فصلنامـه اقتصاد مالی
دورهٔ ۱۸ / شمارهٔ ۳ (پیاپی ٦٨)/ پائیز ۱٤۰۳
صفحه ۳۸۳ تا ٤۰۳

کاربرد مقایسه ای الگوریتم ذرات و الگوریتم ژنتیک در پیش بینی روند بلندمدت و کوتاه مدت بازده سهام

سید جواد کیایی[۱]
زهرا فرشادفر[۲]

تاریخ دریافت: ۱٤۰۳/۰٤/۱٤ تاریخ پذیرش: ۱٤۰۳/۰٦/۲۱

چکیده

عدم وجود قطعیت در روند حرکت بازار سهام پیش بینی آنرا به یک کار پرچالش در حوزه‌ی پیش‌بینی سری‌های زمانی مالی تبدیل کرده است. از سوی دیگر تحلیل داده‌های سری زمانی قیمت های سهام به علت غیرخطی بودن و وجود نویز زیاد آسان نیست. از اینرو هدف این پژوهش پیش بینی روند بلندمدت و کوتاه‌مدت بازار سرمایه است. برای دستیابی به این هدف از الگوریتم های هوش مصنوعی ذرات و ژنتیک بصورت مقایسه‌ای استفاده شده است. متغیر مورد مطالعه شاخص کل قیمت سهام در بورس اوراق بهادار تهران در دوره زمانی ۱۳۹۵ تا ۱۴۰۰ و بصورت ماهانه می‌باشد. داده‌ها پس از گردآوری با استفاده از روش هموارسازی برای روزهای تعطیل بازبینی شده‌اند و به منظور افزایش دقت مدل ها طول پنجره بهینه هر الگوریتم محاسبه شده است. یافته‌های حاصله بیانگر آن است که الگوریتم ژنتیک با به حداقل رساندن خطای پیش بینی یک الگوریتم مناسب برای پیش بینی روند کوتاه مدت و بلند مدت شاخص کل قیمت نسبت به الگوریتم ذرات در دوره زمانی مورد مطالعه است.

واژه‌های کلیدی: الگوریتم ذرات، الگوریتم ژنتیک، هموار سازی داده، یادگیری ماشین، هوش مصنوعی.

طبقه بندی JEL: G17, G32, Z19

۱ گروه حسابداری، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران ایران.sjavadkiaei7@gmail.com
۲ گروه حسابداری، واحد تهران غرب، دانشگاه آزاد اسلامی، تهران ایران.(نویسنده مسئول)zfarshadfar@yahoo.com

نوع مقاله: علمی پژوهشی
Eci@iauctb.ac.ir

Creative Commons – Attribution 4.0 International – CC BY 4.0
Creativecommons.org

۱- مقدمه

توانایی پیش بینی روند بازده سهام یکی از موضوعات مورد مطالعه در بازارهای مالی بوده و هست. اگرچه دیدگاه‌های متفاوتی در مورد کارایی بازارهای مالی وجود دارد اما بسیاری از مطالعات تجربی نشان داده اند که بازارهای مالی تا حدودی قابل پیش بینی هستند (بولرسلو و همکاران[۱]، ۲۰۱۴). پیش بینی های صورت گرفته در بازارهای مالی را می توان در گروه پیش بینی های درون نمونه ای[۲] و پیش بینی های خارج از نمونه [۳] طبقه بندی نمود. بیشتر پژوهش های صورت گرفته در حوزه پیش بینی های درون نمونه ای بازده از دقت خوبی برخوردارند اما در حوزه پیش بینی های خارج از نمونه بازده، این دقت دیده نمی شود. به طور کلی، زمانی که هدف بهبود پیش بینی به ویژه در حوزه داده های خارج از نمونه است اغلب پژوهش های صورت گرفته در صدد دستیابی به این هدف از طریق افزایش متغیرهای پیش بینی کننده هستند و مطالعات کمتری به منظور بهبود پیش بینی بر تغییر مدل متمرکز شده اند و یا در صورت تغییر مدل، اکثر این پژوهش ها بر مدل های خطی تاکید کرده و توجه کمتری به مدل های غیرخطی شده است، این در حالیست که تحقیقات بسیاری نشان داده اند که نوسانات در بازار سهام از الگوی غیر خطی پیروی می کنند (جاسمی و همکاران[۴]، ۲۰۱۱).

در بررسی پژوهش های صورت گرفته در زمینه پیش بینی بازده بورس اوراق بهادار تهران چند نکته قابل تامل است: نخست آنکه پیش بینی روند بازار سرمایه معمولا در یک بستر یکساله انجام شده است و مطالعه ای در زمینه پیش بینی روندکوتاه مدت و بلند مدت بصورت همزمان دیده نمی شود. مساله دوم آن است که در این دسته از پژوهش ها تلاش شده است تا با استفاده از داده های بازارهای موازی چون قیمت نفت و قیمت طلا به پیش بینی روند بازار سهام پرداخته شود. حال آنکه شاخص کل بورس اوراق بهادار به عنوان یک متغیر پروکسی از متغیر های کلان اقتصادی نشاندهنده وضعیت کلی اقتصاد کشور است و تغییرات آن ناشی از تغییرات عوامل درون سازمانی یعنی عوامل درون شرکت ها و عوامل برون سازمانی اثرگذار بر قیمت سهام شرکتها مانند بحرانهای سیاسی و اقتصادی داخلی و بین المللی است (رنجبر ناوی، ۱۴۰۰)، نکته قابل توجه دیگر نبود داده، برای روزهای تعطیل است که برای ایجاد آن باید از روش هموار سازی داده ها استفاده نمود و نهایتا در استفاده از الگوهای غیر خطی وابسته به یادگیری ماشین، تعیین طول پنجره بهینه از اهمیت خاصی در میزان سرعت و دقت پیش بینی برخوردار است که باید مد نظر قرار گیرد (فرشادفر و پروکوپچوک[۵]، ۲۰۱۹).

در این پژوهش تلاش شده است تا به طور همزمان به پیش بینی بلند مدت و کوتاه مدت روند بازار سرمایه پرداخته شود و برای این پیش بینی بجای استفاده از داده های بازارهای موازی از داده های شاخص کل بورس اوراق بهادار استفاده شده است از سوی دیگر برای یافتن داده ها برای روز های تعطیل از روش هموارسازی داده

[1] Bollerslev *et al.*, 2014
[2] In sample forecasting
[3] Out of sample forecasting
[4] Jasemi *et al*., 2011
[5] Farshadfar & Prokopczuk, 2019

ها برای روز های تعطیل استفاده شده و طول پنجره بهینه برای پیش بینی محاسبه شده است که این از نوآوری های پژوهش حاضر است.

از اینرو هدف از انجام این پژوهش آن است که به مقایسه توانمندی الگوریتم ذرات با الگوریتم زنتیک در زمینه پیش بینی روند بلند مدت و کوتاه مدت بازده سهام در بورس اوراق بهادار تهران بپردازد. فرض بر آن است که الگوریتم ژنتیک توانایی پیش بینی بهتری نسبت به الگوریتم ذرات دارد. برای دستیابی به این هدف از داده های شاخص کل بورس اوراق بهادار در دوره زمانی ۱۳۹۵ تا ۱۴۰۰ بصورت ماهانه استفاده شده است.

در ادامه مروری بر پژوهش های صورت گرفته در زمینه پیش بینی بازده انجام شده سپس روش پژوهش توضیح داده شده و نهایتا نتایج و پیشنهادات ارائه شده است.

مبانی نظری و پیشینه پژوهش

پیش بینی نوسان شاخص کل بورس اوراق بهادار تهران برای سرمایه گذارانی که در پی بیشینه کردن منافع خود هستند، بسیار با اهمیت است. با توجه به اهمیت شاخص کل قیمت سهام و همچنین پیش بینی نوسان شاخص کل هدف پژوهش براین بنا نهاده شد که اقدام به پیش بینی روند شاخص کل قیمت سهام با استفاده از الگوریتم های هوش مصنوعی نماید.

