Archive of SID

چکیده

تصمیم‌گیری برای سرمایه‌گذاری، همواره یکی از مهمترین مسائل سرمایه‌گذاران بوده است. سرمایه‌گذاران در بورس تلاش می‌کنند تا از میان طیف وسعی از شرکت‌های عضو و با توجه به شاخص‌های مالی متعدد سرمایه‌گذاری خود را انجام دهد تا ضمن دستیابی به بالاترین بازده، کمترین ریسک را نیز متحمل شود. بر این اساس امروزه، روش‌های متعددی برای تحلیل داده‌های این شرکت‌ها وجود دارد. یکی از روش‌هایی که از میان انبوه داده‌ها، به دسته‌بندی این شرکت‌ها می‌پردازد روش خوشه‌بندی است. بر این اساس، پژوهش حاضر با هدف شناسایی و تفکیک شرکت‌های موفق از ناموفق بازار بورس اوراق بهادار تهران با استفاده از روش خوشه‌بندی K-means و حل این مسأله با کمک الگوریتم‌های فرا ابتکاری انجام گرفته است. نتایج این تحقیق حاکی از آن است که حل این مسأله به روش فراابتکاری در مقایسه با روش‌های معمول، کاراتر بوده و به بهینه‌ی سراسری منجر شده است. همچنین این نتایج با نتایج حاصل از تفکیک شرکت‌های عضو بورس بهادار تهران با روش تعیین ورشکستگی آلتمن مقایسه شده و توسط این روش نیز تأیید شده است.

کلید واژه:
خوشه‌بندی، بازار بورس اوراق بهادار تهران، مدل K-means، الگوریتم‌های فرا ابتکاری، مدل ورشکستگی آلتمن.

ارائه مدل ترکیبی خوشه‌بندی شرکت‌های عضو بورس بهادار تهران: رویکرد الگوریتم‌های فراابتکاری

زهرا صادقی آرانی (نویسنده مسئول)
دکتری مدیریت، عضو هیئت علمی و استادیار گروه مدیریت دانشگاه کاشان، کاشان، ایران
SadeqiArani@kashanu.ac.ir
علی محقر
دکتری مدیریت صنعتی، استاد و عضو هیئت علمی دانشگاه تهران
amohaghar@ut.ac.ir

مقدمه

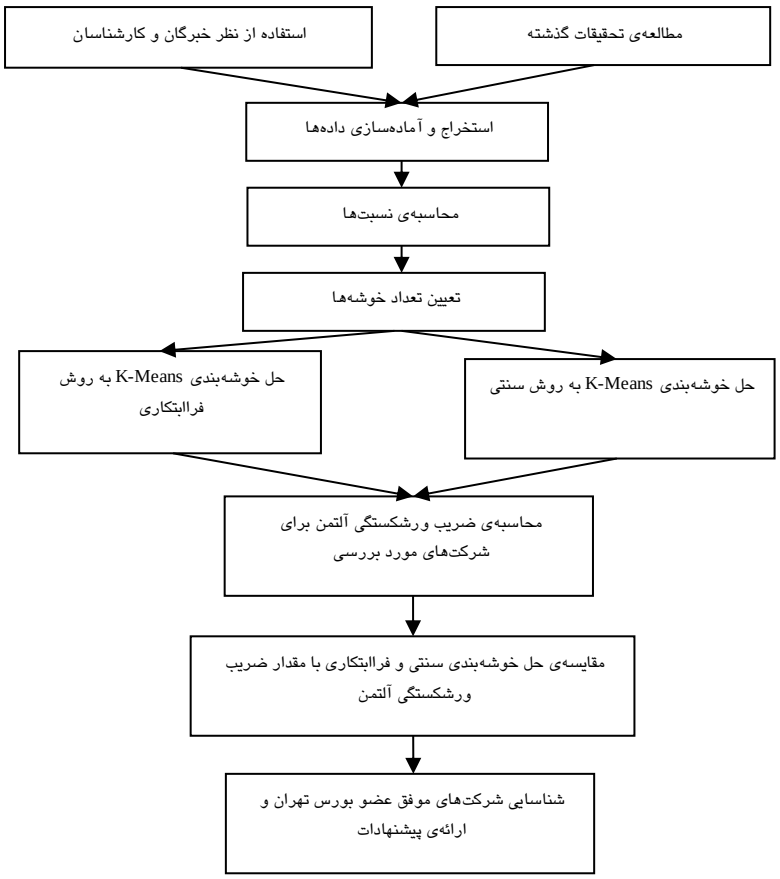
توسعه اقتصادی در جهان امروز، مدیون بورس و فعالیت‌های بازار سرمایه است. وجود بازارهای قوی و کارآمد همراه با سازمان‌های مالی مناسب و فعال از شاخص‌های رشد و توسعه‌ی اقتصادی هر کشور است. بورس اوراق بهادار، بازار متشکل و رسمی برای خرید و فروش اوراق سهام یا اوراق قرضه با ضوابط و مقررات خاص است. رونق بورس اوراق بهادار به عنوان معیار مهمی برای سنجش میزان توسعه‌ی اقتصادی کشور به شمار رفته و باعث تعادل و توازن اقتصادی جامعه است (ودیعی و شکوهی‌زاده، ۱۳۹۱). پیچیدگی بازار سرمایه، به‌ویژه طیف گسترده‌ی ابزارهای سرمایه‌گذاری و عوامل متعدد مؤثر بر آنها، تصمیم‌گیری درخصوص انتخاب سبد دارایی را برای سرمایه‌گذاران را دشوار می‌کند (Zhong & Enke, ۲۰۱۷). سرمایه‌گذاران همواره به دنبال انجام بهترین سرمایه‌گذاری بوده تا نفع بیشتری عایدشان شود، بدین منظور درپی تفکیک شرکت‌های موفق و ناموفق هستند با توجه به طیف وسیع شرکت‌های عضو و تعدد شاخص‌های مالی ارائه شده، ارائه روشی که بتواند این امر را آسان کند و از قابلیت اتکای بالایی برخوردار باشد حائز اهمیت است تا از این طریق، افراد سرمایه‌گذاری خود را در مجموعه‌ای از شرکت‌ها به گونه‌ای انجام دهد تا ضمن دستیابی به بالاترین بازده، کمترین ریسک را نیز متحمل شود (Nagy & Ormos, ۲۰۱۸). تاکنون تکنیک‌های گوناگونی برای تفکیک شرکت‌های موفق از ناموفق ارائه شده است. مدل‌های تعیین ورشکستگی (مهرانی و همکاران، ۱۳۸٤)، تکنیک‌های تصمیم‌گیری چندمعیاره (دانش شکیب و فضلی، ۱۳۸۸)، تحلیل پوششی داده‌ها (قلی‌زاده و ابراهیم پورازبری، ۱۳۸۸؛ تقی‌زاده و پوروربی، ۱۳۸۹؛ خواجوی، ۱۳۸٤) تکنیک‌های فراابتکاری (راعی و همکاران، ۱۳۹۰؛ درخشان و همکاران،

www.SID.ir

Archive of SID

۱۳۹۱) و داده‌کاوی (فرید و پورحیدری، ۱۳۹۱؛ Tekin & Gümüs, ۲۰۱۷؛ Zhong & Enke, ۲۰۱۷) از جمله روش‌هایی است که برای تفکیک شرکت‌های موفق از شرکت‌های ناموفق مورد استفاده قرار گرفته است. وجه مشترک تمامی این روش‌ها استفاده از داده‌های متعدد در صورت‌های مالی این شرکت‌ها می‌باشد. یکی از این تکنیک‌ها مورد استفاده، خوشه‌بندی[۱] است که از جمله روش‌های رایج داده‌کاوی[۲] به‌حساب می‌آید. خوشه‌بندی با ارائه‌ی یک مسأله‌ی بهینه‌سازی (مینیمم‌سازی) و با توجه به شاخص‌های متعدد تلاش می‌کند تا چیزهایی را که دارای ویژگی‌های یکسانی‌هستند در یک خوشه قرار دهد. با ارائه‌ی روش‌های خوشه‌بندی، روش‌های گوناگونی نیز برای حل این مسأله ارائه شده است که هریک دارای مزایا و معایب و محدودیت‌هایی می‌باشد. بسیاری از این روش‌ها به بهینه‌ی محلی[۳] ختم می‌شوند. علاوه براین، مسائل خوشه‌بندی از جمله مسائل NP-Hard [۴] می‌باشد. برای این مسائل تاکنون راه حل سریع و قابل انجام در زمان معقول پیدا نشده‌است از این رو برای حل، به روش‌های فراابتکاری نیاز است (Güngör & Ünler, ۲۰۰۸). در سال‌های اخیر، روش‌های فراابتکاری گوناگونی برای حل مسائل خوشه‌بندی به کار گرفته شده است که مزیت این روش‌ها جستجوی کارا و یافتن بهینه‌ی سراسری است. براین اساس، پژوهش حاضر به منظور، خوشه‌بندی شرکت‌های اوراق بورس بهادار تهران با استفاده از روش K-means و حل این مسأله به‌وسیله‌ی الگوریتم‌های فراابتکاری بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO)، رقابت استعماری (ICA) و جستجوی هارمونی (HS) انجام گرفته است.

در ادامه مقاله بدین صورت ساختاردهی شده است. پس از بیان مقدمه به توضیح روش خوشه‌بندی K-means و الگوریتم‌های فراابتکاری برای حل این مسأله پرداخته می‌شود. در بخش سوم روش تحقیق، جامعه‌ی آماری و متغیرهای تحقیق ارائه خواهد شد. نتایج حاصل از خوشه‌بندی شرکت‌ها با استفاده از الگوریتم‌های مذکور و مقایسه‌ی آن با مدل ورشکستگی آلتمن بخش چهارم این پژوهش را تشکیل می‌دهد. در نهایت نیز پیشنهادات کاربردی و پژوهشی متناسب با نتایج ارائه شده است. فرایند انجام تحقیق در شکل ۱ آمده است.



