



مدیریت سرمایه‌گذاری در بورس اوراق بهادار با استفاده از رویکرد یادگیری ماشین جهت انتخاب

سهام و به‌کارگیری آن به‌عنوان فین تک

کامیار منکچیان شریف‌آباد^۱

فریدون رهنمای رودپشتی^۲

سیدعلیرضا میرعرب^۳

نرگس یزدانیان^۴

تاریخ دریافت مقاله: ۱۴۰۳/۱۲/۲۲ تاریخ پذیرش مقاله: ۱۴۰۴/۰۲/۲۶

چکیده:

این مقاله به بررسی راهکارهای نوین مدیریت سرمایه‌گذاری در بورس اوراق بهادار با تمرکز بر بهره‌گیری از رویکردهای یادگیری ماشین پرداخته تا فرآیند تصمیم‌گیری سرمایه‌گذاران بهبود بخشد. با توجه به رشد بی‌سابقه حجم داده‌های مالی و چالش‌های ناشی از کاهش کارایی روش‌های سنتی پیش‌بینی، استفاده از الگوریتم‌های هوشمند به‌عنوان ابزاری کارآمد جهت پیش‌بینی بازده سهام مطرح گردیده است. در این پژوهش، ابتدا به بررسی مفاهیم اساسی بورس، اهمیت تحلیل داده‌های سری زمانی مالی و نقش داده‌های حجیم در ارائه پیش‌بینی‌های دقیق پرداخته شده است؛ همچنین ضرورت به‌کارگیری فناوری‌های نوین مالی در بهبود عملکرد سیستم‌های سرمایه‌گذاری مورد تأکید قرار گرفته است. سپس الگوریتم جنگل تصادفی به‌عنوان یکی از روش‌های برجسته یادگیری ماشین معرفی و به‌تفصیل روند ترکیب آن با مدل برنامه‌ریزی درجه دوم جهت تشکیل سبد سهام بهینه شرح داده می‌شود. این مدل ترکیبی، ضمن بهره‌گیری از ویژگی‌های انتخابی الگوریتم، به‌منظور تخمین دقیق‌تر بازده آتی سهام و تقسیم بهینه وزن‌ها در پرتفوی، رویکردی نوآورانه ارائه می‌دهد. نتایج حاصل از تحلیل‌های آماری نشان می‌دهد که رویکرد پیشنهادی در مقایسه با روش‌های کلاسیک مانند پرتفوی حداقل واریانس، عملکرد بهتری از نظر کاهش ریسک و افزایش بازده داشته و اعتبار علمی آن از طریق شاخص‌های آماری مورد قبول واقع شده است. این پژوهش علاوه بر ارائه چارچوبی کاربردی برای تحقق سرمایه‌گذاری بهینه، می‌تواند به‌عنوان مرجعی مفید برای پژوهشگران و سرمایه‌گذاران جهت بهره‌مندی از فناوری‌های یادگیری ماشین و تکنیک‌های پیشرفته در حوزه مدیریت سرمایه‌گذاری و همچنین فین تک به شمار آید. ارائه این روش نوین، زمینه‌ساز توسعه سیستم‌های تصمیم‌گیری خودکار، افزایش شفافیت بازار و بهبود توان رقابتی در بازارهای مالی می‌باشد.

کلمات کلیدی:

مدیریت سرمایه‌گذاری، بورس اوراق بهادار، یادگیری ماشین، فین تک

۱- گروه مدیریت و حسابداری، واحد رودهن، دانشگاه آزاد اسلامی، تهران، ایران. kamyar.mn91@gmail.com

۲- گروه مالی و حسابداری، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران. (نویسنده مسئول) rahnama.roodposhti@gmail.com

۳- گروه مدیریت و حسابداری، واحد رودهن، دانشگاه آزاد اسلامی، تهران، ایران. mirarab_alireza@yahoo.com

۴- گروه مدیریت و حسابداری، واحد رودهن، دانشگاه آزاد اسلامی، تهران، ایران. nargesyazdanian@gmail.com

مقدمه:

بورس اوراق بهادار به‌عنوان یکی از ارکان اصلی بازارهای مالی و سرمایه، نقش حیاتی در تأمین نیاز مالی شرکت‌ها و توسعه اقتصادی کشور دارد. تلاش برای استفاده از فناوری‌های نوین در تحلیل روندهای قیمتی و شکل‌دهی سبدهای سرمایه‌گذاری به‌منظور کسب بازدهی مطلوب از دغدغه‌های اصلی سرمایه‌گذاران امروزی محسوب می‌شود. با افزایش پیچیدگی‌ها در داده‌های مالی و تأثیر عوامل بیرونی نظیر سیاست و عوامل روانشناختی سرمایه‌گذاران، رویکردهای سنتی تحلیل اغلب ناکافی به نظر می‌رسند. در همین راستا، استفاده از رویکردهای یادگیری ماشین به‌ویژه الگوریتم جنگل تصادفی، با توجه به توانمندی فراگیر آن در تحلیل داده‌های حجیم و نامنظم، زمینه‌ساز ایجاد سیستم‌های پیش‌بینی دقیق و ارائه راهکارهای بهینه در مدیریت سبد سهام شده است.

این مقاله باهدف ارائه چارچوبی علمی و کاربردی جهت استفاده هم‌زمان از الگوریتم‌های یادگیری ماشین و روش‌های برنامه‌ریزی درجه دوم، قابلیت پیش‌بینی قیمت و بازده سهام و بهینه‌سازی سبد سرمایه‌گذاری را موردبررسی قرار می‌دهد. از مهم‌ترین اهداف این پژوهش می‌توان به بهبود دقت پیش‌بینی، کاهش ریسک سرمایه‌گذاری و ارائه سودمندی بالاتر نسبت به روش‌های کلاسیک مانند مدل مارکویتز اشاره نمود.

با توجه به اهمیت تشکیل سبد بهینه جهت کاهش ریسک و افزایش بازده در سرمایه‌گذاری‌های بورس، مدل‌های سنتی مانند پرتفوی حداقل واریانس مبتنی بر داده‌های تاریخی، محدودیت‌هایی ازجمله عدم انعکاس دقیق روند آینده را دارا می‌باشند. استفاده از داده‌های تاریخی به‌تنهایی برای پیش‌بینی تغییرات آینده سهام ریسک‌های قابل‌توجهی به همراه دارد؛ ازاین‌رو، رویکردهای مبتنی بر یادگیری ماشین با استفاده از داده‌های به‌روزتر و الگوریتم‌های هوشمند می‌توانند به‌عنوان جایگزینی کارآمد معرفی شوند.

با توجه به تحولات سریع فناوری و گسترش فناوری‌های نوین مالی، ضرورت جایگزینی مدل‌های سنتی با رویکردهای دانش‌بنیان بیش‌ازپیش احساس می‌شود. توجه به اهمیت شفافیت اطلاعات و افزایش اعتماد سرمایه‌گذاران ازجمله مواردی است که می‌تواند به توسعه بازارهای سرمایه کمک شایانی نماید. استفاده از یادگیری ماشین به‌عنوان روشی جهت پیش‌بینی روندهای اقتصادی آینده، به سرمایه‌گذاران کمک می‌کند تا تصمیمات خود را با استناد به داده‌های علمی و دقیق اتخاذ نمایند.

اهداف پژوهش از منظر کاربردی شامل:

- پیش‌بینی بازده آتی سهام با استفاده از الگوریتم جنگل تصادفی و داده‌های قیمتی بازار
- ارزیابی عملکرد سبد سهام بهینه با بهره‌گیری از مدل برنامه‌ریزی درجه دوم
- مقایسه عملکرد سبدهای بهینه حاصل از روش ارائه‌شده با روش‌های سنتی مانند پرتفوی حداقل واریانس

• ارائه چارچوبی نوین جهت ترکیب فناوری‌های داده‌کاوی با نظریه‌های کلاسیک مدیریت سرمایه‌گذاری

این اهداف در راستای ارتقای کیفیت تصمیم‌گیری سرمایه‌گذاران و کاهش خطاهای ناشی از نوسانات بازار موردبررسی قرار می‌گیرند.

ادبیات و پیشینه تحقیق

ادبیات موضوع

در هر پژوهش علمی، مرور ادبیات و پیشینه تحقیقات انجام‌شده ابزاری ساختاری برای تعیین چارچوب نظری ارائه می‌دهد. در این بخش به دو محور اصلی پرداخته می‌شود:

تاریخچه و تکامل بازار بورس: بورس اوراق بهادار از قرون وسطی آغاز شد و با گذر زمان به یکی از مهم‌ترین ارکان سیستم‌های مالی تبدیل گردید. نخستین مفاهیم خریدوفروش سهام در اروپا ایجاد شد و سپس با گسترش اقتصادی و فن‌آوری‌های مدرن بهبودهایی صورت گرفت؛ ازجمله ظهور مدل‌های تحلیل بنیادی و تکنیکال که در روند تصمیم‌گیری‌های سرمایه‌گذاران نقش مهمی ایفا کردند.

نظریه‌های مدیریت سبد سرمایه‌گذاری: یکی از مهم‌ترین مدل‌های کلاسیک در این حوزه مدل مارکویتز است که با استفاده از میانگین و واریانس بازده، سبد بهینه‌ای را ارائه می‌دهد. علی‌رغم این‌که مدل مارکویتز از نظر محاسباتی کارآمد است، اما انتقاداتی ازجمله اتکا به داده‌های تاریخی دارد. از سوی دیگر، مدل‌های نوین مبتنی بر فناوری‌های یادگیری ماشین قابلیت انعطاف‌پذیری و دقت بیشتری در تشخیص الگوهای پنهان در داده‌های مالی فراهم می‌کنند.

