فصلنامه مهندسی مالی و مدیریت اوراق بهادار
دوره چهارده، شماره پنجاه و پنج، تابستان ۱۴۰۲
نوع مقاله : علمی پژوهشی
صفحات : ۲۷۰-۲۵۲

مدل بهینه‌سازی سبدسرمایه‌گذاری مبتنی بر پیش‌بینی با استفاده از رگرسیون بردار پشتیبان

محمدامین منادی[۱]
امیرعباس نجفی[۲]

تاریخ دریافت مقاله : ۱۴۰۱/۰۱/۲۴ تاریخ پذیرش مقاله : ۱۴۰۱/۰۶/۰۲

چکیده

هدف از بهینه‌سازی سبد سرمایه‌گذاری، انتخاب ترکیبی بهینه از دارایی‌های مالی است که می‌بایست راهنمای سرمایه‌گذاران برای دستیابی به بالاترین بازده در برابر کمترین ریسک ممکن باشد. از سوی دیگر، یکی از عوامل کلیدی در تصمیم‌گیری‌های سبد سرمایه‌گذاری مربوط به پیش بینی قیمت سهام است. برای این کار بطور متداول از مدل‌های کلاسیک غیرخطی ریاضی و هوشمند مانند رگرسیون استفاده می شود. در مطالعه حاضر برای کاهش خطاهای پیش‌بینی، از مدل غیرخطی رگرسیون بردار پشتیبان با خروجی‌های متعدد استفاده شده است. برای نشان دادن کارایی مدل پیشنهادی از داده‌های شرکت‌های شاخص S&P500 در دوره زمانی ۲۰۱۶/۰۹/۱۲ تا ۲۰۲۱/۰۸/۰۲ استفاده شده است. نتایج نشان می‌دهد که انتخاب سبد سهام مبتنی بر پیش‌بینی به کمک رگرسیون بردار پشتیبان با خروجی چندگانه به دلیل در نظر گرفتن روابط بین خروجی‌ها به صورت همزمان از نظر معیار شارپ، عملکرد بهتری نسبت به انتخاب سبد سرمایه‌گذاری بر اساس پیش‌بینی با استفاده از روش رگرسیون دارد.

کلمات کلیدی

انتخاب سبد، پیش‌بینی قیمت سهام، رگرسیون بردار پشتیبان، خروجی‌چندگانه

۱-گروه مهندسی مالی، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران.
mahdimonadiii@gmail.com
۲-گروه مهندسی مالی، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران. (نویسنده مسئول)
aanajafi@kntu.ac.ir

مقدمه

با رشد و توسعه اقتصاد جهانی و رقابتی شدن آن، تصمیم‌گیری در رابطه با تخصیص بهینه منابع بیش از پیش اهمیت پیدا نموده است. بورس اوراق بهادار علاوه بر آن که آئینه تمام‌نمای وضعیت اقتصادی کشورها بوده، مکانی برای سرمایه‌گذاران است که بتوانند منابع خود را در آن سرمایه‌گذاری کنند. از آنجا که کاهش ریسک و افزایش بازده سرمایه‌گذاری در بازار سهام همواره مهم‌ترین و اساسی‌ترین دغدغه سرمایه‌گذاران بوده، در ادبیات مالی، سرمایه‌گذاری در سبد سهام به منظورکاهش ریسک پیشنهاد گردیده است (زمانی و همکاران، ۱۳۹۳). از سویی دیگر، آگاهی از تصمیمات و دستاوردهای مربوط به آینده، همواره با عدم اطمینان و ابهام روبه رو بوده و کسانی در رقابت به پیروزی رسیده‌اند که بتوانند آینده را پیش‌بینی کنند و حداقل اطلاعاتی در خصوص تصمیم‌گیری داشته باشند (حبیب زاده و ایزدپور، ۱۳۹۹). هم‌چنین انتخاب سبد سهام به عنوان یک برنامه سرمایه‌گذاری علاوه‌بر آن که باید ماحصل گذشته سهام را در بربگیرد، باید پتانسیل آتی سهام را نیز در نظر بگیرد، لذا اهمیت پیش‌بینی قیمت سهام بیش از پیش آشکار می‌گردد (زمانی و همکاران، ۱۳۹۳). به علت اینکه بازار سهام اساساً پویا، پیچیده، غیرخطی و غیرپارامتریک است و پیش‌بینی قیمت سهام از تعداد بسیاری از عوامل و متغیرهای اقتصادی و غیراقتصادی متأثر بوده، همواره از گذشته تاکنون انتخاب بهترین و کارآمدترین مدل به منظور پیش‌بینی ماهیت غیرخطی داده‌های سهام امری ضروری، دشوار و چالش برانگیز بوده است (باجلان و همکاران، ۱۳۹۶). علاوه بر آن، پژوهش‌های گذشته نشان می‌دهند که استفاده از روش‌های پیش‌بینی سنتی، دارای خطای بالاتر و عملکرد ضعیف‌تر در مقایسه با روش‌های جدیدتر و مدل‌های غیرخطی می‌باشند (فلاح‌پور و همکاران، ۱۳۹۲). از این رو، صاحبان علوم حوزه‌ای را با نام یادگیری ماشینی[1] (تئوری یادگیری[2]) مورد استفاده قرار دادند. در این پژوهش نیز از یکی از روش‌های هوش مصنوعی با عنوان رگرسیون بردار پشتیبان با خروجی چندگانه به منظور پیش‌بینی قیمت سهام به کار گرفته شده است و در ادامه با داده‌های پیش‌بینی شده به بهینه‌سازی سبد سهام پرداخته شده است. در مسئله رگرسیون با خروجی چندگانه که عبارت است از آموزش چند تابع به صورت همزمان، هر تابع به یک متغیر خروجی تعلق دارد و زمانی که بین متغیرهای خروجی نوعی وابستگی برقرار باشد، نمی‌توان از چندین دسته‌بندی با خروجی یکتا استفاده نمود و روابط بین متغیرهای خروجی را نادیده گرفت (ژانگ و همکاران[3]، ۲۰۱۲).

ادبیات و پیشینه تحقیق

در این بخش، پژوهش‌ها و تحقیقات پیشین در خصوص بهینه‌سازی و انتخاب سبد سهام مبتنی بر پیش‌بینی مورد بررسی قرار گرفته است. فریتس و همکاران[4] (۲۰۰۹) با استفاده از شبکه عصبی، برای

ارزیابی نرمال بودن خطای‌های پیش‌بینی، به پیش‌بینی بازده قیمت ۸۲ سهام شاخص بورس برزیل و بهینه‌سازی سبد سهام مربوطه پرداختند. نتایج آن‌ها نشان داد که دستیابی به خطاهای پیش‌بینی نرمال با سری‌های زمانی غیرعادی امکان پذیر است. همچنین در مقایسه با مدل میانگین-واریانس مارکویتز، روش مذکور بازده بالاتری با ریسک یکسان دارا می‌باشد. پینتو و همکاران[۵] (۲۰۱۱) در پژوهشی با پیش‌بینی بازده سهام، مدل مارکویتز را مورد بررسی قرار دادند. آن‌ها در مدل پیش‌بینی خود از خود رگرسیونی برون‌زا[۶] استفاده کردند و پس از آن به تعیین اوزان بهینه پرداختند. نتایج آن‌ها نشان داد که این رویکرد دارای بازده بالاتر با ریسک یکسان نسبت به مدل ماکویتز می‌باشد. هائو و همکاران[۷] (۲۰۱۳) در پژوهش خود، از طریق رگرسیون بردار پشتیبان و شبکه عصبی به پیش‌بینی سهام پرداخته، به گونه‌ای که خطای پیش‌بینی در مدل میانگین-واریانس مارکویتز حداقل گردد. برای آزمایش این رویکرد، مجموعه‌ای از سهام شاخص شانگهای را بکار گرفتند. نتایج آن‌ها نشان داد که بهینه‌سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از رگرسیون بردار پشتیبان علاوه بر آن که دارای خطای پیش‌بینی کمتری است، دارای بازده بالاتری با ریسک یکسان نسبت به روش بهینه‌سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از شبکه عصبی می‌باشد.

