
بسم الله الرحمن الرحيم

دانشکده: ریاضی

گروه: ریاضی کاربردی

مدل سازی و پیش بینی بازار بورس به کمک روش‌های ریاضی

فرهاد رامشینی

استاد راهنما:

دکتر احمد نزاکتی

اساتید مشاور:

دکتر حمید حسن پور

مهندس مجید عامری

پایان نامه جهت اخذ درجه کارشناسی ارشد ریاضی کاربردی

دیماه ۱۳۹۰

مدیریت تحصیلات تکمیلی
فرم شماره (۶)

بسمه تعالی

شماره :
تاریخ :
ویرایش :

فرم صورتجلسه دفاع از پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای **فرهاد رامشینی** رشته **ریاضی** گرایش **کاربردی** تحت عنوان **مدل سازی و پیش بینی بازار بورس به کمک روش‌های ریاضی**

که در تاریخ **۱۳۹۰/۱۰/۱۷** با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با درجه :قابل قبول امتیاز ۱۸/۵۷) ☒ دفاع مجدد ☐ مردود ☐

۱ـ عالی (۲۰ ـ ۱۹)
۲ـ بسیار خوب (۱۸/۹۹ ـ ۱۸)
۳ـ خوب (۱۷/۹۹ ـ۱۶)
۴ـ قابل قبول (۱۵/۹۹ ـ ۱۴)
۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استادراهنما	دکتر احمد نزاکتی	دانشیار	
۲ـ۱- استاد مشاور	دکتر حمید حسن پور	دانشیار	
۲-۲- استاد مشاور	مهندس مجید عامری	مربی	
۳ـ نماینده شورای تحصیلات تکمیلی	دکتر علیرضا ناظمی	استادیار	
۴ـ استاد ممتحن	دکتر داود شاهسونی	استادیار	
۵ ـ استاد ممتحن	دکتر محمد علی مولائی	استادیار	

رئیس دانشکده :

تعهد نامه

اینجانب **فرهاد رامشینی** دانشجوی دوره کارشناسی ارشد رشته **ریاضی کاربردی** دانشکده **ریاضی** دانشگاه صنعتی شاهرود نویسنده پایان نامه مدل سازی و پیش بینی بازار بورس به کمک روش های ریاضی تحت راهنمائی **دکتر احمد نزاکتی** متعهد می شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ: **امضای دانشجو:**

تقدیم به مادرم به پاس مهربانی‌هایش

روحش شاد. یادش گرامی

تقدیر و تشکر

سپاس بی‌کران خداوند عزوجل را که این وجود حقیر را شایسته تعلیم قرار داد. امید که همواره این بندهٔ حقیر را شامل لطف خویش بدارد.

با تشکر فراوان از اساتید ارجمند آقایان دکتر احمد نزاکتی، دکتر حمید حسن پور و مجید عامری که در راهنمایی این پایان نامه از هیچ کمکی دریغ نورزیده اند. امید آنکه در تمامی مراحل زندگی راهنمایم باشند.

با تشکر از دوست عزیزم دکتر رئوف غلامی که به حق کمک فراوانی در تدوین این پایان نامه نموده اند، صمیمانه از ایشان سپاسگذارم.

چکیده

امروزه یکی از مهمترین موضوعات مورد علاقه اقتصاددانان و تحلیل‌گران مالی، تبیین چگونگی و روند نوسانات قیمت هاست که راه های متفاوت و دیدگاه‌های گوناگونی را در این باره پدید آورده است. در این میان، با توجه به در دسترس نبودن اطلاعات دقیق درباره عوامل مؤثر بر نوسانات بازار سهام، پیش بینی این تغییرات به سادگی میسر نیست. تغییر درجه تأثیر متغیرها با زمان بر روی شاخص ها و یکدیگر، وجود متغیرهای سیاسی داخلی و جهانی که تأثیر مستقیم بر سیستم های اقتصادی دارد چالش اساسی در مدل سازی بازار بورس است.

طراحی ماشین هایی با توانایی فراگیری و تجربه آموزی مدت ها یکی از بحث های مهم علوم تحقیق بود تا سرانجام با ظهور کامپیوتر های قدرتمند در این راه گام های مثبتی برداشته شد. در ادامه اثبات گردید که این ماشین ها توانایی بالایی جهت فراگیری داده ها دارند و می توانند پس از فراگیری داده های آموزشی، نتایج قابل قبولی را ارائه نمایند. ماشین برداری در اوایل دهه ۹۰ معرفی شد و پس از آن به عنوان روشی نیرومند در حل مسائل طبقه بندی غیر خطی و رگرسیون مورد استفاده قرار گرفت. بر طبق نظریه محققین دلیل موفقیت این روش را می توان در قدرت بالای تعمیم دادن مسائل و توانمندی در مقابله با نویز ها و کمبود داده ها دانست.

در همین راستا هدف از مبحث حاضر، معرفی روشی جدید و قدرتمند در شناسایی الگو تحت عنوان ماشین های برداری پشتیان است که از ترکیب بهینه سازی، آموزه های آماری و هسته های کرنل بهره می جوید. پس از تحقیق در شاخص‌های بازار بورس، تغییرات قیمت سهام شرکت سرمایه گذاری البـرز در سـال‌های ۱۳۸۷ و ۱۳۸۸ به عنوان متغیر وابسته و ۱۴ متغیر اقتصادی و بازار سهام به عنوان متغیرهای مستقل تشخیص داده شدند. پس از نرمال‌سازی، مدل‌سازی به کمک ماشین‌های برداری صورت گرفت. عملکرد کرنل‌های مختلف و پارامترهای مدل نیز مورد آزمون قرار گرفتند و

مدل بهینه استخراج گردید. در انتها شبکه های عصبی رگرسیون مختصر توضیح داده شده و بر روی داده ها مدل شد که نتایج حاصل از ماشین برداری خطای کمتری داشت.

فهرست مطالب

عنوان شماره صفحه

فصل دوم: بازار بورس و بازارکارا 6

2-1- مقدمه 7

2-2- مروری بر بازارهای بورس و جایگاه بورس تهران 7

2-2-1- جایگاه و موقعیت بورس اوراق بهادار تهران در مقایسه با دیگر بورس‌های فدراسیون جهانی بورس 7

2-2-2- ارزش جاری بازار سهام 9

2-2-3- ارزش مبادلات سهام 10

2-2-4- نقش بورس در اقتصاد 10

2-2-5- روش‌های تجزیه و تحلیل بازار بورس 11

2-3- نظریه بازار کارا 13

2-3-1- گونه های سه گانهٔ کارآیی بازارسرمایه 13

2-3-2- آزمون‌های تجربی کارآیی بازار 15

2-3-3- مطالعات تجربی سطح ضعیف کارایی در بازارهای دنیا 16

2-3-4- مطالعات تجربی سطح ضعیف کارایی در بورس اوراق بهادار تهران 18

فصل سوم: ماشین های برداری پشتیبان 20

3-1- مقدمه 21

3-2- آموزش الگو 23

3-2-1- آموزش با راهنما 23

3-2-2- آموزش بدون راهنما 24

3-2-3- آموزش وراثتی 25

3-2-3-1- ظرفیت 25

3-2-3-2- پیچدگی نمونه ها 25

3-3-2-3- پیچیدگی محاسبات ***25***

۳-۳- تئوری یادگیری آماری ۲۶

۳-۴- جداکننده های خطی ۳۰

3-5- ابر صفحه جداساز[1] ***30***

3-6- ماشین برداری پشتیبان خطی ***32***

3-7- ضریب لاگرانژ ***34***

3-8- حل مسئله ماشین های برداری پشتیبان خطی ***39***

3-9- ماشین های برداری پشتیبان در سیستم های خطی جدا ناپذیر ***41***

3-10- آنالیز رگرسیون ***43***

3-10-1- مراحل مختلف آنالیز رگرسیون ***44***

3-10-1-1- بیان مسئله ***44***

3-10-1-2- انتخاب درست متغیرها ***45***

3-10-1-3- جمع آوری داده ***45***

3-10-1-4- ویژگی های مدل ***46***

3-10-1-5- روش های برازش ***47***

3-10-1-6- برازش مدل ***47***

3-10-1-7- تعیین اعتبار و مشخص نمودن نقص های مدل ***48***

3-10-1-8- اهداف آنالیز رگرسیون ***49***

3-11- آنالیز رگرسیون خطی ***49***

3-12- ماشین های برداری پشتیبان برای رگرسیون خطی[2] ***51***

3-12-1- فرموله کردن رگرسیونر خطی ***52***

3-12-2- بهینه سازی کاربردی ماشین برداری رگرسیونگر ***54***

3-13- فضای ویژگی ***55***

3-14- ماشین های برداری غیر خطی ***57***

3-15- حقه های کرنل[3] ***59***

3-16- آنالیز رگرسیون غیر خطی ***60***

3-17- ماشین برداری رگرسیونگر برای سیستم های غیر خطی ***61***

[1] -Hyperplane
[2] - SVR
[3] -Kernel Trick

3-18- قدم به قدم برای طبقه بندی با ماشین های برداری پشتیبان **62**

3-19- قدرتمندی ها و ضعف ماشین های برداری پشتیبان **63**

فصل چهارم: مدل سازی داده‌ها با استفاده از ماشین برداری پشتیبان **64**

4-1- مقدمه **65**

4-2- الگوریتم SMO **65**

4-3- کرنل عمومی پیرسن **67**

4-4- پیاده سازی ماشین برداری پشتیبان **69**

فصل پنجم: شبکه عصبی رگرسیون عمومی **72**

5-1- مقدمه **73**

5-2- قابلیت های شبکه های عصبی مصنوعی **73**

5-3- مسائل مناسب جهت استفاده از شبکه های مصنوعی **74**

5-4- تئوری شبکه های عصبی رگرسیون عمومی **75**

5-5- پیاده سازی شبکه عصبی رگرسیون عمومی **79**

فصل ششم: نتیجه‌گیری و پیشنهادات **81**

6-1- مقدمه **82**

6-2- نتیجه گیری **82**

6-3- پیشنهادات **83**

مراجع **85**

فصل اول

کلیات

۱-۱- مقدمه

ناشناخته بودن دینامیک عوامل تأثیر گذار بر تغییرات قیمت سهام، نبـود مـدل‌هـای کامـل بـرای توصیف هر یک از این عوامل و اثرات متقابل آن‌ها بر یکدیگر، دلیلی برای روی آوردن به پـیش بینـی تغییرات قیمت سهام است. پیش بینی قیمت یا بازده سهام به کمک کشف الگوهـای رفتـاری فرآینـد مولد قیمت سهام امکان پذیر است. میزان موفقیت در کشف این گونه الگوهای رفتاری، میزان کـارایی پیش بینی را مشخص می‌کند. در واقع، فرآیند مولد قیمت سهام را می توان به عنوان یک الگوی پویـا بررسی کرد. فرآیند مزبور ممکن است به صورت مدل‌های خطی، غیرخطی و یا تصادفی به دست آید.

از اواسط دهه ۷۰ و به ویژه از سال ۱۹۸۰ تلاش‌های گسترده‌ای در زمینه پیش‌بینی قیمت سهام با استفاده از روش‌های ریاضی جدید، سریهای زمانی و ابزار پیشرفته‌تری مثل هوش مصنوعی آغاز شد و آزمون‌های زیادی بر روی اطلاعات قیمت و شاخص سهام در کشورهایی مثل انگلستان، آمریکا، کانادا، آلمان و ژاپن صورت گرفت تا وجود یا فقدان ساختاری معین در اطلاعات قیمت سهام نشان داده شود. در همین دوران نظریه آشوب[1] توسط ریاضی‌دانانی چون ادوارد لـورنز و جیمـز یـورک[2] شـکل‌گرفـت. طرفـداران این نظریه بر این بـاورند کـه در میـان الگـــوهای ظـاهراً تصـادفی پدیـده‌هـای مختلـف از سیستم‌های هواشناسی گرفته تا سازمان‌ها و بازارهای بورس نوعی نظم وجود دارد [۱].

تجربه ثابت کرده، داده‌هایی که از آزمایشات گوناگون به‌دست می‌آیند محدود و نمونه‌وار می‌باشـند و فقط درون فضای ورودی و با توزیع پراکنده شکل می‌گیرند. دست‌یابی شبکه‌های عصـبی بـه تولیـد مدل‌هایی که بتواند روی داده‌ها منطبق شود مشکلات زیادی را تحمل کرده است. ماشین‌هـای بـردار پشتیبان[3] روش مؤثری برای مدل سازش داده‌ها می‌باشند. آنها با افزایش ابعاد مسئله و بـا اسـتفاده از

[1] - chaos theory
[2] - Edward Lorenz & James Yorks
[3] -SVM

نگاشت کرنل یک چهارچوب کاری یکپارچه را برای اکثر مدل‌ها فراهم می‌کنند و امکان مقایسه را به وجود می‌آورند [۲].

۱-۲- سابقه و ضرورت انجام تحقیق

روش ماشین برداری پشتیبان در اوایل سال ۱۹۹۰ جهت انجام فرآیندهای طبقه بندی دو کلاسه توسط ولادمیر وپنایک[1] معرفی شد [۳]. پس از آن محققین بسیاری به توسعه الگوریتم‌های این ماشین پرداخته اند [**2**].

افراد زیادی الگوریتم‌های این روش را در فرآیندهای طبقه بندی و رگرسیون با موفقیت به کار گرفتند از جمله: به کارگیری الگوریتم ماشین برداری در طراحی دارو، آنالیز داده‌های ژن انسان، بررسی وضعیت بازار و پیش بینی نرخ اعتبار، پیش بینی نرخ بارش سالانه، پیش بینی میزان ذخیره زغال سنگ و ... اشاره نمود [۴]. تمامی موارد بیان شده به خوبی توانستند بر قابلیت ماشین برداری پشتیبان صحه بگذارند و با مقایسه نمودن نتایج خود با نتایج روش‌های دیگر، قدرت این ماشین را مورد تأیید قرار دادند.

علاوه بر موارد فوق، دنگ چیو و پینگ چن[2] از ماشین‌های برداری جهت پیش‌بینی بازار بورس استفاده کردند و تا ۷۹ درصد به درستی توانستند روند بازار را پیش‌بینی کنند [۵]. هوانگ ناکاموری و وانگ[3] نیز برای پیش بینی دینامیک بازار بورس توسط ۹ شاخص کلان اقتصادی از ماشین‌های برداری استفاده کردند [۶].

[1] -Vladimir Vapnik
[2] - Deng-Yiv Chiu, Ping-Jie Chen
[3] - Huang, Nakamori, and Wang

به هر حال تحقیقات زیادی در رابطه با قدرت پیش‌بینی ماشین‌های برداری در بازار بورس صورت نگرفته است. بررسی منابع منتشره نشان می دهد که تاکنون از روش ماشین برداری جهت پیش بینی بازار بورس در ایران استفاده نشده است.

در به کار گیری این روش در پیش‌بینی بازار بورس سوالات زیر مد نظر می‌باشند:

۱- چگونه می‌توان از روش ماشین‌برداری به عنوان یک روش مناسب‌تر در پیش بینی متغیرهای تأثیرگذار بر روی سهام مورد نظر استفاده کرد؟

۲- آیا این روش می‌تواند روند سهام را پیش‌بینی کند؟

۱-۳- هدف از مطالعه

هدف از مطالعه حاضر، مدل‌سازی بازار بورس تهران با استفاده از روش ماشین بـرداری پشـتیبان، تحقیق در دقت و پیش‌بینی روند قیمت سهام می‌باشد. در ایـن سـیر بـه معرفـی بـازار بـورس، روش ماشین‌های برداری، الگوریتم های وابسته به مدل، نرم افزارهای مربوطه، چگونگی استفاده از داده‌ها و فرمت بندی داده‌ها بحث می‌کنیم.

۱-۴- تعریف متغیرها

پس از جمع آوری و تهیه کلیه اطلاعات مربوط به سهام بازار بورس تهران نسبت به انتخاب سـهام شرکت سرمایه‌گذاری البرز اقدام گردید. پس از مطالعه و تحقیق نسبت بـه متغیرهـای تـأثیرگـذار بـر سهام شرکت البرز، متغیرهای زیر که شامل قیمت نفت، قیمت طـلا، نـرخ ارز، تـورم، شـاخص قیمـت کالاهای صادراتی، شاخص قیمت تولیدکننده، و ۸ شاخص بازار شامل شاخص کل، صـنعت، وابسـتگی پولی و مالی، بازار اول، بازار دوم، قیمت، ۵۰ شرکت برتر و بـازده نقـدی بـورس مـی‌باشـد بـه عنـوان متغیرهای مستقل شناسایی شد. داده‌های مـورد مطالعـه مربـوط بـه مشـاهدات دوره زمـانی ۱۳۸۷ و

۱۳۸۸ می‌باشد و شامل ۷۳۰ داده می‌باشد. لازم به توضیح است که ۳ متغیر کلان اقتصادی (قیمت نفت، قیمت طلا ونرخ ارز) و تورم و شاخص قیمت کالاهای صادراتی و همچنین شاخص قیمت تولیدکننده از بانک اطلاعاتیِ بانک مرکزی جمهوری اسلامی ایران گردآوری شده و ۸ شاخص بازار بورس از کتابخانه مرکزی سازمان بورس و اوراق بهادار تهران اخذ شده است.

پس از استخراج و گردآوری کلیه‌ی نمونه‌ها و متغیرهای مورد نیاز از منابع موثق، و نرمال سازی آن‌ها اقدام به پیاده سازی مدل گردید، که این مرحله با استفاده از نرم افزارهای داده‌کاوی صورت گرفت و از نرم افزار متلب[1] جهت فیت کردن مدل و رسم نمودارها کمک گرفته شد. داده های مربوط به سال ۱۳۸۷ جهت آموزش مدل و داده‌های مربوط به سال ۱۳۸۸ جهت تست مدل مورد استفاده قرار گرفتند.

۱-۵- سازماندهی پایان نامه

با توجه به مطالب بیان شده، در پایان‌نامه حاضر به بررسی عملکرد روش پیش‌بینی ماشین برداری پشتیبان در بازار بورس تهران پرداخته خواهد شد. در فصل دوم به معرفی، اهمیت و نقش بازار بورس در اقتصاد ملی، جایگاه بورس تهران نسبت به بازارهای جهانی، روش‌های تجزیه و تحلیل بازار بورس، نظریه بازار کارا، نظریه آشوب و پیشینه‌ی تحقیقات انجام شده بر بازار بورس تهران و دیگر بورس‌ها خواهیم پرداخت. در فصل سوم به معرفی ماشین های برداری پرداخته می‌شود. در فصل چهارم مدل‌سازی داده‌ها، پیاده سازی روش، معرفی الگوریتم ها و کرنل ها گنجانده شده است. تئوری شبکه عصبی رگرسیون عمومی و مقایسه با ماشین های برداری در فصل پنجم مورد تجزیه و تحلیل قرار می گیرد. نتایج حاصله و پیشنهادات در فصل ششم ذکر می گردد.

[1] -Matlab

فصل دوم

بازار بورس و بازار کارا

۲-۱- مقدمه

به منظور آشنایی با بازار بورس، لزوم تحقیق، چالش های اساسی و قابلیت پیش بینی پـذیری، بـر آن شدیم تا مروری بر این سیستم اقتصادی داشته باشیم. لذا در این فصل به طور مختصر به مطالعه و تحقیق راجع به بازار بورس، جایگاه بورس تهران در فدراسیون جهانی بورس، شاخص‌های بازار، روش-های تجزیه و تحلیل بازار بورس، نظریه بازار کارا، و آزمون‌های کارایی بورس ایران و سـایر کشـورهای عضو بورس خواهیم پرداخت.

۲-۲- مروری بر بازارهای بورس و جایگاه بورس تهران

با شناخت سایر بورس‌های جهان و جایگاه بورس تهران به لزوم تحقیقات بیشتر بر روی این سیستم، علی الخصوص بورس ایران پی می‌بریم.

۲-۲-۱- جایگـاه و موقعیت بـورس اوراق بهادار تهران در مقایسه با دیگر بورس‌های عضو فدراسیون جهانی بورس

هر چند سابقه تأسیس بورس های اوراق بهادار در اکثر کشورهایی کـه در حـال حاضـر از آنـان بـا عنوان کشورهای صنعتی نام برده می شود، به حدود ۳۰۰ سال پیش باز می گردد، ولی با این وجـود، میزان رشد سازوکار اوراق بهادار در کلیـه کشـورها یکسـان نبـوده اسـت. کشـورهای آمریکـا، ژاپـن و انگلستان سمبل کشورهایی هستند که از سازوکار اوراق بهادار در فرآیند توسعه اقتصـادی خـود سـود فراوان برده اند. شاهد مدعا اینکه؛ در حال حاضر کشورهای یاد شـده ۶۰ درصـد ارزش جـاری سـهام تمامی بورس های جهان را به خود اختصاص داده اند. کشورهای در حال توسعه که عملاً پس از پایان جنگ جهانی دوم فرآیند توسعه‌ی اقتصادی خود را آغاز کرده انـد؛ تکیـه بیشـتری بـر سـازوکار اوراق

بهادار داشته اند. شاهد مدعا اینکه؛ در حال حاضر، حجم مبادلات بورس اوراق بهادار در کشورهای کره جنوبی، مالزی و ترکیه، به ترتیب برابر با ۵۰، ۸۵/۵ و ۳۳ درصد تولید ناخالص داخلی می باشد. همان گونه که مشاهده می‌گردد؛ حجم معاملات بورس اوراق بهادار در کشورهای یاد شده، حتـی از پـاره ای کشورهای پیشرفته صنعتی نیز بیشتر می‌باشد. در ایران، علی‌رغم آغاز فعالیت بورس اوراق بهادار در ۳ دهه‌ی پیش، سازوکار اوراق بهادا نتوانسته است، نقش قابل ملاحظه‌ای در ساختار بازار سـرمایه کشـور پیدا کند. به طوری که، شاخص‌های اساسی مانند: نسبت حجـم معـاملات بـه تولیـد ناخـالص ملـی و تشکیـل سرمایه، از طریق سازوکار اوراق بهادار حجم بسیـار پایینی را نشان می‌دهند. البته، به موازات کاستی‌های یاد شده، یادآوری این واقعیت ضروری است که پـس از تجدیـد فعالیـت بـورس در سـال ۱۳۶۸، سهم سازوکار اوراق بهادار در ساختار اقتصادی کشور روند رو به رشدی داشته است؛ به طوری که، می‌توان انتظار داشت در صورت ادامه روند فعلی، بورس تهران بتواند در آینـده‌ی نزدیـک جایگـاه ویژه‌ای برای خود در سطح خاورمیانه بیابد. فدراسیون جهانی بورس اوراق بهادار شامل ۴۹ بازار بورس اوراق بهادار می‌باشد. به منظور نشان دادن جایگاه بـورس تهـران در فدراسـیون جهـانی بـورس اوراق بهادار[1]، سنجشی در بعضی متغیرها به شرح زیر صورت پذیرفته است:

تعداد شرکت‌های عضو، ارزش جاری بازار سـهام، ارزش مبـادلات سـهام، تعـداد روزهـای فعالیـت، متوسط ارزش مبادلات روزانه، نسبت فعالیت، شاخص قیمت سهام، نسبت درآمـد بـه قیمـت، نسـبت بازده سهام و نقش بورس در اقتصاد ملی. بازار بورس تهران در اکثر طبقه بندیهای فوق رتبه‌ی چهلـم به بعد را دارد که در زیر فقط به ذکر چند نمونه بسنده می‌کنیم.

[1] - FIBV

۲-۲-۲- ارزش جاری بازار سهام

ارزش جاری سهام، ارزش بازاری شرکت‌های پذیرفته شده در یک بورس را نشان می‌دهد.

