مجله مهندسی مالی و مدیریت اوراق بهادار
شماره سی‌وچهارم / بهار ۱۳۹۷

تشکیل پرتفوی سهام با استفاده از مدل تحلیل ممیز قطری درجه دو و وزن دهی بر اساس احتمال پسین

سعید فلاح‌پور[۱]
حسین پیرایش شیرازی نژاد[۲]

تاریخ دریافت: ۹۶/۰۶/۱۴ تاریخ پذیرش: ۹۶/۰۹/۱۵

چکیده

پیش‌بینی بازدهی سهام از مهم‌ترین مسائل سرمایه‌گذاری در اوراق بهادار است از طرفی، با توجه به اینکه بازار سهام، سیستمی غیرخطی و آشوب گونه که تحت تأثیر شرایط سیاسی، اقتصادی و غیره هست لذا پیش‌بینی بازدهی نیازمند ابزارهای هوشمند و پیشرفته‌ای همچون ماشین‌های یادگیرنده است. در این تحقیق هدف اصلی تفکیک سهام به دو طبقه پربازده و کم بازده و تشکیل پرتفوی است که بدین منظور از تحلیل ممیز قطری درجه دوم و ماشین بردار پشتیبان و همچنین برای گزینش بهترین متغیرها جهت پیش‌بینی طبقه بازدهی از روش انتخاب ویژگی متوالی استفاده شده است. برای هر مدل در حالتی که وزن سهم‌ها برابر است بر اساس پیش‌بینی طبقه بازدهی هر سهم طی سال‌های ۸۸-۹۱ پرتفوی تشکیل داده شده‌است که نتایج رضایت بخش بوده و همه پرتفوی-های تشکیل شده بازدهی بیشتر از بازدهی پرتفوی معیار داشتند. برای مدل تحلیل ممیز با انتخاب ویژگی، از احتمال پسین جهت وزن دهی استفاده و با پرتفوی معیار مقایسه شد که نتایج دلالت بر وجود تفاوت معنادار بین بازدهی دو پرتفوی و برتری پرتفوی مدل تحلیل ممیز دارد.

واژه‌های کلیدی: تحلیل ممیز، انتخاب ویژگی، احتمال پسین، طبقه‌بندی، ماشین بردار پشتیبان.

۱- استادیار دانشکده مدیریت دانشگاه تهران، تهران، ایران
۲- کارشناس ارشد مهندسی مالی، دانشگاه تهران، تهران، ایران (نویسنده مسئول) pirayesh7@gmail.com

۱- مقدمه

یکی از مهم‌ترین مسائلی که هر سرمایه‌گذار اعم از فردی و نهادی در تشکیل پرتفوی با آن مواجه می‌شود، انتخاب سهم‌هایی با بیشترین بازده است. با توجه به پیچیدگی پیش‌بینی طبقه بازدهی به دلیل وجود تعداد زیاد پارامترهای تأثیرگذار بر آن باید حجم انبوهی از اطلاعات مورد بررسی قرار گیرند. با توسعه سیستم‌های خبره و روش‌های داده‌کاوی در سال‌های اخیر استفاده از این سیستم‌ها و روش‌ها در راستای تجزیه و تحلیل داده‌ها جهت پیش‌بینی بازدهی مفید به نظر می‌رسد. به منظور پیش‌بینی طبقه بازده سهام در افق سرمایه‌گذاری مفروض، روش‌های متنوعی وجود دارد که در این پژوهش تفکیک سهام به گروه‌های متناسب با معیار انتخاب شده با استفاده ماشین‌های یادگیرنده انجام می‌شود. با توجه به این مهم که متغیرهای زیادی برای پیش‌بینی بازدهی سهم در نظر گرفته می‌شوند انتخاب مجموعه‌ای از متغیرها که به بهترین نحو تمیز بین دو طبقه سهام پربازده و کم بازده را ایجاد کنند می‌تواند باعث بهبود عملکرد مدل در پیش‌بینی طبقه سهام شود. یکی از روش‌های رایج در انتخاب متغیرها در مدل‌سازی، روش انتخاب ویژگی متوالی است.

تحلیل ممیز و ماشین بردار پشتیبان جزء ماشین‌های یادگیرنده هستند. ماشین‌های یادگیرنده دانش مربوط به نمونه‌ها و مستندات که داده‌ها هستند را استخراج می‌کنند به عبارتی دیگر این سیستم‌ها رفتار یادگیری انسان را در سیستم‌های هوشمند شبیه‌سازی می‌کنند. ماشین‌های یادگیرنده در سیستم‌های هوشمند برای افزایش دانش و تغییر آن، افزایش کارایی و تصحیح خودکار خطا مورد استفاده قرار می‌گیرند. اخذ دانش یکی از مهم‌ترین کاربردهای ماشین یادگیرنده است که امروزه بسیار مورد توجه قرار می‌گیرد. به این معنی که عمل یادگیری اطلاعات پایه را از محیط استخراج کرده و برای تحلیل حوادث آینده از آن بهره می‌گیرد. (داسل استوارت، نورویج پیتر, ۱۳۹۲)

مسئله دیگری که در زمینه تشکیل پرتفوی مطرح است بحث وزن دهی به سهم‌های انتخاب شده است. در این مقاله با استفاده از احتمال پسین سهم‌های انتخاب شده، وزن هر سهم در پرتفوی تعیین می‌شود.

در این تحقیق دقت این دو مدل در تعیین طبقه بازدهی سهم‌ها بررسی می‌شود و بر اساس طبقه‌بندی انجام شده توسط این دو مدل پرتفوی تشکیل می‌دهیم. همچنین برای مدل تحلیل ممیز بر اساس احتمال پسین وزن دهی انجام شده و بازدهی با پرتفوی معیار مقایسه می‌شود.

