



**بررسی مقایسه‌ای دقت پیش‌بینی مدل‌های ماشین بردار پشتیبان، شبکه بیزین و سی‌فایو در
پیش‌بینی قیمت‌گذاری کمتر از واقع شرکت‌های پذیرفته‌شده در بورس اوراق بهادار و فرابورس**

بی‌تا دهقان خانقاهی^۱

جمال بحری ثالث^۲

سعید جبارزاده کنگرلویی^۳

علی آشتاب^۴

تاریخ دریافت مقاله : ۹۸/۱۱/۱۸ تاریخ پذیرش مقاله : ۹۹/۰۱/۲۹

چکیده

پژوهش‌های پیشین در زمینه عملکرد کوتاه‌مدت عرضه عمومی اولیه، بیانگر این واقعیت است که سهام تازه عرضه شده در کوتاه‌مدت، عملکرد بهتری نسبت به بازار دارد. مدل‌های آماری توانسته‌اند پیش‌بینی‌های خوبی در مورد عملکرد این‌گونه سهام ارائه دهند ولی مفروضات محدودکننده برخی از این مدل‌ها بر اثربخشی‌شان موثر بوده است. بنابراین، روش‌های دیگری برای مقابله با این محدودیت‌ها و بهبود پیش‌بینی‌ها معرفی شدند. از آنجایی که عرضه عمومی اولیه موضوع بااهمیتی در بازار سرمایه است، در این پژوهش به بررسی مدل‌های طبقه‌بندی مختلف برای دستیابی به مدلی که از کارایی و دقت بالایی در پیش‌بینی قیمت‌گذاری کمتر از واقع سهام عرضه عمومی اولیه برخوردار است، پرداخته شد. به‌منظور دستیابی به این هدف، ۸۴ شرکت بورسی و ۵۴ شرکت فرابورسی طی سال‌های ۱۳۸۲ تا ۱۳۹۶ به‌روش حذف سیستماتیک جهت تحلیل در نظر گرفته شد. نتایج نشان داد که مدل‌های ماشین بردار پشتیبان، شبکه بیزین و سی‌فایو از دقت بالایی در پیش‌بینی قیمت‌گذاری کمتر از واقع برخوردارند. همچنین نتایج نشان داد که متغیرهای مهم تاثیرگذار شامل رشد دارایی‌ها، دوره تصدی حسابرس، تخصص حسابرس در صنعت، نسبت فعالیت‌های تامین‌مالی، نسبت قیمت به سود هرسهام، بازده دارایی‌ها، نسبت فعالیت‌های عملیاتی، اندازه موسسه حسابرسی، فرصت‌های رشد و نوسانات قیمت سهام هستند.

کلمات کلیدی

عرضه عمومی اولیه، قیمت‌گذاری کمتر از واقع، مدل‌های طبقه‌بندی، یادگیری ماشین

۱- گروه حسابداری، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران. bita_d14@yahoo.com

۲- گروه حسابداری، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران. (نویسنده مسئول) j.bahri@iaurmia.ac.ir

۳- گروه حسابداری، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران. s.jabbarzadeh@iaurmia.ac.ir

۴- گروه حسابداری، دانشگاه ارومیه، ارومیه، ایران. a.ashtab@urmia.ac.ir

اصطلاح عرضه عمومی اولیه^۱ که رواج آن به بازارهای پررونق ۱۹۹۰ برمی‌گردد، عبارت است از این که شرکت برای اولین بار در طول دوران فعالیت خود اقدام به عرضه سهام به‌عموم مردم نماید [۲۵]. در ایران پس از ابلاغ سیاست‌های اصل ۴۴ توسط مقام معظم رهبری، عرضه عمومی اولیه با ورود شرکت‌های دولتی مشمول این اصل، قدرت گرفت و با توجه به اجرای این اصل و گسترش دامنه خصوصی‌سازی در کشور در دهه اخیر، مورد توجه شرکت‌ها واقع شد. یکی از مسائلی که شرکت‌ها در این زمینه با آن روبرو هستند، قیمت‌گذاری صحیح سهام عرضه‌های عمومی اولیه است [۳۰]. قیمت‌گذاری واقعی سهام عرضه شده با در نظر گرفتن منافع سرمایه‌گذاران و شرکت منتشرکننده اوراق، نیامند دخالت دادن متغیرهای زیادی است. اگر قیمت چنین عرضه‌هایی کمتر از واقع تعیین شود، شرکت به منابع مالی مورد نیاز خود دست پیدا نکرده و چنانچه قیمت عرضه اولیه بیش از واقع تعیین شود، سرمایه‌گذاران از خرید سهام خودداری نموده و در نتیجه، شرکت از دستیابی به منابع مالی مورد نیاز محروم خواهد ماند. قیمت‌گذاری کمتر از واقع^۲ عرضه عمومی، یکی از مهم‌ترین ناهنجاری‌های بازار سرمایه است. در سطح بین‌الملل، پژوهش‌های زیادی نشان داده است که در عرضه‌های عمومی اولیه، شرکت‌ها سهام خود را به قیمتی کمتر از قیمتی که سرمایه‌گذاران تمایل به پرداخت آن دارند، عرضه می‌کنند [۱۳]. به‌عبارت دیگر، شرکت‌ها با فداکردن منافع سهامداران اولیه خود و به اصطلاح، «باقی گذاشتن مبالغ زیادی پول روی میز» هزینه قابل توجهی متحمل شده و برای سرمایه‌گذاران جدید بازده مثبت غیرعادی ایجاد می‌کنند. شکل‌گیری بازده کوتاه‌مدت سهام عرضه اولیه به قیمت‌گذاری کمتر از واقع این سهام نسبت داده می‌شود [۳۱] و [۳۲].

در بیشتر پژوهش‌ها در زمینه قیمت‌گذاری کمتر از واقع، به‌بررسی رویکرد آمار کلاسیک از جمله رگرسیون پرداخته شده است. با وجود این‌که مدل‌های آماری توانسته‌اند پیش‌بینی‌های خوبی را در مورد عملکرد سهام تازه عرضه شده ارائه دهند ولی برای مدل‌سازی روابط بین متغیرهایی که مشاهدات آن‌ها دقیق نیستند، نمی‌توان مدل‌های معمولی آماری را که براساس مشاهدات دقیق و برخی مفروضات توزیعی استوار هستند، به‌کار برد. نیز، مفروضات محدودکننده برخی از این مدل‌ها (ابهام در روابط متغیرهای ورودی و خروجی مدل، هم خطی بین متغیرهای توضیحی، کمبود مشاهدات...) بر اثربخشی‌شان موثر بوده است [۱۸]. بنابراین، پاسخ به این سوال که آیا می‌توان مدلی ارائه داد که بتواند قیمت‌گذاری کمتر از واقع سهام عرضه عمومی اولیه را با دقت بالایی پیش‌بینی کند؟، از موضوعات حائز اهمیت در این زمینه است. پیش‌بینی فرایندی است که رفتاری را بر مبنای رفتارهای گذشته حدس

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

می‌زند. برای انجام پیش‌بینی معمولاً به داده‌های گذشته اتکا می‌شود. بدین صورت که با تجزیه و تحلیل داده‌های گذشته، در پی شناسایی الگویی برمی‌آییم که بتوان آینده را ترسیم کرد. برای بررسی این امر در این پژوهش، از روش‌های یادگیری ماشین^۲ استفاده می‌شود. یادگیری ماشین، مطالعه و ساخت الگوریتم‌هایی را که می‌توانند براساس داده‌ها یادگیری و پیش‌بینی انجام دهند، بررسی می‌کند. چنین الگوریتم‌هایی از دستورهای برنامه پیروی صرف نمی‌کنند و از طریق مدل‌سازی از داده‌های ورودی نمونه، پیش‌بینی یا تصمیم‌گیری می‌کنند. مدل‌های یادگیری ماشین نسبت به مدل‌های آماری به دلایلی همچون در نظر نگرفتن توزیع جوامع، نداشتن مشکل هم‌خطی بین متغیرهای توضیحی و یادگیری تطبیقی، دقت بیشتری دارند. این پژوهش بر مقایسه عملکرد پیش‌بینی الگوریتم‌های درخت‌های تصمیم‌گیری^۴ (سی‌فایو^۵، چاید^۶، کواست^۷، مدل رگرسیون و طبقه‌بندی^۸)، مدل ماشین بردار پشتیبان^۹، شبکه بیزین^{۱۰}، شبکه عصبی^{۱۱}، رگرسیون لجستیک^{۱۲} و تحلیل ممیزی^{۱۳} متمرکز می‌شود. هدف، ارائه بهترین مدل برای پیش‌بینی عملکرد کوتاه‌مدت عرضه‌های عمومی اولیه است تا بتوان از این طریق به تصمیم‌گیری درست سرمایه‌گذاران کمک کرد.

