فصلنامــه علمی پژوهشی دانش سرمایه‌گــذاری
سال هشتم/ شماره بیست‌ونهم/ بهار ۱۳۹۸

طراحی مدل هوشمند ترکیبی جهت پیش‌بینی نقاط طلایی قیمت سهام

محمد مشاری
۱-دکتری تخصصی،گروه مدیریت مالی،واحد علی‌آباد کتول،دانشگاه آزاد اسلامی،علی‌آباد کتول،ایران
m.moshari1356@gmail.com

حسین دیده خانی
۲-استادیار،گروه مهندسی صنایع ،واحد علی‌آباد کتول،دانشگاه آزاد اسلامی،علی‌آباد کتول،ایران (نویسنده مسئول)
h.didehkhani@gmail.com

کاوه خلیلی دامغانی
۳-دانشیار،گروه مهندسی صنایع، واحد تهران جنوب،دانشگاه آزاد اسلامی ،تهران،ایران
kaveh.khalili@gmail.com

ابراهیم عباسی
۴-دانشیار،گروه مدیریت مالی، دانشگاه الزهرا، تهران، ایران
abbasiebrahim2000@yahoo.com

تاریخ دریافت: ۹۶/۰۳/۰۱ تاریخ پذیرش: ۹۶/۰۶/۲۷

چکیده

هدف از این پژوهش ارائه یک مدل هوشمند جهت پیش‌بینی نقاط طلایی در نمودار قیمت سهام ،به عنوان یک سیستم پشتیبان تصمیم می‌باشد.جهت انجام این پژوهش از داده‌های صنعت خودرو، و ساخت قطعات طی سال ۱۳۸۰ تا ۱۳۹۵ استفاده‌شده است.در این پژوهش ابتدا نتایج حاصل از به‌کارگیری مدل های مختلف پیش‌بینی مبتنی بر داده‌کاوی مقایسه قرار گرفت،در مرحله بعد متغیرهای پژوهش به وسیله الگوریتم ژنتیک بهینه‌سازی شد و مجددا مدل‌سازی انجام پذیرفت.نتایج حاصل از این پژوهش نشان داد که نقاط طلایی با دقت قابل قبول پیش‌بینی پذیرند و بهینه‌سازی در همه مدل ها باعث افزایش دقت پیش‌بینینمی‌گردد اما خطای فاحش را به میزان قابل توجهی کاهش می‌دهد.

واژه‌های کلیدی: بازارهای مالی، نقاط طلایی، پیش‌بینی مالی، بهینه‌سازی، داده‌کاوی.

Archive of SID
www.SID.ir

۱- مقدمه

یکی از مهم‌ترین دستاوردهای علم قابلیت پیش‌بینی بخشیدن به متغیرها و پدیده‌هامی‌باشد. پژوهشگران علوم مالی نیز با استفاده از ابزارهای مختلف به دنبال طراحی مدل‌هایی هستند که به‌وسیله آن متغیرها و حوادث مدنظرشان در بازارهای مالی را پیش‌بینی کنند. پیش‌بینی سود (جوگ و مک کنومی،۲۰۰۳) پیش‌بینی ورشکستگی (وندا،۲۰۰۴) و پیش‌بینی جریانات نقدی (براچت، ۲۰۰۷) از این دست پژوهش‌هاست.

اما به جرات می‌توان گفت که پیش‌بینی قیمت سهام از مهم‌ترین اهداف در علوم مالی و سرمایه‌گذاری است (هاموز، ۲۰۱۵). امروزه پیش‌بینی قیمت سهام نه‌تنها خیلی چالش‌برانگیز، بلکه بسیار مورد علاقه سرمایه‌گذاران می‌باشد.(چی ام اچ، ۲۰۱۱) پیش‌بینی قیمت سهام از دو جنبه حائز اهمیت است.

اول آنکه پیش‌بینی دقیق قیمت سهام برای بهینه‌سازی پرتفوهای سرمایه‌گذاری و اجرای راهبردهای سرمایه‌گذاری بسیار تأثیرگذار است(چانگ چین، ۲۰۰۵). دوم آنکه با استفاده از روشهای پیش‌بینی مطلوب می‌توان بازدهی را در سطح مشخصی از ریسک افزایش داد و بالعکس (کیومر، ۲۰۰۶).

علیرغم پیشرفت‌ها و دگرگونی‌های انجام پذیرفته، پژوهش‌هایانجام‌شده در مورد پیش‌بینی قیمت سهام از یک موضوع اساسی رنج می برد و آن طول دوره پیش‌بینی است. مامی‌دانیم اگر پیش‌بینی قیمت برای یک دوره کوتاه مدت انجام پذیرد (مثلا روزانه و هفتگی) به علت خودهمبستگیقیمت‌ها در کوتاه مدت عملا پیش‌بینی فاقد ارزش است و اگر در بازه های بلند مدت انجام پذیرد (مثلا سالانه) به علت تغییرات اساسی در شرایط اقتصادی و سیاسی، تغییرات قیمت بسیار شدید است که دقت عملکرد مدل ها بسیار پایین است.

با پیش‌بینی نقاططلایی (نقاط شروع و پایان روند) به عنوان نقاط خریدوفروشعملاً زمان از مدل پیش‌بینی کننده حذف می‌شود. طراحی مدلی که متغیر زمان از آن حذف شده باشد درنتیجه تغیر در نوع متغیر پیش‌بینی شونده است. به بیان دیگر مسئلهبه‌نوعیتعریف‌شده که پیش‌بینی زمان و ادامه روند فاقد اهمیت باشد، آنچه که اهمیت دارد نقطه شروع یاپایان(تغییر روند) است.

نقاط طلایی به علت تکرارپذیری زیادشان در رفتار قیمت سهام بسیار قابل مقایسه و انطباق‌اند، از طریق مقایسه خصوصیات قیمت در نقاط طلاییمی‌توان خصوصیات مشترک این نقاط را استخراج کرد و از این خصوصیات در پیش‌بینی‌ها استفاده نمود. تکرارپذیریبزرگ‌ترین نقطه اتکاء در این پژوهش است. اینپژوهشبه دنبال چرایی به وجود آمدن نقاط طلایی نیست بلکه به دنبال چگونگی و چیستی نقاط طلایی است.

۲- مبانی نظری و مروری بر پیشینه پژوهش

۲-۱ شبکه عصبی مصنوعی[1]

شبکه عصبی مصنوعی به کمک فرآیند یادگیری و با استفاده از پردازشگرهای ساده به نام نرون تلاش می‌کند با شناخت روابط ذاتی میان داده ها، نگاشتی میان فضای ورودی (لایه ورودی) و فضای مطلوب (لایه خروجی) ارائه دهد. در شبکه عصبی مصنوعی سعی بر این است که ساختاری شبیه ساختار بیولوژیکی مغز انسان و شبکه بدن طراحی شود تا مانند آن، قدرت یادگیری و تصمیم‌گیری داشته باشد. لایه یا لایه های مخفی،

Archive of SID

www.SID.ir

اطلاعات دریافت شده از لایه ورودی را پردازش کرده و در اختیار لایه خروجی قرار می‌دهند. هر شبکه با دریافت مثال‌هایی آموزش می‌بیند. آموزش فرآیندی است که به یادگیری منجر می‌شود.یادگیری شبکه، زمانی انجام می‌شود که وزن ارتباطی بین لایه ها چنان تغییر کند که اختلاف مقادیر پیش‌بینی‌شده و محاسبه‌شده در حد قابل قبولی باشد. (سبزی پور،۱۳۸۹)

۲-۲- نزدیک‌ترین همسایه[۲]

یک الگوریتم برای پیدا کردن نزدیک ترین نقطه در فضاهای متریک است. شیوه کار بدین صورت است که: مجموعه S شامل تعدادی نقطه در یک فضای متریک مانند M و نیز یک نقطه پرس وجو Q€M داده‌شده است. هدف پیدا کردن نزدیک ترین نقطه در S به Q است. در بسیاری از موارد فضای Mیک فضایاقلیدسیdعدی و فاصله بین نقاط با معیار فاصله اقلیدسی یا دیگر فاصله‌های متریک سنجیده می‌شود.

