

بهینه‌سازی سبد سهام بر اساس مدل ترکیبی نسبت امگا و میانگین - واریانس مارکوویتز مبتنی بر یادگیری ماشین جمعی دو سطحی

ساناز فریدی^۱

مهدی معدنچی زاج^۲ ✉

امیر دانشور^۳

شادی شاهوردیانی^۴

فریدون رهنمای رودپشتی^۵

تاریخ پذیرش: ۱۴۰۱/۰۵/۲۳

تاریخ دریافت: ۱۴۰۰/۰۹/۰۲

چکیده

در این مقاله به بهینه‌سازی سبد سهام شرکت‌های فعال پذیرفته شده در اوراق بهادار بورس تهران بر اساس مدل ترکیبی نسبت امگا و میانگین - واریانس مارکوویتز (*MVOF*) پرداخته شده است. برای این امر ۴۸۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۹۰ تا ۱۳۹۹ انتخاب و بر اساس داده‌های ورودی به فیلتر کردن شرکت‌ها پرداخته شد. از این رو یک روش ترکیبی متشکل از روش بهینه‌سازی قواعد معاملاتی مبتنی بر تحلیل تکنیکال (۶ اندیکاتور *ROC*، *RSI*، *SMA*، *EMA*، *WMA* و *MACD*) و ماشین یادگیری جمعی دو سطحی (*SVM*، *RF*، *BN*، *MLP* و *KNN*) جهت آموزش داده‌ها و ارائه سیگنال خرید پرداخته شد. لذا ۸۵ شرکت جهت بهینه‌سازی سبد سهام انتخاب شدند. برای آموزش داده‌های از ۸۵ شرکت فیلتر شده توسط روش ترکیبی استفاده و تعداد طبقات مختلف با ۵۰ یادگیرنده استفاده شده است. نتایج نشان می‌دهد استفاده از مدل *OF* نسبت به مدل *MVF* بالاترین بازده سبد سهام را در طی سال‌های ۱۳۹۵ تا ۱۳۹۹ دارد. در حالی که مدل *MVF* پایین‌ترین میزان ریسک سرمایه‌گذاری را به خود اختصاص داده است. در نتیجه با ترکیب مدل‌های فوق، مشاهده شده بازده سبد سهام در این روش بسیار بالاتر از روش‌های دیگر است. در حالی که ریسک سرمایه‌گذاری آن کمتر بوده است. لذا در صورت استفاده از مدل *MVOF* بازدهی سبد سهام افزایش و ریسک سرمایه‌گذاری در آن کاهش می‌یابد.

واژه‌های کلیدی: مدل نسبت امگا، مدل میانگین - واریانس مارکوویتز، ماشین یادگیری جمعی دو سطحی، الگوریتم فرا ابتکاری

^۱ دانشجوی دکتری گروه مدیریت مالی، دانشکده مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران. sanaz.faridi66@gmail.com

^۲ استادیار گروه مدیریت مالی، دانشکده مدیریت، دانشگاه آزاد اسلامی، واحد الکترونیکی، تهران، ایران (نویسنده مسئول): madanchi@iauec.ac.ir

^۳ استادیار گروه مدیریت فناوری اطلاعات، دانشکده مدیریت، دانشگاه آزاد اسلامی، واحد الکترونیکی، تهران، ایران. daneshvar.amir@gmail.com

^۴ استادیار، گروه مدیریت بازرگانی، دانشگاه آزاد اسلامی، واحد شهرقدس، تهران، ایران. shahverdiyani@gmail.com

^۵ استاد، گروه اقتصاد، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران. Rahnama.roodposhti@gmail.com

۱- مقدمه

پیش‌بینی بازار سهام یک مساله چالش برانگیز پیش‌بینی سری‌های زمانی است، زیرا بازار سهام اساساً یک سیستم غیرخطی، پویا و پرسر و صدا و آشفته است. در واقع، قیمت سهام تحت تأثیر عوامل زیادی مانند رویدادهای سیاسی، سیاست‌ها و اخبار شرکت، شرایط اقتصادی، نرخ بهره و احساسات سرمایه‌گذاران قرار می‌گیرد. اخیراً، بسیاری از محققان انواع مختلفی از مدل‌های یادگیری ماشینی را برای پیش‌بینی بازار سهام به کار برده‌اند و نتایج رضایت‌بخشی مانند رگرسیون بردار پشتیبان و پیش‌بینی تصادفی را تولید کرده‌اند (وانگ و همکاران، ۲۰۲۰). شبکه‌های عصبی مصنوعی به عنوان هسته فناوری یادگیری عمیق نیز به طور گسترده برای پیش‌بینی بازار سهام مورد استفاده قرار گرفته‌اند. در میان تمام فناوری‌های یادگیری عمیق، شبکه عصبی $LSTM$ پرسپترون چند لایه عمیق و شبکه عصبی کانولوشن اغلب در پیش‌بینی سری‌های زمانی مالی استفاده می‌شوند. با توجه به برتری مدل‌های یادگیری ماشینی در پیش‌بینی بازار سهام، بسیاری از محققان این مدل‌ها را قبل از تشکیل سبد سهام در فرآیند پیش‌انتخاب سهام اعمال می‌کنند که نتایج رضایت‌بخشی ایجاد می‌کنند (ژانگ و همکاران، ۲۰۱۸). در واقع، پیش‌انتخاب سهام با کیفیت بالا برای مدیریت موفق پرتفوی بسیار مهم است. در بازار سهام، سرمایه‌گذاران فردی معمولاً سعی می‌کنند بازده آتی سهام سرمایه‌گذاری خود را تعیین کنند و سپس وزن بهینه هر سهم را برای ساختن پرتفوی تعیین کنند (نایک، ۲۰۱۹).

بنابراین، پس از فرآیند پیش‌انتخاب سهام، سرمایه‌گذاران همچنین باید وزن سرمایه‌گذاری بهینه را برای هر سهم انتخابی، قبل از انجام سرمایه‌گذاری تجاری محاسبه کنند (چانگ، ۲۰۱۹). این روش عمدتاً بر اساس تئوری مدرن پورتفولیو است. تئوری مدرن پرتفوی شامل مدل‌های مختلفی برای محاسبه وزن بهینه پرتفوی هر دارایی است. تحت دارایی‌های داده شده، مدل‌های پورتفولیو برای بهینه‌سازی یک یا چند تابع هدف تحت شرایط محدودیت‌های مختلف استفاده می‌شود (قهرمانی

نهر و همکاران، ۲۰۲۱). با حل مسئله بهینه‌سازی پرتفوی، وزن سرمایه‌گذاری بهینه هر دارایی به دست می‌آید. مدل میانگین واریانس مارکوویتز (MVF) به عنوان سرآغاز تئوری مدرن پورتفولیو، یک مدل بهینه‌سازی پرتفوی را با به حداکثر رساندن بازده مورد انتظار پورتفولیو و به حداقل رساندن ریسک سرمایه‌گذاری پورتفولیو ایجاد می‌کند. این مدل یک مرز کارآمد را تشکیل می‌دهد که نشانگر پرتفوی دارایی است به طوری که ریسک کل را تحت بازده مورد انتظار از پیش تعیین شده کاهش می‌دهد. برای هر سطح از بازده مورد انتظار، مرز کارایی استراتژی سرمایه‌گذاری بهینه را ارائه می‌دهد. با این حال، مدل MV محدودیت‌های زیادی برای کاربرد عملی دارد، مانند محدودیت فرضیه‌ها و پیچیدگی محاسباتی برای دارایی‌های مقیاس بزرگ‌تر. بنابراین، مدل‌های متعددی برای حل این مسائل ارائه شده است. به عنوان مثال، کونو و یامازاکی (۱۹۹۱) مدل میانگین انحراف مطلق (MAD) را توسعه دادند، که در آن از انحراف مطلق برای جایگزینی واریانس به عنوان معیار ریسک استفاده کردند. الکساندر و باپتیستا (۲۰۰۲) یک مدل ارزش متوسط در معرض خطر (Var) با ترکیب مدل MV با مدل Var را پیشنهاد کردند که در آن Var ضرر تحمل‌شده را تحت یک سطح اطمینان معین تخمین می‌زند. در واقع، پیش‌انتخاب سهام با کیفیت بالا برای موفقیت مدیریت سبد سهام بسیار مهم است (وانگ و همکاران، ۲۰۲۰).

اهمیت بهینه‌سازی سبد سهام منجر به آن شده است تا در این مقاله یک مدل ترکیبی از مدل MVF و OF به نام $MOVVF$ ارائه گردد. این مدل به طور همزمان به پیشینه‌سازی بازده سبد سهام و کمیته‌سازی ریسک سرمایه‌گذاری با اجتناب از رشد شارپی قیمت سهام می‌پردازد. برای این منظور از داده‌های جمع‌آوری شده شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۹۰ تا ۱۳۹۹ استفاده شده است. داده‌های جمع‌آوری شده شامل سه دسته ورودی متغیرهای تکنیکال، متغیرهای بنیادی و سری زمانی ۱۰۰ روز اخیر قیمت سهام می‌باشد. آموزش داده‌ها و تایین سیگنال خرید سهام توسط یک روش ترکیبی انجام می‌پذیرد. در

مرحله اول فیلتر کردن شرکت‌های فعال بورسی از طریق ۶ اندیکاتور صورت می‌پذیرد و در مرحله دوم، از یک ماشین یادگیری جمعی ۲ سطحی استفاده می‌شود. بهینه‌سازی سبد سهام براساس سیگنال‌های خرید صادر شده از ماشین یادگیری جمعی ۲ سطحی و اندیکاتورهای به کار رفته، با استفاده از ۳ روش *NSGA II*, *H&B* و *MOGWO* انجام می‌پذیرد. در بخش دوم به بررسی پیشینه تحقیقات مرتبط با بهینه‌سازی سبد سهام با روش‌های مختلف پرداخته شده است. در بخش سوم به ارائه مدل‌های بهینه‌سازی سبد سهام شامل مدل میانگین-واریانس مارکوویتز (*MVF*)، نسبت امگا (*OF*) و مدل ترکیبی (*MVOF*) و همچنین روشهای ماشین یادگیری جمعی دو سطحی و الگوریتم‌های *NSGA II* و *MOGWO* نیز تشریح شده است. در بخش چهارم به تجزیه و تحلیل و ارائه نتایج حاصل از پیاده‌سازی رویکرد بر روی داده‌های شرکت‌های بورسی فعال در ایران پرداخته شده است. در بخش پنجم به نتیجه‌گیری و ارائه پیشنهاد‌های اتی تحقیق پرداخته شده است.

۲- پیشینه تحقیق

محققان بیشماری به مدل‌سازی و حل مسئله بهینه سبد پرداخته‌اند. از جمله اولین محققانی که به بهینه‌سازی سبد سهام پرداخت پرلود (۱۹۸۴) بود. پرلود (۱۹۸۴) یک الگوریتم عملی را برای بهینه‌سازی پرتفوی میانگین واریانس در مقیاس بزرگ، با تاکید بر توسعه یک رویکرد محاسباتی کارآمد برای طیف وسیعی از مدل‌های پرتفوی مورد استفاده توسط جامعه سرمایه‌گذاری، توصیف کرد. او نشان داد که چگونه الگوریتم پارامتریک را می‌توان با پراکنده کردن ماتریس کوواریانس با معرفی چند متغیر و محدودیت اضافی، کارآمد کرد و اینکه چگونه می‌توان این دو رویکرد به ظاهر نامرتب را با هم ترکیب کرد. فریتس و همکاران (۲۰۰۹) یک مدل بهینه‌سازی پورتفولیوی مبتنی بر پیش‌بینی جدید را ارائه کردند که می‌تواند فرصت‌های سرمایه‌گذاری کوتاه‌مدت را جذب کند. آنها از پیش‌بینی کننده‌های شبکه عصبی برای پیش‌بینی بازده سهام استفاده کردند و مجموعه بزرگی از آزمایش‌ها را با

داده‌های واقعی از بازار سهام برزیل برای بررسی مدل بهینه‌سازی پرتفوی خود انجام دادند. نتایج نشان داد که مدل بهینه‌سازی پرتفوی مبتنی بر پیش‌بینی از فرصت‌های کوتاه‌مدت استفاده کرده و از مدل میانگین واریانس بهتر عمل کرده و شاخص بازار را شکست داد. اوی (۲۰۱۶) کاربرد تخمین کوواریانس معکوس پراکنده را برای بهینه‌سازی پرتفوی حداقل واریانس مارکوویتز، با استفاده از کمند گرافیکی ار پیشنهاد کرد. نتایج تجربی نشان داد که کمند گرافیکی تمایل دارد تا عناصر مورب ماتریس کوواریانس معکوس برآورد شده را با افزایش منظم‌سازی بیش از حد تخمین بزند. او برای رفع این مشکل، یک روش جدید جهت تنظیم منظم‌سازی بهینه معرفی کرد و عملکرد بر روی داده‌های مصنوعی و واقعی بازار سهام را با برآوردهای کوواریانس موجود در ادبیات مقایسه کرد و نشان داد که استراتژی‌های جدید سبدگردانی عملکرد خوبی دارند. چونگ و همکاران (۲۰۱۷) با استفاده از روش ترکیبی خوشه بندی و الگوریتم ژنتیک به بهینه‌سازی سبد سهام پرداختند. آنها ابتدا سبدهای از سبد سهام را از طریق تجزیه خوشه‌ای اطلاعات سرمایه گذاران انتخاب کردند. سپس با استفاده از الگوریتم ژنتیک به بهینه‌سازی وزن سهام انتخاب شده بر اساس پیشینه‌سازی بازده سبد سهام پرداختند.

