

طراحی سیستم معاملات الگوریتمی ترکیبی بهینه شده با الگوریتم ژنتیک

نوع مقاله: پژوهشی

شیوا قاسم پور^۱

امیررضا کیقبادی^۲

شادی شاهوردیانی^۳

مهدی معدنچی زاج^۴

تاریخ پذیرش: ۱۴۰۴/۹/۲۵

تاریخ دریافت: ۱۴۰۴/۸/۴

چکیده

در سال‌های اخیر، گسترش استفاده از سیستم‌های معاملات الگوریتمی در بازارهای مالی جهان و به‌تازگی در بازار سرمایه ایران، منجر به افزایش نیاز به طراحی و توسعه سامانه‌های هوشمند معاملاتی شده است. مزیت اصلی این سیستم‌ها، افزایش سرعت و دقت تصمیم‌گیری و حذف مؤلفه‌های احساسی در معاملات است. پژوهش حاضر با هدف طراحی یک سیستم معاملات الگوریتمی ترکیبی بهینه، مبتنی بر الگوریتم ژنتیک انجام شده است. در این تحقیق از رویکرد یادگیری جمعی با ترکیب چهار الگوریتم درخت تصمیم (DTree)، K نزدیک‌ترین همسایگی (KNN)، $RUSBoost$ و شبکه عصبی پایه شعاعی (RBN) استفاده شده است تا به‌صورت هم‌زمان انتخاب ویژگی بهینه، تنظیم پارامترهای الگوریتم‌ها و تعیین وزن هر یک انجام گیرد. داده‌های تحقیق شامل اطلاعات قیمتی و اندیکاتورهای تکنیکال سهام منتخب بورس اوراق بهادار تهران است. الگوریتم ژنتیک به‌منظور بهینه‌سازی ۵۱ متغیر تصمیم شامل پارامترهای الگوریتم‌ها، ویژگی‌ها و وزن مدل‌ها به‌کار گرفته شد. نتایج نشان داد مدل ترکیبی پیشنهادی دارای دقت طبقه‌بندی بالا در تشخیص سه کلاس خرید، فروش و نگهداری بوده و نسبت به روش‌های منفرد عملکرد بهتری دارد. بنابراین، مدل ارائه‌شده می‌تواند به عنوان یک سیستم معاملات خودکار بهینه در بازار سرمایه ایران مورد استفاده قرار گیرد.

واژه‌های کلیدی: معاملات الگوریتمی، الگوریتم ژنتیک، یادگیری جمعی، بهینه‌سازی، شبکه‌های عصبی طبقه‌بندی JEL: G17, C61, C63

۱ گروه مدیریت مالی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران
Shiva.ghasempour@iau.ac.ir

۲ گروه مدیریت صنعتی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول)
ar.keyghobadi@iau.ac.ir

۳ گروه مدیریت بازرگانی، واحد شهرقدس، دانشگاه آزاد اسلامی، شهرقدس، ایران
shadishahverdiani@iau.ac.ir

۴ گروه مدیریت مالی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران
ma.madanchi@iau.ac.ir

مقدمه

در حال حاضر، بخش وسیعی از معاملات در بازارهای جهانی و اخیراً در بازار بورس اوراق بهادار تهران توسط سیستم‌های معاملات الگوریتمی و بصورت خودکار انجام می‌شود. رشد حضور چنین سیستم‌های معاملات الگوریتمی خودکار و ارائه الگوریتم‌های جدید با عملکردهای مختلف، تقاضای نهادهای مختلف برای امکان‌سنجی، طراحی و توسعه سیستم‌های معاملاتی خودکار در کشورهای مختلف و اخیراً ایران را افزایش داده است. مزیت بزرگی که چنین سیستم‌هایی برای سرمایه‌گذاران دارند، اینست که باعث افزایش سرعت و دقت آنالیز اطلاعات و تصمیم‌گیری معاملاتی بدون دخالت احساسات می‌شود. به منظور طراحی و استقرار سیستم‌های معاملات الگوریتمی، تکنیک‌های مختلفی بر پایه تحلیل تکنیکال، بنیادی و تحلیل سری‌های زمانی استفاده شده است. این تکنیک‌های مختلف را مطابق با نظر دینگ و کین^۱ (۲۰۲۰) می‌توان به دسته کلان ۱- آماری و ۲- هوش مصنوعی تقسیم کرد. روش‌های آماری شامل مدل‌هایی مانند رگرسیون لجستیک و مدل ARCH است. روش‌های هوش مصنوعی نیز شامل شبکه‌های عصبی مانند پرسپترون^۲، شبکه عصبی بازگشتی^۳، شبکه ناویه - بیز^۴، شبکه پس‌انتشار^۵، شبکه حافظه کوتاه‌مدت - بلندمدت^۶ (LSTM)، ماشین بردار پشتیبان^۷ (SVM) و غیره است. در این میان شبکه‌های عصبی کاربرد گسترده‌ای داشته‌اند. شبکه‌های عصبی توانایی یادگیری با استفاده از روند تغییرات داده‌ها (داده‌های آموزش^۸) و قوانین حاکم بر آن را داشته و سپس بر روی داده‌های دیگر که از قبل ندیده‌اند (داده‌های تست^۹) آنرا تعمیم دهند (مقدم و همکاران^{۱۰}، ۲۰۱۶).

علاوه بر این، عمده سیستم‌های معاملات الگوریتمی پیشنهادی توسط محققین مختلف بر اساس یک یا چند روش مستقل و یا ترکیبی از دو یا چند الگوریتم بوده است که به منظور مقایسه با الگوریتم اصلی پیشنهادی مورد استفاده قرار گرفته‌اند. در تحقیق حاضر بر اساس رویکرد یادگیری جمعی^{۱۱} یک رویکرد ترکیبی مبتنی بر الگوریتم درخت تصمیم (DTree)، K - نزدیکترین همسایگی (KNN)، روش RUSBoost و شبکه عصبی پایه شعاعی^{۱۲} (RBN) پیشنهاد شده است که همزمان به انتخاب

1 Ding & Qin

2 Perceptron

3 Recurrent Neural Network

4 Naive Bayes Network

5 Back Propagation Network

6 Long-Short Term Memory

7 Support Vector Machine

8 Training data

9 Test data

10 Moghaddam

11 Ensemble learning

12 Radial basis network

ویژگی بهینه و بهینه‌سازی پارامترهای الگوریتم‌های دسته‌بندی و وزن هر کدام از آنها منجر می‌شود. به منظور بهینه‌سازی نیز با توجه قدرت بالایی که الگوریتم‌های فراابتکاری و بویژه الگوریتم ژنتیک دارد، در این تحقیق از این الگوریتم برای بهینه‌سازی استفاده می‌شود.

۱- مبانی نظری و مروری بر پیشینه پژوهش

۱-۲- معاملات الگوریتمی

در گذشته معامله‌گران بر اساس تجربه و دانش خود به ایجاد استراتژی‌های معاملاتی در بازارهای مالی می‌پرداختند. در حال حاضر نیز بسیاری از معامله‌گران بر این اساس اقدام به معامله می‌کنند. این روش دارای عیب‌هایی از جمله زمانبر بودن جمع‌آوری داده‌ها و اطلاعات مورد نیاز برای تحلیل، زمانبر بودن فرآیند تحلیل و پردازش اطلاعات، تکیه بر دانش و تجربه فردی معامله‌گر، انجام مکرر آزمون و خطا برای کسب تجربه و عیب‌یابی، محدود بودن تعداد پارامترهایی که همزمان قادر به بررسی هستند، می‌باشد (گامبر و زیرمرمن^۱، ۲۰۱۸). اجرای معامله با سازوکارهای خرید یا فروش حجم مشخصی از یک سهم در ارتباط است. ثبت سفارش‌های بزرگ می‌تواند به شدت بر نوسان‌ها و تغییرهای ناگهانی قیمت اثرگذار باشد. از این رو، استراتژی‌های اجرای معاملات برای برقراری تعادل بین این هزینه‌ها و در نتیجه به حداقل رساندن هزینه کل معاملات، طراحی شده‌اند. در واقع موضوع عملی که برای سرمایه‌گذاران، نهادهای ناظر بازار، قانون‌گذاران و بازارگردانان مطرح می‌شود، این است که چگونه می‌توانند حجم زیادی از یک سهم را به‌صورت بهینه خرید و فروش کنند، به‌طوری که اثر قیمتی سفارش را به حداقل برسانند (یافتیان و رستگار، ۱۳۹۷). همه این موارد باعث شد که روش‌های سنتی برای تحلیل و انجام معامله، قادر به شناسایی کامل همه فرصت‌های بازار نشوند. بنابراین نیاز به یک ابزار خودکار برای انجام معامله است و در این زمینه، می‌توان به الگوریتم‌هایی اتکاء کرد که بر اساس رویکردهای مختلف از جمله تحلیل‌های تکنیکال، قادر به انجام معاملات خودکار هستند و از آنها به عنوان معاملات الگوریتمی یاد می‌شود (یان^۲، ۲۰۱۹).