توضیح الگوریتم جنگل تصادفی به‌عنوان یکی از ابزارهای یادگیری ماشین: الگوریتم جنگل تصادفی ترکیبی از چندین درخت تصمیم‌گیری است که از طریق بگینگ و انتخاب تصادفی ویژگی‌ها عملکرد بهتری در پیش‌بینی متغیرهای پیوسته و طبقه‌بندی ارائه می‌دهد. به دلیل کاهش احتمال بیش‌برازش^۱ و افزایش تعمیم‌پذیری، این الگوریتم در پیش‌بینی بازده سهام قابلیت کاربرد دارد.

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

نظریه‌های متفاوتی در خصوص ارزیابی و پیش‌بینی بورس در بازارهای سازمان‌یافته مطرح شده است. در اوایل قرن بیستم، گروهی از متخصصان صاحب‌تجربه در ارزیابی اوراق بهادار اعتقاد راسخ بر این امر داشتند که می‌توان از طریق مطالعه و تجزیه و تحلیل روند تاریخی تغییرات قیمت سهام، تصویری را برای پیش‌بینی قیمت آینده سهام ارائه نمود. مطالعات علمی‌تر با تأکید بر شناسایی دقیق رفتار قیمت سهام، گرایش به سمت مدل‌های ارزشیابی قیمت سهام را به وجود آورد. در ابتدا نظریه گام‌های تصادفی به‌عنوان یک شروع در تعیین رفتار قیمت سهام مطرح شد. سپس به ویژگی‌ها و ساختار بازار سرمایه توجه شد که نتیجه این مطالعات و بررسی‌ها منجر به فرضیه‌ی بازار کارآی سرمایه شد. این فرضیه به دلیل ترکیب خاص آن، مورد توجه محافل علمی قرار گرفت. در بازار کارآی سرمایه، اعتقاد بر این است که قیمت سهام انعکاسی از اطلاعات جاری مربوط به آن سهم است و تغییرات قیمت سهام دارای الگوی خاص قابل پیش‌بینی نیست. نظریات مطرح‌شده تا دهه‌ی ۱۹۸۰ میلادی به‌خوبی تعیین‌کننده رفتار قیمت سهام در بازار بودند تا اینکه تحولات بازار سهام نیویورک در سال ۱۹۸۷ میلادی، اعتبار فرضیات بازار کارآی سرمایه و مدل‌هایی نظیر تصادفی بودن قیمت‌ها را به شدت زیر سؤال برد. در دهه ۱۹۹۰ میلادی و بعدازآن، بیشتر توجه متخصصان به یک رفتار آشوبگرانه همراه بانظم معطوف شد و تلاش در جهت طراحی مدل‌های غیرخطی به‌منظور پیش‌بینی قیمت سهام اهمیت روزافزونی یافت. با این نظریات، ازجمله تکنیک‌هایی که اهمیت بالایی یافتند، سیستم‌های هوشمند بودند؛ زیرا با فرض خطی بودن ساختار بازار، به‌آسانی می‌توان بسیاری از مدل‌ها را طراحی نمود. با این وجود، بسیار سخت است که بتوان رفتار مجموعه‌های پیچیده‌ای نظیر بازار سرمایه در یک مجموعه اقتصادی مدرن را به‌طور کامل در یک مجموعه معادلات ساده و خطی نشان داد. مزیت عمده سیستم‌های هوشمند مانند یادگیری ماشین که شبکه‌های عصبی مصنوعی و درخت تصمیم نیز از آن جمله هستند، در مدل‌سازی و پیش‌بینی مجموعه‌های نامنظم و غیرخطی است (کامرو، ۱۳۹۷).

پرتفوی

پرتفوی به معنی سبد سرمایه‌گذاری ترجمه‌شده، سبد سرمایه‌گذاری مفهومی فراتر از سبد سهام است و شامل سایر سرمایه‌گذاری‌های غیر سهم نیز می‌شود. به لحاظ فنی یک سبد سرمایه‌گذاری مجموعه کامل دارایی‌های حقیقی و مالی سرمایه‌گذاری را در برمی‌گیرد. لغت پرتفولیو، در عبارت ساده به ترکیبی از دارایی‌ها گفته می‌شود که توسط یک سرمایه‌گذار برای سرمایه‌گذاری تشکیل می‌شود. این سرمایه‌گذار می‌تواند یک فرد یا یک موسسه باشد. از نظر تکنیکی، یک پرتفولیو در برگیرنده مجموعه‌ای از دارایی‌های واقعی و مالی سرمایه‌گذاری شده یک سرمایه‌گذار است (بیات و اسدی، ۱۳۹۶).

بهینه‌سازی پرتفوی

مفاهیم بهینه‌سازی سبد سهام و تنوع‌بخشی به‌مثابه ابزاری در راستای توسعه و فهم بازارهای مالی و تصمیم‌گیری مالی درآمده‌اند. انتشار نظریه پرتفوی سهام هری مارکوویتز، اصلی‌ترین و مهم‌ترین موفقیت در این راستا بود. از زمانی که مارکوویتز مدل خود را منتشر کرد این مدل تغییرات و بهبودهای فراوانی را در شیوه نگرش مردم به سرمایه‌گذاری و سبد سهام ایجاد کرد و به‌عنوان ابزاری کارا برای بهینه‌سازی سبد سهام به کار گرفته شده است. مارکوویتز پیشنهاد کرد که سرمایه‌گذاران ریسک و بازده را به‌صورت توأمان در نظر می‌گیرند و میزان تخصیص سرمایه بین فرصت‌های سرمایه‌گذاری گوناگون را بر اساس تعامل بین این دو انتخاب نمایند یکی از مباحث مهمی که در بازارهای سرمایه مطرح است و باید موردتوجه سرمایه‌گذاران اعم از اشخاص حقیقی یا حقوقی قرار گیرد، بحث انتخاب سبد سرمایه‌گذاری بهینه می‌باشد و در این رابطه، بررسی و مطالعه سرمایه‌گذاران در جهت انتخاب بهترین سبد سرمایه‌گذاری با توجه به میزان ریسک و بازده آن انجام می‌شود. معمولاً فرض بر این است که سرمایه‌گذاران ریسک را دوست ندارند و از آن گریزانند و همواره در پی آن هستند تا در اقلامی از دارایی‌ها سرمایه‌گذاری کنند که بیشترین بازده و کمترین ریسک را داشته باشند. به‌عبارت دیگر، سرمایه‌گذاران به بازده سرمایه‌گذاری به‌عنوان یک عامل مطلوب می‌نگرند و به واریانس بازده‌ها (ریسک) به‌عنوان یک عنصر نامطلوب نظر دارند. در بهینه‌سازی پرتفوی مسئله اصلی انتخاب بهینه دارایی‌ها و اوراق بهاداری است که با مقدار مشخصی سرمایه می‌توان تهیه کرد اگرچه کمینه کردن ریسک و بیشینه نمودن بازده سرمایه‌گذاری به نظر ساده می‌رسد اما در عمل روش‌های متعددی برای تشکیل پرتفوی بهینه به‌کاررفته است. مسئله بهینه‌سازی مارکوویتز و تعیین مرز کارای سرمایه‌گذاری، زمانی که تعداد دارایی‌های قابل سرمایه‌گذاری و محدودیت‌های موجود در بازار کم باشد، توسط مدل‌های ریاضیات حل‌شدنی است؛ اما هنگامی که شرایط و محدودیت‌های دنیای واقعی در نظر گرفته شود، مسئله پیچیده و مشکل خواهد بود، سال‌هاست که در حل چنین مسائل پیچیده‌ای ریاضیات پیشرفته و کامپیوترها به کمک انسان شتافته تا هر چه بیشتر وی را در بیرون آوردن از شرایط عدم اطمینان محیطی و ابهام یاری رسانند. ازجمله روش‌هایی که در سال‌های اخیر در حل بسیاری مسائل بهینه‌سازی، گره‌گشای ابهامات بشر بوده است و در پاسخ به مسائل پیچیده رویکردی موفق داشته است، روش‌ها و الگوریتم‌های موسوم به ابتکاری (فرا ابتکاری نیز نامیده می‌شوند) است. روش‌های ابتکاری که باهدف رفع کاستی‌های روش‌های کلاسیک بهینه‌سازی معرفی شدند با جستجویی جامع و تصادفی، احتمال دستیابی به نتایج بهتر را تا حد زیادی تضمین می‌کنند ازجمله این الگوریتم‌ها می‌توان به الگوریتم

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

ژنتیک، کوچ پرندگان، تبرید فلزات، جستجوی ممنوعه، کولونی مورچگان، رقص زنبورها، رقابت استعماری و ... اشاره کرد (بیات و اسدی، ۱۳۹۶)

بهینه‌سازی کلاسیک

مارکویتز کسی بود که مفهوم تنوع‌بخشی در سبد سهام را معرفی کرد و آن را توسعه داد. او به‌طور کلی نشان داد که چگونه تنوع‌بخشی در سبد سرمایه، ریسک آن را برای سرمایه‌گذار کاهش می‌دهد. سرمایه‌گذاران می‌توانند سبد سهام کارا را به ازای یک بازده معین و از طریق کمینه کردن ریسک سبد سهام به دست آورند. در ادامه فرآیند فوق می‌تواند منجر به تشکیل سبدهای کارا شود که اصطلاحاً مرز کارای میانگین واریانس نامیده می‌شود.

علت این که سهام یک شرکت به‌عنوان یک دارایی ریسکی مطرح می‌شود، ثابت نبودن (تصادفی بودن) نرخ بازدهی کل (هفتگی، ماهیانه، سالانه) آن است. چون این نرخ‌ها با توجه به زمان تغییر می‌کنند، لذا می‌توان برای آن‌ها تابع توزیع احتمال تشکیل داد و معیارهای موردنیاز مدل مارکویتز نظیر، میانگین، انحراف معیار، کواریانس و ... را از آن به دست آورد.

مدل مارکویتز بر مبنای مفروضات زیر بیان شده است:

- سرمایه‌گذاران ریسک‌گریزند و دارای مطلوبیت مورد انتظار افزایشی می‌باشند و منحنی مطلوبیت نهایی ثروت آن‌ها کاهنده است.