بن و همکاران (۲۰۱۸) دو روش یادگیری ماشینی، منظم‌سازی و اعتبارسنجی متقابل را برای بهینه‌سازی پورتفولیو تطبیق دادند. آنها منظم‌سازی مبتنی بر عملکرد (*PBR*) را معرفی کردند که راه‌حل را به سمت راه‌حل مرتبط با خطای برآورد کمتر در عملکرد هدایت می‌کند. آنها نشان می‌دادند که این مدل‌ها را می‌توان به‌عنوان مسائل بهینه‌سازی قوی با مجموعه‌های عدم قطعیت جدید انتخاب کرد. تا و همکاران (۲۰۱۸) در مقاله‌ای برای نشان دادن اینکه چگونه تکنیک‌های یادگیری ماشین می‌توانند معاملات کمی را برآورده کنند، از رگرسیون خطی و مدل‌های رگرسیون برداری پشتیبانی برای پیش‌بینی حرکت سهام و چندین تکنیک بهینه‌سازی برای بهینه‌سازی بازده و کنترل ریسک در معاملات، استفاده کردند. با وجود شکاف بین مدل‌سازی پیش‌بینی و



یادگیری تقویتی مکرر پرداختند. آنها الگوریتم یادگیری تقویتی را با اقدامات معاملاتی مداوم بر روی دارایی‌های متعدد پیاده‌سازی کردند و محدودیت‌های پورتفولیو و فرابارامترهای شبکه خط مشی را به طور همزمان کنترل نمودند. شبکه شکل ساعت شنی به عنوان یک انتخاب طبیعی برای مدیریت دارایی ظاهر شد.

هان و همکاران (۲۰۲۱) در مطالعه‌ای عملکرد یادگیری ماشین و یادگیری عمیق را در پیش انتخاب سهام مورد مقایسه قرار دادند. آنها در تشکیل پورتفولیو، پیش‌بینی بازگشت یادگیری ماشین و یادگیری عمیق را با تاکید بر پیشرفت بهینه‌سازی پورتفولیو همراه با پیش‌بینی بازده، ترکیب کردند. نتایج نشان داد مدل پیشرفته میانگین واریانس با پیش‌بینی تصادفی جنگل بهترین عملکرد را دارد. چن و همکاران (۲۰۲۱) در مطالعه‌ای پیش‌بینی سهام را در انتخاب نمونه کارها ادغام کردند تا ویژگی‌های آینده را به تصویر بکشند. آنها یک مدل ترکیبی به نام *IFAXGBoost* برای پیش‌بینی قیمت سهام توسعه دادند. مدل پیش‌بینی *IFAXGBoost* در مدل میانگین واریانس گنجانده شده است. آزمایش‌های گسترده اثربخشی روش پیشنهادی را نشان داد. کاجمارک و همکاران (۲۰۲۱) در مطالعه‌ای نشان دادند که تکنیک‌های بهینه‌سازی پورتفولیو ارایه شده توسط بهینه‌سازهای میانگین واریانس مارکوویتز و برابری ریسک سلسله مراتبی (*HRP*)، بازده با ریسک تعدیل شده پرتفوی‌های ساخته شده با سهام از پیش انتخاب شده با ابزار یادگیری ماشین را افزایش می‌دهد. آنها از روش جنگل تصادفی برای پیش‌بینی مقطع بازده اضافی مورد انتظار استفاده کردند. سن و همکاران (۲۰۲۱)، تحلیلی از سری زمانی قیمت‌های تاریخی پنج سهام برتر از ۹ بخش مختلف بازار سهام هند را از ۱ ژانویه ۲۰۱۶ تا ۳۱ دسامبر ۲۰۲۰ انجام دادند و پورتفولیو بهینه برای هر یک را ساختند. بازده واقعی و پیش‌بینی‌شده هر یک از پورتفولیوها بالا هستند که نشان‌دهنده دقت بالای مدل است. سامی (۲۰۲۱) در مطالعه‌ای طبقه بندی‌های متعددی را بر اساس فرآیندهای یادگیری ماشین بدون نظارت به منظور تعیین رضایت بخش دارایی‌ها یا اوراق بهادار قابل سرمایه‌گذاری برای

معاملات واقعی، استراتژی معاملاتی پیشنهادی به بازدهی بالاتری دست یافت. پاپویا و همکاران (۲۰۱۹) یک مدل تصمیم‌گیری برای سرمایه‌گذاری‌های تجاری روزانه در بازار سهام پیشنهاد کردند. در این راستا رویکرد تلفیقی بین، ماشین بردار پشتیبان (*SVM*) و انتخاب پورتفولیو نتایج قابل توجهی را نشان داد. از جمله اینکه هر چه سود هدف تعریف شده بیشتر باشد، عملکرد *SVM* بهتر است. هزینه‌های دلالی می‌تواند یک محدودیت قوی برای امکان سنجی مدل پیشنهادی باشد. سولین و همکاران (۲۰۱۹) با استفاده از شبکه عصبی مصنوعی و الگوریتم ژنتیک به بهینه‌سازی سبد سهام پرداختند. هدف آنها پیش‌بینی ارزش سهام آینده با استفاده از شبکه عصبی مصنوعی و سپس بهینه‌سازی سبد سهام با الگوریتم ژنتیک با هدف دستیابی به بیشترین بازده سبد سهام و حداقل ارزش ریسک بود.

پرین و همکاران (۲۰۲۰) در مطالعه‌ای بررسی کردند که چگونه تخصیص پورتفولیو می‌تواند از توسعه الگوریتم‌های بهینه‌سازی پرتفوی در مقیاس بزرگ سود ببرد. آنها برنامه‌های بهینه‌سازی پورتفولیو پیچیده‌تر با تابع هدف غیر درجه دوم، منظم‌سازی با توابع جریمه و محدودیت‌های غیرخطی را در نظر گرفتند و سه مدل اصلی پرتفوی بتای هوشمند را مورد بحث قرار دادند: پرتفوی سهم ریسک برابر، سبد بودجه بندی ریسک، و متنوع‌ترین پرتفوی. ژانگ و همکاران (۲۰۲۰) در مطالعه‌ای مدل‌های یادگیری عمیق را برای بهینه‌سازی مستقیم اتخاذ کردند. آنها مستقیماً وزن پرتفو را با به‌روزرسانی پارامترهای مدل بهینه کردند و به جای انتخاب دارایی‌های منفرد، وجوه قابل مبادله شاخص‌های بازار را برای تشکیل یک سبد سهام معامله کردند. مبادله شاخص‌های طبقات مختلف دارایی، به طور قابل ملاحظه‌ای طیف دارایی‌های موجود را کاهش می‌داد. نتایج نشان داد که این مدل بهترین عملکرد را در دوره آزمایشی ۲۰۱۱ تا پایان آوریل ۲۰۲۰ داشته است. آنها یک تحلیل حساسیت برای روشن کردن ارتباط ویژگی‌های ورودی نیز انجام دادند. ابوصلاح و همکاران (۲۰۲۰) در مقاله‌ای به بررسی گنجاندن محدودیت‌های پورتفولیو در چارچوب

سهام پرتفوی، ارائه کرد. او پورتفولیوی پیشنهادی را با یک نمونه کار تصادفی برای یک بازه زمانی خاص، به منظور تعیین بازده پرتفوی، نسبت شارپ و عملکرد پرتفوی، مقایسه نمود.

شوندنر و همکاران (۲۰۲۱) یک چارچوب مفهومی به نام برابری ریسک متوالی تطبیقی (ASRP) برای گسترش برابری ریسک سلسله مراتبی (HRP) به عنوان یک تخصیص ابتکاری دارایی ارائه کردند. آنها عملکرد HRP استاندارد را با سایر روش‌های مبتنی بر درخت استاتیک و تطبیقی، و همچنین روش‌های مبتنی بر سری‌بندی که به درخت‌ها متکی نیستند، مقایسه کردند. تجزیه و تحلیل‌ها نشان داد که بیشتر جایگزین‌های مبتنی بر درخت برای HRP از خوشه‌بندی تک پیوندی استفاده شده در HRP بر اساس ریسک تعدیل شده بهتر عمل می‌کنند. پراساد و همکاران (۲۰۲۱) در مطالعه‌ای از دارایی‌های بورس اوراق بهادار نپال (NEPSE) برای ساخت پرتفو استفاده کردند و از چارچوب مدرن واریانس میانگین مبتنی بر نظریه پورتفولیو برای یافتن وزن بهینه برای تخصیص به هر دارایی در پرتفو استفاده نمودند. آنها سعی کردند محدودیت‌ها و کاستی‌های نظریه پورتفولیو مدرن را که پایه‌ای برای کارهای تحقیقاتی آینده برای حل مسئله بهینه‌سازی با استفاده از تکنیک‌های یادگیری ماشین خواهد بود، بررسی کنند. کانلن و همکاران (۲۰۲۱) یادگیری ماشین و بهینه‌سازی پورتفولیوی مبتنی بر عامل را بررسی کردند. آنها دریافتند که عوامل مبتنی بر شبکه‌های عصبی رمزگذار خودکار رابطه ضعیف‌تری را با پرتفولیوهای مرتب‌سازی شده با مشخصه‌های رایج نسبت به تکنیک‌های کاهش ابعاد رایج، نشان می‌دهند. جنابی (۲۰۲۱) در مطالعه‌ای یک روش یادگیری ماشینی قوی را برای انتخاب سبد بهینه کالا و در چارچوب ارزش در معرض خطر (LVAR) تعدیل شده با نقدینگی، ارائه کرد. بدین منظور او از الگوریتم مدل‌سازی ریسک بازار و تجزیه و تحلیل سبد سرمایه‌گذاری با استفاده از تکنیک‌های یادگیری ماشینی تقویتی استفاده کرد. ساکار و همکاران (۲۰۲۱) مطالعه‌ای با هدف ایجاد تعادل بین جنبه‌های اقتصاد و هوش محاسباتی انجام دادند. آنها برتری PSO را

برای بهینه‌سازی سبد سهام، پیش‌بینی قیمت و روند سهام، و سایر جنبه‌های مرتبط با بازار سهام به همراه پیامدهای PSO را مورد تجزیه و تحلیل قرار دادند. مین و همکاران (۲۰۲۱) در مقاله‌ای مدل‌های پرتفوی قوی هیبریدی تحت مجموعه‌های عدم قطعیت بیضی مشترک را پیشنهاد کردند و معیارهای ریسک متفاوتی در مدل‌های مقاوم هیبریدی را در نظر گرفتند. آنها XGBoost و LSTM را برای یادگیری فرآیندها اعمال نموده و یک الگوریتم خوشه بندی برای توصیف مجموعه‌های عدم قطعیت توسعه دادند.

بر اساس مطالعات صورت گرفته ویژگی‌های اصلی مقاله به شرح زیر است:

- استفاده از یک مدل ترکیبی بهینه‌سازی سبد سهام
- استفاده از یک روش ترکیبی بر مبنای ماشین یادگیری جمعی در ارائه سیگنال خرید
- استفاده از الگوریتم‌های فرا ابتکاری در حل مسئله

۳- روش تحقیق

در این مقاله به بهینه‌سازی سبد سهام شرکت‌های فعال پذیرفته شده در بورس اوراق بهادار تهران پرداخته شده است. مدل مورد نظر برای بهینه‌سازی سبد سهام متشکل از ترکیب دو مدل مختلف در سرمایه‌گذاری است. از این رو مدل توسعه داده شده برای بهینه‌سازی سبد سهام شامل ترکیب مدل‌های میانگین - واریانس (MV) و نسبت امگا (OF) است که در این مقاله به نام اختصاری (MVOF) شناخته می‌شود. در این مدل توابع هدف به کار رفته شامل بیشینه‌سازی بازده پیش بینی سبد سهام، کمینه‌سازی ریسک سرمایه‌گذاری در سبد سهام و بیشینه‌سازی بازده غیر عادی سبد سهام است. برای این امر، اطلاعات ۴۸۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران جمع‌آوری و جهت تجزیه و تحلیل انتخاب شده است.

با توجه به وجود شرکت‌های مختلف در بورس اوراق بهادار تهران ابتدا شرکت‌های فعال بر اساس فیلترهایی که



ایران طی ۱۰ سال (۱۳۹۰ تا ۱۳۹۹) می‌پردازد، لذا \hat{r}_i بازده پیش‌بینی شده سهم i در دوره زمانی (هر ماه) و $\bar{\varepsilon}_i$ میانگین خطای پیش‌بینی سرمایه‌گذاری در ۶۰ ماه منتهی به سال‌های (۱۳۹۵ تا ۱۳۹۹) تعریف می‌گردد. مقدار خطای پیش‌بینی سرمایه‌گذاری از رابطه $\varepsilon_i = r_i - \hat{r}_i$ بدست می‌آید؛ که در آن r_i مقدار بازده واقعی سهم i در دوره مورد بررسی است. در نتیجه این امر میتوان بیان کرد رابطه (۱) به کمینه‌سازی ریسک سرمایه‌گذاری سبد سهام، رابطه (۲) به بیشینه‌سازی بازده پیش‌بینی شده سبد سهام، رابطه (۳) به بیشینه‌سازی بازده غیر عادی سبد سهام می‌پردازد. لذا رابطه (۲) و (۳) به طور توأم در صد نزدیکی بازده پیش‌بینی شده به بازده واقعی است. رابطه (۴) و (۵) تضمین می‌کند مجموع درصد سرمایه‌گذاری در سبد سهام از مقدار ۱ بیشتر نشود.