جدول ۲-۱: ده بورس برتر جهان در سال ۲۰۰۸ از نظر ارزش بازار و جایگاه بورس تهران

رتبه بورس‌ها	نام بازار	پایان سال ۲۰۰۸ (میلیارد ریال)	پایان سال ۲۰۰۷ (میلیارد ریال)	تغییرات (درصد)
۱	گروه بورس نیویورک	۹۲۰۹	۱۵۶۵۱	-۴۱/۲
۲	گروه بورس توکیو	۳۱۱۶	۴۳۳۱	-۲۸/۱
۳	بورس نزدک	۲۳۹۶	۴۰۱۴	-۴۰/۳
۴	گروه یورونکست	۲۱۰۲	۴۲۲۳	-۵۰/۲
۵	بورس اوراق بهادار لندن	۱۸۶۸	۳۸۵۲	-۵۱/۵
۶	بورس شانگهای	۱۴۲۵	۳۶۹۴	-۶۱/۴
۷	بورس هنگ کنگ	۱۳۲۹	۲۶۵۴	-۴۹/۹
۸	بورس آلمان	۱۱۱۱	۲۱۰۵	-۴۷/۲
۹	گروه بورس تورنتو	۱۰۳۳	۲۱۸۷	-۵۲/۷
۱۰	گروه بورس اسپانیا	۹۴۸	۱۷۸۱	-۴۶/۸
۴۲	بورس اوراق بهادار تهران	۴۸	۴۳	+۱۱

۲-۲-۳- ارزش مبادلات سهام

بالا بودن تعداد شرکت‌های عضو و ارزش جاری سهام، به خودی خود نمی‌تواند امتیازی ارزشمند برای بازار سرمایه یک کشور محسوب شود؛ بلکه گردش نقدینگی در چهارچوب خرید و فروش اوراق بهادار و داشتن فعالیتی متناسب با ابعاد بازار، به مراتب اهمیت بیشتری دارد.

جدول ۲-۲: ده بورس برتر جهان در سال ۲۰۰۸ از نظر ارزش سهام معامله شده و جایگاه بورس تهران

رتبه بورس‌ها	نام بازار	سال ۲۰۰۸ (میلیارد ریال)	سال ۲۰۰۷ (میلیارد ریال)	تغییرات (درصد)
۱	بورس نزدک	۳۶۴۴۶	۲۸۱۱۶	۲۹/۶
۲	گروه بورس نیویورک	۳۳۶۳۹	۲۹۲۱۰	۱۵/۲
۳	بورس اوراق بهادار لندن	۶۴۷۴	۱۰۳۲۴	-۳۷/۳
۴	گروه بورس توکیو	۵۵۸۶	۶۴۷۶	-۱۳/۷
۵	گروه یورونکست	۴۴۵۴	۵۶۴۸	-۲۱/۱
۶	بورس آلمان	۳۸۸۱	۴۳۲۴	-۱۰/۲
۷	بورس شانگهای	۲۵۸۷	۴۰۷۰	-۳۶/۴
۸	گروه بورس اسپانیا	۲۴۳۹	۲۹۷۱	-۱۷/۹
۹	گروه بورس تورنتو	۱۰۳۳	۲۱۸۷	-۵۲/۷
۱۰	بورس هنگ کنگ	۱۷۳۶	۱۶۴۹	۵/۳
۴۲	بورس اوراق بهادار تهران	۱۵	۸	۸۵

۲-۲-۴- نقش بورس در اقتصاد ملی

به منظور ارزیابی نقش بورس در اقتصاد ملی، مقایسه‌ای بین متغیرهایی از بورس اوراق بهادار و اقتصاد کلان در بین اعضای فدراسیون بورس‌های جهانی صورت گرفته است. در واقع، ارزش جاری بازار سهام نسبت به تولید ناخالص داخلی مقایسه شده است، مأخذ کلیهٔ آمار و ارقام جدول شمارهٔ (۲-۳)، صندوق بین‌المللی پول و بانک جهانی بوده است.

جدول ۲-۳: مقایسه ارزش جاری بازار سهام نسبت به تولید ناخالص داخلی

بورس	۲۰۰۸		
	ارزش بازار (میلیون دلار)	تولید ناخالص داخلی (میلیون دلار)	درصد
بورس هنگ کنگ	۱,۳۲۸,۷۶۸	۲۲۳,۸۰۰	۵۹۴
بورس سنگاپور	۲۶۴,۹۷۴	۱۹۲,۸۰۰	۱۳۷
بورس توکیو	۳,۱۱۵,۸۰۳	۴,۸۴۴,۰۰۰	۶۴
بورس نیویورک	۹,۲۰۸,۹۳۴	۱۴,۳۳۰,۰۰۰	۶۴
بورس تهران	۳۸۲,۰۰۰	۴۸,۷۱۳	۷۸۴

۲-۲-۵- روش‌های تجزیه و تحلیل بازار بورس

از آنجا که اطلاع از روند قیمت‌ها در سیستم‌های اقتصادی منجر به سودهای کلانی می‌شود، تحقیقات زیادی بر روی این بازارها علی الخصوص بازار بورس انجام شده است. روشهای زیادی به کار برده شده است که به دو گروه: نمودارگراها[1] و بنیادگراها[2] تقسیم بندی شده است. نمودارگراها معتقدند که امکان محاسبه ارزش ذاتی سهام و اوراق قرضه وجود ندارد، و باید انتخاب سهام را بر اساس نمودار انجام دهیم. به نظر آنها، بازار دستخوش حالت‌هایی شبه روانی می‌شود. همچنین، آنان معتقدند که با بررسی روند گذشته‌ی قیمت ها، می‌توان روند آینده را پیش بینی کرد. فردی خبره می‌تواند این روند را تشخیص دهد. استادان مالی دانشگاه ها که مُبلغ بررسی‌های بنیادین هستند، به این گروه ایراد می‌گیرند و ادعا می‌کنند که آنان فالگیران حرفه‌ای هستند و معتقدند که هر سهم ارزش ذاتی دارد، و برای تعیین ارزش آن، باید به مطالعه‌ی عمیق و بنیادین شرکت و کل اقتصاد – به اتکای تمامی اطلاعات موجود– پرداخت. نمودارگراها در برابر نظریات و استدلال‌های بنیادگراها موضعی دفاعی‌تر گرفته و می‌گویند: هدف ما تعیین تغییرات بلند مدت ارزش سهام نیست، لیکن باید

[1]- Chartists
[2] - Fundamentalists

به دنبال استفاده از فرصت‌های کوتاه مدت بوده، و سودهای آنی و فوری را دنبال کرد. در تحلیل تکنیکی (روش نمودارگراها) فقط از نمودار قیمت‌ها، حجم معاملات و مقادیر محاسبه شده از قیمت‌ها استفاده می‌شود در حالی که در تحلیل بنیادی از اطلاعات بسیار وسیعی مانند اطلاعات درون شرکتی و اطلاعات برون شرکتی (صادرات و واردات کالاها، تعرفه‌های گمرکی، نرخ سود بانکی، تورم، نرخ ارز، رشد اقتصادی، تحولات سیاسی، قیمت نفت و...) استفاده می‌کند.

مطالعات آکادمیک نشان داده است که با احتساب هزینه‌های کمیسیون خرید و فروش، سود کسی که از نمودار استفاده می‌کند، بیش از کسی نیست که بر حسب تصادف سهام را می‌خرد. البته، هر دو گروه معتقدند که بازار تغییر قیمت دارد و مقداری از این تغییر، لحظه‌ای است و حاصل عدم تعادل بین عرضه و تقاضا در لحظه‌ای خاص می‌باشد و به ارزش بلندمدت سهام ربط ندارد. بقیه‌ی تغییرات، میان مدت و بلند مدت است. نمودارگراها می‌گویند که می‌توان منحنی بالا و پایین رفتن قیمت سهام را ترسیم کرد، و یک راه پولدار شدن این است که معین کنیم در چه زمان باید سهام را خرید و در چه زمانی آن را فروخت. بنیادگراها می‌گویند: مشکل این است که مشخص نیست چه وقت قیمت، پایین‌ترین (برای خرید) و چه وقت بالاترین (برای فروش) است. آنها می‌گویند که میزان بازدهی به‌دست آمده، به میزان پذیرش خطر مربوط می‌شود. آنها نمی‌گویند که نمی‌شود در بورس میلیونر شد، اما می‌گویند که میلیونر شدن از روی نمودار امکان ندارد، چنین کاری در بازار کارا فقط به شانس افراد بستگی دارد. در بازار کارا نمی‌توان به‌طور مداوم و پیوسته، بازدهی بهتر از بازدهی بازار به دست آورد. در کوتاه‌مدت این امر امکان دارد، اما آن هم صرفاً به شانس – و نه به خبرگی و نمودار– بستگی دارد [۷], [۸].

۲-۳- نظریه بازار کارا

اینکه اصلاً آیا بازار بورس قابلیت پیش بینی دارد یا نه، موضوعی است که سال‌ها مورد تحقیق قرار گرفته و در نهایت تحت عنوان نظریه بازارهای کارا مطرح گردید. در ادامه‌ی این فصل به معرفی این نظریه و کارا بودن بازارهای بورس دنیا و بورس تهران که نقش اساسی در تحقیق دارد اشاره می‌کنیم.

بازاری با سیستمی که اجازه می‌دهد تا اجرای معاملات مشتریان به سرعت صورت گیرد، از لحاظ عملیاتی کارا خوانده می‌شود. اگر قیمت‌ها در بازاری پس از رسیدن به شوک‌های مهم، خود را به سرعت تعدیل کنند، بدان بازاری کارا از لحاظ اطلاعاتی گفته می‌شود. از سال ۱۹۶۵، در حوزهٔ مالی، فرضیه‌ای در مورد کارا بودن بازار سرمایه مطرح و بر این اساس، تحقیقات متعددی صورت گرفته است. علت انجام این تحقیقات آزمودن این نظریه است که آیا بازار سهام در کسب و پردازش اطلاعات ورودی به طور معقول عمل می‌کند، و آیا اطلاعات بدون درنگ و بدون تمایل و گرایش خاص در قیمت اوراق بهادار منعکس می شود یا خیر؟

فرضیهٔ بازارهای کارا نمی‌گوید که همهٔ سرمایه گذاران معقول عمل می‌کنند، و یا تک‌تک آنان به جزءجزء اطلاعات دسترسی دارند؛ بلکه صرفاً می‌گوید که همهٔ اطلاعات در قیمت سهام بازتاب دارد [۷]. این نظریه از دیدگاه عمومیت و گستردگیِ دامنهٔ مفروضات و پیامدهای آن، در سه سطح مطرح است که ذیلاً به آنها پرداخته می‌شود.

۲-۳-۱- گونه‌های سه گانهٔ کارآیی بازارسرمایه

مسألهٔ کارآیی، مقوله‌ای صرفاً سیاه یا سفید نیست. بازار نه کاملاً کارآیی دارد و نه یکسره از مفاهیم کارآیی به کنار است. مسأله، مراتب امر و درجهٔ کارآیی است. سؤال این است که بازار تا چه اندازه کارا است؟ یکی از معیارهای سنجش کارآیی بازار، این است که بپرسیم چه نوع اطلاعاتی در قیمت اوراق

بهادار منعکس است و تأثیر دارد؟ در سال ۱۹۷۰، فاما[1] پیشنهاد تقسیم این فرضیه را به سه بخش بازارهای کارای ضعیف (شامل اطلاعات بازار)، بازارهای کارای نیمه قوی (شامل اطلاعات عمومی) و بازارهای کارای قوی (شامل تمام اطلاعات) داد.

اگر شکل ضعیف فرضیه بازار کارا مطرح باشد، تحلیل فنی[2] یا تحلیل نموداری قیمت سهام بی‌اثر می‌شود. شکل ضعیف کارایی بازار می‌گوید که با مطالعهٔ روند تاریخی قیمت سهام، قادر به پیش بینی روند آینده قیمت سهام نمی‌توان بود و قیمت سهام روند خاصی ندارد. به عبارت دیگر، بازار سهام حافظه‌ای ندارد، و قیمت سهام در بازار کارا به شکل تصادفی تغییر می‌کند. این همان نظریه گشت تصادفی یا گردش تصادفی[3] است. اگر این نظریه صحت داشته باشد، آن دسته از سرمایه‌گذاران که بر اساس نمودار سهام را انتخاب می‌کنند، راه خطا می‌روند.

اگر شکل نیمه قوی[4] بازار مورد نظر باشد، تا زمانی که تحلیل‌ها به طور عام منتشر نشده با هیچ تحلیلی نمی‌توان بازدهی بهتر از بقیه به دست آورد. برای مثال، تحلیل صورت‌های حسابداری شرکت، دیگر تحلیلی کارساز نبوده و نمی‌تواند به تمایز بین سرمایه‌گذاری سودآور و غیر سودآور منجر شود. به عبارت دیگر، از آنجایی که این صورت‌ها را قبلاً هزاران تحلیلگر دیگر مورد بررسی قرار داده‌اند، آن تحلیلگران به اتکای آنچه یافته‌اند، عمل نموده‌اند، و قیمت جاری سهام، اکنون بازتاب تمام اطلاعاتی است که در صورت‌های مالی یافت می‌شود. همین موضوع در مورد سایر منابع اطلاعات عمومی منتشر شده نیز صدق می‌کند. در شکل قوی نظریه بازار کارا فرض می‌شود که تمامی اطلاعات مربوط و موجود – اعم از اطلاعات محرمانه و اطلاعات در دسترس عموم (جاری) – در قیمت اوراق بهادار انعکاس دارد. قیمت اوراق بهادار حتی منعکس کننده‌ی تمام اطلاعات محرمانه جاری و تاریخی است. در این شکل، اگر اتفاقی در شرکت بیفتد، دیگر نمی‌توان گفت که فقط مدیرعامل آن را می‌داند و دیگران از آن بی‌اطلاع هستند. در این زمینه، قیمت سهام شرکت مورد نظر، بلافاصله نسبت به این

---

[1]- Fama
[2]- Technical Analysis
[1]- Random Walk
[2]- Semi Strongly Form

اتفاق عکس‌العمل نشان می‌دهد، چرا که دیگر اطلاعات محرمانه‌ای[1] نباید موجود باشد. در چنین شرایطی، فرض می‌شود که سیستم‌های کنترل داخلی شرکت به قدری قوی است که کسی نمی‌تواند اطلاعات محرمانه و یا اطلاعات خاص داشته باشد. برای مثال، وقتی مدیر عامل یا هیئت مدیره از امری باخبر می‌شود، بلافاصله بقیه نیز آن را می‌شنوند و چون همه خبر دارند، قیمت به سرعت متأثر می‌گردد و بنابراین اطلاعات به ظاهر محرمانه، دیگر ارزش چندانی برای مدیر عامل و سایر مدیران عالی رتبه ندارد[2]. شک نیست که این نوع کارآیی فقط در کتاب‌ها یافت می‌شود، و حتی در بورس‌های بسیار معتبر هم واقعیت نمی‌یابد. اگر بازار به شکل قوی کارا نباشد، کسی که سرمایه دارد و اطلاعاتش از بقیه بیشتر است و افراد خبره‌ای برای تحلیل این اطلاعات خاص در دسترس دارد، قاعدتاً بازده بیشتری به دست می‌آورد.

هر چه از سطح ضعیف فرضیه به سمت سطح قوی آن پیشرفت حاصل شود، انواع مختلف تحلیل‌های سرمایه‌گذاری در تعیین مرز بین سرمایه‌گذاری‌های سودآور و غیر سودآور اثر خود را از دست می‌دهند [۹]، [۱۰].

۲-۳-۲- آزمون‌های تجربی کارآیی بازار

هر سه سطح فرضیه بارها مورد آزمون قرار گرفته است. نخست باید از فرضیهٔ گشت تصادفی سخن گفت که به شکل‌های مختلف آزمون شده است. هری رابرتز[3] برای اولین بار در سال ۱۹۵۹ اظهار داشت که: یک رشته از اعداد را که به طور تصادفی استخراج کرده و نمودار آن را ترسیم نموده، به چشم همان‌طور می‌آید که نمودار یک رشته از قیمت‌های سهام نشان می‌دهد. اگر نمودارگر حرفه‌ای

---

[1]- معامله به اتکای اطلاعات محرمانه در بسیاری از کشورها، جرم (و حتی جرم جنایی) است و در صورت اثبات، شامل جرایم سنگین و پرداخت خسارت می‌شود.

[2]- لازم به توضیح است که معمولاً معامله سهام توسط مدیران ممنوع نیست، ولی آن معامله‌ها باید به اطلاع ضابطان و مسئولان رسانده شود. به علاوه، مدیرانی که سهم شرکت خود را می‌خرند یا می‌فروشند، نباید به سرعت این کار را انجام دهند. معمولاً برای این کار فاصله زمانی تعیین می‌شود. مثلاً ۶ ماه، یعنی حداقل بین خرید و فروش، ۶ ماه فاصله زمانی لازم است، و چنین معامله‌هایی مجاز است و باید به اطلاع عموم برسد.

[3]- Harry Roberts

روند خاصی را در هر دو نمودار جستجو نماید؛ چنین نقاطی را در هر نمودار پیدا می‌کند (هم در نمودار اعداد تصادفی و هم در نمودار قیمت سهام). جالب‌تر اینکه، رابرتز اشاره دارد که اگر تغییر قیمت سهام، هم محاسبه گردد و تمام این تغییرات روی نمودار رسم شوند، باز هم دو نمودارِ بسیار شبیه به هم حاصل می‌گردد. در پی آزمون‌های اولیه رابرتز، آزمون‌های مور[1] (۱۹۶۲)، فاما (۱۹۶۵) و گرانجر و مورگن اشترون[2] (۱۹۶۳) تأیید کاملی از نتایج رابرتز به همراه آورد. به طور خلاصه، همه شواهد حاصل از آزمون‌های تجربی، فرضیه گشت تصادفی را تأیید می‌کردند [۹].

۲-۳-۳- مطالعات تجربی سطح ضعیف کارایی در بازارهای دنیا

برای اولین بار، لوئیز باچلر[3] در سال ۱۹۰۰ رفتار قیمت کالاها را در فرانسه مورد بررسی قرار داد. وی فرض نمود که تغییرات قیمت از یک معامله به معامله دیگر، یک متغیر تصادفی مستقل و دارای توزیع یکنواخت با واریانس محدود است. به طور کلی، او به این نتیجه دست یافت که که قیمت جاری کالاها نمی‌تواند در تعیین قیمت آینده آنها راهنمای مناسبی باشد. به دنبال این تحقیق باچلر سعی نمود که رفتار قیمت اوراق قرضه دولتی فرانسه را مورد بررسی قرار دهد. وی با مطالعاتی که در این زمینه انجام داد، به این نتیجه رسید که قیمت اوراق قرضه دولتی از مدل گردش تصادفی تبعیت می‌نماید. لیمن در سال ۱۹۹۰ و کنراد کول و نیمالندرن[4] در سال ۱۹۹۱ مطالعاتی بر روی رفتار قیمت سهام شرکت‌های آمریکایی انجام داده اند. آنها به خود همبستگی قابل توجهی در قیمت اوراق بهادار دست یافته و نشان دادند که قیمت‌ها و بازده گذشته اوراق بهادار حاوی اطلاعات با ارزشی است که می‌تواند جهت پیش‌بینی آینده قیمت‌ها استفاده شود [۱۱]. کیم نلسون و استارتز[5] با استفاده از روش نسبت واریانس و با به کارگیری بازده ماهیانه سهام در بورس نیویورک، در دوره زمانی

3- Moor
1- Granger and Morgenstern
2- Bachelier
3- Lemann, Conrad Kaul and Nimalendran
4- Kim, Nelson and Startz

(۱۹۲۶-۱۹۸۶) عدم کارایی بازار را در سطح ضعیف نشان دادند و اعلام نمودند که تقریباً ۲۵ تا ۴۰ درصد از نوسان بازده سهام در آینده را می‌توان با استفاره از بازده گذشتهٔ سهام پیش بینی نمود [۱۲]. در سال ۱۹۹۶، داکری و کاسانس[۱] کارایی بازار سهام بوستون را به روش رگرسیون ساده مورد مطالعه قرار دادند. داده‌های این بررسی نیز قیمت ماهیانه سهام برای ۷۳ شرکت می‌باشد. نتایج ارائه شده توسط این محققین، فرضیه‌های مربوط به گردش تصادفی قیمت سهام، که شرط ضروری جهت کارایی بازار سهام است، را رد می‌نماید و علاوه بر آن، تمایل به حرکت منظم در طول زمان در قیمت سهام بوستون مشاهده می‌شود [۱۳]. نتایج تحقیقات دهه‌ی آخر قرن بیستم نشان داد که اطلاعات قیمت و بازده گذشته، حاوی اطلاعات با ارزشی جهت پیش‌بینی است؛ به طوری که، با همبستگی غیر خطی موجود بین بازده، می‌توان نوسان بازده سهام را در آینده پیش‌بینی نمود.

کارایی بورس سهام لندن در سطح ضعیف، در سال ۱۹۹۰ توسط دونالدسون[۲] مورد مطالعه قرار گرفت. وی به این نتیجه رسید که تغییرات قیمت سهام قابل پیش‌بینی است [۱۴]. استاک[۳] در سال ۱۹۸۸ روند قیمت سهام شرکت‌های پذیرفته شده در بورس آلمان را مورد بررسی قرار داد که نتیجه عدم کارایی این بازار را در سطح ضعیف نشان داد [۱۵]. میلز[۴] در سال ۱۹۹۴ با به‌کارگیری مدل رگرسیون و با استفاده از بازده سهام اوراق قرضه در انگلستان، قابلیت پیش‌بینی بازده را بررسی نمود. نتایج این تحقیق نشان داد که شواهد قوی در مورد قابلیت پیش‌بینی بازده بخصوصی، برای افق زمانی کوتاه مدت وجود دارد [۱۶]. لاک[۵] با مطالعه شاخص وزنی سهام در بورس سهام تایوان، تغییرات روزانه، هفتگی و ماهیانهٔ قیمت سهام را مورد بررسی قرار داد. داده‌های این بررسی مربوط به دورهٔ زمانی (۱۹۹۴-۱۹۶۷) بوده و روش‌های مورد استفاده: رگرسیون خطی ساده، آزمون گردش‌ها و آزمون نسبت واریانس می‌باشند. نتایج به دست آمده، شواهدی قوی در رد فرضیه های گردش تصادفی در

5- Dockery and Kauussanos
1- Donaldson
2- Stock
3- Mils
4- Lock

بـازار سهام تایوان ارائه می‌نماید [۱۷]. تحقیقات بسیار دیگری در کشورهای مختلف صورت گرفته که از حوصله‌ی این تحقیق خارج است. نتیجه‌ی کلی این مبحث در پاراگراف زیر آمده است:

مطالعات تجربی کشورهای مختلف نشان می‌دهد که انگلستان و فرانسه پیشگامان تحقیق در زمینه کارایی بورس اوراق بهادار بوده‌اند؛ و به طور کلی، دامنهٔ این مطالعات از اوایل دههٔ نهم قرن بیستم بـه سایـر بورس‌های جهان کشیده شده و تحقیقات مختلفـی در آلمان، فنـلاند، کلمبیـا، سوئـد، ترکیـه، ژاپن و ... انجام گرفته است. نتـایج اکثر ایـن تحقیقات، خود همبـستگی منفـی در افـق زمـانی بلنـد مدت و خود همبـستگی مثبـت در افق زمانـی کوتـاه مـدت بین قیمت های متوالی سهـام را نشـان می‌دهد؛ و با به‌کارگیری بازده گذشتهٔ سهام، می‌توان بازده آینده را در کوتاه‌مدت پیش‌بینی نمود.