شکل (۱): فلوچارت انجام تحقیق

www.SID.ir

Archive of SID

۱. مبانی نظری تحقیق

در ابتدای این بخش ابتدا تعریفی از خوشه‌بندی ارائه شده سپس روش خوشه‌بندی K-means به عنوان رایج‌ترین روش خوشه‌بندی ارائه می‌شود. در ادامه نیز الگوریتم‌های فراابتکاری رقابت استعماری (ICA)، بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO) و جستجوی هارمونی (HS) که برای حل مسأله خوشه‌بندی در این پژوهش به‌کار گرفته شده است، معرفی می‌گردد.

۱،۱. خوشه‌بندی

پردازش داده، یکی از شاخص‌های بسیار مهم در دنیای اطلاعات است و تاکنون روش‌های متعددی برای داده‌کاوری معرفی شده است. از میان روش‌های ارائه شده، روش خوشه‌بندی بیش از سایرین رایج بوده به طوری که قابلیت ورود به فضای داده و تشخیص ساختار داده‌ها را امکان‌پذیر می‌نماید (Chen. et. al., ۲۰۱۳). خوشه‌بندی به عمل تقسیم جمعیت ناهمگن به تعدادی از زیر مجموعه‌ها یا خوشه‌های همگن گفته می‌شود (Andrés, et. al., ۲۰۱۱). وجه تمایز خوشه‌بندی از طبقه‌بندی[۵] این است که خوشه‌بندی به دسته‌های از پیش تعیین شده تکیه ندارد. در دسته‌بندی بر اساس یک مدل هر کدام از داده‌ها به دسته‌ای از پیش تعیین شده اختصاص می‌یابد؛ این دسته‌ها یا از ابتدا در طبیعت وجود داشته‌اند (مثل جنسیت، رنگ پوست و...) یا از طریق یافته‌های پژوهش‌های پیشین تعیین گردیده‌اند. در خوشه‌بندی هیچ دسته از پیش تعیین شده‌ای وجود ندارد و داده‌ها صرفاً براساس تشابه گروه‌بندی می‌شوند و عناوین هر گروه نیز توسط کاربر تعیین می‌گردد.

خوشه‌بندی، یافتن ساختاری در مجموعه‌ای از داده‌ها است که طبقه‌بندی نشده‌اند. به بیان دیگر می‌توان گفت که خوشه‌بندی قراردادن داده‌ها در گروه‌هایی است که اعضای هر گروه از زاویه‌ی خاصی شبیه یکدیگرند. در نتیجه، شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل می‌باشد. معیار شباهت در اینجا، فاصله بوده یعنی نمونه‌هایی که به یکدیگر نزدیکترند در یک خوشه قرار می‌گیرند. بر این اساس، محاسبه فاصله بین دو داده در خوشه‌بندی بسیار مهم می‌باشد؛ زیرا کیفیت نتایج نهایی را دستخوش تغیر قرار خواهد داد. فاصله که همان معرف عدم تجانس است حرکت در فضای داده‌ها را می‌سازد و سبب ایجاد خوشه‌ها می‌گردد. با محاسبه فاصله بین دو داده می‌توان فهمید که چقدر این دو داده به هم نزدیک هستند و بر این اساس در یک خوشه قرار داده می‌شود.

یکی از روش‌های متداول برای خوشه‌بندی روش K-means است که برای خوشه‌بندی داده‌های متنوعی مورد استفاده قرار گرفته است. این الگوریتم مجموعهٔ داده را به k زیرمجموعه (خوشه) تقسیم می‌کند به طوری که همهٔ اجزاء هر زیر مجموعه از نزدیک‌ترین فاصله با مرکز آن زیر مجموعه برخوردار باشند. این روش، K شیء (نمونه ورودی) را به صورت تصادفی به عنوان مراکز خوشه‌ها انتخاب می‌کند. آنگاه سایر اشیاء (نمونه‌های ورودی) را بر اساس حداقل فاصله‌ی اقلیدسی با مراکز خوشه‌های تعیین شده به خوشه‌های مناسب انتساب می‌دهد. سپس میانگین هر خوشه مجدداً محاسبه شده و به عنوان مرکز جدید خوشه‌ها در نظر گرفته می‌شوند. این عملیات آنقدر تکرار می‌شود تا مراکز خوشه‌ها دیگر تغییر نکنند. معیاری که باید در روش K-Means کمینه شود عبارت است از:

$$E = \sum_{k=1}^{c} \sum_{x \in Qk} \| x - C_k \|$$

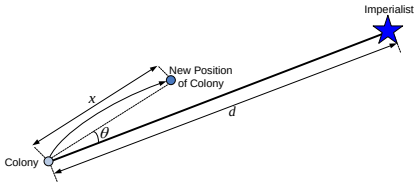
که در رابطهٔ بالا، C تعداد خوشه‌ها، Q_k، k امین خوشه و C_k مرکز خوشه Q_k است.

۲،۱. الگوریتم رقابت استعماری[۶]

الگوریتم رقابت استعماری یک الگوریتم بهینه‌سازی الهام گرفته از تکامل اجتماعی-سیاسی انسان می‌باشد که برای اولین بار توسط آتش‌پزگرگری و کارولوکس (۲۰۰۷) ارائه شد. این الگوریتم با تعدادی جمعیت اولیه تصادفی، که هر کدام از آنها یک «کشور» نامیده می‌شوند، شروع می‌گردد. کشورها به دو دستهٔ مستعمره[۷] و استعمارگر یا امپریالیست[۸] تقسیم می‌شوند. در این الگوریتم، سیاست‌های جذب[۹] و رقابت استعماری[۱۰]، هستهٔ اصلی این الگوریتم راتشکیل می‌دهند .تعدادی از بهترین عناصر جمعیت (معادل نخبه‌ها در الگوریتم ژنتیک) به عنوان امپریالیست انتخاب شده و باقیماندهٔ جمعیت نیز به عنوان مستعمره، در نظر گرفته می‌شوند. استعمارگرها بسته به قدرتشان که با تابع هزینه نسبت عکس دارند، این مستعمرات را با یک روند خاص به سمت خود جلب می‌کنند. شکل ۲ حرکت مستعمرات را به سمت امپریالیست نشان می‌دهد. مطابق این شکل، کشور امپریالیست، کشور مستعمره را در راستای محورهای فرهنگ و زبان به سمت خود جذب می‌کند.

www.SID.ir

Archive of SID



شکل (۲): حرکت مستعمرات به سمت امپریالیست

در این شکل، فاصلهٔ میان استعمارگر و مستعمره با d نشان داده شده است. x و θ نیز اعدادی تصادفی با توزیع یکنواخت به‌صورت رابطه زیر می‌باشند.

$$x \approx U(0, \beta \times d) \qquad ; \qquad \theta \approx U(-\gamma, \gamma)$$

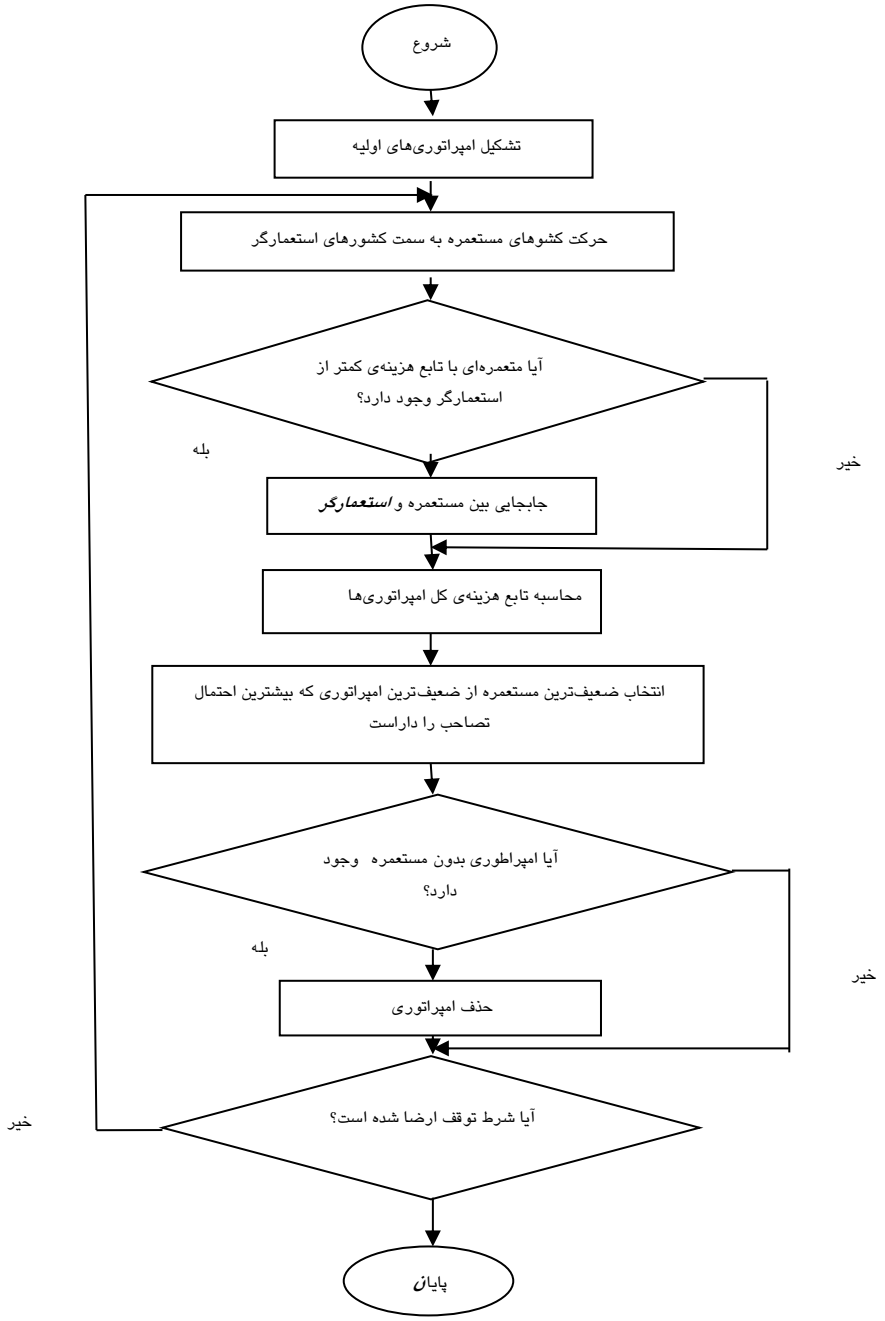
چنانچه β بزرگ‌تر از یک باشد باعث می‌شود تا کشور مستعمره در حین حرکت به سمت کشور استعمارگر، از جهت‌های مختلف به آن نزدیک شود. در این رابطه، γ پارامتری دلخواه است که افزایش آن باعث افزایش جستجوی اطراف امپریالیست می‌شود و کاهش آن نیز سبب می‌گردد تا مستعمرات تا حد ممکن، نزدیک به بردار واصل مستعمره به استعمارگر حرکت کنند. در اکثر پیاده‌سازی‌ها، انتخاب β و γ به ترتیب ۲ و π/٤ انتخاب مناسبی بوده است (Hosseini Nasab, et. al., ۲۰۱۰). قدرت کل هر امپراطوری، به هر دو بخش تشکیل دهنده‌ی آن یعنی کشور امپریالیست (به عنوان هستهٔ مرکزی) و مستعمرات آن، بستگی دارد. در حالت ریاضی، این وابستگی با تعریف قدرت امپراطوری به‌صورت مجموع قدرت کشور امپریالیست، به اضافهٔ درصدی از میانگین قدرت مستعمرات آن (ζ)، مدل شده است.

$$T.C._n = Cost(imperialist_n) + \xi\, mean\{Cost(colonies\, of\, empire_n)\}$$

با شکل‌گیری امپراطوری‌های اولیه، رقابت امپریالیستی میان آنها شروع می‌شود. هر امپراطوری که نتواند در رقابت استعماری، موفق عمل کند و بر قدرت خود بیفزاید (و یا حداقل از کاهش نفوذش جلوگیری کند)، از صحنهٔ رقابت استعماری، حذف خواهد شد. بنابراین، بقای یک امپراطوری، وابسته به قدرت آن در جذب مستعمرات امپراطوری‌های رقیب، و به سیطره درآوردن آنها خواهد بود. در نتیجه، در جریان رقابت‌های امپریالیستی، به تدریج بر قدرت امپراطوری‌های بزرگتر افزوده و امپراطوری‌های ضعیف‌تر، حذف خواهند شد. امپراطوری‌ها برای افزایش قدرت خود، مجبور خواهند شد تا مستعمرات خود را نیز پیشرفت دهند. با گذشت زمان، مستعمرات، از لحاظ قدرت به امپراطوری‌ها نزدیکتر خواهند شد و شاهد یک نوع همگرایی خواهیم بود. حد نهایی رقابت استعماری، زمانی است که یک امپراطوری واحد با مستعمراتی که از لحاظ موقعیت، به خود کشور امپریالیست خیلی نزدیک هستند، در دنیا داشته باشیم. شکل ۳ فلوچارت الگوریتم را نشان می‌دهد.

www.SID.ir

Archive of SID



شکل (۳): الگوریتم رقابت استعماری

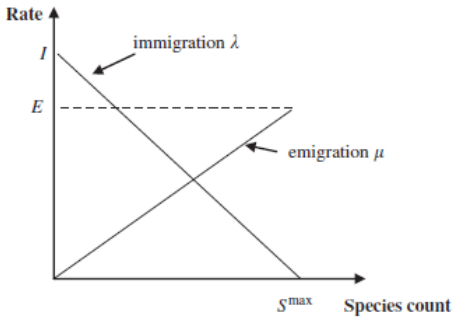
۳,۱. بهینه‌سازی مبتنی بر جغرافیای زیستی

بهینه‌سازی مبتنی بر جغرافیای زیستی[۱۱] (BBO) یکی از الگوریتم‌های نسبتاً جدید در حوزه بهینه‌سازی هوشمند است که در سال ۲۰۰۸ توسط دَن سایمون[۱۲] معرفی شد. در این الگوریتم، از چگونگی انتشار گونه‌های[۱۳] جانداران در زیستگاه‌های متعدد الهام گرفته شده است. با ارائه مدلی احتمالی برای چگونگی مهاجرت گونه در زیستگاه‌ها، مدل ریاضی استخراج شده که در نهایت منجر به ایجاد مدل بهینه‌سازی جدیدی شده است. ایده‌ی اصلی این الگوریم از تمایل جانوران به انحصار منابع و در نتیجه تصمیم به مهاجرت به سکونت‌گاه‌هایی با جمعیت کمتر شکل گرفته است. تابع هدف این الگوریتم با شاخص تناسب سکونت[۱۴] (HSI) اندازه‌گیری می‌شود. برای یک سکونت‌گاه خاص، مهاجرت به[۱۵] این مکان (λ) با جمعیت رابطه‌ی معکوس دارد و با مهاجرت از[۱۶] این مکان (μ) رابطه‌ی مثبت دارد. به عبارت دیگر، هر چه جمعیت یک سکونت‌گاه کمتر باشد نرخ مهاجرپذیری آن بیشتر و مهاجرت از آن کم‌تر است. حداکثر میزان مهاجرت‌پذیری یک مکان با

www.SID.ir

Archive of SID

S^{max} و حداکثر نرخ مهاجرت‌پذیری آن نیز با E نشان‌داده شده می‌شود. به همین ترتیب، I هم بیانگر حداکثر نرخ مهاجرت از یک سکونت‌گاه است (شکل ٤). معمولاً، در تئوری این الگوریتم مجموع I و E برابر یک عدد ثابت است که در این پژوهش برابر یک درنظر گرفته شده است. متغیرهای تصمیم در این الگوریتم (معادل ژن در الگوریتم ژنتیک) با متغیر شاخص تناسب[۱۷] (SIV) تعیین می‌شود که عبارت است از مجموعه عوامل که باعث برازندگی یک سکونت‌گاه نسبت به سکونت‌گاه دیگر شده یا سکونت‌پذیری یک مکان را تعیین می‌کنند. متغیر شاخص تناسب به عنوان متغیر مستقل و شاخص تناسب سکونت به عنوان متغیر وابسته شناخته می‌شود.



شکل (۴): مدل مهاجرت گونه‌ها

در این الگوریتم، سکونت‌گاه‌هایی که شاخص تناسب سکونت کمتری دارند دارای مهاجرت‌پذیری بیشتری هستند. در این سکونت‌گاه‌ها، عواملی که مهاجرت‌پذیری آن‌ها را تشکیل می‌دهند یا همان متغیر شاخص تناسب (SIV)، تحت تأثیر تغییر احتمالی عوامل سایر سکونت‌گاه‌ها با مهاجرت‌پذیری بیشتر قرار می‌گیرند و سایر سکونت‌گاه‌ها، با یک احتمال مشخص، متغیر شاخص تناسب (عامل) را مهاجرت می‌دهند. به عبارت دیگر، هر عامل در سکونت‌گاه مهاجرپذیر i با نرخ λ_i تحت تأثیر قرار گرفته و منبع تغییر با نمونه‌برداری تصادفی و با نرخ μ_j از سایر سکونت‌گاه‌ها تغییر می‌کند. به عنوان مثال، اگر به‌طور تصادفی منبع تغییر (مهاجرت) سکونت‌گاه jام باشد، آن متغیر (متغیر شاخص تناسب)، به احتمال μ_j از سکونت‌گاه jام ($j\neq i$) به سکونت‌گاه مهاجرپذیر منتقل می‌شود. این احتمال (μ_j) براساس شاخص تناسب سکونت (HSI) تعیین می‌شود. یعنی سکونت‌گاهی که شاخص تناسب سکونت بیشتری داشته باشد نرخ مهاجردهی بالاتری نیز داشته و شانس بیشتری در تغییر متغیر شاخص تناسب سکونت‌گاه مهاجرپذیر دارد. به عبارت دیگر، مشخصه‌ها از سکونت‌گاه‌هایی که محل سکونت بهتری هستند، به سکونت‌گاه‌های بدتر (با تابع هدف HSI کمتر) انتقال داده می‌شود. سایمون برای افزایش تنوع[۱۸] در این الگوریتم، جهش[۱۹] را هم برای گونه‌ای که برای تغییر انتخاب شده است را نیز پیشنهاد می‌دهد. به طور کلی، مراحل این الگوریتم به صورت زیر است:

۱- تولید مجموعه جمعیت اولیه (سکونت‌گاه‌ها) به صورت تصادفی و ارزیابی آن‌ها: این ارزیابی براساس شاخص تناسب سکونت (HSI) و با استفاده از متغیرهای شاخص تناسب (SIV) که همان متغیرهای تصمیم مسأله است صورت می‌گیرد.

۲- تعین پارامترهای نرخ مهاجرت‌پذیری به این مکان (λ) و مهاجرت از[۲۰] این مکان (μ)، این دو نرخ بر اساس شاخص تناسب سکونت هر سکونت‌گاه تعیین می‌شود به طوری که سکونت‌گاهی که دارای شاخص تناسب سکونت بیشتری بوده دارای نرخ λ کمتر و μ بیشتر می‌باشد.

۳- به ازای هر سکونت‌گاه مانند i مرحله‌ی ۳ انجام می‌گیرد:

۳-۱ به ازای متغیر kام در سکونت‌گاه i مراحل زیر انجام می‌گیرد:

۳-۱-۱ با احتمال λ_i در x_{ik} (SIV) تغییرات طبق مراحل زیر را اعمال می‌شود:

انجام می‌شود x_{ik} به x_{jk}) را با توجه به مقادیر μ و به صورت تصادفی تعیین گردیده و مهاجرت از متغیر j۲-۱-۳ مبداء مهاجرت ($x_{ik} = x_{ik} + \alpha(x_{jk} - x_{ik})$)

۳-۱-۳ با احتمال معین بر روی مؤلفه‌ی x_{ik} تغییرات تصادفی (جهش) صورت می‌گیرد.

٤- مجموعه‌ی پاسخ‌های جدید بر اساس شاخص تناسب سکونت به دست آمده ارزیابی می‌شود.

٥- جمعیت قدیمی و جمعیت ناشی از مهاجرت ترکیب شده و با استفاده از روش‌های انتخاب، جمعیت مرحله‌ی بَعد ایجاد می‌شود.

٦- در صورت برآورده نشدن شرط خاتمه به مرحله‌ی ۳ برمی‌گردیم.

www.SID.ir

Archive of SID

٤،١. الگوریتم جستجوی هارمونی

الگوریتم جستجوی هارمونی[٢١] (HS) برای اولین بار توسط گیم و همکارانش (٢٠٠١) میلادی ارائه گردید. این الگوریتم با الهام از فرایند بهبود و تکامل هارمونی (هم‌سازی یا هم‌آهنگی) به وسیله‌ی مجموعه‌ای از نوازندگان موسیقی ایجاد شده است. در این روش، هر نوازنده یا ابزار موسیقی[٢٢] یک متغیر تصمیم را نشان می‌دهد. در طول اجرای الگوریتم هر نوت[٢٣] که هر نوازنده می‌نوازد بیانگر مقداری است که به متغیر تصمیم تخصیص داده می‌شود. هدف از تکرار الگوریتم یافتم بهترین هارمونی[٢٤] بین نوازنده‌گان یا نقطه‌ی بهینه‌ی سراسری است. این الگوریتم برای متغیرهای گسسته و پیوسته کارا بوده و به دلیل تصادفی بودن عملیات در مراحل مختلف، احتمال متوق شدن الگوریتم در بهینه‌ی محلی اندک است. مراحل الگوریتم جستجوی هارمونی به صورت زیر است:

١- مقداردهی اولیه مسأله بهینه‌سازی و پارامترهای اولیه: در این مرحله تابع هدف مسأله و پارامترهای الگوریتم مشخص می‌شود. این پارامترها شامل اندازه‌ی حافظه‌ی هارمونی یا تعداد بردارهای جواب[٢٥] (HMS)، نرخ انتخاب از حافظه‌ی هارمونی[٢٦] (HMCR) و نرخ تنظیم گام[٢٧] (PAR) برای بهبود بردارهای جواب و همچنین پهنای باند[٢٨] (BW) برای تنظیم فرکانس و تعداد تکرار الگوریتم (Max It.) به‌عنوان شرط خاتمه است. مقادیر دو نرخ انتخاب از حافظه‌ی هارمونی (HMCR) و نرخ تنظیم گام (PAR) بین صفر تا یک تعیین می‌شوند.

٢- تولید حافظه هارمونی: در این مرحله ماتریس حافظه‌ی هارمونی با انتخاب تصادفی مقادیر متغیرها در هر بردار پاسخ تولید شده و بر اساس مقدار تابع هدف مرتب می‌شوند:

$$HM = \begin{bmatrix} x^1 \\ x^2 \\ . \\ . \\ . \\ x^{HMS} \end{bmatrix} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_N^1 \\ x_1^2 & x_2^2 & \dots & x_2^1 \\ . & & & . \\ . & & & . \\ . & & & . \\ x_1^{HMS} & x_2^{HMS} & \dots & x_N^{HMS} \end{bmatrix}$$

در حافظه هارمونی فرض می‌شود که K هارمونی توسط n نوازنده ساخته شده است، در واقع اندازه حافظه نوازندگان برابر HMS هارمونی است. بنابر این، ماتریس تولیدی شامل HMS سطر و n ستون یا تعداد نوازنده (تعداد متغیرهای تاثیر گذار در مسئله) است.

٣- ایجاد یک هارمونی جدید بهبود یافته:

بردار هارمونی جدید $x' = (x'_1, x'_2, \dots, x'_N)$ برمبنای سه قانون انجام می‌گیرد: ١- انتخاب مؤلفه‌ها از حافظه‌ی قبلی؛ ٢- تنظیم گام و ٣ انتخاب تصادفی. در انتخاب مؤلفه‌های هارمونی جدید، با احتمال HMCR از مؤلفه‌های حافظه‌ی قبلی استفاده شده و با احتمال HMCR -١ به صورت تصادفی ساخته می‌شود.

در ادامه برای بهبود جواب یا به عبارتی تغییر فرکانس، با احتمال PAR تغییرات جزئی در جواب جدید به صورت زیر ایجاد می‌شود.

If (rand()<PAR)

$$x_i^{'} \leftarrow x_i^{'} + rand() \times BW$$

Else

$$x_i^{'} \leftarrow x_i^{'}$$

end

٤- به روز کردن حافظه هارمونی

اگر مقدار تابع هدف بردار هارمونی جدید از بدترین هارمونی در حافظه‌ی هارمونی بهتر باشد به حافظه اضافه شده و بدترین هارمونی موجود از حافظه حذف می‌شود.

٥- گام های ٣و٤ تا زمانی که شرط پایانی ارضا شود یا تکرارها تمام شود ادامه می‌یابد.

٥،١. مروری بر مطالعات قبلی انجام شده

استفاده از الگوریتم‌های فراابتکاری برای تعیین سبد سهام یا ارزیابی عملکرد شرکت‌های فعال در بورس اوراق بهادار در مطالعات گوناگونی انجام گرفته که به برخی از آنها در ادامه اشاره می‌شود:

www.SID.ir

Archive of SID

در پژوهشی انتخاب و بهینه‌سازی سبد سهام با استفاده از الگوریتم ژنتیک بر اساس تعاریف متفاوتی از ریسک انجام شده و ۱٤٦ شرکت فعال در بورس طی سال‌های ۱۳۸۰ تا ۱۳۸۷ مورد بررسی قرار گرفته است. نتایج تحقیق نشان داده است که الگوریتم ژنتیک در تکرارهای مختلف از بهینگی و ثبات بالایی جهت انتخاب سبد سهام برخوردار است. همچنین اختلافی بین دو مدل متفاوت از ریسک (مدل میانگین-واریانس و مدل میانگین-نیم واریانس) مشاهده نشده است (گرگز و همکاران، ۱۳۸۹).

در تحقیقی دیگر از الگوریتم جستجوی هارمونی (HS) به منظور حل مساله بهینه سازی مقید پرتفوی سهام استفاده شده است. در این پژوهش به منظور حل مساله بهینه‌سازی پرتفوی سهام با استفاده از اطلاعات قیمت ۲۰ سهم پذیرفته شده در بورس اوراق بهادار تهران از مهر ۱۳۸۵ تا اسفند ۱۳۸۷، مرز کارای سرمایه‌گذاری برای دو الگو با عامل خطرپذیری واریانس و نیم واریانس رسم می‌گردد. نتایج این تحقیق نشان می‌دهد که روش جستجوی هارمونی در بهینه‌سازی مقید پرتفوی سهام، موفق عمل کرده و در یافتن جواب‌های بهینه در تمامی سطوح خطرپذیری و بازده از دقت قابل قبولی برخوردار است (راعی و همکاران، ۱۳۹۰).

در پژوهشی دیگر از الگوریتم ژنتیک در انتخاب یک مجموعه دارایی از سهام بورس اوراق بهادار استفاده شده است. در این تحقیق، دو الگوریتم ژنتیک برای یافتن پاسخی نزدیک به بهینه طراحی شده است. اولین الگوریتم، مجموعه دارایی با بالاترین بازده و کمترین ریسک و نیز کمترین ضریب همبستگی با سایر دارایی‌ها را انتخاب و الگوریتم ژنتیک دوم، وزن هر یک از دارایی‌ها را در مجموعه دارایی تعیین می‌کند. در نهایت، این دو الگوریتم بر روی سهام عرضه شده در بورس اوراق بهادار تهران با بیش از ۲۰۰ سهام پیاده و نتایج پیشنهادی برای سرمایه‌گذاران رائه شده است (عبدالعلی‌زاده شهیر و عشقی، ۱۳۸۲).

در پژوهشی دیگر با هدف انتخاب بهترین مجموعه از سهام، اهداف چندگانه و تعدد گزینه‌ها لحاظ شده است. در این مقاله پس از توسعه مدل انتخاب سبد سهام مارکویتز، روشی مبنی بر ترکیب دو روش بهینه‌یابی اجتماع مورچگان و شبیه سازی تبرید-تدریجی پارتو پیشنهاد گردیده است. به منظور اعتبارسنجی این روش، عملکرد آن در بورس اوراق بهادار تهران، با عملکرد چند روش فراابتکاری دیگر مقایسه شده است. نتایج بدست آمده حاکی از برتری روش پیشنهادی نسبت به روش‌های مذکور (از نقطه نظر معیارهای قابل طرح در این حوزه) بوده است (درخشان و همکاران، ۱۳۹۱).

در تحقیق دیگر مسأله انتخاب سبد سهام مارکوویتز را در نظر گرفته و در پی رهگیری مرز کارای مورد نظر مدل مارکوویتز تحت شرایط وجود محدودیت‌های عدد صحیح تعداد سهام می‌باشد. بدین منظور به وسیله الگوریتم ژنتیک پیشنهادی، مسأله مقید با استفاده از داده‌های واقعی شرکت‌های داخلی و نیز خارجی حل نموده و با مساله نامقید مارکوویتز مقایسه شده است. نتایج به دست آمده حاکی از آن است که الگوریتم پیشنهادی در هر دو نمونه توانسته است در فضای جستجوی موجه، اقدام به بهینه‌سازی نموده و در نتیجه مساله سبد سهام مقید را به خوبی حل نماید (تقوی‌فرد و همکاران، ۱۳۸٦).

استفاده از تکنیک خوشه‌بندی برای بخش‌بندی سهام شرکت‌ها پذیرفته شده در بورس اوراق بهادار تهران به پژوهش فرید و پورحیدری (۱۳۹۱) مربوط می‌شود. جامعه آماری این تحقیق را تمامی شرکت‌های فعال در بورس در فاصله‌ی سال‌های ۱۳۸٤ تا ۱۳۸۸ تشکیل داده است. شرکت‌های فوق بر اساس شاخص‌های مالی و بر اساس تحلیل خوشه‌ای فازی در سه خوشه قرار گرفته و بر این اساس پیشنهاداتی برای ترکیب سبد سهام سرمایه‌گذاران ارائه شده است (فرید و پورحیدری، ۱۳۹۱).

از روش‌های فراابتکاری برای حل مسأله خوشه‌بندی نیز به وفور استفاده شده است. در مطالعه‌ای که توسط گونگور و آنلر[۲۹] (۲۰۰۸) در کشور ترکیه انجام گرفت برای حل مسائل خوشه‌بندی از نوع K-Harmonic means از روش فراابتکاری جستجوی ممنوع استفاده کرده است. داده‌های این پژوهش از داده‌های استاندارد UCI بوده است. نتایج این تحقیق نشان داده است که استفاده از روش فراابتکاری به جواب مناسب‌تری نسبت به حل خوشه‌بندی معمولی و حتی فازی (به خصوص در اندازه‌های بزرگ مسأله) منجر شده است. همچنین در مطالعه‌ی دیگری از این دو نویسنده در سال ۲۰۰۷ از شبیه‌سازی تبرید برای حل مسائل خوشه‌بندی از نوع K-Harmonic means استفاده شده است. نتایج این تحقیق نیز حاکی از ارائه‌ی بهینه سراسری و همچنین استواری[۳۰] در جواب بوده است.

در مطالعه‌ی دیگری نیز برای حل مسائل خوشه‌بندی K-Harmonic means از روش فراابتکاری جستجوی همسایگی متغیر[۳۱] (VNS) استفاده شده است. در این مطالعه با استفاده از داده‌های استاندارد مشابه پژوهش گونگور و آنلر (۲۰۰۸) نشان داده شده است که روش VNS جواب بهتری را در مقایسه با روش جستجوی ممنوع، شبیه‌سازی تبرید و روش ابتکاری حل مسائل خوشه‌بندی ارائه می‌کند (Alguwaizani, et. al., ۲۰۱۱).

در پژوهشی دیگری در کشور تایوان از روش‌های بهینه‌سازی اجتماع ذرات (PSO) برای تنظیم پارامترهای خوشه‌بندی و همچنین از ANFIS برای پیش‌بینی شکست کسب و کار استفاده شده است. جامعه‌ی آماری این پژوهش را ۱٦۰ شرکت الکترونیک عضو بورس اوراق

www.SID.ir

Archive of SID

بهادار تایوان تشکیل می‌دهند. نتایج این تحقیق حاکی از آن است که روش ترکیبی فوق، کارایی بیشتری در پیش‌بینی‌های بلندمدت دارد (Chen, ۲۰۱۳).

در مطالعه‌ی دیگری نیز از روش الگوریتم ژنتیک الهام‌گرفته شده از کوانتم[۳۲] برای حل مسائل خوشه‌بندی K- means استفاده شده است. نتایج این تحقیق برای داده‌های شبیه‌سازی شده و داده‌های واقعی نشان داده که روش پیشنهادی علاوه بر ارائه‌ی بهینه‌ی سراسری، در مقایسه با سایر روش‌های مرسوم اثربخشی بیشتری داشته است (Xiao, et. al., ۲۰۱۰).

در تحقیقی دیگری برای پیش‌بینی فروش شرکت‌های تخته مدار چاپی[۳۳] در کشور تایوان از روش ترکیبی خوشه‌بندی K- means و شبکه عصبی مصنوعی فازی استفاده شده است. در این پژوهش، ابتدا از خوشه‌بندی برای کاهش و پردازش آسان‌تر داده‌ها و بخش‌بندی بازار بهره برده شده سپس با استفاده از داده‌های ۵ ساله پیش‌بینی فروش صورت گرفته است. نتایج تحقیق نشان داده که استفاده از روش ترکیبی در مقایسه با سایر رویکردها اثربخشی بیشتری داشته است (Chang, et. al., ۲۰۰۹).

در پژوهشی دیگر از روش K- means و نقشه‌ی خود سازماندهی[۳۴] SOM برای خوشه‌بندی رفتارهای مشتریان مخابرات ایران استفاده شده است. در مرحله‌ی بعدی از درخت تصمیم برای طبقه‌بندی این مشتریان و از شبکه عصبی مصنوعی برای پیش‌بینی این رفتار استفاده شده است. نتایج این تحقیق حاکی از دقت بالای این روش ترکیبی نسبت به سایر روش‌های پیش‌بینی بدون خوشه‌بندی است (Farvaresh & Sepehr, ۲۰۱۱).

در پژوهشی دیگر از روش فراابتکاری الگوریتم رقابت استعماری برای حل مسأله‌ی خوشه‌بندی K- means استفاده شده است و نتایج این روش با سایر روش‌های فراابتکاری از جمله الگوریتم ژنتیک، جستجوی ممنوع، بهینه‌سازی ازدحام ذرات، کلونی مورچگان و شبیه‌سازی تبرید مقایسه شده است. نتایج به دست آمده از الگوریتم رقابت استعماری در مقایسه با سایر روش‌های مذکور استوارتر و مناسب‌تر بوده است (Niknam, et. al, ۲۰۱۱).

۲. روش تحقیق، جامعه‌ی آماری و متغیرهای مورد بررسی

جامعه‌ی آماری این پژوهش را ٦٦ شرکت پذیرفته شده در اوراق بورس بهادار تهران تشکیل می‌دهد که صورت‌های مالی آن‌ها منتهی به تاریخ ۹۱/۱۲/۳۰ حسابرسی شده است. این اطلاعات از سایت مرکز بورس اوراق بهادار تهران گرفته و پس از پالایش، صورت‌های مالی نسبت‌های مورد بررسی از این صورت‌ها استخراج شده است. برای بررسی این شرکت‌ها، از ۵ نسبت که در مطالعات آلتمن[۳۵] (۱۹٦۸) مورد استفاده قرار گرفته بهره برده شده است. استفاده از این نسبت‌ها به این دلیل است که اولاً داده‌ها و اطلاعات مربوط به این متغیرها برای تمامی شرکت‌ها موجود و در دسترس بوده است. ثانیاً مطالعات متعددی وجود دارد که از این ۵ نسبت برای ارزیابی یک شرکت و پیش‌بینی ورشکستگی آن‌ها با استفاده از مدل‌های آماری و غیرآماری استفاده کرده است. همچنین نویسندگان فراوانی در حوزه‌های متعدد و در دوره‌های زمانی متفاوت اعتبار شاخص‌های ارائه شده توسط آلتمن را تأیید کرده‌اند (Andrés, et, al, ۲۰۱۱). این نسبت‌ها عبارت‌اند:

X_1: نسبت سرمایه‌ی در گردش به کل دارایی‌ها

X_2: نسبت سود انباشته به کل دارایی‌ها

X_3: نسبت سود قبل از مالیات (EBIT) به کل دارایی‌ها

X_4: نسبت ارزش بازار سهام به ارزش دفتری کل بدهی‌ها

X_5: نسبت فروش کل به کل دارایی‌ها

آلتمن در این مدل به منظور بررسی ورشکستگی یا سلامت یک شرکت معادله‌ی زیر را ارائه کرده است:

$Z= ۰,۷۱۷X_1 + ۰,۸٤X_2 +۳,۱□_۳+۰,٤۲□_٤+۰,۹۹۸□_۵$

چنانچه مقدار Z به دست آمده کمتر از ۱/۲ باشد شرکت ورشکسته، بیش از ۲/۹ سالم و بین این دو مقدار، شرکت حالتی مابین ورشکستگی و غیر ورشکستگی خواهد داشت.

به منظور استفاده از این نسبت‌ها و همسان‌سازی آن‌ها، ابتدا این داده‌ها با استفاده از فرمول زیر نرمالیزه شده و سپس برای حل مسأله‌ی خوشه‌بندی مورد استفاده قرار گرفته است (Alguwaizani, et. al., ۲۰۱۱):

$$x_i = \frac{x_i - Minx_i}{Maxx_i - Minx_i} * 100 \qquad \forall i = 1,2,..., N$$

www.SID.ir

Archive of SID

لازم به ذکر است هر یک از الگوریتم‌های فراابتکاری مورد استفاده در این پژوهش دارای پارامترهای متعددی هستند. برای تنظیم پارامتر هر یک از این الگوریتم‌ها، از مطالعات قبلی انجام شده استفاده گردیده و با اجرای مکرر الگوریتم‌ها، بهترین پارامتر برای هر کدام درنظر گرفته شده که در جدول ۱ آمده است.

جدول (۱): تنظیم پارامترهای الگوریتم‌های فراابتکاری مورد استفاده

الگوریتم	رقابت استعماری (ICA)		بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO)		جستجوی هارمونی (HS)	
پارامترها	۲۰۰	MaxIt.	۲۰۰	MaxIt.	۳۰۰	MaxIt.
	۵۰	nPop	۵۰	nPop	۲۰	HMS
	۱۰	nImp	۰,۲	Keep Rate	۰,۹۹	HMCR
	nPop-nImp	nCol	Linspace(۱۰,nPop)*	μ	۰,۲	PAR
	۲	B	۱-μ	λ	۰,۰۵×(VarMax-VarMin)	BW
	۰,۱	pRevolution	۰,۹	α		
	۰,۰۵	Revolution Rate	۰,۱	pMutation		
	۰,۱	Z				

*: تساعد حسابی بین یک تا صفر به تعداد جمعیت

پیش از خوشه‌بندی شرکت‌های مورد بررسی با استفاده از الگوریتم‌های فراابتکاری یافتن تعداد بهینه‌ی خوشه مفید خواهد بود. بهترین تعداد خوشه آن است که مجموع تشابه بین مرکز هر خوشه و اعضای آن خوشه حداکثر و مجموع تشابه بین مراکز خوشه‌ها حداقل باشد. در این پژوهش به منظور انتخاب بهترین تعداد خوشه از شاخص دیویس-بولدین (DBI) استفاده شده است. در این شاخص تلاش می‌شود تا علاوه بر فاصله‌ی درون خوشه‌ای فاصله‌ی بین خوشه‌ها را نیز بررسی می‌کند. چنانچه $O = \{C^n | n = ۱,۲, \dots, K\}$ مجموع مراکز خوشه‌ها و C^n مراکز خوشه‌ها باشد و $O^n = \{C_i | i = ۱,۲, \dots, \|T^c - O\|\}$ مجموعه باقیمانده چیزهایی که به عنوان مرکز خوشه انتخاب نشده‌اند و T^n مجموعه‌ی تمامی اشیائی است که خوشه‌بندی روی آن صورت گرفته است کیفیت نتایج خوشه‌بندی با K خوشه به صورت R(K) تعریف می‌شود که کمترین مقدار R(K) به عنوان تعداد بهینه خوشه مورد استفاده قرار می‌گیرد. فرمول‌های اندازه‌گیری این شاخص به صورت زیر تعریف می‌شود:

$$R(K) = \frac{۱}{K}\sum_{n=۱}^{K}\left(\max_{۱\le m\le K, m\neq n}\left\{\frac{s_n + s_m}{d_{nm}}\right\}\right)$$

$$s_n = \frac{۱}{\|O^n\|}\sum_{ci\in O^n}\|c_i - c^n\|$$

$$s_m = \frac{۱}{\|O^m\|}\sum_{cj\in O^m}\|c_j - c^m\|$$

$$d_{nm} = \|c^n - c^m\|$$

در این معادله، S^n میانگین شباهت بین مرکز خوشه‌ی C^n و همه‌ی اعضای خوشه‌ی O^n و S^m میانگین شباهت بین مرکز خوشه‌ی C^m و همه‌ی اعضای خوشه‌ی O^m را نشان می‌دهد. همچنین شباهت مرکز دو خوشه با d_{nm} را نشان می‌دهد. هر چه مقدار شاخص دیویس-بولدین (DBI) کمتر باشد نشانگر مطلوبیت بیشتر خوشه‌بندی است.

۳. نتایج و یافته ها

نتایج حاصل از شاخص دیویس-بولدین حاکی از آن است که بهترین تعداد خوشه برابر ۳ می‌باشد (جدول ۲). بنابراین، ۳ خوشه برای خوشه‌بندی شرکت‌های بورس اوراق بهادار تهران در نظر گرفته شده است. سپس با استفاده از الگوریتم‌های فراابتکاری بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO)، کرم شب‌تاب (FA)، رقابت استعماری (ICA) و جستجوی هارمونی (HS) مسأله‌ی K- means به منظور خوشه‌بندی شرکت‌های بورس بهادار تهران حل شده است.

www.SID.ir

Archive of SID

جدول (۲): مقادیر شاخص دیویس-بولدین بر اساس تعداد خوشه

تعداد خوشه (K)	شاخص دیویس-بولدین (DBI)
۲	۰/۹٤۸
۳	**۰/۸۳۷**
٤	۱/۰۰۰
٥	۰/۸۹۹
٦	۰/۹٦۱
۷	۰/۹۳۲
۸	۱/۰۰٦
۹	۰/۹۹٥

نتایج حاصل از ۲۰ بار اجرای برنامه برای هر الگوریتم در قالب بیشترین (Max)، کمترین (Min) و میانگین (Mean) جواب‌ها در جدول ۳ آمده است. این نتایج حاکی از آن است که حل مسأله‌ی خوشه‌بندی K-Means به وسیله‌ی الگوریتم‌های فراابتکاری در مقایسه با روش مرسوم دارای جواب بهتری می‌باشد. علاوه براین، اگر چه نتایج جواب برای سه الگوریتم فراابتکاری مورد استفاده دارای تفاوت اندکی است.

جدول (۳): نتایج حل مسأله‌ی خوشه‌بندی K-Means با استفاده از الگوریتم‌های فرا ابتکاری

الگوریتم فراابتکاری / جواب	K- means	رقابت استعماری (ICA)	بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO)	جستجوی هارمونی (HS)
Max	۱۷۱٤,۰۰۷	۱٥٦۳,٥۰۳	۱٥٦۳,٥۱۳	۱٥٦۲,۹۱
Mean	۱۷۱٤,۰۰۷	۱٥٦۲,۹۷۳	۱٥٦۲,۹٤٤	**۱٥۶۲,۸۹٥**
Min	۱۷۱٤,۰۰۷	۱٥٦۲,۸٦٥	۱٥٦۲,۸۷۱	۱٥٦۲,۸۸۲

جدول (۴): تعداد شرکت‌های موجود در هر خوشه به تفکیک الگوریتم‌های مورد استفاده

الگوریتم فراابتکاری / خوشه	K- means	رقابت استعماری (ICA)	بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO)	جستجوی هارمونی (HS)
خوشه ۱	٤۰	۱٦	۱٦	۱٦
خوشه ۲	۱	۲۲	۲۲	۲۲
خوشه ۳	۲٥	۲۸	۲۸	۲۸
جمع کل	**۶۶**	**۶۶**	**۶۶**	**۶۶**

نتایج نشان می‌دهد در روش خوشه‌بندی K-Means یک شرکت در خوشه‌ی اول و ۲٥ و ٤۰ شرکت به ترتیب در خوشه‌های دوم و سوم قرار گرفت است. این نتایج برای خوشه‌بندی از طریق هر سه الگوریتم‌های فراابتکاری به نتایج یکسانی رسیده که به ترتیب ۱٦، ۲۲ و ۲۸ شرکت در خوشه‌های اول تا سوم قرار می‌گیرد (جدول ٤).

www.SID.ir

Archive of SID

جدول (۵): ویژگی‌های آماری ۵ نسبت مورد بررسی برای شرکت‌های مذکور به تفکیک هر خوشه

خوشه / مقادیر آماری		ماکزیمم (Max)	مینمم (Min)	میانگین (Mean)	انحراف معیار (Std.dev.)
خوشه ۱	X_1	۰,۵۷٤	۰,۱٦٥	۰,۳۲٥	۰,۱۲۳
	X_2	۰,٥٥٨	۰,۲۱۹	۰,٤۲۸	۰,۰۸۸
	X_3	۰,٥٦٤	۰,۱٦٦	۰,٤۲۳	۰,۰۹۳
	X_4	۲٥,۰٥٦	۱,۹۲۱	۸,٤۲۲	٤,۹٦۲
	X_5	۱,۸٦٥	۰,٤۳۷	۰,۹۸۰	۰,۳٦٦
خوشه ۲	X_1	۰,۳۲۱	-۰,۱٥٦	۰,۰۹۱	۰,۱۲۲
	X_2	۰,٤۲۲	۰,۱۷٤	۰,۲۹۳	۰,۰۷٤
	X_3	۰,۳٤۸	۰,۰٦٥	۰,۲۲٥	۰,۰۷۰
	X_4	۹,۸۰۱	۱,۳٥٤	۳,٦۸۹	۲,۱۷۳
	X_5	۱,٤۲۹	۰,۱۸٦	۰,٥۹۲	۰,۲۹٤
خوشه ۳	X_1	۰,٥۱۸	-۰,۲۲۳	۰,۱٥٦	۰,۱٤۸
	X_2	۰,۲۲۱	۰,۰۰٦	۰,۱۱۲	۰,۰٥٦
	X_3	۰,۲۹٦	۰,۰۰٤	۰,۱۱۷	۰,۰٦٤
	X_4	٥,۹۸۰	۰,۲۰٤	۱,۳٦۱	۱,۲٦٤
	X_5	۲,٦۷۲	۰,۱٦۳	۱,۰۹۳	۰,٤۹۱

ویژگی‌های آماری ۵ نسبت مورد بررسی برای ٦٦ شرکت مذکور و به تفکیک هر خوشه در جدول ۵ نشان داده شده است. همچنین، نتایج حاصل از آزمون تحلیل واریانس به منظور بررسی برابری ۵ نسبت (ویژگی) در ۳ خوشه در جدول ٦ آمده است. نتایج این آزمون نشان دهنده‌ی آن است که با توجه به مقدار معنی‌داری (sig.) صفر برای تمامی نسبت‌ها، در سطح اطمینان ۹۵ درصد میانگین این ویرگی‌ها در ۳ خوشه برابر نبوده و به عبارت دیگر تفکیک بین خوشه‌ها به درستی صورت گرفته است.

جدول (۶): آزمون برابری میانگین متغیرهای مورد بررسی در ۳ خوشه

		Sum of Squares	Df	Mean Square	F	Sig.
X_1	Between Groups	.٥۲٤	۲	.۲٦۲	۱۳,۹۰۹	.۰۰۰
	Within Groups	۱,۱۸۷	٦۳	.۰۱۹		
	Total	۱,۷۱۱	٦٥			
X_2	Between Groups	۱,۰۸۱	۲	.٥٤۰	۱۰۱,٤۹٤	.۰۰۰
	Within Groups	.۳۳٥	٦۳	.۰۰٥		
	Total	۱,٤۱٦	٦٥			
X_3	Between Groups	.۹٥٤	۲	.٤۷۷	۸۳,۲۱۹	.۰۰۰
	Within Groups	.۳٦۱	٦۳	.۰۰٦		
	Total	۱,۳۱٦	٦٥			
X_4	Between Groups	٥۰۸,٤۳۹	۲	۲٥٤,۲۱۹	۲۹,٥۱۹	.۰۰۰
	Within Groups	٥٤۲,٥٥٤	٦۳	۸,٦۱۲		
	Total	۱۰٥۰,۹۹۳	٦٥			
X_5	Between Groups	۳,۲۳٥	۲	۱,٦۱۷	۹,٤٤۰	.۰۰۰
	Within Groups	۱۰,۷۹٤	٦۳	.۱۷۱		
	Total	۱٤,۰۲۹	٦٥			

در ادامه، نتایج حاصل از خوشه‌بندی به روش الگوریتم‌های فراابتکاری با نتایج مدل ورشکستگی آلتمن برای شرکت‌های مورد بررسی عضو بورس بهادار تهران مقایسه می‌شود. نتایج این مقایسه نشان می‌دهد که خوشه‌ی اول تماماً شرکت‌های سالمی هستند که میانگین Z آنها بیش از ۲/۹ بوده است. ٦۰ درصد از خوشه‌ی دوم را شرکت‌های سالم و ٤۰ درصد را شرکت‌های نیمه سالم تشکیل می‌دهند. بیش از ۸۲ درصد از خوشه‌ی سوم نیز شرکت‌های غیر سالم هستند (جدول ۷).

www.SID.ir

Archive of SID

جدول (۷): مقایسه‌ی نتایج خوشه‌بندی با روش تعیین ورشکستگی آلتمن

وضعیت شرکت / خوشه	سالم		نیمه سالم		ورشکسته		جمع کل
	تعداد	درصد	تعداد	درصد	تعداد	درصد	
خوشه اول	۱۶	۱۰۰	–	–	–	–	۱۶
خوشه دوم	۱۳	۶۰	۹	۴۰	–	–	۲۲
خوشه سوم	۵	۱۷/۸۶	۲۰	۷۱/۴۳	۳	۱۰/۷۱	۲۸
جمع کل	۳۴		۲۹		۳		۶۶

لیست شرکت‌های هر خوشه و مقدار ضریب ورشکستگی آلتمن (Z) در جدول ۸ آمده است. شرکت‌های داروسازی فارابی و سینا، شرکت‌های معدنی املاح ایران، فرآوری مواد معدنی ایران، معادن بافق، معدنی و صنعتی چادرملو و صنایع خاک چینی ایران، شرکت‌های فولاد خراسان و فولاد خوزستان، صنایع پتروشیمی کرمانشاه، سیمان بهبهان، کالسیمین و باما جزء شرکت‌های خوشه‌ی اول بوده که بر اساس مدل آلتمن شرکت‌های سالم و پربازده هستند.

جدول (۸): شرکت‌های مورد بررسی و مقدار ضریب ورشکستگی آلتمن (Z) به تفکیک خوشه

خوشه‌ی ۱		خوشه‌ی ۲		خوشه‌ی ۳	
نماد شرکت	مقدار Z	نماد شرکت	مقدار Z	نماد شرکت	مقدار Z
دسینا	۵,۱۷۴	اخابر	۳,۲۰۴	بکام	۱,۰۵۸
دفرا	۵,۶۵۴	داسو	۳,۰۲۷	پکرمان	۲,۵۲۳
زمگسا	۷,۰۹۲	دالبر	۲,۸۴۶	چکارن	۲,۰۸۹
سبهان	۷,۲۴۴	درازک	۲,۷۷۸	ختراک	۱,۶۹۳
سنیر	۶,۰۱۳	رتاپ(ذرپارس)	۱,۹۳۵	خشرق	۱,۱۹۹
شاملا	۶,۵۷۱	رکیش	۳,۷۲۹	خفتر	۱,۳۰۶
شفارا	۴,۱۰۲	ساراب	۳,۲۶۷	خموتور	۱,۸۲۳
فاسمین	۷,۳۷۵	ستران	۲,۱۳۷	دابور	۱,۶۶۹
فخاس	۷,۱۲۱	سخاش	۳,۰۲۰	دارو	۱,۴۲۷
فخوز	۵,۱۹۳	سشمال	۲,۳۲۴	دامین	۲,۰۳۱
فراور	۵,۹۷۵	سغرب	۳,۲۷۶	دلر	۱,۶۱۸
کاما	۶,۹۶۸	سکرما	۳,۹۴۸	دیران	۱,۶۱۹
کبافق	۱۲,۳۴۷	سمازن	۳,۵۵۲	رتکو	۲,۱۲۸
کچاد	۶,۵۳۵	شاراک	۳,۴۷۹	سدور	۱,۸۱۷
کخاک	۵,۹۵۸	شیراز	۱,۸۲۳	شدوص	۴,۵۰۷
کرماشا	۳,۳۷۲	غپینو	۲,۲۱۰	شسینا	۲,۴۷۲
		فجر	۵,۰۰۹	شفارس	۱,۵۹۷
		فملی	۳,۵۹۸	شلعاب	۲,۱۵۵
		کرازی	۲,۷۸۰	شنفت	۲,۴۸۴
		کروی	۵,۵۱۰	غاذر	۴,۴۵۵
		کطبس	۱,۸۰۱	غشصفا	۳,۹۲۷
		والبر	۴,۰۵۷	فسرب	۳,۸۸۳
				قزوین	۲,۲۲۷
				قصفها	۳,۹۶۱
				کپشیر	۲,۲۶۶
				کلوند	۱,۴۴۱
				مداران	۲,۳۳۵
				ولیز	۰,۷۲۲

نتیجه‌گیری

تصمیم‌گیری یکی از حیاتی‌ترین وظایف مدیران است که برخی مدیریت را با تصمیم‌گیری‌هایش می‌شناسند. این امر، به خصوص در مسائل مالی با حساسیت‌های بیشتری همراه است. بر این اساس، یکی از دغدغه‌ی اساسی سرمایه‌گذارن در بازار، یافتن شرکت‌های موفقی است که سرمایه‌گذاری در آن سودآور باشد. اهمیت بالای سرمایه‌گذاری در بورس و بازدهی سرمایه در این زمینه، رقابت شدید و گسترده، تعدد شاخص‌های ارزیابی شرکت‌ها و تنوع ابزارهای اندازه‌گیری نیاز به یک روش کارا برای این پایش را الزامی می‌سازد.

www.SID.ir

Archive of SID

معمولاً شناسایی این شرکت‌ها براساس معیارها و از طرق مختلفی صورت می‌گیرد. یکی از این روش‌ها خوشه‌بندی است که در مطالعات جدید، از روش‌های فراابتکاری برای حل آن استفاده شده است. در این پژوهش، استفاده از روش‌های فراابتکاری رقابت استعماری (ICA)، بهینه‌سازی مبتنی بر جغرافیای زیستی (BBO) و جستجوی هارمونی (HS) در حل مسأله‌ی K-Means در مقایسه با روش‌های معمول کاراتر بوده و به بهینه‌ی سراسری منجر شده است. همچنین این نتایج با نتایج حاصل از تفکیک شرکت‌های عضو بورس بهادار تهران با روش تعیین ورشکستگی آلتمن مقایسه شده و توسط این روش نیز تأیید شده است. نتایج حاکی از آن است که ۱۶ شرکت از ٦٦ شرکت مورد بررسی در خوشه‌ی اول قرار گرفته که بر اساس روش آلتمن تمامی این شرکت‌ها جزء شرکت‌های سالم و پربازده بوده و معرفی این شرکت‌ها سرمایه‌گذاران را در امر سرمایه‌گذاری‌های کم‌خطر یاری می‌رساند.

در ادامه به منظور بهبود تفکیک شرکت‌های بورس، استفاده از داده‌های مختلط، روش‌های دیگر تفکیک یا خوشه‌بندی و همچنین خوشه‌بندی فازی پیشنهاد می‌شود. همچنین استفاده از سایر روش‌های فراابتکاری جدید و تنظیم پارامتر با استفاده از روش‌های تاگوچی، طراحی آزمایشات و ... می‌تواند به توسعه‌ی حل مسأله‌ی خوشه‌بندی کمک نماید.

منابع

تقوی فرد محمدتقی، منصوری طاها و خوش طینت محسن. (۱۳۸٦). ارایه یک الگوریتم فراابتکاری جهت انتخاب سبد سهام با در نظر گرفتن محدودیت‌های عدد صحیح پژوهش‌های اقتصادی، ۷(٤): ٦۹-٤۹.

تقی‌زاده هوشنگ و پوراربی میروحید (۱۳۸۹). ارزیابی کارایی شرکت‌های تولیدکننده سیمان در بورس اوراق بهادار تهران با استفاده از مدل تحلیل پوششی داده‌ها (۱۳۸۱-۱۳۸۷)، پژوهشنامه اقتصادی، ۷: ۱۹۹-۲۲۰.

خواجوی شکراله، سلیمی فرد علیرضا و ربیعه مسعود (۱۳۸٤)، کاربرد تحلیل پوششی داده ها (DEA) در تعیین پرتفویی از کارآترین شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، علوم اجتماعی و انسانی دانشگاه شیراز، ۲۲ (۲): ۷۵-۸۹.

دانش شکیب، معصومه و فضلی، صفر (۱۳۸۸). رتبه بندی شرکت‌های سیمان بورس اوراق بهادار تهران با استفاده (AHP–TOPSIS) از رویکرد ترکیبی، چشم‌انداز مدیریت، ۳۲، ۱۲۹-۱۰۹.

درخشان مجتبی، گل مکانی حمیدرضا و حنفی زاده پیام (۱۳۹۱). رویکردی فراابتکاری برای انتخاب سبد سهام با اهداف چندگانه در بورس اوراق بهادار تهران نشریه بین المللی مهندسی صنایع و مدیریت تولید (فارسی) (نشریه بین المللی علوم مهندسی)، ۲۳(۳): ۳۳۱-۳۱۸.

راعی، رضا، محمدی، شاپور و علی بیکی، هدایت (۱۳۹۰). بهینه‌سازی سبد سهام با رویکرد «میانگین–نیم واریانس» و با استفاده از روش «جستجوی هارمونی»، پژوهش‌های مدیریت در ایران (مدرس علوم انسانی)، ۱۵(۳): ۱۲۸-۱۰۵.

عبدالعلی‌زاده شهیر، سیمین و عشقی، کوروش (۱۳۸۲). کاربرد الگوریتم ژنتیک در انتخاب یک مجموعه دارایی از سهام بورس اوراق بهادار. فصلنامه پژوهش‌های اقتصادی ایران،۱۷، ۱۹۲-۱۷۵.

فرید داریوش پورحمیدی مسعود (۱۳۹۱). بخش‌بندی سهام شرکت‌ها پذیرفته شده در بورس اوراق بهادار تهران با استفاده از تحلیل خوشه‌ای فازی، پژوهش‌های حسابداری مالی، ٤ (۳)، ۱۰۵-۱۲۸.

قلی‌زاده محمدحسن و ابراهیم پورازبری مصطفی (۱۳۸۸). طرح مدل رتبه‌بندی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از تحلیل پوششی داده‌ها (مورد: صنعت مواد غذایی و آشامیدنی). پژوهشنامه اقتصادی، ۵: ۱۷۳-۱۹۳.

گرگز، منصور، عباسی، ابراهیم و مقدسی، مطهره (۱۳۸۹). انتخاب و بهینه‌سازی سبد سهام با استفاده از الگوریتم ژنتیک بر اساس تعاریف متفاوتی از ریسک. مدیریت صنعتی دانشکده علوم انسانی دانشگاه آزاد اسلامی واحد سنندج، ۵ (۱۱)، ۱۱۵-۱۳٦.

مهرانی، ساسان، مهرانی، کاوه، منصفی، یاشار و کرمی، غلامرضا (۱۳۸٤). بررسی کاربردی الگوهای پیش‌بینی ورشکستگی زیمسکی و شیراتا در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، بررسی‌های حسابداری و حسابرسی، ۱۲ (٤)، ۱۳۱- ۱۰۵.

ودیعی محمدحسین و شکوهی‌زاده، محمود (۱۳۹۱). بررسی معیارهای مالی مؤثر بر تصمیم‌گیری سرمایه‌گذاران در بورس اوراق بهادار، مجله‌ی دانش حسابداری، ۳ (۸)، ۱۵۱-۱۷۱.

Alguwaizani, A., Hansen, P., Mladenovic, N. Ngai, E. (۲۰۱۱). Variable neighborhood search for harmonic means clustering, Applied Mathematical Modelling, ۳۵, ۲٦۸۸–۲٦۹٤

Andrés, J.D., Lorca, P., Juez, F. Sánchez-Lasheras, F. (۲۰۱۱). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS) Expert Systems with Applications, ۳۸, ۱۸٦٦–۱۸۷۵

Atashpaz-Gargari, E., Lucas, C. (۲۰۰۷). Imperialist Competitive Algorithm: An algorithm for optimization inspired by imperialistic competition. IEEE Congress on Evolutionary Computation, ٤٦٦۱–٤٦٦۷.

www.SID.ir

Archive of SID

Chang, P.Ch., Liu, Ch.H., Fan, Ch.Y. (۲۰۰۹). Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry Knowledge-Based Systems, ۲۲, ۳٤٤–۳٥٥
Chen N., Ribeiro, B., Vieira, A. & Chen, A. (۲۰۱۳) Clustering and visualization of bankruptcy trajectory using self-organizing map Expert Systems with Applications, ٤٠, ۳۸٥–۳۹۳
Chen, M.Y. (۲۰۱۳). A hybrid ANFIS model for business failure prediction utilizing particle swarm optimization and subtractive clustering, Information Sciences, ۲۲۰, ۱۸۰–۱۹٥
Farvaresh, H. & Sepehri, M. M. (۲۰۱۱). A data mining framework for detecting subscription fraud in telecommunication. Engineering Applications of Artificial Intelligence, ۲٤, ۱۸۲–۱۹٤
Güngör, Z., Ünler, A. (۲۰۰۷). K-harmonic means data clustering with simulated annealing heuristic Applied Mathematics and Computation, ۱۸٤, ۱۹۹–۲۰۹
Güngör, Z., Ünler, A. (۲۰۰۸). K-Harmonic means data clustering with tabu-search method, Applied Mathematical Modelling, ۳۲, ۱۱۱٥–۱۱۲٥

Hosseini Nasab, E., Khezri, M., Khodamoradi, M. S. & Atashpaz Gargari E. (۲۰۱۰). An application of Imperialist Competitive Algorithm to Simulation of Energy Demand Based on Economic Indicators: Evidence from Iran, European Journal of Scientific Research, ٤۳ (٤), ٤۹٥-٥۰٦.
Nagy, L. & Ormos, M. (۲۰۱۸), Friendship of Stock Market Indices: A Cluster-Based Investigation of Stock Markets, Journal of Risk Financial Management, ۱۱ (۸۸): ۲-۱٦.
Niknam, T., TaherianFard, E., Narges Pourjafarian, N. & Alireza Rousta, A. (۲۰۱۱). An efficient hybrid algorithm based on modified imperialist competitive algorithm and K-means for data clustering Engineering, Applications of Artificial Intelligence, ۲٤, ۳۰٦–۳۱۷
Simon, D. (۲۰۰۸). Biogeography-Based Optimization, IEEE Transactions on Evolutionary Computation, ۱۲ (٦), ۷۰۲–۷۱۳.
Tekin, B. & Gümüs, F. B.(۲۰۱۷). The Classification of Stocks with Basic Financial Indicators: An Application of Cluster Analysis on the BIST ۱۰۰ Index, International Journal of Academic Research in Business and Social Sciences, ۷(٥): ۱۰٤-۱۳۱.
Xiao, J., Yan, Y., Zhang, J. Tang, Y. (۲۰۱۰). A quantum-inspired genetic algorithm for k-means clustering Expert Systems with Applications, ۳۷, ٤۹٦٦–٤۹۷۳ .

Zhong, X. & Enke, D. (۲۰۱۷). A Comprehensive Cluster and Classification Mining Procedure for Daily Stock Market Return Forecasting, Neurocomputing, ۲٦۷: ۱٥۲-۱٦۸.

۱. ***Clustering***
۲. ***Data mining***
۳. ***Local optimum***
٤. ***Non-deterministic Polynomial-time hard***
٥. ***Classification***
٦. ***Imperialist Competitive Algorithm (ICA)***
۷. ***Colony***
۸. ***Imperialist***
۹. ***Assimilation***
۱۰. ***Imperialistic Competition***
۱۱. ***Biogeography-Based Optimization***
۱۲. ***Dan Simon***
۱۳. ***Species***
۱٤. ***Habitat Suitability Index***
۱٥. ***Immigration Rate***
۱٦. ***Emigration Rate***
۱۷. ***Suitability Index Variable***
۱۸. ***Diversity***
۱۹. ***Mutation***
۲۰. ***Emigration Rate***
۲۱ ***Harmony search***
۲۲. ***Musician***
۲۳. ***Note***
۲٤. ***Best Harmony***
۲٥. ***Harmony Memory size***
۲٦. ***Harmony Memory Consideration Rate (HMCR)***
۲۷. ***Pitch Adjustment Rate (PAR)***
۲۸. ***Band Width***

www.SID.ir

Archive of SID

۲۹. *Güngör & Ünler*
۳۰. *Robustness*
۳۱. *Variable Neighborhood Search (VNS)*
۳۲. *quantum-inspired genetic algorithm*
۳۳. *Printed Circuit Board*
۳۴. *Self-organizing map (SOM)*
۳۵. *Altman*

www.SID.ir