مدل امگا با پیش‌بینی

نسبت امگا (OF) اولین بار توسط کیتینگ و شادویک (۲۰۰۲) معرفی شد. سپس، به طور گسترده‌ای برای ایجاد سبد سهام استفاده می‌شود، زیرا از محدودیت‌های شناخته شده نسبت شارپ اجتناب می‌کند. نسبت امگا به صورت زیر تعریف می‌شود:

$$\omega = \frac{E(y_i) - \tau}{E[\tau - y_i]^+} + 1 \quad (6)$$

در رابطه فوق τ آستانه تقسیم بازده به مورد انتظار (درآمد) و غیر منتظره (زیان) را نشان می‌دهد و اغلب توسط سرمایه‌گذاران تصمیم می‌گیرد. y_i بازگشت تصادفی سهم i است. از آنجایی که نسبت امگا به توزیع احتمال بازده دارایی نیاز دارد، راه حل به‌دست‌آمده زمانی که این توزیع احتمال نادقیق باشد، به سوگیری و بیش از حد خوش‌بینانه تبدیل می‌شود (کاپسوس و همکاران، ۲۰۱۴). بنابراین، کپسوس و همکاران. (۲۰۱۴) بدترین نسبت امگا را برای حل این مشکل معرفی کرد و مدل امگا را به شرح زیر می‌باشد:

$$\max \psi \quad (7)$$

s. t.:

سیگنال خرید از آنها صادر شده است، انتخاب و سپس بر اساس مدل $MVOF$ به بهینه‌سازی سبد سهام پرداخته خواهد شد. در ادامه به تشریح مدل‌های میانگین واریانس (MV)، مدل نسبت امگا (OF) و مدل ترکیبی ($MVOF$) پرداخته شده است.

۳-۱- مدل‌های بهینه‌سازی سبد سهام

مدل میانگین واریانس با پیش‌بینی

مارکوویتز (۱۹۵۹) به عنوان پیشرو نظریه مالی مدرن، مدل میانگین واریانس (MV) را معرفی کرد که یک راه حل ریاضی برای بهینه‌سازی بین حداکثر کردن بازده سبد مورد انتظار و حداقل رساندن ریسک سرمایه‌گذاری در سبد سهام بود. با توسعه مدل مارکوویتز مدل‌های مختلفی پدید آمد که در آن یک چارچوب جدید توسط یو و همکاران (۲۰۲۰) ارائه شد. در این مدل توسعه یافته پیش‌بینی بازگشت سرمایه نیز در مدل MV در نظر گرفته شده است که به مدل (MVF) شهرت یافته است. در نتیجه مدل بهینه‌سازی چند هدفه MVF به شرح زیر ارائه شده است:

$$\min \sum_{i=1}^n \sum_{j=1}^n X_i X_j \sigma_{i,j} \quad (1)$$

$$\max \sum_{i=1}^n X_i \hat{r}_i \quad (2)$$

$$\max \sum_{i=1}^n X_i \bar{\varepsilon}_i \quad (3)$$

s. t.:

$$\sum_{i=1}^n X_i = 1 \quad (4)$$

$$0 \leq X_i \leq 1, \quad i = 1, 2, \dots, n \quad (5)$$

در مدل فوق X_j درصد سرمایه‌گذاری در سهم i و n تعداد کل سهام موجود برای سرمایه‌گذاری است. همچنین $\sigma_{i,j}$ کواریانس بین دو سهم i و j ، \hat{r}_i بازده پیش‌بینی شده سهم i و در نهایت $\bar{\varepsilon}_i$ میانگین خطای پیش‌بینی شده از سرمایه‌گذاری در سهم i است. از آنجایی که این مقاله به بررسی سرمایه‌گذاری در سهام شرکت‌های فعال بورس



۱۳۹۹ بهره گرفته شده است.

۳-۲- فیلتر کردن شرکت‌های فعال بورسی

در این روش انتخاب سبد سهام و پیش‌بینی روند بازار در روز آینده (روند بازار نزولی، یکنواخت، صعودی) بر اساس سه نوع داده ورودی خواهد بود:

۱- سری زمانی بازدهی هر شرکت در ۱۰۰ روز

منتهی به روز جاری

۲- متغیرهای تکنیکال در روز جاری (قیمت پایانی،

تعداد خریداران، تعداد دفعات معامله، حجم

معاملات، ارزش معاملات روزانه، ارزش روز

شرکت، نسبت P/E و تعداد سهام هر شرکت)

۳- متغیرهای بنیادی هر شرکت در روز جاری

(قیمت سکه طرح قدیم، سکه طرح جدید، دلار،

شاخص، طلا، نفت، یورو)

بر اساس اطلاعات فوق، شرکت‌هایی که سهام آنها

حداقل یک روز در ماه معامله شده باشد، جهت تجزیه و

تحلیل انتخاب می‌گردد.

۳-۳- آموزش داده‌ها و تعیین سیگنال خرید

بر اساس سه نوع اطلاعات ورودی در بخش قبل، از یک

روش ترکیبی جهت پیش‌بینی روند بازار و همچنین

تعیین سیگنال خرید، فروش و یا نگهداری استفاده شده

است. این روش‌ها شامل:

- بهینه‌سازی قواعد معاملاتی مبتنی بر تحلیل

تکنیکال

- استفاده از ماشین یادگیری جمعی (دو سطحی

$(HHEL)$

مطابق با شکل (۱) و در مرحله اول که بر اساس ۶

اندیکاتور $MACD$ و WMA ، EMA ، SMA ، ROC ، RSI به

فیلتر کردن شرکت‌های فعال بورسی پرداخته می‌شود، در

صورتی که سهام شرکتی حداقل در سه اندیکاتور فوق

سیگنال خرید را صادر کند، آن سهم برای خرید در روز

آتی مورد توجه قرار خواهد گرفت. مطابق با شکل (۲) و در

مرحله دوم نیز از ماشین یادگیری جمعی بر اساس

همجنس و غیر همجنس بودن به پیش‌بینی روند بازار و

$$\delta \left(\sum_{j=1}^n X_j \bar{r}_j^i - \tau \right) - (1 - \delta) \frac{1}{T^i} \sum_{t=1}^{T^i} \eta_t^i \geq \psi \quad (8)$$

$$\eta_t^i \geq - \sum_{j=1}^n X_j \bar{r}_j^i + \tau, \quad \forall t = 1, 2, \dots, T^i, i = 1, 2, \dots, l \quad (9)$$

$$\sum_{j=1}^n X_j = 1 \quad (10)$$

$$\eta_t^i \geq 0, \quad \forall t = 1, 2, \dots, T^i, i = 1, 2, \dots, l \quad (11)$$

$$0 \leq X_j \leq 1, \quad \forall j = 1, 2, \dots, n \quad (12)$$

در روابط فوق X_j درصدی از سبد سهام تخصیص یافته

به سهم i ، η_t^i متغیر کمکی جهت خطی‌سازی مدل سبد

سهام، δ نشان دهنده ترجیح بازده ریسک مدل است. در

این مقاله مقادیر δ ، T^i و τ به ترتیب برابر با ۰.۵، ۰.۶ و

ماه، ۰ و ۱ در نظر گرفته شده است. با ترکیب دو مدل

فوق، می‌توان مدل نهایی را به شرح زیر بیان کرد:

$$\max \psi \quad (13)$$

$$\max \sum_{i=1}^n X_i \hat{r}_i \quad (14)$$

$$\max \sum_{i=1}^n X_i \bar{\varepsilon}_i \quad (15)$$

s. t.:

$$0.5 \left(\sum_{i=1}^n X_j \hat{r}_i \right) - \frac{1}{120} \sum_{t=1}^{60} \eta_t \geq \psi \quad (16)$$

$$\eta_t \geq - \sum_{i=1}^n X_j \hat{r}_i, \quad \forall t = 1, 2, \dots, 60 \quad (17)$$

$$\sum_{i=1}^n X_i = 1 \quad (18)$$

$$\eta_t \geq 0, \quad \forall t = 1, 2, \dots, 60 \quad (19)$$

$$0 \leq X_i \leq 1, \quad \forall i = 1, 2, \dots, n \quad (20)$$

با توجه به مدل فوق، در این مقاله به بهینه‌سازی سبد

سهام بر اساس مدل ترکیبی میانگین- واریانس - امگا

$(MVOF)$ برای شرکت‌های فعال پذیرفته شده در بورس

اوراق بآدار تهران پرداخته شده است. در ادامه به منظور

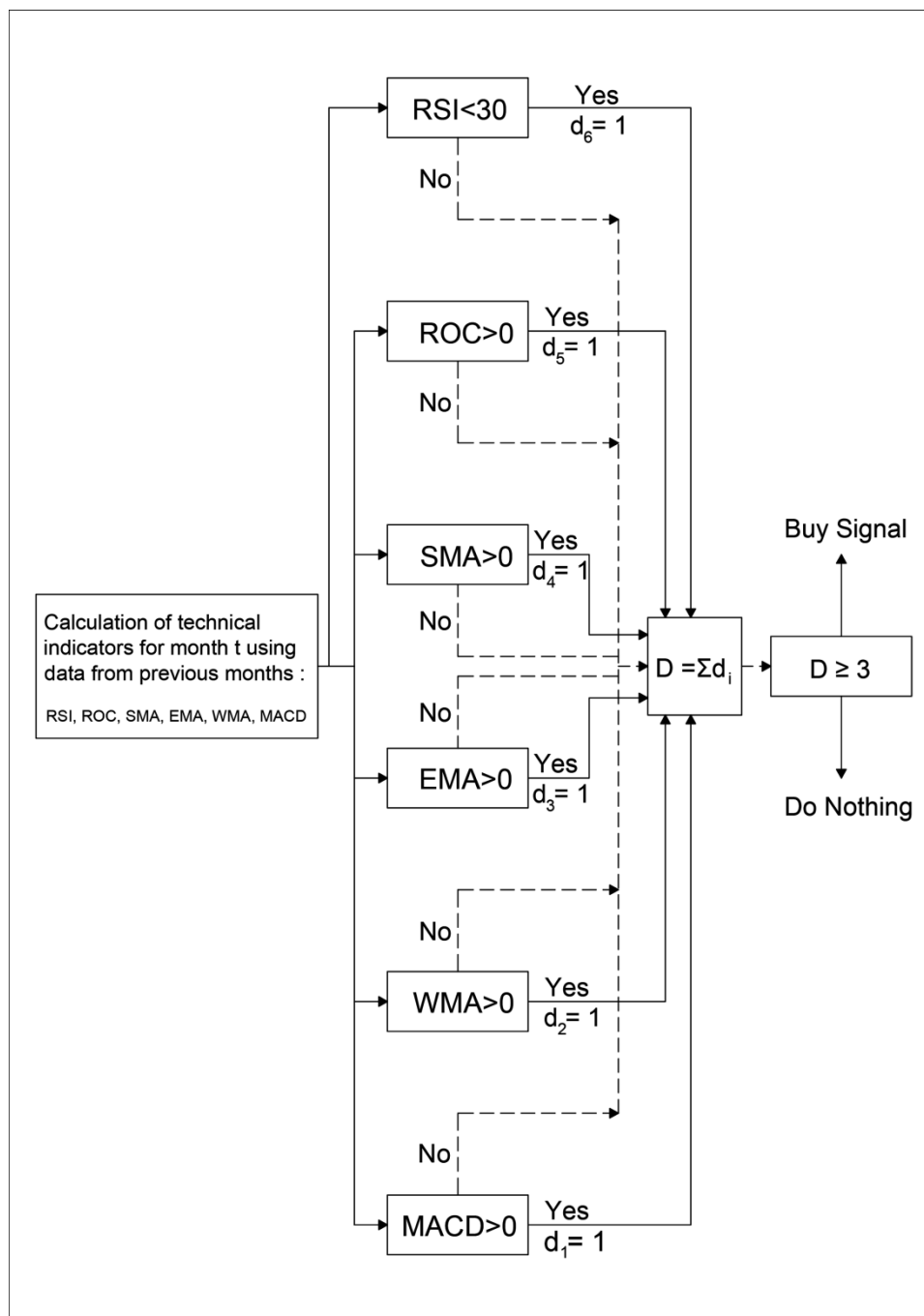
فیلتر کردن شرکت‌های فعال بورس اوراق بهادا تهران از

