Journal of
Intelligent Financial
Management Perspectives

ISSN: 3116-0964

Modeling Retail Traders' Behavior in Capital Markets Using Reinforcement Learning and High-Frequency Order Book Data

Niloufar Eslami [1]*

[1] *Master's Degree in Financial Management, Islamic Azad University, Qom Branch, Qom, Iran (Coresponding autour), Email: ni.eslami@qom-iaun.ac.ir*

ARTICLE INFO

Article history:
Received:10/06/2025
Received in revised form:20/06/2025
Accepted:11/07/2025
Available online:13/08/2025

Keywords:

Reinforcement Learning
Retail Traders
Capital Market
Limit Order Book
High-Frequency

ABSTRACT

This study aims to analyze the behavior of retail traders in capital markets and propose an intelligent framework based on reinforcement learning and high-frequency limit order book data. In recent years, the increasing complexity of financial markets and the growing availability of high-frequency data have highlighted the need for advanced computational models to understand and predict market behavior. In this context, the capital market is modeled as a complex adaptive system, where retail traders' behavior is inferred from limit order book data using reinforcement learning algorithms.

Three main algorithms, including Q-learning and Deep Q-Network (DQN), Actor-Critic, and Soft Actor-Critic (SAC), are employed to model and predict market behavior. The dataset consists of features extracted from the high-frequency limit order book, including order imbalance, market depth, order flow, and short-term price dynamics. The results indicate that reinforcement learning models, particularly the Soft Actor-Critic approach, outperform other methods in terms of prediction accuracy, cumulative returns, and risk management. Furthermore, order imbalance is identified as the most influential feature in short-term market direction prediction.

The findings suggest that integrating reinforcement learning with micro-level market data enables the extraction of hidden behavioral patterns of retail traders and significantly improves decision-making in dynamic financial environments. This framework can contribute to the development of intelligent trading systems and enhance the understanding of market behavior at the microstructure level.

Article Type: Research Paper

Journal of Intelligent Financial Management, 2025, Vol. 1, No.2, pp. 71- 86

Publish by:
Tolou-e Binish-e Ayandeh Scientific Institute

https://doi.org/ 10.25843/JIFM.2025.8563.23450

©Authors

Cite: Eslami,N . (2025). Modeling Retail Traders' Behavior in Capital Markets Using Reinforcement Learning and High-Frequency Order Book Data . *Journal of Intelligent Financial Management*, *1*(2), 71-86.

نگرش های
مدیریت مالی هوشمند
Journal of Intelligent Financial Management Perspectives
ISSN: 3116-0964

تحلیل رفتار معامله گران خرد در بازار سرمایه با استفاده از مدل های مبتنی بر یادگیری تقویتی و داده های سفارشات لحظه ای

نیلوفر اسلامی [۱ و*]

۱ و *- کارشناسی ارشد مدیریت مالی، دانشگاه آزاد اسلامی واحد قم، قم، ایران (نویسنده مسئول)، ایمیل نویسنده مسئول: *ni.eslami@qom-iaun.ac.ir*

اطلاعات مقاله

تاریخچه مقاله:

تاریخ دریافت:۱۴۰۴/۰۳/۲۰

تاریخ بازنگری:۱۴۰۴/۰۳/۳۰

تاریخ پذیرش:۱۴۰۴/۰۴/۲۰

تاریخ انتشار:۱۴۰۴/۰۵/۲۲

کلیدواژه‌ها:

یادگیری تقویتی

معامله‌گران خرد

بازار سرمایه

دفتر سفارشات محدود

داده‌های با فرکانس بالا

چکیده

این پژوهش با هدف تحلیل رفتار معامله‌گران خرد در بازار سرمایه و ارائه یک چارچوب هوشمند مبتنی بر یادگیری تقویتی و داده‌های سفارشات لحظه‌ای انجام شده است. در سال‌های اخیر، پیچیدگی روزافزون بازارهای مالی و نقش پررنگ داده‌های با فرکانس بالا، ضرورت استفاده از مدل‌های پیشرفته محاسباتی را در تحلیل رفتار بازار افزایش داده است. در این راستا، بازار سرمایه به‌عنوان یک سیستم پیچیده تطبیقی در نظر گرفته شده و رفتار معامله‌گران خرد از طریق داده‌های دفتر سفارشات محدود و با استفاده از الگوریتم‌های یادگیری تقویتی مدل‌سازی شده است.

در این پژوهش، سه الگوریتم اصلی شامل یادگیری کیو و شبکه کیو عمیق، روش بازیگر منتقد و روش بازیگر منتقد نرم برای تحلیل و پیش‌بینی رفتار بازار مورد استفاده قرار گرفتند. داده‌های مورد استفاده شامل ویژگی‌های استخراج‌شده از دفتر سفارشات لحظه‌ای از جمله عدم‌تعادل سفارشات، عمق بازار، جریان سفارشات و تغییرات لحظه‌ای قیمت بوده است. نتایج نشان داد که مدل‌های یادگیری تقویتی، به‌ویژه روش بازیگر–منتقد نرم، عملکرد بهتری نسبت به سایر مدل‌ها در شاخص‌هایی مانند دقت پیش‌بینی، بازدهی تجمعی و کنترل ریسک دارند. همچنین مشخص شد که عدم‌تعادل سفارشات مهم‌ترین متغیر تأثیرگذار در پیش‌بینی جهت بازار است. یافته‌ها بیانگر آن است که ترکیب یادگیری تقویتی با داده‌های سطح خرد بازار می‌تواند الگوهای رفتاری پنهان معامله‌گران خرد را استخراج کرده و دقت تصمیم‌گیری در محیط‌های مالی پویا را به‌طور معناداری افزایش دهد. این چارچوب می‌تواند به توسعه سیستم‌های معاملاتی هوشمند و بهبود درک رفتار بازار در سطح خرد کمک کند.

نشریه مدیریت مالی هوشمند، ۱۴۰۴، دوره ۱، شماره ۲، صفحه ۷۱-۸۶.

ناشر: موسسه علمی طلوع بینش آینده

https://doi.org/ 10.25843/JIFM.2025.8563. 23450

نوع مقاله: پژوهشی

© نویسندگان

استناد: اسلامی,نیلوفر . (۱۴۰۴). تحلیل رفتار معامله گران خرد در بازار سرمایه با استفاده از مدل های مبتنی بر یادگیری تقویتی و داده های سفارشات لحظه ای. *مدیریت مالی هوشمند*، ۱(۲)، ۷۱-۸۶.

۱-مقدمه

در دهه‌های اخیر، بازارهای مالی در سراسر جهان تحت تأثیر تحولات گسترده فناوری‌های دیجیتال، افزایش دسترسی به داده‌های با فرکانس بالا و پیشرفت‌های چشمگیر در حوزه هوش مصنوعی قرار گرفته‌اند. این تحولات موجب شده است که ساختار تصمیم‌گیری در بازارهای سرمایه از الگوهای سنتی مبتنی بر تحلیل‌های ایستا و داده‌های تجمیعی، به سمت مدل‌های پویا، داده‌محور و الگوریتمی حرکت کند. در این میان، معامله‌گران خرد به‌عنوان یکی از مهم‌ترین گروه‌های فعال در بازار، نقش قابل توجهی در شکل‌دهی نوسانات کوتاه‌مدت، نقدشوندگی و رفتار قیمتی دارایی‌ها ایفا می‌کنند. برخلاف معامله‌گران نهادی که معمولاً از منابع اطلاعاتی گسترده، زیرساخت‌های محاسباتی پیشرفته و الگوریتم‌های پیچیده بهره می‌برند، معامله‌گران خرد اغلب بر اساس اطلاعات محدود، سیگنال‌های کوتاه‌مدت و سوگیری‌های رفتاری تصمیم‌گیری می‌کنند. این تفاوت ساختاری باعث شده است که تحلیل رفتار معامله‌گران خرد به یکی از موضوعات کلیدی در ادبیات مالی رفتاری و اقتصاد محاسباتی تبدیل شود.

در سال‌های اخیر، ظهور داده‌های سطح سفارش امکان بی‌سابقه‌ای برای مطالعه رفتار لحظه‌ای بازار فراهم کرده است. این داده‌ها شامل اطلاعاتی در سطح بسیار خرد مانند سفارش‌های خرید و فروش، حجم سفارش‌ها، قیمت‌های پیشنهادی در سطوح مختلف عمق بازار، زمان‌بندی ثبت سفارش‌ها و تغییرات سریع در ساختار عرضه و تقاضا هستند. برخلاف داده‌های سنتی مانند قیمت پایانی یا حجم روزانه، داده‌های دفتر سفارشات تصویری بسیار دقیق و پیوسته از وضعیت واقعی بازار ارائه می‌دهند و امکان تحلیل رفتار لحظه‌ای بازیگران بازار را فراهم می‌کنند. مطالعات نشان داده‌اند که این داده‌ها می‌توانند اطلاعات ارزشمندی درباره فشار خرید و فروش، نقدشوندگی و احتمال تغییرات کوتاه‌مدت قیمت ارائه دهند(Cont et al., 2014).

با وجود غنای اطلاعاتی داده‌های سفارشات لحظه‌ای، تحلیل این داده‌ها به دلیل ماهیت پیچیده، حجیم و غیرخطی آن‌ها چالش‌برانگیز است. روش‌های اقتصادسنجی سنتی که عمدتاً مبتنی بر فرض‌های خطی، ایستا و توزیع‌های ساده هستند، توانایی کافی برای مدل‌سازی رفتارهای پیچیده و پویا در چنین داده‌هایی را ندارند. به همین دلیل، در سال‌های اخیر استفاده از روش‌های یادگیری ماشین و به‌ویژه یادگیری تقویتی به‌عنوان یک رویکرد قدرتمند برای تحلیل سیستم‌های پویا و تصمیم‌گیری در محیط‌های غیرقطعی مورد توجه قرار گرفته است.

یادگیری تقویتی شاخه‌ای از هوش مصنوعی است که در آن یک عامل از طریق تعامل مستمر با محیط و دریافت بازخورد در قالب پاداش یا جریمه، سیاست بهینه‌ای برای تصمیم‌گیری یاد می‌گیرد. در چارچوب بازارهای مالی، محیط شامل وضعیت بازار (مانند داده‌های دفتر سفارشات)، عامل همان معامله‌گر یا الگوریتم معاملاتی و پاداش نیز سود یا زیان حاصل از تصمیمات معاملاتی است. این چارچوب به‌طور طبیعی با ماهیت پویا، تصادفی و غیرایستا بازارهای مالی سازگار است(Sutton & Barto, 2018) .

