Archive of SID

فصلنامه علمی پژوهشی حسابداری مالی/سال هفتم/ شماره ۲۸ /زمستان ۹۴/صفحات ۳۱- ۱

بررسی سودمندی روش غیرخطی رگرسیون بردارهای پشتیبان و روش‌های کاهش متغیرهای پیش‌بین در پیش‌بینی بازده سهام

محمد حسین ستایش *
مصطفی کاظم‌نژاد **

تاریخ دریافت: ۹۴/۰۹/۱۸ تاریخ پذیرش: ۹۴/۱۲/۱۰

چکیده:

هدف این پژوهش، بررسی سودمندی روش‌های کاهش متغیرها و روش غیرخطی رگرسیون بردارهای پشتیبان، در پیش‌بینی بازده سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است. با استفاده از روش‌های مبتنی بر همبستگی و تحلیل عاملی، متغیرهای بهینه از بین ۵۲ متغیر اولیه، انتخاب یا استخراج شده است. در ادامه، با استفاده از روش‌های غیرخطی رگرسیون بردارهای پشتیبان و شبکه‌های عصبی مصنوعی و همچنین رگرسیون خطی، بازده سهام ۱۰۱ شرکت پذیرفته شده در بورس اوراق بهادار تهران در بازه زمانی ۱۳۸۳ الی ۱۳۹۲ پیش‌بینی شده است. یافته‌های تجربی این پژوهش حاکی از عملکرد بهتر رگرسیون بردارهای پشتیبان نسبت به دو روش دیگر پیش‌بینی و عملکرد بهتر هر دو روش غیرخطی پیش‌بینی نسبت به رگرسیون خطی است. افزون براین، یافته‌های پژوهش، بیانگر سودمندی روش‌های کاهش متغیرها و وجود تفاوت معنادار بین میزان سودمندی دو روش مبتنی بر همبستگی و تحلیل عاملی و همچنین برتری روش مبتنی بر همبستگی است.

* دانشیارحسابداری، دانشگاه شیراز، گروه حسابداری، شیراز، ایران
Email; setayesh@shirazu.ac.ir

** دانشجوی دکترای حسابداری، دانشگاه شیراز، گروه حسابداری، شیراز، ایران (نویسنده مسئول)
Email: mkazemi5166@gmail.com

www.SID.ir

Archive of SID

واژه‌های کلیدی: پیش‌بینی بازده سهام، رگرسیون بردارهای پشتیبان، روش‌های کاهش متغیرهای پیش‌بین.

۱- مقدمه

با توجه به اهمیت قیمت و بازده سهام، تاکنون مدل‌ها و نظریه‌های مختلفی شامل مدل قیمت‌گذاری دارایی‌های سرمایه‌ای[۱] مدل‌های عاملی[۲]، مدل آربیتراژ[۳]، تحلیل‌های فنی[۴] و بنیادی[۵] برای تبیین و پیش‌بینی آن ارائه شده است. یافته‌های پژوهش‌های تجربی نیز حاکی از آن است که پیش‌بینی قیمت و بازده سهام، بخش عمده‌ای از پژوهش‌های انجام شده را به خود اختصاص داده است (کوثاری[۶] ۲۰۰۱ و نمازی و ناظمی، ۱۳۸۴). تاکنون در اغلب پژوهش‌های انجام شده در بورس اوراق بهادار تهران، از روش‌های خطی برای پیش‌بینی بازده سهام استفاده شده است (به‌عنوان نمونه، ثقفی و شعری، ۱۳۸۳؛ نمازی و محمدتبار کاسگری، ۱۳۸، هاشمی و ساعدی، ۱۳۸۸، فروغی و نادم، ۱۳۸۹ و دستگیر و همکاران، ۱۳۹۱). در چند پژوهش نیز از روش غیرخطی شبکه‌های عصبی مصنوعی[۷] استفاده شده که یافته‌های این پژوهش‌ها حاکی از عملکرد بهتر این روش نسبت به روش خطی است (به‌عنوان نمونه، آذر و کریمی، ۱۳۸۸و همت‌فر و همکاران،۱۳۹۰). با این وجود، علی‌رغم عملکرد بهتر روش غیرخطی رگرسیون بردارهای پشتیبان[۸] نسبت به شبکه‌های عصبی مصنوعی در پیش‌بینی برخی متغیرهای حسابداری (به‌عنوان نمونه، مین و لی[۹]، ۲۰۰۴ و لی و تاو[۱۰]،۲۰۱۰)، تاکنون پژوهشی در بورس اوراق بهادار تهران در زمینه پیش‌بینی بازده با استفاده از این روش انجام نشده است. افزون‌بر این، در پژوهش‌های انجام شده در زمینه پیش‌بینی بازده، هدف و تأکید اصلی، ارائه مدل‌های مناسب و دقیق برای پیش‌بینی بوده و کمتر به کاهش متغیرها[۱۱] (انتخاب متغیرها[۱۲] و استخراج

1 Capital Asset Pricing Model (CAPM)
2 Factor Model
3 Arbitrage
4 Technical Analysis
5 Fundamental Analysis
6 Kothari
7 Artificial Neural Network (ANN)
8 Support Vector Regression (SVR)
9 Lee and Min
10 Lee and To
11 Variable Reduction
12 Feature (Variable) Selection

www.SID.ir

Archive of SID

متغیرها یا عامل‌ها[1]) و روش‌های مناسب آن پرداخته شده است. در مقابل، در بسیاری از پژوهش‌های داخلی و خارجی انجام شده در این زمینه، مرحله کاهش متغیرهای پیش‌بین، نادیده گرفته شده و متغیرهای پیش‌بین، بدون ضابطه و صرفاً با توجه به پژوهش‌های گذشته انتخاب شده است که این امر به انتخاب متغیرهای پیش‌بین غیربهینه و در برخی موارد، متغیرهای پیش‌بین نامناسب منجر می‌شود (تسای[2]، ۲۰۰۹ و ستایش[3] و همکاران، ۲۰۱۲). یافته‌های پژوهش لو[4] (۲۰۱۰) نیز حاکی از آن است که انتخاب متغیرها و استخراج متغیرهای (عامل‌های) پیش‌بین و روش‌های آن، نسبت به انتخاب مدل پیش‌بینی، تأثیر بیشتری بر عملکرد پیش‌بینی دارد.

با توجه به اهمیت بازده سهام در تصمیم‌گیری سرمایه‌گذاران و اعتباردهندگان، عملکرد بهتر روش‌های غیرخطی و نقش با اهمیت مرحله انتخاب و استخراج متغیرهای پیش‌بین بهینه در پیش‌بینی، پژوهش حاضر به بررسی عملکرد دو روش غیرخطی رگرسیون بردارهای پشتیبان و شبکه‌های عصبی مصنوعی و مقایسه آن‌ها با رگرسیون خطی در پیش‌بینی بازده سهام می‌پردازد. هم‌چنین، سودمندی روش تحلیل عاملی[5] در استخراج متغیرها و روش مبتنی بر همبستگی در انتخاب متغیرهای پیش‌بین بهینه، بررسی و مقایسه می‌شود.

۲- مبانی نظری و پیشینه پژوهش

تاکنون مدل‌ها و نظریه‌های مختلفی شامل مدل قیمت‌گذاری دارایی‌های سرمایه‌ای، مدل‌های عاملی، مدل آربیتراژ، تحلیل‌های فنی و بنیادی برای تبیین و پیش‌بینی آن ارائه شده است. مدل میانگین- واریانس مارکویتز، مشهورترین و متداول‌ترین رویکرد در مسئله انتخاب سرمایه گذاری و کاراترین ابزار برای انتخاب پرتفوی بهینه است. با توجه به مشکلات این مدل، شارپ[6]، لینتر[7] و ... در دهه ۱۹۶۰ مدل قیمت‌گذاری دارایی‌های سرمایه‌ای را ارائه کردند. اعتبار این مدل که برای چند دهه از مدل‌های کاربردی در انتخاب طرح‌های سرمایه‌گذاری بود از اواسط دهه ۱۹۸۰ مورد تردید قرار گرفت و راس در دهه ۱۹۷۰، نظریه قیمت‌گذاری آربیتراژ را پایه‌گذاری کرد. مسأله موجود در نظریه مزبور، نامشخص بودن (یا حداقل مربوط به آینده

1 Variable (Factor) Extraction
2 Tsai
3 Setayesh
4 Lo
5 Factor Analysis (FA)
6 Sharpe
7 Lintner

www.SID.ir

Archive of SID

بودن) عوامل مؤثر در مدل و تعداد آن است. بنابراین انتخاب عوامل نسبتا اختیاری و در معرض انتقاد است. یک روش پیشنهادی برای حل این مسأله، استفاده از روش‌هایی جهت کاهش متغیرها (از قبیل روش مبتنی بر همبستگی و تحلیل عاملی) است (تهرانی و نوربخش، ۱۳۸۶ و فاما، ۱۹۹۱). پژوهش برای انتخاب طرح‌های سرمایه‌گذاری و سبد اوراق بهادار در مسیر متفاوتی ادامه یافت و سرانجام نتایج پژوهش فاما و فرنچ در دهه ۱۹۹۰ در قالب مدل سه‌عاملی (عامل بازار، عامل اندازه و عامل ارزش) برای پیش‌بینی بازده سهام ارائه شد. اگر چه مدل سه عاملی بهتر از مدل قیمت‌گذاری دارایی‌های سرمایه‌ای، بازده مورد انتظار را تبیین می‌کرد، ولی پژوهش‌های زیادی در زمینه یافتن عوامل مؤثر دیگری در تبیین و پیش‌بینی بازده سهام جهت دستیابی به مدل جامع‌تر انجام شده است (به عنوان نمونه، مدل چهار عاملی کارهارت[1] با اضافه کردن عامل مومنتوم به مدل سه عاملی و مدل پنج عاملی فاما و فرنچ با اضافه کردن عامل‌های سودآوری و سرمایه‌گذاری به مدل سه عاملی). اغلب پژوهش‌های انجام شده تا این زمان با فرض وجود رابطه خطی بین بازده سهام و متغیرهای پیش‌بین انجام شد.