الگوریتم‌های معاملاتی مبتنی بر پارامترهای سفارش مانند زمان، قیمت و یا حجم سفارش هستند. در بازارهای مالی الکترونیکی، معاملات الگوریتمی به معنای استفاده از برنامه‌های کامپیوتری برای ورود سفارش‌های معاملاتی است که سیگنال‌های معاملاتی آن توسط بخشی دیگر از سیستم تولید شده است (سیفرت^۳، ۲۰۱۸). در تعریفی دیگر، معاملات الگوریتمی را نوعی معاملات الکترونیکی می‌نامند که پارامترهای آن از قواعد از پیش تعریف شده‌ای پیروی کرده و به دنبال رسیدن به نتایج

1 Gomber & Zimmermann

2 Yan

3 Seyfert

مشخصی است. برخی از این پارمترها حجم، قیمت، ابزار معاملاتی، زمان و اخبار هستند (نوتی^۱ و همکاران، ۲۰۱۱).

۲-۲- دسته‌بندی معاملات الگوریتمی

الگوریتم‌های معاملاتی را می‌توان به سه دسته الگوریتم‌های اجرا^۲، معاملات کسب سود^۳ و الگوریتم‌های معاملاتی پرسامد^۴ تقسیم کرد (کیسل^۵، ۲۰۲۱):

- الگوریتم‌های اجرا: الگوریتم اجرا وظیفه اجرای تصمیم سرمایه‌گذاری اتخاذشده توسط سرمایه‌گذاران یا مدیریت سبد سرمایه‌گذاری را بر عهده دارد. سرمایه‌گذار بر اساس سبک و هدف سرمایه‌گذاری تعیین می‌کند که چه چیزی را بخرد یا بفروشد و سپس ترتیب را در الگوریتم وارد می‌کند. س الگوریتم دستور را اجرا می‌کند و تصمیم را بر اساس مجموعه‌ای از قوانین تنظیم‌شده توسط سرمایه‌گذار یا مدیریت سبد سرمایه‌گذاری اجرا می‌کند.
- الگوریتم‌های جستجوی سود: الگوریتمی است که هم خرید و یا فروش در بازار را تعیین می‌کند و هم آن تصمیمات را بدون تعامل با سرمایه‌گذار انجام می‌دهد. به‌عنوان مثال، این الگوریتم‌ها از اطلاعات قیمت واقعی و داده‌های بازار مانند قیمت‌ها، حجم، نوسانات و کمسیون قیمت‌ها برای تعیین خرید و فروش استفاده خواهند کرد و در صورت مساعد بودن شرایط برای سرمایه‌گذار، معامله را انجام می‌دهند. الگوریتم جستجوی سود بر اساس مدل‌ها کمی، قیمت‌گذاری اشتباه در بازار، یا یک استراتژی مشخص نسبت به انجام معامله و کسب سود اقدام می‌کند.
- معاملات پرسامد: معاملات پرسامد (HFT) نوعی از الگوریتم جستجوی سود است که به جستجوی سود معاملاتی کوتاه‌مدت است. غالباً دوره زمانی نگهداری برای معامله HFT تنها برای چند ثانیه یا کمتر است و معاملات HFT به‌واسطه افزایش شدید ریسک تمایلی به نگهداری موقعیت معامله ندارند. شاخصه الگوریتم‌های HFT با نرخ گردش مالی بسیار بالا است و با بهره‌گیری از قیمت‌گذاری نادرست بازار و نقدینگی در بازارهای مختلف به دنبال سودآوری هستند. علاوه بر این، الگوریتم‌های HFT در تلاش برای کشف اهداف خرید و/یا فروش سرمایه‌گذاری بلندمدت از طریق پردازش داده‌های بازار، قیمت‌ها، و

Nuti^۱

Execution algorithm^۲

Profit seeking algorithm^۳

High frequency algorithm^۴

Kissell^۵

مظنه‌ها و سپس استفاده از این اطلاعات به نفع خود برای کسب سود هستند. تمامی الگوریتم‌های HFT الگوریتم‌های جستجوی سود هستند، اما همه الگوریتم‌های جستجوی سود، HFT نیستند.

۳-۲- الگوریتم ژنتیک

الگوریتم ژنتیک (GA) ابزاری هست که توسط آن ماشین می‌تواند مکانیسم انتخاب طبیعی را شبیه‌سازی کند. این عمل با جستجو در فضای مسئله جهت یافتن جواب برتر و نه الزاماً بهینه صورت می‌پذیرد. الگوریتم ژنتیک را می‌توان یک روش جستجوی کلی قلمداد کرد که از قوانین تکامل بیولوژیک طبیعی تقلید می‌کند؛ درواقع الگوریتم‌های ژنتیک از اصول انتخاب طبیعی داروین برای یافتن فرمول بهینه جهت پیش‌بینی یا تطبیق الگو استفاده می‌کنند. الگوریتم‌های ژنتیک اغلب گزینه خوبی برای تکنیک‌های پیش‌بینی بر مبنای رگرسیون هستند (دیویس^۱، ۱۹۹۱).

الگوریتم ژنتیک به‌عنوان یک الگوریتم محاسباتی بهینه‌سازی با در نظر گرفتن مجموعه‌ای از نقاط فضای جواب در هر تکرار محاسباتی به نحو مؤثری نواحی مختلف فضای جواب را جستجو می‌کند. در مکانیسم جستجو گرچه مقدار تابع هدف تمام فضای جواب محاسبه نمی‌شود ولی مقدار محاسبه‌شده تابع هدف برای هر نقطه، در متوسط‌گیری آماری تابع هدف برای هر نقطه، در متوسط‌گیری آماری تابع هدف در کلیه زیر فضاهایی که آن نقطه به آن‌ها وابسته بوده دخالت داده می‌شود و این زیر فضاها به‌طور موازی از نظر تابع هدف متوسط‌گیری آماری می‌شوند. این مکانیسم را توازی ضمنی^۲ می‌گویند. این روند باعث می‌شود که جستجوی فضا به نواحی از آن که متوسط آماری تابع هدف در آن‌ها زیاد بوده و امکان وجود نقطه بهینه مطلق در آن‌ها بیشتر است سوق پیدا کند. چون در این روش برخلاف روش‌های تک‌مسیری فضای جواب به‌طور همه‌جانبه جستجو می‌شود، امکان کمتری برای همگرایی به یک نقطه بهینه محلی وجود خواهد داشت (میرجلیلی^۳، ۲۰۱۹).

در الگوریتم ژنتیک، مجموعه‌ای از متغیرهای طراحی را توسط رشته‌هایی با طول ثابت^۴ یا متغیر^۵ کدگذاری می‌کنند که در سیستم‌های بیولوژیکی آن‌ها را کروموزوم یا فرد^۶ می‌نامند. هر رشته یا کروموزوم یک نقطه پاسخ در فضای جستجو را نشان می‌دهد. به ساختمان رشته‌ها یعنی مجموعه‌ای

1 Davis

2 Implicit Parallelism

3 Mirjalili

4 Fixed Length

5 variable

6 Individual

از پارامترها که توسط یک کروموزوم خاص نمایش داده می‌شود ژنوتیپ^۱ و به مقدار رمزگشایی آن فنوتیپ^۲ می‌گویند. الگوریتم‌های وراثتی فرآیندهای تکراری هستند، که هر مرحله تکراری را نسل و مجموعه‌هایی از پاسخ‌ها در هر نسل را جمعیت نامیده‌اند (میرجلیلی، ۲۰۱۹).

الگوریتم‌های ژنتیک، جستجوی اصلی را در فضای پاسخ به اجرا می‌گذارند. این الگوریتم‌ها با تولید نسل^۳ آغاز می‌شوند که وظیفه ایجاد مجموعه نقاط جستجوی اولیه به نام «جمعیت اولیه»^۴ را بر عهده دارند و به‌طور انتخابی یا تصادفی تعیین می‌شوند. از آنجایی که الگوریتم‌های ژنتیک برای هدایت عملیات جستجو به‌طرف نقطه بهینه از روش‌های آماری استفاده می‌کنند، در فرآیندی که به انتخاب طبیعی وابسته است، جمعیت موجود به تناسب برازندگی افراد آن برای نسل بعد انتخاب می‌شود. سپس عملگرهای ژنتیکی شامل انتخاب^۵، پیوند (ترکیب)، جهش و دیگر عملگرهای احتمالی اعمال شده و جمعیت جدید به وجود می‌آید. پس از آن جمعیت جدیدی جایگزین جمعیت پیشین می‌شود و این چرخه ادامه می‌یابد (میرجلیلی، ۲۰۱۹).

۳-۲- پیشینه تحقیق

تحقیقات خارجی اخیر در زمینه معاملات الگوریتمی را می‌توان به تحقیقات مبتنی بر شبکه‌های عصبی غیرعمیق و عمیق تقسیم کرد. در زمینه شبکه‌های عصبی عمیق، شبکه LSTM کاربرد گسترده‌ای داشته است. لی و همکاران^۶ (۲۰۱۷) از ترکیب مدل LSTM با روش بیزین به منظور استخراج عوامل احساسی بازار استفاده کرده است تا قدرت پیش‌بینی مدل LSTM را افزایش دهد. بیک و کیم^۸ (۲۰۱۸) روش جدیدی برای پیش‌بینی شاخص بورس سهام ارائه کرده‌اند که شامل دو ماژول است: ماژول LSTM پیشگیری از بیش‌برازش و ماژول LSTM پیش‌بینی. چانگ و شین^۹ (۲۰۱۸) برای پیش‌بینی قیمت سهام از یک روش ترکیبی LSTM و الگوریتم ژنتیک (GA) استفاده کرده است. لی و همکاران^{۱۰} (۲۰۱۹) یک مدل LSTM مبتنی بر توجه تکاملی با جستجوی تصادفی رقابتی برای پیش‌بینی سری زمانی چندمتغیره ارائه کرده‌اند.

