فصلنامه مهندسی مالی و مدیریت اوراق بهادار
شماره سی و هشتم/ بهار ۱۳۹۸

بررسی دقت ماشین‌های یادگیر در پیش‌بینی بازده حاصل از تغییر قیمت سهام با استفاده از مدل رافست، نزدیک‌ترین همسایه و درخت تصمیم‌گیری.

محمدرضا کریمی پویا[۱]
مهرداد قنبری[۲]
بابک جمشیدی نوید[۳]
منصور اسماعیل پور[۴]

تاریخ دریافت مقاله :۹۷/۰۲/۱۵ تاریخ پذیرش مقاله :۹۷/۰۳/۲۴

چکیده

پیش‌بینی یکی از مولفه‌های مهم و ضروری در برنامه‌ریزی‌های کوتاه‌مدت و میان‌مدت در هر کسب وکاری است. یک پیش‌بینی دقیق می‌تواند در کسب بازده، مدیریت جریان‌های نقدی و تخصیص منابع نقش موثری داشته باشد و به سرمایه‌گذار این امکان را می‌دهد که در یک بازه زمانی مشخص حدود درآمد کسب‌وکار و بازده خود را تخمین بزند. محققان در این اندیشه‌اند که روش‌های قدیمی، هزینه‌بر و زمان‌بر را کنار گذاشته و روش‌هایی جدید همچون استفاده از ماشین‌های یادگیر را پیاده‌سازی نمایند. این پژوهش از نظر نوع پژوهش، تحلیلی-تجربی، از نظر طرح پژوهش، پس رویدادادی، از نظر هدف، کاربردی، از نظر منطق اجرا، قیاسی و از نظر زمان انجام، طولی و از نوع آینده‌نگر می‌باشد. در این پژوهش، از الگوی الگوریتم نزدیک‌ترین همسایه، روش رافست و درخت تصمیم‌گیری برای بهبود قدرت پیش‌بینی، کاهش هزینه و زمان پیش‌بینی بازده سهام استفاده شده است. بدین منظور نمونه‌ای متشکل از ۱۱۳شرکت موجود در بورس اوراق بهادار تهران در طی یک دوره ۱۰ ساله (۱۳۹۴-۱۳۸۵) از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران انتخاب شده است نتایج پژوهش نشان داد که تمام فرضیه‌های این پژوهش مبنی بر وجود تفاوت در دقت تخمین این مدل‌ها در پیش‌بینی سه متغیر وابسته می‌باشند.

کلمات کلیدی

تخمین(پیش‌بینی)، درخت تصمیم، رافست، نزدیک‌ترین همسایه، بازده‌آتی سهام

۱ دانشجوی دکترای حسابداری، دانشکده علوم انسانی، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران.(نویسنده مسئول)
mkp1386@google.com
۲ استادیار گروه حسابداری، دانشکده علوم انسانی، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران.
mehrdadghanbary@yahoo.com
۳ استادیار گروه حسابداری، دانشکده علوم انسانی، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران.
jamshidinavid@yahoo.com
۴ استادیار گروه کامپیوتر، دانشکده فنی مهندسی، واحد همدان، دانشگاه آزاد اسلامی، همدان، ایران. esmaeilpour@iauh.ac.ir

مقدمه

بازده یک سرمایه‌گذاری شامل مواردی مانند ۱-افزایش قیمت سهام درآخر سال مالی نسبت به ابتدای سال مالی مورد محاسبه۲- سود نقدی هر سهم۳- مزایای ناشی ازحق تقدم خرید سهام ۴- مزایای ناشی از سود سهمی یا سهام جایزه می‌باشد و بطور خلاصه بیانگر منافع حاصل از سرمایه‌گذاری است و سرمایه‌گذاران در پی فرصت‌های سرمایه‌گذاری هستند که بازده سرمایه آنها را حداکثر کند. برای رسیدن به این هدف سرمایه‌گذاران باید عوامل زیادی را مد نظر قرار دهند، زیرا نقدترین دارائی خود را به اوراق بهادار تبدیل می‌کنند. اگر سرمایه‌گذاران بدون توجه به یک سری عوامل، اقدام به سرمایه‌گذاری نمایند، نتایج مطلوبی از سرمایه‌گذاری عاید آنها نخواهد شد.

امروزه، سرمایه گذاری در بورس، بخش مهمی از اقتصاد کشور را تشکیل می دهد. به همین دلیل پیش‌بینی بازده سهام برای سهامداران از اهمیت خاصی برخوردار شده است تا بتوانند بالاترین بازده را از سرمایه گذاری خود کسب کنند. از سوی دیگر، شاخص قیمت سهام نشان دهنده وضعیت کلی بازار سهام است و می‌تواند به پیش‌بینی سهامداران جهت سرمایه گذاری کمک کند(رضایی و پیری، ۱۳۹۴).

شوک‌های ناگهانی بازار و سقوط قیمت‌ها تعداد زیادی از سرمایه‌گذاران را از بازار خارج می‌کند. تخصیص صحیح منابع، موجب افزایش اطمینان سرمایه‌گذاران و کارائی بازار خواهد شد. افزایش ابزارهای مرتبط با شاخص‌مالی، دامنه فرصت‌های سرمایه-گذاری جهانی را برای سرمایه‌گذاران گسترش داده است. دو دلیل اصلی برای پیشرفت این ابزارها وجود دارد: اول اینکه، آنها ابزارهای موثری را برای سرمایه‌گذاران به منظور محافظت از ریسک‌های بالقوه بازار فراهم می‌آورند و دوم اینکه، فرصت‌های کسب سود جدیدی را برای کسانی که از موقعیت‌های زمانی و مکانی بازار استفاده می‌برند، ایجاد می‌کنند. بنابراین ارائه مدل مناسب برای پیش‌بینی بازده سهام از اهمیت و کاربرد بسیار بالائی برخوردار است (چن[1] و همکاران، ۲۰۱۰).

رویکرد عمده برای پیش‌بینی بازده سهام، روش‌های آماری مبتنی بر سری‌های زمانی مانند: میانگین متحرک، هموارسازی، آریما، رگرسیون و یا ترکیبی از آنها می‌باشد. این روش‌ها

بنا به ماهیت خود، در محیط‌های با تغییرات کم با تقریب خوبی قادر به پیش‌بینی هستند اما در مواردی چون پیش‌بینی بازار سهام که شرایط محیطی همواره در حال تغییر است نمی‌توانند تقریب خوبی از تغییرات محیطی را تخمین بزنند. به این ترتیب نیاز به استفاده از ابزارها و مدل‌های نوین جهت پیش‌بینی این دوره‌ها ضروری می‌گردد(حقیقت منفرد و همکاران، ۱۳۹۱).

