مدلسازی اقتصادی
Economic Modelling
دوره ۱۶. شماره ۵۸. تابستان ۱۴۰۱ صفحه‌های ۵۱ تا ۶۶.

علمی

بهینه‌سازی پرتفوی سهام در بورس اوراق بهادار تهران (کاربرد رهیافت یادگیری تقویتی)[1]

مهدی اسفندیار*، محمدعلی کرامتی**، رضا غلامی جمکرانی+، محمدرضا کاشفی نیشابوری×

DOI: 10.30495/ECO.2022.1965665.2687

تاریخ دریافت:
۱۴۰۱/۰۵/۲۵
تاریخ پذیرش:
۱۴۰۱/۰۷/۲۴

طبقه‌بندی JEL:
G11, G17, P45

واژگان کلیدی:
بهینه‌سازی پرتفوی، معاملات الگوریتمی، یادگیری تقویتی، یادگیری کیو، بورس اوراق بهادار تهران

چکیده

هدف این مقاله کاربرد معاملات الگوریتمی با تمرکز بر رویکرد یادگیری تقویتی برای بهینه‌سازی پرتفوی سهام‌های منتخب است. این پژوهش از حیث هدف، کاربردی و از نظر نوع داده، کمّی و از لحاظ روش، توصیفی – اکتشافی و از منظر طرح تحقیق، پس‌رویدادی است. جامعه آماری این پژوهش، ۶۷۲ شرکت بورسی است که از این تعداد، داده‌های پنج شرکت (نمونه آماری) طی دوره زمانی ۱۳۹۶–۱۴۰۰ بررسی شده است. یافته‌های تحقیق در دوره‌های صعودی و نزولی بازار نشان داد که رویکرد یادگیری تقویتی در بازارهای صعودی و نزولی به صورت معناداری بر رویکرد خرید و نگهداری برتری دارد و عملکرد بهتری ارائه داده است و نتایج با عملکرد الگوریتم‌ها در بازارهای بورس سازگار است. نتایج آشکار کرد که از دیدگاه سودآوری، رویکرد یادگیری تقویتی نسبت به رهیافت خرید و نگه‌داری، عملکرد بهتر و موثرتری داشته است؛ بنابراین، به‌کارگیری روش یادگیری تقویتی پیشنهاد می‌شود.

[1] این مقاله مستخرج از رساله دکتری آقای مهدی اسفندیار به راهنمایی دکتر محمدعلی کرامتی و دکتر رضا غلامی جمکرانی و مشاوره دکتر محمدرضا نیشابوری در دانشگاه آزاد اسلامی واحد قم است.

*دانشجوی دکتری، گروه مدیریت صنعتی، واحد قم، دانشگاه آزاد اسلامی، قم، ایران، پست الکترونیکی: esfandiar.mehdi@yahoo.com

** دانشیار گروه مدیریت صنعتی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول)، پست الکترونیکی: mohammadalikeramati@yahoo.com

+استادیار، گروه حسابداری، واحد قم، دانشگاه آزاد اسلامی، قم، ایران، پست الکترونیکی: gholami@qom-iau.ac.ir

×استادیار گروه مدیریت صنعتی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران، پست الکترونیکی: R.kashery@iauctb.ac.ir

۱. مقدمه

امروزه، مفاهیمی چون «بهینه‌سازی سبدسهام»[1] و «تنوع‌بخشی»[2] به‌مثابه ابزاری برای توسعه و فهم بازارهای مالی و تصمیم‌گیری مالی به‌کارگرفته شده‌اند؛ در این میان، رواج نظریه پرتفوی سهام مارکوییتز[3] (۱۹۵۲) اصلی‌ترین و مهم‌ترین موفقیت تلقی شده است. این مدل توانسته است در شیوه نگرش افراد به سرمایه‌گذاری و سبد سهام تغییرات بسیاری ایجاد کند و جایگاه برجسته‌ای به‌عنوان ابزاری کارا برای بهینه‌سازی سبدسهام کسب نماید (فو، چانگ و چانگ،[4] ۲۰۱۳).

بهینه‌سازی پورتفوی، مولفه اصلی نظام معاملاتی است. هدف بهینه‌سازی، انتخاب بهترین توزیع دارایی در یک پورتفوی برای بیشینه‌کردن بازده‌ها در سطح ریسکی معین است. مارکوئیتز نخستین اقتصاددانی بود که این نظریه را مطرح کرد و این نظریه با عنوان «نظریه پورتفولوی مدرن» (MPT)[5] قبول عام یافت. مزیت اصلی چنین پورتفولیویی مبتنی‌بر ترویج تنوع‌بخشی است که «منحنی سرمایه صاحبان سهام»[6] را هموار می‌سازد؛ یعنی، نسبت به معاملات دارایی انفرادی به بازده بیشتری منجر می‌شود. این امر بدین‌معناست که ریسک (نوسان) پورتفولیوی طولانی‌مدت همواره کمتر از ریسک دارایی انفرادی است (ژانگ، زهرن و رابرتس،[7] ۲۰۲۰).

به‌باور مارکوییتز، سرمایه‌گذاران، ریسک و بازده را باهم درنظرمی‌گیرند و میزان تخصیص سرمایه به فرصت‌های سرمایه‌گذاری گوناگون را براساس تعامل ریسک و بازده انتخاب می‌کنند (ریوز، موز و وِنما[8]، ۲۰۱۴).

امروزه، کامپیوترها، معاملات را مدیریت می‌کنند؛ در صنعت مالی، پیوسته تلاش می‌شود فناوری‌های جدید به‌کار گرفته شوند. در این‌راستا، نظام‌ها و روش‌های مختلفی در کارکردهای بازارهای مالی استفاده می‌شود؛ برای مثال، می‌توان به الگوریتم‌های کنترل، پردازش سیگنال و روش‌های ریاضیاتی – آماری اشاره کرد.

پس از پیشرفت‌های برجسته‌ای که در فناوری رخ داد، فناوری در تمام صنایع و عرصه‌های اقتصادی دیده شد و در این‌راستا، بازارهای مالی نیز بهره‌مند شد و پس از برنامه‌نویسی برنامه‌های تحلیلی و سامانه‌های معاملاتی آنلاین، ورود فناوری به بازارهای مالی بیش از پیش مورد استقبال سرمایه‌گذاران و کارگزاران قرار گرفت. استفاده از ابزارهای گوناگون برای کسب سود از بازارهای مالی سبب شد برنامه‌نویسان به طراحی سیستم‌هایی اقدام کنند که به‌صورت خودکار به اجرای معاملات می‌پردازد. این سیستم‌های خودکار «معاملات الگوریتمی»[9] نام گرفت.

با توجه به اینکه از سال ۱۳۹۴ امکان به‌کارگیری معاملات الگوریتمی در بورس اوراق بهادار تهران فراهم شده است، هدف اصلی این مقاله بررسی امکان بهینه‌سازی پرتفوی سهام با بهره‌گیری از معاملات الگوریتمی با رویکرد

---

[1] Stock Portfolio Optimization
[2] Diversification
[3] Markowitz
[4] Fu, Chung & Chung
[5] وی به بسط نظریه پورتفولیوی مدرن (Modern Portfolio Theory (MPT مبادرت ورزید. نظریه وی با عنوان نظریه انتخاب پورتفولیو (Theory of Portfolio Choice) مشهور شده است.
[6] The Equity Curve
[7] Zhang, Zohren & Roberts
[8] Reeves, Moose & Venema
[9] Algorithmic Trading

یادگیری تقویتی است؛ بدین‌منظور، دو رهیافت «یادگیری تقویتی»[1] و «خرید و نگهداری»[2] با یکدیگر مقایسه می‌شود. می‌شود.