۲- مبانی نظری و مروری بر پیشینه پژوهش

همان‌طور که در مقدمه گفته شد، اولین مسئله‌ای که هر سرمایه‌گذار با آن مواجه هست تعیین اوراق بهادار ریسک دارای است که مایل به سرمایه‌گذاری در آن‌ها است. یکی از روش‌های حل این مسئله در سال ۱۹۵۲ توسط هری مارکویتز ارائه شد در روش مارکویتز فرض می‌شود که سرمایه‌گذاران در زمان حال مبلغ معینی پول دارند و این پول برای مدت معینی سرمایه‌گذاری می‌شود. در پایان این مدت سهام خریداری شده به فروش می‌رسد و وجوه حاصل به مصرف رسیده یا مجدداً سرمایه‌گذاری می‌شوند. از این رو به روش مارکویتز می‌توان به عنوان روش تک دوره‌ای نگاه کرد. از دیدگاه مارکویتز، پرتفوی‌های کارا، پرتفوی‌هایی هستند که بالاترین بازدهی مورد انتظار را به ازای سطح مشخصی از ریسک و همچنین پایین‌ترین درجه ریسک را به ازای سطح مشخصی از بازدهی مورد انتظار داشته باشند. با توجه به هزینه‌های قابل توجه و نیاز مدل مارکویتز به انجام محاسبات زیاد، وترستون با پیگیری کارهای مارکویتز مدل ارزش در معرض ریسک را ارائه داد. (ویلیام اف شارپ، گوردون جی. الکساندر، جفری وی بیلی, ۱۳۸۹)

ارزش در معرض ریسک نیز از طبقه معیارهای اندازه نامطلوب ریسک هست ارزش در معرض ریسک که سرمایه در معرض خطر نیز نامیده می‌شود، به عنوان معیاری آماری، حداکثر زیان احتمالی پرتفوی را در یک دوره زمانی مشخص با بیان کمی ارائه می‌دهد. در روش ارزش در معرض خطر برای انتخاب پرتفوی بهینه، اصول کار شبیه به مدل مارکویتز است، با این تفاوت که سرمایه‌گذار به دنبال Var کمتر و بازده بیشتر هست. نتایج تحقیقی که به منظور مقایسه این دو مدل در ایران طی سال‌های ۱۳۸۰- ۱۳۸۷ بیانگر آن بود که این دو مدل تفاوت معناداری ندارند. (قدرت اله طالب نیا، مریم فتحی, ۱۳۸۹)

در حال حاضر مدل‌های چند دوره‌ای نیز توسعه پیدا کرده‌اند که اثرات تصمیمات بر دوره‌های بعدی را نیز مورد توجه قرار می‌دهند و با استفاده از الگوریتم‌های مختلف ما نسبت به بهینه‌سازی سبد سهام اقدام می‌کنند. (امیرعباس نجفی، سیامک موشخیان, ۹۳)

مطالعات مختلفی جهت سنجیدن کارایی مدل‌های هوش مصنوعی نسبت به مدل های کلاسیک جهت پیش‌بینی بازدهی انجام شده است. در تحقیقی که به منظور مقایسه عملکرد مدل فاما و فرنچ و شبکه‌های عصبی مصنوعی در پیش‌بینی بازده سهام در بورس تهران در بازه زمانی سال‌های ۱۳۷۸-۱۳۷۹ انجام شد نشان دهنده آن بود که بین میانگین خطای مدل‌ها در پیش‌بینی بازدهی سهام شرکت‌ها و پرتفوی‌های تشکیل شده اختلاف معنی‌داری وجود دارد، که این اختلاف حاکی از برتری مدل شبکه عصبی رگرسیون عمومی بر مدل فاما و فرنچ در پیش‌بینی بازدهی سهام شرکت‌ها و پرتفوی‌ها هست. (پارسائیان, ۱۳۹۱)

سیکورا[۱] و همکاران (۲۰۱۲) روشی را برای انتخاب و ارزیابی سهام بر اساس تحلیل ممیز ارائه کردند و بدین منظور رابطه مقطعی[۲] بین متغیرهای بنیادی و مالی، ضریب بتا و میانگین بازده سهام را در نظر گرفته‌اند. در این مقاله آن‌ها سهم‌هایی که در بورس سائوپائولو در سال‌های ۲۰۰۶ تا ۲۰۱۰ معامله می‌شدند را مورد مطالعه قرار دادند. آن‌ها به منظور طبقه‌بندی سهام ۲۰٪ از سهم‌هایی که بیشترین بازده را داشتند در طبقه پربازده و ۲۰٪ که کمترین بازده را داشتند در طبقه کم بازده، طبقه‌بندی کردند. آن‌ها در این مقاله برای به دست آوردن بهترین تمیز بین گروها در ابتدا با استفاده از روش گام‌به‌گام[۳] به انتخاب متغیرها پرداختند. در این مقاله از آزمون باکس-ام[۴] جهت آزمودن همسانی ماتریکس واریانس_کواریانس بهره گرفته شده است.

پالانیسوامی[۵] و همکاران (۲۰۰۱) با چاپ مقاله‌ای با عنوان انتخاب سهام با استفاده از ماشین بردار پشتیبان به مطالعه در بورس استرالیا و استفاده از این روش جهت انتخاب سهام پرداختند آن‌ها بدین منظور از اطلاعات حسابداری که در گزارش‌های سالیانه شرکت‌ها وجود داشت استفاده کردند و به منظور جلوگیری از ایجاد خطا گزارش شرکت‌هایی که بیش از یک متغیر را نداشت از لیست داده‌های خود حذف نمودند. طبقه‌بندی شرکت‌ها در مقاله آن‌ها بدین صورت بود که ۲۵٪ اول شرکت‌ها از لحاظ بازده را جزو گروه پربازده و بقیه شرکت‌ها را در گروه کم بازده قرار دادند و توانستند در طی ۵ سال با پرتفوی با وزن یکسان به بازده ۲۰۸٪ که بسیار بهتر از بازده ۷۱٪ بنچ مارک شده است دست یابند.

چیا هوی هانگ[۶] و همکاران (۲۰۰۸) با چاپ مقاله‌ای با عنوان یک سیستم پشتیبان تصمیم‌گیری برای سرمایه‌گذاری در شرکت‌های پذیرفته شده در تایوان به بررسی یک روش ترکیبی جهت انتخاب استراتژی‌های سرمایه‌گذاری پرداختند. در این مقاله دو روش تحلیل ممیز و درخت تصمیم‌گیری به عنوان مدل‌های پایه در نظر گرفته شدند آن‌ها توانستند بازده سالیانه بین ۵۹٪ تا ۳۷٪ را به دست آورند که بهتر از ٪۱۲/۷ شاخص تایفکس است. آن‌ها برای دسته‌بندی سهام به پربازده و کم بازده از حداقل بازده مورد قبول استفاده کردند و آن را ۲۸٪ در نظر گرفتند.

رامون هرتا[۷] و همکاران (۲۰۱۳) با چاپ مقاله‌ای بر اساس مدل یادگیرنده به بررسی یک استراتژی معامله‌ای سودآور جهت شناسایی سهام پربازده و کم بازده پرداختند. آن‌ها در مقاله خود از متغیرهای تکنیکال و فاندمنتال بهره برده و مدل را یک‌بار در هر ماه آموزش دادند. آن‌ها جهت خرید سهام از ۲۵٪ اول و جهت فروش از ۲۵٪ آخر سهامی که توسط مدل ماشین بردار پشتیبان غیرخطی شناسایی شده بود استفاده نمودند و توانستند بدون احتساب هزینه‌های معاملاتی به بازده سالانه ۱۵٪ با نوسان پذیری زیر ۸٪ دست یابند.آن‌ها همچنین با مقایسه بین ماشین بردار پشتیبان خطی و غیرخطی به این نتیجه رسیدند که با وجود آن که استفاده از مدل خطی نیز

منجر به کسب بازدهی می‌شود اما مدل غیرخطی بازدهی بیشتری را کسب می‌کند.در این مقاله اثر سه مجموعه از ویژگی‌ها تکنیکال، فاندمنتال و ترکیبی مورد بررسی قرار گرفت و نتایج حاصله دلالت بر این نکته دارند که ویژگی‌های مبتنی بر تکنیکال عملکرد بهتری نسبت به فاندمنتال دارند و بهترین عملکرد با ترکیب این دو ویژگی حاصل می‌شود.