بیان مسأله پژوهش

بدون شک بورس یکی از مناسب‌ترین جایگاه‌ها جهت جذب سرمایه‌های کوچک و استفاده از آن‌ها جهت رشد یک شرکت، در سطح کلان و نیز رشد شخصی فرد سرمایه‌گذار است. از آن‌جایی که هدف و تعریف سرمایه‌گذاری، به تعویق انداختن مصرف جهت مصرف بیشتر و بهتر در آینده است، افراد با سرمایه‌گذاری انتظار دستیابی به سود مورد انتظار خود را دارند. دوران گشایش بازارهای اوراق بهادار همواره این فکر وجود داشته است که به کمک روشی، عوامل موثر اقتصادی، حسابداری و مالی را پیش‌بینی کنند و دراین راه سخت افزارها و نرم افزارها، تحلیل‌های متفاوت ابداع شده و مورد استفاده قرارگرفت. متخصصان بازار سرمایه برای سالیان متمادی بازار را مطالعه نموده‌اند و الگوهایی را فرا گرفته‌اند و پیش‌بینی‌ها را براساس آن انجام می‌دهند (طلوعی اشلقی و حق دوست،۱۳۸۹).

آنچه در پژوهش‌های فوق که به آن پرداخته نشده استفاده از سه ماشین رافست، نزدیک‌ترین همسایه و درخت تصمیم‌گیری و مقایسه توانایی‌های آن‌ها در پیش‌بینی بازده آتی سهام در بازار آشوب‌ناک سرمایه ایران است.

در سال‌های اخیر یادگیری ماشین در مسائلی مانند آنالیز داده و داده‌کاوی کاربرد گسترده‌ای پیدا نموده و از این‌رو پژوهش‌های روی یادگیری نیمه نظارتی، هم در زمینه تئوری و هم در کاربرد عملی توسعه زیادی یافته است. امروزه داده‌های بدون برچسب به‌راحتی جمع‌آوری می‌شوند اما راهی برای استفاده از آن‌ها نیست زیرا یادگیری با نظارت از داده‌های آموزشی که برچسب کلاس آن‌ها مشخص است برای تعیین برچسب کلاس برای داده‌های تست استفاده می‌کند. (افتخاری وعارفیان،۱۳۹۲).

مبانی نظری و پیشینه پژوهش

بورس اوراق بهادار: بورس اوراق بهادار یک بازار رسمی و همیشگی است که در یک محل معین تشکیل می‌شود و در آن، اوراق بهادار مورد معامله قرار می‌گیرد. مشهورترین اوراق بهادار «پول کشورها» است که که در بورس معامله پول، خرید و فروش می‌شوند. افزون بر پول کشورهای مختلف، سهام شرکت‌ها، اوراق‌قرضه و اسناد خزانه مهم‌ترین اوراق بهادار هستند که در این‌گونه بورس‌ها مورد معامله قرار می‌گیرند. تا اواخر سده نوزدهم لیون، آمستردام، پاریس و لندن به ترتیب عمده‌ترین بازارهای مالی بودند و نقش آنها در مرکز نظام سرمایه‌داری، فراهم آوردن محلی برای دیدار صاحبان اندوخته، افراد و بانک‌ها، و سرمایه‌گذاران، دولت‌ها و سوداگران بود (حسینی، ۱۳۸۹).

بازده سهام: بازده یک سرمایه‌گذاری بیانگر منافع حاصل از سرمایه‌گذاری است و سرمایه‌گذاران در پی فرصت‌های سرمایه‌گذاری هستند که بازده سرمایه آنها را حداکثر کند. برای رسیدن به این هدف سرمایه‌گذاران باید عوامل زیادی را مد نظر قرار دهند، زیرا نقدترین دارائی خود را به اوراق بهادار تبدیل می‌کنند. اگر سرمایه‌گذاران بدون توجه به یک سری عوامل، اقدام به سرمایه‌گذاری نمایند، نتایج مطلوبی از سرمایه‌گذاری عاید آنها نخواهد شد. اصلی ترین عاملی که هر سرمایه‌گذار در تصمیم گیری‌های خود مورد توجه خاص قرار می‌دهد بازده است یعنی سرمایه‌گذاران به دنبال پر بازده‌ترین فرصت‌ها برای سرمایه‌گذاری منابع مازاد خود در بازارهای سرمایه هستند.

پیش‌بینی: پیش‌بینی عنصر کلیدی در تصمیم‌گیری است. زیرا بخش نهایی هر تصمیم بستگی به پیامدهای مربوط به رویدادهای بعد از تصمیم‌گیری دارد. توانایی پیش‌بینی جنبه‌های غیرقابل کنترل این رویدادها قبل از اتخاذ تصمیم می‌تواند به انتخاب بهتری منجر شود. به‌همین دلیل سیستم‌های مدیریتی برای برنامه‌ریزی و اداره عملیات، دارای مدل‌های پیش‌بینی در ابعاد مختلف است. هدف پیش‌بینی کاهش ریسک در تصمیم‌گیری است(چاردولی، ۱۳۹۴). با وجود آن که پیش‌بینی‌ها معمولا" دقیق نیست، اما میزان خطای پیش‌بینی به روش مورد

استفاده، بستگی دارد. با تخصیص بیشتر منابع برای پیش‌بینی می‌توان دقت آن را افزایش داد ودرنتیجه، برخی از زیان‌های‌ناشی از عدم‌اطمینان را در فرآیند تصمیم‌گیری حذف‌یاکاهش‌داد.

الگوریتم نزدیک‌ترین همسایه: در سال ۱۹۵۶ بنیاد راکفلر[۲] برگزاری کنفرانسی را بر عهده داشت که چشم اندازش به قرار زیر بود: امکان استفاده از کامپیوترها و شبیه‌سازی در هر زمینه از یادگیری و سایر حوزه‌های هوش مصنوعی. در همین کنفرانس بود که اصطلاح هوش مصنوعی مورد استفاده عمومی قرار گرفت. به طور کلی هوش مصنوعی را به این صورت می‌توان تعریف کرد: فرآیندهای کامپیوتری که سعی دارند فرآیندهای تفکر انسان را تقلید نمایند که این فرآیندها با فعالیت‌های که نیاز به استفاده از هوش دارند در ارتباط هستند. الگوریتم‌های یادگیری مانند نزدیک‌ترین همسایه و رافست و فناوری‌های مرتبط به آنها به عنوان بخش‌هایی از هوش مصنوعی می‌باشند. اخیرا" افزایش فعالیت‌های پژوهشی در زمینه هوش مصنوعی که رشد روزافزونی داشته است بیانگر این مطلب است که مدل‌های هوش مصنوعی مانند نزدیک‌ترین همسایه و رافست از نظر قابلیت پیش‌گویی و طبقه بندی الگو قوی می‌باشند. نزدیک‌ترین همسایه نیز به طور موفقیت آمیزی در زمینه های مختلف بازرگانی، صنعتی و علوم دیگر به کار رفته است(ابریشمی، ۱۳۸۷).