یادگیری تقویتی، یکی از روش‌های اجرای معاملات الگوریتمی است. در سال‌های اخیر، یادگیری تقویتی به ابزاری مهم در توسعه نظریه بهینه‌سازی پورتفولیوی متوالی و پویا تبدیل شده است. بهینه‌سازی پرتفوی از طریق یادگیری تقویتی تلاش می‌کند با انجام معاملات بیشتر در پاسخ به پیش‌بینی‌های نوسانات بازار، سودآوری پرتفوی را افزایش دهد؛ با تحقق این هدف، می‌توان گفت سبد انتخابی بهینه شده است؛ زیرا با اجرای معاملات بیشتر و تبدیل سود حاصل از یک سهم مشخص به پول نقد و خرید مجدد سهام در فرصت مناسب، در عمل سرمایه‌گذار به بیشینه کردن سود حاصل از پرتفوی خود می‌پردازد.

در این مقاله پورتفولیو مستقیما با استفاده از مدل‌های یادگیری عمیق بررسی می‌شود. برخلاف روش‌های سنتی که در آن، در وهله نخست، معمولا با بهره‌گیری از مدل‌های اقتصادسنجی، بازده مورد انتظار پیش‌بینی می‌شود، این گام پیش‌بینی‌کننده کنار گذاشته می‌شود تا مستقیما تخصیص‌های دارایی به‌دست آید. بنابراین، با بهینه‌سازی مستقیم «نسبت شارپ[3]» بازده در واحد ریسک به حداکثر می‌رسد. از این‌رو، هدف این مقاله بررسی امکان و توان کاربرد یادگیری تقویتی برای بهینه‌سازی پورتفوی با توجه به محدودیت‌های تعیین‌شده توسط بازار سهام، مانند نقدینگی و هزینه‌های مبادله است. عوامل یادگیری تقویتی عمیق (DRL) در دو محیط مختلف آموزش می‌بینند تا توانایی یادگیری بهترین راهبردهای معاملاتی برای تخصیص دارایی‌ها آزمون شود (یعنی، ایجاد بازده‌های تجمعی موردانتظار بیشتر). همه عوامل «بدون مدل»[4] هستند و قبلا برای معضلات مالی بهینه شده‌اند؛ بنابراین، فضا – حالت همان‌طور که در محیط‌های بازارهای مالی یافت می‌شود از بعد زیادی برخوردار است.

برای این‌منظور، مقاله در پنج بخش سازمان‌دهی می‌شود: بعد از مقدمه، در بخش دوم، ادبیات پژوهش بررسی می‌شود؛ در بخش سوم، روش پژوهش بیان می‌شود و بخش‌های چهارم و پنجم نیز به یافته‌ها و نتیجه‌گیری و پیشنهادها اختصاص می‌یابد.

۲. مروری بر ادبیات پژوهش

در این بخش، برای برقراری ارتباط بهتر مفاهیم و نظریات و درنهایت، دست‌یابی به بینشی عمیق نسبت به مسائل اقتصادی، اصطلاحات و مفاهیم اساسی تعریف و تحدید می‌شود.

از منظر اصطلاح‌شناسی[5]، پورتفولیو (پرتفوی)[6] شامل مجموعه‌ای از دارایی‌های واقعی و مالی سرمایه‌گذاری‌شده سرمایه‌گذار است (پی جونز، ۱۳۹۱). دارایی‌ها[7]، اقلامی است که از ارزش اقتصادی برخوردار است. سهام‌ها[8]، وجوه

[1] Reinforcement Learning
[2] Buy-and-Hold Strategy
[3] Sharpe Ratio
[4] Model-Free
[5] Terminology
[6] Portfolio
[7] Assets
[8] Stocks

نقد[1] (صندوق و بانک)، وام‌ها و مساعده‌ها[2] و درآمدهای ثبت‌نشده[3] جزو این اقلام به شمار می‌آیند (جانستون و ژوریک[4]، ۲۰۱۱). الگوریتم، فهرستی از دستورالعمل‌های بخوبی تعریف‌شده است که برای اجرای یک دستور (وظیفه)[5] (وظیفه)[5] لازم است. در الگوریتم، ترتیب و توالی انجام عملیات اهمیت دارد. همچنین، هرالگوریتم، نقاط آغاز و پایان معینی دارد. در حوزه معاملات الگوریتمی، انجام معامله به کامپیوتر و سامانه کامپیوتری واگذار می‌شود (ترِلیون، گالاس و لالچند[6]، ۲۰۱۳).

اگرچه، کاربرد معاملات الگوریتمی به اوایل دهه ۱۹۷۰ میلادی برمی‌گردد؛ با آغاز هزاره سوم، به‌کارگیری آن، رونق یافته است. پیشرفت فناوری در بازارهای مالی و شیوه معاملات ابزارهای مالی، دو تغییر مهم و مرتبط ایجاد کرده است که یکی خودکارسازی فرایندهای معاملاتی و دیگری، تجهیز بازارهای مالی با سامانه ثبت الکترونیکی محدوده سفارش‌هاست. سرعت و کیفیت دسترسی به بازارهای مالی موجب برانگیختن سرمایه‌گذاران به بهره‌گرفتن از معاملات الگوریتمی شده است (کریلنکو و لو[7]، ۲۰۱۳). امروزه، یکی از مباحث مهم بازارهای مالی، به‌کارگیری معاملات الگوریتمی در مدیریت پورتفولیو یا سبد سرمایه‌گذاری است.

بهینه‌سازی پورتفولیو، عبارت است از انتخاب بهترین ترکیب از دارایی‌های مالی به‌نحوی که باعث شود، تاحد ممکن بازده پورتفولیو، بیشینه و ریسک آن، کمینه شود. ایده اساسی نظریه مدرن پورتفولیو این است که اگر در دارایی‌هایی که به‌طورکامل هم‌بستگی ندارند، سرمایه‌گذاری شود؛ ریسک آن دارایی‌ها یکدیگر را خنثی کرده و می‌توان یک بازده ثابت با ریسک کمتر به دست آورد (مارکوئیتز، ۱۹۵۲). در بهینه‌سازی پورتفولیو، مساله اصلی، انتخاب بهینه دارایی‌ها و اوراق بهاداری است که با مقدار مشخصی سرمایه می‌توان به‌دست آورد؛ اگرچه، کمینه‌کردن ریسک و بیشینه‌کردن بازده سرمایه‌گذاری ساده به‌نظر می‌رسد؛ روش‌های متعددی برای تشکیل پورتفولیوی بهینه به‌کار رفته است. الگوریتم‌های معاملاتی، مدل‌های ریاضی است که معاملات بازار بورس را تحلیل کرده و براساس آن، فرصت‌های نقدینگی را شناسایی کرده و اطلاعات را به تصمیم‌های معاملاتی هوشمند تبدیل می‌کند. این الگوریتم‌ها در تمامی مراحل مختلف چرخه معاملات (قبل، حین و بعد) به‌کار گرفته می‌شود (ترلیون و همکاران، ۲۰۱۳).

- یادگیری تقویتی

یادگیری تقویتی، نوعی تکنیک یادگیری ماشینی است که در هنگام عرضه نشدن پارامترهای شبکه عصبی از طریق تعامل متقابل با محیط مشخص می‌سازد. یادگیری تقویتی با این مسأله مرتبط است که شبکه عصبی چگونه باید در مقابل محیط واکنش نشان دهد تا پاداش بلندمدت را افزایش دهد. این یادگیری در کاربرد خود در مسائل گوناگون از قبیل کنترل ربات، ارتباط از راه دور و بازی‌هایی نظیر شطرنج و نیز برای تصمیم‌گیری‌های متوالی موفق عمل کرده است (میهاچ و نانیه[8]، ۲۰۰۲)

[1] Cash
[2] Loans & Advances
[3] Accrued Incomes
[4] Johnston & Djuric´
[5] Order/Task
[6] Treleaven, Galas & Lalchand
[7] Kirilenko & Lo
[8] Mihatsch & Neuneier

یادگیری تقویتی بر رفتارهایی تمرکز دارد که عامل باید برای بیشینه کردن پاداش انجام دهد. طبق تعریف، رفتار عبارت از نگاشتی است که از مجموعه اداراکات عامل به تصمیم‌گیری برای اجرای مناسب منجر می‌شود و رفتار عقلانی، رفتاری است که با اطلاعات دردسترس به‌دنبال بیشتر کردن کارآمدی است. یادگیری تقویتی روش یادگیری برای انتخاب کردن رفتار مناسب براساس پاداش و تنبیه است، بدون اینکه لازم باشد نحوه اجرای عمل برای عامل مشخص شود (حاج‌رسولی‌ها، ۱۳۹۳).

