پیش‌بینی قیمت سهام در بازارهای مالی با استفاده از الگوریتم ترکیبی GSO–SVR

نوع مقاله: پژوهشی

جواد سالاری اسکری[1]
حسین شفیعی[2]*
عباس شیبانی[3]

تاریخ دریافت: تاریخ پذیرش:

چکیده

هدف: *هدف این پژوهش بررسی کارایی مدل ترکیبی الگوریتم جستجوی طلایی و رگرسیون بردار پشتیبان (GSO–SVR) در پیش‌بینی قیمت پایانی روزانه سهام است.*

روش‌شناسی : *این پژوهش از نظر هدف کاربردی و از نظر روش کمی است. داده‌های قیمت پایانی روزانه مربوط به هشت شرکت سهامی عام (اپل، گوگل، فیس‌بوک، ای‌بی‌ام، مایکروسافت، توییتر، یاهو و فورد) گردآوری شد. طرح تجربی برای حفظ ترتیب زمانی و کاهش نشت اطلاعات، داده‌ها را به‌صورت زمانی تفکیک کرده و انتخاب فرابارامترها را به مجموعه اعتبارسنجی محدود کرد. الگوریتم جستجوی طلایی برای تعیین مقادیر مناسب پارامتر جریمه SVR و ضریب هسته تابع پایه شعاعی با کمینه‌سازی خطای میانگین مربعات ریشه (RMSE) به کار رفت و عملکرد آن با SVR سنتی و SVR بهینه‌شده با الگوریتم ژنتیک (GA–SVR) مقایسه شد. معیارهای ارزیابی شامل خطای میانگین مربعات نرمال‌شده (NMSE)، خطای مطلق میانگین (MAE) و تقارن جهتی (DS) بودند.*

یافته‌ها: *نتایج نشان داد که مدل پیشنهادی GSO–SVR در تمامی هشت شرکت توانست کمترین میزان NMSE و MAE را به دست آورد و در ۷ شرکت از ۸ شرکت بالاترین تقارن جهتی (DS) را ثبت کند؛ این امر حاکی از بهبود چشمگیر دقت پیش‌بینی عددی و توافق جهتی قیمت‌ها نسبت به مدل‌های بنچمارک است.*

نتیجه‌گیری: *بهینه‌سازی فرابارامترها بر مبنای الگوریتم جستجوی طلایی می‌تواند عملکرد پیش‌بینی سری‌های زمانی مالی را به‌طور محسوسی ارتقا دهد، هرچند اثربخشی آن باید برای هر سهم و معیار ارزیابی به‌صورت تجربی بررسی شود.*

کلیدواژه‌ها: پیش‌بینی قیمت سهام؛ رگرسیون بردار پشتیبان؛ الگوریتم جستجوی طلایی؛ الگوریتم ژنتیک؛ بهینه‌سازی فرابارامترها؛ سری‌های زمانی مالی
طبقه‌بندی JEL: C45, G17, C61, M41

۱ دانشجوی دکتری گروه حسابداری، واحد سیرجان، دانشگاه آزاد اسلامی، سیرجان، ایران
3131527528@iau.ir
۲ گروه حسابداری، واحد سیرجان، دانشگاه آزاد اسلامی، سیرجان، ایران (نویسنده مسئول)
hosseinshafiee@iau.ac.ir
۳ گروه حسابداری، واحد سیرجان، دانشگاه آزاد اسلامی، سیرجان، ایران
Sheybani.abbas@iau.ac.ir

[Downloaded from mieaoi.ir on 2026-11-19]

مقدمه

پیش‌بینی قیمت سهام از مسائل مهم در اقتصادسنجی مالی، مالی کمّی و تصمیم‌گیری داده‌محور است. سرمایه‌گذاران و تحلیلگران مالی از پیش‌بینی‌ها برای پشتیبانی از تخصیص پرتفوی، ارزیابی ریسک و سنجش شرایط بازار بهره می‌گیرند. بااین‌حال، پیش‌بینی سری‌های زمانی مالی دشوار است؛ زیرا ویژگی‌های آماری این سری‌ها ممکن است در گذر زمان تغییر کنند و حرکات مشاهده‌شده حاصل برهم‌کنش اطلاعات بازار، رفتار سرمایه‌گذاران، شرایط کلان اقتصادی و تحولات مختص هر شرکت باشند. فرضیه بازار کارا، با مرتبط‌کردن قابلیت پیش‌بینی قیمت دارایی‌ها با اطلاعات منعکس‌شده در قیمت بازار، دیدگاهی نظری و بنیادین درباره این دشواری ارائه می‌کند(فاما[1]، ۱۹۷۰).این دیدگاه لزوماً به معنای شکست همه مدل‌های پیش‌بینی نیست، اما بر اهمیت تمایز میان عملکرد واقعی خارج از نمونه و الگوهایی تأکید می‌کند که ممکن است از انتخاب مدل یا بیش‌برازش داده‌های تاریخی ناشی شوند. ازاین‌رو، ارزیابی پیش‌بینی قیمت سهام باید بر آزمایش‌هایی دقیق استوار باشد که ترتیب زمانی مشاهدات را حفظ کنند و از به‌کارگیری اطلاعات آینده در فرایند توسعه مدل جلوگیری کنند.

روش‌های آماری متعارف، از جمله مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و مدل ناهمسانی واریانس شرطی خودرگرسیو تعمیم‌یافته (GARCH)، ابزارهایی جاافتاده برای تحلیل وابستگی زمانی و نوسان‌پذیری شرطی‌اند. بااین‌حال، عملکرد پیش‌بینی هر مدل آماری به ویژگی‌های فرایند مولد داده و هدف پیش‌بینی وابسته است. دسترس‌پذیری روزافزون داده‌های مالی، کاربرد روش‌های یادگیری ماشین را نیز گسترش داده است؛ روش‌هایی که توانایی بازنمایی روابط غیرخطی و برهم‌کنش‌های پیچیده میان پیش‌بین‌ها را دارند. مرورهای نظام‌مند پژوهش‌های پیش‌بینی بازار سهام، ماشین‌های بردار پشتیبان، شبکه‌های عصبی مصنوعی، روش‌های تجمیعی و معماری‌های یادگیری عمیق را از رویکردهای اصلی این حوزه معرفی می‌کنند(کومبوره و همکاران[2]، ۲۰۲۲؛ لین و مارکز[3]، ۲۰۲۴؛ مینتاریا و همکاران[4]، ۲۰۲۳؛ سونکاوده و همکاران[5]، ۲۰۲۳). مرورهای جدیدتر نیز بر تنوع مجموعه‌داده‌ها، اهداف پیش‌بینی، معیارهای ارزیابی و طراحی‌های تجربی در این ادبیات تأکید دارند؛ تنوعی که مقایسه مستقیم نتایج مطالعات را دشوار می‌سازد .شواهد تجربی در قیمت‌گذاری دارایی‌ها نیز نشان می‌دهد که روش‌های یادگیری ماشین می‌توانند برهم‌کنش‌های غیرخطی میان پیش‌بین‌ها را استخراج کنند؛ هرچند اثربخشی آن‌ها به اطلاعات در دسترس، رویه مدل‌سازی و چارچوب ارزیابی وابسته است(گیلیو و همکاران[6]، ۲۰۲۲؛ گو و همکاران[7]، ۲۰۲۰). این شواهد، بررسی مستمر مدل‌های انعطاف‌پذیر را توجیه می‌کند و هم‌زمان ضرورت مقایسه‌های تجربی شفاف و بازتولیدپذیر را یادآور می‌شود.