۳- مدل‌های پژوهش

۳-۱- تحلیل ممیز قطری در جه دوم

تحلیل ممیز یک تکنیک آماری است که این امکان را می‌دهد تا تفاوت‌های بین دو گروه یا بیشتر را برحسب چند متغیر به صورت همزمان مورد مطالعه قرار گیرند و از لحاظ فنی درواقع بسط یافته‌ی تحلیل واریانس چند متغیری است.

در این تحلیل تمرکز روی داده‌های مربوط به موارد است که این موارد می‌تواند انسان‌ها، حیوانات، وضعیت اقتصادی در زمان‌های مختلف و خیلی موارد دیگر باشد. گروه‌ها باید طوری تعریف شوند که هر مورد فقط و فقط به یک گروه تعلق داشته باشد. زمان‌هایی که نمی‌توان عضویت موارد در گروها را تشخیص داد باید این موارد از تحلیل کنار گذاشته شوند تا به نتایج خطا منجر نشود. این موارد بعد از رسیدن به معادلات تشخیصی می‌توان عضویت گروهی‌شان را مشخص کرد. (داداش زاده ناهیده ، احیایی پویان, ۱۳۹۳)

دو کلاس از مشاهدات را با میانگین‌های μ_y و کوواریانس‌های Σ_y در نظر بگیرید. می‌توان افتراق دهنده فیشر بین این دو توزیع را به صورت نسبت واریانس بین دو کلاس به واریانس درون دو کلاس تعریف کرد:

(۱)

$$S = \frac{\sigma^2_{\text{between}}}{\sigma^2_{\text{within}}} = \frac{(\vec{w}\cdot\vec{\mu}_{y=1} - \vec{w}\cdot\vec{\mu}_{y=0})^2}{\vec{w}^T\Sigma_{y=1}\vec{w} + \vec{w}^T\Sigma_{y=0}\vec{w}} = \frac{(\vec{w}\cdot(\vec{\mu}_{y=1} - \vec{\mu}_{y=0}))^2}{\vec{w}^T(\Sigma_{y=0} + \Sigma_{y=1})\vec{w}}$$

به عبارت دیگر این مقدار، مقیاسی از نسبت سیگنال به نویز برای برچسب‌گذاری کلاس است. می‌توان نشان داد که حداکثر جداسازی زمانی اتفاق می‌افتد که:

$$\vec{w} \propto (\Sigma_{y=0} + \Sigma_{y=1})^{-1}(\vec{\mu}_{y=1} - \vec{\mu}_{y=0}) \qquad (۲)$$

وقتی که فرضیات LDA ارضا شد، معادله بالا معادل با LDA خواهد بود. حتماً به یاد داشته باشید که بردار نرمال ابر صفحه جداکننده است. به عنوان یک مثال، در یک مسئله دوبعدی، خطی که دو گروه را به بهترین شکل تقسیم می‌کند عمود منصف است. به طور کلی، نقاط داده‌ای که باید جدا شوند باید بر روی بردار $\vec{w}$تصویر شوند. پس آستانه‌ای که به بهترین وجه داده را جداسازی می‌کند از تحلیل توزیع یک‌بعدی انتخاب می‌شود. .به هر حال، اگر تصاویر نقاط از هر دو کلاس تقریباً یک توزیع را نشان دهد، ابر صفحه وسط تصاویر دو مرکز یک انتخاب مناسب خواهد بود. در این مورد پارامتر c در شرط آستانه صریحاً به صورت زیر خواهد بود. (خلیل میرزایی،۱۳۸۸)

(۳)

$$c = \vec{w} \cdot \frac{1}{2}(\vec{\mu}_{y=0} + \vec{\mu}_{y=1}) = \frac{1}{2}\vec{\mu}_{y=1}^{t}\Sigma^{-1}\vec{\mu}_{y=1} - \frac{1}{2}\vec{\mu}_{y=0}^{t}\Sigma^{-1}\vec{\mu}_{y=0}$$

تحلیل ممیز درجه دوم تفاوت زیادی با تحلیل ممیز خطی ندارد بجز آنکه فرض می‌شود ماتریس کواریانس برای هر طبقه می‌تواند متفاوت باشد و بنابراین ماتریس کواریانس Σ_k را برای هر گروه K , ... , 2, 1= ,k ,k متفاوت در نظر گرفته می‌شود که در نتیجه تابع ممیز درجه دوم به صورت ذیل است:

(۴)

$$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k) + \log\pi_k$$

در تابع ممیز درجه دوم بر خلاف تابع ممیز خطی ماتریس کواریانس یکتا نیست و نمی‌توان عبارت درجه دوم را حذف کنید بنابراین تابع ممیز از درجه دوم است و قانون ممیز به شکل زیر است:

(۵)

$$\hat{G}(x) = \arg\max_k \delta_k(x)$$

مرزهای تصمیم[۸] به صورت معادلات درجه دو از x هستند.

به دلیل آنکه تحلیل ممیز درجه دوم نسبت به حالت خطی در ماتریس کواریانس انعطاف‌پذیرتر است بهتر به داده‌ها فیت می‌شود هر چند تعداد متغیرها در تحلیل ممیز درجه دوم نسبت به حالت خطی به طور قابل ملاحظه‌ای افزایش می‌یابد که این نیز خود به دلیل آن است که برای هر گروه یک ماتریس کواریانس جدا وجود دارد.

تحلیل ممیز قطری درجه دوم شبیه تحلیل ممیز درجه دوم است با این تفاوت که یک ماتریس کواریانس قطری به صورت زیر تخمین زده می‌شود. (ThePennsylvaniaStateUniversity)

$$C(x) = \text{argmin}_k \sum_{i=1}^{G} \left\{ \frac{(x_i - \mu_{ki})^2}{\sigma_{ki}^2} + \log \sigma_{ki}^2 \right\} \qquad (۶)$$

۳-۲- مدل ماشین بردار پشتیبان

ماشین بردار پشتیبان روشی است که برای دسته‌بندی داده‌های خطی و غیرخطی استفاده می‌شود. بطور مختصر یک SVM الگوریتمی است که در ابتدا با کمک یک نگاشت غیرخطی داده‌های آموزشی اولیه را به یک بعد بالاتر تبدیل کرده و در این بعد به دنبال یک ابر صفحه بهینه می‌گردد تا داده‌های یک کلاس را از کلاس دیگر به صورت خطی تفکیک کند. با یک نگاشت غیرخطی مناسب به یک بعد بالای کافی، داده‌های دو کلاس را همیشه می‌توان با کمک یک ابر صفحه تفکیک کرد.