درخت تصمیم‌گیری[۳]: درخت تصمیم یکی از مشهورترین و قدیمی‌ترین روش‌های ساخت مدل رده‌بندی است. در الگوریتم‌های رده‌بندی مبتنی بر درخت تصمیم، دانش خروجی به صورت یک درخت از حالات مختلف مقادیر و ویژگی‌ها ارائه می‌شود. نمایش دانش به شکل درخت سبب شده است که رده‌های مبتنی بر درخت تصمیم کاملا" قابل تفسیر باشند(حقیقت و همکاران، ۱۳۹۱).

در این بخش به مجموعه‌ای از مطالعاتی که در داخل و خارج از کشور در این حوزه انجام گرفته است، اشاره می‌شود.

پیشینه داخلی:

در پژوهش‌های داخلی در زمینه پیش‌بینی قیمت سهام، بیشتر از مدل‌های عصبی، عصبی– فازی و الگوریتم ژنتیک در پژوهش‌هایی مانند زیر استفاده گردیده است و پژوهشی

که در آن از مدل‌های رافست، درخت تصمیم‌گیری ونزدیک‌ترین همسایه استفاده گردیده باشد مشاهده نگردید.

پرچمی وهمکاران(۱۳۹۵) در پژوهشی تحت عنوان پیش‌بینی بازده سهام با متغیرهای مالی با رویکرد شبکه های عصبی به بررسی تأثیر متغیرهای مالی بر بازده سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران برای دوره زمانی ۱۳۸۲تا ۱۳۹۲پرداختند. نتایج پژوهش نشان داد که ۸۹ درصد پیش‌بینی بازده سهام توسط روش رگرسیون و ۹۶ درصد توسط شبکه‌های عصبی مصنوعی می‌باشد و با توجه به نتایج، روش شبکه‌های عصبی مصنوعی بهتر از روش رگرسیون حداقل مربعات می‌تواند بازده سهام را پیش‌بینی کند.

مرادی و همکاران(۱۳۹۵) در پژوهشی تحت عنوان پیش‌بینی سری‌های زمانی کوتاه مدت، میان‌مدت و بلندمدت بازده سهام بورس اوراق بهادار تهران با استفاده از الگوریتم L-CO-R به بررسی الگوریتم هم‌تکاملی L-CO-R، توصیف و عملکرد آن در پیش‌بینی سری‌های زمانی در دوره‌های کوتاه مدت، میان‌مدت‌و بلند مدت پرداختند. در این پژوهش سری‌های زمانی روزانه، ماهانه و سالانه بازده سهام بورس اوراق بهادار تهران از سال ۱۳۸۶ تا سال۱۳۹۲ مورد بررسی قرار گرفت. نتایج تحلیل‌های آماری نشان داد که الگوریتم هم‌تکاملی L-CO-R می‌تواند با استفاده از مقادیر واقعی مشخص شده، افق‌های زمانی کوتاه مدت، میان‌مدت‌و بلندمدت را پیش‌بینی کند. همچنین نتایج پژوهش نشان داد که الگوریتم هم‌تکاملیL-CO-R برای دوره‌های بلندمدت بهتر از روش‌های دیگر عمل می‌کند در حالی‌که روش باکس و جنکینس برای دوره‌های کوتاه-مدت عملکرد بهتری دارند.

رستمی(۱۳۹۵)در پژوهشی به بررسی کارآیی الگوریتم‌های یادگیری ماشین در پیش‌بینی قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخت. نتایج حاصل نشان از آن داشت که مدل‌های رافست و رگرسیون لوجستیک که زیر مجموعه‌های الگوریتم های یادگیری ماشین هستند، توانایی پیش‌بینی شاخص‌سهام را دارند.

صالحی و فرخی پیله رود(۱۳۹۵) در پژوهشی تحت عنوان پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت‌تصمیم در صنایع کشاورزی و نساجی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران به بررسی دقت پیش‌بینی مدیریت سود با استفاده از

شبکه‌های عصبی و درخت تصمیم‌گیری و مقایسه آن با مدل‌های خطی پرداختند. برای این منظور نه متغیر تأثیرگذار بر مدیریت سود به عنوان متغیرهای مستقل و اقلام تعهدی اختیاری، به عنوان متغیر وابسته مورد استفاده قرار دادند. نتایج به‌دست آمده نشان داد که روش شبکه‌عصبی و درخت‌تصمیم‌گیری در پیش‌بینی مدیریت سود نسبت به روش‌های خطی دقیق‌تر و دارای سطح خطای کمتری است. در رابطه با ارتباط بین متغیرهای وابسته با متغیر مستقل نیز به این نتیجه رسیدند که مدیریت سود با متغیرهای اقلام تعهدی اختیاری دوره قبل اقلام تعهدی غیراختیاری دوره قبل یا آستانه عملکرد و حساسیت در پرداخت بابت عملکرد در روش‌های رگرسیون، شبکه عصبی، درخت کارت دارای بیشترین ارتباط است.

سینایی، مرتضوی و تیموری اصل‌دار(۱۳۹۴) به پیش‌بینی شاخص بورس اوراق بهادار تهران به وسیله شبکه عصبی مصنوعی و ارائه شواهدی مبنی بر رفتار آشوب‌گونه شاخص قیمت سهام در بورس اوراق بهادار پرداختند. آنها دو مجموعه از داده‌ها را برای ورودی شبکه عصبی انتخاب نمودند، وقفه‌های مختلفی از شاخص و عوامل کلان اقتصادی به‌عنوان متغیر مستقل انتخاب کردند. در این پژوهش از مدل خطی میانگین متحرک یکپارچه برای پیش‌بینی شاخص بورس در هفته‌های بعدی استفاده شده است. نتایج حاصل از پژوهش نشان می‌دهد شبکه عصبی عملکرد بهتری نسبت به مدل خطی میانگین متحرک یکپارچه برای پیش‌بینی شاخص قیمت دارد.

پیشینه خارجی:

چامبرز[۴] (۲۰۱۶) به توسعه یک روش عمومی برای پیش‌بینی حرکات روزانه قیمت سهام با به‌کارگیری و یکپارچه‌سازی سه مدل پیش‌بینی تحلیلی: سیستم استنتاج عصبی- فازی تطبیقی، شبکه‌های عصبی مصنوعی و پشتیبانی ماشین آلات بردار پرداختند. نتایج پژوهش در بورس استانبول برای یک دوره ۸ ساله نشان می‌دهد که ماشین بردار پشتیبانی بهتر از مدل‌های دیگر پیش‌بینی را انجام می‌دهد.