رگرسیون بردار پشتیبان (SVR) گزینه‌ای مناسب برای مسائل رگرسیون غیرخطی است که در آن‌ها تعمیم‌پذیری پیش‌بینی اهمیت محوری دارد. SVR که در چارچوب یادگیری آماری توسعه

1 Fama
2 Kumbure et al
3 Lin& Marques
4 Mintarya et al
5 Sonkavde et al
6 Giglio et al
7 Gu et al

[Downloaded from mieaoi.ir on 2026-11-19]

یافته است، تابعی را برآورد می‌کند که پیچیدگی مدل را در برابر انحراف از پاسخ‌های مشاهده‌شده، با استفاده از تابع زیانِ نامحسوس نسبت به ε، متوازن می‌سازد(دراکر و همکاران[1]، ۱۹۹۷؛ اسمولا و شولکوپف[2]، ۲۰۰۴). استفاده از هسته تابع پایه شعاعی (RBF) امکان بازنمایی روابط غیرخطی میان مشاهدات تاریخی و مقادیر آینده را فراهم می‌کند، بی‌آنکه شکل تابعی پارامتری از پیش تعیین شود. عملکرد پیش‌بینی مدل RBF-SVR به‌ویژه به پارامتر منظم‌سازی C، ضریب هسته γ و پهنای ناحیه نامحسوس ε وابسته است. پارامتر C موازنه میان تخت‌بودن مدل و جریمه انحراف‌های خارج از ناحیه نامحسوس را کنترل می‌کند؛ γ میزان محلی‌بودن مؤثر هسته را تعیین می‌کند و ε پهنای ناحیه‌ای را مشخص می‌سازد که انحراف‌های درون آن در تابع زیان نامحسوس نسبت به ε وارد نمی‌شوند. ترکیب نامناسب این پارامترها می‌تواند به برازش ناکافی یا حساسیت بیش‌ازحد به مشاهدات آموزشی بینجامد. بنابراین، انتخاب ابرپارامترها بخشی اساسی از توسعه مدل SVR است، نه صرفاً مقدمه‌ای محاسباتی.

انتخاب ابرپارامترهای SVR را می‌توان به‌صورت مسئله‌ای بهینه‌سازی صورت‌بندی کرد که در آن هر بردار پارامتر نامزد بر اساس تابع زیان پیش‌بینی ازپیش‌تعیین‌شده ارزیابی می‌شود. جست‌وجوی شبکه‌ای متعارف روشی نظام‌مند برای پیمایش فضای متناهی پارامترهاست، اما با افزایش تعداد پارامترها و مقادیر نامزد، هزینه محاسباتی آن به‌سرعت افزایش می‌یابد. الگوریتم‌های فراابتکاری مبتنی بر جمعیت، با تولید و به‌روزرسانی تکراری راه‌حل‌های نامزد بر اساس سازوکارهای جست‌وجوی اختصاصی، جایگزینی برای این روش فراهم می‌کنند. الگوریتم ژنتیک (GA) از عملگرهای تکاملی مانند انتخاب، تقاطع و جهش استفاده می‌کند(گلدبرگ[3]،۱۹۸۹). بهینه‌سازی ازدحام ذرات و دیگر روش‌های مبتنی بر جمعیت نیز راه‌حل‌های نامزد را با بهره‌گیری از اطلاعات حاصل از جست‌وجو به‌روزرسانی می‌کنند(کندی و ابرهارت[4]، ۱۹۹۵). این الگوریتم‌ها برای مسائل بهینه‌سازی غیرخطی و مشتق‌ناپذیر قابل‌استفاده‌اند، اما عملکردشان به فضای جست‌وجو، پیکربندی الگوریتم، بودجه محاسباتی و تابع هدف بستگی دارد. افزون بر این، قضیه‌های «عدم وجود ناهار رایگان» در بهینه‌سازی هشدار می‌دهند که نباید انتظار داشت یک الگوریتم در همه رده‌های مسئله بر روش‌های دیگر برتری یابد(ولپرت و مک‌ریدی[5]، ۱۹۹۷). بنابراین، در کاربردهای پیش‌بینی مالی، روش‌های فراابتکاری باید با معیارهای مبنای مناسب و در شرایط آزمایشی هم‌تراز ارزیابی شوند.

مطالعات پیشین، مدل‌های یادگیری ماشین را در کنار رویه‌های بهینه‌سازی برای پیش‌بینی بازار سهام بررسی کرده‌اند.مرورهای ادبیات، استفاده از ماشین‌های بردار پشتیبان و دیگر الگوریتم‌های یادگیری را همراه با انتخاب ویژگی، محاسبات تکاملی و بهینه‌سازی مبتنی بر ازدحام مستند کرده‌اند(کومبوره و همکاران[6]، ۲۰۲۲؛ لین و مارکز، ۲۰۲۴؛ مینتاریا و همکاران[7]، ۲۰۲۳؛ سونکاوده و همکاران[8]، ۲۰۲۳). پژوهش‌های داخلی، تأثیر نماگرهای بازار

1 Drucker et al
2 Smola& Schölkopf
3 Goldberg
4 Kennedy& Eberhart
5 Wolpert & Macready
6 Kumbure et al
7 Mintarya et al
8 Sonkavde et al

[Downloaded from mieaoi.ir on 2026-11-19]

سرمایه بر ارزش و قیمت سهام را با استفاده از داده‌های پنلی بررسی کرده‌اند و نیز کاربرد الگوریتم‌های یادگیری عمیق را در پیش‌بینی قیمت سهام مطالعه کرده‌اند (اشرفی سلطان‌احمدی و همکاران، ۱۴۰۴؛ سعیدی‌اقدم و همکاران، ۱۴۰۱). این مطالعات زمینه‌ای روش‌شناختی برای بررسی این پرسش فراهم می‌کنند که آیا انتخاب مدل مبتنی بر بهینه‌سازی می‌تواند عملکرد پیش‌بینی را بهبود دهد یا خیر. بااین‌حال، نمی‌توان بهبودهای گزارش‌شده را خودبه‌خود به الگوریتم بهینه‌سازی نسبت داد؛ زیرا تفاوت در ساخت پیش‌بین‌ها، افراز داده‌ها، دامنه ابرپارامترها، معیارهای ارزیابی و بودجه محاسباتی نیز ممکن است بر نتایج اثر بگذارد. مقایسه معنادار باید تا حد امکان، وظیفه پیش‌بینی، پیش‌پردازش، افراز داده‌ها و معیارهای ارزیابی را ثابت نگه دارد. همچنین، باید بهبود دقت عددی پیش‌بینی را از بهبود پیش‌بینی جهت حرکت تفکیک کرد؛ زیرا این دو هدف یکسان نیستند.

الگوریتم بهینه‌سازی جست‌وجوی طلایی (GSO) که نوروزی و همکاران (۲۰۲۲) معرفی کرده‌اند، روشی مبتنی بر جمعیت برای بهینه‌سازی مسائل عددی است. این الگوریتم با استفاده از به‌روزرسانی راه‌حل‌های نامزد و سازوکار اندازه گام تطبیقی، فضای جست‌وجو را می‌کاود و در تکرارهای متوالی راه‌حل‌های نامزد را پالایش می‌کند (نوروزی و همکاران، ۲۰۲۲). این سازوکار مبنایی روش‌شناختی برای بررسی GSO به‌عنوان جایگزینی در انتخاب ابرپارامترهای SVR فراهم می‌آورد. بااین‌حال، عملکرد عمومی گزارش‌شده برای GSO اثبات نمی‌کند که این الگوریتم لزوماً پیش‌بینی‌های مالی دقیق‌تری تولید خواهد کرد. اثربخشی آن در پیش‌بینی قیمت سهام باید با تابع هدف مناسب و ارزیابی کنترل‌شده خارج از نمونه سنجیده شود. به‌ویژه، مقایسه با بهینه‌سازی مبتنی بر الگوریتم ژنتیک می‌تواند مشخص کند که آیا عملکرد GSO–SVR، تحت شرایطی قابل‌مقایسه، با عملکرد یک روش جست‌وجوی مبتنی بر جمعیت دیگر تفاوت دارد یا خیر.