شاخص کل نشان دهنده وضعیت کلی اقتصاد کشور است، افزایش این شاخص به طور معمول به معنی رونق و بهبودی در اوضاع و احوال اقتصادی و کاهش آن گویای بحران و رکود است (محمدی، ۱۳۹۴). محاسبه شاخص قیمت سهام مستلزم در اختیار داشتن اطلاعات کامل پایه ای و آخرین اطلاعات از تغییرات قیمتهای سهام و حجم معاملات آنهاست به این جهت این شاخص توسط سازمان کارگزاران بورس اوراق بهادار یا موسسات تخصصی مستقل از بورس، محاسبه میشود. تغییرات شاخص کل قیمت سهام به دنبال تغییرات در مقدار سهام مورد معامله و تغییرات قیمت آنها پدید می آید. این تغییرات ناشی از عوامل درونسازمانی و برونسازمانی است که منظور از عوامل درونی بیشتر به اوضاع و احوال درون شرکتها با توجه به سرمایه گذاریهای جدید، سودآوری، زیاندهی کارایی تصمیمات مجمع عمومی و مدیریت شرکتهاست. منظور از عوامل بیرونی یا برونسازمانی اثرگذار بر قیمت سهام شرکتها، قوانین مالی و مالیاتی، بحرانهای سیاسی و اقتصادی داخلی، ثبات یا بی ثباتی حکومتها، جنگها و تهدیدهای سیاسی منطقه ای یا بین المللی است. بورس تهران از فروردین ماه ۱۳۶۹ اقدام به محاسبه و انتشار شاخص قیمت خود با نام تیپیکس[1] نموده است. این شاخص ۵۲ شرکت که در آن زمان کل شرکتهای پذیرفته شده در بورس را شامل میشدند، در برمیگرفت (کرباسی یزدی و همکاران،۱۳۹۱).

دردنیای رقابتی امروز، استفاده ازالگوی پیش بینی مناسب و کارا درپیش بینی قیمت کالاها و سودهای سهام و تنظیم بازار نقش اساسی دارد. شاخص کل قیمت از عوامل گوناگون و گاه متناقضی تأثیر میپذیرد که تحلیل این رفتار غیرخطی و پر نوسان، مستلزم استفاده از الگوریتمهای پیچیده است (پن و همکاران[2]، ۲۰۱۵).

[1] TEPIX
[2] Phan *et al.*, 2015

هوش مصنوعی[1] شاخه ای از علوم رایانه است که هدف اصلی آن تولید ماشین‌های هوشمند با توانایی انجام وظایف هوش انسانی است. که درحقیقت نوعی شبیه سازی هوش انسانی برای کامپیوتر است در این فرایند در واقع ماشین به گونه ای برنامه نویسی شده که همانند انسان فکر کند و توانایی تقلید از رفتار انسان را داشته باشد . این تعریف می تواند به تمامی ماشین هایی اطلاق شود که بگونه‌ای همانند ذهن انسان عمل می‌کنند و می‌توانند کارهایی مانند حل مسئله و یادگیری داشته باشند.

منظور از یادگیری ماشین[2] طراحی ماشین‌هایی است که با استفاده از مثال‌های داده شده به آن‌ها و تجربیات خودشان، بیاموزند. در این علم تلاش می‌شود تا با بهره‌گیری از الگوریتم‌ها، یک ماشین به شکلی طراحی شود که بدون آنکه صراحتا برنامه‌ریزی و تک تک اقدامات به آن دیکته شود بتواند بیاموزد و عمل کند. در یادگیری ماشین، به جای برنامه‌نویسی همه چیز، داده‌ها به یک الگوریتم عمومی داده می‌شوند و این الگوریتم است که براساس داده‌هایی که به آن داده شده منطق خود را می‌سازد (فرشادفر و پروکوپچوک[3]، ۲۰۱۹).

الگوریتم ها مجموعه فرایندها یا فرمول هایی هستند که به کمک آن‌ها می‌توان بسیاری از مسائل برنامه نویسی را حل نمود؛ به بیانی دیگر الگوریتم به دنباله‌ای از مراحل گفته می شود که برای حل یک مساله مشخص تعریف می شوند.

یکی ازانواع الگوریتم ها که معمولا در فرآیند های جستجو و بهینه سازی مورد استفاده قرار می گیرند، الگوریتم های تکاملی[4] (EA) هستند. الگوریتم های تکاملی مجموعه ای از اکتشافات مدرن متشکل از موفقیت در بسیاری از برنامه های کاربردی با پیچیدگی های زیاد است که تحت عنوان محاسبات تکاملی[5] (EC) شناخته می شوند.

ایده اصلی الگوریتم های تکاملی در سال ۱۹۶۰ توسط ریچنبرگ مطرح گردید. برای حل مسایل بهینه سازی می توان از الگوریتم های تکاملی مختلفی استفاده کرد که از آن جمله می توان به الگوریتم های ژنتیک و ازدحام ذرات اشاره کرد. در ادامه به تعریف ساختار هر یک از این الگوریتم ها پرداخته می شود:

۱-۲-شبکه عصبی

الگوریتم‌های یادگیری از ساختارهای لایه‌ای با نام شبکه‌های عصبی استفاده می‌کنند تا بر اساس تجزیه و تحلیل داده‌ها، برای گرفتن تصمیم خاصی، رفتار انسان را تقلید کنند. طرح این ساختار لایه‌ای، برگرفته از ساختار مغز انسان است. همانطور که مغز انسان به شناسایی الگوهای مختلف داده‌ها و دسته‌بندی انواع اطلاعات می‌پردازد، می‌توان شبکه‌های عصبی را به شیوه‌ای مشابه با رفتار مغز انسان آموزش داد تا به تشخیص الگوها بپردازند و دسته‌بندی داده‌ها را انجام دهند.

[1] Artificial Intelligence
[2] Machine Learning
[3] Farshadfar & Prokopczuk, 2019
[4] Evolutionary Algorithms
[5] Evolutionary Computation

به عبارتی شبکه عصبی یک سامانه پردازشی داده هاست که از مغز انسان ایـده گرفته و پردازش داده ها را بر عهـده پردازنـده هـای کوچک و بسیار زیادی سپرده و به صورت شبکه ای به هم پیوسته و موازی با یکدیگر هستند تا یک مسئله را حل نمایند. در واقع این الگوریتم قابلیـت حـل مـسئله بدون کمک فرد خبـره یـا متخصصان را دارا است و الگوهایی را در اطلاعات و دادهها شناسایی می‌کند. شبکه های عصبی با توانایی قابـل توجـه خـود در استنتاج نتایج از داده های پیچیده، می تواننـد در استخراج الگوها و شناسایی گرایش های مختلفی که شناسایی آنها برای انسانها و کامپیوتر بسیار دشوار است استفاده شوند.

۲-۲-الگوریتم ژنتیک[1]

ایده اصلی الگوریتم های تکاملی در سال ۱۹۶۰ توسط ریچنبرگ مطرح گردید. الگوریتم های ژنتیک که یکی از این نوع الگوریتم ها است، در حقیقت روش جستجوی کامپیوتری بر پایه الگوریتم های بهینه سازی و بر اساس ساختار ژن ها و کروموزومها است که توسط پروفسور هالند در دانشگاه میشیگان مطرح و پس از وی توسط دانشجویانش مثل گولبرگ توسعه یافت. الگوریتم های ژنتیک یک روش جستجوی مؤثر در فضاهای بسیار وسیع و بزرگ است که در نهایت منجر به جهت گیری به سمت پیدا کردن یک جواب بهینه می گردد که شاید نتوان در مدت زمان زندگی یک فرد به آن جواب بهینه دست یافت. بنابراین الگوریتم های ژنتیک با یک سری متغیرهای کد شده کار می کنند. نکته جالب دیگر این است که اصول الگوریتم ژنتیک بر پردازش تصادفی یا به تعبیر صحیح تر، پردازش تصادفی هدایت شده استوار است. بنابراین عملگرهای تصادفی، فضای جستجو را به صورت تطبیقی مورد بررسی قرار می دهند.

برای پیاده سازی، الگوریتم ژنتیک کار خود را با مجموعه ای از کرموزوم های کد گذاری شده (که به آنها جمعیت اولیه گفته می شود) شروع و بر اساس تابع شایستگی، میزان مناسب بودن (شایستگی) هر کرموزوم را برحسب یک عدد بیان می کند. سپس به صورت تصادفی از بین کرموزوم های موجود، والدین انتخاب می شوند و با توجه به ویژگی های والدین و به کمک عملگرهای مختلف الگوریتم ژنتیک مانند ادغام، جهش و سایر عملگرها، تغییراتی در ژن های موجود درکرموزوم های والدین ایجاد می شود تا فرزندان تولید شده و به مرحله بعد منتقل گردند (به هر یک از مراحل الگوریتم ژنتیک، نسل گفته می شود) اصول کار الگوریتم ژنتیک بر این منوال قرار می گیرد که کرموزوم های هر نسل، شایستگی بیشتری نسبت به نسل قبلی داشته باشند. این عمل هزاران یا شاید میلیون ها بارتکرار می شود تا یکی از شرایط توقف الگوریتم حاصل شود و بدین ترتیب، الگوریتم ژنتیک به بهینه سازی سراسری همگرا می شود. در نهایت جواب های بدست آمده باید رمزگشایی گردند تا مقادیر واقعی آنها نشان داده شود.