اطلاعات ۴۸۰ شرکت پذیرفته شده طی سال‌های ۱۳۹۰ تا

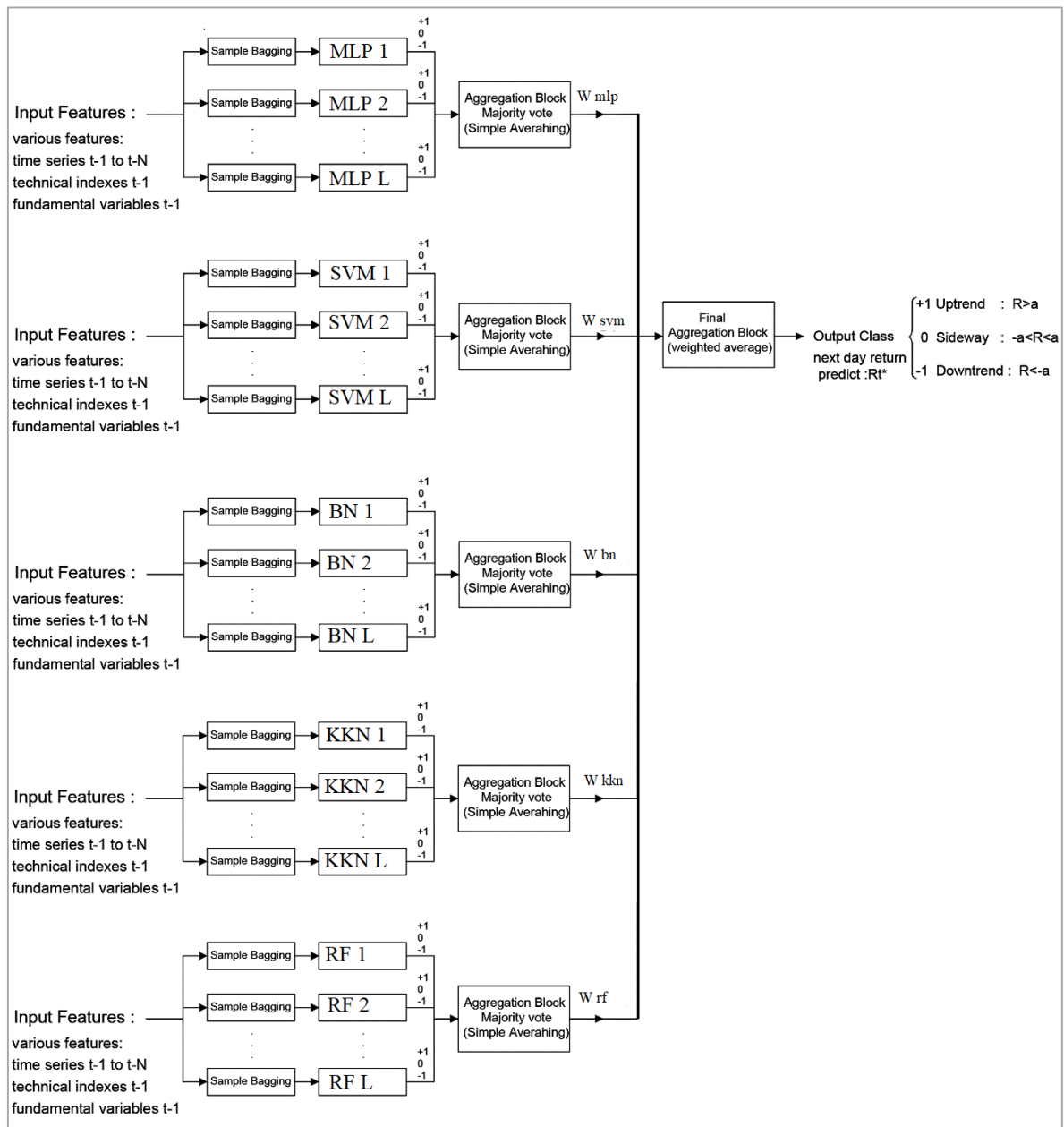


موازی در هر بخش استفاده شده است. این امر منجر به کاهش خطای در ارائه سیگنال خرید می‌گردد. در ادامه بر اساس تجمیع نتایج روش‌های یادگیری، سیگنال نهایی خرید، فروش و یا نگهداری صادر میگردد و در این بخش نیز روش رای اکثریت (میانگین وزنی) به عنوان مدل نهایی یادگیری جمعی دو سطحی استفاده می‌شود.

فیلتر کردن نهایی سهام‌ها جهت ارائه سیگنال خرید پرداخته خواهد شد. در این ماشین یادگیری از ۵ نوع مدل یادگیری غیر همگن نظیر (ماشین بردار پشتیبان، جنگل تصادفی، K - نزدیک‌ترین همسایه، شبکه عصبی MLP و شبکه بیزین) استفاده شده است. برای کاهش خطای محاسباتی در این ماشین یادگیری از L مدل یادگیری



شکل ۱- بهینه سازی قواعد معاملاتی مبتنی بر تحلیل تکنیکال



شکل ۲- استفاده از ماشین یادگیری جمعی (دو سطحی HHEL)

که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده‌ها به وسیله روش‌های QP که روش‌های شناخته شده‌ای در حل مسائل محدودیت‌دار هستند صورت می‌گیرد. قبل از تقسیم خطی برای اینکه ماشین بتواند داده‌های با پیچیدگی بالا را دسته‌بندی کند داده‌ها به وسیله تابع ρ به فضای با ابعاد خیلی بالاتر برده می‌شود.

• روش یادگیری ماشین بردار پشتیبان (SVM)
 ماشین بردار پشتیبان یکی از روش‌های یادگیری بانظارت است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌کنند. این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی نشان داده‌است. مبنای کاری دسته‌بندی کننده SVM دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب گردد

• جنگل تصادفی (RF)

جنگل تصادفی یک روش یادگیری ترکیبی برای دسته‌بندی، رگرسیون می‌باشد، که بر اساس ساختاری متشکل از شمار بسیاری درخت تصمیم، بر روی زمان آموزش و خروجی کلاس‌ها (کلاس‌بندی) یا برای پیش‌بینی‌های هر درخت به شکل مجزا، کار می‌کنند. جنگل‌های تصادفی برای درختان تصمیم که در مجموعه آموزشی دچار بیش برآزش می‌شوند، مناسب هستند. عملکرد جنگل تصادفی معمولاً بهتر از درخت تصمیم است، اما این بهبود عملکرد تا حدی به نوع داده هم بستگی دارد.

از تکنیک یادگیری نظارت شده به نام بازپرداخت برای آموزش استفاده می‌کند. لایه‌های متعدد آن و فعال‌سازی غیر خطی آن MLP را از یک پرسپترون خطی متمایز می‌کند. در واقع می‌تواند داده‌هایی را متمایز کند که به صورت خطی قابل تفکیک نیستند. یادگیری در شبکه عصبی با تغییر وزن اتصال پس از پردازش هر قطعه از داده‌ها، براساس میزان خطا در خروجی در مقایسه با نتیجه مورد انتظار رخ می‌دهد. این نمونه که از یادگیری با نظارت و از طریق بازگشت به عقب و تعمیم الگوریتم حداقل مربعات در پرسپترون خطی انجام می‌شود.

• شبکه بیزین (BN)

ایده کلی و مقدماتی درباره‌ی شبکه‌های بیزی، که در واقع ترکیبی از دو شاخه نظریه گراف و نظریه احتمال هستند، بیان می‌شود. این شبکه‌ها عمدتاً نشان دهنده روابط علی و معلولی میان متغیرها می‌باشند. ساختار گراف یک شبکه بیزی برای صورتبندی توزیع احتمال توأم‌متغیرهای شبکه بکار می‌رود. هنگامیکه ساختار گراف معلوم باشد، مدل‌های احتمالاتی میتوانند برای استدلال و پیش‌بینی در مورد متغیرها بکار روند و در صورت نامشخص بودن ساختار گراف، با استفاده از این مدل‌ها میتوان به یادگیری ساختار مدل پرداخت و آنگاه استدلال و پیش‌بینی در مورد متغیرها را انجام داد. برای یادگیری خودکار ساختار شبکه بیزی در الگوریتم‌های یادگیری ماشین، یکی از روش‌های اصلی بر پایه تعیین وابستگی بین متغیرها بنا نهاده شده است. برای تعیین این وابستگی‌ها معیارهای زیادی مانند معیار آنتروپی طراحی شده اند. با استفاده از این معیارها، وابستگی هر متغیر نسبت به متغیرهای دیگر سنجیده می‌شود و آنها که معیار وابستگی شان از یک آستانه‌ای بیشتر بود، به عنوان والد‌ها انتخاب می‌شود. پس از تعیین ساختار، اگر مقدار همه متغیرها به طور کامل قابل مشاهده باشند، از تخمین احتمال معمولی شامل گرفتن تعدادی نمونه و شمردن تعداد اتفاقات یک رویداد در این مجموعه نمونه استفاده می‌شود. اگر بعضی از متغیرها قابل مشاهده نباشند، با استفاده از آموزش یک شبکه نورونی می‌توان مقادیر جداول احتمالات شرطی را یاد گرفت.

• K- نزدیک‌ترین همسایه (KNN)

K- نزدیک‌ترین همسایه یک متد آمار ناپارامتری است که برای طبقه‌بندی آماری و رگرسیون استفاده می‌شود. در هر دو حالت K شامل نزدیک‌ترین مثال آموزشی در فضای داده‌ای می‌باشد و خروجی آن بسته به نوع مورد استفاده در طبقه‌بندی و رگرسیون متغیر است. در حالت طبقه بندی با توجه به مقدار مشخص شده برای K، به محاسبه فاصله نقطه‌ای که میخواهیم برچسب آن را مشخص کنیم با نزدیک‌ترین نقاط میپردازد و با توجه به تعداد رای حداکثری این نقاط همسایه، در رابطه با برچسب نقطه مورد نظر تصمیم‌گیری می‌کنیم. برای محاسبه این فاصله میتوان از روش‌های مختلفی استفاده کرد که یکی از مطرح‌ترین این روش‌ها، فاصله اقلیدسی است. در حالت رگرسیون نیز میانگین مقادیر بدست آمده از کی خروجی آن می‌باشد. از آنجا که محاسبات این الگوریتم بر اساس فاصله است نرمال‌سازی داده‌ها می‌تواند به بهبود عملکرد آن کمک کند.

• شبکه عصبی MLP

پرسپترون چند لایه، دسته‌ای از شبکه‌های عصبی مصنوعی پیشخور است. یک MLP شامل حداقل سه لایه گره است: یک لایه ورودی، یک لایه پنهان و یک لایه خروجی. به جز گره‌های ورودی، هر گره یک نورون است که از یک تابع فعال‌سازی غیر خطی استفاده می‌کند. MLP



۳-۳- روش‌های بهینه سازی سبد سهام

پس از ارائه سیگنال خرید، فروش و یا نگهداری توسط روش ترکیبی ارائه شده، بایستی بهینه‌سازی سبد سهام بر اساس مدل *MVOF* صورت پذیرد. برای این امر سه روش خرید و نگهداری (*H&B*)، الگوریتم *NSGA II* و الگوریتم *MOGWO* به کار برده شده است.

• روش خرید و نگهداری (*H&B*)

در این روش بدون در نظر گرفتن تغییر در کاهش یا افزایش بازده سهام قیمت، در ابتدای سال اقدام به خرید سهم شده و تا انتهای افق برنامه ریزی مورد مطالعه نگهداری می‌شود.

• الگوریتم *NSGA II*

یکی از کارآمدترین و مشهورترین الگوریتم‌های بهینه‌سازی چندهدفه، الگوریتم ژنتیک مرتب‌سازی نا مغلوب *NSGA II* است. این الگوریتم یکی از سریع‌ترین و توانمندترین الگوریتم‌های بهینه‌سازی است که نسبت به سایر روش‌ها از پیچیدگی عملیاتی کمتری برخوردار بوده و با استفاده از اصل عدم تسلط (عدم غلبه کردن) و محاسبه فاصله ازدحام نقاط بهینه پارتو را به دست می‌آورد که از گستردگی مطلوبی در حوزه تغییرات توابع اهداف برخوردارند و به طراح، آزادی انتخاب طراحی موردنظر خود را از میان طراحی‌های بهینه‌شده می‌دهد. در *NSGA II*، به‌طور هم‌زمان حفظ نخبه‌گرایی و پراکندگی مدنظر قرار گرفته است. انتخاب جمعیت جدید در هر گام این روش بر اصل غلبگی استوار بوده و با استفاده از نخبه‌گرایی و رتبه‌بندی جمعیت در هر گام حل، بهترین جواب‌های نا مغلوب را برگزیده و به گام بعد می‌رود. اگر دو تابع هدف بهینه‌سازی f_1 و f_2 وجود داشته باشند، در آن صورت برای دو جواب x و y ، جواب x ، جواب y را مغلوب می‌کند اگر داشته باشیم $f_1(x) \geq f_1(y)$ و $f_2(x) > f_2(y)$ یا اینکه $f_1(x) > f_1(y)$ و $f_2(x) \geq f_2(y)$ همچنین برای رعایت توزیع مناسب چگالی جواب‌ها در این الگوریتم از مفهومی با عنوان فاصله ازدحام استفاده می‌شود. به‌طور کلی برای مرتب کردن جمعیتی با