ژانگ و همکاران (۲۰۱۲) برای اولین بار به پیش بینی قیمت سهام مبتنی بر رگرسیون بردار پشتیبان حداقل مربعات به صورت چندگانه در پژوهش خود پرداختند. آن‌ها در پژوهش خود از یک روش جدید برای ساخت یک مدل چند خروجی به صورت مستقیم استفاده کردند. نتایج آن‌ها نشان داد که این روش در مقایسه با مدل کلاسیک خود دارای عملکرد بهتری می‌باشد چرا که همبستگی بین خروجی‌ها در نظر گرفته شده است. آن‌ها مسئله خود را به کمک داده‌های ساختگی بر روی شاخص بورس برزیل انجام دادند. ژو و همکاران[۸] (۲۰۱۳) در پژوهش خود روش دیگری را برای توسعه رگرسیون بردار پشتیبان حداقل مربعات به منظور خروجی چندگانه ارائه دادند. علاوه بر آن، ایشان الگوریتم یادگیری جدیدی به منظور افزایش سرعت یادگیری و کاهش محاسبات اراده دادند و نتایج پژوهش آنان حاکی از مؤثر بودن این رویکرد بوده است. میشرا و همکاران[۹] (۲۰۱۶) یک مدل میانگین-واریانس مبتنی بر پیش‌بینی به منظور حل مسائل بهینه‌سازی سبد مقید ارائه دادند. آنان در پژوهش خود ابتدا از شبکه‌های عصبی به منظور پیش‌بینی بازده استفاده کردند و در رابطه با بهینه‌سازی سبد از الگوریتم‌های خود تکاملی چند هدفه به خصوص الگوریتم هوش ازدحامی بهره می‌برند. نتایج آنان نشان داد که بهینه‌سازی سبد با استفاده از الگوریتم ازدحام ذرات و مبتنی بر پیش‌بینی بازده دارای عملکرد بهتری نسبت به مدل مارکویتز و دیگر الگوریتم‌های بهینه‌سازی می‌باشد. چن و همکاران[۱۰] (۲۰۱۹) اذعان داشتند که موفقیت در انتخاب

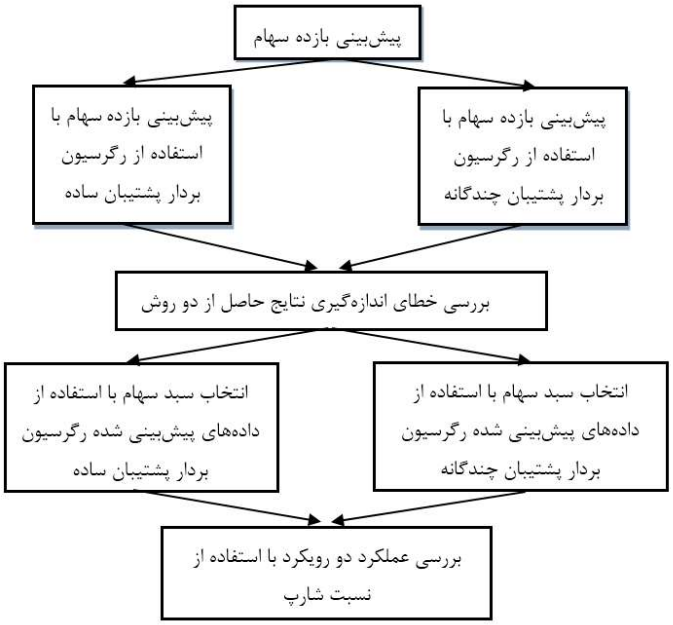
سهام مشروط بر عملکرد آتی بازار سهام است. آنان برای این ادعای خود به پیش‌بینی قیمت سهام چین پرداختند که مرحله اول شامل پیش‌بینی قیمت سهام با استفاده از روش محاسباتی هوشمند از قبیل یادگیری ماشین با قدرت یادگیری بالا و محاسبات سریع و مرحله دوم شامل یک ترکیب خطی از داده‌های پیش‌بینی شده و داده‌های بنیادین می‌باشد. نتایج این محققان نشان می‌دهد عملکرد این روش نسبت به مدل‌های سنتی بدون پیش‌بینی بهتر می‌باشد. یانگ و همکاران[11] (۲۰۲۱) به بهینه‌سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از یادگیری ماشین بر روی سهام شانگهای پرداختند. در این پژوهش در مرحله اول، یک مدل هیبریدی[12] و الگوریتم کرم شب تاب بهبود یافته[13] برای پیش بینی قیمت سهام برای دوره بعدی پیشنهاد شده است. به عبارتی پارامترهای مدل هیبریدی توسط الگوریتم کرم شب تاب بهبود یافته برآورد شده و سپس در مرحله دوم با استفاده از نظریه میانگین-واریانس به بهینه‌سازی سبد سهام به کمک داده‌های پیش‌بینی شده، پرداخته شده است. نتایج این پژوهش نشان می‌دهد که روش پیشنهادی دارای عملکرد بهتری نسبت به روش‌های کلاسیک بدون در نظر گرفتن پیش‌بینی می‌باشد. مآ و همکاران[14] (۲۰۲۱) به منظور پیش‌بینی از الگوریتم‌های یادگیری ماشین و یادگیری عمیق و به منظور بهینه‌سازی از تئوری میانگین-واریانس بهره جستند. آنان در یادگیری ماشین از رگرسیون بردار پشتیبان و جنگل تصادفی[15] و در یادگیری عمیق از روش‌های DMLP[16]، LSTM[17] و CNN[18]استفاده نمودند و نتایج آنان نشان داد که بهینه‌سازی سبد به کمک تئوری میانگین-واریانس و پیش‌بینی بر اساس الگوریتم جنگل تصادفی دارای عملکرد بهتری می‌باشد.