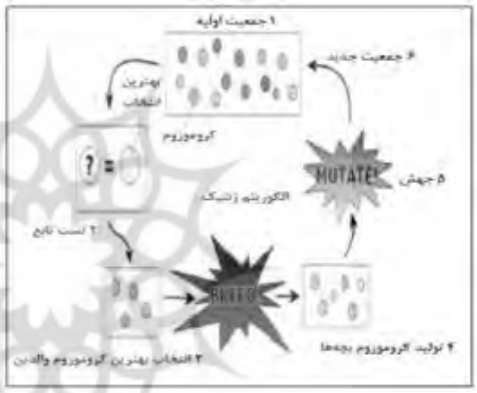
این الگوریتم یک تکنیک برنامه نویسی است که از تکامل ژنتیکی به عنوان یک الگـوی حـل مـسئله استفاده می کنـد. مسئله ای که باید حل شود، ورودی است و راه حل ها طبـق یک الگو کدگذاری می شود و متریک که

[1] Genetic Algorithm

تابع برازش نیز نام دارد، هـر راه حـل کاندید را ارزیابی می کند و اکثر آنهـا بـه صـورت تصادفی انتخـاب مـی شـوند. در الگوریتم هـای ژنتیک ابتـدا بـه طـور تصادفی یا الگوریتمیک، چندین جواب برای مسئله تولید می شود. این مجموعـه جواب جمعیت اولیه است؛ هر جواب یک کروموزوم است. سپس با استفاده از عملگرهای الگوریتم ژنتیک پس از انتخاب کروموزوم هـای بهتر، کروموزومها را بـا هـم ترکیب کرده و جهشی در آنهـا ایجـاد می شود. در نهایت نیز جمعیت فعلی با جمعیت جدیدی که از ترکیب و جهش در کروموزومها حاصل می شود، ترکیب می شود. مراحل انجام کار در این الگوریتم به صورت زیر است:

مرحله اول: یک سری کروموزوم به عنوان جمعیت اولیه به صورت تصادف انتخاب می شود . هر کروموزوم عددی در مبنای دو است.

1.	00001011
2.	00100010
3.	01000000
4.	11100001
5.	01101100
6.	00000111
7.	11001010
8.	11110000
9.	00010101
10.	10000000
11.	11100100



شکل ۱- ساختار ژنتیک

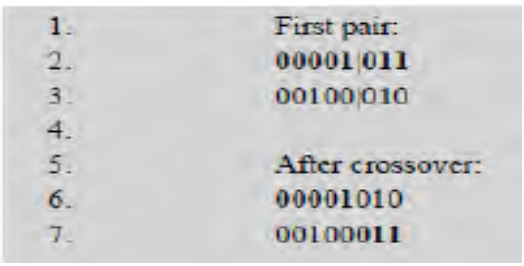
منبع: دو دائو و همکاران، ۲۰۱۷

به عنوان مثال، این کروموزوم ها به عنوان جمعیت اولیه در نظر گرفته شده است : بعد از این که جمعیت اولیه معلوم شد این کروموزوم هـا در تابع برازش[1] امتحان می شوند، بر حسب این که به جواب مـورد نظر نزدیک ترند یا خیر، یک عدد بین صفر تا یک به آنها اختصاص داده می شود که صفر یعنی به درد نمی خـورد و یک یعنی خیلی خوب است.

مرحله دوم، مرحله گونه[2] است که طبق فرایند ادغام، کروموزوم ها با هم ترکیب شده و کروموزوم جدید حاصل می شود.

فرآیند ادغام : در این مرحله کروموزومهای برگزیده دوتا دوتـا انتخـاب شـده و فرایند ادغام روی هر زوج به صورت زیر انجام می شود:

[1] fitness
[2] Breed

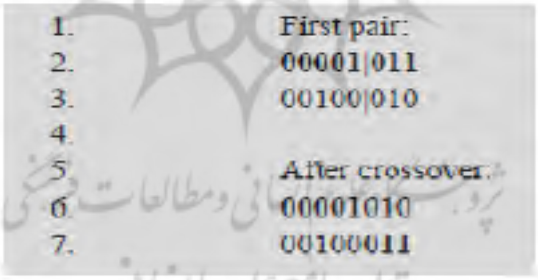


شکل ۲- فرایند ادغام

منبع: دو دائو و همکاران، ۲۰۱۷

در شکل بالا، فرآیند ادغام[1] برای زوج اول مشاهده می شـود. همان طور که مشخص است اول هر کروموزوم از بیت ۱۵ ام به دو قسمت تقسیم شده است و ۵ بیت اول کروموزوم اول با ۳ بیت دوم کروموزوم دوم ترکیب شده و برعکس؛ به این ترتیب دو فرزند جدید به وجود می آید. همین کار برای بقیه کروموزومها نیز انجام می شـود که ممکن است یک کروموزوم دو یا چند بار در فرآینـد ادغام به کار برده شود. احتمال شرکت کروموزومهایی که سلامت بهتـری دارند در فرآیند ادغام بیشتر است.

مرحله سوم مرحله جهش است. بعد از فرآیند ادغام یک مرحله دیگر وجـود دارد که احتمال وقوع آن خیلی کم است، این مرحله جهش[2] است. در این فرآیند یک بیت تصادفی از یک کروموزوم تصادفی عـوض مـی شـود . مثلا اگر بیت چهارم یک کروموزوم انتخاب شود در صورتی که صـفر باشد آن یک می شود یا بالعکس.



شکل شماره ۳- مرحله جهش

منبع: دو دائو و همکاران، ۲۰۱۷

[1] Crossover
[2] Mutation

۲-۳-الگوریتم ازدحام ذرات[1]

در اوایل سال ۱۹۹۰ میلادی، پژوهش‌های گوناگونی پیرامون رفتار اجتماعی گروه‌های حیوانات انجام شد. این پژوهش‌ها حاکی از آن بودند که برخی از حیوانات که به یک گروه خاص متعلق هستند، مانند پرندگان، ماهی‌ها و دیگر موارد، قادر به آن هستند که اطلاعات را در گروه‌های (دسته‌ای یا گله‌ای) خودشان به اشتراک بگذارند و چنین قابلیتی به این حیوانات مزایای قابل توجهی برای بقا اعطا می‌کرد.

با الهام گرفتن از این مطالعات، «کندی » و «اِبرهارت» در سال ۱۹۹۵ الگوریتم بهینه سازی ازدحام ذرات یا الگوریتم PSO را در یک مقاله معرفی کردند. الگوریتم بهینه سازی ازدحام ذرات یا الگوریتم PSO یک الگوریتم فراابتکاری است که برای بهینه‌سازی توابع پیوسته غیر خطی مناسب محسوب می‌شود. نویسندگان مقاله مذکور، الگوریتم بهینه سازی ازدحام ذرات یا الگوریتم PSO را از مفهوم هوش ذرات که معمولا در گروه‌های حیوانات مانند گله‌ها و دسته‌های حیوانات وجود دارد الهام گرفته و ساخته‌اند. اساس حرکت تجمعی ذرات، شبیه سازی یک رفتار دسته جمعی است که از آن برای نشان دادن حرکت گروه پرنده ها و ماهیان استفاده میشود و یکی از رایج ترین تکنیک های بهینه سازی فرا اکتشافی است. حرکت تجمعی ذرات از عامل جمعیت که شامل راه حلهای بالقوه مسأله ی تحت بررسی است، جهت اکتشاف در فضای جستجو استفاده میکند.

این الگوریتم شامل دو معادله سرعت و مکان است که مختصات هر ذره نمایانگر یک جواب ممکن مرتبط با دو بردار است .بردارهای موقعیت (xi) و سرعت (vi) دو بردار وابسته و مرتبط با هر ذره در فضای جستجو N بعدی هستند که به ترتیب به صورت زیر بیان می گردند:

رابطه (۱) : $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}$

رابطه (۲): $v_i = \{v_{i1}, v_{i2}, \dots, v_{in}\}$

یک اجتماع از پرندگان از تعدادی ذره (پاسخ های ممکن) تشکیل شده است که در یک فضای پاسخ ممکن ، برای جستحوی جواب های بهینه پیش میروند. موقعیت هر ذره با توجه به بهترین جستجوی خود ذره ، بهترین تجربه کلی پرواز گروهی و بردار سرعت پیشین خود ذره ، مطابق با رابطه زیر به هنگام میشود .

رابطه (۳) : $v_i^{k+1} = wv_i^k + c_1 r_1 \left(pbest_i^k - x_i^k\right) + c_2 r_2 \left(gbest^k - x_i^k\right)$

رابطه (۴): $x_t^{k+1} = x_t^k + v_t^{k+1}$

که در آن C_1 و C_2 دو ثابت مثبت هستند. r_1 و r_2 دو عدد تصادفی با توزیع یکنواخت بین صفر و یک هستند. w وزن لختی است. $pbest^k_i$ بهترین موقعیت ذره i است که با توجه به تجربه این ذره به دست می آید:

[1] Particle Swarm Optimization

رابطه (۵) : $pbest_i^k = [x_{t1}^{pbest}, x_{t2}^{pbest}, \dots\dots\dots, x_{tN}^{pbest}]$

$Gbest^k$ بهترین موقعیت ذره با توجه به تجربه کلی گروهی است.