ییِنگ وی[۵] (۲۰۱۶) در مقاله‌ای تحت عنوان استفاده از مدل ترکیبی مبتنی بر تجزیه حالت تجربی [۶] برای پیش‌بینی سری زمانی سهام به رسیدگی به این مشکلات جهت بهبود

عملکرد پیش‌بینی مدل‌های سری‌زمانی پرداختند.آنها با استفاده از آن مدل و مدل تجربی ترکیبی[۷] به پیش‌بینی قیمت سهام در بورس تایوان پرداختند. آنها برای مقایسه نتیجه اندازه‌گیری عملکرد پیش‌بینی خود، مدل ارائه شده را با مدل چن، مدل یو، مدل اتورگرسیو[۸]، مدل ترکیبی مبتنی بر تجزیه حالت تجربی و مدل رگرسیون بردار پشتیبان[۹] مقایسه نمودند. نتایج نشان داد که مدل ترکیبی آنها که پیش‌بینی را بر اساس ریشه میانگین ارزش مربعات خطا انجام داد نسبت به مدل‌های دیگر دارای قدرت پیش‌بینی بالاتری بوده است.

پاولاک[۱۰] (۲۰۱۶) پژوهشی تحت عنوان پیش‌بینی تاثیر هزینه‌های بازار بر قیمت سهام با استفاده از مدل‌های ناپارامتریک ماشین‌های یادگیر انجام داد. در این پژوهش مدل‌های ناپارامتری ماشین‌های یادگیر مانند شبکه‌های عصبی، شبکه‌های بیزی عصبی، فرآیندگوسی و رگرسیون بردارپشتیبان، برای پیش‌بینی تاثیر هزینه بازار بر دقت مدل پیش‌بینی با استفاده از داده‌های واقعی بازار سهام ایالات متحده استفاده گردید. در نتیجه، مشخص گردید که علیرغم مشکلاتی که ماشین‌های ناپارامتریک در جداکردن هزینه‌های دائم و موقت با آن روبرو می‌باشند می‌توانند جایگزین خوبی در کاهش هزینه‌های معامله و بهبود عملکرد پیش‌بینی باشند.

زمیجسکی[۱۱](۲۰۱۶) پژوهشی تحت عنوان پیش‌بینی بازار سهام کراچی در روز بسته شدن بازار با استفاده از روش ماشین-های یادگیر انجام داد. او از متغیرهای ورودی با ویژگی‌های مختلف استفاده نمود. متغیرهای مورد استفاده در این مدل شامل نرخ نفت، طلا و نقره نرخ، نرخ بهره، ارز، نوع اخبار و خوراک رسانه‌های اجتماعی بود. نتایج نشان داد که عملکرد شاخص بازار بورس کراچی را می‌توان با روش‌های ماشین یادگیری پیش‌بینی کرد.

تسانگ[۱۲] (۲۰۱۳) پژوهشی تحت عنوان پیش‌بینی جهت حرکت شاخص قیمت سهام با استفاده از روش ماشین‌های یادگیر را در بورس هنگ کنگ انجام داد. در این پژوهش از سه روش ماشین‌های یادگیر، رگرسیون لجستیک[۱۳] ، ماشین بردار پشتیبان[۱۴] و ادراک چند لایه[۱۵] استفاده گردید. همچنین در این پژوهش ازاطلاعات روزانه و فصلی برای یک دوره ۵ ساله از نوامبر ۲۰۰۷ تا نوامبر ۲۰۱۲ استفاده شد. نتایج نشان داد که جهت تغییرات بورس هنگ کنگ با استفاده از داده های آزمایشی دارای دقت بیشتری نسبت به داده های اصلی بوده و طبقه بندی بر اساس رگرسیون لجستیک باعث بهبود عملکرد پیش‌بینی می‌گردد.

روش شناسی تحقیق

این پژوهش، جزء پژوهش‌های تحلیلی- تجربی طبقه‌بندی و روش آزمون فرضیه‌های پژوهش، روش تجزیه‌وتحلیل آماری از نوع تحلیل همبستگی است. طرح پژوهش مورد استفاده از نظر نوع داده کمی و از نظر نوع زمان گذشته نگر و پس‌رویدادی می‌باشد چرا که هر یک از فرضیه‌های پژوهش با استفاده از اطلاعات واقعی که بر مبنای عملکرد واقعی بورس اوراق بهادار تهران در طول دوره زمانی پژوهش(سال‌های ۱۳۹۴-۱۳۸۵) حاصل شده است، آزمون می‌گردند. پژوهش به لحاظ هدف، پژوهشی کاربردی است. از نظر منطق اجرا، نوع پژوهش استقرایی محسوب می‌گردد.

فرضیه‌های پژوهش

در پژوهش حاضر فرضیه‌ها به شرح زیر می‌باشند:

فرضیه اصلی: بین دقت ماشین‌های یادگیر در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تفاوت وجود دارد.

فرضیه فرعی اول: بین دقت مدل‌رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.

فرضیه فرعی دوم: بین دقت مدل‌رافست و درخت‌تصمیم‌گیری در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.

فرضیه فرعی سوم: بین دقت درخت تصمیم‌گیری و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.

جامعه آماری، روش نمونه‌گیری و حجم نمونه

جامعه آماری این پژوهش را کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در برمی‌گیرد.

شرکت‌هایی که قبل از سال ۱۳۸۵ در بورس اوراق بهادار تهران پذیرفته شده باشند و تا

پایان سال ۱۳۹۴ از تابلو بورس خارج نشده باشند. سال مالی شرکت منتهی به پایان اسفندماه باشد و سال مالی طی دورهٔ پژوهش تغییر نکرده باشد. شرکت موردنظر طی دورهٔ پژوهش، فعالیت مستمر داشته و دچار وقفه معاملاتی بیشتر از ۶ ماه نشده باشد.

در نتیجه تعداد ۱۱۳ شرکت مورد بررسی قرار گرفت.