با بررسی نگارنده، پژوهشی در خصوص مقایسه دقت پیش‌بینی مدل‌های یادگیری ماشین در پیش‌بینی پدیده قیمت‌گذاری کمتر از واقع در شرکت‌های پذیرفته شده در بورس اوراق بهادار و فرابورس صورت پذیرفته است. انتظار می‌رود با توجه به توانایی این مدل‌ها، از آن برای موضوع پیچیده عرضه‌های عمومی اولیه استفاده شود، زیرا قیمت‌گذاری سهام در عرضه‌های عمومی اولیه مستلزم فرایند تصمیم‌گیری بسیار دشواری است. به‌علاوه، قیمت‌گذاری یک شرکت به‌منظور تعیین قیمت عرضه عمومی اولیه به متغیرهای زیادی بستگی دارد که خود این متغیرها مستقل از هم بوده و روابط بین آن‌ها مشخص نیست. در چنین شرایطی قیمت‌گذاری عرضه‌های عمومی اولیه برای سرمایه‌گذار و شرکت عرضه‌کننده مشکل خواهد بود.

ساختار مقاله حاضر بدین صورت است که در بخش اول به مبانی نظری و پیشینه پژوهش پرداخته می‌شود. در ادامه، به توصیف روش پژوهش و متغیرهای پژوهش پرداخته و سپس، روش تجزیه و تحلیل داده‌ها و یافته‌های پژوهش مطرح می‌شود. در آخر، نتیجه‌گیری و پیشنهادهای پژوهش بیان می‌گردد.

مبانی نظری و پیشینه تجربی پژوهش

موضوع عرضه عمومی اولیه سهام و بازده کوتاه‌مدت غیرعادی آن، پژوهش‌های متعددی در سطح جهان و در بازارهای سهام را به‌خود اختصاص داده است [۱۳]. پدیده قیمت‌گذاری کمتر از واقع سهام

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

عرضه عمومی اولیه، برای نخستین بار در بازار سرمایه آمریکا توسیله ایبوتسون^{۱۴} (۱۹۷۵) شناسایی شد [۱۹]. طبق نظر وی، شرکت‌هایی که سهام خود را برای اولین بار از طریق بورس اوراق بهادار به عموم مردم عرضه می‌کنند، سهام را به قیمتی پایین‌تر از ارزش ذاتی عرضه می‌کنند؛ به‌گونه‌ای که این شرکت‌ها از تاریخ عرضه تا پایان اولین ماه داد و ستد بورس به‌طور متوسط ۱۱/۴ درصد بازده مثبت ایجاد کرده‌اند [۲۱]. ایبوتسون و جف^{۱۵} (۱۹۷۵) اظهار کردند که قیمت‌گذاری کمتر از واقع سهام عرضه عمومی اولیه پدیده‌ای ادواری است، یعنی در دوره‌هایی که تعداد عرضه‌های عمومی اولیه در بازار زیاد است، قیمت‌گذاری کمتر از واقع بیشتر بوده و در دوره‌هایی که تعداد عرضه‌های عمومی اولیه در بازار کمتر است، میزان قیمت‌گذاری کمتر از واقع (بازده غیرعادی کوتاه‌مدت) کمتر می‌شود [۲۲].

پژوهشگران دلایل متعددی برای بازده‌های غیرعادی سهام عرضه‌های عمومی اولیه ارائه کرده‌اند. این دلایل در قالب تئوری‌های مبتنی بر عدم تقارن اطلاعاتی، عوامل نهادی، عوامل نظارتی و عوامل رفتاری شکل گرفته‌اند [۷] و [۱۱]. عدم تقارن اطلاعاتی^{۱۶} نقش اساسی را بین سه گروه ذینفع در فرایند عرضه عمومی اولیه، یعنی شرکت عرضه‌کننده سهام، بانک سرمایه‌گذار و سرمایه‌گذاران بالقوه در بازار ایفا می‌کند [۲۰] و [۲۶]. عدم تقارن اطلاعاتی، پدیده‌ای منفی است که تصمیم‌گیری‌های نامناسب اقتصادی توسط سرمایه‌گذاران را به‌دنبال دارد و زمانی به‌وجود می‌آید که یکی از طرفین درگیر در عرضه عمومی اولیه سهام، اطلاعات بیشتری راجع به شرکت عرضه‌کننده سهام، نسبت به سایرین داشته باشد، مشروط بر این‌که از این اطلاعات در زمان برقراری ارتباط با طرف دیگر، به‌طور موثر و به‌نفع خود استفاده نماید [۱۶]. بنابراین، قیمت‌گذاری کمتر از واقع برای ایجاد تعادل منافع تمامی شرکت‌کنندگان در عرضه اولیه ضروری به‌نظر می‌رسد. مدل‌هایی که بر محور عدم تقارن اطلاعاتی شکل گرفته‌اند و توضیح‌دهنده پدیده قیمت‌گذاری کمتر از واقع هستند، عبارتند از:

۱. شناخته شده‌ترین مدل، فرضیه نفرین برنده^{۱۷} است که توسط راک^{۱۸} (۱۹۸۶) ارائه شد. وی معتقد بود برخی از سرمایه‌گذاران دارای اطلاعات بهتری درباره ارزش واقعی سهام شرکت نسبت به سایر ذینفان درگیر در فرایند عرضه عمومی اولیه هستند. بر همین اساس، سرمایه‌گذاران بااطلاع، تنها متقاضی خرید سهام عرضه‌های اولیه کمتر از واقع قیمت‌گذاری شده می‌شوند؛ درحالی‌که سرمایه‌گذاران کم اطلاع، به‌دلیل عدم قدرت تشخیص، متقاضی خرید کلیه سهام عرضه‌های اولیه می‌شوند. از آنجایی که تعداد سهام عرضه شده محدود است با تقاضای بیش از حد سرمایه‌گذاران بااطلاع مواجه شده و این امر موجب سهمیه‌بندی سهام برای سرمایه‌گذاران بی‌اطلاع می‌شود. اگر سهام

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

عرضه شده بیش از واقع قیمت‌گذاری شوند، منحصرآ عاید سرمایه‌گذاران کم اطلاع شده و در نتیجه برای آن‌ها بازده منفی ایجاد می‌کند [۲۴].

۲. مدل دیگری از این تئوری‌ها بر عدم تقارن اطلاعاتی بین شرکت عرضه‌کننده سهم و سرمایه‌گذاران بالقوه تاکید می‌کند. براساس این مدل، شرکت عرضه‌کننده سهم اطلاعات بهتری درباره ارزش فعلی و ریسک جریان‌های نقدی واحد تجاری نسبت به بقیه ذینفعان درگیر در فرایند عرضه عمومی دارد. بنابراین قیمت‌گذاری کمتر از واقع به عنوان علامتی برای نشان دادن ارزش واقعی شرکت و جذب سرمایه‌گذاران به کار می‌رود [۲۹].