گروهی از صاحب‌نظران به پدیده آشفتگی (آشوب) توجه می‌کنند. آن‌ها با استفاده از فنون جدید ریاضی، بازار را سیستمی، پیچیده و در حال تحول می‌دانند. این گروه، مانند ریاضی‌دانان و فیزیک‌دانان اعتقاد دارند چنانچه هر پدیده‌ای به‌درستی مشاهده و بررسی شود، وقایع به ظاهر تصادفی آن، تبیین‌پذیر است و اگر در خور پیش‌بینی نباشد، حداقل قابل درک خواهد بود. طرفداران الگوی آشفتگی معتقدند که رفتار تغییر قیمت‌ها و بازده به صورت غیرخطی است. تغییرات قیمت سهام ناشی از عوامل بسیار زیادی است که در هر لحظه از زمان بر آن‌ها تأثیر می‌گذارد. آن‌ها در واقع معتقدند که به تعداد سرمایه‌گذاران، عوامل مؤثر بر قیمت سهام وجود دارد. شواهد زیادی در دست است که بازارهای سرمایه دارای سیستمی غیرخطی است و مادام که این مسئله مهم در نظریه‌های مالی نادیده گرفته شود، اعتبار نتایج به شدت تضعیف می‌شود. نظریه‌های موجود مالی در حال حاضر بر مبنای روابط خطی بنا شده‌اند و رگرسیون خطی بر فرض توزیع نرمال استوار است، در حالی که مردم و طبیعت به طور کلی ماهیتی غیرخطی دارند. علت اصلی مشکلات نظریه‌های موجود، انجام مطالعات در چارچوب روابط خطی و توزیع نرمال می‌باشد (نقل از عبده‌تبریزی و گنابادی، ۱۳۷۵).

1 Carhart

www.SID.ir

Archive of SID

برای اولین بار در ادبیات، پژوهش‌های مرامور[1] و مرامور-کوستا[2] (۱۹۹۷) و مرامور و پاهور[3] (۱۹۹۸) نشان داد که رابطه نسبت‌های مالی و بازده سهام، غیرخطی است. مرامور و مرامور-کوستا (۱۹۹۷) به بررسی رابطه بین بازده سهام و نسبت‌های مالی شرکت‌های اسلوانیایی پرداختند و برخلاف یافته‌های اغلب پژوهش‌های قبلی که به بررسی خطی رابطه بین نسبت‌های مالی و بازده سهام می‌پرداختند، وجود رابطه غیرخطی بین نسبت‌های مالی و بازده سهام را تأیید کردند. یافته‌های پژوهش مرامور و پاهور (۱۹۹۸) در آمریکا و ژاپن نیز وجود رابطه غیرخطی را تأیید کرد. پس از انجام این پژوهش‌ها، مطالعات زیادی به بررسی رابطه خطی و غیرخطی بین متغیرهای حسابداری، بازار، اقتصادی و سایر متغیرها با بازده سهام پرداختند. یافته‌های پژوهش کاناس[4] و یانوپولس[5] (۲۰۰۱) اولسون و موسمن (۲۰۰۳) در شرکت‌های کانادایی، عمران[6] و رجب[7] (۲۰۰۴) در شرکت‌های مصری، مک‌میلان[8] (۲۰۰۷) در کشورهای انگلیس، آمریکا، فرانسه و ژاپن ، بویاسیگلو[9] و آوسی[10] (۲۰۱۰) در ترکیه، جیسن و همکاران[11] (۲۰۱۳) در بورس ژوهانسبورگ نیز حاکی از عملکرد بهتر روش‌های غیرخطی نسبت به روش‌های خطی بود. در بورس اوراق بهادار تهران نیز یافته‌های پژوهش آذر و کریمی (۱۳۸۸)، عباسی و باقری (۱۳۹۰)، همت‌فر و همکاران (۱۳۹۰) و نصیرزاده (۱۳۹۲) این نتایج را تأیید می‌کند.

تاکنون پژوهشی که به بررسی کاهش متغیرهای پیش‌بین و روش‌های آن و هم‌چنین پیش‌بینی بازده سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از روش غیرخطی رگرسیون بردارهای پشتیبان پرداخته باشد، انجام نشده است. در این راستا، پژوهش حاضر به پیش بینی بازده سهام با استفاده از رگرسیون غیرخطی بردارهای پشتیبان و بررسی سودمندی روش‌های کاهش متغیر مبتنی بر همبستگی و تحلیل عاملی بر عملکرد پیش‌بینی می‌پردازد.

1 Mramor
2 Mramor-Kosta
3 Pahor
4 Kanas
5 Yannopoulos
6 Omran
7 Ragab
8 McMillan
9 Boyacioglu
10 Avci
11 Gysen et al.

www.SID.ir

Archive of SID

۱-۲- کاهش متغیرهای پیش‌بین

علاقه اصلی پژوهشگر اغلب به جای این که متوجه آزمون فرضیه یا ارزشیابی اهمیت نسبی متغیرهای پیش‌بین باشد متوجه رسیدن به بهترین پیش‌بینی ممکن متغیر وابسته توسط چند پیش‌بینی کننده است. در این شرایط، تلاش پژوهشگر در جهت کسب بالاترین مجذور همبستگی چندمتغیری ممکن، صرف می‌شود. چون بیشتر متغیرهای علوم رفتاری متقابلا همبسته‌اند، اغلب می‌توان از بین مجموعه کل متغیرها یک مجموعه کوچک‌تر را انتخاب کرد و تقریبا به همان R^2 دست یافت که از مجموعه کل متغیرها حاصل می‌شود (سرایی، ۱۳۸۸). به‌طور کلی، اندک بودن منطقی متغیرهای پیش‌بین و دقت بالای پیش‌بینی از مهم‌ترین معیارهای کیفیت یک مدل پیش‌بینی محسوب می‌شود (ابریشمی، ۱۳۸۷). با توجه به این که در خصوص متغیرهای پیش‌بینی کننده بازده سهام، اجماع وجود ندارد، متغیرهای پیش‌بین استفاده شده در ادبیات، بایستی بررسی و متغیرهای پیش‌بین بهینه در پیش‌بینی بازده سهام، شناسایی و انتخاب شود. بنابراین، عملکرد پیش‌بینی پس از انتخاب متغیرهای بهینه می‌تواند نسبت به عدم انتخاب متغیرهای بهینه بهتر شود (لیانگ و همکاران، ۲۰۱۵).

کاهش متغیرهای پیش‌بین جهت رسیدن به بهترین نتیجه در پیش‌بینی، از مباحث چالش برانگیز در دو دهه اخیر بوده است. از دیدگاه تئوری، یادگیری براساس تعداد متغیرهای پیش-بین بیشتر باعث می‌شود تا دقت پیش‌بینی بالا رود. با این وجود، شواهد تجربی بیانگر آن است که این امر همواره صادق نیست؛ زیرا تمام متغیرها، برای تشخیص و پیش‌بینی مهم نیستند و یا برخی از آن‌ها به‌طور کلی در پیش‌بینی نامربوط هستند (لیندنبام و همکاران[1]، ۲۰۰۴). با توجه به این‌که عامل‌های بسیاری از جمله کیفیت داده‌ها در موفقیت یک الگوریتم یادگیری مؤثر است، اگر داده‌ها حاوی متغیرها و یا اطلاعات تکراری و نامربوط[2] باشند و یا حاوی اطلاعات نامطمئن باشند، اخذ دانش از آن داده‌ها مشکل می‌شود (هال[3]، ۲۰۰۰). افزون براین، کاهش تعداد متغیرهای پیش‌بین نامربوط یا اضافی، علاوه بر کاهش زمان اجرای الگوریتم‌های آموزشی، به مفهومی عمومی‌تر منجر می‌شود. سایر مزایای بالقوه انتخاب متغیرهای پیش‌بین شامل تسهیل درک و تجسم داده‌ها، کاهش الزامات اندازه‌گیری و ذخیره اطلاعات، کاهش اضافه‌بار ابعاد[4] و بهبود عملکرد پیش‌بینی و فراهم کردن بینش بهتر در مورد مفهوم زیربنایی از

1 Lindenbaum
2 Redundant
3 Hall
4 Curse of Dimensionality

www.SID.ir

Archive of SID

پیش‌بینی دنیای واقعی است (تسای، ۲۰۰۹). دو جنبه مهم در روش‌های مختلف کاهش ابعاد (متغیرها) وجود دارد:

استخراج متغیرهای پیش‌بین[1]:

استخراج متغیرهای پیش‌بین یا به عبارت دیگر تبدیل متغیرهای پیش‌بین[2] فرآیندی است که K متغیر جدید حاصل می‌شود که حاصل ترکیب N متغیر پیش‌بین اولیه هستند. شناخته شده ترین الگوریتم‌های استخراج متغیرهای پیش‌بین شامل تحلیل مولفه‌های اصلی و تحلیل عاملی است. با توجه به کاربرد بیشتر روش تحلیل عاملی (به‌عنوان نمونه، گهر[3]، ۱۹۷۸ و رول و راس[4]، ۱۹۸۰) در این پژوهش نیز از این روش، برای استخراج متغیرها (عامل‌ها) استفاده می‌شود.

انتخاب متغیرهای پیش‌بین:

در مقابل الگوریتم‌های استخراج متغیرهای پیش‌بین، الگوریتم‌های انتخاب متغیر پیش‌بین، به انتخاب بهترین K متغیر از بین N متغیر اصلی می‌پردازد و بقیه متغیرهای کم اهمیت‌تر حذف می‌شود.

شایان ذکر است که در انتخاب متغیرها، متغیرهای اصلی بدون تغییر، انتخاب می‌شود اما در استخراج متغیرها، متغیرها به صورت تغییریافته استفاده می‌شوند. روش مبتنی بر همبستگی از مهم‌ترین روش‌های انتخاب متغیر در پیش‌بینی متغیرهای پیوسته محسوب می‌شود که در این پژوهش نیز استفاده شده است. یافته‌های پژوهش‌های انجام شده (از قبیل هال، ۱۹۹۹؛ کارگودا و همکاران، ۲۰۱۰؛ کوپرینسکا و همکاران، ۲۰۱۵) حاکی از عملکرد بهتر روش مزبور نسبت به سایر روش‌های انتخاب متغیر است. افزون براین، به منظور مقایسه با روش استخراج متغیر تحلیل عاملی که در اغلب پژوهش‌های مالی استفاده شده است، از این روش استفاده شده است.

در ادامه به تشریح روش‌های انتخاب متغیر مبتنی بر همبستگی و روش استخراج متغیر تحلیل عاملی پرداخته می‌شود.

1 Feature Extraction
2 Feature Transformation
3 Gehr
4 Roll and Ross

www.SID.ir

Archive of SID

۲-۱-۱- روش مبتنی بر همبستگی

به طور کلی یک متغیر مناسب است اگر به متغیر وابسته (طبقه، در طبقه‌بندی) ارتباط داشته باشد اما نسبت به سایر متغیرهای پیش‌بین مربوط، اضافی یا زائد نباشد. اگر همبستگی بین دو متغیر به عنوان معیار مناسب بودن در نظر گرفته شود، تعریف ذکر شده به این صورت خواهد شد که یک متغیر مناسب است در صورتی که همبستگی بالایی با متغیر وابسته (طبقه) و همبستگی پایینی با سایر متغیرهای پیش‌بین داشته باشد. به عبارت دیگر، اگر همبستگی بین یک متغیر پیش بین و متغیر وابسته (طبقه) به اندازه کافی بالا باشد تا برای پیش‌بینی متغیر وابسته (طبقه) مربوط باشد و همبستگی بین آن و سایر متغیرهای پیش‌بین مربوط، به سطح معینی نرسد، به طوری که نتوان توسط سایر متغیرهای مربوط، متغیر مزبور را پیش‌بینی کرد، آن متغیر به عنوان یک متغیر مناسب برای پیش‌بینی (طبقه‌بندی) قلمداد می‌شود. در این حالت، مساله اصلی انتخاب متغیر، جستجوی معیار مناسب همبستگی بین متغیرها و رویه منطقی برای انتخاب متغیرهای مناسب بر اساس این معیار خواهد بود (یو و لیو[1]، ۲۰۰۳).