رابطه (۶) : $gbest_i^k = [x_1^{gbest}, x_2^{gbest}, \dots\dots\dots, x_N^{gbest}]$

و سرانجام K شاخص تکرار است . روند الگوریتم PSO در حالت کلی به این صورت است که، برای هر ذره بردارهای موقعیت و سرعت به صورت تصادفی مشابه با ابعاد مساله مقداردهی اولیه میشوند. برازندگی برای هر ذره pbest اندازه گیری شده و ذره دارای بهترین مقدار برازندگی gbest ذخیره میگردد. بردارهای سرعت و موقعیت مطابق با روابط بالا برای هر ذره به هنگام میشوند. علیرغم مفهوم ساده و پیاده سازی آسان روش معرفی شده، برتری آن در مقایسه با سایر روشها در بسیاری از زمینه های مختلف کاربردی، به اثبات رسیده است. اگر رابطه تغییر سرعت که در الگوریتم سراسری آمده است به صورت زیر نوشته شود.

رابطه (۷): $v_i^{k+1} - v_i^k = c_1 r_1 (pbest_i^k - x_i^k) + c_2 r_2 (gbest^k - x_i^k)$

مشخص است که سمت چپ رابطه در واقع بیانگر شتاب ذره i ام در زمان t است.

رابطه (۸): $\alpha_t^k = v_t^{k+1} - v_t^k$

در نتیجه به c_1 و c_2 که اعداد ثابت مثبتی هستند، ضرایب شتاب گفته میشود و در بعضی از مطالعات به ترتیب به عنوان ضرایب شناسایی و اجتماعی معروف هستند و کشش نسبی pbest و gbest را تعیین میکنند .ضریب c_1 ، ضریبی است که تعیین میکند ذره چه مقدار از بهترین وضعیت حافظه اش تاثیر میگیرد و c_2 تعیین میکند که ذره چه مقدار از بهترین وضعیت بقیه جمع تاثیر میگیرد. افزایش مقدار پارامتر c_1 باعث تقویت جستجوی فضای پاسخ میگردد، به طوری که هر ذره به سمت بهترین موقعیت خودش، pbest حرکت میکند. همچنین افزایش در مقدار پارامتر c_2 باعث تقویت میزان بهره برداری از حداقل سراسری فرضی میشود؛ کارهای اخیر که در زمینه PSO انجام شده است، مقادیر $C_1 = C_2 = 2$ را به عنوان یک مقدار مناسب برای این الگوریتم پیشنهاد میکنند. ضرایب r_1 و r_2 برای تولید اعداد تصادفی مابین صفر و یک هستند که معمولاً هر دو عدد تابع تصادفی کاملاً متفاوت از هم هستند و در بیشتر پیاده سازی ها، دو عدد تصادفی مذکور به صورت مستقل از هم و تصادفی با توزیع یکنواخت هستند که برای تغییر دادن میزان کشش pbest و gbest استفاده میشوند. استفاده از اعداد تصادفی در الگوریتم بهینه سازی

به منظور شبیه سازی بخش اندکی از رفتار غیرقابل پیشگویی اجتماع ذرات است این الگوریتم در سال ۲۰۰۲ توسط کلرک[1] و کندی[2] برای مسائل چند هدفه توسعه داده شد.

در ارتباط با کاربرد هر یک از این الگوریتم ها در ادبیات اقتصادی پژوهش هایی در داخل و خارج از کشور صورت گرفته که در ادامه به پاره ای از جدید ترین پژوهش ها در این حوزه پرداخته می شود.

ربکا و همکاران (۲۰۲۲) در پژوهشی با عنوان پیش بینی روند صعودی ۱۵ سهم با توجه به قیمت بسته شدن روز گذشته نتیجه می گیرند که افزایش ۸۰ درصدی دقت پیش بینی استفاده از الگوریتم ژنتیک از جنگل تصادفی بیشتر است.

کی شارما و همکاران (۲۰۲۲) در پژوهشی با عنوان پیش بینی بازار سهام نتیجه می گیرند که مدل ترکیبی هیبریدی شبکه عصبی مصنوعی با ژنتیک دقت بیشتری نسبت به مدل منفرد شبکه عصبی مصنوعی در کوتاه مدت و بلند مدت را داراست

چاندار (۲۰۱۹) در پژوهشی با عنوان پیش‌بینی بازار سهام با استفاده از خوشه‌بندی کاهشی (تفریقی) بر رویکرد ترکیبی فازی عصبی نتیجه می گیرد میانگین عملکرد رویکرد فازی عصبی مبتنی بر خوشه‌بندی تفریقی به طور قابل توجهی بهتر از سایر شبکه‌ها بود.

پاره ای از جدیدترین پژوهشهای داخلی صورت گرفته به صورت زیر می باشد:

بابایی و همکاران (۱۴۰۲) در مقاله ای با عنوان پیش بینی ورشکستگی شرکت ها با استفاده از مکانیزم های حاکمیت شرکتی و نسبت های مالی نتیجه می گیرند که الگوریتم ژنتیک نتایج بهینه ی شبکه عصبی را بهبود می بخشد و جواب بهینه تری ارائه می گردد.

مهرانی و رحیمی پور (۱۴۰۱) در پژوهشی با عنوان پیش بینی احتمال تجدید ارائه صورتهای مالی در می یابند که دقت و کارآیی مدل بهبود یافته با الگوریتم ژنتیک ۷۳/۲۱ درصد دقت پیش بینی کل را داشته که دارای بالاترین قدرت پیش بینی در پژوهش بوده است.

رجبی و حاجی یخچالی (۱۴۰۱) در مقاله ای با عنوان بهینه سازی جریان نقدینگی سبد پروژه نتیجه می گیرند که عملکرد الگوریتم بهینه سازی ازدحام ذرات از الگوریتم ژنتیک از نظر سرعت اجرای برنامه و سرعت رسیدن به نتیجه نهایی بهتر است.

القیانی و همکاران (۱۴۰۰) در پژوهشی با عنوان تبیین گزارشگری مالی-مالیاتی متقابلانه شرکت ها نتیجه می گیرند که الگوریتم ازدحام ذرات با بیشترین درصد پیش بینی صحیح, بهینه ترین مدل را حاصل نموده و در بررسی توسط داده های آزمایشی و آموزشی کاراترین الگوریتم است.

[1] Clerc
[2] Kennedy

با توجه به مطالعات صورت گرفته در زمینه پیش بینی بازده فرضیه اصلی این پژوهش آن است که الگوریتم ژنتیک توانایی بهتری نسبت به الگوریتم ذرات در پیش بینی روند بلند مدت و کوتاه مدت بازده دارد فرضیه های فرعی بصورت زیر هستند:

فرضیه اول: الگوریتم ژنتیک توانایی بهتری نسبت به الگوریتم ذرات در پیش بینی روند بلند مدت بازده دارد.

فرضیه دوم: الگوریتم ژنتیک توانایی بهتری نسبت به الگوریتم ذرات در پیش بینی روند کوتاه مدت بازده دارد.

روش پژوهش

پژوهش حاضر پژوهشی کاربردی است که هدف از انجام آن مقایسه توان پیش بینی بازده الگوریتم های شبکه عصبی بازگشتی[1] در بورس اوراق بهادار تهران است. برای دستیابی به این منظور دو دسته از توانمند ترین الگوریتم های این حوزه یعنی الگوریتم های ژنتیک و ازدحام ذرات برای پیش بینی بازده سهام انتخاب شده و کارایی نتایج به دست آمده از این دو الگوریتم با یکدیگر مقایسه شده اند. برای رسیدن به این هدف از داده های شاخص کل بورس اوراق بهادار تهران در دوره پنج ساله ۱۴۰۰ – ۱۳۹۵ و به صورت ماهانه استفاده شده است. نرم افزار مورد استفاده نرم فزار پایتون و کتابخانه های کاربردی کتابخانه کراس[2] و دیپ[3] می باشند. از ۸۰ درصد داده های موجود در دیتاست برای آموزش مدل و ۲۰ درصد باقیمانده برای تست شبکه استفاده شده است. نتایج بدست آمده به شکل جدول گردآوری و نتایج با استفاده از فرمول خطای مربعات میانگین اشتباهات (MSFE) با یکدیگر مقایسه شده اند.

یافته های پژوهش

قبل ازاستفاده از داده های خام از روش تسطیح داده ها که یکی از روش های پیش پردازش داده ها[4]استفاده شده است چرا که داده های بازارهای مالی معمولا در روزهای تعطیل ارایه نمی شوند. به همین دلیل در روزهای تعطیل آخر هفته و همچنین تعطیلات رسمی، در دیتاست های ارایه شده، داده ای وجود ندارد. این مساله می تواند یک مشکل بالقوه در آموزش و آماده سازی شبکه های عصبی باشد. چون در شبکه عصبی بازگشتی توالی داده های و بازه زمانی بین داده های متوالی اهمیت بسزایی دارد.