ارزیابی قابل‌اتکا هنگامی اهمیت ویژه‌ای دارد که مدل پیش‌بینی از طریق بهینه‌سازی مکرر بر اساس مشاهدات تاریخی انتخاب شود. اعتبارسنجی متقاطع سری زمانی و ارزیابی مبدأ غلتان از روش‌های شناخته‌شده ارزیابی پیش‌بینی‌اند که ترتیب زمانی را در فرایند ارزیابی حفظ می‌کنند و اطمینان می‌دهند هر پیش‌بینی فقط با استفاده از اطلاعات موجود پیش از مشاهده مقدار هدف ساخته شود(برگمیر و همکاران[1]، ۲۰۱۸؛ هایندمن و آتاناسوپولوس[2]، ۲۰۲۱؛ تاشمن[3]، ۲۰۰۰). طرح نگهداشت زمانی نیز زمانی مفید است که یک دوره آزمون نهایی در سراسر فرایند توسعه مدل دست‌نخورده باقی بماند. بااین‌حال، انتخاب مکرر مدل‌ها یا پیکربندی‌ها بر اساس عملکرد تاریخی می‌تواند سوگیری انتخاب و بیش‌برازش ایجاد کند؛ حتی اگر معیار عملکرد نهایی بر نمونه‌ای جداگانه محاسبه شود(بیلی و همکاران[4]، ۲۰۱۷؛ وایت[5]، ۲۰۰۰). افزون بر این، تفاوت عددی میان پیش‌بینی‌های رقیب به‌تنهایی نشان‌دهنده تفاوت معنادار آماری در دقت پیش‌بینی نیست. برای مقایسه دقت پیش‌بینی می‌توان، با رعایت مفروضات مربوط، از آزمون‌هایی مانند

1 Bergmeir et al
2 Hyndman& Athanasopoulos
3 Tashman
4 Bailey et al
5 White

[Downloaded from mieaoi.ir on 2026-11-19]

آزمون دیبولد–ماریانو استفاده کرد(دیبولد[1] و ماریانو[2]، ۱۹۹۵) همچنین، در مقایسه چند مدل، آزمون توان پیش‌بینی برتر می‌تواند یکی از ابزارهای آماری قابل‌بررسی باشد(هانسن[3]، ۲۰۰۵).این ملاحظات، طراحی آزمایشی‌ای را ایجاب می‌کنند که برازش مدل، انتخاب ابرپارامتر و ارزیابی نهایی را از یکدیگر جدا سازد و شواهد به‌دست‌آمده را با احتیاط تفسیر کند.

بر این اساس، پژوهش حاضر چارچوب ترکیبی بهینه‌سازی جست‌وجوی طلایی–رگرسیون بردار پشتیبان (GSO–SVR) را برای پیش‌بینی یک‌گام‌به‌جلوی قیمت پایانی سهام بررسی می‌کند. مدل از هسته RBF استفاده می‌کند و GSO با کمینه‌سازی ریشه میانگین مربعات خطا (RMSE) در زیرمجموعه اعتبارسنجی مرتب‌شده بر اساس زمان، مقادیر مناسب پارامتر منظم‌سازی C و ضریب هسته γ را جست‌وجو می‌کند. پارامتر ε ثابت نگه داشته می‌شود تا اثر بهینه‌سازی این دو ابرپارامتر به‌طور جداگانه بررسی شود. روش پیشنهادی با SVR متعارف و SVR بهینه‌شده با الگوریتم ژنتیک (GA–SVR) مقایسه می‌شود. داده‌های مورد استفاده، قیمت‌های پایانی روزانه سهام هشت شرکت اپل، گوگل، فیسبوک، IBM، مایکروسافت، توییتر، یاهو و فورد را در بر می‌گیرند. در طراحی آزمایش، افراز زمانی داده‌ها، وظیفه پیش‌بینی و رویه ارزیابی برای هر سه مدل یکسان است. عملکرد مدل‌ها با معیارهای MAE، NMSE و DS ارزیابی می‌شود که اطلاعات مکملی درباره خطاهای عددی و میزان تطابق جهت تغییرات قیمت ارائه می‌کنند.

سهم پژوهش حاضر، ارزیابی تجربی GSO به‌عنوان رویه‌ای برای انتخاب ابرپارامتر در پیش‌بینی قیمت سهام مبتنی بر SVR است. به‌طور مشخص، GSO–SVR با SVR متعارف و GA–SVR در چارچوبی مشترک مقایسه می‌شود، عملکرد مدل‌ها در چند سهم بررسی می‌شود و معیارهای خطامحور و جهت‌محور در نظر گرفته می‌شوند. هدف، احراز وجود مزیت پیش‌بینی قابل‌اندازه‌گیری در شرایط آزمایش حاضر است، نه فرض‌کردن برتری همگانی GSO. یافته‌ها باید با توجه به شرکت‌های منتخب، دوره مشاهده، افق پیش‌بینی، تنظیمات جست‌وجوی پارامتر و طراحی مجموعه آزمون تفسیر شوند. تصریح این شرایط می‌تواند به مقایسه‌ای بازتولیدپذیر از روش‌های فراابتکاریِ انتخاب ابرپارامتر در پیش‌بینی سری‌های زمانی مالی کمک کند.

۱. مدل پیشنهادیGSO–SVR

الف) الگوریتم بهینه‌سازی جست‌وجوی طلایی

الگوریتم بهینه‌سازی جست‌وجوی طلایی (GSO) فراابتکاری مبتنی بر جمعیت است که نوروزی و همکاران آن را معرفی کردند(نوروزی، ۲۰۲۲). این الگوریتم فضای جست‌وجوی کران‌دار را با به‌روزرسانی راه‌حل‌های نامزد از طریق سازوکار اندازه گام کاوش می‌کند؛ سازوکاری که اطلاعات بهترین راه‌حل‌های فردی و بهترین راه‌حل کلی جمعیت را ترکیب می‌کند. در این پژوهش، GSO برای بهینه‌سازی ابرپارامترهای مدل رگرسیون بردار پشتیبان (SVR) در پیش‌بینی قیمت پایانی سهام در روز بعد به کار می‌رود.

1 Diebold
2 Mariano
3 Hansen

[Downloaded from mieaoi.ir on 2026-11-19]

فرض کنیم N اندازه جمعیت و D بُعد مسئله بهینه‌سازی باشد .موقعیت-i امین راه‌حل نامزد در تکرار t با بردار زیر نمایش داده می‌شود:

$$O_i(t) = [O_{i1},(t), O_{i2},(t) \dots O_{iD}(t)].$$

جمعیت اولیه به‌صورت تصادفی در کران‌های ازپیش‌تعیین‌شده جست‌وجو تولید می‌شود .برای هر نامزد i و هر بُعدd ، موقعیت اولیه به صورت رابطه (۱) تعریف می‌شود:

$$O_{id}(0) = lb_d + r_{id}(ub_d - lb_d), \quad (1)$$

که در آن lb_d و ub_d به‌ترتیب کران‌های پایین و بالای بُعد d و $r_{id} \sim U(0,1)$ عددی تصادفی و مستقل است .سپس برازندگی هر نامزد با تابع هدف تعریف‌شده برای مسئله سنجیده می‌شود.
بهترین موقعیتی که نامزد i تا تکرار t یافته است با $O_{i,best}(t)$ و بهترین موقعیت شناسایی‌شده توسط کل جمعیت با $O_{gbest}(t)$ نمایش داده می‌شود .این موقعیت‌ها هرگاه نامزدی بر اساس تابع هدف کمینه‌سازی مقدار برازندگی بهتری به‌دست آورد، به‌روزرسانی می‌شوند.
حرکت هر نامزد با بردار اندازه گام $S_i(t)$ تعیین می‌شود .بر اساس صورت‌بندی نوروزی و همکاران(۲۰۲۲) به‌روزرسانی اندازه گام به صورت رابطه(۲) است:

$$\begin{aligned} S_i(t+1) = {} & T(t)S_i(t) \\ & + C_1 \cos(r_1)\left[O_{i,\text{best}}(t) - O_i(t)\right] \\ & + C_2 \sin(r_2)\left[O_{\text{gbest}}(t) - O_i(t)\right], \end{aligned} \quad (2)$$