- سرمایه‌گذاران سبد سرمایه خود را بر مبنای میانگین واریانس مورد انتظار بازدهی انتخاب می‌کنند؛ بنابراین منحنی‌های بی‌تفاوتی آن‌ها تابعی از نرخ بازده و واریانس مورد انتظار می‌باشد.

- هر گزینه سرمایه‌گذاری تا بی‌نهایت بار قابل تقسیم است. سرمایه‌گذاران افق زمانی (یک دوره‌ای) داشته و این برای همه سرمایه‌گذاران مشابه است. سرمایه‌گذاران در یک سطح مشخصی از ریسک، بازده بالاتری را ترجیح می‌دهند و بالعکس، سرمایه‌گذاران در انتخاب خود به دو عامل توجه می‌کنند.

الف) "بازده مورد انتظار بالا" که عامل مطلوب است.

ب) "عدم اطمینان بازده" که عامل نامطلوب است. (بیات و اسدی، ۱۳۹۶)

بهینه‌سازی نوین

مبنای این روش بر پایه این استدلال استوار است که احتمال خطر از دست دادن سرمایه یا سود یک نوع سهام در بازار بسیار بیشتر از مجموعه یا ترکیب سهام است. لذا قاعده‌تاً سرمایه‌گذار حرفه‌ای نباید تمام سرمایه خود را در یک‌قلم دارایی سرمایه‌گذاری کند. بلکه بایستی آن را در مجموعه‌ای از

سهام سرمایه‌گذاری کند که این مجموعه به پرتفوی معروف است. پرتفوی یا سبد سهام متشکل از سهام متعدد شرکت‌هایی است که در شرایط عادی احتمال کاهش بازده همه دارایی‌ها تا حدودی به صفر نزدیک است

با توجه به پیچیده شدن و سرعت عوامل تأثیرگذار در تشکیل پرتفوی بهینه، استفاده از روش‌های سنتی تشکیل پرتفوی کار دشواری می‌باشد. رفتار سهام در بازار، رفتار غیرخطی است بنابراین روش‌های خطی نسبت به تشخیص صحیح رفتار غیرخطی ناتوان هستند و استفاده از الگوهای غیرخطی تأثیر به‌سزایی در اتخاذ تصمیمات بهینه در بازار سهام دارد.

پس از مدل‌های کلاسیک، مدل‌های زیادی برای تشکیل پرتفوی بهینه سهام ارائه شده است؛ که اکثر مدل‌ها در کنار روش‌های گذشته، عوامل متعددی را بر تشکیل پرتفوی بهینه سهام اثرگذار می‌دانند و به دنبال راهی برای حداکثر ساختن سود در بازارهای مالی هستند. ایده اساسی نظریه مدرن پرتفوی این است که اگر در دارایی‌هایی که به‌طور کامل باهم همبستگی ندارند سرمایه‌گذاری شود؛ ریسک آن دارایی‌ها یکدیگر را خنثی کرده و می‌توان یک بازده ثابت با ریسک کمتر به دست آورد.

در زمینه برنامه‌نویسی نرم‌افزار بر پایه الگوی تعیین ارزش سهام، اوراق قرضه یا برگه‌ای اختیار معامله سازمان‌های تخصصی متعددی به وجود آمده و راه و روش‌های زیادی برای ترکیب مطلوب سرمایه‌گذاری‌ها ارائه شده است.

یکی از فناوری‌های نوین که اخیر توجه دانش‌پژوهان قرار گرفته است مدل‌ها و رویکردهای مبتنی بر یادگیری ماشین و هوش مصنوعی می‌باشد، این مدل‌ها می‌توانند روش‌های متعددی را برای انتخاب پرتفوی بهینه سهام مورد ارائه نمایند. (سرچمی، پور، محمدی و زینلی، ۱۳۹۹)

بازده سهام:

بازده سهام عبارت است از نسبت کل درآمد (یا زیان) حاصل از سرمایه‌گذاری و میزان سرمایه‌ای که برای کسب درآمد طی یک دوره معین استفاده شده است. این دوره می‌تواند یک روز، یک ماه، یک سال و غیره باشد. برای ارزیابی عملکرد واحدهای تجاری تاکنون ملاک‌های مختلفی ارائه شده است که از جمله آن‌ها مهم‌ترین ملاک برای ارزیابی عملکرد و سودآوری مؤسسات نرخ بازده سهام است که نقش اساسی در سرمایه‌گذاری دارد. این معیار به‌تنهایی دارای محتوای اطلاعاتی برای سرمایه‌گذاران است و از آن برای ارزیابی عملکرد و عملکرد یک واحد تجاری استفاده می‌شود. وقتی این معیار کاهش یابد، این یک علامت هشداردهنده برای شرکت است و نشانه‌ای از عدم عملکرد شرکت است. شاید این معیار

مدیریت سرمایه گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

نسبت به معیارهای عملکرد مبتنی بر حسابداری دارای محتوای اطلاعات بیشتری باشد زیرا ارزیابی عملکرد مبتنی بر بازار منعکس کننده اطلاعات سرمایه گذاران نیست (راستین، ۱۳۹۸)

یادگیری ماشین

در موارد زیر به یادگیری ماشین احتیاج داریم:

- در مواردی که انسان خبره در دسترس نیست؛

- زمانی که مسئله موردنظر در طول زمان تغییر می کند و به شرایط محیط وابسته است؛

- حالتی که حل مسئله به تطبیق با شرایط خاصی وابسته است.

هرچند ممکن است یک مدل دقیق ارائه نگردد، اما می توان یک تقریب خوب و مفید به دست آورد. مدل به دست آمده می تواند برای پیش بینی مورد استفاده قرار گیرد و یا به منظور استخراج دانش از داده ها به کار آید. روش های مختلف یادگیری ماشین به شرح زیر است:

● روش های یادگیری نظارت شده

در این الگوریتم ها باید تابعی ایجاد شود که ورودی ها (متغیرهای مستقل) را گرفته و خروجی موردنظر (متغیر وابسته یا هدف) را تولید کند. نمونه هایی از این الگوریتم ها عبارت اند از رگرسیون، درخت های تصمیم، جنگل های تصادفی، N نزدیک ترین همسایه و رگرسیون لجستیک.

● روش های یادگیری بدون نظارت

یادگیری بدون ناظر متغیر هدف ندارد و خروجی الگوریتم نامشخص است. بهترین مثالی که برای این نوع از الگوریتم ها می توان زد، گروه بندی خودکار (خوشه بندی) یک جمعیت است مثلاً با داشتن اطلاعات شخصی و خریدهای مشتریان، به صورت خودکار آن ها را به گروه های همسان و هم ارز تقسیم کنیم. الگوریتم k -medoid و K -Means از این دسته هستند.

● روش های تقویت شده

یادگیری تقویت شونده نوع سوم از الگوریتم ها که شاید بتوان آن ها را در زمره الگوریتم های بدون ناظر هم دسته بندی کرد. در این نوع از الگوریتم ها؛ یک ماشین برای گرفتن یک تصمیم خاص، آموزش داده می شود و ماشین بر اساس موقعیت فعلی (مجموعه متغیرهای موجود) و عمل های مجاز (مثلاً حرکت به جلو، حرکت به عقب و غیره)، یک تصمیم را می گیرد که در دفعات اول، این تصمیم می تواند کاملاً تصادفی باشد و به ازای هر عمل یا رفتاری که بروز می دهد، سیستم یک بازخورد یا امتیاز به او می دهد و از روی این بازخورد، ماشین متوجه می شود که تصمیم درست را اتخاذ کرده است یا نه که در

دفعات بعد در آن موقعیت، همان عمل را تکرار کند یا عمل و رفتار دیگری را امتحان کند. با توجه به وابسته بودن حالت و رفتار فعلی به حالات و رفتارهای قبلی، فرآیند تصمیم‌گیری مارکوف، یکی از مثال‌های این گروه از الگوریتم‌ها می‌تواند باشد. الگوریتم‌های شبکه‌های عصبی هم می‌توانند در این دسته به حساب آیند. منظور از کلمه تقویت شونده در نام‌گذاری این الگوریتم‌ها هم اشاره به مرحله بازخورد است که باعث تقویت و بهبود عملکرد برنامه و الگوریتم می‌شود.

● روش‌های یادگیری عمیق

نوع جدیدی از یادگیری ماشین برگرفته از نحوه عملکرد مغز هست، مغز انسان ابتدا لبه‌ها را استخراج می‌کند، سپس لکه‌ها، سپس سطوح، سپس اشیاء و غیره، این مشاهدات الهام‌بخش روند اخیری در یادگیری ماشین شناخته‌شده به‌عنوان یادگیری عمیق است که تلاش می‌کند تا این نوع معماری را در یک کامپیوتر تکثیر کند. (سرچمی 1399, et al.)