۲-۳-۴- مطالعات تجربی سطح ضعیف کارایی در بورس اوراق بهادار تهران

در سال ۱۳۶۹، در ایران مطالعه‌ای در رابطه با تابع توزیع نوسانات قیمت سهام بورس اوراق بهادار، با استفاده از شاخص بازار سهام ایران و برای دو دوره‌ی زمانی (۱۳۵۷-۱۳۵۳) و (۱۳۶۸-۱۳۶۴) انجام شد و نتایج تحقیق دلالت بر عدم کارایی بازار بـورس ایـران در شـکل ضـعیف داشـت [۱۸]. تحقیـق دیگری نیز جهت آزمون کارایی در سطح ضعیف در سال ۱۳۷۱ با استفاده از قیمت سهام ۱۷ شرکت و برای دوره زمانی (۱۳۷۹-۱۳۶۸) و با به کارگیری روش آزمون گـردش‌هـا انجـام شـد و مجـدداً عـدم کارایی بازار بورس اوراق بهادار ایران به اثبات رسید [۱۹]. تحقیق دیگری در سال ۱۳۷۴ با استفاده از قیمت‌های هفتگی سهام ۵۰ شرکت پذیرفته شده در بورس اوراق بهادار و بـا بـه‌کـارگیری روش خـود همبستگی و آزمون گردش‌ها جهت آزمون کارایی در سطح ضعیف انجـام شـد و نتـایج تحقیـق عـدم کارایی بازار بورس ایران را نشان داد. در سال ۱۳۷۴ تحقیق دیگری در ایران جهت آزمون کارایی بازار بورس در سطح ضعیف انجام شد که نتایج حاصل از این تحقیق نشان داد که تغییرات قیمت به صورت مستقل و تصادفی نیستند و روند و الگوی خاصی در رفتار قیمت‌ها مشاهده می‌شـود و آگـاهی از ایـن الگو می‌تواند جهت کسب منافع بیشتر به سرمایه گذاران کمک نماید. با توجه به نتایج به دست آمده،

عدم کارایی بازار بورس در سطح ضعیف به اثبات می‌رسد [۱۰]. تحقیق دیگری نیز در سال ۱۳۷۳ صورت گرفت که نتایج به‌دست آمده از این تحقیق هم نیز بیانگر عدم کارایی بازار بورس تهران و رد فرضیه گشت تصادفی می‌باشد و روند قابل پیگیری در قیمت‌ها دیده می‌شود [۲۰].

طبیعی است نتایج بالا بر اساس اطلاعات موجود و محدودیت‌های زمانی و مکانی بورس اوراق بهادار تهران به دست آمده است. انتظار می‌رود با افزایش تعداد شرکت‌های پذیرفته شده، ازدیاد آگاهی مردم از مکانیزم بورس، پیشرفت تخصص‌های علمی و فنی در داخل بورس و مکانیزه شدن اطلاعات مالی، بورس اوراق بهادار تهران به سمت کارایی حرکت نماید و نهایتاً به حد قابل قبولی از درجه بهینه سازی اقتصادی برسد.

در نهایت از این فصل این طور استنباط می‌شود که بازار بورس تهران قابلیت پیش بینی در کوتاه مدت به کمک تحلیل‌های نموداری و در بلند مدت با استفاده از تحلیل‌های تکنیکی را دارد. در فصل بعد به توضیح روش ماشین‌های برداری پشتیبان خواهیم پرداخت.

فصل سوم

ماشین های برداری پشتیبان

۳-۱- مقدمه

از فصل قبل نتیجه شد که بازار بورس تهران با استفاده از تحلیل‌های تکنیکی قابلیت پیش بینی در کوتاه مدت و بلند مدت را دارد. تا حدودی با چالش‌ها، لزوم تحقیق و متغیرهای بازار سهام آشنا شدیم. در این فصل به طور مفصل راجع به روش به کار برده شده در پایان نامه بحث می‌کنیم.

در امور تحلیل داده‌ها، به ناچار به یک سری داده‌های تجربی پراکنده (داده‌هایی که از آزمایشات گوناگون به‌دست می‌آیند) برخورد می‌کنیم. تجربه ثابت کرده، داده‌های به‌دست آمده محدود و نمونه‌وار می‌باشند و فقط درون فضای ورودی و با توزیع پراکنده شکل می‌گیرند. دست‌یابی شبکه‌های عصبی سنتی به تولید مدل‌هایی که بتواند روی داده‌ها منطبق شود مشکلات زیادی را تحمل کرده است.

طراحی ماشین هایی با توانایی فراگیری و تجربه آموزی مدت ها یکی از بحث های مهم علوم تحقیق بود تا سرانجام با ظهور کامپیوتر های قدرتمند در این راه گام های مثبتی برداشته شد. در ادامه اثبات گردید که این ماشین ها توانایی بالایی جهت فراگیری داده ها دارند و می توانند پس از فراگیری داده های آموزشی، نتایج قابل قبولی را ارائه نمایند. لزوم به کارگیری چنین روش هایی نتیجه ضعف تکنیک های برنامه ریزی کلاسیک بود که قادر به حل بسیاری از مشکلات به دلیل عدم ارائه مدل ریاضی اولیه نبودند. بعد ها با ارائه روش هایی که عملکردی مشابه سیستم های عصبی مغز انسان داشتند و از قانون بیز در مدل سازی بهره می جستند، راه حل های نوینی در حل بسیاری از مشکلات برداشته شد. شبکه عصبی یکی از پویاترین حوزهای تحقیق در دوران معاصر بود که توجه افراد بسیاری از رشته های گوناگون علمی را به خود جلب نمود. در ابتدای امر کاربرد این روش بسیار محدود تر از حال بود و شاید تنها در شاخه الکترونیک برای حذف سیگنال های مزاحم که اصطلاحاً نویز نامیده می شوند، مورد استفاده قرار می گرفت، اما کم کم جای خود را به عنوان یک روش قدرتمند در حل مشکلات بسیاری از علوم باز نمود.

حال پس از گذشت سال ها از عمر شبکه های عصبی و توسعه کاربردهای آن، نوبت به ارائه روشی جدید و نیرومندتر با نام ماشین های برداری پشتیبان رسیده است. ماشین برداری در اوایل دهه ۹۰ معرفی شد و پس از آن به عنوان روشی نیرومند در حل مسائل طبقه بندی غیر خطی و رگرسیون مورد استفاده قرار گرفت. بر طبق نظریه محققین دلیل موفقیت این روش را می توان در قدرت بالای تعمیم دادن مسائل و توانمندی در مقابله با نویز ها و کمبود داده ها دانست.

در همین راستا هدف از فصل حاضر، معرفی روشی جدید و قدرتمند در شناسایی الگو تحت عنوان ماشین های برداری پشتیان است که از ترکیب بهینه سازی، آموزه های آماری و هسته های کرنل بهره می جوید و به صورت گسترده در حل بسیاری از مسائل از جمله، شناسایی و طبقه بندی نوشتاری، مشکلات بیولوژیکی و پزشکی، تصاویر روزنانس مغناطیسی، پردازش سیگنال های خطی، شناسایی گفتار، پردازش تصویر، سیستم های هیدرولوژیکی و ... به کارگرفته می شود.

مدل سازی کلاسیک

فرض می شود پارامترهای A، B، C و D سیستمی را تعریف می کنند. این پارامترها می توانند هر عاملی مانند غلظت مواد شیمیایی، پارامترهای فیزیولوژیکی، وضعیت بیماران یا هر چیز دیگری باشند. در کنار هر پارامتری که یک نمونه را تعریف می کند یک پاسخ R وجود دارد که می تواند چیزی مانند پایداری فرمول دارویی یا پارامترهای وضعیت بیمار باشد. مدل های کلاسیک در شرایطی مانند این از نخستین گام مرتکب اشتباه بزرگی می شوند که تنها در سیستم های ساده (خطی یا نزدیک به خطی) قابل صرف نظر است. نخستین گام یک روش کلاسیک در بررسی داده ها، بررسی شاخص های تمایل به مرکز (میانگین) و شاخص های پراکندگی (انحراف معیار) است. پس از این مرحله نمونه ها و اهمیت فردی نمونه ها کنار گذاشته می شوند که این موجب اشتباهات بزرگی خواهد شد. به عنوان مثال، ممکن است پارامتر A از صفر تا ۱۵ درجه روی پایداری دارو اثر مثبت داشته باشد و از ۱۵ درجه به بالا این اثر منفی شود. در این حالت بررسی تاثیر پارامتر A به صورت

فردی ضروری است اما این در حالی است که با گرفتن میانگین از کل پارامترها، اثر همراهی مقادیر A یا B یا هر پارامتر دیگری از صورت مسئله پاک می شود. در هیچ یک از روش های کلاسیک روی اثر تک تک نمونه ها بحث نمی شود و این یک اشکال عمده است. در واقع روش کلاسیک با عملی شبیه به آسیاب کردن، پیچیدگی روابط موجود بین پارامترهای یک مسئله را از بین می برد و از کشف پیچیدگی ها باز می ماند. برای حل مسائل نظیر این به سیستمی نیاز است که با اهمیت دادن به تک تک پارامترها، به ارزیابی آنها بپردازد و بدون پیش داوری در مورد نوع تابع هر پارامتر، آنها را بررسی و ارزیابی نماید [۲۱]، [۲۳].

۳-۲- آموزش الگو

برای طراحی فرآیند آموزش، ابتدا باید مدلی از محیط مورد نظر که شبکه می تواند در آن عملکرد خوبی داشته باشد، طراحی گردد و اطلاعات لازم و در دسترس، فراهم آیند. ثانیاً، باید مشخص گردد که چگونه باید وزن دهی صورت گیرد و چه روشی مناسب تر است [۲۱]، [۲۲]، [۲۴]. در واقع آموزش الگو، روشی است که در آن قوانین آموزشی به کار گرفته می شوند تا وزن دهی به بهترین نحو ممکن صورت پذیرد. انواع روش های آموزش الگو به شرح زیر می باشند.

۱- آموزش با راهنما

۲- آموزش بدون راهنما

۳- آموزش وراثتی

۳-۲-۱- آموزش با راهنما

یادگیری در شبکه های رایج (فازی، عصبی، ماشین های برداری و...) به طریق نظارت شده صورت می‌گیرد. در این نوع از یادگیری به شبکه آموخته می شود که بین دادهای آموزشی و خروجی

ارتباط برقرار کند. این کار از طریق حداقل نمودن اختلاف بین نتایج حاصل از مجموعه آموزشی و خروجی های مطلوب انجام می پذیرد [۲۳]، [۲۵]. این فرآیند شبیه یادگیری یک خردسال است. والدین تصویر حیوانات مختلف را به کودک نشان می دهند و نام هرکدام را به کودک می گویند. به عنوان مثال، می توان سگ را در نظر گرفت. کودک تصاویر مختلف سگ را می بیند و در کنار اطلاعات ورودی (صدا و تصویر) برای هر نمونه، به او گفته می شود که این اطلاعات مربوط به یک نوع سگ است. بدون اینکه به او گفته شود، سیستم مغز او اطلاعات ورودی را تجزیه و تحلیل می کند و به یافته هایی در زمینه هریک از پارامترهای ورودی مانند رنگ، اندازه، صدا و... می رسد، پس از مدتی او قادر خواهد بود تا یک نوع خاص از سگ را که هرگز ندیده است، شناسایی نماید. از آنجایی که در مورد هر جانور در مرحله آموزش به کودک گفته شد که آیا این سگ است یا خیر، به این نوع آموزش با راهنما گفته می شود. در واقع در این نوع از یادگیری، جواب درستی در واکنش به الگوی ورودی ارائه می گردد و وزن ها به گونه ای اختصاص داده می شوند تا شبکه بتواند جوابی را ارائه نماید که تا حد امکان به جواب صحیح نزدیک باشد. به این نوع از آموزش، آموزش با معلم[1] نیز گفته می شود [۲۲].

۳-۲-۲- آموزش بدون راهنما

در این نوع از یادگیری، جواب درستی برای داده های ورودی که به عنوان الگو معرفی می شوند، ارائه نمی گردد. در این روش، شبکه تنها در میان ساختار داده ها به جستجو می پردازد و آنها را بر اساس روابطی که کشف می کند، طبقه بندی می نماید. در واقع، هدف، یافتن الگویی بین داده‌هاست که بتواند بر اساس خصوصیات مشابه، داده ها را کلاس بندی کند. البته این نوع از یادگیری کاربرد کمتری دارد [۲۴].

[1] -Learning with a teacher

۳-۲-۳- آموزش وراثتی

این نوع از آموزش، ترکیبی از آموزش با راهنما و غیر نظارت شده است. بر اساس این نوع از آموزش، قسمتی از وزن ها از طریق یادگیری نظارت شده و گروهی دیگر از طریق فرآیند غیر نظارت شده، تعیین می شوند.

همراه با فرآیندهای یادگیری ذکر شده، سه ساختار بنیادی دیگر نیز باید در نظر گرفته شود. این سه ساختار عبارتند از

۱- ظرفیت

۲- پیچیدگی داده ها

۳- پیچیدگی محاسبات

۳-۲-۳-۱- ظرفیت

ظرفیت مشخص کننده تعداد الگوهای آموزشی قابل ذخیره است و بیان می‌کند که یک شبکه می تواند از چه تابعی استفاده کند یا چه شرایط مرزی را در نظر بگیرد.

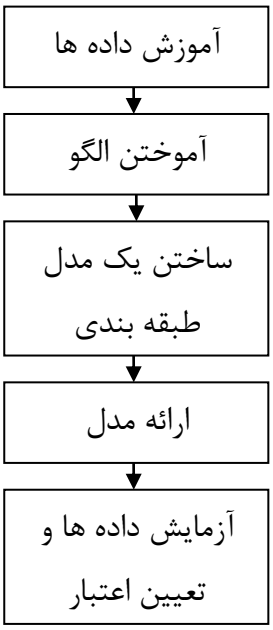
۳-۲-۳-۲- پیچدگی نمونه ها

این ساختار، تعداد الگوهای آموزشی مورد نیاز شبکه برای رسیدن به یک جواب درست را تعیین می کند.

۳-۲-۳-۳- پیچیدگی محاسبات

زمان مورد نیاز برای آموزش الگوریتم یادگیری را مشخص می نماید تا شبکه بتواند راه حل درستی را برای الگوی ارائه شده، معرفی نماید.

الگوریتم های یادگیری بسیاری وجود دارند که پیچیدگی های محاسباتی بالایی نیز دارند. طراحی الگوریتمی مؤثر با راندمان بالا برای شبکه یکی از موضوعات مورد علاقه محققین است. مراحل مختلف یک فرآیند طبقه بندی در شکل ۳-۱ نشان داده شده است [۲۱]، [۲۲]، [۲۳]، [۲۴].



شکل ۳-۱: مراحل مختلف فرآیند طبقه بندی

۳-۳- تئوری یادگیری آماری

هدف از مدل سازی، انتخاب مدلی است که در چهار چوب فضای فرضیه بوده و رابطه نزدیکی با تابع در برگیرنده فضای هدف داشته باشد. خطاهای موجود در مدل سازی عموماً ناشی از دو عامل می باشند.

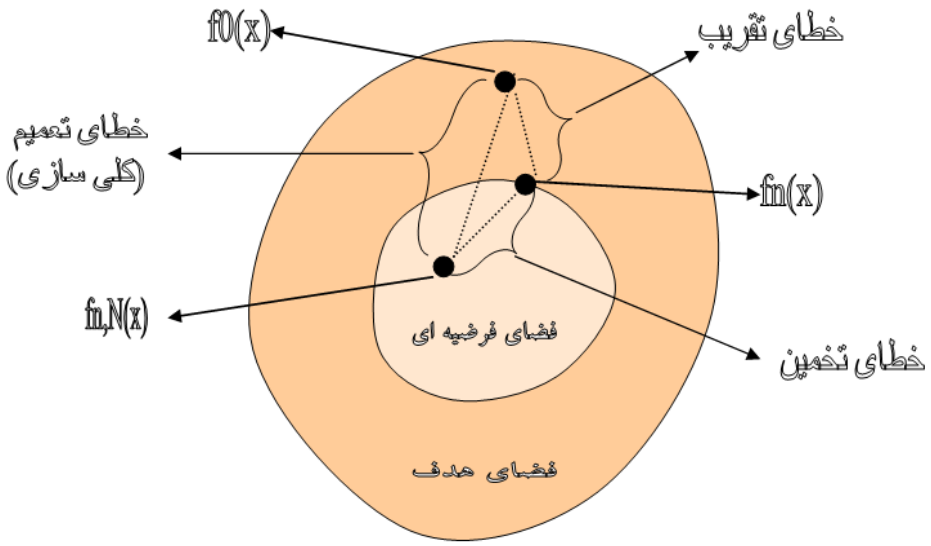
۱- خطای ناشی از تقریب

این خطا ناشی از کوچکتر بودن فضای فرضیه نسبت به فضای هدف می باشد، به این معنی که این خطا زمانی اتفاق می افتد که تابع مد نظر در خارج از فضای فرضیه مربوطه تعریف شود. انتخاب

نامناسب فضای مدل می تواند موجب بروز یک خطای بزرگ تقریب و عدم تطابق مدل گردد. بـه ایـن مدل، مدل نامطابق[1] گویند [۲۲].

۲- خطای تخمین

خطای تخمین خطایی است که در طول پروسه یادگیری رخ می دهد، انتخاب تکنیک مناسب را دچار مشکل می‌سازد و مدل منتجـه از فضای فرضیـه را غیر بهینـه خواهد نمـود اما مـی‌تـوان بـا انتخـاب روش های بهینه سازی مناسب میزان این خطـا را کاهش داد. شمایی کلی از خطای مـدل سـازی در شکل ۳-۲ نشان داده شده است.



شکل ۳-۲: خطای مدل سازی

این خطاها (خطای تقریب و تخمین) در ترکیب با یکدیگر خطای تعمیم[2] را به وجود می آورند.

ماشین فراگیر را می توان به صورت تابعی از $f(x)$ بیان نمود کـه بـه عنـوان ورودی، نمونـه ای مانند x را دریافت می کند و برداری مانند y را معرفـی مـی کنـد. ایـن تابـع یکـی از انـواع تـوابع پارامتری $f(\mathbf{w}, \mathbf{x})$ است و هدف، تخمیـن پارامتر W جهـت تخمـین مقـدار y از بـردار x اسـت. در

[1] -Not Corresponding Model
[2] -Generalization Error

همیـن رابطـه، فرض می شود که زوج های $\{x, y\}$از تابع توزیع $P(x, y)$ پیروی می کنند و بر همین اساس یک تابع لاس[1] به صورت $L(f(w,x),y)$ بین تخمین $f(w, x)$ و خروجی مطلوب تعریـف می گردد. تابع لاسی که معمولاً مورد استفاده قرار می گیرد تابع مربعات لاس است. بهتـرین راه حـل ممکن در این حالت زمانی اتفاق می افتد که تابع خطایی که وابسته به تابع لاس است، حداقل گـردد. این تابع خطا به صورت رابطه (۳-۱) بیان می گردد، که هدف به دست آوردن تابع f به صـورتی مـی باشد که میزان خطا را به حداقل برساند.

$$R[f] = \int_{x \times y} L(y, f(x)) P(x, y) dx dy \qquad (۱-۳)$$

در رابطه (۳-۱)، $P(x,y)$ تابعی ناشناخته است که با توجه به اصل حـداقل سـازی خطـای تجربـی، مقدار تقریبی برای آن در نظر گرفته می شود که با توجه به آن می توان رابطه (۳-۲) را ارائه نمود.

$$R_{emp}[f] = \frac{1}{l} \sum_{i=1}^{l} L(y_i, f(x_i)) \qquad (۲-۳)$$

رابطه (۳-۲) خطای تجربی را به حداقل می رساند و از طرفی بر اساس رابطه (۳-۳)

$$f_{n,l}(x) = \arg\min R_{emp}[f] \qquad (۳-۳)$$

حداقل سازی خطای تجربی به حدی مربوط می شود که به صورت رابطه (۳-۴) بیان می گردد

$$\lim_{l \to \infty} R_{emp}[f] = R[f] \qquad (۴-۳)$$

رابطه (۳-۴) با توجه به قانون اعداد بزرگ، معتبر است. با این حال رابطه (۳-۵) نیز باید برقرار باشد.

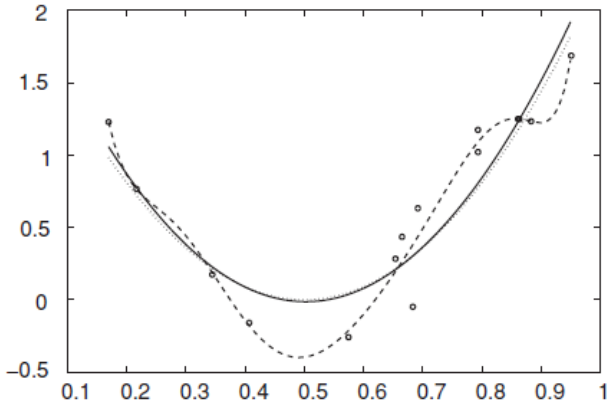
$$\lim_{l \to \infty} \min_{f \in Hn} R_{emp}[f] = \min_{f \in Hn} R[f] \qquad (۵-۳)$$

رابطه (۳-۵) زمانی درست می باشدکه H_n به قدر کافی کوچک باشد. این شرایط به صورت شهودی و عینی وجود ندارد و نیازمند آن است که مقدار مینیمم به صورت همگـرا در آیـد. ایـن خصوصـیات، از جمله ویژگی هایی است که برای ریسک تجربی در نظر گرفته می شود [۱]، [۳]. در اکثر موارد سعی

[1] -Loss function

می شود تا با افزایش میزان آموزش به ماشین تابع خطا به حداقل ممکن برسد اما این امر در اکثر موارد منجر به افزایش خطا می گردد که به بیش برازش[1] معروف است.

یک روش ساده برای کاهش میزان این بیش برازشی، کاهش میزان پیچیدگی ماشین فراگیر است. در این شرایط بهترین کار، انتخاب ساده ترین تابع ممکن است به نحوی که در کنار کم کردن خطا به خوبی نیز بتواند داده ها را توضیح دهد. بنابراین، بهینه سازی ماشین فراگیر (طبقه بندی یا رگرسیون) بستگی زیادی به انتخاب بهترین، مناسب ترین و ساده ترین تابع ممکن دارد. این مطلب صحیح است که یک تابع پیچیده قادر است تا خطای تجربی را به حداقل برساند اما در مواجه با داده های آموزشی که ناشناخته می باشند، تابع پیچیده عملکرد ضعیفی خواهد داشت و در زمان تست داده ها عملکرد ماشین فوق ضعیف خواهد بود و میزان خطا بالا خواهد رفت. این مطلب به صورت ساده در شکل ۳-۳ توضیح داده شده است.



شکل ۳-۳: مثالی از یک مدل رگرسیون در حالتی که پیچیدگی مدل، کنترل نشده است (خط چین) و در حالتی که پیچیدگی مدل در نظر گرفته شد (خط پر)

همان گونه که در شکل ۳-۳ مشاهده می شود، خط پر دارای خطایی بیش از خط چین است اما قدرت تعمیم بالاتری دارد، ساده است و در عین حال خطای آن در مقایسه با خط چین خیلی بالا

[1] -Overfitting

نمی باشد. در مسائلی از این دست که کاهش پیچیدگی مدل و ماشین فراگیر احساس می شود، می توان با تنظیمات[1] خاصی که بر روی توابع صورت می گیرد، مسئله را حل نموده و به اصطلاح بهینه سازی کرد. این تئوری به تئوری وپنایک- چروننکیس[2] معروف است که اساس پایه گذاری ماشین های برداری پشتیبان قرار گرفته است [۲۱]، [۲۲]، [۲۴]،[۲۶].

۳-۴- جداکننده های خطی:

یک جداکننده[3] اغلب به صورت یک تابع $f(x): R^d \rightarrow R$ نشان داده می شود. وقتی که دو کلاس داریم' یک داده به کلاس مثبت نسبت داده می شوداگر $f(x) \geq 0$ باشد و درغیر این صورت به کلاس منفی تعلق دارد.

تابع خطی' تابعی است که از ترکیب خطی ورودی x به صورت زیر تعریف می شود:

$$f(x) = \sum w_{ij}\, x_j + b = w^T x + b$$

مجموعه ای از نقاط (x,y) که $y_i \in \{-1,1\}$ به صورت خطی قابل جداسازی هستند اگر یک جداکننده ی خطی بتوان پیدا کرد به نحوی که برای تمام i ها رابطه ی زیر برقرار باشد:

$$y_i f(x_i) \geq 0$$

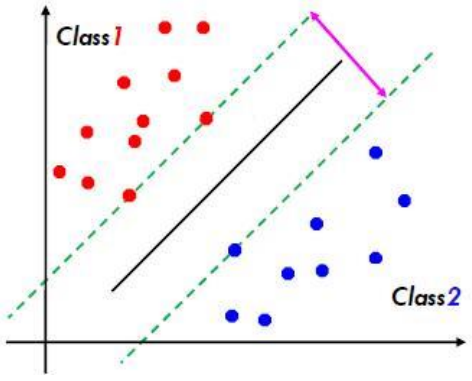
۳-۵- ابر صفحه جداساز

ابرصفحه[4] یک مفهوم در هندسه است' یک تعمیم از یک صفحه در تعداد متفاوتی از ابعاد. یک ابرصفحه یک زیرفضای k بعدی در یک فضای n بعدی تعریف می کند که $k<n$. مثلا خط یک

[1] -Regularization
[2] -Vapnik–Chervonenkis
[3] -Linear classifier
[4] -hyperplane

ابرصفحه یک بعدی در یک فضا با هر تعداد بعد است. در سه بعد' صفحه یک ابرصفحه دوبعدی است و به همین ترتیب برای فضاهای با ابعاد بالاتر ابرصفحه تعریف می شود.

