



بهینه سازی سبد سهام با استفاده از یادگیری تقویتی عمیق مبتنی بر مدل شبکه عصبی کانولوشن

ساناز دلیلی^۱

یاسر رضائی پیتته نوئی^{۲*}

محمدتقی کبیری^۳

مهدی صفری گرایلی^۴

میثم عرب زاده^۵

تاریخ دریافت: ۱۴۰۳/۰۱/۲۷ تاریخ پذیرش: ۱۴۰۳/۰۳/۰۱

چکیده

مدلهای کلاسیک بهینه سازی پرتفوی، زمانی که محدودیت های بازار سرمایه و معیارهای تشکیل پرتفوی کم باشد قابل استفاده هستند. ولی با توجه به واقعیت های بازار سرمایه و ضعف مدل های کلاسیک در کشف روابط پیچیده و غیرخطی میتوان پی برد که مسئله پیچیده تر و فراتر از مدل های کلاسیک است. در حل مسائل این چنین پیچیده، علوم بین رشته ای، کامپیوترها و مدل های مبتنی بر یادگیری ماشین به کمک انسانها آمده اند و رویکردی موفق در حل این مسائل داشته اند. اخیراً تکنیک های یادگیری عمیق، همچون شبکه عصبی بازگشتی برای بهینه سازی سبد سهام طراحی شده اند. در این پژوهش، یک مدل شبکه عصبی کانولوشن یادگیری تقویتی عمیق، برای بهینه سازی سبد سهام بر اساس داده های تاریخی پیشنهاد شده است. بدین منظور، از بازدهی های روزانه 20 شرکت پذیرفته شده در بورس اوراق بهادار تهران که دارای نقدینگی بالا در بازه زمانی 1392 تا 1401 بودند، استفاده شده است. نتایج به دست آمده از این تحقیق نشان می دهد بر اساس ارزیابی بازده پرتفوی، شاخص ریسک و شاخص شارپ و اینکه در همه معیارها نتایج به دست آمده پرتفوی مدل شبکه عصبی کانولوشن یادگیری عمیق در حد مطلوبی بوده است، می توان بیان داشت مدل شبکه عصبی کانولوشن یادگیری عمیق دارای توانایی مطلوبی در تشکیل پرتفوی بهینه سهام می باشد.

کلمات کلیدی: بهینه سازی، سبد سهام، یادگیری تقویتی عمیق، شبکه عصبی کانولوشن.

^۱ دانشجوی دکتری گروه مهندسی مالی، واحد کاشان، دانشگاه آزاد اسلامی، کاشان، ایران. Sanazdalili2626@gmail.com

^۲ دانشیار گروه حسابداری، دانشکده مدیریت و اقتصاد، دانشگاه گیلان، رشت، ایران. (نویسنده مسئول). rezaei.yasser@gmail.com

^۳ استادیار گروه حسابداری، دانشکده علوم اداری و اقتصادی، دانشگاه اراک، اراک، ایران. m-kabiri@araku.ac.ir

^۴ دانشیار گروه حسابداری، واحد بندرگز، دانشگاه آزاد اسلامی، بندرگز، ایران. Mehdi.safari83@yahoo.com

^۵ استادیار گروه حسابداری، واحد کاشان، دانشگاه آزاد اسلامی، کاشان، ایران. Arabzadeh62@yahoo.com

۱-مقدمه

یکی از مسائل اساسی در مدیریت مالی موضوع اداره صحیح فرآیند سرمایه گذاری است که با عنوان بهینه سازی سبد سرمایه گذاری یا بهینه سازی پرتفوی شناخته می شود. منظور از سبد سرمایه گذاری یا اصطلاحاً پرتفوی و یا پرتفولیو، انواع دارایی هایی است که یک سرمایه گذار در آن سرمایه گذاری کرده است (ژانگ^۱، ۲۰۲۳). هراندازه تعداد سهام موجود در سبد سرمایه گذاری یک سرمایه گذار بیشتر باشد ضرورت دارد که سرمایه گذاری زمان بیشتری را برای تحلیل و بررسی این سهام اختصاص بدهد. ازاین رو، مدیریت پرتفوی که اصطلاحاً به آن سبد گردانی نیز گفته میشود برای افزایش مطلوبیت مورد انتظار سرمایه گذار ضروری است (کاراندنت و اسپرلی^۲، ۲۰۲۴). به علاوه، این موضوع برای پیشرفت اقتصادی کشورها به خصوص کشورهای درحال توسعه اهمیت زیادی دارد توسعه اقتصادی بدون سرمایه گذاری امکانپذیر نیست (پیاو^۳، ۲۰۲۳). هدایت درست و تخصیص بهینه منابع مالی سبب افزایش تولید و رشد ناخالص ملی، ایجاد شغل و افزایش رفاه عمومی می شود و بازارهای مالی ازجمله بازار بورس اوراق بهادار نقش مهمی در این زمینه ایفا می کنند. این بازارها مهم ترین جایگاه تخصیص منابع مالی به فعالیت های اقتصادی در سراسر دنیا به شمار می روند. ازاین رو توسعه ساز و کارهای اثربخش برای مدیریت بهینه فرآیند سرمایه گذاری در این بازارها از اهمیت زیادی برخوردار است (شارما و شخاوات^۴، ۲۰۲۲). در این راستا، شرکتهای سرمایه گذار و موسسات مالی نقش بسیار

پررنگ تری ایفا می کنند ازاین رو، بهینه سازی فرآیند سرمایه گذاری شرکت سرمایه گذار در این بازارها باعث گردش بهتر چرخ های تولید ملی و اقتصاد کشورها خواهد شد. علاوه بر این، فعالیتهایی ازاین دست حاشیه سود شرکت سرمایه گذار را نیز افزایش می دهد، اما این کار مستلزم کنترل ریسک موجود در این فرآیندها و مدیریت بهینه آن است (اسید^۵ و همکاران، ۲۰۱۸، سن^۶ و همکاران، ۲۰۲۱). در این راستا، شناسایی و به کارگیری شیوه های بهینه جهت افزایش بازده و کاهش ریسک ناشی از ورود شرکت سرمایه گذار به این قبیل فعالیتهای ضروری و اجتناب ناپذیر است. انتخاب سبد سهام جهت کسب بالاترین بازده یکی از مهمترین دغدغه های سرمایه گذران میباشد. انتخاب پرتفوی تصمیمی است که بهترین مجموعه دارایی های مالی را انتخاب میکند. با وجود اینکه ابزارهای مختلف و تکنیک های پیشرفته برای رسیدگی به مشکلات سرمایه گذاری وجود دارد؛ اما هنوز عدم قطعیت روند را اداره میکند و حتی برای متخصصان واجد شرایط نیز مشکل آفرین است (ژانگ^۷، ۲۰۲۳).

روشهای متعددی از قبیل مدل های کنترل تصادفی، مدل های برنامه ریزی چند هدفه، تحلیل تفکیک کننده، شبکه های عصبی و مدل های بهینه سازی توسط محققان مختلف مورد استفاده قرار گرفته است که وجه تشابه همه، گرایش به مدل های چند معیاره است. منظور از سبد سهام عادی تشکیل سبد سهام بهینه با بازدهی مناسب از بین تمام شرکت ها و صنایع موجود در بازار سرمایه است که

^۱ Zhang

^۲ Carandente & Sperli

^۳ Piao

^۴ Sharma, Meeta & Shekhawat

^۵ Essid

^۶ Sen

^۷ Zhang

محدودیتی در انتخاب شرکت یا صنعت در تشکیل سبد سهام نداریم. تلاش های عمده ای جهت ارائه روشهای نوین تجزیه و تحلیل بازارهای مالی در کنار روشهای گذشته برای تشکیل پرتفوی بهینه انجام شده است. انتخاب سبد سهام جهت کسب بالاترین بازده یکی از مهمترین دغدغه های سرمایه گذاران میباشد. انتخاب پرتفوی تصمیمی است که بهترین مجموعه دارایی های مالی را انتخاب میکند. با وجود اینکه ابزارهای مختلف و تکنیک های پیشرفته برای رسیدگی به مشکلات سرمایه گذاری وجود دارد؛ اما هنوز عدم قطعیت روند را اداره میکند و حتی برای متخصصان واجد شرایط در این پروژه، کاربرد عوامل اقتصاد کلان نظیر قیمت کالا، تاریخچه بازار، و..... برای پیش بینی پرتفوی استفاده خواهیم کرد. هوش مصنوعی واژه ای است که برای بیان طیف وسیعی از موضوعات استفاده میشود و بیش از آنکه یک فناوری خاص باشد، همه فناوری هایی که دارای جنبه ای از هوش باشند را در برمیگیرد. در مقابل یادگیری ماشین به حوزه خاصی اشاره دارد. به عبارت دیگر یادگیری ماشین برای اشاره به مجموعه فناوری های خاصی از هوش مصنوعی استفاده میشود. یادگیری ماشین نیز شامل شاخه های متفاوتی است که یکی از آنها یادگیری عمیق است پیش بینی، گزاره ای درباره یک رویداد در آینده است. پیش بینی قیمت سهام در بازار های مالی برای خرید و یا فروش سهام نقش مهمی را ایفا میکند. ازاین در این پژوهش به ارائه یک مدل علمی برای بهینه سازی سبد سهام با استفاده از یادگیری تقویتی عمیق پرداخته می شود.