جنگل تصادفی^۲ یکی از محبوب‌ترین و پرکاربردترین الگوریتم‌های یادگیری ماشین از نوع نظارت‌شده و غیر پارامتریک است که برای پیشبرد امور طبقه‌بندی و رگرسیون از آن استفاده می‌شود و همچنین یکی از الگوریتم‌های موردتوجه در دوره یادگیری ماشین است. جنگل تصادفی از مجموعه‌ای از درخت‌های تصمیم چندگانه برای تولید پیش‌بینی یا طبقه‌بندی استفاده می‌کند. با ترکیب خروجی‌های درختان، الگوریتم جنگل تصادفی نتیجه تلفیقی و دقیق‌تری ارائه می‌دهد. این الگوریتم قادر است تا به‌طور مؤثری مشکلات طبقه‌بندی و رگرسیون را حل نماید، زیرا قدرت این الگوریتم در توانایی مدیریت مجموعه داده‌های پیچیده و کاهش بیش‌برازش نهفته است. الگوریتم جنگل تصادفی در سال ۲۰۰۳ توسط لیو بریمان در دانشگاه اکسفورد ساخته شده است. RF چندین درخت تصمیم مختلف به‌عنوان طبقه‌بندی‌های پایه تولید می‌کند و رأی اکثریت را برای ترکیب با نتایج درختان اصلی اعمال می‌کند. (سعادت، فرساد، ۱۳۹۵) جنگل تصادفی هم توانایی رگرسیون و هم توانایی کلاس‌بندی را با هم دارد. روش کار در یک جنگل تصادفی چنین است که داده‌های نمونه‌ای با عمل جایگذاری به تعداد n مجموعه نمونه تقسیم می‌شود و از هر دسته، نمونه‌ای برای آموزش یک درخت استفاده می‌شود. تمام درخت‌ها عمق مشخصی دارند و درهرگره تقسیم یک‌ویژگی یا فیچر به تصادف از بین مجموعه ویژگی‌ها انتخاب می‌شود و تقسیم یا شاخه بندی بر اساس آن صورت می‌گیرد. به علت استفاده از چندین دسته نمونه این روش مشکل داده‌های پرت و داده‌های گم‌شده را ندارد. (غلامیان و داودی، ۱۳۹۷)

فینتک

فینتک در لغت به معنای فن‌آوری مالی است که در حال تبدیل شدن به بازوی سیستم مالی فعلی

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

می‌باشد. البته استفاده از فن‌آوری در امور مالی چندان جدید نیست ولی استفاده از بسترهای ترکیبی این فن‌آوری‌ها، نظیر بلاک چین، اینترنت اشیا و هوش مصنوعی این حوزه را تحت تغییر عمده قرار خواهد داد. (صندوق-بازنشستگی-کشوری، ۱۳۹۶)

تکنولوژی‌ها و فناوری‌های نوآورانه در حال ایجاد تغییرات و تحولات مهمی در عرصه خدمات مالی از جمله سپرده‌گذاری و وام‌دهی، سرمایه‌گذاری، بیمه، خدمات پرداخت و خدمات توسعه بازار سرمایه می‌باشد که استفاده از این فناوری‌ها موجب کاهش زمان و هزینه و ارتقاء کیفیت خدمات این حوزه‌ها شده که آثار مثبت و مهمی را نیز در رشد و توسعه اقتصادی، افزایش داده‌ها و اطلاعات جهت تحلیل‌ها و بروز شفافیت بیشتر و همچنین افزایش تاب‌آوری اقتصادی خواهد داشت. (Monokchian, sharifabad, Mohamadi, & Bigvand, 2021)

فناوری مالی (به‌اختصار fintech یا FinTech) فناوری و نوآوری‌ای است که هدف آن رقابت با روش‌های مالی سنتی در ارائه خدمات مالی است. این صنعت در حال ظهور است که از فناوری برای بهبود فعالیت‌ها در امور مالی استفاده می‌کند. استفاده از تلفن‌های هوشمند برای بانکداری همراه، سرمایه‌گذاری، خدمات قرض‌گیری و ارزش‌های رمزنگاری‌شده نمونه‌هایی از فناوری‌هایی است که هدف آن دسترسی بیشتر خدمات مالی به عموم مردم است. شرکت‌های فن‌آوری مالی شامل هر دو استارت‌آپ و مؤسسات مالی تأسیس‌شده و شرکت‌های فناوری است که سعی می‌کنند استفاده از خدمات مالی ارائه‌شده توسط شرکت‌های مالی موجود را جایگزین یا افزایش دهند.

فناوری‌های مالی جدید با استفاده از علوم مختلف مانند فناوری بلاک چین، یادگیری ماشینی، بیومتریک و محاسبات ابری تغییرات بزرگی را در زمینه خدمات مالی ایجاد خواهد کرد. این فناوری‌ها می‌توانند تغییرات قابل‌توجهی در بازارهای مالی و رفتار کنشگران در این بازارها ایجاد کنند. بهره‌مندی از فرصت‌های فناوری‌های مالی جدید منجر به رشد بیشتر و توسعه اقتصادی بیشتر کشور می‌شود.

در شرایط کنونی رشد بیشتر و توسعه اقتصادی بیشتر کشور ضروری است. افزایش کارایی و اثربخشی اکثر فعالیت‌ها، به‌ویژه در زمینه مالی می‌تواند زمینه مناسبی برای این رشد و توسعه باشد. Fintechs در توسعه و ارتقاء خدمات مالی می‌تواند هزینه‌ها را کاهش دهد، سرعت را افزایش دهد، تصمیمات آگاهانه‌تری بگیرد، به داده‌های جدید دسترسی پیدا کند و درک عمیق‌تری از مشتریان و بازارها به دست آورد. در این مطالعه ارزیابی میزان آگاهی جامعه از فناوری‌های جدید مالی (fintech) و ارزیابی انتظارات در این زمینه پرداخته‌شده است. نتایج این تحقیق می‌تواند به‌عنوان مبنایی برای مسائل مربوط به فناوری‌های نوین و فناوری‌های مربوط به خدمات مالی استفاده شود.

(Monokchian-sharifabad, 2021)

به کارگیری فن‌آوری‌های مالی (FinTechs) در راستای دستیابی به راه‌حل بهینه برای تصمیم‌گیری‌های کارآمدتر در خصوص سرمایه‌گذاری در بازار سهام در عصر حاضر ضرورت داشته تا در کنار مدل‌های کلاسیک (سنتی) مانند تئوری‌های مدرن و فرا مدرن پورترفوی از یادگیری ماشینی یا هوش مصنوعی برای کمک به تصمیمات سرمایه‌گذاری استفاده نماییم.

با توجه به نقش مهمی که صنعت خدمات مالی در زندگی روزمره افراد ایفا می‌کند، داشتن چشم‌انداز از تغییرات بالقوه‌ای که می‌تواند این صنعت را دست‌خوش تغییر کند- به‌خصوص برای فعالان این صنعت - امری ضروری می‌باشد.

در حالی که فینتک از بسیاری جهات، مانند ارائه‌ی راه‌حل‌های مشتری محور و هزینه‌های کمتر، بهتر از روش‌های سنتی عمل می‌کنند اما هنوز هم سازمان‌های سنتی مالی در برخی از عرصه‌های مهم هم چون اعتمادسازی موفق‌تر از رقبای فینتکی خود ظاهر می‌شوند با این وجود با توجه به موفقیت‌های و ارزش‌افزوده احتمالی پیش‌بینی می‌گردد سهم بازار فینتک از بازار خدمات مالی در آینده افزایش یابد.

(Monokchian-sharifabad, 2021)

پیشینه پژوهش

مطالعات خارجی

در تحقیقی که توسط شونبرگ در سال ۱۹۹۰ با عنوان پیش‌بینی قیمت سهام با استفاده از شبکه عصبی ارائه شد، امکان پیش‌بینی قیمت سهام به صورت کوتاه‌مدت، روز به روز به کمک شبکه‌های عصبی با مطالعه بر روی سه سهام مهم آلمان که به طور تصادفی انتخاب شدند، مورد بررسی قرار داده شد. در این تحقیق شبکه‌های پرسپترون، آدالین، مادالین و پس انتشار خطا مورد بررسی قرار داده شد. در مدت زمان ۱۰ روزه به درجه بسیار بالای دقت تا ۹۰٪ رسیدند (کامرو، ۱۳۹۷)

در تحقیقی که توسط کوهارا 1 و همکاران در سال ۱۹۹۷ با عنوان پیش‌بینی سهام با استفاده از دانش قبلی و شبکه‌های عصبی انجام شد، روشی برای استفاده هم‌زمان از دانش قبلی و شبکه‌های عصبی برای بهبود توانایی پیش‌بینی ارائه شد. در این پژوهش از دانش قبلی درباره پیش‌بینی قیمت سهام و اطلاعات روزنامه داخلی و خارجی استفاده شده است. چند شاخص اقتصادی نیز بر اساس دانش پیشین انتخاب شده و با دانش موجود به شبکه‌های عصبی وارد شده است. نتایج نشان داد که خطای پیش‌بینی رویکرد ارائه شده کوچک‌تر از تجزیه و تحلیل رگرسیون چندگانه در سطح ۵٪ اهمیت است (کامرو، ۱۳۹۷)

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

در تحقیقی که توسط بالینگز و همکاران در سال ۲۰۱۵ با عنوان ارزیابی چند دسته‌بندی برای پیش‌بینی جهت روند قیمت سهام انجام شد، هدف از انجام آن، ارزیابی روش‌های اجماعی نظیر جنگل تصادفی، AdaBoost و Kernel Factory در برابر مدل‌های طبقه‌بندی واحد نظیر شبکه‌های عصبی، رگرسیون منطقی، ماشین‌های بردار پشتیبانی و K-همسایه نزدیک‌تر عنوان شد. اطلاعات مورد استفاده از ۵۷۶۷ شرکت اروپایی که به‌طور عمومی فهرست شده‌اند جمع‌آوری شده است. نتایج نشان داد که جنگل تصادفی، بهترین الگوریتم است که به دنبال آن ماشین بردار پشتیبان قرار دارد، Kernel Factory، AdaBoost، شبکه عصبی، K همسایه نزدیک و رگرسیون منطقی قرار دارد (کامرو، ۱۳۹۷).

در این تحقیقی که توسط میلوسویچ در سال ۲۰۱۶ با عنوان پیش‌بینی سود: پیش‌بینی حرکت بلندمدت قیمت سهام با استفاده از یادگیری ماشین انجام شد از یک روش یادگیری ماشین برای ارزیابی قیمت آتی سهام در طول زمان استفاده شد. این روش قادر است در ۷۶.۵ درصد موارد به‌درستی پیش‌بینی کند که آیا ارزش یک شرکت می‌تواند ۱۰ درصد بالاتر یا بیشتر در طول یک سال بالا رود (کامرو، ۱۳۹۷).