در یادگیری تقویتی، نوع اقدام عامل از قبل مشخص نمی‌شود؛ بلکه عامل با جستجوی مبتنی‌بر سعی و خطا رفتاری را یاد می‌گیرد که بیشترین پاداش را به‌دست آورده و سود کوتاه‌مدت فدای سود بلندمدت می‌شود. در راهبرد جستجو برای رسیدن به پاداش بیشتر دو رویکرد اصلی وجود دارد: ۱) رویکرد بهره‌مندانه (حریصانه)[1] و ۲) رویکرد اکتشافی (تصادفی)[2]. چالش اصلی، ایجاد تعادل با ترکیبِ دو رویکرد یادشده است؛ ازاین‌رو، لازم است بین جستجوی موارد جدید و استفاده از دانش پیشین تناسب ایجاد شود؛ بر همین اساس، در فضای جستجو با توزیع تصادفی یک‌نواخت لازم است چندبار تکرار شود تا امکان حداکثر نمودن پاداش بلندمدت به‌دست آید.

یادگیری تقویتی در زمینه‌هایی چون اقتصاد و نظریه بازی‌ها بیشتر به بررسی تعادل‌های ایجادشده با عقلانیت محدود عامل‌ها می‌پردازد. در یادگیری تقویتی، داده‌ها براساس تابع هزینه (پاداش یا تنبیه) تحلیل می‌شود. در یادگیری تقویتی، زوج مشترک ورودی و خروجی درکار نیست و رفتارهای ناکارآمد نیز از بیرون اصلاح نمی‌شود و همچنین، تمرکز زیادی بر کارآمدی تعاملی وجود دارد که نیازمند دست‌یابی به تعادل بین اکتشاف‌های جدید و بهره‌برداری از دانش اندوخته‌شده است (اصغری اسکویی، فلاحی، دوستی‌زاده و مشیری، ۱۳۹۷: ۱۴).

یادگیری تقویتی از دومؤلفه اصلی «عامل»[3] و «محیط»[4] تشکیل شده است. عامل در محیط زندگی کرده و با توجه به بازخوردی که از محیط می‌گیرد، تجربه‌های خود را به‌روزرسانی می‌کند. براساس تحلیل صوری[5] (نمودار ۱)، در یادگیری تقویتی، عامل، در هرلحظه، در حالت (S_t) قرار دارد که با انتخاب عملی (A_t) از فضای اعمال خود و اجرای آن به حالت بعدی (S_{t+1}) انتقال می‌یابد و پاداشی (R_t) از محیط دریافت می‌کند. سپس، براساس پاداش، تجربه خود از حضور در حالت (S_{t+1}) و اجرای عمل (A_{t+1}) را به‌روزرسانی می‌کند. به‌روزرسانی تجربه عامل، برآورد ارزشی از حالت موردنظر و عمل اجراشده در آن حالت (برآورد «ارزش حالت ـ عمل»[6]) است. در فرایند به‌روزرسانی برآورد ارزش حالت ـ عمل رابطه (۱) برقرار است.

تخمین جدید ← تخمین قبلی + نرخ یادگیری × پاداش به‌دست‌آمده (۱)

میزان یادگیری (α) عددی بین صفر و یک بوده و با افزایش زمان کاهش می‌یابد. در ابتدای زندگی عامل، مقدار α نزدیک به یک است؛ زیرا عامل در ابتدا تجربه‌ای ندارد و باید میزان اهمیت عامل در برآورد ارزش حالت ـ عمل خود به بازخوردی که از محیط دریافت می‌کند، زیاد باشد. با افزایش زمان، از آنجا که عامل در هرتغییر حالت،

---

[1] Exploiting (Greedy)
[2] Exploring (Random)
[3] Agent
[4] Environment
[5] Formal
[6] A State-Action Value Function

تابع ارزش حالت– عمل تابع Q نیز نامیده می‌شود و مشخص می‌کند که عامل، عملی خاص را در حالتی با یک خط مشی π انجام دهد، چقدر مطلوبیت دارد. تابع Q با Q(s) نشان داده می‌شود. این امر نشان‌دهنده ارزش اجرای یک اقدام در حالتی است که از یک خط مشی π پیروی می‌کند.

بـازخوردی از محیط می‌گیرد، بهتر است به برآورد ارزش حالت ـ عمـل خـود اهمیـت بیشـتری داده و تـأثیر بـازخورد از محـیط را در به‌روزرسانی برآوردهای جدید کاهش دهد. بنابراین، عامل می‌تواند در طول زندگی خود در محیط، رفتار محیط را بخوبی شناسایی کرده و در حالت‌های مختلف، بهترین تصمیم را برای بیشینه کردن پاداش خود اتخاذ کند (فلاح‌پور و حکیمیان، ۱۳۹۸: ۲۶).

- اجزای یادگیری تقویتی

۱) **عامل:** الگوریتم یا مدلی است که اقدامات را انجام می‌دهد و با گذشت زمان فرایند یادگیری را طی می‌کند؛

۲) **محیط:** محیطی است که عامل با آن تعامل متقابل برقرار می‌کند و اعمال خود را در آن فضا به‌انجام می‌رساند؛

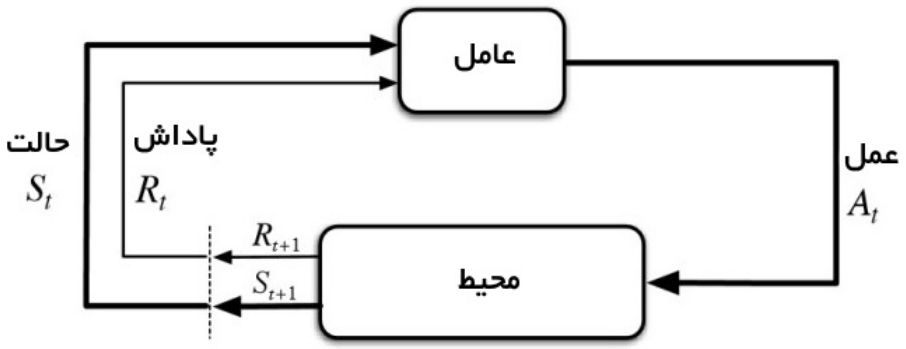
۳) اقدام/عمل[1]: همان عملی است که عامل انجام می‌دهد. اقدامات اساساً فعل‌وانفعالات عامل در یک محیط هستند.

٤) **پاداش**[2]: نتیجه یک عمل است. هرعملی، پاداشی به‌دنبال دارد. پاداش می‌تواند مثبت (تشویق) یا منفی (جریمه) باشد.

۵) **وضعیت/حالت**[3]: وضعیت فعلی عامل در محیط است. اعمالی که عامل اجرا می‌کند می‌تواند وضعیت آن را تغییر دهد.

٦) **سیاست**[4]: راهبرد یا رفتاری است که اقداماتی را دربرمی‌گیرد که عامل برای دست‌یابی به نتیجه موردنیاز باید انجام دهد.