در ادامه به برخی از پژوهش‌های داخلی و خارجی انجام شده در رابطه با موضوع پژوهش پرداخته می‌شود:

بابا و سویل (۲۰۱۹) در پژوهش خود با بیان نقاط ضعف مدل‌های رگرسیون خطی، از مدل درخت تصمیم برای پیش‌بینی بازده‌های سهام عرضه عمومی اولیه در بازه زمانی ۲۰۱۸-۱۹۹۸ در بورس اوراق بهادار ترکیه و بررسی مقایسه‌ای دقت پیش‌بینی این مدل‌ها پرداختند. نتایج نشان داد که روش درخت تصمیم دارای دقت بیشتری در پیش‌بینی بازده‌های عرضه عمومی اولیه دارد. همچنین متغیرهای عایدات عرضه عمومی اولیه و حجم عرضه عمومی اولیه قدرت بیشتری در توضیح این دقت داشتند [۱۳].

کولاک و فو^{۱۹} (۲۰۱۸) پیش‌بینی عملکرد عرضه عمومی اولیه را با روش یادگیری ماشین در دوره زمانی ۲۰۱۶-۱۹۸۵ مورد بررسی قرار دادند. آن‌ها در پژوهش خود ضمن بیان نقایص مدل‌های آماری کلاسیک، به بررسی تطبیقی روش‌های آماری کلاسیک و یادگیری ماشین پرداختند و به این نتیجه رسیدند که مدل‌های یادگیری ماشین توان بیشتری در پیش‌بینی بازده‌های غیرعادی سهام عرضه عمومی اولیه دارند. متغیرهای مهم تاثیرگذار پژوهش آن‌ها اصلاح قیمت در دوره ثبت دفتری، حجم عرضه اولیه و روزهای انتظار از تاریخ ثبت تا تاریخ انتشار بودند [۱۷].

وانگ و همکاران^{۲۰} (۲۰۱۸) در پژوهش خود کاربرد سیستم‌های فازی عصبی، به‌عنوان یکی از روش‌های یادگیری ماشین را برای پیش‌بینی قیمت‌گذاری کمتر از واقع در زمان عرضه عمومی اولیه را بررسی کردند. آن‌ها معتقد بودند که شرکت‌ها در زمان عرضه عمومی اولیه با فداکردن منافع سهامداران اولیه خود، برای سرمایه‌گذاران جدید بازده مثبت غیرعادی ایجاد می‌کنند. نتایج پژوهش آن‌ها نشان داد که مدل‌های فازی نسبت به سایر مدل‌ها از دقت بالاتری برخوردار است [۳۱].

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پاییز ۱۳۹۹

باستی و همکاران^{۲۱} (۲۰۱۵) به بررسی عملکرد کوتاه‌مدت عرضه عمومی اولیه در بین سال‌های ۲۰۱۳-۲۰۰۵ در بورس اوراق بهادار ترکیه پرداختند. برای شناسایی عوامل موثر بر عملکرد کوتاه‌مدت عرضه عمومی اولیه از روش‌های یادگیری ماشین درخت تصمیم‌گیری و ماشین بردار پشتیبان و برای ارزیابی عملکرد مدل‌های پیش‌بینی از روش اعتبارسنجی متقابل استفاده کردند. نتایج نشان داد که پدیده قیمت‌گذاری کمتر از واقع در بورس اوراق بهادار ترکیه وجود دارد و عوامل موثر بر این پدیده تمایلات بازار، حجم فروش سالانه، دوره گردش کل دارایی‌ها، روش‌های فروش سهام عرضه عمومی اولیه، روش‌های پذیره‌نویسی، قیمت پیشنهادی، نسبت بدهی و تعداد سهام عرضه شده دارای بیشترین تاثیر بر عملکرد کوتاه‌مدت عرضه عمومی اولیه هستند. همچنین نتایج نشان داد که مدل‌های مورد استفاده در پژوهش، توان پیش‌بینی عملکرد کوتاه‌مدت عرضه عمومی اولیه را با دقت بالایی دارا هستند [۱۴].

اسماعیلی و همکاران (۱۳۹۷) در پژوهشی به بررسی پیش‌بینی عملکرد کوتاه‌مدت عرضه عمومی اولیه سهام با روش‌های یادگیری ماشین در شصت شرکت دارای عرضه عمومی اولیه در بورس اوراق بهادار تهران طی دوره ۱۳۸۴-۱۳۹۴ پرداختند. آن‌ها برای ارزیابی و کنترل عملکرد مدل‌های طبقه‌بندی نزدیک‌ترین همسایگی، ماشین بردار پشتیبان، درخت تصمیم‌گیری کارت و بیز ساده از روش اعتبارسنجی متقابل ده تایی استفاده کردند. نتایج نشان داد که در بین مدل‌ها، مدل‌های نزدیک‌ترین همسایگی و ماشین بردار پشتیبان از دقت و کارایی بالایی در پیش‌بینی عملکرد کوتاه‌مدت عرضه عمومی اولیه برخوردار هستند [۱].

نجفی‌مقدم (۱۳۹۷) به بررسی تاثیر برخی از متغیرهای ساختاری مانند نوع صنعت، سن شرکت در زمان عرضه، ارزش بازار، نوع عرضه و نیز شرایط بازار بر روی بازدهی‌های غیرمتعارف عرضه‌های اولیه در بازار بورس تهران پرداخت. در بین متغیرهای توضیحی، تنها متغیر نوع صنعت بر بازدهی غیرمتعارف عرضه اولیه در دوره‌های مختلف بی‌تاثیر بود. متغیر سن شرکت و اندازه دارای رابطه منفی با بازدهی غیرمتعارف عرضه اولیه است. در رابطه با متغیر نوع عرضه مشخص شد که عرضه‌های دولتی عموماً بازدهی‌های غیرمتعارف آن کمتر است یا به‌عبارت دیگر، دولت (خصوصی‌سازی)، شرکت‌ها را چندان ارزان به‌عموم مردم واگذار نکرده است [۱۱].

صالحی (۱۳۹۵) به بررسی تاثیر عوامل موثر بر ارزش‌گذاری سهام در عرضه‌های عمومی اولیه با رویکرد معادلات ساختاری پرداخت. نتایج نشان داد که بین شاخص کل بورس، نرخ ارز، تولید ناخالص داخلی، نرخ تورم، قیمت نفت، سیاست داخلی سازمان، رقابتی بودن صنعت، حجم سرمایه‌گذاری صنعت

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

و ارزش‌گذاری سهام در عرضه‌های عمومی اولیه در بورس اوراق بهادار رابطه معنادار وجود دارد و بین نرخ رشد نقدینگی، نرخ سود بانکی و ارزش‌گذاری سهام در عرضه‌های عمومی اولیه در بورس اوراق بهادار ارتباطی یافت نشد [۷].

روش‌شناسی پژوهش

این پژوهش از لحاظ روش، یک پژوهش از نوع همبستگی است. در این پژوهش از الگوریتم‌های مدل‌های ماشین بردار پشتیبان، شبکه بیزین و درخت تصمیم‌گیری سی‌فایو به‌عنوان مدل‌های پیش‌بینی استفاده شد. جامعه آماری این پژوهش کلیه شرکت‌هایی هستند که در بازه زمانی ۱۳۸۲ تا ۱۳۹۶ برای اولین بار اقدام به عرضه عمومی سهام خود در بورس اوراق بهادار تهران و فرابورس ایران نموده‌اند. جدول ۱ روند انتخاب نمونه آماری پژوهش را نشان می‌دهد.

جدول ۱: روند انتخاب نمونه آماری پژوهش

۲۷۶	کلیه شرکت‌هایی که در بازه زمانی ۱۳۸۲ تا ۱۳۹۶ در بورس و فرابورس پذیرش شده‌اند
۲۴	کسر می‌شود: شرکت‌هایی که بعد از پذیرش، عرضه عمومی اولیه نداشتند
۸۰	کسر می‌شود: شرکت‌هایی که عضو بانک‌ها، شرکت‌های سرمایه‌گذاری، هلدینگ‌ها و بیمه‌ها هستند
۳۰	کسر می‌شود: شرکت‌هایی که پایان سال مالی آن‌ها منتهی به ۲۹ اسفند نیست
۴	کسر می‌شود: شرکت‌هایی که در دوره مورد بررسی دارای توقف معاملاتی بیش از یک ماه هستند
۱۳۸	نمونه آماری قابل آزمون با در نظر گرفتن پیش‌فرض‌ها

با اعمال شرایط فوق، تعداد ۱۳۸ شرکت شامل ۸۴ شرکت از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران و ۵۴ شرکت از شرکت‌های پذیرفته شده در فرابورس ایران، پژوهش انتخاب شد.