-
- 1 Genotype
 - 2 Phenotype
 - 3 Seeding
 - 4 Initial Population
 - 5 Selection
 - 6 Mirjalili
 - 7 Li
 - 8 Baek & Kim
 - 9 Chung & Shin
 - 10 Li

ژائو و همکاران^۱ (۲۰۱۷) با استفاده از شبکه LSTM وزن متغیر با زمان اقدام به تخمین قیمتی سهام پرداختند. نتایج آنها نشان داد که رابطه بین اهمیت داده‌ها و سری‌های زمانی آنها ثابت نبوده و به صورت خطی و درجه دوم و نسبتاً شبه خطی می‌باشد.

گودلک و همکاران^۲ (۲۰۱۷) در تحقیقی با بهره‌گیری از شبکه عصبی CNN و استفاده از تکنیک ساخت تصویر بر اساس اندیکاتورهای یک سیستم معاملات الگوریتمی طراحی کرده‌اند. نتایج این روش نشان داد که سوددهی آن نسبت به روش خرید و نگهداری بیشتر است. سزار و اوزبایوقلو^۳ (۲۰۱۸) نیز در تحقیقی مشابه، از ۱۵ اندکاتور تحلیل تکنیکال و داده‌های قیمت سهام‌های داوجونز و تعدادی ETF برای ساخت تصاویر استفاده کردند. نتایج این تحقیق نیز حاکی از برتری روش پیشنهادی به عنوان یک سیستم معاملات الگوریتمی است.

ون و یوان^۴ (۲۰۱۸) در تحقیقی از یک شبکه ترکیبی LSTM-CNN برای طراحی سیستم معاملات الگوریتمی استفاده کرده‌اند. نتایج این تحقیق نشان داد که عملکرد این روش با در نظر گرفتن معیار سنجش F1-Score نسبت به تک‌تک شبکه‌های CNN و LSTM بهتر است.

لی و همکاران^۵ (۲۰۱۹) یک روش جدید معاملاتی بر اساس یادگیری تقویتی عمیق ارائه کرده‌اند که در آن از شبکه Q عمیق مبتنی بر ارزش^۶ (DQN) و شبکه LSTM استفاده شده است. نتایج تجربی این روش نشان داد که منجر به بازدهی تنظیم شده توسط ریسک^۷ در بازار سهام و بازار آتی می‌شود.

روندو^۸ (۲۰۱۹) از شبکه LSTM مبتنی بر یادگیری تقویتی عمیق برای معاملات فرکانس بالا به منظور حداکثر کردن بازدهی استفاده کرده است. نتایج این تحقیق بر روی داده‌های فارکس نسبت ارز EU/USD حاکی از دقت ۸۵ درصدی این روش در تخمین روندهای کوتاه - میان مدتی است.

لی و همکاران^۹ (۲۰۲۰) مدل یادگیری تقویتی عمیق آگاه از ویژگی‌های زمان - محور^{۱۰} (TFJ-DRL) ارائه کرده‌اند که به منظور بهبود یادگیری نمایش سیگنال و اتخاذ تصمیم در معاملات الگوریتمی، مدل یادگیری عمیق را با مدل یادگیری تقویتی یکپارچه ساخته است. این مدل بر روی داده‌های مالی واقعی با روندهای قیمتی مختلف (صعودی، نزولی و خنثی) ارزیابی شده است و نتایج حاکی از شدیداً مقاوم بودن روش و کاربردی بودن جالب توجه آن است.

1 Zhao

2 Gudelek

3 Sezer & Ozbayoglu

4 Wen & Yuan

5 Li

6 Value-based deep Q-Network

7 risk-adjusted return

8 Rundo

9 Lei

10 Time-Driven Feature-Aware Jointly Deep Reinforcement Learning

در داخل کشور نیز تحقیقات مختلفی در این زمینه انجام شده است که در ادامه به اهم این تحقیقات اشاره شده است. شقاقیان فر و شاپور (۱۳۹۵) در تحقیقی به بررسی امکان پیش‌بینی بازدهی سهام در رویدادهای تقویمی و کسب سود با استفاده از یک سیستم معاملات خودکار پرداخته‌اند که در آن با استفاده از جنگل‌های تصادفی بازدهی هر معامله تقویمی در بازار بورس اوراق بهادار تهران پیش‌بینی می‌شود. نتایج حاصل از تحقیق نشان می‌دهد معامله تقویمی بر اساس سیگنال‌های تولید شده توسط یادگیری تجمعی «جنگل‌های تصادفی موزون بر حسب عملکرد اخیر» در مقایسه با استراتژی خرید و نگهداری، معامله فصلی، رگرسیون خطی و جنگل تصادفی ساده عملکرد بهتری دارد.

رستگار و ساعدی‌فر (۱۳۹۶) با استفاده از داده‌های تاریخی سهام در بورس اوراق بهادار تهران، ابتدا احتمال انواع سفارش‌گذاری‌ها شامل سفارش بازار، سفارش در شکاف قیمتی و سفارش محدود را برای سمت خرید و سمت فروش به طور جداگانه محاسبه کرده، سپس استراتژی بهینه معاملاتی را بر اساس معیار قیمت میانگین موزون حجمی (VWAP) بررسی کرده‌اند. نتایج این پژوهش نشان می‌دهد برای هر سفارش بزرگ خرید، استراتژی با استفاده از انواع سفارش می‌تواند بهتر از استراتژی‌هایی با استفاده از تنها یک نوع سفارش باشد.

یافتیان و رستگار (۱۳۹۷) یک سیستم معاملاتی خودکار با استفاده از پردازش تصویر به وسیله شبکه عصبی پیچشی (CNN) ارائه کرده‌اند. بدین منظور، ۲۸ اندیکاتور تحلیل تکنیکال انتخابی به تصاویر ۲ بعدی تبدیل شده و در نتیجه برای هر داده روی سری زمانی قیمت سهم، یک تصویر دو بعدی با ابعاد ۲۸×۲۸ ساخته شده است. نتایج پژوهش نشان می‌دهد که در ۸۰٪ موارد، این روش بازدهی بیشتری نسبت به استراتژی مرسوم خرید و نگهداری کسب کرده است.

رستگار و صداقتی‌پور (۱۳۹۷) یک سیستم معاملات الگوریتمی بر روی قرارداد آتی سکه طلای بورس کالای ارائه کرده‌اند. سیستم معاملاتی ارائه شده در این تحقیق از الگوریتم بهینه‌سازی ذرات چندهدفه^۱ (MOPSO) با هدف بهینه‌سازی دو تابع بازدهی و ارزش در معرض خطر شرطی^۲ (CVaR) استفاده کرده است. نتایج این تحقیق نشان داد که سیستم معاملاتی طراحی شده نسبت به بازدهی به ریسک مطلوب تری نسبت به دیگر استراتژی‌های رقیب مانند خرید و نگهداری و فروش و نگهداری دارد. همچنین چارچوب زمانی ۳۰ دقیقه برای طراحی سیستم‌های معاملاتی بر روی قرارداد آتی سکه طلا مناسب به نظر می‌رسد.

محمدی و همکاران (۱۳۹۸) با محاسبه بازده حاصل از استراتژی معاملات الگوریتمی و بازده حاصل از استراتژی منفعلانه مبتنی بر شاخص و نسبت شارپ عملکرد معاملات زوجی را با رویکرد هم‌انباشتگی در بورس و اوراق بهادار تهران با عملکرد استراتژی منفعل مقایسه کرده‌اند. نتایج این

1 Multi-Objective Particle Swarm Optimization

2 Conditional Value-at-Risk

تحقیق نشان می‌دهد که استفاده از این سیستم به عنوان یک سیستم معاملاتی خنثی نسبت به تغییرات و روندهای بازار، بازدهی چشمگیری نسبت به بازدهی سیستم مبتنی بر شاخص در مدت مشابه دارد.

چهرقانی و رستگار (۱۳۹۸) یک رویکرد ترکیبی جدید مبتنی بر یادگیری تقویتی کیو و رگرسیون شبکه عصبی عمیق در معاملات سهام ارائه کرده‌اند که امکان پیش‌بینی تعداد سهام برای معامله و عمل معاملاتی را فراهم می‌سازد این سیستم از داده‌های قیمت سهام‌های مختلف برای ساخت تصمیم استفاده می‌کند.

صادقی و همکاران (۱۳۹۹) در تحقیقی با استفاده از یک روش دو مرحله‌ای با استفاده از یادگیری جمعی و ماشین بردار پشتیبان (EHSVM) به پیش‌بینی روندهای بازار و انجام معامله در بازار فارکس پرداخته‌اند. در مرحله اول با استفاده از یادگیری جمعی EHSVM به پیش‌بینی روندهای بازار بر اساس صعودی، نزول و بدون نوسان در قالب سه کلاس اقدام کرده‌اند با استفاده از الگوریتم ژنتیک مرتب‌سازی نامغلوب^۱ ۲ (NSGA-II) بهترین استراتژی را انتخاب کرده‌اند. نتایج این تحقیق نشان داد که مدل ترکیبی نسبت به الگوریتم ژنتیک توانایی بیشتری برای انجام معاملات دارد.