۲- مبانی نظری و پیشینه پژوهش

دو مولفه مهم در تصمیم گیری برای سرمایه گذاری، میزان ریسک و بازده دارایی های سرمایه ای است. انتخاب مجموعه دارایی بهینه اغلب با تبادل بین ریسک

و بازده صورت می گیرد و هرچه ریسک مجموعه دارایی بیشتر باشد سرمایه گذاران انتظار دریافت بازده بالاتری خواهند داشت (وو^۱، ۲۰۲۳). شناسایی مرز کارایی مربوط به سبد دارایی ها این امکان را به سرمایه گذاران می دهد که بر اساس تابع مطلوبیت و درجه ریسک گریزی و ریسک پذیری خود، بیشترین بازده مورد انتظار از سرمایه گذاری خود به دست می آورند (اسید^۲ و همکاران، ۲۰۱۸، سن^۳ و همکاران، ۲۰۲۱). هر یک از سرمایه گذاران بر مبنای ریسک پذیری و ریسک گریزی خود، نقطه ای را بر روی مرز کارا انتخاب کرده و ترکیب پرتفوی خود را هدف حداکثر کردن بازده و کمینه کردن ریسک تعیین می کند. بهینه سازی پرتفوی عبارت است از انتخاب بهترین ترکیب از دارایی های مالی به نحوی که باعث شود، تا حد ممکن بازده پرتفوی سرمایه گذاری حداکثر و ریسک پرتفوی حداقل شود. ایده اساسی نظریه مدرن پرتفوی این است که اگر در دارایی هایی که به طور کامل با هم همبستگی ندارند سرمایه گذاری شود؛ ریسک آن دارایی ها یکدیگر را خنثی کرده و می توان یک بازده ثابت با ریسک کمتر به دست آورد (کاراندنت و اسپرلی^۴، ۲۰۲۴).

۲-۱- مفهوم بهینه سازی سبد سهام

مفاهیم بهینه سازی سبد سهام و تنوع بخشی به مثابه ابزاری در راستای توسعه و فهم بازارهای مالی و تصمیم گیری مالی در آمده اند. انتشار نظریه پرتفوی سهام هری مارکوویتز، اصلی ترین 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 سرمایه گذاری و سبد سهام ایجاد کرد و به عنوان ابزاری کارا برای بهینه سازی سبد سهام به کار گرفته شده است (اسید^۵ و همکاران، ۲۰۱۸، سن^۶ و همکاران،

^۴ Carandente & Sperli

^۵ Essid

^۶ Sen

^۱ Wu

^۲ Essid

^۳ Sen

۲۰۲۱، مهلووات^۱ و همکاران، ۲۰۱۸). مارکوویتز پیشنهاد کرد که سرمایه گذاران ریسک و بازده را به صورت توامان در نظر میگیرند و میزان تخصیص سرمایه بین فرصتهای سرمایه گذاری گوناگون را بر اساس تعامل بین این دو انتخاب نمایند. یکی از مباحث مهمی که در بازارهای سرمایه مطرح است و باید مورد توجه سرمایه گذاران اعم از اشخاص حقیقی یا حقوقی قرار گیرد، بحث انتخاب سبد سرمایه گذاری بهینه می باشد و در این رابطه، بررسی و مطالعه سرمایه گذاران در جهت انتخاب بهترین سبد سرمایه گذاری با توجه به میزان ریسک و بازده آن انجام می شود. معمولاً فرض بر این است که سرمایه گذاران ریسک را دوست ندارند و از آن گریزانند و همواره در پی آن هستند تا در اقلامی از دارایی ها سرمایه گذاری کنند که بیشترین بازده و کمترین ریسک را داشته باشند. به عبارت دیگر، سرمایه گذاران به بازده سرمایه گذاری به عنوان یک عامل مطلوب می نگرند و به واریانس بازده ها (ریسک) به عنوان یک عنصر نامطلوب نظر دارند (شارما و شخاوات^۲، ۲۰۲۲، کرکه آبادی و کلهری، ۱۳۹۹). در بهینه سازی پرتفوی مسئله اصلی انتخاب بهینه دارایی ها و اوراق بهاداری است که با مقدار مشخصی سرمایه میتوان تهیه کرد اگر چه کمینه کردن ریسک و بیشینه نمودن بازده سرمایه گذاری به نظر ساده می رسد اما در عمل روش های متعددی برای تشکیل پرتفوی بهینه به کار رفته است مسئله بهینه سازی مارکوویتز و تعیین مرز کارای سرمایه گذاری، زمانی که تعداد دارایی های قابل سرمایه گذاری و محدودیت های موجود در بازار کم باشد، توسط مدل های ریاضیات حل شدنی است اما هنگامی که شرایط و محدودیت های دنیای واقعی در نظر گرفته شود، مسئله پیچیده و مشکل خواهد بود، سال هاست که در حل چنین مسائل پیچیده ای 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 رساند. از جمله روش هایی که در سالهای اخیر در حل بسیاری مسائل بهینه سازی، گره گشای ابهامات بشر بوده است و در پاسخ به مسائل پیچیده رویکردی موفق داشته است، روش ها و الگوریتم های موسوم به ابتکاری (فراابتکاری نیز نامیده میشوند) است (چویی^۳، ۲۰۲۲).

۲-۲- بهینه سازی کلاسیک (مدل مارکوویتز)

مارکوویتز کسی بود که مفهوم تنوع بخشی در سبد سهام را معرفی کرد و آن را توسعه داد. او به طور کلی نشان داد که چگونه تنوع بخشی در سبد سرمایه، ریسک آن را برای سرمایه گذار کاهش می دهد. سرمایه گذاران می توانند سبد سهام کارا را به ازای یک بازده معین و از طریق کمینه کردن ریسک سبد سهام به دست آورند. در ادامه فرآیند فوق می تواند منجر به تشکیل سبدهای کارا شود که اصطلاحاً مرز کارای میانگین واریانس نامیده می شود (ما^۴ و همکاران، ۲۰۲۱).

برای استفاده از مدل مارکوویتز به داده های زیر نیاز است: -بازده مورد انتظار مربوط به سهم i که با $E(R)$ نشان داده می شود.

-انحراف معیار بازده مورد انتظار مربوط به سهم i ام که به عنوان معیاری برای ریسک هر سهم در نظر گرفته شده با S_i نمایش داده می شود.

-کواریانس، به عنوان معیار همراهی و ارتباط حرکتی بین نرخ های بازدهی سهام مختلف که با علامت S_{ij} نشان داده می شود.