که در آن T(t) عملگر انتقال، C_1 و C_2 ضرایب تصادفی با توزیع یکنواخت در بازه [0,2] و r_1 و r_2 مقادیر تصادفی با توزیع یکنواخت در بازه (0,1) هستند .عملگر انتقال به صورت رابطه (۳) تعریف می‌شود:

$$T(t) = 100 \exp\left(-20\frac{t}{t_{\max}}\right), \quad (3)$$

که t_{max} حداکثر تعداد تکرار است .این عملگر با پیشرفت جست‌وجو کاهش می‌یابد و نقش آن تنظیم سهم اندازه گام قبلی در حرکت بعدی است، در حالی که جملات بهترین فردی و بهترین کلی، راه‌حل‌های نامزد را به سوی نواحی امیدبخش فضای جست‌وجو هدایت می‌کنند.
برای محدودکردن بزرگی هر حرکت، بردار اندازه گام به‌صورت مؤلفه‌به‌مؤلفه مقید می‌شود:

$$-S_{d,\max} \le S_{id}(t+1) \le S_{d,\max}, \quad (4)$$

که در آن حداکثر اندازه گام مجاز برابر است با:

$$S_{d,\max} = 0.1(ub_d - lb_d). \quad (5)$$

سپس موقعیت هر نامزد مطابق رابطه (۶) به‌روزرسانی می‌شود:

$$O_i(t+1) = O_i(t) + S_i(t+1). \quad (6)$$

پس از هر به‌روزرسانی، قیدهای فضای جست‌وجو اعمال می‌شوند تا راه‌حل‌های نامزد در کران‌های ازپیش‌تعیین‌شده باقی بمانند .برازندگی هر نامزد به‌روزشده ارزیابی و در صورت مشاهده بهبود، موقعیت‌های بهترین فردی و بهترین کلی بازنگری می‌شوند .ضعیف‌ترین نامزد

[Downloaded from mieaoi.ir on 2026-11-19]

با یک نامزد تصادفی تازه جایگزین می‌شود، مطابق رویه GSO اتخاذشده در مرجع(نوروزی، ۲۰۲۲) الگوریتم با رسیدن به حداکثر تعداد تکرار پایان می‌یابد و بهترین نامزد یافت‌شده در جست‌وجو به‌عنوان راه‌حل بهینه بازگردانده می‌شود .در کاربرد حاضر، اندازه جمعیت = N ۳۰ و حداکثر تکرار = t_{max} ۱۰۰ در نظر گرفته شده‌اند .این تنظیمات بودجه محاسباتی مدل–GSO SVRرا تعیین می‌کند.

ب) بهینه‌سازی ابرپارامترهایSVR

از رگرسیون بردار پشتیبان (SVR) به‌عنوان مدل پیش‌بینی پایه استفاده می‌شود، زیرا متغیر پاسخ (قیمت پایانی روز بعد) پیوسته است SVR .تابعی را برآورد می‌کند که تخت‌بودن مدل را در برابر انحراف از پاسخ‌های مشاهده‌شده و با استفاده از تابع زیان نامحسوس نسبت به ε متوازن می‌سازد(اسمولا[1]، ۲۰۰٤) عملکرد پیش‌بینی آن به انتخاب ابرپارامترها، از جمله پارامتر منظم‌سازیC ، پارامتر هسته γ و نوع تابع هسته، وابسته است.

در این پژوهش از هسته تابع پایه شعاعی (RBF) استفاده می‌شود که به صورت رابطه(۷) تعریف می‌شود:

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2), \quad (7)$$

که در آن x_i و x بردارهای ورودی و $\gamma > 0$ پهنای مؤثر هسته را کنترل می‌کند .پارامتر منظم‌سازی $C > 0$ جریمه اعمال‌شده بر انحراف‌های خارج از ناحیه نامحسوس ε را کنترل می‌کند .مقدار بزرگ C بر کاهش خطاهای آموزشی تأکید بیشتری می‌گذارد، حال آن‌که مقدار کوچک‌تر رواداری بیشتری نسبت به این خطاها فراهم می‌کند.

در مدل پیشنهادیGSO_{SVR} ، الگوریتم GSO ترکیب مؤثری از C و γ را جست‌وجو می‌کند؛ ازاین‌رو هر راه‌حل نامزد با بردار دوبُعدی رابطه (۸) نمایش داده می‌شود:

$$O_i(t) = \begin{bmatrix} C_i(t) \\ \gamma_i(t) \end{bmatrix}. \quad (8)$$

جست‌وجو در کران‌های زیر انجام می‌شود:

$$0.1 \le C \le 100,\ 0.01 \le \gamma \le 10. \quad (9)$$

پارامتر ε برای هر سه مدل مقایسه‌ای برابر ۰,۰۱ ثابت است، به شرط آن‌که این مقدار در پیاده‌سازی به‌طور یکسان به کار رود .هسته RBF نیز ثابت نگه داشته می‌شود تا رویه بهینه‌سازی به‌طور مشخص بر انتخاب C و γ متمرکز باشد.

۱) تابع برازندگی

تابع برازندگی، خطای پیش‌بینی مدل SVR را در زیرمجموعه اعتبارسنجیِ مرتب‌شده زمانی اندازه می‌گیرد .برای بردار پارامتر نامزد(C, γ) = θ ، RMSE اعتبارسنجی به صورت رابطه (۱۰) تعریف می‌شود:

1 Smola

[Downloaded from mieaoi.ir on 2026-11-19]

$$\mathrm{RMSE}_{\mathrm{val}}(\theta) = \sqrt{\frac{1}{n_{\mathrm{val}}}\sum_{j=1}^{n_{\mathrm{val}}}\left(y_j - \hat{y}_j(\theta)\right)^2}, \quad (10)$$

که در آن n_val تعداد مشاهدات اعتبارسنجی، y قیمت پایانی مشاهده‌شده و $\hat{y}(\theta)$ پیش‌بینی متناظر تولیدشده توسط مدل SVR با ابرپارامترهای θ است .مسئله بهینه‌سازی به صورت رابطه (۱۱) صورت‌بندی می‌شود:

$$\theta^* = \mathrm{RMSE}_{\mathrm{val}}(\theta), \quad (11)$$

که فضای پارامترهای شدنی آن $\Omega = [0.1, 100] \times [0.01, 10]$ است.

برای هر نامزد، مدل SVR با استفاده از زیرمجموعه برازش مدل آموزش می‌بیند و RMSE اعتبارسنجی آن به‌عنوان مقدار برازندگی به کار می‌رود .نامزدی که کمترین RMSE اعتبارسنجی را دارد ترجیح داده می‌شود .مشاهدات اعتبارسنجی صرفاً برای انتخاب مدل استفاده می‌شوند و بخشی از مجموعه آزمون نهایی نیستند.