ماشین های برداری پشتیبان جدا کننده را می توان به مسائل طبقه بندی دو کلاسه تا چند کلاسه تقسیم نمود. برای درک ساده موضوع می توان از مسائل طبقه بندی دو کلاس استفاده نمود. هدف از این گونه مسائل ایجاد یک معیار طبقه بندی است که برای نمونه های نادیده به خوبی قابل استفاده باشد و در عین حال از قابلیت تعمیم خوبی برخوردار باشد. برای مثال می توان شکل ۳-۴ را در نظر گرفت. همان گونه که مشاهده می شود برای داده های شکل ۳-۴ تعداد زیادی ابر صفحه جداساز خطی وجود دارد که قادر است تا داده ها را جداسازی نماید اما تنها یکی از آنها دارای حاشیه[1] (فاصله بین صفحه جداساز و نزدیکترین نقاط) ماکزیمم است. این طبقه بندی کننده خطی، ابر صفحه جداساز بهینه نامیده می شود که به صورت شهودی انتظار می رود که بتواند مرز به دست آمده را به تمام محدوده های ممکن تعمیم دهد [۲۱]، [۲۶]، [۲۷]، [۲۸]. ابر صفحه جداساز بهینه نیز در شکل ۳-۴ مشاهده می شود.



شکل ۳-۴ : صفحه جداساز بهینه با حداکثر مقدار حاشیه

[1] -Margin

۳-۶- ماشین برداری پشتیبان خطی

در روش ماشین های برداری یک داده به صورت یک بردار n بعدی دیده می شود و ما می خواهیم بدانیم می توان چنین نقاطی را با یک ابرصفحه n-1 بعدی جدا کرد. این عمل جداسازی خطی نامیده می شود. ابرصفحه های بسیاری وجود دارند که می توانند داده ها را جدا کنند. سوال اینجاست که چه ابرصفحه ای را انتخاب کنیم.

فرض می شود مسئله ای برای جداسازی مجموعه نمونه های آموزشی متعلق به دو کلاس جداگانه که به صورت $D=\{(x_1,y_1),.....(x_n,y_n)\},x\in R^n,y\in\{-1,1\}$ هستند، وجود دارد. عموماً در مسائل جداسازی خطی برای برداری مانند x وزنی مانند w در نظر گرفته شود به گونه ای که این وزن بتواند بردارها را در کلاس صحیح طبقه بندی نماید به گونه ای که برای هر بردار x متعلق به مجموعه داده های D داشته باشیم $w^T x_i+b\begin{cases}\geq +1\\ \leq -1\end{cases}$.

در SVM ماکزیمم کردن حاشیه بین دو کلاس مدنظر است. بنابراین ابرصفحه ای را انتخاب می‌کند که فاصله ی آن از نزدیک ترین داده ها در هر دو طرف جداکننده ی خطی' ماکزیمم باشد. اگر چنین ابرصفحه ای وجود داشته باشد' به عنوان ابرصفحه ماکزیمم حاشیه[1] شناخته می شود.

در این حالت برای صفحه جداساز بهینه معادله (۳-۶) در نظر گرفته می شود

$$w^T x+b=0 \qquad (۳-۶)$$

و رابطه بین بردار x و وزنی مانند w به صورت ضرب داخلی بیان می گردد. اگر رابطه (۳-۶) برای ابر صفحه جداساز در نظر گرفته شود، داده های آموزشی در بالا و پایین این صفحه قرار خواهند گرفت که به ترتیب برای y=1 ، $w^T x+b>0$ و برای y=-1، $w^T x+b<0$ خواهد بود [۲۱]، [۲۶]،

[1] maximum-margin hyperplane

[۲۷]، [۲۸]، [۲۹]. بر اساس شرایط بیان شده در بالا، زمانی گفته می شود یک مجموعـه از نقاط بـه صورت بهیـنه با یـک صفحه جدا کننده، جداسازی شده اند که:

۱- بدون اشتباه در کلاس مربوط به خود قرار گرفته باشند.

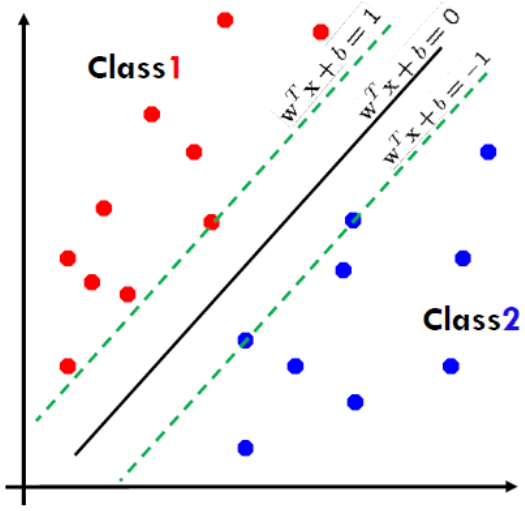
۲- فاصله بین نزدیکترین نقاط هر کلاس داده تا صفحه جدا کننده بیشینه باشد.

بر این اساس، پارامترهای w و b باید به گونه ای محاسبه گردند که دو شرط ذکر شده برقرار باشد.

برای کنترل جداپذیری داده ها رابطه (۳-۷) برای حاشیه بیان می گردد.

$$w^T x + b \begin{cases} \geq 1 & \text{for } y_i = 1 \\ \leq -1 & \text{for } y_i = -1 \end{cases} \qquad (۳-۷)$$

در شکل ۳-۵ معادلات در نظر گرفته شده برای حاشیه ها و صفحه جداسازی بهینه مشاهده می شود.



شکل ۳-۵ : صفحه جداساز و حاشیه ها

جهت معرفی صفحه جداسازی که از بیشترین حاشیه ممکن برخوردار باشد، سعی می شود تـا فاصـله بین دو حاشیه در نظر گرفته شده برای آن ماکزیمم گـردد. بـرای محاسـبه فاصـله ایـن دو حاشـیه و ماکزیمم نمودن آن از رابطه (۳-۸) استفاده می شود.

$$d(w,b;x) = \frac{\left|(w^T x + b - 1) - (w^T x + b + 1)\right|}{\|w\|} = \frac{2}{\|w\|} \qquad (۸-۳)$$

بر اساس خروجی محاسبه شده از این رابطه اگر $\frac{2}{\|w\|}$ ماکزیمم گردد، حاشیه مورد نظر ماکزیمم خواهد شد. اما راه حل ساده تری نیز وجود دارد و برای سادگی کار می توان مقدار به دست آمده را معکوس نموده و آن را مینیمم نمود که در این صورت به فرم $\frac{1}{2}w^T w$ نوشته خواهد شد. بر اساس تمام روابط و شرایط بیان شده در حالت کلی برای به دست آوردن ماکزیمم مقدار حاشیه و بهینه ترین ابر صفحه جداساز رابطه (۳-۹) ارائه می گردد.

$$Min_{w,b} = \frac{1}{2}w^T w$$
$$\text{Subject to } y_i(w^T x + b) \geq 1 \qquad (۹-۳)$$

این رابطه مستقل از b است زیرا $y_i(w^T x + b) \geq 1$ (به صورت یک صفحه جداساز) تحقق پیدا می کند و تغییر b باعث حرکت آن در جهت طبیعی به سوی خود می گردد، بنابراین مرز بدون تغییر باقی می ماند [۲۱]، [۲۷]، [۳۰].

اما سؤالی که پس از معرفی معادله (۳-۹) مطرح می شود، چگونگی مینیمم سازی مسئله بهینه سازی فوق و به دست آوردن مقادیر بهینه w و b با توجه به شرایط و محدودیت های موجود است. راه کار پیشنهادی استفاده از ضریب لاگرانژ است.

۳-۷- ضریب لاگرانژ

ضریب لاگرانژ که گاهی ضریب نامعین نیز خوانده می شود، برای یافتن نقاط خاص یک تابع که دارای چندین متغیر و محدودیت است، مورد استفاده قرار می گیرد.

فرض می شود که هدف، یافتن ماکزیمم مقدار تابع $f(x_1,x_2)$ با در نظر گرفتن محدویتی است که برای x_1 و x_2 وجود دارد. این محدودیت را می توان به صورت معادله (۳-۱۰) بیان نمود.

$$g(x_1,x_2)=0 \qquad (۱۰-۳)$$

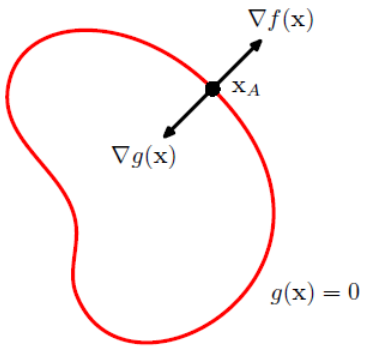
یک رویکرد برای رسیدن به جواب مد نظر می تواند حل معادله محدودیت (۳-۱۰) باشد. بـه ایـن صورت که می توان x_2 را به صورت تابعی از x_1 به صورت $x_2 = h(x_1)$ نوشت. ایـن رابطـه سـپس در تابع $f(x_1,x_2)$ قرار داده می شود تا معادله ای به صورت $f(x_1, h(x_1))$ به دست آید که تنها تابعی از متغیر x_1 است. ماکزیمم مقدار این معادله با در نظر گرفتن x_1 به سادگی از طریق مشتق گیری بـه دست می آید، در حالی که x_2 مطابق با رابطه $x_2 = h(x_1)$ است.

مشکل این روش یافتن راه حل تحلیلی برای معادله محدودیت است به صورتی که اجـازه دهـد x_2 بر حسب x_1 نوشته شود. یافتن چنین معادله ای در صورت پیچیده بودن مسائل بسیار سخت خواهـد بود.

یک روش ساده و خوب، معرفی پارامتری مانند λ است که **ضریب لاگرانژ** نامیـده مـی شـود و در واقع از منظر هندسی بیان می شود. متغیر D بعدی x با مولفه های $x_1, \ldots, x_D$ در نظر گرفته مـی شود. معادله محدودیت $g(\mathbf{x}) = 0$ در این حالت به صورت $(D-1)$ بعد در فضای x به صـورتی کـه در معادله (۳-۱۰) گفته شد، بیان می گردد که در هر نقطه ای از سطح محدویت، گرادیان $\nabla g(\mathbf{x})$ تـابع محدودیت، عمود بر واحد سطح خواهد بود. برای مشاهده شرایط بیان شـده، نقطـه ای ماننـد X روی سطح محدودیت در نظر گرفته می شود و فرض می شود که نقطـه ای ماننـد $x+\varepsilon$ نیـز روی ایـن سطح قرار دارد. اگر بسط تیلور برای متغیر X نوشته شود، رابطه ای به صـورت معادلـه (۳-۱۱) بـه دست می آید.

$$g(x+\varepsilon)=g(x)+\varepsilon^T\nabla g(x) \qquad (۱۱-۳)$$

به خاطر اینکه x و $x+\varepsilon$، هر دو روی سطح محدودیت قرار گرفته اند، $g(x)=g(x+\varepsilon)$ و $\varepsilon^T \nabla g(x)=0$ خواهند بود. در حالت حدی $\|\varepsilon\| \to 0$، بنابراین $\varepsilon^T \nabla g(x)=0$ خواهد بود. بنابراین، ε موازی با سطح محدودیت $g(\mathbf{x}) = 0$ می شود و ∇g بر سطح عمود خواهد شد. این روابط از شکل ۳-۷ به خوبی اثبات می گردد.



شکل ۳-۷: تصویر هندسی از تکنیک ضریب لاگرانژ که در آن، تابع $f(x)$ باید با توجه محدودیت $g(\mathbf{x}) = 0$ ماکزیمم گردد. اگر X، D بعدی باشد، $g(\mathbf{x}) = 0$ مطابق با سطح (D−1) خواهد بود که با منحنی روشن کشیده شده است.

در مرحله بعد، می توان نقطه ای مانند x^* را روی سطح محدویت جستجو نمود به نحوی که $f(\mathbf{x})$ ماکزیمم گردد. این نقطه در محلی قرار خواهد گرفت که در آنجا بردار $\nabla f(\mathbf{x})$ عمود بر سطح محدویت است. این نقطه در شکل ۳-۷ نشان داده شده است. زیرا در غیر این صورت می توان مقدار $f(\mathbf{x})$ را با کمی جابه جایی در سطح محدودیت افزایش داد. بنابراین، ∇f موازی یا غیر موازی با ∇g خواهد شد و نتیجتاً پارامتری مانند λ وجود خواهد داشت که

$$\nabla f + \lambda \nabla g = 0$$

که در اینجا $\lambda \neq 0$ به عنوان ضریب لاگرانژ شناخته می شود. در این نقطه به صورت قراردادی تابع لاگرانژ به صورت معادله (۳-۱۲) بیان می گردد

$$L(x,\lambda) \equiv f(x) + \lambda g(x)) \qquad (۱۲-۳)$$

بنابراین، برای یافتن ماکزیمم مقدار تابع $f(x)$ با توجه به محدودیت $g(x) = 0$ ، می توان تابع لاگرانژ را به صورتی که در معادله (۳-۱۲) گفته شد، تعریف نمود. برای برداری مانند x که D بعدی است، این رابطه معادلات $D+1$ بعدی در اختیار می گذارد که با استفاده از آن می توان دو مقدار x^* و λ را تعیین نمود.

تاکنون، مسائل به صورت ماکزیمم سازی تابع با توجه به محدودیت مساوی با صفر $g(\mathbf{x}) = 0$ ارائه گردیدند. حال می توان توابع را به صورت مینیمم سازی در نظر گرفت در حالی که تابع محدودیت بزرگتر مساوی صفر باشد ($g(x) \geq 0$).

بنابراین، دو نوع راه حل ممکن برای اینگونه از مسائل وجود دارد. در یک حالت، تابع محدودیت در داخل فضای جواب قرار گرفته است که در آنجا $g(x) \geq 0$ ، که به آن محدودیت غیر فعال[1] گویند و در حالتی دیگر، محدودیت فعال[2] قرار دارد که در آن تابع محدودیت روی مرز فضای جواب $(g(\mathbf{x}) = 0)$ قرار گرفته است.

در مورد اول، تابع $g(x)$ نقشی را بازی نمی کند و تنها می توان $\nabla f(x) = 0$ را در نظر گرفت. این موقعیت نیز مطابق با شرایط تابع لاگرانژ است اما با این تفاوت که $\lambda = 0$ است. در مورد دوم، که راه حل روی مرز ها شکل می گیرد، محدودیت از نوع غیر مساوی است و مطابق با تابع لاگرانژ برابر با $\lambda \succ 0$ خواهد بود. بنابراین، علامت تابع لاگرانژ بحرانی و مهم خواهد بود زیرا تابع $f(x)$ تنها زمانی ماکزیمم خواهد گردید که گرادیانش با توجه به $g(x) > 0$ محاسبه گردد. بنابراین برای برخی از مقادیر $\lambda \succ 0$ ، برابر با $\nabla f(\mathbf{x}) = -\lambda \nabla g(\mathbf{x})$ خواهد بود.

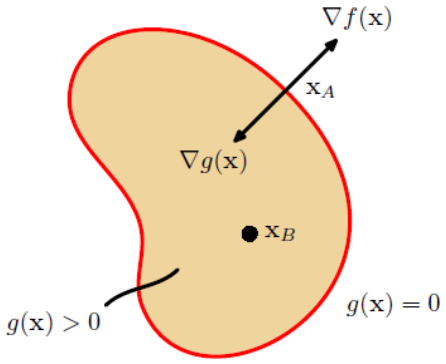
[1] -Passive
[2] -Active

برای هر دوی این موارد $\lambda g(x)=0$ خواهد بود. بنابراین، راه حل مسئله ماکزیمم سازی $f(x)$ با توجه به محدودیت $g(x)\geq 0$ توسط بهینه سازی تابع لاگرانژ (معادله ۳-۱۲) با در نظر گرفتن x و λ به صورت روابط (۳-۱۳) به دست خواهد آمد.

$$\begin{aligned} &g(x)\geq 0 \\ &\lambda\geq 0 \\ &\lambda g(x)=0 \end{aligned} \qquad (۱۳-۳)$$

این روابط با نام شرایط کاروش، کوهن، تاکر [1] (KKT) شناخته می شوند.

باید توجه داشت که اگر هدف مینیمم کردن تابع $f(\mathbf{x})$ با توجه به محدودیت $g(x)\geq 0$ باشد، می‌توان تابع لاگرانژ $L(x,\lambda)\equiv f(x)+\lambda g(x)$ را مینیمم نمود که این با در نظر گرفتن $\lambda\geq 0$ صورت می گیرد [۲۱]، [۳۳]. فضای کلی مسئله شرح داده شده در شکل ۳-۸ مشخص گردیده است.



شکل ۳-۸ : مسئله بهینه سازی ماکزیمم نمودن تابع $f(\mathbf{x})$ با در نظر گرفتن محدودیت $g(x)\geq 0$

[1] -Karush-Kuhn-Tucker

۳-۸- حل مسئله ماشین های برداری

بر اساس آنچه که در بخش قبل بیان شد، می توان مسئله بهینه سازی محدبی که به صورت رابطه (۳-۹) ارائه گردیـده بود، حل نمـود که با در نظر گرفتـن تابع لاگرانـژ به فرم بدون محدودیت رابطـه (۳-۱۴) نوشته می شود.

$$L_p(w,b,\alpha)=\frac{1}{2}w^T w-\sum_{i=1}^{N}\alpha_i\left\{y_i\left[w^T x_i+b\right]-1\right\} \qquad (۱۴-۳)$$

اگر از رابطه (۳-۱۴) نسبت به w و b مشتق جزئی گرفته شده و مساوی صفر قرار داده شود، مقـدار بهینه w به دست خواهد آمد.

$$\frac{\partial L}{\partial w}=0\Rightarrow w_0=\sum_{i=1}^{N}\alpha_i x_i y_i$$

$$\frac{\partial L}{\partial b}=0\Rightarrow\sum_{i=1}^{N}\alpha_i y_i=0$$

حال اگر مقدار w به دست آمده از مشتقات جزئی در رابطه (۳-۱۴) قرار داده شـود، معادلـه اساسـی ماشین های برداری به صورت رابطه (۳-۱۵) به دست خواهد آمد.

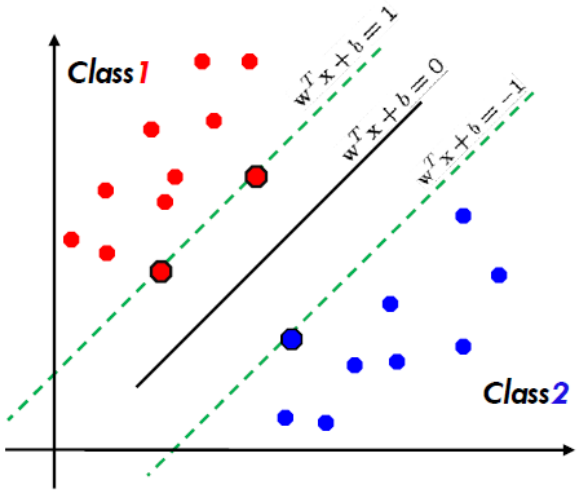
$$L_d(\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{N}y_i y_j\alpha_i\alpha_j x_i{}^T x_j$$

بنابراین هدف در ماشین های برداری حل معادله (۳-۱۵) با توجه به محدودیت های بـه دسـت آمـده است. در این حالت از ماشین های برداری، یعنی در سیستم های خطی جداپذیر، مقدار ضریب لاگرانژ باید بزرگتر از صفر باشد [۳۴].

$$\text{Max} \quad L_d(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j x_i^T x_j$$

$$\text{S.t} \begin{cases} \alpha_i \geq 0 \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{cases} \qquad (۱۵-۳)$$

مقدار بهینه b نیز از طریق رابطه $b = y_i - w^T x_i$ به دست می آید. با حل مسئله بهینه سازی رابطه (۳-۱۵) و استفاده از رابطه (۳-۱۶) می توان به بهینه ترین ابر صفحه جداساز دست یافت. این همان معادله ای است که در ماشین های برداری در نظر گرفته شده است.



شکل ۳-۹ : انتخاب صفحه جداساز بهینه

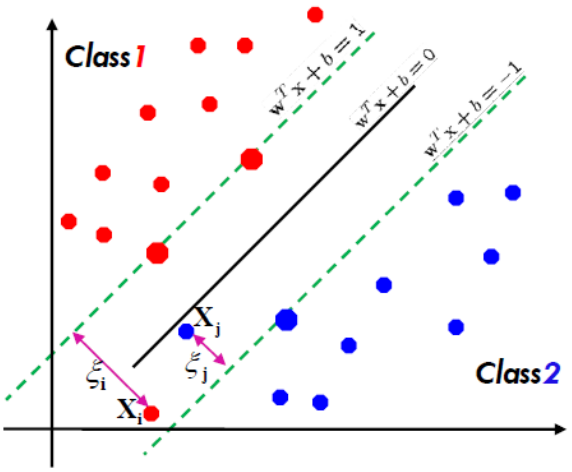
نکته: داده هایی که در شکل ۳-۹ روی حاشیه ها قرار گرفته اند، پشتیبان ها هستند. این داده ها معیاری هستند که ماشین های برداری برای طبقه بندی صحیح داده ها از آنها بهره می جویند.

۳-۹- ماشین های برداری پشتیبان در سیستم های خطی جدا ناپذیر

گاهی اوقات در سیستم های خطی داده هایی حضور دارند که در طبقه اشتباه کلاس‌بندی شده اند (شکل ۳-۱۰). برای استفاده از ابر صفحه جداساز بهینه در این شرایط، باید تابع جریمه را با تعریف پارامتر ξ_i (میزان خطای طبقه بندی) معرفی نمود.

$$F(\xi) = \sum_{i=1}^{N} \xi_i$$

مسئله بهینه سازی هم اکنون برای مینیمم سازی خطای طبقه بندی به کار می رود.



شکل ۳-۱۰ : سیستم های خطی جدا ناپذیر با میزان خطای ξ_i

بنابراین، مسئله بهینه سازی محدب در سیستم های خطی جدا ناپذیر به صورت معادله (۳-۱۷) نوشته می شود. در واقع ابر صفحه جداساز بهینه تعمیم یافته توسط بردار w تعیین می شود که تابع (۳-۱۷) را مینیمم سازد.

$$Min_{w,b} = \frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i \qquad (۱۷-۳)$$

$$\text{S.t} \quad y_i (w^T x_i + b) \geq 1 - \xi_i$$

در رابطه فوق C ضریب موازنه[1] جهت ماکزیمم نمودن حاشیه ها و مینیمم سازی خطاست. در حقیقت C یک تنظیم کننده[2] بین حاشیه ها و خطای روش است. با در نظر گرفتن ضرایب لاگرانژ β,α که بزرگتر از صفر هستند، می توان معادله (۳-۱۷) را به صورت معادله (۳-۱۸) بیان نمود.

$$L_p(w,b,\xi,\alpha,\beta)=\frac{1}{2}w^T w+C\sum_{i=1}^{N}\xi_i-\sum_{i=1}^{N}\alpha_i\left\{y_i\left[w^T x_i+b\right]-1+\xi_i\right\}-\sum_{i=1}^{N}\beta_i\xi_i \qquad (۳-۱۸)$$

در این رابطه β,α ضرایب لاگرانژ هستند که باید نسبت به w، b و x مینیمم شوند. لاگرانژ کلاسیک دوگانه قادر است مسئله اولیه مربوط به معادله (۳-۱۸) را به مسئله دوگانه آن تبدیل کند. مسئله دو گانه آن توسط رابطه (۳-۱۹) ارائه می گردد.

$$MaxW(\alpha,\beta)=Max_{\alpha,\beta}(Min_{w,b,\xi}L(w,b,\alpha,\xi,\beta)) \qquad (۳-۱۹)$$

اگر از رابطه (۳-۱۸) نسبت به w، b و ξ_i مشتق گرفته شده و مساوی صفر قرار داده شود مقادیر زیر به دست می آیند

$$\frac{\partial L}{\partial w}=0\Rightarrow w=\sum_{i=1}^{N}\alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b}=0\Rightarrow\sum_{i=1}^{N}\alpha_i y_i=0$$

$$\frac{\partial L}{\partial \xi_i}=0\Rightarrow\alpha_i+\beta_i=C \qquad i=1,2,...,N$$

با قرار دادن این روابط در رابطه (۳-۱۸)، معادله اساسی ماشین های برداری در حالت خطی جداناپذیر به دست می آید که به صورت معادله (۳-۲۰) معرفی می شود.