ایده اصلی روش مبتنی بر همبستگی این است که متغیرهایی بهینه محسوب می‌شوند که با متغیر وابسته همبستگی بالا و با سایر متغیرهای پیش‌بین همبستگی نداشته یا همبستگی ناچیزی داشته باشد. این روش یک الگوریتم کاملاً خودکار است، یعنی محدودیتی برای شناسایی آستانه[2] یا تعداد متغیرهایی که باید انتخاب شود، اعمال نمی‌کند. با این وجود، در صورت تمایل، اعمال آستانه یا تعداد متغیرها به‌راحتی امکان پذیر است (هال، ۱۹۹۹).

۲-۱-۲- روش تحلیل عاملی

تحلیل عاملی نامی عمومی برای برخی از روش‌های آماری چندمتغیره است که هدف اصلی آن خلاصه کردن داده‌هاست. این روش به بررسی همبستگی درونی تعداد زیادی از متغیرها می‌پردازد و در نهایت آن‌ها را در قالب عامل‌های عمومی محدودی دسته‌بندی کرده و تبیین می‌کند. در تحلیل عاملی کلیه متغیرها به‌طور همزمان مدنظر قرار می‌گیرند و هر یک از متغیرها به عنوان یک متغیر وابسته لحاظ می‌شود. تحلیل عاملی از جمله روش‌های چندمتغیره است که در آن متغیرهای مستقل و وابسته مطرح نیست، زیرا این روش جزء تکنیک‌های هم وابسته محسوب می‌شود و کلیه متغیرها نسبت به هم، وابسته لحاظ شده و سعی می‌شود تا

1 Yu and Liu
2 Threshold

www.SID.ir

Archive of SID

تعداد زیادی متغیر در چند عامل خلاصه شوند (کلانتری، ۱۳۸۵). برای اجرای تحلیل عـاملی، چهار گام اساسی زیر ضرورت دارد (خاکی، ۱۳۹۰):

الف) تهیه یک ماتریس همبستگی از تمام متغیرهای مورد استفاده در تحلیل و برآورد اشتراک

ب) استخراج عامل‌ها

ج) انتخاب و چرخش عامل‌ها برای ساده‌تر ساختن و قابل فهم‌تر کردن ساختار عاملی

د) تفسیر نتایج

۲-۲- روش‌های خطی و غیرخطی پیش‌بینی

روش‌های غیرخطی از قبیل شبکه‌های عصبی، چندین مزیت مهم در مقایسه با مـدل‌هـای آماری از قبیل رگرسیون خطی دارند. معایب رگرسیون خطی نسبت به روش‌های غیرخطی (از قبیل شبکه‌های عصبی) به شرح زیر است (دی‌تاین و همکاران[1]، ۲۰۰۳):

- ماهیت خطی رگرسیون: یک عیب مهم رگرسیون‌های خطی این است که رگرسیون هیچ شاخص مستقیمی مبنی بر این‌که آیا داده‌ها در حالت خطی به بهترین صورت نشان داده می‌شود را ارائه نمی‌کند. بـا توجـه بـه ماهیـت علـوم اجتمـاعی، در بسـیاری از حالت‌ها تحلیل آماری خطی نامناسب است.
- از پیش مشخص کردن مدل: استفاده از مدل‌های رگرسیون، مستلزم از پـیش مشـخص کردن مدل پایه است. این کار باعث حل آسان‌تر مسـأله مـی‌شـود ولـیکن، نیازمنـد حدس‌های زیاد است.
- مفروضات رگرسیون: عملکرد مدل‌های رگرسیون خطی وابسته به مفروضات گوناگونی از قبیل عدم وجود روابط خطی چندگانه و توزیع نرمال باقیمانده‌ها است.
- عدم انطباق پذیری: رگرسیون چند متغیره در حالتی که اجزای مدل را نتوان به‌وسـیله حدس مشخص کرد، دارای خاصیت انطباق‌پذیری با داده‌ها نیست.

شبکه‌های عصبی مصنوعی و رگرسیون بردارهای پشتیبان از مهم‌ترین روش‌های غیرخطـی محسوب می‌شوند که در ادامه به تشریح آن‌ها پرداخته می‌شود.

۱-۲-۲- شبکه‌های عصبی مصنوعی

شبکه‌های عصبی یک تکنیک پـردازش اطلاعـات مبتنـی بـر روش سیسـتم‌هـای عصـبی بیولوژیکی مانند مغز است. طی سال‌های گذشته، شبکه‌های عصبی به‌عنوان یک فناوری ظاهر

1 DeTienne et al.

www.SID.ir

Archive of SID

شده است که می‌تواند الگوی داده‌ها را شناسایی و مدل‌سازی کند، کاری که با روش‌های آماری به سهولت امکان‌پذیر نیست. مفهوم بنیادی شبکه‌های عصبی، ساختار سیستم پردازش اطلاعات است که از تعداد زیادی واحدهای پردازشی (نورون [۱]) مرتبط با شبکه‌ها تشکیل شده‌اند. سلول عصبی بیولوژیکی یا نورون، واحد سازنده سیستم عصبی در انسان است. یک نورون از بخش‌های اصلی زیر تشکیل شده است (جعفریه و همکاران، ۱۳۸۵):

۱) بدنه سلولی [۲] که هسته در آن است و سایر قسمت‌های سلولی از آن منشأ گرفته است.

۲) هسته.

۳) آکسون [۳] که وظیفه آن انتقال اطلاعات از سلول عصبی است.

۴) دندریت [٤] که وظیفه آن انتقال اطلاعات از سلول‌های دیگر به سلول عصبی است.

یکی از مهم‌ترین شبکه‌های عصبی، شبکه‌های عصبی پرسپترون چند لایه است که بر مبنای یادگیری با سرپرست است. ساختار پایه شبکه عصبی بر مبنای پرسپترون چند لایه، یک لایه ورودی، یک یا چند لایه مخفی و یک لایه خروجی است. هرکدام از این لایه‌ها از یک یا چندین گره تشکیل شده است. لایه ورودی به تعداد متغیرهای مستقل دارای گره است و به همین ترتیب نیز لایه خروجی به تعداد متغیرهای وابسته دارای گره می‌باشد. اما مشخص کردن ساختار لایه پنهان، دشوار است (هوگلاند [٥]، ۲۰۱۲). برای آموزش این شبکه عصبی از قانون یادگیری پس‌انتشار خطا استفاده می‌شود. الگوریتم پس انتشار خطا، نوع خاصی از الگوریتم حداقل میانگین مربعات [٦] به شمار می‌رود. مانند قاعده یادگیری حداقل میانگین مربعات، الگوریتم پس‌انتشار نیز بر مبنای تخمین بیشترین کاهش گرادیان عمل می‌کند. شاخص کارایی در این الگوریتم، میانگین مربعات خطاهاست. تنها تفاوت حداقل میانگین مربعات و پس‌انتشار در نحوه محاسبه مشتقات است. در مورد یک شبکه تک لایه خطی، خطا یک تابع خطی صریح از وزن‌های شبکه است و مشتقات بر حسب وزن‌های آن به‌راحتی قابل محاسبه است. در شبکه‌های چند لایه با تابع انتقال غیرخطی، رابطه بین وزن‌های شبکه و خطا بسیار پیچیده‌تر است و در راستای محاسبه مشتقات نیاز به استفاده از قاعده زنجیری [۷] وجود دارد (کیا، ۱۳۸۸).

---

1 Neuron
2 Cell Body
3 Axon
4 Dendrite
5 Hoglund
6 Least Mean Square (LMS)
7 Chain Rule

www.SID.ir

Archive of SID

با توجه به این که روش پرسپترون چند لایه پس انتشار خطا بیشترین کاربرد را در حوزه مالی دارد، در این پژوهش نیز از این روش استفاده شده است.

۲-۲-۲- رگرسیون بردارهای پشتیبان

رگرسیون بردارهای پشتیبان یکی از انواع رگرسیون‌ها به‌شمار می‌رود که تئوری آن به‌صورت زیر تعریف می‌شود (اسمولا[1] و اسکولکوپ[2]، ۲۰۰۴):

فرض کنید داده‌های آموزشی به‌صورت زیر باشد، که X فضای الگوهای ورودی را نشان می‌دهد:

$$\{(x_1, y_1),(x_2, y_2),(x_3, y_3),...,(x_i, y_i)\} \subseteq X \times R \quad (۱)$$

در رگرسیون $\varepsilon - SV$ هدف پیدا کردن تابع f(x) است که حداکثر اختلاف خروجی f(x) یعنی y_i به ازای xi، از مقدار y_i واقعی برای تمام داده‌های آموزشی برابر با ε و تا حد ممکن f(x) مانند خطی بدون انحنا باشد:

$$f(x) = \langle w, x \rangle + b \; with \; w \in \chi, b \in R \quad (۲)$$

که <.,.> ضرب داخلی در فضای X را نشان می‌دهد و عدم انحنا در f(x) به معنی w کوچک می‌باشد. یکی از راه‌ها برای تضمین این امر کمینه کردن نرم w می‌باشدکه این موضوع به صورت بهینه‌سازی یک مسأله convex قابل حل است:

$$\begin{aligned} &minimize && \frac{1}{2}\|w\|^2 \\ &subject\ to && \begin{cases} y_i - \langle w, x_i \rangle - b \le \varepsilon \\ \langle w, x_i \rangle + b - y_i \le \varepsilon \end{cases} \end{aligned} \quad (۳)$$

1 Smola
2 Scholkopf

www.SID.ir

Archive of SID

در اکثر اوقات، انحراف بعضی داده‌ها از مقدار واقعی، بیشتر از ε است و همین امر باعث پیدایش تعمیم جدیدی از تئوری SV شد که با ورود متغیرهای جدیدی به‌نام ξ (متغیرهای کمکی) و با استفاده از فرمول‌های جدید، این مشکل را رفع می‌کرد:

$$\begin{aligned} &minimize && \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{\ell}(\xi_i + \xi_i^*) \\ &subject\ to && \begin{cases} y_i - \langle w, x_i\rangle - b \le \varepsilon + \xi_i \\ \langle w, x_i\rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \end{aligned}$$

(۴)

برای غیرخطی کردن الگوریتم بردارهای پشتیبان از هسته استفاده می‌شود (اسمولا و اسکولکوپ ، ۲۰۰۴). از مشهورترین هسته‌ها، چند جمله‌ای[1]، پایه شعاعی[2] و هلالی[3] هستند. در این پژوهش، مشابه اغلب پژوهش‌ها، با توجه به عملکرد بهتر، از هسته چندجمله‌ای استفاده می‌شود (هانگ و چن، ۲۰۰۹ و ستایش و همکاران، ۲۰۱۲).