زمانی و همکاران (۱۳۹۲) به منظور پیش‌بینی قیمت سهام از طریق شبکه‌های عصبی فازی پرداختند و با استفاده از پیش‌بینی‌های بدست آمده، مدل‌های بهینه‌سازی سبد سهام مبتنی بر معیارهایی مانند میانگین، واریانس و چولگی سبد سهام را با استفاده از الگوریتم ژنتیک حل نمودند. نتایج پژوهش آنان نشان داد که روش آنان در مقایسه با روش‌های سنتی و شاخص بازار، بازدهی بیشتری را فراهم می‌آورد. خواجه‌زاده و همکاران (۱۳۹۹) از الگوریتم رگرسیون لبه به منظور پیش‌بینی قیمت سهام پرداختند و با استفاده از الگوریتم فراابتکاری فرهنگی به انتخاب بهینه سبد به کمک داده‌های پیش‌بینی شده اقدام کردند. نتایج آنان نشان می‌دهد که الگوریتم فراابتکاری فرهنگی توانایی ایجاد سبد سهام بهینه را دارد.

با توجه به ادبیات پژوهش می‌توان پی برد که تاکنون در زمینه انتخاب بهینه سبد سهام مبتنی بر پیش‌بینی قیمت سهام به کمک رگرسیون بردار پشتیبان خروجی چندگانه تحقیقی صورت نگرفته است و از آن می‌توان به عنوان یک خلأ تحقیقاتی نام برد. بر همین اساس پژوهش جاری از ابزار مذکور در زمینه پیش‌بینی قیمت سهام استفاده می‌نماید و در انتها به بهینه‌سازی پرتفوی می‌پردازد.

روش‌شناسی پژوهش

در این پژوهش با استفاده از رگرسیون بردار پشتیبان، مدلی برای پیش‌بینی قیمت سهام ارائه می‌شود تا روابط بین متغیرهای خروجی را در نظر گرفته و خطای پیش‌بینی را کاهش دهد. به منظور بررسی کارایی روش پیشنهادی نیز، بازده پرتفوی حاصل از مدل پیشنهادی را با نتایج مدل‌های سنتی مقایسه خواهیم نمود. شکل ۱ روش‌شناسی پژوهش حاضر را نشان می‌دهد.



شکل ۱- چارچوب روش تحقیق

پرسش‌ها و سؤالات اصلی تحقیق

پژوهش جاری با طرح ایده تشریح داده شده، سعی در یافتن پاسخ به برخی سؤالات مرتبط در این زمینه دارد.

- مدل‌سازی یک ماشین برای رگرسیون خروجی‌ها به طور همزمان به چه طریق ممکن است؟
- روش سنتی که از چندین رگرسیون بردار پشتیبان استفاده می‌کند و بصورت همزمان آنان را در نظر نمی‌گیرد دارای چه معایبی می‌باشد؟
- آیا انتخاب پرتفوی مبتنی بر این روش نسبت به روش سنتی دارای عملکرد بهتری می‌باشد؟

سرمایه‌گذاری

سرمایه‌گذاری اصطلاحی است با چندین معنی مرتبط در مدیریت بازرگانی، مالی و اقتصاد در رابطه با پس‌انداز یا به تعویق انداختن مصرف. سرمایه‌گذاری انتخابی است که فرد سرمایه‌گذار پس‌انداز خود را با امید سود به خطر می‌اندازد. به عبارتی، به جای ذخیره کالا یا معادل پول آن کالا سرمایه‌گذار ترجیح می‌دهد آن کالا ذخیره شده یا پول را به ازای سود یا قسمتی از سود قرض دهد. به طور کلی، سرمایه‌گذاری عبارت است از تبدیل نمودن وجوه مالی به چندین نوع دارایی که برای مدت زمانی نگهداری خواهند شد.

پرتفوی[۱۹]

در نظریه کلاسیک انتخاب سبد، این‌گونه تصور شده است که با تشکیل پرتفوی از سهام صنایع مختلف، به عبارتی تنوع بخشی، می‌توان ریسک پرتفوی را کاهش داد. در این پژوهش در مسئله بهینه‌سازی پرتفوی به جای استفاده از مدل مارکویتز و کمینه‌سازی واریانس بازده پیش‌بینی شده سهام از کمینه‌سازی خطای پیش‌بینی به عنوان تابع هدف استفاده شده است. به عبارت دیگر در این مدل اختلاف بین مقدار واقعی و مقدار پیش‌بینی شده به عنوان ریسک تعریف می‌شود مدل مذکور که همان پایه و اساس آماری مدل میانگین-واریانس مارکویتز را دارا می‌باشد، به صورت مدل ۱ تعریف می‌گردد (فریتس و همکاران، ۲۰۰۹).

$$
\begin{aligned}
&min \quad \hat{V} = \hat{\sigma}_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{ij} \\
&Subject\ to \quad \sum_{i=1}^{N} w_i \mu_i = R^* \\
&\sum_{i=1}^{N} w_i = 1 \\
&0 \le w_i \le 1 \qquad , i = 1, \dots, N \\
&\sigma_{ij} = \frac{1}{n-1}\sum_{t=1}^{n} \varepsilon_{it}\varepsilon_{jt} \\
&\varepsilon_t = r_t - \hat{r}_t
\end{aligned}
\qquad (۱)
$$

در این روابط داریم: N : تعداد سهام موجود

w_i : نسبت سرمایه گذاری مربوط به سهم i ام در سبد سرمایه‌گذاری μ_i : بازدهی سهم i ام

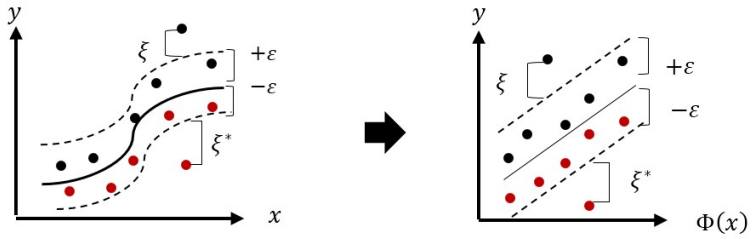
σ_{ij}: کوواریانس بین سهام، به عنوان معیار همراهی و حرکتی بین نرخ‌های بازدهی سهام مختلف

R^* : بازده مورد درخواست سبد

رگرسیون بردار پشتیبان[۲۰]