۲. ادغام GSO و SVR

فرایند بهینه‌سازی GSO–SVR شامل مراحل زیر است:

۱. **مقداردهی اولیه:** ۳۰ راه‌حل کاندید در محدوده‌های پارامتری مشخص‌شده در رابطه (۹) تولید و اندازه‌گام آن‌ها مطابق پیاده‌سازی GSO به‌کاررفته مقداردهی اولیه می‌شود.

۲. **ارزیابی برازش:** برای هر کاندید، یک مدل SVR با استفاده از زیرمجموعه برازش مدل آموزش داده می‌شود و RMSE آن روی زیرمجموعه اعتبارسنجیِ مرتب‌شده بر اساس زمان محاسبه می‌گردد.

۳. **به‌روزرسانی بهترین راه‌حل‌ها:** هرگاه RMSE اعتبارسنجی کاهش یابد، بهترین موقعیت فردی هر کاندید و بهترین موقعیت کل جمعیت به‌روزرسانی می‌شوند.

٤. **محاسبه اندازه گام:** بردارهای اندازه گام با استفاده از رابطه (۲) محاسبه می‌شوند و قیود حرکتیِ روابط (٤) و (٥) بر آن‌ها اعمال می‌گردند.

٥. **به‌روزرسانی موقعیت:** موقعیت کاندیدها با استفاده از رابطه (٦) به‌روزرسانی و محدوده‌های پارامتری رابطه (۹) بر آن‌ها اعمال می‌شود.

٦. **جایگزینی در جمعیت:** مطابق روش GSO به‌کاررفته، کاندیدی که عملکرد ضعیف‌تری دارد با یک کاندید تصادفی جدید جایگزین و جمعیت حاصل ارزیابی می‌شود.

۷. **پایان فرایند:** مراحل ارزیابی و به‌روزرسانی تا تکمیل ۱۰۰ تکرار ادامه می‌یابند. سپس بهترین بردار پارامتریِ مجاز یافت‌شده در فرایند جست‌وجو بازگردانده می‌شود.

۸. **برازش نهایی مدل:** پیکربندی منتخب SVR با استفاده از کل داده‌های دوره توسعه شامل مشاهداتی که برای اعتبارسنجی به‌کار رفته‌اند دوباره برازش می‌شود و عملکرد مدل روی مجموعه آزمونِ زمانیِ دست‌نخورده ارزیابی می‌گردد.

پروتکل اعتبارسنجی برای هر شرکت به‌طور جداگانه اجرا می‌شود. از مجموعه آزمون برای ارزیابی برازش کاندیدها یا انتخاب ابرپارامترها استفاده نمی‌شود. افزون بر این، در مرحله

[Downloaded from mieaoi.ir on 2026-11-19]

اعتبارسنجی، تمام تبدیل‌های پیش‌پردازش فقط با استفاده از داده‌های مربوط به برازش مدل برآورد می‌شوند و سپس، بدون برازش مجدد، بر مشاهدات اعتبارسنجی اعمال می‌گردند. پس از انتخاب مدل، تبدیل نهایی پیش‌پردازش با استفاده از کل دوره توسعه برآورد و بر مجموعه‌آزمون اعمال می‌شود.

عملکرد GSO–SVR با عملکرد SVR متعارف و GA–SVR، تحت شرایط یکسان از نظر افراز داده‌ها، افق پیش‌بینی، ساخت پنجره ورودی، نوع هسته، مقدار ثابت و معیارهای ارزیابی مقایسه می‌شود. مدل‌های GA–SVR و GSO–SVR از محدوده‌های جست‌وجوی یکسان و تابع هدف مبتنی بر RMSE اعتبارسنجی استفاده می‌کنند. این طراحی با کاهش تفاوت‌های ناشی از ناسازگاری در پردازش داده‌ها یا رویه‌های ارزیابی، تفسیر مقایسه را شفاف‌تر می‌کند.

چارچوب پیشنهادی با هدف بررسی تجربی این موضوع ارائه شده است که آیا بهینه‌سازی ابرپارامترها با GSO، پیش‌بینی برون‌نمونه‌ای قیمت سهام را بهبود می‌دهد یا خیر. بنابراین، هرگونه ادعا درباره دقت برتر، استحکام بیشتر یا معناداری آماری، بر نتایج آزمایشی گزارش‌شده در بخش III استوار است و از پیش مفروض گرفته نمی‌شود.

۲ .داده‌ها، طرح تجربی و ارزیابی عملکرد

الف) مجموعه‌داده و طراحی تجربی

در این پژوهش، رویکرد ترکیبی ماشین بردار پشتیبان–بهینه‌سازی جست‌وجوی طلایی (GSO–SVM) برای پیش‌بینی قیمت سهام در روز معاملاتی بعد ارزیابی می‌شود .رویکرد پیشنهادی با مدل متعارف ماشین بردار پشتیبان (SVM) و مدل ماشین بردار پشتیبان بهینه‌شده با الگوریتم ژنتیک (GA–SVM) مقایسه می‌شود .هدف آن است که مشخص شود بهینه‌سازی ابرپارامترهای SVMباGSO ، تحت یک پروتکل تجربی هم‌تراز، عملکرد پیش‌بینی خارج از نمونه را بهبود می‌دهد یا خیر.

تحلیل تجربی هشت شرکت سهامی عام شامل اپل، گوگل، فیس‌بوک،IBM ، مایکروسافت، توییتر، یاهو و فورد را در بر می‌گیرد .مشاهدات قیمت پایانی روزانه برای یک دوره ۱۸۰روزه گردآوری شده است .پس از حذف رکوردهای مفقود یا نامعتبر، مجموعه‌داده نهایی برای هر شرکت شامل مشاهدات روزهای معاملاتی معتبر است .از آنجا که بازارهای سهام در همه روزهای تقویمی باز نیستند، اندازه نمونه بر حسب جلسات معاملاتی مشاهده‌شده تعریف می‌شود، نه روزهای تقویمی.

برای تضمین قابلیت مقایسه میان شرکت‌ها، یک رویه پردازش داده، افق پیش‌بینی، طول پنجره ورودی، پروتکل آموزش و آزمون و معیارهای عملکرد یکسان برای همه مدل‌ها به کار می‌رود. مجموعه‌داده پیش از هر برازش مدل یا بهینه‌سازی ابرپارامتر به‌صورت زمانی مرتب می‌شود.

ب) ساخت پیش‌بین‌ها و متغیر هدف

وظیفه پیش‌بینی به‌صورت مسئله پیش‌بینی یک‌گام‌به‌جلوی صورت‌بندی می‌شود .فرض کنیم P قیمت پایانی مشاهده‌شده در زمان معاملاتی t باشد .برای طول پنجره ورودی ثابت پنج مشاهده، بردار ویژگی به صورت زیر تعریف می‌شود:

$$x_t = (P_{t-5}, P_{t-4}, P_{t-3}, P_{t-2}, P_{t-1}),$$

و متغیر هدف، قیمت پایانی مشاهده‌شده بعدی است:

$$y_t = P_t.$$

[Downloaded from mieaoi.ir on 2026-11-19]

بر این اساس، هر پیش‌بینی تنها از قیمت‌های پایانی استفاده می‌کند که پیش از جلسه معاملاتی هدف در دسترس بوده‌اند و هیچ مشاهده آینده‌ای در بردار پیش‌بین‌ها وارد نمی‌شود.
برای هر شرکت، مشاهدات حاصل به ترتیب زمانی مرتب می‌شوند .هر پارامتر تبدیل یا پیش‌پردازش صرفاً از داده‌های آموزشی برآورد و سپس بدون برازش مجدد بر داده‌های اعتبارسنجی و آزمون اعمال می‌شود .به‌طور خاص، هنگام استفاده از نرمال‌سازی کمینه–بیشینه، مقادیر کمینه و بیشینه تنها از مشاهدات آموزشی محاسبه می‌شوند:

$$P_t^* = \frac{P_t - P_{\min,\text{train}}}{P_{\max,\text{train}} - P_{\min,\text{train}}}.$$

همین تبدیل برگرفته از داده آموزشی بر مشاهدات اعتبارسنجی و آزمون اعمال می‌شود .اگر مقداری خارج از دامنه آموزشی قرار گیرد، با استفاده از کمینه و بیشینه اصلی حاصل از داده‌های آموزشی تبدیل می‌شود و پارامترهای مقیاس‌بندی بازمحاسبه نمی‌گردند .مقادیر پیش‌بینی‌شده پیش از محاسبه معیارهای ارزیابی مبتنی بر قیمت، به مقیاس اصلی قیمت بازگردانده می‌شوند.