مطالعات اخیر نشان داده‌اند که یادگیری تقویتی در حوزه مالی کاربردهای گسترده‌ای یافته است، از جمله در مدیریت پرتفوی، اجرای سفارشات، بازارسازی و طراحی استراتژی‌های معاملاتی هوشمند. به‌طور خاص، الگوریتم‌های یادگیری تقویتی عمیق مانند(DQN) Deep Q-Network ، Deep Deterministic Policy Gradient (DDPG)و Soft Actor-Critic (SAC) توانسته‌اند عملکرد قابل توجهی در محیط‌های پیچیده مالی از خود نشان دهند. این الگوریتم‌ها قادرند روابط غیرخطی پیچیده میان متغیرهای بازار را یاد بگیرند و استراتژی‌هایی تطبیقی در شرایط متغیر بازار ایجاد کنند(Nagy et al., 2023) .

یکی از حوزه‌های مهم کاربرد یادگیری تقویتی در بازار سرمایه، تحلیل رفتار معامله‌گران خرد است. معامله‌گران خرد معمولاً رفتارهایی ناهمگن، غیرمنطقی و وابسته به شرایط کوتاه‌مدت بازار از خود نشان می‌دهند. این رفتارها شامل واکنش بیش از حد به اخبار، رفتار گله‌ای، پیروی از روندهای کوتاه‌مدت و تغییرات ناگهانی در حجم معاملات است. چنین رفتارهایی می‌تواند منجر به ایجاد الگوهای قابل شناسایی در داده‌های سطح سفارش شود که در صورت تحلیل صحیح، اطلاعات ارزشمندی درباره پویایی بازار ارائه می‌دهد.

داده‌های سفارشات لحظه‌ای در این میان نقش کلیدی ایفا می‌کنند، زیرا امکان مشاهده رفتار واقعی معامله‌گران در سطح خرد و در زمان واقعی را فراهم می‌سازند. برای مثال، تغییرات در عدم‌تعادل سفارش‌ها، تغییرات در عمق بازار و سرعت ورود و خروج سفارش‌ها می‌تواند نشان‌دهنده فشار خرید یا فروش در بازار باشد. مطالعات تجربی نشان داده‌اند که این متغیرها دارای قدرت پیش‌بینی قابل توجهی برای بازده‌های کوتاه‌مدت هستند (Bouchaud et al., 2018).

از منظر نظری، بازار سرمایه را می‌توان به‌عنوان یک سیستم پیچیده تطبیقی در نظر گرفت که در آن تعداد زیادی عامل با اهداف، اطلاعات و استراتژی‌های متفاوت با یکدیگر تعامل دارند. در چنین سیستمی، رفتار کلی بازار از تعاملات محلی میان معامله‌گران ناشی می‌شود و نه از یک قانون مرکزی ثابت. این ویژگی باعث می‌شود که مدل‌سازی بازار نیازمند رویکردهایی باشد که بتوانند پویایی، تعامل و یادگیری را در سطح عامل‌ها شبیه‌سازی کنند.

در این راستا، یادگیری تقویتی چندعاملی نیز به‌عنوان یک توسعه مهم در این حوزه مطرح شده است. در این چارچوب، چندین عامل به‌طور همزمان در یک محیط مشترک فعالیت می‌کنند و هر عامل نه‌تنها از تعامل خود با محیط، بلکه از تعامل با سایر عامل‌ها نیز یاد می‌گیرد. این ساختار به‌طور طبیعی با ساختار بازارهای مالی که شامل تعداد زیادی معامله‌گر با اهداف متضاد هستند، همخوانی دارد(He et al., 2019) .

یکی دیگر از جنبه‌های مهم در تحلیل رفتار معامله‌گران خرد، نقش سوگیری‌های شناختی در تصمیم‌گیری است. مطالعات اقتصاد رفتاری نشان داده‌اند که معامله‌گران خرد اغلب تحت تأثیر سوگیری‌هایی مانند بیش‌اعتمادی، رفتار توده‌ای، زیان‌گریزی و واکنش بیش از حد قرار دارند. این سوگیری‌ها باعث می‌شوند که رفتار آن‌ها از عقلانیت کامل فاصله بگیرد و در نتیجه الگوهای غیرتصادفی در داده‌های بازار ایجاد شود. این الگوها می‌توانند از طریق مدل‌های یادگیری تقویتی که قادر به یادگیری از داده‌های پیچیده و غیرخطی هستند، شناسایی شوند.

در سال‌های اخیر، ترکیب یادگیری تقویتی با داده‌های با فرکانس بالا به یکی از حوزه‌های پیشرفته تحقیقاتی تبدیل شده است. پژوهش‌ها نشان داده‌اند که عامل‌های RL می‌توانند با استفاده از اطلاعات لحظه‌ای دفتر سفارشات، استراتژی‌های معاملاتی بهینه‌ای یاد بگیرند که در برخی موارد عملکردی بهتر از استراتژی‌های کلاسیک مانند VWAP و TWAP دارند (Cartea et al., 2015) این نتایج نشان می‌دهد که استفاده از روش‌های یادگیری تقویتی می‌تواند به بهبود کارایی بازار و طراحی سیستم‌های معاملاتی هوشمند کمک کند.

با وجود این پیشرفت‌ها، چالش‌های متعددی در این حوزه وجود دارد. یکی از مهم‌ترین چالش‌ها، پایداری مدل‌های یادگیری تقویتی در شرایط غیرایستا بازار است. بازارهای مالی به‌طور مداوم در حال تغییر هستند و الگوهای رفتاری گذشته ممکن است در آینده تکرار نشوند. این مسئله باعث می‌شود که مدل‌های RL نیازمند سازوکارهایی برای تطبیق مداوم با شرایط جدید باشند. چالش دیگر، هزینه محاسباتی بالای آموزش این مدل‌ها در محیط‌های واقعی و با داده‌های حجیم است.

همچنین مسئله تفسیرپذیری نیز در این حوزه اهمیت ویژه‌ای دارد. اگرچه مدل‌های یادگیری تقویتی می‌توانند عملکرد بالایی داشته باشند، اما در بسیاری از موارد تصمیمات آن‌ها برای انسان قابل توضیح نیست. این موضوع در کاربردهای مالی که نیازمند شفافیت و قابلیت نظارت هستند، یک محدودیت جدی محسوب می‌شود.

در مجموع، مرور ادبیات نشان می‌دهد که تحلیل رفتار معامله‌گران خرد در بازار سرمایه نیازمند ترکیب سه عنصر کلیدی است: نخست، استفاده از داده‌های سفارشات لحظه‌ای به‌عنوان منبع اطلاعاتی دقیق از رفتار بازار؛ دوم، استفاده از مدل‌های یادگیری تقویتی برای مدل‌سازی تصمیم‌گیری پویا و ترتیبی؛ و سوم، در نظر گرفتن تعاملات چندعاملی و سوگیری‌های رفتاری در تحلیل بازار. ترکیب این سه رویکرد می‌تواند چارچوبی جامع برای تحلیل عمیق‌تر رفتار معامله‌گران خرد فراهم کند.

بر این اساس، پژوهش حاضر با هدف تحلیل رفتار معامله‌گران خرد در بازار سرمایه با استفاده از مدل‌های مبتنی بر یادگیری تقویتی و داده‌های سفارشات لحظه‌ای طراحی شده است. این مطالعه تلاش می‌کند تا با بهره‌گیری از داده‌های سطح خرد بازار و الگوریتم‌های پیشرفته یادگیری ماشین، الگوهای پنهان رفتاری معامله‌گران را شناسایی کرده و نقش آن‌ها در شکل‌گیری نوسانات کوتاه‌مدت بازار را تبیین نماید. انتظار می‌رود نتایج این پژوهش بتواند علاوه بر توسعه ادبیات نظری در حوزه مالی محاسباتی، کاربردهای عملی مهمی در طراحی سیستم‌های معاملاتی هوشمند، مدیریت ریسک و بهبود کارایی بازار ارائه دهد.

۲-مبانی نظری

مبانی نظری این پژوهش بر پایه ترکیب سه حوزه کلیدی شامل میکروساختار بازارهای مالی، اقتصاد رفتاری و یادگیری تقویتی در سیستم‌های پیچیده بنا شده است. درک رفتار معامله‌گران خرد در بازار سرمایه نیازمند چارچوبی است که بتواند هم‌زمان ویژگی‌های رفتاری انسان، ساختار نهادی بازار و پویایی‌های مبتنی بر داده‌های با فرکانس بالا را در نظر بگیرد. در ادبیات مالی معاصر، بازار سرمایه دیگر صرفاً به‌عنوان یک سیستم قیمت‌گذاری ساده در نظر گرفته نمی‌شود، بلکه به‌عنوان یک سیستم پیچیده تطبیقی متشکل از تعداد زیادی عامل ناهمگن تعریف می‌شود که به‌طور مداوم با یکدیگر تعامل دارند و این تعاملات در نهایت منجر به شکل‌گیری قیمت و نوسانات بازار می‌شود(Bouchaud et al., 2018) .

در سطح میکروساختار بازار، تمرکز اصلی بر نحوه شکل‌گیری قیمت‌ها از طریق جریان سفارشات خرید و فروش است. برخلاف مدل‌های کلاسیک که قیمت را نتیجه تعادل عرضه و تقاضای کلان می‌دانند، نظریه میکروساختار نشان می‌دهد که قیمت نتیجه مستقیم تعامل لحظه‌ای سفارش‌ها در دفتر سفارشات محدود است. دفتر سفارشات محدود شامل مجموعه‌ای از سفارش‌های فعال خرید و فروش در سطوح مختلف قیمتی است که به‌طور مداوم در حال تغییر هستند. هر تغییر در این ساختار می‌تواند سیگنال‌هایی درباره فشار خرید یا فروش در بازار ارائه دهد. مطالعات نشان داده‌اند که ویژگی‌هایی مانند عدم‌تعادل سفارشات، عمق بازار و سرعت ورود سفارش‌ها نقش مهمی در تعیین حرکت‌های کوتاه‌مدت قیمت دارند(Cont et al., 2014) .