تجزیه‌وتحلیل نزدیک‌ترین همسایه

برداشت ما آن است که سری‌های زمانی را می‌توان به‌صورت ترکیبی از مؤلفه‌های مجزا در مقیاس و سطوح نوسان متفاوت دانست چرا که این امر برای سری‌های زمانی مالی چندان عجیب نیست. زمانی که سری‌های زمانی مشاهده شده ترکیبی از چنین فرآیندهای پیچیده‌ای باشد آنگاه فرد پیش‌بینی کننده‌ای که قادر نیست مؤلفه‌های وابسته به مقیاس مجزای سری زمانی را تشخیص دهد، قادر نخواهد بود تا مدلی مناسب برای پیش‌بینی صحیح بسازد. از طرف دیگر اگر بتوان سری‌های زمانی اولیه را به مؤلفه‌های وابسته به مقیاس و یا فرکانس خاص تجزیه کرده و هرکدام از این مؤلفه‌ها را جداگانه مدل بندی کرد، آنگاه می‌توان مدل‌های با دقت بیشتری ساخت. درواقع می‌توان گفت که فعالیت‌های اقتصادی دارای دوره زمانی متفاوت دارای مقیاسی متفاوتی است. این امر بدان معنی است که حرکت دوره تجاری چیزی جز مجموع این مؤلفه‌های مجزا که در افق‌های زمانی متفاوت نمایان می‌شوند، نیست. ازاین‌رو با توجه به این موضوع از طریق نرم‌افزارMATLAB دست به ساختن یک مدل نزدیک‌ترین همسایه پیش‌بینی کننده جهت پیش‌بینی متغییرهای وابسته در بورس اوراق بهادار خواهیم زد.

روش درخت تصمیم‌گیری

در اجرای درخت تصمیم‌گیری روش یادگیری درخت تصمیم رگرسیون به کار می‌رود و تنظیمات داده آموزش و آزمایش همانند روش نزدیک‌ترین همسایه است.

رافست[۱۶]

مجموعه آزمون و آموزش: در اجرای این مدل از هوش مصنوعی نیازی به نرمال سازی داده‌ها نبوده و در ابتدا داده‌ها را کلاسه بندی کرده که تبدیل به داده‌های کیفی شوند. زیرا

مدل رافست برای داده‌های کیفی است و نتیجه و تجزیه‌وتحلیل پایانی را می‌دهد. در این مدل داده‌ها به دو قسمت ۸۰ درصد آموزش و ۲۰ درصد تست تبدیل خواهند گردید. در قسمت آموزش داده‌ها دیسکریت شده یا به زبان ساده‌تر بازه بندی می‌شوند که در قسمت آموزش استفاده می‌شود. در مدل رافست از سه روش ژنتیک الگوریتم[۱۷]، جانسون الگوریتم[۱۸] و هولتس[۱۹]استفاده می‌شود. در روش ژنتیک الگوریتم تمام متغیرها و عوامل موثر بر قیمت سهام در استخراج رول‌ها (قوانین) دخالت داده می‌شوند. در روش جانسون الگوریتم متغیرها را بررسی و متغیرهایی که بیشترین تاثیر را در تغییرات قیمت سهام دارند دخالت می‌دهد و قوانین را استخراج می‌کند و در روش هولتس با دقت بسیار بالا به بررسی تغییرات قیمت سهام پرداخته و تک تک عوامل را مورد بررسی قرار داده و تعداد تکرارهای موجود در قیمت سهام بررسی می‌کند و قوانین را با میزان اطمینان و خطا نشان می‌دهد. مدل رافست توسط نرم افزار رافست اجرا می‌گردد. تمامی نتایج به دست آمده از رافست بر مبنای در نظر گرفتن میزان تورم و تمامی عوامل موثر بر صنایع بوده است. اما با استفاده از مقاله‌های مختلف تورم در میان متغیر آورده نشده زیرا امکان خروج نتایج اشتباه وجود داشته، زیرا تورم در میزان شاخص سهام تاثیر گذاشته و بعد شاخص بر قیمت سهام موثر بوده و سپس بر روی تابلوهای معاملاتی ظاهر می‌شود.

مقایسه روش‌ها

در این بخش قصد داریم مقایسه روش‌های پیاده سازی شده را از لحاظ RMSE ارائه دهیم. در جدول زیر روش‌ها با هم مقایسه گردیده‌اند.

جدول ۱ مقایسه خطای الگوریتم‌ها

نام روش	خطای بازده سالانه تغییر قیمت سهام
رافست (ژنتیک)	۱/۰۱۵۲
رافست (جانسون)	۲/۲۵۶۴
رافست (هولت)	۱/۱۱۲۵
نزدیک‌ترین همسایه	۱/۲۴۶۳
درخت تصمیم	۰/۴۸۳۷

جدول فوق نشان می دهد که در پیش‌بینی بازده حاصل از تغییر قیمت سهام ماشین درخت تصمیم با کمترین خطا، دارای بهترین پیش‌بینی می باشد.

آزمون میانگین زوجی برای فرضیه‌های پژوهش

آزمون t برای نمونه‌های جفت شده یا زوجی برای تجزیه‌وتحلیل آزمون‌هایی به‌کار می‌رود که هر فرد دو بار در دو وضعیت متفاوت مورد مشاهده قرار می‌گیرد. در اینجا نیز بازده سهام با استفاده از سه ماشین مختلف مورد پیش‌بینی واقع شده و فرضیه H0 در طرح داده‌های جفت شده این است که اختلافی بین مقادیر میانگین بازده حاصل از تغییر قیمت سهام در دو نمونه‌های جفت شده از جامعه وجود ندارد. بنابراین آزمون آماری آن به‌شرح زیر می‌باشد:

$$\begin{cases} H0: \mu1=\mu2=\mu3 \\ H1: \mu1\neq\mu2\neq\mu3 \end{cases}$$

در آزمون آماری فوق منظور از μ۱ میانگین بازده در ماشین رافست، μ۲ میانگین بازده در ماشین درخت تصمیم‌گیری و منظور از μ۳ میانگین بازده در مدل نزدیکترین همسایه می‌باشد. چنانچه سطح معناداری مدل کمتر از ۵٪ باشد در سطح اطمینان ۹۵٪ هیچ دلیلی برای رد اختلاف بین توانایی ماشین‌های یادگیر در پیش‌بینی بازده حاصل از سهام وجود ندارد.

برای انجام این آزمون باید مقادیر ویژگی عددی باشند، داده‌های هر دو متغیر دارای توزیع نرمال باشد و $n\geq30$ ، نمونه‌ها به صورت مستقل و تصادفی انتخاب شده باشند، تعداد نمونه‌ها برای هر دو متغیر برابر باشد و تفاوت در مقدار واریانس دو متغیر اشکالی ندارد(ری بد، ۱۳۹۳).