برای برطرف شدن این مشکل از روش تسطیح داده ها استفاده شده است. هدف از کاربرد این روش، خواندن داده های خام موجود برای شاخص سهام (مقادیر ارایه شده در بازه زمانی ۵ ساله) و ایجاد یک داده یکدست با اختلاف دقیق یک روز بین هر دو داده می باشد. بدین منظور یک کد دستوری تهیه شده که ابتدا داده ها را از

[1] RNN
[2] Keras
[3] Deap: Distributed Evolutionary Algorithms in Python
[4] PreProcessing

فایل ورودی خوانده و سپس به شناسایی روزهای تعطیل و ایجاد داده برای آنها بر اساس آخرین قیمت در آخرین روز غیرتعطیل می پردازد و در نهایت داده های اصلاح شده را ذخیره می کند.

۱-۴-انتخاب طول پنجره بهینه

در شبکه‌های عصبی بازگشتی، معمولاً از یک پنجره یا نوار زمانی[1] برای ورودی به شبکه استفاده می‌شود. پنجره زمانی به معنی یک بازه زمانی مشخص در دنباله ورودی است که در هر مرحله به شبکه داده می‌شود. این بازه زمانی می‌تواند شامل چندین نمونه داده بوده و به صورت گذرا[2] یا غیر گذرا[3] باشد. می‌توان طول پنجره زمانی را در شبکه‌های عصبی بازگشتی بهینه‌سازی کرد. انتخاب طول پنجره زمانی به دقت وابسته است و به عنوان یک پارامتر مهم در شبکه‌های عصبی بازگشتی در نظر گرفته می‌شود. به‌طور کلی، اگر پنجره زمانی بسیار کوتاه باشد، اطلاعات زیادی از داده‌های گذشته در نظر گرفته نمی‌شود و اگر پنجره زمانی بسیار طولانی باشد، ممکن است باعث ایجاد مشکلاتی مانند بیش‌برازش شود. بنابراین، بهترین طول پنجره زمانی باید به گونه‌ای باشد که بهترین عملکرد شبکه را ارائه دهد و برای هر مسئله ممکن است متفاوت باشد.

مطالعات نشان می دهد که طول پنجره در پیش بینی سری زمانی بازار سهام مهم است. چرا که علیرغم این فرض کلی که یک مدل هوش مصنوعی با داده‌های بیشتر بهتر عمل می‌کند، شناسایی اینکه کدام داده‌ها به عملکرد آن کمک می‌کنند نیز مهم است. داده های بسیار کم، عملکرد مدل را فلج می کند در حالی که داده های زیاد فرآیند یادگیری آن را کند می کند یا بدتر از آن، عملکرد را کاهش می دهد.

طول پنجره زمانی مربوط به داده‌های این پژوهش با استفاده از روش های الگوریتم ژنتیک، ازدحام ذرات بهینه سازی شده است. به این صورت که برای بهینه‌سازی طول پنجره زمانی در RNN با استفاده از الگوریتم ژنتیک، پارامترهایی مانند طول پنجره زمانی به عنوان جمله‌های ژنتیکی در نظر گرفته شده و برای هر جمله ژنتیکی یک مقدار مناسب برای طول پنجره زمانی در نظر گرفته شده است سپس با استفاده از عملیات انتخاب، یافتن بهترین طول پنجره زمانی صورت گرفته است.

روش کار چنین است که کتابخانه های مورد استفاده لود شده و داده های اصلاح شده در بخش های قبلی از فایل csv خوانده می شود. همچنین در این بخش ۸۰ درصد داده های موجود در دیتاست برای آموزش مدل و ۲۰ درصد باقیمانده برای تست شبکه استفاده شده است. برای تعیین تعداد روزهای پیش بینی در کد های دستوری از پارامتر y_delay در فرآیند آموزش استفاده شده که مقادیر آن همانطور که در جدول شماره یک نشان داده شده است برای دوره ی یک ماهه ۳۰ برای دوره سه ماهه ۹۰ برای دوره شش ماهه ۱۸۲ و برای یک ساله ۳۶۴ در نظر گرفته شده است.

[1] Time window
[2] overlapping
[3] Non overlapping

جدول ۱- تعداد روزهای پیش بینی در فرآیند آموزش

y_delay	بازه پیش بینی
۳۰	یک ماهه
۹۰	سه ماهه
۱۸۲	شش ماهه
۳۶۴	یک ساله

منبع: یافته های پژوهشگر

در کد مذکور تنظیمات الگوریتم ژنتیک با استفاده از کتابخانه DEAP اعمال شده است و الگوریتم برای بهینه سازی آماده شده است. اندازه جمعیت در الگوریتم برابر ۲۰ و تعداد نسل ها برابر ۱۰۰۰ نسل در نظر گرفته شده است.

در این فرایند تابع train_evaluate به عنوان تابع خسارت تعریف شده است. کار اصلی این تابع این است که طول پنجره زمانی ارایه شده توسط الگوریتم را دریافت کرده و داده های آموزش را بر اساس آن ایجاد می کند سپس مدل توسط کتابخانه Keras آموزش داده شده و با استفاده از داده های تست، دقت آن سنجیده می شود. شاخص mean_squared_error به عنوان شاخص دقت در نظر گرفته شده است. دقت محاسبه شده به عنوان خروجی تابع خسارت برگردانده میشود. لازم به ذکر است که مدل بهینه سازی برای ماکزیمم سازی تنظیم شده است. تابع generate_xy برای آماده سازی داده ها و نرمالیزه کردن آنها طراحی شده است . همچنین داده ها را در قالب موردنیاز Keras برای آموزش تنظیم می کند. بعد از انجام بهینه سازی، بهترین نتایج بدست آمده از بهینه سازی برای طول پنجره زمانی، مشخص می شود که در جدول شماره ۲ آورده شده است.

همانطور که در جدول شماره دو دیده می شود مقدار بهینه طول پنجره زمانی بدست آمده برای الگوریتم ژنتیک و الگوریتم ذرات بهینه سازی شده و فواصل پیش بینی های مختلف ارایه شده است:

به این صورت که برای دوره پیش بینی یک ماهه در الگوریتم ژنتیک طول پنجره بهینه ۳۸ و بازه سه ماهه ۲۶ و در دوره شش ماهه ۴۷ و در دوره پیش بینی یک ساله طول پنجره بهینه ۲۹ می باشد.

به همین صورت برای الگوریتم ذرات طول پنجره بهینه برای دوره های پیش بینی یک ماهه و سه ماهه و شش ماه و یک ساله به ترتیب ۲۷و۷۷ و ۵۰ و ۶۷ محاسبه شده است.

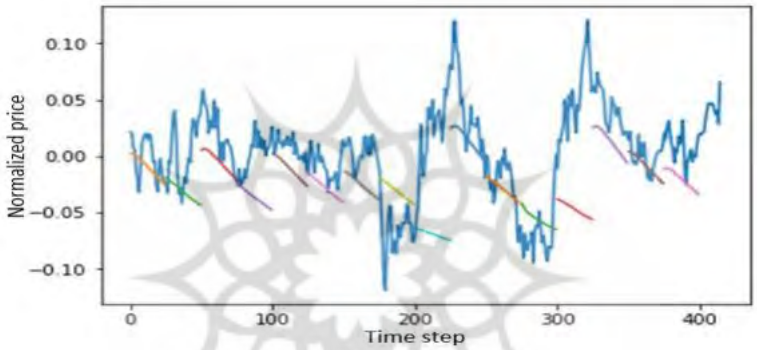
جدول ۲- بررسی طول پنجره با استفاده از الگوریتم ذرات و الگوریتم ژنتیک

مقدار بهینه پنجره زمانی	بازه پیش بینی	روش بهینه سازی
۳۸	یک ماهه	الگوریتم ژنتیک
۲۶	سه ماهه	
۴۷	شش ماهه	
۲۹	یک ساله	
۲۷	یک ماهه	

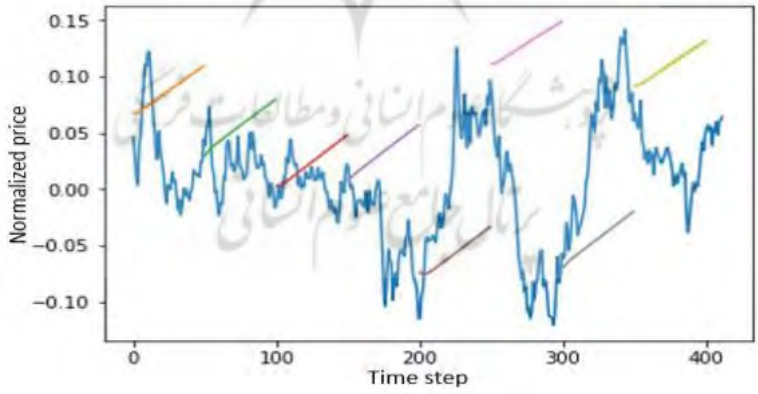
مقدار بهینه پنجره زمانی	بازه پیش بینی	روش بهینه سازی
۷۷	سه ماهه	الگوریتم ازحام ذرات
۵۰	شش ماهه	
۶۷	یک ساله	