۲-۳- درخت تصمیم[۳]

درخت تصمیم یکی از روشهای ناپارامتریک رده‌بندی کردن است که با توجه به نوع متغیر وابسته به دو دسته رده‌بندی درخت برای متغیر رسته ای و رگرسیون درختی برای متغیر پیوسته تقسیم می‌شود. رده‌بندی درختی در راستای روشهایی نظیر تحلیل ممیزی (تابع تشخیص) و رگرسیون لجستیک است. در این روش مجموعه‌ای از شرط‌های منطقی به صورت یک الگوریتم با ساختار درختی برای رده‌بندی یا پیش‌بینی یک پیامد به کار می رود (بهنام پور و همکاران، ۱۳۹۲)

۲-۴- جنگل تصادفی[۴]

جنگل‌های تصادفی یک روش یادگیری ترکیبی برای دسته‌بندی و رگرسیون می باشد که بر اساس ساختاری متشکل از شمار بسیاری درخت تصمیم، بر روی زمان آموزش و خروجی کلاس ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا کار می کنند. جنگل‌های تصادفی برای درختان تصمیم که در مجموعه آموزش دچار بیش برازش می شوند مناسب هستند.

۲-۵- شبکه‌های بیزی[۵]

تئوری بیز یکی از روش های آماری برای رده‌بندی به شمار می‌آید. در این روش کلاس های مختلف، هر کدام به شکل یک فرضیه دارای احتمال در نظر گرفته می شوند. هر رکورد آموزشی جدید، احتمال درست بودن فرضیه‌های پیشین را افزایش و یا کاهش می‌دهد و در نهایت، فرضیاتی که دارای بالاترین احتمال شوند، به عنوان یک کلاس در نظر گرفته‌شده و برچسبی بر آن‌ها زده می‌شود. این تکنیک با ترکیب تئوری بیز و رابطه سببی بین داده ها، به طبقه‌بندی می‌پردازد.

Archive of SID

www.SID.ir

۲-۶- الگوریتم ژنتیک[۶]

این الگوریتم یکی از الگوریتم های جستجو است که بر اساس ژنتیک ارگانیزم های زنده پایه‌گذاری شده است. این الگوریتم اصل "انتخاب یا بقای اصلح" داروین را با یک سری اطلاعات تصادفی ساخت‌یافته ادغام نموده و یک الگوریتم جستجو با خصوصیت روش های تکامل طبیعی ایجاد می‌نماید. در هر نسل مجموعه جدیدی از رشته ها با استفاده از مناسب ترین اجزای نسل قبل ایجاد می‌شود و اجزای جدید برای تناسب مورد آزمون قرار می‌گیرند (شوهنگ ۲۰۰۲). الگوریتم ژنتیک انعطاف‌پذیری بالایی را جهت تلفیق با تکنیک های ابتکاری فراهم می‌سازد و از این طریق حل کارا و موثر یک مساله را میسر می‌سازد.(گلدبرگ، ۱۹۸۹)

۲-۷- پیشینه پژوهش

پنج گروه از پژوهشگران معتقد هستند که پیش‌بینی قیمت سهام امکان‌پذیر نسیت. اولین گروه کسانی هستند که به فرضیه بازار کارا[۷]اعتقاددارند.در بازار کارای سرمایه، اعتقاد بر این است که قیمت سهام انعکاسی از تمام اطلاعات مربوط به آن سهم است و تغییرات قیمت سهام دارای الگوی خاص قابل پیش‌بینی نیست.

دومین گروه کسانی هستند که به علت موثر بودنعوامل متعدد بر تغییرات قیمت سهام و رفتار آشوب گونه و غیر خطی تغییرات قیمت سهام پیش‌بینی قیمت سهام را امری غیر ممکن می‌دانند (منجمی و همکاران ۱۳۸۸). در حقیقت پراکندگی قیمت سهام تحت تاثیر عوامل کلان اقتصادی مانند نرخ بهره و نرخ تورم، وقایع سیاسی مانند جنگ وتهدیدات بین‌المللیو منطقه‌ای، وقایع اجتماعی مانند اعتصاب‌ها و آشوب‌ها و عوامل رفتاری و روانی سرمایه‌گذاران قرار دارد.(چی ام اچ، ۲۰۱۱).اگر قیمت یا بازده در بازارهای مالی با دقت بالایی قابل پیش‌بینی باشد، سیستم‌های پیش‌بینی کننده قیمت تبدیل به یک دستگاه چاپ پول شده که ثروت زیادی را نصیب سرمایه‌گذارانمی‌کند که این امر، در یک اقتصاد پایدار امکان‌پذیرنیست (گرنجر و تیمرمن، ۲۰۰۴). به بیان دیگر اگر اگر تغییرات قیمت سهام در بازار قابل پیش‌بینی باشد (به صورت تصادفی نباشد) احتمال وجود یک ثروت نامحدود برای سرمایه‌گذاران وجود خواهد داشت که با توجه به وضعیت حاضر و نتایج پژوهش‌های موجود این امکان وجود ندارد(گرانگر،۱۹۹۱). سومین گروه نوع رویدادها، در بازارهای مالیرا دلیل اصلی عدم پیش‌بینی پذیری قیمت می‌دانند. هر رویدادی دارای دو ویژگی احتمال و شدت است. مدل‌هایپیش‌بینی کننده قیمت، مبتنی بر احتمالات است و اساسا محققان به شدت رویدادها توجهی نمی‌کنند. طبق نظر این گروه به دلایل رفتاری، پیش‌بینی تغییرات قیمت با احتمال پایین ولی شدت بالا غیر قابل امکان است (نیکلاس تالب،۲۰۱۲).

چهارمین گروه ماهیت بازارهای مالی را دلیل اصلی عدم پیش‌بینی پذیری قیمت سهام می‌دانستند. به نظر این گروه، ماهیت بازی با جمع صفر[۸]بودن در بازارهای مالی دلیل اصلی دشواری پیش‌بینی در این حوزه است. پیش‌بینیدقیق‌تر هوا باعث تغییر رفتار هوا نمی‌شود ولی پیش‌بینی دقیق قیمت سهام بر خود قیمت سهام تاثیر گذار است. یک اقتصاددان به ناملوکاس چیزی را مطرح کرده بود که به نام خودش به عنوان نقد لوکاس معروف شد. نقد لوکاس اظهار داشت که شاید پیش‌گویی اقتصاددانان بر روند اقتصاد اثر بگذارد که اثر آن پیش‌گویی را

Archive of SID

www.SID.ir

خنثی سازد. فرض کنیم اقتصاددانان تورم را پیش‌بینی کنند، خزانه‌داری و بانک مرکزی در واکنش به گفته‌های ایشان وارد عمل شده و با اعمال سیاست‌های پولی و مالی تورم را پایین می‌آورند.(تانگ و لین، ۲۰۰۷)

پنجمین گروه عامل مهمی به نام "اقتصاد اطلاعات" را مطرح می‌سازند. به نظر این گروه شکل‌گیری "اطلاعات نامتقارن" عامل مهمی در پیش‌بینی‌ناپذیری در بازارهای مالی است. همیشهکسانی (مدیران ارشد شرکت ها) هستند که دارای اطلاعات محرمانه باشند.(نیکواقبال و همکاران، ۱۳۹۲)

با این حال گروهی از پژوهشگران با توجه به نتایج حاصل از تحقیقات شان پیش‌بینی در بازارهای مالی و قیمت سهام را امری بسیار دشوار ولی مکان پذیر می‌دانند (مالکی، ۲۰۰۳). ماهیت پیچیده، تکاملی، داشتن خاصیت دینامیکی و غیر خطی که به دلیل تعامل حوادث و شرایط اقتصادی به وجود می‌آید، و انتظارات غیر عقلایی سرمایه‌گذارانپیش‌بینی قیمت سهام را به امری دشوار مبدل می‌سازد نه امری محال (ناکاموری، ۲۰۰۵). کشف و بهبود الگوریتمهای پیش‌بینی کننده و امکان انجام محاسبات پیچیده به وسیله رایانه‌ها، پژوهشگران را در این امر دشوار یاری می رساند (هانگ یانگ، ۲۰۰۹). پژوهشگران در جهت بهبود عملکرد سیستمهای پیش‌بینی کننده سه تحول اساسی را در روند تحقیقاتشان به وجود آوردند.

