Thermophys**, 31 (2010), 55-69
- [26]. D. Gonzalez-Salgado, J. Troncoso, F. Plantier, “Study of the volumetric properties of weakly associated alcohols by means of high-pressure speed of sound measurements” **J. Chem. Thermodyn**, 38 (2006), 893-899
- [27]. A. Zuniga-Moreno, L. A. Galicia-Luna, “Densities, Isothermal Compressibilities, and Isobaric Thermal Expansivities of Hexan-2-ol, Octan-1-ol, and Decan-1-ol from

- (313 to 363) K and Pressures up to 22 MPa” **J. Chem. Eng. Data**, 52 (2007), 1773-1783
- [28]. A. Zuniga-Moreno, L. A. Galicia-Luna, F. F. Betancourt-cardenas, “Densities and derived thermodynamic properties of 1-heptanol and 2-heptanol at temperatures from 313 K to 363 K and pressures up to 22 MPa” **J. Chem. Thermodyn**, 40 (2008), 96-106
- [29]. S. Matsuo, T. Makita, “Volumetric Properties of 1-Alkanols at Temperatures in the Rang 298-348 K and Pressures up to 40 MPa” **Int. J. Thermophys**, 10 (1989), 885-897
- [30]. E. Zorebski, M. Dzida, M. Cempa, “Study of the Effects of Temperature and Pressure on the Acoustic and Thermodynamic Properties of 2-Methyl-2,4-pentanediol” **J. Chem. Eng. Data**, 53 (2008), 1950-1955
- [31]. R. Span, W. Wagner, “Equations of State for Technical Applicationa. I. Simultaneously Optimized Functional Fprms for Nonpolar and Polar Fluids”, **Int. J. Thermophys**, 24 (2003), 1-109
- [32]. D. Pecar, V. Dolecek, “Isothermal compressibilities and isobaric expansibilities of pentane, hexane, heptane and their binary and ternary mixtures from density measurements”, **Fluid. phase. Equilibr**, 211 (2003), 109-127
- [33]. J. L. Daridon, B. Lagourette, J. P. E. Grolier, “Experimental Measurements of the Speed of Sound in *n*-Hexane from 293 to 373 K and up to 150 MPa”, **Int. J. Thermophys**, 19 (1998), 145-159
- [34]. T. S. Khasanshin, A. P. Shchamialiou, O. G. Poddubskij, “Thermodynamic Properties of Heavy *n*-Alkanes in the Liquid State: *n*-Dodecane”, **Int. J. Thermophys**, 24 (2003), 1277-1289
- [35]. D. Bessières, H. Saint-Guirons, J. L. Daridon, “Thermophysical Properties of *n*-Tridecane from 313.15 to 373.15 K and up to 100 MPa from Heat Capacity and Density data”, **J. Therm. Anal. Cal**, 62 (2000), 621-632



- [36]. J. L. Daridon, H. Carrier, B. Lagourette, "Pressure Dependence of the Thermophysical Properties of *n*-Pentadecane and *n*-Heptadecane", **Int. J. Thermophys**, 23 (2002), 697-708
- [37]. S. Dutour, J. L. Daridon, B. Lagourette, "Pressure and Temperature Dependence of the Speed of Sound and Related Properties in Normal Octadecane and Nonadecane", **Int. J. Thermophys**, 21 (2000), 173- 184
- [38]. S. Dutour, B. Lagourette, J. L. Daridon, "High-Pressure speed of sound and compressibility in heavy normal hydrocarbons:  $n\text{-C}_{23}\text{H}_{48}$  and  $n\text{-C}_{24}\text{H}_{50}$ ", **J. Chem. Thermodyn**, 33 (2001), 765-774
- [39]. S. Dutour, B. Lagourette, J. L. Daridon, "High-Pressure speed of sound, density and compressibility of heavy normal paraffins:  $\text{C}_{28}\text{H}_{58}$  and  $\text{C}_{36}\text{H}_{74}$ ", **J. Chem. Thermodyn**, 34 (2002), 475-484
- [40]. <http://www.Chemometrics.blogfa.com>
- [41]. [http://www.strandls.com/sarchitect/documents/manual\\_html/desctheory.html](http://www.strandls.com/sarchitect/documents/manual_html/desctheory.html)
- [42]. R. Todeschini, V. Consonni, "Handbook of Molecular Descriptors", John Wiley-VCH, New York, pp 84
- [43]. R. Todeschini, V. Consonni, "Handbook of Molecular Descriptors", John Wiley-VCH, New York, pp 513-514

## Abstract

The main aim of this work is development of a quantitative structure–property relation (QSPR) method using an artificial neural network (ANN) for predicting the isothermal compressibility ( $\kappa_T$ ) of different organic compounds. The data set consisted of 56 molecules in various temperatures and pressures, which form 4297 data point. The best set of calculated descriptors was selected by two procedures: stepwise regression method and combined data splitting feature selection (CDFS) strategy. Modeling between selected descriptors and  $\kappa_T$  was performed using ANN and after optimization and training of the network, it was used for the prediction of isothermal compressibility of seven external set compounds, which did not have contribute in model development steps. A comparison was made between two models. It was found that the CDFS model have lower number of descriptors and higher prediction ability. The mean square error of the test set obtained by MLR-ANN and CDFS-ANN methods were 0.4456 and 0.3012, respectively.

In the next step, artificial neural network and concept of group contribution method was simultaneously used to correlate and predict the  $\kappa_T$  of organic compounds. To do so, a correlation of functional group as well as temperature and pressure was used as input variables. After optimization the network, the model was used for prediction of  $\kappa_T$  of the same external set. The mean square error of the test set obtained by GCM-ANN method was 0.2303. Comparison of GCM-ANN mean square error of prediction values with those obtained using MLR-ANN and CDFS-ANN show the superiority of GCM-ANN over that MLR and CDFS models.

**Keywords:** quantitative structure–property relation (QSPR), artificial neural network (ANN), isothermal compressibility ( $\kappa_T$ ), MLR, combined data splitting feature selection (CDFS), group contribution method (GCM)

Prediction of thermodynamic properties of some hydrofluoroether refrigerants using a new equation of state



Shahrood University of Technology  
Faculty of Chemistry  
M.Sc. Thesis in Physical Chemistry

**Prediction of isothermal compressibility of organic compound  
from their molecular structure descriptors  
over a wide range of temperature and pressure**

**Zahra Kazem Nadi**

Supervisors:

**Dr. Z. Kalantar**

Advisor:

**Dr. Z. Moosavi**

February 2013