





دانشکده: شیمی

مطالعه ارتباط کمی ساختار - خاصیت خطی و غیر خطی ویسکوزیته آلکانها در محدوده وسیعی از دما و فشار

دانشجو: نیکو فخاری

استاد راهنما:

دکتر زهرا کلانتر

استاد مشاور:

دکتر ناصر گودرزی

پایان نامه کارشناسی ارشد جهت اخذ درجه کارشناسی ارشد

بهمن ماه ۹۱

تقدیم به:

دو ستاره همیشه پر فروغ آسمان زندگی ام

پدر و مادر

دریای محبت و ایثار که عطر دعای سجاده اشان همیشه بدرقه راه پر فراز و نشیب زندگی ام بوده و من تمام مجال زندگی و تحصیل را مدیون جان فشانی ها و دل سوزی های مدبرانه آنان می باشم.

سپاس خدای مهربان را که اندیشه‌ام داد.

حمد و ستایش بی‌قیاس خدایی را سزااست که از الطاف خود در انسان دمید و او را اشرف مخلوقات خود قرار داد. حال که به لطف او توفیق تحصیل علم و کسب دانش را پیدا نمودم، از خداوند متعال می‌خواهم که قدم‌هایم را در راه خدمت به جامعه استوار گرداند تا بتوانم از آنچه در این سال‌ها آموخته‌ام در مسیر پیشرفت و آبادانی کشور عزیزم استفاده نمایم.

با تقدیر و تشکر از استاد ارجمند سرکار خانم دکتر زهرا کلانتر و جناب آقای دکتر ناصر گودرزی که با سعه صدر و وسعت نظر مرا در انجام این پروژه تحقیقاتی صمیمانه یاری نمودند.

همچنین از **پدر و مادر عزیزم** که همواره باعث ایجاد انگیزه در من برای ادامه تحصیل و تلاش در این راه بوده‌اند و همه اساتید گرانقدر تشکر و قدردانی نموده توفیق و عزت همه را از خداوند متعال مسئلت می‌نمایم.

از زحمات بی‌دریغ جناب آقای رئوف غلامی نیز در طول انجام پایان نامه صمیمانه تشکر می‌کنم.

چکیده

در این تحقیق از شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای مدلسازی و پیش‌بینی ویسکوزیته‌ی آلکانها در محدوده‌ی وسیعی از دما و فشار استفاده گردید. تعداد زیادی توصیف‌کننده با استفاده از نرم‌افزار دراگون محاسبه شدند و توصیف‌کننده‌های انتخاب شده زیر مجموعه‌ای از ۱۸ دسته توصیف‌کننده‌های دراگون توسط روش رگرسیون خطی چندگانه می‌باشد. ۴ توصیف‌کننده تئوری و ۲ توصیف‌کننده تجربی دما و فشار به عنوان ورودی شبکه مورد استفاده قرار گرفتند. توصیف‌کننده‌های انتخاب شده برای ایجاد یک شبکه ۱-۵-۵ و ماشین بردار پشتیبان استفاده شد. داده‌ها در شبکه عصبی به سه سری: آموزش، ارزیابی و تست تقسیم شدند و داده‌ها در ماشین بردار پشتیبان به دو سری: آموزش و تست تقسیم شدند. پس از آموزش شبکه و بهینه کردن پارامترهای شبکه عصبی و ماشین بردار پشتیبان، کارایی مدل توسط سری تست مورد بررسی قرار گرفت و نتایج حاصل از شبکه عصبی و ماشین بردار پشتیبان برای این سری‌ها با نتایج تجربی مقایسه گردید.

واژه‌های کلیدی: QSPR، ویسکوزیته، شبکه عصبی مصنوعی، ماشین بردار پشتیبان

نتایج حاصل از این پایان نامه در سه پوستر تحت عناوین:

" پیش بینی ویسکوزیته گازهای طبیعی با استفاده از توصیف کننده های تئوری در محدوده وسیعی از دما و فشار بوسیله مدل رگرسیون خطی "

" مدل سازی شبکه عصبی مصنوعی برای پیش بینی ویسکوزیته آلکانها با استفاده از توصیف کننده های تئوری در محدوده وسیعی از دما و فشار "

در هفتمین سمینار شیمی فیزیک ایران در دانشگاه پیام نور به شهر پذیرفته شد.

" GA-ANN modeling for viscosity prediction of *n*-alkanes over a wide range of temperatures and pressures using theoretically derived descriptors "

در پنجمین سمینار بین المللی ریاضی شیمی ایران در دانشگاه پیام نور یزد پذیرفته شد.

فصل اول

۲	۱-۱- مقدمه
۲	۲-۱- ویسکوزیته
۴	۱-۲-۱- ارتباط شار اندازه حرکت و ویسکوزیته
۵	۳-۱- مروری بر روش‌های پیش‌بینی ویسکوزیته
۶	۱-۳-۱- معادله‌ی ماکسول بولتزمن
۷	۲-۳-۱- روش چاپمن - انسکوگ
۷	۳-۳-۱- نظریه‌ی انسکوگ توسعه یافته (MET)
۹	۴-۳-۱- نظریه‌ی رین واتر - فرند
۹	۵-۳-۱- توابع همبستگی
۱۰	۴-۱- روش‌های مختلف پیش‌بینی ویسکوزیته ترکیبات آلی
۱۱	۱-۴-۱- توابع همبستگی ارائه شده برای پیش‌بینی ویسکوزیته ترکیبات
۱۱	۱-۴-۱- الف- روش تاداس
۱۲	۱-۴-۱- ب- روش چانگ
۱۶	۱-۴-۱- ج- روش هانلی و الی
۱۷	۲-۴-۱- پیش‌بینی ویسکوزیته ترکیبات آلی براساس روش سهم گروه‌ها
۱۸	۳-۴-۱- پیش‌بینی ویسکوزیته ترکیبات آلی با شبکه عصبی

فصل دوم

۲۱	۱-۲- مقدمه
۲۱	۲-۲- تاریخچه‌ی کمومتریکس
۲۲	۳-۲- روش‌های پارامتری
۲۲	۱-۳-۲- رگرسیون خطی چندگانه
۲۳	۱-۱-۳-۲- انتخاب سری داده‌ها
۲۳	۲-۱-۳-۲- محاسبه‌ی توصیف‌کننده‌ها
۲۵	۳-۱-۳-۲- تجزیه، تحلیل و ارزیابی توصیف‌کننده‌ها
۲۵	۴-۱-۳-۲- انتخاب بهترین توصیف‌کننده
۲۶	۵-۱-۳-۲- ارزیابی مدل به دست آمده
۲۶	۶-۱-۳-۲- مدل سازی با استفاده از روش MLR
۲۷	۲-۳-۲- شبکه‌های عصبی مصنوعی
۲۷	۱-۲-۳-۲- مقدمه‌ای بر شبکه‌های عصبی
۲۷	۲-۲-۳-۲- تاریخچه
۲۸	۳-۲-۳-۲- نرون بیولوژیکی
۲۹	۴-۲-۳-۲- نرون محاسباتی
۳۱	۵-۲-۳-۲- ساختار لایه‌ای شبکه عصبی مصنوعی
۳۲	۶-۲-۳-۲- انواع شبکه‌ها از نظر ارتباطات بین نرونی
۳۳	۷-۲-۳-۲- آموزش شبکه عصبی

۳۴	۸-۲-۳-۲- معرفی روش‌های مختلف بهینه سازی عملکرد شبکه عصبی
۳۵	۳-۳-۲- ماشین برداری پشتیبان
۳۵	۱-۳-۳-۲- مقدمه
۳۶	۲-۳-۳-۲- طبقه‌بندی خطی با ماشین‌های برداری پشتیبان
۳۶	۲-۳-۳-۲- الف- آشنایی با مفاهیم اولیه دسته‌بندی کننده خطی SVM
۳۹	۲-۳-۳-۲- ب- ریاضیات مسئله دسته‌بندی کننده خطی SVM
۳۹	۲-۳-۳-۲- ج- حاشیه‌ی نرم
۴۶	۳-۳-۳-۲- طبقه‌بندی غیرخطی با ماشین‌های برداری پشتیبان
۴۶	۳-۳-۳-۲- الف- فضای ویژگی
۴۷	۳-۳-۳-۲- ب- ماشین برداری غیر خطی
۵۰	۴-۳-۳-۲- آنالیز رگرسیون
۵۰	۴-۳-۳-۲- الف- آنالیز رگرسیون خطی
۵۰	۴-۳-۳-۲- الف-۱- ماشین برداری رگرسیونگر برای سیستم‌های خطی
۵۲	۴-۳-۳-۲- الف-۲- فرموله کردن SVR
۵۳	۴-۳-۳-۲- الف-۳- بهینه سازی کاربرد ماشین بردار رگرسیونگر SVR
۵۴	۴-۳-۳-۲- ب- آنالیز رگرسیون غیرخطی
۵۴	۴-۳-۳-۲- ب-۱- ماشین برداری رگرسیونگر برای سیستم‌های غیر خطی
۵۶	۴-۲- ارزیابی مدل

فصل سوم

۶۱	۱-۳-۱- مدل سازی QSPR ویسکوزیته
۶۱	۳-۲- مراحل مدل سازی
۶۱	۳-۲-۱- انتخاب سری داده‌ها
۶۲	۳-۲-۲- رسم و بهینه سازی مولکول‌ها
۶۳	۳-۲-۳- محاسبه‌ی توصیف‌کننده‌های مولکولی
۶۳	۳-۲-۴- انتخاب توصیف‌کننده‌های مناسب
۶۴	۳-۲-۵- مدل سازی توسط روش رگرسیون خطی چندگانه
۶۶	۳-۲-۶- مدل سازی با روش‌های غیرخطی
۶۶	۳-۲-۶-۱- مدل سازی توسط شبکه عصبی
۶۷	۳-۲-۶-۱- الف- انتخاب تعداد گره‌های لایه پنهان، تعداد متغیر ورودی، نوع تابع آموزش و نوع تابع انتقال
۷۳	۳-۲-۶-۱- ب- انتخاب مقدار بهینه μ
۷۴	۳-۲-۶-۱- ج- ساختار شبکه عصبی مصنوعی بهینه شده
۷۵	۳-۲-۶-۲- مدل سازی به روش SVM
۷۷	۳-۲-۷- بررسی نتایج
۷۷	۳-۲-۷-۱- ارزیابی مدل‌های خطی و غیرخطی با استفاده از سری تست
۸۵	۳-۲-۷-۲- ارزیابی مدل‌های برتر خطی و غیرخطی با استفاده از پارامترهای آماری
۸۵	۳-۲-۷-۳- ارزیابی مدل ارائه شده توسط شبکه عصبی با استفاده از آزمون χ^2 - تصادفی
۸۶	۳-۲-۸- بررسی ارتباط توصیفگرهای منتخب با ویسکوزیته

۸۶.....	۳-۲-۱-۳D- MORSE توصیف‌کننده‌های
۸۷.....	۳-۲-۲-۳ MLOGP توصیف‌کننده‌های
۸۸.....	۳-۳-۱-۳ IC توصیف‌کننده‌های
۸۹.....	۳-۳- نتیجه‌گیری
۹۰.....	آینده‌نگری
۹۱.....	منابع

فهرست اشکال

- شکل (۱-۱): جریان سیال بین دو صفحه‌ی موازی، وقتی که صفحه‌ی بالایی ثابت است..... ۳
- شکل (۱-۲) - ساختار نرون طبیعی..... ۲۹
- شکل (۲-۲) - ساختمان یک نرون محاسباتی..... ۳۰
- شکل (۳-۲) - ساختار یک شبکه‌ی سه لایه..... ۳۲
- شکل (۴-۲) - نمایش دو دسته داده‌ی آموزشی..... ۳۶
- شکل (۵-۲) - خط‌های جداساز که داده‌ها را جداسازی می‌کند..... ۳۷
- شکل (۶-۲) - خط جداساز بهینه با حداکثر مقدار حاشیه..... ۳۷
- شکل (۷-۲) - نمایش بردارهای پشتیبان روی ابرصفحه‌های موازی مرزی..... ۳۹
- شکل (۸-۲) - صفحه‌ی جداساز و حاشیه‌ها..... ۴۰
- شکل (۹-۲) - سیستم‌های خطی جدا ناپذیر با میزان خطای ξ ۴۴
- شکل (۱۰-۲) - داده‌های ورودی ارجاع داده شده به فضای بالاتر..... ۴۶
- شکل (۱۱-۲) - مدل ریاضی برازش یافته در معادله رگرسیون خطی با روش کمترین مربعات..... ۵۱
- شکل (۱۲-۲) - نظریه‌ی اصلی حساسیت به مقدار ϵ نمونه‌های خارج از حاشیه $\pm \epsilon$ متغیرهای خطای غیر صفر (ξ) هستند که قسمتی از راه‌حل ارائه شده خواهند بود..... ۵۲
- شکل (۱۳-۲) - مدل ارائه شده توسط ماشین برداری برای سیستم غیرخطی..... ۵۵
- شکل (۱-۳) - نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) MSE سری ارزیابی (ط) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است، برای شبکه‌هایی با الگوریتم آموزشی بایزین و تابع انتقال لگاریتم سیگموئیدی..... ۷۰
- شکل (۲-۳) - نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) MSE سری ارزیابی (ط) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است، برای شبکه‌هایی با الگوریتم آموزشی لونیبرگ-مارکوارت و تابع انتقال لگاریتم سیگموئیدی..... ۷۲
- شکل (۳-۳) - نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) MSE سری ارزیابی (ط) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است، برای شبکه‌هایی با الگوریتم آموزشی بایزین و تابع انتقال تانژانت سیگموئیدی..... ۷۴
- شکل (۴-۳) - نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) MSE سری ارزیابی (ط) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است، برای شبکه‌هایی با الگوریتم آموزشی لونیبرگ-مارکوارت و تابع انتقال تانژانت سیگموئیدی..... ۷۶
- شکل (۵-۳) - میانگین مربع خطای سری ارزیابی بر حسب پارامتر μ ۷۷
- شکل (۶-۳) - تصویر شماتیک ساختار شبکه‌ی عصبی به دست آمده پس از بهینه سازی برای پیش‌بینی ویسکوزیته آلکان‌های خطی..... ۷۸
- شکل (۷-۳) - نمودار مقادیر پیش‌بینی شده‌ی ویسکوزیته بر حسب مقادیر تجربی برای سری تست الف: مدل MLR ب: مدل ANN ج: مدل SVM..... ۸۶

شکل (۸-۳) - نمودار باقیمانده‌ها بر حسب مقدار تجربی ویسکوزیته برای سری تست الف: مدل MLR ب: مدل ANN ج: مدل SVM ۸۷

فهرست جداول

- جدول (۱-۱) - پارامترهای پیوستگی برای چند ترکیب آلی..... ۱۴
- جدول (۲-۱) - ضرایب معادله ی (۲۲-۱)..... ۱۴
- جدول (۳-۱) - ضرایب معادله ی (۲۹-۱)..... ۱۶
- جدول (۱-۲) - انواع توصیف کننده های محاسبه شده توسط نرم افزار Dragon..... ۲۴
- جدول (۲-۲) - توابع کرنل در فضای ویژگی..... ۴۸
- جدول (۱-۳) - داده های ویسکوزیته در محدوده دما و فشار..... ۶۲
- جدول (۲-۳) - پارامترهای آماری برای مدل های به دست آمده از روش MLR..... ۶۵
- جدول (۳-۳) - کل توصیف کننده های مولکولی انتخاب شده توسط روش رگرسیون خطی چندگانه مرحله ای..... ۶۵
- جدول (۴-۳) - ماتریس همبستگی کل توصیف کننده های انتخاب شده توسط رگرسیون خطی چندگانه مرحله ای..... ۶۶
- جدول (۵-۳) : پارامترهای بهینه برای چهار شبکه عصبی مختلف..... ۶۸
- جدول (۶-۳) - میانگین مربع خطا برای سری پیش بینی در مقادیر متفاوت μ ۷۷
- جدول (۷-۳) - توابع و مقادیر پارامترهای بهینه شده شبکه عصبی..... ۷۸
- جدول (۸-۳) - جدول بهینه سازی پارامترهای ماشین بردار پشتیبان..... ۸۰
- جدول (۹-۳) - نتایج حاصل از ارزیابی مدل های MLR و ANN و SVM با استفاده از سری تست..... ۸۱
- جدول (۱۰-۳) - پارامترهای آماری برای سری تست مدل های بهینه شده توسط سه روش MLR، ANN و SVM..... ۸۸
- جدول (۱۱-۳) - مقادیر R^2 برای سری تست با استفاده از آزمون χ^2 -تصادفی..... ۸۹

فصل اول

مروری بر ویسکوزیته و روش‌های پیش‌بینی آن

۱-۱- مقدمه

هرگاه سیالی از لحاظ یکی از خواص ترکیب، دما و سرعت مولکول‌ها غیریکنواخت باشد پدیده‌های انتقالی برای یکنواخت شدن آن خاصیت در سیستم اتفاق می‌افتد. به طور کلی انتقال یک خاصیت معین مثل جرم، اندازه‌ی حرکت و یا انرژی از یک قسمت سیستم به قسمت دیگر را فرآیند انتقالی می‌گویند و میزان انتقال آن خاصیت معین را با شار^۱ آن خاصیت بیان می‌کنند. شار مقداری از خاصیت است که از واحد سطح در واحد زمان می‌گذرد. مشاهدات تجربی خواص انتقالی نشان می‌دهد که شار یک خاصیت (J) با گرادیان خاصیت مربوطه (∇Y) متناسب می‌باشد:

$$J \propto \nabla Y \quad (1-1)$$

امروزه دانستن خواص انتقالی سیالات، از اهمیت خاصی در طراحی صنعتی برخوردار است. به عنوان مثال دانستن اطلاعات کامل و دقیق از ویسکوزیته‌ی سیالات در طراحی بهینه‌ی دستگاه‌هایی مثل مبدل‌های حرارتی، مبردها^۲ و در هر جا که سیالی در حال حرکت است، مهم است. همچنین در صنایع نفتی، مقدار نشتی مواد آلاینده از شکاف‌های ایجاد شده در خطوط انتقال و یا مخازن حمل و نقل با ویسکوزیته‌ی سیال در ارتباط است. از لحاظ نظری نیز ارتباط ویسکوزیته با نیروهای بین مولکولی که به شناخت هر چه بیشتر ماهیت و رفتار مواد برمی‌گردد، همیشه مورد توجه بوده است.

[۱].

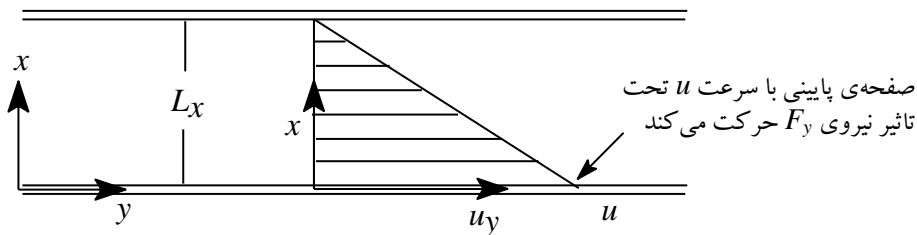
۱-۲- ویسکوزیته

خاصیتی است که سختی جاری شدن یک سیال را مشخص می‌کند و با η نشان داده می‌شود. به منظور دسترسی به تعریف دقیقی از ویسکوزیته، سیالی را در نظر بگیرید که بین دو صفحه‌ی موازی و نزدیک به هم قرار دارد، شکل (۱-۱). این صفحات به قدری بزرگ هستند که می‌توان از اثرات لبه‌های

^۱ - flux

^۲ - refrigerant

آنها صرف نظر کرد. صفحه‌ی بالایی ثابت است و صفحه‌ی پایینی با مساحت A توسط نیروی F_y کشیده می‌شود. این نیروی F_y باعث می‌شود صفحه‌ی پایینی با سرعت u_y حرکت کند. اگر سیال را به شکل لایه‌هایی بر روی یکدیگر فرض کنیم، لایه‌های سیالی که در مجاورت صفحات قرار دارند، سرعتی برابر با سرعت صفحات دارند. لذا سرعت سیال در سطح صفحه‌ی پایینی، u_y و در سطح صفحه‌ی بالایی صفر خواهد بود. با گذشت زمان سرعت لایه‌های میانی افزایش می‌یابد و آنها نیز به موازات صفحات حرکت می‌کنند به طوری که سرعت لایه‌های متوالی میانی به نسبت فاصله از صفحه‌ی پایینی به طور یکنواخت تغییر می‌کند. به این ترتیب، لایه‌های افقی مجاور هم در سیال با سرعت‌های متفاوت حرکت می‌کنند و روی یکدیگر می‌غلطند. هنگام عبور این لایه‌ها از روی یکدیگر، هر لایه توسط لایه‌ی بالایی که کندتر حرکت می‌کند عقب می‌افتد و از آن طرف هر لایه با لایه‌ی پایینی که سریع‌تر حرکت می‌کند شتاب می‌گیرد. این تبادل اندازه حرکت بین لایه‌های مجاور هنگام عبور از روی هم باعث می‌شود که لایه‌ها یک نیروی مقاومت اصطکاکی به یکدیگر وارد کنند و همین اصطکاک بین لایه‌های سیال عامل ایجاد ویسکوزیته‌ی سیال است [۱ و ۲]. بنابراین ویسکوزیته در سیالات حقیقی، ناشی از انتقال اندازه‌ی حرکت بین لایه‌های سیال می‌باشد.



شکل (۱-۱): جریان سیال بین دو صفحه‌ی موازی، وقتی که صفحه‌ی بالایی ثابت است.

اما از دیدگاه مولکولی، اندازه حرکت چگونه بین لایه‌های سیال انتقال می‌یابد؟ انتقال اندازه حرکت بین لایه‌های سیال به دو صورت می‌تواند صورت گیرد: همرفت و رسانش. همان طور که توضیح داده شد اگر سیال را به شکل لایه‌هایی بر روی یکدیگر فرض کنیم، چون لایه‌های مجاور با

سرعت‌های مختلف حرکت می‌کنند مولکول‌های موجود در این لایه‌ها دارای مقادیر متفاوت از متوسط اندازه حرکت هستند. در گازها حرکات اتفاقی مولکول‌ها، باعث انتقال مولکول‌های لایه‌ی سریع‌تر به لایه‌ی کندتر شده و این اثر باعث افزایش اندازه‌ی حرکت لایه‌ی کندتر می‌گردد. این پدیده را انتقال اندازه حرکت در اثر پدیده‌ی همرفت می‌نامند. مبادله‌ی اندازه‌ی حرکت بین مولکول‌های لایه‌های مجاور می‌تواند بدون انتقال مولکول‌ها از یک لایه به لایه‌ی دیگر و فقط در اثر برخورد ذرات صورت پذیرد. این روش انتقال اندازه حرکت را پدیده‌ی رسانش می‌نامند. در سیالات رقیق انتقال اندازه حرکت، بیشتر در اثر پدیده‌ی همرفت صورت می‌گیرد و با افزایش چگالی سیال، سهم رسانش در انتقال اندازه حرکت بیشتر می‌شود. با توجه به این توضیحات به خوبی می‌توان فهمید که چرا ویسکوزیته‌ی گازها با افزایش دما، افزایش ولی ویسکوزیته‌ی مایعات با افزایش دما کاهش می‌یابد. افزایش دما در گاز باعث افزایش متوسط سرعت ذرات در جهات مختلف شده و این عامل انتقال اندازه حرکت در اثر پدیده‌ی همرفت را افزایش می‌دهد، در نتیجه ویسکوزیته‌ی گاز افزایش می‌یابد. افزایش دما در مایعات باعث کاهش چسبندگی لایه‌ها و در نتیجه کاهش ویسکوزیته‌ی آن‌ها می‌گردد.

۱-۲-۱- ارتباط شار اندازه حرکت و ویسکوزیته

همان طور که توضیح داده شد اگر سیال به شکل لایه‌هایی بر روی یکدیگر فرض شود، اعمال نیروی F_y باعث ایجاد اندازه حرکت در راستای y می‌شود. در این صورت، شار اندازه حرکت در

راستای y (τ_{xy}) و در واحد سطح با $\frac{\partial u_y}{\partial x}$ متناسب است، بنابراین می‌توانیم رابطه‌ی زیر را بنویسیم:

$$\tau_{xy} \propto \frac{\partial u_y}{\partial x} \quad (۲-۱)$$

که در آن u_y مؤلفه‌ی سرعت لایه‌ها در راستای y می‌باشد. لازم به تذکر است که اولین زیروند τ_{xy} (شار اندازه حرکت) به سطحی اشاره دارد که اندازه حرکت از آن منتقل می‌شود و دومین زیروند به

راستای اندازه حرکت انتقالی اشاره می‌کند. این عبارت با در نظر گرفتن یک ضریب تناسب می‌تواند به تساوی تبدیل شود:

$$\tau_{xy} = -\eta \frac{\partial u_y}{\partial x} \quad (3-1)$$

که در آن، ثابت η به ضریب ویسکوزیته یا ویسکوزیته معروف است. علت این که ضریب تناسب در عبارت شار اندازه حرکت، منفی در نظر گرفته شده است این است که شار، مؤلفه‌ی یک بردار است. اگر $\tau_{xy} > 0$ باشد، شار به سمت افزایش x است و برعکس اگر $\tau_{xy} < 0$ باشد، شار به سمت کاهش x می‌باشد. از طرف دیگر، جریان اندازه حرکت در جهت نزولی گرادیان سرعت برقرار می‌شود. بنابراین، اگر $\frac{\partial u_y}{\partial x} < 0$ باشد یعنی سرعت با افزایش x کاهش یابد، آنگاه $\tau_{xy} > 0$ است. بنابراین، ضریب تناسب در عبارت شار (3-1) باید منفی باشد که این ضریب با $-\eta$ مشخص شده است [3].

معادله‌ی (3-1) قانون ویسکوزیته‌ی نیوتون است و سیالاتی که از معادله‌ی فوق پیروی می‌کنند سیالات نیوتنی^۳ نامیده می‌شوند. ویسکوزیته‌ی بسیاری از سیالات مهم در صنعت و در طبیعت از معادله‌ی (3-1) پیروی نمی‌کنند که به آن‌ها سیالات غیرنیوتنی می‌گویند [1].

۳-۱- مروری بر روش‌های پیش‌بینی ویسکوزیته

تاریخچه‌ی مطالعات نظری روی ویسکوزیته و سایر پدیده‌های انتقالی به نوعی ریشه در نظریه‌های اولیه ارائه شده برای سینتیک گازها دارد که با کارهای برنولی در قرن هجدهم شروع شد و با نظریات افرادی چون کلازیوس^۴، ماکسول^۵ و بولتزمن^۶ توسعه یافت.

^۳ - Newtonian fluids

^۴ - Clausius

از آن جا که یکی از اهداف این پایان نامه محاسبه‌ی ویسکوزیته‌ی ترکیبات آلی چگال می‌باشد، در ابتدا نگاهی اجمالی به روش‌های محاسباتی ارائه شده برای ویسکوزیته‌ی گازها خواهیم داشت و سپس روش‌های مختلف پیش‌بینی ویسکوزیته‌ی ترکیبات آلی مورد بررسی قرار خواهند گرفت.

۱-۳-۱- معادله‌ی ماکسول - بولتزمن

اولین مقاله‌ی ماکسول با عنوان « شرح نظریه‌ی دینامیکی گازها » در سال ۱۸۶۰ ارائه گردید. او برای اولین بار حرکت مولکول‌ها را تصادفی فرض کرد و برای یک سیستم در حال تعادل گرمایی، تابع توزیع سرعت را چنین به دست آورد:

$$f_0(u) = n \left(\frac{m}{2\pi k_B T} \right)^{3/2} \exp\left(-\frac{mu^2}{2k_B T} \right) \quad (۴-۱)$$

که u سرعت، n چگالی ذرات، m جرم ذرات، k_B ثابت بولتزمن و T دمای کلین است. سپس او با استفاده از تابع توزیع ارائه شده، ویسکوزیته‌ی گازهای رقیق، η_0 را محاسبه کرد:

$$\eta_0 = \frac{1}{3} n m l \bar{u} \quad (۵-۱)$$

اما از آن جا که ویسکوزیته‌ی یک سیال در دمای ثابت با افزایش چگالی افزایش می‌یابد، معادله‌ی (۵-۱) با چالش اساسی مواجه گردید. بنابراین، بولتزمن در سال ۱۸۷۲، یک رهیافت عمومی از نظریه‌ی سینتیک گازها را برای فرایندهای غیرتعادلی ارائه داد اما به دلیل پیچیدگی‌های ریاضی هرگز موفق به حل آن نشد [۲].

^۵ - Maxwell

^۶ - Boltzman

۱-۳-۲- روش چاپمن - انسکوگ

چاپمن^۷ و انسکوگ^۸ به طور مستقل در سال‌های ۱۹۱۶ و ۱۹۱۷ موفق به ارائه‌ی روشی برای حل معادله‌ی بولتزمن شدند [۴]. روش چاپمن - انسکوگ برای ویسکوزیته‌ی یک گاز در حد چگالی صفر به معادله‌ی (۶-۱) منجر گردید:

$$\eta_0 = \frac{5}{16} \left(\frac{mkT}{\pi} \right)^{1/2} \frac{f_\eta}{\sigma^2 \Omega^{(2,2)*}(T^*)} \quad (6-1)$$

که در آن $\Omega^{(2,2)*}(T^*)$ ، انتگرال برخورد کاهش‌یافته بوده و تابعی از دما و برهم‌کنش‌های بین مولکولی است و f_η فاکتور تصحیح است که در چگالی‌های کم تأثیر ناچیزی داشته و مقدار نزدیک به یک دارد [۵].