اندازه n بر اساس سطوح نا مغلوب‌ها، هر جواب با تمام جواب‌های دیگر موجود در جمعیت مقایسه می‌شود تا مشخص شود که آن جواب مغلوب می‌شود یا خیر. در نهایت یک تعداد حل وجود دارد که هیچ‌کدام غالب و مغلوب همدیگر نمی‌شوند لذا این جواب‌ها، اولین مرز از مرزهای نا مغلوب را تشکیل می‌دهند. این جواب‌ها به مجموعه F_1 انتقال داده می‌شوند. برای تعیین جواب‌های موجود در مرزهای بعدی، جواب‌های موجود در مرز اول به‌طور موقت نادیده گرفته می‌شود و فرایند فوق دوباره تکرار می‌شود و این بار جواب‌ها به مجموعه F_2 منتقل شده و رتبه دوم را کسب می‌کنند. این روند برای همه جواب‌های نا مغلوب جمعیت ادامه دارد. یکی از معیارهای موردنظر الگوریتم تکاملی در راه رسیدن به مرز بهینه پارتو، حفظ تنوع و گستردگی جواب‌ها در مجموعه جواب‌های به‌دست‌آمده می‌باشد. در واقع مرتب کردن غیر مغلوب‌ها رویه‌ای است در جهت رسیدن به جواب‌های بهتر و مکانیسم تنوع هم درصدد حفظ تنوع و گستردگی در این جواب‌ها می‌باشد. در این الگوریتم این کار توسط فاصله ازدحام به این صورت انجام می‌شود. مقدار کمتر فاصله ازدحامی یک جواب، بیان‌کننده تراکم بیشتر جواب‌ها در اطراف آن است. برای مرحله بعد، جواب‌هایی که در ناحیه‌ای با تراکم کم‌تر یا به عبارت دیگر بافاصله ازدحامی بیشتر هستند، انتخاب شوند. با این کار تنوع و پراکندگی در جواب‌های به‌دست‌آمده بیشتر می‌شود. هدف از بکار گیری فاصله ازدحام در *NSGA II*، ایجاد تنوع در جواب‌های جمعیت بوده و نشان‌دهنده میزان تراکم جواب‌ها در کنار یک جواب مشخص است. فاصله ازدحام برای جواب‌های مرتب‌شده به‌صورت صعودی و مختص به مجموعه F از رابطه (۲۱) به دست می‌آید.

$$CD(X^1) = CD(X^S) = \infty$$

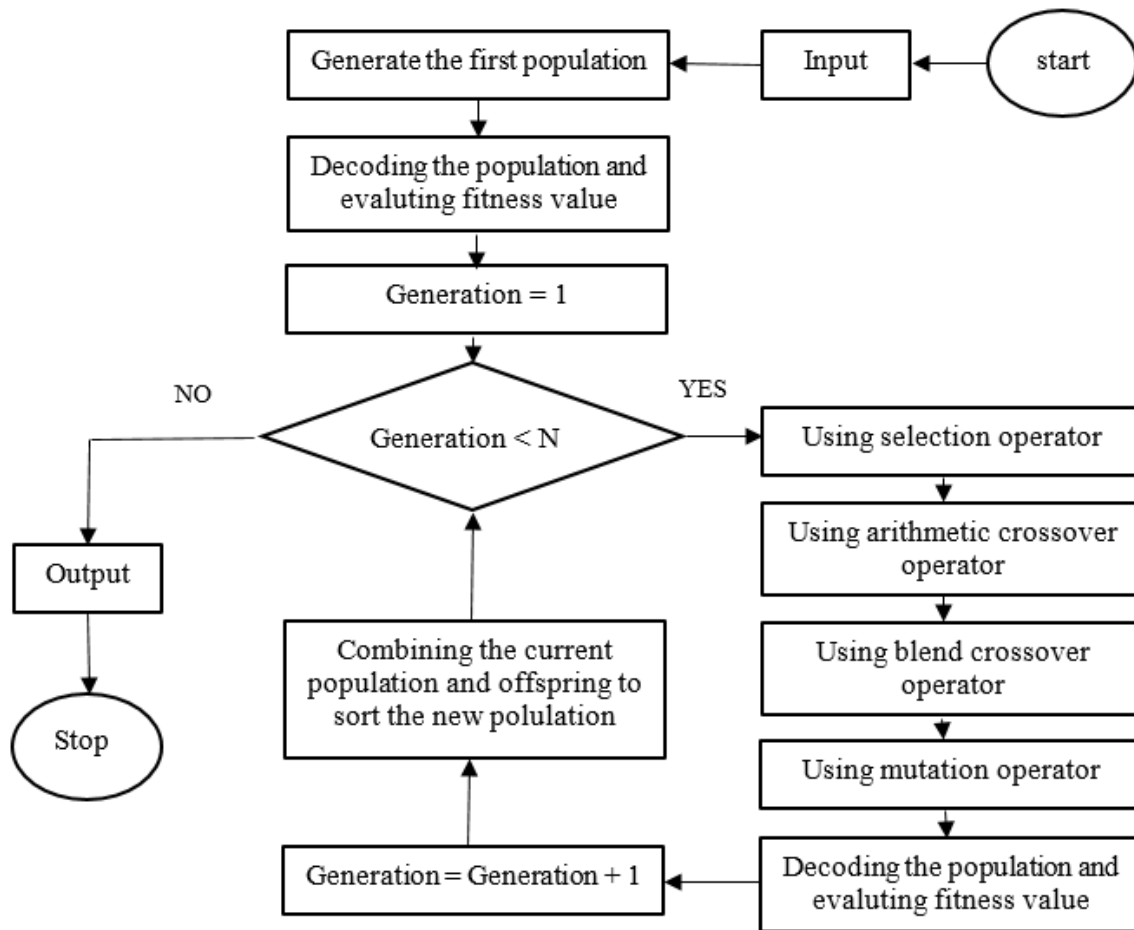
$$CD(X^i) = \left[\frac{Z_1(X^{i+1}) - Z_1(X^{i-1})}{Z_1(X^S) - Z_1(X^1)} \right] + \left[\frac{Z_2(X^{i+1}) - Z_2(X^{i-1})}{Z_2(X^S) - Z_2(X^1)} \right], i = 2, \dots, S - 1 \quad (21)$$

در رابطه بالا، $CD(X^i)$ میزان فاصله ازدحام برای جواب X^i است. پس از ادغام جمعیت والدین و فرزندان،



گام ۵: قرار دهید $R_t = P_t \cup Q_t$
 گام ۶: از روش رتبه بندی نا مغلوب برای تعیین مجموعه های پارتو F_i در جمعیت R_t استفاده کنید،
 گام ۷: قرار دهید $P_{t+1} = \emptyset$ و $i = 1$
 گام ۸: تا زمانی که $|P_{t+1}| + |F_i| < N$
 الف. جواب های مجموعه F_i را به جمعیت P_{t+1} اضافه کنید،
 ب. قرار دهید $i = i + 1$
 گام ۹: جواب های مجموعه F_i را بر اساس فاصله ازدحام و به صورت نزولی مرتب کنید،
 گام ۱۰: به اندازه $N - |P_{t+1}|$ از اولین جواب های F_i به جمعیت P_{t+1} انتقال دهید، و
 گام ۱۱: قرار دهید $t = t + 1$ و به گام ۲ بازگردید.

مرتب سازی نا مغلوب انجام شده و گام های ۷ و ۸ که در زیر توضیح داده می شوند اجرا می شود. بر اساس گام ۱۰، برای ایجاد یک زیرمجموعه از آخرین مجموعه نا مغلوب و به علت زیاد شدن اندازه جمعیت بعدی، از معیار فاصله ازدحام استفاده می شود:
 گام ۱: جمعیت اولیه P_0 به اندازه N با جواب های تصادفی ایجاد کنید و $t = 0$ قرار دهید،
 گام ۲: در صورت عدم برقراری شرط توقف، به P_t بازگردید،
 گام ۳: با استفاده از عملگر انتخاب مسابقه ای دودویی، N والد از جمعیت P_t انتخاب کنید،
 گام ۴: با اعمال عملگرهای تقاطع و جهش بر روی جمعیت P_t ، جمعیت فرزندان Q_t به اندازه N ایجاد کنید،



شکل ۳- شبه کد الگوریتم NSGA II

• الگوریتم *MOGWO*

راه حل‌های مناسب هستند. بقیه راه حل‌های کاندید به صورت $\Omega(X)$ فرض می‌شوند. برای انجام شکار، گرگ‌های خاکستری باید طعمه را پیدا و محاصره کنند. لذا معادلات زیر موقعیت‌های گرگ‌ها را در اطراف طعمه به روز می‌کند.

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \quad (22)$$

$$\vec{X}(t+1) = \vec{X}(t) - \vec{A} \cdot \vec{D} \quad (23)$$

در رواب فوق \vec{C} و \vec{A} بردارهای ضرایب هستند. \vec{X}_p بردار موقعیت شکار و \vec{X} بردار موقعیت گرگ‌های خاکستری است. این یک معادله تعادلی بین محاصره و شکار است. بنابراین شعاع جستجو باید طی فرایند بهینه‌سازی شود، برای این منظور معادلات مربوط به دو ضریب استفاده شده در روابط فوق به شرح زیر است.

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (24)$$

$$\vec{C} = 2\vec{r}_2 \quad (25)$$

معادلات فوق، گرگ‌های خاکستری را قادر می‌سازد موقعیت خود را در اطراف طعمه به روز کنند. در نتیجه برای انجام شکار، معادلات زیر استفاده می‌شود.

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \quad (26)$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha, \vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta, \vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \quad (27)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (28)$$

۴- تجزیه و تحلیل نتایج

پس از ارائه روش ترکیبی ماشین یادگیری جمعی و همچنین روش‌های بهینه سازی سبد سهام بر اساس مدل ترکیبی *MVOF*، در این بخش به تجزیه و تحلیل نتایج در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخته شده است. بر این اساس ۴۸۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۹۰ تا ۱۳۹۹ انتخاب و داده‌های مرتبط با سه دسته متغیر (سری زمانی بازدهی هر شرکت در ۱۰۰ روز منتهی به روز جاری، متغیرهای تکنیکال در روز جاری (قیمت پایانی، تعداد خریداران، تعداد دفعات معامله، حجم معاملات،

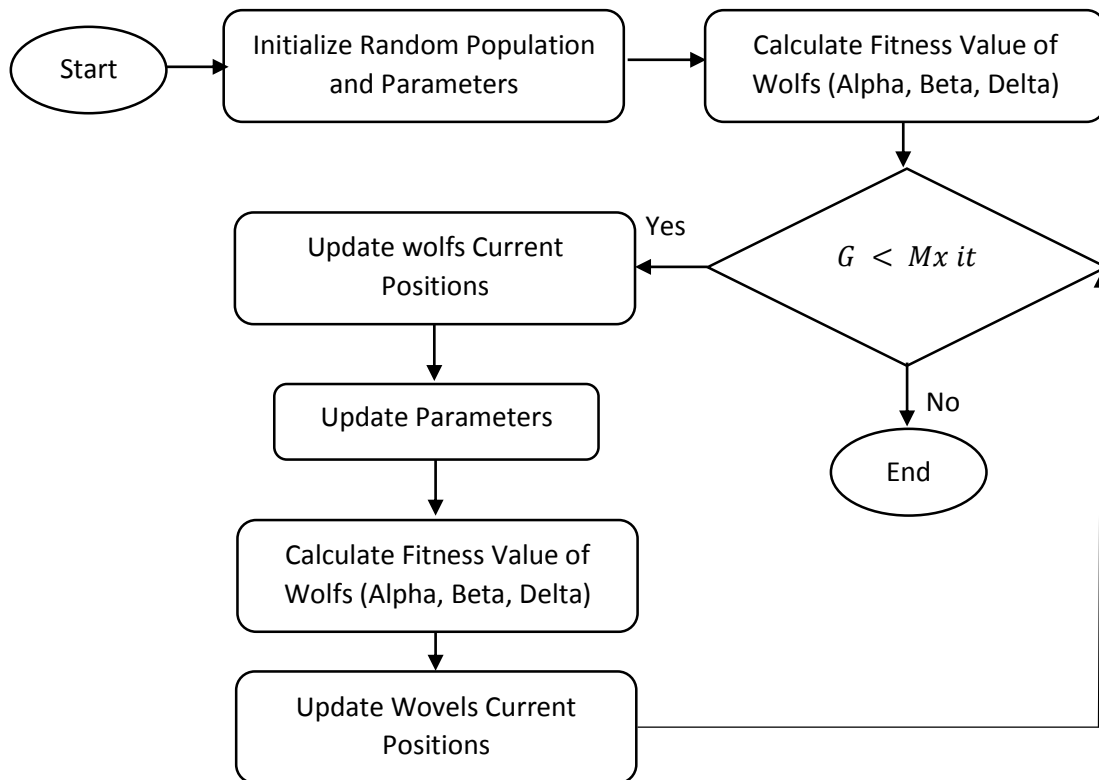
گرگ خاکستری *Canis Iupus* متعلق به خانواده *Candidae* است. گرگ‌های خاکستری شکارچیان در راس هرم غذایی می‌باشند، به این معنی که آنها در راس زنجیره غذایی می‌باشند. گرگ‌های خاکستری عمدتاً ترجیح می‌دهند تا به طور گروهی زندگی کنند. اندازه گروه به طور متوسط ۵-۱۲ گرگ است. رهبران شامل یک نر و یک ماده می‌باشند که به آنها *Alpha* گفته می‌شود. *Alpha* مسئول اصلی تصمیم‌گیری‌ها در مورد شکار، محل خواب، زمان بیدار شدن و ... است. تصمیمات *Alpha* به گروه اعلام می‌شود؛ با این حال برخی از رفتارهای دموکراتیک نیز مشاهده شده است که در آن یک *Alpha* از سایر گرگ‌ها در گروه تبعیت می‌کند. در اجتماعات، کل گله با پایین نگه داشتن خود، *Alpha* را تایید می‌کنند. گرگ *Alpha* مرسوم به گرگ غالب نیز است، زیرا دستورات رو باید توسط گروه اجرا شود. گرگ‌های *alpha* تنها مجاز به جفت‌گیری در گله می‌باشند. نکته قابل توجه این است که *Alpha* لزوماً قوی‌ترین عضو گله نیست، بلکه بهترین عضو از نظر مدیریت در گله است. دومین سطح در سلسله مراتب گرگ‌های خاکستری *Beta* است. *Beta* گرگ‌های زیر دستی هستند که به *Alpha* در تصمیم‌گیری یا سایر تصمیمات گله کمک می‌کنند. گرگ *Beta* میتواند نر یا ماده باشد و او بهترین جایگزین *Alpha* در صورت مرگ یا پیر شدن او است. *Beta* دستورات و فرمان‌های *Alpha* را در سراسر گله اجرا کرده و بازخوردها را به *Alpha* می‌دهد. گرگ *Omega* پای‌ترین طبقه در سلسله مراتب گرگ خاکستری است. گرگ *Omega* نقش قربانی را ایفا می‌کند. معمولاً گرگ‌های *Omega* باید از همه گرگ‌های سطح بالا و غالب پیروی کنند. آنها آخرین گرگ‌هایی هستند که اجازه غذا خوردن دارند. در صورتی که گرگ یک *Alpha* یا *Omega* نباشد، به او *Delta* می‌گویند. گرگ‌های *Delta* باید تابع *Alpha* و *Beta* باشند. با این حال آنها بر *Omega* غالب هستند.

