Archive of SID

نشریه علمی مطالعات مدیریت کسب و کار هوشمند- سال هشتم ـ شماره ۲۹ـ پاییز۱۳۹۸
صفحات ۳۴- ۵

پیش‌بینی شاخص کل بورس اوراق بهادار تهران با رویکرد سیستم استنتاج عصبی– فازی انطباق پذیر و الگوریتم رقابت استعماری

مجتبی صالحی *
فاطمه گرشاسبی **

چکیده

در این پژوهش یک روش نوین ترکیبی برای پیش‌بینی شاخص کل بورس اوراق بهادار تهران ارائه شده است که هم‌زمان از الگوریتم رقابت استعماری به‌عنوان روش انتخاب ویژگی و شبکه فازی عصبی انطباق‌پذیر به‌عنوان تابع پیش‌بینی کننده استفاده می‌نماید. برای انجام این امر از ۶۸ ویژگی مؤثر بر بازار بورس اوراق بهادار؛ شامل شاخص‌های اقتصادی، شاخص‌های بورس ایران و سایر کشورها، شاخص‌های تحلیل فنی و شاخص‌های شمعدان ژاپنی به‌صورت روزانه در بازه زمانی ۱۳۸۹-۱۳۹۵ به‌عنوان ورودی مدل استفاده شده است. همچنین، شاخص کل بورس اوراق بهادار روز آتی به‌عنوان متغیر هدف مسأله در نظر گرفته شده است. نتایج به‌دست‌آمده نشان می‌دهد که مدل ترکیبی شبکه عصبی- فازی انطباق‌پذیر و الگوریتم رقابت استعماری پیش‌بینی‌های بسیار مناسب‌تری داشته و به نسبت شبکه‌های عصبی منفرد از سرعت بالاتر و توانایی تقریب قوی‌تری برای پیش‌بینی شاخص کل بورس اوراق بهادار برخوردار بوده است.

کلیدواژگان: پیش‌بینی بازار بورس، الگوریتم رقابت استعماری، شبکه فازی عصبی انطباق پذیر، انتخاب ویژگی، سری‌های زمانی

---

* عضو هیأت علمی، گروه مهندسی صنایع، دانشکده فنی و مهندسی، دانشگاه پیام نور، تهران.(نویسنده مسئول)؛
m.salehi61@chmail.ir
** کارشناسی ارشد، مهندسی صنایع، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیر

تاریخ دریافت:۱۳۹٦/۰٤/۰۹ تاریخ پذیرش:۱۳۹۸/۰٦/۱۱

www.SID.ir

Archive of SID

مقدمه

امروزه دستیابی به رشد بلندمدت و مداوم اقتصادی نیازمند تجهیز و تخصیص بهینه منابع در سطح اقتصاد ملی است و این مهم بدون کمک بازارهای مالی؛ به‌ویژه بازار سرمایه گسترده و کارآمد، به سهولت امکان‌پذیر نیست (منجمی و همکاران، ۱۳۸۸). از مهم‌ترین عوامل مؤثر در افزایش سرمایه‌گذاری، رشد و توسعه اقتصادی، داشتن بازارهای مالی کارآمد همراه با سازمانهای مالی مناسب و فعال در این بازارهاست (بن ناصر و غازیونی[1]، ۲۰۰۷). در این راستا فعالیت‌هایی چون تأمین نقدشوندگی سهام، افزایش شفافیت اطلاعات، امکان جمع‌کردن و سرمایه‌های کوچک اهمیت ویژه‌ای را دارا هستند که این وظیفه مهم به سازمان بورس اوراق بهادار واگذار شده است. تحقق هدف فوق توسط بورس اوراق بهادار، منوط به ایجاد امنیت سرمایه‌گذاری برای سرمایه‌گذاران است؛ زیرا نگرانی سرمایه‌گذاران تا حد زیادی ناشی از نوسان‌های شدید شاخص بورس اوراق بهادار و به تبع آن، سهام موجود در بازار است. سرمایه‌گذاران با پیش‌بینی تغییرات شاخص کل سهام و قیمت سهام و ارزیابی سهام خود و دیگر سهام موجود در بازار می‌توانند بهترین سهام را انتخاب و درنتیجه اقدام به نگهداری، فروش و یا بعضاً جایگزینی سهام خود با سهام دیگری نمایند (اسدی،۲۰۱۲). در این راستا، سرمایه‌گذار نیازمند ابزارهای قدرتمند و قابل‌اعتماد است تا از طریق آن به پیش‌بینی شاخص سهام و قیمت سهام بپردازد(ساندرا رامناش و همکاران[2]). وجود الگو در تغییر قیمت‌ها در بازار سرمایه و کشف روابط بین قیمت و سایر متغیرها با استفاده از مدل‌های مناسب، یکی از دغدغه‌های بازار سرمایه است. از طرفی، مدل‌های خطی موجود در یک محیط پویای غیرخطی پرآشوب چندان جوابگو نخواهد بود. پس لازم است از تکنیک‌ها و ابزارهایی استفاده شود که توانایی پاسخگویی شرایط مذکور را داشته باشد (حیدری زارع و کردلوئی، ۱۳۸۹).

در سال‌های اخیر، مدل‌های ساختاری و سنتی که در تبیین وضع موجود به‌طور نسبی موفق بوده‌اند، سابقه چندان موفقی در زمینه پیش‌بینی نداشته‌اندو همین امر موجب توجه

1. Ben Naser and Ghazuani
2. Sundaresh Ramnath et al.

www.SID.ir

Archive of SID

اقتصاددانان و سایر پژوهشگران به روش‌های نوین از قبیل مدل‌های شبکه عصبی[1] و منطق فازی[2] شده است. این توجه به دلیل نتایج دقیق‌تر و صحیح‌تر این مدل‌ها در قیاس با سایر مدل‌هاست. شبکه‌های عصبی برخلاف روش‌های سنتی، این قابلیت را دارند که به‌آسانی با اطلاعات موجود تطابق پیدا کنند و بدون اینکه نیازی به پیش‌شرط خاصی باشد، می‌توانند از روی یک سری اطلاعات آموزش ببینند و رابطه مناسب بین اطلاعات برقرار نمایند؛ بنابراین بخصوص در مسائلی که ایجاد ارتباط بین اطلاعات موجود مشکل است، شبکه‌های عصبی بسیار کارا هستند. هرچند استفاده از مدل‌های شبکه عصبی و منطق فازی نتایج خوبی را در پیش‌بینی نشان داده‌اند ولیکن دستیابی به نتایج دقیق در به‌کارگیری این مدل‌های نوین، مستلزم داشتن اطلاعات دقیق از عوامل تأثیرگذار بر کمیت مورد بررسی است.

با توجه به اینکه عوامل مؤثر با بازارهای مالی، به‌طور خاص بازار بورس اوراق بهادار دارای گستردگی بالایی است و جمع‌آوری و مدل کردن این عوامل، کاری زمان‌بر و سخت است، انتخاب مهم‌ترین عوامل مؤثر بر بازار امری اجتناب‌ناپذیر خواهد بود. در بیشتر موارد بسیاری از ویژگی‌های کاندید در فرایند یادگیری نامربوط و زاید هستند و چنانچه عملیات انتخاب ویژگی صورت نگیرد، ممکن است دقت الگوریتم یادگیری و سرعت آموزش به میزان قابل‌توجهی کاهش یابد؛ بنابراین، انتخاب ویژگی‌های مرتبط و ضروری در مرحله پیش‌پردازش از اهمیتی بنیادین برخوردار است.

هدف اصلی این تحقیق، پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه فازی- عصبی انطباق پذیر[3] به‌عنوان مدل پیش‌بینی کننده و الگوریتم رقابت استعماری برای انتخاب ویژگی‌های بهینه است. در این تحقیق، از ۶۸ ویژگی مؤثر بر بازار بورس؛ شامل متغیرهای اقتصادی، شاخص‌های بازار بورس اوراق بهادار، شاخص‌های بازار بورس‌های بین‌المللی، شاخص‌های تحلیل فنی و شاخص‌های شمعدان ژاپنی استفاده شده است.

1. Neural Network
2. Fuzzy Logic
3. ANFIS (Adaptive Neural Fuzzy Inference System)

www.SID.ir

Archive of SID

مبانی نظری و پیشینه پژوهش

در این بخش ابتدا مبانی نظری تحقیق بررسی و سپس ادبیات حوزه تحقیق مرور شده است.