مسئله رگرسیون بردار پشتیبان، تعمیمی از مسئله دسته‌بندی می‌باشد. خروجی مسئله دسته‌بندی به صورت گسسته و متناهی است این در حالی است که خروجی رگسیون بردار پشتیبان به صورت پیوسته تعریف می‌شود. به عبارت دیگر رگرسیون بردار پشتیبان یک تابع چند متغیره پیوسته تخمین می‌زند. ماشین‌های بردار پشتیبان، مسائل طبقه‌بندی دوتایی را با فرموله کردن آن‌ها به صورت بهینه‌سازی محدب حل می‌کنند (دراکر و همکاران[۲۱]، ۱۹۹۷). این مسائل بهینه‌سازی مستلزم یافتن حداکثر حاشیه ابر صفحه جداساز می‌باشند، در حالی که تا حد امکان بسیاری از نقاط آموزشی را به طور صحیح طبقه‌بندی می‌کنند. این ابر صفحه‌ها به کمک بردارهای پشتیبان شناسایی می‌گردند. تعمیم‌پذیری بالا و محدود و پراکنده بودن جواب‌های این مدل، موجب توسعه و معرفی رگرسیون بردار پشتیبان گردید. تعمیم SVM به SVR با معرفی یک محدوده غیر حساس خطا[۲۲] در اطراف تابع که لوله یا مجرا خطا[۲۳] نامیده می‌شود، موفق بوده است. در این مدل هدف از بهینه‌سازی، یافتن لوله‌ای است که بهترین تخمین را از تابع پیوسته با در نظر گرفتن مقدار خطا و پیچیدگی مدل داشته باشد. در واقع SVR به صورت یک تابع محدب غیرحساس به خطا به عنوان مدل بهینه‌سازی تعریف می‌گردد و به دنبال باریک‌ترین لوله منحصر به فرد است که تعداد زیادی نمونه‌های آموزشی را در بر بگیرد. همانطور که در SVM ابر صفحه بهینه بر حسب بردارهای پشتیبان بدست می‌آید در SVR نیز مجرای بهینه بر حسب بردارهای پشتیبان برآورد می‌گردد (اواد و خانا[۲۴]، ۲۰۱۵). ایدهٔ اصلی رگرسیون بردار پشتیبان، مطابق شکل ۲ نگاشت داده‌ها در فضایی با ابعاد بالا از طریق نگاشت غیرخطی Φ و سپس رسم رگرسیون خطی در فضای ویژگی می‌باشد.

$$f(x) = w.\Phi(x) + b \qquad (۲)$$



شکل۲- رگرسیون بردار پشتیبان (تیسن و همکاران، ۲۰۰۲)

رگرسیون بردار پشتیبان، مسئله برآورد تابع را به صورت یک مسئله بهینه‌سازی طراحی می‌کند بطوریکه سعی در حداقل‌سازی خطای پیش‌بینی دارد. به فاصله بین خروجی‌های پیش‌بینی شده و خروجی‌های موردنظر، خطای پیش‌بینی می‌گویند. این مدل به صورت روابط ۳ تعریف می‌گردد.

$$\begin{aligned} &\min_{w,b} \frac{1}{2}\|w\|^2 \\ &s.t \ \Big(\big(w.\Phi(x_i)\big)+b\Big) - y_i \le \varepsilon, i = 1,2,\dots,l \\ &y_i - \Big(\big(w.\Phi(x_i)\big)+b\Big) \le \varepsilon, i = 1,2,\dots,l \\ &w \in R^n, b \in R \end{aligned} \qquad (۳)$$

فرض ضمنی در مسئله بهینه‌سازی بالا به این صورت است که نمونه‌های آزمایشی می‌توانند با دقت به وسیله تابع خطی در فضای ویژگی تقریب زده شوند. اگرچه در عمل، همیشه نویزها و عدم دقت‌هایی وجود دارد، بنابراین درک این مدل ایده آل دشوار خواهد بود. برای بهبود عملکرد این مدل رگرسیون حاشیه نرم بصورت روابط ۴ مدل می‌گردد.

$$\begin{aligned} &\min_{w,b,\xi_i,\xi_i^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i+\xi_i^*) \\ &s.t \ \Big(\big(w.\Phi(x_i)\big)+b\Big) - y_i \le \varepsilon + \xi_i, i = 1,2,\dots,l \\ &y_i - \Big(\big(w.\Phi(x_i)\big)+b\Big) \le \varepsilon + \xi_i^*, i = 1,2,\dots,l \xi_i \ge 0, i = 1,2,\dots,l \\ &\xi_i^{(*)} \ge 0, i = 1,2,\dots,l \\ &w \in R^n, b \in R \end{aligned} \qquad (۴)$$

ξ_i و ξ_i^* متغیرهای کمبود هستند و C متغیری است که به صورت دستی تنظیم می‌شود و رابطه‌ای بین یکنواختی و مقدار خطا برقرار می‌کند. هر چه مقدار C افزایش یابد مسئله وزن بیشتری به حداقل‌سازی مقدار خطا پیش‌بینی می‌دهد و به ماشین بردار با حاشیه سخت نزدیکتر می‌شود حاشیه بین طبقات کاهش می‌یابد و هر چه C کمتر باشد، مسئله وزن بیش‌تری را به حداکثرسازی حاشیه بین طبقات می‌دهد (لین و همکاران[۲۵]، ۲۰۰۷).

با اعمال شرایط KKT (برابر صفر قرار دادن حاصلضرب قیود در ضرایب لاگرانژ) مسئله دوگان مدل بالا بر اساس نگاشت به صورت مدل ۵ نوشته می‌شود:

$$\begin{aligned} &\max_{a,a^*} -\frac{1}{2}\sum_{ij}(a_i-a_i^*)\left(a_j-a_j^*\right)\Phi(x_i)'\Phi(x_i) - \varepsilon\sum_i(a_i-a_i^*) + y_i\sum_i(a_i-a_i^*) \\ &\sum_i(a_i-a_i^*) = 0, a, a^* \epsilon [0, C] \end{aligned} \qquad (۵)$$

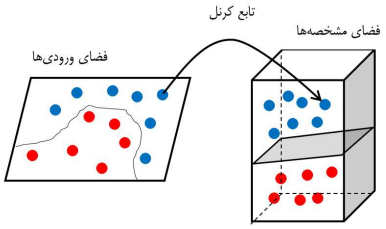
$$K(x_i, x) = \phi(x_i).\phi(x)$$

$$w = \sum_{i=1}^{l}(a_i - a_i^*)\phi(x_i)$$

$$f(x) = \sum_{i=1}^{l}(a_i - a_i^*)\ K(x_i, x) + b$$

a^* و a ضرایب لاگرانژ بوده و دارای مقداری غیرصفر و برابر C در خارج از مرز لوله هستند و همزمان نمی‌توانند غیر صفر باشند؛ چرا که بدین معنی است یک داده آموزشی هم در مرز بالایی و هم در مرز پایینی قرار دارد که غیرممکن است. بردارهای پشتیبان دارای $a, a^* \epsilon (0, C)$ می‌باشند.

با این نگاشت فرم غیرخطی ماشین بردار پشتیبان تعریف می‌شود و با استفاده از توابع کرنل[۲۶] صورت می‌پذیرد. در واقع به ترکیب تابع کرنل و ماشین بردار پشتیبان خطی، ماشین بردار پشتیبان غیرخطی اطلاق می‌شود (شکل ۳).