علت این که سهام یک شرکت به عنوان یک دارایی ریسکی مطرح می شود، ثابت نبودن (تصادفی بودن) نرخ بازدهی کل (هفتگی، ماهیانه، سالانه) آن است. چون این

^۳ Cui

^۴ Ma

^۱ Mehlatat

^۲ Sharma, Meeta & Shekhawat

آوردن استشنا تابع Q بصورت ذیل حداکثر می کند (کاراندنت و اسپرلی، ۲۰۲۴):

رابطه (۱)

$$Q(s_t, a_t) = E[r(s_t, a_t, s_{t+1}) + \mu \max Q(s_{t+1}, a_{t+1})]$$

الگوریتم یادگیری تقویتی عمیق که از یک شبکه عصبی استفاده می کند و تابع Q بالا را تقریب می زند، شبکه عمیق (DQN) نامیده می شود. با DQN ، حالت به عنوان یک تابع مقدار کدگذاری می شود. با این حال، استفاده از رویکرد DQN دشوار است زیرا مقدار محاسبات زمانی که فضای عمل بزرگ است افزایش می یابد. وزن پرتفوی در بورس گسسته نیست و دارای فضای اکشن به صورت یک جعبه پیوسته است که منجر به نفرین می شود که تعداد فضاهای اکشن را بی نهایت افزایش می دهد. بنابراین، DQN برای بهینه سازی پورتفوی مناسب نیست. بسیاری از الگوریتم های تقویتی یادگیری در مدیریت پورتفوی استفاده شده است. برخی از مطالعات از یادگیری عمیق Q استفاده کرده اند (چن^۴ و همکاران، ۲۰۲۰). برخی از مطالعات از بهینه سازی خط مشی پروگزیمال استفاده کرده اند. در میان آنها، شیب سیاست قطعی عمیق^۵ به طور گسترده مورد استفاده قرار می گیرد. شیب سیاست قطعی عمیق یک الگوریتم یادگیری تقویتی است که یادگیری Q و گرادیان خط مشی را ترکیب می کند. شیب سیاست قطعی عمیق یک نسخه بهبود یافته از الگوریتم گرادیان خط مشی قطعی است. اولاً، شیب سیاست جبری الگوریتمی است که خط مشی بهینه را با ترکیب یادگیری کیو با شیب خط مشی پیدا می کند. در یافتن چنین خط مشی، شیب سیاست قطعی عمیق از یک شبکه عصبی برای به روز رسانی پارامتر با استفاده از تقریب مداوم استفاده می کند (ما و همکاران، ۲۰۲۰).

نرخ ها با توجه به زمان تغییر می کنند، لذا می توان برای آن ها تابع توزیع احتمال تشکیل داد و معیارهای مورد نیاز مدل مارکویتز نظیر، میانگین، انحراف معیار، کواریانس و را از آن به دست آورد (شارما و سخاوات^۱، ۲۰۲۲).

مدل مارکویتز بر مبنای مفروضات زیر بیان شده است: سرمایه گذاران ریسک گریزند و دارای مطلوبیت مورد انتظار افزایشی می باشند و منحنی مطلوبیت نهایی ثروت آن ها کاهنده است. سرمایه گذاران سبد سرمایه خود را بر مبنای میانگین واریانس مورد انتظار بازدهی انتخاب می کنند بنابراین منحنی های بی تفاوتی آن ها تابعی از نرخ بازده و واریانس مورد انتظار می باشد (کاراندنت و اسپرلی^۲، ۲۰۲۴، فتاحی نافچی و همکاران، ۱۳۹۸).

۲-۳- یادگیری تقویتی عمیق

یادگیری تقویتی عمیق برای یادگیری انسان استفاده شده است (اسید و همکاران، ۲۰۱۸). هنگامی که یک فرد برای اولین بار دوچرخه سواری را یاد می گیرد، می آموزد که چگونه هنگام دوچرخه سواری بدون در نظر گرفتن تکانه زاویه ای، تکانه و نیرو، دوچرخه سواری کند. به طور مشابه، یادگیری تقویتی عمیق یک مسئله بهینه سازی را طبق یک قانون خاص حل نمی کند، اما عاملی که در یک محیط خاص تعریف شده است وضعیت فعلی را تشخیص می دهد و پاداش را در بین اقدامات به حداکثر می رساند (وودساید - اوریاخی^۳ و همکاران، ۲۰۱۱).

یادگیری کیو (Q) اساس یادگیری تقویتی است. یادگیری Q بر اساس تابع Q است. تابع Q اشاره دارد به تابع پاداشی که عامل زمانی که یک عامل عملی را در یک حالت خاص انجام می دهد دریافت می کند. برای به روز رسانی تابع Q در یادگیری Q ، از عمل حریصانه $a(t+1)$ استفاده می شود که $Q(s_{t+1})$ را برای بدست

^۴ Chen

^۵ DDPG

^۱ Sharma, Meeta & Shekhawat

^۲ Carandente & Sperli

^۳ Woodside-Oriakhi

شیب سیاست قطعی عمیق روش خارج از سیاست را برای به روز رسانی شبکه های بازیگر و انتقادی به طور جداگانه اتخاذ کرده و استفاده می کند. شبکه کنشگر $\mu(s|\theta\mu)$ حالت ها را به اقداماتی که $\theta\mu$ مجموعه پارامترهای شبکه عامل است، نگاشت می کند، و شبکه انتقادی $Q(s, a|\theta)$ مقدار عمل را در آن حالت، که θ و Q مجموعه پارامترهای شبکه بحرانی هستند. از آنجایی که کاوش برای اقدام بهتر در یادگیری تقویتی ضروری است، نويز تصادفی N اضافه شده و در شبکه بازیگر استفاده می شود. در شبکه منتقد، از بافر تکرار R برای صرفه جویی در آموزش و به روز رسانی موثر مدل استفاده می شود.

علاوه بر این، شیب سیاست قطعی عمیق از به روز رسانی های هدف نرم استفاده می کند. به روز رسانی هدف نرم به آرامی پارامترهای شبکه را در به روز رسانی شبکه های بازیگر و منتقد به روز می کند. در یک روش یادگیری عمیق عمومی، پارامترهای هر لایه یک بار آموزش داده می شوند و به طور مداوم بدون ذخیره به روز می شوند، اما شیب سیاست قطعی عمیق با یک بار استفاده مجدد از پارامترهای آموخته شده، به آرامی پارامترهای لایه شبکه را تغییر می دهد. این یادگیری مزیت افزایش ثبات یادگیری را دارد. وظیفه بعدی بهینه سازی وزن هر نام برای دستیابی به حداقل ریسک پیش فرض و در عین حال به حداکثر رساندن بازده پرتفوی هدف است. اگرچه راه حل سنتی شکل بسته بسیار کارآمد است، اما زمانی که محدودیت های غیر محدب اعمال می شود، با مشکلاتی مواجه می شود (تهرانی و همکاران، ۱۳۹۷، آرل^۱ و همکاران، ۲۰۱۰). و محدودیت ها معمولاً در بازار مالی واقعی استفاده می شوند. بنابراین، روش اکتشافی، یک رویکرد عددی مهم که از یک کلاس از الگوریتم ها تشکیل شده است، به یک راه حل معقول تبدیل می شود. چنین روشهای اکتشافی شامل استفاده از الگوریتم ژنتیک،

جستجوی تابو، بازپخت شبیه سازی شده، الگوریتم کلونی مورچه ها و بهینه سازی ازدحام ذرات است. هر الگوریتم با تکامل طبیعی تقویت می شود. اگرچه همه روش ها در پایان به یک راه حل تقریباً بهینه منتهی می شوند، اما هر الگوریتم در حل مسائل خاص مزیتی دارد. با این حال، کارایی کد نویسی و سرعت کامپیوتر نیز از عوامل حیاتی هستند و گاهی حتی از انتخاب الگوریتم نیز تاثیرگذارتر هستند (کاراندنت و اسپرلی^۲، ۲۰۲۴).

مارکوویتز توانست برای نخستین بار مفهوم ریسک را به صورت کمی ارا نه کند. این مدل سرآغاز تحول عظیم در نظریه سرمایه گذاری بود، اما ضعف هایی داشت که در مدل های توسعه یافته بعدی برطرف شد (دنگ و یو^۳، ۲۰۱۳). اولین تغییر مشهور و مطرح شده در این رابطه، شاخصی بود که ویلیام شارپ، برای اندازه گیری ریسک با رویکرد حساسیت محور مطرح کرد. این شاخص که با نام ضریب حساسیت یا بتا شناخته می شود، به شارپ امکان معرفی مدل تک عاملی را داد (شارپ، ۱۹۸۷). پس از مدل شارپ، مدل قیمت گذاری دارایی های سرمایه ای راس معرفی شد که مشکل در نظر گرفتن منابع غیربازاری ریسک در مدل شارپ را برطرف کرد. پس از این مدل ها، عمده مطالعات در زمینه مدل سازی بهینه سازی سبد سهام، وارد مرحله جدیدی شد و بیشتر افراد تلاش کردند با افزودن محدودیت هایی به مدل های یادشده، آنها را به دنیای واقعی نزدیک تر کنند. از اصلی ترین محدودیت های اضافه شده به مدل، محدودیت تعداد مشخص دارایی در سبد سهام بود (وودساید اوریاخی^۴ و همکاران، ۲۰۱۱). بیل و فارس (۱۹۷۶) نیز محدودیت حداقل و حداکثر میزان سرمایه در یک دارایی را مطرح کردند. امروزه بیشتر این محدودیت ها به مدل اولیه مارکوویتز اضافه شده و مدلی را شکل داده است که مدل میانگین -

^۳ Deng & Yu

^۴ Woodside-Oriakhi

^۱ Arel

^۲ Carandente & Sperli

واریانس با مؤلفه های مقید نام دارد و در بسیاری از پژوهش ها استفاده می شود.