منبع: یافته های پژوهشگر

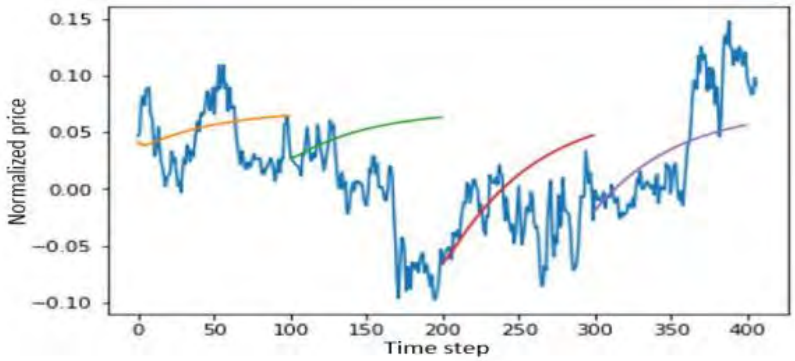
به عنوان نمونه در نمودار های زیر نیز اثر طول پنجره بر دقت پیش بینی شاخص سهام در سه بازه ۲۵ و ۵۰ و ۱۰۰ روزه با استفاده از الگوریتم ژنتیک مقایسه شده اند.



نمودار ۱- طول پنجره زمانی ۲۵ روزه



نمودار ۲- طول پنجره زمانی ۵۰ روزه



نمودار ۳- طول پنجره زمانی ۱۰۰ روزه

با توجه به اینکه از شبکه عصبی المان به عنوان پایه بهینه سازی بر اساس هر الگوریتم استفاده شده پس با مقدار بهینه پنجره زمانی مواجه خواهیم شد .

۲-۴-پیش بینی روند بلند مدت و کوتاه مدت بازده سهام

بعد از مشخص شدن طول پنجره بهینه پیش بینی روند کوتاه مدت (دوره یک ماهه و سه ماهه) و روند بلند مدت (شش ماهه و یکساله) بازده سهام انجام شده است که نتایج آن در جدول شماره ۳ آورده شده است.

برای مقایسه الگوریتم ها در ۴ دوره ی مختلف، از داده ها و نتایج ۵ نمونه برای هر دوره استفاده شده است ، لازم به ذکر است مقادیر پیش بینی شده در این جدول، بهترین نتایج بدست آمده از هر الگوریتم می باشند که بر اساس طول بهینه پنجره زمانی مربوط به هر الگوریتم و هر دوره محاسبه شده اند. مقادیر پیش بینی شده مربوط به هر تاریخ آورده شده توسط هر الگوریتم در کنار مقدار شاخص واقعی در همان تاریخ قرار گرفته، هرکدام از مقادیر پیش بینی شده که نزدیک تر به شاخص واقعی می باشند نشان دهنده عملکرد بهتر آن الگوریتم می باشند.

همانطور که در جدول ۳ دیده می شود در پیش بینی روند کوتاه مدت در دوره پیش بینی یک ماهه به غیر از تاریخ ۱۳۹۵/۰۳/۳۱ که الگوریتم ازدحام ذرات عددی نزدیک تر به شاخص واقعی را نشان می دهد در سایر موارد الگوریتم ژنتیک عددی نزدیک تر به شاخص واقعی را نشان می دهد. در دوره پیش بینی سه ماهه در تمامی نمونه ها الگوریتم زنتیک پیش بینی بهتری از روند کوتاه مدت شاخص سهام دارد.

در پیش بینی روند بلند مدت یعنی در دوره پیش بینی شش ماهه به غیر از مورخ ۱۳۹۶/۱۰/۳۰ که الگوریتم ازدحام ذرات عددی نزدیک تر به شاخص واقعی را نشان می دهد در سایر نمونه ها الگوریتم ژنتیک عددی نزدیک تر به شاخص واقعی را نشان می دهد. در دوره پیش بینی یک ساله نیز به غیر از نمونه مورخ ۱۳۹۶/۱۲/۲۸ که الگوریتم ازدحام ذرات عددی نزدیک تر به شاخص واقعی را بدست آورد. در سایر نمونه ها الگوریتم ژنتیک عملکرد بهتری در پیش بینی روند بلند مدت بازده سهام دارد.

جدول ۳- بررسی طول پنجره با استفاده از الگوریتم ذرات و الگوریتم ژنتیک

شاخص پیش بینی شده	شاخص واقعی	تاریخ مورد پیش بینی	دوره پیش بینی	روش بهینه سازی
۷۲۸۱۳	۷۲۶۱۵	۱۳۹۵/۰۳/۳۱	۱ ماه	الگوریتم ژنتیک
۷۲۷۶۳	۷۲۶۱۵	۱۳۹۵/۰۳/۳۱	۱ ماه	الگوریتم ازدحام ذرات
۷۳۷۰۵	۷۶۴۵۱	۱۳۹۵/۰۶/۳۱	۳ ماه	الگوریتم ژنتیک
۷۳۹۷۸	۷۶۴۵۱	۱۳۹۵/۰۶/۳۱	۳ ماه	الگوریتم ازدحام ذرات
۸۰۰۰۴	۸۰۱۲۳	۱۳۹۵/۰۹/۳۰	۶ ماه	الگوریتم ژنتیک
۷۹۵۹۸	۸۰۱۲۳	۱۳۹۵/۰۹/۳۰	۶ ماه	الگوریتم ازدحام ذرات
۹۶۳۱۳	۹۶۲۹۰	۱۳۹۶/۱۲/۲۸	۱ سال	الگوریتم ژنتیک
۹۶۲۹۰	۹۶۲۹۰	۱۳۹۶/۱۲/۲۸	۱ سال	الگوریتم ازدحام ذرات
۸۶۳۸۰	۸۶۴۸۰	۱۳۹۶/۰۷/۳۰	۱ ماه	الگوریتم ژنتیک
۸۶۲۵۳	۸۶۴۸۰	۱۳۹۶/۰۷/۳۰	۱ ماه	الگوریتم ازدحام ذرات
۸۸۸۱۵	۸۸۷۷۵	۱۳۹۶/۰۸/۳۰	۳ ماه	الگوریتم ژنتیک
۸۸۶۱۷	۸۸۷۷۵	۱۳۹۶/۰۸/۳۰	۳ ماه	الگوریتم ازدحام ذرات
۹۸۸۴۷	۹۸۸۱۷	۱۳۹۶/۱۰/۳۰	۶ ماه	الگوریتم ژنتیک
۹۸۸۰۲	۹۸۸۱۷	۱۳۹۶/۱۰/۳۰	۶ ماه	الگوریتم ازدحام ذرات
۱۷۸۸۲۶	۱۷۸۶۵۹	۱۳۹۷/۱۲/۲۸	۱ سال	الگوریتم ژنتیک
۱۷۸۲۹۷	۱۷۸۶۵۹	۱۳۹۷/۱۲/۲۸	۱ سال	الگوریتم ازدحام ذرات
۱۳۶۶۲۷	۱۳۶۹۱۱	۱۳۹۷/۰۵/۳۰	۱ ماه	الگوریتم ژنتیک
۱۳۷۵۷۴	۱۳۶۹۱۱	۱۳۹۷/۰۵/۳۰	۱ ماه	الگوریتم ازدحام ذرات
۱۸۸۹۲۵	۱۸۷۷۷۹	۱۳۹۷/۰۷/۳۰	۳ ماه	الگوریتم ژنتیک
۱۸۳۶۲۸	۱۸۷۷۷۹	۱۳۹۷/۰۷/۳۰	۳ ماه	الگوریتم ازدحام ذرات
۱۰۸۹۰۷	۱۰۸۸۳۱	۱۳۹۷/۰۴/۳۱	۶ ماه	الگوریتم ژنتیک
۱۰۸۷۳۸	۱۰۸۸۳۱	۱۳۹۷/۰۴/۳۱	۶ ماه	الگوریتم ازدحام ذرات
۳۰۱۸۵۷	۳۰۲۱۰۴	۱۳۹۸/۰۶/۳۱	۱ سال	الگوریتم ژنتیک
۳۰۱۳۰۳	۳۰۲۱۰۴	۱۳۹۸/۰۶/۳۱	۱ سال	الگوریتم ازدحام ذرات
۳۵۳۸۰۵	۳۵۳۹۹۷	۱۳۹۸/۰۹/۳۰	۱ ماه	الگوریتم ژنتیک
۳۵۵۴۵۳	۳۵۳۹۹۷	۱۳۹۸/۰۹/۳۰	۱ ماه	الگوریتم ازدحام ذرات
۲۱۵۶۱۳	۲۱۵۰۹۲	۱۳۹۸/۰۲/۳۱	۳ ماه	الگوریتم ژنتیک
۲۱۳۲۰۵۴	۲۱۵۰۹۲	۱۳۹۸/۰۲/۳۱	۳ ماه	الگوریتم ازدحام ذرات
۲۶۶۱۱۳	۲۶۶۱۲۷	۱۳۹۸/۰۵/۳۰	۶ ماه	الگوریتم ژنتیک
۲۶۵۹۳۹	۲۶۶۱۲۷	۱۳۹۸/۰۵/۳۰	۶ ماه	الگوریتم ازدحام ذرات
۵۱۳۰۱۸	۵۱۲۹۰۱	۱۳۹۸/۱۲/۲۸	۱ سال	الگوریتم ژنتیک
۵۱۱۸۷۷	۵۱۲۹۰۱	۱۳۹۸/۱۲/۲۸	۱ سال	الگوریتم ازدحام ذرات
۶۰۲۰۴۱	۶۹۰۰۳۷	۱۳۹۹/۰۱/۳۱	۱ ماه	الگوریتم ژنتیک