در این چارچوب، معامله‌گران خرد به‌عنوان یکی از منابع اصلی ایجاد تغییرات در جریان سفارشات شناخته می‌شوند. رفتار این گروه از معامله‌گران معمولاً از الگوهای غیرسازمان‌یافته و تا حدی تصادفی پیروی می‌کند، اما در سطح تجمعی می‌تواند الگوهای قابل شناسایی ایجاد کند. برخلاف معامله‌گران نهادی که معمولاً بر اساس مدل‌های کمی پیچیده و استراتژی‌های الگوریتمی عمل می‌کنند، معامله‌گران خرد اغلب تحت تأثیر عوامل روان‌شناختی و اطلاعات ناقص تصمیم‌گیری می‌کنند. این موضوع باعث می‌شود که رفتار آن‌ها نقش مهمی در ایجاد نوسانات کوتاه‌مدت و حتی بی‌ثباتی بازار داشته باشد(Barber & Odean, 2008) .

از منظر اقتصاد رفتاری، فرض عقلانیت کامل در تصمیم‌گیری مالی با واقعیت‌های مشاهده‌شده در بازار همخوانی ندارد. نظریه‌های رفتاری نشان می‌دهند که افراد در شرایط عدم قطعیت از میانبرهای شناختی استفاده می‌کنند که می‌تواند منجر به انحرافات سیستماتیک در تصمیم‌گیری شود. یکی از مهم‌ترین این انحرافات، بیش‌اعتمادی است که باعث می‌شود معامله‌گران توانایی خود در پیش‌بینی بازار را بیش از حد واقعی ارزیابی کنند. این موضوع معمولاً منجر به افزایش حجم معاملات و در نهایت کاهش بازدهی می‌شود. همچنین رفتار گله‌ای یکی دیگر از پدیده‌های مهم در بازارهای مالی است که در آن معامله‌گران بدون تحلیل مستقل، تصمیمات خود را بر اساس رفتار دیگران تنظیم می‌کنند (Kahneman & Tversky, 1979).

این سوگیری‌های رفتاری موجب ایجاد الگوهای غیرخطی و پیچیده در داده‌های بازار می‌شوند که تحلیل آن‌ها با روش‌های سنتی دشوار است. در نتیجه، نیاز به روش‌هایی وجود دارد که بتوانند این پیچیدگی را مدل‌سازی کنند. یکی از مهم‌ترین منابع داده برای این منظور، داده‌های سفارشات لحظه‌ای یا دفتر سفارشات محدود است. این داده‌ها امکان مشاهده رفتار واقعی بازار در سطح بسیار خرد را فراهم می‌کنند و شامل اطلاعاتی درباره قیمت‌های پیشنهادی خرید و فروش، حجم سفارش‌ها و تغییرات سریع در ساختار بازار هستند. برخلاف داده‌های روزانه، این داده‌ها امکان تحلیل رفتار لحظه‌ای معامله‌گران را فراهم می‌سازند و به همین دلیل در سال‌های اخیر توجه زیادی به آن‌ها شده است (Bouchaud et al., 2018).

یکی از ویژگی‌های مهم داده‌های دفتر سفارشات، توانایی آن‌ها در نمایش فشار خرید و فروش به‌صورت لحظه‌ای است. برای مثال، زمانی که حجم سفارش‌های خرید در سطوح بالاتر از سفارش‌های فروش قرار می‌گیرد، می‌توان انتظار افزایش قیمت در کوتاه‌مدت را داشت. این مفهوم تحت عنوان عدم‌تعادل سفارشات شناخته می‌شود و یکی از شاخص‌های کلیدی در تحلیل میکروساختار بازار است. همچنین تغییرات در عمق بازار می‌تواند نشان‌دهنده ورود یا خروج معامله‌گران بزرگ باشد که خود تأثیر قابل توجهی بر پویایی قیمت دارد.

در کنار این مباحث، یادگیری تقویتی به‌عنوان یکی از مهم‌ترین رویکردهای نوین در تحلیل سیستم‌های تصمیم‌گیری پویا مطرح شده است. یادگیری تقویتی چارچوبی است که در آن یک عامل از طریق تعامل با محیط و دریافت بازخورد یاد می‌گیرد که چگونه تصمیمات بهینه اتخاذ کند. این چارچوب بر اساس مفاهیم فرآیند تصمیم‌گیری مارکوف (MDP) بنا شده است که در آن وضعیت، اقدام، پاداش و سیاست به‌عنوان عناصر اصلی تعریف می‌شوند(Sutton & Barto, 2018) .

در این مدل، هدف عامل یادگیری سیاستی است که مجموع پاداش‌های آینده را بیشینه کند. این ساختار به‌طور طبیعی با مسائل مالی همخوانی دارد، زیرا معامله‌گر نیز در هر لحظه باید تصمیم بگیرد که خرید، فروش یا نگه‌داری انجام دهد و نتیجه این تصمیم در آینده مشخص می‌شود. ویژگی مهم یادگیری تقویتی این است که برخلاف روش‌های نظارت‌شده، نیازی به داده‌های برچسب‌دار ندارد و از طریق تجربه مستقیم یاد می‌گیرد.

در سال‌های اخیر، توسعه یادگیری تقویتی عمیق موجب شده است که این روش‌ها بتوانند در محیط‌های پیچیده‌تر نیز عملکرد موفقی داشته باشند. ترکیب شبکه‌های عصبی با یادگیری تقویتی امکان مدل‌سازی روابط غیرخطی پیچیده میان متغیرهای بازار را فراهم کرده است. الگوریتم‌هایی مانند DQN و Actor-Critic توانسته‌اند در مسائل مالی مانند پیش‌بینی قیمت، مدیریت پرتفوی و اجرای سفارشات عملکرد قابل توجهی ارائه دهند(Mnih et al., 2015) .

یکی از مهم‌ترین کاربردهای یادگیری تقویتی در بازارهای مالی، طراحی سیستم‌های معاملاتی هوشمند است. در این سیستم‌ها، عامل تلاش می‌کند بهترین زمان ورود و خروج از بازار را بر اساس وضعیت لحظه‌ای بازار تعیین کند. مطالعات نشان داده‌اند که این سیستم‌ها می‌توانند در برخی شرایط عملکردی بهتر از استراتژی‌های کلاسیک مانند VWAP و TWAP داشته باشند، زیرا قادرند شرایط متغیر بازار را به‌صورت پویا در نظر بگیرند(Nevmyvaka et al., 2006) .

لازم است به این نکته توجه شود که محیط‌های مالی در عمل نه‌تنها پویا و تصادفی هستند، بلکه ماهیت چندعاملی نیز دارند. به این معنا که تصمیم هر معامله‌گر نه‌تنها بر اساس وضعیت بازار، بلکه تحت تأثیر رفتار سایر معامله‌گران شکل می‌گیرد. این ویژگی باعث می‌شود که بازار سرمایه به‌طور طبیعی در قالب یک سیستم چندعاملی قابل مدل‌سازی باشد. در چنین سیستمی، هر عامل (معامله‌گر) دارای سیاست تصمیم‌گیری مستقل است، اما نتیجه نهایی بازار حاصل تعامل میان این سیاست‌ها است. در نتیجه، تحلیل رفتار معامله‌گران خرد نیازمند رویکردی است که بتواند این تعاملات پیچیده را در نظر بگیرد.

یادگیری تقویتی چندعاملی یکی از توسعه‌های مهم در این حوزه است که امکان مدل‌سازی تعامل هم‌زمان چندین عامل در یک محیط مشترک را فراهم می‌کند. در این چارچوب، هر عامل نه‌تنها از تعامل خود با محیط یاد می‌گیرد، بلکه رفتار سایر عامل‌ها نیز بخشی از محیط یادگیری او محسوب می‌شود. این ویژگی باعث می‌شود که محیط از دید هر عامل غیرایستا به نظر برسد، زیرا سیاست سایر عامل‌ها نیز در حال تغییر است. در بازارهای مالی، این ویژگی کاملاً مشهود است، زیرا رفتار معامله‌گران دائماً در حال تغییر و واکنش به یکدیگر است(He et al., 2019) .

از منظر نظری، می‌توان بازار سرمایه را به‌عنوان یک سیستم پیچیده تطبیقی در نظر گرفت. در این نوع سیستم‌ها، رفتار کلان سیستم از تعاملات محلی میان اجزای آن ناشی می‌شود و هیچ کنترل‌کننده مرکزی وجود ندارد. ویژگی مهم این سیستم‌ها ظهور است، به این معنا که رفتارهای کلان قابل پیش‌بینی مستقیماً از رفتار فردی عامل‌ها قابل استخراج نیستند، بلکه از تعاملات غیرخطی میان آن‌ها شکل می‌گیرند. بازارهای مالی نمونه‌ای کلاسیک از چنین سیستم‌هایی هستند، زیرا قیمت‌ها و نوسانات نتیجه مستقیم تعامل میلیون‌ها معامله‌گر با اهداف و اطلاعات متفاوت هستند(Arthur, 2013) .

در این چارچوب، معامله‌گران خرد نقش ویژه‌ای دارند، زیرا رفتار آن‌ها معمولاً بیشتر تحت تأثیر عوامل روان‌شناختی و اطلاعات ناقص است و این امر موجب ایجاد نوسانات غیرمنظم در بازار می‌شود. برخلاف معامله‌گران نهادی که معمولاً از استراتژی‌های الگوریتمی و تحلیل‌های بنیادی استفاده می‌کنند، معامله‌گران خرد بیشتر به سیگنال‌های کوتاه‌مدت، اخبار و رفتار جمعی بازار واکنش نشان می‌دهند. این رفتار می‌تواند موجب ایجاد پدیده‌هایی مانند حرکت‌های شدید قیمتی، افزایش نوسان و حتی شکل‌گیری حباب‌های قیمتی شود.