هروانی و خلیلی‌عراقی (۱۳۹۹) یک الگوریتم معاملاتی بر مبنای میانگین متحرک تعدیل‌پذیر AMA جهت پیش‌بینی حرکت سهام در آینده با استفاده از بهینه‌سازی شمارشی ارائه کرده‌اند. نتایج این تحقیق نشان می‌دهد که می‌توان از استراتژی معاملات الگوریتمی به وسیله اندیکاتور AMA جهت پیش‌بینی حرکت قیمت سهام در آینده بازار سرمایه ایران استفاده کرد.

موسوی انزهایی و نیکومرام (۱۳۹۹) با استفاده از مدلی ترکیبی از آینده‌پژوهشی، تحلیل بنیادی، قواعد معاملاتی خبرگان و الگوریتم‌های یادگیری ماشین، الگویی برای اتخاذ راهبردهای معاملاتی مناسب پیشنهاد داده‌اند. نتایج این تحقیق نشان داد که استفاده از روش پیشنهادی برای سرمایه‌گذاری در بازار سهام بازدهی بالاتری را نسبت به شاخص بورس ایجاد می‌کند.

اسماعیلی و همکاران (۱۴۰۰) یک سیستم معاملاتی خودکار بر روی رمزارز اتریوم ارائه کرده‌اند. بدین منظور در ابتدا برای پیش‌بینی روند آتی قیمت اتریوم در یک دقیقه پیش رو از شش اندیکاتور تکنیکال به همراه قیمت بیت‌کوین در تایم فریم یک دقیقه‌ای و چهار مدل یادگیری ماشین استفاده شده است. نتایج نشان می‌دهد که مدل جنگل تصادفی بهترین عملکرد با دقت ۹۱ درصد را داشته است.

۲- روش تحقیق

روش پیشنهادی تحقیق حاضر دارای ۴ مرحله زیر است:

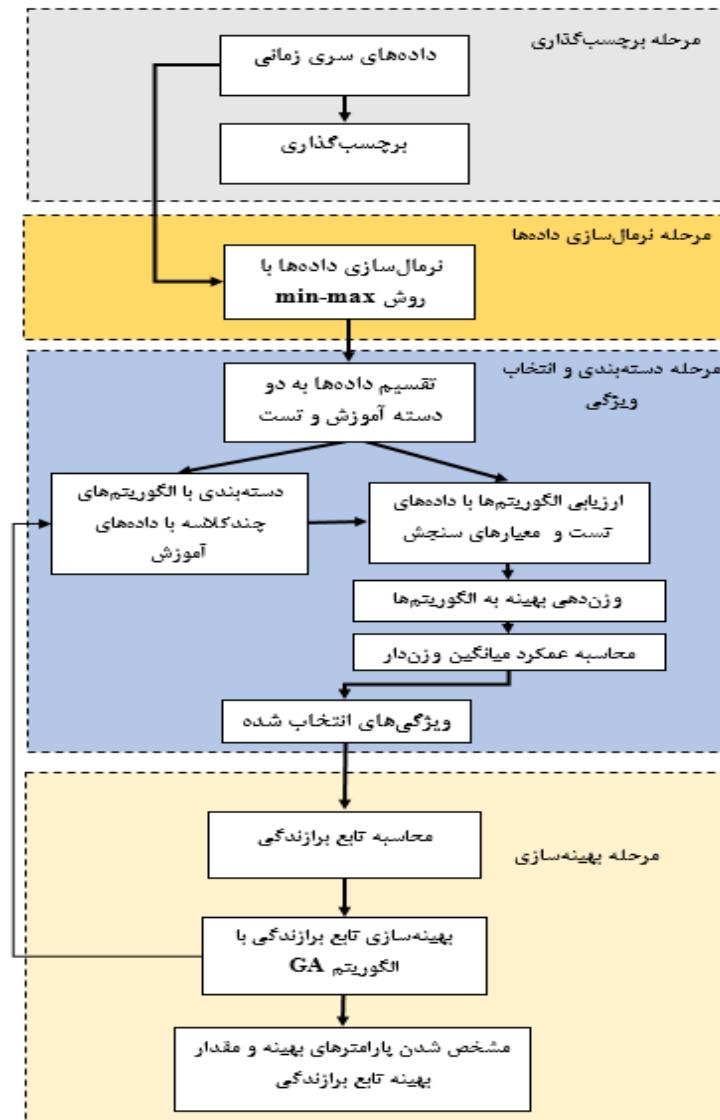
۱- برچسب‌گذاری داده‌های سری زمانی

۲- نرمال‌سازی داده‌ها

۳- دسته‌بندی و انتخاب ویژگی

۴- بهینه‌سازی

شماتیک روش پیشنهادی در شکل ۱ نشان داده شده است.



شکل ۱: شماتیک روش پیشنهادی (یافته‌های محقق)

۳-۱- برچسب‌گذاری داده‌های سری زمانی

داده‌های مورد استفاده در این تحقیق، معاملات روزانه برخی از سهام منتخب بازار اوراق بهادار تهران است. این داده‌ها از تاریخ فروردین سال ۱۳۹۱ تا انتهای اسفند سال ۱۳۹۹ جمع‌آوری می‌شود. متغیرهای ورودی مورد بررسی در تحقیق حاضر عبارتند از داده‌های قیمتی (باز شدن (Open)، بیشترین (High)، کمترین (Low) و بسته شدن (Close)) به همراه حجم معاملات و اندیکاتورهای غیرحجمی و حجمی. این متغیرها به عنوان ویژگی برای فرایند دسته‌بندی و تخمین قیمت استفاده می‌شوند. اندیکاتورهای غیرحجمی وابسته به داده‌های قیمتی و اندیکاتورهای حجمی وابسته به داده‌های حجم و (یا) قیمت معاملات هستند.

در این تحقیق اسیلاتور^۲ها، استوکستیک^۳ها، اندیس^۴ها و اندیکاتورها را بصورت خلاصه اندیکاتور می‌نامیم. اندیکاتورهای غیرحجمی در جدول ۱ نشان داده شده است. در مجموع ۱۵ اندیکاتور غیرحجمی در نظر گرفته شده است. لازم به ذکر است که برخی از اندیکاتورها دارای دوره زمانی هستند و مقادیر آنها متأثر از این دوره زمانی است. معمولاً این دوره زمانی ۹ روزه و ۱۲ روزه در نظر گرفته می‌شوند. با در نظر گرفتن این دوره‌های زمانی، در مجموع تعداد ۲۹ اندیکاتور غیرحجمی وجود دارد. اندیکاتورهای حجمی نیز بصورت جدول ۲ است. در مجموع تعداد ۶ اندیکاتور حجمی وجود دارد. بطور کلی با در نظر گرفتن قیمت‌های باز شدن، بیشترین، کمترین، بسته شدن و حجم معاملات، همراه با ۲۹ اندیکاتور غیرحجمی و ۶ اندیکاتور حجمی، در مجموع تعداد ۴۰ ویژگی در این تحقیق در نظر گرفته شده است.

برای انجام برچسب‌گذاری روی داده‌های سری زمانی در معاملات الگوریتمی، از سری زمانی قیمت بسته شدن استفاده می‌شود. در این تحقیق با استفاده از روش پنجره لغزشی^۵ برچسب‌گذاری نقاط خرید، فروش و نگهداری را انجام می‌دهیم. در تکنیک پنجره لغزشی با توجه به طول پنجره، در هر لغزش یک گام (در اینجا با توجه به اینکه داده‌های مورد استفاده بصورت روزانه است، یک روز) به جلوتر می‌رود و این عمل تا رسیدن به پنجره آخر ادامه پیدا می‌کند.

هزینه تراکنش در بازار بورس اوراق بهادار تهران یا درواقع کارمزد کارگزاری برای خرید برابر ۰/۳۷۱۲ درصد و برای فروش برابر ۰/۸۸ درصد و در مجموع برای یک معامله خرید و فروش برابر با ۱/۲۵۱۲ درصد است. با توجه به اینکه هدف ما از خرید و فروش یک سهم کسب سود است، بنابراین حداقل سود ۱ درصد را

1 Indicator

2 Oscillator

3 Stochastic

4 Index

5 Sliding window

به عنوان کمترین میزان سود قابل قبول برای معاملات الگوریتمی در نظر می‌گیریم. بنابراین در فرایند برچسب‌گذاری باید قیدی برابر با ۲/۲۵۱۲ درصد در نظر گرفته شود، چنانچه نقطه خرید و فروش (انجام یک معامله) دارای سودی بیشتر از این آستانه باشد، برچسب‌گذاری می‌شود و در غیر اینصورت وارد پنجره بعدی می‌شویم تا زمانی که نقطه فروش مناسب پیدا شود.