۱-۳-۳- نظریه‌ی انسکوگ بهبودیافته^۹ (MET)

ویسکوزیته‌ی سیال کروی سخت چگال طبق نظریه‌ی انسکوگ عبارت است از:

$$\eta = \eta_0 b_0 \rho \left(\frac{1}{Y} + 0.800 + 0.761Y \right) \quad (7-1)$$

که در آن η ویسکوزیته‌ی سیال کروی سخت، η_0 ویسکوزیته‌ی گاز کروی سخت رقیق، b_0 حجم ممنوعه^{۱۰} بوده و با $(2/3)\pi\sigma^3$ برابر است که σ قطر کروی سخت می‌باشد و $Y = b_0 \rho g(\sigma)$ که $g(\sigma)$ تابع توزیع شعاعی کروی سخت در $r = \sigma$ است. مقدار Y به صورت زیر به معادله‌ی حالت کروی سخت مربوط است:

^۷- Chapman

^۸- Enskog

^۹- modified Enskog theory (MET)

^{۱۰}- prohibited-volume

^{۱۱}- number density

$$Y = \frac{P}{\rho kT} - 1 \quad (8-1)$$

که در آن p ، فشار سیال است. خود انسکوگ در مرحله‌ی بعد، نظریه‌ی فوق را با دو اصلاحیه به سیالات حقیقی چگال توسعه داد که به آن نظریه‌ی انسکوگ بهبودیافته می‌گویند. او به جای ارتباط Y با فشار سیال طبق معادله‌ی (8-1)، آن را به فشار حرارتی سیال نسبت داد. علت چنین پیشنهادی این بود که فشاری که یک مولکول احساس می‌کند از دو بخش تشکیل شده است: فشار ناشی از دیواره‌ی ظرف p و فشار داخلی، $(\partial E/\partial V)_T$ ، که نیروی چسبندگی بین مولکول‌ها را نشان می‌دهد و حاصل جمع این دو، طبق رابطه‌ی دقیق ترمودینامیکی برابر با فشار حرارتی، $T\left(\frac{\partial P}{\partial T}\right)_\rho$ است. که

ضریب فشار حرارتی نامیده می‌شود. $\left(\frac{\partial P}{\partial T}\right)_\rho$

$$T\left(\frac{\partial P}{\partial T}\right)_\rho = P + \left(\frac{\partial E}{\partial V}\right)_T \quad (9-1)$$

با استفاده از فشار حرارتی به جای فشار سیال، معادله‌ی (8-1) به معادله‌ی (10-1) تبدیل گردید:

$$Y = \frac{1}{\rho kT} \left[T\left(\frac{\partial p}{\partial T}\right)_\rho \right] - 1 \quad (10-1)$$

همچنین با توجه به این که در حد دانسیته‌ی صفر، $Y/\rho \rightarrow b_0$ ، را برحسب ضرایب دوم (B) و سوم ویریا (C) و مشتق‌های آن‌ها نسبت به دما چنین تعریف نمود:

$$b_0 = B(T) + T \frac{dB}{dT} + C(T) + T \frac{dC}{dT} \quad (11-1)$$

که در آن $B(T)$ و $C(T)$ به ترتیب ضرایب دوم و سوم ویریا هستند. به این ترتیب با معلوم بودن معادله‌ی حالت دقیق و مقادیر ضرایب ویریا برحسب دما برای هر سیال و با استفاده از معادله‌های

فوق می‌توان ویسکوزیته‌ی سیالات چگال را محاسبه نمود [۴]. در مرجع [۶] با استفاده از نظریه‌ی MET، ویسکوزیته‌ی سیالات فوق بحرانی محاسبه شده است.

۱-۳-۴- نظریه‌ی رین واتر - فرند

در نظریه‌ی رین واتر^{۱۲} و فرند^{۱۳} که در سال ۱۹۸۴ ارائه شد ضریب دوم ویریا انتقالی برای گازها، B_{η} ، براساس دیدگاه میکروسکوپی توسعه داده شد [۷ و ۸]. در مدل رین واتر و فرند یک گاز در چگالی‌های متوسط (حدود ۲ مول بر لیتر) به صورت مخلوطی از مونومرها و دی‌مرها در نظر گرفته می‌شود. سپس وابستگی ویسکوزیته به چگالی به شکل یک بسط ویریا نوشته می‌شود:

$$\eta = \eta_0(1 + B_{\eta}\rho + \dots) \quad (1-12)$$

که η_0 ویسکوزیته‌ی چاپمن-انسکوگ، B_{η} ضریب دوم ویریا ویسکوزیته و ρ چگالی است. در این صورت B_{η} را می‌توان بر حسب مجموع سه جمله‌ی سهم دو مونومر، سهم مونومر - دایمر و اثر سهم ذره‌ی سوم بر برخوردهای دو مونومر نوشت. آن‌ها به این ترتیب موفق شدند B_{η} را با توافق خوبی در مقایسه با داده‌های تجربی به دست آورند و ویسکوزیته را در چگالی‌های متوسط محاسبه کنند.

۱-۳-۵- توابع همبستگی^{۱۴}

نظریه‌های پایه‌ای مثل نظریه‌ی سینتیک گازها و مکانیک آماری قادر به محاسبه‌ی خواص ماکروسکوپی مجموعه‌ای از مولکول‌ها براساس خواص مولکولی و نیروهای بین مولکولی هستند. ولی واقعیت این است که به جز در نواحی گازی رقیق که برهم‌کنش‌ها شناخته شده‌اند یک نظریه‌ی مبسوط برای پدیده‌های انتقالی وجود ندارد. بنابراین مراجعه به اندازه‌گیری‌های آزمایشگاهی تنها راه

^{۱۲}- Rain water

^{۱۳}- Friend

^{۱۴}- correlation functions

است ولی با توجه به این که برای دوازده اندازه‌گیری ترموشیمی فیزیکی در ده دما و ده فشار متفاوت برای مواد شیمیایی مورد مصرف در صنایع و مخلوط آن‌ها به تعداد بسیار زیادی اندازه‌گیری نیاز خواهد بود اهمیت بسزای توابع همبستگی مشخص می‌شود. یک تابع همبستگی همانند یک معادله ریاضی است که با مشخص بودن حالت ترمودینامیکی سیستم، متغیرهای ترمودینامیکی مورد نظر را می‌توان با استفاده از آن محاسبه کرد.

استفاده از توابع همبستگی برای محاسبه‌ی خواص انتقالی، اهمیت ویژه‌ای در روش‌های کاربردی دارد. اگر چه ممکن است برخی از این توابع، از دقت بالایی برخوردار نباشند اما بر آورد اولیه از یک خاصیت که با استفاده از توابع همبستگی به دست می‌آید کمک ارزشمندی به طراحی اولیه در صنایع می‌کند [۹].

به طور مثال، یکی از توابع همبستگی که برای پیش‌بینی ویسکوزیته‌ی گازها در محدوده‌ی وسیعی از دما و فشار ارائه شده است معادله‌ی (۱۳-۱) می‌باشد. با استفاده از این معادله‌ی پنج پارامتری می‌توان ویسکوزیته‌ی گازها را پیش‌بینی کرد [۵].

$$\eta(\rho, T) = \eta_0(T) + (1 + cT)^2 \times \left(\frac{a_1\rho + a_2\rho^2}{1 + b_1\rho + b_2\rho^2} \right) \quad (13-1)$$

سپس معادله‌ی (۱۴-۱) با سه پارامتر برای محاسبه راحت‌تر ویسکوزیته‌ی گازها معرفی گردید [۱۰].

$$\eta(\rho, T) = \eta_0(T) + \exp\left(\frac{-1}{2T}\right) \times \left(\frac{a_0\rho}{1 + a_1\rho + a_2\rho^2} \right) \quad (14-1)$$

این معادله شامل سه پارامتر a_0 و a_1 و a_2 می‌باشد که این پارامترها ضرایب مستقل از دما و دانسیته بوده و به نوع سیال بستگی دارند.

۴-۱- روش‌های مختلف پیش‌بینی ویسکوزیته‌ی ترکیبات آلی

با وجود مطالعات گسترده در زمینه‌ی روش‌های پیش‌بینی ویسکوزیته‌ی سیالات رقیق، خصوصاً گازهای تک اتمی و مبردها، کارهای تحقیقاتی انجام شده در زمینه‌ی پیش‌بینی ویسکوزیته‌ی ترکیبات آلی چگال بسیار اندک می‌باشد. شاید یکی از علل این موضوع را باید در عدم شناخت دقیق برهم‌کنش‌های موجود در این سیستم‌ها جستجو کرد. البته عدم وجود داده‌های آزمایشگاهی برای بیشتر این ترکیبات نیز خود مزید بر علت می‌باشد. از آنجا که یکی از اهداف این پایان نامه، پیش‌بینی ویسکوزیته‌ی ترکیبات آلی چگال می‌باشد، در ادامه روش‌های محاسباتی ارائه شده برای پیش‌بینی ویسکوزیته‌ی ترکیبات آلی ارائه خواهد شد.

۱-۴-۱- توابع همبستگی ارائه شده برای پیش‌بینی ویسکوزیته‌ی ترکیبات آلی

۱-۴-۱-الف- روش تاداس^{۱۵}

در سال ۱۹۵۲، رابطه‌ی همبستگی (۱-۱۵) بین هدایت گرمایی باقیمانده‌ی یک مایع و دانسیته‌ی مربوطه‌اش ارائه گردید [۱۱]:

$$\lambda - \lambda_0 = f(\rho) \quad (1-15)$$

که در آن λ هدایت گرمایی سیال در چگالی مورد نظر، λ_0 هدایت گرمایی در حد چگالی صفر برای سیال مورد نظر و $f(\rho)$ تابعی از دانسیته می‌باشد.

تاداس و همکارانش با بررسی خواص انتقالی مواد مختلف در محدوده‌ی وسیعی از دما و فشار [۱۲] و [۱۳] نشان دادند که معادله‌ی (۱-۱۵) و عبارت مشابه آن برای ویسکوزیته‌ی باقیمانده در هر دو فاز گاز و مایع توافق بسیار عالی با داده‌های آزمایشگاهی دارد. بر این اساس آن‌ها تصمیم گرفتند با استفاده از مقادیر ویسکوزیته در فشار اتمسفر و نمودارهای خواص انتقالی باقیمانده برحسب دانسیته،

^{۱۵}- Thodos

ویسکوزیته‌ی ترکیبات آلی چگال را به دانسیته و ثابت‌های بحرانی این ترکیبات مرتبط کند. آن‌ها سرانجام موفق شدند تابع همبستگی (۱۶-۱) را برای محاسبه‌ی ویسکوزیته‌ی این ترکیبات برحسب دانسیته‌ی کاهش‌یافته ارائه کنند:

$$\begin{aligned} [(\eta - \eta_0)\xi + 10^{-4}]^{0.25} = & 0.10230 + 0.023364\rho_r + 0.058533\rho_r^2 \\ & - 0.040758\rho_r^3 + 0.0093324\rho_r^4 \end{aligned} \quad (16-1)$$

که در آن η ، ویسکوزیته‌ی سیال چگال در چگالی مورد نظر و η_0 ، ویسکوزیته در حد چگالی صفر برای سیال مورد نظر می‌باشد که هر دو بر حسب $\mu\text{Pa}\cdot\text{s}$ می‌باشند، ρ_r دانسیته‌ی کاهش یافته است و ξ یک مشخصه‌ی ماده است که از معادله‌ی (۱۷-۱) به دست آید:

$$\xi = \frac{T_c^{1/6}}{M^{1/2} p_c^{2/3}} \quad (17-1)$$

که در آن T_c دمای بحرانی، p_c فشار بحرانی و M جرم مولکولی سیال مورد نظر است. آن‌ها با این روش توانستند ویسکوزیته‌ی یازده ترکیب آلی از اتان تا n -دکان، سیکلوپنتان و سیکلوهگزان را با متوسط خطای نسبی $13/8\%$ پیش‌بینی کنند [۱۴].

۱-۴-۱-ب- روش چانگ^{۱۶}

در سال ۱۹۸۰، چانگ بر مبنای نظریه‌ی انسکوگ برای گازهای چگال [۴] و مفهوم خاصیت انتقالی باقیمانده، که طبق آن خاصیت انتقالی باقیمانده با دانسیته همبسته^{۱۷} می‌شود [۱۱ و ۱۳]، معادله‌ی (۱۸-۱) را برای پیش‌بینی ویسکوزیته‌ی ترکیبات آلی چگال ارائه کرد [۱۵]:

$$\eta = \left[\eta_0^* \left(\frac{1}{G_1(\rho)} + A_6 Y \right) + A_7 Y^2 G_1(\rho) \exp C(T) \right] \left(2.6693 \frac{\sqrt{M \cdot \varepsilon/k}}{\sigma^2} \right) \quad (18-1)$$

^{۱۶}- Chang

^{۱۷}- correlate

در عبارت (۱۸-۱)، η ویسکوزیته برحسب $\mu\text{Pa.s}$ و η_0^* ویسکوزیته‌ی کاهش‌یافته در حد چگالی صفر است که از معادله‌ی (۱۹-۱) به دست می‌آید:

$$\eta_0^* = \eta_{\text{CE}}^* (1 - 0.275\omega + 0.059035\mu^{*4} + \chi) \quad (19-1)$$

که در آن ω ، فاکتور شکل است و μ^* ، گشتاور دو قطبی کاهش‌یافته بوده و از رابطه‌ی (۲۰-۱) به دست می‌آید. χ پارامتر پیوستگی^{۱۸} نامیده می‌شود و فقط برای مواد قطبی مقدار دارد زیرا این مواد می‌توانند تجمع‌های مولکولی مختلف تشکیل دهند. مقادیر χ برای چند ترکیب در جدول (۱-۳) آمده است و برای سایر الکل‌ها می‌توان با استفاده از معادله‌ی (۲۰-۱)، χ را به دست آورد.

$$\mu^* = \frac{100\mu}{\sqrt{1.380535\sigma^3 \varepsilon/k}} \quad (20-1)$$

$$\chi = 0.0682 + 0.276659 \times \left(\frac{17 \times \text{no. of OH groups}}{\text{molar mass}} \right) \quad (21-1)$$

که $\sigma^3 = 0.3189/\rho_c$ و $\varepsilon/k = T_c/1.2593$ است. در این عبارت‌ها ρ_c ، دانسیته‌ی بحرانی برحسب $\text{molecule}/\text{\AA}^3$ و T_c ، دمای بحرانی برحسب K است. η_{CE}^* نیز از رابطه‌ی (۲۲-۱) محاسبه می‌شود:

$$\eta_{\text{CE}}^* = \frac{T^{*1/2}}{\Omega^{(2,2)*}} \quad (22-1)$$

که در آن $\Omega^{(2,2)*}$ انتگرال برخورد کاهش یافته است و از رابطه‌ی (۲۳-۱) محاسبه می‌شود [۱۰۱] و T^* دمای کاهش یافته است و از معادله‌ی (۲۴-۱) به دست می‌آید:

$$\Omega^{(2,2)*} = \frac{A}{T^{*B}} + \frac{C}{\exp(DT^*)} + \frac{E}{\exp(FT^*)} + GT^{*B} \sin(ST^{*W} - P) \quad (23-1)$$

$$T^* = \frac{T}{\varepsilon/k} = 1.2593 \frac{T}{T_c} \quad (24-1)$$

^{۱۸} - association parameter

که ضرایب این معادله در جدول (۲-۱) آمده است.

جدول (۱-۱): پارامترهای پیوستگی برای چند ترکیب آلی.

ماده	χ
متانول	۰/۲۱۵۱۷۵
اتانول	۰/۱۷۴۸۲۳
پروپانول	۰/۱۴۳۴۵۳
بوتانول	۰/۱۳۱۶۷۱
استیک اسید	۰/۰۹۱۵۴۹
استیک اسید اتیل استر	۰/۰۳۰۹۲۴

جدول (۲-۱): ضرایب معادله‌ی (۲۲-۱).

A	۱/۱۶۱۴۵
B	۰/۱۴۸۷۴
C	۰/۵۲۴۸۷
D	۰/۷۷۳۲۰
E	۲/۱۶۱۷۸
F	۲/۴۳۷۸۷
P	$-۶/۴۳۵۰۰ \times ۱۰^{-۴}$
S	۷/۲۷۳۷۱
W	

	۱۸/۰۳۲۳۰
	-۰/۷۶۸۳۰

تابع $G_1(\rho)$ در عبارت (۱۸-۱) نیز چنین معین می‌شود:

$$G_1(\rho) = \frac{A_1 \left(\frac{1 - \exp(A_4 Y)}{Y} \right) + A_2 G \exp(A_5 Y) + A_3 G}{A_1 A_4 + A_2 + A_3} \quad (25-1)$$

که توابع G و Y در این عبارت، از معادلات (۲۶-۱) و (۲۷-۱) به دست می‌آیند.

$$G = \frac{1 - 0.5Y}{(1 - Y)^3} \quad (26-1)$$

$$Y = \frac{\pi \rho^*}{6} = \frac{\pi \rho \sigma^3}{6} = \frac{0.3189 \pi \rho}{6 \rho_c} \quad (27-1)$$

تابع $C(T)$ در معادله‌ی (۱۸-۱) نیز از معادله‌ی (۲۸-۱) معین می‌شود:

$$C(T) = A_8 + \frac{A_9}{T^*} + \frac{A_{10}}{T^{*2}} \quad (28-1)$$

که در آن T^* دمای کاهش یافته از معادله‌ی (۲۴-۱) به دست می‌آید. ضرایب A_1 تا A_{10} از معادله‌ی

(۲۹-۱) به دست می‌آیند:

$$A_i = \sum_{i=1}^{10} (a_{0i} + a_{1i} \omega + a_{2i} \mu^{*4} + a_{3i} \chi) \quad (29-1)$$

ضرایب a_{0i} تا a_{3i} در جدول (۳-۱) آمده است.

جدول (۳-۱): ضرایب معادله‌ی (۱-۲۹).

i	a_{0i}	a_{1i}	a_{2i}	a_{3i}
۱	۶/۳۲۴۰۲	۵۰/۴۱۱۹۰	-۵۱/۶۸۰۱۰	۱۱۸۹/۰۲۰۰۰۰
۲	$۱/۲۱۰۲۰ \times ۱۰^{-۳}$	$-۱/۲۱۰۲۰ \times ۱۰^{-۳}$	$-۶/۲۵۷۱۰ \times ۱۰^{-۳}$	$۳/۷۲۸۳۲ \times ۱۰^{-۲}$
۳	۵/۲۸۳۴۶	۲۵۴/۲۰۹۰۰	-۱۶۸/۴۸۱۰۰	۳۸۹۸/۲۷۰۰۰
۴	۶/۶۲۲۶۳	۳۸/۰۹۵۷۰	-۸/۴۶۴۱۴	۳۱/۴۱۷۸۰
۵	۱۹/۷۴۵۴۰	۷/۶۳۰۳۴	-۱۴/۳۵۴۴۰	۳۱/۵۲۶۷۰
۶	-۱/۸۹۹۹۲	-۱۲/۵۳۶۷۰	۴/۹۸۵۲۹	-۱۸/۱۵۰۷۰
۷	۲۴/۲۷۴۵۰	۳/۴۴۹۴۵	-۱۱/۲۹۱۳۰	۶۹/۳۴۶۶۰
۸	$۷/۹۷۱۶۳ \times ۱۰^{-۱}$	۱/۱۱۷۶۴	$۱/۲۳۴۸۳ \times ۱۰^{-۲}$	-۴/۱۱۶۶۱
۹	$-۲/۳۸۱۶۵ \times ۱۰^{-۱}$	$۶/۷۶۹۵۴ \times ۱۰^{-۲}$	$-۸/۱۶۲۹۷ \times ۱۰^{-۱}$	۴/۰۲۵۲۸
۱۰	$۶/۸۶۲۹۵ \times ۱۰^{-۲}$	$۳/۴۷۹۲۸ \times ۱۰^{-۱}$	$۵/۹۲۵۶۱ \times ۱۰^{-۱}$	$-۷/۲۶۶۳۳ \times ۱۰^{-۱}$

چانگ با این تابع همبستگی پیچیده توانست ویسکوزیته‌ی آلکان‌های خطی از متان تا n -دودکان، بنزن، تولوئن، اتیل بنزن، ۱-هگزن، ۱-هپتن، ۱-اکتن، سیکلوهگزان، ایزوبوتان را با متوسط خطای نسبی ۱۳/۶٪ پیش‌بینی کند [۱۵].

۱-۴-۱-ج- روش هانلی و الی

در سال ۱۹۸۱، هانلی^{۱۹} و الی^{۲۰} بر مبنای اصل حالت‌های متناظر و حل تطبیقی^{۲۱} توانستند روشی را برای پیش‌بینی ویسکوزیته‌ی سیالات آلی شامل آلکان‌های خطی از اتان تا n -دودکان، ایزوبوتان، سیکلوهگزان، متیل سیکلوهگزان، اتیل سیکلوهگزان، ۱-هگزن، ۱-هپتن، ۱-اکتن، بنزن، تولوئن در محدوده‌ی وسیعی از دما و فشار ارائه دهند [۱۶]. آن‌ها برای انجام این کار نیاز به یک سیال مرجع^{۲۲} داشتند تا بتوانند بر مبنای آن محاسبات خود را انجام دهند. چون هدف آن‌ها پیش‌بینی ویسکوزیته‌ی ترکیبات آلی به خصوص آلکان‌های خطی بود، از این رو ترکیب متان را به عنوان مرجع انتخاب کردند، که با ترکیبات مورد بررسی هماهنگی داشته باشد. آن‌ها توانستند ویسکوزیته‌ی ۲۰ ترکیب آلی را در محدوده‌ی وسیعی از دما و فشار با متوسط خطای نسبی $10/14\%$ پیش‌بینی کردند [۱۶].

۱-۴-۲- پیش‌بینی ویسکوزیته‌ی ترکیبات آلی بر اساس روش سهم گروهها

در متون علمی، فقط یک روش سهم گروهها برای پیش‌بینی ویسکوزیته‌ی ترکیبات آلی مایع برحسب دما گزارش شده است. در این کار ساستری^{۲۳} و رائو^{۲۴} برای تعیین سهم گروهها در ویسکوزیته، از کار قبلی خود ایده گرفتند [۱۷]. نتایج کار قبلی آن‌ها نشان داده بود که نمودار ویسکوزیته‌ی مایعات نسبت به فشار بخار در همان دما در مقیاس لگاریتمی خطی است و این ارتباط را می‌توان از رابطه‌ی (۱-۳۰) مشاهده کرد:

$$\eta = Ap^{-N} \quad (30-1)$$

^{۱۹}- Hanley

^{۲۰}- Ely

^{۲۱}- conformal solution

^{۲۲}- reference

^{۲۳}- Sastri

^{۲۴}- Rao

به علاوه، یک تغییر تیز در نقطه‌ی جوش نرمال ترکیبات مورد بررسی مشاهده کردند که نشان

می‌داد در نقطه‌ی جوش نرمال، رابطه‌ی فوق به صورت رابطه‌ی (۳۱-۱) درمی‌آید:

$$A = \eta_b \quad (31-1)$$

که در آن η_b ویسکوزیته‌ی مایع در نقطه‌ی جوش نرمال است. با جایگذاری عبارت (۳۱-۱) در

(۳۰-۱)، عبارت زیر برای ویسکوزیته به دست می‌آید:

$$\eta = \eta_b p^{-N} \quad (32-1)$$

سپس با بررسی‌های بیشتر مشخص شد که برای اعضای یک سری همولوگ با طبیعت مشابه در فاز

مایع، η_b و N در عبارت (۲۰-۱) بسیار کم تغییر می‌کند. آن‌ها از این موضوع، برای ارتباط

ویسکوزیته‌ی مایعات آلی با ساختار شیمیایی آن‌ها استفاده کردند. برای این منظور، ابتدا با استفاده از

تعداد زیادی داده‌ی تجربی برای ویسکوزیته و فشار بخار مایعات آلی بر حسب دما، مقادیر η_b و N را

به دست آوردند. سپس با استفاده از این مقادیر سهم گروه‌های عاملی مختلف در η_b و N را بر اساس

معادلات (۳۳-۱) و (۳۴-۱) استخراج نمودند:

$$\eta_b = \sum \Delta \eta_b + \sum \Delta \eta_b^{\text{cor}} \quad (33-1)$$

$$N = 0.2 + \sum \Delta N + \sum \Delta N^{\text{cor}} \quad (34-1)$$

آن‌ها با استفاده از سهم گروه‌های عاملی در η_b و N و معادلات (۲۰-۱) و (۲۱-۱) مقادیر η_b و N را

برای ۱۲۰ هیدروکربن مختلف شامل n -آلکان‌ها، آلکان‌های شاخه‌ای، سیکلوآلکان‌ها، ترکیبات

آلیفانیک اشباع نشده و ترکیبات آروماتیک محاسبه کردند. سپس با استفاده از مقادیر η_b و N

محاسباتی و مقادیر فشار بخار تجربی این ترکیبات در هر دما، ویسکوزیته‌ی این هیدروکربن‌ها را بر

حسب دما پیش‌بینی کردند. متوسط خطاهای نسبی آن‌ها در پیش‌بینی ویسکوزیته‌ی آلکان‌های خطی

بر حسب دما $7/5\%$ ، آلکان‌های شاخه‌ای $8/8\%$ ، سیکلوآلکان‌ها 9% ، ترکیبات آلیفاتیک اشباع نشده $5/7\%$ و ترکیبات آروماتیک $6/6\%$ می‌باشد [۱۸].

۱-۴-۳- پیش‌بینی ویسکوزیته‌ی ترکیبات آلی با شبکه‌های عصبی

در متون علمی یک گزارش برای پیش‌بینی ویسکوزیته‌ی ترکیبات آلی بر حسب دما به وسیله‌ی شبکه‌های عصبی موجود می‌باشد که توسط هومر^{۲۵} و همکارانش در سال ۱۹۹۹ ارائه شده است. آن‌ها برای این منظور از یک شبکه‌ی عصبی چند لایه‌ی پیش‌خور با آلفا-۱ آموزش لونی-مارکوارت استفاده کردند. ۱۱ ورودی برای طراحی شبکه به کار رفته است که شامل دما، پنج خاصیت فیزیکی (وزن مولکولی، فاکتور شکل پیتزر، نقطه‌ی جوش نرمال، دانسیته‌ی مایعات بر حسب دما، گرمای تبخیر) و پنج خاصیت ساختاری (تعداد حلقه‌های آلیفاتیک، تعداد حلقه‌های آروماتیک، تعداد پیوندهای دوگانه، تعداد پیوندهای سه‌گانه، تعداد پیوندهای دوگانه در حلقه‌های آلیفاتیک) می‌باشد. پس از انتقال اطلاعات به شبکه و آموزش آن، پارامترهای بهینه‌ی شبکه تعیین گردید. سپس آن‌ها با استفاده از این شبکه‌ی آموزش دیده توانستند ویسکوزیته‌ی هیدروکربن‌های مختلف را بر حسب دما $(0/7 < T_r < 0/45)$ با متوسط خطای نسبی $8/9\%$ پیش‌بینی کنند [۱۹].

^{۲۵}- Homer

فصل دوم

کمومتریکس

۲-۱- مقدمه

با توجه به گسترش سیستم‌ها و فرآیندهای شیمیایی و از طرفی هزینه‌های بالای آنها، استفاده از روش‌های محاسباتی برای بهبود بخشیدن به فرآیندهای اندازه‌گیری و استخراج اطلاعات شیمیایی بیشتر از داده‌های موجود بسیار مفید می‌باشد. تجزیه و تحلیل آماری و ریاضی داده‌های شیمیایی کمومتریکس نامیده می‌شود [۲۰].

۲-۲- تاریخچه‌ی کمومتریکس^{۲۶}

اصطلاح کمومتریکس برای اولین بار توسط اسوانت ولد^{۲۷} دانشمند جوان سوئدی که در زمینه شیمی فیزیک آلی فعالیت داشت، مطرح گردید. همکاری ولد با بروس آر. کووالسکی^{۲۸} که در دانشگاه واشنگتن بر روی الگوشناسی در شیمی تجزیه مطالعه می‌کرد، منجر به تأسیس انجمن بین‌المللی کمومتریکس (ICS)^{۲۹} در سال ۱۹۷۴ گردید. بنا به تعریف ICS، کمومتریکس عبارت است از کاربرد

^{۲۶} - Chemometrics

^{۲۷} - Svante wold

^{۲۸} -Bruce R .Kowaski

^{۲۹} -International Chemometrics Society

روش‌های ریاضی، آماری برای برقراری ارتباط بین سنجش‌های انجام شده روی یک سیستم یا فرایند شیمیایی به منظور درک بهتر اطلاعات شیمیایی.

در حال حاضر با توجه به رشد سریع تجهیزات مورد استفاده در شیمی و حجم بسیار زیاد داده‌ها و اطلاعات به منظور جمع‌آوری، پردازش، تفسیر اطلاعات و استخراج نتایج مفید از آنها، نیاز به کامپیوتر امری اجتناب ناپذیر است. از طرف دیگر شیمیدانان، گاهی اوقات با موادی سر و کار دارند که بسیار گران، سمی و خطرناک بوده و یا در مواردی به راحتی قابل دسترس نیستند. در این موارد، می‌توان از روش‌های ریاضی و آمار به طور مؤثری جهت توصیف و توجیه نتایج آزمایش‌های مختلف استفاده نمود [۲۱]. به عبارت بهتر می‌توان گفت که موارد فوق از مهم‌ترین دلایل افزایش کاربرد روش‌های کمومتریکس نزد شیمیدانان است.