فرض کنیم مجموعه نقاط داده $\{(x_1, c_1), (x_2, c_2), \dots, (x_n . c_n)\}$ را در اختیار داریم و می‌خواهیم آن‌ها را به دو طبقه $c_i = \{-1,1\}$ تفکیک کنیم. هر x_i یک بردار p بعدی از اعداد حقیقی است.

ایده اصلی این است که یک جداکننده مناسب انتخاب شود. منظور، جداکننده‌ای است که بیشترین فاصله را با نقاط همسایه از هر دو طبقه دارد. این جواب درواقع بیشترین مرز را با نقاط مربوط به دو طبقه مختلف دارد و می‌تواند با دو ابر سطح موازی که حداقل از یکی از نقاط دو طبقه عبور می‌کنند، کران‌دار شود. این بردارها، بردارهای پشتیبان نام دارند. فرمول ریاضی این دو ابر سطح موازی که مرز جداکننده را تشکیل می‌دهند.

$$w.x - b = \pm 1 \qquad (۷)$$

نکته قابل توجه این است که اگر داده‌های تعلیمی به صورت خطی تفکیک‌پذیر باشند، می‌توان دو ابر سطح مرزی را به گونه‌ای انتخاب کرد که هیچ داده‌ای بین آن‌ها نباشد و سپس فاصله بین این دو ابر سطح موازی را به حداکثر رساند. با به کارگیری قضایای هندسی، فاصله این دو ابر سطح عبارت است از $2/|w|$ ، پس باید $|w|$ را به حداقل رساند. همچنین باید از قرار گرفتن نقاط داده در ناحیه درون مرز جلوگیری کرد، برای این کار یک محدودیت ریاضی به تعریف فرمول اضافه می‌شود. برای هر i ، با اعمال محدودیت‌های زیر اِطمینان حاصل می‌شود که هیچ نقطه‌ای در مرز قرار نمی‌گیرد:

برای داده‌های مربوط به طبقه اول

$$w.x_i - b \geq 1 \qquad (۸)$$

برای داده‌های مربوط به طبقه دوم

$$w.x_i - b = -1 \qquad (۹)$$

می‌توان این محدودیت را به صورت رابطه زیر نشان داد:

$$c_i(w.x_i - b) \geq ۱ \qquad , 1 \leq i \leq n \qquad (۱۰)$$

حل یک مسئله بهینه‌سازی غیر محدب[۹] بسیار دشوارتر از حل یک مسئله بهینه‌سازی محدب است. خوشبختانه می‌توان مسئله را با جایگزینی $\frac{1}{2}\|w\|^2$ به جای $|w|$ بدون تغییری در جواب، حل کرد. این مسئله از نوع مسائل برنامه‌نویسی درجه دوم[۱۰] است. پس مسئله بهینه‌سازی به شکل زیر تبدیل می‌شود:

به حداقل رساندن $\frac{1}{2}\|w\|^2$ با در نظر گرفت محدودیت $c_i(w.\ x - b) \geq 1 \quad , 1 \leq i \leq n$

(هان ژیان، کمبر میشلین، پی ژان, ۱۳۹۳)، (هنردوست پارسا ،۱۳۹۳)، (منیری آرش ،۱۳۸۵)، (راعی رضا، فلاح پور سعید, ۱۳۸۷)

۳-۳- انتخاب ویژگی متوالی رو به عقب[۱۱] (SBS)

با اعمال تکنیک‌های کاهش داده‌ها، نمایش کوچک‌تری از مجموعه داده‌ها وجود دارد، در حالیکه جامعیت داده‌های اولیه حفظ می‌شود. با تولید نتایج تقریباً یکسان از این داده‌های کاهش یافته، روش کارآمدتر و سریع‌تر عمل خواهد کرد. (هان ژیان، کمبر میشلین، پی ژان, ۱۳۹۳)

در انتخاب ویژگی متوالی رو به عقب، الگوریتم با مجموعه‌ای شامل تمام ویژگی‌ها شروع می‌کند و در هر بار تکرار الگوریتم، ویژگی که بوسیله تابع ارزیابی انتخاب می‌شود، را از مجموعه مورد نظر حذف می‌کند. این کار را تا زمانی ادامه می‌دهد که شرط خاتمه ارضا شود. مشکل این روش این است که ویژگی حذف شده را دیگر به مجموعه اضافه نمی‌کند، حتی اگر مناسب باشد. (Burges, Christopher, 1998)

۴- روش‌شناسی پژوهش

برخی را عقیده بر این است که علم همان روش است. در هر حال می‌توان به خوبی پذیرفت که هیچ علمی فاقد روش نیست و دست آوردهای هر پژوهش علمی به همان نسبت حائز ارزشند که با روش‌هایی درست اخذ شده باشند. (ساروخانی ۱۳۸۲)

تعاریف متعدد و متفاوتی برای تحقیق بیان شده است که یکی از آن‌ها تحقیق را بدین گونه تعریف کرده است. فرآیند تنظیم طرح مشخص و اجرای آن برای دستیابی به نتایج، بر اساس سؤال اصلی در راستای اهداف تعیین شده است (احمدی ۱۳۹۱)

انواع تحقیقات علمی را می‌توان برحسب هدف طبقه‌بندی کرد. در رایج‌ترین نوع، تحقیقات به دو دسته تحقیقات بنیادی و تحقیقات کاربردی طبقه‌بندی می‌شوند. (سی ۱۳۸۶) هدف تحقیقات کاربردی، توسعه دانش کاربردی در یک زمینه خاص است و به سمت کاربرد عملی دانش هدایت می‌شود. (همکاران ۱۳۷۸) این تحقیقات با استفاده از زمینه و بستر شناختی و معلوماتی که از طریق تحقیقات بنیادی فراهم شده، برای رفع نیازمندی‌های بشر و بهبود و بهینه‌سازی ابزارها، روش‌ها، اشیا مورد استفاده قرار می‌گیرد. (نیا ۱۳۸۴) پژوهشی که در این مقاله انجام شده است جزء دسته تحقیقات کاربردی دسته‌بندی می‌شود.