پ)پروتکل آموزش، اعتبارسنجی و آزمون

برای حفظ ساختار زمانی داده‌های سری زمانی مالی، مشاهدات به‌صورت زمانی و نه تصادفی افراز می‌شوند .۷۵ درصد اول مشاهدات معتبر به دوره توسعه اختصاص می‌یابد و ۲۵ درصد نهایی برای دوره آزمون نگه داشته می‌شود .اگر N تعداد مشاهدات معتبر باشد، افراز بر اساس شاخص زمانی زیر تعیین می‌شود:

$$N_{\text{train}} = \lfloor 0.75N \rfloor,\ N_{\text{test}} = N - N_{\text{train}}.$$

دوره توسعه نیز به‌صورت زمانی به زیرمجموعه برازش مدل و زیرمجموعه اعتبارسنجی تقسیم می‌شود .زیرمجموعه اعتبارسنجی شامل ۱۰ درصد آخر مشاهدات دوره توسعه و مشاهدات پیشین برای برازش مدل‌های نامزد به کار می‌رود.
زیرمجموعه اعتبارسنجی برای انتخاب ابرپارامترهای مدل‌های مبتنی بر SVM استفاده می‌شود . مجموعه آزمون نهایی برای انتخاب ابرپارامتر، تعیین شرایط توقف، انتخاب مدل یا اصلاح طرح تجربی به کار نمی‌رود .پس از انتخاب ابرپارامترها، پیکربندی منتخب با داده‌های در دسترس دوره توسعه برازش مجدد می‌شود و عملکرد آن در دوره آزمون دست‌نخورده ارزیابی می‌گردد.
این پروتکل با هدف کاهش نشت اطلاعات و ارائه ارزیابی خارج از نمونه‌ای که ماهیت زمانی پیش‌بینی قیمت سهام را بازتاب دهد، طراحی شده است .هر سه مدل با همان مشاهدات آزمون هر شرکت ارزیابی می‌شوند.

ت) مدل ماشین بردار پشتیبان

رگرسیون بردار پشتیبان (SVR) به‌عنوان مدل پیش‌بینی پایه به کار می‌رود .با توجه به یک نمونه آموزشی{n}^{i=1}_{(x_i, y_i)} ، مدل SVR تابعی به صورت زیر برآورد می‌کند:

$$f(x) = w^T \phi(x) + b,$$

که در آن φ(·) نگاشت ویژگی، w بردار ضرایب و b عرض از مبدأ است .برای استخراج روابط بالقوه غیرخطی میان قیمت‌های پایانی تاریخی و متغیر هدف از هسته تابع پایه شعاعی (RBF) استفاده می‌شود:

[Downloaded from mieaoi.ir on 2026-11-19]

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2),$$

که γ پهنای هسته را کنترل می‌کند .پارامتر منظم‌سازی C موازنه میان پیچیدگی مدل و جریمه خطاهای پیش‌بینی را تعیین می‌کند و ε پهنای ناحیه زیان نامحسوس در SVR را مشخص می‌سازد.

مدل متعارف SVM به‌عنوان مبنای مقایسه به کار می‌رود .ابرپارامترهای آن با همان پروتکل اعتبارسنجی رویکردهای فراابتکاری انتخاب می‌شود تا تفاوت‌های عملکردی صرفاً به رویه‌های تنظیم نابرابر نسبت داده نشود.

ث)رویه‌های بهینه‌سازی GA–SVM وGSO–SVM

GA–SVM(۱

از الگوریتم ژنتیک برای تعیین مقادیر مناسب ابرپارامترهای SVM استفاده می‌شود .هر راه‌حل نامزد ترکیبی از C و γ را نشان می‌دهد .ارزیابی نامزد با برازش مدل SVR روی زیرمجموعه برازش مدل و محاسبه ریشه میانگین مربعات خطا (RMSE) آن در زیرمجموعه اعتبارسنجیِ زمانی انجام می‌شود.

پیکربندی الگوریتم ژنتیک به شرح زیر است :اندازه جمعیت ۵۰، حداکثر ۱۰۰ نسل، احتمال تقاطع ۰,۸۰ و احتمال جهش ۰,۱۰ .دامنه‌های جست‌وجو عبارت‌اند از:

$$C \in [0.1, 100],\ \gamma \in [0.01, 10].$$

تابع برازندگی به صورت زیر تعریف می‌شود و هدف، کمینه‌سازی RMSE اعتبارسنجی است:

$$\text{Fitness}(\theta) = \text{RMSE}_{\text{val}}(\theta),\ \theta = (C, \gamma),$$

GSO–SVM(۲

رویکرد ترکیبی پیشنهادی از الگوریتم بهینه‌سازی جست‌وجوی طلایی (GSO) برای جست‌وجوی ابرپارامترهای مناسب SVM بهره می‌گیرد..هر شیء نشان‌دهنده یک بردار پارامتر نامزد است:

$$z_i = (C_i, \gamma_i).$$

تابع هدف، RMSE اعتبارسنجی مدل SVR متناظر با نامزد است .در رویهGSO ، اشیا موقعیت خود را با عملگر اندازه گام تطبیقی St_i به‌روزرسانی می‌کنند که شامل عملگر انتقالT ، موقعیت‌های بهترین فردی و موقعیت بهترین کلی است .به‌روزرسانی اندازه گام به صورت زیر است:

$$St_i(t+1) = T \cdot St_i(t) + C_1 \cdot \cos(r_1) \cdot \left(Obest_i - x_i(t)\right) + C_2 \cdot \sin(r_2) \cdot \left(Ogbest_i - x_i(t)\right),$$

که در آن C_1 و C_2 ضرایب تصادفی در بازه [0,2] و $r_1, r_2 \in (0,1)$ هستند .عملگر انتقال چنین تعریف می‌شود:

$$T = 100 \times \exp\left(-20 \times \frac{t}{t_{\max}}\right).$$

اندازه گام نیز به‌صورت زیر مقید می‌شود:

$$-St_{i_{\max}} \le St_i \le St_{i_{\max}},\ St_{i_{\max}} = 0.1 \times (ub_i - lb_i).$$

[Downloaded from mieaoi.ir on 2026-11-19]

پیکربندی GSO شامل جمعیتی ۳۰ عضوی و حداکثر ۱۰۰ تکرار است .ابرپارامترهای منتخب همان مقادیر مرتبط با کمترین RMSE اعتبارسنجی در طول جست‌وجو هستند .راه‌حل‌های نامزد به دامنه‌های پارامتری ازپیش‌تعیین‌شده محدود می‌شوند .برای هر دو روش GA–SVM و GSO–SVM، رویه ارزیابی نامزدها، مشاهدات اعتبارسنجی، پیاده‌سازی SVM و کران‌های ابرپارامتر یکسان نگه داشته شده و مقدار ε برای هر سه مدل برابر ۰,۰۱ ثابت است.