در مدل‌سازی ریاضی سلسله مراتب اجتماعی گرگ‌ها، α (*Alpha*) به عنوان مناسب‌ترین راه حل در نظر گرفته می‌شود. متعاقباً β (*Beta*) و δ (*Delta*) دومین و سومین



اخیر بوده و حداقل یک معامله در سهام شرکت در طول یک ماه انجام شده است؛ برای آموزش داده‌ها انتخاب گردیده است. از این رو ۲۰۸ شرکت دارای شرایط مذکور بوده و جهت آموزش داده‌ها انتخاب شده اند.

ارزش معاملات روزانه، ارزش روز شرکت، نسبت P/E و تعداد سهام هر شرکت) و متغیرهای بنیادی هر شرکت در روز جاری (قیمت سکه طرح قدیم، سکه طرح جدید، دلار، شاخص، طلا، نفت، یورو) جمع‌آوری شده است. از این بین شرکتهایی که دارای بیشترین اطلاعات طی ۱۰ سال



شکل ۴- شبه کد الگوریتم MOGWO

فعال بورسی، ۸۵ شرکت فعال نهایی شده انتخاب شده اند. برای دستیابی به تعداد طبقات مناسب روش‌های یادگیری، تعداد طبقات ۳، ۴، ۵، ۶ و ۷ در نظر گرفته شده و تعداد یادگیرنده‌ها در هر طبقه ۵۰ لحاظ گردیده است. همچنین مقدار آستانه تعیین روند صعودی و نزولی بودن به ترتیب ۲٪ سود و ۲٪ ضرر در نظر گرفته شده است. جدول (۱) همبستگی بین داده‌های آموزش دیده و پیش بینی شده به دست آمده از یادگیری ماشین جمعی را در تعداد طبقات مختلف نشان می‌دهد.

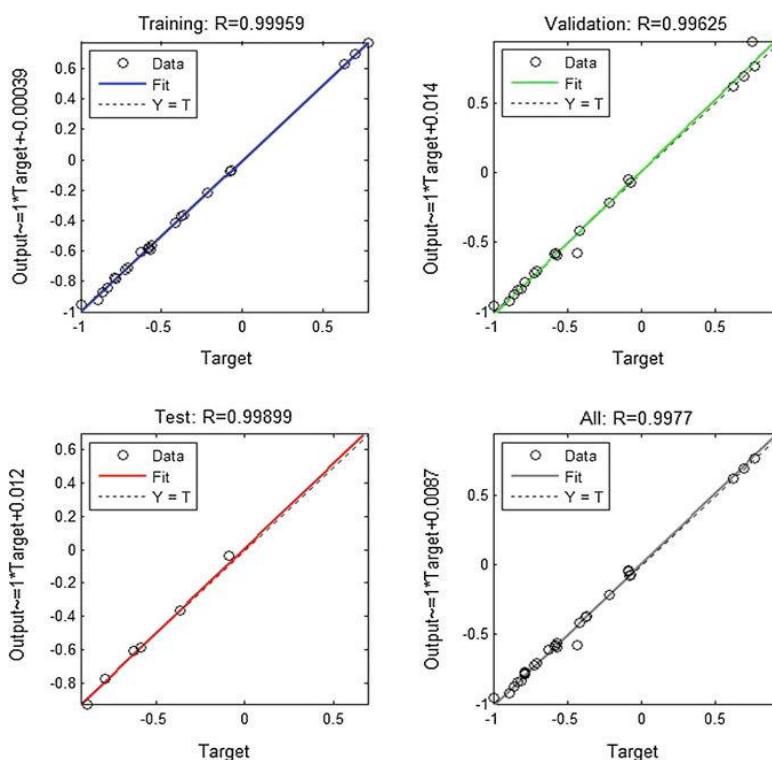
بر اساس جدول (۱) مشاهده می‌شود زمانی که تعداد طبقات برابر با ۷ باشد، همبستگی بین داده‌های آموزش دیده و پیش‌بینی شده به دست آمده از یادگیری ماشین

آموزش داده‌ها و ارائه سیگنال خرید توسط شرکت‌ها برای ۵ سال اول (۱۳۹۰ تا ۱۳۹۴) صورت پذیرفته است. در این بخش یک روش ترکیبی برای این امر به کار رفته است. در مرحله اول، به فیلتر کردن ۲۰۸ شرکت بر اساس روش بهینه سازی قواعد معاملاتی مبتنی بر تحلیل تکنیکال (۶ اندیکاتور RSI ، ROC ، SMA ، EMA ، WMA و $MACD$) پرداخته شده است. شرکتهایی که حداقل در سه اندیکاتور فوق سیگنال خرید را صادر کرده باشند، جهت آموزش انتخاب می‌گردند. از سویی یک ماشین یادگیری جمعی دو سطحی ($HHEL$) متشکل از L روش یادگیری (SVM ، RF ، BN و KNN) جهت آموزش داده‌ها به کار برده شده است. با فیلتر کردن شرکت‌های

جمعیتی دو سطحی در بهترین حالت خود و برابر با ۰,۹۹۷۷ قرار می‌گیرد. شکل (۵) نتایج یادگیری ماشین بر اساس روش *HHEL* نشان می‌دهد.

جدول ۱- مقایسه همبستگی بین داده‌های آموزش دیده و پیش بینی شده در طبقات مختلف یادگیری

طبقه	Validation	Training	Test	All
۳	۰,۷۵۷۶۵	۰,۹۱۳۸۱	۰,۵۰۷۵۷	۰,۸۵۳۶
۴	۱	۰,۹۹۶۳۹	۰,۸۳۳۶	۰,۹۷۹۲
۵	۰,۹۹۰۱۷	۰,۹۹۶۹۸	۰,۹۵۹۶۶	۰,۹۹۲۱
۶	۰,۹۹۰۵۹	۰,۹۹۴۹۱	۰,۹۸۹۲۹	۰,۹۹۳۱
۷	۰,۹۹۶۲۵	۰,۹۹۹۵۹	۰,۹۹۸۹۹	۰,۹۹۷۷



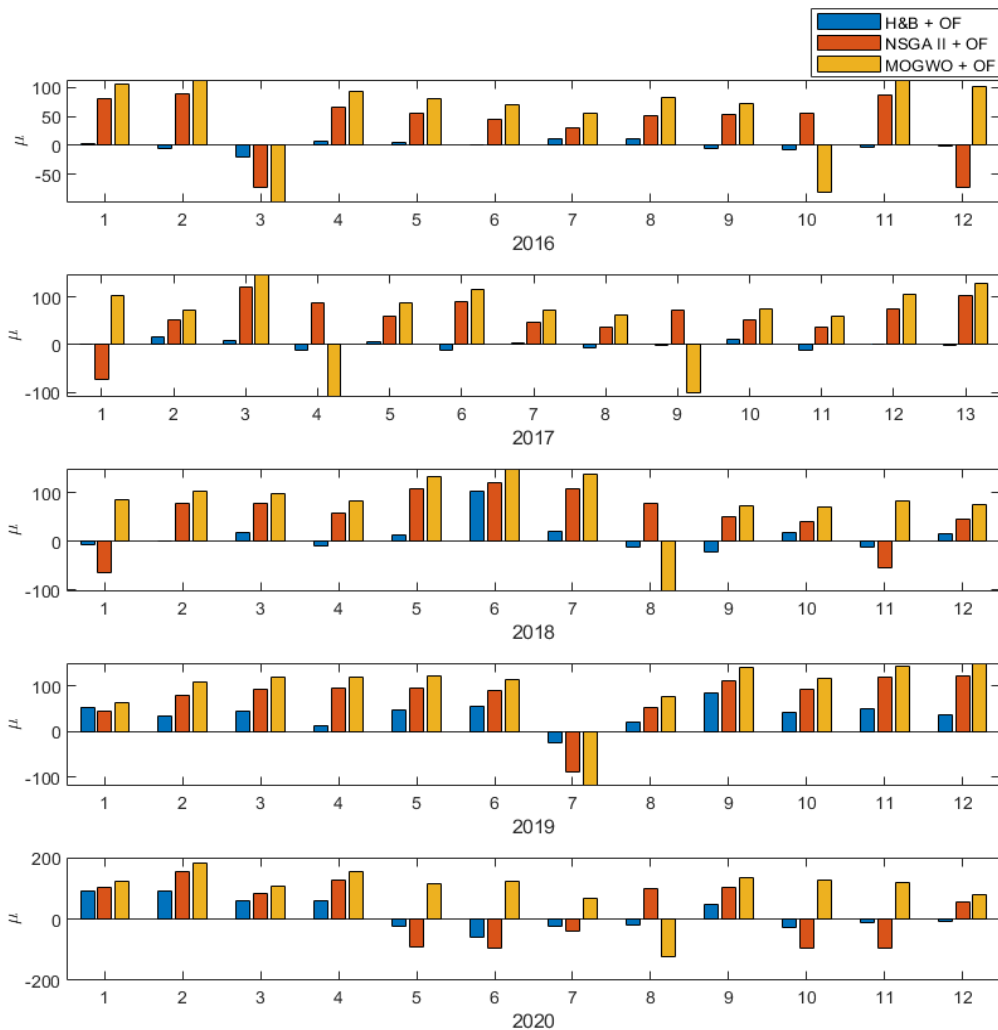
شکل ۵- نمودار رگرسیون خطی برای آموزش، ارزیابی و تست داده‌ها با روش *HHEL*

با توجه به حجم داده‌های ورودی، داده‌ها به خوبی آموزش دیده اند و یادگیری ماشین نتایج بسیار دقیقی نیز ارائه می‌دهد. به طوری که خطای آن در حدود صفر متقارن است و همچنین همبستگی بین داده‌های آموزش دیده و پیش‌بینی شده به دست آمده از یادگیری ماشین جمعی دو سطحی مطابق با شکل (۵) تقریباً برابر با ۰,۹۹۷۷ است. پس از آموزش داده‌های ۸۵ شرکت بورسی فیلتر شده در این بخش، به بهینه سازی سبد سهام مطابق با مدل ترکیبی *MVOF* پرداخته شده است. از این رو شکل (۶) تا (۸) میانگین بازده سبد سهام را در ماه‌های مختلف هر سال توسط سه روش

با توجه به حجم داده‌های ورودی، داده‌ها به خوبی آموزش دیده اند و یادگیری ماشین نتایج بسیار دقیقی نیز ارائه می‌دهد. به طوری که خطای آن در حدود صفر متقارن است و همچنین همبستگی بین داده‌های آموزش دیده و پیش‌بینی شده به دست آمده از یادگیری ماشین جمعی دو سطحی مطابق با شکل (۵) تقریباً برابر با ۰,۹۹۷۷ است.