جدول ۲ نتایج آزمون میانگین زوجی برای برای آزمون فرضیات فرعی

نوع ماشین	آزمون t برای مقایسه ی میانگین ها							
	میانگین	انحراف استاندارد نمونه	انحراف استاندارد میانگین	فاصله اطمینان ۹۵ درصدی برای انحراف		آماره t	درجه آزادی	معناداری دو طرفه
				حد پایین	حد بالا			
رافست و نزدیکترین همسایه	./۰۵۲۰۲	./۶۶۵۳۱	./۱۴۸۷۷	-./۲۵۹۳۶	./۳۶۳۴۰	./۳۵۰	۲۲۴	./۷۳۰
رافست و درخت تصمیم	./۸۶۱۹	./۴۹۲۷۹	./۱۱۰۱۹	./۶۳۱۳۲	۱/۰۹۲۵۸	۷/۸۲۲	۲۲۴	۰/۰۰
درخت تصمیم و نزدیکترین همسایه	-./۸۰۹۹۳	./۵۳۰۴۹	./۱۱۸۶۲	-۱/۰۵۸۲۱	-./۵۶۱۶۵	-۶/۸۲۸	۲۲۴	./۰۰

همانطور که جدول فوق نشان می‌دهد:

فرضیه فرعی اول تایید نمی‌گردد. چرا که مقدار t کوچک بوده و مقدار سطح معناداری دوطرفه بیش از ۵٪ می‌باشد. در نتیجه فرضیه H0 رد نمی‌گردد و می‌توان گفت هیچ اختلاف معناداری بین دقت مدل رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران وجود ندارد و دو ماشین رافست و نزدیکترین همسایه، توانایی مشابهی در پیش‌بینی بازده حاصل از تغییر قیمت سهام در بورس اوراق بهادار تهران دارند.

فرضیه فرعی دوم تایید می‌گردد. چرا که مقدار t بزرگ بوده و مقدار سطح معناداری دوطرفه کمتر از ۵٪ می‌باشد. در نتیجه فرضیه H0 تایید نمی‌گردد و می‌توان گفت اختلاف معناداری بین دقت مدل رافست و درخت تصمیم گیری در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران وجود دارد و دو ماشین رافست و درخت تصمیم، توانایی متفاوتی را در پیش‌بینی بازده حاصل از تغییر قیمت سهام در بورس اوراق بهادار تهران دارند.

فرضیه فرعی سوم نیز تایید می‌گردد. چرا که مقدار t بزرگ بوده و مقدار سطح معناداری دوطرفه کمتر از ۵٪ می‌باشد. در نتیجه فرضیه H0 تایید نمی‌گردد و می‌توان گفت اختلاف معناداری بین دقت درخت تصمیم و مدل نزدیکترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران وجود دارد و دو ماشین نزدیکترین همسایه و درخت تصمیم، توانایی متفاوتی را در پیش‌بینی بازده حاصل از تغییر قیمت سهام در بورس اوراق بهادار تهران دارند.

نتیجه آزمون فرضیه اصلی

باتوجه به عدم پذیرش فرضیه فرعی اول که مشخص شد هیچ اختلاف معناداری بین دقت مدل رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران وجود ندارد و همچنین نتایج حاصل از فرضیه‌های فرعی دوم و سوم که از یک طرف مشخص شد بین توانایی ماشین‌های رافست و درخت تصمیم

گیری و و از طرف دیگر بین درخت تصمیم گیری و نزدیکترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام اختلاف وجود دارد، چنین می‌توان نتیجه گرفت که در فرضیه اصلی اول، H0 تایید نمی‌گردد و به‌عبارت دیگر نتایج آزمون نشان می دهد که بین دقت ماشین‌های یادگیر در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تفاوت وجود دارد.

جدول ۳ خلاصه نتیجه حاصل از آزمون فرضیه‌ها

فرضیه	شرح فرضیه	نتیجه فرضیات
اصلی	بین دقت ماشین‌های یادگیر در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تفاوت وجود دارد.	پذیرش فرضیه
فرعی اول	بین دقت مدل رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.	رد فرضیه
فرعی دوم	بین دقت مدل رافست و درخت تصمیم‌گیری در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.	پذیرش فرضیه
فرعی سوم	بین دقت درخت تصمیم‌گیری و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد.	پذیرش فرضیه

نتیجه آزمون فرضیه فرعی اول

در فرضیه فرعی اول مدل رافست در تخمین بازده حاصل از تغییر قیمت سهام کاراتر از مدل نزدیک‌ترین همسایه است که در فاصله زمانی ۱۳۸۵ تا ۱۳۹۴ موردسنجش قرار گرفت. با توجه به نتایج مشخص شد که مدل رافست در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از الگوریتم جانسون، ژنتیک و هولت به ترتیب خطای۲/۲۵۶۴ ، ۱/۰۱۵۲ و ۱/۱۱۲۵ تولید کرد که روش ژنتیک نسبت به الگوریتم نزدیک‌ترین همسایه که ۱/۲۴۶۳ بود دارای خطای کم‌تر است و بنابراین بنابراین فرضیه H0 تایید نمی‌شود. بعبارت دیگر بین دقت مدل رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد و روش رافست نسبت به روش نزدیک‌ترین همسایه دارای درصد

خطای کمتری است. بنابراین فرضیه H0 تأیید نمی‌شود. نتایج این پژوهش با نتیجه پژوهش رستمی(۱۳۹۵) و چاردولی(۱۳۹۴) که به بررسی کارآیی الگوریتم‌های یادگیری ماشین در پیش‌بینی قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخت همخوانی دارد. همچنین نتایج این پژوهش با نتیجه پژوهش فالمر (۲۰۱۲) که ازسیستم استنتاج عصبی- فازی سازگار مدل روش یکپارچه انتخاب ویژگی غیرخطی برای پیش‌بینی بازده سهام پرداخت و به منظور ارزیابی عملکرد مدل پیشنهادی، داده‌های معامله در بازار بورس سهام سال‌های ۲۰۰۶-۱۹۹۸ را به عنوان مجموعه داده‌های تجربی جمع‌آوری و با سایر مدل‌ها مقایسه نمود تطبیق دارد.

نتیجه آزمون فرضیه فرعی دوم

با توجه به نتایج مشخص شد که در فاصله زمانی ۱۳۸۵ تا ۱۳۹۴ دقت مدل رافست از درخت تصمیم در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران کم‌تر است. همچنین با توجه به نتایج تحقیق مشخص شد که مدل درخت تصمیم این متغیر را در ۱۰ سال با خطای ۰/۴۸۳۷ تخمین زد در حالیکه در الگوریتم ژنتیک درصد خطا ۱/۰۱۵۲ می‌باشد که بسیار بیشتر از میزان خطای پیش‌بینی در مدل درخت تصمیم می‌باشد. بعبارت دیگر بین دقت مدل رافست و درخت تصمیم‌گیری در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد و میزان دقت پیش‌بینی در روش درخت تصمیم بیشتر است. بنابراین فرضیه H0 تأیید نمی‌شود. نتایج این پژوهش با نتیجه پژوهش چاردولی(۱۳۹۴) که به بررسی دقت الگوی رافست و درخت تصمیم‌گیری در پیش‌بینی قیمت سهام، در بورس اوراق بهادار تهران پرداخت همخوانی دارد.