[1] -Trade-Off
[2] Regularization

$$\text{Max} \quad L_d(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j x_i{}^T x_j$$

$$\text{S.t} \quad \begin{cases} 0 \le \alpha_i \le C \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{cases} \qquad (۲۰-۳)$$

همان گونه که مشاهده می شود، تابع هدف سیستم های جداناپـذیر خطـی مشـابه بـا سیسـتم هـای جداپذیر خطی است و تنها تفاوت این دو، اصلاح کران های ضرایب لاگرانژ می باشد. پارامتر C نیز که قابلیت کنترل ظرفیت اضافی در طبقه بندی کننده را فراهم می آورد، باید تعیـین شـود [۲۱]، [۲۸]، [۳۵]، [۳۶].

۳-۱۰- آنالیز رگرسیون

آنالیز رگرسیون یکی از پرکابردترین ابزارهای آماری جهت آنالیز داده ها است. این یک روش ساده برای ارزیابی روابط بین متغیرهای مدل به صورت یک تـابع اسـت. هـدف اصـلی در آنـالیز رگرسـیون برازش مدل بر داده های در نظر گرفته شده و ارزیابی میزان برازش با استفاده از توابـع آمـاری اسـت. روابط موجـود به صورت معادلـه یا مدل که مرتبط کننده متغیـر وابسـته (پاسخ[1]) و متغیـرهـای غیر وابسته (پیش بینی کننده[2]) است، بیان می شوند. به عنوان مثـالی از آنـالیز رگرسـیون مـی تـوان بـه مشخص نمودن ارتباط بین قیمت خانه با ویژگی هـای آن (اتـاق هـای خـواب، بزرگـی و...) و میـزان پرداخت مالیات اشاره کرد. در ایـن مثال، متغیـر وابسته، قیمت فروش خانه و متغیـرهای غیر وابسته، ویژگی های خانه و میزان مالیات پرداختی برای یک خانه خواهد بود.

عمومـاً متغیـر وابسته را با Y و مجموعه متغیـــرهای پـیش بینـــی کننـده را بـا $x_1, x_2, \ldots, x_p$ نشـان می دهند که p مشخص کننده تعداد متغیرهای پیش بینی کننده (غیر مستقل) است. ارتبـاط بین Y و $x_1, x_2, \ldots, x_p$ می تواند با استفاده از مدل رگرسیون رابطه (۳-۲۱) تخمین زده شود.

[1] -Response
[2] -Predictor

$$Y = f(x_1, x_2, ..., x_p) + \varepsilon \qquad (۲۱-۳)$$

در این معادله ε خطای تصادفی تخمین است که میزان خطـای برازش داده هـا را نشان می دهـد و تابع $f(x_1, x_2, ..., x_p) + \varepsilon$ ارتباط بین Y و $x_1, x_2, ..., x_p$ را توضیح می دهد. به عنوان مثـالی از یک مدل رگرسیون خطی می توان به معادله زیر اشاره نمود.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon$$

$\beta_0, \beta_1, ..., \beta_p$ ضرایب رگرسیون نامیده می شوند و ثابتهای مجهول هستند که با استفاده از داده ها تعیین می گردند.

۳-۱۰-۱- مراحل مختلف آنالیز رگرسیون

مراحل مختلف آنالیز رگرسیون را می توان در غالب ۹ مرحله بیان نمود.

۱- بیان مسئله ۲- انتخاب درست متغیرها ۳- جمع آوری داده ها ۴- ویژگی های مـدل ۵- روش های برازش یک مدل ۶- برازش مدل ۷- معیارهای مدل ۸- تعیین اعتبار و ارزیابی نقص هـای مـدل ۹- اهداف آنالیز رگرسیون

۳-۱۰-۱-۱- بیان مسئله

آنالیز رگرسیون معمولا با فرموله کردن مسئله آغاز می شـود. ایـن مرحلـه شـامل پاسـخ دادن بـه سؤالاتی است که معمولاً توسط آنالیز بیان می شوند. بیان مسئله اولین و شاید مهمترین قدم در آنالیز رگرسیون است. این قدم بسیار مهم است زیرا مسئله ای که به صورت ناقص[1] تعریف می شـود یـا بـد

[1] -Ill-Defined

فرموله می شود، می تواند منجر به اتلاف زمان، هزینه و نتیجه گیری اشتباه شود. این کـار همچنـین می تواند منجر به انتخاب متغیرهای نادرست یا یک مدل آماری اشتباه گردد.

۳-۱۰-۱-۲- انتخاب درست متغیرها

قدم بعدی پس از بیان مسئله، انتخاب مجموعه متغیرهایی است که توسط متخصص مطالعه کننده مسئله برای توضیح یا پیش بینی متغیر وابسته استفاده می شود. در این مرحله، متغیر وابسته بـا Y و متغیرهای غیر وابسته با $X_1, X_2, ..., X_p$ نشان داده می شوند.

۳-۱۰-۱-۳- جمع آوری داده

گام بعدی پس از انتخاب متغیرهای مناسب ، جمع آوری داده ها از محیط مورد مطالعه برای آنالیز است. برخی اوقات، داده هـا در شرایط کنترل شده جمع آوری می شوند تا بـتوان فاکتورهایـی کـه مد نظر نیستند را ثابت نگه داشت، اما بیشتر اوقات، جمع آوری داده ها در شـرایط غیـر آزمایشـگاهی انجام می گیرد و به سختی می توان شرایط را تحت کنترل قرار داد. در مـوارد بسـیاری، جمـع آوری داده ها شامل n داده مشاهده ای از موضوع مـورد نظـر اسـت کـه هـر کـدام از ایـن n مشـاهده، بـه اندازه‌گیـری از متغیـرهای مناسب محـدود می شوند. داده ها معمـولاً به صورت جـدول ۳-۱ ثبـت می شوند. یک ستون در جدول ۳-۱ به متغیرها و ردیفی نیز داده های مشاهده ای را نشان می دهـد که تعداد p+1 مقدار برای موضوع مورد نظر هستند؛ یک مقدار به متغیر وابسته و یک مقدار به متغیر غیر وابسته اختصاص می یابد. به عنوان مثال در جدول ۳-۱، $x_{i,j}$ به i امین مقـدار و j امـین متغیـر اشاره دارد. این مقادیر در جدول ۳-۱ ارائه شده است.

جدول ۳-۱ : نماد داده های مورد استفاده برای آنالیز رگرسیون

X_p	...	X_2	X_1	متغیر وابسته (Y)	تعداد
X_{1p}	...	X_{12}	X_{11}	Y_1	۱
X_{2p}	...	X_{22}	X_{21}	Y_2	۲
X_{3p}	...	X_{32}	X_{31}	Y_3	۳
...	...	...	...	...	...
X_{np}	...	X_{n2}	X_{n1}	Y_n	N

هر متغیر جدول ۳-۱ می تواند یک متغیر کیفی یا کمی باشد. اگر تمام متغیرهای غیر وابسته کیفی باشند، روش آنالیزی که برای داده ها استفاده می شود، تکینک آنالیز واریانس نامیده می شود. اما اگر قسمتی از متغیرها کیفی و قسمت دیگر کمی باشند، از روش آنالیزی کواریانس استفاده خواهد شد.

۳-۱۰-۱-۴- ویژگی های مدل

شکل مدلی که برای مرتبط ساختن متغیرهای غیر وابسته با متغیر وابسته استفاده می شود، توسط متخصص بر اساس اطلاعات، اهداف یا قضاوتش مشخص می شود. مدل فرضی ارائه شده سپس از طریق آنالیز داده ها تایید یا رد خواهد شد. باید به این نکته نیز توجه داشت که در این مرحله تنها شکل مدل بر اساس پارامترهای مورد نظر مشخص می گردد و به صورت تابع $f(x_1,x_2,...,x_p)$ خواهد بود. این تابع می تواند در دوکلاس خطی یا غیر خطی طبقه بندی گردد. به عنوان مثالی از یک تابع خطی و غیر خطی می توان به ترتیب به توابع $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ و $Y = \beta_0 + e^{\beta_1 X_1} + \varepsilon$ اشاره نمود.

یک معادله رگرسیون که تنها شامل یک متغیر غیر وابسته است، معادله رگرسیون ساده نامیده می شود و معادله ای که حاوی بیش از یک متغیر باشد، معادله رگرسیون چند گانه نامیده می شود. اما در حالتی که تنها یک متغیر وابسته (Y) در مسئله مورد بررسی قرار می گیرد، آنالیز رگرسیون، آنالیز تک متغیره نامیده خواهد شد در حالی که اگر دو یا چند متغیر وابسته مورد بررسی قرار گیرد، آنالیز از نوع رگرسیون چند متغیره خواهد بود.

۳-۱۰-۱-۵- روش های برازش

پس از اینکه مدل تعریف گردید و داده ها جمع آوری شدند، قدم بعدی، تخمین پارامترهای مدل بر اساس داده های برداشت شده است که برازش مدل[1] نامیده می شود. متداول ترین روش تخمینی که در این فرآیند استفاده می شود، روش کمترین مربعات است. در شرایط خاص با فرضیات قطعی، روش کمترین مربعات، تخمین هایی با خصوصیات مورد نظر را تولید خواهد نمود. البته روش های دیگری نیز برای تخمین وجود دارد که از میان آنها می توان به روش درستنمایی ماکزیمم[2]، روش ریج[3]، روش تجزیه و تحلیل مؤلفه های اصلی[4] اشاره کرد.

۳-۱۰-۱-۶- برازش مدل

قدم بعدی در آنالیز، تخمین پارامترهای رگرسیون یا برازش مدلی بر روی داده های جمع آوری شده با استفاده از روش های تخمین انتخاب شده است. پارامترهای رگرسیون تخمین زده شده $\beta_0,\beta_1,...,\beta_p$ در معادله رگرسیون $Y=f(x_1,x_2,...,x_p)+\varepsilon$، به صورت $\hat{\beta}_0,\hat{\beta}_1,...,\hat{\beta}_p$ بیان می گردند. معادله رگرسیون تخمینی نهایی به صورت رابطه (۳-۲۲) خواهد بود.

---

[1] -Parameter estimation or model fitting
[2] -Maximum likelihood method
[3] -Ridge method
[4] -Principle component method

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_p x_p \qquad (۲۲-۳)$$

ارائه می گردد. کلاه بالای پارامترها نشان دهنده تخمینی بودن پارامترهای مورد نظر است و مقدار $\hat{Y}$ مقدار برازش یافته نامیده می شود.

توجه به این نکته ضروری است که بر اساس معادله رگرسیون می توان مقدار متغیر وابسته را برای مقادیر مختلف متغیرهای غیر وابسته که در داده های مشاهده ای وجود ندارد، پیش بینی نمود که در چنین مواردی، مقدار $\hat{Y}$، مقدار پیش بینی شده نامیده می شود.

۳-۱۰-۱-۷- تعیین اعتبار و مشخص نمودن نقص های مدل

اعتبار مدل آماری مانند آنالیز رگرسیون بستگی به فرضیات مشخصی دارد که معمولا برای داده ها و مدل در نظر گرفته می شوند. صحت آنالیز و نتیجه ای که از یک آنالیز به دست می آید، به صورت مستقیم به اعتبار این فرضیات بستگی دارد. بنابراین، قبل از استفاده از رابطه $Y = f(x_1, x_2, ..., x_p) + \varepsilon$ برای هر هدف خاصی، ابتدا باید فرضیات مشخصی در نظر گرفته شوند. در این موارد، عموماً به سوالات زیر باید پاسخ داده شود.

۱- چه فرضیاتی باید در نظر گرفته شود؟

۲- فرض در نظر گرفته شده را چگونه می توان تعیین اعتبار نمود؟

۳- اگر تعدادی از فرض ها نقض شوند، چه اقدامی باید صورت گیرد ؟

در حالت کلی، اعتبار فرضیات باید قبل از نتیجه گیری به اثبات برسد. آنالیز رگرسیون در واقع یک فرآیند چند مرحله ای است که در آن از منطق و اعتبار برای محاسبه مقادیر خروجی و اصلاح آنها

استفاده می شود. این فرآیند تکرار می شود تا نتایج رضایت بخشی ارائه گردد که فرضیات به خوبی در آن شکل گرفته باشند و به خوبی بر داده ها برازش داشته باشد.

۳-۱۰-۱-۸- اهداف آنالیز رگرسیون

ارائه یک معادله رگرسیون، یکی از مهمترین نتایج آنالیز و تجزیه و تحلیل داده می باشد. این معادله می تواند رابطه بین Y و مجموعه متغیرهای غیر وابسته $x_1,x_2,...,x_p$ را ساده و قابل فهم سازد و برای اهداف گوناگونی مورد استفاده قرار گیرد. به عنوان مثال می توان از معادله رگرسیون برای ارزیابی اهمیت متغیرهای غیروابسته، آنالیز تاثیر تغییرات مقادیر متغیر غیر وابسته بر روی متغیر وابسته یا برای پیش بینی مقادیر متغیر وابسته استفاده نمود. آنالیز رگرسیون در واقع آنالیز مجموعه ای از داده هاست که به شناخت ارتباط بین متغیرها در محیط های مشخص کمک می کند.

۳-۱۱- آنالیز رگرسیون خطی

در آنالیز رگرسیون خطی، رابطه خطی بین متغیر وابسته و غیر وابسته وجود دارد. در این آنالیز روی ضرایب همبستگی و کواریانس بحث می گردد و میزان خطی بودن رابطه بین متغیرها اندازه گیری می گردد.

مقادیر $\overline{Y}=\frac{\sum_{i=1}^{N} y_i}{n}$ و $\overline{X}=\frac{\sum_{i=1}^{N} x_i}{n}$ میانگین مقادیر X ،Y هستند و مقادیر

- $y_i-\overline{y}$، انحراف هر داده مشاهده ای وابسته را از میانگین متغیر وابسته نشان می دهد.
- $x_i-\overline{x}$، انحراف هر داده مشاهده ای غیروابسته را از میانگین متغیر غیر وابسته نشان می دهد.

- حاصل دو معادله فوق را می توان به صورت معادله $(y_i - \overline{y})(x_i - \overline{x})$ بیان نمود.

بر اساس روابط بیان شده در بالا می توان رابطه (۳-۲۳) را معرفی نمود که با نام کواریانس شناخته می شود و رابطه بین متغیرها را بیان می کند.

$$Cov(Y, X) = \frac{\sum_{i=1}^{N}(y_i - \overline{y})(x_i - \overline{x})}{n-1} \qquad (۲۳-۳)$$

اگر کواریانس بزرگتر از صفر باشد، ارتباط بین متغیر وابسته و غیر وابسته مثبت است و اگر کواریانس کوچکتر از صفر باشد، این ارتباط منفی خواهد بود.

معادله رگرسیون سیستم های خطی ساده به صورت $Y = \beta_0 + \beta_1 X + \varepsilon$ بیان می گردد. در این رابطه مقدار β_1 از رابطه (۳-۲۴) به دست می آید.

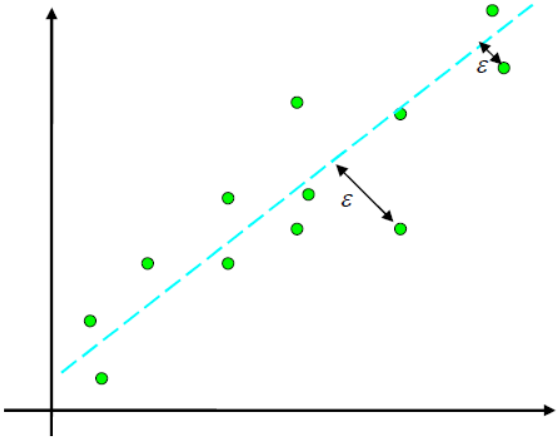
$$\beta_1 = \frac{\sum_{i=1}^{N}(y_i - \overline{y})(x_i - \overline{x})}{\sum(x_i - \overline{x})^2} \qquad (۲۴-۳)$$

و مقدار β_0 از معادله $\beta_0 = \overline{Y} - \beta_1 \overline{X}$ محاسبه خواهد شد. مقدار خطا نیز در این معادله از طریق رابطه (۳-۲۵) محاسبه خواهد شد.

$$\sum_{i=1}^{N} \varepsilon_i^{\,2} = \sum_{i=1}^{N}(Y_i - \beta_0 - \beta_1 X_i)^2 \qquad (۲۵-۳)$$

روشی که برای ارائه یک مدل ریاضی جهت برازش بر داده ها در روش های معمول (کمترین مربعات، ...) استفاده می شود، یافتن بهترین مدلی است که به صورت کلی بر داده ها برازش خوبی داشته باشد و اگر داده ای انحراف بسیاری از داده های دیگر داشته باشد، مدل مجبور خواهد بود، به

گونه ای برازش یابد که این داده نیز تحت پوشش قرار گیرد. این مدل به همراه خطـای آن در شـکل ۳-۱۱ نشان داده شده است [۳۹].



شکل ۳-۱۱ : مدل ریاضی برازش یافته در معادله رگرسیون خطی با روش کمترین مربعات

۳-۱۲- ماشین های برداری پشتیبان برای رگرسیون خطی

یک رگرسیون خطی تابعی مانند $f(x)=w^T x+b$ است که با استفاده از مجموعه بردارهـایی مانند x، مقدار اسکالری مانند y را تخمین می زند. در مسـائل طبـــقه بنـدی، رگرسـیون خطـــی بـه صورت سنتـی با استفاده از روش کمترین مربعات حل می شود. به عبارت دیگر، تابع رگرسـیون یـک ابر صفحه جداساز است که بر روی داده ها با در نظر گرفتن کمترین مربع خطا بـین ابرصـفحه و داده ها، برازش می یابد. اما ماشین برداری رگرسیونگر[1](SVR) هدف دیگری دارد. ایده اصـلی یـک SVR پیدا نمودن تابعی است که بر داده با کمترین انحراف از کمیتی مانند ε برای هر جفت x_i , y_i بـرازش یابد. در عین حال سعی آن بر این مطلب است که کمترین مقدار ممکن نیز برای نرم w به دست آید. این بدان معنی است که SVR، به خطاهای کمتر از ε کاری ندارد و سعی می کند تا خطاهای بزرگتر

[1] - Support Vector Regressor

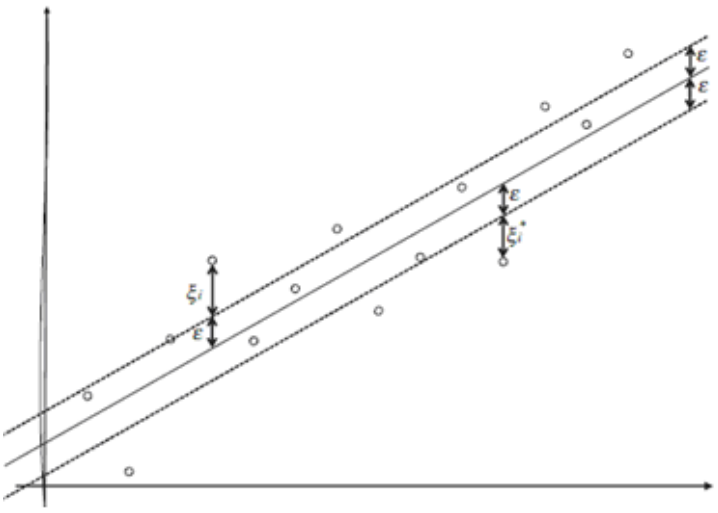
از آن را کوچک نماید. انجام چنین کاری اجازه می دهد تا ماشینی ساخته شود که پارامترهای آن، ترکیبی خطی از نمونه هایی باشد که دارای خطای بیشتر یا مساوی ε هستند. مینیمم سازی نرم w، یک راه حل هموار[1] را تولید خواهد نمود که خطای بیش برازش را مینیمم می سازد. ایده اصلی، یافتن راه حلی است که از کمترین انرژی ممکن داده استفاده کند [۲۶]، [۴۰].

۳-۱۲-۱- فرموله کردن SVR

بر اساس مطالب بیان شده، رگرسیون خطی برای ماشین های برداری می تواند به صورت تابع اولیه (۳-۲۶) فرموله گردد که در آن هدف، مینیمم سازی خطای کلی و نرم w است.

$$L_p = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i') \qquad (۳-۲۶)$$

$$\text{Subject to } \begin{cases} y_i - w^T x - b \le \xi_i + \varepsilon \\ y_i + w^T x + b \le \xi_i + \varepsilon \\ \xi_i, x_i \ge 0 \end{cases}$$



شکل ۳-۱۲: نظریه اصلی حساسیت به مقدار ε. نمونه های خارج از حاشیه $\pm\varepsilon$ متغیرهای خطای غیر صفر (ξ) هستند که قسمتی از راه حل ارائه شده خواهند بود.

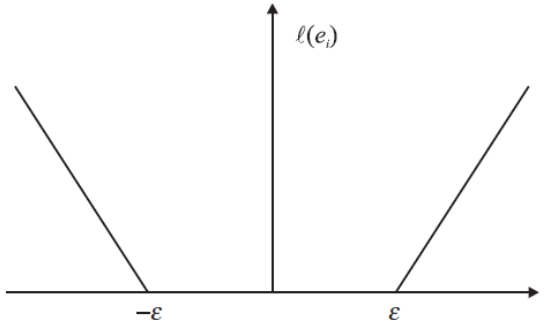
[1] - smooth solution

محدودیت های معادله (۳-۲۶) بدین معنی هستند که خطای بیش تر از ε به کمتر از $\xi_i + \varepsilon$ تبدیل خواهد شد و اگر مقدار خطا کمتر از ε باشد، متغیرهای خطا (ξ_i) مساوی صفر در نظر گرفته می شوند. این، نظریه حساسیت به ε[1] است. این نظریه در شکل ۳-۱۲ نشان داده شده است.

در این نوع از مدل های رگرسیون، تابع هدف مجموع مقادیر خطا را مینیمم خواهد ساخت و تنها نمونه هایی مورد استفاده قرار خواهند گرفت که خطای آنها بیش از ε باشد، بنابراین راه حل، تابعی از این نمونه ها خواهد بود. تابع هزینه[2] (خطا) به کارگرفته شده در این حالت یک تابع خطی است و معادل کاربرد تابع هزینه حساس به ε خواهد بود (شکل ۳-۱۳). تابع هزینه (خطا) در بحث ماشین های برداری به صورت رابطه (۳-۲۷) بیان می گردد.

$$L(e_i) = \begin{cases} 0 & |e_i| < \varepsilon \\ |e_i| - \varepsilon & |e_i| > \varepsilon \end{cases} \qquad (۲۷-۳)$$

این روش مشابه کاربرد ماشین برداری طبقه بندی کننده است. به صورت خاص، خطا باید کمتر از ε شود و نرم پارامترها مینیمم سازی گردد. برای حل این تابع، متغیرهای خطایی به صورت ξ_i در قسمت محدودیت تعریف می شوند تا بتوان به کمک آنها نرم پارامترها را مینیمم سازی نمود [۲۶]، [۴۰]، [۴۱].