تاکور و همکاران (۲۰۱۷) انتخاب پرتفوی سهام را با استفاده از تئوری شواهد دمپستر شیفر در بورس اوراق بهادار بمبئی با استفاده از متغیرهای نسبت قیمت به سود، نسبت قیمت به ارزش دفتری، نسبت قیمت به فروش، نسبت بدهی‌های بلندمدت به حقوق صاحبان سهام، سود هر سهم و نسبت پرداخت موردبررسی قرار داده‌اند. از مدل دلفی-فازی برای شناسایی عوامل ضروری استفاده شده سپس عوامل ضروری و داده‌های تاریخی با استفاده از تئوری شواهد دمپستر شیفر رتبه‌بندی سهام انجام شده و سرانجام یک مدل پرتفوی که سهام با رتبه بالا را پیشنهاد می‌دهد انتخاب گردیده است. نتیجه عملکرد در مقایسه با عملکرد اخیر دارایی‌ها بسیار رضایت‌بخش می‌باشد. (سرچمی 1399 et al.)

پاتاری و همکاران (۲۰۱۸) مقایسه روش‌های تصمیم‌گیری چند معیاره برای انتخاب پرتفوی سهام در بورس نزدیک آمریکا طی یک دوره چهل و سه ساله از ۱۹۶۹ تا ۲۰۱۱ را با موردبررسی قرار دادند. از چهار روش مقیاس متوسط، روند سلسله‌مراتب تحلیلی، تکنیک برای ترجیح سفارش با شباهت به راه‌حل ایده آل و تجزیه و تحلیل پوشش افزایشی داده‌ها استفاده شده است. متغیرهای تحقیق شامل ارزش بازار سهام، ارزش شرکت، سود، ارزش دفتری سهام، فروش، سود سهام عادی، جریان نقد عملیاتی، جریان نقد آزاد، سود قبل از بهره، مالیات و استهلاک و سود قبل از بهره و مالیات هستند. نتایج حاکی از این است که سرمایه‌گذارانی که سبک‌های سرمایه‌گذاری مشخص را دنبال می‌کنند، این سبک‌های متفاوت را هنگام انتخاب روش‌های تصمیم‌گیری چند معیاره می‌توانند داشته باشند که

بهترین سازگاری را باهدف‌های انتخاب پرتفوی آن‌ها به همراه دارد. (سرچمی et al., 1399)

جین و همکاران (۲۰۱۹) به بررسی رویکرد برابری ریسک سلسله‌مراتبی (HRP) تخصیص سبد سهام، معرفی شده توسط لویز د پرادو (۲۰۱۶) پرداختند و از نظریه نمودار (گراف) و یادگیری ماشینی برای ایجاد یک پرتفوی متنوع استفاده نمودند؛ مانند روش‌های سنتی تخصیص مبتنی بر ریسک، HRP نیز تابعی از برآورد ماتریس کوواریانس است، باین‌حال، به برگشت‌ناپذیری^۳ آن نیاز ندارد. در آن مقاله، ابتدا تأثیر نادرست تشخیص کوواریانس بر عملکرد روش‌های مختلف تخصیص را مطالعه شده و در مرحله بعد، تحت یک مدل پیش‌بینی کوواریانس مناسب مطالعه صورت گرفته که آیا HRP مبتنی بر یادگیری ماشین از سبدهای سنتی مبتنی بر ریسک بهتر عمل می‌کند یا خیر. نتایج پژوهش حاکی از این بوده است که وقتی برآورد کوواریانس خام^۴ است، پرتفوی‌های معیار نوسانات معکوس^۵ قوی‌تر هستند و پس از آن پرتفوی‌های مبتنی بر یادگیری ماشین. (Jain & Jain, 2019)

کازمارک و همکاران (۲۰۲۱) نشان دادند که تکنیک‌های بهینه‌سازی سبد دارایی که توسط میانگین‌های واریانس مارکوویتز و برابری ریسک سلسله‌مراتبی (HRP) ارائه می‌شوند، بازدهی تعدیل‌شده سهام با استفاده از سهام از پیش انتخاب‌شده با ابزار یادگیری ماشین را افزایش می‌دهد. آن‌ها از روش جنگل تصادفی برای پیش‌بینی مقطع بازده اضافی مورد انتظار استفاده نموده و n سهام با بیشترین پیش‌بینی ماهانه را انتخاب می‌کردند با سه تکنیک مختلف-میانگین واریانس، HRP و را برای ایجاد وزن سبد با استفاده از بازده سهام از S & P500 و STOXX 600 برای استحکام مقایسه نمودند. نتایج خارج از نمونه^۶ نشان می‌دهد که هر دو بهینه‌کننده میانگین واریانس و HRP از قانون 1/N بهتر عمل می‌کنند. این نتیجه‌گیری در تقابل با انتقاد رایج از کارایی بهینه‌سازان است و روش جدیدی را در استفاده عملی بالقوه آن‌ها ارائه می‌دهد. (Kaczmarek & Perez, 2021)

مطالعات داخلی

مدرس و استخری (۱۳۸۶) سبد سهام در بورس اوراق بهادار تهران را با الگوریتم ژنتیکی انتخاب می‌کنند. نتایج تحقیقات آن‌ها همچنین نشان می‌دهد که بهینه‌سازی پرتفوی با استفاده از ژنتیک نسبت به راه‌حل آن با مدل‌های کلاسیک کارآمدتر است و با این رویکرد ابتکاری می‌توان حدود عملکرد بهتری نسبت به راه‌حل مدل سنتی مارکوویتز به دست آورد. (راستین، ۱۳۹۸)

رهنمای رود پشته و همکاران (۱۳۸۹) در پژوهش حاضر به معرفی مدل نوینی از مدل‌های قیمت‌گذاری دارایی سرمایه‌ای به نام پرداخته می‌شود. برای این (R-CAPM) مدل قیمت‌گذاری دارایی سرمایه‌ای تجدیدنظر شده منظور، ابتدا درجه اهرم اقتصادی، یعنی اثرات اختلالات اقتصادی که

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

شامل پنج متغیر کلان اقتصادی نرخ تورم، نرخ ارز، نرخ بیکاری، صادرات و هزینه مالی است، بر روی فروش شرکت‌ها اندازه‌گیری و از این طریق ضریب حساسیت بتا را محاسبه کرده و به دنبال آن با دستیابی به مدل قیمت‌گذاری دارایی سرمایه‌ای تجدیدنظر شده، بازده مورد انتظار به صورت واقعی محاسبه، سپس پیش‌بینی می‌شود و جهت سنجش توان تبیین این مدل را با سایر مدل‌های قیمت‌گذاری اعم از مدل قیمت‌گذاری دارایی سرمایه‌ای، مدل قیمت‌گذاری دارایی سرمایه‌ای کاهشی و مدل قیمت‌گذاری دارایی سرمایه‌ای تعدیل شده برای ۶۷ شرکت پذیرفته شده در بورس اوراق بهادار تهران در دوره زمانی ۸ ساله مورد مقایسه قرار داده و با استفاده از تحلیل همبستگی و رگرسیون و آزمون‌هایی نظیر فریدمن، ویلکاکسون داده‌ها مورد تجزیه و تحلیل قرار می‌گیرد. نتیجه پژوهش بیانگر آن است که مدل قیمت‌گذاری دارایی سرمایه‌ای تجدیدنظر شده با توجه به شرایط حاکم بر جامعه ما در مقایسه با سایر روش‌های قیمت‌گذاری در پیش‌بینی ریسک و بازده دارای توان تبیین بالاتری است. (رهنمای-رودپشتی و حسینی، ۱۳۸۹)

رهنمای رود پشتی و همکاران (۱۳۹۳) بهینه‌سازی پرتفوی متشکل از سهام صندوق‌های سرمایه‌گذاری مشترک بورس اوراق بهادار تهران را با رویکرد الگوریتم ژنتیک در سی صندوق سرمایه‌گذاری مشترک فعال در بازار سرمایه ایران در طی سال‌های ۱۳۸۹ الی ۱۳۹۱ با سه متغیر بازده مورد انتظار، ریسک پرتفوی و کوواریانس مورد بررسی قرار دادند. نتایج حاکی از آن است که الگوریتم ژنتیک می‌تواند جهت انتخاب سبد متشکل از سهام صندوق‌های مشترک به کار رود و با استفاده از آزمون زوجی مشخص شد که سبدهای تشکیل شده با استفاده از الگوریتم ژنتیک نسبت به روش سنتی، مطلوب‌تر می‌باشند. همچنین اندازه سبد تأثیر چندانی بر نتایج نداشته و در تمام سطوح، الگوریتم ژنتیک دارای عملکرد بهتری است و هر چه تنوع سبد بیشتر و بزرگ‌تر باشد، برتری الگوریتم ژنتیک بر روش خطی قابل ملاحظه‌تر می‌شود. (سرچمی 1399, et al.)

بیات، شکری (۱۳۹۴)، در پژوهشی به بررسی فرایند انتخاب پرتفوی بهینه به روش ارزش در معرض ریسک پرداختند و در آن به انواع روش‌های بهینه‌سازی به روش ارزش در معرض ریسک مانند روش واریانس کواریانس، شبیه‌سازی تاریخی و روش شبیه‌سازی مونت کارلو اشاره کردند و در نهایت به این - نتیجه رسیدند که در انتخاب پرتفوی بهینه مدل ارزش در معرض ریسک بهترین مدل پیشنهادی می‌باشد (بیات و اسدی، ۱۳۹۶)

کوهبنانی نژاد و همکاران (۱۳۹۷) انتخاب پرتفوی بهینه با استفاده از سیستم خبره در محیط فازی ممدانی را در صنایع شیمیایی، فراورده نفتی، کانی فلزی و فلزات اساسی بورس اوراق بهادار تهران در

بازه زمانی ۱۳۸۷ تا ۱۳۹۰ انجام دادند. با در نظر گرفتن متغیرهای حاشیه سود خالص، سود قبل از کسر مالیات، بازده حقوق صاحبان سهام، بدهی به حقوق صاحبان سهام، ارزش بازار، نسبت قیمت به سود، سود تقسیمی، نسبت ارزش بازار به ارزش دفتری، مومنتوم قیمت و شاخص فنی، شرکت‌ها را رتبه‌بندی کردند و پس از بهینه‌سازی پرتفوی با معیارهای بازده، بازده مرکب، ترینر و جنسن عملکرد پرتفوی تشکیل‌شده با استفاده از سیستم خبره در محیط فازی ممدانی را موردسنجش قرار دادند. نتایج ارزیابی عملکرد پرتفوی تشکیل‌شده برای سه حالت سرمایه‌گذار ریسک‌گریز، ریسک‌خنثی و ریسک‌پذیر نشان می‌دهد که عملکرد پرتفوی پیشنهادی مثبت بوده؛ اما در مقیاسی دقیق‌تر پرتفوی تشکیل‌شده برای سرمایه‌گذار ریسک‌گریز در وضعیت مطلوب‌تری قرار دارد. (سرچمی et al., 1399)

(اقتصاد و محمدی، ۱۴۰۲) پیش‌بینی بازده در تشکیل پورتفولیو را با مدل یادگیری ماشین، یعنی جنگل تصادفی و مدل یادگیری عمیق حافظه کوتاه‌مدت طولانی ترکیب می‌کند. به‌منظور ارزیابی عملکرد مدل پیشنهادی، داده‌های تاریخی ۵ ساله از سال ۱۳۹۶ تا ۱۴۰۱ از شاخص ۵ صنعت بانکی، خودرویی، دارویی، فلزی و نفتی است. نتایج تجربی نشان می‌دهد که مدل‌های بهینه‌سازی میانگین واریانس با پیش‌بینی بازدهی به‌وسیله جنگل تصادفی، بهتر عمل می‌کنند.

روش‌شناسی پژوهش

پژوهش حاضر از نظر نوع داده کمی بوده و با استفاده از داده‌های واقعی از شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران انجام‌شده است. در ادامه، گام‌های اجرایی و فرآیند تحلیل داده‌ها تشریح می‌شود. جامعه آماری و نمونه پژوهش جامعه آماری شامل تمامی شرکت‌های پذیرفته‌شده در بورس تهران است که پس از اعمال معیارهای انتخاب، نمونه نهایی شامل شرکت‌هایی می‌شود که:

شرکت‌های هلدینگ و سرمایه‌گذاری نباشند

بانک‌ها به دلیل ماهیت خاص فعالیت حذف شوند

تغییر سال مالی یا توقف عملیات نداشته باشند

اطلاعات موردنیاز به‌صورت کامل در دسترس باشند دوره زمانی موردبررسی از سال ۱۳۹۵ تا ۱۴۰۳ بوده و داده‌های مربوط به قیمت پایان روز، حجم معاملات و مبلغ معاملات روزانه ثبت‌شده‌اند.

فرآیند پیش‌بینی قیمت سهام

الف) گردآوری داده‌ها: داده‌های جمع‌آوری‌شده از منابع معتبر مانند نرم‌افزار دریافت اطلاعات بورس و بانک‌های اطلاعاتی سازمان بورس، ابتدا در نرم‌افزار Excel جهت تحلیل اولیه ذخیره و بررسی شدند.

(ب) پیاده‌سازی الگوریتم جنگل تصادفی:

- الگوریتم یادگیری ماشین جنگل تصادفی با استفاده از نرم‌افزار R پیاده‌سازی شده است.
- فرآیند برگزاری آموزش مدل شامل انتخاب تصادفی زیرمجموعه‌های داده (با تکنیک Bootstrap Sampling) و ساخت مجموعه‌ای از درختان تصمیم‌گیری انجام شده و برای هر نمونه جدید پیش‌بینی بر مبنای میانگین خروجی درختان صورت می‌گیرد.
- معیارهای ارزیابی عملکرد مدل شامل میانگین مربعات خطا (RMSE)، میانگین خطای مطلق (MAE) و میانگین نسبی خطای مطلق برای سنجش دقت پیش‌بینی به کار گرفته شده‌اند.

- تشکیل سبد بهینه سهام:

پس از پیش‌بینی بازده آتی سهام، فرآیند بهینه‌سازی سبد سرمایه‌گذاری با استفاده از روش برنامه‌ریزی درجه دوم انجام می‌گیرد. در این بخش مدل برنامه‌ریزی درجه دوم جهت تعیین وزن هر سهم در سبد بهینه به کاررفته است و هدف ارائه سبدی است که برای سطح معینی از بازده، حداقل ریسک محتمله را تأمین کند.

با ترکیب خروجی‌های الگوریتم جنگل تصادفی و مدل برنامه‌ریزی درجه دوم، سبد بهینه سهام شناخته شده و از لحاظ عملکردی از طریق معیارهایی نظیر نسبت شارپ ارزیابی شده است.

الگوریتم جنگل تصادفی^۷ یکی از قدرتمندترین الگوریتم‌های یادگیری ماشین است که برای مسائل طبقه‌بندی^۸ و رگرسیون^۹ مورد استفاده قرار می‌گیرد. این الگوریتم بر پایه‌ی درختان تصمیم‌گیری^{۱۰} ساخته شده است و با ایجاد مجموعه‌ای از درختان تصمیم، پیش‌بینی‌های دقیقی ارائه می‌دهد. جنگل تصادفی به زبان ساده از ترکیب چندین درخت تصمیم‌گیری استفاده می‌کند تا دقت و پایداری مدل را بهبود ببخشد. این رویکرد بر اساس دو تکنیک اصلی عمل می‌کند:

۱. بگینگ^{۱۱} (تقویت داده با مجموع‌گیری بوت‌استرپ): در جنگل تصادفی، داده‌های آموزشی به صورت تصادفی و با جایگذاری^{۱۲} انتخاب می‌شوند و برای ساخت درختان مختلف مورد استفاده قرار می‌گیرند. به این ترتیب، هر درخت تصمیم روی یک زیرمجموعه تصادفی از داده‌ها آموزش داده می‌شود.

۲. انتخاب تصادفی ویژگی‌ها: در هر گره^{۱۳} از هر درخت، تنها یک زیرمجموعه تصادفی از ویژگی‌ها^{۱۴} برای تقسیم‌بندی داده‌ها انتخاب می‌شود. این امر باعث تنوع بیشتر میان درختان می‌شود و مشکلاتی نظیر تسلط ویژگی‌های قوی^{۱۵} بر مدل را کاهش می‌دهد.

به دلیل استفاده از بگینگ و انتخاب تصادفی ویژگی‌ها، احتمال بیش‌برازش^{۱۶} مدل کاهش یافته و

توان تعمیم مدل افزایش می‌یابد. مراحل کار الگوریتم جنگل تصادفی به شرح زیر می‌باشد:

۱- انتخاب زیرمجموعه‌ای از داده‌ها^{۱۷}:

به صورت پیش فرض، برای هر درخت یک زیرمجموعه تصادفی از داده‌ها با جایگذاری^{۱۸} انتخاب می‌شود.

داده‌هایی که در زیرمجموعه آموزشی یک درخت گنجانده نشده‌اند، به عنوان داده‌های خارج از بسته^{۱۹} عمل می‌کنند. این داده‌ها برای ارزیابی عملکرد مدل استفاده می‌شوند.

۲- ایجاد درختان تصمیم‌گیری: برای هر زیرمجموعه، یک درخت تصمیم‌گیری جداگانه ساخته می‌شود. در هر گره درخت، زیرمجموعه تصادفی از ویژگی‌ها برای یافتن بهترین تقسیم‌بندی بررسی می‌شود، این ویژگی باعث کاهش وابستگی به یک ویژگی خاص و افزایش تنوع میان درخت‌ها می‌شود.

۳- پیش‌بینی با هر درخت: هر درخت ساخته شده، پیش‌بینی خود را بر روی داده‌های جدید ارائه می‌دهد.

۴- در مسائل رگرسیون: از میانگین مقادیر پیش‌بینی شده توسط درختان به شرح زیر استفاده می‌شود.

در مسائل رگرسیون، مدل به صورت زیر پیش‌بینی انجام می‌دهد:

۱. مدل تعداد زیادی درخت تصمیم‌گیری می‌سازد. هر درخت به صورت جداگانه روی یک زیرمجموعه تصادفی از داده‌ها که از طریق بوت‌استرپینگ انتخاب شده‌اند، آموزش داده می‌شود. هر زیرمجموعه شامل نمونه‌هایی از داده‌های اصلی و معمولاً کمتر از کل داده‌ها است.

۲. وقتی می‌خواهیم مقدار یک نمونه جدید را پیش‌بینی کنیم:

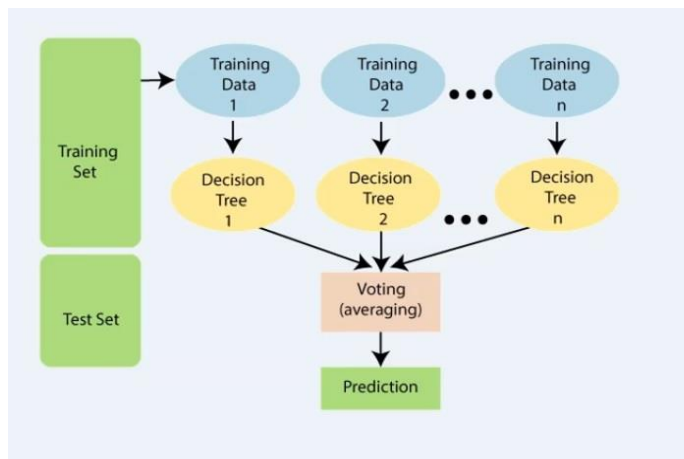
○ هر درخت تصمیم‌گیری یک پیش‌بینی مستقل می‌دهد (مقداری را که برای آن نمونه مناسب می‌داند انتخاب می‌کند). این پیش‌بینی بر اساس اصول تقسیم‌بندی است که در طول ساخت درخت‌ها آموخته شده است.