در این پژوهش برای تجزیه و تحلیل داده‌ها و انجام آزمون‌های آماری از نرم‌افزار SPSS Modeler18 استفاده شد. الگوریتم‌های یادگیری ماشین مورد استفاده در پژوهش به شرح زیر است:

الگوریتم ماشین بردار پشتیبان: ماشین بردار پشتیبان، یکی از قوی‌ترین و دقیق‌ترین الگوریتم‌های یادگیری ماشین است که جهت طبقه‌بندی و تفکیک گروه‌ها به کار می‌رود. این الگوریتم روش‌های آماری و یادگیری ماشین را باهم ترکیب می‌کند و اساس نظری آن بر پایه تئوری یادگیری آماری قرار دارد [۱۴]. این الگوریتم از تئوری بهینه‌سازی برای طبقه‌بندی استفاده می‌کند و براساس تئوری یادگیری آماری، خطای طبقه‌بندی را به حداقل می‌رساند [۱]. با مشخص شده داده‌های پژوهش، مدل ماشین بردار پشتیبان داده‌ها را به گروه‌های متمایزی تقسیم می‌کند. این مدل‌ها دارای خواص کلی طبقه‌بندی داده‌ها با حداکثر قابلیت تعمیم، رسیدن به نقطه بهینه تفکیک داده‌ها، تعیین

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

خودکار ساختار بهینه برای طبقه‌بندی کننده و امکان مدل کردن داده‌های غیرخطی با استفاده از تجزیه و تحلیل اجزای اصلی هستند [۹].

الگوریتم درخت تصمیم‌گیری: روش مبتنی بر درخت تصمیم، یکی از ابزارهای قوی برای دسته‌بندی و اعتباردهی محسوب می‌شود. درخت تصمیم از نظریه اطلاع و مقدار آنتروپی جهت انتخاب بهترین متغیر شروع پیمایش استفاده می‌کند. هر راس درخت به‌عنوان یک کلاس یا قاعده، نمایشگر یک آزمایش یا تصمیم یکتاست. یال‌های هر راس، متناظر با احتمال‌های حاصل از آزمایش روی راس است. پیشامد داده‌ها را به چندین زیرمجموعه افراز می‌کند که توسط برگ‌های این درخت شناخته می‌شود. درخت تصمیم برخلاف سایر فنون به تولید قانون می‌پردازد و پیش‌بینی خود را در قالب قوانین توضیح می‌دهد، در حالی که در سایر فنون، پیش‌بینی نهایی بدون چگونگی اجرا بیان می‌شود [۱۰]. قوانین ایجاد شده در درخت تصمیم به‌صورت اگر و آن‌گاه بیان می‌شوند. رایج‌ترین الگوریتم‌های درخت تصمیم شامل چاید، سی‌فایو، رگرسیون و طبقه‌بندی و کواست هستند [۱۴]. یکی از پرکاربردترین درخت‌های تصمیم، سی‌فایو است. چرا که این الگوریتم با توجه به اریب کمتر نسبت به الگوریتم‌های مشابه برای داده‌هایی که دارای نویز هستند، مناسب بوده و استفاده از روش بوستینگ^{۲۲} باعث افزایش عملکرد آن می‌شود. بوستینگ، تکنیکی برای بهبود دقت پیش‌بینی است. مدل درخت تصمیم مزایای بسیاری همانند قابل فهم بودن قوانین ایجاد شده توسط درخت تصمیم، برخورداری از کارایی بالاتر، مناسب برای نمونه‌هایی با حجم زیاد داده و دارا بودن دقت طبقه‌بندی بالاتر دارد [۱۹].

الگوریتم شبکه بیزین: شبکه بیزین یک الگوی نموداری احتمالاتی است که مجموعه‌ای از متغیرها و احتمالات مربوط به هریک را نشان می‌دهد. ساختار شبکه بیزین در واقع یک نمایش نموداری از اثرات متقابل متغیرهایی است که باید الگوبندی شوند و علاوه بر این که کیفیت رابطه بین متغیرهای مساله را نشان می‌دهد، کمیت ارتباط بین این متغیرها را نیز به نمایش می‌گذارد که به‌صورت عددی از توزیع احتمالات مشترک آن‌ها استفاده می‌کند. این روش بر مبنای احتمالات شرطی (قانون بیز) قرار دارد (به شرح زیر).

$$P(C/A) = \frac{P(C/A) * P(C)}{P(A)} : I$$

ویژگی بسیار مثبت الگوریتم بیزین این است که امکان اثبات بهینگی دارد [۱۰]. شبکه بیزین امکان محاسبات رو به جلو و رو به عقب را برای تحلیل‌گر ایجاد می‌کند. در واقع نه تنها از تجمیع وضعیت پارامترهای علت می‌توان به وضعیت معلول رسید، بلکه در این روش با در اختیار داشتن وضعیت معلول

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

یا همان پارامتر پیش‌بینی شونده با یک فرایند بازگشت به عقب، امکان محاسبه وضعیت پارامترهای تاثیرگذار به وجود خواهد آمد.

روش‌های ذکر شده، با پیاده‌سازی سریع و آسان، دقت پیش‌بینی بسیار بالایی دارند و می‌توانند تعداد زیادی از متغیرهای ورودی را بدون پیش‌برازش، اداره کنند.

متغیرهای پژوهش

متغیرهای مورد استفاده در پژوهش در جدول ۲ بیان شده است.

جدول ۲: تعریف عملیاتی متغیرهای پژوهش

متغیر وابسته	توصیف
قیمت‌گذاری کمتر از واقع (اگروال و همکاران، ۱۹۹۳)	از طریق محاسبه بازده غیرعادی اولیه ماهانه محاسبه می‌شود. $MAARit = Rit - Rmt$ $MAARit =$ بازده کوتاه‌مدت تعدیل شده سهم i در زمان t بعد از عرضه اولیه
متغیرهای مستقل	توصیف
عدم تقارن اطلاعاتی (حیدری و همکاران، ۱۳۹۶)	کیفیت اقلام تعهدی: از طریق محاسبه انحراف معیار باقیمانده رگرسیون اقلام تعهدی، بر جریان نقدی عملیاتی گذشته، حال، آینده و تغییرات در درآمد اموال، ماشین‌آلات و تجهیزات از رابطه زیر به دست می‌آید. $TCA_{jt} = \beta_0 + \beta_1 \text{Cashflow}_{t-1} + \beta_2 \text{Cashflow}_t + \beta_3 \text{Cashflow}_{t+1} + \beta_4 \Delta \text{Revenuet} + \beta_5 \text{Fixed Assetst} + \varepsilon_t$
	نوسانات قیمت سهام: انحراف معیار درصد تغییرات ماهانه قیمت سهام
	خطای پیش‌بینی سود: حاصل تقسیم قدرمطلق تفاوت سود واقعی و سود پیش‌بینی شده هر سهم بر قدرمطلق سود پیش‌بینی شده هر سهم
	اندازه شرکت: لگاریتم طبیعی جمع دارایی‌ها
حاکمیت شرکتی (پنگ و همکاران، ۲۰۱۸)	فرصت‌های رشد: حاصل تقسیم ارزش دفتری شرکت بر ارزش بازاری آن
	اندازه هیات مدیره: تعداد کل مدیران حاضر در هیات مدیره که در گزارش سالانه شرکت‌ها ذکر شده است. دوگانگی: با استفاده از متغیر مجازی صفر-یک نشان داده می‌شود. اگر مدیرعامل، رئیس یا نایب رئیس هیئت مدیره باشد عدد یک و در غیر این صورت عدد صفر را به خود اختصاص خواهد داد.
	استقلال هیات مدیره: نسبت اعضای غیرموظف هیات مدیره به کل اعضای هیات مدیره
محافظه‌کاری (لین و تیان، ۲۰۱۲)	ترکیب مالکیت: در این پژوهش، درصد سهامداری که بیش از ۵ درصد یا بیش از ۵ میلیارد ریال ارزش اسمی کل سهام منتشره شرکت را در اختیار داشته باشند، به‌عنوان تمرکز مالکیت محسوب می‌شوند.
	درصد سهامداران نهادی: مجموع درصد سهام شرکت که متعلق به بانک‌ها، بیمه‌ها، نهادهای مالی، هلدینگ‌ها، سازمان‌ها، نهادها و شرکت‌های دولتی است.
محافظه‌کاری (لین و تیان، ۲۰۱۲)	برای اندازه‌گیری محافظه‌کاری، از مدل گیولی و هاین (۲۰۰۰) که مبتنی بر متغیرهای حسابداری است، استفاده می‌شود.