شکل ۳-۱۳ : تابع خطای حساس به ε

[1] - ε-insensitivity
[2] - Cost function

۳-۱۲-۲- بهینه سازی کاربردی ماشین برداری رگرسیونگر (SVR)

برای حل مسئله بهینه سازی دارای محدودیت (۳-۲۶)، باید از بهینه سازی لاگرانژ برای تبدیل این معادله به یک معادله بدون محدودیت استفاده نمود. با در نظر گرفتن تابع لاگرانژ و مشتق گیری از تابع هدف بدون محدودیت، نسبت به دو پارامتر w و b، دو معادله به صورت (۳-۲۸) به دست می آیند.

$$w = \sum_{i=1}^{N}(\alpha_i - \alpha_i')x_i$$
$$\sum_{i=1}^{N}(\alpha_i - \alpha_i') = 0 \qquad (۲۸-۳)$$

با قرار دادن رابطه به دست آمده از معادله (۳-۲۸) برای مقدار w، در معادله به دست آمده از تابع لاگرانژ (معادله ۳-۱۴)، معادله اساسی ماشین های برداری رگرسیونگر به صورت معادله (۳-۲۹) نوشته خواهد شد.

$$L_d = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i')x_i^T x_j(\alpha_i - \alpha_j') + \sum_{i=1}^{N}((\alpha_i - \alpha_i')y_i - (\alpha_i + \alpha_i')\varepsilon) \qquad (۲۹-۳)$$
$$\text{Subject to } 0 \le (\alpha_i - \alpha_i') \le C$$

برای یافتن مقدار b نیز می توان از یکی از دو رابطه (۳-۳۰) استفاده نمود.

$$-y_i + w^T x_i + b + \varepsilon = 0$$
$$y_i - w^T x_i - b + \varepsilon = 0 \qquad (۳۰-۳)$$
$$\alpha_i, \alpha_j \prec C$$

۳-۱۳- فضای ویژگی

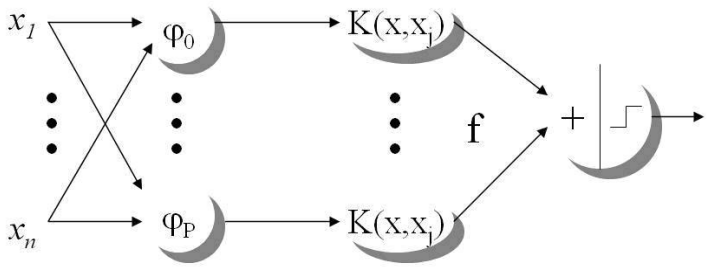
کاربرد اصلی ماشین‌های برداری خاص سیستم های خطی است و صفحه جداساز بهینه تنها برای حالت خطی وجود دارد، اما مشکل اینجاست که در بسیاری از مسائل طبقه بندی و رگرسیون، راه حل خطی جواب مناسبی را ارائه نمی کند و دربسیاری از موارد راه حل غیر خطی نیاز است. خوشبختانه، تئوری مرسر[1] در اوایل سال ۱۹۰۰ ارائه گردید که توانست ماشین های برداری را در سیستم های غیر خطی بسط دهد. ایده اصلی این قضیه، بردن برداری مانند x موجود در فضای محدود (که به فضای ورودی[2] مشهور است) به فضای بالاتر (شاید فضای بینهایت) و طبقه بندی در فضای بالاتر است. این تئوری، تئوری مرسر نام دارد که در آن برداری مانند x در فضای بالاتر به صورت $\varphi(x)$ نوشته می شود. با استفاده از این قضیه، ماشین برداری خطی می تواند در فضای بالاتر ساخته شود (که اغلب فضای ویژگی نامیده می شود) و این در حالی است که فضای ورودی در حالت غیر خطی باقی می‌ماند. این حالت با توجه به خاصیت $\int K(x,y)g(x)g(y)dxdy \succ 0$ قضیه مرسر امکان پذیر می شود. البته در عمل، اغلب این جابجایی ها ناشناخته می مانند و تنها ضرب داخلی فضای منطبق بر آنها می تواند به صورت تابعی از بردارهای ورودی به صورت $\varphi(x_i, x_j) = K(x_i, x_j)$ بیان گردد. این فضا، فضای هیلبرت کرنل[3] (RKHS) نامیده می شود و ضرب داخلی[4] آن کرنل مرسر نام گرفته است. خوشبختانه لازم نیست که بردارها در فضای ویژگی نشان داده شوند و تنها محاسبه ضرب داخلی آنها کافی خواهد بود، بنابراین می توان به راحتی یک ماشین برداری را در فضای هیلبرت نشان داد. در واقع، تئوری مرسر ضرب داخلی $K(x_i, x_j)$ را در فضای ویژگی معرفی می کند. بنابراین، تابعی مانند $\varphi: R_n \to H$ و ضرب داخلی $\varphi(x_i, x_j) = K(x_i, x_j)$ وجود دارد اگر و تنها اگر برای هر تابعی مانند $g(x)$، $\int g(x)dx \prec \infty$، رابطه $\int K(x,y)g(x)g(y)dxdy \geq 0$ برقرار باشد. کرنل های مختلفی

[1] -Mercer
[2] -Input Space
[3] -Reproducing Kernel Hilbert Spaces
[4] -Dot Product

برای استفاده در فضای ویژگی معرفی شده اند که بسته به شـرایط مختلـف مـورد اسـتفاده قـرار مـی گیـرند. این کرنل ها که در واقع ارتبـاط دهنده بین پارامـترهای مدل و تابع هدف هستند در جـدول ۳-۲ آمده اند. عملکرد کرنل در فضای بالاتر نیز در شکل ۳-۱۳ نشان داده شـده اسـت [۲۱]، [۲۶]، [۴۲]، [۴۳]، [۴۴].

جدول ۳-۲ : توابع کرنل در فضای ویژگی

نوع طبقه بندی	تابع کرنل
خطی	$K(x_i, x_j) = (x_i^{\mathrm{T}} x_j)^{\rho}$
چند جمله ای از درجه ρ	$K(x_i, x_j) = (x_i^{\mathrm{T}} x_j + 1)^{\rho}$
گوسین یا نمایی	$K(x_i, x_j) = e^{\frac{\|x_i - x_j\|^2}{2\sigma^2}}$
پیرسن	$K(x_i, x_j) = \frac{1}{\left[1 + \left(\frac{2(x_i - x_j)\sqrt{2^{(\frac{1}{\omega})} - 1}}{\sigma}\right)^2\right]^{\omega}}$
پرسپترون چند لایه	$K(x_i, x_j) = \tanh(\gamma x_i^{\mathrm{T}} x_j + \mu)$
دریکله برای مسائل شرایط مرزی	$K(x_i, x_j) = \frac{\sin((n+1/2)(x_i - x_j))}{2\sin((x_i - x_j)/2)}$



شکل ۳-۱۳: عملکرد کرنل در فضای بالاتر

۳-۱۴- ماشین های برداری غیر خطی

راه حلی که برای ماشین های برداری خطی ارائه شده است، بهینه سازی وزن های ارائه شده برای داده ها با استفاده از رابطه $\omega = \sum_{i=1}^{N} y_i \alpha_i x_i$ است، تا داده ها به درستی در گروه مربوط به خود طبقه بندی شوند. اما در سیستم های غیر خطی برای بردن داده ها به فضای بالاتر، برداری مانند x به صورت $\varphi(x)$ نوشته می شود، در نتیجه برای به دست آوردن وزن مناسب برای داده ها از رابطه $\omega = \sum_{i=1}^{N} y_i \alpha_i \varphi(x_i)$ استفاده می شود. در واقع داده ها در این حالت به فضای بالاتر که به فضای هیلبرت نیز معروف است، برده می شوند. مشکلی که وجود دارد ضرب داخلی بردارهاست که باید محاسبه گردد، بنابراین نمی توان از رابطه $y_i = \omega^{\mathrm{T}} \varphi(x_i) + b$ استفاده نمود زیرا پارامتر w در این حالت، بعدی بی نهایت دارد. در جهت حل این مسئله از توابع کرنل استفاده می شود. بنابراین از رابطه (۳-۳۱) در حالت غیر خطی استفاده می شود.

$$y_i = \sum_{i=1}^{N} y_i \alpha_i \varphi(x_i)^{\mathrm{T}} \varphi(x_j) + b = \sum_{i=1}^{N} y_i \alpha_i K(x_i, x_j) + b \qquad (۳۱-۳)$$

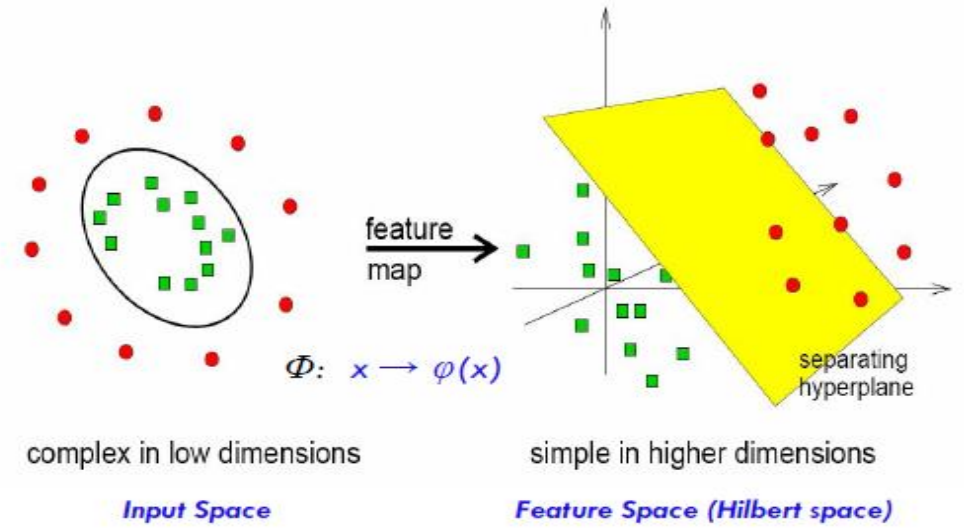
حالا ماشین قادر است تا با استفاده از ضریب لاگرانژ و ضرب داخلی که به صورت توابع کرنل بیان می گردد، شروع به کار کند. بنابراین برای حل توابع دوال، نیاز به ضرایب لاگرانژ و ماتریس کرنل است. با در نظر گرفتن این پارامترها، ماشین های برداری غیر خطی نیز می توانند به سادگی ماشین های برداری خطی مورد استفاده قرار گیرند.

برای محاسبه b، نیز می توان از رابطه موجود در سیستم های خطی که به صورت $y_i(\omega^{\mathrm{T}} x_i + b) - 1 = 0$ بیان می شود، استفاده نمود. این رابطه در حالت غیر خطی به صورت معادلات (۳-۳۲) بیان می گردد.

$$y_i(\sum_{i=1}^{N} y_i \alpha_i \varphi(x_i)^{\mathrm{T}} \varphi(x_j) + b) - 1 = 0 \qquad (۳۲-۳)$$

$$y_i(\sum_{i=1}^{N} y_i \alpha_i K(x_i, x_j) + b) - 1 = 0$$

که در معادله بالا برای تمام x_j باید، $\alpha_j \prec C$ باشد [۲۱]، [۲۶]، [۴۵]، [۴۶]. همان گونه که بیان گردید، برای به دست آوردن مقدار بهینه b می توان از رابطه (۳-۳۲) استفاده نمود، اما از تمام مقادیر به دست آمده باید میانگین گیری شود. چگونگی جداسازی داده ها در فضای ویژگی در شکل ۳-۱۴ مشاده می شود.



شکل ۳-۱۴ : داده های ورودی ارجاع داده شده به فضای بالاتر

بر اساس مطالب بیان شده، معادله اساسی که در اکثر مراجع برای ماشین های برداری غیر خطی ارائه شده است به صورت معادله (۳-۳۳) است. این معادله همان معادله اساسی در سیستم های خطی است و تنها کرنل جای $(x_i^T x_j)$ را گرفته است.

$$\text{Max } L_d(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$

$$\text{Subject to } \begin{cases} 0 \le \alpha_i \le C \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{cases} \qquad (۳۳-۳)$$

۳-۱۵- حقه های کرنل[1]

سوالی که عموما در مبحث های فضای هیلبرت پرسیده می شود محاسبه مقدار $\varphi(x)$ است. بر اساس مطالب بیان شده، مقدار بهینه w در مسائل غیر خطی بر اساس رابطه $\omega = \sum_{i=1}^{N} y_i \alpha_i \varphi(x_i)$ به دست می آید که در این رابطه $\varphi(x)$ مقداری مجهول است. در چنین شرایطی برای به دست آوردن جواب بهینه از حقه کرنل استفاده می شود [۴۷] که در ادامه تشریح خواهد شد.

همان گونه که بیان گردید مقدار بهینه w برای ماشین های برداری در سیستم های غیر خطی از رابطه $\omega = \sum_{i=1}^{N} y_i \alpha_i \varphi(x_i)$ به دست می آید. با در نظر گرفتن این رابطه و قرار دادن آن در رابطه $b = y_i - w^T \varphi(x_i)$ می توان به معادله (۳-۳۴) برای محاسبه مقدار بهینه b دست یافت.

$$b = y_i - \sum_{i,j=1}^{N_{SV}} \alpha_i y_i K(x_i, x_j) \qquad (۳۴-۳)$$

حال می توان معادلات ارائه شده برای w و b را در معادله اصلی صفحه جداساز در سیستم های غیر خطی که به صورت $d(x) = w^T \varphi(x) + b$ بیان می گردد، قرار داد. با در نظر گرفتن شرایط بیان شده، معادله اساسی صفحه جداساز بهینه در سیستم های غیر خطی به صورت معادله (۳-۳۵) بیان می گردد.

$$d(x) = \sum_{i=1}^{N} y_i \alpha_i K(x, x_i) + b \qquad (۳۵-۳)$$

همان گونه که مشاهده می شود، مقدار معادله صفحه جداساز را می توان بدون محاسبه $\varphi(x)$ در سیستم‌های غیر خطی به دست آورد، تنها کافیست که از کرنل مناسب برای حل معادله صفحه استفاده شود [۲۱]، [۲۶]، [۴۸]، [۴۹].

[1] -Kernel Trick

۳-۱۶- آنالیز رگرسیون غیر خطی

رگرسیون غیر خطی، روشی برای یافتن یک مدل غیر خطی مرتبط کننده متغیر وابسته و مجموعه متغیرهای غیر وابسته است. برخلاف رگرسیون های خطی مرسوم که مدل های خطی را تخمین می زنند، رگرسیون غیر خطی می تواند مدلی را با ارتباط دلخواه بین متغیرهای وابسته و غیر وابسته تخمین بزند. انجام چنین کاری توسط الگوریتم های تخمین گری که به صورت مداوم استفاده می شوند، میسر می گردد. باید توجه داشت که این روش برای یک مدل چندجمله ای ساده مانند $Y = A + BX^2$ لازم نیست زیرا با تعریف نمودن $W = X^2$، یک مدل ساده خطی مانند $Y = A + BW$ به دست می آید که می تواند با استفاده از روش های سنتی رگرسیون خطی تخمین زده شود. در هر تکرار آنالیز رگرسیون غیر خطی، پارامترها تخمین زده می شوند و مجموع مجذورات خطا محاسبه می گردد. این خطا با نام های دیگری مانند مجموع مجذورات برای رگرسیون، خطای پس ماند، مجموع خطای عدم صحت، خطای استاندارد نیز عنوان می گردد. در این نوع از رگرسیون، متغیرهای وابسته و غیروابسته باید کمی باشند. متغیرهای اصلی مدل نیز باید به صورت باینری یا دیگر انواع متغیرهای آشکار ساز بیان گردند.

نتایج این آنالیز تنها در زمانی اعتبار لازم را خواهند داشت که تابع بیان شده بتواند به طور دقیق ارتباط بین متغیرهای وابسته و غیر وابسته را توضیح دهد. علاوه براین، انتخاب یک مقدار درست برای آغاز کار بسیار مهم است، زیرا حتی اگر تابع درستی از مدل مورد نظر ساخته شود، اما مقدار در نقطه آغازین اشتباه انتخاب شود، مدل قادر نخواهد بود که به خوبی داده ها را پوشش دهد و راه حل بهینه محلی به جای بهنیه کلی به دست خواهد آمد [۳۹].

۳-۱۷- ماشین برداری رگرسیونگر برای سیستم های غیر خطی

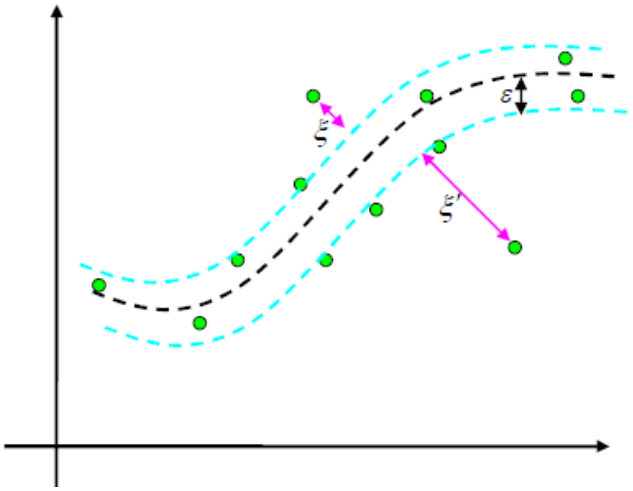
همان گونه که در بحث ماشین های برداری رگرسیونگر در سیستم های خطی گفته شد، مقدار بهینه w از رابطه $w = \sum_{i=1}^{N}(\alpha_i - \alpha_i')x_i$ به دست می آید که این رابطه در سیستم های غیر خطی تبدیل به رابطه $w = \sum_{i=1}^{N}(\alpha_i - \alpha_i')\varphi(x_i)$ می شود. مسئله ای که در این حالت ایجاد می شود، همان مشکلی است که در سیستم های غیر خطی 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 کلاس بندی داده ها از حقه کرنل استفاده می شود تا بتوان بدون محاسبه مقدار $\varphi(x_i)$ و تنها با استفاده از کرنل های موجود، بهینه ترین مدل ریاضی را بر داده ها برازش نمود. بر اساس مطالب بیان شده، رابطه صفحه جداساز در سیستم های غیر خطی به صورت $y_i = \sum_{i=1}^{N} w^T \varphi(x_i) + b$ بیان می گردد. با قراردادن رابطه به دست آمده برای w در رابطه صفحه می توان معادله اساسی (۳-۳۶) را برای ماشین های برداری رگرسیونگر در سیستم های غیر خطی معرفی نمود.

$$y_i = \sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i')\varphi(x_i)^T \varphi(x_j) + b = \sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i - \alpha_i')K(x_i, x_j) + b \qquad (۳۶-۳)$$

با در نظر گرفتن رابطه (۳-۳۶) نیازی به محاسبه مقدار $\varphi(x_i)$ نیست و مقدار b نیز از رابطه ای که در کلاس‌بندی بیان شد، به صورت $b = y_i - \sum_{i,j=1}^{N_{SV}} \alpha_i y_i K(x_i, x_j)$ به دست می آید.

عملکرد ماشین های برداری پشتیبان در قیاس با مدل های دیگری که در سیستم های غیر خطی وجود دارد، حاکی از قدرت و عملکرد بالای این ماشین ها در اکثر سیستم ها به کارگرفته شده،

خصوصا در سیستم های غیر خطی است [۲۱]، [۲۶]، [۵۱]. مدل ریاضی ارائـه شـده توسـط ماشـین برداری برای یک رگرسیون غیر خطی در شکل ۳-۱۵ نشان داده شده است.



شکل ۳-۱۵ : مدل ارائه شده توسط ماشین برداری برای سیستم غیر خطی

۳-۱۸- قدم به قدم برای طبقه بندی با ماشین های برداری پشتیبان

بر اساس آنچه که بیان گردید، می توان مراحل طبقه بنـدی داده هـا بـا اسـتفاده از ماشـین هـای برداری را در ۵ مرحله خلاصه نمود.

۱. ارائه ماتریس الگوی داده ها
۲. انتخاب کرنل مناسب بر اساس مطالعات انجام شده
۳. انتخاب پارامترهای مشخص شده برای هر کرنل و مقدار پارامتر موازنه C
۴. برای مشخص کردن این پارامترها می توان از مقادیر پیشـنهاده شـده توسـط نـرم افزارهـای SVM (LIBSVM, LWSVM,Weka) استفاده نمود یا اینکه این مقادیر را با انجام تعیین اعتبار به دست آورد.
۵. پیاده سازی الگوریتم یادگیری و به دست آوردن مقادیر α_i

۳-۱۹- قدرتمندی ها و ضعف ماشین های برداری پشتیبان

توانمندی های ماشین های برداری که باعث شهرت سریع این روش شده است در ۳ گـروه عمـده بیان می شود:

۱- فرآیند آموزش این نوع از ماشین ها بسیار آسان است (نسبت به شبکه عصبی و سیسـتم فـازی) و ماشین مانند شبکه عصبی در نقطه بهینه محلی گیر نمی‌کند.

۲- در انتقال داده ها به فضای بالاتر به خوبی عمل نموده و عملکرد خوبی در سیستم های غیر خطی دارد.

۳- موازنه بین پیچیدگی طبقه بندی کننده و میزان خطا قابل کنترل است.

تنها مشکل ماشین های برداری که به عنوان نقطه ضعف این ماشین ها بیان شده است، نیاز به انتخاب یک کرنل مناسب است زیرا در صورت عدم مناسب بودن آن، نتایج رضایت بخش نخواهند بـود [۲۱]، [۵۲].

فصل چهارم

مدل سازی داده‌ها با استفاده از

ماشین برداری پشتیبان

۴-۱- مقدمه

در این فصل الگوریتم و کرنل استفاده شده در روش را مورد بحث قرار می‌دهیم و در انتها به پیاده سازی ماشین برداری می‌پردازیم.

۴-۲- الگوریتم SMO

الگوریتم SMO[1] برای آموزش ماشین‌های بـرداری پشـتیبان در سـال ۱۹۹۹ توسـط جـان پلـت[2] تشریح شد. فرآیند آموزش ماشین‌های برداری نیازمند به حل یک مسـئله‌ی بـزرگ درجـه دو (QP)[3] می باشد. الگوریتم SMO با شکستن این مسئله به یک سری از مسائل کوچکتر که به صورت آنالیزی قابل حل هستند، QP را بهینه می‌کند [۵۳]. موضـوع مـورد بحـث در الگـوریتم SMO حـل معادلـه اساسی ماشین های برداری است:

$$\text{Max} \quad L_d(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} y_i y_j \alpha_i \alpha_j x_i^T x_j$$
$$\text{S.t} \begin{cases} \alpha_i \geq 0 \\ \sum_{i=1}^{N} \alpha_i y_i = 0 \end{cases} \qquad (۱-۴)$$

یک نقطه در (۴-۱)، بهینه است اگر و فقط اگر شـرایط KKT را فـراهم آورد [۵۴]. چنـین نقطـه‌ای ممکن است منحصر به فرد نباشد. برای این مسئله شرایط KKT به صورت زیر است:

$$\begin{aligned} &\alpha_i = 0 \Rightarrow y_i(w^T x_i + b) \geq 1 \\ &\alpha_i = C \Rightarrow y_i(w^T x_i + b) \leq 1 \\ &0 < \alpha_i < C \Rightarrow y_i(w^T x_i + b) = 1 \end{aligned} \qquad (۲-۴)$$

[1] Sequential Minimal Optimization
[2] John Platt
[3] Quadratic Problem

بدون از دست دادن کلیت مسئله، در اولین گام الگوریتم به محاسبه‌ی دومین ضربگر لاگرانژ α_2 می‌پردازد. ابتدا α_2 به صورت تصادفی انتخاب می‌شود، اگر y_1 برابر با y_2 نشد، مرزهای زیر برای α_2 به کار برده می شوند.

$$L = \max(0, \alpha_2^{(old)} - \alpha_1^{(old)}), \qquad H = \min(C, C + \alpha_2^{(old)} - \alpha_1^{(old)}) \tag{۳-۴}$$

و اگر y_1 برابر با y_2 شد، مرزهای زیر برای α_2 در نظر گرفته می‌شوند.

$$L = \max(0, \alpha_1^{(old)} + \alpha_2^{(old)} - C), \quad H = \min(C, \alpha_1^{(old)} + \alpha_2^{(old)}) \tag{۴-۴}$$

سپس متغیر η را به صورت زیر تعریف می‌کنیم:

$$\eta = 2k(x_1, x_2) - k(x_1, x_1) - k(x_2, x_2) \tag{۵-۴}$$

در مرحله بعد SMO به محاسبه‌ی ماکزیمم نامقید تابع هدف در معادله (۴-۱) می‌پردازد:

$$\alpha_2^{new} = \alpha_2^{old} - \frac{y^2(E_1 - E_2)}{\eta} \tag{۶-۴}$$

که در آن $E_i = f^{old}(x_i)$ و y_i خطای i مین نمونه‌ی آموزشی است.