شکل ۳- ماشین بردار پشتیبان غیرخطی

تابع کرنل مورد استفاده در این پژوهش، به نام تابع کرنل RBF (گوسین[۲۷]) به شرح رابطه ۶ است.

$$\kappa(x_i, x) = exp(-\gamma\|x - x_i\|^2) \qquad (۶)$$

ماشین بردار پشتیبان با خروجی چندگانه

در نظریه یادگیری ماشین، پیش‌بینی با خروجی چندگانه به مسئله‌ای گفته می‌شود که بیش از یک متغیر خروجی داشته باشد. به عبارتی دیگر، نتیجه حل این مسئله، برداری از متغیرهای خروجی خواهد بود. لذا هدف، یافتن مدلی برای نگاشت متغیرهای ورودی به یک بردار خروجی است. در دنیای واقعی، چالش‌های بسیاری از قبیل از دست رفتن داده[۲۸] و وجود نویز بدلیل چند متغیره بودن خروجی وجود دارد. برای مقابله با این چالش‌ها، ثابت شده است که روش‌های رگرسیونی با خروجی چندگانه عملکرد پیش‌بینی بالاتری در مقایسه با روش‌های رگرسیونی با خروجی یکتا دارند (کورتس و وپنیک[۲۹]، ۱۹۹۵). روش‌های رگرسیون چند خروجی با در نظرگرفتن نه تنها روابط بین متغیرهای ورودی و خروجی

بلکه روابط بین اهداف وخروجی‌ها، تضمین کننده نمایش بهتر هستند و قابلیت تفسیر بالاتری را دارا می‌باشند (تیوا و همکاران[۳۰]،۲۰۱۱). مزیت دیگر رویکردهای چند هدفی این است که می‌توانند مدل‌هایی ساده‌تر با کارایی محاسباتی بهتر تولید کنند (کسوف و همکاران[۳۱]، ۲۰۰۹).

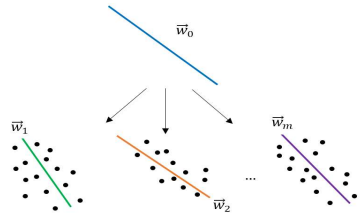
روش‌های موجود برای رگرسیون چند خروجی را می‌توان بدین صورت طبقه‌بندی کرد: تکنیک تبدیل[۳۲] و تکنیک تطبیق الگوریتم[۳۳]. تکنیک تبدیل، به عنوان مدل سنتی و محلی شناخته می‌شود و مسئله خروجی چندگانه را به چندین مسئله تک خروجی به صورت مستقل تبدیل می‌کند و سپس با حل هر کدام از تک خروجی‌ها و تجمیع آن‌ها جواب مسئله اصلی را تعیین خواهد نمود. تکنیک تطبیق الگوریتم، با اصلاح برخی از الگوریتم‌های در دسترس در مدل تک خروجی، اقدام به حل مسئله چند خروجی کرده و از تبدیل آن صرفنظر می‌کند. از این الگوریتم علاوه بر پیش‌بینی خروجی‌ها، برای مدل کردن و تفسیر وابستگی بین خروجی‌ها نیز استفاده می‌شود. مسائلی که به طور معمول در این الگوریتم مورد بررسی قرار می‌گیرد، مدل‌سازی روابط بین خروجی‌ها به صورت همزمان می‌باشد به عبارت دیگر در صورتی‌که بین متغیرهای خروجی همبستگی برقرار نباشد و قابل صرف نظر باشد، مسئله معادل یک مدل با چندین خروجی یکتا است. اما اگر بین متغیرهای خروجی همبستگی قابل صرف نظر نباشد لازم می‌باشد از تکنیک‌های تطبیق استفاده نمود (برچانی و همکاران[۳۴]، ۲۰۱۵). ژو و همکاران[۳۵] (۲۰۱۳) به این نوع خروجی اشاره می‌کنند و برای حل آن یک الگوریتم ارائه دادند که در ادامه تشریح می‌شود.

رگرسیون بردار پشتیبان حداقل مربعات[۳۶]

با تغییر محدودیت‌های نابرابر در رگرسیون بردار پشتیبان با محدودیت‌های برابر، رگرسیون بردار پشتیبان حداقل مربعات، مسئله برنامه‌ریزی غیرخطی محدب را به یک مسئله برنامه‌ریزی خطی محدب تبدیل خواهد کرد.. آنان در طی یک مطالعه تجربی دقیق دریافتند که در این الگوریتم سرعت یادگیری داده‌های آموزشی در مقایسه با رگرسیون بردار پشتیبان بسیار بالاتر می‌باشد (سویکنس و همکاران، ۲۰۱۰). بنابراین این الگوریتم مورد توجه بسیاری بوده است. با وجود کارایی این روش باید بیان کرد که روش مذکور قابلیت نگاشت به یک خروجی چندگانه را ندارد. حال مشکل پیش رو، نحوه مدلسازی خروجی‌های مختلف به صورت همزمان می‌باشد (ژو و همکاران،۲۰۱۳).

رگرسیون بردار پشتیبان حداقل مربعات با خروجی چندگانه[۳۷]

ابتدا فرض کنید تمام شیب ابرصفحه‌ها $w_i \in R^{n_h}$ به صورت $w_i = w_0 + v_i$ نوشته می‌شوند که بردارهای $v_i \in R^{n_h}$ دارای اطلاعات ویژه و خاص و $w_0 \in R^{n_h}$ دارای اطلاعات مشترک هستند که بردار میانگین نام دارند (ژو و همکاران،۲۰۱۳).



شکل ۴- مثالی از رگرسیون بردار پشتیبان با خروجی چندگانه (ژو و همکاران، ۲۰۱۳)

برای حل $w_0 \in R^{n_h}$، $V = (v_1, v_2, \dots, v_m) \in R^{n_h * m}$ و $b = (b_1, b_2, \dots, b_m)^T \epsilon R^m$ به صورت همزمان، می‌توان مدل ریاضی ۷ را حداقل نمود.

$$\min_{w_0,V,b} \mathcal{J}(w_0, V, b) = \frac{1}{2} w_0^T w_0 + \frac{1}{2}\frac{\lambda}{m} trace(V^T V) + \frac{\gamma}{2m} trace(\Xi^T \Xi) + repmat(b^T, l, 1) \quad (۷)$$
$$s.t.\ \ Y = Z^T W + repmat(b^T, l, 1) + \Xi$$

طوریکه $\Xi = (\xi_1, \xi_2, \dots, \xi_m) \in R^{l*m}$ $W = (w_0 + v_1, w_0 + v_2, \dots, w_0 + v_m) \in R^{n_h * m}$، m برابر با تعداد سطرهای ورودی l برابر با تعداد متغیرهای خروجی و $Z = \big(\varphi(x_1), \varphi(x_1), \dots, \varphi(x_1)\big) \in R^{n_h * l}$ و λ همان پارامتر جریمه، γ پارامتر تابع لاگرانژ است که به صورت $\lambda = 2.\hat{}(0:2:15)$ و $\gamma = 2.\hat{}(0:2:10)$ همانند پژوهش ژو و همکاران (۲۰۱۴) تعریف می‌شوند. مدل لاگرانژ رابطه فوق به صورت رابطه ۸ خواهد بود:

$$\mathcal{L}(w_0, V, b, \Xi, \mathrm{A}) = \mathcal{J}(w_0, V, b) - trace(A^T(Z^T W) + repmat(b^T, l, 1) + \Xi - \mathrm{Y}) \quad (۸)$$

در جایی‌که $A = (a_1, a_2, \dots, a_m) \in R^{l*m}$ ماتریسی از ضرایب لاگرانژ می‌باشد. شرایط KKT به صورت روابط ۹ تا ۱۳ اعمال می‌شوند.