همان‌طور که گفته شد در این تحقیق سعی بر آن است که با استفاده از مدل ماشین بردار پشتیبان و تحلیل ممیز به پیش‌بینی طبقه بازده هر سهم پرداخته و بر اساس آن پرتفوی تشکیل دهیم. به منظور بررسی مؤثر بودن استفاده از مدل‌های منتخب در کسب بازدهی بیشتر برای مدل ماشین بردار پشتیبان با کرنل درجه دوم، دو پرتفوی و برای تحلیل ممیز قطری درجه دوم سه پرتفوی تشکیل داده و بازدهی پرتفوی‌های تشکیل داده شده با پرتفوی معیار مقایسه شده است.

برای پیش‌بینی طبقه بازدهی سهام هر سال ابتدا با استفاده از اطلاعات سال‌های مالی قبل و الگوریتم انتخاب ویژگی متوالی رو به عقب متغیرهای لازم جهت پیش‌بینی طبقه بازدهی سال بعد انتخاب شده و با استفاده از این متغیرهای سال قبل پیش‌بینی طبقه بازدهی در سال بعد با استفاده از دو مدل تحلیل ممیز قطری درجه دوم و تحلیل ممیز قطری درجه دوم انجام می‌شود. لازم به ذکر است که در این تحقیق سهم‌هایی که در هر سال از منظر بازدهی جزء ۲۰٪ برتر بودند سهام با طبقه پربازده و سهم‌هایی که به لحاظ بازدهی جزء ۲۰٪ کمترین بازدهی بودند، سهام کم بازده در نظر گرفته شده‌اند. برای بررسی وجود تفاوت معنادار بین دو مدل از آزمون مقایسه زوجی t استفاده شده است که در ادامه نتایج آن آورده می‌شود.

در این مقاله ۴۹ شرکت از شرکت‌های پذیرفته شده در شرکت بورس اوراق بهادار تهران در دوره‌های زمانی ۱۳۸۸ تا سال ۱۳۹۱بر اساس معیارهای زیر انتخاب شدند.

۱) اطلاعات آن‌ها بصورت کامل در دوره مورد بررسی وجود داشته باشد.
۲) جزء خدمات مالی نباشند.
۳) زیان ده نباشند.
۴) سال مالی منتهی به ۱۲/۲۹ باشد.
۵) در دوره مورد بررسی سال مالی تغییر نکرده باشد.

به منظور گردآوری اطلاعات مورد نیاز تحقیق، از سامانه جامع اطلاع‌رسانی ناشران موسوم به کدال، سایت شرکت مدیریت فناوری بورس تهران، سایت بورس اوراق بهادار تهران و نرم‌افزار ره‌آورد نوین استفاده شده است. برای انجام محاسبات مرتبط با مدل‌ها و انتخاب ویژگی از نرم‌افزار MATLAB2013b استفاده شده و پس آزمایی، تشکیل و محاسبه بازده پرتفوی‌ها با استفاده از نرم‌افزار EXCEL2013 انجام شده است. محاسبات مرتبط با آزمون مقایسه زوجی t در نرم‌افزار IBM SPSS Statistics22 انجام شده است.

۵- متغیرهای پژوهش و نحوه اندازه‌گیری آن‌ها

با توجه به پیچیدگی‌های ذاتی مدل کردن بازدهی و تعدد متغیرهای تأثیرگذار بر آن در این پژوهش سعی شد تا حد امکان همه‌ی متغیرهای قابل اندازه‌گیری تأثیرگذار بر بازدهی با توجه به تحقیقات پیشین در نظر گرفته شود تا مدل‌ها بتوانند با دقت بیشتری به پیش‌بینی طبقه بازدهی بپردازند. در نتیجه ۴۰ متغیر که در جدول زیر آورده شده است بدین منظور محاسبه شد.

جدول۱- متغیرهای پژوهش

نسبت بار مالی وام	گردش سرمایه جاری	نسبت جاری	سود خالص به فروش
رتبه نقد شوندگی	گردش دارایی‌های ثابت	نسبت آنی	سود ناخالص به فروش
EPS واقعی	گردش مجموع دارایی‌ها	نسبت نقدینگی	سود عملیاتی
درصد رشد EPS	نسبت بدهی به مجموع دارایی‌ها	نسبت دارایی‌های جاری	سود ناویژه به فروش
درصد رشد جمع درآمدها	نسبت بدهی به ارزش ویژه	نسبت کفایت نقدی	سود به سود ناویژه
درصد رشد حاشیه سود	نسبت دارایی ثابت به ارزش ویژه	نسبت گردش نقدی	بازده دارایی‌ها ROA
درصد پوشش EPS	نسبت بدهی بلندمدت به ارزش ویژه	سرمایه در گردش خالص	بازدهی سرمایه ROE
امتیاز کل تعدیلی	نسبت بدهی جاری به ارزش ویژه	گردش موجودی کالا	بازده سرمایه در گردش
ضریب β	نسبت مالکانه	دوره وصول مطالبات	بازده دارایی ثابت
نسبت شارپ	نسبت پوشش بدهی	نسبت موجودی کالا به سرمایه در گردش	سنجش سودمندی وام

برای محاسبه میانگین انحراف EPS متوسط انحراف EPS پیش‌بینی از EPS واقعی در ۳ سال مالی قبل لحاظ گردید.

برای محاسبه امتیاز کل تعدیلی از نظریه عوامل ذاتی رشد استفاده شد بر اساس نظریه عوامل ذاتی رشد بیش از ۸۰٪ از عواملی که باعث افزایش یا کاهش سودآوری سهام شرکت‌ها می‌شوند تحت تأثیر عواملی نظیر رشد EPS، رشد درآمدها و رشد حاشیه سود قرار دارند، بنابراین شرکت‌هایی که امتیاز کل بیشتری دارند بر اساس عوامل ذاتی رشد بهترین هستند. لازم به ذکر است که ضرایب یک در نظر گرفته شده‌اند.

(۱۱)

$$\frac{\left(\left(\text{ضریب انحراف EPS}\times\text{میانگین انحراف EPS}\right)+\left(\text{ضریب پوشش EPS}\times\text{درصد پوشش EPS}+1\right)+\left(1\times\text{امتیاز کل}\right)\right)}{\text{ضریب انحراف EPS}+\text{ضریب پوشش EPS}+1}\times 100$$

برای محاسبه نسبت شارپ و بتا بازدهی بدون ریسک ۲۰٪ در نظر گرفته شده است همچنین برای محاسبه بازدهی مورد انتظار که میانگین ساده بازدهی‌ها است مقاطع زمانی روزانه در نظر گرفته شده‌اند.

$$\frac{\text{بازدهی بدون ریسک}-\text{بازدهی انتظاری}}{\sigma} \qquad (۱۲)$$

ضریب یا عامل بتا وسیله‌ای برای ارزیابی عملکرد یک سهم خاص یا گروهی از سهام در جریان حرکت کلی بازار است که به صورت زیر محاسبه شد.

$$\frac{COV\left(\text{بازدهی بازار}\times\text{بازدهی سهم}\right)}{Var\left(\text{بازدهی بازار}\right)} \qquad (۱۳)$$

رتبه نقد شوندگی عددی است که رتبه نقد شوندگی یک سهم در بازار را نشان می‌دهد و با استفاده از فرمول زیر محاسبه شد.