اولین تحول استفاده از الگوریتمهای ترکیبی به‌جای الگوریتمهای منفرد بود. زمانی که پژوهشگران در حوزه‌های علوم غیر مالی مانند (برودی،۲۰۰۶) در ابهام‌زدایی از احساس کلمات (دی بک، ۲۰۱۲) در پیش‌بینی ریزش مشتری در تحقیقات شان از روش های ترکیبی جهت پیش‌بینی استفاده کردند، این روشها در میان پژوهشگران مالی نیز پرطرفدار شد.عرب مازار و همکاران (۱۳۸۸)از ترکیب شبکه های عصبی و الگوریتم ژنتیک برای پیش بینی قیمت سهام عرضه های اولیه استفاده کردندکه نتایج پژوهش آنها نشان داد ترکیب شبکه های عصبی با الگوریتم ژنتیک قدرت پیش بینی را به طور محسوسی افزایش می دهد. لازم به ذکر است که برتری مطلق الگوریتمهای ترکیبی بر الگوریتمهای منفرد در پیش‌بینی قیمت سهام هنوز به اثبات نرسیده است و نیاز به تحقیقات بیشتر در این زمینه مشهود است (کیومار، ۲۰۰۶).

دومین تحول استفاده از منطق فازی به جای منطق ارسطویی در مدل های پیش‌بینی کننده قیمت سهام بود. قابلیت اتکاءقیمت‌های پیش‌بینی‌شده با استفاده از مدل های مبتنی بر منطق فازی به طرز قابل ملاحظه‌ای، نسب به مدل های مبتنی بر منطق ارسطویی، بیشتراست (سوتومایور،۲۰۰۶).

سومین و مهم‌ترین تحول تغیر در متغیر پیش‌بینی شونده بود. تا قبل‌تر از این همه پژوهش‌ها سعی در پیش‌بینی مطلق قیمت سهام داشتند که نتایج امیدوار کننده نبود. پیش‌بینی بازده قیمتی بجای قیمت مطلق که توسطلیونگانجام پذیرفتاولین ابتکار بود. (بالینگز، ۲۰۱۵) بجایپیش‌بینی مطلق قیمت دامنه‌ای از قیمت شامل حداکثر قیمت و حداقل قیمت را برای پیش‌بینی پیشنهاد کرد.چودری و گری(۲۰۰۸) در پژوهشی با ترکیب ماشین بردار پشتیبان و الگوریتم ژنتیک ،به پیش بینی روند حرکتی قیمت سهام پرداختند.

Archive of SID

www.SID.ir

۳- روش‌شناسی پژوهش

این پژوهش کاربردی و از نوع پژوهش‌های میدانی است.یعنی بر مبنای اطلاعات جمع‌آوری‌شده از بورس اوراق بهادار تهران، از الگوریتم‌های متفاوت جهت پیش‌بینیاستفاده‌شده است. در این پژوهش ابتدا نقاط طلایی نمودار قیمت در ۲۶ شرکت خودرو و ساخت قطعات طی یک دوره ۱۶ ساله (از ابتدای ۱۳۸۰ تا پایان سال۱۳۹۵) از طریق مشاهده استخراج شد. نقاط طلایی، نقاطی هستند که هستند که روند حرکت قیمت در آن نقاط تغییر می‌کند. در شکل شماره (۱) می‌توان نقاط طلایی کف و سقف را مشاهده کرد.

جدول ۱- خلاصه پیشینه پژوهش

محققان	سال	افق سرمایه‌گذاری	نوع و تعداد متغیرهای ورودی	منطق محاسباتی	الگوریتم مورد استفاده	متغیرپیش‌بینی شونده
نصیر زاده و همکاران	۱۳۸۸	روزانه هفتگی ماهانه	تکنیکی۶ روانی۳	فازی قطعی	SVM LARS ANFIS	مطلق قیمت
منجمی و همکاران	۱۳۸۸	روزانه	تکنیکی۶ روانی۲ اقتصادی ۳	فازی قطعی	GFNN ANN ANFIS	مطلق قیمت
عباسی و همکاران	۱۳۹۰	سالانه فصلی	بنیادی ۲ تکنیکی ۲	فازی	ANFIS	مطلق قیمت
نیکواقبال و همکاران	۱۳۹۲	روزانه	اقتصادی ۵	فازی قطعی	ANFIS MFNN NNARX	بازده قیمت
فلاح پور و همکاران	۱۳۹۲	روزانه	تکنیکی ۴۱ روانی ۳	قطعی	SVM GA	روند حرکت قیمت
کوچوهوانگ	۲۰۰۱	روزانه	تکنیکی ۳ روانی ۲	فازی قطعی	ANN GA.ANFIS	مطلق قیمت
اسگابار و کلوته	۲۰۰۲	روزانه	تکنیکی ۱۲	قطعی	ANN GA	نقاط خریدوفروش
ژی یانگ ژانگ	۲۰۰۶	ماهانه	نظرات خبرگان	قطعی	SVM	روند حرکت قیمت
ام تی سانگ و همکاران	۲۰۰۷	روزانه	تکنیکی ۷	قطعی	ANN	مطلق قیمت
یاکوب و همکاران	۲۰۱۱	روزانه	تکنیکی ۱۰	قطعی فازی	SVM ANFIS	روند حرکت قیمت
وانگ و گوپتا	۲۰۱۳	روزانه	تکنیکی ۹	قطعی	ANN ARIMA	بازده قیمت
پژوهش حاضر	**۲۰۱۷**	******	**تکنیکی ۴۸**	**قطعی**	**ANN,DT,RF,KNN, NB**	**نقاط طلایی**

Archive of SID

www.SID.ir

نمودار ۱- نمودار شمعی قیمت سهام شرکت سایپا

منظور از نقطه کف آن سطحی از قیمت است که پس از آن قیمت سهام در یک روند کاملا صعودی یک افزایش قیمت حداقل ۱۰٪ را تجربه کند و منظور از قیمت سقف آن سطحی از قیمت است که پس از آن قیمت سهام در یک روند کاملا نزولی یا خنثی قرار می گیرد به بیان دیگر نقطه سقف نقطه پایان روند صعودی است.

مدل‌سازیبه وسیله ۴۸متغیر ورودی انجام گرفت. از مجموع ۴۸متغیر۵ متغیراولیه یا اصلی (متغیرهای ۱ تا ۵ که در جدول متغیرها فهرست شده اند.) و ۴۳متغیر دیگر که شامل اندیکاتورها[۹] و اسلایتورهای[۱۰] تحلیل تکنیکال[۱۱] هستند که با استفاده از فرمول های ریاضی به دست آمده اند و هر اندیکاتور و اسلایتور با توجه به دوره متداول استفاده توسط تحلیل گران محاسبه‌شده است.

در نهایت برای هر نقطه یک ماتریس متشکل از ۴۸متغیر ورودی و یک متغیر خروجی که یکی از سه حالت کف، سقف و هیچ‌کدام (نه سقف و نه کف) است به دست آمده است. حالت نه سقف و نه کف برای آموزش بهتر مدل تعریف شده است.

جامعه آماری این پژوهش، شرکت های پذیرفته‌شده در بورس اوراق بهادار تهران هستند. که به علت تعدد زیاد شرکتها و فراوانی بیشتر نقاط طلایی در صنعت خودرو و ساخت قطعات این صنعت به عنوان نمونه انتخاب‌شده است. اسامی نمادهای شرکتهای این صنعتبه همراه تعداد نقاط شناسایی‌شده در پیوست یک درج گردیده است.

به طور کلی چهار گام زیر جهت انجام پژوهش انجام‌شده است:

گام اول: شناسایی و جمع‌آوری متغیرهای مسئله

اطلاعات مربوط به متغیرهای این پژوهش از طریق سامانه سازمان بورس و اوراق بهادار استخراج گردید و با استفاده از این اطلاعات اندیکاتورها و اسلایتورها تکنیکی محاسبه شد. در پیوست شماره ۲ تعریف خلاصه‌ای از همه متغیرهای پژوهش ارایه شده است.