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

اندازه موسسه حسابرسی: با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود. اگر شرکت توسط سازمان حسابرسی مورد رسیدگی قرار گرفته باشد عدد یک و در غیر این صورت عدد صفر اختیار می‌شود.	کیفیت حسابرسی (چانگ و همکاران، ۲۰۰۸)
دوه تصدی حسابرس: با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود. اگر حسابرس سال بعد، حسابرس شرکت صاحبکار باشد، عدد یک و در غیر این صورت عدد صفر اختیار می‌شود.	
درصد سهامداران نهادی: مجموع درصد سهام شرکت که متعلق به بانک‌ها، بیمه‌ها، نهادهای مالی، شرکت‌های هلدینگ، سازمان‌ها، نهادها و شرکت‌های دولتی است.	
نوع اظهارنظر حسابرس: با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود. اگر شرکت دارای اظهارنظر مقبول باشد عدد یک و در غیر این صورت عدد صفر را به خود اختصاص می‌دهد.	
تخصیص حسابرس در صنعت: در این پژوهش از معیار سهم بازار به عنوان شاخصی برای اندازه‌گیری تخصص صنعت موسسه حسابرسی استفاده شد که از حاصل تقسیم مجموع دارایی‌های تمام صاحبکاران یک موسسه حسابرسی خاص در یک صنعت خاص به مجموع دارایی‌های صاحبکاران در این صنعت محاسبه می‌شود. این معیار با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود. اگر شرکت توسط موسسه حسابرسی متخصص صنعت حسابرسی شده باشد عدد یک و در غیر این صورت عدد صفر را به خود اختصاص می‌دهد.	
نسبت جاری: حاصل تقسیم دارایی‌های جاری بر بدهی‌های جاری	نسبت‌های مالی (سیاه‌تیری، ۱۳۸۹)
نسبت آنی: حاصل تقسیم دارایی‌های نقد بر بدهی‌های جاری	
نسبت نقدینگی: حاصل تقسیم وجه نقد و سرمایه گذاری‌های کوتاه مدت بر بدهی‌های جاری	
نسبت بدهی: حاصل تقسیم کل بدهی‌ها بر کل دارایی‌ها	
نسبت بدهی به حقوق صاحبان سهام: حاصل تقسیم کل بدهی‌ها بر حقوق صاحبان سهام	
نسبت مالکانه: حاصل تقسیم حقوق صاحبان سهام بر کل دارایی‌ها	
نسبت فعالیت‌های عملیاتی: حاصل تقسیم مابه التفاوت جریان وجه نقد فعالیت‌های عملیاتی در پایان و ابتدای دوره بر خالص جریان وجه نقد فعالیت‌های عملیاتی در ابتدای دوره	
نسبت فعالیت‌های سرمایه‌گذاری: حاصل تقسیم مابه التفاوت جریان وجه نقد فعالیت‌های سرمایه‌گذاری در پایان و ابتدای دوره بر خالص جریان وجه نقد فعالیت‌های سرمایه‌گذاری در ابتدای دوره	
نسبت فعالیت‌های تامین مالی: حاصل تقسیم مابه التفاوت جریان وجه نقد فعالیت‌های تامین مالی در پایان و ابتدای دوره بر خالص جریان وجه نقد فعالیت‌های تامین مالی در ابتدای دوره	
نسبت‌های فعالیت: شامل نسبت گردش مجموع دارایی‌ها و نسبت گردش موجودی‌ها نسبت‌های سودآوری: شامل نسبت حاشیه سود خالص، نسبت بازده دارایی‌ها، نسبت بازده سرمایه، سود هر سهم و نسبت قیمت به سود هر سهم	
ضریب تغییرات فروش: این معیار، ویژه واحد تجاری است که با استفاده از اطلاعات در طول یک دوره ۵ ساله محاسبه می‌شود.	عدم اطمینان محیطی (شهاب‌نویسی، ۱۳۹۶)
ضریب تغییرات سود قبل از مالیات: این معیار نیز، ویژه واحد تجاری است که با استفاده از اطلاعات در طول یک دوره ۵ ساله محاسبه می‌شود.	

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

مقایسه سود پیش‌بینی شده و سود واقعی مدیریت: با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود. اگر سود پیش‌بینی شده از سود واقعی بیشتر باشد عدد یک (مدیر بیش اعتماد) و در غیر این صورت عدد صفر را به خود اختصاص خواهد داد.	
نسبت مخارج سرمایه‌ای به جمع دارایی‌های ابتدای دوره: با استفاده از متغیر مجازی صفر - یک نشان داده می‌شود که از طریق محاسبه میانه مخارج سرمایه‌ای حاصل می‌شود. اگر مخارج سرمایه‌ای تقسیم بر مجموع دارایی‌ها در یک سال مشخص، بزرگتر از میانه به سطوح مخارج سرمایه‌ای بر مجموع دارایی‌ها باشد عدد یک و در غیر این صورت عدد صفر را به خود اختصاص خواهد داد.	
رشد دارایی‌ها: سومین معیار براساس پژوهش اسپراند و رخن (۲۰۱۲) مقدار سرمایه‌گذاری مازاد در دارایی‌ها از باقیمانده رگرسیون رشد مجموع دارایی‌ها بر رشد فروش، بدست می‌آید. بدین معنا که بیش سرمایه‌گذاری مازاد برابر یک خواهد بود اگر مازاد حاصل از رگرسیون سرمایه‌گذاری مازاد بیشتر از صفر باشد و در غیر این صورت عدد صفر خواهد گرفت.	بیش اعتمادی مدیران (حق شناس، ۱۳۹۷)
برای اندازه‌گیری این معیار از شاخص گرایش‌های احساسی بازار سرمایه استفاده می‌شود. این شاخص توسط جونز (۲۰۰۵) و با تعدیل مدل ارائه شده توسط پرساود (۱۹۹۶) بسط داده شد.	گرایش احساسی سرمایه‌گذار (جونز، ۲۰۰۵)
نرخ تورم: منبع این معیار، آمار اطلاعات بانک مرکزی است.	
نرخ رشد تولید ناخالص داخلی: اطلاعات لازم برای محاسبه این معیار از سایت بانک مرکزی و مرکز آمار ایران قابل دسترسی است.	متغیرهای کلان اقتصادی (صالحی، ۱۳۹۵)
تغییرات شاخص کل بورس و شاخص کل فرابورس: منبع دسترسی به این معیار، سایت بورس اوراق بهادار تهران و فرابورس ایران است.	

تحلیل یافته‌ها

آمار توصیفی متغیرهای پژوهش پژوهش در جدول ۳ آورده شده است.