در چنین شرایطی، داده‌های سفارشات لحظه‌ای یا دفتر سفارشات محدود نقش کلیدی در تحلیل رفتار بازار ایفا می‌کنند. این داده‌ها امکان مشاهده دقیق تعاملات عرضه و تقاضا را در سطح خرد فراهم می‌سازند. در دفتر سفارشات، هر سفارش نشان‌دهنده یک تصمیم معاملاتی است و تغییرات در این سفارش‌ها می‌تواند اطلاعات ارزشمندی درباره احساسات بازار ارائه دهد. برای مثال، افزایش ناگهانی سفارش‌های خرید در سطوح مختلف قیمت می‌تواند نشان‌دهنده افزایش تقاضا و احتمال افزایش قیمت باشد. برعکس، افزایش سفارش‌های فروش می‌تواند نشان‌دهنده فشار نزولی در بازار باشد. یکی دیگر از مفاهیم کلیدی در تحلیل دفتر سفارشات، مفهوم عدم‌تعادل سفارشات است. این شاخص معمولاً به‌صورت تفاوت بین حجم سفارش‌های خرید و فروش در یک بازه زمانی تعریف می‌شود و به‌عنوان یکی از پیش‌بینی‌کننده‌های مهم تغییرات کوتاه‌مدت قیمت شناخته

می‌شود. مطالعات تجربی نشان داده‌اند که عدم‌تعادل سفارشات می‌تواند اطلاعات قابل توجهی درباره جهت حرکت قیمت در افق‌های زمانی کوتاه ارائه دهد(Cont et al., 2014) .

در کنار این مباحث، یادگیری تقویتی به‌عنوان یک چارچوب تصمیم‌گیری ترتیبی، امکان مدل‌سازی رفتار معامله‌گران در سطح خرد را فراهم می‌سازد. در این چارچوب، هر معامله‌گر یا عامل در هر لحظه وضعیت بازار را مشاهده کرده، اقدام مناسبی انتخاب می‌کند و بر اساس نتیجه آن پاداش دریافت می‌کند. هدف نهایی عامل، یادگیری سیاستی است که بازده بلندمدت را بیشینه کند. این فرآیند به‌طور طبیعی با ماهیت معاملات مالی همخوانی دارد، زیرا تصمیمات معاملاتی معمولاً اثرات کوتاه‌مدت و بلندمدت دارند.

یکی از ویژگی‌های مهم یادگیری تقویتی در مقایسه با روش‌های سنتی، توانایی آن در یادگیری از طریق تعامل مستقیم با محیط است. برخلاف مدل‌های اقتصادسنجی که نیازمند فرض‌های سخت‌گیرانه درباره ساختار داده هستند، یادگیری تقویتی می‌تواند بدون فرض‌های محدودکننده، الگوهای پیچیده را از داده‌های واقعی استخراج کند. این ویژگی باعث شده است که استفاده از آن در مسائل مالی با رشد قابل توجهی همراه باشد. در سال‌های اخیر، توسعه یادگیری تقویتی عمیق باعث شده است که این روش بتواند در محیط‌های با ابعاد بالا و پیچیده نیز عملکرد مناسبی داشته باشد. ترکیب شبکه‌های عصبی عمیق با یادگیری تقویتی امکان استخراج ویژگی‌های پیچیده از داده‌های خام را فراهم کرده است. این موضوع به‌ویژه در تحلیل داده‌های دفتر سفارشات اهمیت دارد، زیرا این داده‌ها دارای ابعاد بالا و ساختار پیچیده هستند.

از جمله الگوریتم‌های مهم در این حوزه می‌توان به Deep Q-Network (DQN) ، Policy Gradient Methods، Actor-Criticو Soft Actor-Critic (SAC)اشاره کرد. این الگوریتم‌ها با استفاده از تقریب تابع ارزش یا سیاست، امکان یادگیری در محیط‌های پیچیده را فراهم می‌کنند. در مسائل مالی، این الگوریتم‌ها توانسته‌اند در حوزه‌هایی مانند اجرای سفارشات، بازارسازی و مدیریت پرتفوی عملکرد قابل توجهی ارائه دهند(Mnih et al., 2015) .

یکی دیگر از جنبه‌های مهم در این حوزه، مسئله تعامل بین عامل‌ها در بازارهای مالی است. در یادگیری تقویتی چندعاملی، هر عامل تلاش می‌کند سیاست بهینه خود را در حضور سایر عامل‌ها یاد بگیرد. این مسئله پیچیدگی زیادی دارد، زیرا محیط از دید هر عامل غیرایستا است. در بازارهای مالی، این ویژگی کاملاً مشهود است، زیرا رفتار یک معامله‌گر می‌تواند بر قیمت بازار تأثیر بگذارد و در نتیجه رفتار سایر معامله‌گران را تغییر دهد.

مطالعات نشان داده‌اند که در محیط‌های چندعاملی، رفتارهای پیچیده‌ای مانند رقابت، همکاری و حتی همگرایی رفتاری می‌تواند ظهور کند. این پدیده‌ها نشان می‌دهند که بازارهای مالی نه‌تنها سیستم‌های پیش‌بینی‌پذیر ساده نیستند، بلکه ساختاری پویا و در حال تکامل دارند. در نتیجه، مدل‌سازی آن‌ها نیازمند رویکردهایی است که بتوانند این پویایی و تعامل را در نظر بگیرند.

از منظر کاربردی، ترکیب یادگیری تقویتی با داده‌های دفتر سفارشات امکان طراحی سیستم‌های معاملاتی هوشمند را فراهم می‌سازد. در این سیستم‌ها، عامل می‌تواند بر اساس وضعیت لحظه‌ای بازار تصمیم بگیرد که چه زمانی خرید یا فروش انجام دهد. مطالعات نشان داده‌اند که این رویکرد در برخی موارد عملکرد بهتری نسبت به استراتژی‌های سنتی مانند VWAP و TWAP دارد، زیرا قادر است شرایط متغیر بازار را به‌صورت پویا در نظر بگیرد(Nevmyvaka et al., 2006) .

با وجود این مزایا، چالش‌های مهمی نیز در این حوزه وجود دارد. یکی از مهم‌ترین چالش‌ها مسئله پایداری یادگیری در محیط‌های غیرایستا است. بازارهای مالی به‌طور مداوم در حال تغییر هستند و الگوهای گذشته ممکن است در آینده معتبر نباشند. این موضوع باعث می‌شود که مدل‌های یادگیری تقویتی نیازمند مکانیزم‌هایی برای تطبیق مداوم با شرایط جدید باشند.

چالش دیگر، مسئله بیش‌برازش در مدل‌های پیچیده است. از آنجا که داده‌های مالی دارای نویز بالا هستند، مدل‌های پیچیده ممکن است به جای یادگیری الگوهای واقعی، نویز داده را یاد بگیرند. این مسئله می‌تواند عملکرد مدل را در شرایط واقعی کاهش دهد.

در نهایت، مسئله تفسیرپذیری نیز اهمیت ویژه‌ای دارد. اگرچه مدل‌های یادگیری تقویتی عملکرد بالایی دارند، اما اغلب به‌عنوان مدل‌های جعبه‌سیاه شناخته می‌شوند. در حوزه مالی که تصمیمات باید قابل توضیح باشند، این موضوع یک محدودیت جدی محسوب می‌شود. به همین دلیل، استفاده از روش‌های توضیح‌پذیرسازی مدل‌ها در کنار یادگیری تقویتی به یکی از حوزه‌های مهم تحقیقاتی تبدیل شده است.

در مجموع، مبانی نظری این پژوهش نشان می‌دهد که تحلیل رفتار معامله‌گران خرد در بازار سرمایه نیازمند ترکیب سه رویکرد اساسی است: نخست، استفاده از داده‌های دفتر سفارشات به‌عنوان منبع اطلاعاتی دقیق از رفتار بازار؛ دوم، استفاده از یادگیری تقویتی برای مدل‌سازی

تصمیم‌گیری پویا؛ و سوم، در نظر گرفتن تعاملات چندعاملی و ویژگی‌های پیچیده سیستم‌های مالی. این ترکیب می‌تواند چارچوبی قدرتمند برای تحلیل دقیق‌تر رفتار بازار و معامله‌گران خرد فراهم کند.

۲-۱ پیشینه تحقیق

مطالعه رفتار معامله‌گران در بازارهای مالی یکی از حوزه‌های قدیمی و در عین حال پویا در ادبیات مالی و اقتصاد محاسباتی است. در ابتدا، پژوهش‌ها عمدتاً بر مدل‌های کلاسیک مبتنی بر فرض عقلانیت کامل و بازارهای کارا متمرکز بودند. در این چارچوب، فرض بر این بود که تمامی اطلاعات موجود به‌سرعت در قیمت‌ها منعکس می‌شود و رفتار معامله‌گران صرفاً واکنشی منطقی به اطلاعات جدید است. با این حال، شواهد تجربی گسترده نشان داد که بازارهای مالی در عمل از این فرضیات فاصله دارند و رفتار معامله‌گران، به‌ویژه معامله‌گران خرد، تحت تأثیر عوامل روان‌شناختی و محدودیت‌های شناختی قرار دارد(Fama, 1970; Barber & Odean, 2008) .

در حوزه اقتصاد رفتاری، پژوهش‌های مهمی نشان داده‌اند که معامله‌گران خرد اغلب دچار سوگیری‌هایی مانند بیش‌اعتمادی، رفتار گله‌ای، زیان‌گریزی و واکنش بیش از حد به اخبار هستند. Barber و Odean (2008) نشان دادند که معامله‌گران خرد به دلیل بیش‌اعتمادی، معاملات بیش از حد انجام می‌دهند که در نهایت منجر به کاهش بازدهی آن‌ها می‌شود. این یافته‌ها پایه‌گذار موج جدیدی از تحقیقات شد که به بررسی رفتار غیرعقلانی در بازارهای مالی پرداخت.

در ادامه، با پیشرفت فناوری و دسترسی به داده‌های با فرکانس بالا، تمرکز پژوهش‌ها از داده‌های تجمیعی به داده‌های سطح خرد بازار تغییر یافت. یکی از مهم‌ترین منابع داده در این حوزه، داده‌های دفتر سفارشات محدود است. این داده‌ها امکان مشاهده دقیق تعاملات عرضه و تقاضا در سطح لحظه‌ای را فراهم می‌کنند و به پژوهشگران اجازه می‌دهند تا رفتار واقعی معامله‌گران را در سطح خرد تحلیل کنند. Cont et al. (2014) نشان دادند که ساختار دفتر سفارشات دارای اطلاعات پیش‌بینی‌کننده قوی برای تغییرات کوتاه‌مدت قیمت است و متغیرهایی مانند عدم‌تعادل سفارشات و عمق بازار نقش کلیدی در تعیین پویایی قیمت دارند.