شاخص کل بازار بورس اوراق بهادار

شاخص قیمت سهام بورس اوراق بهادار، شاخصی از نوع میانگین حسابی با وزن‌هایی برابر ارزش بازار سهام شرکت‌هاست و با عنوان تپیکس[1] شناخته می‌شود؛ شامل تمامی شرکت‌های پذیرفته شده در بورس است و درصورتی‌که نماد شرکتی بسته باشد یا برای مدتی معامله نشود، قیمت آخرین معامله آن در شاخص لحاظ می‌گردد. همان‌گونه که از فرمول (۱) مشخص است، تعداد سهام منتشرشده شرکت‌ها، معیار وزن دهی در شاخص مذبور است. این امر به تأثیر بیشتر شرکت‌های بزرگ در شاخص منجر می‌شود. این شاخص به‌طور روزانه محاسبه و اعلام می‌شود (سینائی و همکاران، ۱۳۸۴).

$$TEPIX_t = \frac{\sum_{i=1}^{n} P_{it} Q_{it}}{D_t} \times 100 \qquad (۱)$$

P_{it}= قیمت شرکت i ام در زمان t

Q_{it}= تعداد سهام منتشرشده شرکت i ام در زمان t

D_t= عدد پایه در زمان t که در زمان مبدأ برابر $\sum q_{i0}p_{i0}$ است.

p_{i0}= قیمت شرکت i ام در زمان مبدأ

q_{i0}= تعداد سهام منتشرشده شرکت i ام در زمان مبدأ

n = تعداد شرکت‌های مشمول شاخص

این شاخص، نمایانگر تغییرات کل بازار و همچنین، نشان‌دهنده وضعیت کلی اقتصاد کشور است. افزایش این شاخص به معنی رونق و بهبودی در اوضاع و احوال اقتصادی و کاهش آن، گویای بحران و رکود است (هرمزی، ۱۳۸۸). بورس اوراق بهادار به معنی یک بازار متشکل و رسمی سرمایه است که در آن خریدوفروش سهام شرکت‌ها و اوراق مشارکتی

1. TEPIX

www.SID.ir

Archive of SID

دولتی و مؤسسات معتبر خصوصی، تحت ضوابط و مقررات خاصی انجام می‌شود و سابقه آن در کشور ما به اواخر دهه ۱۳۴۰ می‌رسد.

شاخص قیمت سهام، شاخصی است که بیان‌کننده سطح عمومی قیمت سهام شرکت‌های پذیرفته شده در بورس است. در بازارهای مالی قابل‌مبادله، محور تصمیم‌گیری عاملان بازار، قیمت سهام است. سرمایه‌گذاران در این بازار به اخبار آخرین وضعیت قیمت‌ها و تحلیل‌های پیش‌بینی نوسان های آتی قیمت چشم دوخته‌اندو چون تحلیل قیمت‌ها برای همگان با سهولت بیشتری انجام می‌پذیرد، استفاده از قیمت سهام متداول‌تر است (هیبتی و موسوی، ۱۳۸۷).

انتخاب ویژگی

مسأله انتخاب ویژگی، یکی از مسائلی است که در مبحث یادگیری ماشین و در بررسی مسائل شناخت الگو از دیدگاه آمار مطرح است. کاهش تعداد متغیرها می‌تواند به کاهش خطای جداسازی منجر شود؛ چون با زیاد شدن تعداد متغیرها، تعداد پارامترهای جداساز که باید تخمین زده شوند و در نتیجه، پیچیدگی جداساز افزایش می‌یابد. همچنین افزایش تعداد پارامترها می‌تواند به قابلیت تعمیم دهی ضعیف منجر شود (صالحی و کردکتولی، ۱۳۹۶). در بیشتر موارد، بسیاری از ویژگی‌های کاندید در فرایند یادگیری نامربوط و زاید هستند و چنانچه عملیات انتخاب ویژگی صورت نگیرد، ممکن است دقت الگوریتم یادگیری و سرعت آموزش به میزان قابل‌توجهی کاهش یابد؛ بنابراین انتخاب ویژگی‌های مرتبط و ضروری در مرحله پیش‌پردازش از اهمیتی بنیادین برخوردار است. موضوع اصلی در ساخت الگوریتم‌های انتخاب ویژگی، ارزیابی کیفیت ویژگی‌های کاندید است (هو و همکاران[1]، ۲۰۰۸). روش‌های انتخاب ویژگی با دو هدف تعریف می‌شوند:

هدف اول: به حداکثر رساندن دقت روش؛

هدف دوم: به حداقل رساندن تعداد ویژگی‌هایی که استفاده می‌شود (وریا و همکاران[2]، ۲۰۱۲).

1. Hu et al.
2. Vieira et al.

www.SID.ir

Archive of SID

الگوریتم رقابت استعماری

الگوریتم رقابت استعماری[1] برای نخستین بار در سال ۲۰۰۷ توسط آتش‌پز گرگری و لوکاس[2] با هدف بهینه‌سازی و با رویکرد نوین مدل‌سازی فرایندهای سیاسی- اجتماعی مطرح شد. الگوریتم جدید معرفی شده با الهام‌گیری از یک فرایند اجتماعی- سیاسی، نسبت به روش‌های مطرح‌شده دارای توانایی بالایی بوده و تا حد بسیار زیادی نیز سریع است.

همانند دیگر الگوریتم‌های تکاملی، این الگوریتم نیز با تعدادی جمعیت اولیه تصادفی که هرکدام از آن‌ها یک «کشور» نامیده می‌شوند، شروع می‌شود. تعدادی از بهترین عناصر جمعیت (دارای کمترین هزینه) به‌عنوان امپریالیست انتخاب و باقی‌مانده جمعیت نیز به‌عنوان مستعمره[3] در نظر گرفته می‌شوند.

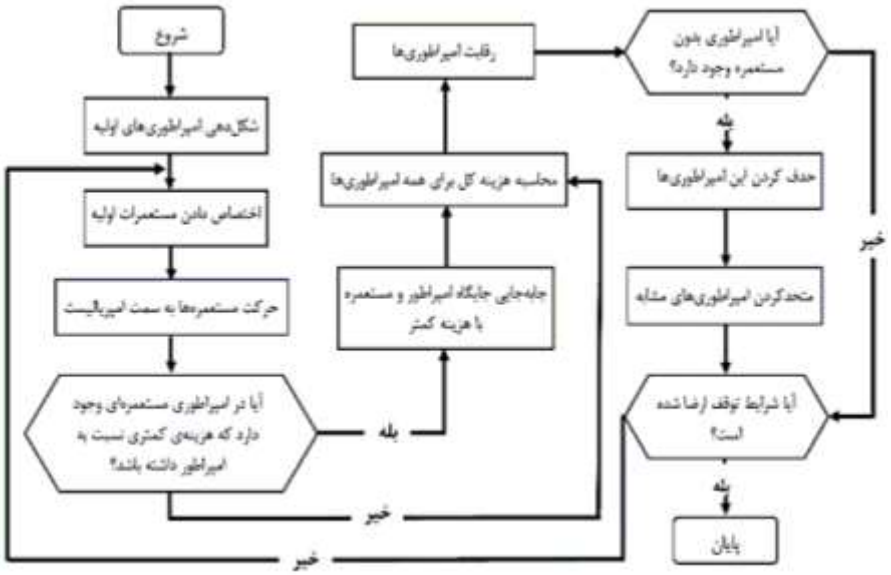
استعمارگران بسته به قدرتشان، این مستعمرات را با روندی خاص، به سمت خود می‌کشند. قدرت کل هر امپراتوری، به هر دو بخش تشکیل‌دهنده آن؛ یعنی کشور امپریالیست (به‌عنوان هسته مرکزی) و مستعمرات آن بستگی دارد. در حالت ریاضی، این وابستگی با تعریف قدرت امپراتوری به‌صورت مجموع قدرت کشور امپریالیست، با اضافه درصدی از میانگین قدرت مستعمرات آن، مدل شده است. با شکل‌گیری امپراتوری‌های اولیه، رقابت امپریالیستی میان آن‌ها شروع می‌شود. هر امپراتوری که نتواند در رقابت استعماری، بر قدرت خود بیفزاید (حداقل از کاهش نفوذش جلوگیری کند)، از صحنه رقابت استعماری حذف خواهد شد؛ بنابراین، بقای یک امپراتوری، وابسته به قدرت آن در جذب مستعمرات امپراتوری‌های رقیب و به سیطره درآوردن آن‌ها خواهد بود. در نتیجه، در جریان رقابت‌های امپریالیستی، به‌تدریج بر قدرت امپراتوری‌های بزرگ‌تر افزوده شده و امپراتوری‌های ضعیف‌تر، حذف خواهند شد. امپراتوری‌ها برای افزایش قدرت خود، مجبور خواهند شد تا مستعمرات خود را نیز پیشرفت دهند. با گذشت زمان، مستعمرات، از لحاظ قدرت به امپراتوری‌ها نزدیک‌تر خواهند شد و شاهد یک نوع همگرایی خواهیم بود. حد نهایی رقابت

1. Imperialist Competitive Algorithm
2. Atashpaz-Gargari and Lucas
3. Colony

www.SID.ir

Archive of SID

استعماری، زمانی است که یک امپراتوری واحد در دنیا داشته باشیم با مستعمراتی که از لحاظ موقعیت، به خود کشور امپریالیست خیلی نزدیک هستند. درنهایت، پاسخ بهینه الگوریتم امپریالیست، امپراتوری باقی‌مانده خواهد بود. شکل (۲) ساختار الگوریتم رقابت استعماری را نشان می‌دهد.