پس از آموزش داده‌های ۸۵ شرکت بورسی فیلتر شده در این بخش، به بهینه سازی سبد سهام مطابق با مدل ترکیبی *MVOF* پرداخته شده است. از این رو شکل (۶) تا (۸) میانگین بازده سبد سهام را در ماه‌های مختلف هر سال توسط سه روش





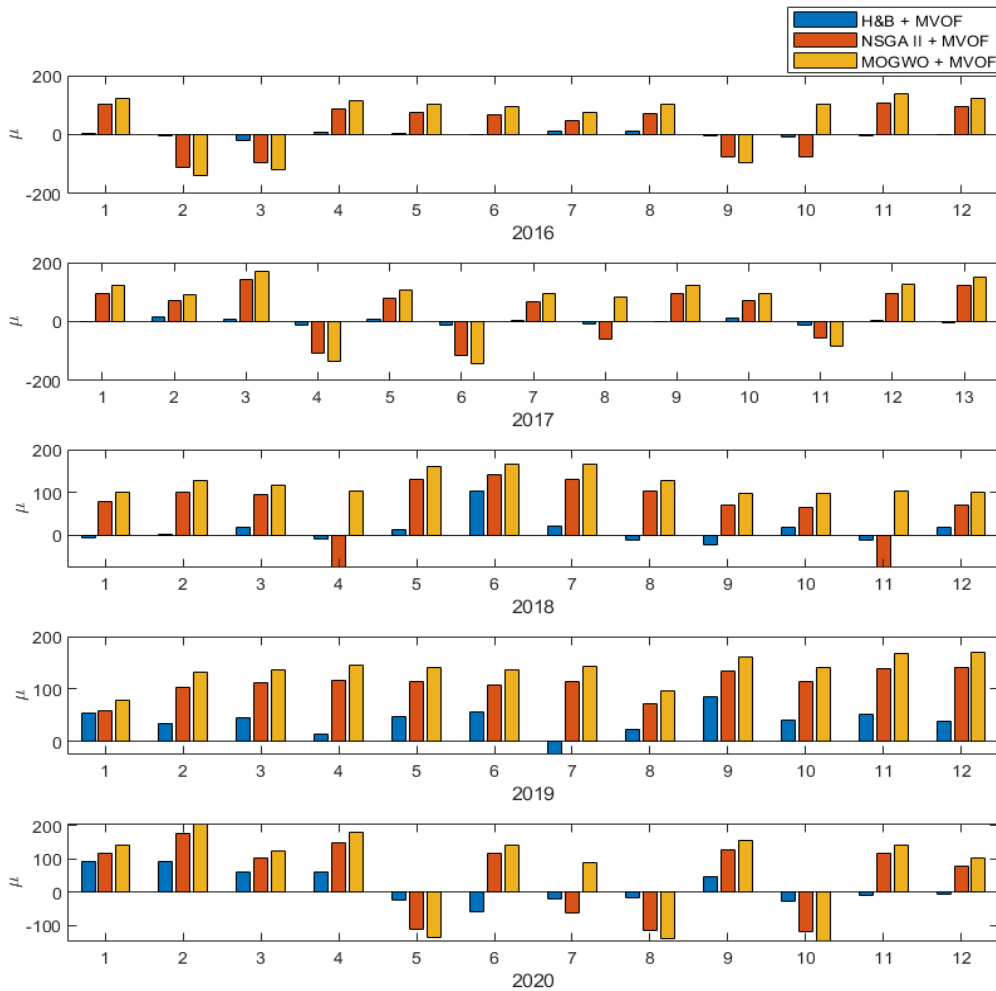
شکل ۶- میانگین بازده سبد سهام در سال‌های ۱۳۹۵ تا ۱۳۹۹ در مدل *OF*

بالاتر از روش *OF* دارد. همچنین میانگین بازده سبد سهام در استفاده از مدل *MVF* برای الگوریتم *NSGA II* مقدار ۵۲,۲۳٪ و الگوریتم *MOGWO* مقدار ۸۰,۵۶ درصد بوده است. این نشان می‌دهد مدل *OF* در بهینه‌سازی سبد سهام از منظر بازده سبد سهام کارایی بالاتری دارد. در این مقاله از روش ترکیبی *MVOF* جهت بهینه‌سازی سبد سهام استفاده شده است که مزیت‌های دو مدل را در بر دارد. بر این اساس میانگین بازده کل سبد سهام در روش *H&B* برابر با ۱۲,۹۸٪، الگوریتم *NSGA II* برابر با ۵۷,۵۹٪ و الگوریتم *MOGWO* برابر با ۸۸,۰۰٪ می‌باشد. با مقایسه بازده سبد سهام در سه مدل می‌توان بیان کرد کارایی مدل *MVOF* بسیار بالاتر از دو مدل دیگر است.

با بررسی مدل *OF* می‌توان بیان کرد، میانگین کل سبد سهام طی ۵ سال در روش *H&B* برابر با ۱۲,۵۲٪ با بازگشت سرمایه ۱۰۲,۴۵٪ ثبت شده است. در این حالت سرمایه‌گذار اقدام به خرید و نگهداری سهام شرکت‌های فیلتر شده کرده و به‌طور مساوی از تمامی سهام‌ها خریداری و نگهداری کرده است. با استفاده از الگوریتم *NSGA II* میانگین نرخ بازده سبد سهام برابر با ۵۲,۸۲٪ و با استفاده از الگوریتم *MOGWO* میانگین بازده سبد سهام برابر با ۷۹,۵۴٪ بدست آمده است. در شکل (۷) و زمانی که استفاده از مدل *MVF* مد نظر است، میانگین کل سبد سهام طی ۵ سال در روش *H&B* برابر با ۱۲,۶۲٪ بدست آمده است که مقداری

(۲) میانگین بازده سبد سهام در روش‌های مختلف بهینه‌سازی به همراه میانگین ریسک و بازده سرمایه نشان داده شده است.

از آنجایی که مدل *OF* در بیشینه سازی بازده سبد سهام از کارایی بالاتری برخوردار است، روش *MVF* نسبت به آن دارای ریسک سرمایه‌گذاری کمتری است. در جدول



شکل ۸- میانگین بازده سبد سهام در سال‌های ۱۳۹۵ تا ۱۳۹۹ در مدل *MVOF*

جدول ۲- میانگین بازده سبد سهام، ریسک سرمایه‌گذاری و بازگشت سرمایه در سال‌های مختلف

<i>MVOF</i>	<i>MVF</i>	<i>OF</i>	روش بهینه‌سازی	مدل
٪ ۱۲٫۹۴	٪ ۱۲٫۶۲	٪ ۱۲٫۵۲	<i>H&B</i>	میانگین بازده سبد سهام
٪ ۵۷٫۵۹	٪ ۵۲٫۲۳	٪ ۵۲٫۸۲	<i>NSGA II</i>	
٪ ۸۸٫۰۰	٪ ۸۰٫۵۶	٪ ۵۴٫۷۹	<i>MOGWO</i>	
٪ ۰٫۹۸	٪ ۱٫۰۴	٪ ۱٫۰۴	<i>H&B</i>	میانگین ریسک سرمایه‌گذاری
٪ ۱٫۶۹	٪ ۱٫۷۰	٪ ۱٫۶۰	<i>NSGA II</i>	
٪ ۳٫۴۲	٪ ۳٫۹۶	٪ ۳٫۴۳	<i>MOGWO</i>	
۱۰۰٫۷۷	۱۰۰٫۷۳	۱۰۲٫۴۵	<i>H&B</i>	میانگین بازگشت سرمایه
۱۱۶٫۴۸	۱۱۵٫۶۴	۱۱۸٫۹۷	<i>NSGA II</i>	
۱۲۸٫۵۹	۱۲۶٫۵۸	۱۲۵٫۰۷	<i>MOGWO</i>	

سبد سهام نیز بالاتر از دیگر سالها بوده است. جهت مقایسه بازده سبد سهام و همچنین ریسک سرکایه گذاری در بین سالهای ۱۳۹۵ تا ۱۳۹۹ شکل های (۱۰) ارائه شده است.

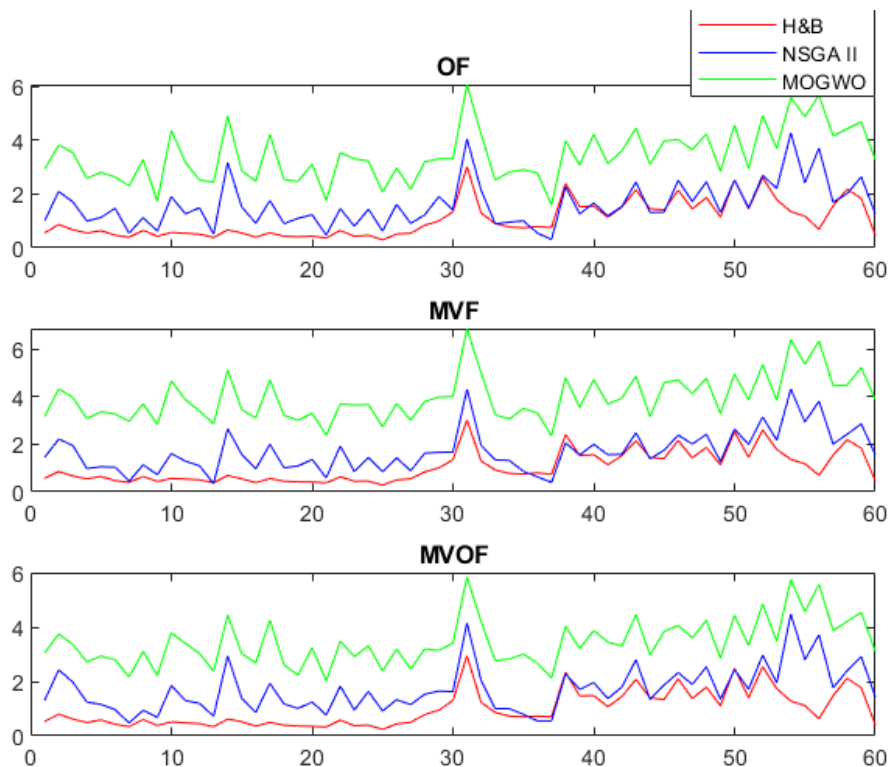
مطابق با نتایج شکل (۱۰) الگوریتم *MOGWO* بالاترین بازده سبد سهام را در کلیه مدل های مورد بررسی شامل *OF*، *MVF* و *MVOF* دارد. این در حالی است که ریسک سرمایه گذاری در استفاده از این روش بالاتر است. جهت انتخاب یکی از سه روش بهینه سازی سبد سهام از روش تصمیم گیری چند شاخصه تاپسیس استفاده شده است. مطابق با نتایج کسب شده و استفاده از روش تصمیم گیری چند هدفه تاپسیس، وزن مطلوبیت هر یک از روش ها در جدول (۳) نشان داده شده است.

مطابق با تحلیل تاپسیس، الگوریتم *NSGA II* کاراترین روش بهینه سازی سبد سهام برای مدل *MVOF* می باشد.

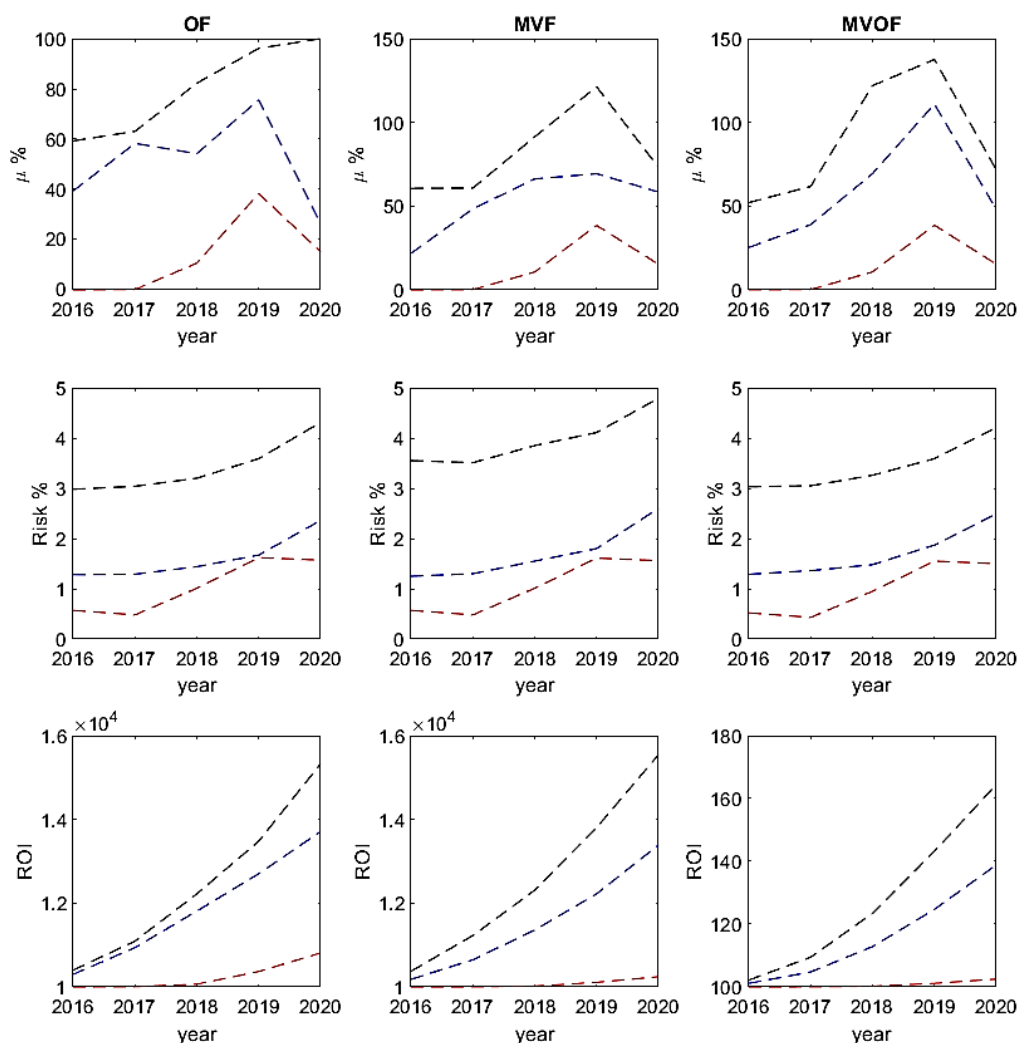
مطابق با جدول (۲) در صورت استفاده از مدل *MVOF* بازدهی سبد سهام افزایش و ریسک سرمایه گذاری در آن کاهش می یابد. این مدل مزیت دو مدل دیگر را با یکدیگر همراه دارد. هم چنین بر اساس نتایج جدول فوق، الگوریتم *MOGWO* بالاتری بازده سبد سهام را بدست آورده است. این در حالی است که ریسک سرمایه گذاری در این حالت برابر با ۳،۴۲٪ است. با بررسی نتایج سه مدل ارائه شده الگوریتم *MOGWO* بالاتری بازده سبد سهام را با بالاترین ریسک سرمایه گذاری بدست آورده است.

شکل (۹) میانگین ریسک سرمایه گذاری را در مدل های مختلف و توسط روش های مختلف بهینه سازی نشان می دهد.