نتیجه آزمون فرضیه فرعی سوم

با توجه به نتایج مشخص شد که در فاصله زمانی ۱۳۸۵ تا ۱۳۹۴ دقت مدل نزدیکترین همسایه از درخت تصمیم در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران کم‌تر است. همچنین مشخص شد که میزان خطای پیش‌بینی

در بازده حاصل از تغییر قیمت سهام در الگوریتم نزدیک‌ترین همسایه ۱/۲۴۶۳ و در مدل درخت تصمیم ۰/۴۸۳۷ است و میزان دقت در الگوریتم نزدیک‌ترین همسایه کمتر است. بعبارت دیگر بین دقت درخت تصمیم‌گیری و روش نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران تفاوت وجود دارد و بنابراین فرضیه H0 تایید نمی‌شود. نتایج این پژوهش با نتیجه پژوهش کاظمی(۱۳۹۳) که به اندازه‌گیری ریسک اعتباری مشتریان حقیقی بانک با استفاده از روش شبکه عصبی و نزدیک‌ترین همسایه در مقابل روش پارامتریک آماری رگرسیون‌لجستیک پرداخت همخوانی دارد.

نتیجه‌گیری

جایگزین کردن تحلیل‌های علمی به‌جای یافته‌های ذهنی نباید مورد بی‌توجهی قرار گیرد. تصمیم‌گیری مناسب و منطبق با اصول علمی فرصت مناسبی است که یک سرمایه‌گذار را قادر می‌سازد تا بهترین بهره‌برداری را کرده و با شناخت راه‌کارهای علمی منابع خود را به نحو مناسب‌تری تخصیص دهد. بنابراین لازم است تا با روش‌های علمی و جدید تخمین‌هایی جهت کاهش ریسک سرمایه‌گذاری انجام شود. با توجه به تحلیل ارائه شده این نتیجه حاصل شد که فرضیه اصلی با درجه بالایی از اطمینان به اثبات رسید. بنابراین می‌توان بیان کرد که روش‌های الگوریتم یادگیری ماشین می‌تواند به عنوان یک روش قابل قبول و مطمئن جهت پیش‌بینی بازده سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران بکار گرفته شود و نتیجه پژوهش حاضر حاکی از قابل‌استفاده بودن روش‌های هوشمند در برابر روش‌های سنتی می‌باشد.

از سوی دیگر با استفاده از روش‌های نزدیک‌ترین همسایه، رافست و درخت تصمیم‌گیری می‌توان با تجزیه‌وتحلیل‌هایی از طریق تولید قوانین و شناسایی روابط پیچیده و غیرخطی بین متغیرهای مختلف، به بهبود تصمیم‌گیری خصوصا تصمیمات بازار سرمایه کمک کرد. در این پژوهش به میزان کارایی مدل‌های نزدیک‌ترین همسایه، رافست و درخت تصمیم در تخمین بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخته شد. نتیجه حاصل از تجزیه‌وتحلیل‌های فوق نشان می‌دهد که:

مدل رافست این توانایی را دارد که حجم کثیری از ورودی‌ها و خروجی‌ها را دریافت کرده و قوانین به‌دست آمده از مدل‌ها را با دقت بالایی نشان دهد. البته این مدل توانایی نشان دادن اعداد پیش‌بینی و نمودارها را ندارد و مدل مبتنی بر قانون است که فقط قوانین را استخراج می‌نماید.

بر اساس نتایج پژوهش می‌توان بازده حاصل از تغییر قیمت سهام را با استفاده از این روش‌ها پیش‌بینی کرد. بنابراین یک سرمایه‌گذار با استفاده از ماشین های یادگیر مورد بحث می‌تواند بازده حاصل از تغییر قیمت سهام را با کمترین درصد خطا برآورد نماید.

بر اساس نتایج این پژوهش این نتیجه حاصل شد که در پیش‌بینی بازده سالانه حاصل از تغییر قیمت سهام، روش درخت تصمیم کم‌ترین خطا را دارد و دقیق‌ترین تخمین را ارائه می‌دهد. پس از درخت تصمیم روش رافست با استفاده از الگوریتم ژنتیک، رافست با استفاده از الگوریتم هولت و مدل نزدیک‌ترین همسایه به ترتیب بیشترین اولویت را ازلحاظ دقت قرار دارند.

در سطح اطمینان ۹۵٪ هیچ اختلاف معناداری بین دقت مدل رافست و نزدیک‌ترین همسایه در پیش‌بینی بازده حاصل از تغییر قیمت سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران وجود ندارد و دو ماشین رافست و نزدیکترین همسایه، توانایی مشابهی در پیش‌بینی بازده حاصل از تغییر قیمت سهام در بورس اوراق بهادار تهران دارند.

با توجه به نتایج به‌دست‌آمده فرضیه اصلی این پژوهش مبنی بر وجود تفاوت بین توانایی ماشین های یادگیر در پیش‌بینی بازده حاصل از تغییر در قیمت سهام توسط متغیرهای سالانه ارزش‌افزوده اقتصادی، نسبت بدهی به دارایی‌ها، نرخ بازده دارایی‌ها، نرخ بازده سرمایه‌گذاری‌ها، سود هر سهم، سود باقیمانده، جریان نقدینه عملیاتی، ارزش‌افزوده نقدی، نسبت قیمت به درآمد و متغیرهای کنترلی نرخ تورم، نرخ ارز، تغییرات قیمت طلا و نفت، مورد تایید قرار می‌گیرد.

پیشنهادهای کاربردی

با توجه به فرضیه‌های پژوهش که شامل یک فرضیه اصلی و سه فرضیه فرعی می باشد، پیشنهادهای کاربردی زیر را می‌توان ارائه کرد:

با توجه به نتایج حاصل از فرضیه اصلی پیشنهاد می‌گردد تا سرمایه گذاران، تحلیل‌گران بورس اوراق بهادار، مدیران شرکت‌های سرمایه‌گذاری و شرکت‌های موجود در بورس اوراق بهادار تهران بتوانند با استفاده از ماشین‌های یادگیر عواملی را که در شکل‌گیری بازده حاصل از تغییر قیمت سهام نقش دارند، شناسایی و با اطمینان بیشتری نسبت به برنامه‌ریزی و تصمیمات سرمایه‌گذاری آینده خود اقدام نمایند. همچنین پیشنهاد می‌گردد تا هنگام استفاده از ماشینهای یادگیر در پیش‌بینی بازده حاصل از تغییر قیمت سهام، در انتخاب متغیرهای موثر بر تغییر قیمت سهام از تمام متغیرهای موثر مانند تورم،تغییرات نرخ ارز، واردات، تحریم ها، تغییرات سیاستهای دولت و...استفاده نمایند.