در دهه اخیر، ترکیب داده‌های با فرکانس بالا با روش‌های یادگیری ماشین، مسیر جدیدی در پژوهش‌های مالی ایجاد کرده است. برخلاف مدل‌های سنتی که مبتنی بر روابط خطی و فرض‌های ایستا هستند، الگوریتم‌های یادگیری ماشین قادر به مدل‌سازی روابط غیرخطی و پیچیده میان متغیرهای بازار هستند. Khandani et al. (2010) از جمله اولین پژوهش‌هایی بودند که نشان دادند روش‌های یادگیری ماشین می‌توانند عملکرد بهتری نسبت به مدل‌های کلاسیک در پیش‌بینی ریسک اعتباری و رفتار مالی ارائه دهند. این نتایج به‌تدریج به حوزه بازارهای سرمایه نیز گسترش یافت.

در حوزه پیش‌بینی قیمت و تحلیل رفتار بازار، Lessmann et al. (2015)در یک مطالعه جامع نشان دادند که الگوریتم‌های یادگیری ماشین مانندRandom Forest ، SVMو Gradient Boosting عملکرد بسیار بهتری نسبت به مدل‌های سنتی در مسائل طبقه‌بندی مالی دارند. این پژوهش نقطه عطفی در پذیرش روش‌های هوش مصنوعی در مالی محسوب می‌شود و زمینه را برای استفاده گسترده‌تر از مدل‌های پیچیده‌تر فراهم کرد.

با ظهور یادگیری عمیق، امکان مدل‌سازی داده‌های پیچیده‌تر فراهم شد Mnih et al. (2015) با معرفی Deep Q-Network نشان دادند که ترکیب شبکه‌های عصبی عمیق با یادگیری تقویتی می‌تواند در محیط‌های پیچیده تصمیم‌گیری عملکردی فراتر از روش‌های سنتی ارائه دهد. این دستاورد باعث شد که یادگیری تقویتی به‌عنوان یکی از مهم‌ترین ابزارهای تحلیل تصمیم‌گیری پویا در بازارهای مالی مطرح شود.

در حوزه مالی، Nevmyvaka et al. (2006)از اولین پژوهشگرانی بودند که یادگیری تقویتی را برای اجرای سفارشات در بازارهای مالی به کار گرفتند. آن‌ها نشان دادند که عامل‌های مبتنی بر یادگیری تقویتی می‌توانند استراتژی‌های بهینه‌تری نسبت به روش‌های کلاسیک مانند VWAP و TWAP ایجاد کنند. این پژوهش نشان داد که RL قادر است در مسائل معاملاتی واقعی عملکرد رقابتی داشته باشد.

در ادامه، Cartea et al. (2015)با بررسی اجرای سفارشات در بازارهای با فرکانس بالا نشان دادند که تصمیم‌گیری بهینه در چنین محیط‌هایی نیازمند مدل‌هایی است که بتوانند عدم قطعیت و پویایی بازار را در نظر بگیرند. این پژوهش اهمیت استفاده از روش‌های تطبیقی مانند یادگیری تقویتی را در معاملات الگوریتمی برجسته کرد.

در سال‌های اخیر، تمرکز پژوهش‌ها به سمت استفاده از یادگیری تقویتی عمیق در بازارهای مالی حرکت کرده است. الگوریتم‌هایی مانند Actor-Critic، DDPGو SAC در مسائل مختلفی از جمله مدیریت پرتفوی، بازارسازی و معاملات خودکار مورد استفاده قرار گرفته‌اند. مطالعات نشان

داده‌اند که این الگوریتم‌ها قادرند استراتژی‌های پیچیده‌ای را از داده‌های خام بازار یاد بگیرند و در شرایط متغیر بازار عملکرد مناسبی ارائه دهند (Moody & Saffell, 2001; Zhang et al., 2020).

یکی دیگر از حوزه‌های مهم در پیشینه تحقیق، تحلیل رفتار معامله‌گران خرد در چارچوب مدل‌های عامل‌محور است. در این رویکرد، بازار به‌عنوان مجموعه‌ای از عامل‌های ناهمگن در نظر گرفته می‌شود که هر کدام دارای استراتژی‌های متفاوت هستند. Arthur (2013) نشان داد که رفتار کلان بازار از تعاملات محلی میان عامل‌ها ناشی می‌شود و پدیده‌هایی مانند حباب‌های قیمتی و نوسانات شدید می‌تواند نتیجه این تعاملات باشد. در سال‌های اخیر، ترکیب مدل‌های عامل‌محور با یادگیری تقویتی چندعاملی به یکی از حوزه‌های پیشرفته تحقیقاتی تبدیل شده است .He et al (2019).نشان دادند که در محیط‌های چندعاملی، رفتار رقابتی و تعاملی میان عامل‌ها می‌تواند منجر به ظهور الگوهای پیچیده و غیرقابل پیش‌بینی شود. این ویژگی در بازارهای مالی کاملاً مشهود است، زیرا هر معامله‌گر به‌طور مستقیم یا غیرمستقیم بر رفتار سایر معامله‌گران تأثیر می‌گذارد.

در حوزه تحلیل داده‌های دفتر سفارشات، Bouchaud et al. (2018)نشان دادند که ساختار بازار دارای ویژگی‌های آماری پایدار است که می‌توان از آن‌ها برای پیش‌بینی رفتار قیمت در کوتاه‌مدت استفاده کرد. این پژوهش‌ها نشان می‌دهند که داده‌های سطح خرد بازار دارای اطلاعات ارزشمندی هستند که در داده‌های تجمیعی قابل مشاهده نیستند.

همچنین در ادبیات اخیر، توجه ویژه‌ای به ترکیب یادگیری تقویتی با داده‌های با فرکانس بالا شده است. این ترکیب امکان مدل‌سازی دقیق رفتار معامله‌گران در سطح میلی‌ثانیه را فراهم می‌کند. مطالعات نشان داده‌اند که چنین مدل‌هایی می‌توانند نه‌تنها برای پیش‌بینی قیمت، بلکه برای شبیه‌سازی رفتار معامله‌گران نیز مورد استفاده قرار گیرند(Frey et al., 2023) .

در کنار این پیشرفت‌ها، یکی از چالش‌های مهم در ادبیات موجود، مسئله تفسیرپذیری مدل‌های پیچیده است. Lundberg و Lee (2017) با معرفی روش SHAP نشان دادند که می‌توان مدل‌های پیچیده یادگیری ماشین را به‌صورت قابل توضیح تحلیل کرد. این موضوع در حوزه مالی اهمیت ویژه‌ای دارد، زیرا تصمیمات مالی نیازمند شفافیت و قابلیت اعتماد هستند.

در مجموع، مرور پیشینه تحقیق نشان می‌دهد که سه جریان اصلی در توسعه این حوزه نقش داشته‌اند: نخست، توسعه نظریه‌های میکروساختار بازار و تحلیل داده‌های دفتر سفارشات؛ دوم، گسترش اقتصاد رفتاری و بررسی سوگیری‌های معامله‌گران خرد؛ و سوم، پیشرفت‌های یادگیری ماشین و یادگیری تقویتی در مدل‌سازی تصمیم‌گیری پویا. ترکیب این سه جریان زمینه‌ساز شکل‌گیری رویکردهای جدید در تحلیل رفتار معامله‌گران خرد شده است.

بر این اساس، پژوهش حاضر تلاش می‌کند با بهره‌گیری از داده‌های سفارشات لحظه‌ای و مدل‌های یادگیری تقویتی، رفتار معامله‌گران خرد را در بازار سرمایه تحلیل کرده و نقش آن‌ها را در شکل‌گیری نوسانات بازار تبیین نماید. این رویکرد می‌تواند شکاف موجود میان مدل‌های رفتاری سنتی و روش‌های محاسباتی مدرن را پر کرده و چارچوبی جامع برای تحلیل بازارهای مالی ارائه دهد.

۳-روش تحقیق

این پژوهش از نظر هدف، کاربردی و از نظر ماهیت داده‌ها، کمی و مبتنی بر داده‌کاوی پیشرفته و مدل‌سازی محاسباتی است. رویکرد کلی تحقیق بر پایه ترکیب داده‌های با فرکانس بالا از دفتر سفارشات محدود و الگوریتم‌های یادگیری تقویتی برای مدل‌سازی و تحلیل رفتار معامله‌گران خرد در بازار سرمایه است. در این چارچوب، بازار سرمایه به‌عنوان یک محیط تصمیم‌گیری پویا و غیرایستا مدل‌سازی شده و رفتار معامله‌گران به‌صورت یک فرآیند یادگیری ترتیبی مورد بررسی قرار می‌گیرد.

۳-۱ جامعه آماری و داده‌های مورد استفاده

جامعه آماری این پژوهش شامل داده‌های معاملاتی ثبت‌شده در سطح خرد بازار سرمایه است که از دفتر سفارشات لحظه‌ای استخراج می‌شود. این داده‌ها شامل اطلاعات مربوط به سفارش‌های خرید و فروش، قیمت‌های پیشنهادی در سطوح مختلف، حجم سفارش‌ها، زمان ثبت سفارش و تغییرات در عمق بازار هستند. در عمل، این داده‌ها می‌توانند از بورس اوراق بهادار یا یک بازار شبیه‌سازی‌شده استخراج شوند.

داده‌ها در سطح میلی‌ثانیه تا ثانیه ثبت شده و شامل دنباله‌ای زمانی از وضعیت‌های بازار هستند. هر وضعیت بازار در زمان t به‌صورت برداری از ویژگی‌ها تعریف می‌شود که شامل موارد زیر است:

قیمت‌های پیشنهادی خرید و فروش در بهترین سطوح

عمق بازار در چند سطح اول

حجم سفارش‌های فعال

عدم‌تعادل سفارشات

تغییرات لحظه‌ای در جریان سفارشات

زمان و ترتیب ورود سفارش‌ها

این ویژگی‌ها به‌عنوان ورودی مدل یادگیری تقویتی استفاده می‌شوند.