ج (معیارهای ارزیابی عملکرد

عملکرد پیش‌بینی با میانگین قدرمطلق خطا(MAE) ، خطای مربعات میانگین نرمال‌شده (NMSE)و تقارن جهت‌گیری (DS) ارزیابی می‌شود .همه معیارها بر مجموعه آزمون و با پیش‌بینی‌های بازگردانده‌شده به مقیاس اصلی قیمت پایانی، در موارد لزوم، محاسبه می‌شوند.

۱) میانگین قدرمطلق خطا

MAEمیانگین قدرمطلق تفاوت میان قیمت‌های پایانی مشاهده‌شده و پیش‌بینی‌شده را می‌سنجد:

$$\mathrm{MAE} = \frac{1}{n_{\mathrm{test}}} \sum_{t=1}^{n_{\mathrm{test}}} | y_t - \hat{y}_t |.$$

مقادیر کمتر MAE بیانگر خطاهای مطلق پیش‌بینی کمتر در میانگین است.

۲)خطای مربعات میانگین نرمال‌شده

NMSEخطاهای مربعی پیش‌بینی را نسبت به یک جمله نرمال‌سازی مشخص می‌سنجد .در این پژوهش چنین تعریف می‌شود:

$$\mathrm{NMSE} = \frac{\sum_{t=1}^{n_{\mathrm{test}}} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{n_{\mathrm{test}}} (y_t - \bar{y}_{\mathrm{test}})^2},$$

که در آن $\bar{y}$_test میانگین مقادیر هدف مشاهده‌شده در مجموعه آزمون است .مقادیر کمتر نشان‌دهنده عملکرد بهتر است.

۳)تقارن جهت‌گیری

تقارن جهت‌گیری درصد مشاهداتی را می‌سنجد که در آن‌ها جهت تغییر پیش‌بینی‌شده و جهت تغییر واقعی یکسان است:

$$\mathrm{DS} = \frac{100}{n_{\mathrm{test}}} \sum_{t=1}^{n_{\mathrm{test}}} I\left[(\hat{y}_t - y_{t-1})(y_t - y_{t-1}) > 0\right],$$

که در آن I[·] تابع نشانگر است که در برقراری شرط مقدار یک و در غیر این صورت صفر می‌گیرد .مقدار بالاتر DS نشان‌دهنده تطابق بیشتر میان جهت تغییرات پیش‌بینی‌شده و مشاهده‌شده قیمت است.

[Downloaded from mieaoi.ir on 2026-11-19]

چ) نتایج تجربی

جدول ۱ عملکرد خارج از نمونهSVM ، GA–SVM و GSO–SVM را برای هشت شرکت نشان می‌دهد .مقادیر کمتر NMSE و MAE نشان‌دهنده دقت عددی بهتر و مقدار بالاتر DS نشان‌دهنده تطابق جهت‌گیری بهتر است.

جدول ۱ .عملکرد پیش‌بینی خارج از نمونهSVM ، GA–SVM وGSO–SVM

شرکت	مدل	NMSE	MAE	DS (٪)
اپل	SVM	0.873	18.42	56.2
اپل	GA–SVM	0.781	14.86	58.7
اپل	GSO–SVM	**0.652**	**9.43**	**62.1**
گوگل	SVM	0.805	35.35	75.0
گوگل	GA–SVM	0.379	7.41	83.3
گوگل	GSO–SVM	**0.312**	**5.28**	**86.7**
بوک‌فیس	SVM	0.824	21.24	45.8
بوک‌فیس	GA–SVM	0.512	12.37	79.2
بوک‌فیس	GSO–SVM	**0.398**	**8.15**	**87.5**
IBM	SVM	0.507	12.59	66.7
IBM	GA–SVM	0.318	9.62	83.3
IBM	GSO–SVM	**0.254**	**6.85**	**91.7**
مایکروسافت	SVM	0.829	9.01	75.0
مایکروسافت	GA–SVM	0.346	0.278	83.3
مایکروسافت	GSO–SVM	**0.287**	**0.194**	**88.5**
توییتر	SVM	0.336	15.29	41.7
توییتر	GA–SVM	0.152	8.92	83.3
توییتر	GSO–SVM	**0.118**	**5.76**	**87.5**
یاهو	SVM	0.812	92.45	50.0
یاهو	GA–SVM	0.598	85.61	75.0
یاهو	GSO–SVM	**0.472**	**72.34**	**79.2**
فورد	SVM	0.800	18.64	66.7
فورد	GA–SVM	0.636	5.21	25.8
فورد	GSO–SVM	**0.521**	**3.87**	**58.3**

منبع: یافته های پژوهش

یادداشت :مقادیر کمتر NMSE و MAE بهتر و مقادیر بالاتر DS بهتر است .جدول تنها نتایج مجموعه آزمون را گزارش می‌کند .همه مدل‌ها از دوره‌های آزمون یکسان و تعاریف یکسان NMSE، MAE و DS استفاده کرده‌اند.

[Downloaded from mieaoi.ir on 2026-11-19]

نتایج نشان می‌دهد GSO–SVM در هر هشت شرکت کمترین NMSE و MAE و در هفت شرکت از هشت شرکت بیشترین DS را به‌دست آورده است .برای نمونه در اپل، GSO–SVM مقدار NMSE را از ۰,۸۷۳ (SVM) و ۰,۷۸۱ (GA–SVM) به ۰,۶۵۲ و مقدار MAE را از ۱۸,۴۲ و ۱۴,۸۶ به ۹,۴۳ کاهش داده است .بهبودهای مشابهی برای گوگل، فیس‌بوک، IBM، مایکروسافت، توییتر و یاهو مشاهده می‌شود .در فورد، GSO–SVM کمترین NMSE (۰,۵۲۱) و MAE (۳,۸۷) را ارائه کرده، اما مدل SVM مقدار DS بالاتری (۶۶,۷ درصد) نسبت به GSO–SVM (۵۸,۳ درصد) داشته است .این یافته‌ها نشان می‌دهند بهینه‌سازی ابرپارامتر مبتنی بر GSO عموماً دقت عددی مدل SVR را بهبود می‌دهد و در بیشتر موارد عملکرد جهت‌گیری نیز بهتر می‌شود.

ح) مقایسه آماری و ارزیابی استحکام

تفاوت عددی میان دو مدل به‌تنهایی معناداری آماری را اثبات نمی‌کند .ازاین‌رو نتایج جدول ۱ عمدتاً به‌عنوان مقایسه عملکرد خارج از نمونه تحت یک پروتکل تجربی هم‌تراز تفسیر می‌شوند . نتیجه‌گیری‌ها به شرکت‌ها، دوره نمونه، افق پیش‌بینی و پروتکل ارزیابی بررسی‌شده در این پژوهش محدود است .ادعاهای کلی‌تر درباره تعمیم به دیگر بازارهای مالی یا دوره‌های زمانی نیازمند ارزیابی‌های اضافی خارج از نمونه است.

خ) جمع‌بندی یافته‌های تجربی

نتایج مقایسه‌ای نشان می‌دهند بهینه‌سازی ابرپارامتر مبتنی بر GSO در بیشتر موارد ارزیابی‌شده، عملکرد پیش‌بینی مدل SVR پایه را نسبت به SVM متعارف و GA–SVM بهبود می‌دهد .رویکرد GSO–SVM دقت عددی پیش‌بینی بهتری، با مقادیر کمتر NMSE وMAE ، و عموماً تطابق جهت‌گیری بهتری، با مقادیر بالاترDS ، ارائه می‌کند .تحلیل حاضر هیچ بهبود مشاهده‌شده‌ای را صرفاً به یک سازوکار جست‌وجوی خاص نسبت نمی‌دهد، مگر آنکه شواهد تجربی اضافی از این تبیین پشتیبانی کند .بنابراین نتیجه‌گیری نهایی بازتاب نتایج تجربی است، نه رفتار مورد انتظار الگوریتم بهینه‌سازی.