۳- فرضیه‌های پژوهش

براساس مبانی نظری و پیشینه پژوهش، فرضیه‌هایی به شرح زیر تدوین شده است:

۱- تفاوت معناداری بین عملکرد روش‌های غیرخطی رگرسیون بردارهای پشتیبان و شبکه‌های عصبی مصنوعی، در پیش‌بینی بازده سهام شرکت‌ها وجود دارد.

۲- تفاوت معناداری بین عملکرد روش غیرخطی رگرسیون بردارهای پشتیبان و رگرسیون خطی، در پیش‌بینی بازده سهام شرکت‌ها وجود دارد.

۳- تفاوت معناداری بین عملکرد روش غیرخطی شبکه‌های عصبی مصنوعی و رگرسیون خطی، در پیش‌بینی بازده سهام شرکت‌ها وجود دارد.

۴- کاهش متغیرهای پیش‌بین بهینه، تأثیر معناداری بر عملکرد روش‌های پیش‌بینی دارد.

۴-۱- متغیرهای منتخب روش انتخاب متغیر مبتنی بر همبستگی، تأثیر معناداری بر عملکرد پیش‌بینی رگرسیون بردارهای پشتیبان دارد.

۴-۲- عامل‌های روش استخراج متغیر تحلیل عاملی، تأثیر معناداری بر عملکرد پیش‌بینی رگرسیون بردارهای پشتیبان دارد.

1 Polynomial
2 Radial Basis
3 Sigmoid

www.SID.ir

Archive of SID

۴-۳- متغیرهای منتخب روش انتخاب متغیر مبتنی بر همبستگی، تأثیر معناداری بر عملکرد پیش‌بینی رگرسیون خطی دارد.
۴-۴- عامل‌های روش استخراج متغیر تحلیل عاملی، تأثیر معناداری بر عملکرد پیش‌بینی رگرسیون خطی دارد.
۴-۵- متغیرهای منتخب روش انتخاب متغیر مبتنی بر همبستگی، تأثیر معناداری بر عملکرد پیش‌بینی شبکه‌های عصبی مصنوعی دارد.
۴-۶- عامل‌های روش استخراج متغیر تحلیل عاملی، تأثیر معناداری بر عملکرد پیش‌بینی شبکه‌های عصبی مصنوعی دارد.

۴- روش پژوهش

این پژوهش از نوع پژوهش‌های کمّی است که از روش علمی ساخت و اثبات تجربی استفاده می‌کند و بر اساس فرضیه‌ها و طرح‌های پژوهش از قبل تعیین شده انجام می‌شود. از این دسته پژوهش‌ها زمانی استفاده می‌شود که معیار اندازه‌گیری داده‌ها کمّی است و برای استخراج نتیجه‌ها از فن‌های آماری استفاده می‌شود (نمازی، ۱۳۸۲).

۴-۱- روش گردآوری داده‌ها و اطلاعات

در این پژوهش برای جمع‌آوری داده‌ها و اطلاعات از روش‌های کتابخانه‌ای و میدانی استفاده شده است. مبانی نظری پژوهش از کتب، مجلات و سایت‌های تخصصی فارسی و لاتین گردآوری شده و داده‌های مالی مورد نیاز با مراجعه به سایت سازمان بورس اوراق بهادار تهران، صورت‌های مالی شرکت‌ها و همچنین با استفاده از نرم‌افزارهای تدبیرپرداز و ره‌آورد نوین گردآوری شده است. در مرحله اول با بررسی ادبیات و پیشینه پژوهش (شامل حدود ۲۵۰ مقاله فارسی و انگلیسی)، ۲۰۰ متغیر پیش‌بین اولیه شناسایی شد. از بین متغیرهای شناسایی شده، ۵۲ متغیری که بیشتر در ادبیات مربوط به تبیین یا پیش‌بینی بازده سهام استفاده شده و داده‌های مورد نیاز برای سنجش آن‌ها از طریق پایگاه‌های اطلاعاتی سازمان بورس و اوراق بهادار و همچنین نرم‌افزارهای تدبیرپرداز و ره‌آورد نوین در دسترس بود، انتخاب شد. در این راستا، سعی شد که هم متغیرهای حسابداری (اقلام مبتنی بر ترازنامه، صورت سود و زیان و صورت جریان وجوه نقد) و هم متغیرهای بازار استفاده و محتوای اطلاعاتی آن‌ها مدنظر قرار گیرد. بنابراین، ابعاد سودآوری، کارایی، اهرم مالی، نقدینگی، نسبت‌های مبتنی بر هر سهم، نسبت‌های مبتنی بر جریان وجوه نقد و نسبت‌های بازار در نظر گرفته شده است. جدول‌های شماره (۱) و (۲)، این متغیرها و میانگین آن را در شرکت‌های مورد بررسی نشان می‌دهد. در این جدول‌ها،

www.SID.ir

Archive of SID

متغیرهایی که به صورت تغییرات (Δ) نشان داده شده است، بیانگر درصد رشد سالانه متغیر مزبور می‌باشد. بتای هر سهم نیز از تقسیم کوواریانس بازده سهم و بازده بازار به واریانس بازده بازار به‌دست آمده است. برای محاسبه بتا در سال t از اطلاعات ماهانه بازده سهام و بازده بازار برای پنج ساله منتهی به سال t استفاده شده است. اقلام تعهدی نیز تفاوت بین سود عملیاتی و جریان نقد حاصل از فعالیت‌های عملیاتی است.

در ادامه، با استفاده از روش انتخاب متغیر مبتنی بر همبستگی در نرم‌افزار Weka، از بین ۵۲ متغیر ذکر شده، متغیرهای بهینه، انتخاب شده است. در جدول شماره (۳)، هفت متغیر بهینه انتخاب شده در روش مبتنی بر همبستگی ارائه شده است. همچنین، با استفاده از نرم‌افزار SPSS، روش تحلیل عاملی منجر به استخراج ۲۰ عامل گردید. شایان ذکر است که در انجام تحلیل عاملی ابتدا باید از این مسأله اطمینان حاصل شود که آیا می‌توان داده‌های موجود را برای تحلیل عاملی، مورد استفاده قرار داد یا خیر. به عبارت دیگر آیا تعداد داده‌های مورد نظر برای تحلیل عاملی مناسب هستند یا خیر. بدین منظور از شاخص KMO[1] و آزمون بارتلت استفاده می‌شود. مقادیر بزرگ شاخص KMO عموما نشان می‌دهند که تحلیل عاملی ممکن است برای داده ها سودمند باشد. اگر مقدار این شاخص کمتر از ۰/۵ باشد، نتایج تحلیل عامل احتمالا خیلی سودمند نیست (مؤمنی و فعال قیومی، ۱۳۸۶). اگر مقدار احتمال آزمون بارتلت کوچکتر از ۵ درصد باشد تحلیل عاملی برای شناسایی ساختار (مدل عاملی) مناسب است؛ زیرا فرض شناخته شده بودن ماتریس همبستگی رد می‌شود (مؤمنی و فعال قیومی، ۱۳۸۶). در این پژوهش، شاخص KMO برابر ۰/۶۱۱ و مقدار احتمال آزمون بارتلت کمتر از ۰/۰۵ و بنابراین، اجرای تحلیل عاملی در داده‌های مورد بررسی، امکان پذیر بود.

۲-۴- متغیرهای پژوهش

۱-۲-۴- متغیرهای مستقل (پیش‌بین)

کلیه متغیرهای مستقل این پژوهش که نحوه انتخاب آن در بخش قبل توضیح داده شد در جدول‌های شماره (۱) و (۲) ارائه شده است. متغیرهای مستقل بهینه انتخاب شده در روش مبتنی بر همبستگی نیز در جدول شماره (۳) ارائه شده است. همچنین، روش تحلیل عاملی منجر به استخراج ۲۰ عامل گردید.

1 Kaiser-Meyer-Olkin (KMO)

www.SID.ir

Archive of SID

جدول ۱. متغیرهای پیش‌بین اولیه

ردیف	نام متغیر	نام اختصاری	ردیف	نام متغیر	نام اختصاری
۱	اندازه	Size	۲۷	رشد نسبت حاشیه سود عملیاتی	ΔOPM
۲	نسبت سود به قیمت	E/P_t	۲۸	نسبت حاشیه سود خالص	NPM
۳	نسبت ارزش دفتری به بازار سهام	BV/MV	۲۹	رشد حاشیه سود خالص	ΔNPM
٤	نسبت بدهی	Lev	۳۰	بازده سهام سال جاری	R_t
٥	نسبت بدهی به حقوق صاحبان سهام	TL/E	۳۱	نسبت پوشش هزینه‌های بهره	Cov
٦	بازده دارایی‌ها	ROA	۳۲	جریان‌های نقدی فعالیت‌های عملیاتی	OCF
۷	رشد بازده دارایی‌ها	ΔROA	۳۳	رشد جریان‌های نقدی فعالیت‌های عملیاتی	ΔOCF
۸	بازده حقوق صاحبان سهام	ROE	۳٤	خالص وجه نقد بازده سرمایه‌گذاری‌ها و سود پرداختی بابت تامین مالی	IRFEF
۹	رشد بازده حقوق صاحبان سهام	ΔROE	۳٥	خالص وجه نقد فعالیت‌های سرمایه‌گذاری	ICF
۱۰	ریسک سیستماتیک	Beta	۳٦	خالص وجه نقد فعالیت‌های تامین مالی	FCF
۱۱	سود هر سهم	EPS	۳۷	رشد فروش خالص	ΔS
۱۲	سود تقسیمی هر سهم	DPS	۳۸	رشد دارایی‌ها	ΔA
۱۳	نسبت سود هر سهم به قیمت اول دوره هر سهم	$EPS/P_{(t-1)}$	۳۹	رشد حقوق صاحبان سهام	ΔE
۱٤	نسبت سود تقسیمی هر سهم به قیمت اول دوره هر سهم	$DPS/P_{(t-1)}$	٤۰	نسبت سود عملیاتی به مجموع دارایی‌ها	OI / TA
۱٥	رشد سود تقسیمی هر سهم	ΔDPS	٤۱	رشد نسبت سود عملیاتی به دارایی‌ها	ΔOI / TA
۱٦	نسبت جاری	CR	٤۲	اقلام تعهدی	ACC
۱۷	نسبت آنی	QR	٤۳	نسبت گردش سهام	Turn
۱۸	گردش موجودی کالا	Inv Turn	٤٤	تعداد دفعات مبادلات	NT
۱۹	رشد گردش موجودی کالا	ΔInv Turn	٤٥	درصد سود تقسیمی	DPR
۲۰	نسبت گردش دارایی‌های ثابت	FA Turn	٤٦	سود عملیاتی	OI

www.SID.ir

Archive of SID

۲۱	رشد نسبت گردش دارایی‌های ثابت	Δ FA Turn	٤۷	رشد سود عملیاتی	Δ OI
۲۲	گردش مجموع دارایی‌ها	TA turn	٤۸	سود خالص	NI
۲۳	رشد گردش مجموع دارایی‌ها	Δ TA turn	٤۹	رشد سود خالص	Δ NI
۲٤	نسبت حاشیه سود ناخالص	GPM	۵۰	سود جامع	CI
۲۵	رشد نسبت حاشیه سود ناخالص	Δ GPM	۵۱	رشد سود (زیان) جامع	Δ CI
۲٦	نسبت حاشیه سود عملیاتی	OPM	۵۲	نسبت خالص سرمایه در گردش به دارایی‌ها	WC/TA