محدودیت های اضافه شده به مدل، به غیرخطی شدن فضای مسئله بهینه سازی سبد سهام منجر شد و حل آن را بسیار مشکل کرد؛ به نحوی که در موارد بسیاری، این مسئله تبدیل به مسئله ان - پی. سخت شد و حل آن توسط روش های مبتنی بر مشتق و ریاضیات امکان پذیر نبود. در این راستا، می توان به مطالعه شاولیو و کوپمن^۱ (۲۰۰۸)، اشاره کرد که تلاش کردند باروش لاگرانژ، مسئله یاد شده را حل کنند، اما با شکست مواجه شدند. تحقیقات ویلما، احمد و ناماسر (۲۰۰۸)، برتسیما و شیودا^۲ (۲۰۰۹) و گولپینار، آن و مویینی^۳ (۲۰۱۰) از جمله تحقیقات شکست خورده در زمینه حل مسئله بهینه سازی سبد سهام با روش های دقیق هستند. از این رو برای حل مدل های محدودیت دار، افراد بسیاری به سراغ الگوریتم های فراابتکاری رفتند. در تحقیقات اولیه در این زمینه، چنگ^۴ و همکارانش (۲۰۰۰)، برای نخستین بار الگوریتم هایفراابتکاری ژنتیک، تبرید شبیه سازی شده و جستجوی ممنوعه را برای حل مسئله به کار بردند. پژوهشگران دیگری مانند سلیمانی، گلمکانی و سلیمی (۲۰۰۹)، آگنوتوپولوس و مامانیس^۵ (۲۰۱۰) و هان هونگ (۲۰۱۱)، مدل های مختلفی از مسئله را با استفاده از الگوریتم ژنتیک حل کردند (قدوسی و همکاران، ۱۳۹۴). فرناندز و گومز^۶ (۲۰۰۷) درپژوهش خود الگوریتمی برمبنای شبکه های عصبی برای حل این مسئله پیشنهاد دادند و کورا (۲۰۰۹) الگوریتم اجتماع پزندگان را برای حل مسئله پیشنهاد کرد. مهلووات^۷ و همکاران (۲۰۱۸)، مدل انتخاب پرتفوی با

استفاده از تحلیل پوششی داده ها بر مبنای فازی چندمنظوره را در بورس اوراق بهادار هندوستان در دوره یکساله (پنجاه و دو هفته) ۲۰۱۷ تا ۲۰۱۸ با متغیرهای گردش دارایی، سود هر سهم، نرخ رشد سود هر سهم، نسبت اهرم، نسبت قیمت به سود و بتا موردبررسی قرارداداند. نتایج به دستآمده حاکی از آن است که با این مدل سرمایه گذار قادر است طرح وزن دهی پرتفوی را طوری انتخاب کند که رضایت بیشتری را به همراه داشته باشد. ژو^۸ و همکاران (۲۰۱۸) انتخاب پرتفوی تحت نگرشهای مختلف در محیط فازی را در بورس اوراق بهادار چین موردبررسی قرارداداند. پرتفوی را به صورت ده سهمی تشکیل دادند. نتایج تحقیق نشان میدهد که میتوان مرزهای گوناگون در نگرشها، سطح اطمینانها و ارزش در معرض خطر مختلف، به دست آورد. اسید و همکاران (۲۰۱۸) انتخاب پرتفوی با استفاده از روش بازده متوسط بازی معنیدار، در بورس اوراق بهادار پاریس از سال ۲۰۱۰ تا ۲۰۱۴ را موردبررسی قرارداداند. متغیرهای واریانس و کشیدگی به عنوان متغیر ورودی و متغیرهای متوسط بازده سالانه و چولگی به عنوان متغیر خروجی در نظر گرفته شدند. نتایج تحقیق نشان میدهد که پرتفوی تشکیل شده با روش بازده متوسط بازی معنیدار نسبت به سایر پرتفوی های تشکیل شده دارای عملکرد خوبی میباشد. اریو و همکاران (۲۰۱۷) انتخاب پرتفوی را با استفاده از روش خوشه بندی مبتنی بر *P-spline* در هشتادوشش شرکت بورس اوراق بهادار ایتالیا در بازه زمانی ۲۰۱۳ تا ۲۰۱۴ موردبررسی قرارداداند و نتایج نشان میدهد که استراتژی استفاده شده برای کمک به تصمیمات سرمایه گذاری متخصصین مالی مفید میباشد. در ایران نیز یزدانی و همکاران (۱۴۰۰) به بررسی انتخاب

^۵ Anagnostopoulos & Mamanis

^۶ Fernandez & Gomez

^۷ Mehlatat

^۸ Zhou

^۱ Shaw, D.X., Liu, S. & Kopman

^۲ Bertsimas, D., Shioda

^۳ Gulpinar, An & Moeini

^۴ Chang

بهینه سازی سبد سهام با استفاده از الگوریتم ترکیبی فرا ابتکاری پرداختند. در این پژوهش، از الگوریتم شبکه های عصبی و برنامه نویسی شبکه ژنتیک، برای تشخیص ویژگی های موثر و از درخت تصمیم $3ID$ بهبود یافته به عنوان روش پیشنهادی جهت پیشبینی قیمت و روند تغییر قیمت سهام برای انتخاب سبد بهینه استفاده شده است. نتایج تحقیق نشان می دهد، روش پیشنهادی علاوه بر کاهش سر بار محاسباتی و حافظه ای، قادر است تا با دقت بالایی به پیشبینی نوسانات شدید با الگوهای غیر خطی پرداخته و نسبت به روش های روز دنیا چون جستجوی نزدیک ترین همسایگی، رگرسیون خطی، میانگین متحرک خود همبسته و الگوریتم پیش بینی سری زمانی، بهتر عمل نماید. کرکه آبادی و کلهری (۱۳۹۹) به شناسایی و رتبه بندی عوامل موثر بر انتخاب بهینه سبد سهام با استفاده از روش AHP فازی پرداختند. با توجه به نتایج پژوهش در بین عوامل به ترتیب شاخص اهرمی (وزن مربوطه $16/0$)، ریسک پذیری $13/0$ ، سرمایه در گردش $11/0$ ، بازار $9/0$ نقدینگی، چشم انداز شرکت، توسعه صنعت، شخصیتی، سود آوری، رشد، ساختار شرکت دارای بیشترین درجه اهمیت بودند. نیکو و همکاران (۱۳۹۹) به مدل سازی انتخاب سبد بهینه سهام بر مبنای ارزیابی ریسک و رویکرد مالی رفتاری (حسابداری ذهنی) در بورس اوراق بهادار پرداختند. در این راستا پس از نظر سنجی چند مرحله ای از ۲۸ خبره و تدوین مدل پایه بر اساس روشهای کیفی، با محاسبه بازده صنایع مختلف بورسی طی ده سال به ارزیابی مقایسه ای روشهای بهینه سازی پرتفوی با توسعه روشهای محاسبه ریسک از واریانس به ارزش در معرض خطر و وارد نمودن متغیر حسابداری ذهنی سرمایه گذاران در قالب دو دوره آموزش هشت ساله و آزمایش دو ساله پرداخته است. نوآوری اصلی تحقیق در تدوین مدل ریاضی بهینه سازی با درج محدودیتهای حسابداری ذهنی و نقدشوندگی بوده و ارزیابی نهایی مدل توسعه یافته نشان دهنده کارکرد مناسب مدلسازی مذکور در عملکرد پرتفوی می باشد. نتایج حاصله نشان می دهد که عملکرد پرتفویهای حاصل از مدل مبتنی بر حسابداری ذهنی $DMSS$ و