$$\frac{\partial \mathcal{L}}{\partial w_0} = 0 \ \Rightarrow \ w_0 = \sum_{i=1}^{m} Z a_i \quad (۹)$$
$$\frac{\partial \mathcal{L}}{\partial V} = 0 \ \Rightarrow \ V = \frac{m}{\lambda} ZA \quad (۱۰)$$
$$\frac{\partial \mathcal{L}}{\partial b} = 0 \ \Rightarrow \ A^T 1_l = 0_l \quad (۱۱)$$
$$\frac{\partial \mathcal{L}}{\partial \Xi} = 0 \ \Rightarrow \ A = \gamma \Xi \quad (۱۲)$$
$$\frac{\partial \mathcal{L}}{\partial A} = 0 \ \Rightarrow \ (Z^T W) + repmat(b^T, l, 1) + \Xi - Y = 0_{l*m} \quad (۱۳)$$

از روابط فوق می‌توان پی برد که w_0 یک ترکیب خطی از v_i می‌باشد، یا به عبارتی دیگر:

$$w_0 = \frac{\lambda}{m}\sum_{i=1}^{m} v_i \quad (۱۴)$$

به کمک رابطه ۱۴، مدل ۷ را به صورت زیر ساده می‌شوند:

$$\min_{V,b} \mathcal{J}(V,b) = \frac{1}{2}\frac{\lambda^2}{m^2} V 1_m 1_m^T V^T + \frac{1}{2}\frac{\lambda}{m} trace(V^T V) + + \frac{\gamma}{2m} trace(\Xi^T \Xi)$$
$$s.t.\ \ Y = Z^T V + repmat\left(\frac{\lambda}{m} Z^T V 1_m, 1, m\right) + + repmat(b^T, l, 1) + \Xi \quad (۱۵)$$

در اینجا می‌توان با استفاده از رویکرد سایکینز و همکاران[۳۸] (۲۰۰۲) W و Ξ را حذف نمود و سیستم خطی شرایط KKT را به صورت زیر بدست آورد.

$$\begin{bmatrix} 0_{ml*m} & P^T \\ P & H \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0_m \\ y \end{bmatrix} \quad (۱۶)$$

به‌طوری‌که:

$$P = blockdiag\left(\overbrace{1_l, 1_l, \dots, 1_l}^{m}\right) \quad (۱۷)$$

$$H = \Omega + \gamma^{-1} I_{ml} + \frac{m}{\lambda} B \quad (۱۸)$$

$$B = blockdiag(\Omega_1, \Omega_2, \dots, \Omega_m) \in \mathbb{R}^{n*n} \quad (۱۹)$$

$$\Omega = Z^T Z \quad (۲۰)$$

$$a = (a_1^T, a_2^T, \dots, a_m^T)^T \quad (۲۱)$$

$$y = (y_1^T, y_2^T, \dots, y_m^T)^T \quad (۲۲)$$

اکنون با بدست آوردن a^* و b^* می‌توان تابع تصمیم دارای همبستگی را با در نظر گرفتن خروجی چندگانه به صورت رابطه ۲۲ مدلسازی کرد:

$$f(x) = \varphi(x)^T W^* + b^{*T} = \varphi(x)^T repmat(w_0^*, 1, m) + + \varphi(x)^T V^* + b^{*T} \quad (۲۳)$$
$$= \varphi(x)^T repmat\left(\sum_{i=1}^{m} Z a_i^*, 1, m\right) + \frac{m}{\lambda}\varphi(x)^T Z A^* + b^{*T}$$
$$= repmat\left(\sum_{i=1}^{m}\sum_{j}^{l} a_{ij}^* K(x, x_j)\right) + \frac{m}{\lambda}\sum_{j}^{l} a_j^* K(x, x_j) + b^{*T}$$

با توجه به بررسی رابطه ۱۶ مشخص می‌شود که این ماتریس یک ماتریس مثبت معین نیست، لذا

حل آن به صورت مستقیم دشوار می‌باشد. برای حل این مسئله ژو و همکاران (۲۰۱۴) یک مدل به صورت زیر ارائه دادند.

$$\begin{bmatrix} S & 0_{ml*ml} \\ 0_{m*m} & H \end{bmatrix} \begin{bmatrix} b \\ H^{-1}Pb + a \end{bmatrix} = \begin{bmatrix} P^T H^{-1} y \\ y \end{bmatrix} \quad (۲۳)$$

به‌طوری‌که $S = P^T H^{-1} P$ یک ماتریس مثبت معین می‌باشد. b و a به ترتیب مراحل زیر بدست می‌آیند.

ابتدا حل η و vاز طریق رابطه ذیل:

$$H\eta = P, Hv = y \quad (۲۴)$$

سپس محاسبه S به صورت زیر:

$$S = P^T \eta \quad (۲۵)$$

و در انتها محاسبه b و a به صورت زیر خواهد بود.

$$b = S^{-1}\eta^T y, a = v - \eta b \quad (۲۶)$$

با استفاده از روابط فوق می‌توان فرآیند آموزش را با حل دو مجموعه معادله خطی با ضرایب ماتریس مثبت معین یعنی H انجام داد. از آنجا که H یک ماتریس با ضرایب مثبت و معین می‌باشد می‌توان از بسیاری از روش‌های بهینه‌سازی سریع و کارآمد در این روش استفاده نمود. از طرفی از آنجا که تعداد خروجی‌ها نسبت به تعداد نمونه‌ها بسیار کمتر است، به راحتی می‌توان معکوس ماتریس S را تنها با ضرب ماتریس‌ها استخراج نمود (ژو و همکاران، ۲۰۱۴).

یافته‌های پژوهش

پس از طراحی و پیاده سازی الگوریتم، با جمع‌آوری داده‌های مورد نظر به پیش‌بینی بازده سهام و تشکیل پرتفوی پرداخته می‌شود. بدین منظور و جهت پیش‌بینی با رگرسیون بردار پشتیبان از نرم افزار MATLAB استفاده می‌شود. یافته‌های پژوهش در دو بخش تشریح می‌شود: در بخش اول، بهینه سازی سبد سهام با استفاده از رگرسیون بردار پشتیبان خروجی همزمان چندگانه و خروجی یکتا تشریح می‌شود و در بخش دوم، مقایسه دقت پیش‌بینی و عملکرد پرتفوی با استفاده از این دو ابزار ارایه می‌شود. بدین منظور، ۲۰۰ داده روزانه به عنوان داده آموزشی، ۵۵ داده به عنوان داده تست مد نظر قرار گرفته‌اند. همچنین متغیرهای اولین قیمت، آخرین قیمت و قیمت پایانی روز قبل به عنوان متغیرهای ورودی و قیمت پایانی روز بعد به عنوان متغیر خروجی رگرسیون بردار پشتیبان در نظر گرفته شده‌اند. به منظور بررسی خطای پیش‌بینی از میانگین خطا[۳۹]، جذر میانگین مربع خطا[۴۰] و درصد میانگین قدرمطلق خطا[۴۱]و

برای‌عملکرد پرتفوی از نسبت شارپ استفاده‌شده‌است. نتایج حاصل در جدول ۱ نشان داده شده‌است.