شکل ۲: ساختار الگوریتم رقابت استعماری

شبکه- فازی عصبی انطباق‌پذیر

پروفسور جانگ[1] در سال ۱۹۹۳ سیستم استدلال عصبی- فازی تطبیقی را معرفی نمود که برای اجرا از سیستم فازی تاکاگی سوگنو در ساختار شبکه عصبی استفاده می‌کند تا نگاشتی غیرخطی بین فضای ورودی و خروجی ایجاد نماید. در این نوع از سیستم‌های استنتاجی، از شبکه‌های عصبی برای جبران نقاط ضعف منطق فازی بهره می‌برند. بزرگ‌ترین مزیت آن‌ها این است که می‌توانند از قابلیت یادگیری شبکه‌های عصبی استفاده و از زمان‌بری هزینه‌زای

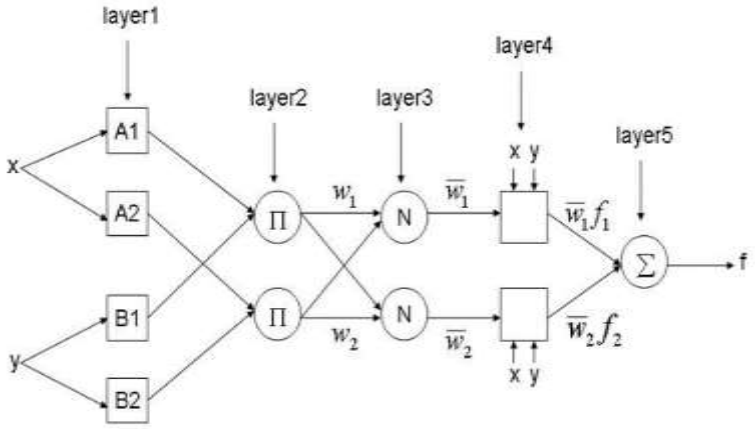
1. Jang

www.SID.ir

Archive of SID

تنظیم قوانین یک موتور استنتاج در سیستم منطق فازی مرسوم جلوگیری کنند. به‌طور عملی، هیچ محدودیتی در توابع گره‌های شبکه‌های تطبیقی وجود ندارد؛ مگر آنکه تنها نیاز است آن‌ها مشتق‌پذیر باشند. تنها محدودیت ساختاری شکل شبکه‌ها، نوع پیش‌خور بودن آن‌هاست و در برابر این محدودیت کوچک، کاربرد شبکه‌های تطبیقی به‌صورت روزافزون و وسیعی در حیطه‌های کاربردی گسترش یافته است (جانگ، ۱۹۹۳).

این سیستم به این صورت عمل می‌کند: در هر دور آموزش، هنگام حرکت روبه‌جلو خروجی‌های نورون‌ها به‌صورت عادی تا لایه آخر محاسبه می‌شوند و سپس پارامترهای نتیجه توسط روش کمترین مربعات خطا محاسبه می‌شوند. در ادامه، پس از محاسبه خطا در بازگشت رو به عقب، نسبت خطا بر روی پارامترهای شرط پخش شده و با استفاده از روش شیب نزولی خطا، مقدار آن‌ها تصحیح می‌شود. معماری شبکه ANFIS در شکل (۳) نشان داده شده است.



شکل ۳: معماری شبکه عصبی- فازی تطبیقی

مروری بر ادبیات پیشین

پیمنتا و همکاران [1] (۲۰۱۴)، از ۱۲ شاخص تحلیل فنی برای پیش‌بینی توسط درخت‌های

1. Pimenta et al.

www.SID.ir

Archive of SID

تصمیم استفاده کردند. آن‌ها با استفاده از برنامه‌نویسی ژنتیکی[1] و شاخص‌های تحلیل فنی بر مبنای داده‌های آموزش، سعی در تولید و انتخاب بهترین درخت‌های تصمیم داشتند. در این مدل هر نسل از کروموزوم‌ها در الگوریتم ژنتیک؛ شامل درخت‌های تصمیم خریدوفروش بر اساس میزان شاخص‌های فنی مختلف است و در انتها یک جفت درخت تصمیم خریدوفروش به‌عنوان بهترین راهکار برای یک شاخص انتخاب می‌شوند. این دو درخت تصمیم برای تصمیم‌سازی خریدوفروش روی داده‌های آزمون استفاده می‌شوند. آن‌ها نشان دادند این روش از روش خرید و نگهداری سهام بهتر عمل می‌کند.

لینسی[2] (۲۰۱۶)، با استفاده از روش یک سیستم استنتاج فازی چندگانه به پیش‌بینی روند روزانه سهام می‌پردازد. او از ۱۲ شاخص فنی به‌عنوان ورودی مدل استفاده کرده است. نوآوری مدل استفاده از روش‌های پیش‌پردازش داده‌ها، قبل از استفاده از روش استنتاج فازی قلمداد می‌شود. روند شاخص برای سه تئوری وضعیت خرید، فروش و نگه‌داشتن سهام بررسی و تصمیم‌گیری می‌شود. لینسی برای آزمون مدل به پیش‌بینی سهام ۲۵ شرکت موجود در بازار سهام آمریکا و بخش سهام نزدک پرداخته است.

گوچن[3] (۲۰۱۶)، مدل ترکیبی از الگوریتم‌های فرا ابتکاری و شبکه عصبی مصنوعی برای پیش‌بینی شاخص کل بازار سهام ترکیه ارائه کرده است. ورودی مدل ۴۵ شاخص فنی بر روی قیمت بازگشایی در دوره زمانی ۲۰۰۵-۲۰۱۳ است. او از دو الگوریتم فرا ابتکاری ژنتیک و جستجوی هارمونی[4] برای بهینه‌سازی شبکه عصبی استفاده کرده است. نتایج به‌دست‌آمده نشان می‌دهد که ترکیب شبکه عصبی چندلایه پیش‌خور با الگوریتم فرا ابتکاری جستجوی هارمونی دقت بالاتری در پیش‌بینی دارد.

کوون[5] (۲۰۱۶)، از ۱۷ شاخص تحلیل فنی به‌عنوان ورودی مدل شبکه عصبی خود استفاده کرده و برای بهبود عملکرد شبکه عصبی از الگوریتم ژنتیک برای بهبود وزن یال‌ها کمک

1. Genetic Programming
2. Lincy
3. Goçken
4. Harmony Search Algorithm
5. kwon

www.SID.ir

Archive of SID

گرفته است. نتایج نشان داد که این مدل از روش خرید و نگهداری سهام نسبت به مدل‌های مشابه بهتر عمل کرده و دقت بالاتری دارد.

سودهاکار[1] (۲۰۱۶) با استفاده از شبکه عصبی مصنوعی به پیش‌بینی شاخص سهام نزدک می‌پردازد. برای بهبود بخشیدن به روند پیش‌بینی از روش داده‌کاوی برای پیش‌پردازش داده‌ها استفاده شده است. داده‌های ورودی، ۸ داده مربوط به شاخص‌های بورس و ۵ شاخص کلان اقتصادی است. سودهاکار با استفاده از روش داده‌کاوی با الگوریتم رگرسیون به انتخاب ویژگی پرداخته و با استفاده از شبکه عصبی پس انتشار خطا مقدار شاخص بورس را پیش‌بینی می‌کند. با استفاده از این روش، ۷ ویژگی بهینه برای پیش‌بینی انتخاب شده که نتایج نشان‌دهنده بهبود در میزان دقت و عملکرد مدل است.

کیو و همکاران[2] (۲۰۱۶)، با استفاده از شبکه عصبی به پیش‌بینی شاخص کل بازار بورس ژاپن پرداختند. آن‌ها از ۱۸ شاخص فنی به همراه چند شاخص دیگر به‌عنوان ورودی شبکه عصبی تحلیل بنیادین کرده‌اند. این ۱۸ شاخص به‌وسیله روش سطوح فازی[3] از بین ۷۱ شاخص در نظر گرفته شده انتخاب شده‌اند. هدف از انجام این تحقیق، نشان دادن عملکرد مناسب الگوریتم ژنتیک و شبیه‌سازی تبرید[4] برای پرهیز از کمینه‌های محلی در آموزش شبکه عصبی و کارکرد مناسب ورودی‌های انتخاب‌شده (شاخص‌های فنی و بنیادین) برای پیش‌بینی بود.

پرما[5] و همکاران (۲۰۱۶)، با استفاده از شبکه عصبی مصنوعی به پیش‌بینی شاخص بورس آلمان پرداختند. آنها در این پژوهش برای بهبود عملکرد و پیش‌بینی شبکه عصبی از الگوریتم ژنتیک برای بهینه‌سازی وزن‌های شبکه عصبی کمک گرفتند (در حالت معمولی با روش انتشار به عقب انجام می‌شود). آنها مدل خود را با معیار میانگین نرمال شده مربعات خطا با روش‌های آماری مقایسه کردند و نتایج نیز برتری هوش مصنوعی را نشان می‌داد.

با توجه به مقالات ارائه شده و سایر مقالات مورد بررسی در این پژوهش، می‌توان

1. Sudhakar
2. Qiu
3. Fuzzy Surfaces
4. Simulated Annealing
5. Prema

www.SID.ir

Archive of SID

روش‌های پیش‌بینی در حوزه بازار بورس را به چهار دسته: ۱- روش‌های اقتصادسنجی؛ ۲- روش‌های دینامیک؛ ۳- روش‌های هوش مصنوعی و ۴- روش‌های داده‌کاوی تقسیم‌بندی نمود. پژوهش‌های انجام‌شده در سال‌های اخیر توجه بسیار زیادی به روش‌های هوش مصنوعی داشته‌اند که نشان‌دهنده توانایی بالای این دسته از الگوریتم‌ها در پیش‌بینی بازار شاخص‌های بازار بورس اوراق بهادار و به‌طورکلی، پیش‌بینی روندهای غیرخطی و آشوبناک است. همچنین، روش‌های بهینه برای روش‌های پیش‌بینی هوش مصنوعی به دو دسته: ۱- روش‌های بهینه کننده پارامترهای مدل و وزن‌های الگوریتم و ۲- روش‌های انتخاب ویژگی و پیش‌پردازش داده‌ها تقسیم‌بندی می‌شوند. روش‌های دسته اول سال‌های بسیاری است که استفاده شده است؛ اما روش‌های انتخاب ویژگی اخیراً وارد ادبیات تحقیق این حوزه شده است. همان‌طور که در فصل قبل عنوان شد، با توجه به اینکه بازارهای پولی و مالی در تعامل با یکدیگر و با دیگر متغیرهای داخلی کشور و بین‌المللی قرار دارند و نظر به اهمیت انتخاب ویژگی‌های ورودی در پیش‌بینی و همچنین، زمان‌بر و سخت بودن جمع‌آوری تمامی عوامل مؤثر بر بازار بورس برای هر بار پیش‌بینی؛ در این پژوهش روش‌های انتخاب ویژگی استفاده شده‌اند.