شاخص پیش بینی شده	شاخص واقعی	تاریخ مورد پیش بینی	دوره پیش بینی	روش بهینه سازی
۶۹۴۹۳۰	۶۹۰۰۳۷	۱۳۹۹/۰۱/۳۱	۱ ماه	الگوریتم ازدحام ذرات
۱۱۶۷۶۱۲	۱۱۵۰۷۱۸	۱۳۹۹/۱۰/۳۰	۳ ماه	الگوریتم ژنتیک
۱۱۸۰۲۳۵	۱۱۵۰۷۱۸	۱۳۹۹/۱۰/۳۰	۳ ماه	الگوریتم ازدحام ذرات
۱۲۰۴۱۳۹	۱۲۳۸۳۵۷	۱۳۹۹/۱۱/۲۹	۶ ماه	الگوریتم ژنتیک
۱۱۸۲۶۰۲	۱۲۳۸۳۵۷	۱۳۹۹/۱۱/۲۹	۶ ماه	الگوریتم ازدحام ذرات
۱۲۹۲۳۶۲	۱۲۹۰۴۷	۱۳۹۹/۱۲/۲۶	۱ سال	الگوریتم ژنتیک
۱۲۸۷۵۲۷	۱۲۹۰۴۷	۱۳۹۹/۱۲/۲۶	۱ سال	الگوریتم ازدحام ذرات

منبع: یافته های پژوهشگر

۵- مقایسه قدرت پیش بینی مدل ها

در این مقاله مقدار شاخص بورس اوراق بهادار تهران طی سالهای ۱۳۹۵ تا ۱۴۰۰ پس ازهموار سازی داده ها با استفاده از الگوریتم ژنتیک و الگوریتم ذرات در قالب برنامه پایتون و با تعیین طول پنجره بهینه مورد آزمون قرار گرفت و نتایج آن ها با یکدیگر در دوره های کوتاه مدت و بلند مدت مقایسه شد اکنون لازم است تا دقت پیش بینی مدل ها در قالب معیار MSFE مقایسه شود، در ادبیات اقتصادی متداولترین معیار برای ارزیابی آماری دقت پیش بینی بین مدل ها، **معیار** MSFE [1]است که به طور گسترده ای در ادبیات پیش بینی بازده سهام استفاده شده است و به صورت زیر محاسبه می شود:

$$\mathrm{MSFE} = \frac{1}{\mathrm{p}}\sum_{\mathrm{j}=1}^{\mathrm{p}}\left(\mathrm{r}_{\mathrm{T}_{0+\mathrm{j}}} - \hat{\mathrm{r}}_{\mathrm{T}_{0+\mathrm{j}}}\right)^2$$ فرمول شماره(۹)

که در آن به تبعیت از رپچ و همکاران (۲۰۱۰) و ولش و گویال (۲۰۰۸) از مدل میانگین تاریخی به عنوان مدل معیار یا مبنای مقایسه به صورت زیر استفاده شده است :

فرمول شماره (۱۰)

$$\hat{\mathrm{r}}_{\mathrm{T}_0+1} = \frac{1}{\mathrm{T}_0}\sum_{\mathrm{j}=1}^{\mathrm{T}_0}\mathrm{r}_{\mathrm{j}}$$

که در آن $\hat{\mathbf{r}}_{\mathbf{T_{0+j}}}$ بازده تخمینی، $\mathbf{r}_{\mathbf{T_{0+j}}}$ بازده واقعی $\boldsymbol{T_0}$ و **P** تعداد مشاهدات در نمونه و خارج از نمونه است. برای تابع خطا از MSFE استفاده شده که دو پارامتر ستون داده های واقعی و ستون داده های پیش بینی شده به آن داده شده است، کار تابع این است که در هر نسل این مقدار را بدست آورده و به عنوان MSFE ذخیره می کند تا به هدف خود یعنی حداقل نمودن MSFE دست یابد.

[1] Master of Science in Finance and Economics

جدول ۴- مقدار MSFE برای نتایج کلی هر الگوریتم :

MSFE Ga 30	۱۵۲۵۶۰۳۶۸۵
MSFE Ga 90	۱۱۶۲۴۰۱۳۹۴
MSFE Ga 180	۴۲۹۳۹۹۵۷۱
MSFE Ga 364	۶۴۶۶۰۶
MSFE PSO 30	۱۶۴۶۶۷۲۱۵۱
MSFE PSO 90	۱۵۴۴۰۴۳۱۵۷
MSFE PSO 180	۸۷۲۰۰۳۳۳۱
MSFE PSO 364	۱۴۰۰۶۸۹۵

منبع: یافته های پژوهشگر

جدول ۴ بیانگر میزان خطای هر الگوریتم در پیش بینی شاخص سهام با استفاده از معیار MSFE می باشد. همانطور که در جدول دیده می شود مقدار این شاخص برای الگوریتم ژنتیک در دوره کوتاه مدت یک ماهه برابر با ۱۵۲۵۶۰۳۶۸۵ و در یک دوره بلند مدت یکساله برابر با ۶۴۶۶۰۶ می باشد. به همین صورت نتایج برای الگوریت ذرات دریک دوره کوتاه مدت یک ماهه برابر با ۱۶۴۶۶۷۲۱۵۱ و برای یک دوره بلند مدت یکساله برابر با ۱۴۰۰۶۸۹۵ بدست آمده است.

همانطور که در جدول دیده می شود در دوره یک ماهه و سه ماهه یعنی در کوتاه مدت مقدار MSFE بدست آمده برای الگوریتم ژنتیک کمتر از الگوریتم ازدحام ذرات است. مقدار MSFE بدست آمده دوره شش ماهه و یک ساله نیز برای الگوریتم ژنتیک کمتر از الگوریتم ازدحام ذرات است.

۵-نتیجه گیری

پیش بینی روند بازار از جمله پژوهشهای کاربردی در حوزه بازارهای مالی است. از آنجایی که مدیریت سبد سرمایه گذاری در جهت کسب بیشترین بازده یک فرآیند پیچیده است و اغلب از ساختار خاصی تبعیت نمی‌کند توانایی پیش بینی روند بلند مدت و کوتاه مدت بازده می تواند تا حد زیادی سرمایه گذاران را در دستیابی به این هدف یاری دهد.

از اینرو در این پژوهش تلاش شد تا با استفاده از الگوریتم ژنتیک و الگوریتم ذرات در قالب برنامه پایتون به پیش بینی روند بلند مدت و کوتاه مدت بازده سهام پرداخته شود. برای انجام این کار از داده های شاخص بورس اوراق بهادار تهران طی سالهای ۱۳۹۵ تا ۱۴۰۰ پس ازهموار سازی داده ها برای روز های تعطیل استفاده شد که این روش با پژوهش های فرشادفر و پروکوپچوک (۲۰۱۹) همسویی دارد.

پس از آن برای افزایش دقت پیش بینی شبکه طول پنجره بهینه با استفاده از روش های الگوریتم ژنتیک و ازدحام ذرات بدست آمد و در شبکه مورد استفاده قرار گرفت نتایج گویای آن بود که مقادیر نمونه پیش بینی شده

توسط الگوریتم ژنتیک نسبت به الگوریتم ذرات دفعات بیشتری به مقدار شاخص واقعی نزدیک تر بوده است. در ارزیابی آماری دقت پیش بینی نیز مقایسه نتایج پیش بینی طی ۵ سال، عملکرد بهتر الگوریتم ژنتیک نسبت به الگوریتم ذرات را نشان داده است. بنابراین با توجه به نتایج به دست آمده فرضیه های پژوهش مبنی بر عملکرد بهتر الگوریتم ژنتیک نسبت به الگوریتم ذرات در پیش بین بازده سهام در دوره های کوتاه مدت یعنی دوره ی یک ماهه و سه ماه تایید می شود. هم چنین بر اساس نتایج بدست آمده عملکرد بهتر الگوریتم ژنتیک نسبت به الگوریتم ذرات در پیش بینی بلند مدت بازده سهام رد نمی شود. این نتایج با یافته های آبراهام و همکاران (۲۰۲۲) و کی شارما و همکاران (۲۰۲۲) همسویی دارد.