بر اساس شکل (۹) می توان بیان کرد ریسک سرمایه گذاری در بین سالهای ۱۳۹۸ تا ۱۳۹۹ به شدت افزایش یافته و این امر به دلیل رشد شارپی قیمت سهام در شرکت های بورس ایران بوده است. در این سالها بازده



شکل ۹- میانگین ریسک سبد سرمایه گذاری در مدل های مختلف طی سال های ۱۳۹۵ تا ۱۳۹۹



شکل ۱۰- میانگین بازده سبد سهام، ریسک سرمایه گذاری و بازگشت سرمایه در مدل های مختلف

جدول ۳- وزن مطلوبیت روش های بهینه سازی سبد سهام

روش بهینه سازی	میانگین بازده سبد سهام	میانگین ریسک سرمایه گذاری	میانگین بازگشت سرمایه	وزن مطلوبیت
H&B	٪ ۱۲,۹۴	٪ ۰,۹۸	۱۰۰,۷۷	۰,۴۶۱۹
NSGA II	٪ ۵۷,۵۹	٪ ۱,۶۹	۱۱۶,۴۸	۰,۶۴۰۶
MOGWO	٪ ۸۸,۰۰	٪ ۳,۴۲	۱۲۸,۵۹	۰,۵۳۸۱

۵- نتیجه گیری

در این مقاله به بهینه سازی سبد سهام شرکت های فعال پذیرفته شده در اوراق بهادار بورس تهران بر اساس مدل ترکیبی نسبت امگا و میانگین - واریانس مارکوویتز (MVOF) پرداخته شده است. برای این امر ۴۸۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران طی سال های ۱۳۹۰ تا ۱۳۹۹ انتخاب و بر اساس داده های ورودی به

فیلتر کردن شرکت ها پرداخته شد. از این رو یک روش ترکیبی متشکل از روش بهینه سازی قواعد معاملاتی مبتنی بر تحلیل تکنیکال (۶ اندیکاتور RSI , ROC , SMA , EMA , WMA و $MACD$) و ماشین یادگیری جمعی دو سطحی (SVM , RF , BN , MLP و KNN) جهت آموزش داده ها و ارائه سیگنال خرید پرداخته شد. لذا ۸۵ شرکت جهت بهینه سازی سبد سهام انتخاب شدند. برای آموزش



جدید به منظور دستیابی به بازده سبد بالاتر با ریسک پایین تر نیز پیشنهاد می شود.

منابع

- * Aboussalah, A. M., & Lee, C. G. (2020). *Continuous control with stacked deep dynamic recurrent reinforcement learning for portfolio optimization. Expert Systems with Applications, 140, 112891.*
- * Al Janabi, M. A. (2021). *Optimization algorithms and investment portfolio analytics with machine learning techniques under time-varying liquidity constraints. Journal of Modelling in Management.*
- * Alexander, G. J., & Baptista, A. M. (2002). *Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis. Journal of Economic Dynamics and Control, 26, 1159-1193.*
- * Awoye, O. A. (2016). *Markowitz minimum variance portfolio optimization using new machine learning methods (Doctoral dissertation, (UCL) University College London.*
- * Ban, G. Y., El Karoui, N., & Lim, A. E. (2018). *Machine learning and portfolio optimization. Management Science, 64(3), 1136-1154.*
- * Chang, K. H., & Young, M. N. (2019). *Behavioral stock portfolio optimization considering holding periods of B-stocks with short-selling. Computers & Operations Research, 112, 104773.*
- * Chen, W., Zhang, H., Mehlawat, M. K., & Jia, L. (2021). *Mean-variance portfolio optimization using machine learning-based stock price prediction. Applied Soft Computing, 100, 106943.*
- * Cheong, D., Kim, Y. M., Byun, H. W., Oh, K. J., & Kim, T. Y. (2017). *Using genetic algorithm to support clustering-based portfolio optimization by investor information. Applied Soft Computing, 61, 593-602.*
- * Conlon, T., Cotter, J., & Kynigakis, I. (2021). *Machine Learning and Factor-Based Portfolio Optimization. Available at SSRN 3889459.*
- * Freitas, F. D., De Souza, A. F., & de Almeida, A. R. (2009). *Prediction-based portfolio optimization model using neural networks. Neurocomputing, 72(10-12), 2155-2170.*
- * Ghahremani-Nahr, J., Nozari, H., & Sadeghi, M. E. (2021). *Investment modeling to study the performance of dynamic networks of insurance*

داده‌های از ۸۵ شرکت فیلتر شده توسط روش ترکیبی استفاده و تعداد طبقات مختلف با ۵۰ یادگیرنده استفاده شد. در نتیجه این تحلیل مشخص گردید زمانی که تعداد طبقات ۷ باشد همبستگی بین داده‌های آموزش دیده و پیش‌بینی شده بسیار بالا و خطای محاسباتی آن کمتر می‌باشد. لذا بعد از آموزش داده‌های از داده‌های سال‌های ۱۳۹۵ تا ۱۳۹۹ جهت تست داده‌ها و بهینه سازی سبد سهام با مدل *MVOF* استفاده گردید. ابزارهای به کار رفته برای بهینه سازی سبد سهام شامل روش نگهداری و خرید (*H&B*)، الگوریتم *NSGA II* و الگوریتم *MOGWO* بود. نتایج نشان می‌دهد استفاده از مدل *OF* نسبت به مدل *MVF* بالاترین بازده سبد سهام را در طی سال‌های ۱۳۹۵ تا ۱۳۹۹ دارد. در حالی که مدل *MVF* پایین‌ترین میزان ریسک سرمایه‌گذاری را به خود اختصاص داده است. در نتیجه با ترکیب مدل‌های فوق، مشاهده شده بازده سبد سهام در این روش بسیار بالاتر از روش‌های دیگر است. در حالی که ریسک سرمایه‌گذاری آن کمتر بوده است. لذا در صورت استفاده از مدل *MVOF* بازدهی سبد سهام افزایش و ریسک سرمایه‌گذاری در آن کاهش می‌یابد. این مدل مزیت دو مدل دیگر را با یکدیگر همراه دارد. هم چنین بر اساس نتایج جدول فوق، الگوریتم *MOGWO* بالاتری بازده سبد سهام را بدست آورده است. این در حالی است که ریسک سرمایه‌گذاری در این حالت برابر با ۳،۴۲٪ است. همچنین مطابق با نتایج، مشاهده شد الگوریتم *MOGWO* بالاترین بازده سبد سهام را در کلیه مدل‌های مورد بررسی شامل *OF*، *MVF* و *MVOF* دارد. این در حالی است که ریسک سرمایه‌گذاری در استفاده از این روش بالاتر است. جهت انتخاب یکی از سه روش بهینه‌سازی سبد سهام از روش تصمیم‌گیری چند شاخصه تاپسیس استفاده شد. مطابق با تحلیل تاپسیس، الگوریتم *NSGA II* کاراترین روش بهینه‌سازی سبد سهام برای مدل *MVOF* با وزن مطلوبیت ۰،۶۴۰۶ است.

به‌منظور بهبود مقاله پیشنهاد می‌شود از روش‌های یادگیری ماشین متفاوت و استفاده از زنجیره مارکوف جهت پیش‌بینی قیمت سهام و انتخاب شرکت‌های فعال بورسی استفاده شود. همچنین الگوریتم‌های فرا ابتکاری



- companies in Iran. *Modern Research in Performance Evaluation*.
- * Kaczmarek, T., & Perez, K. (2021). Building portfolios based on machine learning predictions. *Economic Research-Ekonomska Istraživanja*, 1-19.
 - * Kapsos, M., Christofides, N., & Rustem, B. (2014). Worst-case robust Omega ratio. *European Journal of Operational Research*, 234, 499-507.
 - * Keating, C., & Shadwick, W. F. (2002). A universal performance measure. *Journal of Performance Measurement*, 6, 59-84.
 - * Konno, H., & Yamazaki, H. (1991). Mean-absolute Deviation Portfolio Optimization model and its applications to Tokyo Stock Market. *Management Science*, 37(5), 519-531.
 - * Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems with Applications*, 165, 113973.
 - * Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. New York: John Wiley Sons Inc
 - * Min, L., Dong, J., Liu, J., & Gong, X. (2021). Robust mean-risk portfolio optimization using machine learning-based trade-off parameter. *Applied Soft Computing*, 113, 107948.
 - * Naik, N., & Mohan, B. R. (2019, February). Optimal feature selection of technical indicator and stock prediction using machine learning technique. In *International Conference on Emerging Technologies in Computer Engineering* (pp. 261-268). Springer, Singapore.
 - * Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. *Expert Systems with Applications*, 115, 635-655.
 - * Perold, A. F. (1984). Large-scale portfolio optimization. *Management science*, 30(10), 1143-1160.
 - * Perrin, S., & Roncalli, T. (2020). Machine learning optimization algorithms & portfolio allocation. *Machine Learning for Asset Management: New Developments and Financial Applications*, 261-328.
 - * Prasad, P. C., Jaiswal, A., Shakya, S., & Singh, S. (2021). Portfolio Optimization: A Study of Nepal Stock Exchange. In *Proceedings of International Conference on Sustainable Expert Systems* (pp. 659-672). Springer, Singapore.
 - * Sami, H. M. (2021). Portfolio Construction Using Financial Ratio Indicators and Classification through Machine Learning. *Int. J. Manag. Account*, 3(4), 83-90.
 - * Schwendner, P., Papenbrock, J., Jaeger, M., & Krügel, S. (2021). Adaptive Serial Risk Parity and Other Extensions for Heuristic Portfolio Construction Using Machine Learning and Graph Theory. *The Journal of Financial Data Science*, 3(4), 65-83.
 - * Sen, J., Dutta, A., & Mehtab, S. (2021). Stock Portfolio Optimization Using a Deep Learning LSTM Model. *arXiv preprint arXiv: 2111.04709*.
 - * Solin, M. M., Alamsyah, A., Rikumahu, B., & Saputra, M. A. A. (2019, July). Forecasting Portfolio Optimization using Artificial Neural Network and Genetic Algorithm. In *2019 7th International Conference on Information and Communication Technology (ICoICT)* (pp. 1-7). IEEE.
 - * Ta, V. D., Liu, C. M., & Addis, D. (2018, December). Prediction and portfolio optimization in quantitative trading using machine learning techniques. In *Proceedings of the Ninth International Symposium on Information and Communication Technology* (pp. 98-105).
 - * Thakkar, A., & Chaudhari, K. (2021). A comprehensive survey on portfolio optimization, stock price and trend prediction using particle swarm optimization. *Archives of Computational Methods in Engineering*, 28(4), 2133-2164.
 - * Wang, W., Li, W., Zhang, N., & Liu, K. (2020). Portfolio formation with preselection using deep learning from long-term financial data. *Expert Systems with Applications*, 143, 113042.
 - * Yu, J. R., Chiou, W. J. P., Lee, W. Y., & Lin, S. J. (2020). Portfolio models with return forecasting and transaction costs. *International Review of Economics & Finance*, 66, 118-130, 2020.
 - * Zhang, Y., Li, X., & Guo, S. (2018). Portfolio selection problems with Markowitz's mean-variance framework: a review of literature. *Fuzzy Optimization and Decision Making*, 17(2), 125-158.
 - * Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep learning for portfolio optimization. *The Journal of Financial Data Science*, 2(4), 8-20.



Abstract

Stock portfolio optimization based on the combined model of omega ratio and mean-variance Markowitz based on two-level ensemble machine learning

Sanaz Faridi ¹

Mahdi Madanchi Zaj ^{*2}

Amir Daneshvar ³

Shadi Shahverdi Ani ⁴

Fereydon Rahnama Roodposhti ⁵

Abstract

In this paper, the stock portfolio of active companies listed on the Tehran Stock Exchange is optimized based on the combined model of omega ratio and mean-variance Markowitz (MVOF). For this purpose, 480 companies listed on the Tehran Stock Exchange during the years 1390 to 1399 were selected and based on the input data, the companies were filtered. Hence a combined method consisting of trading rules optimization method based on technical analysis (6 indicators RSI, ROC, SMA, EMA, WMA and MACD) and two-level collective learning machine (SVM, RF, BN, MLP and KNN) for Data training and purchase signal presentation were addressed. Therefore, 85 companies were selected to optimize the stock portfolio. To teach the data, 85 companies filtered by the combined method were used and the number of different classes with 50 learners was used. The results show that using the OF model compared to the MVF model has the highest stock portfolio returns during the years 1395 to 1399. While the MVF model has the lowest investment risk. As a result, by combining the above models, the stock portfolio return in this method is much higher than other methods. While the investment risk was lower. Therefore, if the MVOF model is used, the return on the stock portfolio will increase and the investment risk in it will decrease.

Keywords: Omega Ratio Model, Markowitz Mean-Variance Model, Two-Level Collective Learning Machine, Ultra-Innovative Algorithm.

¹ Department of Financial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran. sanaz.faridi66@gmail.com

² Department of Financial Management, Electronic Branch, Islamic Azad University, Tehran, Iran (Corresponding Author): madanchi@iauec.ac.ir

³ Department of Information Technology Management, Electronic Branch, Islamic Azad University, Tehran, Iran. daneshvar.amir@gmail.com

⁴ Department of Financial Management, Human Sciences Faculty, Islamic Azad university, Shahr-e-Qods branch, Tehran, Iran. shshahverdiani@gmail.com



⁵ Department of Accounting, Islamic Azad University, Science and Research Branch, Tehran, Iran.
rahnama.roodposhti@gmail.com