منبع: یافته‌های پژوهش

جدول ۲. میانگین متغیرهای مستقل (پیش‌بین)

ردیف	نام متغیر	نحوه سنجش	میانگین	ردیف	نام متغیر	نحوه سنجش	میانگین
۱	Size	Log $MV_{(t-1)}$	۱۲/٤٦	۲۷	Δ OPM	Δ (OI/S)	۰/۰۵٤
۲	E/P	NI/P	۰/۲٤۱	۲۸	NPM	NI/S	۰/۱۷
۳	BV/MV	BV/MV	۰/۱۸٤	۲۹	Δ NPM	Δ (NI/S)	۰/۰۰۵
٤	Lev	TL/TA	۰/٦۷۳	۳۰	R_t	Rt	۰/۲٦۱
۵	TL/E	TL/E	۲/۰۵۸	۳۱	Cov	EBIT/Int	۹۵/۹۵۹
٦	ROA	NI /TA	۰/۱٤۵	۳۲	OCF	$OCF/MV_{(t-1)}$	۰/۱۱۲
۷	Δ ROA	Δ (NI /TA)	-۱۲/۸۰۱	۳۳	Δ OCF	$\Delta OCF/MV_{(t-1)}$	۰/۰۳۱
۸	ROE	NI/E	۰/۳۱۵	۳٤	IRFEF	$IRFEF /MV_{(t-1)}$	۰/۰۰۲
۹	Δ ROE	Δ (NI/E)	-۷/٤٤۵	۳۵	ICF	$ICF/MV_{(t-1)}$	-۰/۰٦۹
۱۰	Beta	Beta	۰/۳٤۵	۳٦	FCF	$FCF/MV_{(t-1)}$	-۰/۳۵٤
۱۱	EPS	EPS	۷۹۹/۵۷۹	۳۷	Δ S	Δ S	۲٤/۲۹٤
۱۲	DPS	DPS	٤۷۹/۱۱٤	۳۸	Δ A	Δ A	۲۱/۹٦۰
۱۳	$EPS/P_{(t-1)}$	$EPS/p_{(t-1)}$	۰/۳۱٦	۳۹	Δ E	Δ E	۲٤/۸۷۲
۱٤	$DPS/P_{(t-1)}$	$DPS/p_{(t-1)}$	۰/۱۹۱	٤۰	OI / TA	OI / TA	۰/۱۷۱
۱۵	Δ DPS	Δ DPS	۰/۲۹٦	٤۱	Δ OI / TA	Δ (OI / TA)	۰/۱۲۵
۱٦	CR	CA/CL	۱/٦۷٤	٤۲	ACC	$ACC / MV_{(t-1)}$	۰/۰۸۲
۱۷	QR	(CA-Inv) / CL	۰/۸۷۵	٤۳	Turn	TV/ OS	۰/۱٤۲
۱۸	Inv Turn	$COGS/ Inv_{(ave)}$	٤/۱۰۸	٤٤	NT	NT	۸٦۷۲
۱۹	Δ Inv Turn	$\Delta COGS/ Inv_{(ave)}$	۰/۱۵٤	٤۵	DPR	DPS / EPS	۰/٦۲۱

www.SID.ir

Archive of SID

۲۰	FA Turn	S/FA	۵/۱۴۳	٤٦	OI	OI / $MV_{(t-1)}$	۰/۱٤۳
۲۱	Δ FA Turn	Δ S/FA	۰/۰۹٤	٤۷	Δ OI	Δ OI	۱۱/ ٦۹۹
۲۲	TA turn	S/TA	۰/۸۰۸	٤۸	NI	NI / $MV_{(t-1)}$	۰/۱۲۷
۲۳	Δ TA turn	Δ S/TA	۰/۱٤۷	٤۹	Δ NI	Δ NI	۲۲/۳۲۱
۲٤	GPM	GP/S	۰/۲٦٤	۵۰	CI	CI / $MV_{(t-1)}$	۰/۱۵٦
۲۵	Δ GPM	Δ GP/S	۰/۰٦٤	۵۱	Δ CI	Δ CI	۲٦/۲۵٤
۲٦	OPM	OI/S	۰/۲۱۱	۵۲	WC/TA	WC/TA	۰/۰٦۲

ریسک :Beta سود خالص :NI حقوق صاحبان سهام:E کل دارایی‌ها:TA کل بدهی‌ها :TL ارزش دفتری:BV قیمت :P ارزش بازار شرکت:MV
بهای تمام شده :COGS موجودی کالا :Inv بدهی‌های جاری:CL دارایی‌های جاری:CA سود تقسیمی هر سهم :DPS سود هر سهم :EPS سیستماتیک
سود قبل بهره و مالیات :EBIT هزینه‌های مالی :Int بازده سهام :R سود عملیاتی :OI سود ناخالص :GP دارایی‌های ثابت :FA فروش :S کالای فروش رفته
جریان نقد :FCF جریان نقد فعالیت‌های سرمایه‌گذاری :ICF بازده سرمایه گذاری‌ها و سود پرداختی بابت تامین مالی :IRFEF جریان نقد عملیاتی :OCF
سرمایه در گردش :WC سود جامع :CI تعداد دفعات مبادله :NT تعداد سهام منتشره :OS حجم مبادلات :TV اقلام تعهدی:ACC فعالیت‌های تامین مالی

منبع: یافته‌های پژوهش

جدول ۳. متغیرهای بهینه انتخاب شده در روش مبتنی بر همبستگی

ردیف	نام متغیر	ردیف	نام متغیر
۱	Beta	۵	ΔS
۲	Δ DPS	٦	Δ OI
۳	Lev	۷	R_t
٤	GPM		

منبع: یافته‌های پژوهش

۲-۲-۴- متغیر وابسته

متغیر وابسته در این پژوهش، بازده سهام است که مشابه پژوهش‌های پیشین (به‌عنوان نمونه، دستگیر و خدابنده، ۱۳۸۲ و ثقفی و شعری، ۱۳۸۳) از تقسیم مجموع عایدات یک سهم در طول دوره مالی بر قیمت سهم در ابتدای دوره، محاسبه می‌شود. عایدات شامل افزایش قیمت سهم و سود نقدی پرداخت شده است و با توجه به تأثیر افزایش سرمایه، بازده سهام تعدیل خواهد شد.

$$\text{بازده سهام} = \frac{\text{مزایای سود سهمی} + \text{مزایای حق تقدم} + \text{سود نقدی خالص} + \text{تفاوت قیمت سهام آخر و اول دوره}}{\text{قیمت سهام اول دوره}}$$

(۵)

www.SID.ir

Archive of SID

به منظور افزایش دقت محاسبه، بازده سهام به‌صورت ماهانه محاسبه و میانگین هندسی بازده ماهانه برای یک سال، به‌عنوان بازده سالانه درنظر گرفته شده است. بازده میانگین هندسی، بازده‌های تجمعی و ترکیبی مربوط به گذشته را اندازه‌گیری می‌کند. در طول چند دوره، میانگین هندسی، میانگین نرخ رشد صحیح‌تری نسبت به میانگین حسابی را نشان می‌دهد (فبوزی و مارکویتز، ۲۰۰۲؛ پینتو و همکاران، ۲۰۱۰ و تهرانی و نوربخش، ۱۳۸۶).

۳-۴- جامعه و نمونه آماری

جامعه آماری این پژوهش، کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی دوره زمانی ۱۳۸۳ الی ۱۳۹۲ می‌باشد. از این جامعه، بدون انجام نمونه‌گیری، شرکت‌های حائز شرایط زیر به عنوان نمونه انتخاب شده است:

۱- تا پایان اسفندماه سال ۱۳۸۲ در بورس اوراق بهادار تهران پذیرفته شده باشند.

۲- سال مالی آن‌ها منتهی به پایان اسفندماه بوده و در دوره زمانی مورد بررسی، تغییری در آن ایجاد نشده باشد.

۳- ارزش دفتری حقوق صاحبان سهام آن‌ها در دوره مورد بررسی، مثبت باشد.

۴- شرکت‌ها طی دوره مورد بررسی، فعالیت مستمر داشته و سهام آن‌ها بدون وقفه با اهمیت مورد معامله قرار گرفته باشد (حداقل ۴۰ مبادله در سال).

۵- جزء شرکت‌های تولیدی باشد.

۶- اطلاعات مالی مورد نیاز برای انجام این پژوهش را در دوره زمانی ۱۳۸۳ الی ۱۳۹۲ به‌طور کامل ارائه کرده باشند.

با توجه به بررسی‌های انجام شده، تعداد ۱۰۱ شرکت در دوره زمانی ۱۳۸۳ الی ۱۳۹۲ حائز شرایط فوق بوده و مورد بررسی قرار گرفته است.

۴-۴- روش آزمون فرضیه‌ها

پس از مشخص شدن متغیرها و عامل‌های بهینه به‌وسیله هر یک از روش‌های کاهش متغیر، از روش رگرسیون خطی و روش‌های غیرخطی رگرسیون بردارهای پشتیبان و شبکه‌های عصبی مصنوعی برای پیش‌بینی بازده سهام استفاده شده است. به‌منظور ارزیابی عملکرد رگرسیون بردارهای پشتیبان در پیش‌بینی بازده سهام، معیارهای ارزیابی (شامل میانگین قدرمطلق درصد خطا[1]، مجذور مربع میانگین خطا[2] و ضریب تعیین) مربوط به پیش‌بینی این

1 Mean Absolute Percentage Error
2 Root Mean Squared Error

www.SID.ir

Archive of SID

روش با روش‌های متداول شبکه‌های عصبی مصنوعی و رگرسیون خطی مقایسه می‌شود. معیارهای مزبور، متداول‌ترین معیارهای ارزیابی عملکرد در مسائل پیش‌بینی محسوب می‌شود که نحوه محاسبه آن در جدول شماره (۴) ارائه شده است. شایان ذکر است که در این پژوهش، از داده‌های یک سال قبل شرکت‌ها برای پیش‌بینی بازده سهام استفاده شده است. بالاتر بودن ضریب تعیین و پایین‌تر بودن دو معیار دیگر بیانگر عملکرد بهتر پیش‌بینی است.