جدول ۱: اندیکاتورهای غیر حجمی مورد مطالعه

ردیف	اندیکاتور	توضیح	داده‌های مورد نیاز	دوره زمانی
۱	A/D.O	Accumulation/Distribution Oscillator	Open, High, Low, Close	ندارد
۲	S.O	Stochastic oscillator	High, Low, Close	دارد
۳	Acceleration	Acceleration Between Times	Open, High, Low, Close	دارد
۴	MOM	Momentum Between Times	Open, High, Low, Close	دارد
۵	SMA	Simple Moving Average	Close	دارد
۶	EMA	Exponential Moving Average	Close	دارد
۷	MMA	Modified Moving Average	Close	دارد
۸	MACD	Moving Average Convergence/Divergence	Open, High, Low, Close	ندارد
۹	RSI	Relative Strength Index	Close	دارد
۱۰	Williams %R	Williams %R	High, Low, Close	دارد
۱۱	Bollinger	Middle, Upper, and Lower Bands	Close	دارد
۱۲	ROC	Price Rate of Change	Close	دارد
۱۳	WADLine	Williams Accumulation/Distribution line	High, Low, Close	ندارد
۱۴	Typical Price	Average of the high, low, and closing prices	High, Low, Close	ندارد
۱۵	Chaikin Volatility	Chaikin Volatility	High, Low	دارد

منبع: نتایج تحقیق

جدول ۲: اندیکاتورهای حجمی مورد مطالعه

ردیف	اندیکاتور	توضیح	داده‌های مورد نیاز	دوره زمانی
۱	NVI	Negative Volume Index	Close, Volume	ندارد
۲	PVI	Positive Volume Index	Close, Volume	ندارد
۳	OBV	On-Balance Volume	Close, Volume	ندارد
۴	PVT	Price and Volume Trend	Volume	ندارد
۵	VROC	Volume Rate of Change	Volume	دارد

منبع: نتایج تحقیق

به توجه به اینکه انتخاب طول پنجره مناسب از اهمیت بالایی برخوردار است، در این تحقیق با استفاده از روش آنالیز حساسیت به انتخاب طول پنجره مناسب اقدام می‌کنیم. آنالیز حساسیت را با براساس طول پنجره و تعداد نقاط خرید (یا فروش) انجام می‌دهیم و آن طول پنجره‌ای که تعداد نقاط خرید و فروش بیشتری را حاصل کند، انتخاب می‌شود. چراکه در معاملات الگوریتمی، نقاط نگه‌داشتن خیلی بیشتر از نقاط خرید و فروش است و این امر در مبحث یادگیری الگوریتم‌ها برای انجام دسته‌بندی نقطه ضعیف محسوب می‌شود، زیرا الگوریتم بیشتر کلاس نگه‌داشتن را یاد می‌گیرد و دقت دسته‌بندی کلاس‌های خرید و فروش پایین می‌آید. برای رفع این مشکل نیز روش‌های مختلفی ارائه شده است که از جمله آنها می‌توان به تکنیک شافلینگ^۱ (بچلور و وان^۲، ۲۰۰۷؛ هولیس^۳ و همکاران، ۲۰۱۸) و یادگیری بر اساس پنجره لغزشی (هوتا^۴ و همکاران، ۲۰۱۷) و سایر روش‌ها اشاره کرد. در اینجا ما از تکنیک یادگیری با پنجره لغزشی استفاده خواهیم کرد.

۳-۲- نرمال‌سازی داده‌ها

با توجه به اینکه داده‌ها از نظر بزرگی دارای محدوده‌های مختلفی هستند، ابتدا باید از نظر محدوده همسان‌سازی شوند. به این عمل همسان‌سازی، نرمال‌کردن داده گفته می‌شود. یکی از روش‌های پرکاربرد برای نرمال‌سازی داده‌ها، روش min-max است که داده‌ها را در محدوده بین ۰ تا ۱ قرار می‌دهد:

$$X_i^N = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

1 shuffling

2 Batchelor & Kwan

3 Hollis

4 Hota

که در آن، X_i^N متغیر نرمال شده، $X_{\min}$ مینیمم مقدار آن و $X_{\max}$ ماکزیمم مقدار آن است.

۳-۳- دسته‌بندی و انتخاب ویژگی بهینه

به منظور انجام عمل دسته‌بندی و در نهایت انتخاب ویژگی بر اساس معیارهای سنجش عملکرد الگوریتم‌های دسته‌بندی، ابتدا باید الگوریتم‌های دسته‌بندی آموزش ببینند. بنابراین، ابتدا لازم است داده‌ها را به دو دسته داده‌های آموزش و تست تقسیم کرد. برای انجام این عمل و با توجه به اینکه انتخاب داده‌های تست و آموزش بصورت تصادفی است، از روش اعتبارسنجی مقطعی^۱ k-fold استفاده می‌شود. در روش k-fold داده‌ها به k دسته تقسیم می‌شوند و به ازای k بار داده‌های تست و آموزش انتخاب می‌شوند و k بار الگوریتم‌های دسته‌بندی اجرا شده و نتایج عملکردی آنها محاسبه می‌شود و در نهایت میانگین این k بار به عنوان عملکرد نهایی در نظر گرفته می‌شود. در اینجا ما به ازای k برابر ۵ اینکار را انجام خواهیم داد. ۸۰٪ داده‌ها را در هر fold به عنوان داده‌های آموزش و ۲۰٪ مابقی را به عنوان داده‌های تست برای بررسی عملکرد الگوریتم‌های دسته‌بندی در نظر می‌گیریم. با توجه به اینکه در معاملات الگوریتم با سه کلاس خرید، فروش و نگهداری سروکار داریم، بنابراین باید از روش‌های دسته‌بندی چندکلاسه استفاده شود. روش‌های مختلفی برای دسته‌بندی چندکلاسه وجود دارد که در اینجا ما از ۴ الگوریتم درخت تصمیم (DTree)، K - نزدیکترین همسایگی (KNN)، روش RUSBoost و شبکه عصبی پایه شعاعی (RBN) استفاده می‌کنیم. پس از تست این الگوریتم‌ها با داده‌های تست، وزن بدست آمده از فرایند بهینه‌سازی برای هر کدام از الگوریتم‌ها در نظر گرفته شده و میانگین وزن‌دار معیارهای عملکردی دسته‌بندی محاسبه می‌شود. غالباً در معاملات الگوریتمی، ۴ معیار عملکردی زیر در نظر گرفته می‌شود:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۲)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۳)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (۴)$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (۵)$$

1 Cross-validation

که در آن TP^1 بیانگر کلاس مثبتی است که بدرستی مثبت تشخیص داده شده است. TN کلاس منفی هست که بدرستی منفی تشخیص داده شده است. FP بیانگر کلاس منفی است که مثبت تشخیص داده شده است و FN نیز بیانگر کلاس مثبت است که به اشتباه منفی تشخیص داده شده است. مقادیر هر یک از معیارهای چهارگانه بین ۰ و ۱ قرار دارد، و هرچقدر این مقدار به ۱ نزدیک باشد، بیانگر عملکرد خوب الگوریتم‌های دسته‌بندی است.

با در نظر گرفتن اینکه در معاملات الگوریتم با سه کلاس روبرو هستیم، بنابراین چهار معیار فوق بطور جداگانه برای هر کلاس محاسبه شده و در نهایت مقدار میانگین وزن‌دار آنها به عنوان مقدار نهایی بدست می‌آید. روند اینکار در جدول ۳ نشان داده شده است. در این جدول $\overline{Accuracy}$ بیانگر مقدار میانگین دقت هر سه کلاس است و برای سه معیار عملکردی هم به همین ترتیب می‌باشد.

جدول ۳: روند محاسبه مقادیر میانگین وزن‌دار معیارهای عملکردی

مقدار میانگین	الگوریتم ۴ ERBN	الگوریتم ۳ RUSBOOST	الگوریتم ۲ KNN	الگوریتم ۱ DTREE	
$\sum_{i=1}^4 W_i \overline{Accuracy}_i$	W_4	W_3	W_2	W_1	$\overline{Accuracy}$
$\sum_{i=1}^4 W_i \overline{Precision}_i$	W_4	W_3	W_2	W_1	$\overline{Precision}$
$\sum_{i=1}^4 W_i \overline{Recall}_i$	W_4	W_3	W_2	W_1	$\overline{Recall}$
$\sum_{i=1}^4 W_i \overline{F1-Score}_i$	W_4	W_3	W_2	W_1	$\overline{F1-Score}$

منبع: نتایج تحقیق

۵-۳- فرایند بهینه‌سازی

الگوریتم‌ها دارای یک سری پارامتر به عنوان تنظیمات الگوریتم‌ها هستند که مقادیر آنها به عنوان هایپرپارامتر از طریق فرایند بهینه‌سازی بدست می‌آیند. علاوه بر این، در این تحقیق ۴۰ ویژگی مورد مطالعه قرار گرفته است، بنابراین برای انتخاب ویژگی بهینه نیز این ویژگی‌ها به عنوان هایپرپارامتر در نظر گرفته شده‌اند. در جدول ۴ پارامترهای بهینه‌سازی به همراه کران بالا و پایین آنها آورده شده است. برای ویژگی‌ها، کران پایین برابر ۰ و کران بالا برابر ۱ است و این متغیرها بصورت باینری در نظر گرفته می‌شوند، بطوری که مقدار ۰ به معنای انتخاب نشدن آن ویژگی و مقدار ۱ به معنای انتخاب شدن آن است. با توجه به جدول ۴، ۷ متغیر تصمیم مربوط به الگوریتم‌های دسته‌بندی، ۴۰ متغیر تصمیم مربوط به ویژگی‌ها و ۴ متغیر تصمیم مربوط به الگوریتم‌های دسته‌بندی است. بنابراین در مجموع، تعداد ۵۱ متغیر از فرایند بهینه‌سازی بدست خواهند آمد.