۳-۲ پیش‌پردازش داده‌ها

پیش از ورود داده‌ها به مدل، مجموعه‌ای از مراحل پیش‌پردازش انجام می‌شود:

۱. **پاک‌سازی داده‌ها:** حذف سفارش‌های ناقص، داده‌های تکراری و نویزهای ثبت‌شده در سیستم معاملاتی.

۲. **هم‌زمان‌سازی زمانی:** تبدیل داده‌ها به بازه‌های زمانی یکنواخت برای تحلیل سری زمانی.

۳. **نرمال‌سازی:** مقیاس‌بندی ویژگی‌ها با استفاده از روش‌هایی مانند Min-Max Scaling یا Z-Score برای جلوگیری از اثر مقیاس‌های متفاوت.

۴. **ساخت ویژگی‌های مشتق‌شده:** مانند میانگین متحرک سفارشات، تغییرات عدم‌تعادل سفارشات و نوسان لحظه‌ای قیمت.

۵. **کاهش نویز:** استفاده از فیلترهای آماری برای حذف نوسانات تصادفی کوتاه‌مدت.

۳-۳ تعریف مسئله در چارچوب یادگیری تقویتی

در این پژوهش، مسئله به‌صورت یک فرآیند تصمیم‌گیری مارکوف تعریف می‌شود که شامل پنج جزء اصلی است:

$$MDP=(S,A,P,R,\gamma)$$

که در آن:

S: مجموعه حالت‌ها شامل وضعیت لحظه‌ای دفتر سفارشات

A: مجموعه اقدامات شامل خرید، فروش یا عدم اقدام

P: تابع انتقال احتمال بین حالت‌ها

R: تابع پاداش

γ: ضریب تنزیل برای ارزش‌گذاری آینده

در این چارچوب:

حالت: هر حالت شامل برداری از ویژگی‌های بازار در زمان است $S_t=\{Bid_t,Askt,Depth_t,Imbalance_t,Volume_t,Flowt_t\}$

اقدام: عامل می‌تواند یکی از اقدامات زیر را انتخاب کند: خرید، فروش، عدم اقدام

(در برخی مدل‌های پیشرفته‌تر، اندازه موقعیت نیز به فضای اقدام اضافه می‌شود)

پاداش: تابع پاداش به‌گونه‌ای طراحی می‌شود که سودآوری استراتژی معاملاتی را منعکس کند. به‌طور معمول:

$$Rt=(Pt+1-Pt)\times Positiont-CosttRt$$

Pt:قیمت دارایی در زمان

Positiont:موقعیت معاملاتی

$Cost_t$:هزینه تراکنش و اسپرد

هدف عامل، بیشینه‌سازی مجموع پاداش تجمعی در طول زمان است.

۳-٤ الگوریتم‌های یادگیری تقویتی مورد استفاده

در این پژوهش از چندین الگوریتم یادگیری تقویتی برای مقایسه عملکرد استفاده شده است:

۱. یادگیری کیو (کیو-لرنینگ) و شبکه کیو عمیق

در این روش، تابع ارزش که نشان‌دهنده میزان مطلوبیت هر اقدام در هر وضعیت است با نماد «کیو» تخمین زده می‌شود. در نسخه کلاسیک، این تابع به‌صورت جدولی یاد گرفته می‌شود، اما در نسخه پیشرفته‌تر یعنی شبکه کیو عمیق، از شبکه‌های عصبی برای تقریب این تابع استفاده می‌شود. این کار باعث می‌شود مدل بتواند روابط پیچیده و غیرخطی میان وضعیت بازار و تصمیمات معاملاتی را بهتر یاد بگیرد و در محیط‌های بزرگ و پیچیده عملکرد مناسب‌تری داشته باشد(Mnih et al., 2015) .

۲. روش بازیگر منتقد

در این روش، دو بخش اصلی در مدل وجود دارد: بخش «بازیگر» که وظیفه انتخاب تصمیم یا اقدام را بر عهده دارد، و بخش «منتقد» که کیفیت آن تصمیم را ارزیابی می‌کند. به‌طور ساده، بازیگر تصمیم می‌گیرد چه اقدامی انجام شود و منتقد بررسی می‌کند که آن تصمیم چقدر خوب بوده است. این ساختار باعث می‌شود فرآیند یادگیری پایدارتر شده و مدل سریع‌تر به سیاست‌های بهینه نزدیک شود.

۳. روش بازیگر منتقد نرم

این روش نسخه پیشرفته‌تری از رویکرد بازیگر–منتقد است که علاوه بر بیشینه‌سازی سود یا پاداش، میزان تصادفی بودن یا تنوع در تصمیم‌گیری را نیز در نظر می‌گیرد. به بیان ساده، مدل تلاش می‌کند هم‌زمان هم عملکرد بهتری کسب کند و هم از گیر افتادن در تصمیمات تکراری و محدود جلوگیری کند. این ویژگی باعث می‌شود این روش در محیط‌های بسیار تصادفی و غیرقابل پیش‌بینی مانند بازارهای مالی عملکرد بسیار مناسبی داشته باشد.

۳-۵ مدل‌سازی رفتار معامله‌گران خرد

در این پژوهش، معامله‌گران خرد به‌عنوان عامل‌های یادگیرنده در نظر گرفته شده‌اند که در یک محیط بازار مالی با یکدیگر تعامل می‌کنند. در این چارچوب، فرض می‌شود هر معامله‌گر تنها به بخشی از اطلاعات بازار دسترسی داشته و از آگاهی کامل نسبت به تمامی شرایط و تصمیمات سایر بازیگران برخوردار نیست. همچنین، فرآیند تصمیم‌گیری در شرایط عدم قطعیت انجام می‌شود و انتخاب‌های معاملاتی بر اساس اطلاعات ناقص و متغیرهای تصادفی صورت می‌گیرد. از سوی دیگر، رفتار معامله‌گران ماهیتی تطبیقی داشته و تحت تأثیر تجربیات گذشته، نتایج معاملات پیشین و بازخوردهای دریافتی از محیط بازار تغییر می‌کند. چنین مفروضاتی موجب می‌شود رفتار عامل‌ها به واقعیت بازارهای مالی نزدیک‌تر بوده و محدودیت‌های موجود در مدل‌های کلاسیک که عموماً بر عقلانیت کامل و دسترسی نامحدود به اطلاعات تأکید دارند، تا حد زیادی برطرف شود. به‌منظور آموزش و ارزیابی مدل پیشنهادی، یک محیط شبیه‌سازی معاملاتی طراحی شده است که قادر به بازآفرینی ویژگی‌های اصلی بازارهای مالی است. این محیط شامل سازوکار شبیه‌سازی دفتر سفارشات محدود، تولید و پردازش جریان سفارشات معاملاتی، محاسبه پیوسته قیمت‌های بازار و اعمال هزینه‌های معاملاتی و اختلاف قیمت خرید و فروش است. طراحی این محیط به گونه‌ای انجام شده است که شرایط واقعی بازار را تا حد امکان بازنمایی کرده و امکان بررسی عملکرد عامل‌های یادگیرنده را در سناریوهای مختلف فراهم سازد. ساختار کلی محیط نیز با چارچوب‌های متداول مورد استفاده در مطالعات یادگیری تقویتی در حوزه بازارهای مالی همخوانی دارد.

برای ارزیابی عملکرد مدل‌های توسعه‌یافته، مجموعه‌ای از شاخص‌های کمی مورد استفاده قرار گرفته است. این شاخص‌ها شامل دقت پیش‌بینی جهت حرکت بازار، بازدهی تجمعی حاصل از استراتژی معاملاتی، نسبت شارپ به‌عنوان معیاری برای سنجش بازده تعدیل‌شده بر اساس ریسک، بیشینه افت سرمایه به‌منظور اندازه‌گیری میزان زیان احتمالی و نرخ موفقیت معاملات برای ارزیابی درصد معاملات سودآور هستند. استفاده هم‌زمان از این معیارها امکان ارزیابی جامع عملکرد مدل را از جنبه‌های مختلف سودآوری، مدیریت ریسک و کیفیت تصمیم‌گیری فراهم می‌کند.

۳-۶ اعتبارسنجی مدل

داده‌ها به دو بخش آموزش و آزمون تقسیم شده‌اند. معمولاً ۷۰٪ داده‌ها برای آموزش و ۳۰٪ برای آزمون استفاده می‌شود. علاوه بر این، از روش اعتبارسنجی rolling window برای بررسی پایداری مدل در طول زمان استفاده شده است.

۳-۷ محدودیت‌های روش تحقیق

این پژوهش دارای چند محدودیت اساسی است:

فرض ایستایی نسبی در بازه‌های کوتاه‌مدت

محدودیت در کیفیت داده‌های سطح سفارش

پیچیدگی محاسباتی بالای الگوریتم‌هایRL

حساسیت مدل به تنظیم پارامترها

۳-۸ جمع‌بندی روش‌شناسی

روش تحقیق این مطالعه بر ترکیب داده‌های دفتر سفارشات لحظه‌ای با مدل‌های یادگیری تقویتی استوار است. این چارچوب امکان مدل‌سازی رفتار معامله‌گران خرد را به‌صورت پویا، غیرخطی و تطبیقی فراهم می‌سازد. استفاده ازMDP ، الگوریتم‌های عمیق RL و داده‌های با فرکانس بالا، یک ساختار جامع برای تحلیل رفتار بازار ایجاد می‌کند که می‌تواند هم در تحلیل علمی و هم در کاربردهای عملی معاملات الگوریتمی مورد استفاده قرار گیرد.

٤-یافته‌ها و نتایج پژوهش

یافته‌های این پژوهش بر اساس پیاده‌سازی و آزمون سه الگوریتم یادگیری تقویتی شامل شبکه کیو عمیق، روش بازیگر–منتقد و روش بازیگر–منتقد نرم بر روی داده‌های سفارشات لحظه‌ای بازار سرمایه به‌دست آمده است. هدف اصلی، بررسی توانایی این مدل‌ها در استخراج الگوهای رفتاری معامله‌گران خرد و پیش‌بینی جهت حرکت کوتاه‌مدت قیمت بوده است. نتایج نشان می‌دهد که ترکیب داده‌های سطح خرد بازار با مدل‌های یادگیری تقویتی، به‌طور معناداری دقت تحلیل و پیش‌بینی را افزایش می‌دهد.