پیشنهادهایی در ارتباط با پژوهش‌های آتی

با توجه به نتایج حاصل از این پژوهش، پیشنهادهایی به شرح زیر ارائه می‌شود:

- تخمین متغیرهای این پژوهش برای دوره‌های مالی طولانی‌تر و مقایسه توانایی‌های ماشین‌های یادگیر در پیش‌بینی بازده سهام شرکت‌های صنایع مختلف.

- بررسی تاثیر متغیرهای غیر اقتصادی بر متغیرهای اقتصادی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران

- استفاده از روش‌های غیرخطی تخمین مانند شبکه‌های عصبی جدید، مدل‌های پارامتریک و روش‌های تکاملی و یادگیری عمیق در پیش‌بینی متغیرهای اقتصادی و مالی در سطح شرکت‌های پذیرفته شده در بورس اوراق بهادر و در سطح متغیرهای کلان اقتصاد کشور.

فهرست منابع

۱) ابریشمی، حمید(۱۳۸۷). مبانی اقتصاد سنجی. تهران: نورعلم.

۲) افتخاری، مهدی و عارفیان، فاطمه(۱۳۹۲). روش جدید K نزدیکترین همسایه فازی و ناهموار برای طبقه‌بندی نیمه‌نظارتی، همایش ملی مهندسی کامپیوتر و توسعه پایدار با محوریت شبکه‌های کامپیوتری، ۲۵-۱۵.

۳) پرچمی, بیتا؛ حسین نعمت زاده و رضا شاهوردی(۱۳۹۵). پیش‌بینی بازده سهام با متغیرهای مالی با رویکرد شبکه‌های عصبی. کنفرانس بین المللی مهندسی کامپیوتر و فناوری اطلاعات، تهران، دبیرخانه دایمی کنفرانس،

۴) چاردولی، مریم.(۱۳۹۴). بررسی کارایی روش رافست و نزدیک ترین همسایه در تخمین قیمت سهام پذیرفته شده در بورس اوراق بهادار تهران. پایان نامه کارشناسی ارشد، دانشکده علوم انسانی. دانشگاه آزاد اسلامی واحد همدان.

۵) حسینی، حمید(۱۳۸۹). تاریخچه بورس در ایران. تهران: انتشارات بورس.

۶) حقیقت منفرد، جلال؛ علی نژاد، محمود و متقالچی، سارا(۱۳۹۱). مقایسه مدل شبکه‌های عصبی با مدل باکس جنکینز در پیش‌بینی شاخص کل قیمت سهام بورس اوراق بهادار تهران. مجله مهندسی مالی و مدیریت اوراق بهادار، شماره ۱۱، ۱۶- ۱.

۷) رستمی، ستیلا (۱۳۹۵). بررسی کارایی الگوریتم یادگیری ماشین در پیش‌بینی شاخص بورس و اوراق بهادار تهران. پایان نامه کارشناسی ارشد. گروه حسابداری. دانشکده علوم انسانی. دانشگاه آزاد اسلامی واحد همدان.

۸) رضایی، فرزین و پیری، عبدالله(۱۳۹۴). تاثیر ساختار سرمایه و نقدینگی بر ارزش بازار شرکت‌ها. پژوهش‌نامه حسابداری مالی و حسابرسی، شماره۱۱، ۱۷۴-۱۵۵.

۹) سینایی، حسنعلی، مرتضوی، سعید و تیموری، یاسر(۱۳۹۴). پیش‌بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه‌های عصبی مصنوعی. بررسی های حسابداری و حسابرسی، شماره۴۱، ۸۳-۵۹.

۱۰) صالحی، مهدی و فرخی پیله رود، لاله(۱۳۹۵). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در صنایع کشاورزی و نساجی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. مجله اقتصاد کشاورزی دانشگاه فردوسی مشهد، دوره۸(۴)، ۱۰۵-۸۹ .

۱۱) طلوعی اشلقی،عباس و حق دوست، شادی(۱۳۸۹). مدل‌سازی پیش‌بینی قیمت سهام با استفاده از شبکه عصبی و مقایسه آن با روش‌های پیش‌بینی ریاضی. پژوهش‌نامه اقتصادی، دوره۷ (۲)، ۲۵۱-۲۳۷.

۱۲) مرادی, مهدی؛ مهدی جباری نوقابی و محمدمهدی رونقی(۱۳۹۵). پیش‌بینی سری‌های زمانی کوتاه مدت، میان‌مدت و بلند-مدت بازده سهام بورس اوراق بهادار تهران با استفاده از الگوریتم L-CO-R. پنجمین کنفرانس بین‌المللی حسابداری و مدیریت و دومین کنفرانس کارآفرینی و نوآوری‌های باز(ص۱۵-۱)، تهران، تیرماه ۱۳۹۵، همایش‌گران مهر اشراق.

۱۳) Chambers,L.(2016).” Daily stock movements prediction using and integrating three models of analytical prediction of adaptive-fuzzy inductive inference system, artificial neural networks and supporting vector machines” London ,Champan&Hall : 113_119.

۱۴) Chen-Fong Tsai; Yen- Jiun Chiou, (2010). “Earnings Management Prediction: A pilot Study of combining Neural Networks and Decision Trees”, Expert systems with Application, 36/ 7183-7191.

۱۵) Pawlak Z,(2016). ”Estimating the effect of market costs on stock prices using nonparametric models of learning machines”, vol. 1, pp. 1-58.

۱۶) Tsang,P,M.Kwok,P,Choy,S.O.,Kwan,R,Ng,S.C,Mak,J.,Tsang,J.Koong,K,Wong,T,L.(2013). Forecasting the direction of the stock price index using the machine learning method on the Hong Kong Stock Exchange, 39: 202_268.

۱۷) Ying Wei ,D,E,(2016).” Forecast of stock markets based on fuzzy time series model”,New York: Addison_Wesley,USA: 180_187.

۱۸) Zmijewski Mark.E. (2016)” Karachi stock market prediction on the day of market closure using the learning machine method” Journal of Accounting Research Vol 22 Supplement.

یادداشت ها

[۱] Chen

[۲] Rakfeler

[۳] **Decision Trees**

[٤] Chambers

[٥] Ying Wei

[٦] adaptive network-based fuzzy inference system(ANFIS)

[۷] empirical mode decomposition

[۸] autoregressive

[۹] support vector regression (SVR)

[۱۰] Pawlak

[۱۱] Zmijewski

[۱۲] Tsang

[۱۳] Logistic Regression (LR)

[۱٤] Support Vector Machines (SVMs)

[۱٥] Multilayer Perceptron (MLP)

[۱٦] Rough Set

[۱۷]Genetic Algorithm

[۱۸]Johnson’s Algorithm

[۱۹] Holtes’s