MVO در حالت ساده یکسان عمل کرده و اختلاف معناداری ندارند. ولی پرتفوی هایی که بر اساس مدلهای مبتنی بر حسابداری ذهنی با در نظر گرفتن محدودیتهای ارزش در معرض ریسک و نقدشوندگی تشکیل شده اند عملکردی به مراتب بهتر از مدل MVO دارند. فتاحی نافچی و همکاران (۱۳۹۸) در بررسی انتخاب سبد بهینه سهام با بکارگیری اطلاعات حسابداری، اطلاعات مبتنی بر ارزش و اطلاعات کارت ارزیابی متوازن، در راستای دستیابی به اهداف پژوهش برای تشکیل سبد بهینه سهام از رویکرد کاهش ابعاد و روشهای تحلیل پوششی داده ها و ماشین بردار پشتیبان و الگوریتمهای خوشه بندی استفاده نمودند. مدل فوق در چهار مرحله باز اجرا گردید و در هر مرحله علاوه بر مولفه ریسک و بازده، معیارهای حسابداری، معیارهای مبتنی بر ارزش و معیارهای مالی و سپس غیر مالی کارت ارزیابی متوازن به عنوان ورودی مدل تهیه پرتفوی به صورت مرحله به مرحله مورد استفاده قرار گرفت. یافته های حاصل از پژوهش حاکی از آن است که معیارهای مورد استفاده در پژوهش در تهیه سبد بهینه سهام دارای محتوای اطلاعاتی بوده و اضافه شدن هر دسته معیار منجر به افزایش مطلوبیت سبد سهام می گردد. این محتوای اطلاعاتی در مورد کارت ارزیابی متوازن چشمگیرتر نیز می باشد. به طور کلی به کارگیری همزمان روشهای ترکیبی بهینه سازی و معیارهای جامع استخراج شده از گزارش های مالی سبد سهام بهینه تر و مطلوبیت بیشتر نسبت به ریسک و بازده ادبیات مارکویتز را به همراه داشت.

همان طور که بیان شد، در تحقیقات زیادی از الگوریتم های فرا ابتکاری برای حل مسئله بهینه سازی سبد سهام و پیدا کردن مرز کارا استفاده شده است؛ اما تا کنون در هیچ تحقیقی، این مسئله به کمک الگوریتم بهینه سازی مبتنی بر آموزش و یادگیری حل نشده است. با توجه به پیچیدگی مدل، در نتیجه افزودن محدودیتهای دنیای واقعی به آن، بهره مندی از الگوریتمهای فرا ابتکاری اجتناب ناپذیر است. با توجه به برتری الگوریتم بهینه سازی مبتنی بر آموزش و یادگیری

۳-۱- معیارهای سنجش ریسک

بازدهی سهام در دوره های مختلف، متغیر است و روند ثابت و یکنواختی ندارد. بنابراین، نوسان و تغییرپذیری، جزء جدانشدنی بازدهی سهام در طی زمان است. با توجه به تغییرپذیری و نوسان، بازده دوره های آتی نیز مطمئن نیستند. عدم اطمینان نسبت به بازده-های آتی سهام، سرمایه گذاری را با ریسک همراه می کند. این پژوهش تلاش می نماید با به کارگیری روش ها و سنجه های مختلف اندازه گیری ریسک، تحقیق جامعی انجام دهد. بدین منظور از روش ها و سنجه های موجود برای اندازه گیری ریسک، واریانس، انحرافات مطلق از میانگین، نیم واریانس و همچنین ارزش در معرض ریسک مشروط استفاده شده است.

۳-۱-۱- واریانس

نخستین و بدیهی ترین معیاری که برای اندازه گیری ریسک کلی سهام از آن استفاده می شود، واریانس نام دارد که به صورت زیر محاسبه می شود.

رابطه (۲)

$$\sigma_i^2 = \frac{\sum_{i=1}^N (r_i - \bar{r})^2}{N}$$

که در آن، σ_i انحراف معیار سهم i ام، r_i بازدهی سهم i ، $\bar{r}$ میانگین بازدهی ها و N تعداد دوره ها است.

۳-۱-۲- انحراف مطلق از میانگین

این شاخص، به دنبال این است که از یک سو پراکندگی را نشان دهد و از سوی دیگر، کلیه مشاهدات را در نظر داشته باشد. این شاخص ریسک به صورت زیر محاسبه می شود.

رابطه (۳)

$$MAD = \sum \frac{|r_i - \bar{r}|}{N}$$

که در آن، r_i بازدهی سهم i ام، $\bar{r}$ میانگین بازدهی ها و N تعداد دوره ها است.

تقویتی عمیق نسبت به سایر الگوریتم های فراابتکاری، در این پژوهش تلاش شده است مدل ارائه شده با استفاده از این الگوریتم حل شود.

در فرآیند پژوهش حاضر، هدف یافتن پاسخ برای پرسش زیر است:

-بهینه سازی سبد سهام با استفاده از یادگیری تقویتی عمیق مبتنی بر مدل شبکه عصبی کانولوشن چگونه است؟

۳- روش شناسی تحقیق

این پژوهش از نظر نوع، مدل سازی و از نظر روش (نحوه گردآوری داده ها)، توصیفی و از نظر هدف، کاربردی میباشد. قلمرو موضوعی این تحقیق کاربرد مدل یادگیری تقویتی عمیق در تشکیل پرتفوی بهینه میباشد. جامعه آماری این پژوهش، شرکتهای موردبررسی در بورس اوراق بهادار تهران در بازه زمانی 1392 تا 1401 میباشد. پس از گردآوری داده های موردنیاز و انجام محاسبات لازم در اکسل، مدل یادگیری تقویتی عمیق در نرم افزار آناکوندا زبان برنامه نویسی پای تون مورد آزمون قرار گرفت و سپس توانایی مدل در تشکیل پرتفوی بهینه سهام توسط معیارهای مدنظر تعیین گردید. لازم به ذکر است الگوریتم بهینه سازی بر پایه یادگیری تقویتی عمیق با چهار رویکرد اندازه گیری ریسک واریانس، انحرافات مطلق از میانگین، نیم واریانس و ارزش در معرض ریسک مشروط اجرا شده است.

در این پژوهش، داده های روزانه مربوط به قیمت های ۲۰ شرکت فعال تر بورس اوراق بهادار تهران که توسط سازمان بورس بر مبنای معیارهایی مبتنی بر نقدشوندگی، حجم و تعداد معاملات و تأثیر بر بازار منتشر می شود، استخراج شده و پس از محاسبه بازدهی های تعدیل شده آنها، به بهینه سازی سبد سهام با استفاده از الگوریتم بهینه سازی مبتنی بر آموزش و یادگیری اقدام می شود.

که در ادامه، نتایج به دست آمده از این الگوریتم ارائه شده است:

۳-۱-۳- نیم واریانس

شاخص ریسک نامطلوب نیم واریانس از مجموع انحرافات نامطلوب یا انحرافات بازدهی های کمتر از نرخ بازدهی میانگین، به دست می آید.

رابطه (۴)

$$SV = E\{\min[r_i - \bar{r}, 0]^2\}$$

که در آن، r_i بازدهی سهم i ام، $\bar{r}$ میانگین بازدهی ها و N تعداد دوره ها و T تعداد دوره هایی است که در آنها بازدهی کمتر از میانگین است.

۳-۱-۴- ارزش در معرض ریسک مشروط

ارزش در معرض ریسک، حداکثر زیان احتمالی پرتفوی را در یک دوره زمانی مشخص به صورت کمی و در قالب عدد بیان می کند. به بیان دیگر، ارزش در معرض ریسک مبلغی از ارزش پرتفوی را که انتظار می رود ظرف مدت یک دوره زمانی مشخص و با میزان احتمال معین (سطح اطمینان $1-\alpha$ درصد) از دست برود، مشخص می کند.

با توجه به کاستی ها و ضعف های ارزش در معرض ریسک، آرتزرنر با معرفی معیار ارزش در معرض ریسک احتمالی، معیاری را معرفی کرد که نارسایی های ارزش در معرض ریسک را پوشش دهد. این الگو و معیار که به نام های ریسک مورد انتظار و واریانس دنباله دار نیز مشهور است، تمام ویژگی هایی که ارزش در معرض ریسک را با کاستی هایی مواجه می کرد، به خوبی دربرمی گیرد. این معیار بدین شرح تعریف شده است: میانگین وقوع ریسک هایی که بزرگ تر و فراتر از ارزش در معرض ریسک هستند. به بیان دیگر، α درصد از میانگین توزیع بازده متغیر تصادفی بزرگ تر از ارزش در معرض ریسک ارزش در معرض ریسک احتمالی از طریق رابطه زیر برای دوره های کوتاه مدت اندازه گیری می شود.