بنابراین، در این پژوهش، روش‌های هوش مصنوعی برای پیش‌بینی شاخص کل بازار بورس اوراق بهادار و روش‌های انتخاب ویژگی برای بهبود دقت پیش‌بینی استفاده شده است.

روش‌شناسی پژوهش

این مطالعه مدلی برای پیش‌بینی مبتنی بر ANFIS با الگوریتم انتخاب ویژگی بهینه بوده، که بر مبنای استفاده از متدولوژی رپر طراحی شده است. در این مدل، مجموعه نتایج خروجی از ANFIS به‌عنوان تابع هزینه در الگوریتم رقابت استعماری عمل می‌کند؛ بدین‌صورت که الگوریتم رقابت استعماری در فضای N بعدی مسأله، ، شبکه عصبی را در هر تکرار با تعدادی از راه‌حل‌ها آموزش می‌دهد. الگوریتم رقابت استعماری با ایجاد تعداد $N_{country}$ کشور اولیه شروع شده و N_{imp} تا از بهترین اعضای این جمعیت (کشورهایی با کمترین

www.SID.ir

Archive of SID

مقدار تابع هزینه) به‌عنوان استعمارگر انتخاب می‌شوند. هر کشور نشان‌دهنده یک راه‌حل است که تعدادی از ویژگی‌ها را در بر دارد. کشورهای استعمارگر با اعمال سیاست جذب (همگون سازی) در راستای محورهای مختلف بهینه‌سازی، کشورهای مستعمره را به سمت خود می‌کشند. راه‌حل بهینه، راه‌حلی با بهترین مقدار از تابع هزینه در همه تکرارها خواهد بود. این جواب، همان ویژگی‌های انتخاب‌شده از بین ویژگی‌های ورودی مدل است. خطای مقدار پیش‌بینی شبکه عصبی فازی ه عنوان تابع هدف مدل استفاده شده است. شکل شماره (۴) ساختار کلی مدل پیشنهادی را نمایش می‌دهد.

متغیرهای زیر در مدل‌سازی مورد استفاده قرار گرفته‌اند؛

β: نرخ جذب

ξ: درصد تأثیر هزینه مستعمره‌ها بر هزینه امپراتوری

N_{pop}: تعداد کشورهای اولیه (جمعیت اولیه)

N_{imp}: تعداد امپراتوری‌ها

N_{col}: تعداد مستعمره‌ها $(N_{col} = N_{pop} - N_{imp})$

NF: تعداد ویژگی‌های بهینه انتخاب‌شده توسط الگوریتم

$Maxit$: حداکثر تعداد تکرارها

$P_{revolve}$: احتمال انقلاب

μ: نرخ انقلاب

N_{var}: تعداد متغیرهای ورودی

$Imp.Cost(i)$: هزینه امپراتوری i ام

$Imp.Col.Cost(i)$: هزینه مستعمره‌های امپراتوری i ام

$\propto$: فشار انتخاب

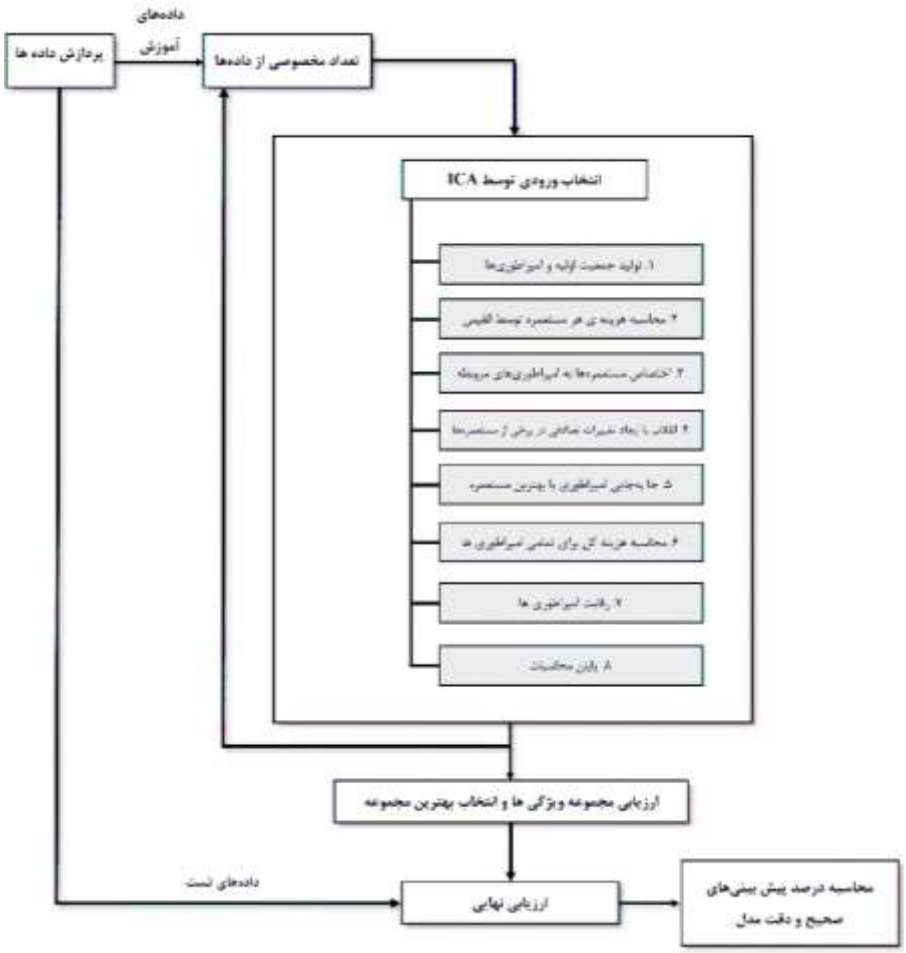
$Epoach_n$: تعداد دوره‌ها

www.SID.ir

Archive of SID

$Num\ MFs$: تعداد توابع عضویت برای FCM

P_Train: درصد داده‌های انتخابی برای آزمون و آموزش



شکل ٤: ساختار مدل پیشنهادی

شکل‌گیری امپراتوری‌های اولیه

الگوریتم مورد استفاده برای ایجاد کشورهای اولیه، ابتدا یک ماتریس صفر به ابعاد $N_{var} \times NF$ ایجاد می‌کند که این کشورها، پاسخ‌های اولیه برای مسأله می‌باشند. در

www.SID.ir

Archive of SID

زمان ایجاد کشورهای اولیه، به‌صورت تصادفی به تعداد NF از ستون‌های هر سطر مقدار یک اختصاص می‌دهد؛ به عبارتی هر کشور دارای NF ویژگی خواهد بود.

هر کشور برای ارزیابی و برآورد هزینه خود، به‌صورت جداگانه به شبکه ANFIS وارد می‌شود. در ادامه با استفاده از داده‌های آموزش ، شبکه را آموزش داده و پس از آن به برآورد هزینه کشور در قالب معیار مجذور میانگین مربعات خطا با استفاده از داده‌های آزمون می-پردازد. ANFIS از تابع FCM[1] به منظور اعمال قوانین و تعداد توابع عضویت، بهره می‌گیرد. همچنین تابع FCM به ترتیب از تابع گوسین و تابع خطی برای توابع عضویت ورودی‌های سیستم و خروجی آن استفاده می‌کند. به همین صورت، تمام هزینه‌های کشورهای اولیه محاسبه می‌شود. تعداد N_{imp} از کشورهایی با کمترین هزینه به‌عنوان امپراتورهای (استعمارگرها) مدل انتخاب می‌شوند. بعد از انتخاب امپراتورها، مابقی کشورهای اولیه به‌عنوان مستعمره در نظر گرفته می‌شوند. به مظور تقسیم مستعمرات اولیه بین استعمارگرها، به هر استعمارگر بر اساس قدرت آن تعدادی از مستعمرات اختصاص داده می‌شود. برای تعریف احتمال انتخاب شدن هر امپراتوری، از روش بولتزمان[2] استفاده شده است. به این صورت که احتمال هر استعمارگر متناسب با $e^{-\alpha c_j}$ بوده که در آن، c_j هزینه‌امپراتوری j ام و α به عنوان پارامتر فشار انتخاب[3] تعریف می‌شود. بدین ترتیب، احتمال هر استعمارگر با استفاده از معادله (۲) محاسبه می‌شود.

$$P_i = \frac{e^{-\alpha c_i}}{\sum_j e^{-\alpha c_j}} \qquad j = 1, \dots, N_{imp} \tag{۲}$$

مقداری از پارامتر فشار انتخاب مطلوب است که مجموع احتمالات نصف بهتر جمعیت در حد ۰/۸ باشد. احتمال‌های ایجادشده برای اختصاص مستعمره‌ها به استعمارگرها، بر اساس روش نمونه‌برداری رولت- ویل[4] محاسبه شده است.