۷) **تابع ارزش**[5]: این تابع، اساساً عامل را از حداکثر پاداشی که برای هریک از وضعیت‌ها در آینده دریافت می‌کند، مطلع می‌کند. تابع ارزش می‌تواند عامل را در انتخاب عملکرد بهینه راهنمایی کند (ساتون و بارتو[6]، ۲۰۱۸).



نمودار ۱. الگوریتم یادگیری تقویتی

منبع: شکری، ۱٤۰۰

[1] Action
[2] Reward
[3] State
[4] Policy
[5] Value Function
[6] Sutton & Barto

- یادگیری کیو[1]

یادگیری کیو، رویکردی در یادگیری تقویتی است که با کاربرد مفهوم تابع ارزش کیو[2] به یادگیری تابع سیاست بهینه کمک می‌کند. یکی از تکنیک‌های یادگیری تقویتی است که با یادگیری یک تابع اقدام/مقدار، سیاست مشخصی را برای اجرای اقدامات مختلف در وضعیت‌های متفاوت دنبال می‌کند. یکی از نقاط قوت این روش، توانایی یادگیری تابع بدون داشتن مدل معینی از محیط است. یادگیری کیو، تلاش می‌کند با توجه به شرایط فعلی، بهترین اقدامات را اجرا کند؛ به‌طورکلی، یادگیری کیو به‌دنبال یادگیری سیاست است تا مجموع پاداش را بیشینه کند (وزان، ۱۳۹۹).

- پیشینه تحقیق

پژوهش‌های زیادی با موضوع بهینه‌سازی پرتفوی انجام شده است که در آنها از روش‌های مختلفی استفاده شده است. از نظر تاریخی، پژوهش اسکابار و کلوته[3] (۲۰۰۲) نخستین اثری است که سعی کرده با ترکیب تحلیل تکنیکال و یادگیری تقویتی بهینه سیستم معاملاتی خودکاری طراحی کرده و گسترش دهد. پارک، سیم و چویی[4] (۲۰۲۰) بیان کرده‌اند که عامل هوشمند برای شناسایی اقدام معاملاتی بهینه برپایه یادگیری کیوی عمیق آموزش می‌بیند و نتایج بهتری از راهبردهای استاندارد کسب می‌شود. ژانگ و همکاران (۲۰۲۰) در تحقیقی بیان کرده‌اند که روش آنها از مدل‌های پایه بهتر عمل کرده و با وجود هزینه‌های سنگین تراکنش، سود مثبتی به‌دست آمده است.

رستگار و دستپاک (۱۳۹۷) در پژوهش خود رویکرد بهره‌گیری از خودمعامله‌گرها برای پیش‌بینی روند آتی سهم و روش یادگیری تقویتی برای مدیریت پویای سبد سهام را مطرح کرده‌اند. نتایج نشان داد که یادگیری تقویتی عملکرد مدل را بهبود بخشیده است.

فلاح‌پور و حکیمیان (۱۳۹۸) در تحقیقی با به‌کارگیری داده‌های معاملات روزانه (درون‌روزی)[5] در بورس اوراق بهادار تهران بیان کرده‌اند که معاملات زوجی، معروف‌ترین و قدیمی‌ترین نوع سیستم‌های معاملات الگوریتمی است که کارآیی و سودآوری آن در بسیاری از پژوهش‌های بازارهای مالی آشکار شده است. در سال‌های اخیر، تحقیقات مهمی درباره معاملات الگوریتمی براساس یادگیری ماشینی صورت گرفته است. نتایج آزمایش روی داده‌های معاملات روزانه زوجی سهام‌های منتخب نشان داد که کاربرد روش یادگیری تقویتی در طراحی سیستم معاملات در معامله‌های زوجی نسبت‌به سایر روش‌ها برتری دارد.

گل‌ارضی و انصاری (۱۴۰۱) در تحقیقی به مقایسه عملکرد الگوریتم ژنتیک مرتب‌سازی نامغلوب با الگوریتم تکاملی قدرت پارتو بهبودیافته در انتخاب پورتفولیوی بهینه در بورس اوراق بهادار تهران پرداختند. نتایج نشان داد که در مدل، میانگین انحراف معیار عملکرد الگوریتم ژنتیک مرتب‌سازی نامغلوب نسبت‌به عملکرد الگوریتم تکاملی

[1] Q-Learning
[2] Q-Value
[3] Skabar & Cloete
[4] Park, Sim & Choi
[5] معاملات درون‌روزی (Intraday Trading) که به‌عنوان معاملات روزانه نیز شناخته می‌شود به خرید و فروش سهام در همان روز اشاره دارد. بازار سهام در معرض نوساناتی است که به تغییر قیمت سهام در طول روز منجر می‌شود.

قدرت پارتو مطلوب‌تر و در مدل میانگین واریانس و میانگین نیم‌واریانس عملکرد الگوریتم تکاملی قدرت پارتو نسبت به عملکرد الگوریتم ژنتیک مرتب‌سازی نامغلوب مطلوب‌تر است.

۳. روش پژوهش

هدف این مقاله کاربرد معاملات الگوریتمی با تمرکز بر روش یادگیری تقویتی[1] به‌منظور بهینه‌سازی پرتفوی سهام‌های منتخب است. این پژوهش از حیث هدف، کاربردی و از نظر نوع داده، کمّی و از لحاظ روش، توصیفی – اکتشافی و از منظر طرح تحقیق، پس‌رویدادی است. جامعه آماری این پژوهش، ۶۷۲ شرکت بورسی در اسفندماه سال ۱۴۰۰ بوده که از میان آنها، پنج‌شرکت (نمونه آماری) انتخاب شده‌اند. روش انتخاب نمونه به‌صورت خوشه‌ای یک‌مرحله‌ای و سپس، انتخاب هدفمند یک سهم از داخل هرخوشه است؛ به این‌صورت‌که انتخاب خوشه‌ها و نیز انتخاب یک سهم از میان سهم‌های (شرکت‌های) هرخوشه به صورت هدفمند انجام شده است. این سهام‌ها (شرکت و نماد اختصاصی) در جدول (۱) عرضه شده است.

محدودیت انتخاب سهام مورد بررسی، براساس موجود بودن سوابق معاملاتی سهام انتخابی در طول دوره مورد بررسی می‌باشد و در هرکدام از پنج گروه صنعتی موردنظر در تحقیق، یک سهم انتخاب شده است. این داده‌ها به صورت داده‌های روزانه است که از سایت رسمی بورس اوراق بهادار تهران (آرشیو داده‌های بورسی)[2] استخراج شده است.

جدول ۱. سهام‌های منتخب

ردیف	نام شرکت	نماد	صنعت
۱	نفت سپاهان	شسپا	نفت و پتروشیمی
۲	همکاران سیستم	سیستم	خدمات
۳	مخابرات ایران	اخابر	مخابرات
۴	کیمیدارو	دکیمی	دارویی
۵	سیمان شرق	سشرق	سیمان

منبع: تارنمای بورس اوراق بهادار تهران (اسفندماه ۱۴۰۰)

داده‌های این پژوهش، ترکیبی از داده‌های اسنادی (کتابخانه‌ای= کیفی) و مشاهده‌ای (میدانی= کمّی) است. داده‌های کیفی، ادبیات نظری و چارچوب نظری را پوشش می‌دهد. نتایج بررسی‌ها، با توجه به گروه‌های صنایع انتخابی، درنهایت به انتخاب دوره زمانی پنج‌ساله از ابتدای سال ۱۳۹۶ – ۱۴۰۰/۱۲/۰۲ منجر شده است. برای آزمون الگوریتم طراحی‌شده به روش یادگیری تقویتی با الگوی یادگیری کیو، براساس نمودار روند شاخص کل بازار بورس، و

[1] تاکنون روش‌های مختلفی برای پیش‌بینی قیمت سهام طراحی شده است که عبارتند از: ۱) تجزیه و تحلیل فنی؛ ۲) تجزیه و تحلیل بنیادی؛ ۳) سری‌های زمانی سنتی و ۴) روش‌های یادگیری ماشین (Kalyvas, 2001). در این مقاله از روش یادگیری ماشینی، از نوع تقویتی با الگوی یادگیری کیو (Q) بهره گرفته شده است.