Archive of SID

www.SID.ir

گام دوم: استفاده از مدل های منفرد و ترکیبی

در این گام پس از استفاده از الگوریتمهای منفرد و استخراج و مقایسه نتایج، از الگوریتمهای ترکیبی جهت امر پیش‌بینیاستفاده‌شده است. در مدل های ترکیبی ابتدا بهینه‌سازی به کمک الگوریتم ژنتیک انجام پذیرفته است و سپس متغییرهای بهینه‌شده به عنوان ورودی مدل مورد استفاده قرار گرفته است.

گام سوم: پیش‌بینی و ارزیابی اولیه

در این گام داده ها به دو دسته آموزش[۱۲]و آزمایش[۱۳] به نسبت ۷۰٪ به ۳۰٪ تقسیم‌بندی شدند. دانشی که در مرحله یادگیری مدل یادگیری مدل تولید می‌شود، می‌بایست در مرحله ارزیابی مورد تحلیل قرار گیرد تا بتوان ارزش آن را تعیین نمود و در پی آن کارایی الگوریتم یادگیرنده مدل را نیز مشخص کرد.

سنجش اولیه هر مدل به وسیله ماتریس اختلال[۱۴]انجام پذیرفته است. در این ماتریس از دو معیار،طبقه دقت[۱۵]و طبقه اطمینان[۱۶]استفاده شده است.

$$\text{Precision(positive)} = \text{positive predictive value(PPV)} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Precision(negative)} = \text{negative predictive value(NPV)} = \frac{\text{TN}}{\text{TN} + \text{FN}}$$

$$\text{Recall(positive)} = \text{true positive rate} = \text{(TPR)} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Recall(negative)} = \text{true negative rate(TNR)} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

گام چهارم: ارزیابی و مقایسه نهایی مدل ها

جهت ارزیابی نهایی مدل ها از معیارهای چهارگانه زیراستفاده گردید.

- دقت مدل

$$\text{Accuracy(ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

- میانگین موزون اطمینان:

$$\text{Weighted} - \text{mean} - \text{recall(WMR)} = (\frac{\text{TP}}{\text{TP} + \text{FN}}) \div \text{n}$$

- میانگین موزون دقت:

$$\text{Weighted} - \text{mean} - \text{Precision(WMP)} = (\frac{\text{TP}}{\text{TP} + \text{FP}}) \div \text{n}$$

- خطای طبقه‌بندی:

$$\text{classification error} = 1 - \text{Accuracy(ACC)}$$

Archive of SID

www.SID.ir

۳-۱- متغیرهای پژوهش

در جدول ۲ فهرست متغیرهای ورودی مورد استفاده در پژوهش آورده شده است.تعریف عملیاتی متغیرها در پیوست شماره یک آمده است.

جدول ۲- متغیرهای مورد استفاده در پژوهش

نام متغیر	نام در مدل	نام متغیر	نام در مدل
قیمت آغازین	PO	میانگین متحرک موزون	WMA
بالاترین قیمت	PH	میانگین متحرک نمایی	EMA
پایین ترین قیمت	PL	میانگین متحرک نمایی سه گانه	TEMA
قیمت پایانی	PC	میانگین متحرک نمایی دوگانه	DEMA
حجم معاملات	Volume	متوسط شاخص جهت حرکت	ADX
شاخص جریان پول	MFI	دامنه متوسط واقعی	ATR
شاخص قدرت نسبی	RSI	باند پایینی بولینگر	Bbandbot
شاخص جهت حرکت مثبت	PDI	باند بالایی بلوینگر	Bbandbot top
شاخص جهت حرکت منفی	MDI	شاخص کانال کالا با قیمت‌های ابتدایی	CCIa
شاخص کانال کالا	CCI	بالاترین سقف در میان سقف‌های قیمتی	HHV
میانگین متحرک	MA	دورهای فاصله ازHHV	HHVbars
چایکین	chaikin	خط رگرسیون خطی	Line Regression
میانگین متحرک هال	HMA	خط رگرسیون منقطع	Reg Intercept
دورهای فاصله از LLV	LLV bars	شیب خط رگرسیون	LinRegSlop
واگرایی و هم‌گرایی متحرک	MACD	پایین ترین کف قیمتی در میان کف‌های قیمتی	LLV
نرخ میانه	Median Price	شاخص قدرت نسبی با قیمت‌های ابتدایی	RSIa
اندیکاتور قیمت	OSCP	شاخص گام تصادفی	RWI
اندیکاتور حجم	OSCV	شاخص گام تصادفی بالا	RWHI
مرجع گذشته	REF	شاخص گام تصادفی پایین	RWILO
شاخص حرکت نسبی	RMI	هموارکننده وایلدر	WWS
توقف و برگشت	SAR	نوسان نمای فوق‌العاده	Ultimtea
اخطارمکدی	MACDS	تغییرات میانگین متحرک نمایی	TRIX
خطای استاندارد	StdErr	استوکاستیک(D%)	D%
انحراف معیار	StdEvd	استوکاستیک(K%)	K%

Archive of SID

www.SID.ir

۴- مدل‌سازی و تحلیل یافته‌های پژوهش

جهت ارزیابی اولیه از ماتریس اختلال استفاده گردیده است.ماتریس اختلال توانایی و صحت عملکرد مدل ها را ارزیابی می کند.در این ماتریس عملکرد مدل ها در مجموعه داده های آزمایش سنجیده می شود،به این صورت که مدل ها با توجه یادگیری انجام شده در مجموعه داد ه های آموزش متغیر هدف (نقاط کف ،سقف و هیچکدام)را چه چیزی پیش بینی کرده اند؟و در واقعیت این نقاط به چه صورن بوده اند؟

جدول ۳- نتایج اولیه مدل های نزدیک‌ترین همسایه، درخت تصمیم و جنگل تصادفی

واقعیت / پیش‌بینی	نزدیک‌ترین همسایه			درخت تصمیم			جنگل تصادفی			طبقه دقت (٪)		
	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	نزدیک ترین همسایه	درخت تصمیم	جنگل تصادفی
هیچ‌کدام	۹۳	۵۴	۵۴	۲۹	۱۲	۳	۱۵۸	۱۰۶	۵	۴۶,۳	۶۵,۹	۵۸,۷
کف	۶۰	۷۱	۵۶	۱۵۳	۱۵۶	۵	۲۴	۵۸	۴	۳۷,۸	۴۹,۷	۶۷,۴
سقف	۴۲	۵۷	۷۹	۱۳	۱۴	۱۸۱	۱۳	۱۴	۱۸۱	۴۴,۳	۸۷	۸۷
طبقه اطمینان(٪)	۴۷,۷	۳۹	۴۱,۸	۱۴,۹	۸۵,۷	۹۵,۸	۸۱	۳۱,۹	۹۵,۲			

جدول ۴- نتایج اولیه مدل های نایبویز و شبکه عصبی

واقعیت / پیش‌بینی	نایبویز			شبکه عصبی			طبقه دقت (٪)	
	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	نایبویز	شبکه عصبی
هیچ‌کدام	۱۵۲	۸۲	۲	۱۴۱	۴۳	۳	۶۴,۴	۷۵,۴
کف	۳۰	۷۳	۱۱	۴۷	۱۲۸	۷	۶۴	۷۰,۳
سقف	۱۳	۲۷	۱۷۶	۱	۱۱	۱۷۹	۸۱,۵	۹۰,۹
طبقه اطمینان (٪)	۷۷,۹	۴۰,۱	۹۳,۱	۷۲,۳	۷۰,۳	۹۴,۷		

www.SID.ir

تحلیل نتایج ماتریس اختلال نشان می دهد که همه الگوریتم ها به غیر از نزدیک‌ترین همسایه بهترین عملکرد را در پیش‌بینی نقطه سقف دارند این بدان معناست که تقریبا الگوریتم های متفاوت عملکرد مشابهی در پیش‌بینی نقاط سقف دارند. همچنین تمام الگوریتم ها به غیر درخت تصمیم بدترین عملکرد را در پیش‌بینی نقطه کف داشته اند.