در فرآیند تصمیم، انتخاب تابع برازندگی مناسب از اهمیت بالایی برخوردار است. در اینجا ما کزیم مقدار F1-Score به عنوان تابع برازندگی در نظر گرفته می‌شود. با توجه به آنچه که گفته شد، مسئله بهینه تحقیق حاضر بصورت زیر تعریف می‌شود:

$$\max: \sum_{i=1}^4 W_i \overline{\text{F1-Score}_i}$$

s.t.

$$LB \leq X_{GA} \leq UB$$

$$\sum_{i=1}^4 W_i = 1$$

(۶)

در معادله (۶)، مقدار تابع $\sum_{i=1}^4 W_i \overline{\text{F1-Score}_i}$ برابر با مقدار میانگین F1-Score هر سه کلاس

و هر ۴ الگوریتم است. علاوه بر این، X_{GA} در این معادله، برداری شامل تمامی ۵۱ متغیر تصمیم است. همچنین قیودی مبنی بر اینکه مجموع وزن‌های الگوریتم‌های دسته‌بندی و مجموعه وزن‌های الگوریتم‌های رگرسیون برابر ۱ نیز برقرار است.

جدول ۴: هایپر پارامترها و کران پایین و بالای آنها

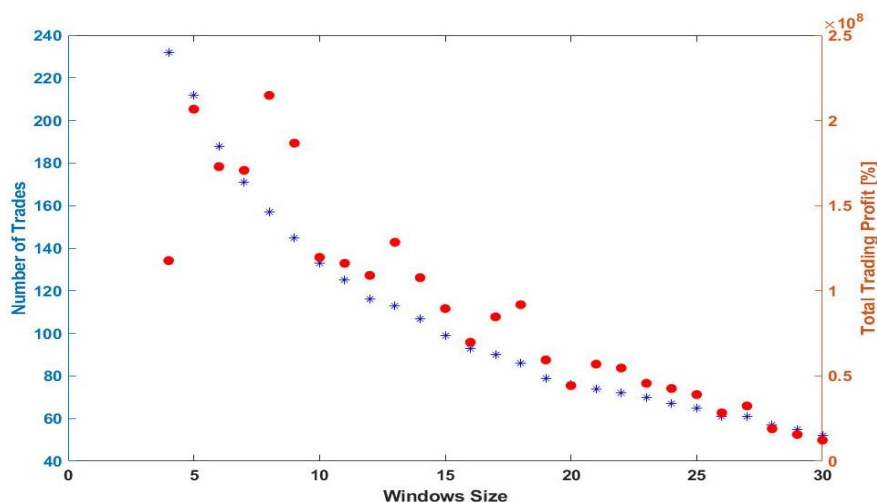
الگوریتم	پارامتر	کران پایین	کران بالا
DTREE	ماکزیمم تعداد شاخه‌ها	۱	۱۰۰
	مینیمم تعداد هر برگ در شاخه	۱	۲۰
KNN	تعداد همسایگی	۱	۲۰
RUSBOOST	تعداد سیکل یادگیری	۱۰۰	۱۰۰۰
	نرخ یادگیری	۰.۱	۰.۲
RBN	مقدار گسترش ^۱ تابع شعاعی	۰.۱	۱
	ماکزیمم تعداد نرون‌ها	۱	۳۰
--	ویژگی‌ها	۰	۱
--	وزن الگوریتم‌های دسته‌بندی	۱/۰	۱

منبع: نتایج تحقیق

۳- یافته‌های تحقیق

۳-۱- برچسب‌گذاری

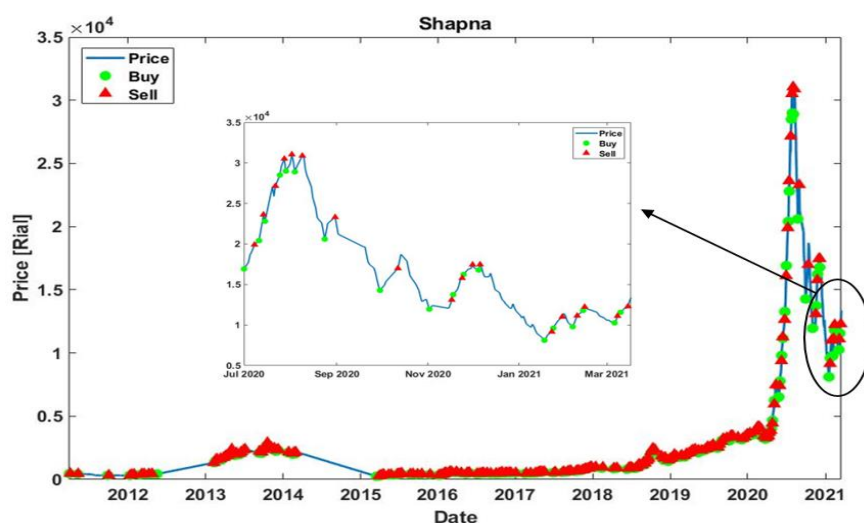
به منظور برچسب‌گذاری نقاط خرید و فروش روی نمودار قیمتی سهام، از روش پنجره لغزشی استفاده شده است. طول پنجره در این روش نقش بسیار مهمی در تعداد نقاط خرید و فروش دارد؛ بگونه‌ای که هرچه قدر طول پنجره کم باشد، تعداد نقاط خرید و فروش بیشتر می‌شود و بالعکس، هرچه قدر طول پنجره بیشتر باشد، تعداد نقاط خرید و فروش کمتر می‌شود. در این بخش، بر اساس آنالیز حساسیت طول پنجره، مقدار مناسب آن را پیدا می‌کنیم. بدین منظور، کمترین مقدار طول را برابر ۴ و بیشترین آن را برابر ۳۰ در نظر می‌گیریم. لازم به ذکر است، حداقل طول پنجره باید بگونه‌ای باشد که علاوه بر نقاط ماکزیمم (فروش) و مینیمم (خرید) در آن طول پنجره، یک یا نقطه نگهداری هم وجود داشته باشد. بدین خاطر هست که مینیمم طول پنجره را برابر ۴ در نظر گرفته‌ایم. نکته مهمی که باید در نظر گرفته شود این است که طول پنجره مناسب باید بگونه‌ای باشد که بیشترین بازدهی خرید و فروش را نیز سبب شود. بنابراین در انتخاب طول پنجره مناسب، دو پارامتر تعداد نقاط خرید (فروش) و حداکثر بازدهی ممکن در نظر گرفته شده است.



شکل ۲: آنالیز حساسیت تعداد معاملات و بازدهی مرکب آن
بر حسب طول پنجره برای سهم شپنا (خروجی نرم افزار)

نمودار آنالیز حساسیت طول پنجره بر حسب تعداد نقاط خرید و فروش و همچنین بازدهی مرکب حاصل از آنها برای سهم شپنا، در شکل ۲ ترسیم شده است. همانطور که از این شکل مشخص است، با در نظر گرفتن مصالحه‌ای بین این دو پارامتر، طول پنجره ۵ به عنوان بهترین انتخاب در نظر گرفته می‌شود. در شکل ۳ نقاط خرید و فروش برچسب‌گذاری شده بر اساس فرایند برچسب‌گذاری ذکر شده در این بخش، به ازای طول پنجره برابر ۵ ترسیم شده است. همانطور که از بخش زوم شده شکل مشخص است، الگوریتم برچسب‌گذاری بخوبی نقاط خرید و فروش را شناسایی و برچسب گذاشته است. برای سایر سهام نیز طول پنجره ۵ حاصل گردید.

لازم به ذکر است که فرایند برچسب‌گذاری در الگوریتم کلی پیشنهادی تحقیق مدام بر اساس قیمت‌های تخمینی بدست آمده از الگوریتم، انجام می‌شود و در نهایت بر اساس بهترین نتیجه حاصل از الگوریتم، برچسب‌گذاری بهینه (برای قیمت بسته شدن تخمینی بهینه) نیز حاصل می‌شود. بر اساس نتایج حاصل از برچسب‌گذاری، تعداد نقاط خرید و فروش (تعداد معاملات)، تعداد دوره‌های (روزهای) نگهداشت، بازده معاملات الگوریتمی و بازده خرید ابتدای دوره (ابتدای سال ۱۳۹۰) و فروش انتهای دوره (انتهای سال ۱۳۹۹) مطابق با جدول ۵ بدست آمد. همانطور که از این جدول مشاهده می‌شود، بازدهی‌های حاصل از معاملات الگوریتمی، بسیار بیشتر از حد تصور.