(۱۴)

$$\text{فرمول رتبه}: \frac{1}{\frac{1}{\text{متوسط ارزش روز}}+\frac{1}{\text{حجم معاملات}}+\frac{1}{\text{تعداد سهام معامله شده}}+\frac{1}{\text{تعداد روز معامله شده}}+\frac{1}{\text{تعداد دفعات معامله}}+\frac{1}{\text{تعداد خریداران}}}$$

برای محاسبه بازدهی از فرمول زیر استفاده کردیم

(۱۵)

$$\text{فرمول درصد بازدهی با احتساب آورده}: \frac{\text{سهام جایزه}+\text{حق تقدم}+DPS+(\text{قیمت روز}-\text{قیمت پایه})}{\text{قیمت پایه}+(1000\times\text{(درصد افزایش سرمایه از محل آورده)})}\times 100$$

۶- فرضیه پژوهش

استفاده از مدل تحلیل ممیز قطری درجه دوم باعث بهبود بازدهی سرمایه‌گذاری می‌شود.

۷-یافته‌های پژوهش

در این پژوهش طبقه بازدهی هر سهم در هر سال با استفاده از اطلاعات متغیرهای سال قبل آن در حالتی که تمام متغیرها استفاده شدند و هنگامی که این متغیرها با استفاده از انتخاب ویژگی‌ها فیلتر شده بودند پیش‌بینی شد. نتایج پیش‌بینی با طبقه واقعی بازدهی انطباق داده و درصد پیش‌بینی درست طبقه بازدهی به عنوان دقت مدل‌ها در نظر گرفته شد.

درصد پیش‌بینی درست گروه‌ها توسط مدل ماشین بردار پشتیبان و بدون استفاده از انتخاب ویژگی در جدول زیر آورده شده است.

جدول۲- دقت مدل بردار پشتیبان بدون استفاده از انتخاب ویژگی

تعداد متغیرها	میانگین	کم بازده	پربازده	سال
40	43%	60%	26%	88
40	60%	46%	73%	89
40	50%	٪53	74%	90
40	47%	60%	33%	91

در جدول زیر درصد پیش‌بینی درست گروه‌ها با استفاده از انتخاب ویژگی را به همراه تعداد متغیرهای انتخاب شده مشاهده می‌کنید.

جدول ۳- دقت مدل بردار پشتیبان با استفاده از انتخاب ویژگی

تعداد متغیرها	میانگین	کم بازده	پربازده	سال
4	67%	53%	80%	88
21	60%	40%	80%	89
16	60%	60%	60%	90
15	57%	60%	53%	91

همان‌طور که در جداول فوق مشاهده می‌کنید صحت مدل ماشین بردار پشتیبان با استفاده از انتخاب ویژگی در تمامی موارد برابر یا بهتر از حالت بدون انتخاب ویژگی بوده است.

در جدول ذیل به اطلاعات دقت مدل تحلیل ممیز در پیش‌بینی صحیح طبقه بازدهی بدون استفاده از انتخاب ویژگی پرداخته شده است.

جدول ۴- دقت مدل تحلیل ممیز بدون استفاده از انتخاب ویژگی

سال	پربازده	کم بازده	میانگین	تعداد متغیرها
88	20%	60%	40%	40
89	60%	33%	47%	40
90	73%	53%	63%	40
91	60%	47%	53%	40

صحت این مدل در حالتی که متغیرهای ورودی توسط انتخاب ویژگی فیلتر شدند به شرح جدول ذیل است.

جدول ۵: دقت تحلیل ممیز با استفاده از انتخاب ویژگی

سال	پربازده	کم بازده	میانگین	تعداد متغیرها
88	33%	73%	53%	12
89	67%	40%	54%	12
90	60%	40%	50%	5
91	53%	67%	60%	6

برای آزمودن فرضیه کسب بازدهی بیشتر با استفاده از مدل‌های تحلیل ممیز و ماشین بردار پشتیبان سبد سهام را برای مدل‌ها با استفاده از انتخاب ویژگی و بدون استفاده از انتخاب ویژگی با وزن برابر برای هر سهم محاسبه کردیم.

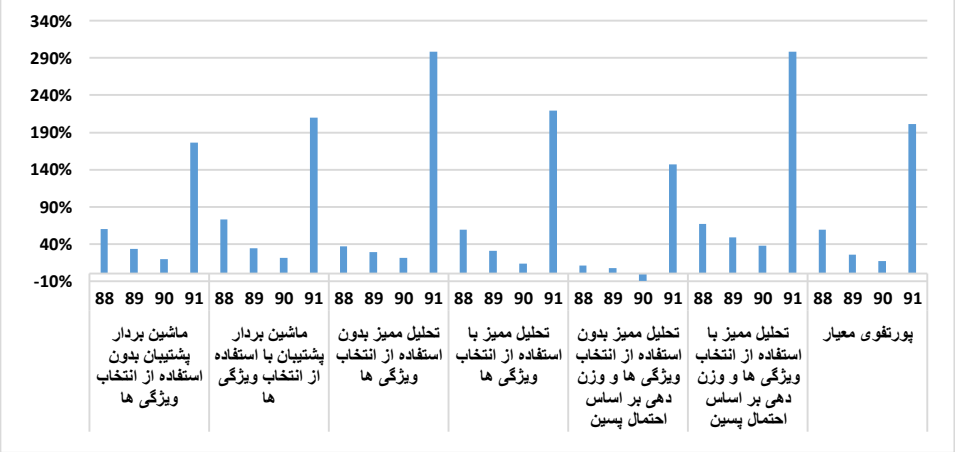
برای تحلیل ممیز با توجه به امکان محاسبه احتمال پسین، بازدهی سبد هنگامی که وزن‌ها بر اساس این احتمال داده شده باشند نیز محاسبه شده که نتایج آن را در جدول ذیل آورده می‌شود. برای محاسبه بازده کل از رابطه زیر استفاده شده است.