سپس $\alpha_2^{new,clipped}$ با ماکزیمم کردن معادله‌ی اساسی، به صورت زیر به‌دست می‌آید:

$$\alpha_2^{new,clipped} = \begin{cases} H, & if\ \alpha_2^{new} \geq H \\ \alpha_2^{new}, & if\ L < \alpha_2^{new} < H \\ L, & if\ \alpha_2^{new} \leq L \end{cases} \tag{۷-۴}$$

اکنون برای محاسبه‌ی α_1 با استفاده از $\alpha_2^{new,clipped}$ و تعریف $s = y_1 y_2$ به صورت زیر عمل می‌کنیم:

$$\alpha_1^{new} = \alpha_1^{old} + s(\alpha_2^{old} - \alpha_2^{new,clipped}) \tag{۸-۴}$$

SMO ضرایب لاگرانژ را به سمت بیشترین مقدار تابع هدف سوق می‌دهد [۵۳], [۵۴].

۴-۳- کرنل عمومی پیرسون

تابع پیرسن در سال ۱۹۸۵ توسط کارل پیرسن[1] گسترش یافت [۵۵]. فرم کلی تابع پیرسن برای برازش منحنی‌های هدف به صورت زیر است :

$$f(x)=\frac{H}{\left[1+\left(\frac{2(x-x_0)\sqrt{2^{(\frac{1}{\omega})}-1}}{\sigma}\right)^2\right]^{\omega}} \qquad (۹-۴)$$

که H ارتفاع رأس در x_0 ، و x متغیر مستقل است. پارامترهای σ و ω به ترتیب نیم-عرض[2] (یا عرض پیرسن) و فاکتور دامنه‌ی[3] رأس نامیده می‌شوند. انعطاف‌پذیری نسبت به تغییر پارامتر ω از شکل گوسین[4] (ω تقریباً نامتناهی) به سمت لورنتزین[5] (ω برابر یک)، دلیل اصلی استفاده از تابع پیرسن برای برازش منحنی‌ها است. به طور کلی سایر مقادیر ω یک گستره‌ی وسیع از اشکال رأسی تولید می‌کند. به طور کلی هر تابع که در شرایط قضیه مرسر[6] صدق کند، متعلق به کلاس توابع کرنل‌های معتبر است [۵۵]. قضیه مرسر ایجاب می‌کند که یک تابع کرنل معتبر باید متقارن باشد، یعنی:

$$K(x,z)=\langle\varphi(x).\varphi(z)\rangle=\langle\varphi(z).\varphi(x)\rangle=K(z,x) \qquad (۱۰-۴)$$

---

[1] -Karl Pearson
[2] -half-width
[3] -tailing factor
[4] -Gaussian shape
[5]- Lorentzian shape
[6]- Mercer's theorem condition

نتیجه می‌شود که ماتریس نمایش کرنل باید نیمه معین[1] مثبت باشد و این یعنی مقادیر ویژه ماتریس کرنل باید نامنفی باشند. با بازنویسی تابع پیرسن به صورت زیر نشان می‌دهیم که این تابع در شرایط قضیه مرسر صادق است و در نتیجه یک کرنل است.

$$K(x_i, x_j) = \frac{1}{\left[1+\left(\frac{2\left(x_i - x_j\right)\sqrt{2^{\left(\frac{1}{\omega}\right)}-1}}{\sigma}\right)^2\right]^{\omega}} \qquad (۱۱-۴)$$

که در آن، بدون از دست دادن کلیت مسئله، متغیر x با دو بردار مستقل x_i و x_j جایگذاری شده و متر اقلیدسی بین این دو بردار معرفی شده است، x_0 برداشته شده و ارتفاع رأس H به طول واحد در نظر گرفته شده است. در این روش برای هر جفت x_i و x_j، تابع پیرسن منجر به تولید یک ماتریس متقارن با درایه های یک و سایر درایه‌ها بین ۰ و ۱ تغییر می‌کنند. بنابراین تابع پیرسن به صورت ماتریس متقارن نمایش داده شد. ماتریس متقارن ($m \times m$) A نیمه معین مثبت نامیده می‌شود، اگر برای هر بردار x $x^T Ax \geq 0$ باشد (تمام مقادیر ویژه A نامنفی باشند). در عمل، چک کردن این تعریف به طور مستقیم بسیار مشکل و با محاسبات بسیار زیاد همراه است. رسیدگی کردن به دو شرط ریاضی زیر که در شرایط ماتریس‌های متقارن حفظ می‌شوند بسیار ساده‌تر و مفیدتر است[۵۶], [۵۷].

۱. کهادهای اصلی پیشتاز[2] نامنفی هستند، یعنی

$$a_{11} \geq 0, \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \geq 0, \ldots, \det A \geq 0$$

که det نشان‌دهنده دترمینان ماتریس است.

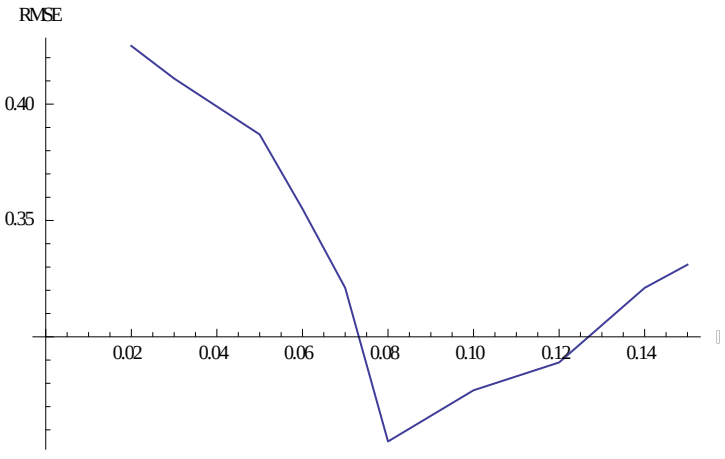
[1] -semi-definite
[2] -leading principal minors

۲. $tr(A) \geq 0$ که tr نشان‌دهنده اثر[1] ماتریس است.

اگر یکی از این شرایط اتفاق بیافتد، شرایط قضیه مرسر حفظ می‌شوند. به طور آنالیزی اثبات شده است که معادله (۵-۳) این خاصیت‌ها را برآورده می‌کند[۵۷], [۵۸].

۴-۴- پیاده سازی ماشین برداری پشتیبان

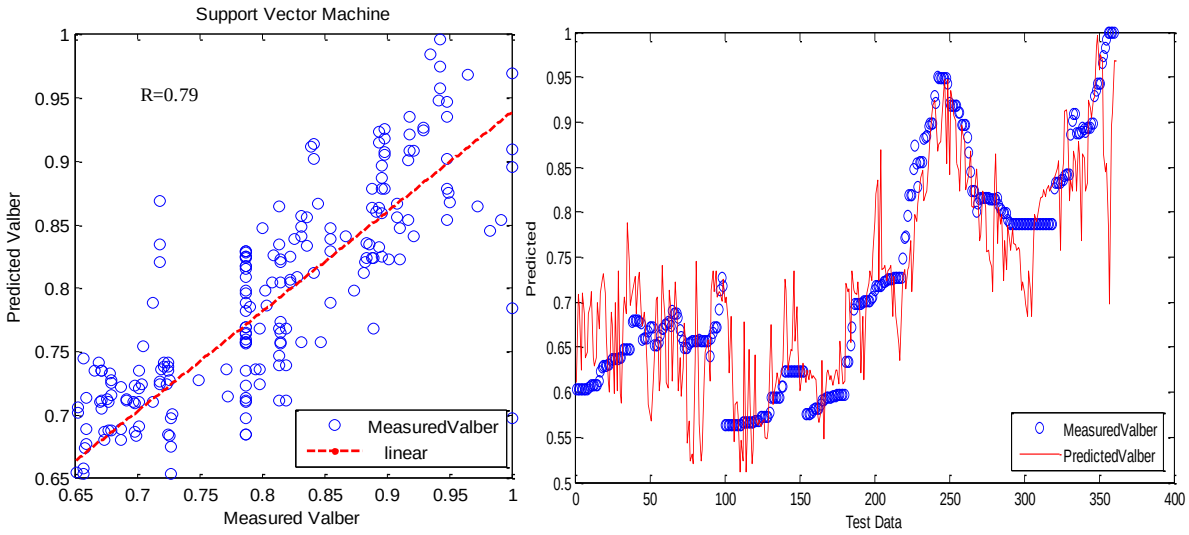
همانند سایر روش‌های آماری چند متغیره، عملکرد SVM برای رگرسیون به ترکیب چندین پارامتر بستگی دارد. آنها در SVR پارامترهای ظرفیت C و پارامتر تابع جریمه (ε) هستند. پارامتر C برای تنظیم حاشیه‌ها و کاهش خطای آموزش داده‌ها مورد استفاده قرار می‌گیرد که در این مطالعه برای آن مقدار ۱۰۰ در نظر گرفته شد. این مقدار توسط بسیاری از محققین توصیه شده است، بنابراین، ما نیز از این مقدار استفاده کردیم. مقدار ضریب ε از قرار گرفتن داده‌های آموزشی بر روی شرایط مرزی جلوگیری می‌کند و در نتیجه امکان پراکندگی در مسئله دوگان را فراهم می‌کند. بنابراین انتخاب یک مقدار مناسب از ε یک امر حیاتی برای الگوریتم است. مقدار این پارامتر نیز به روش آزمون و خطا، ۰/۰۸ مشخص گردید. شکل ۴-۱ این فرآیند را نشان می دهد.



شکل ۴-۱: چگونگی تعیین ε در فرآیند سعی و خطا

[1] - trace

همانگونه که در شکل ۴-۱ مشاهده می شود، ما در بازه ۰ تا ۱، در فواصل ۰,۰۱ به پیش رفتیم تـا بهترین مقدار ε برابر ۰,۰۸ تعیین گردید که دارای حداقل خطای ممکن است. پـس از بهینـه سـازی پارامترهای کرنل پیرسن به روش آزمون و خطا ω و σ به ترتیب ۱۰۰ و ۱,۲ بدست آمد و در نهایـت نتایج ماشین برداری پشتیبان به صورت شکل ۴-۲ به دست آمد.



شکل ۴-۲ : عملکرد ماشین برداری پشتیبان در پیش بینی شاخص بورس

همان گونه که مشاهده می شود ماشین برداری پشتیبان عملکرد خوبی در پیش بینی نرخ بازار بورس دارد. در واقع این مقدار پیش بینی بهترین پیش بینی ممکنی است که تاکنون به دست آمده است که حاکی از قدرت بالای ماشین برداری پشتیبان با به کارگیری الگوریتم SMO و کرنل پیرسن است. نتایج حاصل از کرنل‌های مختلف این روش در جدول ۴-۱ آمده است.

جدول ۴-۱: نتایج بدست آمده از کرنل های روش

خطا (RMSE)	تابع کرنل
0.96	$K(x_i, x_j) = (x_i^{\mathrm{T}} x_j)^{\rho}$
0.91	$K(x_i, x_j) = (x_i^{\mathrm{T}} x_j + 1)^{\rho}$
0.77	$K(x_i, x_j) = e^{\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}}$
0.65	$K(x_i, x_j) = \frac{1}{\left[1 + \left(\frac{2(x_i - x_j)\sqrt{2^{(\frac{1}{\omega})} - 1}}{\sigma}\right)^2\right]^{\omega}}$
0.76	$K(x_i, x_j) = \tanh(\gamma x_i^{\mathrm{T}} x_j + \mu)$
0.82	$K(x_i, x_j) = \frac{\sin((n+1/2)(x_i - x_j))}{2\sin((x_i - x_j)/2)}$

فصل پنجم

شبکه عصبی رگرسیون عمومی

۵-۱- مقدمه

همان گونه که در فصل قبل بیان گردید، جهت بارزسازی قابلیت ماشین برداری پشتیبان در پیش بینی شاخص بازار بورس، از شبکه عصبی استفاده خواهد شد. شبکه های عصبی مصنوعی با الهام از سیستم عصبی انسان به وجود آمده اند. این شبکه ها توانایی یادگیری دارند و قادرند روابط (الگوهای) بین ورودی ها وخروجی (ها) را تشخیص دهند. پس از آموزش مناسب، می توان ورودی های جدید به شبکه عصبی ارائه نمود و خروجی را تخمین زد. این تکنیک می تواند روابط غیر خطی بین ورودی ها و خروجی را در هر نوع مساله تخمینی بشناسد.

شبکه عصبی استفاده شده در اینجا شبکه عصبی رگرسیون عمومی[1] است که با توجه به کاربردها و قابلیت هایی که از خود نشان داده، یکی از شبکه های قدرتمند به حساب می آید. در ادامه مختصری از قابلیت ها و مسائل مناسب جهت استفاده از شبکه های مصنوعی می آوریم و در انتها پس از معرفی شبکه عصبی رگرسیون عمومی، به بررسی قابلیت این شبکه در پیش بینی شاخص بازار بورس خواهیم پرداخت.

۵-۲- قابلیت های شبکه های عصبی مصنوعی

۱- محاسبه یک تابع معلوم

۲- تقریب یک تابع مجهول

۳- شناسایی الگوها[2]

۴- پردازش سیگنال

۵- یادگیری انجام مراحل فوق

[1] -General regression Neural network
[2] -Pattern Recognition

در مرحله آموزش ممکن است دو دسته داده در اختیار شبکه قرار داده شود، یک دسته، داده هایی است که ویژگی آنها شناخته شده است و تابع مشخصی برای آن وجـود دارد کـه ایـن تـابع را شـبکه محاسبه خواهد نمود. اما در دسته دوم، ویژگی داده هایی که در اختیار شبکه قرار می گیرند شناخته شده نیست، در این حالت شبکه یک تابع را برای این داده ها پیش بینی می کند که به تابع مجهـول معروف است. شناسـایی و طبقه بنـدی الگوها یکی از مهمتریـن ویژگـی های شبکه هـای مصنوعـی است. هدف اصلـی شناسایـی الگوها، طبقه بندی آنها است. شنـاسایـی الگو را می توان در دو مرحلـه خلاصه نمود. در مرحله اول، استخراج مشخصه ها صورت می پـذیرد و در مرحلـه دوم، مشـخه هـای استخراج شده طبقه بندی می شوند. مشخصه[1] کمیتی است که سیستم مورد نظر را مـی سـازد و بـر اساس آن می توان سیستم مورد نظر را طبقه بندی نمود. در شناسایی الگـو، عمومـاً مشخصـه هـایی مورد جستجو قرار می گیرند که قادر باشند تا سیستم را به صورت مطمـئن طبقـه بنـدی کننـد. بـه عنوان مثال اگر مسئله شناسایی نوشتار مد نظر باشد. برای تشـخیص حـرف (F) از حـرف (E) تعـداد خطوط عمودی و افقی مورد مقایسه قرار می گیرد. البته، استخراج مشخصه هـا و ویژگـی هـای یـک سیستم به ندرت به این سادگی خواهد بود و اغلب بخش عمده طرح شناسایی را به خود اختصاص می دهد. مشخصه ها در اختیار دستگاه های طبقه بندی کننده قرار می گیرند. وظیفه دستگاه های طبقه بندی کننده انعکاس این مشخصه ها در فضای طبقه بندی است. به عبارت دیگر بـا داشـتن مشخصـه های ورودی، شبکه باید تصمیم بگیرد که سیستم های ارائه شده با کدام کلاس تطابق بیشتری دارنـد [۲۱]، [۲۵].

۵-۳- مسائل مناسب جهت استفاده از شبکه های مصنوعی

۱- خطا در داده های آموزشی وجود داشته باشد. به عنوان مثال می توان بـه داده هـای آموزشـی دارای نویز حاصل از داده های سنسورد نظیر میکروفون و دوربین اشاره نمود.

[1] -Feature

۲- مواردی که در آنها نمونه ها دارای تعداد زیادی زوج ویژگی- مقدار باشند.

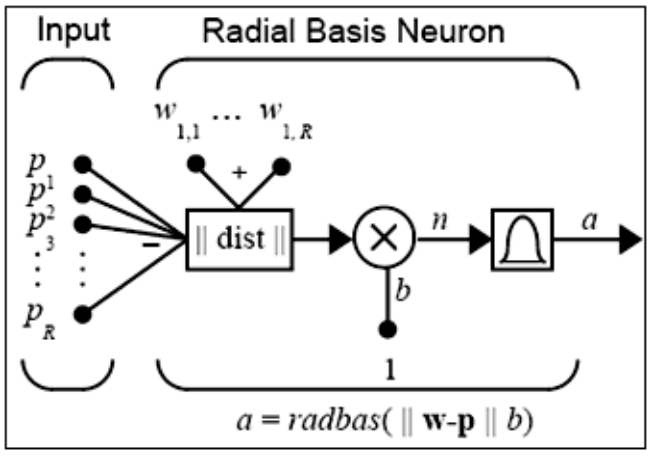
۳- تابع هدف دارای پیوستگی باشد.

۴- زمان کافی برای یادگیری شبکه وجود داشته باشد.

۵- نیازی به تعبیر و تفسیر تابع هدف نباشد زیرا به سختی می توان اوزان یادگرفته شده توسط شبکه را تفسیر نمود [۲۱]، [۲۲].

۵-۴- تئوری شبکه های عصبی رگرسیون عمومی

شبکه های عصبی مصنوعی در واقع تقلید بسیار ساده ای از رفتار سلول های بیولوژیکی می باشند که از واحدهایی به نام نرون تشکیل شده اند. یک شبکه ساده شامل لایه ورودی، لایه میانی، لایه خروجی و تابع محرک است. لایه ورودی سیگنال ها را از محیط خارج دریافت می کند. تابع محرک سیگنال های ورودی را جمع و پردازش می کند. به محض ورود نمونه های آموزشی به داخل شبکه ، نرون ها در داخل شبکه مقادیر تابع محرک را محاسبه کرده و این مقادیر را بر اساس دستوری که به الگوریتم مورد استفاده بستگی دارد به نرون های دیگر منتقل می کنند. شبکه های شعاعی زمانی که بردارهای ورودی بیشتری وجود داشته باشد بهتر کار می کنند .در شکل ۵-۱ شبکه ای شعاعی با ورودی نشان داده شده است.



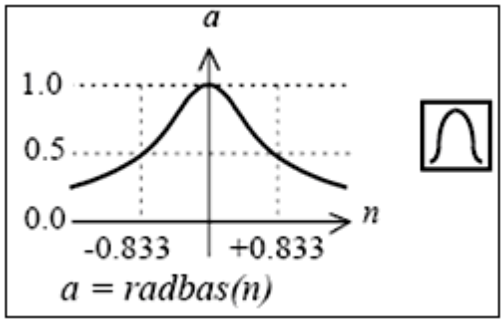
شکل ۵-۱: مدل نرون شبکه عصبی شعاعی

در اینجا ورودی شبکه برای تابع محرک شعاعی، فاصله برداری بین بردار وزن w و بردار ورودی p ضرب در بـایاس اسـت. جعبـه ورودی ||dist|| در این شکل بردار ورودی p و ماتریس تک ردیفه وزن w را می‌گیـرد و تولید ضرب نقطـه‌ای از آن دو می‌کند. تابـع محرک نرون شعاعـی به صـورت زیـر می باشد.

$$Radbas(n) = e^{-n^2} \qquad (۱-۵)$$

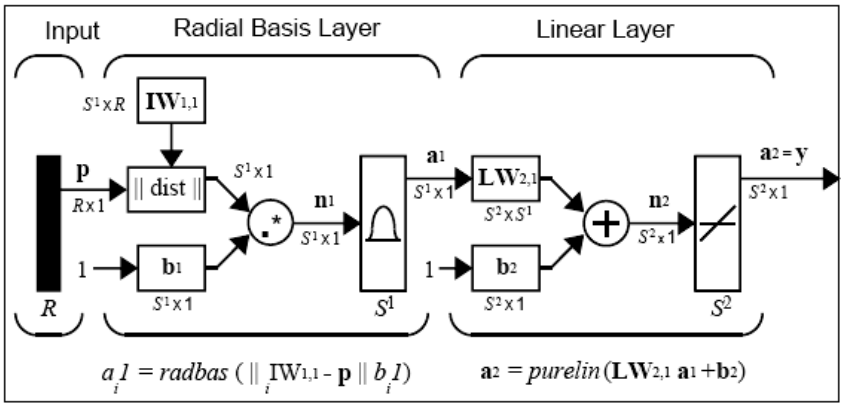
که در آن n ورودی نرون و e تابع نمایی است.

در شکل ۵-۲ نمودار تابع محرک شعاعی نشان داده شده است.



شکل ۵-۲: تابع محرک استفاده شده در شبکه عصبی

وقتی مقدار ورودی فاصله بردار ورودی و بردار وزن برابر صفر باشد، خروجی تابع محرک شعاعی مقدار ماکزیمم ۱ است. وقتی فاصله w و p کاهش می‌یابد، مقدار خروجی افزایش می‌یابد. بنابراین یک نرون شعاعی شبیه یک آشکارساز، وقتـی ورودی p کاملا با برابر با وزنـش شود، مقدار ۱ را تولیـد می‌کند. مقدار بایاس b حساسیت نرون شعاعی را تعدیل می‌کند. شبکه های شعاعی شامل دو لایه هستند: یک لایه شعاعی پنهان با S1 نرون و یک لایه خطی خروجی با S2 نرون (شکل ۵-۳).



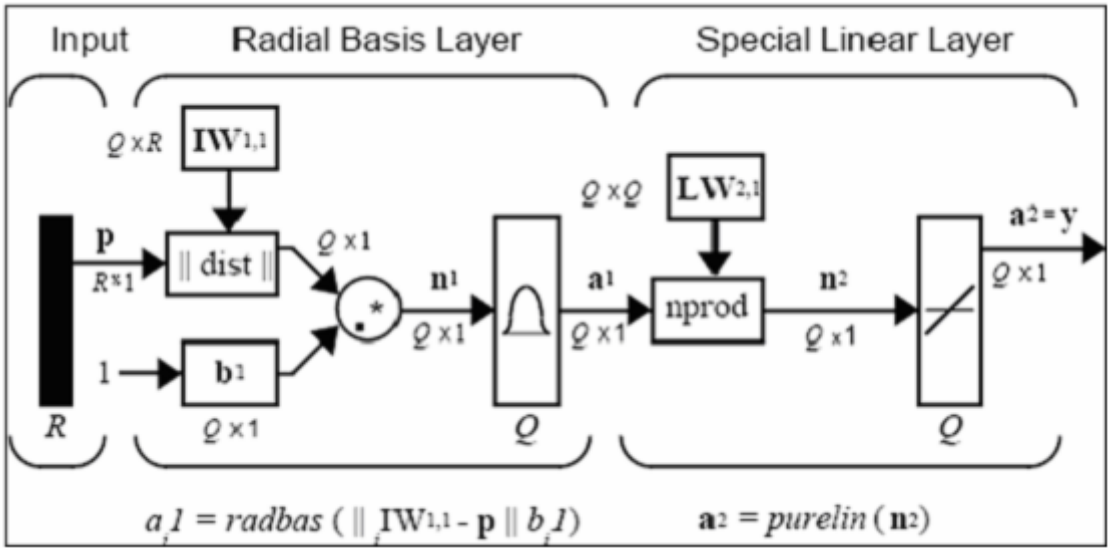
شکل ۵-۳: ساختار شبکه عصبی شعاعی

که R تعداد عناصر بردار ورودی، S1 تعداد نرون‌ها در لایه ۱، S2 تعداد نرون ها در لایه ۲، a_i^1: iامین عنصر از a1 و $i^{WI_{1,1}}$: iامین سطر از بردار IW1,1 است. جعبه ||dist|| در این شکل بردار ورودی p و بردار وزنی ورودی IW1,1 را دریافت و برداری تولید می‌کند که دارای S1 است. این عناصر، فواصل بین بردار ورودی و عناصر $i^{WI_{1,1}}$ تشکیل شده از سطرهای ماتریس وزن ورودی است. بردار بایاس b1 و خروجی ||dist|| نیز از طریق ضرب داخلی با هم ترکیب می‌شوند [۵۹]. خروجی لایه اول برای یک شبکه پیش خور net، از طریق کد زیر حاصل می‌شود:

a{1} = radbas (netprod (dist (net.IW{1,1},p), net.b{1}))

خوشبختانه مجبور نیستیم کدهایی مثل خط فوق را بنویسیم. تمام جزئیات این شبکه به صورت پیش برنامه وجود دارند. اگر یک بردار ورودی به چنین شبکه ای داده شود، هر نرون در لایه پنهان شعاعی، خروجی را با توجه به نزدیکی بردار ورودی به بردار وزن، ارائه می‌دهد. نرون های شعاعی دارای بردارهای وزن کاملاً متفاوت از بردار ورودی p، خروجی نزدیک به صفر را می‌دهند. بنابراین ورودی های کوچک فقط تأثیر جزئی بر روی نرون های خروجی خطی دارند. در مقابل نرون شعاعی، با بردار وزن نزدیک به بردار ورودی p، مقدار خروجی نزدیک به ۱ تولید می‌کند [۶۰].