1. Fuzzy C–Means Clustering Algorithm
2. Bultzmann
3. Selection Pressure
4. Roulette wheel selection

www.SID.ir

Archive of SID

سیاست جذب

سیاست جذب، اصلی‌ترین بخش الگوریتم بوده و مهم‌ترین حلقه را ایجاد می‌کند. در این مرحله از الگوریتم، هر مستعمره باید به سمت استعمارگر خود یا به عبارتی به سمت نقطه بهینگی در امپراتوری خود حرکت کند. بدین منظور، الگوریتم در هر امپراتوری فاصله مستعمره نسبت به وضعیت بهترین پاسخ (استعمارگر) را در نرخ β (نرخ جذب) ضرب کرده و با وضعیت مستعمره جمع می‌کند. نرخ β مقداری بین ۱ و ۲ بوده و مقدار بهینه آن با استفاده از روش طراحی آزمایش‌های تاگوچی تعیین می‌شود. همچنین برداری تصادفی با عنوان $\vec{r}$ ایجاد شده که باعث می‌شود انحرافی[1] از حدفاصل بین مستعمره و استعمارگر ایجاد شده و پاسخ‌های بیشتری مورد بررسی قرار گیرد.

$$x' = x + \beta * \vec{r} * (T - x) \qquad (۳)$$

در رابطه بالا، x وضعیت مستعمره و T وضعیت استعمارگر را نشان می‌دهد.

انقلاب

تغییر تصادفی در پاسخ‌ها باعث می‌شود که بتوانیم پاسخ‌های متفاوتی را ایجاد کرده و از بین آن ها پاسخ‌های بهتری را انتخاب کنیم. باید توجه داشت فقط انقلابی در مستعمره‌ها پذیرفته می‌شود که موقعیت بهتری برای مستعمره ایجاد کند.

به ازای هر مستعمره، ابتدا مقداری تصادفی بین صفر و یک تولید می‌کنیم؛ درصورتی‌که این عدد کمتر یا مساوی احتمال مقدار پارامتر احتمال انقلاب باشد، به تعداد μ از مؤلفه‌های وضعیت آن به‌اندازه ۰/۱ دچار انقلاب می‌گردد. بعد از ایجاد مستعمره جدید، هزینه مستعمره محاسبه شده و در صورت بهبود (کاهش) انقلاب پذیرفته می‌شود؛ در غیر این صورت، وضعیت جدید مستعمره رد خواهد شد.

1. Deviation

www.SID.ir

Archive of SID

جابجایی موقعیت مستعمره و امپریالیست

پس از اعمال سیاست جذب و انقلاب، باید به ازای هر امپراتوری با استفاده از بازخوانی تابع هزینه(تابع ANFIS)، هزینه تمام استعمارگرها و مستعمرات آن‌ها محاسبه گردد. درصورتی‌که در هر امپراتوری مستعمره‌ای وجود داشته باشد که هزینه آن مستعمره از استعمارگر کمتر باشد، استعمارگر سقوط کرده و جایگاه استعمارگر و مستعمره عوض می‌شود و الگوریتم با کشور استعمارگر در موقعیت جدید ادامه خواهد یافت.

محاسبه قدرت کل یک امپراتوری

قدرت یک امپراتوری برابر با قدرت کشور استعمارگر، بعلاوه درصدی از قدرت کل مستعمرات آن خواهد بود. بدین ترتیب برای هزینه کل یک امپراتوری داریم:

$$T.C._n = Cost(imperialist_n) + \xi * mean(colonies\ of\ empire_n) \quad (۴)$$

که در آن $T.C._n$ هزینه کل امپراتوری n ام است و ξ عددی مثبت بوده که معمولاً بین صفر و یک است. در این پژوهش مقدار بهینه ξ با استفاده از روش طراحی آزمایش‌های تاگوچی محاسبه می‌شود.

رقابت استعماری

در این مرحله، ابتدا تمامی امپراتورها بر اساس هزینه مرتب شده و ضعیف‌ترین امپراتوری انتخاب می‌شود. درصورتی‌که این امپراتوری دارای مستعمره باشد ضعیف‌ترین مستعمره انتخاب شده و با استفاده از روش بولتزمان و روش نمونه‌برداری رولت–ویل که در بخش ایجاد کشور اولیه توضیح داده شد، به امپراتورهای دیگر منتقل می‌شود. حذف مستعمرات تا زمانی که هیچ مستعمره‌ای در امپراتوری ضعیف وجود نداشته باشد، ادامه می‌یابد. سپس، استعمارگر ضعیف‌ترین امپراتوری نیز به‌عنوان مستعمره به امپراتورهای قوی‌تر منتقل می‌شود.

www.SID.ir

Archive of SID

خاتمه

در این مدل شرایط خاتمه، رسیدن به حداکثر تکرارها تعریف شده است. درصورتی‌که شرایط خاتمه حاصل شده باشد، بهترین امپراتوری به‌عنوان پاسخ انتخاب شده و وضعیت استعمارگر پاسخ بهینه مسئله خواهد بود.

معیارهای ارزیابی مدل

پنج معادله‌ای که در ادامه آورده شده‌اند معمول‌ترین توابع محاسبه خطا در شبکه‌های عصبی هستند که به ترتیب میانگین مربعات خطا (MSE[1])، مجذور میانگین مربعات خطا (RMSE[2])، میانگین مطلق درصد خطا (MAPE[3]) میانگین مطلق خطا (MAE[4]) نام دارند.

$$MSE = \frac{\sum(y_t - y'_t)^2}{N} = \frac{\sum(e_t)^2}{N} \quad (۵)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum(e_t)^2}{N}} \quad (۶)$$

$$MAPE = \frac{1}{N}\sum\left|\frac{y_t - y'_t}{y_t}\right| = \frac{1}{N}\sum\left|\frac{e_t}{y_t}\right| \quad (۷)$$

$$MAE = \frac{1}{N}\sum|y_t - y'_t| = \frac{1}{N}\sum|e_t| \quad (۸)$$

y_t مقدار واقعی خروجی و y'_t مقدار تخمین زده‌شده توسط مدل تعریف شده است. e_t مقدار خطای پیش‌بینی مدل برای t امین گروه داده‌ها است. در مدل پیشنهادی، مقدار پیش‌بینی‌شده توسط ANFIS با استفاده از معیار RMSE مورد ارزیابی قرار می‌گیرد و مقدار آن به‌عنوان سیگنال تحلیل‌کننده به الگوریتم رقابت استعماری داده می‌شود.

1. Mean of Squared Error
2. Root Mean Squared Error
3. Mean Absolute Percentage Error
4. Mean Absolute Error

www.SID.ir

Archive of SID

داده‌های مورد استفاده

داده‌های مورد استفاده در این تحقیق به‌صورت روزانه بوده و در بازه زمانی فروردین سال ۱۳۸۹ تا اسفندماه ۱۳۹۴ برای ۱۴۴۰ روز کاری بورس تهران جمع‌آوری شده است. از نکات قابل‌توجه در این تحقیق ، جمع‌آوری و تدوین این داده‌ها به‌صورت روزانه به شمار می‌رود که باعث افزایش دقت در پیش‌بینی نتایج در مدل‌های شبکه عصبی و هوش مصنوعی خواهد شد. بر همین اساس، برای هریک از ویژگی‌ها ۱۴۴۰ داده مورد استفاده قرار گرفته و با احتساب ۶۸ ویژگی برای حل مدل پیشنهادی جمعاً ۹۷۹۲۰ داده جمع‌آوری شده و مورد استفاده قرار گرفته است.

عوامل مؤثر در چهار دسته؛ شامل شاخص‌های بورس، شاخص‌های اقتصادی، شاخص-های تحلیل فنی و شاخص‌های شمعدان ژاپنی تقسیم‌بندی شده‌اند.

شاخص‌های بورس عبارت‌اند از: شاخص کل سهام در بازار بورس هامبورگ[1]، شاخص کل سهام در بازار بورس توکیو[2]، شاخص کل سهام در بازار بورس نیویورک[3]، ۵۰۰ سهام برتر در بازار بورس نیویورک[4]، تعداد افراد خریدار و فروشنده حاضر در هرروز کاری، تعداد دفعات خریدوفروش کل سهام در یک روز کاری، شاخص کل بورس اوراق بهادار همان روز، شاخص کل بورس اوراق بهادار یک روز قبل، شاخص کل بورس اوراق بهادار دو روز قبل، قیمت بازگشایی روزانه شاخص، قیمت بسته شدن روزانه شاخص، بالاترین قیمت روزانه شاخص، پایین‌ترین قیمت روزانه شاخص، حجم روزانه خریدوفروش روزانه کلیه سهام و گردش مالی روزانه کلیه سهام.

شاخص‌های اقتصادی عبارت‌اند از: نرخ قیمت دلار ایالات‌متحده آمریکا[5]، قیمت فروش طلای جهانی[6]، قیمت هر بشکه نفت خام اوپک[7]، تورم، نقدینگی، نرخ بهره، قیمت یک گرم

1. German Stock (DAX)
2. Japan Stock (NIKEI 225)
3. New York Stock (NASDAQ)
4. S&P 500
5. Exchange rate: Dollar
6. Republic Gold selling Price
7. Oil Barrel Cost

www.SID.ir

Archive of SID

طلای ۱۸ عیار، قیمت سکه تمام بهار آزادی و شاخص بهای کالاها و خدمات مصرفی.

شاخص‌های تحلیل فنی عبارت‌اند از: شاخص آماری K٪، شاخص آماری D٪، نشانگر R٪ ویلیامز، میانگین متحرک ساده ۶ روزه و شاخص میانگین متحرک نمایی.