۳. مقدار پیش‌بینی شده نهایی برابر با میانگین مقادیر پیش‌بینی شده توسط همه درخت‌ها خواهد

$$\text{بود: فرمول } \hat{y} = \frac{1}{n} \sum y_n$$

که در آن n تعداد درخت‌ها و y_i مقدار بازگشتی پیش‌بینی شده توسط درخت i ام است.

که شماتیک آن نیز به شرح شکل زیر می‌باشد:



جنگل تصادفی مورد استفاده در این پژوهش بر مبنای الگوریتم جنگل تصادفی Breiman (بر اساس کد فرتن اصلی بریمن و کاتلر) که برای طبقه بندی و رگرسیون پیاده سازی می شود همچنین می تواند برای پیش بینی و در حالت بدون نظارت برای ارزیابی نزدیکی بین نقاط داده استفاده شود. (L Breiman, 2001; L Breiman, 2002)

و برای سنجش عملکرد مدل در پیش بینی از دو معیار میانگین مربعات خطا^{۲۰} (RMSE)، میانگین خطای مطلق^{۲۱} (MAE) و میانگین نسبی خطای مطلق به شرح زیر استفاده شده است:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (r_t - \hat{r}_{t-1})^2}$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |r_t - \hat{r}_{t-1}|$$

که در آن r و $\hat{r}$ به ترتیب بازده واقعی و بازده پیش بینی شده را در زمان t نشان می دهد.

بهینه سازی سبد سهام در حالت عمومی که شامل برخی محدودیت ها روی وزن ها است یک مسئله برنامه ریزی درجه دوم است و تکنیک های حل مسائل برنامه ریزی درجه دو^{۲۲} (QP) یک روش قدرتمند برای مدیریت ریسک و بازده در سرمایه گذاری های مالی است. بهینه سازی سبد سهام با استفاده از برنامه ریزی درجه دو به ما این امکان را می دهد که با در نظر گرفتن ریسک و بازده مورد نظر، سبد سهام بهینه ای را برای سرمایه گذاری مشخص کنیم. در این روش، تابع هدف معمولاً به شکل یک تابع درجه دو که شامل ریسک (واریانس یا انحراف معیار) و بازده مورد انتظار سبد است، تعریف می شود.

توابع نرم افزار R، برای حل مسائل بهینه سازی برنامه ریزی درجه دو یکی از ابزارهای پرکاربرد برای

بهینه‌سازی پرتفوی است که در آن هدف کمینه کردن ریسک (با استفاده از ماتریس واریانس و کواریانس بازده‌ها) است درحالی‌که بازده مورد انتظار مشخص شده است.

الگوریتم و تابع مورد استفاده برای حل مسائل بهینه‌سازی درجه دو با استفاده از روش‌های عددی پیشرفته طراحی شده است. مسائل بهینه‌سازی درجه دو به‌طور کلی به شکل زیر هستند:

$$(\frac{1}{2}X^T QX + C^T X) \min_X$$

در آن:

X : یک بردار از متغیرهای تصمیم است (وزن‌های دارایی‌های مختلف در یک پرتفوی)، بردار وزن‌های سهام در سبد سرمایه‌گذاری که نشان‌دهنده درصد سرمایه‌گذاری در هر سهم است. به‌عنوان مثال، اگر x_1 نشان‌دهنده وزن سهم اول باشد، x_2 وزن سهم دوم و غیره، باید مجموع وزن‌ها برابر با ۱ باشد:

$$\sum_{i=1}^n x_i = 1$$

Q : یک ماتریس مربعی مثبت (ماتریس واریانس و کواریانس است) که این ماتریس واریانس و کواریانس بین سهام‌ها را شامل می‌شود و ریسک سبد سهام را به تصویر می‌کشد. هر عنصر σ_{ij} در ماتریس نشان‌دهنده کواریانس بین بازده‌های سهام i و j است. این ماتریس به ما کمک می‌کند تا چگونگی تداخل (ارتباط) ریسک بین سهام‌ها را بفهمیم.

C : یک بردار از مقادیر خطی است (برای بازده مورد انتظار دارایی‌ها تعریف می‌شود) و بردار شامل بازده‌های مورد انتظار برای هر سهم است.

در کنار این تابع هدف، محدودیت‌های خطی اعمال می‌شوند:

$$AX \leq b$$

که در اینجا:

A ماتریسی است که محدودیت‌های غیر منفی و مساوی را بیان می‌کنند.

b برداری است که محدودیت‌های مورد نظر را تعریف می‌کنند.

روش عددی به‌کاررفته برای حل این نوع مسائل برنامه‌ریزی درجه دوم، روش عددی پایدار الگوریتم‌های بهینه‌سازی محدود (Constrained Optimization Algorithms) بر مبنای تکنیک متغیرهای دوگان (Dual Variable Methods) برای حل مسائل برنامه‌ریزی درجه دو می‌باشد. الگوریتم پیشنهادی بر اساس روش دوگان عمل می‌کند و با استفاده از تکنیک‌های عددی پایدار برای حل مسائل بهینه‌سازی طراحی شده است.

مدیریت سرمایه گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

در روش دوگان (Dual Methods) تابع بهینه سازی به دو بخش تقسیم می شود، تابع هدف اصلی و تابع هدف دوگان که محدودیت ها را در نظر می گیرد. با حل دو مسئله ی جداگانه و تطبیق مقادیر دوگانه، می توان به حل بهینه دست یافت که فرآیند حل آن به شرح ذیل می باشد:

۱- تعریف مسئله: ابتدا مسئله ی بهینه سازی را از طریق تابع مربوطه و در قالب تابع هدف و محدودیت ها تعریف می نماییم.

۲- محدودیت ها: محدودیت ها به صورت ماتریس ها و بردارهایی برای تضمین شرایط سرمایه گذاری مانند محدودیت های غیر منفی بودن وزن ها (در صورتی که خرید کوتاه (short selling) مجاز نباشد) یا محدودیت هایی مثل مجموع وزن ها برابر یک باشد، وارد می شوند.

۳- بهینه سازی: الگوریتم پس از دریافت داده های ورودی، شروع به بهینه سازی تابع هدف می کند. این کار با استفاده از یک روش عددی انجام می شود که از روش های دوگان برای یافتن پاسخ بهینه استفاده می کند.

۴- توقف الگوریتم: فرآیند بهینه سازی زمانی متوقف می شود که تغییرات بین تکرارهای متوالی کمتر از یک آستانه معین باشد یا تعداد تکرارها به حداکثر برسد.

بهینه سازی سهام بر مبنای تکنیک برنامه ریزی درجه دوم روش دوگانه گلدفارب و ایدنانی (۱۹۸۲، ۱۹۸۳) برای حل مسائل برنامه ریزی درجه دوم به شکل $\min(-d^T b + \frac{1}{2} b^T D b)$ با محدودیت های $A^T b \geq b_0$ پیاده ساز شده است. (Goldfarb & Idnani, 1982, 1983)

در این روش ابتدا بردار بازدهی هدف را به عنوان قید مساوی در بهینه سازی مارکویتز استفاده نموده و سپس ماتریس واریانس و کواریانس دارایی ها مشخص می گردد و سپس با استفاده از نرم افزار مدل سازی صورت گرفته و وزن های بهینه و مقدار تابع هدف (واریانس بازده پرتفوی) در قالب لیست به عنوان خروجی داده می شود.

نتایج و تجزیه و تحلیل داده ها و ارزیابی عملکرد سبد سهام

پس از اجرای مدل های پیشنهادی در بازه زمانی مورد بررسی، نتایج به دست آمده نشان می دهد که دقت پیش بینی بازده سهام با استفاده از الگوریتم جنگل تصادفی از اطمینان قابل قبولی برخوردار بوده و همچنین معیارهای ارزیابی مدل (مانند RMSE و MAE) نشانگر همگرایی مناسب مدل به داده های واقعی می باشد.

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / دوره ۱۶ / شماره ۶۵ / زمستان ۱۴۰۴

میانگین نسبی خطای مطلق	MAE	RMSE	
۰.۰۵۲	۵۲۵.۶	۵۹۸.۸	Average

MAE	RMSE	
۴۵۷.۶	۴۸۷.۲	STDEV

سبد سهامی بهینه که از طریق ترکیب خروجی پیش‌بینی بازده و برنامه‌ریزی درجه دوم تشکیل شده، نسبت به سبدهای حاصل از رویکردهای کلاسیک دارای نسبت شارپ بهتری بوده و ریسک ناشی از نوسانات بازار را کاهش داده است.

نتایج حاصل از تحلیل‌های آماری اعتبار رویکرد ارائه شده را در شرایط واقعی بازار تأیید می‌کند و نشان می‌دهد رویکرد یادگیری ماشین در کنار مدل‌های بهینه‌سازی می‌تواند ابزار قدرتمندی برای سرمایه‌گذاران فراهم آورد.

به‌طور کلی عملکرد سبد سهام از جنبه‌های مختلفی از جمله نسبت شارپ و شاخص‌های بازدهی ارزیابی شده است. در این پژوهش، نسبت شارپ برای سنجش رابطه بین بازده اضافی و ریسک استفاده شده است. مدل پیشنهادی نسبت به سبدهای حاصل از روش‌های کلاسیک دارای نسبت شارپ بهینه‌تر و ریسک کمتری بوده است.

• شاخص‌های عملکرد از جمله معیارهای میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE) نیز در ارزیابی دقت پیش‌بینی به‌کاررفته و به نفع مدل هوشمند عمل کرده‌اند.