جدول ٤. معیارهای ارزیابی عملکرد پیش‌بینی مورد استفاده

نام معیار	نحوه سنجش
مربع مجذور میانگین خطا (RMSE)	$\sqrt{\frac{\sum_{P=1}^{P}(dp - zp)^2}{P}}$
ضریب تعیین (R^2)	$1 - \frac{\sum_{P=1}^{P}(dp - zp)^2}{\sum_{P=1}^{P}(dp - \bar{d}p)^2}$
میانگین قدرمطلق درصد خطا (MAPE)	$\frac{100}{P} \times \sum_{P=1}^{P}\left\lvert\frac{dp - zp}{dp}\right\rvert$
Z_p: مقدار پیش‌بینی شده، d_p: مقدار واقعی $\bar{d}$: میانگین مقادیر	

اقتباس از: اسمیت و گوپتا، ۲۰۰۲ و آذر و کریمی، ۱۳۸۸

افزون‌بر این، به‌منظور ارزیابی عملکرد روش انتخاب متغیر مبتنی بر همبستگی و روش استخراج متغیر تحلیل عاملی، معیارهای ارزیابی (میانگین قدرمطلق درصد خطا، مجذور مربع میانگین خطا و ضریب تعیین) حاصل از پیش‌بینی با متغیرهای انتخاب و استخراج شده در هر یک از این روش‌ها با یکدیگر و همچنین با معیارهای ارزیابی حاصل از عدم انجام مرحله انتخاب متغیرهای پیش‌بین در هر یک از روش‌های خطی و غیرخطی مقایسه می‌شود. منظور از عدم انجام مرحله انتخاب متغیرهای پیش‌بین، پیش‌بینی با استفاده از ۵۲ متغیر پیش‌بین (قبل از کاهش تعداد متغیرها) است. انتخاب متغیرهای بهینه پیش‌بین (روش مبتنی بر همبستگی) و پیش‌بینی بازده سهام با استفاده از روش‌های خطی و غیرخطی مختلف به‌وسیله نرم‌افزارهای Weka نسخه ۳-۷، انجام شده است. اجرای روش تحلیل عاملی نیز با استفاده از نرم‌افزار SPSS انجام شده است. به‌منظور آزمون فرضیه‌های پژوهش نیز از آزمون t زوجی (براساس

www.SID.ir

Archive of SID

صد دقت حاصل از اجرای روایی متقابل ۱۰ بخشی[1] با ۱۰ بار تکرار در هر روش پیش‌بینی) در نرم‌افزار SPSS نسخه ۲۱ استفاده شده است.

۵-۴- روایی متقابل

به منظور بررسی تعمیم‌پذیری پیش‌بینی‌های انجام شده از روایی متقابل ۱۰ بخشی استفاده می‌شود. روایی متقابل ۱۰ بخشی برای برآورد نرخ خطای واقعی کاملاً قابل‌اتکا و کافی است (هو[2]، ۲۰۱۰). در این روش، نمونه اصلی به ۱۰ دسته نمونه فرعی مختلف تقسیم می‌شود. ۹ نمونه فرعی به‌عنوان نمونه‌های آموزشی استفاده می‌شود و نمونه فرعی باقی‌مانده به‌عنوان نمونه آزمایشی، مورد آزمون قرار می‌گیرد. این شیوه تا حدی تکرار می‌شود که هر یک از ۱۰ نمونه فرعی به‌عنوان نمونه آزمایشی مورد آزمون قرار گیرد. در این پژوهش، روایی متقابل ۱۰ بخشی، با استفاده از اجزای مختلف مجموعه داده‌ها، به‌طور مستقل، ۱۰ بار انجام خواهد شد (روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار).

۵- یافته‌های تجربی پژوهش

جدول شماره (۵) میانگین معیارهای ارزیابی (شامل میانگین قدرمطلق درصد خطا، مجذور مربع میانگین خطا و ضریب تعیین) مربوط به پیش‌بینی بازده سهام را بر اساس روش‌های رگرسیون بردارهای پشتیبان (SVR)، شبکه‌های عصبی مصنوعی (ANN) و رگرسیون خطی (LR) در سه حالت- با استفاده از ۵۲ متغیر پیش‌بین (All)، با متغیرهای انتخابی در روش مبتنی بر همبستگی (Corr) و عامل‌های استخراجی تحلیل عاملی (FA)- نشان می‌دهد. با توجه به نتایج این جدول، در هر سه حالت، عملکرد رگرسیون بردارهای پشتیبان بهتر از شبکه‌های عصبی مصنوعی و رگرسیون خطی است. همچنین عملکرد شبکه‌های عصبی مصنوعی بهتر از رگرسیون خطی است.

جدول ۵. میانگین عملکرد روش‌های مختلف پیش‌بینی با استفاده از سه روش کاهش متغیر

	All			Corr			FA		
	ANN	SVR	LR	ANN	SVR	LR	ANN	SVR	LR
RMSE	٦٣/٨٥٧	٥٣/١٢٨	٨٩/٢٤٧	٥٣/٨٥٧	٤٢/٢٤٧	٨٤/٨٩٥	٥٧/٦٢٤	٤٥/٣٧١	٨٦/٨٩٥
MAPE	٠/٢٤٥	٠/١٩٥	٠/٥١٢	٠/٢١٤	٠/١٣٤	٠/٤٦٧	٠/٢٢٧	٠/١٤٧	٠/٤٧٦
R^2	٠/٢٣٨	٠/٣٢٤	٠/٠٤١	٠/٣٨٧	٠/٥٢٣	٠/١١٤	٠/٣٢٥	٠/٥٠٨	٠/١٠٢

منبع: یافته‌های پژوهش

1 10-Fold Cross Validation
2 Hu

www.SID.ir

Archive of SID

در ادامه، برای آزمون فرضیه اول الی سوم و بررسی وجود تفاوت معنادار بین عملکرد پیش-بینی میان هر جفت روش پیش‌بینی، از آزمون t زوجی استفاده شده است. در این راستا، از دقت‌های حاصل از ۱۰ بار تکرارِ روایی متقابل ۱۰ بخشی (روایی مقطع ۱۰ بخشی با ۱۰ بار تکرار) استفاده شد که منجر به ایجاد ۱۰۰ دقت در مورد هر روش پیش‌بینی می‌شود. آماره آزمون و مقدار احتمال مربوطه حاکی از وجود تفاوت معنادار در مورد کلیه معیارهای عملکرد بین هر جفت روش پیش‌بینی مورد مقایسه در هر سه حالت است. جدول‌های شماره (۶) الی (۸) نتایج این آزمون و مقدار احتمال مربوطه را برای معیار ضریب تعیین نشان می‌دهد. با توجه به آماره t و مقدار احتمال مربوطه ($p<۰/۰۵$)، تفاوت معناداری بین میانگین ضریب تعیین هر جفت روش پیش‌بینی مورد مقایسه در هر سه حالت مورد بررسی وجود دارد. در نتیجه، فرضیه های اول الی سوم پژوهش، تأیید می‌شود. بنابراین، با توجه به بهتر بودن معیارهای ارزیابی رگرسیون بردارهای پشتیبان نسبت به شبکه‌های عصبی مصنوعی و رگرسیون خطی و معنادار بودن آن، می‌توان نتیجه گرفت که روش غیرخطی مزبور، به طور معناداری بهتر از رگرسیون خطی و شبکه‌های عصبی مصنوعی، بازده سهام را پیش‌بینی می‌کند. همچنین، با توجه به بهتر بودن معیارهای ارزیابی شبکه‌های عصبی مصنوعی نسبت به رگرسیون خطی و معنادار بودن تفاوت آن، چنین استنباط می‌شود که این روش غیرخطی به طور معناداری، بهتر از روش خطی، بازده سهام را پیش‌بینی می‌کند.

جدول (٦): نتایج آزمون t و مقدار احتمال مربوطه در حالت استفاده از کلیه متغیرها

	SVR	ANN	LR
SVR	-	-	-
ANN	٤/٧٥١ (۰/۰۰۰)	-	-
LR	٧/٦٣٢ (۰/۰۰۰)	٧/٥٤٨ (۰/۰۰۰)	-

منبع: یافته‌های پژوهش

www.SID.ir

Archive of SID

جدول ۷. نتایج آزمون t و مقدار احتمال مربوطه در حالت استفاده از روش انتخاب متغیر مبتنی بر همبستگی

	SVR	ANN	LR
SVR	–	–	–
ANN	۵/۸٦۵ (۰/۰۰۰)	–	–
LR	۸/۲٤٦ (۰/۰۰۰)	۷/۸۹٤ (۰/۰۰۰)	–

منبع: یافته‌های پژوهش

جدول ۸. نتایج آزمون t و مقدار احتمال مربوطه در حالت استفاده از روش استخراج متغیر تحلیل عاملی

	SVR	ANN	LR
SVR	–	–	–
ANN	٦/۲٤۷ (۰/۰۰۰)	–	–
LR	۸/۲۲٤ (۰/۰۰۰)	٦/۵۱۱ (۰/۰۰۰)	–

منبع: یافته‌های پژوهش

به منظور آزمون فرضیه چهارم، معیارهای عملکرد هر روش پیش‌بینی در حالت استفاده از ۵۲ متغیر پیش‌بین (All) با حالت استفاده از متغیرهای انتخابی توسط روش‌های مبتنی بر همبستگی (Corr) و عامل‌های روش تحلیل عاملی (FA) مقایسه می‌شود. اگر عملکرد هر روش پیش‌بینی (بر اساس جدول ۵) در حالت استفاده از روش‌های مبتنی بر همبستگی و تحلیل عاملی نسبت به استفاده از ۵۲ متغیر، بهتر و تفاوت مربوط از نظر آماری (بر اساس جدول‌های ۹ الی ۱۱) معنادار باشد، می‌توان استنباط کرد که روش‌های کاهش متغیر تأثیر مثبت و معناداری بر عملکرد روش پیش‌بینی دارد. نتایج مربوط به مقایسه معیار میانگین ضریب تعیین هر روش پیش‌بینی در حالت استفاده از ۵۲ متغیر و استفاده از دو روش کاهش متغیر در جدول‌های ۹ الی ۱۱ ارائه شده است. با توجه به آماره t و مقدار احتمال مربوطه ($p<۰/۰۵$)، هر دو روش کاهش متغیر، تأثیر مثبت و معناداری بر ضریب تعیین پیش‌بینی دارد، زیرا با توجه به جدول ۵ در صورت استفاده از متغیرها و عامل‌های روش‌های مبتنی بر همبستگی و تحلیل عاملی، عملکرد پیش‌بینی بهتر از استفاده از ۵۲ متغیر اولیه و با توجه به جدول‌های ۹ الی ۱۱ این تفاوت از نظر آماری معنادار است. با این وجود فقط در حالت پیش‌بینی با شبکه‌های عصبی مصنوعی ($p<۰/۰۵$)، تفاوت معناداری بین عملکرد (میزان سودمندی) این دو روش کاهش متغیر وجود دارد.