شاخص‌های شمعدان ژاپنی عبارت‌اند از: Row approach $\frac{O_i}{C_1}$ ، i=۲،۳،۴،۵،۶،۷ $\frac{C_i}{C_1}$ ، i=۵،۶،۷ $\frac{H_i}{C_1}$ ، i=۵،۶،۷ و $\frac{L_i}{C_1}$ i=۵،۶،۷ و Signal Approach: $\frac{O_i}{C_i}$، i=۲،۳،۴،۵،۶،۷ $\frac{C_i}{C_1}$ ، i=۵،۶،۷ $\frac{H_7}{\mathbf{Max(O_7,C_7)}}$، $\frac{\mathbf{Min(O_7,C_7)}}{L_7}$، $\frac{\mathbf{Max(O_7,C_7)}}{\mathbf{Max(O_6,C_6)}}$، $\frac{\mathbf{Min(O_7,C_7)}}{\mathbf{Min(O_6,C_6)}}$، $\frac{Min(O_6,C_6)}{Max(O_5,C_5)}$، $\frac{Min(O_7,C_7)}{Min(O_5,C_5)}$، $\frac{Max(O_7,C_7)}{Max(O_5,C_5)}$، $\frac{Max(O_6,C_6)}{Min(O_5,C_5)}$، $\frac{H_7}{H_6}$، $\frac{H_7}{H_5}$، $\frac{L_6}{L_5}$، $\frac{L_7}{L_5}$، $\frac{O_7}{H_6}$، $\frac{L_6}{O_7}$ و $\frac{C_7}{O_6}$.

یافته‌های پژوهش

تنظیم پارامتر مناسب و دقیق الگوریتم‌های فرا ابتکاری، اثرات قابل‌توجهی بر عملکرد آن‌ها دارد که باید با توجه به مسئله مسأله تحت بررسی انجام گیرد. در این پژوهش از روش طراحی آزمایش‌های تاگوچی برای تنظیم پارامترهای الگوریتم رقابت استعماری استفاده شده است. پارامترهای موردنظر و سطوح مختلف آن‌ها بر اساس نرم‌های مشابه موجود در ادبیات تحقیق و با توجه به داده‌های تحقیق و ساختار آن، انتخاب شده‌اند. جدول (۱) نشان‌دهنده پارامترهای مورد آزمایش و سطوح آن‌ها است.

جدول ۱: تعیین سطوح سه‌گانه پارامترها

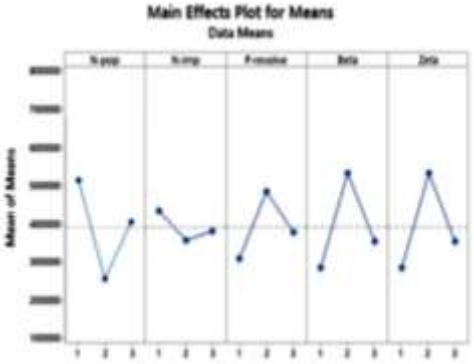
	سطح ۱	سطح ۲	سطح ۳
کشورهای اولیه	۳۰	۴۰	۵۰
امپراتورها	۵	۸	۱۰
نرخ انقلاب	۰/۱	۰/۳	۰/۵
β	۱	۱/۵	۲
ξ	۰/۰۲	۰/۰۵	۰/۱

نتایج تحلیل نرم‌افزار در نمودارهای ۵ و ۶ نشان داده

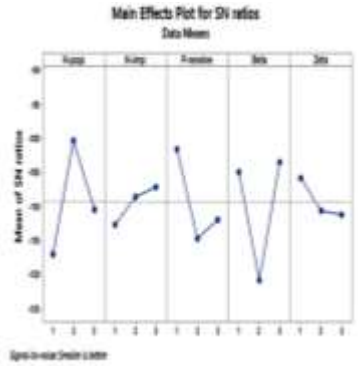
www.SID.ir

Archive of SID

شده است.

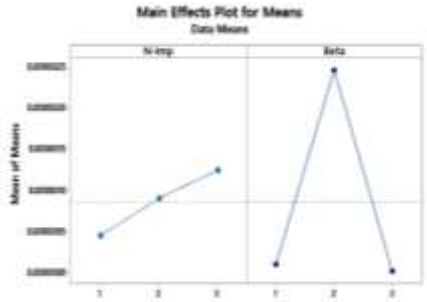


شکل ۵: نمودار Means **به ازای آزمون ۵ پارامتر**

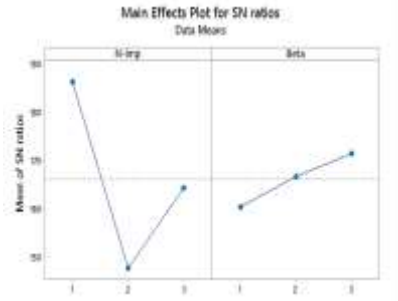


شکل ٦: نمودار SN Ratio **به ازای آزمون ۵ پارامتر**

با توجه به دو نمودار بالا، می توان مقدار بهینه نرخ تعداد جمعیت، نرخ انقلاب و ع را تعیین کرد. N-pop در سطح سوم، نرخ انقلاب سطح دوم و ع سطح اول، مقدار بهینه را اختیار می‌کنند. مقدار بهینه دو پارامتر دیگر نیز با طراحی دوباره در سطح‌های متفاوت تعیین می‌شوند.



شکل ۷: نمودار Means **به ازای آزمون ۲ پارامتر**



شکل ۸: نمودار SN Ratio **به ازای آزمون ۵ پارامتر**

www.SID.ir

Archive of SID

با توجه به نمودارهای به‌دست‌آمده، پارامتر N_{imp} مقدار بهینه خود را در سطح اول و متغیر β در سطح سوم اتخاذ می‌کند. مقدار بهینه تمام پارامترهای مورد بررسی در جدول (۲) نشان داده شده است.

جدول ۲: مقدار بهینه پارامترهای ICA

پارامتر	کشورهای اولیه	استعمارگرها	نرخ انقلاب	β	ξ
مقدار بهینه	۴۰	۵	۰/۱	۲	۰/۰۲

بهینه‌سازی مقدار NF[1]

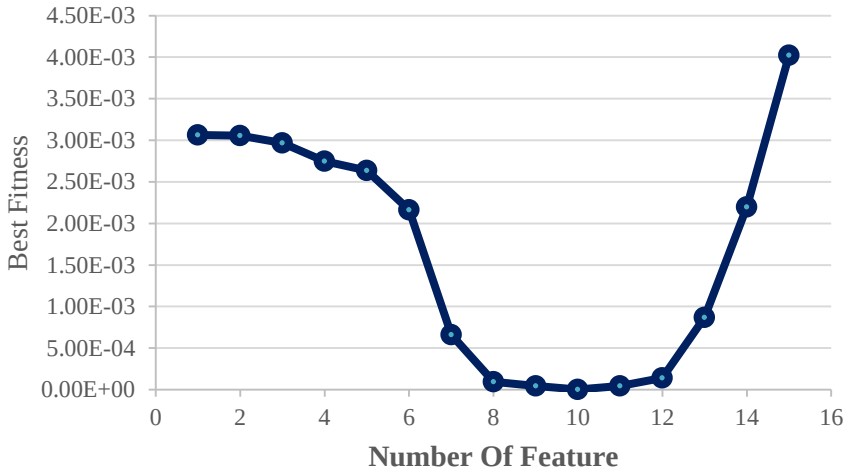
با استفاده از الگوریتم ICA، تعداد بهینه ویژگی‌ها به دست می‌آید. بدین‌صورت برای تعیین تعداد بهینه این ویژگی‌ها، برای مقدار پیشنهادی ۱ تا ۱۵ ویژگی الگوریتم ICA اجرا شده و RMSE های متناظر به‌دست‌آمده از هر اجرا با یکدیگر مقایسه می شوند.

همان‌طور که در شکل ۹ ملاحظه می‌شود، RMSE در NF=۱۰ نسبت به سایر مقادیر NF کمترین مقدار را دارا است. ازاین‌رو مقدار NF=۱۰ به‌عنوان مقدار بهینه در انتخاب تعداد ویژگی‌های مؤثر بر پیش‌بینی شاخص کل بورس اوراق بهادار تهران در الگوریتم رقابت استعماری استفاده می شود.

1. Number Of Feature

www.SID.ir

Archive of SID



شکل ۹: نمودار مقدار بهینه NF

تفکیک و تشخیص میزان اثرگذاری انواع ویژگی‌ها در پیش‌بینی شاخص کل سهام

چنانچه در جدول (۳) مشاهده می‌شود، کل ویژگی‌های مورد استفاده به‌عنوان داده‌های ورودی، به پنج گروه اصلی طبقه‌بندی شده است. برای یافتن اثرگذاری گروه‌های مذکور، با در نظر گرفتن هر یک به‌صورت جداگانه به‌عنوان داده‌های ورودی، پیش‌بینی شاخص کل سهام توسط ANFIS انجام گرفت.

معیارهای ارزیابی خطا MSE، RMSE، MAE و MAPE برای هر بار پیش‌بینی شاخص کل سهام، با داده‌های ورودی مختلف استخراج شد. معیار ارزیابی خطا، برای شاخص‌های اقتصادی کمترین خطا و برای شاخص‌های جهانی بیشترین خطا را دارد. شاخص‌های اقتصادی را می‌توان نمایانگر بازارهای جایگزین بورس نامید. به عبارتی سرمایه‌گذاران اصلی بورس سبد سرمایه‌گذاری خود را بر روی انواع بازارهای پولی و مالی و کالا تقسیم‌بندی کرده‌اند؛ که با توجه به شرایط بازارها در مورد سرمایه‌گذاری خود تصمیم‌گیری می‌نمایند.

www.SID.ir

Archive of SID

برای این گروه از ویژگی‌ها به دلیل عدم وجود وابستگی مستقیم و ارتباط تنگاتنگ بین بازار بورس تهران با سایر بورس‌های بین‌المللی، به دست آمدن معیار ارزیابی خطای بالا ، توجیه‌پذیر است.