جدول ۳: آمار توصیفی متغیرهای پژوهش

متغیر پنهان	متغیر مشاهده‌پذیر	میانگین	انحراف معیار	حداقل	حداکثر
قیمت‌گذاری کمتر از واقع	بازده غیرعادی تعدیل شده ماهانه	۰/۲۳۱	۰/۰۰۴	۰/۲۲۳	۰/۲۴۳
حاکمیت شرکتی	اندازه هیات مدیره	۱/۶۱۸	۰/۶۳۹	۱/۳۸۶	۱/۹۴۵
	استقلال هیات مدیره	۰/۵۵۳	۰/۲۳۰	۰/۲۰	۱
	تمرکز مالکیت	۰/۸۱۵	۰/۱۱۱	۰/۵۴۱	۰/۹۶۲
	درصد سهامداران نهادی	۰/۸۴۷	۰/۹۷۶	۰/۵۹۱	۰/۹۶۶
	حاشیه سود خالص	۰/۵۲۷	۰/۵۸۱	۰/۲۸۲	۲/۴۸۹
	بازده دارایی‌ها	۰/۱۷۱	۰/۱۰۹	-۰/۸۱۴	۰/۴۹۹
	بازده سرمایه	۰/۴۰۱	۰/۱۹۷	۰/۵۳۲	۰/۸۰۶
	دوره گردش موجودی	۲/۶۸	۲/۰۲۲	۰/۰۱	۷/۸۴

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

۲/۰۸	۰/۲۴۱	۰/۴۶۵	۰/۹۱۸	دوره گردش دارایی‌ها	تداوم فعالیت
۱۵۲۹	۲۶	۳۷۱/۳۲	۶۴۴/۶۹	سود هر سهم	
۰/۲۵۷	۰,۱۵۲	۰,۰۶۴	۰,۱۱۰	نسبت قیمت به سود	
۱/۸۵۱	۰/۰۰۵	۰/۱۹۱	۰/۱۳۱	کیفیت اقلام تعهدی	عدم تقارن اطلاعاتی
۰/۳۷۱	۰/۰۲۸	۰/۰۷۸	۱/۳۶۸	نوسانات قیمت سهام	
۰/۹۷۶	۰/۰۰۶	۰/۲۱۶	۰/۲۱۱	خطای پیش‌بینی سود	
۱۸/۴۳۱	۹/۷۷۷	۱/۹۲۳	۱۴/۰۰۹	اندازه شرکت	
۱/۲۶۲	۰,۳۹۹	۰/۲۱۷	۰,۳۷۱	فرصت های رشد	محافظه‌کاری
۰/۳۵۳	-۰/۷۸۶	۰/۲۷۴	-۱/۷۷	معیار گیولی-هاین	
۰/۹۶۶	۰/۵۹۱	۰/۹۷۶	۰/۸۴۷	درصد سهامداران نهادی	کیفیت حسابرسی
۲/۴۷۸	۰/۵۲۲	۰,۴۵۰	۱/۲۵۹	نسبت جاری	نسبت‌های مالی
۲/۱۵	۰,۲۴۰	۰/۴۷۹	۰/۹۰۲	نسبت آنی	
۰/۷۱۱	۰/۰۱	۰/۱۷۵	۰/۱۵۵	نسبت نقدینگی	
۱	۰/۱۷۲	۱/۹۲۶	۰/۵۵۷	نسبت بدهی	
۳/۲۹۸	۰/۲۲۴	۰/۸۸۵	۱/۴۴۶	نسبت بدهی به حقوق صاحبان سهام	
۳/۶۸۸	۰/۱۶۲	۱/۰۵۵	۱/۰۹۴	نسبت مالکانه	
۷/۸۷۵	-۴/۷۵۶	۱/۳۲۹	۰/۲۸۶	نسبت فعالیت‌های عملیاتی	عدم اطمینان محیطی
۳/۶۷	-۱/۵۴۱	۱/۲۶۴	۰/۳۱۳	نسبت فعالیت‌های سرمایه‌گذاری	
۱/۵۹	-۳/۷۹۲	۱/۰۹۴	-۰/۴۶۹	نسبت فعالیت‌های تامین مالی	
۱/۴۹۹	۰/۲۸۳	۰/۳۲۱	۰/۷۶۱	ضریب تغییرات فروش	
۲/۳۳	۰/۳۴۶	۰/۴۱۵	۰/۹۶۹	ضریب تغییرات سود قبل مالیات	گرایش احساسی سرمایه‌گذار
۰/۶۸	-۰/۲۳۹	۰/۲۶۷	۰/۲۴	شاخص گرایش احساسی بازار سرمایه	
۵۰/۵۴	-۲۹/۴	۱۶/۲۴۳	۳/۳۲۴	نرخ تورم	متغیرهای کلان اقتصادی
۶/۵۵۱	-۵/۸۵۹	۲/۷۸۳	-۰/۱۶۸	نرخ رشد تولید ناخالص داخلی	
۴/۳۵۷	-۵/۵۳۸	۲/۴۳۴	۰/۲۵۷	تغییرات شاخص کل بورس	
	یک		صفر		متغیرهای دو وجهی
تعداد	درصد	تعداد	درصد		
۳۴	۲۴/۶	۱۰۴	۷۵/۴	دوگانگی	حاکمیت شرکتی
۴۵	۳۲/۶	۹۳	۶۷/۴	اندازه موسسه حسابرسی	کیفیت حسابرسی
۱۱۳	۸۱/۹	۲۵	۱۸/۱	دوره تصدی حسابرس	
۴۳	۳۱/۲	۹۵	۶۸/۸	نوع اظهار نظر حسابرس	

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

۵۹	۴۲/۸	۷۹	۵۷/۲	تخصص حسابر در صنعت	بیش اعتمادی مدیران
۷۰	۵۰/۷	۶۸	۴۹/۳	مقایسه سود و سود واقعی	
۷۳	۵۲/۹	۶۵	۴۷/۱	نسبت مخارج سرمایه‌ای به جمع دارایی‌ها	
۶۴	۴۶/۴	۷۴	۵۳/۶	رشد دارایی‌ها	

برای بررسی دقت پیش‌بینی مدل‌های طبقه‌بندی، داده‌های شرکت‌های نمونه آماری به‌عنوان ورودی وارد نرم‌افزار و نتایج پیش‌بینی برای مدل‌ها محاسبه شد. برای ارزیابی همه‌جانبه توانمندی مدل‌ها معیار دقت کل^{۲۳} و سطح زیرمنحنی^{۲۴} محاسبه شد که این امر برای هر یک از مدل‌ها به‌طور جداگانه محاسبه و در جدول ۴ آورده شده است. معیار دقت کل، به‌صورت درصدی از اطلاعاتی که به‌طور صحیح توسط مدل پیش‌بینی می‌شوند، تعریف شده است. انتظار می‌رود هر چه سطح زیرمنحنی به عدد یک نزدیک‌تر باشد، دقت پیش‌بینی مدل بیشتر شود [۸].

جدول ۴: نتایج دقت پیش‌بینی برای مدل‌های طبقه‌بندی

سطح زیرمنحنی	دقت کل (درصد)	مدل طبقه بندی
۰/۹۹۹	۹۷/۱۰۱	مدل ماشین بردار پشتیبان (SVM)
۰/۹۹۷	۹۵/۶۵۲	مدل شبکه بیزین (Bayesian Network)
۰/۹۷۳	۹۳/۴۷۸	مدل درخت تصمیم سی‌فایو (C5)
۰/۸۸۷	۸۲/۶۰۹	مدل درخت تصمیم چاید (CHAID)
۰/۸۱۳	۷۵/۳۶۲	مدل رگرسیون لجستیک (Logistic Regression)
۰/۸۱۲	۷۳/۱۸۸	مدل تحلیل ممیزی (Discriminant)
۰/۵۹۴	۵۹/۴۲	مدل درخت تصمیم کواست (Quest)
۰/۵۹۵	۵۶/۵۲۲	مدل شبکه عصبی (Neural Net)
۰/۵۰	۵۰	مدل درخت رگرسیون و طبقه‌بندی (CART)

با توجه به نتایج خروجی مدل‌ها، متغیرهای مهم تاثیرگذار شامل رشد دارایی‌ها، دوره تصدی حسابر، تخصص حسابر در صنعت، نسبت فعالیت‌های تامین‌مالی، نسبت قیمت به سود هر سهم، بازده دارایی‌ها، نسبت فعالیت‌های عملیاتی، اندازه موسسه حسابر، فرصت‌های رشد و نوسانات قیمت سهام هستند.