www.SID.ir

Archive of SID

جدول ۹. نتایج آزمون t و مقدار احتمال مربوطه در زمان پیش‌بینی با SVR

	All	Corr	FA
All	–	–	–
Corr	٦/۲۸٦ (۰/۰۰۰)	–	–
FA	٦/۱۰۷ (۰/۰۰۰)	۱/۳۷۹ (۰/۱۷۱)	–

منبع: یافته‌های پژوهش

جدول ۱۰. نتایج آزمون t و مقدار احتمال مربوطه در زمان پیش‌بینی با ANN

	All	Corr	FA
All	–	–	–
Corr	٥/٥۲۱ (۰/۰۰۰)	–	–
FA	٤/۱۲۸ (۰/۰۰۰)	۳/۱۸٥ (۰/۰۰۰)	–

منبع: یافته‌های پژوهش

جدول ۱۱. نتایج آزمون t و مقدار احتمال مربوطه در زمان پیش‌بینی با LR

	All	Corr	FA
All	–	–	–
Corr	۳/۱٦۳ (۰/۰۰۰)	–	–
FA	۳/٥۲۸ (۰/۰۰۰)	۱/۳٦٤ (۰/۱۷٦)	–

منبع: یافته‌های پژوهش

۶- بحث و نتیجه‌گیری

یافته‌های تجربی بررسی ۱۰۱ شرکت پذیرفته شده در بورس اوراق بهادار تهران در دوره زمانی ۱۳۸۳ الی ۱۳۹۲ حاکی از آن بود که عملکرد رگرسیون بردارهای پشتیبان (شامل میانگین قدرمطلق درصد خطا، مجذور مربع میانگین خطا و ضریب تعیین) به طور معناداری بهتر از رگرسیون خطی و شبکه‌های عصبی مصنوعی است. این موضوع حتی بدون استفاده از

www.SID.ir

Archive of SID

روش‌های انتخاب متغیر (و در صورت استفاده از ۵۲ متغیر پیش‌بین) نیز صادق است. عملکرد شبکه‌های عصبی مصنوعی نیز به طور معناداری بهتر از رگرسیون خطی بوده است. بنابراین، روش‌های غیرخطی به طور معناداری بازده سهام را بهتر از رگرسیون خطی پیش‌بینی می‌کند. به‌نظر می‌رسد یکی از دلایل برتری عملکرد روش‌های غیرخطی نسبت به رگرسیون خطی، احراز نشدن مفروضات رگرسیون خطی در داده‌های مورد بررسی و همچنین ماهیت غیرخطی روابط پیچیده بین متغیرهای بررسی شده است. این نتیجه، با نظریه آشوب (بی‌نظمی) و پویایی‌های غیرخطی نیز سازگار است. رگرسیون خطی بر فرض توزیع نرمال و مفروضات دیگری استوار است، در حالی که رفتار مردم و طبیعت به طور کلی ماهیتی غیرخطی دارد. یافته‌های بررسی سودمندی روش‌های انتخاب و استخراج متغیرهای پیش‌بین نیز حاکی از آن است که عملکرد پیش‌بینی بازده با استفاده از متغیرهای انتخاب و استخراج شده توسط روش‌های مبتنی بر همبستگی و تحلیل عاملی به طور معناداری بهتر از پیش‌بینی با استفاده از کلیه متغیرها (و عدم استفاده از روش‌های انتخاب متغیرهای بهینه) است. در این راستا، متغیرهای انتخاب شده در روش مبتنی بر همبستگی در زمان پیش‌بینی با شبکه‌های عصبی، عملکرد پیش‌بینی بازده را بیشتر از روش تحلیل عاملی بهبود می‌دهد، ولی در زمان پیش‌بینی با روش‌های رگرسیون بردار پشتیبان و خطی، تفاوت معناداری بین سودمندی روش تحلیل عاملی و مبتنی بر همبستگی وجود ندارد. بنابراین، زمان پیش‌بینی با روش غیرخطی رگرسیون بردارهای پشتیبان در هر سه حالت (با استفاده از ۵۲ متغیر پیش‌بین و با متغیرهای انتخابی در روش مبتنی بر همبستگی و عامل‌های استخراجی تحلیل عاملی)، عملکرد پیش‌بینی به طور معناداری بهتر است، که این موضوع حاکی از سودمندی روش غیرخطی مزبور نسبت به شبکه‌های عصبی و رگرسیون خطی است. این نتیجه، مطابق با یافته‌های پژوهش‌های مین و لی (۲۰۰۵)، لی و تاو (۲۰۱۰) و راعی و فلاح پور (۱۳۸۷) است. همچنین، در صورت استفاده از متغیرهای پیش‌بین انتخاب یا استخراج شده نسبت به استفاده از ۵۲ متغیر اولیه، عملکرد پیش‌بینی به طور معناداری افزایش می‌یابد که این موضوع بیانگر سودمندی روش‌های کاهش متغیرهاست. دلیل این امر، مسأله اضافه‌بار ابعاد می‌باشد. به نظر می‌رسد اضافه کردن متغیرهای بیشتر، پارازیت و در نتیجه خطا را افزایش می‌دهد و اضافه کردن متغیرها تنها تا یک حد معین می‌تواند به بهبود پیش‌بینی کمک کند و اضافه کردن بیشتر متغیرها منجر به مسأله اضافه‌بار ابعاد می‌شود. این نتیجه، مطابق با یافته‌های پژوهش‌های ستایش و همکاران (۲۰۱۲)، لیانگ و همکاران (۲۰۱۵) و نمازی و همکاران (۱۳۹۴) است.

www.SID.ir

Archive of SID

۷- پیشنهادهای پژوهش

با توجه به یافته‌های این پژوهش، پیشنهادهای زیر ارائه می‌شود:

- با توجه به متغیرهای بهینه انتخاب شده (جدول ۳)، سرمایه‌گذاران و سایر استفاده‌کنندگان توجه چندانی به اطلاعات صورت جریان وجوه نقد ندارند. بنابراین، به استفاده‌کنندگان از اطلاعات حسابداری، پیشنهاد می‌شود که به اطلاعات این صورت مالی اساسی توجه بیشتری داشته باشند.
- به سرمایه‌گذاران و سایر استفاده کنندگان پیشنهاد می‌شود در تصمیم‌گیری‌های سرمایه‌گذاری مبنی بر خرید و فروش سهام، کاهش خطر سبد سرمایه‌گذاری و ارزیابی ریسک شرکت و سایر تصمیم‌گیری‌ها از رگرسیون بردارهای پشتیبان استفاده کنند.
- با توجه به تأثیر مثبت استفاده از روش‌های انتخاب متغیرهای پیش‌بین نسبت به عدم استفاده از این روش‌ها بر عملکرد پیش‌بینی بازده سهام، به سرمایه‌گذاران و سایر استفاده‌کنندگان پیشنهاد می‌شود که در پیش‌بینی بازده سهام، مرحله انتخاب متغیرهای پیش‌بین را انجام دهند و صرفاً بر اساس پژوهش‌های گذشته متغیرهای پیش‌بین را انتخاب نکنند.

به سایر پژوهشگران نیز پیشنهاد می‌شود که از سایر روش‌های غیرخطی پیش‌بینی و سایر روش‌های انتخاب متغیر برای پیش‌بینی بازده سهام استفاده کنند.

۸- محدودیت‌های پژوهش

در انجام این پژوهش، محدودیت‌هایی وجود داشته که ممکن است بر نتایج و یافته‌های پژوهش تأثیرگذار باشد:

۱. بسیاری از شرایط سیاسی، اقتصادی و اجتماعی ایران (به‌ویژه شرایط تورمی کشور و عدم تهیه صورت‌های مالی تعدیل شده) بر یافته‌های پژوهش مؤثر است که کنترل آن‌ها از حیطه توان پژوهشگر خارج بوده است.

۲. اثرات ناشی از تفاوت در رویه‌های حسابداری شرکت‌ها در اندازه‌گیری، شناسایی و افشای مبادلات و رویدادهای مالی ممکن است نتایج پژوهش را تحت تأثیر قرار دهد. به دلیل دسترسی نداشتن به اطلاعات شرکت‌ها، از این بابت، تعدیلی در متغیرهای استفاده شده صورت نگرفته است.

www.SID.ir

Archive of SID

منابع و مآخذ

۱.ابریشمی، حمید (۱۳۸۷). **مبانی اقتصادسنجی**، جلد دوم، تهران: انتشارات دانشگاه تهران.

۲.آذر، عادل، کریمی، سیروس (۱۳۸۸)؛ "پیش‌بینی بازده سهام با استفاده از نسبت‌های حسابداری با رویکرد شبکه‌های عصبی"، **تحقیقات مالی**، دوره ۱۱، شماره ۲۸، ص‌ص. ۳-۲۰.

۳.پی جونز، چارلز. (۱۳۸۶)؛ **مدیریت سرمایه‌گذاری**، تهرانی، رضا، نوربخش، عسگر. تهران: انتشارات نگاه دانش.

۴.ثقفی، علی، شعری، صابر (۱۳۸۳)؛ "نقش اطلاعات بنیادی حسابداری در پیش‌بینی بازده سهام"، **فصلنامه مطالعات حسابداری**، شماره ۸، ص‌ص. ۸۷-۱۲۰.

۵.جعفریه، حمیدرضا، معتمدی، نگار، ملایی، الهه (۱۳۸۵)؛ "شبکه‌های عصبی و الگوریتم‌های ژنتیک در تجارت"، **ماهنامه تدبیر**، سال ۱۸، شماره ۱۷۷، ص‌ص. ۶۲-۶۸.

۶.خاکی، غلامرضا (۱۳۹۰)؛ **روش تحقیق با رویکرد پایان‌نامه‌نویسی**، تهران: انتشارات بازتاب.

۷.دستگیر، محسن، تاجی، ندا، ساعدی، رحمان (۱۳۹۱)؛ "رابطه بین متغیرهای حسابداری با بازده سهام با استفاده از مدل بازده ژانگ"، **پژوهشنامه حسابداری مالی و حسابرسی**، سال چهارم، شماره ۱۳، ص‌ص. ۴۳-۶۴.

۸.دستگیر، محسن، خدابنده، رامین (۱۳۸۲)؛ "بررسی ارتباط بین محتوای اطلاعاتی اجزای اصلی صورت گردش وجه نقد با بازده سهام"، **مجله علوم انسانی و اجتماعی دانشگاه شیراز**، دوره ۱۹، شماره ۲، پیاپی ۳۸، ص‌ص. ۱۰۰-۱۱۲.

۹.راعی، رضا، فلاح‌پور، سعید (۱۳۸۷)، "کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از نسبت‌های مالی"، **بررسی‌های حسابداری و حسابرسی**، دوره ۱۵، شماره ۵۳: ص‌ص. ۱۷-۳۴.