[2] http://irbours.com

شرکت‌های مورد بررسی، دو بازه زمانی انتخاب شد که شامل ۴۰ دوره معاملاتی روزانه به صورت صعودی و نزولی است.

دوره زمانی مورد بررسی سال‌های ۱۳۹۶-۱۴۰۰ است و در داخل این دوره زمانی، دو دوره ۴۰ روزه معاملاتی به‌عنوان بازارهای صعودی و نزولی انتخاب شده‌اند که روند صعودی از تاریخ ۱۳۹۹/۰۱/۰۶ - ۱۳۹۹/۰۳/۱۰ و روند نزولی از تاریخ ۱۴۰۰/۱۰/۰۵ - ۱۴۰۰/۱۲/۰۲ است. دوره‌های صعودی و نزولی برای بررسی سیستم خودمعامله‌گر طراحی‌شده در وضعیت‌های بازار صعودی و نزولی انتخاب شده است. سیستم خودمعامله‌گر در نرم‌افزار متلب کدنویسی شده است و براساس داده‌های بورسی پردازش‌شده در نرم‌افزار اکسل، محیط معاملاتی را شبیه‌سازی و درخصوص سیگنال‌دهی برای انجام معاملات اقدام می‌کند. روش اجرایی خودمعامله‌گر بدین شکل است که فایل اکسل مربوط به دوره معاملاتی را فرامی‌خواند و داده‌های قیمت سهام را وارد حافظه خودمعامله‌گر می‌کند که یک ماتریس به‌صورت ۱ × ۴۵ (در نظر داشتن پنج دوره مازاد بر دوره موردنظر) است. ماتریس قیمت‌های سهام، در نقش داده‌های قیمت سهام در بازار واقعی عمل می‌کند و خودمعامله‌گر، قادر است با کمک تابع مطلوبتی که برای خودمعامله‌گر طراحی شده است، قیمت‌های روز جاری و قیمت‌های آتی را که توسط الگوریتم مورد بررسی قرار می‌گیرد، درک نموده به تصمیم‌گیری برای سیگنال‌دهی بپردازد. درنهایت، داده‌های معاملاتی خودمعامله‌گر در قالب فایل اکسل ثبت می‌شود تا بتوان با تحلیل خروجی آن، تفسیر بهتری از روند معاملات و تصمیم‌گیری الگوریتم به‌دست آورد. محاسبات تغییرات رشد یا بازده سهام، از طریق فرمول زیر به‌دست می‌آیند.

$$R_n = (P_{t+n} - P_t)/ P_t$$ رابطه (۱)

- معرفی الگوریتم سیگنال‌دهی

الگوریتم سیگنال معاملاتی، براساس مقایسه نرخ بازده آتی محاسبه و ایجاد می‌شود. در این پژوهش، نرخ بازده آتی، حاصل بررسی و مقایسه یک دوره آتی و پنج دوره آتی است. پس از عرضه سیگنال اولیه، الگوریتم یادگیری تقویتی، برمبنای معادله بلمن، سودآوری تصمیم اخذشده را با سایر گزینه‌های محتمل قابلِ انتخاب، سنجیده و مقدار محاسبه‌شده جدید را در ماتریس کیو، اصلاح می‌نماید. در مدل طراحی‌شده، روند تغییرات رشد روزانه محاسبه‌شده است که به محاسبه میانگین تغییرات رشد منجرمی‌شود که همان بازده سهام طی دوره مورد بررسی است.

الگوریتم سیگنال‌دهی، پنج سیگنال خرید قوی، خرید ضعیف، فروش قوی، فروش ضعیف و عدم‌معامله را ایجاد می‌کند که در حالت قوی، ۱۰۰ درصد و در حالت ضعیف، ۵۰ درصد سرمایه نقدی و یا سهام موجود در سبد برای معاملات خرید و یا فروش استفاده می‌کنند.

الگوریتم محاسبه تصمیم‌گیری و سیگنال‌دهی در مورد یک سهم، به صورت زیر است.

```
If(R1 >= 0.015) % Checking The Trade Profit Margin with the Trading Expenses
Return +1 % Strong Buying Signal
Elseif(R5>=0.015)
Return +0.5 % Weak Buying Signal
Elseif(R1 <= -0.015)
Return -1.0 % Strong Selling Signal
Elseif (R5 <= -0.015)
Return -0.5 % Weak Selling Signal
Else
```

Return 0.0 % Holding Signal

- الگوریتم یادگیری کیو و معادله بلمن

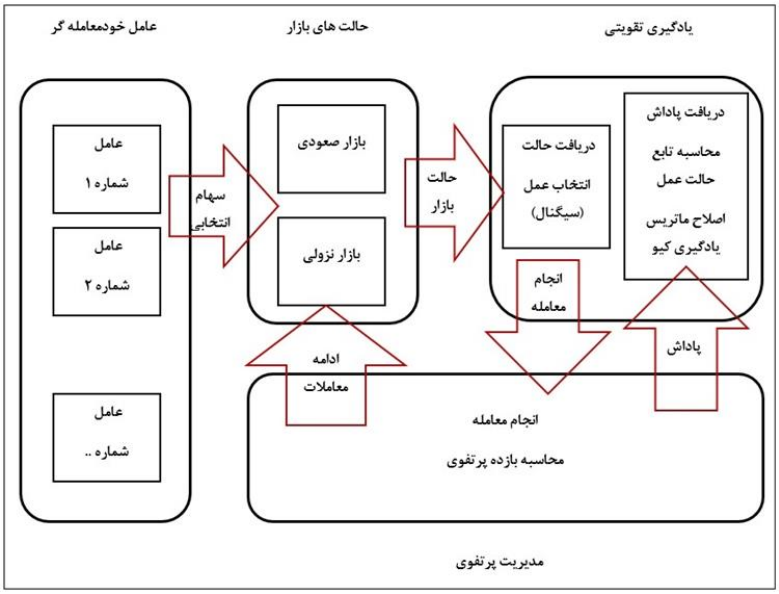
یادگیری تقویتی براساس معادله بلمن[1] بنا شده است،

$$NewQ(s,a) = Q(s,a) + \alpha\,[r(s,a) + \gamma \max Q'(s',a') - Q(s,a)] \quad \text{رابطه (۲)}$$

در رابطه (۲) متغیر نرخ یادگیری α، تعیین می‌کند که تا چه میزان اطلاعات جدید بر اطلاعات قدیمی ترجیح داده شود. مقدار صفر باعث می‌شود که عامل، چیزی یاد نگیرد و مقدار یک باعث می‌شود که عامل فقط اطلاعات جدید را ملاک قرار دهد. همچنین، متغیر نرخ تنزیل (گاما) γ اهمیت پاداش‌های آینده را تعیین می‌کند. مقدار صفر باعث می‌شود که عامل، ماهیت فرصت‌طلبانه گرفته و فقط پاداش‌های فعلی را مدنظر قرار دهد؛ درحالی‌که مقدار یک عامل را ترغیب می‌کند، برای یک دوره زمانی طولانی برای پاداش تقلا کند.

- طراحی مدل

در این مقاله از یک مدل خودمعاله‌گر استفاده می‌شود که برمبنای الگوریتم منطقی مورد استفاده، در چند گام، نسبت‌به اجرای معاملات به‌صورت خودکار اقدام می‌کند و با دریافت داده‌های موردنیاز نسبت‌به تولید سیگنال معاملاتی متناسب با وضعیت بازار اقدام می‌کند و با استفاده از معادله بلمن و تابع کیو، ارقام ماتریس کیو را بررسی کرده و در مورد انتخاب بهترین گزینه تصمیم‌گیری می‌کند. مدل نظری پژوهش در شکل (۱) ترسیم شده است.