$$R_{total}=(1+R_{88})\times(1+R_{89})\times(1+R_{90})\times(1+R_{91})-1$$

جدول ۶- بازدهی مدل‌ها

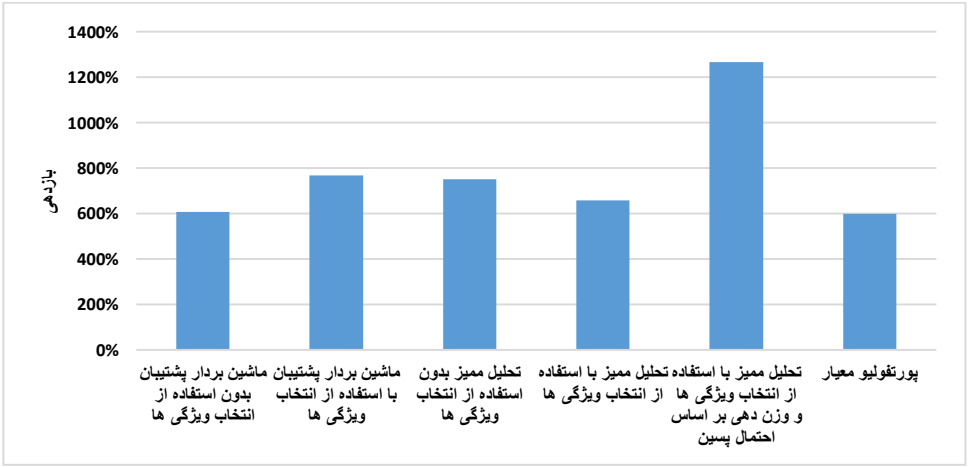
مدل	سال	بازدهی
ماشین بردار پشتیبان بدون استفاده از انتخاب ویژگی	88	60%
	89	33%
	90	20%
	91	176%
	کل	607%
ماشین بردار پشتیبان با استفاده از انتخاب ویژگی	88	73%
	89	34%
	90	21%
	91	210%
	کل	769%
تحلیل ممیز بدون استفاده از انتخاب ویژگی	88	37%
	89	29%
	90	21%
	91	298%
	کل	751%
تحلیل ممیز با استفاده از انتخاب ویژگی	88	59%
	89	31%
	90	14%
	91	219%
	کل	657%
تحلیل ممیز با استفاده از انتخاب ویژگی و وزن دهی بر اساس احتمال پسین	88	67%
	89	49%
	90	38%
	91	298%
	کل	1266%
پرتفوی معیار	88	59%
	89	26%
	90	۱۶٪
	91	201%
	کل	599%

به منظور مقایسه بازدهی در سال‌های ۸۸-۹۱ نمودار ذیل آورده شده است.



نمودار ۱- بازدهی مدل‌ها در هر سال

در نمودار زیر بازدهی کل کسب شده توسط مدل‌ها و همچنین بازدهی پرتفوی معیار که بازدهی ۴۹ سهم منتخب با وزن یکسان است، آورده شده است.



نمودار ۲- بازدهی کل پرتفوی مدل‌ها

همان‌طور که در نمودار فوق مشاهده می‌کنید بازدهی کل پرتفوی در مدل‌های مورد بررسی در همه‌ی موارد از بازدهی پرتفوی معیار بیشتر است بنابراین فرضیه کسب بازدهی بیشتر با استفاده از مدل‌های تحلیل ممیز و ماشین بردار پشتیبان تأیید می‌شود

به منظور بررسی وجود تفاوت معنادار دو پرتفوی معیار و تحلیل ممیز با استفاده از انتخاب ویژگی و وزن دهی بر اساس احتمال پسین را انتخاب می‌کنیم.

H_o :بازدهی پرتفوی معیار ≤ بازدهی پرتفوی تحلیل ممیز با استفاده از انتخاب ویژگی و وزن دهی بر اساس احتمال پسین

H_a بازدهی پرتفوی معیار < بازدهی پرتفوی تحلیل ممیز با استفاده از انتخاب ویژگی و وزن : دهی بر اساس احتمال پسین

از آزمون مقایسه زوجی t برای آزمودن فرضیه وجود تفاوت معنادار بین بازدهی پرتفوی‌ها استفاده می‌کنیم. آزمون را برای بازدهی مدل‌ها طی سال‌های ۸۸-۹۱ انجام داده و سطح معناداری را ۰/۰۵ تعیین کردیم که نتایج حاصله به شرح جدول زیر است.

جدول۷- نتایج آزمون بازدهی پرتفوی‌ها

سطح معناداری	آماره
0/029	3/957

مقدار معناداری در آزمون انجام شده برابر ۰/۰۲۹بوده که از مقدار ۰/۰۵ سطح معناداری تعیین شده کمتر است بنابراین فرض H_o رد شده و فرض تحقیق پذیرفته می‌شود پس تفاوت معناداری بین بازدهی پرتفوی معیار و پرتفوی تحلیل ممیز با استفاده از انتخاب ویژگی و وزن دهی بر اساس احتمال پسین وجود داشته است. این تفاوت در سطح اطمینان ۹۵٪ تصادفی نیست و بازدهی این پرتفوی از بازدهی معیار به طور معناداری بیشتر است. که این امر با توجه به این نکته که پرتفوی تشکیل شده با استفاده از مدل تحلیل ممیز بازدهی کل تقریباً دو برابری نسبت به بازدهی پرتفوی معیار داشته و متغیرها قبل از ورود به مدل به منظور حصول نتیجه بهتر فیلتر شده‌اند و وزن دهی بر اساس احتمال درست بودن پیش‌بینی (احتمال پسین) برای تشکیل پرتفوی صورت گرفته دور از انتظار نیست و می‌توان از این مدل برای تشکیل پرتفویی با بازدهی بیشتر از بازدهی پرتفوی معیار استفاده کرد.

۸- نتیجه‌گیری و بحث

در این مقاله راهکاری جهت پیش‌بینی طبقه بازدهی هر سهم و وزن دهی آن برای تشکیل پرتفوی بر اساس تحلیل ممیز قطری درجه دوم ارائه شد. در این مقاله برای بهبود نتایج با استفاده از انتخاب ویژگی متوالی ابتدا متغیرها برای پیش‌بینی طبقه بازدهی انتخاب شده و با استفاده از دو مدل تحلیل ممیز قطری درجه دوم و ماشین بردار پشتیبان با کرنل درجه دوم طبقه بازدهی هر سهم پیش‌بینی شد. همه‌ی پرتفوی‌های تشکیل شده بر اساس پیش‌بینی طبقه بازدهی سهم‌ها توسط دو مدل ماشین بردار پشتیبان و تحلیل ممیز قطری درجه دوم بازدهی بالاتری نسبت به بازدهی پرتفوی معیار کسب کردند همچنین برای وزن دهی پرتفوی تشکیل شده بر اساس پیش‌بینی طبقه پربازده توسط مدل تحلیل ممیز درجه دوم از احتمال پسین استفاده شد که نتایج حاصله رضایت‌بخش بوده و دلالت بر وجود تفاوت معنادار بین بازدهی این پرتفوی و پرتفوی معیار دارند و توانستیم به بازدهی کل تقریباً دو برابری نسبت به بازدهی پرتفوی معیار دست یابیم بنابراین می‌توان نتیجه گرفت استراتژی انتخاب سهم‌ها برای خرید بر اساس تحلیل ممیز قطری درجه دوم و استفاده از احتمال پسین حاصل از این مدل برای وزن دهی استراتژی مناسبی برای تشکیل پرتفوی است. همچنان که افزایش بازدهی پرتفوی با استفاده از الگوریتم‌های هوش مصنوعی طبقه‌بندی کننده استفاده شده در این مقاله را در مقالات چیا هوی هانگ[۱۲] و همکاران (۲۰۰۸) و پالانیسوامی[۱۳] و همکاران (۲۰۰۱) که در گذشته به آن‌ها پرداختیم نیز می‌توان مشاهده کرد که دلالت بر سودمند بودن استفاده از این الگوریتم‌ها برای تشکیل پرتفوی دارند. همچنین، می‌توان در راستای پیشبرد تحقیقات آتی در این حوزه پیشنهادهای زیر را ارائه نمود:

- ✓ کم کردن افق زمانی پیش‌بینی و استفاده از گزارش‌های سه ماهه شرکت‌ها برای محاسبه متغیرهایی که مورد نیاز مدل است.
- ✓ ترکیب سیستم معاملاتی ارائه شده در این تحقیق با مدل‌های معاملات پرفراوانی[۱۴] به منظور استفاده از نوسانات روزانه قیمت سهام جهت افزایش بازدهی
- ✓ متغیرهایی که برای پیش‌بینی بازدهی در نظر گرفتیم از اطلاعات صورت‌های مالی محاسبه شدند در حالی که اخبار می‌توانند نقش بسزایی در بازدهی شرکت ایفا کنند بنابراین به نظر می‌رسد وارد کردن متغیرهایی در مدل که توسط اخبار بصورت پویا به روز می‌شود در کسب بازدهی بیشتر مؤثر است.

فهرست منابع

* احمدی مسعود، (۱۳۹۱)، روش تحقیق اصول، مفاهیم با رویکرد پایان نامه نویسی. چاپ اول. ساری پژوهش های فرهنگی.
* پارسائیان ناصر ، شمس سمیرا . (۱۳۹۱)، مقایسه عملکرد مدل فاما و فرنچ و شبکه های عصبی مصنوعی. مجله مهندسی مالی و مدیریت اوراق بهادار .
* داسل استوارت ، نورویج پیتر، (۱۳۹۲)، هوش مصنوعی رهیافتی نوین. رهنمون رامین ، چاپ پنجم، تهران: انتشارات ناقوس.
* راعی رضا ، فلاح‌پور سعید. کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از نسبت‌های مالی. بررسی‌های حسابداری و حسابرسی. شماره ۵۳: صفحه ۱۷-۳۸.
* ساروخانی باقر، (۱۳۸۲)، روشهای تحقیق در علوم اجتماعی چاپ اول. تهران :پژوهشگاه علوم انسانی و مطالعات اقتصادی.
* قدرت اله طالب نیا، مریم فتحی، (۱۳۸۹)، ارزیابی مقایسه ای انتخاب پرتفوی بهینه سهام در بورس اوراق بهادار تهران از طریق مدلهای مارکویتز و ارزش در معرض خطر. شماره۶, مطالعات مالی.
* نجفی امیرعباس ، موشخیان سیامک، (۱۳۹۳)، مدل سازی و ارائه راه حل بهینه برای بهینه سازی سبد سرمایه گذاری چند دوره ای با الگوریتم ژنتیک. مجله مهندسی مالی و مدیریت اوراق بهادار.
* هان ژیان، کمبر میشلین، پی ژان. (۱۳۹۳)، داده‌کاوی مفاهیم و تکنیک‌ها. اسماعیلی مهدی, چاپ اول، تهران: انتشارات نیاز دانش.
* ویلیام اف شارپ، گوردون جی. الکساندر، جفری وی بیلی. (۱۳۸۹)، مدیریت سرمایه گذاری. [مترجم] مجید جعفری شریعت پناهی، ابولفضل جعفری. تهران: اتحاد.
* Burges, Christopher J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. 2, pp. 121-167. Retrieved 2015
* Chia Hui Huang, Wen Ching Liou,Berlin Wu. (2008). A Decision Support System for Stock Investment of Listed Companies in Taiwan. Biomedical Soft Computing and Human Sciences, 13, 31-36.
* Dagher, I. (2008). Quadratic kernel-free non- linear support vector machine. Journal of Global Optimization, 41(1), 15-30.
* Elisa Siqueira, Thiago Otuki and Newton da Costa Jr. (2012). Stock Return and Fundamental Variables:A Discriminant Analysis Approach. Applied Mathematical Sciences, 5719-5733.

* Fan, A., & Palaniswami, M. (2001). Stock selection using support vector machines. In Neural Networks, 2001. Proceedings. IJCNN'01. International Joint Conference on (Vol. 3, pp. 1793-1798). IEEE.
* Lee JW, Lee JB, Park M, Song SH. (2005). An extensive comparison of recent classification tools applied to microarray data. pp. 869–885.
* M. Dash, H. Liu. (1997). Feature Selection for Classification. Intelligent Data Analysis, 131-156.
* Quadratic Discriminant Analysis. (2014). (The Pennsylvania State University) Retrieved 2015, from onlinecourses.science.psu.edu: https://onlinecourses.science.psu.edu/stat557/node/43
* Ramon Huertaa, Fernando Corbachob,and Charles Elkan. (2013). Nonlinear support vector machines can systematically identify stocks with high and low future returns. Algorithmic Finance, 2, 45-58.
* Rencher, A. c. (2002). methods of multivariate analysis. CANADA: A JOHN WILEY & SONS, INC, PUBLICATION.
* Shieh GS, Jiang YC, Shih YS. (2006). Comparison of support vector machines to other classifiers using gene expression data. Communications in Statistics. pp. 241–256.
* webpages: iust.ac.ir/parsa/Handouts/SVM.doc. (n.d.).
* Ye J,Li T,Xiong T, Janardan R. (2004). Using uncorrelated discriminant analysis for tissue classification with gene expression data. pp. 90-181.

یادداشت‌ها

[1] Siqueira
[2] cross-sectional
[3] Stepwise
[4] Box's M
[5] Palaniswami
[6] Chia Hui Huang
[7] Ramon Huerta
[8] decision boundaries
[9] Non-convex optimization problem
[10] quadratic programming (QP) optimization
[11] Sequential Backward Selection
[12] Chia Hui Huang
[13] Palaniswami
[14] High Frequency Trading