خروجی این مدل‌ها نشان داد که الگوریتم‌های ماشین بردار پشتیبان، شبکه بیزین و سی‌فایو نسبت به سایر مدل‌های پژوهش، دارای دقت بالاتری در پیش‌بینی قیمت‌گذاری کمتر از واقع هستند. بنابراین

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

این سه مدل طبقه‌بندی در تجزیه و تحلیل نهایی مورد استفاده قرار گرفت. مقایسه این سه مدل در جدول ۵ ارائه شده است.

جدول ۵: مقایسه مدل‌های پیش‌بینی برتر

مدل طبقه بندی	دقت کل (درصد)
مدل ماشین بردار پشتیبان	۹۷,۱
مدل شبکه بیزین	۹۵/۶
مدل درخت تصمیم سی-فایو	۹۳/۰۴

برای دستیابی به نتیجه بهتر، از روش انتخاب ویژگی^{۲۵} استفاده شد. انتخاب ویژگی فرآیند شناسایی ویژگی‌های مرتبط و حذف ویژگی‌های غیرمرتبط و تکراری با هدف مشاهده زیرمجموعه‌ای از ویژگی‌ها که مساله را به خوبی و با حداقل کاهش درجه کارایی تشریح می‌کنند، تعریف شده است. نتایج انتخاب ویژگی نشان داد که متغیرهای مهم بااهمیت متغیرهای مهم تاثیرگذار شامل دوره تصدی حسابرس، رشد دارایی‌ها و نسبت فعالیت‌های تامین مالی هستند.

نتایج خروجی مدل با استفاده از روش انتخاب ویژگی در جدول ۶ ارائه شده است.

جدول ۶: نتایج دقت پیش‌بینی برای مدل‌های طبقه‌بندی برتر با روش انتخاب ویژگی

مدل طبقه بندی	دقت کل (درصد)	سطح زیر منحنی
مدل ماشین بردار پشتیبان	۶۸,۱۱۶	۰/۷۱۶
مدل شبکه بیزین	۶۷/۳۹۱	۰/۷۵۹
مدل درخت تصمیم سی-فایو	۶۸/۸۴۱	۰/۷۱۵

بحث و نتیجه‌گیری

هدف این پژوهش، دستیابی به مدلی است که بتواند پدیده قیمت‌گذاری کمتر از واقع سهام عرضه عمومی اولیه را با دقت بالایی پیش‌بینی کند. بازده کوتاه‌مدت سهام جدید با استفاده از تغییرات ماهانه قیمت سهام شرکت‌های دارای عرضه عمومی اولیه در بورس اوراق بهادار تهران و فرابورس ایران در بازه زمانی ۱۳۹۶-۱۳۸۲ محاسبه شد و با مقایسه آن با بازده کوتاه‌مدت بازار طی دوره مشابه پژوهش، بازده غیرعادی محاسبه و مبنای تعیین قیمت‌گذاری کمتر از واقع شد. در بیشتر پژوهش‌های قبلی، قیمت‌گذاری کمتر از واقع با رویکرد آمار کلاسیک همانند رگرسیون مورد بررسی قرار گرفت. با وجود این که مدل‌های آماری توانسته‌اند پیش‌بینی‌های خوبی را در مورد عملکرد سهام تازه عرضه شده ارائه

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

دهند ولی برای مدل‌سازی روابط بین متغیرهایی که مشاهدات آن‌ها دقیق نیستند، نمی‌توان مدل‌های معمولی آماری را که براساس مشاهدات دقیق و برخی مفروضات توزیعی استوار هستند، به‌کار برد. مدل‌های یادگیری ماشین نسبت به مدل‌های آماری به‌دلایلی همچون در نظر نگرفتن توزیع جوامع، نداشتن مشکل هم‌خطی بین متغیرهای توضیحی و یادگیری تطبیقی، دقت بیشتری دارند. نتایج نشان داد که مدل‌های مورد استفاده در پژوهش، توان پیش‌بینی عملکرد کوتاه‌مدت عرضه عمومی اولیه را با دقت بالایی دارا هستند. مقایسه دقت کل پیش‌بینی برای الگوریتم‌های یادگیری ماشین نشان داد که مدل‌های ماشین بردار پشتیبان، شبکه بیزین و درخت تصمیم سی‌فایو قوی‌تر از سایر مدل‌های موجود در این پژوهش عمل کرده‌اند و پیش‌بینی را با دقت بالاتری انجام داده‌اند. نتایج مقایسه مدل‌های به‌کار رفته نشان می‌دهد که بعد از مدل ماشین بردار پشتیبان، مدل شبکه بیزین و مدل درخت تصمیم سی‌فایو، مدل‌های مناسبی برای پیش‌بینی قیمت‌گذاری کمتر از واقع سهام عرضه عمومی اولیه هستند. نتایج به‌دست آمده از پژوهش تا حدودی سازگار با نتایج پژوهش‌های بابا و سویل (۲۰۱۹)، باستی و همکاران (۲۰۱۵) و اسماعیلی و همکاران (۱۳۹۷) است. در پژوهش باستی و همکاران، به‌ترتیب مدل‌های درخت سی‌فایو، درخت رگرسیون و طبقه‌بندی و ماشین بردار پشتیبان دارای بالاترین دقت پیش‌بینی بودند. درحالی‌که در این پژوهش مدل درخت رگرسیون و طبقه‌بندی دارای حداقل دقت لازم بود. در بیان متفاوت بودن این نتیجه می‌توان گفت که در پژوهش باستی از متغیرهایی استفاده شد که در درخت تصمیم بیشتر از سایر متغیرها قابلیت طبقه‌بندی دارد.

با توجه به نتایج به‌دست آمده، به فعالان بازار سرمایه پیشنهاد می‌شود که از این الگوریتم‌ها در جهت تصمیم‌گیری بهتر استفاده شود. زیرا انتخاب این عوامل مهم منجر به انتخاب سبد سرمایه‌گذاری با کمینه مخاطره و بیشترین بازدهی می‌گردد، ضمن آن که شفافیت محیط تصمیم‌گیری و نتایج حاصله را نیز دو چندان می‌نماید. پیشنهاد می‌شود در پژوهش‌های آتی از سایر ویژگی‌های تاثیرگذار بر عملکرد عرضه عمومی اولیه مانند اثر صنعت استفاده شود. همچنین مقایسه تطبیقی با سایر الگوریتم‌ها مانند الگوریتم ژنتیک و غیره، به‌منظور دست‌یابی به سطح خطای کمتر در برآورد قیمت سهام در زمان عرضه عمومی اولیه صورت گیرد.