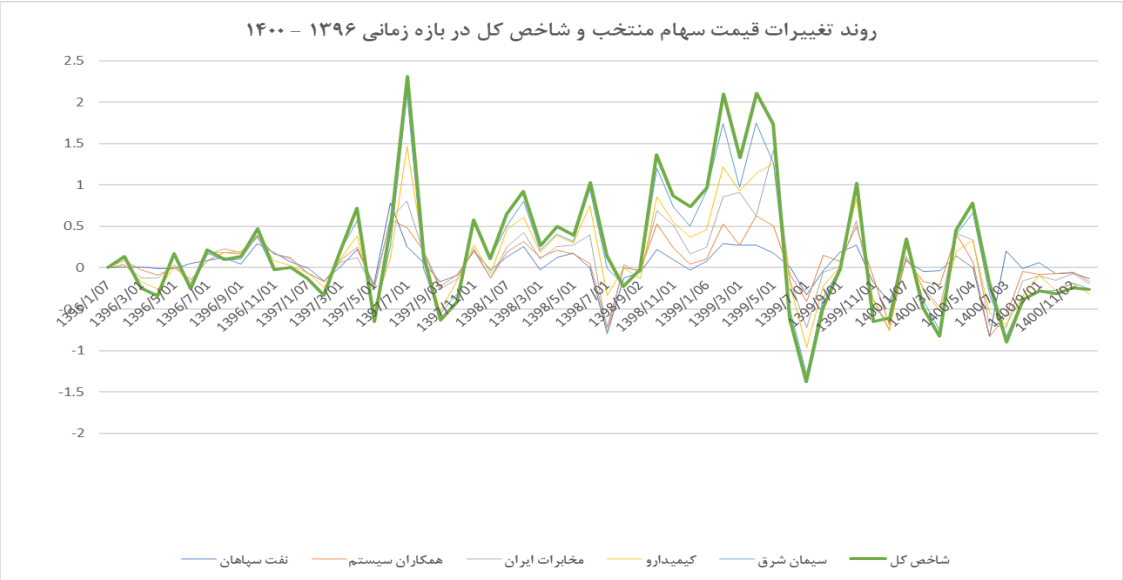
نمودار ۲. مدل نظری تحقیق برای انجام معاملات الگوریتمی به روش یادگیری تقویتی (یادگیری کیو)

منبع: یافته‌های پژوهش

[1] Bellman Equation

با توجه به نمودار (۲) عامل خودمعامله‌گر با محیط (حالت‌های صعودی و نزولی) بازار در تعامل است که براساس سهام انتخابی، هرسهم، سیستمِ خودمعامله‌گرِ اختصاصی (الگوریتم‌های رایانه‌ای) دارد. براساس این مدل، عامل خودمعامله‌گر با دریافت حالت بازار، عمل مناسب را از بین پنج سیگنال محتمل برای اجرا در بازار، براساس قیمت سهام در یک و پنج دوره آتی انتخاب می‌کند. سپس، این عامل، معامله را انجام داده و پاداش را محاسبه می‌کند. عامل خودمعامله‌گر با استفاده از معادله بلمن، تابع حالت- عمل را محاسبه کرده و رقم محاسبه‌شده را در ماتریس حالت- عمل (= ماتریس کیو) اصلاح می‌کند. در ادامه روند معاملاتی، سیستم مدیریت پرتفوی، با توجه به حالت بازار، همین بخش از الگوریتم را تکرار می‌کند.

٤. یافته‌های پژوهش



نمودار ۳. نمودار تغییرات قیمت سهام و شاخص کل در بازه زمانی ۱۳۹٦ - ۱٤۰۰

منبع: یافته‌های پژوهش

در نمودار (۳) روند تغییرات قیمت پنج سهم مورد بررسی در بازه پنج‌ساله عرضه شده است. با اضافه نمودن روند تغییرات شاخص کل بورس، با هدف مقایسه روند پنج سهم مورد بررسی و شاخص کل، بخوبی می‌توان شباهت‌های روند تغییرات قیمت سهام را درک نمود و اقدام به بررسی و مقایسه روند تغییرات قیمت کرد. تشابه روند سهام مورد بررسی با روند تغییرات شاخص کل بورس، در نمودار (۳) بخوبی قابل ملاحظه است. بازده سالانه سهام مورد بررسی و شاخص کل که براساس داده‌های ماهیانه محاسبه شده است، در جدول (۲) ارائه شده است (در محاسبه داده‌های ماهیانه نیز از داده‌های روزانه بورسی استفاده شده است).

جدول ۲. بازده سالانه سهام‌های مورد بررسی

سهام / سال	نفت سپاهان	همکاران سیستم	مخابرات ایران	کیمیدارو	سیمان شرق	شاخص کل
۱۳۹۶	۱/۱۹۳	۰/۰۴۵	-۰/۰۶۱	-۰/۳۵۱	-۰/۳۱۳	۰/۲۶۶
۱۳۹۷	۰/۷۶۹	۰/۳۲۷	۰/۰۴۹	۰/۵۷۳	۰/۸۱۴	۰/۶۵۱
۱۳۹۸	۱/۰۲۸	۱/۷۵۱	۱/۵۸۰	۲/۸۱۲	۲/۲۴۹	۱/۷۴۸
۱۳۹۹	۰/۸۵۹	۱/۵۹۹	۰/۴۴۲	۰/۲۷۰	۰/۱۷۴	۱/۳۶۵
۱۴۰۰	۰/۰۳۳	۰/۰۲۰	-۰/۳۶۹	-۰/۵۳۰	-۰/۰۵۱	-۰/۰۲۷
میانگین	۰/۷۷۶	۰/۷۴۸	۰/۳۲۸	۰/۵۵۵	۰/۵۷۵	۰/۸۰۱

منبع: یافته‌های پژوهش

همان‌طور که در جدول (۲) ملاحظه می‌شود، شرکت نفت سپاهان، در سال ۱۳۹۶، بازدهی ۱۱۹/۳ درصدی را کسب کرده است؛ درحالی‌که شرکت سیمان شرق، در سال ۱۳۹۶، ۳۱/۳ درصد افت ارزش را تجربه کرده است و بازده سالانه شاخص کل در همان سال ۲۶/۶ درصد بوده است. داده‌های سالانه، هرچند اهمیت بسیار زیادی در تصمیم‌گیری‌های سرمایه‌گذاران دارند و بخش عمده بررسی‌های بورسی، برمبنای داده‌های سالانه بنا می‌شود؛ بااین‌همه نمی‌تواند روند تغییرات قیمت را بخوبی منعکس نمایند که با نگاهی به داده‌های ماهیانه در نمودار (۳) می‌توان به نوسانات بازده یا تغییرات قیمت درخصوص سهام موردبررسی و شاخص کل پی برد.

- بررسی عملکرد پورتفولیوها

در مقایسه عملکرد دو رویکرد یادگیری تقویتی و خرید- نگهداری، با توجه به این نکته مهم که افزایش تعداد معاملات در رویکرد یادگیری تقویتی میزان ریسک را در مقایسه با سطح ریسک در رویکرد خرید- نگهداری افزایش می‌دهد و علاوه‌بر این، برطبق نظریه مدرن پرتفوی افزایش سودآوری با افزایش سطح ریسک همراه است، برای سرمایه‌گذارانی که تحمل افزایش ریسک را ندارند، نمی‌توان رویکرد یادگیری تقویتی را رویکردی مناسب برای بهینه کردن پرتفوی سهام دانست.