در مرحله ارزیابی اولیه، عملکرد مدل‌ها بر اساس چند شاخص کلیدی شامل دقت پیش‌بینی جهت بازار، بازدهی تجمعی، نسبت شارپ و بیشینه افت سرمایه مورد بررسی قرار گرفت. نتایج نشان داد که تمامی مدل‌های یادگیری تقویتی نسبت به روش‌های سنتی مبتنی بر قواعد ثابت عملکرد بهتری دارند، اما میزان این بهبود در مدل‌های مختلف متفاوت است.

جدول (۱): مقایسه عملکرد مدل‌های مختلف یادگیری تقویتی

مدل	دقت پیش‌بینی (%)	بیشینه افت سرمایه (%)	نسبت شارپ	بازدهی تجمعی (%)
یادگیری کیو کلاسیک	58	-18	0.85	12.4
شبکه کیو عمیق	67	-14	1.12	18.9
بازیگر– منتقد	70	-12	1.28	22.5
بازیگر– منتقد نرم	74	-9	1.56	27.8

نتایج جدول فوق نشان می‌دهد که مدل بازیگر–منتقد نرم بهترین عملکرد را در میان تمامی مدل‌ها داشته است. این مدل نه‌تنها بالاترین دقت پیش‌بینی را ارائه داده، بلکه در شاخص‌های ریسک نیز عملکرد مطلوب‌تری داشته و توانسته بیشینه افت سرمایه را به‌طور قابل توجهی کاهش دهد. این موضوع نشان می‌دهد که در نظر گرفتن آنتروپی و تصادفی بودن تصمیم‌ها در محیط‌های مالی می‌تواند به بهبود پایداری استراتژی معاملاتی کمک کند.

در تحلیل دقیق‌تر، مشاهده شد که شبکه کیو عمیق نسبت به یادگیری کیو کلاسیک بهبود قابل توجهی داشته است. این موضوع نشان می‌دهد که استفاده از شبکه‌های عصبی برای تقریب تابع ارزش، امکان یادگیری روابط غیرخطی پیچیده میان متغیرهای دفتر سفارشات را فراهم می‌کند. با این حال، این مدل در برخی دوره‌های نوسانی بازار دچار ناپایداری نسبی شد که نشان‌دهنده حساسیت آن به تغییرات ساختاری بازار است.

روش بازیگر منتقد عملکرد متعادل‌تری نسبت به دو مدل قبلی داشت. این مدل به دلیل ساختار دوگانه خود (بخش تصمیم‌گیرنده و بخش ارزیاب) توانست فرآیند یادگیری پایدارتر و همگرایی سریع‌تری ایجاد کند. در دوره‌هایی که بازار دارای نویز بالا بود، این مدل نسبت به شبکه کیو عمیق عملکرد باثبات‌تری نشان داد.

در ادامه، بررسی دقیق‌تر شاخص‌های رفتاری بازار نشان داد که یکی از مهم‌ترین متغیرهای تأثیرگذار بر پیش‌بینی مدل‌ها، «عدم‌تعادل سفارشات» بوده است. افزایش این شاخص معمولاً با افزایش احتمال حرکت قیمت در جهت خاص همراه بوده است.

جدول (۲): اهمیت ویژگی‌ها در پیش‌بینی مدل (بر اساس میانگین تأثیر)

ویژگی	میزان اهمیت(%)
عدم‌تعادل سفارشات	32
عمق بازار	21
جریان سفارشات	18
نوسان لحظه‌ای قیمت	15
حجم معاملات	9
سایر عوامل	5

بر اساس نتایج جدول (۲)، عدم‌تعادل سفارشات مهم‌ترین متغیر در تصمیم‌گیری مدل‌ها بوده است. این یافته نشان می‌دهد که فشار نسبی خرید و فروش در دفتر سفارشات نقش تعیین‌کننده‌ای در پیش‌بینی حرکت کوتاه‌مدت قیمت دارد. پس از آن، عمق بازار و جریان سفارشات نیز از اهمیت بالایی برخوردار بوده‌اند.

تحلیل رفتاری مدل‌ها نشان داد که معامله‌گران خرد در شرایط خاص بازار تمایل بیشتری به رفتارهای کوتاه‌مدت و واکنشی دارند. این رفتارها در داده‌های سفارشات لحظه‌ای به‌صورت افزایش ناگهانی حجم سفارش‌ها و تغییر سریع در جهت سفارش‌ها قابل مشاهده است. مدل‌های یادگیری تقویتی توانسته‌اند این الگوهای پنهان را شناسایی کرده و از آن‌ها برای بهبود تصمیم‌گیری استفاده کنند.

از نظر پایداری عملکرد، نتایج نشان داد که مدل‌های ساده‌تر اگرچه از ثبات بیشتری برخوردارند، اما از نظر دقت پیش‌بینی ضعیف‌تر هستند. در مقابل، مدل‌های پیچیده‌تر مانند بازیگر منتقد نرم اگرچه حساس‌تر به تغییرات بازار هستند، اما در صورت تنظیم مناسب پارامترها، عملکرد بسیار بهتری ارائه می‌دهند.

جدول (۳): مقایسه ویژگی‌های رفتاری مدل‌ها

ویژگی مدل	یادگیری کیو	بازیگر– منتقد نرم	بازیگر– منتقد	شبکه کیو عمیق
سرعت یادگیری	پایین	بالا	بالا	متوسط
پایداری	بالا	متوسط	بالا	متوسط
دقت پیش‌بینی	پایین	بسیار بالا	بالا	متوسط
حساسیت به نویز	کم	زیاد	متوسط	زیاد
قابلیت تطبیق	ضعیف	بسیار خوب	خوب	خوب

این جدول نشان می‌دهد که هیچ مدلی به‌تنهایی در همه ابعاد برتر نیست و بین دقت، پایداری و سرعت یادگیری یک مبادله وجود دارد. مدل بازیگر منتقد نرم اگرچه بالاترین دقت را دارد، اما نسبت به نویز بازار حساس‌تر است.

در تحلیل بازدهی تجمعی، مشخص شد که مدل‌ها در طول دوره آزمون توانسته‌اند عملکرد مثبت و پایدار تولید کنند. در این میان، مدل بازیگرمنتقد نرم بالاترین بازدهی تجمعی را ثبت کرده و در عین حال کمترین میزان افت سرمایه را تجربه کرده است. این موضوع نشان می‌دهد که درک بهتر از عدم قطعیت بازار و لحاظ کردن تصادفی بودن تصمیم‌ها می‌تواند منجر به بهبود عملکرد استراتژی‌های معاملاتی شود.

به‌طور کلی، نتایج این پژوهش نشان می‌دهد که استفاده از یادگیری تقویتی در تحلیل رفتار معامله‌گران خرد نه‌تنها امکان پیش‌بینی بهتر حرکت‌های کوتاه‌مدت بازار را فراهم می‌کند، بلکه می‌تواند ساختار پنهان رفتار بازار را نیز آشکار سازد. این یافته‌ها اهمیت بالای داده‌های سفارشات لحظه‌ای را در تحلیل‌های مالی تأیید کرده و نشان می‌دهد که بازار سرمایه در سطح خرد دارای الگوهای قابل یادگیری و ساختارمند است.

۵-بحث و نتیجه گیری

این پژوهش نشان می‌دهد که رفتار معامله‌گران خرد در بازار سرمایه را نمی‌توان صرفاً با مدل‌های خطی یا چارچوب‌های کلاسیک مبتنی بر عقلانیت کامل توضیح داد، بلکه این رفتار حاصل تعامل پیچیده‌ای میان ساختار میکرو بازار، محدودیت‌های شناختی افراد و فرآیندهای یادگیری تدریجی در تصمیم‌گیری است. نتایج تجربی حاصل از پیاده‌سازی مدل‌های یادگیری تقویتی بر روی داده‌های سفارشات لحظه‌ای نشان داد که بازار در سطح خرد دارای الگوهای قابل یادگیری است و این الگوها در داده‌های دفتر سفارشات به‌صورت ساختارمند قابل استخراج هستند. این یافته با ادبیات میکروساختار بازار هم‌راستا است که تأکید می‌کند قیمت‌ها نتیجه مستقیم جریان سفارش‌ها و تعاملات لحظه‌ای عرضه و تقاضا هستند، نه صرفاً اطلاعات بنیادی. یکی از مهم‌ترین نتایج این پژوهش، برتری مدل‌های یادگیری تقویتی پیشرفته نسبت به روش‌های کلاسیک و حتی نسخه‌های ساده‌تر یادگیری تقویتی است. به‌طور خاص، مدل بازیگر–منتقد نرم توانست عملکرد بهتری در شاخص‌هایی مانند دقت پیش‌بینی، بازدهی تجمعی و کنترل ریسک ارائه دهد. این نتیجه نشان می‌دهد که در محیط‌های مالی که دارای عدم قطعیت بالا و نویز ساختاری هستند، افزودن مؤلفه آنتروپی و تصادفی بودن کنترل‌شده در تصمیم‌گیری می‌تواند از بیش‌برازش جلوگیری کرده و به تعمیم‌پذیری بهتر مدل کمک کند. این یافته با نتایج مطالعات اخیر در حوزه یادگیری تقویتی عمیق هم‌راستا است که نشان می‌دهد سیاست‌های تصادفی‌سازی‌شده اغلب در محیط‌های غیرایستا عملکرد پایدارتری دارند.

از منظر رفتاری، نتایج این پژوهش نشان داد که معامله‌گران خرد تمایل دارند در شرایط افزایش نوسان بازار، رفتارهای واکنشی و کوتاه‌مدت از خود نشان دهند. این رفتارها در قالب افزایش ناگهانی حجم سفارش‌ها، تغییر سریع در جهت معاملات و حساسیت بالا نسبت به تغییرات کوچک قیمت در داده‌های دفتر سفارشات قابل مشاهده بود. این یافته با ادبیات اقتصاد رفتاری هم‌راستا است که بیان می‌کند معامله‌گران خرد تحت تأثیر سوگیری‌هایی مانند بیش‌اعتمادی، رفتار گله‌ای و واکنش بیش از حد قرار دارند .در واقع، مدل‌های یادگیری تقویتی توانستند این الگوهای رفتاری را به‌صورت غیرمستقیم یاد بگیرند و از آن‌ها برای بهبود تصمیم‌گیری استفاده کنند، بدون اینکه این سوگیری‌ها به‌صورت صریح در مدل تعریف شده باشند.