رابطه (۵)

$$CVaR = \frac{e^{-\frac{z_\alpha^2}{2}}}{\alpha\sqrt{2\pi}}\delta_p - \bar{r}_p$$

که در آن α حداکثر خطای مد نظر، Z_α مقدار متناظر احتمال α در توزیع نرمال و p واریانس سبد مد نظر است. برای دوره های بلندمدت از رابطه ۵ استفاده می شود؛ یعنی همانند ارزش در معرض ریسک برای دوره های بلندمدت $\mu \neq 0$ را در نظر می گیریم. بنابراین داریم:

رابطه (۶)

$$CVaR = \frac{e^{-\frac{z_\alpha^2}{2}}}{\alpha\sqrt{2\pi}}\delta_p - \mu w_i = \frac{e^{-\frac{z_\alpha^2}{2}}}{\alpha\sqrt{2\pi}}\delta_p - \bar{r}_p$$

این مقدار بیان می کند که احتمال زیان در یک دوره T روزه بیش از ارزش در معرض ریسک احتمالی باشد، α درصد است (ملائی و همکاران، ۱۳۹۰).

۳-۲- مدل شبکه عصبی کانولوشن یادگیری تقویتی

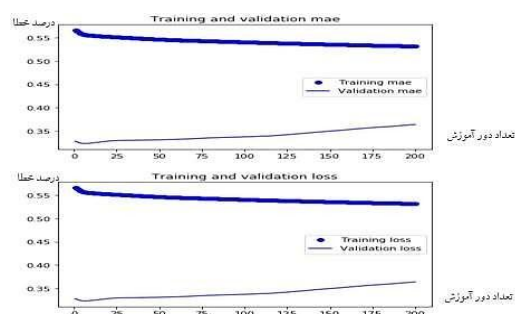
عمیق

الگوریتم بهینه سازی مبتنی بر یادگیری تقویتی عمیق، بر خلاف سایر روش های فراابتکاری که معمولاً از طبیعت الگو گرفته اند، از فرایند یادگیری در کلاس درس الگو گرفته است. عامل یادگیرنده در فرآیند یادگیری تقویتی نیاز دارد تا تابع ارزش را محاسبه و ذخیره نماید. اما در مواقعی که ابعاد مساله بزرگ باشد، ذخیره تمامی ارزشهای حالت های مختلف سیستم عملاً غیرممکن خواهد بود. در اینگونه مواقع به سراغ روشهای تخمین تابع ارزش^۱ خواهیم رفت. در روش یادگیری تقویتی، برای تخمین تابع ارزش از شبکه عصبی استفاده شود، شاخه جدیدی به نام یادگیری تقویتی عمیق^۲ شکل می گیرد که در سالهای اخیر کاربردهای یادگیری تقویتی در طیف وسیعی از حوزه ها آزمایش شده است.

² Deep Reinforcement Learning)

¹ Approximate value function

در این پژوهش پس از تغذیه داده های آموزش و تست نتیجه به صورت شکل زیر حاصل گردیده است.



شکل (۱). ارزیابی عملکرد مدل شبکه عصبی کانولوشن یادگیری تقویتی عمیق

همانگونه که در شکل های ارزیابی نیز مشخص است میزان خطا الگوریتم نزدیک به صفر میباشد که این نتیجه، یادگیری و پیش بینی دقیق الگوریتم را نشان می دهد. شبکه های عصبی یادگیری عمیق با استفاده از الگوریتم بهینه سازی شیب نزولی تصادفی آموزش داده می شوند. به عنوان بخشی از الگوریتم بهینه سازی، خطا برای وضعیت فعلی مدل باید بارها تخمین زده شود. این امر

نیاز به انتخاب یک تابع خطا دارد، تا بتوان وزنهای را به روز کرد تا خطا را کاهش دهد. برای ارزیابی عملکرد این دو مدل از تابع خطای میانگین قدر مطلق خطا استفاده شده است؛ که این مقدار خطا برای شبکه عصبی کانولوشن یادگیری عمیق 0/32.5 درصد میباشد و بسیار پایین و نزدیک به صفر میباشد.

۴- یافته های پژوهش

در این پژوهش، داده های روزانه مربوط به قیمت های شرکتهای فعال تر بورس اوراق بهادار تهران که توسط سازمان بورس بر مبنای معیارهایی مبتنی بر نقدشوندگی، حجم و تعداد معاملات و تأثیر بر بازار منتشر می شود، استخراج شده و پس از محاسبه بازدهی های تعدیل شده آنها، به بهینه سازی سبد سهام با استفاده از الگوریتم بهینه سازی مبتنی بر یادگیری تقویتی اقدام می شود.

۴-۱- آمار توصیفی

جدول (۱)، شرکت های انتخاب شده و آمار توصیفی داده های پنج ساله ای که در این پژوهش استفاده شده است را نشان می دهد.

جدول (۱). آمار توصیفی نمونه آماری پژوهش

ردیف	نماد	نام شرکت	میانگین بازدهی ها	انحراف معیار بازدهی ها
۱	بترانس	ایران ترانسفو	0.009287558	0.032070167
۲	خودرو	ایران خودرو	0.00425184	0.019303643
۳	وبصادر	بانک صادرات	0.001322262	0.016102411
۴	وبملت	بانک ملت	0.010126326	0.016016805
۵	ومعادن	توسعه معادن و فلزات	0.027382114	0.017267351
۶	خسپا	سایپا	0.008686653	0.017533661
۷	وبانک	سرمایه گذاری توسعه ملی	0.004914894	0.024499759
۸	کگل	معدنی و صنعتی گل گهر	0.078161365	0.020153504
۹	وصندوق	سرمایه گذاری صندوق بازنشستگی	0.015288538	0.024682401
۱۰	وغدیر	سرمایه گذاری غدیر	0.006221198	0.015743141
۱۱	ونفت	سرمایه گذاری صنعت نفت	0.005683537	0.023895372
۱۲	فادر	صنایع آذرب	-0.001330959	0.017081822
۱۳	فخوز	فولاد خوزستان	-0.000266248	0.01913976
۱۴	فولاد	فولاد مبارکه اصفهان	0.003971245	0.023587188

0.027841371	0.006090308	گروه بهمن	خبهمن	۱۵
0.01641613	0.001662213	گروه مینا	رمینا	۱۶
0.024007114	0.006159641	لیزینگ رایان سایپا	ولسایپا	۱۷
0.022289053	0.001240076	مخابرات ایران	اخابر	۱۸
0.024358258	0.004631849	معدنی و صنعتی چادرملو	کچاد	۱۹
0.026646099	0.013210863	ملی صنایع مس ایران	فملی	۲۰

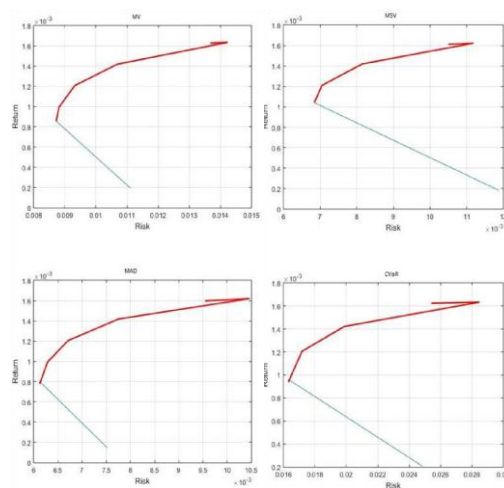
۴-۲- نتایج پیش بینی مدل

پس از آموزش مدل یادگیری عمیق، داده های تست به عنوان ورودی به مدل تغذیه شده تا هرکدام مقدار بازده سهام برای هر نمونه را پیش بینی نمایند.

۴-۲-۱- سبد دارایی بهینه به دست آمده از مدل

شبکه عصبی کانولوشن یادگیری تقویتی عمیق

با اجرای الگوریتم بهینه سازی مبتنی بر آموزش و یادگیری، یک مرز کارا شامل 10 سبد روی آن به دست آمد. با اندازه گیری شاخص شارپ این 10 سبد، بهترین سبد از لحاظ شاخص شارپ، به عنوان سبد بهینه معرفی شد. (شکل ۲)، مرز کارای به دست آمده از طریق الگوریتم بهینه سازی مبتنی بر یادگیری تقویتی عمیق را توسط هر چهار مدل مفروض نشان می دهد.