انتخاب ویژگی‌های بهینه با استفاده الگوریتم ICA

با توجه به تعداد بهینه NF، باید تعداد ۱۰ ویژگی انتخاب شوند که بیشترین تأثیر را در پیش‌بینی شاخص کل سهام بورس دارند. بر این اساس، برای انتخاب ویژگی‌های بهینه از الگوریتم ICA استفاده می‌شود. با توجه به تعداد کل ویژگی‌ها، مقدار N_{Var} را برابر با ۶۸ در نظر گرفته شد.

جدول ۳: مقایسه اثرگذاری گروه‌های مختلف از ویژگی‌ها روی پیش‌بینی شاخص کل سهام

انواع ویژگی‌ها	MSE	RMSE	MAE	MAPE
شاخص‌های اقتصادی	۰/۰۰۰۰۰۰۳۶۱۸	۰/۰۰۰۶۰۱۵۲	۰/۰۰۰۳۵۲۵۶	۰/۰۰۰۵۹۱۵۸
شاخص‌های بورس	۰/۰۰۰۰۳۱۳۸۸	۰/۰۰۱۷۷۱۷	۰/۰۰۱۰۶۱۶	۰/۰۰۱۷۶۶۶
شاخص‌های بورس جهانی	۰/۰۰۱۰۱۱۴	۰/۰۳۱۸۰۳	۰/۰۱۲۵۷۶	۰/۰۲۱۴۵۸
شاخص‌های شمعدان ژاپنی	۰/۰۰۰۶۶۱۶	۰/۰۱۷۰۵	۰/۰۱۰۰۰۲	۰/۰۱۷۰۴۸
شاخص‌های تحلیل فنی	۰/۰۰۰۶۶۶۹	۰/۰۲۵۸	۰/۰۱۰۰۷۸	۰/۰۱۶۹۱۳

برای به دست آوردن تعداد تکرارها باید همگرا شدن الگوریتم ICA برای مدل پیشنهادی بررسی شود. هم‌گرائی در حل مدل توسط الگوریتم ICA در ۲۰ مرتبه تکرار اتفاق افتاده است. از طرفی به دلیل این‌که در الگوریتم ICA پاسخ‌های اولیه (وضعیت کشورهای اولیه) به‌صورت تصادفی انتخاب می‌شود، نتیجتاً الگوریتم در دفعات بعدی اجرا، در مرتبه تکرارهای مختلفی به هم‌گرائی می‌رسد. با تناوب اجرا در تکرارهای مختلف ملاحظه شد که هم‌گرائی در بازه تکرار ۲۰ تا ۳۰ اتفاق خواهد می افتد و ازاین‌رو برای حل الگوریتم حداکثر تعداد تکراربرابر ۵۰ قرار داده شد.

پس از قرار دادن مقادیر پارامترهای مربوطه برای یافتن ویژگی‌های بهینه، مدل ۸۰ مرتبه

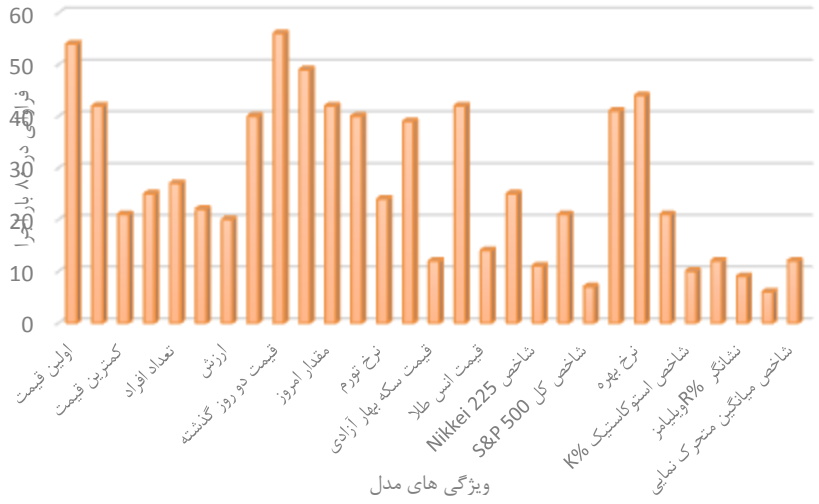
www.SID.ir

Archive of SID

توسط الگوریتم ICA اجرا می‌شود. با توجه به تعیین مقدار NF پس از هر مرتبه اجرا، ۱۰ ویژگی بهینه با توجه به بهترین پاسخ ارائه شده توسط الگوریتم رقابت استعماری انتخاب می شود.

با در نظر گرفتن فراوانی تکرار هر ویژگی در هر دفعه اجرا و ترسیم نمودار توزیع فراوانی مربوطه (شکل ۱۰)، مشخص می شود فراوانی ویژگی نخستین قیمت در ابتدای روز کاری بورس، بیشترین قیمت در همان روز، قیمت دو روز گذشته، قیمت روز گذشته، قیمت همان روز کاری، شاخص بهای کالا و خدمات مصرفی، نرخ ارز آزاد، قیمت هر بشکه نفت خام سبد اوپک، قیمت هر گرم طلای ۱۸ عیار و نرخ بهره به ترتیب، دارای بیشترین فراوانی در ۸۰ دفعه اجرای مدل است.

به‌عبارتی‌دیگر، ویژگی‌های مذکور، همان ۱۰ ویژگی مؤثر برای پیش‌بینی بهینه شاخص کل بورس برای روز آینده می‌باشند که از میان ۶۸ ویژگی با استفاده از الگوریتم ICA انتخاب شدند.



شکل۱۰: فراوانی ویژگی‌های مؤثر بر بازار بورس در ۸۰ دفعه اجرای مدل توسط الگوریتم ICA

www.SID.ir

Archive of SID

پیش‌بینی شاخص بورس اوراق بهادار

در مدل شبکه عصبی فازی، از دو دسته داده آموزش و آزمون استفاده شده است. داده‌های آموزش برای یادگیری و کسب دانش شبکه عصبی فازی از نوع ANFIS و داده‌های آزمون نیز برای آزمون کارایی و مقدار دقت پیش‌بینی مدل به کار رفته است.

با توجه به ادبیات تحقیق و همچنین نیاز بیشتر مدل به داده‌های آموزش برای یافتن روابط بین متغیرهای ورودی و متغیر پاسخ، ۷۵٪ از داده‌های ورودی به مدل به عنوان داده‌های آموزش در نظر گرفته شده و ۲۵٪ باقی‌مانده به عنوان داده‌های آزمون استفاده می‌شوند. داده‌های ورودی در این مدل، ۱۰ ویژگی بهینه تعیین‌شده توسط الگوریتم انتخاب ویژگی رقابت استعماری است. ANFIS از تابع FCM[1] برای اعمال قوانین و تعداد توابع عضویت بهره گرفته است. همچنین تابع FCM به ترتیب از تابع گوسی و تابع خطی برای توابع عضویت ورودی‌های سیستم و خروجی آن استفاده می‌کند. تعداد توابع عضویت ۱۰ در نظر گرفته شده است. تعداد حداکثر تکرارها ۵۰ و حداکثر Epoch نیز برابر با ۱۰۰ است.

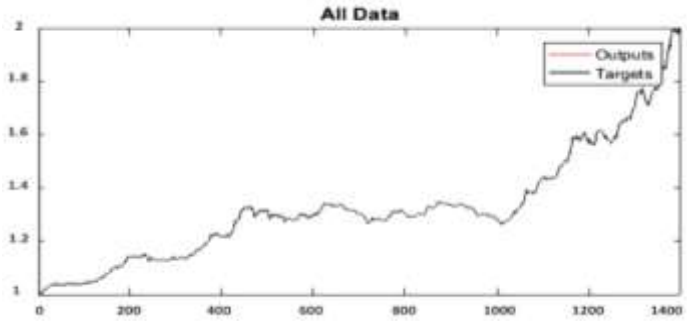
نمودارهای شماره ۱۱ و ۱۲ نتایج پیش‌بینی شبکه ANFIS به ازای داده‌های آزمون و تمام داده‌هارا نشان می دهند . خروجی ها در نمودار مقدار پیش‌بینی‌شده شاخص کل بورس اوراق بهادار، توسط شبکه و اهداف، مقدار واقعی شاخص کل بورس اوراق بهادار تعریف شده است. همان‌طور که در شکل نشان داده شده است، به علت دقت بالای مدل در پیش‌بینی و مینیمم کردن خطا، دو مقدار با تقریب بسیار بالایی بر هم منطبق هستند.

1. Fuzzy C–Means Clustering Algorithm

www.SID.ir

Archive of SID

شکل ۱۱: پیش‌بینی شاخص بورس تهران برای داده‌های آزمون



شکل ۱۲: پیش‌بینی شاخص بورس تهران برای تمام داده‌ها

مقایسه نتایج عملکرد مدل پیشنهادی و شبکه عصبی فازی انطباق پذیر

درصورتی‌که پیش‌بینی شاخص سهام با استفاده از کلیه ۶۸ ویژگی مؤثر توسط شبکه عصبی انطباق پذیر ANFIS صورت بپذیرد، مطابق شکل ۱۳ انطباق‌پذیری داده‌های خروجی از برنامه با نتایج واقعی مطلوب نیست.