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره چهل و چهارم / پائیز ۱۳۹۹

منابع

- ۱) اسماعیلی زهرا، عباسی ابراهیم، فلاح میرفیض (۱۳۹۷). پیش‌بینی عملکرد کوتاه‌مدت عرضه عمومی اولیه سهام با استفاده از مدل‌های نزدیک‌ترین همسایگی و ماشین بردار پشتیبان، چشم‌انداز مدیریت مالی، شماره ۲۱، ۲۷-۹.
- ۲) حق‌شناس نوید (۱۳۹۷). ارتباط بین اطمینان بیش از حد مدیریتی و کمتر قیمت‌گذاری سهام در عرضه‌های عمومی اولیه، پایان‌نامه کارشناسی ارشد حسابداری، موسسه آموزش عالی مولوی
- ۳) حیدری مهدی، قادری بهمن، هم‌خانی سعید (۱۳۹۶). بررسی نقش میانجی عدم تقارن اطلاعاتی در تبیین رابطه بین کیفیت حسابرسی و سیاست‌های تامین مالی. *فصلنامه راهبرد مدیریت مالی*، سال پنجم، شماره ۱۹، ۱۲۶-۹۳.
- ۴) رهنمای رودپشتی فریدون، سیم‌بر فرشید، طوطیان صدیقه (۱۳۸۴). تاثیر متغیرهای کلان اقتصادی بر بازده سهام شرکت‌های سرمایه‌گذاری پذیرفته شده در بورس اوراق بهادار تهران. *پژوهشنامه اقتصادی*، دوره پنجم، شماره ۱۷، ۲، ۲۳۶-۲۰۹.
- ۵) سیاه‌تیری محسن (۱۳۸۹). بررسی تاثیر نسبت‌های مالی موثر بر بازدهی غیرعادی کوتاه‌مدت عرضه‌های اولیه سهام در بورس اوراق بهادار تهران، پایان‌نامه کارشناسی ارشد مدیریت بازرگانی، دانشگاه سیستان و بلوچستان.
- ۶) شهاب‌نویی شاهد (۱۳۹۶). بررسی تاثیر عدم اطمینان محیطی بر قیمت‌گذاری کمتر از واقع بر عرضه‌های عمومی اولیه سهام در شرکت‌های بورس اوراق بهادار ایران، پایان‌نامه کارشناسی ارشد حسابداری، موسسه آموزش عالی آمل.
- ۷) صالحی حمید (۱۳۹۵). تاثیر عوامل موثر بر ارزش‌گذاری سهام در عرضه‌های عمومی اولیه در بورس اوراق بهادار تهران با رویکرد معادلات ساختاری، پایان‌نامه کارشناسی ارشد مدیریت مالی، دانشگاه آزاد اسلامی شهرقدس.
- ۸) صنیعی آباذه محمد، محمودی سینا، طاهرپرور محدثه (۱۳۹۴). داده‌کاوی کاربردی (ویراست دوم)، انتشارات نیاز دانش.
- ۹) فلاح شمس میرفیض، کردلویی حمیدرضا، رشنو مهدی (۱۳۹۱). بررسی دست‌کاری قیمت‌ها در بورس اوراق بهادار تهران با استفاده از مدل ماشین بردار پشتیبان، تحقیقات مالی، دوره چهاردهم، شماره ۱، ۸۴-۶۹.

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

۱۰) مفاخری سعید (۱۳۹۷). مقایسه تکنیک‌های درخت تصمیم و شبکه‌های عصبی در طبقه‌بندی و ارزش‌گذاری مشتریان بانکی و انتخاب عوامل بهینه با استفاده از داده‌کاوی، هشتمین همایش سالانه بانکداری الکترونیکی و نظام‌های پرداخت.

۱۱) نجفی‌مقدم علی (۱۳۹۷). بررسی عوامل ساختاری بر بازدهی غیرمتعارف عرضه‌های اولیه در بورس اوراق بهادار تهران. فصلنامه علمی پژوهشی دانش سرمایه‌گذاری، سال هفتم، شماره ۲۵، ۱۵-۱.

12) Aggarwal, R., Leal, R., and Hernandez, L. (1993). "The aftermarket performance of initial public offerings in Latin America". *Financial Management*, 22(1), PP: 42-53.

13) Baba, Boubekeur. Sevil, Guven. (2019). "Predicting IPO initial returns using random forest". Journal Pre-proof, <https://doi.org/10.1016/j.bir.2019.08.001>.

14) Basti, Eyub. Kuzey, Cemil. Dursun, Delen. (2015). "Analyzing initial public offering's short-term performance using decision trees and SVMs". *Decision Support System*, 73, pp 15-27.

15) Chang, X., Gyax, A. F., Oon, E., & Zhang, H. F. (2008). "Audit Quality, Auditor Compensation and Initial Public Offering Underpricing". *Accounting and Finance*, 48(3), 391-416.

16) Clarkson, G., Jacobsen, T. E., & Batcheller, A. L. (2007). "Information asymmetry & information sharing". *Government Information Quarterly*. 24(4), PP. 827-839.

17) Colak, Gonul, Fu, Mengechun. (2018). "Predicting IPO failure using machine learning technique".

18) Esfahanipour, Akbar, Goodarzi, Milad, Jahanbin, Reza. (2015). "Analysis and forecasting of IPO underpricing". *The Neural Computing Applications Forum*, DOI: 10.1007/s00521-015-1884-1.

19) Han, Ting, (2016). "The application if data mining in the analysis of high underpricing rate of IPO".

20) Haung, Wei, Jinxian Li, Qiang Zhang, (2019). "Information asymmetry, legal environment, and family firm governance: Evidence from IPO underpricing in China", *Pacific-Basin Finance Journal*, Accepted Manuscript.

21) Ibbotson, R. G. (1975). "Price performance of common stock new issues". *Journal of Financial Economics*, 2.

22) Ibbotson, R.G., & Jaffe, J. (1975). "Hot issue markets". *Journal of Finance*, 4, PP: 1027-1042.

- 23) Jones, A. L. (2005). "Measuring investor sentiment in equity markets", *Journal of Asset Management*, 7 (3/4), 208-215.
- 24) Lin, Z. J., and Tian, Z. (2012). "Accounting conservatism and IPO underpricing: China evidence". *Journal of International Accounting, Auditing and Taxation*. 21(2), PP: 127-144.
- 25) Ljungvist, Alexander, (2004). "Handbooks in Finance: Empirical Corporate Finance", chapter III.4: IPO Underpricing.
- 26) Nawadali Rathnayake, Dilesha, Louembo, Pierre Axel., Francois Kassi, Diby, Sun, Gang, Ning, Ding. (2019). "Are IPOs underpriced or overpriced? Evidence from an emerging market". Accepted Manuscript. *Research in International Business and finance*.
- 27) Peng, X., Xiongyuan, W., Kam, C. (2018). "Does customer concentration disclouse affect IPO pricing?" *Finance Research Letters*, Article in press, Elsevier, 1-7.
- 28) Rock, K., 1986, "Why new issues are underpriced". *Journal of Financial Economics*, 15, PP: 187-212.
- 29) Su, D., & Fleisher, B.M. (1999). "An empirical investigation of underpricing in Chinese IPOs". *Pacific-Basin Finance Journal*, 7(2), 173-202.
- 30) Yu, Ji, limg Tue, Da Wu. (2019). "Industry peer firm's Earnings Quality and IPO Underpricing", *Journal of corporate accounting & finance*, published online in Wiley Online Library (wileyonlinelibrary.com).
- 31) Wang, Di, Qian, Xiaolin, Quek, Chai, Tan, Ah-Hwee., Miao, Chunyan, Zhang, Xiaofeng, See NG, Geok, Zhou, YOU. (2018). "An interpretable neural fuzzy inference systems for predictions of underpricing in IPOs". Accepted Manuscript, <https://doi.org/10.1016/j.neucom.2018.07.036>
- 32) Zhou, Lu (Jolliy) & Sadeghi, Mehdi. (2018). "The impact of innovation on IPO short-term performance: Evidence from the Chinese markets". Accepted Manuscript. *Pacific Basin Finance Journal*.

بررسی مقایسه‌ای دقت پیش‌بینی.../دهقان خانقاهی، بحری ثالث، جبارزاده کنگرلویی و آشاب

یادداشت‌ها

1. Initial Public Offering (IPO)
2. Underpricing
3. Machine Learning
4. Decision Trees
5. CHAID
6. C5
7. Classification and Regression Tree
8. Quest
9. Bayesian Network
10. Support Vector Machine (SVM)
11. Neural Net
12. Logistic Regression
13. Discriminant
14. Ibbotson
15. Ibbotson and Jaffe
16. Information Asymmetry
17. Winner Curse Hypothesis
18. Rock
19. Colak and Fu
20. Wang et al
21. Basti et al
22. Boosting
23. Overall Accuracy
24. Area Under Curve (AUC)
25. Feature Selection