شکل (۲). مرز کارای الگوی بهینه مبتنی بر شبکه عصبی کانولوشن یادگیری تقویتی عمیق

پس از به دست آمدن مرز کارای الگوریتم برای هر یک از مدل های چهارگانه این پژوهش، شاخص شارپ سبد های ده گانه هر یک از مرز های کارا محاسبه شد و سبدهای بهینه به صورت جدول ۲ به دست آمد. همان طور که مشاهده می شود، بازدهی مورد انتظار هر چهار سبد، کمابیش یکسان است و عامل تفاوت این چهار رویکرد، به نوعی بازدهی مورد انتظار هر چهار سبد، کمابیش یکسان است و عامل تفاوت این چهار رویکرد، به نوعی ریسک اندازه گیری شده برای سبدهای به دست آمده است که با توجه به تفاوت معیار اندازه گیری ریسک در هر چهار مدل، این تفاوت مورد انتظار بود.

همانگونه که مشاهده میشود در پرتفوی مدل شبکه عصبی کانولوشن یادگیری عمیق؛ نرخ بازده بازدهی مورد انتظار در هر چهار مدل میانگین واریانس، میانگین نیم واریانس، میانگین-قدرمطلق انحرافات، و همچنین ارزش در معرض ریسک مشروط ۰,۰۰۱۲ می باشد. شاخص ریسک مدل میانگین واریانس ۰,۰۱، میانگین نیم واریانس ۰,۰۱، میانگین-قدرمطلق انحرافات ۰,۰۱ و همچنین ارزش در معرض ریسک مشروط ۰,۰۳ می باشد. شاخص شارپ برای مدل میانگین واریانس ۰,۱۴، میانگین نیم واریانس ۰,۱۶، میانگین-قدرمطلق انحرافات ۰,۱۹ و همچنین ارزش در معرض ریسک مشروط ۰,۰۸ می باشد. سرمایه گذاران با تشکیل پرتفوی به دنبال سودآوری بیشتر هستند بنابراین میتوان نتیجه گرفت که مدل شبکه عصبی کانولوشن یادگیری عمیق توانایی مطلوبی در تشکیل پرتفوی بهینه دارد.

جدول (۲). نتیجه پرتفوی مدل شبکه عصبی کانولوشن یادگیری تقویتی عمیق

ردیف	نام شرکت	مدل میانگین واریانس	مدل میانگین-نیم واریانس	مدل میانگین- قدرمطلق انحرافات	مدل ارزش در معرض ریسک مشروط
1	ایران ترانسفو	0.06	0.04	0.01	0.01
2	ایران خودرو	0.03	0.10	0.05	0.11
3	بانک صادرات	0.01	0.01	0.00	0.00
4	بانک ملت	0.00	0.00	0.00	0.00
5	توسعه معادن و فلزات	0.11	0.10	0.13	0.09
6	سایپا	0.00	0.00	0.00	0.00
7	سرمایه گذاری توسعه ملی	0.00	0.00	0.01	0.00
8	معدنی و صنعتی گل گهر	0.00	0.00	0.00	0.00
9	سرمایه گذاری صندوق بازنشستگی	0.00	0.00	0.00	0.00
10	سرمایه گذاری غدیر	0.00	0.00	0.00	0.00
11	سرمایه گذاری صنعت نفت	0.02	0.01	0.00	0.00
12	صنایع آذراب	0.14	0.13	0.15	0.16
13	فولاد خوزستان	0.00	0.00	0.01	0.00
14	فولاد مبارکه اصفهان	0.07	0.07	0.18	0.06
15	گروه بهمن	0.00	0.01	0.00	0.00
16	گروه مپنا	0.00	0.01	0.01	0.00
17	لیزینگ رایان سایپا	0.04	0.05	0.03	0.09
18	مخابرات ایران	0.18	0.21	0.19	0.21
19	معدنی و صنعتی چادرملو	0.15	0.06	0.03	0.09
20	ملی صنایع مس ایران	0.18	0.20	0.19	0.17
	بازدهی مورد انتظار	0.0012	0.0012	0.0012	0.0012
	شاخص ریسک هر مدل	0.01	0.01	0.01	0.03
	شاخص شارپ	0.14	0.16	0.19	0.08

۵- بحث و نتیجه گیری

در این مطالعه به بهینه سازی سبد سهام با استفاده از یادگیری تقویتی عمیق مبتنی بر مدل شبکه عصبی کانولوشن پرداخته شد. با توجه به نتایج ارزیابی بازده پرتفوی، شاخص ریسک و شاخص شارپ و اینکه در همه معیارها نتایج به دست آمده پرتفوی مدل شبکه عصبی کانولوشن یادگیری عمیق در حد مطلوبی بوده است، میتوان گفت مدل شبکه عصبی کانولوشن یادگیری عمیق دارای توانایی مطلوبی در تشکیل پرتفوی بهینه سهام می باشد.

با توجه به پیچیدگی و تنوع عوامل مؤثر بر تعیین پرتفوی شرکتهای، تصمیم گیری در این مورد مسئله بسیار پیچیده ای میباشد و فرآیند بهینه سازی کاملاً بدون ساختار بوده و مجموعه ثابتی از معیارها برای ارزیابی وجود ندارد. اگرچه مدل های کلاسیک تا حدودی توانسته اند پاسخگوی نیازهای اطلاعاتی سرمایه گذاران باشند اما زمانی که عوامل اثرگذار بر بازار سرمایه محدود باشند میتوانند مفید واقع شوند. ولی درواقع عوامل متعددی بر بازار سرمایه تأثیر میگذارند که غیرخطی و پیچیده هستند و با استفاده از معادلات ریاضی و کامپیوترهای پیشرفته قابلحل میباشد. صدها سهم برای انتخاب در بازار سرمایه

منابع

تهرانی، رضا، فلاح تفتی، سیما و آصفی، سپهر. (۱۳۹۷). بهینه سازی سبد سهام به کمک الگوریتم فراابتکاری دسته های میگو با استفاده از معیارهای مختلف از ریسک در بورس اوراق بهادار تهران، فصلنامه تحقیقات مالی، ۲۰(۴): ۸۷-۱۰۱.

حیدری، محمدسعید و ولیدی، جواد و ابراهیمی، سیدبابک. (۱۴۰۰). بهینه سازی سبد سهام مبتنی بر مدل برنامه ریزی امکانی استوار با استفاده از الگوریتم های ژنتیک و جهش قورباغه مخلوط شده، مهندسی مالی و مدیریت اوراق بهادار، ۱۲(۴۷): ۱۱۲-۱۲۷.

فناحی نافچی، حسن، عرب صالحی، مهدی، اسماعیلیان، مجید. (۱۳۹۸). انتخاب سبد بهینه سهام با بکارگیری اطلاعات حسابداری، اطلاعات مبتنی بر ارزش و اطلاعات کارت ارزیابی متوازن -مورد مطالعه: شرکت های پذیرفته شده در بورس اوراق بهادار تهران. پیشرفت های حسابداری، ۱۱(۲): ۲۸۵-۳۲۰.

کرکه آبادی، مهدیه و کلهری، معصومه. (۱۳۹۹). شناسایی و رتبه بندی عوامل موثر بر انتخاب بهینه سبد سهام با استفاده از روش *AHP* فازی (مطالعه موردی: بورس اوراق بهادار)، هشتمین همایش ملی پژوهش های مدیریت و علوم انسانی در ایران، تهران.

نیکو، سیده فرخ، شمس، شهاب الدین، صیقلی، محسن. (۱۳۹۹). مدل سازی انتخاب سبد بهینه سهام بر مبنای ارزیابی ریسک و رویکرد مالی رفتاری (حسابداری ذهنی) در بورس اوراق بهادار تهران. نشریه چشم انداز مدیریت مالی، ۱۰(۳۱): ۷۵-۱۰۱.