۳ .نتیجه‌گیری

این پژوهش رویکرد ترکیبی بهینه‌سازی جست‌وجوی طلایی–رگرسیون بردار پشتیبان (GSO–SVR) را برای پیش‌بینی قیمت پایانی سهام در روز معاملاتی بعد بررسی کرد .چارچوب پیشنهادی توان مدل‌سازی غیرخطی SVR را با یک رویه بهینه‌سازی فرااابتکاری برای انتخاب ابرپارامترهای کلیدی آن، یعنی پارامتر منظم‌سازی و ضریب هسته تابع پایه شعاعی، ترکیب می‌کند .عملکرد این رویکرد در مقایسه با SVR متعارف و SVR بهینه‌شده با الگوریتم ژنتیک (GA–SVR)، در هشت سهم از شرکت‌ها و بخش‌های بازار مختلف بررسی شد .چارچوب تجربی برای حفظ ساختار زمانی داده‌های مالی، با استفاده از افراز زمانی و نگه‌داشتن مشاهدات آزمون نهایی برای ارزیابی خارج از نمونه، طراحی شد .انتخاب ابرپارامترها بر خطای اعتبارسنجی استوار بود و عملکرد پیش‌بینی با خطای مربعات میانگین نرمال‌شده(NMSE) ، میانگین قدرمطلق خطا (MAE) و تقارن جهت‌گیری (DS) سنجیده شد .این معیارهای مکمل، هم دقت عددی پیش‌بینی و هم توانایی مدل در پیش‌بینی جهت حرکت قیمت را نشان می‌دهند.

سهم اصلی این پژوهش بررسی GSO به‌عنوان راهبرد بهینه‌سازی ابرپارامتر در پیش‌بینی قیمت سهام مبتنی برSVR ، در چارچوب مقایسه‌ای هم‌تراز است .بااین‌حال، اثربخشی رویکرد

[Downloaded from mieaoi.ir on 2026-11-19]

پیشنهادی باید برای هر سهم و هر معیار ارزیابی به‌صورت تجربی سنجیده شود و نباید برتری همگانی برای آن فرض گرفت .به‌طور خاص، بهبود در معیارهای خطامحور لزوماً به معنای پیش‌بینی جهت‌گیری بهتر نیست و تفاوت‌های عملکرد مشاهده‌شده در یک دوره آزمون واحد را نمی‌توان بدون آزمون آماری مناسب شاهد معناداری آماری دانست.

چند محدودیت باید مورد توجه قرار گیرد .نخست، تحلیل بر قیمت‌های پایانی تاریخی استوار است و متغیرهای احتمالاً اطلاع‌رسان دیگر مانند حجم معاملات، شاخص‌های مالی یا اطلاعات کلان بازار را در بر نمی‌گیرد .دوم، یافته‌ها ممکن است به سهم‌های منتخب، دوره مشاهده، افق پیش‌بینی و تنظیمات جست‌وجوی ابرپارامتر وابسته باشند .سوم، ارزیابی در یک دوره نگهداشت واحد ممکن است تغییرات شرایط بازار را به‌طور کامل پوشش ندهد .پژوهش‌های آینده می‌توانند با افزودن پیش‌بین‌های بیشتر، ارزیابی در دوره‌های آزمون طولانی‌تر و متعدد، اجرای اعتبارسنجی پنجره غلتان یا پیوسته‌رو و به‌کارگیری آزمون‌های آماری مناسب برای مقایسه عملکرد پیش‌بینی، این محدودیت‌ها را برطرف کنند .در مجموع، این پژوهش چارچوبی برای ارزیابی بهینه‌سازی جست‌وجوی طلایی به‌عنوان ابزار تنظیم مدل‌های SVR در پیش‌بینی سری‌های زمانی مالی ارائه می‌دهد .ارزش عملی آن در نهایت به بازتولیدپذیری بهبودهای مشاهده‌شده در سهم‌ها و دوره‌های ارزیابی مختلف و معناداربودن آن‌ها در ارزیابی سخت‌گیرانه خارج از نمونه وابسته است.

[Downloaded from mieaoi.ir on 2026-11-19]

منابع

۱. اشرفی سلطان احمدی، حسین، جبارزاده کنگرلویی، سعید، بحری ثالث، جمال، و آشتاب، علی. (۱۴۰۴). ارزش‌گذاری سهام بر مبنای نماگرهای بازار سرمایه با رویکرد داده‌های پنلی. *فصلنامه اقتصاد بانکداری اسلامی، ۱۴* (۵۲).

۲. سعیدی‌اقدم، مهران، صادقی، احمد، بحیرایی، علیرضا، و حاجی‌اصغری، سید یوسف. (۱۴۰۱). ارائه مدل پیش‌بینی قیمت سهام با استفاده از الگوریتم‌های یادگیری عمیق و کاربرد آن در قیمت‌گذاری سهام بانک‌های اسلامی. *فصلنامه اقتصاد بانکداری اسلامی، ۱۱* (۴۱).

3. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance, 25*(2), 383–417. https://doi.org/10.2307/2325486
4. Lin, C. Y., & Marques, J. A. L. (2024). Stock market prediction using artificial intelligence: A systematic review of systematic reviews. *Social Sciences & Humanities Open, 9*, 100864. https://doi.org/10.1016/j.ssaho.2024.100864
5. Mintarya, L. N., Halim, J. N. M., Angie, C., Achmad, S., & others. (2023). Machine learning approaches in stock market prediction: A systematic literature review. *Procedia Computer Science, 216*, 96–102. https://doi.org/10.1016/j.procs.2022.12.115
6. Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications, 197*, 116659. https://doi.org/10.1016/j.eswa.2022.116659
7. Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies, 11*(3), Article 94. https://doi.org/10.3390/ijfs11030094
8. Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *WIREs Data Mining and Knowledge Discovery, 14*(1), e1519. https://doi.org/10.1002/widm.1519
9. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009
10. Giglio, S., Kelly, B., & Xiu, D. (2022). Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics, 14*, 337–368. https://doi.org/10.1146/annurev-financial-101521-104735
11. Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing, 14*, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[Downloaded from mieaoi.ir on 2026-11-19]

12. Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support vector regression machines. In *Advances in Neural Information Processing Systems 9* (pp. 155–161). MIT Press.
13. Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks* (Vol. 4, pp. 1942–1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968
14. Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley.
15. Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation, 1*(1), 67–82. https://doi.org/10.1109/4235.585893
16. Noroozi, M., Mohammadi, H., Efatinasab, E., Lashgari, A., Eslami, M., & Khan, B. (2022). Golden search optimization algorithm. *IEEE Access, 10*, 37515–37532. https://doi.org/10.1109/ACCESS.2022.3162853
17. Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. https://otexts.com/fpp3/
18. Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting, 16*(4), 437–450.
19. Bailey, D. H., Borwein, J. M., López de Prado, M., & Zhu, Q. J. (2017). The probability of backtest overfitting. *The Journal of Computational Finance, 20*(4), 39–69. https://doi.org/10.21314/JCF.2016.322
20. Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599
21. White, H. (2000). A reality check for data snooping. *Econometrica, 68*(5), 1097–1126. https://doi.org/10.1111/1468-0262.00152
22. Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis, 120*, 70–83. https://doi.org/10.1016/j.csda.2017.11.003
23. Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics, 23*(4), 365–380. https://doi.org/10.1198/073500105000000063

[Downloaded from mieaoi.ir on 2026-11-19]