این شبکه را می‌توان به عنوان یک شبکه شعاعی نرمال شده در نظر گرفت که برای هر واحد آموزشی یک نرون پنهان دارد که در سال ۱۹۹۰ توسط اسپچت[1] ابداع شد و قادر به تولید خروجی‌های پیوسته می‌باشد [۶۱], [۶۲]. این شبکه ها بر اساس تابع چگالی احتمال پایه گذاری شده و از ویژگی های بارز آن، زمان آموزش سریع و مدل سازی توابع غیرخطی است.GRNN حتی با داده های پراکنده در فضای اندازه‌گیری چند بعدی، تغییرات همواری از داده‌ها فراهم می‌کند. فرم الگوریتم این شبکه برای هر مسئله رگرسیونی در جایی که هیچ گونه فرضیاتی برای قضاوت خطی بودن وجود نداشته باشد می‌توانند مورد استفاده قرار گیرد. این شبکه پارامترهای پس انتشار خطا را ندارد ولی در عوض فاکتور هموار ساز[2] دارد که مقدار بهینه آن در طی چندین اجرا با توجه به میانگین مربعات خطا بدست می آید [۶۲]. همانطور که در شکل (۵-۴) نشان داده شده است، ساختار این شبکه شبیه به ساختار کلی شبکه شعاعی است فقط تفاوت جزئی در لایه دوم دارد [۶۰].

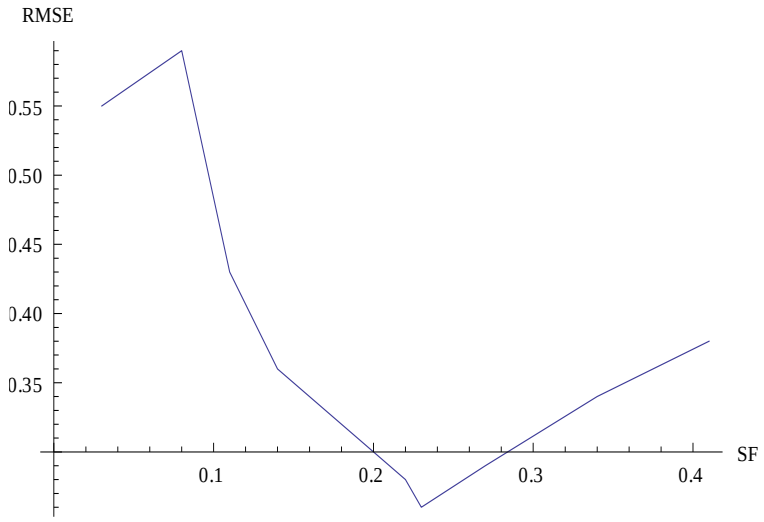


شکل ۵-۴: ساختار شبکه عصبی رگرسیون عمومی

[1] - Specht
[2] - Smooth factor

۵-۵- پیاده سازی شبکه عصبی رگرسیون عمومی

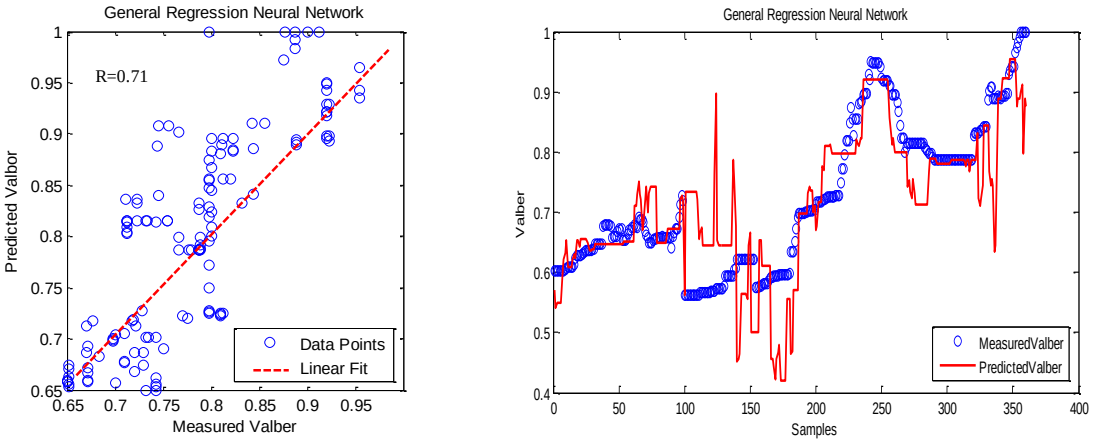
در قسمت قبل، به بررسی و توضیح خصوصیات شبکه رگرسیون عمومی پرداخته شد. در این بخش به پیاده سازی این شبکه پرداخته می شود. همان گونه که در فصل پیاده سازی ماشین برداری پشتیبان توضیح داده شد، از داده های سال ۸۷ برای آموزش شبکه استفاده شد و پس از رسیدن به قدرت پیش بینی ۰,۹۶ در آموزش به فرآیند آزمون شبکه پرداخته شد. لازم به توضیح است که در علوم مالی تحقیقات بر روی سال مالی انجام می شود، به همین دلیل عملیات آموزش شبکه بر روی سال ۸۷ و تست بر روی داده های سال ۸۸ انجام گردید. نکته قابل ذکر در فرآیند آموزش این است که این شبکه نیاز به تعیین فاکتور هموار ساز[1] که همان $\|\text{dist}\|$ است، دارد. آموزش شبکه برای پیش بینی شاخص سهام شرکت سرمایه گذاری البرز با فاکتورهای مختلف هموارساز انجام گرفت. با مقدار دهی این فاکتور وزن هـای هـر ورودی برای تخمـین خروجی مشخص می‌گردند، که در نهایت بعد از فرآیند سعی و خطا ۰,۲۳ تعیین گردید.



شکل ۵-۵: تعیین فاکتور هموارساز در فرآیند سعی و خطا

[1] -Smooth Factor

پس از این مرحله، از شبکه ساخته شده برای پیش بینی شاخص بازار بورس سال ۸۸ استفاده شد. قدرت پیش بینی و عملکرد این شبکه در فرآیند آزمون در شکل ۵-۶ مشاهده می شود.



شکل ۵-۶ : عملکرد شبکه عصبی رگرسیون عمومی در پیش بینی شاخص بورس

همان گونه که در شکل مشاهده می شود، شبکه عصبی رگرسیون عمومی قابلیت خوبی در پیش بینی شاخص بورس دارد اما قدرت آن به خوبی ماشین برداری پشتیبان نیست. علاوه بر این، ماشین برداری این پیش بینی را در طی ۲ ثانیه انجام داد. حال آنکه شبکه عصبی رگرسیون عمومی این فرآیند را در ۴ ثانیه انجام داد که نشان از ضعف الگوریتم روش دارد.

در فصل بعد که فصل نتیجه گیری است با مقایسه خطای دو روش در فرآیند پیش بینی، قابلیت این دو شبکه را بهتر با یکدیگر مقایسه خواهیم کرد.

فصل ششم

نتیجه‌گیری و پیشنهادات

۶-۱- مقدمه

همان گونه که در طی این مطالعه بیان گردید، قابلیت دو شبکه مصنوعی قدرتمند با نام های ماشین برداری پشتیبان و شبکه عصبی رگرسیون عمومی در پیش بینی شاخص بازار بورس با یکدیگر مقایسه شدند. در ادامه در این فصل نتایج و پیشنهادات ارائه گردیده است.

۶-۲- نتیجه گیری

ماشین برداری پشتیبان یکی از قویترین روش های طبقه بندی و پیش بینی است که بر اساس الگوریتم های بهینه سازی، مدل سازی آماری و توابع کرنل پایه گذاری شده است. در این پایان نامه سعی شد تا بر اساس قابلیت های این روش از آن در پیش بینی شاخص بازار بورس استفاده شود. کرنل‌های بسیاری بررسی و تحلیل شدند، که کرنل پیرسون بهترین نتیجه‌ها را در پی داشت. نتایج این بررسی نشان می دهد که ماشین برداری توانایی بالایی در پیش بینی شاخص بورس دارد و همبستگی در حدود ۰,۷۹ را برای داده های نمونه ارائه می کند. برای انجام مقایسه نیز از شبکه عصبی رگرسیون عمومی استفاده شد که یکی دیگر از شبکه های قدرتمند است. نتایج این بررسی نشان داد که شبکه عصبی رگرسیون عمومی نیز قادر به ارائه پیش بینی مناسبی است اما این پیش بینی به خوبی پیش بینی ماشین برداری پشتیبان نبود. جدول ۶-۱ میزان خطای این دو روش را در فرآیند آموزش و آزمون نشان می دهد.

جدول ۶-۱ میزان خطای دو شبکه استفاده شده در فرآیند آموزش و آزمون

نوع شبکه	R (آموزش)	R (آزمون)	خطای آموزش شبکه (RMSE)	خطای آزمون شبکه (RMSE)
ماشین برداری پشتیبان	۰,۹۹	۰,۷۹	۰,۲۲	۰,۶۵
شبکه عصبی رگرسیون عمومی	۰,۹۶	۰,۷۱	۰,۳۴	۰,۷۷

همان گونه که در جدول ۶-۱ مشاهده می شود، ماشین برداری پشتیبان خطای کمتر و عملکرد بهتری در فرآیند پیش بینی دارد.

۶-۳- پیشنهادات

روشهای نوین پیش‌بینی در مدل‌هایی با رفتار آشوبگونه استفاده و تلفیق دو یا چند روش است. از روش کلاس‌بندی ماشین‌های برداری (SVC[1]) هم می‌توان به عنوان یک مدل استفاده کرد. البته با توجه به محدودیت‌های این روش بر روی داده ها می‌توان از تلفیق این روش و سایر روش‌های هوش مصنوعی از جمله الگوریتم ژنتیک استفاده کرد. بدین ترتیب که با استفاده از الگوریتم ژنتیک به داده‌ها درجه تأثیر نسبت می‌دهیم، سپس با استفاده از روش کلاس‌بندی ماشین‌های برداری داده‌ها را در طبقه‌های مفروض بنشانیم. از بُعد مسائل مالی و اقتصادی می‌توان متغیرهای بیشتری به سیستم داد و برای کاهش ریسک بازار سبد سهام تشکیل داد که در این صورت می توان به مقایسه سبدهای سهام و میزان سود دهی آنها پرداخت. لازم به توضیح است که این پایان نامه بر روی سال ۸۷ و ۸۸ انجام شد که برای تحقیق بر روی میزان تعمیم روش می توان SVM را بر روی سال های قبل و حتی

[1] Support Vector Classification

برای هر ماه انجام داد و نتایج را تحلیل کرد. در انتها پیشنهاد می شود که از روش مورد استفاده در این تحقیق که قابـلیت انعطـاف بالایی دارد، و با تغییر در کرنل‌ها و الگوریتم های جدیدی که ابداع می‌شود، در پیش بینی و مدل سازی سیستم های اقتصادی و سیستم های مشابه استفاده گردد.

Abstract

Designing a machine which is capable of learning and experiencing was one of the most important discussions of scientific researches. it was proved that this type of machine is able to learn the tend and variation of data and can provide an acceptable result. Support Vector Machine (SVM)was introduced in 1990s and successfully used to solve regression and classification problems. based on the theory mentioned, the reasons of success in this machine are: 1) generalization ability, 2) robustness against the noise and, 3) solving the problems with limited number of data.

منابع:

[1] Weiss, Gray (1992). Chaos Hits wall Street-the Theory, that is, Business week November. PP. 138-140.

[2] Bishop C. M. (2006). Pattern Recognition and Machine Learning, Springer.

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

[4] Burbidge R, Trotter M, Buxton B, Holden S. (2001). Drug design by machine larning: support vector machines for pharmaceutical data analysis. Computers and Chemistry 26, 5–14

[5] Deng-Yiv Chiu, Ping-Jie Chen (2009). Dynamically exploring internal mechanism of stock market by fuzzy-based support vector machines with high dimension input space and genetic algorithm Department of Information Management, ChungHua University, No. 707, Section 2, WuFu Road, Hsin-Chu 300, Taiwan, ROC

[6] Huang, W., Nakamori, Y., & Wang, S. Y. (2004). Forecasting stock We propose a model integrating fuzzy theory, GA and market movement direction with support vector machine.Computers SVM to explore movements in stock market in Taiwan.and Operations Research (Vol. 32). Elsevier SciOxford, pp. 2513–2522.

[7] Jones, P. (1999). Investments; Analysis and management, Jane Wiley and Sons, Inc 7th Edition, New York. PP. 300-380

[۸] انواری رستمی، علی اصغر (۱۳۷۸). مدیریت مالی و سرمایه گذاری، طراحان نشر، چاپ اول، ص ۲۳۹.

[9] Fama. E .F(1970). Efficient Capital Market : A Review of Theory and Empirical Work. The journal of Finance, No.2, PP. 383-417.

[۱۰] نمازی، محمد و شوشتریان، زکیه.(۱۳۷۵). بررسی کارایی بازار بورس اوراق بهادار ایران. مجله تحقیقات مالی، شماره‌های: ۷و۸، بهار و تابستان، صفحات: ۱۰۴-۸۲.

[11] Conard, J.S. Hameed, A.& Nidlen, C.(1994). Volume and Auto Covariances in Short Horizon Individual Security Returns. Journal of Finance, No. 4, PP.1305-1330.

[12] Kim, M. J. Nelson, C.R. & Startz, R.(1991). Mean Reversion in Stock Prices? A Reapptaisal of the Empirical Evidence. Review of Economic Studies, Vol.58(3),

[13] Dockery , E. Kauussanos, M.G(1996). Testing the Efficient Maeket Hypothesis Using Panel Data, with Application to the Athens Stock Market. Applied Economics Letters, No.3, PP.121-123.

[14] Donaldson, R.G(1990). Is the London Stock Exchange a Rationally Efficient Market. Economic Organization and Public Policy, 58, Octobr.

[15] Stock, D,(1988). Emirical Tests of the Over Reaction Hypothesis for the German Stock Market. Discussion Paper, Universal Bonn Onderforschungsbereich, 303, November.

[16] Mills, C.(1994) Interpreting Evidence of Predictable Variation in Stock and Bond Returns. Durham: Duke University.

[17] Lock, D.B.(1995). Short Horizon Returns on the Taiwan Market and the Random Walk: An Emprical Study. Dissertation, The University of Alaboma.

[۱۸] در امامی، علی اصغر.(۱۳۶۹). بررسی نوسان‌پذیری و ریسک سهام پذیرفته شده در بورس تهران. پایان نامه کارشناسی ارشد مدیریت بازرگانی، دانشگاه تهران، دانشکده مدیریت.

[۱۹] نصراللهی، زهرا.(۱۳۷۱). تجزیه و تحلیل عملکرد بورس اوراق بهادار ایران. پایان نامه کارشناسی ارشد اقتصاد، دانشگاه تربیت مدرس.

[۲۰] فدایی نژاد، محمد اسماعیل.(۱۳۷۴). آزمون شکل ضعیف کارای سرمایه و بورس اوراق بهادار تهران. مجله تحقیقات مالی، شماره‌های: ۵و۶ سال دوم، زمستان و بهار، صفحات: ۲۶-۳.
[21] Duda R. A., Hart P. E. & Stork D. G. (2002). Pattern Classification, Springer.
[22] Bishop C. M. (2006). Pattern Recognition and Machine Learning, Springer.
[23] van der Heijden, F., Duin, R.P.W., de Ridder, D., Tax D.M.J. (2004). Classification, parameter Estimation and State Estimation, John Wiley & Sons Ltd.
[24] Yu Hen H., Jenq-Neng H. (2002). Handbook of NEURAL NETWORK SIGNAL PROCESSING, CRC PRESS.
[25] Theodoridis S., Koutroumbas K. (2006). Pattern recognition, Elsevier.
[26] Martinez-Ramon, M. (2006). Support Vector Machines for Antenna Array Processing and Electromagnetic, Universidad Carlos III de Madrid, Spain, Morgan & Claypool, USA.
[27] Cristianini N., Shawe-Taylor J. (2000). An Introduction to Support Vector Machines (and other kernel-based learning methods), Cambridge University Press, UK.
[28] Abe S, (2008). Support Vector Machines for Pattern Classification, Kobe University, Kobe, Japan.
[29] Wang L. (2005). Support Vector Machines: Theory and Applications, Nanyang Technological University, School of Electrical & Electronic Engineering, Springer Berlin Heidelberg New York.
[30] Steinwart I. (2008). Support Vector Machines, Los Alamos National Laboratory, information Sciences Group (CCS-3), Springer Science+Business Media, LLC.
[31] Law M. (2009). A Simple Introduction to Support Vector Machines, Lecture for CSE 802, Department of Computer Science and Engineering, Michigan State University.
[32] Park Ch. (2008). Support Vector Machine, Advanced Metrology Lab, Texas A&M University.
[33] Burges Ch. J.C. (1998). A tutorial on support vector machine for pattern recognition, bell laboratories Data Mining and Knowledge Discovery, Kluwer Academic Publishers.
[34] Brown M. Lectures 11&12: Generalization, Regularization and Support Vector Machines, http://www.eee.manchester.ac.uk/intranet/pg/coursematerial/
[35] Breckon T. P. (2007). Support Vector Machines, School of Engineering Cranfield University, UK.
[36] Moore A. W. (2001). Support Vector Machines, Associate Professor School of Computer Science, Carnegie Mellon University.
[37] Weston J. Support Vector Machine (and Statistical Learning Theory) Tutorial, NEC Labs America 4 Independence Way, Princeton, USA. jasonw@nec-labs.com
[38] Gunn S. R. (1998). Support Vector Machines for Classification and Regression, Faculty of Engineering, Science and Mathematics, School of Electronics and Computer Science, University of Southampton.
[39] Chattererjee S. (2006). Regression analysis by example, John wily & Sons.
[40] Quang-Anh, T., Xing L., Haixin D., (2005). Efficient performance estimate for one-class support vector machine, Pattern Recognition Letters 26, pp1174–1182.
[41] Sanchez D. V. (2003). Advanced support vector machines and kernel methods, Neurocomputing 55, pp 5 – 20.
[42] Huang Te-Ming., Kecman V. (2006). Kernel Based Algorithms for Mining Huge Data Sets, Faculty of Engineering The University of Auckland, Springer-Verlag Berlin Heidelberg.

[43] Stefano M., Giuseppe J. (2006). Terminated Ramp–Support Vector Machines: A nonparametric data dependent kernel, Neural Networks 19, pp 1597–1611.
[44] Suykens J. (2006). Engineering Kernel Machines, K.U. Leuven, ESAT-SCD/SISTA, Belgium.
[45] Chih-Hung W., Gwo-Hshiung T., Rong-Ho L. (2009). A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression, Expert Systems with Applications 36, pp 4725–4735.
[46] Agarwala S., Vijaya Saradhib V., Karnick H. (2008). Kernel-based online machine learning and support vector reduction, Neurocomputing 71, pp 1230–1237.
[47] Hwei-Jen L., Jih Pin Y. (2009). Optimal reduction of solutions for support vector machines, Applied Mathematics and Computation 214, pp 329–335.
[48] Eryarsoy E., Koehler, Gary J., Aytug H. (2009). Using domain-specific knowledge in generalization error bounds for support vector machine learning, Decision Support Systems 46, pp 481–491.
[49] Lia Q., Licheng J., Yingjuan H. (2007). Adaptive simplification of solution for support vector machine, Pattern Recognition 40, pp 972 – 980.
[50] Kikuchi T., Abe S. (2005). Comparison between error correcting output codes and fuzzy support vector machines, Pattern Recognition Letters 26, pp 1937–1945.
[51] Markowetz F. (2004). Classification by Nearest Shrunken Centroids and Support Vector Machines , Max Planck Institute for Molecular Genetics, Dept. Computational Molecular Biology , Computational Diagnostics Group, Berlin.
[52] Arun Kumar M., Gopal M. (2008). Application of smoothing technique on twin support vector machines, Pattern Recognition Letters 29, pp 1842–1848.
[53] John C. Platt Microsoft Research (1992). Fast Training of Support Vector Machines using Sequential Minimal Optimization.
[54] Michael Vogt (2002). SMO Algorithms for Support Vector Machines without Bias Term, Institute of Automatic Control.
[55] J. Mercer. (1909). Functions of positive and negative type and their connection With the theory of integral equations, Philos. Trans. Roy. Soc. London pp 415–446.
[56] P.I. Davies, N.J. Higham. (2000). Numerically stable generation of correlation Matrices and their factors, BIT 40 ,pp 640–651.
[57] H.Lutkepohl. (1996). Handbook of Matrices, John Willey & Sons, Berlin.
[58] J. Gilbert, L. Gilbert. (1995). Linear Algebra and Matrix Theory, Academic Press.
[59] Balan, B., Mohaghegh, S., and Ameri, S., (1995). State-Of-The-Art in Permeability Determination from Well Log Data, Part 1, A Comparative Study, Model Development. SPE 30978, SPE Eastern Regional Conference and Exhibition, Morgantown, West Virginia.
[60] Demuth, H., and Beale, M., (2002). Neural Network Toolbox For Use with MATLAB. User's Guide Version 4.
[61] Rolon, L., (2004). Developing Intelligent Synthetic Logs: Application to Upper Devonian Units in PA. M.Sc thesis, West Virginia University, Morgantown, West Virginia.
[62] Artun, E., Mohaghegh, S., and Toro, J., (2005). Reservoir Characterization Using Intelligent Seismic Inversion. SPE98012, West Virginia University.

Abstract

Nowadays, one of the most interesting subjects of economist and financial analyzer is to evaluate the variation and trend of price made many different aspects and perspectives. Because of the unavailability of sufficient information on the effective parameters of stock marketing, predicting of variation is not possible. Effect of time on the variation of variables involved and internal and external political variables which directly influence on the economic systems are by far the most significant challenge of modeling the stock marketing.

Designing a machine which is capable of learning and experiencing was one of the most important discussions of scientific researches. It was proved that this type of machine is able to learn the trend and variation of data and can provide an acceptable result. Support Vector Machine (SVM) was introduced in 1990s and successfully used to solve regression and classification problems. Based on the theory mentioned, the reasons of success in this machine are: 1) generalization ability, 2) robustness against the noise and, 3) solving the problems with limited number of data.

In this regard, the aim of this research work is to introduce a novel and robust methodology in pattern recognition field called SVM, utilizing optimization, statistical approaches and kernels. Doing some research on stock marketing index, variation of price of Alborz Company in 1387 and 1388 was taken as the dependent variables and 14 economic variables was considered as the independent variables. After normalization process, modeling was carried out using SVM. Performance of various kernels and network parameters were evaluated and an optimum network was introduced to solve the model constructed. At the end, General regression neural network (GRNN) was described and also used to solve the problem mentioned. Results obtained have shown that the SVM is a more accurate and precise method compared to GRNN.

Shahrood
University of Technology

Modeling and Prediction of stock marketing index using Mathematical methods

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in Applied Mathematics

By:
Farhad Ramshini

Supervisor:
Dr. Ahmad Nezakati

Advisors:
Dr. Hamid Hasanpour
Eng. Majid Ameri

December ۲۰۱۱