شکل ۳: نقاط خرید و فروش کل دوره برای سهم شپنا. (خروجی نرم افزار)

است و چنانچه فردی از سال ۱۳۹۰ با در نظر گرفتن آستانه سود ۱ درصد اقدام به انجام معاملات الگوریتمی تا پایان سال ۱۳۹۹ می‌کرد، به چنین سودهای نجومی دست می‌یافت. در حالی که سود حاصل از خرید و نگهداشت بسیار کمتر از این سودهای نجومی است. مورد دیگر، اختلاف بازدهی ناشی از کارمزد است که نشان می‌دهد، کارمزد کارگزاری‌ها برای هر معامله، کارمزدی زیاد محسوب می‌شود. به منظور درک بازدهی‌های حاصل از معاملات الگوریتمی، بازدهی انباشته سالانه برای هر سه سهم که از طریق این معاملات بدست آمده است، در جدول ۶ نشان داده شده است.

جدول ۵: نتایج حاصل از برچسب‌گذاری

سهم	تعداد معاملات	تعداد روزهای نگهداشت	بازده معاملات الگوریتمی	بازده معاملات الگوریتمی (بدون کارمزد)	بازده خرید و نگهداشت
شپنا	۲۱۲	۷۳۴	۱۶۹۹۳۶۶۵	۲۰۶۶۳۰۴۰۳	۲۷۷۹
خودرو	۲۲۸	۹۱۰	۵۱۷۲۵۱۰۸۶	۱۷۲۴۶۸۱۹۶	۲۶۰۸
فجر	۱۳۰	۶۲۷	۹۶۶۳۶۱	۴۳۷۲۰۲۹	۱۲۷۲۵

منبع: نتایج تحقیق

جدول ۶: بازدهی سالانه انباشته حاصل از معاملات الگوریتمی

سال	شپنا	خودرو	فجر
۹۰	۸/۰۴	۵۹/۹۵	۱۰۰/۰۸
۹۱	۱۰۸۸/۰۶	۳۹۰/۲۳	۲۲۰/۵۸
۹۲	۳۶۸۰/۹۲	۶۱۶۱/۲۵	۶۳۹/۷۹
۹۳	۷۵۵۶/۳۲	۱۸۹۶۷/۷۱	۱۹۷۹/۸۰
۹۴	۱۳۳۳۹/۱۳	۹۱۰۶۱/۲۱	۲۵۸۶/۳۱
۹۵	۲۸۷۴۸/۸۸	۲۷۰۹۶۷/۰۶	۴۳۹۵/۷۶
۹۶	۱۳۸۳۳۹/۰۸	۵۵۸۹۳۰/۱۴	۹۶۱۰/۲۱
۹۷	۴۰۲۹۴۱/۴۵	۲۰۴۳۳۷۳/۹۸	۱۹۳۰۲/۱۸
۹۸	۱۵۴۲۶۰۶/۰۲	۱۵۷۲۸۲۳۵/۰۲	۱۸۷۴۲۲/۲۰
۹۹	۱۶۹۹۴۶۶۴/۹۹	۵۱۷۲۵۱۰۸۶/۶۶	۹۶۶۳۶۱/۴۱

منبع: نتایج تحقیق

۴-۲- فرایند بهینه سازی دسته بندی

با انجام فرایند بهینه سازی، نتایج بهینه حاصل گردید. پارامترهای بهینه الگوریتم ها به همراه وزن بهینه و ویژگی های بهینه انتخابی به ترتیب در جدول ۷ تا ۹ آورده شده است. همانطور که از این جداول مشخص است، برای سهام مختلف و الگوریتم های مختلف، مقادیر بهینه متفاوتی به دست آمده است. با توجه به جدول ۹، تعداد ۱۹ ویژگی بهینه برای سهم شپنا و تعداد ۱۸ ویژگی بهینه برای خودرو و ۱۶ ویژگی بهینه برای فجر انتخاب شده است. از این میان، تعداد ۱۱ ویژگی مشترک وجود دارد.

نتایج دسته بندی کلاس های خرید، فروش و نگهداری (معاملات الگوریتمی) بر اساس معیارهای مورد مطالعه در جدول ۱۰ آورده شده است. مطابق با این جدول، همانطور که انتظار می رفت، کلاس نگهداری نسبت به دو کلاس خرید و فروش دارای عملکرد دسته بندی بهتری است. بصورت میانگین، نتایج عملکردی بدست آمده از روش پیشنهادی حاکی از یک عملکرد دسته بندی بسیار مناسب است و نشان می دهد که روند بهینه سازی بخوبی قادر به بهینه سازی روند دسته بندی کلاس ها بوده است.

جدول ۷: مقادیر بهینه پارامتر الگوریتم‌ها

الگوریتم	پارامتر	کران پایین	کران بالا	مقدار بهینه		
				شپنا	خودرو	فجر
DTREE	ماکزیمم تعداد شاخه‌ها	۱	۱۰۰	۱۰	۱۳	۹
	مینیمم تعداد هر برگ در شاخه	۱	۲۰	۲	۲	۲
KNN	تعداد همسایگی	۱	۲۰	۴	۲	۳
RUSBOOST	تعداد سیکل یادگیری	۱۰۰	۱۰۰۰	۳۱۹	۲۵۷	۱۸۵
	نرخ یادگیری	۰.۱	۲۰	۰.۳۶	۰.۴۱	۰.۳۹
RBN	مقدار گسترش ^۱ تابع شعاعی	۰.۱	۱	۰.۱۱	۰.۱۷	۰.۲۳
	ماکزیمم تعداد نرون‌ها	۱	۳۰	۲۳	۱۹	۱۴

منبع: نتایج تحقیق

جدول ۸: وزن بهینه الگوریتم‌ها

الگوریتم	کران پایین	کران بالا	مقدار بهینه		
			شپنا	خودرو	فجر
DTREE	۰	۱	۲۰.۷	۲۳.۳	۲۰.۲
KNN	۰	۱	۱۹.۹	۲۰.۴	۱۷.۰
RUSBOOST	۰	۱	۲۵.۶	۲۶.۲	۲۷.۲
ERBN	۰	۱	۳۳.۷	۳۰.۱	۳۵.۶

منبع: نتایج تحقیق

جدول ۹: ویژگی‌های بهینه منتخب

مقدار بهینه			ویژگی
فجر	خودرو	شپنا	
OPEN	OPEN	OPEN	قیمت
HIGH	HIGH	HIGH	
LOW	LOW	LOW	
CLOSE	CLOSE	CLOSE	
VOLUME	VOLUME	VOLUME	حجم
ADO	ADO	ADO	اندیکاتورهای غیر حجمی
SO9	SO9	SO12	
SO12	SO12	ACC12	
ACC9	ACC12	MOM12	
ACC12	MOM12	MMA9	
MOM9	MACD	MMA12	
MACD	RSI9	MACD	
RSI12	RSI12	RSI12	
ROC9	BOLLINGER UPPER9	BOLLINGER UPPER12	
ROC12	ROC12	ROC9	
		ROC9	اندیکاتورهای حجمی
		WILLIAMS9	
	NVI	PVI	
	OBV		
VROC12	VROC12	VROC12	
۱۶	۱۸	۱۹	مجموع

منبع: نتایج تحقیق

جدول ۱۰: مقادیر بهینه معیارهای عملکردی دسته‌بندی کلاس‌ها

مقدار			معیار	کلاس
فجر	خودرو	شپنا		
۸۱۱/۰	۸۸۸/۰	۸۶۶/۰	ACC	خرید
۸۳۶/۰	۹۲۹/۰	۹۱۱/۰	PRECISION	
۸۷۴/۰	۹۲۱/۰	۹۰۲/۰	RECALL	
۸۵۵/۰	۹۱۹/۰	۹۰۶/۰	F1-SCORE	
۸۶۱/۰	۸۷۷/۰	۸۹۲/۰	ACC	فروش
۸۶۸/۰	۸۹۲/۰	۹۳۳/۰	PRECISION	
۸۷۳/۰	۹۱۱/۰	۹۲۸/۰	RECALL	
۸۷۰/۰	۹۰۹/۰	۹۳۰/۰	F1-SCORE	
۸۹۲/۰	۹۴۴/۰	۹۱۹/۰	ACC	نگهداری
۹۲۸/۰	۹۴۸/۰	۹۴۶/۰	PRECISION	
۹۰۱/۰	۹۳۴/۰	۹۲۹/۰	RECALL	
۹۱۴/۰	۹۴۱/۰	۹۳۹/۰	F1-SCORE	
۸۵۵/۰	۹۰۳/۰	۸۹۳/۰	ACC	میانگین
۸۷۷/۰	۹۲۶/۰	۹۲۹/۰	PRECISION	
۸۸۳/۰	۹۲۲/۰	۹۲۰/۰	RECALL	
۸۷۹/۰	۹۲۸/۰	۹۲۶/۰	F1-SCORE	

منبع: نتایج تحقیق

۴- نتیجه‌گیری و پیشنهادات

در این پژوهش، یک سیستم معاملات الگوریتمی بهینه طراحی شد که قادر به دسته‌بندی سه کلاس خرید، فروش و نگهداری است. مدل‌های دسته‌بندی وابسته به تعداد و نوع ویژگی‌های مسئله هستند. با توجه به اینکه برای مسائل و داده‌های زمانی مختلف، ویژگی‌های مختلفی وجود دارد و نمی‌توان برای مسائل دسته‌بندی یک سری ویژگی‌های کلی برای تمامی مسائل در نظر گرفت، لذا ویژگی‌های انتخابی بسته به نتایج و تفسیر محقق دارد. در این پژوهش، تعداد ۴۰ ویژگی قیمتی، حجمی و اندیکاتورهای حجمی و غیرحجمی در نظر گرفته شده است. علاوه بر این، الگوریتم‌های دسته‌بندی نیز دارای یک سری پارامتر هستند که تنظیم مقادیر آنها نیاز به تجربه و سعی و خطای مختلف دارد. بنابراین طراحی یک سیستمی که بصورت اتوماتیک این مقادیر را تنظیم کند، بسیار مناسب و جالب توجه است. یکی از بهترین رویکردها برای چنین مسائلی، استفاده از فرایند بهینه‌سازی با الگوریتم‌های فرا ابتکاری و در رأس آن الگوریتم ژنتیک است. این الگوریتم با داشتن فضای جستجو و تابع برازندگی، قادر است که مقادیر بهینه‌های پارامترها را بدست آورد. بنابراین، بطور کلی می‌توان گفت که در صورت عدم استفاده از فرایند بهینه‌سازی (در اینجا با الگوریتم ژنتیک)، باید از همان تعداد ۴۰ ویژگی به اضافه پارامترهای از پیش تنظیم شده الگوریتم‌ها استفاده می‌شد که منجر به نتایج متفاوتی و گوناگونی (بسته به مجموعه مقادیر مختلف پارامترها) می‌شد. نتایج حاصل از این تحقیق، گویای آن است که مدل پیشنهادی دارای عملکرد دسته‌بندی بسیار مناسبی است و می‌تواند به عنوان یک ماژول معاملاتی خودکار در بازار ایران مورد استفاده قرار گیرد.

منابع

۱. اسماعیلی، وحید. ۱۴۰۰. ارائه سیستم معاملات الگوریتمی برای رمزارزها بر پایه یادگیری ماشین، پایان‌نامه کارشناسی ارشد. دانشکده مهندسی صنایع و سیستم‌ها. دانشگاه تربیت مدرس تهران.
۲. چهرقانی، مجید. ۱۳۹۸. طراحی سیستم معاملاتی خودکار همراه با تعیین تعداد بهینه سهام جهت انجام معامله با استفاده از یادگیری Q عمیق. پایان‌نامه کارشناسی ارشد. دانشکده مهندسی صنایع و سیستم‌ها. دانشگاه تربیت مدرس.
۳. رستگار، محمدعلی. و ساعدی‌فر، خاطره. ۱۳۹۶. استراتژی بهینه اجرای معاملات بزرگ با رویکرد شبیه‌سازی عامل‌گرا. تحقیقات مالی. ۱۹(۲): ۲۶۲-۲۳۹.
۴. رستگار، محمدعلی. و صداقتی‌پور، امین. ۱۳۹۷. ارائه سیستم معاملات الگوریتمی برای قرارداد آتی سکه طلا مبتنی بر داده‌های درون-روزی. دانش سرمایه‌گذاری. ۷(۲۸): ۶۷-۴۹.
۵. ساعدی‌فر، خاطره. ۱۳۹۵. استراتژی انجام معاملات بهینه: رویکرد شبیه‌سازی عامل‌گرا. پایان‌نامه کارشناسی ارشد. دانشکده علوم مالی. دانشگاه خوارزمی.
۶. صادقی، علیرضا. ۱۳۹۹. بهینه‌سازی راهبردهای سرمایه‌گذاری در بازار فارکس مبتنی بر ترکیب الگوریتم‌های فراابتکاری و یادگیری ماشین، رساله دکتری تخصصی. رشته مالی. دانشکده مدیریت و اقتصاد. دانشگاه آزاد. واحد علوم و تحقیقات تهران.
۷. غلامیان، الهام. داورودی، سیدمحمد رضا. ۱۳۹۷. پیش‌بینی روند قیمت در بازار سهام با استفاده از الگوریتم جنگل تصادفی. مجله مهندسی مالی و مدیریت اوراق بهادار. ۹(۳۵): ۳۲۲-۳۰۱.
۸. محمدی، صفورا. ۱۳۹۸. مقایسه عملکرد معاملات الگوریتمی با استراتژی منفعلانه. پایان‌نامه کارشناسی ارشد. دانشکده مهندسی صنایع و مدیریت. دانشگاه صنعتی شاهرود.
۹. موسوی‌انزهایی، سیدمجید. نیکومرام، هاشم. ۱۳۹۹. طراحی الگوی تعیین راهبردهای معاملاتی سهام با رویکرد مبتنی بر آینده‌پژوهی. تحلیل بنیادی، مهندسی ویژگی‌ها و الگوریتم‌های یادگیری ماشین. مهندسی مالی و مدیریت اوراق بهادار. ۱۱(۴۵): ۵۱۷-۴۹۹.
۱۰. هروانی، محمد. و خلیلی‌عراقی، مریم. ۱۳۹۹. طراحی استراتژی معاملات الگوریتمی با معرفی اندیکاتور میانگین متحرک تعدیل‌پذیر (AMA) جهت پیش‌بینی حرکت قیمت سهام در آینده در بازار سرمایه ایران. اولین کنفرانس بین‌المللی چالش‌ها و راهکارهای نوین در مهندسی صنایع و مدیریت و حسابداری.
۱۱. یافتیان، امیرحسین. ۱۳۹۷. طراحی یک سیستم معاملات الگوریتمی با استفاده از پردازش تصویر به وسیله شبکه عصبی کانولوشن، پایان‌نامه کارشناسی ارشد. دانشکده مهندسی صنایع و سیستم‌ها. دانشگاه تربیت مدرس.

12. Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, 113, 457-480.
13. Batchelor, R., & Kwan, T. Y. (2007). Judgemental bootstrapping of technical traders in the bond market. *International Journal of Forecasting*, 23(3), 427-445.
14. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
15. Chung, H., & Shin, K. S. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, 10(10), 3765.
16. Ding, G., & Qin, L. (2020). Study on the prediction of stock price based on the associated network model of LSTM. *International Journal of Machine Learning and Cybernetics*, 11(6), 1307-1317.
17. Dixon, M., Klabjan, D., & Bang, J. H. (2015, November). Implementing deep neural networks for financial market prediction on the Intel Xeon Phi. In *Proceedings of the 8th Workshop on High Performance Computational Finance* (pp. 1-6).
18. Gomber, P., & Zimmermann, K. (2018). Algorithmic trading in practice. In *The Oxford Handbook of Computational Economics and Finance*. Oxford University Press.
19. Gudelek, M. U., Boluk, S. A., & Ozbayoglu, A. M. (2017, November). A deep learning based stock trading model with 2-D CNN trend detection. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-8). IEEE.
20. Hollis, T., Viscardi, A., & Yi, S. E. (2018). A comparison of LSTMs and attention mechanisms for forecasting financial time series. *arXiv preprint arXiv:1812.07699*.
21. Hota, H. S., Handa, R., & Shrivastava, A. K. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research*, 13(5), 1145-1156.
22. Kissell, R. (2021). *Algorithmic trading methods: Applications using advanced statistics, optimization, and machine learning techniques*. Academic Press.

23. Lei, K., Zhang, B., Li, Y., Yang, M., & Shen, Y. (2020). Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. *Expert Systems with Applications*, 140, 112872.
24. Li, J., Bu, H., & Wu, J. (2017, June). Sentiment-aware stock market prediction: A deep learning method. In 2017 international conference on service systems and service management (pp. 1-6). IEEE.
25. Li, Y., Zheng, W., & Zheng, Z. (2019). Deep robust reinforcement learning for practical algorithmic trading. *IEEE Access*, 7, 108014-108022.
26. Li, Y., Zhu, Z., Kong, D., Han, H., & Zhao, Y. (2019). EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowledge-Based Systems*, 181, 104785.
27. Moghaddam, A. H., Moghaddam, M. H., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, 21(41), 89-93.
28. Nuti, G., Mirghaemi, M., Treleaven, P., & Yingsaeree, C. (2011). Algorithmic trading. *Computer*, 44(11), 61-69.
29. Rundo, F. (2019). Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems. *Applied Sciences*, 9(20), 4460.
30. Seyfert, R. (2018). Automation and affect: A study of algorithmic trading. In *Affect in Relation* (pp. 197-217). Routledge.
31. Sezer, O. B., & Ozbayoglu, A. M. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. *Applied Soft Computing*, 70, 525-538.
32. Wen, Y., & Yuan, B. (2018, March). Use CNN-LSTM network to analyze secondary market data. In *Proceedings of the 2nd International Conference on Innovation in Artificial Intelligence* (pp. 54-58).
33. Yan, H. (2019). The real effects of algorithmic trading (Doctoral dissertation, Duke University).
34. Yoshihara, A., Fujikawa, K., Seki, K., & Uehara, K. (2014, December). Predicting stock market trends by recurrent deep neural networks. In *Pacific rim international conference on artificial intelligence* (pp. 759-769). Springer, Cham
35. Zhao, Z., Rao, R., Tu, S., & Shi, J. (2017, November). Time-weighted LSTM model with redefined labeling for stock trend prediction. In 2017 IEEE 29th international conference on tools with artificial intelligence (ICTAI) (pp. 1210-1217). IEEE.