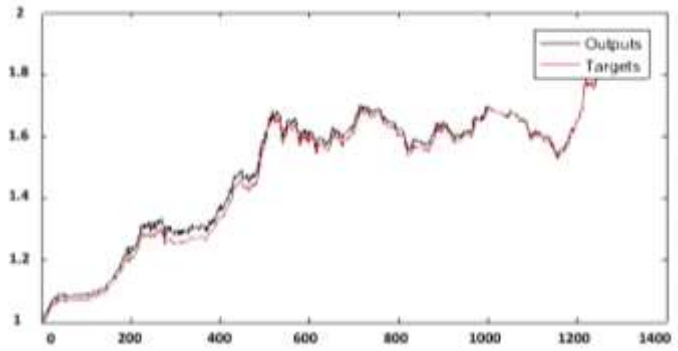
برای توجیه این عدم تطابق‌پذیری باوجود استفاده از کلیه ویژگی‌های مؤثر بر بازار سهام، می‌توان اینگونه عنوان کرد زمانی که تعداد ویژگی‌ها زیاد می‌شود به‌تبع آن تعداد توابع عضویت و قوانین عضویت در شبکه ANFIS افزایش می‌یابد. در این صورت، علاوه بر افزایش زمان اجرا و پیچیده‌تر شدن حل، کسب دانش برای شبکه عصبی به‌درستی صورت نگرفته و در نتیجه با افزایش معیارهای ارزیابی خطا، دقت پیش‌بینی شبکه عصبی به‌طور

www.SID.ir

Archive of SID

محسوسی کاهش می‌یابد.

همچنین می‌توان به این نکته اشاره کرد که به علت وجود همبستگی و ارتباط بین بعضی از ویژگی‌ها و میزان تأثیر مختلف هریک، در صورت استفاده از کلیه ویژگی‌ها برای پیش‌بینی شاخص کل سهام، از این همبستگی و میزان تأثیر متفاوت بین ویژگی‌ها چشم‌پوشی شده و در نتیجه میزان خطا افزایش می‌یابد.



شکل ۱۳: نتایج پیش‌بینی با استفاده از شبکه عصبی فازی بدون انتخاب ویژگی

معیارهای ارزیابی دو مدل در جدول (۴) با یکدیگر مقایسه شده‌اند. نتایج نشان‌دهنده برتری مدل پیشنهادی چه ازلحاظ زمان اجرا و دقت پیش‌بینی است.

جدول ٤: مقایسه عملکرد مدل پیشنهادی با شبکه عصبی فازی ANFIS

	ANFIS-ICA		**ANFIS**	
	Train data	**Test data**	**Train data**	**Test data**
MSE	۱.۲۸e-۱۰	۴.۷e-۱۲	۰/۰۰۰۴۴۸	۰/۰۰۰۱۷
RMSE	۱.۱۳e-۰۵	۲.۱e-۰۶	۰/۰۲۱۱	۰/۰۱۳۰۶
MAPE	۶.۲۵e-۰۶	۱.۲e-۰۶	۰/۰۰۹۷۷	۰/۰۰۸۳
MAE	۱.۰۶e-۰۵	۱.۹۸e-۰۶	۰/۰۱۸۹۵	۰/۰۱۰۶۳
Time	۵۹۹/۵۹۷		۲۴۱۶/۲۷۳	

www.SID.ir

Archive of SID

نتیجه‌گیری

با گروه‌بندی کلیه ویژگی‌های مؤثر بر شاخص کل سهام به پنج دسته و اجرای مدل با در نظر گرفتن هر گروه،این نتیجه حاصل شد که بیشترین تأثیر مربوط به شاخص‌های اقتصادی (شاخص بهای مصرفی، تورم، نرخ بهره، قیمت طلا، قیمت نفت، نقدینگی، قیمت سکه بهار آزادی و ...) و کمترین تأثیر مربوط به شاخص بورس بازارهای بین‌المللی است.

همچنین مشاهده شد که انتخاب بهینه ویژگی از میان کل ویژگی‌ها، به طرز چشمگیری مقدار معیارهای ارزیابی خطای مدل پیش‌بینی را کاهش داده و به‌تبع آن تطابق با نتایج واقعی را بهبود می‌بخشد. تطابق بین نتایج خروجی حاصل از مدل پیشنهادی با نتایج واقعی شاخص کل بورس، به‌مراتب بیشتر از نتایج خروجی حاصل از مدل شبکه عصبی ANFIS است. درصورتی‌که تعداد ورودی ویژگی‌ها به مدل پیش‌بینی شبکه عصبی ANFIS افزایش یابد، به دلیل افزایش توابع عضویت و همچنین قوانین عضویت، پیچیدگی حل افزایش یافته و به تبع آن زمان اجرای مدل طولانی‌تر می‌شود.

www.SID.ir

Archive of SID

منابع

حیدری زارع، ب. کردلوئی، ج. (۱۳۸۹). پیش‌بینی بازده سهام با استفاده از مدل‌های غیرخطی آستانه‌ای و بررسی نقش حجم معاملات در بهبود عملکرد این مدل‌ها. *نشریه علمی‌پژوهشی تحقیقات مالی*، شماره ۳۴، صص. ۹۱-۱۰۸.

سینائی، حسنعلی؛ مرتضوی، سعید ا...؛ تیموری اصل، یاسر. (۱۳۸۴). پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه عصبی مصنوعی. *مجله بررسی‌های حسابداری و حسابرسی*، شماره ۴۱، پاییز، صص. ۸۳-۵۹.

صالحی، مجتبی، کردکتولی، علیرضا، (۱۳۹۶). انتخاب ویژگی‌های بهینه به‌منظور تعیین ریسک اعتباری مشتریان بانکی. *مطالعات مدیریت کسب‌وکار هوشمند*، دوره ۶، شماره ۲۲، زمستان ۱۳۹۶، صفحه ۱۲۹-۱۵۴

منجمی، ا.، ابزاری، م.، و رعیتی شوازی، ع.، (۱۳۸۸). پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار با استفاده از شبکه عصبی فازی و الگوریتم‌های ژنتیک و مقایسه آن با شبکه عصبی مصنوعی. *فصلنامه اقتصادی مقداری*، دوره ۶، شماره ۳، پاییز ۱۳۸۸، صص.۱-۲۶.

هرمزی شیرکو، (۱۳۸۸). *بررسی رابطه بین بازده غیرعادی سهام و متغیرهای کلان اقتصادی در بورس اوراق بهادار تهران*. پایان‌نامه کارشناسی ارشد، دانشگاه آزاد واحد تبریز، دانشکده اقتصاد و مدیریت.

هیبتی، فرشاد؛ موسوی، سیدمصطفی. (۱۳۸۹). پیش‌بینی شاخص بورس سهام با استفاده از مدل‌سازی شبکه عصبی فازی. *پژوهشنامه اقتصادی*، شماره ۷ (ویژه‌نامه بازار سرمایه)، صص. ۷۲-۶۱.

Asadi, S., Hadavandi, E., Mehmanpazir, F. and Nakhostin, M. M. (2012) Hybridization of evolutionary Levenberg-Marquardt neural networks and data pre-processing for stock market prediction, *Knowledge-Based Systems*. Elsevier B.V., 35(5), pp. 245–258.

Ben Naser, Samy and Ghazuani, Samir. (2007). Stock markets, Banks and economic growth Empirical Evidence from the MENA region. *Research in International Business anFinance*, 21(2), pp. 297-315.

Göçken Mustafa, Özçalıcı Mehmet, Boru Aslı, Ayse Tugba Dosdogru.

www.SID.ir

Archive of SID

(2016). Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Systems With Applications*, 44(1), pp. 320–331.

Jang J.S. (1993). ANFIS: Adaptive-Network-based Fuzzy Inference Systems. *IEEE Transactions on Systems Man, and Cybernetics*, 23(3), pp. 665–685.

Huang, Y., Capretz, L. F., & Ho, D. (2019). Neural Network Models for Stock Selection Based on Fundamental Analysis. *arXiv preprint arXiv*:1906.05327

Lincy Rubell Marion G, Jessy John C. (2016). A multiple fuzzy inference systems frame work for daily stock trading with application to NADAQ stock exchange. *Expert Systems with Applications,* 44(1), pp. 13-22.

M. Sudhakar, J. Albert Mayan and N. Srinivasan. (2016). Intelligent Data Prediction System Using Data Mining and Neural Networks. *Proceedings of the International Conference on Soft Computing Systems*, *Advances in Intelligent Systems and Computing 398*. India 2016.

Pimenta, A., Guimaraes, F. G., Carrano, E. G., Nametala, C. A. L. and Takahashi, R. H. C. (2014) GoldMiner: A genetic programming based algorithm applied to Brazilian Stock Market, in *2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pp. 397–402.

Prema, K. V., Manish Agarwal, N., Krishna, M. and Agarwal, V. (2016) Stock Market Prediction using Neuro-Genetic Model, *Indian Journal of Science and Technology*, 8(35), pp. 876–891.

Qiu, M., Song, Y. and Akagi, F. (2016) Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market, Chaos, Solitons & Fractals. *Elsevier Ltd*, 85(3), pp. 1–7.

Vieira, S. M., Sousa, J. M. C., & Kaymak, U. (2012). Fuzzy Criteria for Feature Selection. *Fuzzy Sets and Systems*, 189(1), pp. 1–18.

Yung-Keun Kwon and Byung-Ro Moon. (2007). A Hybrid Neurogenetic Approach for Stock Forecasting, *IEEE Transactions on Neural Networks*, 18(3), pp. 851–864.

www.SID.ir