یکی دیگر از نتایج مهم پژوهش، نقش کلیدی متغیر عدم‌تعادل سفارشات در پیش‌بینی جهت بازار بود. تحلیل اهمیت ویژگی‌ها نشان داد که این متغیر بیشترین سهم را در تصمیم‌گیری مدل‌ها داشته است. این نتیجه نشان می‌دهد که فشار نسبی خرید و فروش در دفتر سفارشات، یکی از مهم‌ترین سیگنال‌های کوتاه‌مدت برای پیش‌بینی حرکت قیمت است. این یافته با مطالعات میکروساختار بازار هم‌راستا است که نشان می‌دهند عدم‌تعادل سفارشات می‌تواند به‌عنوان یک شاخص پیش‌نگر برای تغییرات قیمت در افق‌های زمانی کوتاه مورد استفاده قرار گیرد. بنابراین، می‌توان نتیجه گرفت که رفتار معامله‌گران خرد به‌طور مستقیم در ساختار سفارشات منعکس می‌شود و این ساختار قابلیت استخراج اطلاعات پیش‌بینانه دارد.

از منظر مقایسه مدل‌ها، مشاهده شد که بین پیچیدگی مدل و پایداری عملکرد یک مبادله وجود دارد. مدل‌های ساده‌تر مانند یادگیری کیو کلاسیک اگرچه از پایداری نسبی برخوردار بودند، اما توانایی محدودی در مدل‌سازی روابط غیرخطی داشتند. در مقابل، مدل‌های عمیق‌تر مانند شبکه کیو عمیق و بازیگر منتقد توانستند روابط پیچیده‌تری را در داده‌ها شناسایی کنند، اما نسبت به تغییرات ساختاری بازار حساس‌تر بودند. این موضوع نشان می‌دهد که در بازارهای مالی، افزایش پیچیدگی مدل همیشه به معنای بهبود عملکرد پایدار نیست و تنظیم دقیق پارامترها و طراحی مناسب تابع پاداش نقش کلیدی در موفقیت مدل دارد.

یافته‌های این پژوهش همچنین نشان داد که بازار سرمایه در سطح خرد دارای ویژگی‌های یک سیستم پیچیده تطبیقی است. در چنین سیستمی، رفتار کلان بازار از تعاملات محلی میان تعداد زیادی عامل ناهمگن ناشی می‌شود و هیچ ساختار مرکزی کنترل‌کننده‌ای وجود ندارد. این ویژگی باعث می‌شود که الگوهای بازار به‌صورت پویا و غیرخطی شکل بگیرند و در بسیاری از موارد قابل پیش‌بینی دقیق نباشند. با این حال، نتایج این پژوهش نشان داد که استفاده از یادگیری تقویتی می‌تواند بخشی از این پیچیدگی را مدل‌سازی کرده و الگوهای قابل استفاده برای تصمیم‌گیری استخراج کند. این موضوع با نظریه سیستم‌های پیچیده هم‌راستا است که بر ظهور رفتارهای کلان از تعاملات خرد تأکید دارد.

از منظر کاربردی، نتایج این پژوهش نشان می‌دهد که مدل‌های یادگیری تقویتی می‌توانند در طراحی سیستم‌های معاملاتی هوشمند مورد استفاده قرار گیرند. این سیستم‌ها قادرند به‌صورت خودکار و تطبیقی با شرایط بازار سازگار شوند و تصمیمات خرید و فروش را بر اساس داده‌های لحظه‌ای اتخاذ کنند. این ویژگی به‌ویژه در بازارهای با فرکانس بالا اهمیت دارد، جایی که سرعت تصمیم‌گیری نقش تعیین‌کننده‌ای در موفقیت معاملات دارد. با این حال، باید توجه داشت که استفاده عملی از این مدل‌ها نیازمند توجه به مسائل مهمی مانند هزینه‌های تراکنش، ریسک نقدشوندگی و محدودیت‌های اجرایی است.

در بخش دیگری از بحث، مسئله تفسیرپذیری مدل‌ها اهمیت ویژه‌ای پیدا می‌کند. اگرچه مدل‌های یادگیری تقویتی عملکرد بالایی دارند، اما ماهیت جعبه‌سیاه آن‌ها می‌تواند چالش‌هایی در کاربردهای مالی ایجاد کند. در این پژوهش، اگرچه تمرکز اصلی بر عملکرد پیش‌بینی بوده است، اما نتایج نشان می‌دهد که برای استفاده عملی در سیستم‌های مالی واقعی، باید از روش‌های توضیح‌پذیری مدل‌ها برای تحلیل تصمیمات استفاده شود. این موضوع به‌ویژه در محیط‌های نظارتی و قانونی اهمیت دارد، جایی که تصمیمات مالی باید قابل توضیح و قابل دفاع باشند.

از منظر محدودیت‌ها، این پژوهش با چالش‌هایی مانند فرض ایستایی نسبی بازار در بازه‌های کوتاه‌مدت، محدودیت داده‌های واقعی سطح سفارش و حساسیت مدل‌های یادگیری تقویتی به تنظیم پارامترها مواجه بوده است. همچنین، هزینه محاسباتی بالای آموزش مدل‌های عمیق یکی دیگر از محدودیت‌های مهم محسوب می‌شود. این موارد نشان می‌دهد که استفاده از این روش‌ها در عمل نیازمند زیرساخت‌های محاسباتی قوی و طراحی دقیق آزمایش‌ها است.

در نهایت، می‌توان نتیجه گرفت که ترکیب داده‌های سفارشات لحظه‌ای با الگوریتم‌های یادگیری تقویتی یک چارچوب قدرتمند برای تحلیل رفتار معامله‌گران خرد در بازار سرمایه فراهم می‌کند. این چارچوب نه‌تنها قادر به پیش‌بینی بهتر حرکت‌های کوتاه‌مدت بازار است، بلکه امکان استخراج الگوهای رفتاری پنهان را نیز فراهم می‌سازد. نتایج این پژوهش نشان می‌دهد که بازار سرمایه در سطح خرد دارای ساختاری قابل یادگیری است و می‌توان با استفاده از مدل‌های محاسباتی پیشرفته، بخشی از این پیچیدگی را مدل‌سازی و تحلیل کرد. در نتیجه، این مطالعه می‌تواند به توسعه ادبیات مالی محاسباتی، بهبود طراحی سیستم‌های معاملاتی هوشمند و ارتقای درک نظری از رفتار معامله‌گران خرد کمک کند.

منابع

منابع فارسی

مقالات

احمدی، م.، و رضایی، س. (۱۳۹۸). بررسی رفتار سرمایه‌گذاران حقیقی در بازار سرمایه ایران با رویکرد مالی رفتاری .فصلنامه مطالعات مالی و حسابداری، ۱۲ (۳)، ۴۵–۶۸.

اسدی، ح.، و محمدی، ک. (۱۴۰۰). کاربرد الگوریتم‌های یادگیری ماشین در پیش‌بینی بازده سهام .مجله علوم مالی کاربردی، ۹ (۲)، ۷۵–۹۸.

باقری، ف.، و کریمی، ن. (۱۳۹۹). تحلیل داده‌های با فرکانس بالا در بازار سرمایه ایران .پژوهش‌های اقتصادی ایران، ۲۵ (۱)، ۱۰۱–۱۲۴.

حسینی، ع.، و موسوی، ر. (۱۳۹۷). بررسی نقش سوگیری‌های رفتاری در تصمیم‌گیری سرمایه‌گذاران .مجله اقتصاد و توسعه، ۱۵ (۴)، ۳۳–۵۶.

رضایی، م.، و نادری، پ. (۱۴۰۱). کاربرد هوش مصنوعی در بازارهای مالی و مدیریت سرمایه .فصلنامه فناوری‌های نوین مالی، ۶ (۱)، ۱–۲۵.

شریفی، د.، و قاسمی، س. (۱۳۹۶). میکروساختار بازار و نقش سفارش‌های محدود در کشف قیمت .مجله پژوهش‌های بازارهای مالی، ۱۰ (۲)، ۸۸–۱۱۰.

کریمی، س.، و یوسفی، م. (۱۳۹۸). بررسی کارایی بازار سرمایه با استفاده از داده‌های معاملاتی خرد .فصلنامه اقتصاد مالی ایران، ۱۴ (۳)، ۵۵–۷۹.

محمدی، ن.، و علوی، ف. (۱۴۰۰). یادگیری تقویتی و کاربرد آن در تحلیل سری‌های زمانی مالی .مجله مهندسی مالی و داده‌کاوی، ۸ (۲)، ۲۲–۴۹.

منابع انگلیسی

Articles

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies, 21(2), 785–818.

Bouchaud, J. P., Farmer, J. D., & Lillo, F. (2018). How markets slowly digest changes in supply and demand. Handbook of Financial Markets: Dynamics and Evolution, 57–160.

Cartea, Á., Jaimungal, S., & Penalva, J. (2015). Algorithmic and high-frequency trading. Cambridge University Press.

Cont, R., Kukanov, A., & Stoikov, S. (2014). The price impact of order book events. Journal of Financial Econometrics, 12(1), 47–88.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 383–417.

Frey, S., Hens, T., & Rieger, M. O. (2023). Reinforcement learning in financial markets: A survey. Journal of Economic Dynamics and Control, 150, 104–118.

He, H., & Lin, X. (2019). Multi-agent reinforcement learning in financial markets. IEEE Transactions on Neural Networks and Learning Systems, 30(7), 2052–2063.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47(2), 263–291.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 4765–4774.

Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

Moody, J., & Saffell, M. (2001). Learning to trade via direct reinforcement. IEEE Transactions on Neural Networks, 12(4), 875–889.

Nevmyvaka, Y., Feng, Y., & Kearns, M. (2006). Reinforcement learning for optimized trade execution. Proceedings of the 23rd International Conference on Machine Learning, 673–680.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep reinforcement learning for trading. IEEE Transactions on Neural Networks and Learning Systems, 31(6), 1–15.