جدول ۱- مقایسه دقت نتایج به ازای روش‌های مختلف

سهام	COST	CHRW	AWK	AMZN	AIZ	ABBV	AAPL
پیش‌بینی با استفاده از رگرسیون بردار پشتیبان با خروجی چندگانه							
MAD	۰,۰۱۴۷	۰,۰۱۵۶	۰,۰۱۵۴	۰,۰۱۹۸	۰,۰۱۶۲	۰,۰۱۶۲	۰,۰۲۲۳
MSE	۰,۰۰۰۳	۰,۰۰۰۵	۰,۰۰۰۴	۰,۰۰۰۸	۰,۰۰۰۸	۰,۰۰۰۴	۰,۰۰۰۸
RMSE	۰,۰۱۸۲	۰,۰۲۱۸	۰,۰۱۹۸	۰,۰۲۸۰	۰,۰۲۷۵	۰,۰۲۰۷	۰,۰۲۸۲
پیش‌بینی با استفاده از رگرسیون بردار پشتیبان با خروجی یکتا							
MAD	۰,۰۲۳۹	۰,۰۳۵۷	۰,۰۲۳۳	۰,۰۳۳۹	۰,۰۲۰۹	۰,۰۲۰۱	۰,۰۳۰۴
MSE	۰,۰۰۱۰	۰,۰۰۱۷	۰,۰۰۰۸	۰,۰۰۱۶	۰,۰۰۰۹	۰,۰۰۰۷	۰,۰۰۱۶
RMSE	۰,۰۳۱۰	۰,۰۴۱۵	۰,۰۲۸۵	۰,۰۴۰۳	۰,۰۳۰۴	۰,۰۲۵۸	۰,۰۳۹۶

این جدول نشان می‌دهد که پیش‌بینی بازده سهام با استفاده از ماشین بردار پشتیبان چندگانه در مقایسه با ماشین بردار پشتیبان یکتا خطای کمتری دارد.

در ادامه به بررسی بازده مازاد پرتفوی به ازای هر واحد ریسک بر اساس نسبت شارپ پرداخته می‌شود. از معیار شارپ به منظور ارزیابی عملکرد پرتفوی استفاده می‌گردد. در این مرحله برای ارزیابی، معیار ریسک و بازده به صورت همزمان در نظر گرفته می‌شود.

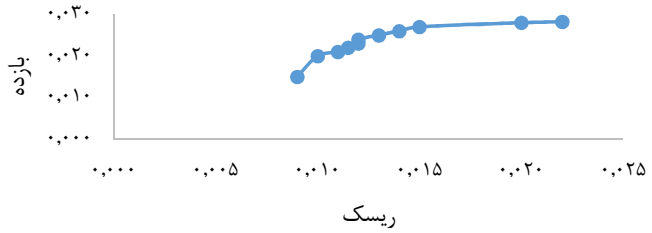
به منظور بررسی بهتر عملکرد الگوریتم توسعه داده شده، دو گروه پرتفوی به نام A و B از روی سطوح ریسک مختلف انتخاب می‌شوند. در پرتفوی A وزن‌های سرمایه‌گذاری مبتنی بر ماشین بردار پشتیبان با خروجی چندگانه و در پرتفوی B وزن‌های سرمایه‌گذاری مبتنی بر ماشین بردار پشتیبان با خروجی یکتا تعیین شده‌اند. سپس بازده موردانتظار پرتفوی و ریسک آنها محاسبه شده و بر اساس معیار شارپ به مقایسه پرتفوها پرداخته می‌شود. لازم بذکر است که بازده بدون ریسک برای محاسبه معیار شارپ یک درصد درنظر گرفته شده است. نتایج در جدول ۲ خلاصه شده است.

جدول ۲- بازده مورد انتظار، ریسک و معیار شارپ پرتفوهای A و B

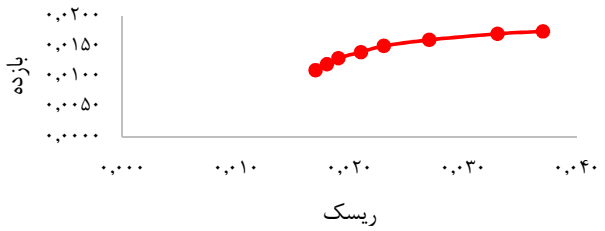
نسبت شارپ پرتفوهای نوع A	نسبت شارپ پرتفوهای نوع B
۰,۰۱۱	۰,۱
۰,۲	۰,۹
۰,۲۲	۰,۸۲

همانطور که مشاهده می‌شود که معیار شارپ محاسبه شده پرتفوی A از پرتفوی B بیشتر است و

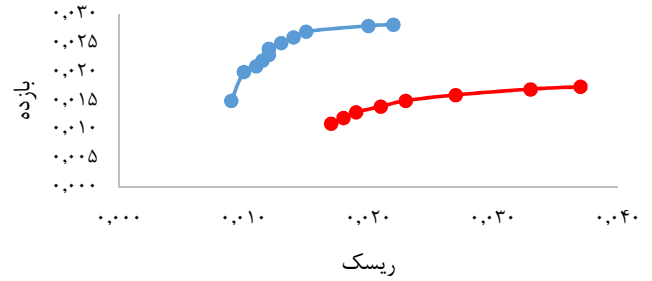
پرتفوی A طبق معیار شارپ عملکرد بالاتری را برای سرمایه‌گذار به دنبال دارد. همچنین نمودارهای مرز کارا استخراج شده (شکل‌های ۵ تا ۷) نشان می‌دهند که بهینه سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از رگرسیون بردار پشتیبان چندگانه در مقایسه با بهینه سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از رگرسیون بردار پشتیبان یکتا به دلیل در نظر گرفتن روابط و همبستگی بین سهام، بازده بالاتری را نسبت به ریسک یکسان و حتی کمتر دارا می‌باشد.



شکل۵- نمودار مرزکارا با استفاده از پیش‌بینی به کمک ماشین بردار پشتیبان



شکل۶- نمودار مرزکارا با استفاده از پیش‌بینی به کمک ماشین بردار پشتیبان یکتا



شکل۷- نمودار مرز کارا برای مقایسه دو حالت مورد بررسی

نتیجه‌گیری

در این پژوهش عملکرد ماشین بردار پشتیبان چندگانه در مقایسه با ماشین بردار یکتا به منظور پیش‌بینی و بهینه سازی سبد سهام مورد بررسی قرار گرفت. نتایج نشان می‌دهد که ماشین بردار پشتیبان چندگانه پیش‌بینی را با خطای کمتری انجام داده است. بنابراین زمانی که بین متغیرهای خروجی نوعی وابستگی برقرار باشد، نمی‌توان از چندین پیش‌بینی با خروجی یکتا استفاده نمود چرا که این روش شامل بکارگیری چندین ماشین بردار پشتیبان برای چنین مسائلی است که نقص آن نادیده گرفتن روابط و وابستگی بین خروجی‌ها می‌باشد. لذا می‌توان به منظور پیش‌بینی جهت افزایش دقت این ابزار را بکار گرفت. همچنین بهینه سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از ماشین بردار پشتیبان چندگانه در مقایسه با بهینه سازی سبد سهام مبتنی بر پیش‌بینی با استفاده از ماشین بردار پشتیبان یکتا به دلیل در نظر گرفتن روابط و همبستگی بین سهام، دارای نسبت شارپ بالاتر و عملکرد بالاتر می‌باشد. طبق نتایج پژوهش می‌توان گفت که بهینه سازی سبد سهام با مدل مذکور به عنوان یک مدل نوین می‌تواند در انتخاب پرتفوی، مؤثر واقع شده و به عملکرد بهتری منتهی گردد.