- در بازار صعودی

جدول ۳. داده‌های تجمیعی پنج‌سهم مورد بررسی (در بازار صعودی) و دو راهبرد معاملاتی (خرید و نگهداری/ یادگیری تقویتی)

ردیف	نام شرکت	نماد	بازده خرید- نگهداری (B-H)	بازده معاملات الگوریتمی (یادگیری تقویتی) (R-L)	مقایسه بازده
۱	نفت سپاهان	شسپا	۰/۸۱۱	۱/۲۸۹	R-L > B-H
۲	همکاران سیستم	سیستم	۰/۲۰۶	۰/۶۰۶	R-L > B-H
۳	مخابرات ایران	اخابر	۰/۹۹۸	۱/۲۱۱	R-L > B-H
۴	کیمیدارو	دکیمی	۰/۳۱۲	۰/۶۰۴	R-L > B-H
۵	سیمان شرق	سشرق	۰/۶۴۰	۱/۳۳۳	R-L > B-H

منبع: یافته‌های پژوهش

در جدول (۳) نتایج هردو راهبرد معاملاتی خرید و نگهداری و معاملات به روش یادگیری تقویتی به‌صورت تجمیعی عرضه شده است. روش یادگیری تقویتی در بازار صعودی، برتری قابل ملاحظه‌ای را نشان می‌دهد و در همه سهم‌ها، این برتری را با نرخ‌های سودآوری مثبت متفاوتی از ۶۰/۴ درصد تا ۱۳۳ درصد حفظ نموده است.

- در بازار نزولی

جدول ٤. داده‌های تجمیعی پنج سهم مورد بررسی (در بازار نزولی) و دو راهبرد معاملاتی (خرید و نگهداری / یادگیری تقویتی)

ردیف	نام شرکت	نماد	بازده خرید-نگهداری (B-H)	بازده معاملات الگوریتمی (یادگیری تقویتی) (R-L)	مقایسه بازده
۱	نفت سپاهان	شسپا	-۰/۲۰۶	۰/۰۵۰	R-L > B-H
۲	همکاران سیستم	سیستم	-۰/۰۱۶	۰/۱۰۱	R-L > B-H
۳	مخابرات ایران	اخابر	-۰/۰۹۸	۰/۱۰۹	R-L > B-H
٤	کیمیدارو	دکیمی	-۰/۲۰۹	۰/۰۹۰	R-L > B-H
۵	سیمان شرق	سشرق	-۰/۰۰۵	۰/۱٦٤	R-L > B-H

منبع: یافته‌های پژوهش

در جدول (٤) هردو راهبرد معاملاتی خرید و نگهداری و معاملات به روش یادگیری تقویتی تا انتهای دوره در جدول (٤) به صورت تجمیعی ارائه شده است. براساس جدول (٤) روش یادگیری تقویتی در بازار نزولی برتری قابل ملاحظه‌ای را نشان می‌دهد و در همه سهم‌ها، این برتری را با نرخ‌های سودآوری مثبت متفاوتی از ۵ درصد تا ۱٦/٤ درصد حفظ نموده است.

۵. نتیجه‌گیری و پیشنهادها

در این پژوهش، برای رسیدن به بینشی عمیق‌تر از بازار بورس و عملکرد آن، داده‌های روزانه پنج شرکت بورسی (مناسب دانسته شد) با استفاده از روش یادگیری تقویتی عمیق طی دوره زمانی ۱۳۹٦ - ۱٤۰۰ بررسی شد. بررسی پنج سهم (سهم‌های انتخابی) نفت سپاهان، همکاران سیستم، مخابرات ایران، کیمیدارو، سیمان شرق نشان داد که روند تغییرات قیمتی این سهام‌ها، همگونی زیادی با هم دارند؛ صرفنظر از تفاوت‌های مختصر در روند تغییرات؛ در برخی از زمان‌ها، این همگونی در نمودارهای سهام‌های مورد بررسی بیشتر است؛ اما، اساسا، می‌توان روند تغییرات را همگون دانست. همچنین، یافته‌ها نشان داد که سطح تغییرات برای هرسهم متفاوت بوده یا تغییرات با اندکی تأخیر نسبت به دیگر سهم‌ها روی داده است.

نتایج آشکار کرد که از دیدگاه سودآوری، رویکرد یادگیری تقویتی نسبت به رهیافت خرید و نگه‌داری، عملکرد بهتر و موثرتری داشته است. مقایسه نتایج این دو رویکرد روشن می‌کند که کاربرد یادگیری تقویتی برای سرمایه‌گذارهایی که توان ریسک‌پذیری بالای رهیافت خرید و نگه‌داری را ندارند، مناسب‌تر است.

برطبق نظر مارکوییتز افزایش سودآوری، ازدیاد میزان ریسک را درپی دارد. در این راستا، سرمایه‌گذاری که در انجام معاملات خود، پس از انتخاب سهم، توان پذیرش ریسک بالاتر را داشته باشد، با کاربرد یادگیری تقویتی در بازار صعودی، منفعت بیشتری تحصیل می‌کند و پیامد آن، کاهش زیان در بازار نزولی است؛ بنابراین، به‌کارگیری این رویکرد توصیه می‌شود؛ برعکس، سرمایه‌گذاری که توان ریسک‌پذیری انجام معاملات بیشتر یک سهم مشخص را ندارد، کاربرد رویکرد یادگیری تقویتی توصیه نمی‌شود.

نوسان‌گیری در بورس، نوعی راهبرد معاملاتی است که پیامد استفاده از تغییرات سریع قیمت‌هاست. کسب سود از این نوع معاملات، روش دشواری است؛ ولی به‌دلیل جذابیت آن برای کسب سریع سود، معامله‌گران زیادی در بازار سرمایه این روش را برای کسب بازده می‌پسندند. نکات مهم برای موفقیت در این معاملات عبارتند از شناسایی سهام مناسب برای نوسان‌گیری و تعیین نقطه مناسب برای خرید و فروش. در این راهبرد، معامله‌گر باید آمادگی نوسان‌گیری از اصلاح و بازگشت روند (براساس تحلیل روان‌شناختی رفتار معامله‌گران)، نمودار سهم، حجم معاملات و سفارش‌ها و در نهایت، روند چرخش نقدینگی را داشته باشد.

سیاست راهبردی در معاملات الگوریتمی بهره‌برداری از نوسانات قیمت سهام است و در رویکرد یادگیری تقویتی بر بهره‌برداری از نوسانات قیمتی یک سهم مشخص تاکید می‌شود. از این‌رو، می‌توان ماهیت بهره‌برداری از نوسانات در رویکرد یادگیری تقویتی را متفاوت از نوسان‌گیری در بازار سهام دانست. افزون‌براین، رویکرد یادگیری تقویتی به‌دنبال استفاده از فرصت‌های معاملاتی برای سهم معین است که پیامد آن، افزایش سودآوری پرتفوی و رونق بیشتر بازار سهام خواهد بود.

براساس یافته‌ها (در جداول ۳ و ٤)، در بازار صعودی، رهیافت خرید و نگهداری، به سودآوری منجر شده است. اما، مبتنی‌بر یافته‌ها، معاملات الگوریتمی سودآوری بیشتری را نسبت به روش خرید و نگهداری تجربه کرده‌اند. اولین دلیل عملکرد متفاوت دو راهبرد مورد بررسی، وجود نوسان در حالت‌های صعودی و نزولی بازار است و الگوریتم‌ها، توان بهره‌برداری از نوسان‌ها را دارند. در بازار صعودی نیز نوسان‌هایی وجود دارد که موجب کاهش قیمت‌ها می‌شود و در بازار صعودی، موفقیت معاملات الگوریتمی، در مقایسه با راهبرد خرید و نگهداری، تشخیص نوسان‌های بازار و فروش سهام و خرید مجدد قیمتی است که می‌تواند سودآوری حداقلی را در کوتاه‌مدت (در این مقاله طی یک و یا پنج دوره آتی) میسر نماید.