۲-۳- روش‌های پارامتری

مجموعه تکنیک‌هایی که برای مدل‌سازی و پیش‌بینی خاصیت‌های مولکولی مورد استفاده قرار می‌گیرند روش‌های پارامتری نام دارند. با استفاده از این روش‌ها می‌توان ارتباط منطقی بین فعالیت یا خاصیت مولکول‌ها با تعدادی از توصیف‌کننده‌های مولکولی برقرار نمود. این توصیف‌کننده‌ها، متغیرهای مستقلی هستند که ویژگی‌های مولکول را به طور کمی بیان می‌کنند. جهت ایجاد این روابط می‌توان هم از روش‌های خطی مثل رگرسیون خطی چندگانه^{۳۰} و هم از روش‌های غیر خطی مثل شبکه عصبی مصنوعی^{۳۱} و ماشین برداری پشتیبان^{۳۲} استفاده نمود.

۲-۳-۱- رگرسیون خطی چندگانه

^{۳۰} -Multiple linear regression

^{۳۱} -Artificial neural networks

^{۳۲} -Support vector machine

در رگرسیون خطی چندگانه، ارتباط بین توصیف‌کننده‌ها به عنوان متغیر مستقل و خاصیت به عنوان یک متغیر وابسته به صورت یک معادله‌ی خطی بیان می‌شود. برای ایجاد این ارتباط، روش‌های آماری کاربرد وسیعی داشته و نرم‌افزارهای آماری متفاوتی جهت تسریع و سهولت انجام عملیات محاسباتی مورد استفاده قرار می‌گیرند. پس از معرفی متغیرهای مستقل و وابسته، براساس روش‌های رگرسیون خطی چندگانه، معادلات خطی ارتباط دهنده‌ی بین این دو نوع متغیر ایجاد می‌شود [۲۲]. قابل ذکر است که متغیرهای مستقل براساس اصول علمی که انتظار می‌رود با متغیر وابسته در ارتباط باشند، انتخاب می‌شوند. مراحل کلی مدل‌سازی به روش پارامتری به ترتیب زیر می‌باشد:

۱- انتخاب سری داده‌ها

۲- انتخاب و محاسبه توصیف‌کننده‌ها

۳- تجزیه و تحلیل آماری توصیف‌کننده‌ها

۴- آنالیز مدل‌های آماری و انتخاب مدل مناسب

۵- ارزیابی اعتبار مدل‌های انتخاب شده

برای آشنایی بیشتر توضیحاتی در مورد هر یک از مراحل فوق ارائه می‌گردد.

۲-۳-۱-۱- انتخاب سری داده‌ها

در مدل‌سازی یافتن مجموعه‌ای از مولکول‌ها که مقادیر تجربی خاصیت مورد بررسی برای آنها در شرایط یکسانی به دست آمده باشد، در اولویت قرار دارد. سپس داده‌ها به سه قسمت سری آموزش، سری آزمون و سری ارزیابی تقسیم می‌شوند. سری آزمون و سری ارزیابی به نحوی انتخاب می‌شوند که نماینده‌ی جمعیت مولکول‌های سری آموزشی باشند. لازم به ذکر است که مولکول‌های سری ارزیابی در هیچ یک از مراحل مدل‌سازی شرکت ندارند. در واقع سری ارزیابی برای ارزیابی اعتبار

مدل‌های حاصله است. هر قدر که سری داده‌ها بزرگ‌تر و متنوع‌تر باشد مدل حاصل از آن معتبرتر بوده و قدرت پیش‌بینی بالاتری خواهد داشت.

۲-۳-۱-۲- محاسبه‌ی توصیف‌کننده‌ها

توصیف‌کننده‌ها پارامترهای عددی هستند که ویژگی‌های متفاوتی از یک مولکول را بیان می‌کنند. تعیین و انتخاب توصیف‌کننده‌ها از اهمیت بسزایی برخوردار است. در صورتی که توصیف‌کننده‌های انتخاب شده مناسب نباشند عمل مدل‌سازی با مشکل مواجه شده و مدل مناسب به دست نخواهد آمد. توصیف‌کننده‌ها از نظر بیان چگونگی خصوصیات مولکول به دسته ۱۸ متفاوت تقسیم می‌شوند. که در جدول (۱-۲) نام این توصیف‌کننده‌ها آورده شده است.

جدول (۱-۲): انواع توصیف‌کننده‌های محاسبه شده توسط نرم افزار Dragon

۱- توصیف‌کننده‌های زیر ساختاری ^{۳۳}	۲- توصیف‌کننده‌های توپولوژی ^{۳۴}
۳- شمارنده‌ی مولکولی مورس ^{۳۵}	۴- توصیف‌کننده‌های BCUT ^{۳۶}
۵- شاخص بار ^{۳۷}	۶- خود ارتباطی‌های دو بعدی ^{۳۸}
۷- توصیف‌کننده‌های بار ^{۳۹}	۸- شاخص‌های آروماتیسیت ^{۴۰}
۹- پروفایل‌های مولکولی راندیک ^{۴۱}	۱۰- توصیف‌کننده‌های هندسی ^{۴۲}
۱۱- توصیف‌کننده‌های RDF ^{۴۳}	۱۲- توصیف‌کننده‌های سه بعدی ^{۴۴}

33 - Constitutional Descriptors

34 - Topological Descriptors

35 - Molecular Walk Counts

36 - BCUT Descriptors

37 - Charge Indices

38 - 2D Autocorrelations

39 - Charge Descriptors

40 - Aromatic Indices

41 - Rancid Molecular Profiles

42 - Geometrical Descriptors

43 - Radial Distribution Function Descriptors

44 - 3D Molecule Representation of Structure Backed on Electron Diffraction Descriptors(3D- MORSE Descriptors)

۱۳- توصیف‌کننده‌های WHIM ^{۴۵}	۱۴- توصیف‌کننده‌های GETAWAY ^{۴۶}
۱۵- گروه‌های عاملی ^{۴۷}	۱۶- اجزای میان اتمی ^{۴۸}
۱۷- توصیف‌کننده‌های تجربی ^{۴۹}	۱۸- خصوصیات مولکولی ^{۵۰}

۳-۱-۳-۲- تجزیه، تحلیل و ارزیابی آماری توصیف‌کننده‌ها

تا این مرحله انتخاب توصیف‌کننده‌ها براساس احتمالات و تشخیص نظری استوار است. بنابراین احتمال اینکه برخی از توصیف‌کننده‌ها برای مدل‌سازی مناسب نباشند وجود دارد و باید آنها را حذف کرد. توصیف‌کننده‌هایی که حذف می‌شوند معمولاً یک یا چند ویژگی زیر را دارند:

۱- توصیف‌کننده‌هایی که کمتر از ۱۰٪ مقادیر غیر صفر دارند یا دارای بیش از ۹۰٪ مقادیر یکسان باشند.

۲- توصیف‌کننده‌هایی که با توصیف‌کننده دیگر همبستگی بالایی دارند ($r > 0.9$)

۳- توصیف‌کننده‌هایی که با متغیر وابسته، همبستگی کمی دارند و محاسبه‌ی آنها مشکل است.

پس از حذف توصیف‌کننده‌ها طبق یکی از موارد فوق، تعداد توصیف‌کننده‌ها کاهش یافته و از پیچیدگی محاسبات کاسته می‌شود.

۳-۱-۳-۲- انتخاب بهترین توصیف‌کننده

45 - Weighted Holistic Invariant Molecular Descriptors
46 - Geometry, Topology and Atom Weighted Assembly
47 - Functional Groups
48 - Atom- Centered Fragments
49 - Empirical Descriptors
50 - Properties

مرحله بعدی در مطالعات QSPR گزینش بهترین توصیف‌کننده‌های محاسبه شده، می‌باشد. یکی از بهترین روش‌های انتخاب توصیف‌کننده‌های مناسب، روش رگرسیون مرحله‌ای^{۵۱} برای ساخت مدل است. مدل، در واقع یک رابطه‌ی ریاضی است که بیان‌کننده‌ی رابطه‌ی بین متغیر وابسته و مستقل می‌باشد و به کمک آن می‌توان با داشتن مقادیر متغیرهای مستقل، متغیر وابسته را ارزیابی کرد. در روش رگرسیون مرحله‌ای ابتدا قوی‌ترین متغیر مستقل؛ یعنی متغیری که ضریب همبستگی آن با متغیر وابسته از بقیه متغیرها بیشتر باشد، شناسایی شده و یک مدل یک متغیره ایجاد می‌گردد. سپس با توجه به ضریب همبستگی، متغیر بعدی وارد مدل می‌شود و در هر مرحله برای متغیرهای وارد شده در مدل در مرحله‌ی قبل آماره‌ی F محاسبه می‌شود. سپس متغیر با کوچکترین آماره‌ی F حذف شده و مدل مناسب به دست می‌آید. لازم به ذکر است در این پایان‌نامه برای انتخاب متغیر از این روش استفاده شده است.

۲-۳-۱-۵- ارزیابی مدل به دست آمده

برای اطمینان از اینکه مدل به دست آمده، مدلی مناسب می‌باشد باید یک سری از پارامترها در نظر گرفته شود. به طور کلی می‌توان گفت مدلی مناسب‌تر است که دارای ویژگی‌های زیر باشد.

- تعداد متغیرهای مستقل آن از سایر مدل‌ها کمتر باشد.
- محاسبه متغیرهای مستقل آن ساده‌تر از سایر مدل‌ها باشد.
- حتی‌الامکان توصیف‌کننده‌هایی که مقدار آن‌ها باید به روش تجربی اندازه‌گیری شود در مدل وارد نشده باشند.
- ضریب همبستگی بین توصیف‌کننده‌هایی که در مدل شرکت دارد کم باشد.
- خطای استاندارد آن کم باشد.
- آماره‌ی F بزرگ بوده و ضریب همبستگی آن به یک نزدیک باشد.

^{۵۱} - Stepwise

- قدرت پیش‌بینی بالایی داشته باشد.

۲-۳-۱-۶- مدل سازی با استفاده از روش MLR

MLR روشی است که برای مدل‌سازی ارتباط خطی بین متغیر وابسته و چند متغیر مستقل به

کار می‌رود. معادله‌ی این نوع رگرسیون به صورت زیر است:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n + e \quad (1-2)$$

که X_i ها متغیرهای مستقل (توصیف‌کننده‌های عددی) و مقادیر b_0 تا b_n ضرایب متغیرهای مستقل‌اند و e نیز مقدار خطای رگرسیون است که باقی‌مانده نامیده می‌شود. Y نیز همان متغیر وابسته یا به عبارت دیگر خاصیت یا فعالیت مورد اندازه‌گیری است. مقدار ضریب متغیر مستقل در معادله، بیانی از میزان اهمیت آن متغیر در پیش‌بینی مقادیر متغیر وابسته است. این روش تنها زمانی می‌تواند به کار رود که تعداد مولکول‌های مورد بررسی بیش از تعداد توصیف‌کننده‌ها باشد و متغیرها با یکدیگر همبستگی زیادی نداشته باشند.

۲-۳-۲- شبکه‌های عصبی مصنوعی

۲-۳-۲-۱- مقدمه‌ای بر شبکه‌های عصبی

شبکه‌ی عصبی یک برنامه نرم‌افزاری است که می‌تواند همانند مغز انسان عمل کند. در واقع

شبکه عصبی مصنوعی ایده‌ای برای پردازش اطلاعات است که از سیستم عصبی زیستی الهام گرفته

شده و مانند مغز انسان به پردازش اطلاعات می‌پردازد.

این سیستم‌ها از تعداد زیادی عنصر پردازش به نام نرون^{۵۲} تشکیل شده‌اند که برای یک مسئله به صورت هماهنگ با هم عمل می‌کنند. شبکه‌های عصبی مصنوعی نظیر مغز انسان، با مثال یاد می‌گیرند و با پردازش روی داده‌های تجربی، دانش یا قانون نهفته در ورای داده‌ها را کشف می‌کنند. به همین خاطر به این سیستم‌ها، هوشمند گفته می‌شود زیرا شبکه‌ها براساس محاسبات روی داده‌های عددی یا مثال‌ها، قوانین کلی را فرا می‌گیرند.

۲-۳-۲- تاریخچه

اولین کوشش‌ها در شبیه‌سازی ساختار دینامیکی مغز در سال ۱۹۴۳ توسط مک کلوج^{۵۳} و پیتز^{۵۴} مطرح شد [۲۴] و نخستین کاربرد عملی شبکه‌های عصبی برای شناسایی الگوها در اواخر دهه‌ی ۵۰، زمانی که روزنبلات^{۵۵} در سال ۱۹۵۸ شبکه پرسپترون^{۵۶} را معرفی کرد، مطرح گردید [۲۵]. سال ۱۹۶۹ آغاز افول موقت شبکه‌های عصبی بود، زیرا عدم توانایی شبکه‌های عصبی در حل مسائل غیر خطی آشکار شد. در سال ۱۹۸۲ هاپفیلد^{۵۷} با معرفی شبکه‌های چند لایه و برگشتی، راه حلی برای حل مسائل غیر خطی ارائه کرد [۲۶] و سپس دیوید راملهارت^{۵۸} و جیمز مک‌کلند^{۵۹} الگوریتم پس انتشار خطا^{۶۰} را در سال ۱۹۸۶ مطرح کردند [۲۷]. امروزه با افزوده شدن قابلیت‌های بیشتر به شبکه‌های عصبی مصنوعی، این شبکه‌ها توانسته‌اند به عنوان ابزاری قدرتمند در مسائل بسیار متنوعی از قبیل تشخیص الگو، شناسایی گفتار، پردازش تصویر و حتی در کنترل یا مدل‌سازی سیستم‌هایی که ساختار داخلی ناشناخته یا پیچیده دارند، به کار روند.

52- Neuron

53- McCluch

54- Pitts

55- Rosenblat

56- Perceptron

57- Hopfield

58- David Rummelhart

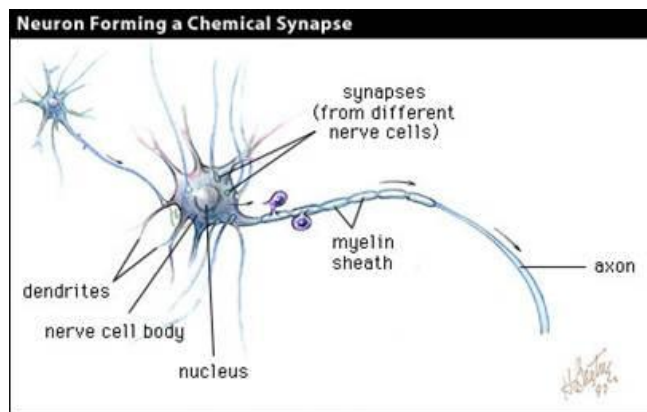
59- James Mcland

60- Back Propagation

۲-۳-۲-۳- نرون بیولوژیکی

هر نرون طبیعی دارای سه قسمت اصلی دندریت^{۶۱}، جسم سلولی^{۶۲} و آکسون^{۶۳} است. محل تلاقی آکسون هر سلول به دندریتهای سلول دیگر را سیناپس می‌گویند که ارتباط مابین نرون‌ها توسط سیناپس‌ها برقرار می‌شود. در حقیقت دندریتهای به عنوان ورودی نرون و آکسون به عنوان خروجی و فضای سیناپسی محل اتصال این دو می‌باشد. زمانی که سیگنال عصبی از آکسون به نرون‌ها و یا عناصر دیگر بدن مثل ماهیچه‌ها می‌رسد، باعث تحریک آنها می‌شود. نرون‌ها از هر یک از اتصالات ورودی خود یک ولتاژ کم دریافت می‌کنند و آنها را با هم جمع می‌زنند. اگر این حاصل جمع به یک مقدار آستانه برسد، نرون روی آکسون خود یک ولتاژ خروجی ارسال می‌کند. این ولتاژ به دندریتهایی که به این آکسون متصلند رسیده و باعث یکسری فعل و انفعال‌های شیمیایی در اتصالات سیناپسی می‌شود و می‌تواند باعث ارسال پیک‌های فعالیت الکتریکی بین نرون‌های دیگر گردد [۲۸].

شکل (۱-۲) نمای کلی یک نرون طبیعی را نشان می‌دهد.



شکل (۱-۲) - ساختار نرون طبیعی

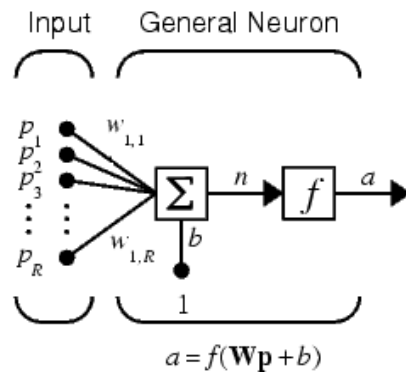
⁶¹ -Dendrite

⁶² -Cell body, Soma

⁶³ -Axon

۲-۳-۲-۴- نرون محاسباتی

مهمترین جزء تشکیل دهنده‌ی هر شبکه‌ی عصبی، نرون می‌باشد. نرون که یک واحد محاسباتی- پردازشی است در حقیقت مدل ساده شده‌ای از نرون طبیعی است. نرون از تعدادی ورودی، یک خروجی، وزن‌های مربوط به سیگنال تحریک و تابع تبدیل^{۶۴} یا تابع تحریک تشکیل شده است. ارتباط بین خروجی و سیگنال تحریک در قسمتی به نام تابع تحریک به وجود می‌آید. بنابراین یک نرون از دو بخش تشکیل شده است. بخش اول وظیفه‌ی دریافت ورودی‌ها، وزن‌دهی آنها و محاسبه‌ی مجموع آنها به عنوان سیگنال تحریک را به عهده دارد و بخش دوم وظیفه‌ی اعمال تابع تحریک روی سیگنال تحریک و ساختن پیغام خروجی نرون به لایه‌ی بعد و یا خروجی شبکه را به عهده دارد. شکل (۲-۲) ساختمان یک نرون محاسباتی را نشان می‌دهد. بیان رابطه‌ی ورودی‌ها و خروجی‌های شبکه در یک مدل ریاضی مطابق فرمولاسیون شکل (۲-۲)، پرسپترون نامیده می‌شود.



شکل (۲-۲)- ساختمان یک نرون محاسباتی

⁶⁴ -Transfer function

اگر به شکل دقت کنید، علاوه بر تمام ورودی‌های وزن‌دار، ورودی دیگری نیز در نظر گرفته شده و آن عدد یک است. این عدد یک، در پیش‌قدر^{۶۵} ضرب شده و با حاصلضرب ورودی‌ها در وزن‌ها جمع می‌شود تا ورودی خالص را ایجاد کند.

$$n = \mathbf{W} \cdot \mathbf{p} + b = W_{1,1}p_1 + W_{1,2}p_2 + \dots + W_{1,R}p_R + b \quad (2-2)$$

سپس تابع محرک روی این ورودی خالص n اثر می‌کند و خروجی نرون را ایجاد می‌کند.

$$a = f(n) = f(\mathbf{W} \cdot \mathbf{p} + b) \quad (3-2)$$

این پیش‌قدر شبیه یک وزن است با این تفاوت که روی ورودی ثابت یک عمل می‌کند و مزیت آن افزودن یک متغیر اضافی به شبکه است. از این رو، شبکه‌ای که پیش‌قدر دارد قوی‌تر از شبکه‌ی فاقد پیش‌قدر می‌باشد. به علاوه، یک نرون بدون پیش‌قدر همیشه در اثر ورودی صفر، خروجی را صفر می‌دهد که این حالت مطلوب نیست و با اعمال پیش‌قدر برطرف می‌شود [۲۹].

۲-۳-۲-۵- ساختار لایه‌ای شبکه عصبی مصنوعی

در این نوع تقسیم‌بندی، شبکه‌های عصبی مصنوعی به دو صورت تک لایه و چند لایه تفکیک می‌شوند. در شبکه‌های تک لایه، نرون‌های ورودی و خروجی یکی هستند بنابراین به تعداد خروجی، نرون در شبکه وجود دارد که همگی در یک لایه و به موازات هم کار می‌کنند. گاهی دینامیک و قوانین شبکه‌های عصبی آنقدر پیچیده است که به پارامترهای قابل تنظیم بیشتر نیاز است، در این موارد از شبکه‌های چند لایه استفاده می‌شود. شبکه‌های چند لایه شامل لایه‌ی ورودی^{۶۶}، لایه‌ی خروجی^{۶۷} و یک یا چند لایه‌ی پنهان^{۶۸} یا لایه‌ی میانی می‌باشند. در این توپولوژی، نرون‌ها علاوه بر این که در لایه‌ی خود به شکل محدودی به یکدیگر اتصال داده شده‌اند، از

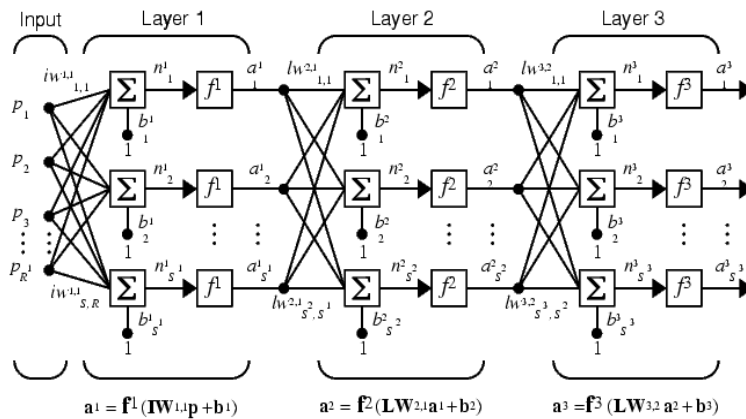
⁶⁵ -Bias

⁶⁶ -Input layer

⁶⁷ -Output layer

⁶⁸ -Hidden layer

طریق اتصال بین لایه‌ها نیز با یکدیگر ارتباط دارند. در این حالت گروهی از نرون‌ها از طریق ورودی‌های خود با جهان واقعی ارتباط دارند، گروهی دیگر از نرون‌ها نیز از طریق خروجی‌های خود، جهان خارج را می‌سازند که لایه‌ی خروجی نام دارند و بقیه نرون‌ها که از دید پنهان هستند لایه‌های پنهان را تشکیل می‌دهند [۳۰]. در یک شبکه‌ی چند لایه، هر لایه، ماتریس وزن، بردار پیش‌قدر، بردار ورودی خالص و بردار خروجی مخصوص به خود را دارد، که برای مشخص کردن لایه‌های مختلف از بالانویس^{۶۹} استفاده می‌شود. در شکل (۲-۳) ساختار یک شبکه‌ی سه لایه نشان داده شده است [۲۹].



شکل (۲-۳): ساختار یک شبکه‌ی سه لایه

۲-۳-۶- 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^{۷۰}

۲- شبکه‌های پس‌خور^{۷۱}

^{۶۹} -Superscript

^{۷۰} -Feed forward

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 ارتباط 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 [۳۰]. بنابراین این نوع از شبکه‌ها می‌توانند با استفاده از حلقه‌های برگشتی، سیگنال‌هایی داشته باشند که در هر دو مسیر ورودی به خروجی و بالعکس حرکت کنند.

۲-۳-۲-۷- 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 شبکه عصبی می‌نامند. به طور کلی برای آموزش شبکه دو روش وجود دارد:

۱- یادگیری با ناظر^{۷۲}

⁷¹ -Feed Back

⁷² -Supervised learning

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 خود سامان‌ده 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.

رفتار سیستم‌های یادگیر توسط الگوریتم‌های بازگشتی بیان می‌شود که به آنها قوانین یادگیری می‌گویند. انواع مختلفی از قوانین یادگیری وجود دارد که یادگیری عملکردی^{۷۴} 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^{۷۵} 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^{۷۶} (MSE) می‌باشد [۲۹].

۲-۳-۲-۸- معرفی روش‌های مختلف بهینه‌سازی عملکرد شبکه عصبی

⁷³ -Unsupervised learning

⁷⁴ -Performance learning

⁷⁵ -Performance

⁷⁶ -Mean square error

در تمام الگوریتم‌های بهینه‌سازی یک مقدار اولیه برای تابع حدس زده می‌شود و بعد این مقدار

اولیه طی مراحل طی که از قاعده‌ی زیر پیروی می‌کند بهینه می‌شود [۳۱]:

$$X_{K+1} = X_K + \alpha P_K \quad (4-2)$$

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 مرحله‌ی بعد بدست می‌آید. α یک مقدار اسکالر بین صفر و یک بوده و سرعت یادگیری^{۷۷} نامیده می‌شود. سرعت یادگیری، طول قدم در هر تکرار را در هر مسیر P_K تعیین می‌کند. انتخاب α و P_K روش‌های مختلف مینیمم‌سازی را از هم متمایز می‌کند که بعضی از آنها عبارتند از:

۱- روش نزول گرادینتی^{۷۸}

۲- الگوریتم پس انتشار خطا^{۷۹}

۳- روش لوبنبرگ-مارکوارت^{۸۰}

۴- الگوریتم تنظیم بایزین^{۸۱}

۲-۳-۳- ماشین بردار پشتیبان

۲-۳-۳-۱- مقدمه

ماشین‌های بردار پشتیبان^{۸۲} (SVM) یکی از روش‌های یادگیری با ناظر هستند که از

الگوریتم‌های آموزش دیده توسط آنها می‌توان هم در طبقه‌بندی و هم در آنالیز رگرسیون استفاده

⁷⁷ -Learning rate

⁷⁸ -Gradient descent algorithm

⁷⁹ -Error back propagation algorithm

⁸⁰ -Levenberg-Marquardt

⁸¹ -Bayesian regularization

⁸² - Support vector machine

نمود. این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر مثل شبکه‌های عصبی برای طبقه‌بندی، نشان داده است.

ماشین‌های برداری پشتیبان اولیه فقط در طبقه‌بندی دو کلاسه به کار می‌رفتند و به آن‌ها جداکننده‌های خطی دوتایی^{۸۳} می‌گفتند اما بعداً از این روش برای طبقه‌بندی چند کلاسه نیز استفاده شد. هدف از این گونه بررسی‌ها ایجاد یک معیار طبقه‌بندی است به گونه‌ای که بتواند برای نمونه‌های جدید به خوبی قابل استفاده بوده و از قدرت تعمیم خوبی برخوردار باشد.

اگر چه SVMها در ابتدا برپایه‌ی مسائل طبقه‌بندی بنا نهاده شده بودند اما امروزه توانسته‌اند با استفاده از تابع کرنل در طبقه‌بندی غیرخطی و آنالیز رگرسیون نیز کاربرد وسیعی پیدا کنند. ماشین‌های بردار پشتیبان در مقایسه با شبکه‌های عصبی دارای دو مزیت عمده هستند که همین امر موجب توجه خاص محققین به این ماشین‌های فراگیر شده است:

- ۱- قابلیت مدل‌سازی توسط SVM با تعداد داده‌های کم نسبت به شبکه‌های عصبی بیشتر است.
- ۲- تئوری قوی ماشین بردار پشتیبان قابلیت تعمیم بالایی به آن داده و از قرار گرفتن آن در بهینه‌ی محلی جلوگیری می‌کنند. در آموزش شبکه‌های عصبی از حداقل‌سازی خطای تجربی برای یافتن بهترین مدل استفاده می‌کنند؛ یعنی، شبکه خطای داده‌های آموزشی را حداقل می‌کند اما در SVM از حداقل‌سازی خطای ساختاری برای یافتن بهترین مدل استفاده می‌شود که در آن خطای تعمیم حداقل می‌شود.

۲-۳-۳-۲- طبقه‌بندی خطی با ماشین‌های برداری پشتیبان

۲-۳-۳-۲- الف- آشنایی با مفاهیم اولیه دسته‌بندی کننده‌ی خطی SVM^{۸۴}

^{۸۳} -Binary linear classifier

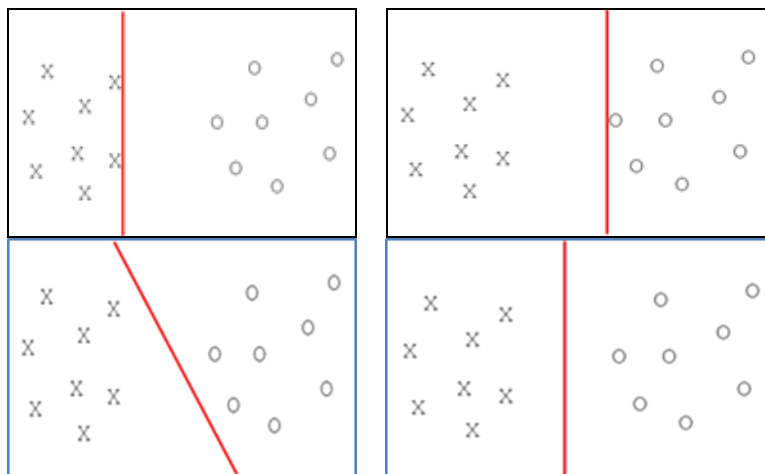
^{۸۴} -Linear classifier

فرض کنید دو دسته داده‌ی آموزشی مطابق شکل (۴-۲) وجود داشته باشند که به صورت خطی از هم جداپذیر باشند.



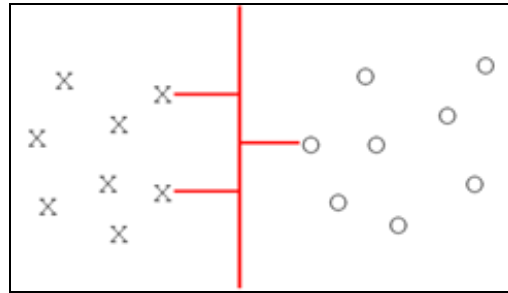
شکل (۴-۲): نمایش دو دسته داده‌ی آموزشی

هدف یافتن بهترین خطی است که بتواند دو دسته را از هم جدا کند. در شکل (۴-۲) انواع خط‌های جداکننده نمایش داده شده است.



شکل (۵-۲): خط‌های جداسازی که داده‌ها را جداسازی می‌کند.

همانطور که در شکل (۵-۲) نشان داده شده است، برای این داده‌ها تعداد زیادی خط جداکننده وجود دارد که قادر هستند داده‌ها را به صورت خطی جدا کنند اما همان‌طور که در شکل (۶-۲) نشان داده شده است، تنها یکی از آن‌ها دارای ماکزیمم فاصله بین خط جداکننده و نزدیکترین نقاط آموزشی در هر طرف خط هستند که به آن خط جداساز بهینه می‌گویند.



شکل (۶-۲): خط جداساز بهینه با حداکثر مقدار حاشیه

در حالت دو بعدی معادله‌ی این خط به صورت زیر نوشته می‌شود:

$$W_1 X_1 + W_2 X_2 + b = 0 \quad (۵-۲)$$

اکنون موضوع دسته بندی خطی داده‌ها را به فضای n بعدی گسترش می‌دهیم. در این فضا هدف، جداسازی نقاط دو دسته توسط یک ابرصفحه‌ی جداساز^{۸۵} $(n-1)$ بعدی است. در این حالت معادله‌ی ابرصفحه‌ی جداکننده با رابطه‌ی زیر بیان می‌شود:

$$\sum_{i=1}^N \mathbf{w}_i \mathbf{X}_i + b = 0 \quad (۶-۲)$$

یا در فرم کلی

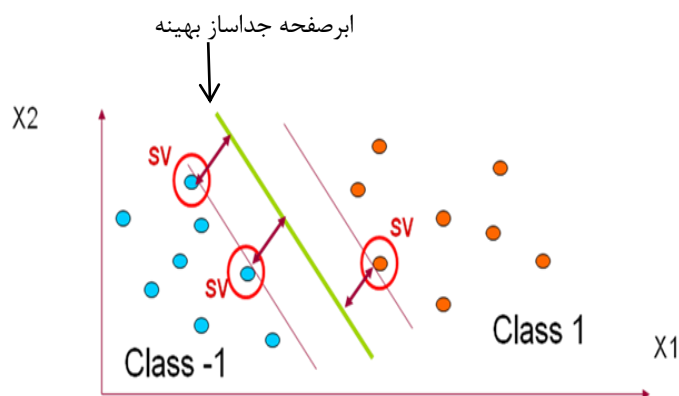
$$\mathbf{W}^T \cdot \mathbf{X} + b = 0 \quad (۷-۲)$$

که در آن علامت " " بین بردارهای \mathbf{X} و \mathbf{W} به معنی ضرب داخلی این دو بردار بوده و علامت T بالای بردار \mathbf{W} ، ترانهاده‌ی این بردار می‌باشد [۳۲].

در این حالت ابرصفحه‌ی جداساز بهینه، ابرصفحه‌ای است که حداکثر حاشیه ممکن را از نزدیکترین نقاط آموزشی در هر طرف ابرصفحه داشته باشد. برای به دست آوردن معادله‌ی این ابرصفحه سعی

⁸⁵ -Separating hyperplane

می‌شود تا با روش‌های بهینه‌سازی شناخته شده در ریاضیات فاصله‌ی بین دو ابرصفحه‌ی مرزی ماکزیمم شود. به نزدیک‌ترین داده‌های آموزشی به ابرصفحه‌ی جداساز بهینه که در واقع روی ابرصفحه‌های موازی مرزی قرار گرفته‌اند، بردارهای پشتیبان^{۸۶} می‌گویند. این داده‌ها که در شکل (۲-۷) با SV نشان داده شده‌اند، معیاری هستند که ماشین برداری پشتیبان از آن‌ها برای طبقه‌بندی صحیح داده‌ها استفاده می‌کند [۳۳]. وقتی ماشین برداری مثال‌های آموزشی دو کلاس را به وسیله‌ی یک شکاف واضح و تا حد امکان بزرگ به دو دسته تقسیم کرد، مثال‌های جدید را نیز در همان فضا ترسیم می‌کند. سپس براساس این‌که این داده‌ها در کدام طرف شکاف قرار گرفته‌اند پیش‌بینی می‌کند که به کدامیک از دو دسته تعلق می‌گیرند.



شکل (۲-۷): نمایش بردارهای پشتیبان روی ابرصفحه‌های موازی مرزی

۲-۳-۳-۲-ب-ریاضیات مسئله دسته‌بندی کننده‌ی خطی SVM

همان‌طور که در قسمت قبل توضیح داده شد مبنای کار در دسته‌بندی کننده‌ی خطی SVM، انتخاب ابرصفحه‌ای با بیشترین حاشیه‌ی اطمینان می‌باشد. برای درک ریاضیات موضوع از یک

⁸⁶ -Support vectors (SV)

طبقه‌بندی دو کلاسه استفاده می‌کنیم و مبنای ریاضی دسته‌بندی کننده‌ی SVM را در مورد این مسئله شرح می‌دهیم.

فرض کنید مجموعه‌ای از داده‌های آموزشی در اختیار داریم که به صورت زیر تعریف می‌شوند:

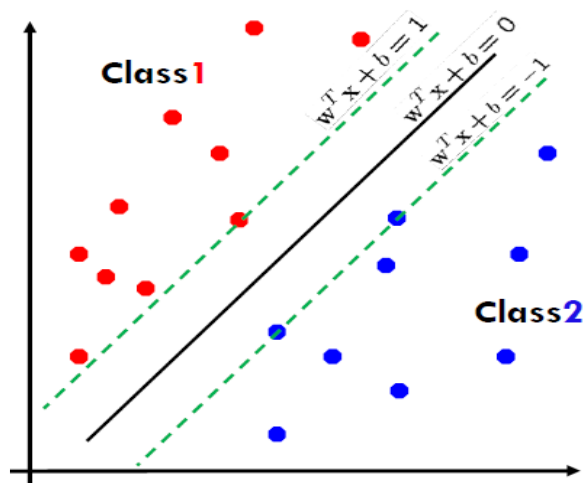
$$D = \{(X_i, y_i) | x_i \in R^n, y_i \in [-1, 1]\}_{i=1}^n \quad (8-2)$$

که در آن X_i یک بردار حقیقی n بعدی است و y_i برابر ۱ و -۱ است و نشان می‌دهد که هر یک از نقاط x_i به کدام طبقه تعلق دارند. هدف یافتن ابرصفحه‌ی جداساز با بیشترین فاصله از نقاط حاشیه‌ای است به گونه‌ای که بتواند نقاط با $y_i = 1$ را از نقاط با $y_i = -1$ جدا کند. همان‌طور که در قسمت قبل بیان شد معادله‌ی ابرصفحه‌ی جداساز در فضای n بعدی به صورت $W^T \cdot X + b = 0$ می‌باشد که باید W و b به گونه‌ای پیدا شوند که اولاً، نمونه‌های آموزشی بدون اشتباه در کلاس خود دسته‌بندی شوند ثانیاً، فاصله‌ی بین نزدیک‌ترین نقاط هر کلاس داده تا ابرصفحه‌ی جداساز، ماکزیمم باشد؛ یعنی، حداکثر حاشیه‌ی ممکن بین ابرصفحه‌های مرزی موازی که داده‌ها را از هم جدا می‌کنند ایجاد شود. برای توصیف این ابرصفحه‌های مرزی موازی از روابط (۹-۲) و (۱۰-۲) استفاده می‌شود:

$$W^T \cdot X + b = -1 \quad (9-2)$$

$$W^T \cdot X + b = +1 \quad (10-2)$$

در شکل (۸-۲) معادلات در نظر گرفته شده برای ابرصفحه‌های مرزی و صفحه‌ی جداساز بهینه نشان داده شده است.



شکل (۸-۲): صفحه‌ی جداساز و حاشیه‌ها

فاصله بین دو ابرصفحه‌ی مرزی را می‌توان به کمک هندسه به صورت زیر به دست آورد:

$$d = \frac{|(W^T \cdot X + b - 1) - (W^T \cdot X + b + 1)|}{\|W\|} = \frac{2}{\|W\|} \quad (۱۱-۲)$$

براساس این رابطه اگر تابع $\frac{2}{\|W\|}$ را با درنظر گرفتن قیدها ماکزیم کنیم حاشیه‌ی مورد نظر

ماکزیم خواهد شد. برای این که از ورود نقاط به حاشیه جلوگیری شود، شرایط زیر نیز اضافه می‌گردد:

برای X_i های متعلق به کلاس ۱ با $y_i=1$:

$$W^T \cdot X + b \geq 1 \quad (۱۲-۲)$$

برای X_i های متعلق به کلاس ۲ با $y_i=-1$:

$$W^T \cdot X + b \leq -1 \quad (۱۳-۲)$$

این دو شرط می‌تواند به صورت یک شرط واحد به شکل زیر بازنویسی شود:

$$y_i(\mathbf{W}^T \cdot \mathbf{X} + b) \geq 1 \quad 1 \leq i \leq n \quad (14-2)$$

بنابراین باید تابع $\frac{2}{\|\mathbf{W}\|}$ را با در نظر گرفتن قید فوق ماکزیمم کنیم تا مقادیر بهینه b و W به دست آیند. اما حل این مسئله بهینه‌سازی به دلیل وابستگی به $\|\mathbf{W}\|$ سخت است، لذا برای سادگی کار می‌توان بدون تغییر در مسئله به جای تابع $\frac{2}{\|\mathbf{W}\|}$ ، تابع $\frac{1}{2}\|\mathbf{W}\|$ را با در نظر گرفتن قید (2-18) مینیمم کرد که این تابع به صورت $\frac{1}{2}\mathbf{W}^T \cdot \mathbf{W}$ نوشته می‌شود.

راه کار پیشنهادی، استفاده از ضرایب نامعین لاگرانژ است. تابع لاگرانژ برای این مسئله به صورت زیر نوشته می‌شود :

$$L_p(\mathbf{W}, b, \alpha) = \frac{1}{2}\mathbf{W}^T \cdot \mathbf{W} - \sum_{i=1}^n \alpha_i [y_i(\mathbf{W}^T \cdot \mathbf{X}_i + b) - 1] \quad (15-2)$$

که در آن α_i ها ضرایب نامعین لاگرانژ هستند. اکنون هدف مینیمم‌سازی تابع لاگرانژ با در نظر گرفتن $\alpha_i > 0$ است. اگر از رابطه‌ی (2-15) نسبت به b و \mathbf{W} مشتق بگیریم و مساوی صفر قرار دهیم مقدار بهینه‌ی W به دست می‌آید.

$$\frac{\partial L}{\partial \mathbf{W}} = 0 \Rightarrow \mathbf{W} = \sum_{i=1}^N \alpha_i \mathbf{X}_i y_i \quad (16-2)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (17-2)$$

رابطه‌ی (2-16) نشان می‌دهد که پارامتر W به صورت یک ترکیب خطی از نقاط آموزشی بیان شده است.

حال اگر مقدار W به دست آمده از مشتقات جزئی رابطه‌ی (2-16) را در خود رابطه‌ی (2-15) قرار دهیم معادله‌ی اساسی ماشین‌های برداری به صورت زیر به دست خواهد آمد:

$$L_d(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j \mathbf{X}_i^T \mathbf{X}_j \quad (18-2)$$

بنابراین هدف در ماشین‌های برداری حل معادله‌ی (۱۸-۲) با توجه به دو محدودیت (۱۹-۲) و (۲۰-۲) می‌باشد:

$$\alpha_i \geq 0 \quad (19-2)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (20-2)$$

مقدار بهینه‌ی b نیز به کمک قید زیر به دست می‌آید:

$$\mathbf{W}^T \cdot \mathbf{X}_i + b = \frac{1}{y_i} = y_i \Rightarrow b = y_i - \mathbf{W}^T \cdot \mathbf{X}_i \quad (21-2)$$

البته بهتر است که با متوسط‌گیری روی همه‌ی بردارهای پشتیبان، مقدار بهینه‌ی b را محاسبه کرد تا در عمل الگوریتم مقاوم‌تری به دست آید:

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (y_i - \mathbf{W}^T \cdot \mathbf{X}_i) \quad (22-2)$$

با حل مسئله‌ی بهینه‌سازی (۱۸-۲) و استفاده از رابطه‌ی (۲۲-۲) می‌توان به بهینه‌ترین ابرصفحه‌ی جداساز دست یافت و سپس از این ابرصفحه جداساز برای طبقه‌بندی نمونه‌های جدید استفاده نمود [۳۲-۳۵].

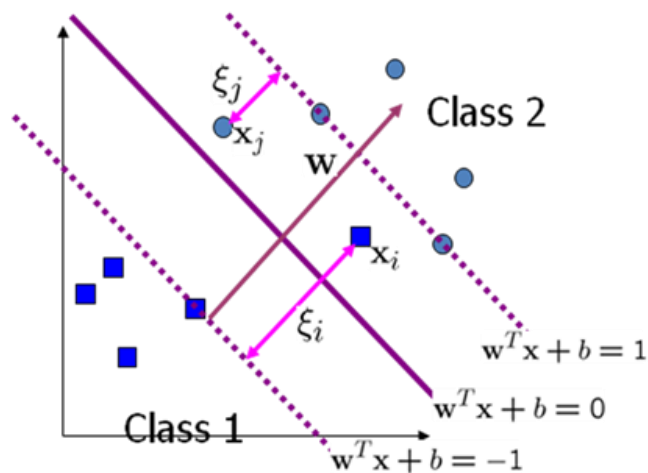
۲-۳-۲-ج-حاشیه‌ی نرم^{۸۷}

⁸⁷-Soft margin

گاهی اوقات در سیستم‌های خطی همان‌طور که در شکل (۹-۲) نشان داده شده است، داده‌هایی حضور دارند که در کلاس اشتباه طبقه‌بندی شده‌اند. برای چنین شرایطی ایده‌ی حاشیه‌ی نرم توسط واپنیک^{۸۸} و کورتینا^{۸۹} در سال ۱۹۹۵ مطرح شد. در این روش یک متغیر ξ_i معرفی می‌شود که میزان خطای طبقه‌بندی اشتباه برای هر داده‌ی X_i را نشان می‌دهد و قید (۲-۱۴) را به صورت زیر در می‌آورد:

$$y_i(\mathbf{W}^T \cdot \mathbf{X}_i + b) \geq 1 - \xi_i \quad (۲۳-۲)$$

همچنین در این شرایط برای به دست آوردن ابرصفحه‌ی جداساز بهینه، یک تابع جریمه به صورت زیر تعریف می‌شود:



شکل (۹-۲): سیستم‌های خطی جدا ناپذیر با میزان خطای ξ_i

$$F(\xi) = \sum_{i=1}^N \xi_i \quad (۲۴-۲)$$

⁸⁸ -Vapnik

⁸⁹ -Cortinna

و سپس مسئله‌ی بهینه‌سازی برای مینیمم‌سازی خطای طبقه‌بندی به کار می‌رود. در این حالت باید تابع (۲۵-۲) با توجه به دو قید (۲۶-۲) و (۲۷-۲) بهینه شود:

$$Min_{W,b} = \frac{1}{2} \mathbf{W}^T \cdot \mathbf{W} + C \sum_{i=1}^N \xi_i \quad (25-2)$$

$$y_i (\mathbf{W}^T \cdot \mathbf{X}_i + b) \geq 1 - \xi_i \quad (26-2)$$

$$F(\xi) = \sum_{i=1}^N \xi_i \quad (27-2)$$

که در رابطه‌ی فوق C ضریب تنظیم جهت ماکزیمم کردن حاشیه‌ها و مینیمم کردن خطاست. با در نظر گرفتن ضرایب لاگرانژ α و β که بزرگتر از صفر هستند می‌توان تابع لاگرانژ را چنین نوشت:

$$L_p(\mathbf{W}, b, \xi, \alpha, \beta) = \frac{1}{2} \mathbf{W}^T \cdot \mathbf{W} + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i \{y_i (\mathbf{W}^T \cdot \mathbf{X}_i + b) - 1 + \xi_i\} - \sum_{i=1}^N \beta_i \xi_i \quad (28-2)$$

اگر از رابطه‌ی (۲۸-۲) نسبت به \mathbf{W} ، b و ξ_i مشتق گرفته و مساوی صفر قرار داده شود، مقادیر زیر به دست می‌آیند:

$$\frac{\partial L}{\partial \mathbf{W}} = 0 \Rightarrow \mathbf{W} = \sum_{i=1}^N \alpha_i y_i \mathbf{X}_i \quad (29-2)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (30-2)$$

$$\frac{\partial L}{\partial \xi} = 0 \Rightarrow \alpha_i + \beta_i = C \quad (31-2)$$

با قرار دادن این روابط در رابطه (۳۲-۲)، معادله‌ی اساسی ماشین‌های برداری در حالت خطی جداناپذیر به دست می‌آید که به صورت معادله‌ی (۲۸-۲) خواهد بود.

$$L_d(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N y_i y_j \alpha_i \alpha_j \mathbf{X}_i^T \mathbf{X}_j \quad (32-2)$$

در این حالت هدف حل معادله‌ی (۳۶-۲) با توجه به دو محدودیت (۳۳-۲) و (۳۴-۲) می‌باشد :

$$0 \leq \alpha_i \leq C \quad (33-2)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (34-2)$$

همان‌گونه که مشاهده می‌شود، تابع هدف سیستم‌های جداناپذیر خطی مشابه با سیستم‌های جداناپذیر خطی است با این تفاوت که ضرایب لاگرانژ α_i دارای کران بالای C هستند. در این حالت پارامتر C نیز که قابلیت کنترل ظرفیت اضافی در طبقه‌بندی کننده را فراهم می‌آورد، باید تعیین شود [۳۲ و ۳۴].

۳-۳-۳-۲ - طبقه بندی غیرخطی^{۹۰} با ماشین‌های برداری پشتیبان

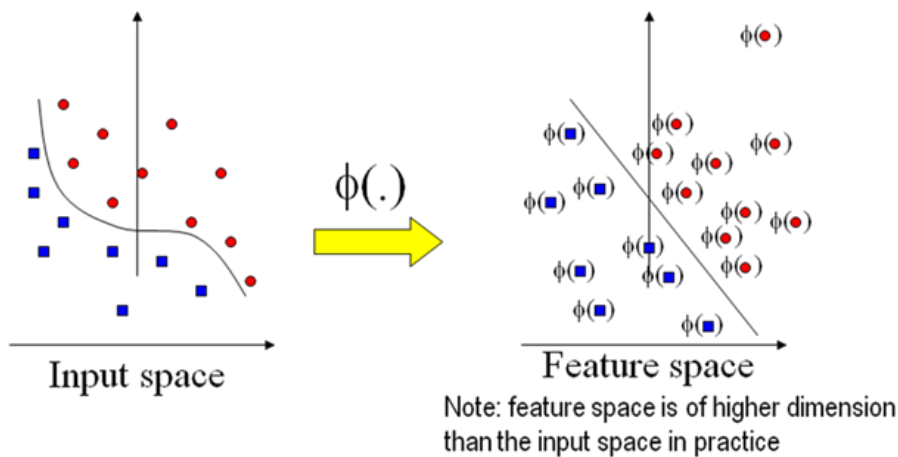
۳-۳-۳-۲-الف - فضای ویژگی

در قسمت‌های قبلی چگونگی طبقه‌بندی خطی با استفاده از ماشین برداری پشتیبان توضیح داده شد. اما مشکل اینجاست که در بسیاری مسائل طبقه‌بندی راه‌حل خطی جواب مناسبی نمی‌دهد و به راه‌حل غیرخطی نیاز است، در نتیجه نمی‌توان از طبقه‌بندی کننده‌ی خطی SVM استفاده کرد. خوشبختانه به کمک تئوری مرسر^{۹۱} که در اوایل سال ۱۹۰۰ ارائه گردیده بود ماشین برداری پشتیبان توانست به سیستم‌های غیرخطی گسترش یابد. برای استفاده از این تئوری، ابتدا هر بردار \mathbf{X}_i موجود

^{۹۰} -Nonlinear classification

^{۹۱} -Mercer

در فضای ورودی^{۹۲} به فضایی با ابعاد بالاتر که فضای ویژگی^{۹۳} نام دارد نگاشت می‌شود. در این فضای جدید، نگاشت بردار \mathbf{X}_i به صورت $\varphi(\mathbf{X}_i)$ نمایش داده می‌شود. سپس در این فضای بالاتر ماشین بردار پشتیبان می‌تواند به جداسازی داده‌ها بپردازد در حالی که فضای ورودی در همان حالت غیرخطی باقی مانده است. شماتیک ساده‌ای از این فرایند در شکل (۲-۱۰) نشان داده شده است [۳۳].



شکل (۲-۱۰): داده‌های ورودی ارجاع داده شده به فضای بالاتر

۲-۳-۳-۳-ب- ماشین‌های برداری غیرخطی

همان‌طور که در قسمت (۲-۲-۳-۲) بیان گردید برای بهینه‌سازی وزن‌های ارئه شده برای داده‌ها باید از رابطه‌ی (۲-۲۰) استفاده گردد تا داده‌ها به درستی در گروه مربوط به خود طبقه‌بندی شوند. اما در سیستم‌های غیرخطی با بردن داده‌ها به فضای بالاتر، همان‌طور که در شکل (۲-۹) نمایش داده شده است برداری مانند \mathbf{X}_i به صورت $\varphi(\mathbf{X}_i)$ نوشته می‌شود.

بنابراین، برای به دست آوردن وزن مناسب باید از رابطه‌ی (۲-۳۵) استفاده نمود:

$$\mathbf{W} = \sum_{i=1}^N \alpha_i y_i \varphi(\mathbf{X}_i) \quad (۲-۳۵)$$

^{۹۲} -Input space

^{۹۳} -Feature space

همچنین رابطه‌ی قید $y_i = \mathbf{W}^T \cdot \mathbf{X}_i + b$ را باید در فضای بالاتر به صورت زیر بازنویسی کرد:

$$y_i = \mathbf{W}^T \cdot \varphi(\mathbf{X}_i) + b \quad (۳۶-۲)$$

که از جایگذاری رابطه‌ی (۳۵-۲) در (۳۶-۲)، به دست می‌آید:

$$y_i = \sum_{i=1}^N \sum_{j=1}^N y_i \alpha_i \varphi(\mathbf{X}_i)^T \varphi(\mathbf{X}_j) + b \quad (۳۷-۲)$$

تئوری مرسر ضرب داخلی $\varphi(\mathbf{X}_i) \cdot \varphi(\mathbf{X}_j)$ را به صورت تابع کرنل $K(\mathbf{X}_i, \mathbf{X}_j)$ در فضای ویژگی معرفی می‌کند که به آن کرنل مرسر می‌گویند [۳۲ و ۳۴]. در نتیجه رابطه‌ی (۳۷-۲) به صورت زیر نوشته می‌شود:

$$y_i = \sum_{i=1}^N \sum_{j=1}^N y_i \alpha_i K(\mathbf{X}_i, \mathbf{X}_j) + b \quad (۳۸-۲)$$

کرنل‌های مختلفی در ریاضیات برای استفاده در فضای ویژگی معرفی شده‌اند که بسته به شرایط، مورد استفاده قرار می‌گیرند. بعضی از این توابع کرنل که در واقع ارتباط دهنده‌ی بین پارامترهای مدل و هدف هستند در جدول (۲-۲) گردآوری شده‌اند [۳۵].

جدول (۲-۲): توابع کرنل در فضای ویژگی

تابع کرنل	نوع طبقه‌بندی
$K(\mathbf{X}_i, \mathbf{X}_j) = (\mathbf{X}_i^T \mathbf{X}_j)^\rho$	خطی
$K(\mathbf{X}_i, \mathbf{X}_j) = (\mathbf{X}_i^T \mathbf{X}_j + 1)^\rho$	چند جمله‌ای از درجه ρ
$K(\mathbf{X}_i, \mathbf{X}_j) = e^{-\frac{\ \mathbf{X}_i - \mathbf{X}_j\ ^2}{2\sigma^2}}$	گوسین یا نمایی

$K(\mathbf{X}_i, \mathbf{X}_j) = \tanh(\gamma \mathbf{X}_i^T \mathbf{X}_j + \mu)$	پرسپترون چند لایه
$K(\mathbf{X}_i, \mathbf{X}_j) = \frac{\sin((n+1/2)(\mathbf{X}_i - \mathbf{X}_j))}{2 \sin((\mathbf{X}_i - \mathbf{X}_j)/2)}$	دریکله برای مسائل شرایط مرزی

برای محاسبه b نیز می‌توان از رابطه‌ی موجود در سیستم‌های خطی؛ یعنی، $b = y_i - W^T \cdot X_i$ استفاده نمود. این رابطه در حالت غیرخطی چنین نوشته می‌شود:

$$b = y_i - \sum_{i=1}^N \sum_{j=1}^N \alpha_i y_i \phi(\mathbf{X}_i)^T \cdot \phi(\mathbf{X}_j) \quad (39-2)$$

$$b = y_i - \sum_{i=1}^N \sum_{j=1}^N \alpha_i y_i K(\mathbf{X}_i, \mathbf{X}_j) \quad (40-2)$$

که در آن $K(\mathbf{X}_i, \mathbf{X}_j)$ همان تابع کرنل می‌باشد و به جای ضرب نقطه‌ای $\phi(\mathbf{X}_i) \cdot \phi(\mathbf{X}_j)$ در فضای ویژگی به کار رفته است.

براساس مطالب بیان شده، معادله‌ی اساسی ماشین‌های برداری غیرخطی به صورت معادله‌ی (41-2) می‌باشد. این معادله، همان معادله‌ی اساسی ماشین برداری پشتیبان در سیستم‌های خطی؛ یعنی معادله‌ی (32-2) است و تنها تابع کرنل به جای ضرب نقطه‌ای $(\mathbf{X}_i^T \cdot \mathbf{X}_j)$ قرار گرفته است.

$$L_d(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \alpha_i \alpha_j K(\mathbf{X}_i, \mathbf{X}_j) \quad (41-2)$$

که در آن:

$$0 \leq \alpha_i \leq C \quad (42-2)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad (43-2)$$

اکنون سوالی که مطرح می‌شود چگونگی محاسبه‌ی $\phi(\mathbf{X}_i)$ است که مقداری مجهول می‌باشد. اما روابط (35-2) تا (41-2) نشان می‌دهد که می‌توان تمام محاسبات را بدون محاسبه‌ی $\phi(\mathbf{X})$ ها و

فقط با استفاده از توابع کرنل انجام داد بنابراین در عمل محاسبه‌ی نگاشت‌های بردارهای ورودی در فضای بالاتر لازم نیست و تنها ضرب داخلی فضای منطبق بر آنها که به صورت تابعی از بردارهای ورودی؛ یعنی، $\varphi(\mathbf{X}_i)^T \cdot \varphi(\mathbf{X}_j) = K(\mathbf{X}_i, \mathbf{X}_j)$ بیان می‌گردد کافی خواهد بود.

در الگوریتم ماشین‌های فراگیر، به نگاشت بردارهای ورودی با ابعاد محدود به فضای ضرب داخلی با ابعاد بالاتر بدون آن که به محاسبه‌ی دقیق مسیر انجام نگاشت نیاز باشد حقه‌ی کرنل^{۹۴} می‌گویند که این کار با هدف به دست آوردن ساختار خطی معنی‌دار در فضای ضرب داخلی صورت می‌گیرد. فضای ضرب داخلی نیز فضای هیلبرت نامیده می‌شود. بنابراین با استفاده از حقه‌ی کرنل می‌توان به راحتی یک ماشین برداری را در فضای هیلبرت به کار گرفت بدون آن که به محاسبه‌ی نگاشت بردارهای ورودی در فضای ویژگی نیاز باشد.

با استفاده از حقه‌ی کرنل می‌توان معادله‌ی اساسی ابرصفحه‌ی جداساز بهینه در سیستم‌های غیرخطی را به صورت معادله‌ی (۲-۴۴) نوشت :

$$d(x) = \sum_{i=1}^N \sum_{j=1}^N y_i \alpha_i K(\mathbf{X}_i, \mathbf{X}_j) + b \quad (۲-۴۴)$$

همان‌گونه که مشاهده می‌شود، مقدار معادله‌ی صفحه‌ی جداساز را می‌توان بدون محاسبه $\varphi(\mathbf{X}_i)$ در سیستم‌های غیرخطی به دست آورد و تنها کافیست که از کرنل مناسب برای حل معادله‌ی صفحه استفاده شود [۳۲ و ۳۴].

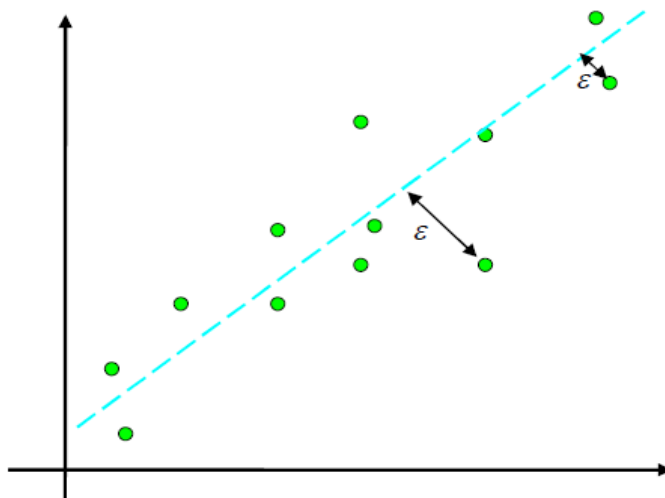
۲-۳-۳-۴- آنالیز رگرسیون

۲-۳-۳-۴- الف- آنالیز رگرسیون خطی

۲-۳-۳-۴- الف- ۱- ماشین برداری رگرسیونگر برای سیستم‌های خطی

⁹⁴ -Kernel trick

در آنالیز رگرسیون خطی، که رابطه‌ی خطی بین متغیر وابسته و غیروابسته وجود دارد میزان خطی بودن رابطه‌ی بین متغیرها اندازه‌گیری می‌گردد. روشی که برای ارائه‌ی یک مدل ریاضی جهت برازش بر داده‌ها استفاده می‌شود، سعی می‌کند بهترین مدلی را بیابد که به صورت کلی بر داده‌ها برازش خوبی داشته باشد و اگر داده‌ای انحراف بسیاری از داده‌های دیگر داشته باشد، مدل مجبور خواهد بود، به گونه‌ای برازش یابد که این داده را نیز تحت پوشش قرار دهد. این مدل به همراه خطای آن در شکل (۱۱-۲) نشان داده شده است [۳۶].



شکل (۱۱-۲) : مدل ریاضی برازش یافته در معادله رگرسیون خطی با روش کمترین مربعات

در رگرسیون خطی تابعی مانند $f(x) = \mathbf{W}^T \cdot \mathbf{X} + b$ با استفاده از مجموعه بردارهایی مانند \mathbf{X} ، مقدار اسکالری مانند y را تخمین می‌زند. در مسائل طبقه‌بندی، رگرسیون خطی به صورت سنتی با استفاده از روش کمترین مربعات حل می‌شود. به عبارت دیگر، تابع رگرسیون یک ابر صفحه جداساز است که بر روی داده‌ها با در نظر گرفتن کمترین مربع خطا بین ابر صفحه و داده‌ها، برازش می‌یابد. اما ماشین برداری رگرسیونگر^{۹۵} (SVR) هدف دیگری دارد. ایده اصلی یک SVR پیدا نمودن تابعی است که بر داده با کمترین انحراف از کمیتی مانند ϵ برای هر جفت X_i, y_i برازش یابد. در عین حال تلاش

^{۹۵} -Support Vector Regression

می‌شود که کمترین مقدار ممکن نیز برای $\|W\|$ به دست آید. این بدان معنی است که SVR، به خطاهای کمتر از ε کاری ندارد و سعی می‌کند تا خطاهای بزرگتر از آن را کوچک نماید. انجام چنین کاری اجازه می‌دهد تا ماشینی ساخته شود که پارامترهای آن، ترکیبی خطی از نمونه‌هایی باشد که دارای خطای بیشتر یا مساوی ε هستند.

۲-۳-۳-۴-الف-۲- فرموله کردن SVR

براساس مطالب بیان شده، رگرسیون خطی برای ماشین‌های برداری می‌تواند به صورت تابع اولیه

(۴۵-۲) نوشته شود که در آن هدف، مینیمم‌سازی خطای کلی و $\|W\|$ است.

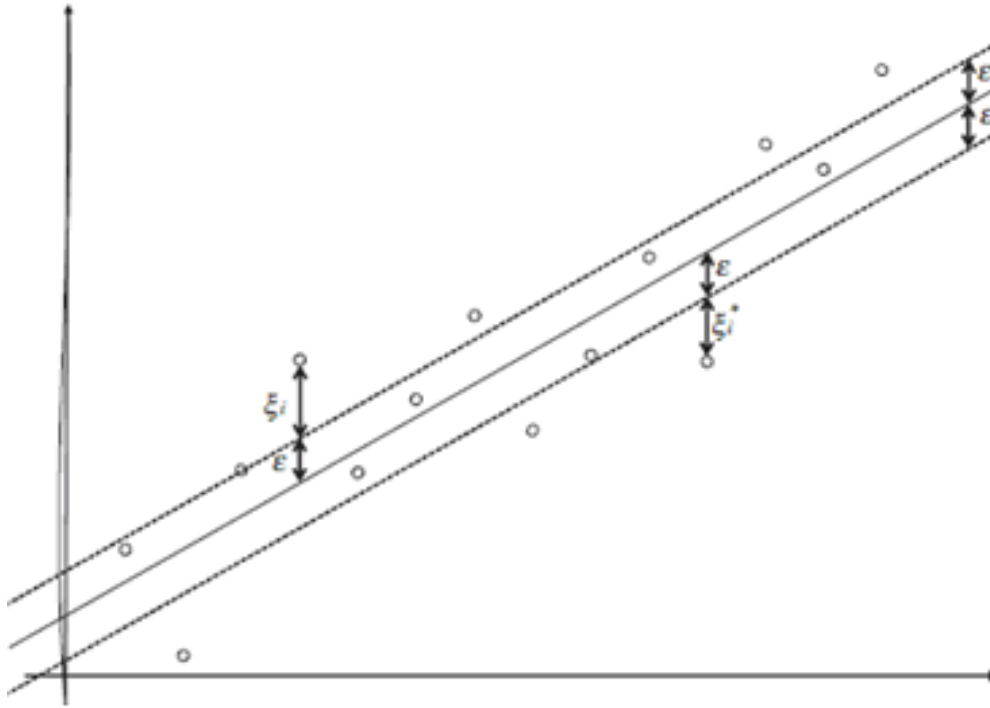
$$L_p = \frac{1}{2} \|W\| + C \sum_{i=1}^N (\xi_i + \xi'_i) \quad (۴۵-۲)$$

با توجه به این که :

$$y_i - W^T \cdot X - b \leq \xi_i + \varepsilon \quad (۴۶-۲)$$

$$-y_i + W^T \cdot X + b \leq \xi'_i + \varepsilon \quad (۴۷-۲)$$

$$\xi_i, X_i \geq 0 \quad (۴۸-۲)$$



شکل (۲-۱۲): نظریه‌ی اصلی حساسیت به مقدار ϵ . نمونه‌های خارج از حاشیه $\pm \epsilon$ متغیرهای خطای غیر صفر (ξ_i) هستند که قسمتی از راه‌حل ارائه شده خواهند بود.

محدودیت‌های معادله (۲-۴۹) بدین معنی هستند که خطای بیش‌تر از ϵ به کمتر از $\epsilon + \xi_i$ تبدیل خواهد شد و اگر مقدار خطا کمتر از ϵ باشد، متغیرهای خطا (ξ_i) مساوی صفر در نظر گرفته می‌شوند. این، نظریه حساسیت به ϵ ^{۹۶} است. این نظریه در شکل (۲-۱۲) نشان داده شده است [۳۲].

در این نوع از مدل‌های رگرسیون، تابع هدف مجموع مقادیر خطا را مینیمم خواهد ساخت و تنها نمونه‌هایی مورد استفاده قرار خواهند گرفت که خطای آنها بیش از ϵ باشد، بنابراین راه‌حل، تابعی از این نمونه‌ها خواهد بود.

۲-۳-۳-۴-الف-۳- بهینه‌سازی کاربردی ماشین برداری رگرسیونگر (SVR)

^{۹۶} - ϵ -sensitivity

برای حل مسئله‌ی بهینه‌سازی دارای محدودیت (۲-۴۹)، باید از بهینه‌سازی لاگرانژ برای تبدیل این معادله به یک معادله‌ی بدون محدودیت استفاده نمود. با در نظر گرفتن تابع لاگرانژ و مشتق‌گیری از تابع هدف بدون محدودیت نسبت به دو پارامتر W و b ، دو معادله به صورت (۲-۴۹) و (۲-۵۰) به دست می‌آیند.

$$W = \sum_{i=1}^N (\alpha_i - \alpha'_i) X_i \quad (۲-۴۹)$$

$$\sum_{i=1}^N (\alpha_i - \alpha'_i) = 0 \quad (۲-۵۰)$$

با قرار دادن رابطه‌ی به دست آمده از معادله‌ی (۲-۳۹) برای مقدار W ، در معادله‌ی به دست آمده از تابع لاگرانژ، معادله‌ی اساسی ماشین‌های برداری رگرسیونگر به صورت معادله‌ی (۲-۵۱) نوشته خواهد شد:

$$L_d = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha'_i) X_i^T \cdot X_j (\alpha_i - \alpha'_i) + \sum_{i=1}^N ((\alpha_i - \alpha'_i) y_i - (\alpha_i + \alpha'_i) \varepsilon) \quad (۲-۵۱)$$

با توجه به این که:

$$0 \leq (\alpha_i - \alpha'_i) \leq C \quad (۲-۵۲)$$

برای یافتن مقدار b نیز می‌توان از رابطه‌ی (۲-۵۳) استفاده نمود.

$$-y_i + W^T \cdot X_i + b + \varepsilon = 0 \quad (۲-۵۳)$$

که باید:

$$\alpha_i, \alpha_j < C \quad (۲-۵۴)$$

۲-۳-۳-۴-ب-آنالیز رگرسیون غیرخطی

۲-۳-۳-۴-ب-۱- ماشین برداری رگرسیونگر برای سیستم‌های غیرخطی

همان‌گونه که در بحث ماشین‌های برداری رگرسیونگر در سیستم‌های خطی بیان شد، مقدار

بهینه W از رابطه‌ی $W = \sum_{i=1}^N (\alpha_i - \alpha'_i) X_i$ به دست می‌آید که این رابطه در سیستم‌های غیرخطی

تبدیل به رابطه‌ی $W = \sum_{i=1}^N (\alpha_i - \alpha'_i) \varphi(X_i)$ می‌شود. مسئله‌ای که در این حالت ایجاد می‌شود، همان

مشکلی است که در سیستم‌های غیرخطی وجود دارد و آن بردن به فضای بالاتر و به دست آوردن

مقدار $\varphi(X_i)$ است که مقدار آن ناشناخته است. در این شرایط مانند حالت طبقه‌بندی داده‌ها از

حقیقی کرنل استفاده می‌شود تا بتوان بدون محاسبه‌ی مقدار $\varphi(X_i)$ و تنها با استفاده از کرنل‌های

موجود، بهینه‌ترین مدل ریاضی را بر داده‌ها برازش نمود. براساس مطالب بیان شده، رابطه‌ی صفحه‌ی

جداساز در سیستم‌های غیرخطی به صورت $y_i = \sum_{i=1}^N W^T \varphi(X_i) + b$ بیان می‌گردد. با قراردادن

رابطه‌ی به دست آمده برای W در رابطه‌ی صفحه، می‌توان معادله‌ی اساسی (۲-۵۹) را برای

ماشین‌های برداری رگرسیونگر در سیستم‌های غیرخطی به صورت زیر معرفی نمود.

$$y_i = \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha'_i) \varphi(X_i)^T \varphi(X_j) + b = \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha'_i) K(X_i, X_j) + b \quad (۲-۵۵)$$

با در نظر گرفتن رابطه‌ی (۲-۵۵) نیازی به محاسبه مقدار $\varphi(X_i)$ نیست و مقدار b نیز از رابطه‌ای که

در بخش طبقه‌بندی بیان شد، به صورت $b = y_i - \sum_{i=1}^{N_{SV}} \sum_{j=1}^{N_{SV}} \alpha_i y_i K(X_i, X_j)$ محاسبه می‌شود [۳۲ و

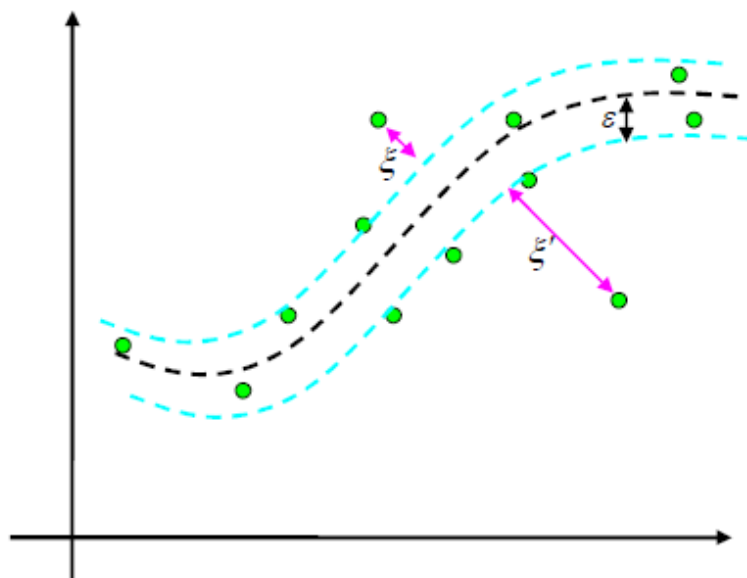
[۳۴].

عملکرد ماشین‌های برداری پشتیبان در قیاس با مدل‌های دیگری که در سیستم‌های غیرخطی

وجود دارد، حاکی از قدرت و عملکرد بالای این ماشین‌ها در اکثر سیستم‌های به کار گرفته شده،

خصوصاً در سیستم‌های غیرخطی است. مدل ریاضی ارائه شده توسط ماشین برداری برای یک

رگرسیون غیرخطی در شکل (۲-۱۳) نشان داده شده است.



شکل (۲-۱۳): مدل ارائه شده توسط ماشین برداری برای سیستم غیرخطی

۲-۴- ارزیابی مدل

برای اطمینان از اینکه مدل به دست آمده، مدل مناسبی است که توانایی پیشگویی نمونه‌های مختلفی از یک جمعیت را داراست، باید مدل را ارزیابی کرد. این ارزیابی از طریق شاخص‌های کمی است که به وسیله آن‌ها صحت نتایج ارائه شده توسط مدل مورد سنجش قرار می‌گیرند. برخی از این شاخص‌های کمی عبارتند از:

ضریب همبستگی^{۹۷}: ساده‌ترین راه برای بررسی میزان همبستگی دو یا چند متغیر، محاسبه آماره‌ی ضریب همبستگی آنهاست. ضریب همبستگی دو متغیر x, y با رابطه (۲-۲۰) تعریف می‌شود. مقدار این آماره بین ۱ تا -۱ متغیر است. مقدار بزرگ‌تر آن نشان دهنده‌ی این است که ارتباط خطی بیشتری میان متغیر وابسته و متغیرهای مستقل وجود دارد.

⁹⁷ - Correlation coefficient

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (56-2)$$

ضریب تعیین^{۹۸}: به عنوان یک شاخص برای بیان دقت خط رگرسیون برآورد شده، به کار می‌رود و نشان‌دهنده‌ی نسبت تغییرات متغیر وابسته توضیح داده شده توسط متغیر مستقل است. به عنوان مثال R^2 برابر با ۰/۹۲۴۳/۹۲۴۳ نشان می‌دهد که ۹۲/۴۳ درصد تغییرات در متغیر وابسته می‌تواند توسط متغیر مستقل توضیح داده شود. رابطه ریاضی مربوط به ضریب تعیین به صورت زیر است:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (57-2)$$

که SSR ^{۹۹} طبق رابطه (۲۲-۲) بیانگر مجموع مربعات انحراف مقادیر پیش‌بینی شده‌ی متغیر وابسته از میانگین مقادیر آن است:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (58-2)$$

SST ^{۱۰۰} طبق رابطه (۲۳-۲) نشانگر مجموع مربعات انحراف مقادیر واقعی متغیر وابسته از میانگین مقادیر آن است.

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (59-2)$$

SSE ^{۱۰۱} نیز مبین مجموع مربعات انحراف مقادیر واقعی متغیر وابسته از مقادیر پیش‌بینی شده برای آن است:

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (60-2)$$

بنابراین با توجه به روابط فوق می‌توان نوشت:

⁹⁸ - Determination coefficient

⁹⁹ - Sum Square Regression

¹⁰⁰ - Sum Square Total

¹⁰¹ - Sum Square Error

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (۶۱-۲)$$

طبق رابطه (۶۱-۲) اگر تمام مشاهدات بر روی خط برازش شده قرار گرفته باشند، یعنی به ازای تمام نقاط $y_i = \hat{y}_i$ باشد، مقدار R^2 برابر یک می‌شود و هرگونه انحرافی از این حالت باعث می‌شود که مقدار R^2 از یک کوچکتر شود.

ضریب تعیین تصحیح شده^{۱۰۲}: ضریب تعیین تصحیح شده یکی از پارامترهایی است که در ارزیابی و مقایسه مدل‌های مختلف با تعداد متفاوت متغیرهای پیشگو استفاده می‌شود و از آنجایی که R^2 (ضریب تعیین)، تعداد پارامترهای موجود در مدل را به حساب نمی‌آورد، بدین منظور از ضریب تعیین تصحیح شده طبق رابطه زیر برای انتخاب مدل برتر استفاده می‌شود:

$$R_{adj}^2 = 1 - \frac{n-1}{n-p-1} \cdot \frac{SSE}{SST} = 1 - (1-R^2) \frac{n-1}{n-p-1} \quad (۶۲-۲)$$

که در این رابطه p تعداد متغیرهای مستقل و n تعداد ترکیبات مورد بررسی می‌باشد. آماره F : آزمون F یا آزمون فیشر در واقع آزمون معنی‌دار بودن آماری در تحلیل رگرسیون ساده و چند متغیره است و برابر با نسبت میانگین مربعات رگرسیون (MSR) به میانگین مربعات باقیمانده‌ها (MSE) است. بیان ریاضی آن به صورت زیر می‌باشد:

$$F = \frac{MSR}{MSE} = \frac{SSR / df_m}{SSE / df_{res}} \quad (۶۳-۲)$$

$$df_m = p \quad , \quad df_{res} = n - p - 1$$

df_m درجه آزادی مدل، p تعداد متغیرهای مستقل مدل، n تعداد کل ترکیبات مربوط به مدل، df_{res} درجه آزادی باقیمانده‌ها است. و SSE ، SSR در روابط قبل توضیح داده شد.

¹⁰² - Adjusted determination coefficient

مجموع مربع باقیمانده ها^{۱۰۳} (PRESS): برابر مجموع مربعات تفاوت بین مقدار کمیت مشاهده شده (y_i) و مقدار تخمین زده شده (\hat{y}_i) است.

$$PRESS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (۶۴-۲)$$

خطای استاندارد پیش بینی^{۱۰۴} (SEP):

$$SEP = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (۶۵-۲)$$

میانگین مربع خطاها^{۱۰۵} (MSE)

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (۶۶-۲)$$

FIT: به عنوان یک پارامتر آماری برای مقایسه مدل‌های مختلف با تعداد متفاوت توصیفگر به کار می‌رود که هر چه مقدار این پارامتر برای یک مدل بیشتر باشد، آن مدل مناسب‌تر می‌باشد. مقدار آن از رابطه زیر به دست می‌آید.

$$FIT = \frac{R^2(n - p - 1)}{(n + p^2)(1 - R^2)} \quad (۶۷-۲)$$

که R^2 مجذور ضریب همبستگی مدل، p تعداد توصیفگرهای مدل و n تعداد ترکیبات مربوط به مدل است.

¹⁰³ - Predictive Residual Sum of Squares

¹⁰⁴ - Standard Error of Prediction

¹⁰⁵ - Mean Square Error

فصل سوم

مطالعه ارتباط کمی ساختار - خاصیت خطی و غیر خطی

ویسکوزیته آلکانها در محدوده وسیعی از دما و فشار

۳-۱- مدل سازی QSPR ویسکوز η خطی در محدوده وسیعی از دما و

فشار

در این مرحله از تحقیق، از رگرسیون خطی چندگانه، شبکه عصبی و ماشین بردار پشتیبان جهت مدل سازی و پیش بینی ویسکوزیته یک همولوگ از آلکان های خطی از n -هگزان تا n -نونادکان استفاده شده است. آنچه که در ادامه می آید چگونگی مراحل مدل سازی با استفاده از این سه روش می باشد. همچنین توانایی مدل های حاصله از سه روش در پیش بینی این پارامتر با یکدیگر مورد مقایسه و ارزیابی قرار گرفته است.

۳-۲- مراحل مدل سازی

۳-۲-۱- انتخاب سری داده‌ها

در این کار چون هدف پیش‌بینی ویسکوزیته آلکان‌های خطی از n -هگزان تا n -نونادکان بود لذا داده‌های ویسکوزیته برای این ترکیبات در محدوده‌ی وسیعی از دما و فشار جمع آوری گردید [۳۷-۴۴]. نام این ترکیبات به همراه محدوده دما و فشار و مرجع مورد استفاده برای هر ترکیب در جدول (۳-۱) آمده است. لازم به ذکر است که این داده‌ها به سه دسته آموزش، ارزیابی و تست تقسیم شدند. به صورتی که ۳ ترکیب را برای تست و ۱۱ ترکیب باقیمانده به طور تصادفی به دو دسته آموزش و ارزیابی تقسیم شدند که تعداد نقاط استفاده شده از هر ترکیب برای این سه سری نیز در جدول (۳-۱) گزارش شده است.

جدول (۳-۱): داده‌های ویسکوزیته در محدوده دما و فشار

نام مولکول	ΔT (K)	ΔP (Mpa)	تعداد نقاط مورد استفاده در سری آموزش	تعداد نقاط مورد استفاده در سری ارزیابی	تعداد نقاط مورد استفاده در سری تست	مرجع
n -هگزان	۲۹۸/۲۹-۳۷۳/۳۶	۵/۰۱-۲۰۴/۵	۳۷	۹	-	۳۸
n -هپتان	۳۰۳-۳۴۸	۵/۸-۲۰۲/۰۵	۵۰	۱۵	-	۳۹-۴۰
n -اکتان	۲۸۳/۱۵-۳۵۳/۱۵	۱۱/۹۶-۱۵۱/۳۷	-	-	۲۹	۳۸-۴۱-۴۲
n -نونان	۳۰۳/۱۵-۳۲۳/۱۵	۵/۴۷-۶۹	۱۹	۶	-	۴۰

n-دکان	۳۰۳/۱۵-۳۴۸/۱۵	۴/۶۷-۲۰۰/۸	۳۲	۷	-	۴۵
n-آن دکان	۳۰۳/۱۵-۳۲۳/۱۵	۵/۳۲-۶۲/۴۲	۲۱	۳	-	۴۰
n-دودکان	۲۹۸/۱۲-۴۷۳/۱۵	۹/۱-۱۶۳/۲۴	-	-	۴۸	۴۱-۴۲-۴۳
n-تری دکان	۲۹۳/۱۵-۳۵۳/۱۵	۲۰-۱۰۰	۲۷	۸	-	۴۵
n-تترادکان	۳۱۳/۲-۳۹۳/۲	۵-۶۰	۲۶	۹	-	۴۵
n-پنتادکان	۳۰۳/۹۵-۵۱۷/۷۵	۹/۸-۴۹	۲۶	۶	-	۴۵
n-هگزادکان	۲۹۸/۰۸-۳۷۳/۲۴	۴/۱-۱۵۴/۲	-	-	۲۵	۴۱-۴۴
n-هپتادکان	۳۲۴/۷-۵۲۲/۷	۲/۴۵۲-۳۹/۲۴	۵۰	۱۳	-	۴۵
n-اکتادکان	۳۲۳/۱۵-۴۷۳/۱۵	۲۰/۷۸-۸۲/۴۷	۲۳	۴	-	۴۵
n-نونادکان	۳۴۶/۴۵-۵۲۳/۱۵	۲/۴۵۲-۴۹/۰۵	۴۵	۶	-	۴۵

۳-۲-۲- رسم و بهینه سازی مولکول‌ها

پس از انتخاب سری داده‌ها اولین مرحله جهت مدل‌سازی، رسم و بهینه‌سازی ساختار مولکول‌های مورد بررسی می‌باشد. بدین منظور ساختار شیمیایی تمام ترکیبات با استفاده از نرم‌افزار Hyperchem رسم و سپس با احتساب اتم‌های هیدروژن، ساختار سه بعدی ترکیبات با استفاده از روش نیمه تجربی AM1 بهینه گردید. لازم به ذکر است که بهینه‌سازی تا حد گرادیانی ۰/۰۰۱ کیلو کالری بر انگستروم مول ادامه یافت.

۳-۲-۳- محاسبه‌ی توصیف‌کننده‌های مولکولی

توصیف‌کننده‌های مولکولی مقادیر عددی هستند که ویژگی‌های مختلف ساختاری یا شکل مولکولی ترکیبات را توصیف می‌کنند. گروه‌های مختلفی از این توصیف‌کننده‌ها تاکنون معرفی شده‌اند که هر کدام جنبه‌های خاصی از ترکیبات را بررسی و توصیف می‌کنند. ساختارهای سه بعدی بهینه شده از مرحله قبل به عنوان ورودی به نرم‌افزار Dragon وارد شده و بوسیله‌ی آن ۱۴۸۱ توصیف‌کننده‌ی مختلف در ۱۸ طبقه مختلف برای هر یک از ترکیبات محاسبه شد. پس از محاسبه‌ی

توصیف‌کننده‌ها با نرم‌افزار Dragon می‌توان توصیف‌کننده‌های با مقدار ثابت در کل ترکیبات را حذف نمود.

۳-۲-۴- حذف توصیف‌کننده‌های نامناسب

مهمترین مرحله در مطالعات QSPR، انتخاب توصیف‌کننده‌های مناسب است که بتواند ویسکوزیته را به ساختار مولکول ارتباط دهند. با توجه به این موضوع و اینکه تعداد زیاد توصیف‌کننده باعث پیچیدگی محاسبات می‌شود و همچنین تعدادی از توصیف‌کننده‌ها حاوی اطلاعات یکسان و در نتیجه اثر یکسان بر روی ویسکوزیته هستند، لذا باید روش کاهش متغیرها به کار گرفته شود. به همین منظور، ابتدا توصیف‌کننده‌هایی که دارای مقادیر یکسانی برای کل مولکول‌ها بودند از فهرست توصیف‌کننده‌ها حذف شدند. سپس با بررسی همبستگی بین متغیرها از میان هر دو توصیف‌کننده‌ای که ضریب همبستگی بزرگتر از ۰/۹ داشتند، آنکه همبستگی کمتری با متغیر وابسته داشت، حذف گردید. برای این کار برنامه‌ای در نرم‌افزار MATLAB نوشته و اجرا شد که تا این مرحله ۲۳ توصیف‌کننده باقی ماند. در نهایت با استفاده از روش برازش مرحله‌ای در انتخاب متغیر، تعدادی از توصیف‌کننده‌ها که بیشترین همبستگی را با پارامتر وابسته داشتند، انتخاب گردیدند. این مراحل به این صورت انجام گرفت که در منوی آنالیز در نرم‌افزار SPSS، گزینه رگرسیون خطی انتخاب شد و بعد از وارد کردن متغیرهای مستقل و متغیر وابسته، روش مرحله‌ای (stepwise) برای مدل‌سازی استفاده شد. چندین مدل مختلف به دست آمد که در بین مدل‌های ایجاد شده مدلی که تعداد توصیف‌کننده‌های کمتر و ضریب تعیین R^2 ، FIT و R^2_{adj} بیشتری داشت، به عنوان بهترین مدل انتخاب شد.

۳-۲-۵- مدل‌سازی توسط مدل رگرسیون خطی چندگانه

مدل‌های مختلفی که توسط روش رگرسیون خطی چندگانه مرحله‌ای به دست آمد، مورد بررسی قرار گرفتند. برای این کار داده‌ها به طور تصادفی به سه سری آموزش، تست و ارزیابی تقسیم شدند. سری آموزش ۳۵۶، سری تست ۱۰۲ و سری ارزیابی ۸۶ نقطه داده را شامل می‌شوند. سپس قدرت پیش‌بینی مدل‌های مختلف مورد سنجش قرار گرفتند. در روش رگرسیون مرحله‌ای، متغیرها یکی پس از دیگری وارد مدل شده و ابتدا متغیری وارد مدل می‌شود که بالاترین میزان همبستگی را با متغیر وابسته دارد. با ورود هر متغیر جدید، کلیه متغیرهای موجود در معادله بررسی شده و اگر هر کدام از آنها سطح معنی‌دار بودن خود را از دست داده باشند، قبل از ورود متغیر جدید از مدل خارج می‌شوند. به همین منظور، داده‌های ویسکوزیته به عنوان متغیر وابسته و توصیف‌کننده‌ها به عنوان متغیر مستقل در نظر گرفته شدند. در بین مدل‌های ایجاد شده، مدلی که تعداد توصیف‌کننده و MSE کمتر، R^2 ، FIT و ضریب تصحیح شده‌ی (R^2_{adj}) بیشتری دارد، به عنوان مدل برتر انتخاب می‌شود. توصیف‌کننده‌های انتخاب شده جهت ایجاد مدل MLR به همراه پارامترهای آماری مربوط در جدول (۲-۳) آمده است.

جدول (۲-۳) : پارامترهای آماری برای مدل‌های به دست آمده از روش MLR برای سری ارزیابی

مدل	نام توصیف‌کننده‌ها	FIT	R^2_{adj}	R^2	MSE
۱	IC5	۰/۲۰۹۴	۰/۱۷۰۹	۰/۱۷۴	۰/۵۲۳
۲	IC5, P	۰/۳۶۸۸	۰/۲۶۷۹	۰/۲۷۲	۰/۴۶۲
۳	IC5, P, RDF135u	۰/۶۰۳۵	۰/۳۷۷۴	۰/۳۸۲	۰/۳۹۰
۴	IC5, P, RDF135u, T	۱/۶۰۴۱	۰/۶۲۱۵	۰/۶۲۵	۰/۲۳۸
۵	IC5, P, RDF135u, T, MLOGP	۲/۲۱۶۷	۰/۶۹۷۶	۰/۷۰۱	۰/۲۲۸

۶	IC5, P, RDF135u, T, MLOGP, Mor08p	۳/۰۴۷۷	۰/۷۶۳۹	۰/۷۶۷	۰/۱۵۷
۷	IC5, P, T, MLOGP, Mor08p	۳/۰۹۵۱	۰/۷۶۳۳	۰/۷۶۶	۰/۱۶۲
۸	IC5, P, T, MLOGP, Mor08p, E1u	۳/۰۴۷۷	۰/۷۶۳۹	۰/۷۶۷	۰/۱۹۰
۹	IC5, P, T, MLOGP, Mor08p, E1u, RDF060v	۲/۶۲۶۸	۰/۷۴۰۱	۰/۷۴۴	۱۹۸/۷۷

به این ترتیب مدل ۷ بعنوان بهترین مدل برگزیده شده که معادله‌ی آن به صورت زیر می‌باشد:

$$Vis = 1/948 + 0/36IC5 + 0/008P - 0/012T + 0/915MLOGP - 13/874Mor08p$$

در جدول (۳-۳) نام توصیف‌کننده‌های انتخاب شده توسط مدل رگرسیون خطی چندگانه همراه با طبقه و اثرمتوسط آن‌ها آورده شده است:

جدول (۳-۳): توصیف‌کننده‌های مولکولی انتخاب شده توسط روش رگرسیون خطی چندگانه مرحله‌ای در مدل ۷

N0	College	Class	Meaning	Mean Effect
۱	IC5	Information indices	Information Content index	۴/۰۵۸
۲	MLOGP	Molecular properties	Moriguchi octanol-water partition coeff	۲۵/۶۵۷
۳	Mor08p	3D MORSE	3D-MORSE-signal 8	-۶/۷۵

همچنین ماتریس همبستگی بین این توصیف‌کننده‌ها در جدول (۴-۳) ارائه شده که این ماتریس عدم همبستگی بین متغیرها را نشان می‌دهد.

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

	IC5	P	T	MLOGP	Mor08p
IC5	۱				

P	۰/۵۵	۱			
T	-۰/۵۲۴	-۰/۲۷۲	۱		
MLOGP	-۰/۳۵۲	-۰/۴۰۶	۰/۶۸۵	۱	
Mor08p	-۰/۰۷۵	-۰/۳۹۲	۰/۵۴۳	۰/۸۲۲	۱

۳-۲-۶- مدل سازی با روش های غیر خطی

در این پایان نامه از بین روش های غیرخطی موجود در مراجع علمی، روش شبکه ی عصبی مصنوعی (ANN) به عنوان پرکاربردترین تکنیک در مطالعات QSPR و روش ماشین برداری پشتیبان (SVM) به عنوان روشی جدید مورد استفاده قرار گرفتند. در ادامه هر یک از این روش های مدل سازی به تفصیل مورد بحث قرار می گیرد.

۳-۲-۶-۱- مدل سازی توسط شبکه عصبی

با توجه به پیچیدگی داده ها و احتمال وجود روابط غیرخطی بین متغیرهای مستقل (توصیف کننده ها) و متغیر وابسته (ویسکوزیته) علاوه بر مدل سازی به روش رگرسیون خطی چندگانه، روش های غیرخطی نیز برای مدل سازی ویسکوزیته ترکیبات مورد استفاده قرار گرفتند. از میان روش های غیرخطی، شبکه ی عصبی به عنوان یک تکنیک برتر برای مطالعه ارتباط ساختار-ویسکوزیته ی ترکیبات انتخاب گردید. شبکه عصبی مورد استفاده در این مطالعه یک شبکه پیشخور با الگوریتم آموزشی پس انتشار می باشد که الگوریتم آن با استفاده از نرم افزار مطلب ۷/۸ نوشته شد. مقدار توصیف کننده های وارد شده در مدل خطی به عنوان ورودی به شبکه ی عصبی داده شدند. مقدار تجربی ویسکوزیته نیز به عنوان هدف به دنبال ورودی ها وارد شبکه شدند تا پاسخ شبکه با آنها سنجیده شود. برای ایجاد یک مدل ANN پیشگو با درصد خطای پیش بینی پایین، باید ابتدا پارامترهای مؤثر بر شبکه بهینه شوند که این پارامترها عبارتند از:

تعداد متغیرهای ورودی (تعداد توصیف کننده ها)

تعداد گره‌های لایه پنهان

نوع تابع انتقال

نوع تابع آموزش

تعداد دوره‌های آموزش

پارامتر μ (مقدار اولیه برای μ)

۳-۲-۶-۱-الف- انتخاب تعداد گره‌های لایه‌های پنهان، تعداد متغیر ورودی، نوع

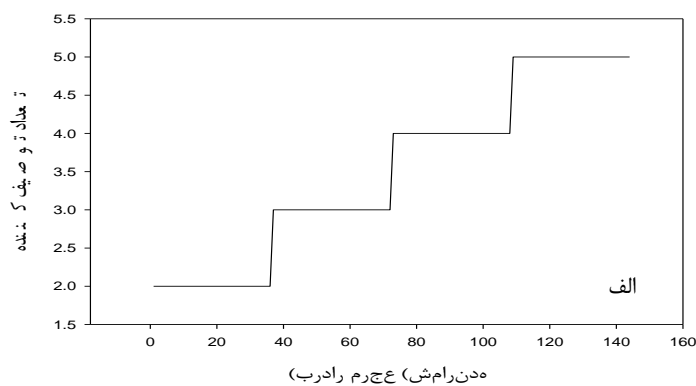
تابع آموزش و نوع تابع انتقال

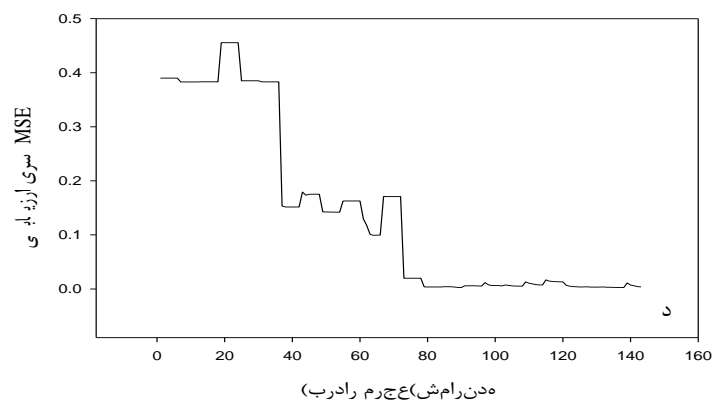
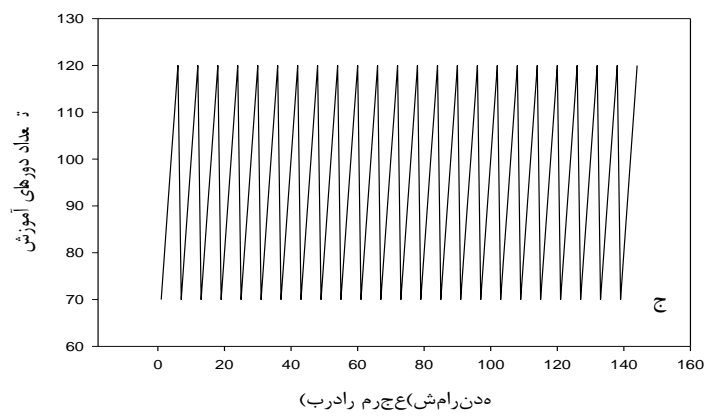
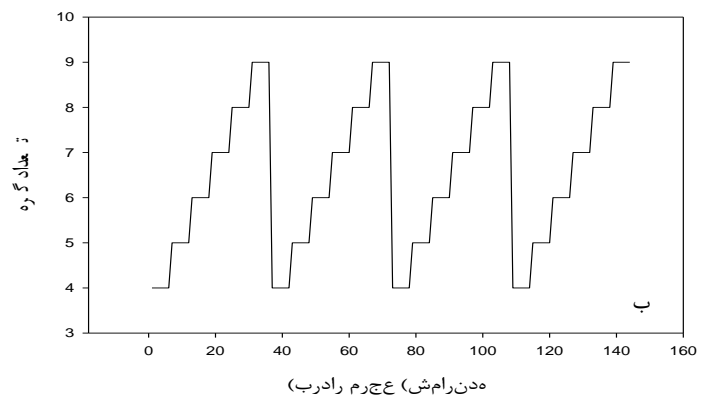
پس از طراحی اولیه شبکه، برای اینکه شبکه قدرت پیش‌بینی مناسب داشته باشد باید به طور مناسب آموزش داده شود. با توجه به اینکه فرایند آموزش، وقت‌گیر است لذا انتخاب الگوریتم آموزشی سریع‌تر امری مهم است. از میان الگوریتم‌های آموزشی متفاوت که در جعبه ابزار مطلب وجود دارد، الگوریتم لونبرگ - مارکوارت و تنظیم بایزین دارای سرعت مناسب، قابلیت پیش‌بینی و تعمیم‌پذیری بهتری نسبت به سایر روش‌های معمول مانند روش نزول گرادیانی می‌باشند. بنابراین این دو الگوریتم آموزشی برای آموزش شبکه انتخاب شدند. همچنین توابع لگاریتم سیگموئید (logsig) و تانژانت سیگموئید (tansig)، به عنوان توابع انتقال لایه‌ی پنهان، انتخاب گردیدند. به این ترتیب با استفاده از دو الگوریتم آموزشی و دو تابع انتقال انتخاب شده، چهار شبکه عصبی مصنوعی مختلف طراحی شدند. سپس برای هر یک از این شبکه‌ها، از روش تغییر همزمان برای تغییر پارامترهای شبکه شامل تعداد توصیف‌کننده‌های ورودی، تعداد گره‌های لایه پنهان و تعداد دوره‌های آموزش جهت آموزش شبکه استفاده گردید و به ازای همه ترکیب‌های ممکن از این سه پارامتر شبکه‌های طراحی شده بهینه‌سازی شدند. برای این منظور تعداد نرون‌های لایه‌ی ورودی از ۲ تا ۵ با گام ۱، تعداد گره‌های لایه‌ی پنهان از ۲ تا ۱۰ با گام ۱ و تعداد دور آموزش از ۱۰ تا ۱۰۰ با گام ۱۰ در نظر گرفته شدند. در تمامی این

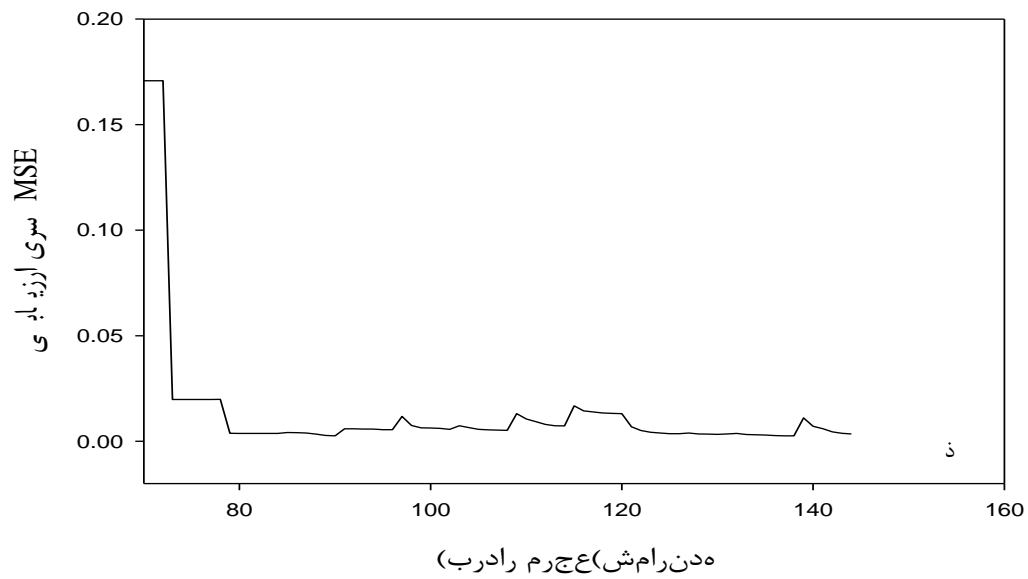
شبکه‌ها از تابع انتقال خطی (purelin) در لایه‌ی خروجی استفاده شد. برای مشخص شدن مقدار بهینه‌ی این پارامترها، مقدار MSE سری ارزیابی برای هر ترکیب ممکن از سه پارامتر شبکه ثابت و برحسب یک بردار مرجع رسم می‌گردد تا پارامترهای با کمترین خطای سری ارزیابی به عنوان مقادیر بهینه برای هر شبکه انتخاب شدند. در شکل‌های (۳-۱) تا (۳-۴) بخشی از این محاسبات برای چهار شبکه عصبی مختلف آورده شده است. لازم به ذکر است که در جدول (۳-۵) مقادیر بهینه برای هر چهار شبکه بهینه شده آورده شده است.

جدول (۳-۵): پارامترهای بهینه برای چهار شبکه عصبی مختلف

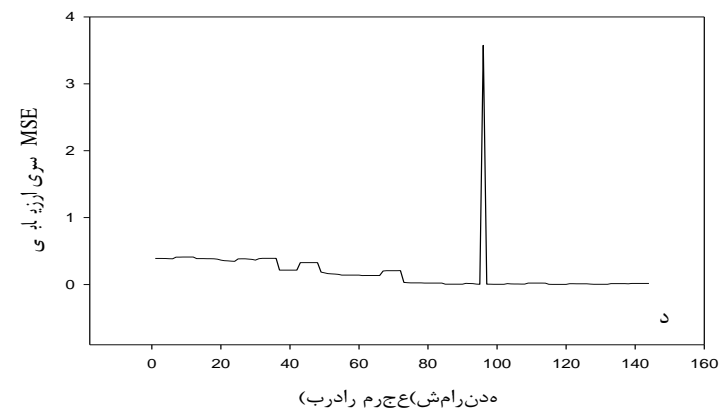
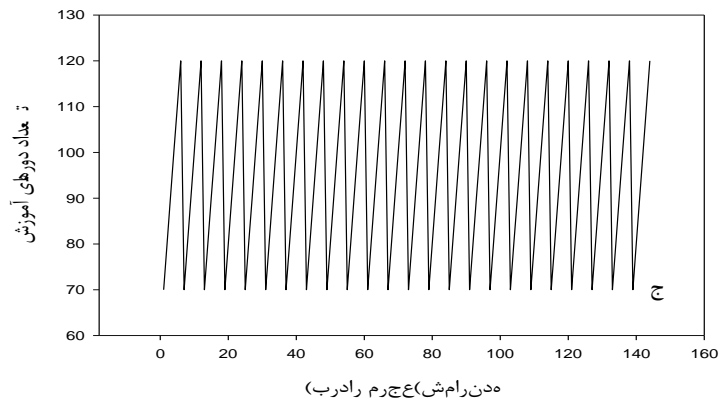
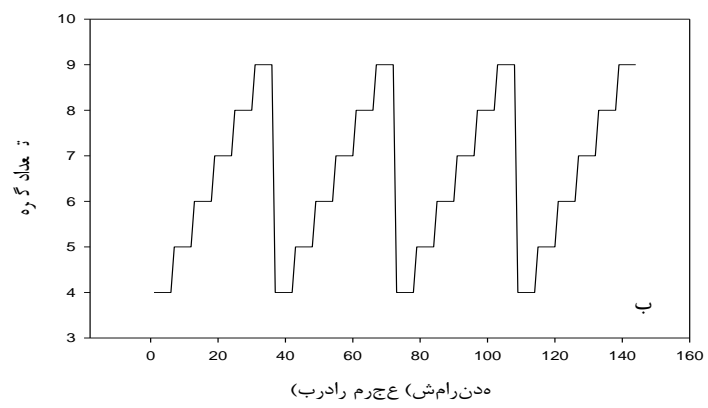
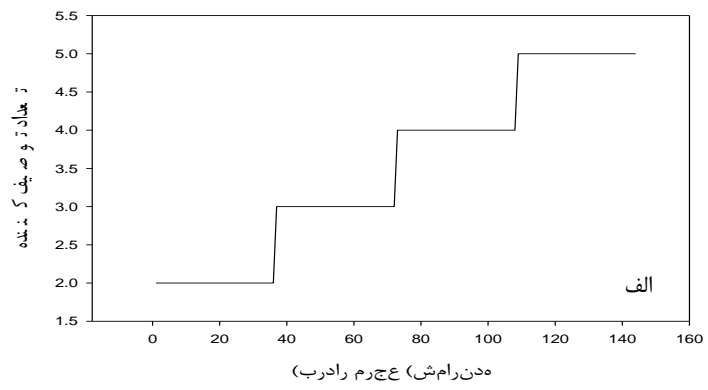
شماره مدل	۱	۲	۳	۴
تابع آموزش	Train br	Train lm	Train br	Train lm
تابع انتقال لایه پنهان	log sig	log sig	tan sig	tan sig
تعداد نرون‌های لایه پنهان	۸	۵	۵	۷
تعداد متغیرهای ورودی	۵	۵	۵	۵
تعداد دورهای آموزش	۱۱۰	۱۰۰	۱۰۰	۱۲۰
میانگین مربع خطا	۰/۰۰۲۶	۰/۰۰۲۲	۰/۰۰۳۳	۰/۰۰۲۴

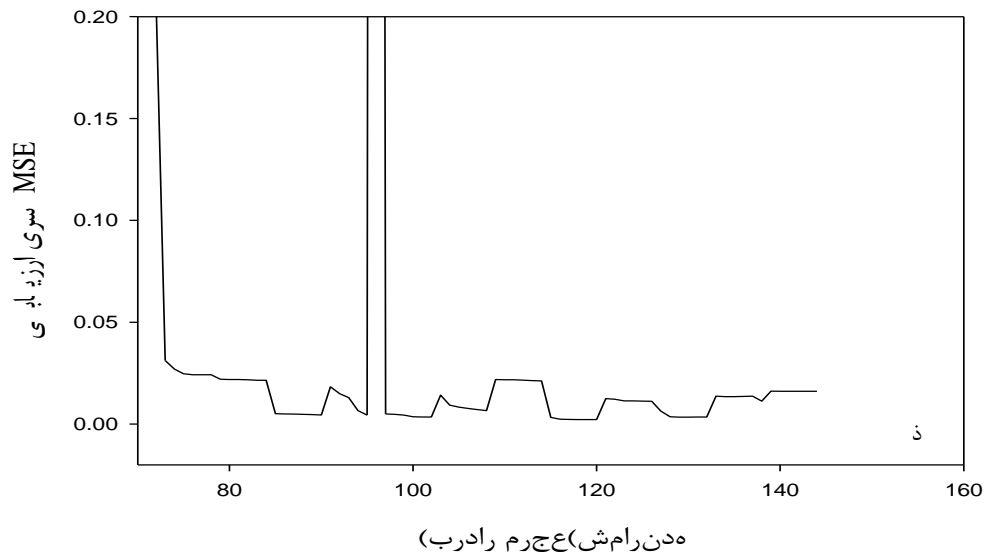




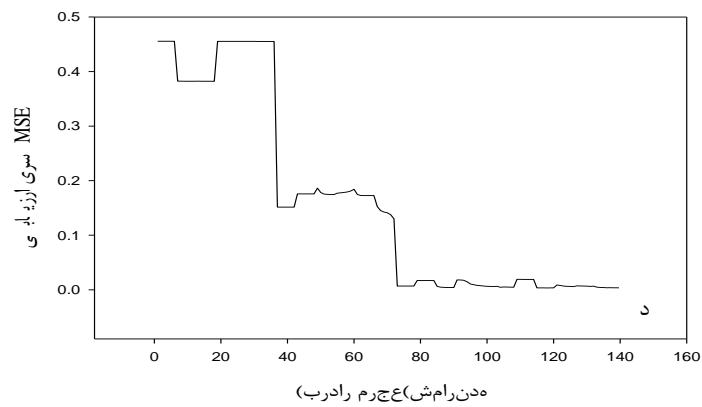
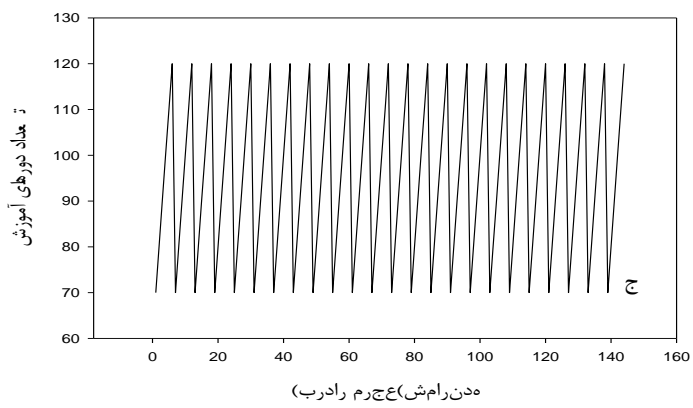
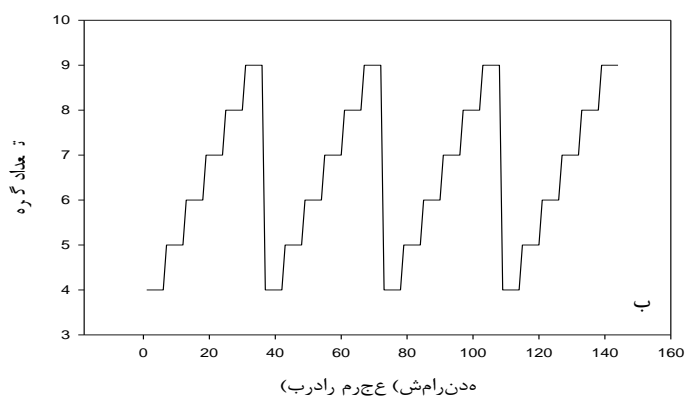
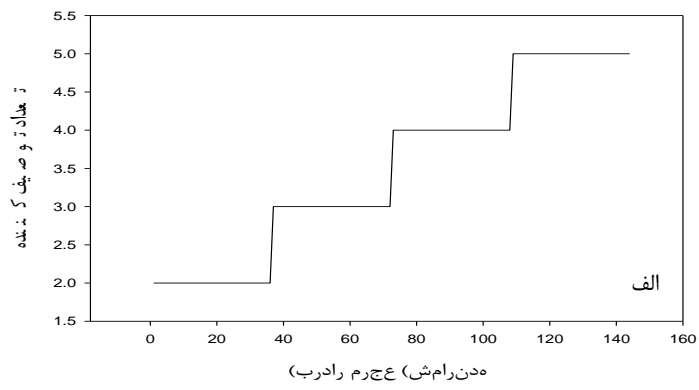


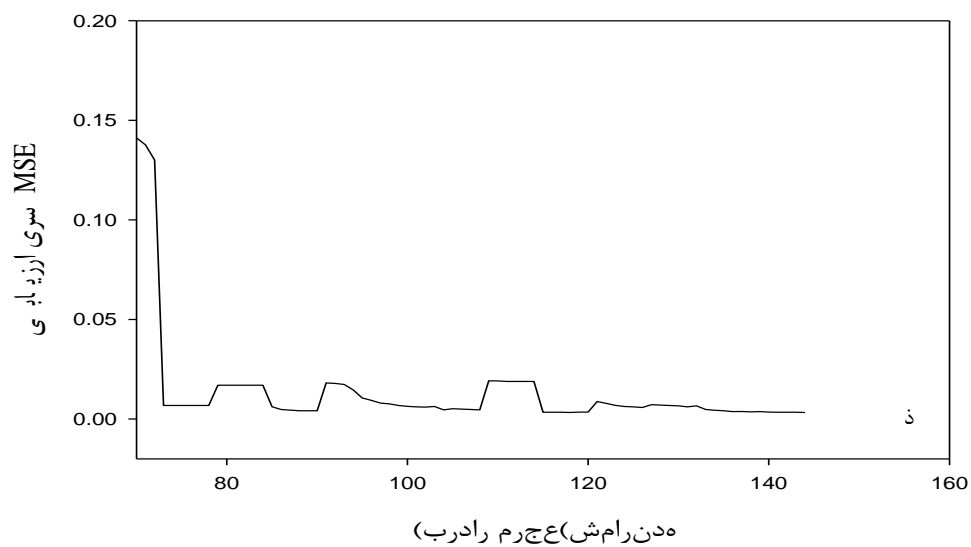
شکل (۳-۱): نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) سری ارزیابی MSE از ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است. برای شبکه‌هایی با الگوریتم آموزشی بایزین و تابع انتقال لگاریتم سیگموئیدی



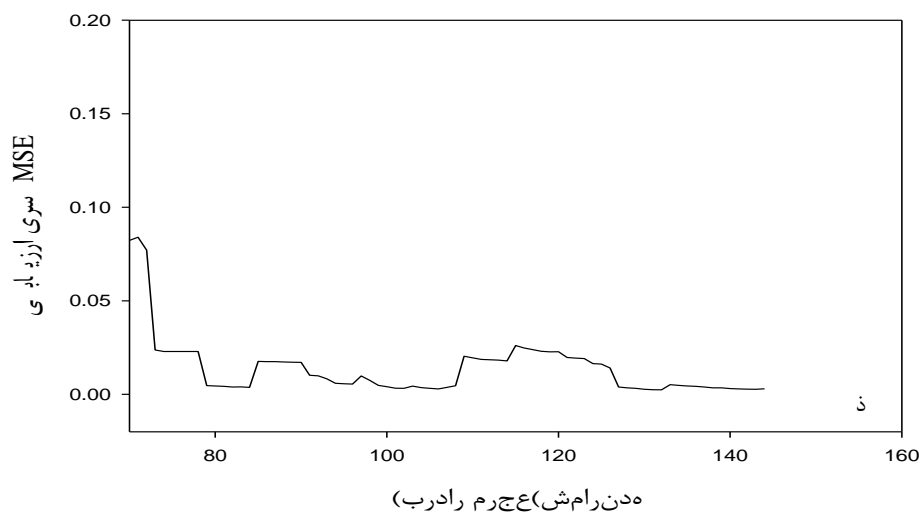


شکل (۳-۲) : نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) سری ارزیابی (ذ) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است. برای شبکه‌هایی با الگوریتم آموزشی لونبرگ- مارکوارت و تابع انتقال لگاریتم سیگموئیدی





شکل (۳-۳): نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) سری ارزیابی MSE از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است. برای شبکه‌هایی با الگوریتم آموزشی بایزین و تابع انتقال تانژانت سیگموئیدی



شکل (۳-۴) : نمودارهای (الف) تعداد توصیف‌کننده (ب) تعداد نرون لایه مخفی (ج) تعداد دور آموزش (د) MSE سری ارزیابی (ذ) ناحیه‌ای از نمودار MSE سری ارزیابی که مینیمم مقدار آن بهتر نشان داده شده است. برای شبکه‌هایی با الگوریتم آموزشی لونبرگ-مارکوارت و تابع انتقال تانژانت سیگموئیدی

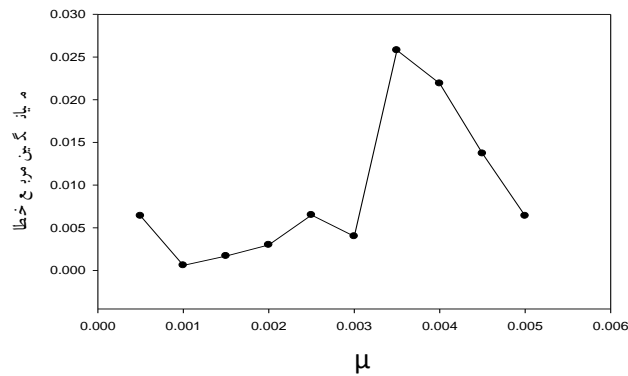
با توجه به نتایج به دست آمده، الگوریتم آموزشی لونبرگ-مارکوارت با تابع انتقال لگاریتم سیگموئیدی، تعداد نرون ۵، تعداد گره لایه‌ی پنهان ۵ و تعداد دور آموزش ۱۰۰ به عنوان شبکه بهینه برای پیش‌بینی ویسکوزیته آلکان‌های خطی انتخاب گردید.

۳-۲-۶-۱-ب- انتخاب مقدار بهینه μ

جهت بهینه کردن μ ، ساختار شبکه با ۵ متغیر ورودی، ۵ گره در لایه‌ی پنهان و الگوریتم آموزشی تنظیم لونبرگ-مارکوارت و تعداد دور ۱۰۰ طراحی گردید. سپس مقدار μ از ۰/۰۰۰۵ تا ۰/۰۰۵ با گام‌های ۰/۰۰۰۵ تغییر داده شد و آنگاه برای هر مورد مقدار میانگین مربع خطای سری پیش‌بینی محاسبه گردید. که نتایج آن در جدول (۳-۶) و شکل (۳-۵) آمده است. نتایج این بررسی نشان داد که بهترین مقدار μ برابر ۰/۰۰۱۰ می‌باشد.

جدول (۳-۶) : میانگین مربع خطا برای سری پیش‌بینی در مقادیر متفاوت μ

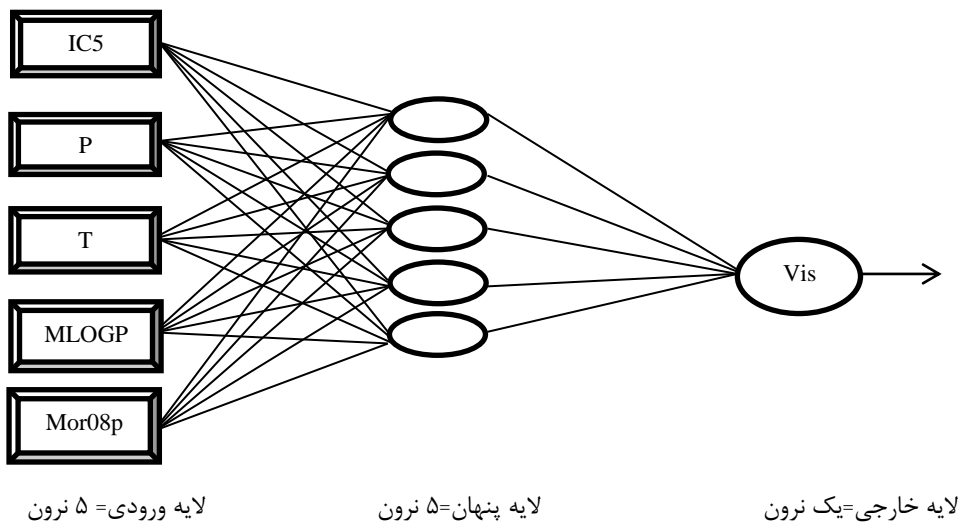
مقدار μ	میانگین مربع خطا
۰/۰۰۰۵	۰/۰۰۳۸
۰/۰۰۱۰	۰/۰۰۲۲
۰/۰۰۱۵	۰/۰۱۲۲
۰/۰۰۲۰	۰/۰۱۳۳
۰/۰۰۲۵	۰/۰۰۳۵
۰/۰۰۳۰	۰/۱۸۴۶
۰/۰۰۳۵	۰/۱۱۳۰
۰/۰۰۴۰	۰/۱۲۹۹
۰/۰۰۴۵	۰/۰۰۴۸
۰/۰۰۵۰	۰/۰۰۳۸



شکل (۳-۵) : میانگین مربع خطای سری پیش‌بینی بر حسب پارامتر μ

۳-۲-۶-۱-ج- ساختار شبکه عصبی مصنوعی بهینه شده

با توجه به نتایج به دست آمده در بهینه سازی شبکه، شبکه‌ی عصبی سه لایه‌ی پس انتشار با الگوریتم آموزشی تنظیم بایزین و تابع انتقال لگاریتم سیگموئیدی، به عنوان شبکه‌ی بهینه برای پیش‌بینی ویسکوزیته‌ی آلکان‌های خطی انتخاب شد که ساختار آن در شکل (۳-۶) نشان داده شده است.



شکل (۳-۶) : تصویر شماتیک ساختار شبکه‌ی عصبی به دست آمده پس از بهینه سازی برای پیش‌بینی ویسکوزیته‌ی آلکان‌های خطی

توابع و مقادیر پارامترهای بهینه شده‌ی شبکه‌ی عصبی برای پیش‌بینی ویسکوزیته آلکان‌ها در جدول (۳-۷) ارائه شده است.

جدول (۳-۷) : توابع و مقادیر پارامترهای بهینه شده شبکه عصبی

Train br	تابع آموزش
log sig	تابع انتقال لایه پنهان
Pure line	تابع انتقال لایه خروجی
۵	تعداد نرون لایه پنهان
۵	تعداد متغیرهای ورودی
۱۰۰	تعداد دوره‌های آموزش
۰/۰۰۱۰	پارامتر μ

۳-۲-۶-۲-مدلسازی به روش SVM

دومین روش برای مدل‌سازی غیرخطی ویسکوزیته استفاده از ماشین بردار پشتیبان می‌باشد. با توجه به توضیحات ارائه شده در (۲-۳-۳-۴-ب)، برنامه‌ای رایانه‌ای در محیط برنامه نویسی MATLAB جهت به کارگیری ماشین بردار پشتیبان نوشته شد. سپس صحت برنامه نوشته شده با استفاده از داده‌های مرجع [۴۵] مورد ارزیابی قرار گرفت. مطابقت نتایج به دست آمده از اجرای برنامه نوشته شده با نتایج به دست آمده از این مرجع درستی برنامه نوشته شده را تایید نمود. جهت استفاده از این روش از میان ۱۴ ترکیب موجود مورد مطالعه تعداد ۱۱ ترکیب جهت آموزش و ۳ ترکیب برای تست استفاده گردید. برای آموزش ماشین برداری پشتیبان باید پارامترهای آن شامل تابع حساسیت وپنایک، پارامتر موازنه‌ی C و پارامتر کرنل بهینه گردد. برای این منظور تابع حساسیت وپنایک در بازه‌ی ۰/۰۰۰۱ تا ۱۰، پارامتر موازنه‌ی C در بازه‌ی ۵۰۰ تا ۷۰۰۰۰ و پارامتر کرنل در بازه‌ی ۰/۰۰۰۰۱ تا ۰/۵ انتخاب گردید. پس از اجرای برنامه مقدار بهینه‌ی پارامترها براساس کمترین مقدار MSE برای سری تست انتخاب می‌شود. نتایج این بررسی نشان داد که ماشین برداری پشتیبان با مقدار ۱ برای تابع حساسیت وپنایک و مقادیر ۵۰۰۰۰ و ۰/۰۰۰۱ به ترتیب برای پارامتر موازنه‌ی C و پارامتر کرنل به عنوان ماشین بهینه برای پیش‌بینی ویسکوزیته‌ی آلکان‌های خطی می‌باشد که در جدول (۳-۸) این مقادیر آورده شده است. در ماشین برداری پشتیبانی که برای مطالعه‌ی حاضر استفاده شد از کرنل گوسین استفاده گردید. این کرنل در بسیاری از مسائل مورد استفاده قرار گرفته و عملکرد بالایی آن در حل این مسائل به اثبات رسیده است [۴۶].

جدول (۳-۸) : جدول بهینه سازی پارامترهای ماشین بردار پشتیبان

MSE	پارامتر کرنل	MSE	پارامتر موازنه‌ی C	MSE	تابع حساسیت وپنایک
۱/۵	۰/۰۰۰۰۱	۰/۷۲	۵۰۰	۵/۹	۰/۰۰۰۱
۱/۳	۰/۰۰۰۰۵	۰/۵۹	۱۰۰۰	۵/۱	۰/۰۰۰۵
۰/۲۳	۰/۰۰۰۰۱	۰/۳۰۲	۵۰۰۰	۴/۸	۰/۰۰۱
۰/۸۳	۰/۰۰۰۰۵	۰/۲۱	۱۰۰۰۰	۴/۵	۰/۰۰۵
۲/۷	۰/۰۰۱	۰/۱۵	۲۰۰۰۰	۴/۳۲	۰/۰۱
۱/۹	۰/۰۰۵	۰/۱۳	۳۰۰۰۰	۳/۳	۰/۰۵
۱/۴	۰/۰۱	۰/۱۲	۴۰۰۰۰	۰/۸۲	۰/۵
۱/۳	۰/۰۵	۰/۰۷۸	۵۰۰۰۰	۰/۵۹	۱
۱/۴	۰/۱	۰/۰۹۴	۶۰۰۰۰	۲/۷	۵
۱/۷	۰/۵	۰/۱۰۱	۷۰۰۰۰	۲/۸	۱۰

۳-۲-۷- بررسی نتایج

۳-۲-۷-۱- ارزیابی مدل‌های خطی و غیرخطی با استفاده از سری تست

اعتبار و اهمیت مدل‌های پیش‌بینی‌کننده وقتی مشخص می‌شود که بتواند خواص مولکول‌های مجهولی که در دسته‌ی آموزش موجود نیست را پیش‌بینی کند. بدین منظور هم از مدل منتخب خطی و هم از دو مدل غیرخطی بهینه شده توسط شبکه‌ی عصبی و ماشین بردار پشتیبان، جهت پیش‌بینی ویسکوزیته‌ی ۱۰۲ نقطه داده که در مدل‌سازی استفاده نشده‌اند، به کار گرفته شدند که نتایج مربوط به پیش‌بینی ویسکوزیته‌ی این ترکیبات در جدول (۳-۹) نشان داده شده است. همچنین در شکل (۳-۷) ضریب تعیین مقدار پیش‌بینی شده برحسب مقدار تجربی برای سری تست با استفاده از هر سه مدل بهینه شده نشان داده شده است. ضریب تعیین موجود در نمودارهای شکل (۳-۷) بیانگر این است که این روش خطی قادر به پیش‌بینی ویسکوزیته نمی‌باشد چون مقدار خطای

آن بسیار بالا و حتی در بعضی داده‌ها مقدار ویسکوزیته را منفی پیش‌بینی می‌کند که ویسکوزیته به هیچ عنوان مقدار منفی نمی‌شود. بنابراین برای پیش‌بینی ویسکوزیته باید از روش‌های غیرخطی استفاده نمود.

جدول (۹-۳): نتایج حاصل از ارزیابی مدل‌های MLR و ANN و SVM با استفاده از سری تست

شماره ترکیب	Vis						
	مقدار تجربی	مقدار پیش‌بینی شده			درصد خطا $\left(\frac{cal - Exp}{Exp} \times 100\right)$		
		MLR	ANN	SVM	MLR	ANN	SVM
۱	۰/۶۹۲۰	۰/۵۲۹۱	۰/۵۹۸۶	۰/۷۷۳۱	-۲۳/۵۴۰۵	-۱۳/۴۹۷۱	۱۱/۷۱۹۶
۲	۰/۷۹۱۰	۰/۶۳۷۳	۰/۶۹۳۱	۰/۹۰۱۵	-۱۹/۴۱۸۵	-۱۲/۳۷۶۷	۱۳/۹۸۲۳
۳	۰/۹۸۵۰	۰/۸۲۱۱	۰/۸۵۷۵	۱/۱۴۶۳	-۱۶/۶۳۹۶	-۱۲/۹۴۴۲	۱۶/۳۷۵۶
۴	۱/۲۰۶۰	۱/۰۲۱۴	۱/۰۴۴۹	۱/۴۲۶۲	-۱۵/۳۰۶۸	-۱۳/۳۵۸۲	۱۸/۲۶۷
۵	۱/۴۵۵۰	۱/۲۲۱۴	۱/۲۴۴۷	۱/۶۷۸۰	-۱۶/۰۵۵	-۱۴/۴۵۳۶	۱۵/۳۲۶۴
۶	۲/۰۵۶۰	۱/۶۱۸۲	۱/۷۰۱۸	۲/۰۲۰۷	-۲۱/۲۹۳۸	-۱۷/۲۲۷۶	-۱/۳۹۱۰
۷	۰/۸۵۲۰	۰/۶۸۳۰	۰/۷۹۴۷	۱/۰۰۸۰	-۱۹/۸۲۳۹	-۶/۷۲۵۳	۱۸/۳۲۱۶
۸	۱/۲۰۴۰	۱/۰۷۴۲	۱/۱۲۶۴	۱/۴۵۲۳	-۱۰/۷۷۲۴	-۶/۴۴۵۱	۲۰/۶۳۱۲
۹	۱/۶۶۶۰	۱/۴۴۹۴	۱/۴۹۵۷	۱/۸۰۵۶	-۱۲/۹۹۵۲	-۱۰/۲۲۲۱	۸/۳۷۹۳
۱۰	۰/۵۸۷۰	۰/۳۵۴۸	۰/۵۳۷۴	۰/۷۰۹۴	-۳۹/۵۵۷۱	-۸/۴۴۹۷	۲۰/۸۵۱۷
۱۱	۰/۶۵۵۰	۰/۴۳۷۰	۰/۶۰۱۰	۰/۷۷۱۰	-۳۳/۲۶۷۲	-۸/۲۴۴۲	۱۷/۷۲۵۱
۱۲	۰/۸۱۴۰	۰/۶۴۶۶	۰/۷۶۵۷	۰/۹۶۸۵	-۲۰/۵۵۲۸	-۵/۹۳۳۶	۱۸/۹۹۲۶
۱۳	۱/۱۸۳۰	۱/۰۴۴۱	۱/۰۹۹۵	۱/۴۱۹۸	-۱۱/۷۳۲۹	-۷/۰۵۸۳	۲۰/۰۱۶۹
۱۴	۱/۶۴۰۰	۱/۴۴۱۶	۱/۴۸۷۳	۱/۷۹۷۸	-۱۲/۰۹۷۶	-۹/۳۱۰۹	۹/۶۲۱۹
۱۵	۰/۴۴۵۰	۰/۰۵۶۳	۰/۴۵۰۶	۰/۴۷۸۴	-۸۷/۳۴۸۳	۱/۲۵۸۴	۷/۵۲۸۰

ادامه جدول (۹-۳)

١٦	٠/٤٩٦٠	٠/١٤٧٢	٠/٥٠٨٢	٠/٥٠٨٩	-٧٠/٣٢٢٦	٢/٤٥٩٦	٢/٦٠٠٨
١٧	٠/٦١٣٠	٠/٣٤٦٣	٠/٦٣٥٦	٠/٦١٣٧	-٤٣/٥٠٧٣	٣/٦٨٦٧	٠/١٣٠٥
١٨	٠/٨٧١٠	٠/٧٤٥٠	٠/٩٠٤٧	٠/٩٥٥٠	-١٤/٤٦٦١	٣/٨٦٩١	٩/٦٤٤٠
١٩	١/١٨٢٠	١/١٤١٦	١/٢١٠٩	١/٣١١٢	-٣/٤١٧٩	٢/٤٤٥٠	١٠/٩٣٠٦
٢٠	٠/٣٥٥٠	-٠/٢٤٠٠	٠/٣٨٢١	٠/٢٨٧٩	-١٦٧/٦٠٦	٧/٦٣٣٨	-١٨/٩٠١٤
٢١	٠/٤٠٦٠	-٠/١٢٥٤	٠/٤٤١٨	٠/٣١٢٨	-١٣٠/٩١١	٨/٨١٧٧	-٢٢/٩٣١
٢٢	٠/٤٩٩٠	٠/٠٧٢٥٢	٠/٥٤٥٥	٠/٣٨٢٣	-٨٥/٤٧٠٩	٩/٣١٨٦	-٢٣/٣٨٦٨
٢٣	٠/٥٩٧٠	٠/٢٧٤٨	٠/٦٥٤٥	٠/٤٩٩١	-٥٣/٩٦٩٨	٩/٦٣١٤	-١٦/٣٩٨٧
٢٤	٠/٧٠٢٠	٠/٤٦٧٨	٠/٧٦٣٦	٠/٦٣٦١	-٣٣/٣٦١٨	٨/٧٧٤٩	-٩/٣٧٣٢
٢٥	٠/٩٣٩٠	٠/٨٧١٩	١/٠١٨٨	٠/٩١٢١	-٧/١٣٥٢	٨/٤٩٨٤	-٢/٨٥٤١
٢٦	٠/٣٨٨٠	-٠/٣٠٧٣	٠/٣٦٩١	٠/٢٦١٢	-١٧٩/٢٢٧	-٤/٨٧١١	-٣٢/٦٨٠٤
٢٧	٠/٤٦٦٠	-٠/٠١٧١	٠/٥١٣٨	٠/٣٤٢٠	-١٠٣/٦٧	١٠/٢٥٧٥	-٢٦/٥٨٨
٢٨	٠/٦٥٨٠	٠/٣٨٣٧	٠/٧٢٣٧	٠/٥٧٦٣	-٤١/٦٧١٧	٩/٩٨٤٨	١٢/٤٠١٢
٢٩	٠/٨٧٨٠	٠/٧٨٤١	٠/٩٦٣١	٠/٨٢٥٦	-١٠/٦٨٣٤	٩/٦٩٢٤	-٥/٩٥٦٧
٣٠	٢/٤٣٠٠	٠/٨٨١٦	٢/٤٦٠٤	٢/٤١٤٤	-٦٣/٧٢٠٢	١/٢٥١٠	-٠/٦٣٧٨
٣١	١/٧٠٧٠	٠/٦١٤٠	١/٧٨٩٥	١/٨٤٥٦	-٦٤/٠٣٠٥	٤/٨٣٣٠	٨/١١٩٥
٣٢	٢/١٤٤٠	٠/٧٧٥٦	٢/١٧٢٢	٢/١٧٢٠	-٦٣/٨٢٤٦	١/٣١٥٢	١/٣١٠٦
٣٣	٢/٣٨٨٠	٠/٨٥٨٠	٢/٣٩٣٢	٢/٣٥٨٨	-٦٤/٠٧٠٤	٠/٢١٧٧	-١/٢١٨٥
٣٤	٣/٢٣٦٠	١/١٠٦٨	٣/٢٠٤٩	٢/٩٧٤٩	-٦٥/٧٩٧٣	-٠/٩٦١٠	-٨/٠٦٨٦
٣٥	٣/٥٦٧٠	١/١٩٧٢	٣/٥٦٩٦	٣/٢٠٤٨	-٦٦/٤٣٦٨	٠/٠٧٢٨	-١٠/١٥١٤
٣٦	١/١٥٩٠	٠/٣٢٤٤	١/٢٩٣٩	١/٢٨٨٩	-٧٢/٠١٠٤	١١/٦٣٩٣	١١/٢١٦٥
٣٧	١/٥٥٣٠	٠/٥٥٥٦	١/٦٥٧١	١/٦١٩٦	-٦٤/٢٢٤١	٦/٧٠٣١	٤/٢٩٤٩

ادامه جدول (٩-٣)

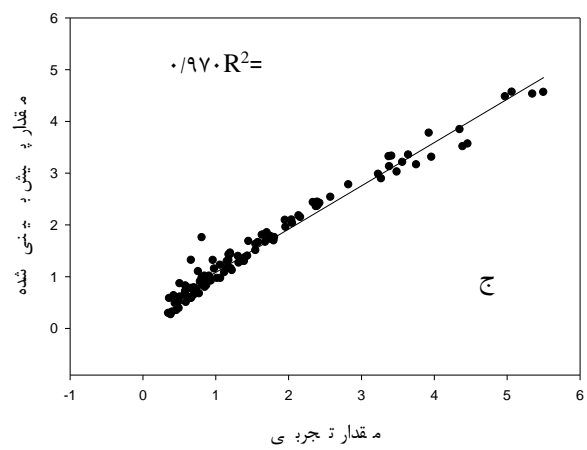
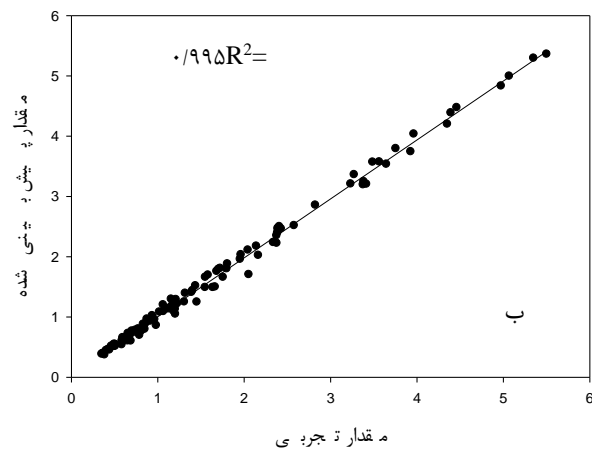
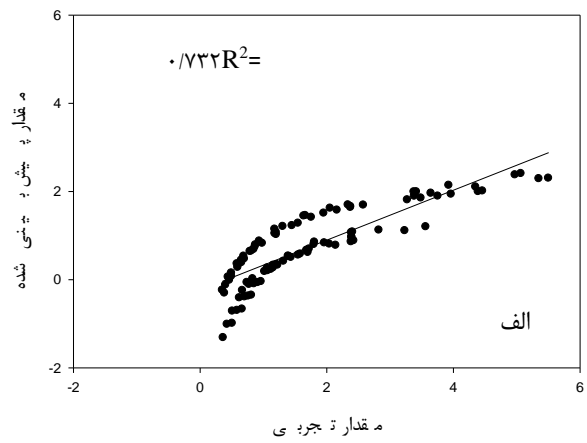
٣٨	٢/٠٤٦٠	٠/٨٠٢٨	٢/١٠٧٦	٢/٠٩١٦	-٦٠/٧٦٢٥	٣/٠١٠٧	٢/٢٣٣٦
٣٩	٢/٣٩٧٠	٠/٩٥٩٦	٢/٤٤٥٧	٢/٤٢٥٣	-٥٩/٩٦٦٦	٢/٠٣١٧	١/١٨٤٨
٤٠	٢/٨٢٧٠	١/١٢٠٤	٢/٨٥٣٩	٢/٧٧١١	-٦٠/٣٦٧٩	٠/٩٥١٥	-١/٩٧٧٣
٤١	١/٠٦٥٠	٠/٢٦٠٦	١/٢٠٠٠	١/٢١٧٨	-٧٥/٥٣٠٥	١٢/٦٧٦٠	١٤/٣٤٧٤
٤٢	١/٥٨٥٠	٠/٥٧٦٦	١/٦٩٢٤	١/٦٥٥٢	-٦٣/٦٢١٥	٦/٧٧٦٠	٤/٤٣٥٣
٤٣	٢/٣٩٨٠	٠/٩٦٥٤	٢/٤٥٨٨	٢/٤٣٧٣	-٥٩/٧٤١٥	٢/٥٣٥٤	١/٦٤٣٠
٤٤	٠/٧٤٣٠	-٠/٠٦٦٧	-٠/٧٧٧٥	-٠/٧٥٨٧	-١٠٨/٩٩١	٤/٦٤٣٣	٢/١١٣٠
٤٥	١/٣٢١٠	٠/٤١٥٦	١/٣٩١٨	١/٢٦٠٤	-٦٨/٥٣٩	٥/٣٥٩٥	-٤/٥٧٩٨
٤٦	١/٧٢٦٠	٠/٦٩٢٤	١/٨٠٤١	١/٧١٤٤	-٥٩/٨٨٤١	٤/٥٢٤٩	-٠/٦٧٢٠
٤٧	٢/٤١٠٠	١/٠٧٧٢	٢/٤٩٧١	٢/٣٧٨١	-٥٥/٣٠٢٩	٣/٦١٤١	-١/٣١٩٥
٤٨	٠/٨٣٤٢	٠/٠١٨٠	-٠/٨٧٧١	-٠/٨١٧١	-٩٧/٨٤٢٢	٥/١٤٢٦	-٢/٠٣٧٨
٤٩	١/٠٢٢٠	٠/١٨٢٨	١/٠٨٠٠	-٠/٩٦٣٤	-٨٢/١١٣٥	٥/٦٧٥١	-٥/٧٢٤٠
٥٠	١/٢١٣٠	٠/٣٣٩٦	١/٢٨٦٢	١/١٥١٧	-٧٢/٠٠٣٣	٦/٠٣٤٦	-٥/٠٤٥٣
٥١	١/٤٣٩٠	٠/٥٠٢٨	١/٥١٥٤	١/٣٩٤٦	-٦٥/٠٥٩١	٥/٣٠٩٢	-٣/٠٧٨٥
٥٢	١/٦٨٥٠	٠/٦٦٢٠	١/٧٥٥١	١/٦٦٠٩	-٦٠/٧١٢٢	٤/١٦٠٢	-١/٤٣٠٢
٥٣	١/٩٦٥٠	٠/٨٣٠٠	٢/٠٣٠١	١/٩٥١٦	-٥٧/٧٦٠٨	٣/٣١٢٩	-٠/٦٧٦٨
٥٤	٢/٣٩٣٠	١/٠٦٤٤	٢/٤٦٨٩	٢/٣٥٣٤	-٥٥/٥٢٠٣	٣/١٧١٧	-١/٦٥٤٨
٥٥	٠/٦٧١٠	-٠/٢٤٨٠	-٠/٦٨٢٣	-٠/٥٧٩٢	-١٣٦/٩٧٥	١/٦٨٤٠	-١٣/٦٦٦٢
٥٦	١/٢٢٨٠	٠/٣٢٦٣	١/٢١٦٢	١/١١٦٧	-٧٣/٤٢٨٣	-٠/٩٦٠٩	-٩/٠٦٣٥
٥٧	١/٨٠٠٠	٠/٧٩٤٣	١/٧٩٦٧	١/٦٨٩١	-٥٥/٨٧٢٢	-٠/١٨٣٣	-٦/١٦١١
٥٨	٠/٧٧٧٠	-٠/١١١٥	-٠/٧٩٩٧	-٠/٦٦٦٤	-١١٤/٣٦٣	٢/٩٢١٤	-١٤/٢٢١٤
٥٩	١/٠٦٩٠	٠/١٩٩٦	١/٠٨٦٩	-٠/٩٦٣٢	-٨١/٣٢٨٣	١/٦٧٤٤	-٩/٨٨٧٧

ادامه جدول (٩-٣)

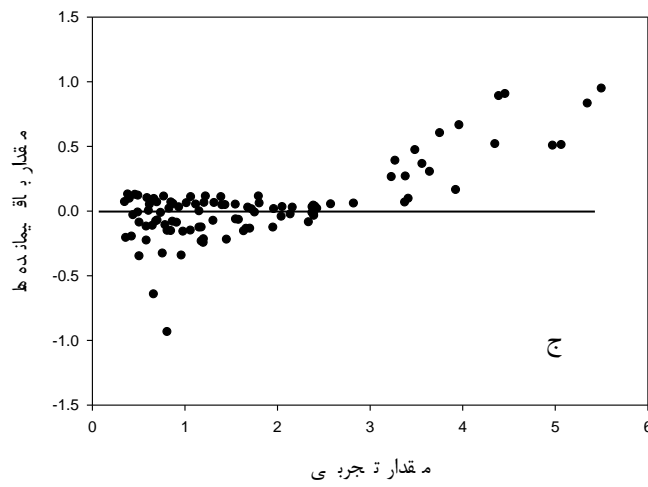
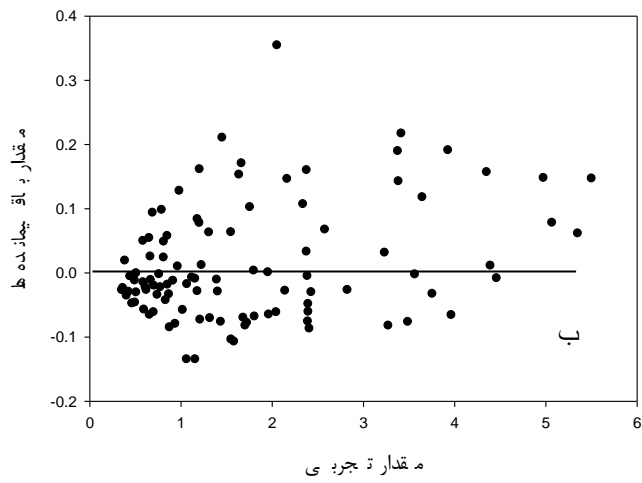
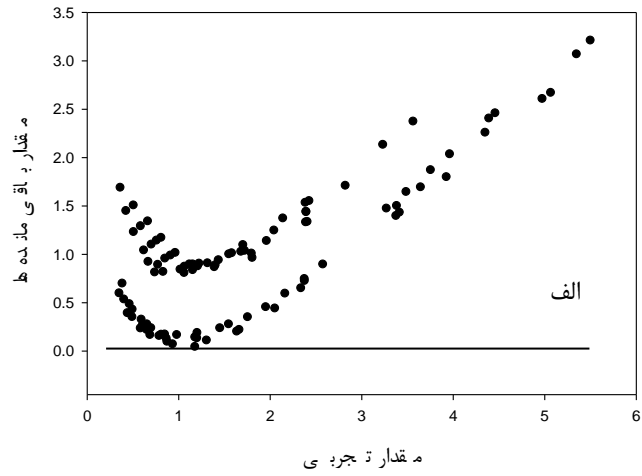
٦٠	١/٤٠٦٠	٠/٥٢١٥	١/٤٣٥٤	١/٣٦٣٠	-٦٢/٩٠١٨	٢/٠٩١٠	-٣/٠٥١٢
٦١	١/٨٠٩٠	٠/٨٤٨٧	١/٨٧٧٣	١/٧٥١٧	-٥٣/٠٧٩	٣/٧٧٥٥	-٣/١٦١٩
٦٢	٠/٦٢٣٠	-٠/٤١٥٤	٠/٦٤٩٩	٠/٥٧٦١	-١٦٦/٦٧٧	٤/٣١٧٨	-٧/٥٢٨٠
٦٣	٠/٨٥٥٠	-٠/٠٩٧٨	٠/٨٧٣٧	٠/٧٩٠٢	-١١١/٤٥	٢/١٨٧١	-٧/٥٦٧٢
٦٤	١/١٢٣٠	٠/٢٢٩٤	١/١٣٠٥	٠/٠٧٤٦	-٧٩/٥٧٢٦	٠/٦٦٧٨	-٤/٣٠٠٩
٦٥	١/٣٩٥٠	٠/٥٣٠٧	١/٤٠٥٨	١/٢٨٩٠	-٦١/٩٤٩٨	٠/٧٧٤١	-٧/٥٩١٤
٦٦	٠/٥١٢٠	-٠/٧١٦٠	٠/٥٤٢٧	٠/٦٠٣١	-٢٣٩/٨٤٤	٥/٩٩٦٠	١٧/٧٩٢٩
٦٧	٠/٧٠٦٠	-٠/٣٩٢٢	٠/٧٢٦٢	٠/٧٨٠٦	-١٥٥/٥٥٢	٢/٨٦١١	١٠/٥٨٠٧
٦٨	٠/٩١٥٠	-٠/٠٦٩٤	٠/٩٢٧٥	١/٠٠٧٠	-١٠٧/٥٩٦	١/٣٦٦١	١٠/٠٦٥٥
٦٩	١/١٥٨٠	٠/٢٦٤١	١/١٦٧٥	١/١٦١٩	-٧٧/١٩٣٤	٠/٨٢٠٣	٠/٣٤٥٤
٧٠	٠/٤٢٩٠	-١/٠١٦٢	٠/٤٥٨٧	٠/٦٢٨٣	-٣٦٦/٨٧٦	٦/٩٢٣٠	٤٦/٤٨٠١
٧١	٠/٥٨٨٠	-٠/٧٠٠٨	٠/٦٠٢٩	٠/٨١٨٠	-٢١٩/١٨٤	٢/٥٣٤٠	٣٩/١٣٢٦
٧٢	٠/٧٦٣٠	-٠/٣٧٦٧	٠/٧٦٥٤	١/٠٩٣٢	-١٤٩/٣٨٤	٠/٣١٤٥	٤٣/٢٨٩٦
٧٣	٠/٩٦٦٠	-٠/٠٤٥٩	٠/٩٥٦٣	١/٣١١٨	-١٠٤/٧٦٢	-١/٠٠٤١	٣٥/٧٩٧١
٧٤	٠/٣٦٧٠	-١/٣١٨٤	٠/٣٩٠٩	٠/٥٧٦٣	-٤٥٩/٢٣٧	٦/٥١٢٢	٥٧/٠٢٩٩
٧٥	٠/٥١٠٠	-٠/٩٩٣١	٠/٥١٠٩	٠/٨٦٢٢	-٢٩٤/٧٤٥	٠/١٧٦٤	٦٩/٠٥٨٨
٧٦	٠/٦٦٧٠	-٠/٦٧٠٦	٠/٦٤١٨	١/٣١٢٥	-٢٠٠/٥٥٥	-٣/٧٧٨١	٩٦/٧٩١٦
٧٧	٠/٨١٣٠	-٠/٣٥٤٤	٠/٧٨٩٤	١/٧٥٠٦	-١٤٣/٥٩٢	-٢/٩٠٢٨	١١٥/٣٣٨٣
٧٨	٣/٢٧٦٠	١/٨٠٦٢	٣/٣٥٨٥	٢/٨٨٨٨	-٤٤/٨٦٥٧	٢/٥١٨٣	-١١/٨١٦٢
٧٩	٣/٧٥٨٠	١/٨٩١٠	٣/٧٩٠٩	٣/١٥٧٤	-٤٩/٦٨٠٧	٠/٨٧٥٤	-١٥/٩٨١٩
٨٠	٤/٣٩٥٠	١/٩٩٣٤	٤/٣٨٤١	٣/٥٠٨٤	-٥٤/٦٤٣٩	-٠/٢٤٨٠	-٢٠/١٧٢٩
٨١	٣/٤٩٠٠	١/٨٤٨٥	٣/٥٦٦٦	٣/٠٢٠٩	-٤٧/٠٣١٥	٢/١٩٤٨	-١٣/٤٤١٣

ادامه جدول (٩-٣)

82	3/967.	1/9349	4/0.331	3/3.47	-51/2226	1/6662	-16/6927
83	4/463.	2/0.77	4/4716	3/56.2	-55/0.123	0/1926	-2/2285
84	1/956.	1/5.44	1/9554	2/0.86.	-23/0.828	-0/0.306	6/6513
85	2/581.	1/6884	2/5138	2/53.7	-34/5796	-2/6.36	-1/9488
86	3/387.	1/8892	3/2446	3/1218	-44/2191	-4/2.43	-7/8299
87	4/354.	2/0.988	4/1974	3/838.	-51/7938	-3/5966	-11/8512
88	5/50.4.	2/2972	5/3572	4/5587	-58/2613	-2/6671	-17/173
89	2/379.	1/6365	2/3463	2/3959	-31/2.64	-1/3745	0/7145
90	3/649.	1/9581	3/5315	3/3479	-46/336	-3/220.	-8/2488
91	5/353.	2/2877	5/2918	4/5235	-57/2613	-1/1432	-15/496
92	1/310.	1/2.26	1/2473	1/3879	-8/1984	-4/7862	5/9541
93	2/166.	1/5746	2/0.199	2/1414	-27/3.38	-6/7451	-1/1357
94	3/418.	1/9898	3/2.11	3/3250.	-41/7847	-6/3458	-2/720.8
95	5/0.70.	2/4.34	4/9922	4/5613	-52/5957	-1/5345	-10/0.316
96	1/759.	1/4117	1/6568	1/7725	-19/7385	-5/810.1	0/7674
97	2/380.	1/6565	2/22.2	2/3517	-30/395	-6/7142	-1/1848
98	3/380.	1/9869	3/19.6	3/3159	-41/213	-5/6.35	-1/8934
99	3/977.	2/1357	3/74.2	3/770.1	-45/6678	-4/8511	-4/0.905
100	4/977.	2/3733	4/8295	4/4733	-52/3126	-2/9636	-10/1185
101	1/551.	1/2770.	1/4880.	1/50.31	-17/6596	-4/0.619	-3/0.818
102	2/340.	1/6930.	2/2333	2/4290.	-27/6453	-4/5598	3/8.34



شکل (۳-۷): نمودار مقادیر پیش‌بینی شده ویسکوزیته برحسب مقادیر تجربی برای تست الف: مدل MLR ب: مدل ANN ج: مدل SVM



شکل (۳ - ۸) - نمودار باقیمانده‌ها بر حسب مقدار تجربی ویسکوزیته برای سری تست توسط الف: مدل MLR ب: مدل ANN ج: مدل SVM

۳-۲-۷-۲- ارزیابی مدل‌های شبکه‌ی عصبی و رگرسیون خطی چندگانه با

استفاده از پارامترهای آماری

برای بررسی میزان کارایی مدل‌های بهینه حاصل از سه روش ماشین بردار پشتیبان (SVM) و شبکه‌ی عصبی مصنوعی (ANN) و رگرسیون خطی چندگانه (MLR)، پارامترهای آماری حاصل از این سه مدل در پیش‌بینی داده‌های سری تست با هم مقایسه گردید که نتایج آن در جدول (۳-۱۰) گردآوری شده است.

جدول (۳-۱۰) - پارامترهای آماری برای سری تست مدل‌های بهینه شده توسط سه روش MLR و ANN و SVM

سری تست	R ²	PRESS	MSE	SEP
MLR	۰/۷۳۲	۱۵۰/۶۳۳۸	۱/۴۷۶۸	۰/۲۶۹۴
ANN	۰/۹۹۵	۰/۸۱۶۰	۰/۰۰۸۰	۰/۰۰۱۴
SVM	۰/۹۷۰	۸/۰۳۷۶	۰/۰۷۸۸	۰/۰۱۴۳

با توجه به پارامترهای آماری بالا، می‌توان بیان کرد که شبکه عصبی مصنوعی و ماشین بردار پشتیبان مدل‌های خوبی در پیش‌بینی ویسکوزیته می‌باشند.

۳-۲-۷-۳- ارزیابی مدل ارائه شده توسط شبکه عصبی با استفاده از آزمون Y

تصادفی

این تکنیک ارزیابی با هدف بررسی هرگونه ارتباط تصادفی بین داده‌ها انجام شد. در این آزمون مقادیر تصادفی از متغیر وابسته تولید گردید. مدل QSPR جدید با استفاده از ماتریس متغیرهای مستقل اصلی و مقادیر تصادفی از متغیر وابسته توسعه یافت. اگر مدل اصلی هیچ‌گونه ارتباط تصادفی

نداشته باشد، تفاوت قابل توجهی بین مقدار ضریب تعیین مدل اصلی و مدل QSPR که با پاسخ‌های تصادفی توسعه یافته، وجود دارد. نتایج حاصل از چندین بار اجرای آزمون Y- تصادفی در جدول (۱۱-۳) نشان داده شده است. مقادیر کوچک ضریب تعیین (R^2) بیانگر عدم وجود ارتباط شانس‌ی یا وابستگی ساختاری به سری آموزش در مدل توسعه یافته توسط شبکه می‌باشد.

جدول (۱۱-۳) : مقادیر R^2 برای سری تست با استفاده از آزمون Y- تصادفی

تکرار	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
R^2	۰/۰۲۶	۰/۰۱۳	۰/۰۱۲	۰/۰۶۶	۰/۰۷۹	۰/۰۱۰	۰/۰۲۷	۰/۰۱۵	۰/۰۱۸	۰/۰۳۰

۳-۲-۸- بررسی ارتباط توصیف‌کننده‌های منتخب با ویسکوزیته

۳-۲-۸-۱- توصیف‌کننده‌های 3D- MORSE

توصیف‌کننده‌های $3D - Morse^{106}$ (نمایش سه بعدی ساختار مولکول براساس تفرق الکترون) از

طریق معادله تبدیلی زیر که در پراش الکترون استفاده می‌شود، محاسبه می‌گردند:

$$I(s) = \sum_{i=2}^N \sum_{j=1}^{i-1} A_i A_j \frac{\sin(sr_{ij})}{sr_{ij}} \quad (۸-۳)$$

که در این معادله، I شدت الکترون پراکنده شده، A_i و A_j خاصیت اتمی اتم i و j ، s زاویه

پراکندگی، r_{ij} فاصله بین اتم‌های i و j ، N تعداد کل اتم‌ها را نشان می‌دهد. این روش باعث می‌شود

که ساختار سه بعدی مولکول به یک کد ثابت تبدیل شود.

برخی از این توصیف‌کننده‌های سه بعدی مانند $Mor08p^{107}$ در مدل برتر دیده می‌شوند. چون این

توصیف‌کننده‌ها آرایش سه بعدی اتم‌ها را بدون اینکه به اندازه مولکول ارتباط داشته باشند، بیان

می‌کنند بنابراین برای تعداد زیادی مولکول با تفاوت‌های ساختاری زیاد قابل کاربرد است [۴۸]. اثر

¹⁰⁶ 3D- Molecular Representation of Structure based on Electron diffraction

¹⁰⁷ 3D -MORSE-signal 8 /Weighted by atomic polarizabilities

متوسط منفی این توصیف‌کننده (۶/۷۵-) در مدل برتر بیان‌کننده این است که هر چه مقدار این توصیف‌کننده زیاده‌تر شود ویسکوزیته کاهش پیدا می‌کند.

۲-۸-۲-۳- توصیف‌کننده‌های MLOGP

MLOGP یا ضریب توزیع آب-اکتانول بیانگر ضریب توزیع مورینگوچی^{۱۰۸} می‌باشد [۴۹]. این توصیف‌کننده، از گروه توصیف‌کننده‌های ویژگی^{۱۰۹} است که به طور خاص، در این جا ویژگی چربی دوستی را مطرح می‌کند. چربی دوستی، بیانی از توزیع یک ترکیب بین دو فاز آلی و آبی بوده و معمولاً با علامت P نمایش داده می‌شود. P به عنوان نسبت غلظت‌های یک ترکیب در فازهای آلی و آبی یک سیستم دو جزئی تحت شرایط تعادل تعریف می‌شود:

$$P = \frac{[C]_{org}}{[C]_{aq}} \quad (۹-۳)$$

که در آن، $[C]_{org}$ و $[C]_{aq}$ به ترتیب غلظت‌های حل‌شونده در فازهای آلی و آبی هستند. ضرایب توزیع معمولاً به صورت لگاریتمی بیان می‌شوند:

$$\log P = \log \frac{[C]_{org}}{[C]_{aq}} = \log [C]_{org} - \log [C]_{aq} \quad (۱۰-۳)$$

به منظور اجتناب از تجمعات ممکن حل‌شونده در فاز آلی، ضریب توزیع باید در غلظت‌های پایین یا به صورت برون‌یابی شده در یک غلظت رقیق شده از حل‌شونده اندازه‌گیری شوند. اندازه‌گیری‌های بسیاری از تمایل نسبی مولکول، با توجه به دو فاز و بسته به پدیده‌های جذب، انتقال و توزیع انجام گرفته است. ترکیباتی که $P > 1$ یا $\log > 0$ چربی دوست یا لیپوفول نام دارند؛ و ترکیباتی که $P < 1$ یا $\log < 0$ ، هیدروفیل یا آب دوست می‌باشند. به طور خاص، چربی دوستی به اندازه‌ی مولکول حل

^{۱۰۸} - Muriguchi Octanol- water Partition Coefficient(logP)

^{۱۰۹} - Properties

شونده، قطبیت آن و اثرات پیوند هیدروژنی بستگی دارد. توصیف‌کننده‌ی مولکولی که بیشترین استفاده را در مورد این خاصیت دارد، ضریب توزیع آب-اکتانول است که با K_{OW} نشان داده می‌شود. به این ترتیب، ضریب توزیع میان ۱-اکتانول و آب به شکل زیر بیان می‌گردد:

$$\log P = \log K_{OW} = \log \frac{[C]_{1-octanol}}{[C]_{water}} = \log [C]_{1-octanol} - \log [C]_{water} \quad (11-3)$$

با توجه به رابطه بیان شده برای این توصیف‌کننده هرچه جرم مولکولی زیاد شود مقدار این توصیف‌کننده افزایش می‌یابد. و مقدار مثبت اثر متوسط این توصیف‌کننده (۲۵/۶۵۷) در مدل برتر بیان کننده این است که هر چه مقدار این توصیف‌کننده زیادتر شود خاصیت آب‌گریزی مولکول افزایش پیدا می‌کند و با افزایش خاصیت آب‌گریزی ویسکوزیته زیاد می‌شود.

۳-۲-۸-۳- توصیف‌کننده‌های IC

این توصیف‌کننده‌ها بر اساس تئوری اطلاعات شانون^{۱۱۰} شکل گرفته است و اصل تساوی اتم‌های مولکول‌های هیدروکربوری را بیان می‌کند. بر اساس این تئوری، دو اتم مساوی هستند اگر دارای عنصر یکسانی باشند. این توصیف‌کننده‌ها به صورت زیر بیان می‌شوند.

$$IC = -\sum \frac{ni}{n} \log_2 \frac{ni}{n} \quad (12-3)$$

در رابطه فوق، ni تعداد اتمها در i مین کلاس و n تعداد کل اتمها در مولکول می باشد. برخی از این توصیف‌کننده‌ها مانند $^{111}IC5$ در مدل برتر دیده می‌شوند. با توجه به رابطه بالا برای این توصیف‌کننده هرچه تعداد اتمها زیاد شود مقدار این توصیف‌کننده افزایش می‌یابد. و اثر متوسط مثبت این

^{۱۱۰} - shanon

^{۱۱۱} - Neighborhood symmetry of 5-order

توصیف‌کننده (۴/۰۵۸) در مدل برتر بیان‌کننده این است که هر چه مقدار این توصیف‌کننده زیادت‌تر شود ویسکوزیته افزایش پیدا می‌کند.

۳-۳- نتیجه‌گیری:

در این پایان‌نامه از دو روش خطی و غیرخطی برای مدل‌سازی ویسکوزیته‌ی آلکان‌های خطی از n -هگزان تا n -نونادکان در محدوده وسیعی از دما (۵۲۳/۱۵-۲۹۸/۱۲ K) و فشار (۲۰۴/۵ mPa - ۲/۴۵) استفاده گردید. نتایج به دست آمده نشان داد که مدل خطی قادر به پیش‌بینی ویسکوزیته‌ی این سری از آلکان‌های خطی نمی‌باشد. همچنین نتایج مربوط به مدل‌سازی به روش‌های غیرخطی شبکه‌ی عصبی مصنوعی و ماشین برداری پشتیبان توانایی این مدل‌ها را در پیش‌بینی ویسکوزیته تایید نمود. لازم به ذکر است که در ماشین برداری پشتیبان محدوده باقی‌مانده‌ها از ۱ تا ۱- و در شبکه عصبی این محدوده از ۳/۵ تا ۱/۵- می‌باشد. که این بیانگر کارایی خوب ماشین برداری پشتیبان است. قابل ذکر است که با روش‌های قبلی پیش‌بینی ویسکوزیته که در بخش (۱-۴) آلکان‌های خطی با متوسط خطای نسبی بالاتر از ۰/۸٪ پیش‌بینی شده است. این در حالی است که ویسکوزیته این ترکیبات با دو روش غیرخطی انجام شده در این پایان‌نامه با متوسط خطای نسبی حدود ۰/۰۷ برای ماشین برداری پشتیبان و حدود ۰/۰۰۸ برای شبکه عصبی مصنوعی پیش‌بینی گردید که این موضوع کارایی عالی این دو روش غیرخطی را تایید می‌نماید. همچنین برای استفاده از معادله‌ی انسکوگ بهبودیافته برای پیش‌بینی ویسکوزیته سیالات چگال باید مقادیر فشارحرارتی و حجم ممنوعه برای ترکیبات موردنظر موجود باشد که با توجه به دردسترس نبودن این دو کمیت برای آلکان‌های خطی چگال نمی‌توان از این معادله استفاده نمود.

آینده‌نگری:

❖ از آن جایی که در تحقیق حاضر توصیف‌کننده‌های به دست آمده از روش رگرسیون خطی چندگانه به عنوان ورودی شبکه عصبی مصنوعی مورد استفاده قرار گرفته است و چون در روشهای رگرسیون خطی چندگانه تنها برهم‌کنشهای خطی بین پارامترها لحاظ می‌شود، لذا می‌توان از روشهایی غیرخطی همچون ژنتیک الگوریتم، بهینه‌سازی اجتماع مورچگان^{۱۱۲} و الگوریتم طرح‌ریزی متوالی^{۱۱۳} بهره برد.

❖ روشهایی مانند RVM^{۱۱۴} را می‌توان به جای شبکه عصبی مصنوعی و SVM برای مدل‌سازی غیرخطی استفاده کرده و نتایج را با هم مقایسه نمود.

^{۱۱۲} - Ant Colony Optimization

^{۱۱۳} - Successive Projection Algorithm

^{۱۱۴} - Relevance vector machine

منابع

- [۱] - لوین، ا، (۱۳۸۲) "شیمی فیزیک"، ویرایش چهارم، ترجمه‌ی اسلامپور، غ، پارسافر، غ، مقاری، ع، نجفی، ب، انتشارات فاطمی، چاپ سوم.
- [۲]- Fahien, R. W., (1983) "Fundamental of transport phenomenon", Mc Grow-Hill Book Company USA.
- [۳] - اتکینز، ب، دیلو، (۱۳۷۸) "شیمی فیزیک"، ویرایش چهارم، ترجمه‌ی زینی اصفهانی، ا، خالقی، ش، خدادادی، ق، مرکز نشر دانشگاهی، چاپ اول.
- [۴]- Hirschfelder, J. O., Curtiss, C. F., Bird, R. B., (1964) "Molecular theory of gases and liquids", John Wiley and Sons, New York.
- [۵]- Najafi, B., Ghayeb, Y., Parsafar, G. A., (2000) "New correlation function for viscosity calculation of gases over wide temperature and pressure ranges", Int.J.Thermophys., 21, pp :1011-1031.
- [۶]- Ghayeb, Y., Najafi, B., Moeini, V., Parsafar, G. A., (2003/2004) "Calculation of the viscosity of supercritical fluids based on the modified Enskog theory", High Temp. High pres., 35/36, pp: 217-226.
- [۷]- Friend, D. G., Rainwater, J. C., (1984) "Transport properties of a moderately dense gas", Chem. Phys. Lett., 107, 6, pp: 590-594.
- [۸]- Rainwater, J. C., Friend, D. G., (1987) "Second viscosity and thermal conductivity virial coefficients of gases: Extension to low reduced temperatures", Phys. Rev. A, 36, 8, pp: 4062-4066.
- [۹]- Millat, J., Dymond, J., Nieto de Castro, C. A., (1996) "Transport properties of fluids", Cambridge University press, USA.
- [۱۰] - طوقانی، غ، (۱۳۸۰)، پایان‌نامه کارشناسی ارشد "ارائه توابع همبستگی جهت محاسبه‌ی ویسکوزیته گازها و مایعات در محدوده‌ای از دما و فشار"، دانشکده شیمی، دانشگاه صنعتی اصفهان.
- [۱۱]- Abas-Zade, A. K., Zhur. Ekspl. i., (1952)., Teoret. Fiz., 23, pp: 60.
- [۱۲]- Groiner, W. S., Thodos, G., (1961) "Viscosity and Thermal Conductivity of Ammonia in the Gaseous and Liquid states", J. Chem. Eng, Data, 6, pp: 240.
- [۱۳]- Schaefer, C. A., Thodos, G., J., (1959) "Thermal Conductivity of diatomic gases" AIChE. J, 5, pp: 367.
- [۱۴]- Jossi, J. A., Stiel, L. I., Thodos, G., J., (1962) "The viscosity of pure substances in the dense gaseous and liquid phase", AIChEJ, 8, pp: 59-63.

[۱۵]- Stephan, K., Hechenberger, T., (1089) “*Thermal conductivity and viscosity data of fluid mixture*”, Published by DECHMA, 10.

[۱۶]- Ely, J. F., Hanley, J. M., (1981) “*Prediction of transport properties. I- Viscosity of fluids and mixtures*”, Ind. Eng. Chem. Fundam., 20, pp: 323-332.

[۱۷]- Sastri, S. R. S., Rao, M. V. R., (1970)., Chem. Age of India, 21, pp: 675.

[۱۸]- Sastri, S. R. S., Rao, K. K., (1992) “*A new group contribution method for predicting viscosity of Organic liquids*”, Chem. Eng. J., 50, 1, pp: 9-25.

[۱۹]- Blas, F. J., (2000) “*Excess thermodynamic properties of chainlike mixture*”, J. Phys. Chem. B, 104, pp: 9239-9248.

[۲۰]- فردوسی، م.ع، (۱۳۸۹)، پایان نامه کارشناسی ارشد "پیش بینی ثابت های هنری بعضی از ترکیبات آلی با استفاده از روش های خطی و غیرخطی QSPR"، دانشکده شیمی، دانشگاه صنعتی شاهرود.

[۲۱]- Manssnat D. L., Vandeginste B. G., Deming S. N. and Kaufman L. (1998) “*Chemometrics, A Text Book*”, Elsevier, Amesterdom.

[۲۲]- یامی، ف، (۱۳۸۸)، پایان نامه کارشناسی ارشد "مدل سازی QSPR اندیس بازداري هیدروکربن های آروماتیک چند حلقه ای با روش رگرسیون خطی چندگانه و شبکه عصبی مصنوعی"، دانشکده شیمی، دانشگاه صنعتی شاهرود.

[۲۳]- <http://hassangolmohamadi.persianblog.ir>.

[۲۴]- منهاج، م، (۱۳۸۷) "مبانی شبکه های عصبی (هوش محاسباتی)" جلد اول، چاپ پنجم، مرکز نشر دانشگاه صنعتی امیرکبیر، تهران.

[۲۵]- McCulloch W.W. and Pitts W. (1943) “*A Logical Calculus of the Ideas Imminent in Nervous Activity*”, Bull. Match. Biophys., 5, pp: 115-133.

[۲۶]- Rosenblatt F. (1958) “*The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*”, Cornell Aeronautical Lab, Psychological Review, 65, 6, pp: 386-408.

[۲۷]- Hopfield J. J. (1982) “*Neural networks and physical systems with emergent collective computational abilities*”, Proceeding of the National Academy of Sciences of the USA, 79, 8, pp: 2554-2558.

[۲۸]- Rumelhart D. E. and McCielland J. L. (1986) “*Parallell Distributed Processing: Exploration in the Microstructure of Cognition*”, 1.

[۲۹]، جلد اول، دانشگاه صنعتی امیرکبیر (واحد تفرش)، چاپ اول. - منهاج، م.ب، (۱۳۷۹) "هوش محاسباتی

[۳۰]- سعیدی، م، (۱۳۸۳) "عملیات شبکه های عصبی (قسمت دوم)"، ماهنامه شبکه، شماره ۵۲.

[۳۱]- شیرپور، ا، (۱۳۸۸)، پایان‌نامه کارشناسی ارشد "کاربرد مطالعات رابطه کمی ساختار - فعالیت برای پیش‌بینی و مدل‌سازی فعالیت دارویی برخی از داروهای سنتز شده جدید" دانشکده شیمی، دانشگاه صنعتی شاهرود.

[۳۲]- Martinez-Ramon, M. (2006) "*Support Vector Machines for Antenna Array Processing and Electromagnetic*", Universidad Carlos III de Madrid, Spain, Morgan & Claypool, USA.

[۳۳]- Chen N., Yang J., Lu W., (2004) "*Support Vector Machines in chemistry*", Word scientific publishing Co. Pte. Ltd.

[۳۴]- Abe S, (2008) "*Support Vector Machines for Pattern Classification*", Kobe University, Kobe, Japan.

[۳۵]- Wang L. (2005) "*Support Vector Machines: Theory and Applications*", Nanyang Technological University, School of Electrical & Electronic Engineering, Springer Berlin Heidelberg New York.

[۳۶]- Cristianini N., Shawe-Taylor J. (2000) "*An Introduction to Support Vector Machines (and other kernel-based learning methods)*", Cambridge University Press, UK.

[۳۷]- Oliveira C. M. B. P., Wakeham W. A., (1992) "*The Viscosity of Five Liquid Hydrocarbons at pressure up to 250 MPa*" International Journal of Thermophysics, 13, 5.

[۳۸]- Assael M. J., Oliveira C. P., Papadaki M., and Wakeham W. A., (1992) "*Vibrating-Wire Viscometers for Liquids at High Pressure*", International Journal of Thermophysics, 13, 4.

[۳۹]- Assael M. J., and Papadaki M., (1991) "*Measurements of the Viscosity of n-Heptane, n-Nonane, and n-Undecane at Pressure up to 70 MPa*", International Journal of Thermophysics, 12, 5.

[۴۰]- Tanaka Y., Hosokawa H., Kubota H., and Makita T., (1991) "*Viscosity and Density of Binary Mixture of Cyclohexane with n-Octane, n-Dodecane, and n-Hexadecane Under High Pressure*", International Journal of Thermophysics, 12, 2.

[۴۱]- Dymond J. H., Robertson J., and J. Isdale D., (1981) "*Transport Properties of Nonelectrolyte Liquid Mixture -III. Viscosity Coefficients for n-Octane, n-Dodecane and Equimolar Mixture of n-Octane + n-Dodecane and n-Hexane + n-Dodecane from 250 to 100°C at Pressure Up to the Freezing Pressure or 500 MPa*", International Journal of Thermophysics, 2, 2.

[۴۲]- Caudwell D. R., Trusler J. P. M., Vesovic V., and Wakeham W. A., (2004) "*The Viscosity and Density of n-Dodecane and n-Octadecane at pressure up to 200 MPa and Temperature up to 473K*", International Journal of Thermophysics, 25, 5.

[٤٣]- Dymond J. H., Robertson J., and J. Isdale., (1980) “*Transport Properties of Nonelectrolyte Liquid Mixture –III. Viscosity Coefficients for n- Octane, n-Dodecane and Equimolar Mixture of n-Octane+ n-Dodecane and n-Hexane +nDodecane from 250 to 100°C at Pressure Up to the Freezing Pressure or 500 MPa*”, International Journal of Thermophysics, 1, 4.

[٤٤]- <http://fiz.chemie.de/infotherm/>.

[٤٥]- Chun. Y., Yuan. C., Jun. D., Bin T., (2012) “*SVM Method Used To Study Gender Differences Based On Microelement*”, International Conference on Medical Physics and Biomedical Engineering, pp: 212-215.

[٤٦]- Arun Kumar M., Gopal M. (2008) “*Application of smoothing technique on twin support vector machines, Pattern Recognition Letters 29*”, pp: 1842–1848.

[٤٧]- Hemmer M. C., Steinhauer V., Gasteiger J., (1999) ”*The prediction of the 3D structure of organic molecules from their infrared spectra*”, J. Vib. Spectrosc. , 19, pp: 151-164.

[٤٨]- Schuur J. H., Selzer P., Gasteiger J., (1996) “*The coding of the three dimensional structure of molecules by molecular transform and its application to structure- spectra correlation and studies of biological activity*”, J. Chem. Inf. Comput. Sci , 36, pp: 334-344.

[٤٩]- Hirono S., Liu Q., Nakagome I., Matushita Y., (1992) “*Simple Method of Calculating Octanol/Water Partition Coefficient*” , Chem, Pharm. Bull., 40, pp: 127-130.

[٥٠]- Basak, S. C and Magnuson, V. R., (1983) “*Molecular topology and narcosis. A Quantitative Structure –Activity Relationship (QSAR) Study of Alcohols using complementary information content*”. Arznein. Forsch, 33, pp: 501-503.

Abstract

Artificial neural network (ANN) and support vector machine (SVM) were successfully used for the modeling and prediction of viscosity of n-alkanes over a wide range of temperature and pressure. A large number of descriptors were calculated with Dragon software and a subset of calculated descriptors was selected from 18 classes of Dragon descriptors with a stepwise multiple linear regression (MLR) as a feature selection technique. Four calculated and two experimental descriptors, pressure and temperature, were selected as the most feasible descriptors. The selected descriptors were used as input nodes for generated 5-5-1 networks and support vector machine. The data set in ANN was divided into three data sets: training, prediction and test set and the data set in SVM was divided into two data set: training and test. After training and optimization of the ANN and SVM parameters, the performance of the model was investigated by the test set. The results obtained using ANN and SVM were compared with the experimental values .

Keyword: QSAR, Viscosity, Artificial neural network, Multiple Linear Regression, Support vector machine



Shahrood University of Technology

Faculty of Chemistry

**Linear and non-linear QSPR study of n-alkanes
viscosity over a wide range of temperature and pressure**

Nikoo fakhari

Supervisor:

Dr. Z. Kalantar

Advisor:

Dr. N. Godarzi

Date:Dec – 2012