نتایج ارزیابی عملکرد نشان می‌دهد که در شرایط متغیر و نوسانی بازار، استفاده از رویکرد ترکیبی یادگیری ماشین و برنامه‌ریزی درجه دوم، عملکرد مطلوب‌تری را نسبت به روش‌های کلاسیک ارائه می‌دهد.

بحث و نتیجه‌گیری

نتایج به‌دست‌آمده از این پژوهش حاکی از آن است که رویکرد ترکیبی مبتنی بر الگوریتم جنگل تصادفی و برنامه‌ریزی درجه دوم، نسبت به روش‌های کلاسیک مانند پرتفوی حداقل واریانس عملکرد بهتری از نظر کاهش ریسک و افزایش بازده دارد. نکات مهمی که در این بخش مورد بحث قرار می‌گیرند عبارت‌اند از:

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

- اهمیت استفاده از داده‌های به‌روز و رویکردهای نوین در پیش‌بینی روندهای آتی سهام
- کاهش وابستگی به داده‌های تاریخی که در مدل‌های سنتی به چشم می‌خورد و افزایش انعطاف‌پذیری در مواجهه با نوسانات بازار

- استفاده از الگوریتم‌های هوشمند جهت استخراج الگوهای پنهان و بهبود دقت پیش‌بینی
- اهمیت تعیین وزن بهینه در ترکیب سهام منتخب جهت کاهش ریسک غیر سیستماتیک سبد و افزایش پایداری بازدهی

با توجه به نتایج به‌دست‌آمده، می‌توان نتیجه گرفت که رویکرد ارائه‌شده نه تنها از نظر فنی بلکه از منظر کاربرد عملی نیز نقش مهمی در تسهیل فرآیندهای تصمیم‌گیری سرمایه‌گذاران ایفا خواهد کرد. نتایج پژوهش حاضر نشان‌دهنده اهمیت و کارایی استفاده از فناوری‌های نوین مالی به‌ویژه الگوریتم‌های یادگیری ماشین در مدیریت سرمایه‌گذاری است. نتیجه‌گیری کلی به شرح زیر بیان می‌شود:

الگوریتم جنگل تصادفی با استفاده از ویژگی‌های انتخابی و تکنیک‌های نمونه‌گیری، توانسته است پیش‌بینی دقیقی از تغییرات و بازده سهام ارائه دهد.

ترکیب خروجی مدل پیش‌بینی با برنامه‌ریزی درجه دوم در تعیین وزن بهینه سهام، منجر به کاهش ریسک و بهبود عملکرد سبد سرمایه‌گذاری گردیده است.

مقایسه با روش‌های سنتی نشان می‌دهد که رویکرد هوشمند پیشنهادی از نظر سوددهی و کاهش ریسک، عملکرد بهتری از خود به جای گذاشته است.

با توجه به نتایج حاصل، استفاده از این رویکرد می‌تواند به‌عنوان یک ابزار تصمیم‌گیری علمی برای سرمایه‌گذاران در بازارهای بورس به کار رود.

با وجود محدودیت‌های موجود، پژوهش ارائه‌شده زمینه‌های توسعه بیشتر و کاربرد گسترده‌تر این روش‌ها را فراهم می‌آورد و می‌تواند گامی نوین در توسعه سیستم‌های مدیریت سرمایه‌گذاری باشد.

- 1) Breiman, L. (2001). Random Forests. Machine Learning, 45.
- 2) Breiman, L. (2002). Manual On Setting Up, Using, And Understanding Random Forests V3.1.
- 3) Goldfarb, D., & Idnani, A. (1982). Dual and Primal-Dual Methods for Solving Strictly Convex Quadratic Programs. Numerical Analysis, Springer-Verlag, Berlin, 226–239.
- 4) Goldfarb, D., & Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. Mathematical Programming, 27.
- 5) Jain, P., & Jain, S. (2019). Can Machine Learning-Based Portfolios Outperform Traditional Risk-Based Portfolios? The Need to Account for Covariance Misspecification Risks.
- 6) Kaczmarek, T., & Perez, K. (2021). Building portfolios based on machine learning predictions. Economic Research-Ekonomska Istraživanja.
- 7) Monokchian-sharifabad, k. (2021). Survey and analysis of the public knowledge of new financial technologies (fintechs) and expectations in this area. International Journal of Financial Technology (IJFT), 2.
- 8) Monokchian-sharifabad, K., Mohamadi, A & Bigvand, A. (2021). A review and analysis of conducted researches in the field of economic resilience and the role of Fin-Tech in this field. International Journal of Financial Technology (IJFT), 1.
- ۹) اقتصاد، ا. ع. و محمدی، ع. (۱۴۰۲). بهینه‌سازی سبد سرمایه‌گذاری به کمک پیش‌بینی بازده مورد انتظار با استفاده از روش‌های شبکه عصبی LSTM، جنگل تصادفی و ARIMA چشم‌انداز مدیریت مالی، ۴۳.
- ۱۰) بیات، ع. و اسدی، ل. (۱۳۹۶). بهینه‌سازی پرتفوی سهام: سودمندی الگوریتم پرندگان و مدل مارکویتز. مهندسی مالی و مدیریت اوراق بهادار، ۳۲.
- ۱۱) راستین، ف. (۱۳۹۸). مقایسه مدل‌های انتخاب سبد بهینه سهام مارکویتز و مدل محدودیت ال با روش یادگیری ماشین. (کارشناسی ارشد)، موسسه آموزش عالی حکیم طوس.

مدیریت سرمایه‌گذاری در بورس.../منکچیان شریف آباد، رهنمای رودپشتی، میرعرب و یزدانیان

۱۲) رهنمای-رودپشتی، ف. و حسینی، ز. ا. (۱۳۸۹). تبیین قیمت‌گذاری دارایی سرمایه‌ای: مقایسه تطبیقی مدل‌ها. بررسی‌های حسابداری و حسابرسی، ۶۲.

۱۳) سرچمی، م.، پور، ا. خ.، محمدی، م. و زینلی، ح. (۱۳۹۹). بهکارگیری مدل‌های یادگیری ماشین در تشکیل پرتفوی بهینه سهام و مقایسه کارایی آن‌ها. مهندسی مالی و مدیریت اوراق بهادار، ۴۵.

۱۴) صندوق-بازنشستگی-کشوری. (۱۳۹۶). آینده بازارهای خدمات مالی: هلدینگ بانک، بیمه و سهام صندوق بازنشستگی کشوری.

۱۵) غلامیان، ا. و داودی، س. م. ر. (۱۳۹۷). پیش‌بینی روند قیمت در بازار سهام با استفاده از الگوریتم جنگل تصادفی. مهندسی مالی و مدیریت اوراق بهادار، ۳۵.

۱۶) کامرو، س. (۱۳۹۷). ارائه استراتژی ترکیبی معاملات سهام و تشکیل سبد سهام در بورس اوراق بهادار با استفاده از ترکیب شیوه‌های نوین داده‌کاوی و تحلیل فنی. (کارشناسی ارشد)، دانشگاه صنعتی شریف

- 1 overfitting
- 2 Random Forest
- 3 invertibility
- 4 crude
- 5 inverse volatility
- 6 out-of-sample
- 7 Random Forest
- 8 Classification
- 9 Regression
- 10 Decision Trees
- 11 Bagging
- 12 Sampling with Replacement
- 13 Node
- 14 Features
- 15 Dominant Features
- 16 Overfitting
- 17 Bootstrap Sampling
- 18 Sampling with Replacement
- 19 Out-of-Bag - OOB
- 20 Root Mean Squared Error
- 21 Mean Absolute Error
- 22 Quadratic Programming

Investment Management in the Stock Market Using a Machine Learning Approach for Stock Selection and Its Application as FinTech

Kamyar monokchian sharif abad¹

Fereydoon Rahnama Roodposhti²

Receipt: 12/03/2025 Acceptance: 16/05/2025 Seyyed Alireza Mirarab Baygi³

Narges Yazdanian⁴

Abstract

This study explores innovative investment management strategies in the stock market, with a particular focus on leveraging machine learning approaches to enhance investors' decision-making processes. Given the unprecedented growth in the volume of financial data and the challenges arising from the declining efficiency of traditional forecasting methods, the adoption of intelligent algorithms has emerged as an effective tool for predicting stock returns. This research begins by examining the fundamental concepts of the stock market, the significance of financial time series data analysis, and the critical role of big data in providing accurate predictions. Additionally, the necessity of employing cutting-edge financial technologies to improve the performance of investment systems is emphasized. Subsequently, the Random Forest algorithm, recognized as a prominent machine learning technique, is introduced, and the process of integrating it with a quadratic programming model for constructing an optimal stock portfolio is described in detail. This hybrid model, which capitalizes on the selective features of the algorithm, offers an innovative approach for more precise estimation of future stock returns and optimal weight allocation within the portfolio. The results of statistical analyses demonstrate that the proposed approach outperforms classical methods, such as the minimum variance portfolio, in terms of risk reduction and return enhancement. The scientific validity of the model is further confirmed through widely accepted statistical metrics. Beyond providing a practical framework for achieving optimal investment outcomes, this research serves as a valuable reference for researchers and investors seeking to harness machine learning technologies and advanced techniques in the fields of investment management and fintech. The introduction of this novel methodology paves the way for the development of automated decision-making systems, enhances market transparency, and strengthens competitive capabilities in financial markets.

Keywords:

Investment Management, Stock Exchange, Machine Learning, FinTech

1-Department of Management and Accounting, Roudehen Branch, Islamic Azad University, Roudehen, Iran. kamyar.mn91@gmail.com

2-Department of Finance and Accounting, Science and Research Branch, Islamic Azad University, Tehran, Iran. (Corresponding Author) rahnama.roodposhti@gmail.com

3-Department of Management and Accounting, Roudehen Branch, Islamic Azad University, Roudehen, Iran. mirarab_alireza@yahoo.com

4-Department of Management and Accounting, Roudehen Branch, Islamic Azad University, Roudehen, Iran. nargesyazdanian@gmail.com