در معاملات به روش یادگیری تقویتی، تشخیص درست حالت بازار، اهمیت زیادی دارد و مهم‌ترین گام، پیش‌بینی حالت بازار است و در مرحله بعدی، انتخاب اقدام صحیح قرار دارد که توسط سیستم سیگنال‌دهی انجام می‌شود. در بازار نزولی، در رویکرد معاملاتی خرید و نگهداری زیان رخ می‌دهد؛ زیرا در انتهای دوره و در طول دوره‌ای که سهم موردنظر خریداری و نگهداری می‌شود، روند کلی قیمت به صورت نزولی است و این روند نزولی، موجب زیان است.

دلیل سودآوری معاملات الگوریتمی در بازارهای نزولی بهره‌برداری از نوسان‌های بازار است که به محض فراهم‌شدن وضعیت فروش، به فروش و تحصیل سود مبادرت می‌ورزند و در ادامه با قیمت مناسب‌تری، خرید سهم را تکرار می‌کنند و با اجتناب از زیان و از دست رفتن ارزش سبد، آن را در وضعیت بهتری قرار می‌دهند که منجر به بهینه‌سازی سبد و افزایش سودآوری واقعی و نسبی سبد می‌شود.

در برخی موارد، معاملات الگوریتمی، ضررهایی نیز به صورت محدود ایجاد می‌کند و با توجه به محاسباتی که انجام می‌دهد، برای کاهش زیان و افزایش سود حرکت می‌کند و نوسان‌های بازار سهام به‌گونه‌ای است که نمی‌توان ضررهای معاملاتی را به صورت کامل حذف کرد و یا در مواردی که کاهش قیمت سهام به میزانی باشد که سیگنال ایجادشده، نگهداری باشد، ناچار باید شاهد ضررهای ناشی از کاهش قیمت سهام بود.

الگوریتم‌ها در بازارهای نزولی نیز سودآوری داشته‌اند و این موجب محبوبیت معاملات الگوریتمی شده است. البته، در بازارهای نزولی، کاربرد راهبرد خرید و نگهداری، پذیرفته نشده است و انجام معاملات بیشتر موجب کاهش ضرر خواهد بود؛ حتی اگر سودآوری مثبت ایجاد نکند، در کاهش ضررها مؤثر خواهد بود. درنتیجه، در بازار نزولی، به صورت منطقی افزایش تعداد معاملات، در بدترین وضعیت، ضررها را می‌کاهد و بهره‌برداری از نوسان‌های بازار، مهم‌ترین نقش را در سودآوری بیشتر دارد و برای دست‌یابی به چنین امکانی، پیش‌بینی قیمت‌های آتی با استفاده از ابزارهای تحلیل تکنیکال، ضروری است.

مهم‌ترین نکته در الگوریتم‌ها توجه به هزینه‌های معاملاتی یا کارمزدهای معاملاتی است. به همین دلیل، در بورس اوراق بهادار تهران، در بازه ۱/۵ درصدی (بازه‌ای است که سهام خریدنی و فروختنی محسوب می‌شود) سهام موجود در سبد، قابلیت معامله ندارد و سرمایه‌گذار، ناگزیر به پذیرش زیان‌هایی است که از سوی کاهش قیمت در این محدوده خاص ایجاد می‌شود و در صورتی که افزایش قیمت نیز در همین محدوده باشد؛ درواقع، به‌دلیل وجود هزینه‌های کارمزدی، چنین سودهایی، زمینه‌ای برای انجام معاملات و کسب سود فراهم نمی‌کند. به همین دلیل، در معاملات الگوریتمی نیز از ناحیه سودهای اندک، منفعتی کسب نمی‌شود و این به دلیل نرخ‌های کارمزد است و به عبارت دیگر، با توجه به مقررات بورس اوراق بهادار ایران، امکان بهره بردن از نوسان‌های کم‌تر از ۱/۵ درصدی مهیا نیست و معاملاتی که اجراشدنی است، به نوسان‌هایی بیشتر از این بازه نیاز دارد.

منابع

- اصغری اسکوئی، محمد رضا، فلاحی، فرهاد، دوستی‌زاده، میثم، مشیری، سعید (۱۳۹۷). کاربرد یادگیری تقویتی در یک مدل‌سازی عامل‌محور برای بازار عمده‌فروشی برق ایران. *پژوهشنامه اقتصاد انرژی ایران*، ۷(۲۵): ۱-۴۰.
- جونز، چارلز پی. (۱۳۹۱). مدیریت سرمایه‌گذاری. ترجمه: تهرانی، رضا، نوربخش، عسگر، تهران: نشر نگاه دانش.
- رستگار، محمدعلی، دستپاک، محسن (۱۳۹۷). ارائه مدل معاملاتی با فراوانی زیاد همراه با مدیریت پویای سبد سهام به روش یادگیری تقویتی در بورس اوراق بهادار تهران. *فصلنامه تحقیقات مالی*، ۲۰(۱): ۱۶ -۱.
- فلاح‌پور، سعید، حکیمیان، حسن (۱۳۹۸). بهینه‌سازی استراتژی معاملات زوجی با استفاده از روش یادگیری تقویتی، با به‌کارگیری دیتاهای درون‌روزی در بورس اوراق بهادار تهران، *فصلنامه تحقیقات مالی*، ۲۱ (۱): ۳۴-۱۹.
- گل‌ارضی، غلامحسین، انصاری، حمیدرضا (۱۴۰۱). مقایسه عملکرد الگوریتم‌های تکاملی NSGAII و SPEA2 در انتخاب پرتفولیوی بهینه در بورس اوراق بهادار تهران. *فصلنامه تحقیقات مالی*، ۲۴ (۳): ۴۱۰-۴۳۰.
- نورویگ، پیتر و راسل، استوارت ج. (۱۳۹۳) هوش مصنوعی پیشرفته. مترجم: حسین حاج‌رسولی‌ها، تهران: نشر نیازدانش.
- Fu, TC., Chung, CP. & Chung, FL. (2013). Adopting genetic algorithms for technical analysis and portfolio management. *Computers and Mathematics with Applications*, 66 (10), 1743–1757. https://doi.org/ 10.1016/j.camwa. 2013.08.012.
- Johnston, D. E. & Djurić, P. M. (2011). The science behind risk management. IEEE Signal Processing Magazine, 28(5), 26-36.
- Kirilenko, A. A., & Lo, A. W. (2013). Moore's law versus Murphy's Law: Algorithmic trading and its discontents. *Journal of Economic Perspectives*, 27(2): 51-72.
- Markowitz H.M. (1952) . Portfolio Selection. *Journal of Finance*, 7 (1): 77-91.
- Mihatsch, O. & Neuneier, R. (2002). Risk-sensitive reinforcement learning. Machine learning, 49(2): 267-290.
- Park, H., Sim, M. K., & Choi, D. G. (2020). An intelligent financial portfolio trading strategy using deep Q-learning. *Expert Systems with Applications*, 158.
- Reeves, M., Moose, S., & Venema, T. (2014). The growth share matrix. BCG–The Boston Consulting Group.
- Skabar, A., & Cloete, I. (2002). Neural networks, financial trading and the efficient markets hypothesis. In ACSC: 241-249.
- Sutton, R. S. & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT Press.
- Treleaven, P., Galas, M. & Lalchand, V. (2013). Algorithmic trading review. Communications of the ACM, 56(11): 76-85.
- Vogiatzis, G. & Garcia, N. (2019). Learning non-metric visual similarity for image retrieval. *Image and Vision Computing*, 82: 18-25.
- Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep reinforcement learning for trading. *The Journal of Financial Data Science*, 2(2): 25-40.