به طور خلاصه می‌توان گفت نقاط سقف (نقاط پایان روند صعودی) از قابلیت پیش‌بینی بالاتری برخوردارند و نقاط کف (نقاط شروع روند صعودی) از قابلیت پیش‌بینی پایینی برخوردارند. این بدان معناست که فارغ از الگوریتم مورد استفاده پیش‌بینی نقاط سقف با خطای کمتریو پیش‌بینی نقاط کف با خطای بیشتری امکان‌پذیراست.در جدول شماره پنج بهترین و بدترین عملکرد هر الگوریتم در پیش بینی نقاط طلایی درج گردیده است.

جدول ۵- مقایسه توانایی مدل ها

الگوریتم	بهترین عملکرد	بدترین عملکرد
نزدیک‌ترین همسایه	هیچ‌کدام	کف
درخت تصمیم	سقف	هیچ‌کدام
جنگل تصادفی	سقف	کف
نایبویز	سقف	کف
شبکه عصبی	سقف	کف

۴-۲- نتایج مقایسه‌ای مدل ها با استفاده از معیارهای چهار گانه:

با توجه به معیارهای سه گانه ارزیابی بهترین عملکرد را مدل مبتنی بر الگوریتم شبکه عصبی دارد. نتایج ارزیابی و مقایسه نهایی نشان می دهد که بطور کلی شبکه عصبی بالاترین میزان دقت و پایین ترین میزان خطا و الگوریتم نزدیکترین همسایه پایین ترین میزان دقت و بالاترین میزان خطا را دارد.

جدول ۶- نتایج نهایی سنجش عملکرد مدل ها

الگوریتم / معیار	شبکه عصبی	نایبویز	جنگل تصادفی	درخت تصمیم	نزدیک‌ترین همسایه
دقت	۷۹,۱۵	۷۰,۸۵	۶۹,۹۶	۶۴,۶۶	۴۲,۹۳
خطای طبقه‌بندی	۲۰,۸۵	۲۹,۱۵	۳۰,۰۴	۳۵,۳۴	۵۷,۰۷
میانگین موزون اطمینان	۷۹,۱۲	۷۰,۳۹	۶۹,۳۸	۶۵,۴۵	۴۲,۸۳
میانگین موزون دقت	۷۸,۸۶	۶۹,۹۷	۷۰,۵	۶۷,۵۴	۴۲,۸۷

Archive of SID

www.SID.ir

۴-۳- نتایج مدل‌های بهینه‌شده

در این قسمت ابتدا متغیرهای پژوهشبه وسیله الگوریتم ژنتیک بهینه‌سازی گردیدند و بعد از آن متغیرهای بهینه‌شده به عنوان ورودی مدل‌هامورد استفاده قرار گرفتند.این مرحله نیز مانند مرحله قبل جهت ارزیابی اولیه از ماتریس اختلال استفاده شده است.

جدول ۷- نتایج اولیه مدل های نزدیک‌ترین همسایه، درخت تصمیم و جنگل تصادفی بهینه‌شده

واقعیت / پیش‌بینی	نزدیک‌ترین همسایه			درخت تصمیم			جنگل تصادفی			طبقه دقت (٪)		
	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	نزدیک ترین همسایه	درخت تصمیم	جنگل تصادفی
هیچ‌کدام	۱۴۹	۳۰	۵	۰	۰	۰	۱۸۹	۱۸۵	۱۹۲	۸۱	۰	۳۳,۴
کف	۵۷	۱۲۹	۹	۰	۰	۰	۰	۰	۰	۶۶,۱	۰	۰
سقف	۱۴	۹	۱۶۴	۱۸۹	۱۸۷	۱۹۰	۰	۰	۰	۸۷,۷	۳۳,۴	۰
طبقه اطمینان(٪)	۶۷,۷	۷۶,۸	۹۲,۱	۰	۰	۱۰۰	۱۰۰	۰	۰			

جدول ۸- نتایج اولیه مدل های نایبویز و شبکه عصبی بهینه‌شده

واقعیت / پیش‌بینی	نایبویز			شبکه عصبی			طبقه دقت (٪)	
	هیچ‌کدام	کف	سقف	هیچ‌کدام	کف	سقف	نایبویز	شبکه عصبی
هیچ‌کدام	۱۳۹	۶۱	۳	۱۲۹	۴۰	۴	۶۸,۵	۷۴,۶
کف	۵۲	۱۱۷	۱۵	۳	۱۴۹	۳	۶۳,۶	۹۶,۱
سقف	۹	۱۸	۱۵۲	۳	۶	۱۷۷	۸۴,۹	۹۵,۲
طبقه اطمینان(٪)	۶۹,۵	۵۹,۷	۸۹,۴	۹۵,۶	۷۶,۴	۹۶,۲		

بررسی نتایج نشان می‌دهد که بهینه‌سازی به شدت بر عملکرد الگوریتم نزدیک‌ترین همسایه تاثیر مثبت دارد. می‌توان ادعا کرد که بهینه‌سازی باعث افزایش دقت عملکرد به میزان تقریبی دو برابر شده است. نکته قابل توجه اینکه بهینه سازی بهترین و بدترین عملکرد الگوریتم ها را نیز دستخوش تغیر کرد ،بطور مثال :الگوریتم نزدیک‌ترین همسایه بهینه‌شده بهترین عملکرد را در پیش‌بینی نقاط سقف دارد (در حالی در مدل بهینه نشده

Archive of SID

www.SID.ir

بهترین عملکرد در نقطه نه سقف نه کف بود)والگوریتم بدترین عملکرد را در پیش‌بینی نقاط نه سقف نه کفدارد (در حالی که در مدل بهینه نشده نقطه کف بود).

نتایج نشان می‌دهد که بهینه‌سازی در دو مدل درخت تصمیمو جنگل تصادفی باعث تضعیف عملکرد مدل می گردد،به نحوی که بهینه سازی در مدل درخت تصمیم باعث شده مدل کلیه نقاط را نقطه سقف و در مدل جنگل صادفی کلیه نقاط را نه سقف و نه کف پیش بینی کند. می توان ادعا کرد که بهینه سازی در مدل جنگل تصادفی و درخت تصمیم باعث از بین رفتن خاصیت پیشگویی مدل ها گردیده است.

بهینه‌سازی باعث بهبود عملکرد کلیمدل نایبویز گردیده است اما تاثیر آن بر کلیه نقاط یکسان نیست. بهینه‌سازی بیشترین تاثیر را بر بهبود عملکرد مدل در پیش‌بینی نقطه کف گذاشته است ولی کماکان مدل بهینه‌شده همانند مدل بهینه نشده بهترین عملکرد را در پیش‌بینی نقاط سقف و ضعیف ترین عملکرد را در پیش‌بینی نقاط کف دارد. بهینه‌سازی بهترین تاثیر را بر مدل شبکه عصبی گذاشته است به نحوی که باعث بهبود عملکرد مدل در پیش‌بینی کلیه نقاط شده است.

۴-۴- نتایج نهایی مدل های بهینه‌شده

بررسی نتایج حاصله بیانگر این موضوع است کهشبکه عصبی بهینه‌شده بهترین عملکرد را از همه جهات در پیش‌بینی نقاط طلایی قیمت سهام دارد.

جدول ۹- نتایج نهایی مدل های بهینه‌شده

الگوریتم / معیار	شبکه عصبی	نایبویز	جنگل تصادفی	درخت تصمیم	نزدیک‌ترین همسایه
دقت	۸۰,۳۹	۷۲,۰۸	۳۳,۳۹	۳۳,۵۷	۷۸,۰۹
خطای طبقه‌بندی	۱۹,۶۱	۲۷,۹۲	۶۶,۶۱	۶۶,۴۳	۲۱,۹۱
میانگین موزون اطمینان	۸۰,۵۳	۷۲,۸۷	۳۳,۳۳	۳۳,۳۳	۷۸,۲۸
میانگین موزون دقت	۸۰,۵۷	۷۲,۳۳	۱۱,۱۳	۱۱,۱۹	۷۸,۸۸

۵-۵- بررسی نتایج خطای فاحش مدل ها

کیفیت اشتباه پیش‌بینی مدل های طراحی‌شده نیز حائز اهمیت است،یعنیتعداد نقاطسقف که مدل به اشتباه آنرا کف پیش‌بینی کرده و یا بلعکس نقاط کفی که به اشتباه سقف پیش‌بینی‌شدهمی‌تواند باعث تصمیم گیرهای بسیار اشتباه و زیان‌آور گردد از این‌رو این اشتباهات در پیش‌بینی اشتباه فاحش نامیده شده و در جدول ۱۵. میزان اشتباهات فاحش هر مدل قبل و بعد از بهینه‌سازی آورده شده است.

Archive of SID

www.SID.ir

جدول ۱۰- تعداد خطای فاحش

خطای فاحش / مدل	قبل از بهینه‌سازی	بعد از بهینه‌سازی
نزدیک‌ترین همسایه	۱۱۶	۱۸
درخت تصمیم	۱۹	۱۸۷
جنگل تصادفی	۲۲	۰
نایبویز	۳۸	۳۳
شبکه عصبی	۱۸	۹

۶- بحث و نتیجه‌گیری

نقاط طلایی با دقت قابل قبولی قابل پیش‌بینی هستند. ترکیب شبکه عصبی و الگوریتم ژنتیک این مهم را به بهتر از سایر مدل های منفرد و ترکیبی انجام می‌دهد. مدل مبتنی بر الگوریتم شبکه عصبی با دقت ۸۰٪ می‌تواند نقاط شروع یک‌روند صعودی و پایان یک‌روند را پیش‌بینی کند. مدل‌های ترکیبی و بهینه‌سازی توان بهبود عملکرد بعضی از مدل های شبکه عصبی، نایوبیز و نزدیک‌ترین همسایه را دارد که این موضوع نتایج تحقیقات چینزو (۲۰۰۵) را تایید کرد و در بعضی از مدل ها(درخت تصمیم و جنگل تصادفی) نه تنها باعث بهبود عملکرد نشد بلکه به شدت تاثیر منفی بر صحت عملکرد مدل ها گذاشت که این موضوع نتایج تحقیقات کیومار (۲۰۰۶) را تایید کرد.بطور کلی می توان گفت:بهینه سازی در مدل های مبتنی بر الگوریتم های رده بندی تاثیر منفی دارد.

نکته قابل توجه دیگر که از تحلیل ماتریس اختلال قابل فهم است عملکرد بهتر مدل ها در پیش‌بینی نقاط پایان روند صعودی (جایی که قیمت شروع به ریزش می‌کند) می‌باشد. بااینکهمتغیرهای مورد بررسی در نقاط شروع و پایان روند یکسان است ولی نقاط شروع روند (جایی که قیمت شروع به رشد می‌کند) عملکرد مدل ها ضعیف‌تر است.به نظر می رسد دقت بهتر مدل ها در پیش بینی نقاط سقف به علت مشابهت بالای خصوصیات قیمت در این نقطه باشد که این موضوع باعث آموزش بهتر مدل ها در این نقطه می شود،و این آموزش بهتر باعث عملکرد بهتر مدل ها می گردد.بهینه سازی تاثیر متفاوتی بر مدل های مختلف دارد اما باعث کاهش میزان خطای فاحش در تمامی مدل ها به غیر درخت تصمیم شده است.

فهرست منابع

* امیری، بیگلری کامی، مقصود، مهدی (۱۳۹۳)، پیش‌بینی رفتار سهام با استفاده از مدل زنجیر مارکوف، مجله مهندسی مالی و مدیریت اوراق بهادار،شماره ۲۰،پاییز ۱۳۹۳
* بهنام پور،ناصر،حاجی زاده،ابراهیم.سمنانی،شهریار،زایری، (۱۳۹۲).معرفی الگوریتم های مدل رده‌بندی درختی و کاربرد آن در تعیین عوامل موثر بر ابتلاء به سرطان مری در استان گلستان.جورجانی،۲-۴۶-۵۵.

Archive of SID
www.SID.ir

* تهرانی،رضا،مرادپورسعید (1391).پیش‌بینی بازده شاخص بورس اوراق بهادار با استفاده از مدلهای شبکه‌های عصبی مصنوعی شعاع پایه.مجله مهندسی مالی و مدیریت اوراق بهادار،شماره دهم،بهار ۱۳۹۱
* راعی رضا (۱۳۸۰)،شبکه‌های عصبی،رویکردی نوین در تصمیم‌گیری‌های مدیریت،مدرس،دوره پنجم،ص ۱۵۴،۱۳۸۰-۱۳۳
* سبزی پور،علی اکبر،بیات ورکشی،مریم (۱۳۸۹).ارزیابی دقت روش های شبکه عصبی مصنوعی و عصبی فازی در شبیه‌سازی تابش کل خورشیدی،مجله پژوهش فیزیک ایران،زمستان،دوره ۱۰،شماره ۴،ص ۳۴۷-۳۵۷
* عباسی ابراهیم، (۱۳۹۴) تنظیمپارامتر اندیکاتورهای تحلیل تکنیکال با استفاده از بهینه‌سازی چند هدفه گروه ذرات و سیستم استنتاج فازی-عصبی،فصلنامه علمی پژوهشی دانش سرمایه‌گذاری سال چهارم،شماره پانزدهم،پاییز ۱۳۹۴
* عباسی،ابوئی مهریزی،ابراهیم،امیر (۱۳۹۰)،کاربرد شبکه عصبی -فازی انطباقی در پیش‌بینی قیمت سهام شرکت ایران خودرو،مجله مهندسی مالی و مدیریت اوراق بهادار شماره هفتم،تابستان ۱۳۹۰
* فلاح پور و همکاران، (۱۳۹۲)،پیش‌بینی روند حرکتی قیمت سهام با استفاده از ماشین بردار پشتیبان بر پایه الگوریتم ژنتیک در بورس اوراق بهادار تهران،مجله علمی پژوهشی تحقیقات مالی،دوره ۱۵،شماره ۲،پاییز و زمستان ۱۳۹۲،ص ۲۸۸-۲۶۹
* منجمی سید امیرحسین و همکاران (۱۳۸۸)،پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار با استفاده از شبکه‌های عصبی فازی و الگوریتم ژنتیک و مقایسه آن با شبکه عصبی مصنوعی،فصلنامه اقتصاد مقداری،دوره ۶،شماره ۳،پاییز ۱۳۸۸،ص ۱-۲۶.
* مهدوی غلامحسین و محمدرضا بهمنش، (۱۳۸۴)،طراحی مدل پیش‌بینی قیمت سهام شرکت های سرمایه‌گذاری با استفاده از شبکه عصبی مصنوعی،فصلنامهپژوهشنامه اقتصادی،شماره ۱۹،ص ۲۱۱-۲۳۳.
* نصیر زاده،نیک‌روش،فرزانه،زهرا،ارزیابی توانایی مدل های داده‌کاوی در پیش‌بینی قیمت سهام،پایان نامه کارشناسی ارشد
* نیکو اقبال و همکاران (1393)،ارزیابی مدل های شبکه عصبی مصنوعی ایستا و پویا در پیش‌بینی قیمت سهام،فصلنامه علمی پژوهشی دانش مالی تحلیل اوراق بهادار، سال هفتم،شماره بیست و دوم تابستان ۱۳۹۳،ص ۲۱-۲۹.
* عرب مازار یزدی،محمد و مهسا قاسمی(۱۳۸۸)،قیمت گذاری عرضه های اولیه:ترکیب شبکه های عصبی مصنوعی و الگوریتم ژنتیک ،بررسی های حسابداری و حسابرسی،دوره ۱۶،ص۸۷-۱۰۲.
* نیکو اقبال،علی اکبر نادیا گندلی علیخانی اسماعیل نادری (۱۳۹۳)،ارزیابی مدل های شبکه عصبی مصنوعی ایستا و پویا در پیش‌بینی قیمت سهام،فصلنامه علمی پژوهشی دانش مالی تحلیل اوراق بهادار سال هفتم،شماره بیستم و دوم تابستان۱۳۹۳

www.SID.ir

* Tsang,P.M.,Kwok,P.,Choy,S.O.,Kwan,R.,Ng,S.C.,Mak,j.,Koong,K.,Wong,TL.(2007)Design and implementation of NN5 for Hong Kong stock price foreecasting,Engineering Application of Artificial Intelligence 20:453-461
* Al-Hmouz, R., Pedrycz, W., & Balamash, A. (2015). Description and prediction of time series: a general framework of granular computing. Expert Systems with Applications, 42(10), 4830-4839.
* Ballings, M., Van den Poel, D., Hespeels, N., & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction.Expert Systems with Applications, 42(20), 7046-7056.
* Brochet, F., Nam, S., & Ronen, J. (2007).Accruals and the prediction of future cash flows.International Business & Economics Research Journal, 2(3), 55-82.
* Brody, S., Navigli, R., &Lapata, M. (2006, July).Ensemble methods for unsupervised WSD.In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics (pp. 97-104).Association for Computational Linguistics.
* Browne,C.H(2007).The little book of value investing (vol.5).john wiley&sons
* Cheung, Y. W., Chinn, M. D., &Pascual, A. G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive?.Journal of international money and finance, 24(7), 1150-1175.
* Choudry,R&Grag K.(2008).A Hybrid Machin Learning System for Stock Market Forecasting.Word Academy of Science,Engineering and Technology,39.
* De Bock, K. W., & Van den Poel, D. (2012). Reconciling performance and interpretability in customer churn prediction using ensemble learning based on generalized additive models. Expert Systems with Applications, 39(8), 6816-6826.
* Goldberg, D. E., & Holland, J. H. (1988). Genetic algorithms and machine learning.Machine learning, 3(2), 95-99.
* Granger, C. W. J. (1991).Forecasting Stock market prices, Lessons for casters (pp. 178-179). Working paper ،university of California, San Diego, Department of Economics
* Hamel, L. H. (2011). Knowledge discovery with support vector machines (Vol. 3). John Wiley & Sons
* Hsu, C. M. (2011).A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming.Expert Systems with Applications, 38(11), 14026-14036.
* Huang, C. J., Yang, D. X., & Chuang, Y. T. (2008).Application of wrapper approach and composite classifier to the stock trend prediction.Expert Systems with Applications, 34(4), 2870-2878.
* Huang, W., Nakamori, Y., & Wang, S. Y. (2005).Forecasting stock market movement direction with support vector machine.Computers & Operations Research, 32(10), 2513-2522.
* Jog, V., & McConomy, B. J. (2003). Voluntary disclosure of management earnings forecasts in IPO prospectuses. Journal of Business Finance & Accounting, 30(1-2), 125-168.
* Kara, Y., Boyacioglu, M. A., &Baykan, Ö. K. (2011).Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange.Expert systems with Applications, 38(5), 5311-5319.
* Kumar, M., &Thenmozhi, M. (2006). Forecasting stock index movement: A comparison of support vector machines and random forest.
* Kumar, M., &Thenmozhi, M. (2006). Forecasting stock index movement: A comparison of support vector machines and random forest.

www.SID.ir

* Kuo, R. J., Chen, C. H., & Hwang, Y. C. (2001). An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. Fuzzy sets and systems, 118(1), 21-45.
* Lin, L., Cao, L., Wang, J., & Zhang, C. (2004). The applications of genetic algorithms in stock market data mining optimisation. In Conference on Data Mining, Text Mining and Their Business Application. Wessex Institute of Technology Press.
* Malkiel, B. G. (2003). The efficient market hypothesis and its critics.The Journal of Economic Perspectives, 17(1), 59-82.
* Nassim, N. T. (2007). The black swan: the impact of the highly improbable. NY: Random House.
* Skabar, A., & Cloete, I. (2002).Neural networks, financial trading and the efficient markets hypothesis.Australian Computer Science Communications, 24(1), 241-249
* Souto-Maior, C. D., Murcia, F. D., Borba, J. A., & da Costa, N. C. (2006).Forecasting IBOVESPA Index with Fuzzy Logic.
* Tang, C. F., & Lean, H. H. (2007). Is the Phillips curve stable for Malaysia? New empirical evidence.Malaysian Journal of Economic Studies, 44(2), 95
* Timmermann, A., & Granger, C. W. (2004).Efficient market hypothesis and forecasting.International Journal of forecasting, 20(1), 15-27.
* Wallace Wanda, A. (2004). Risk assessment by internal auditors using past research on bankruptcy applying bankruptcy models.The IIA Research Foundation, Florida
* Zhang, Z. Y., Shi, C., Zhang, S. L., & Shi, Z. Z. (2006). Stock time series forecasting using support vector machines employing analyst recommendations. Advances in Neural Networks-ISNN 2006, 452-457

Archive of SID

www.SID.ir

پیوست‌ها

پیوست یک

ردیف	نام شرکت	نماد معاملاتی	تعداد نقاط کف	تعداد نقاط سقف	تعداد نقاط نه کف نه سقف	مجموع
۱	ایران خودرو	خودرو	۳۰	۳۰	۲۹	۸۹
۲	پارس خودرو	خپارس	۲۷	۲۷	۲۲	۷۶
۳	سایپا دیزل	خکاوه	۲۷	۲۷	۵۵	۱۰۹
۴	سایپا آذین	خاذین	۴۵	۴۶	۴۳	۱۳۴
۵	زامیاد	خزامیا	۴۳	۴۳	۱۵	۱۰۱
۶	سرمایه‌گذاری رنا	ورنا	۱۹	۱۹	۲۴	۶۲
۷	صنایع ریخته‌گری ایران	خریخت	۱۲	۱۲	۴	۲۸
۸	رینگ سازی مشهد	خرینگ	۳۱	۳۱	۱۸	۸۰
۹	گروه بهمن	خبهمن	۲۵	۲۵	۱۵	۶۵
۱۰	قطعات اتومبیل ایران	ختوقا	۳۳	۳۳	۱۹	۸۵
۱۱	چرخشگر	خچرخش	۱۵	۱۵	۱۰	۴۰
۱۲	گسترش سرمایه گذاری ایران خودرو	خگستر	۲۴	۲۴	۱۴	۶۲
۱۳	ایران خودرو دیزل	خاور	۳۱	۳۱	۳۰	۹۲
۱۴	نیرو محرکه	خمحرکه	۲۰	۲۰	۶۱	۱۰۱
۱۵	مهرکام پارس	خمهر	۱۹	۱۹	۴۲	۸۰
۱۶	آهنگری تراکتور سازی ایران	خاهن	۲۱	۲۲	۱۹	۶۲
۱۷	مهندسی صنعتی روان فن آور	خفناور	۸	۹	۶	۲۳
۱۸	ایرکاپارت صنعت	خکار	۲۲	۲۲	۲۰	۶۴
۱۹	لنت ترمز ایران	خلنت	۱۱	۱۱	۳۱	۵۳
۲۰	محور سازان ایران خودرو	خوساز	۲۵	۲۵	۲۳	۷۳
۲۱	ریخنه گری سازی تراکتور سازی ایران	ختراک	۳۱	۳۱	۱۰	۷۲
۲۲	تولید محور خودرو	خمهور	۱۷	۱۷	۳۳	۶۷
۲۳	فنرسازی زر	خزر	۱۴	۱۴	۱۸	۴۶
۲۴	موتورسازانتراکتورسازی ایران	خموتور	۲۲	۲۲	۳	۴۷
۲۵	الکتریک خودرو شرق	خشرق	۲۱	۲۰	۲۹	۷۰
۲۶	سایپا	خساپا	۳۶	۳۵	۵۰	۱۲۱
مجموع			۶۲۹	۶۳۰	۶۴۳	۱۹۰۲

Archive of SID

www.SID.ir

پیوست دو

نام متغیر	تعریف
قیمت آغازین(PO)	قیمتی است که در شروع معاملات برای هر سهم در تابلو بورس اعلام می‌شود.
بالاترین قیمت(PH)	بیشترین قیمتی که در یک روز معاملاتی سهم با آن معامله می‌شود.
پایین ترین قیمت(PL)	کمترین قیمتی که در یک روز معاملاتی سهم با آن معامله می‌شود.
قیمت پایانی(PC)	قیمتی که در پایان معاملات برای هر سهم در تابلو بورس اعلام می‌شود.
حجم معاملات(Volume)	تعداد سهام معامله شده در یک روز معاملاتی
شاخص جریان پول(MFI)	میزان قدرت پولی را که وارد بازار می‌شود و یا از آن خارج می‌شود اندازه می‌گیرد.
شاخص قدرت نسبی(RSI)	نقاط اشباع خرید(بیش خرید) و اشباع فروش(بیش فروش) را مشخص میکند.
شاخص جهت حرکت مثبت(PDI)	قدرت حرکت قیمت در جهت مثبت را نشان می دهد.
شاخص جهت حرکت منفی(MDI)	قدرت حرکت قیمت در جهت منفی را نشان می دهد.
شاخص کانال کالا(CCI)	تغییرات نرخ بازار را نسبت به میانگین آماری آن بازار می‌سنجد.مقادیر بالا نشان دهنده این هستند که نرخ در مقایسه با نرخ‌های میانگین به‌گونه‌ای غیر عادی بالاست.
میانگین متحرک(MA)	مقدار میانگین نرخ سهام را در یک دوره زمانی به ما می‌دهد.
میانگین متحرک موزون(WMA)	میانگین متحرک موزون را برای دورهای مشخص حساب می‌کند.
میانگین متحرک نمایی(EMA)	یک نوع میانگین متحرک شبیه به میانگین متحرک ساده است .این نوع میانگین وزن بیتری به اطلاعات جدید می‌دهد.
میانگین متحرک نمایی سه گانه(TEMA)	ترکیبی از یک میانگین متحرک نمایی ،یک میانگین متحرک نمایی دو برابر شده و یک میانگین متحرک نمایی سه برابر شده است.
میانگین متحرک نمایی دوگانه(DEMA)	ترکیبی از یک میانگین متحرک نمایی و یک میانگین متحرک نمایی دو برابر شده می‌باشد.
متوسط شاخص جهت حرکت(ADX)	یک شاخص روند است که میزان قدرت روند جاری را اندازه‌گیریمی‌کند.
دامنه متوسط واقعی(ATR)	مقیاسی برای سنجش پویایی بازار است.مقادیر بالای ATR در کف بازار رخ می‌دهد و بالعکس
باند پایینی بولینگر(Bbandbot)	نوار پایینی باند بلوینگر که ترکیبی از میانگین متحرک قیمت با مضروب انحراف معیار است.باند پایین به عنوان نقطه حمایت یا بیش فروش در نظر گرفته می شود.
باند بالایی بلوینگر(Bbandbot top)	نوار بالایی باند بلوینگر که ترکیبی از میانگین متحرک قیمت با مضروب انحراف معیار است.باند بالای به عنوان نقطه مقاومت یا بیش خریددر نظر گرفته می شود.
شاخص کانال کالا باقیمت‌های ابتدایی(CCIa)	همان شاخص کانال کالا است که در محاسبات بجای قیمت پایانی از قیمت آغازین استفاده‌شده است.
چایکین(chaikin)	یک نوسان‌نمای میانگین متحرکی است که بر پایه تجمع/توزیع می‌باشد.
بالاترین سقف در میان سقف‌های قیمتی(HHV)	بالاترین نقطه قیمتی را در میان چندین سقف قیمتی در یک دوره مشخص را نشان می دهد.
دورهای فاصله از(HHV(HHVbars	تعداد روزها یا دورهایی که از HHV گذشته است را محاسبه می‌کند.
میانگین متحرک هال(HMA)	نوعی میانگین متحرک پویاست که مشکل تاخیر در سیگنال دهی سایر میانگین متحرک‌ها را ندارد.
خط رگرسیون خطی (Line Regression)	ابزاری آماری برای پیش‌بینی مقادیرآینده با توجه به مقادیر گذشته می‌باشد.
خط رگرسیون منقطع(Lin Reg Intercept)	خط رگرسیون را در روندها و دورهای کوتاه مدت (مثلا چهارده روزه)محاسبه و مقایسه

Archive of SID

www.SID.ir

نام متغیر	تعریف
	می‌کند.
شیب خط رگرسیون(LinRegSlop)	شیب خط رگرسیون را برای دورهای مشخص محاسبه می‌کند.
پایین ترین کف قیمتی در میان کف‌های قیمتی(LLV)	پایین ترین نقطه قیمتی را در میان چندین کف قیمتی در یک دوره مشخص را نشان می دهد.
دورهای فاصله از LLV(LLV bars)	تعداد روزهایا دورهایی که از LLV گذشته است را محاسبه می‌کند.
واگرایی و هم‌گرایی متحرک(MACD)	وقوع واگرایی و هم‌گرایی اندیکاتور مک دی با نمودار قیمت را مشخص می‌سازد.
نرخ میانه((Median Price	نقطه میانی نرخ هر روز می باشد.
اندیکاتور قیمت(OSCP)	اندیکاتور قیمت را بر مبنای میانگین متحرک نمایی محاسبه می‌کند.
اندیکاتور حجم(OSCV)	اندیکاتور حجم را بر مبنای میانگین متحرک نمایی محاسبه می‌کند.
مرجع گذشته(REF)	در این اندیکاتور ارزش‌های آینده یک ناحیه زمانی به یک ارزش گذشته ارجاع داده می شود.
شاخص حرکت نسبی(RMI)	همان شاخص قدرت نسبی است با این تفاوت که RSIدر این است که RSI قدرت حرکت را تنها نسبت به یک روز گذشته در نظر می‌گیرد و RMI نسبت به چند روز گذشته.
شاخص قدرت نسبی باقیمت‌های ابتدایی(RSIa)	همان شاخص قدرت نسبی است که در محاسباتش از قیمت ابتدایی استفاده‌شده است.
شاخص گام تصادفی(RWI)	اندیکاتوری است که اختلاف بین بالاترین گام تصادفی کشف‌شده و پایین ترین گام تصادفی کشف‌شده را محاسبه می‌کند.
شاخص گام تصادفی بالا(RWHI)	اندیکاتور است که بالترین گام تصادفی کشف‌شده را نشان می دهد.
شاخص گام تصادفی پایین(RWILO)	اندیکاتوری است که پایین ترین گام تصادفی کشف‌شده را نشان می دهد.
توقف و برگشت(SAR)	اندیکاتوری برای تشخیص روند و به دست آوردن نقاط کلیدی اتمام یک‌روند است.
اخطارمکدی(MACDS)	اختلاف دو میانگین متحرک نمایی ۲۶ روزه و ۱۲ روزه می باشد.
خطای استاندارد(StdErr)	میزان اختلاف مقادیر قیمت از خط رگرسیون تخمینی را نشان می دهد.
انحراف معیار(StdEvd)	میزان انحراف از قیمت‌های میانگین برای دورهای ۱۴ روزه را نشان می دهد.
استوکاستیک(D%)	موقعیت‌های اشباع خریدوفروش را بر اساس K% در دوره ۵ روزه تعیین می‌کند
استوکاستیک(K%)	موقعیت‌های اشباع خریدوفروش را در دوره ۵ روزه مشخص می‌کند.
تغییرات میانگین متحرک نمایی(TRIX)	درصد میزان تغییر یک میانگین متحرک نمایی سه بار هموارسازی شده از نرخ بسته شدن بازار را نشان می دهد وحول و حوش صفر نوسان میکند.
نوسان‌نمایفوق‌العاده(Ultimtea)	مجموع وزنی سه نوسان‌نما که هر کدام دوره زمانی متفاوتی دارند را نشان می دهد.
هموار کننده وایلدر(WWS)	این اندیکاتور قیمت‌ها را جهت شناسایی نقاط نزولی و شروع روند صعودی هموار می کند.

Archive of SID

www.SID.ir

یادداشت‌ها

[1] Artificial Neural Network(ANN)
[2] k-nearest neighbors(KNN)
[3] Decision Tree(DT)
[4] random forest(RF)
[5] Bayesian network(BN)
[6] Genetic Algorithm
[7] Efficient market hypothesis
[8] zero sum game
[9] indicators
[10] Oscillator
[11] technical analysis
[12] train
[13] Confusion matrix
[14] Confusion matrix
[15] class precision
[16] Class recall

Archive of SID

www.SID.ir