۱۰. سرایی، حسن (برگرداننده) (۱۳۸۸). **رگرسیون چندمتغیری در پژوهش رفتاری**، تهران: انتشارات سمت.

۱۱. عباسی، ابراهیم، باقری، سحر (۱۳۹۱)؛ "پیش‌بینی بازده سهام با استفاده از مدل‌های غیرخطی آستانه‌ای و بررسی نقش حجم معاملات در بهبود عملکرد این مدل‌ها"، **تحقیقات مالی**، دوره سیزدهم، شماره ۳۲، ص‌ص. ۹۱-۱۰۸.

www.SID.ir

Archive of SID

۱۲. عبده‌تبریزی، حسین، گنابادی، محمود (۱۳۷۵)؛ "تردید در اعتبار مدل‌های مالی"، **مجله حسابدار**، شماره ۱۱۵، ص‌ص. ۱۳-۲۰.

۱۳. فروغی، داریوش و مسعود نادم (۱۳۸۹)".بررسی تأثیر ویژگی های خاص سازمانی بر رابطه ی سود و بازده"، **فصلنامه علمی پژوهشی حسابداری مالی**، سال دوم، شماره ۷، ص‌ص. ۷۳-۹۶.

۱۴. کلانتری، خلیل (۱۳۸۵)؛ **پردازش و تحلیل داده‌ها در تحقیقات اجتماعی - اقتصادی**، تهران: انتشارات شریف.

۱۵. مؤمنی، منصور، فعال‌قیومی، علی (۱۳۸۶)؛ تحلیل‌های آماری با استفاده از SPSS. چاپ اول، تهران: انتشارات کتاب نو.

۱۶. نصیرزاده، فرزانه (۱۳۹۲)؛ "ارزیابی توانایی مدل‌های داده‌کاوی در پیش‌بینی قیمت سهام"، یازدهمین همایش ملی حسابداری ایران، ۱۷ و ۱۸ مهرماه ۱۳۹۲، دانشگاه فردوسی مشهد.

۱۷. نمازی، محمد (۱۳۸۲)؛ **بررسی عملکرد اقتصادی بازار بورس اوراق بهادار در ایران**، تهران: انتشارات معاونت اقتصادی وزارت امور اقتصادی و دارایی.

۱۸. نمازی، محمد، ناظمی، امین (۱۳۸۴)؛ "بررسی تحلیلی تحقیقات انجام شده در بورس اوراق بهادار تهران"، **تحقیقات مالی**، شماره ۱۹، ص‌ص. ۱۳۵-۱۶۶.

۱۹. نمازی، محمد، محمدتبارکاسگری، حسن (۱۳۸۶)؛ "به‌کارگیری مدل چندعاملی برای توضیح بازده سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران"، **مجله علوم انسانی و اجتماعی دانشگاه شیراز**، دوره ۲۶، شماره ۱، پیاپی ۵۰، ص‌ص. ۱۵۷-۱۸۰.

۲۰. نمازی، محمد؛ کاظم‌نژاد، مصطفی و نعمت‌اللهی، محمدمهدی (۱۳۹۴). "مقایسه روش های مختلف انتخاب متغیرهای پیش بین در پیش‌بینی بحران مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران"، **مجله مهندسی مالی و مدیریت اوراق بهادار**، زیر چاپ.

۲۱. هاشمی، سیدعباس و رحمان ساعدی (۱۳۸۸). «بررسی تأثیر ویژگی‌های خاص شرکت‌ها بر محتوای اطلاعاتی سود و جریان‌های نقدی عملیاتی در توضیح بازده سهام»، **فصلنامه علمی پژوهشی حسابداری مالی**، شماره ۴، ص‌ص. ۱۰۸-۱۳۰.

۲۲. هاگان، مارتین تی، دیموث، هاوارد بی، بیل، مارک (۱۳۸۸)؛ **طراحی شبکه‌های عصبی**، کیا، مصطفی، تهران: کیان رایانه سبز.

www.SID.ir

Archive of SID

۲۳. همت‌فر، محمود، حسینی، سید علی‌اکبر، شاه ویسی، فرهاد، نجفی، یوسف (۱۳۹۰)؛ "روابط خطی و غیرخطی بین متغیرهای حسابداری و بازده سهام شرکت‌های صنعت خودرو و ساخت قطعات"، **پژوهشنامه حسابداری مالی و حسابرسی**، سال سوم، شماره ۱۲، ص‌ص. ۱۳۷-۱۵۴.

24. Boyacioglu, M. A., and Avci, D. (2010). "An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the Prediction of Stock Market Return: The Case of the Istanbul Stock Exchange". **Expert Systems with Applications**, 37(12): 7908–7912.

25. DeTienne, K. B.; DeTienne, D. H.; and Joshi, S. A. (2003). "Neural Networks as Statistical Tools for Business Researchers". **Organizational Research Methods,** 6(2): 236-265.

26. Fabozzi, F. J., and Markowitz, H. M. (2002). **The Theory and Practice of Investment Management,** 2nd Edition. John Wiley & Sons.

27. Fama, E. F. (1991). "Efficient Capital Markets: II". **The Journal of Finance**, 46(5): 1575–1617.

28. Gehr, A. (1978). "Some Tests of the Arbitrage Pricing Theory". **Journal of the Midwest Finance Association**, 7: 91–105.

29. Gysen, M.; Huang, C. S.; and Kruger, R. (2013). "The Performance of Linear Versus Non-Linear Models in Forecasting Returns on the Johannesburg Stock Exchange". **International Business & Economics Research Journal**, 12(8): 985-994.

30. Hall, M. A. (1999). "Correlation-based Feature Selection for Machine Learning". **Ph.D Thesis**, University of Waikato.

31. Hall, M. A. (2000). "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning". **In Proceedings of the Seventeenth International Conference on Machine Learning (June 29 - July 02). P. Langley, Ed. Morgan Kaufmann Publishers, San Francisco, CA**, 359-366.

32. Hoglund, H. (2012). "Detecting Earnings Management with Neural Networks". **Expert Systems with Applications**, 39(10): 9564-9570.

www.SID.ir

Archive of SID

33. Hu, Y. C. (2010). "Analytic Network Process for Pattern Classification Problems Using Genetic Algorithms". **Information Sciences**, 180(13): 2528–2539.

34. Hung, C., and Chen, J. H. (2009). "A Selective Ensemble based on Expected Probabilities for Bankruptcy Prediction". **Expert Systems with Applications**, 36(3): 5297–5303.

35. Kanas, A., and A. Yannopoulos (2001). "Comparing Linear and Nonlinear Forecasts for Stock Returns". **International Review of Economics and Finance**, 10(4): 383-398.

36. Karegowda, A. G., Manjunath, A. S., and Jayaram, M. A. (2010). "Comparative Study of Attribute Selection Using Gain Ratio and Correlation Based Feature Selection". **International Journal of Information Technology and Knowledge Management**, 2(2): 271-277.

37. Koprinska, I., Rana, M., and Agelidis, V. G. (2015). "Correlation and Instance Based Feature Selection for Electricity Load Forecasting". **Knowledge-Based Systems**, 82: 29–40.

38. Kothari, S. P. (2001). "Capital Markets Research in Accounting". **Journal of Accounting and Economics**, 31(1-3): 105-231.

39. Lee, M. C., and To, C. (2010). "Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress", **International Journal of Artificial Intelligence & Applications,** Vol. 1, No. 3, pp. 31-43.

40. Min, J. H., and Lee, Y. (2005). "Bankruptcy Prediction Using Support Vector Machine with Optimal Choice of Kernel Function Parameters", **Expert Systems with Applications**, Vol. 28, pp. 603- 614.

41. Lee, M. C., and To, C. (2010). "Comparison of Support Vector Machine and Back Propagation Neural Network in Evaluating the Enterprise Financial Distress". **International Journal of Artificial Intelligence & Applications**, 1(3): 31-43.

42. Liang, D.; Tsai, C. H.; and Wu, H. T. (2015). "The Effect of Feature Selection on Financial Distress Prediction", **Knowledge-Based Systems**, 73: 289–297.

www.SID.ir

Archive of SID

43. Lindenbaum, M., Markovitch, S., and Rusakov, D. (2004). "Selective Sampling for Nearest Neighbor Classifiers". **Machine Learning**, 54(2): 125-152.

44. Lo, S. C. (2010). "The Effects of Feature Selection and Model Selection on the Correctness of Classification". **Proceedings of the 2010 IEEE IEEM**, 989-993.

45. McMillan, D. G. (2007). "Non-linear Forecasting of Stock Returns: Does Volume Help?". **International Journal of Forecasting**, 23(1): 115-126.

46. Min, J. and Lee, Y. (2004). "Business failure prediction with support vector machines and neural networks: A comparative study", The 9th Asia-Pacific Decision Sciences Institute Conference.

47. Mramor, D. and Pahor, M. (1998). "Testing Nonlinear Relationship between Excess Rate of Return on Equity and Financial Rations". **23rd Meeting of the EURO Working Group on Financial Modelling**, 119-134.

48. Mramor, D. and Mramor-Kosta, N. (1997). "Accounting Ratios as Factors of Rate of Return on Equity". **New Operational Approaches for Financial Modelling, Physica-Verlag Heidelberg**, 335-348.

49. Olson, D. and Mossman, C. (2003). "Neural Network Forecasts of Canadian Stock Returns Using Accounting Ratios". **International Journal of Forecasting**, 19(3): 453–465.

50. Omran, M. and Ragab, A. (2004). "Linear Versus Non-linear Relationships between Financial Ratios and Stock Returns: Empirical Evidence from Egyptian Firms". **Review of Accounting and Finance**, 3(2): 84-102.

51. Pinto, J.; Henry, E.; Robinson, T. R.; and J. D. Stowe (2010). **Equity Asset Valuation**, 2th edition. John Wiley & Sons.

52. Roll, R. and Ross, S. A. (1980). "An Empirical Investigation of the Arbitrage Pricing Theory". **Journal of Finance**, 35(5): 1073–1103.

www.SID.ir

Archive of SID

53. Setayesh, M. H., Kazemnezhad, M., Nikouei, M. A., and Azadi, S. (2012). "The Effectiveness of Fuzzy-Rough Set Feature Selection in the Prediction of Financial Distress: A Case of Iranian Context". **Wulfenia Journal**, 19(10): 268-287.

54. Smith, K., and Gupta, J. (2002). **Neural Networks in Business: Techniques and Applications**. Idea Group Publishing.

55. Smola, A. J., and Scholkopf, B. (2004). "A Tutorial on Support Vector Regression". **Statistics and computing**, 14: 199-222.

56. Tsai, C. F. (2009). "Feature Selection in Bankruptcy Prediction". **Knowledge-Based Systems**, 22(2): 120–127.

57. Yu, L. and Liu, H. (2003). "Feature Selection for High-Dimensional Data: A Fast Correlation-based Filter Solution". **Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003),** Washington DC, 856-863.

www.SID.ir