یزدانی خداشهری محمدباقر، نسل موسوی سید حسین، حسینی شیروانی میرسعید. (۱۴۰۰). انتخاب بهینه سبد

وجود دارد و تجزیه و تحلیل آنها کار بسیار دشواری است؛ اما با استفاده از مدل یادگیری عمیق میتوان درمدتزمان بسیار کوتاهی دادهها را جمعآوری کرد و صدها سهم را در یک بازه زمانی واقعی رصد کرد و بهترین ایده ها را از این فرآیند دریافت کرد. یادگیری عمیق میتواند میلیونها داده ای را که در زمان واقعی اتفاق میافتد، جمع کند و بسیاری از همبستگی ها که با روش سنتی نمیتوان به دست آورد، با استفاده از آن به راحتی به دست آورد. مزیت مدل یادگیری تقویتی عمیق آن است که می تواند قواعد را بدون آنکه فرمول بندی شده باشد، استخراج نماید. با توجه به اینکه مدل یادگیری عمیق توانایی تشکیل پرتفوی بهینه را دارا میباشد سرمایه گذاران می توانند با استفاده از این الگو به تشکیل پرتفوی و کسب سود در بازار سرمایه اقدام نمایند. نتیجه این پژوهش با نتایج پژوهش های مهلوات و همکاران (۲۰۱۸)، ژو و همکاران (۲۰۱۸)، اسید و همکاران (۲۰۱۸) و اریو و همکاران (۲۰۱۷) همسو می باشد. بر اساس نتایج کسب شده می توان نتیجه گرفت که الگوریتم بهینه سازی مبتنی بر مدل شبکه عصبی کانولوشن یادگیری تقویتی عمیق، دارای کارایی و توانایی مناسب برای حل این مسئله است. به محققان آتی پیشنهاد می شود مدل بهینه سازی یادگیری تقویتی عمیق بر اساس سایر الگوریتم های این روش همچون شبکه های حافظه طولانی کوتاه- مدت^۱، شبکه های عصبی بازگشتی^۲، شبکه های مولد تخاصمی^۳، شبکه های تابع پایه شعاعی^۴، پرسپترون های چندلایه^۵، نگاشت های خودسازمان دهنده^۶، شبکه های باور عمیق^۷، ماشین های بولتزمن محدود شده^۸ و اتوانکدرها یا خودرمزگذار^۹ پیاده سازی و اجرا نمایند.

6 Self-Organizing Maps

7 Deep Belief Networks

8 Restricted Boltzmann Machines

9 AutoEncoders

¹ Long Short Term Memory Networks

² Recurrent Neural Networks

³ Generative Adversarial Networks)

⁴ Radial Basis Function Networks)

⁵ Multilayer Perceptrons

- annealing algorithm. *Journal of Financial Research*, 17(9), 141-158.
- Gulpinar, N., An, L.T.H., Moeini, M. (2010). Robust investment strategies with discrete asset choice constraints using DC programming. *Optimization*, 59(1), 45-62.
- Ma, Yong & Jiang, Hao & Xiao, Weilin. (2021). Tax evasion, audits with memory, and portfolio choice. *International Review of Economics & Finance*. 71. 896-909.
- Mehlawat, M. Kumar, A. Yadav, S. Chen, W. (2018). "Data envelopment analysis based fuzzy multi-objective portfolio selection model involving higher moments". *Information Sciences*, 460-461:128-150.
- Piao, Jinze. (2023). Portfolio Optimization Based on Deep Learning and Factor Constraints. *Advances in Economics, Management and Political Sciences*. 48(5). 264-273.
- Sen, Jaydip & Dutta, Abhishek & Mehtab, Sidra. (2021). Portfolio Optimization Using Deep Learning Models -A Comparative Study of Risk-Based Portfolio Design Approaches. *3rd International Workshop on Machine Learning for Finance*, Ca' Foscari University of Venice, Italy, December 17, 2021At: Venice, ItalyAffiliation: Università Ca' Foscari Venezia 10.
- Sharma, Meeta & Shekhawat, Hardayal. (2022). Portfolio optimization and return prediction by integrating modified deep belief network and recurrent neural network. *Knowledge-Based Systems*. 4(25):128-139.
- Shaw, D.X., Liu, S. & Kopman, L. (2008). Lagrangian relaxation procedure for cardinality-constrained portfolio optimization. *Optimization Methods & Software*, 23(3), 411-420.
- Woodside-Oriakhi, M., Lucas, C., & Beasley, J. E. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213, 538e550.
- Wu, Songwei. (2023). Dynamic Optimization of Investment Portfolio Based on Self-Attention Mechanism - LSTM Model Prediction. *Advances in Economics, Management and Political Sciences*. 61(12). 110-119
- سهام با استفاده از الگوریتم ترکیبی فرا ابتکاری. تحقیق در عملیات در کاربردهای آن. ۱۸ (۱): ۱۰۱-۱۲۴.
- Anagnostopoulos, K.P., Mamanis, G. (2010). A portfolio optimization model with three objectives and discrete variables. *Computers & Operations Research*, 37 (7), 1285-1297.
- Arel, I. Rose, D.C. Karnowski, T.p. (2010). "Deep Machine Learning - A New Frontier in Artificial Intelligence Research [Research Frontier]". *IEEE Computational Intelligence Magazine*, 5: 13-81
- Bertsimas, D., Shioda, R. (2009). Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1), 1-22.
- Carandente, Vincenzo & Sperli, Giancarlo. (2024). Harnessing Cognitively Inspired Predictive Models to Improve Investment Decision-Making. *Cognitive Computation*18(6). 1-16.
- Chang, T. J., Meade, N., Beasley, J. E. & Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimization. *Computers & Operations Research*, 27 (13), 1271-1302.
- Chen CH. Lu CH. Lin CH. (2020). "An intelligence approach for group stock portfolio optimization with a trading mechanism". *Knowledge and Information Systems*, 62: 287-613
- Cui, Tianxiang & Ding, Shusheng & Jin, Huan & Zhang, Yongmin. (2022). Portfolio constructions in cryptocurrency market: A CVaR-based deep reinforcement learning approach. *Economic Modelling*. 11(9). 89-116.
- Deng, D. Yu, D. (2013). "Deep Learning: Methods and Applications Found". *Trends Signal Process*,7(4): 197-387.
- Essid, H. Ganouati, J. Vigeant, S. (2018). "A mean-maverick game crossefficiency approach to portfolio selection: an application to Paris stock exchange". *Expert system with applications*,1-45.
- Fernandez, A. & Gomez, S. (2007). Portfolio Selection Using Neural Networks. *Computer & Operation Research*, 34(4), 1177-1191.
- Ghodusi, S., Tehrani, R. & Bashiri, M. (2015). Portfolio optimization with simulated

Zhou, X. Wang, J. Yang, X. Lev, V. Tu, Y. Wang, SH. (2018). "Portfolio selection under different attitudes in fuzzy environment". *Information Sciences*, :264 278-2.

Zhang, Erchuan. (2023). *Portfolio Optimization Strategy Based on Four Deep Learning Models*. *Advances in Economics, Management and Political Sciences*. 47(3). 295-302.

Stock portfolio optimization using deep reinforcement learning based on convolutional neural network model

[\]Sanaz Delili

[\]Yasser Rezaei Pitenoei*

[\]Mohammad Taghi Kabiri

[\]Mehdi Safari Gerayli

[\]Meysam Arabzadeh

Abstract

Classic portfolio optimization models can be used when capital market constraints and portfolio formation criteria are low. But considering the realities of the capital market and the weakness of classical models in discovering complex and non-linear relationships, it can be understood that the problem is more complex and beyond the classical models. In solving such complex problems, interdisciplinary sciences, computers and models based on machine learning have come to the aid of humans and have had a successful approach in solving these problems. Recently, deep learning techniques, such as recurrent neural network, have been designed for portfolio optimization. In this research, a deep reinforcement learning convolutional neural network model is proposed for stock portfolio optimization based on historical data. In this sense, the daily returns of 20 companies admitted to the Tehran Stock Exchange, which had high liquidity in the period from 1392 to 1401, have been used. The results obtained from this research show that based on the evaluation of the portfolio return, risk index and Sharpe index and that in all criteria the results of the deep learning convolution neural network model portfolio were optimal, it can be said that the neural

[\] PhD student, Department of Financial Engineering, Kashan Branch, Islamic Azad University, Kashan, Iran
Sanazdalili2626@gmail.com.

[\] Associate Professor , Department of Accounting, University of Guilan, Rasht, Iran. (Responsible author).
rezaei.yasser@gmail.com

[\] Assistant Professor, Department of Accounting, Faculty of Administrative and Economic Sciences, Arak University, Arak, Iran. m-kabiri@araku.ac.ir

[\] Associate Professor Department of Accounting, Bandargaz Branch, Islamic Azad University, Bandargaz, Iran.
Mehdi.safari83@yahoo.com

[\] Assistant Professor, Department of Accounting, Kashan Branch, Islamic Azad University, Kashan, Iran.
Arabzadeh62@yahoo.com

Journal of Financial Knowledge of Securities Analysis
Vol. 18 / No. 65, Spring 2025

network model Deep learning convolution has a good ability to form an optimal stock portfolio.

key words: *Optimization, stock portfolio, deep reinforcement learning, convolutional neural network.*