منابع

۱) حبیب زاده، ملیحه و ایزدپور، مصطفی (۱۳۹۹). پیش‌بینی سودآوری با رویکرد شبکه عصبی و مقایسه آن با ماشین‌بردارپشتیبان (SVM) ودرخت‌تصمیم C5. دانش‌مالی‌تحلیل‌اوراق‌بهادار، ۱۳(۴۶)، ۳۹-۵۶.

۲) خواجه زاده، سامیران، شاهوردیانی، شادی، دانشور، امیر و معدنچی زاج، مهدی (۱۳۹۹). پیش بینی سبد بهینه سهام رویکرد الگوریتم فراابتکاری و فرآیند تصمیم مارکوف. تصمیم گیری و تحقیق در عملیات، ۵(۴)، ۴۲۶-۴۴۵.

۳) زمانی، محسن، افسر، امیر، ثقفی، سید وحید، بیات، الهام. (۱۳۹۳). سیستم خبره پیش بینی قیمت سهام و بهینه سازی سبد سهام با استفاده از شبکه‌های عصبی فازی، مدل سازی فازی و الگوریتم ژنتیک. مهندسی مالی و مدیریت اوراق بهادار، ۵(۲۱)، ۱۰۷-۱۳۰.

۴) باجلان، سعید، فلاحپور، سعید و دانا، ناهید (۱۳۹۶). پیش‌بینی روند قیمت سهام با استفاده از ماشین بردار پشتیبان تعدیل‌یافته همراه با انتخاب ویژگی هیبرید. چشم انداز مدیریت مالی، ۷(۱۷)، ۶۹-۸۶.

۵) فلاح پور، سعید، گل ارضی، غلامحسین و فتوره چیان، ناصر (۱۳۹۲). پیش بینی روند حرکتی قیمت سهام با استفاده از ماشین بردار پشتیبان بر پایه ژنتیک در بورس اوراق بهادار تهران. تحقیقات مالی، ۲(۱۵)، ۲۶۹-۲۸۸.

6) Awad, M., & Khanna, R. (2015). Efficient learning machines: theories, concepts, and applications for engineers and system designers. Springer nature.

7) Borchani, H., Varando, G., Bielza, C., & Larranaga, P. (2015). A survey on multi-output regression. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 5(5), 216-233

8) Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

9) Chen, W., Zhang, H., Mehlawat, M. K., & Jia, L. (2021). Mean–variance portfolio optimization using machine learning-based stock price prediction. Applied Soft Computing, 100, 1-18.

10) Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. Advances in neural information processing systems, 9, 155-161.

11) Freitas, F.D., De Souza, A.F., & De Almeida, A.R. (2009). Prediction-based portfolio optimization model using neural networks. Neurocomputing, 72, 2155-2170.

12) Hao, C., Wang, J., Xu, W., & Xiao, Y. (2013). Prediction-based portfolio selection model using support vector machines. In 2013 Sixth International Conference on Business Intelligence and Financial Engineering, 567-571.

13) Kocev, D., Džeroski, S., White, M. D., Newell, G. R., & Griffioen, P. (2009). Using single-and multi-target regression trees and ensembles to model a compound index of vegetation condition. Ecological Modelling, 220(8), 1159-1168.
14) Lin, K., Lin, Q., Zhou, C., & Yao, J. (2007). Time series prediction based on linear regression and SVR. In Third International Conference on Natural Computation (ICNC 2007), pp. 688-691.
15) Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and machine learning. Expert Systems with Applications, 165, 113973.
16) Mishra, S. K., Panda, G., & Majhi, B. (2016). Prediction based mean-variance model for constrained portfolio assets selection using multiobjective evolutionary algorithms. Swarm and Evolutionary Computation, 28, 117-130.
17) Pinto, D. D. D., Monteiro, J. G. M. S., & Nakao, E. H. (2011). An approach to portfolio selection using an ARX predictor for securities' risk and return. Expert Systems with Applications, 38(12), 15009-15013.
18) Suykens, J. A., Van Gestel, T., & De Brabanter, J. (2002). Least squares support vector machines. World Scientific.
19) Thissen, U.V.B.R., Van Brakel, R., De Weijer, A.P., Melssen, W.J., & Buydens, L.M.C. (2003). Using support vector machines for time series prediction. Chemometrics and intelligent laboratory systems, 69, 35-49.
20) Tuia, D., Verrelst, J., Alonso, L., Pérez-Cruz, F., & Camps-Valls, G. (2011). Multioutput support vector regression for remote sensing biophysical parameter estimation. IEEE Geoscience and Remote Sensing Letters, 8(4), 804-808.
21) Xu, S., An, X., Qiao, X., & Zhu, L. (2014). Multi-task least-squares support vector machines. Multimedia tools and applications, 71(2), 699-715.
22) Xu, S., An, X., Qiao, X., Zhu, L., & Li, L. (2013). Multi-output least-squares support vector regression machines. Pattern Recognition Letters, 34(9), 1078-1084.
23) Yang, F., Chen, Z., Li, J., & Tang, L. (2019). A novel hybrid stock selection method with stock prediction. Applied Soft Computing, 80, 820-831.
24) Zhang, W., Liu, X., Ding, Y., & Shi, D. (2012). Multi-output LS-SVR machine in extended feature space. In 2012 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 130-134.

یادداشت‌ها :

1 Machine Learning
2 Learning Theory
3 Zhang et al.
4 Freitas et al.

5 Pinto et al.
6 Autoregressive exogenous (ATX)
7 Hao et al.
8 Xu et al.
9 Mishra et al.
10 Yang et al.
11 Chen et al.
12 eXtreme Gradient Boosting (XGBoost)
13 Improved firefly algorithm (IFA)
14 Ma et al.
15 Random forest
16 Deep multilayer perceptron
17 Long short-term memory
18 Convolutional neural network
19 Portfolio
20 Support Vector Regression (SVR)
21 Drucker et al.
22 ε-insensitiveregion
23 ε-Tube
24 Awad and Khanna
25 Lin et al.
26 Kernel
27 Radial Basis Function (Gaussian) kernel
28 Missing Data
29 Cortes and Vapnik
30 Tuia et al.
31 Kosev et al.
32 Transformation
33 Algorithm adaptation
34 Borchani et al.
35 Xu et al.
36 Least_squares support vector regression machine (LS_SVR)
37 Multi-output Least-squares support vector regression machine (MLS_SVR)
38 Suykens et al.
39 Mean Error (ME)
40 Root Mean Square Error (RMSE)
41 Mean Absolute Percentage Error (MAPE)