از اینرو به سیاست گزاران پیشنهاد می شود تا برای بررسی دلایل اقتصادی موثر بر رشد ناگهانی بورس تهران در دو دوره های مختلف همجنین بررسی و تفاوت وضعیت اقتصادی ایران در زمان رشد و کود بی سابقه بازار بورس تهران از این الگوریتم استفاده نمایند.

به پژوهشگران نیز پیشنهاد می شود تا به مطالعات تکمیلی در زمینه پیش بینی میزان بازده صنایع مختلف بازار بورس تهران به صورت جداگانه با استفاده از سایر الگوریتم های هوش مصنوعی که کمتر مورد استفاده قرار گرفته اند اهتمام ورزند.

فهرست منابع

امام وردی، قدرت اله، صفرزاده بیجاربنه، سمانه. (۱۳۹۴). ازمون اشوبی و غیر خطی بودن شاخص قیمت سهام در بورس تهران. اقتصاد مالی، ۹ (۳۳)،۵۵ – ۷۴.

بابانژاد باقری، سیده مریم، پورآقاجان، عباسعلی، عباسیان، محمد مهدی. (۱۴۰۲). پیش بینی ارزش شرکت مبتنی بر روش‌های یادگیری عمیق. فصلنامه اقتصاد مالی، ۱۷(۶۴)، ۲۹۱-۳۱۸.

رجبی، رضا، حاجی یخچالی، سیامک. (۱۴۰۱). بهینه سازی جریان نقدینگی سبد پروژه با درنظر گرفتن شاخص های بازار با استفاده از الگوریتم های ژنتیک و ازدحام ذرات. مهندسی سازه و ساخت، ۹(۹)، ۱۰۳-۱۲۰.

علی بابایی، غزاله، خان محمدی، محمد حامد. (۱۴۰۲). ارائه مدلی به منظور پیش بینی ورشکستگی شرکت ها با استفاده از مکانیزم های حاکمیت شرکتی و نسبت های مالی. دانش سرمایه‌گذاری، ۱۲(۴۸)، ۶۷-۹۸.

رنجبر ناوی، رستم، ارشدی، علی، چناری، حسن. (۱۴۰۰). پیش بینی شاخص کل قیمت سهام با استفاده از الگوی خاکستری، بورس اوراق بهادار تهران،۱۴ (۵۳)، ۱۱۵- ۱۳۷.

مصطفائی درمیان، سبحان، دعائی، میثم. (۱۴۰۰). ارائه رویکردی مبتنی بر بهینه سازی تصادفی به منظور حل مساله انتخاب سبد سهام در بازار سرمایه ایران با استفاده از الگوریتم های فراابتکاری. نظریه های کاربردی اقتصاد، ۸ (۴)، ۲۵۳-۲۸۴.

نور احمدی، مرضیه، صادقی، حجت الله. (۱۴۰۲). کاربرد شبکه‌های فیلترشده برمبنای آستانه در انتخاب سبد سهام و ارزیابی عملکرد آن. فصلنامه اقتصاد مالی، ۱۷(۶۴)، ۱-۲۶.

مرادی، فریدون، یعقوب نژاد، احمد، جهانشاد، آزیتا، (۱۴۰۲). کاربرد الگوریتم های فراابتکاری در پیش بینی درماندگی مالی با استفاده از متغیرهای مالی و غیرمالی درون شرکتی و اقتصادی (الگوریتم های بهینه سازی مـلخ و کلونی مورچگان). فصلنامه اقتصاد مالی ،۱۷(۶۴)، ۷۱-۱۰۴.

مهرانی، ساسان، رحیمی پور، اکبر. (۱۴۰۱). پیش بینی احتمال تجدید ارائه صورتهای مالی با استفاده از مدل بنیش و بهبود مدل از طریق رگرسیون لاجیت و الگوریتم ژنتیک. نشریه پژوهش های حسابداری مالی، ۱۴(۲) ، ۹۱-۱۱۶.

یوخنه القیانی، ماریام، بحری ثالث، جلال، جبارزاده کنگلوئی، سعید، زواری رضایی، اکبر. (۱۴۰۰). تبیین گزارشگری مالی-مالیاتی متقلبانه شرکت ها: رویکرد ترکیبی داده کاوی کلاسیک ANFIS و الگوریتم های فراابتکاری. مطالعات تجربی حسابداری مالی، ۱۸ (۷۱)، ۸۹-۱۱۱.

Abraham R, Samad ME, Bakhach AM, El-Chaarani, H., Sardouk, A., Nemar, SE., Jaber, D., (2022). Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. Journal of Risk and Financial Management.;15(5):188.
https://doi.org/10.3390/jrfm15050188

Bollerslev, T., Marrone, J., Xu, L., Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence. Journal of Financial and Quantitative Analysis, 49 (03), 633–661.

Chandar, S.K. (2019) .Stock market prediction using subtractive clustering for a neuro fuzzy hybrid approach. Cluster Comput, 22 (6), 13159–13166.

Duy Dao, Son, Abhary, Kazem, Marian, Romeo. (2017). An innovative framework for designing genetic algorithm structures, Expert Systems with Applications, 90(30),196-208.

Farshadfar, Z., Prokopczuk, M., Nonlinear Model Improves Stock Return Out of Sample
Forecasting (Case Study: United State Stock Market), (2019). International Journal of Finance and Managerial Accounting, 3(12), 1-13.

Farshadfar, Z., Prokopczuk, M., (2019). Improving Stock Return Forecasting by Deep Learning Algorithm, Advances in mathematical finance& applications, 4 (3), 1-13.
DOI: 10.22034/amfa.2019.584494.1173

Ge lliu, Wenping Ma (2022). A quantum artificial neural network for stock closing price prediction. Information Sciences, Vol 598, Pages 75-85.

Jasemi, M., Kimiagari, A., Memariani, A. (2011). A conceptual model for portfolio management sensitive to mass psychology of market. International Journal of Industrial Engineering Theory Application and Practice, 18 (1), 1–15.

Kumar, G., Singh, U.P, Jain, S. (2022). An adaptive particle swarm optimization-based hybrid long short-term memory model for stock price time series forecasting. Soft Comput 26, 12115–12135.
https://doi.org/10.1007/s00500-022-07451-8

Phan, D., Sharma, S., Narayan, P. (2015). Stock return forecasting: Some new evidence. International Review of Financial Analysis, 40, 38–51. E.

Sharma, D.K., Hota, H.S., Brown, K. (2022). Integration of genetic algorithm with artificial neural network for stock market forecasting. International journal of system, 13, 828–841. https://doi.org/10.1007/s13198-021-.1209-5

Financial Economics
Vol. (18) Issue (68) Autumn 2024

Abstract

Comparative Application of Particle Algorithm and Genetic Algorithm in Predicting Long-Term and Short-Term Trend of Stock Returns

Seyed Javad Kiaei[1]
Zahra Farshadfar[2]

Receive: 04/July/2024 Acceptance: 11/ September/2024

Abstract

The lack of certainty in the movement of the stock market has made forecasting a challenging task in the field of financial time series forecasting. On the other hand, it is not easy to analyze the time series data of stock prices due to non-linearity and high noise. Therefore, the aim of this research is to predict the long-term and short-term trend of the capital market. To achieve this goal, artificial intelligence algorithms of particles and genetics have been used in a comparative manner. The studied variable is the total stock price index in Tehran Stock Exchange in the period of 2016 to 2021 and on a monthly basis.

The data have been reviewed after collection using the smoothing method for holidays, and in order to increase the accuracy of the models, the optimal window length of each algorithm has been calculated. The findings indicate that the genetic algorithm by minimizing the prediction error is a suitable algorithm for predicting the short-term and long-term trend of the total price index compared to the particle algorithm in the studied time period.

Keywords: Particle Algorithm, Genetic Algorithm, Data leveling, Machine learning, Artificial Intelligence
JEL Classification: G17, G32, Z19

[1] Department of Accounting, West Tehran Branch, Islamic Azad University, Tehran, Iran. sjavadkiaei7@gmail.com
[2] Department of Accounting, West Tehran Branch, Islamic Azad University, Tehran, Iran. (Corresponding author) zfarshadfar@yahoo.com

Creative Commons – Attribution 4.0 International – CC BY 4.0
Creativecommons.org

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی