



مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست و طبقه‌بندی احتمالی بیزین در پیش‌بینی بیش‌اطمینانی مدیران شرکت‌های بازار سرمایه ایران

شکوفه اعتبار^۱

رویا دارابی^۲

محسن حمیدیان^۳

سیده محبوبه جعفری^۴

تاریخ دریافت مقاله: ۹۷/۰۳/۰۳ تاریخ پذیرش مقاله: ۹۷/۰۶/۱۴

چکیده

بیش‌اطمینانی که یکی از ویژگی‌های شخصیتی افراد می‌باشد، که ممکن است بر تصمیم‌گیری‌های مدیران شرکت‌ها تأثیر داشته باشد، مدیریت با رفتار غیرمنطقی می‌تواند بر عملکرد شرکت در بلندمدت تأثیرگذار باشد. هدف اصلی از پژوهش کاربردی حاضر «مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست و طبقه‌بندی احتمالی بیزین در پیش‌بینی بیش‌اطمینانی مدیران شرکت‌های بازار سرمایه ایران طی سال‌های ۱۳۸۷ تا ۱۳۹۵» می‌باشد که در این راستا الگوریتم هوش مصنوعی آدابوست و احتمالی بیزین به‌منظور ارائه مدل جهت پیش‌بینی بیش‌اطمینانی مدیریت در شرکت‌های پذیرفته شده بورس اوراق بهادار تهران، برای سال جاری و یک سال آتی اعتبارسنجی شدند. در مجموع، تعداد ۳۴ صنعت مختلف به‌عنوان نمونه آماری انتخاب شد (اطلاعات مالی تعداد ۳۱۴۵ شرکت-سال در مجموع). روش گردآوری اطلاعات در بخش مبانی نظری از روش کتابخانه‌ای و برای انجام محاسبات و طبقه‌بندی آماری داده‌های مالی، از نرم‌افزارهای Excel و برای آزمون فرضیه‌های پژوهش از نرم‌افزار matlab ۲۰۱۷ استفاده شده است. یافته‌های پژوهش نشان می‌دهد که، مدل‌های غیرخطی پیش‌بینی الگوریتم آدابوست و الگوریتم احتمالی بیزین به دست آمده، هر دو توانایی پیش‌بینی بیش‌اطمینانی مدیریت برای سال جاری و یک سال بعد را دارند؛ اما مدل پیش‌بینی آدابوست در مقایسه با مدل پیش‌بینی بیزین نتایج بهتری را برای پیش‌بینی اعتماد به نفس کاذب مدیریت دارد که نشان‌دهنده قدرت بیشتر در یادگیری و کارایی بهتر این مدل به‌منظور پیش‌بینی بیش‌اطمینانی مدیریت می‌باشد.

واژه‌های کلیدی

بیش‌اطمینانی مدیریت، الگوریتم یادگیری ماشین آدابوست، الگوریتم طبقه‌بندی احتمالی بیزین

۱ دانشجوی دکتری تخصصی حسابداری، دانشکده اقتصاد و حسابداری، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران، shokoufeh.etebar@gmail.com

۲ عضو هیات علمی، دانشکده اقتصاد و حسابداری، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران (نویسنده مسئول)، royadarabi110@yahoo.com

۳ عضو هیات علمی، دانشکده اقتصاد و حسابداری، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران hamidian_2002@yahoo.com

۴ عضو هیات علمی، دانشکده اقتصاد و حسابداری، دانشگاه آزاد اسلامی، واحد تهران جنوب، تهران، ایران majafari22@gmail.com

۱- مقدمه

پیشرفت در فن‌آوری اطلاعات و ارتباطات و ایجاد پایگاه‌های داده الکترونیکی در سال‌های اخیر، تغییرات مهمی در روابط اقتصادی شرکت‌ها ایجاد کرده است، مدیران و خصوصیات رفتاری آن‌ها تأثیر بسزایی در تداوم فعالیت و موفقیت شرکت‌ها در محیط رقابتی امروز فراهم می‌آورد (آروکینامول، لوکس^۱، ۲۰۱۷، ۴). با توجه به اهمیت موضوع تصمیمات مدیریت و تأثیر این تصمیم‌گیری‌ها بر عملکرد آتی شرکت‌ها در بازار سرمایه می‌توان به ضروری بودن موضوع توانایی مدیران در استفاده اثربخش از منابع کمیاب شرکت‌ها در جهت درک بهتر از محصول، فناوری، کارایی کارکنان، رویه‌های صنعت، وضعیت کلان اقتصادی و سرمایه‌گذاری در پروژه‌های باارزش‌تر شرکت پی برد که منجر به اثربخشی عملیات روزانه سازمان‌ها و بهبود ارزش فعلی خالص سازمان می‌شود (برک، استنتون^۲، ۲۰۰۷، ۵۳۲). اهمیت موضوع به دلیل اینکه اعتماد به نفس کاذب مدیر می‌تواند بر تصمیمات تأمین مالی شرکت اثرگذار باشد و بنا بر تحقیقات انجام شده، تأثیر بیش‌اطمینانی مدیریت در بلندمدت بر ریسک گزارشگری مالی شرکت‌ها نشان داده شده است (الشاندایدی و همکاران^۳، ۲۰۱۸، ۱). سیستم اطلاعات حسابداری به‌عنوان یک سیستم فرعی سیستم اطلاعات مدیریت، جزو ابزارهای مهم تهیه‌کننده اطلاعات مناسب جهت اتخاذ تصمیم می‌باشد. کار این سیستم از یک طرف، به کیفیت اطلاعات ارائه‌شده توسط آن و از طرف دیگر، به بهره‌برداری مناسب مدیریت از آن بستگی دارد (دمریجان و همکاران^۴، ۲۰۱۲، ۱۳۰۱) و از آنجایی که مدیران شرکت‌ها توجه بیشتری به اثرهای بلندمدت معاملات دارند، احتمالاً آن نوع معاملاتی را بر می‌گزینند که اثر بهتری بر عملکرد آتی شرکت داشته باشند (عظیمی، مشایخی، ۱۳۹۵، ۲۷۱). همچنین (هیتون^۴، ۲۰۱۶، ۳۳) در پژوهش خود، تلاش کرد تا مدلی برای بررسی فرآیند تصمیم‌گیری مدیران با خوش‌بینی بالا، بدون در نظر گرفتن هزینه‌های نمایندگی و عدم تقارن اطلاعاتی ارائه نماید. مدیران خوش‌بین، اغلب جریان‌های نقدی را بیشتر از واقع پیش‌بینی می‌کنند. در نتیجه فرصت‌های سرمایه‌گذاری شرکت نیز بیش از واقع ارزش‌گذاری خواهد شد، پژوهش‌های انجام‌یافته خارجی در دو سال گذشته بر کاربرد ترکیبی روش‌های یادگیری ماشین و شبکه‌های عصبی مبتنی بر الگوریتم و بررسی‌های الگوی رفتاری انسانی (عملکرد اعتماد به نفس بیش از حد مدیریت) برای مدل‌سازی آن با استفاده از متغیرهای حسابداری رونق یافته است (ماروکی‌ولمن، کرنس، لهتو^۵، ۲۰۱۷، ۱). به‌منظور انجام پیش‌بینی‌های کوتاه‌مدت خود از شبکه‌های عصبی استفاده نمودند (آبناو و همکاران^۶، ۲۰۱۷، ۲)، همچنین (بایگ و همکاران^۷، ۲۰۱۷، ۳) در پژوهش

مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست.../ اعتبار، دارایی، حمیدیان و جعفری

خود از الگوریتم آدابوست برای یادگیری شبکه عصبی مصنوعی استفاده کردند. ضرورت این پژوهش در این است که با توجه به اینکه گروه‌های مختلف ذینفعان و سایر استفاده‌کنندگان از اطلاعات مالی شرکت، خواستار اطمینان یافتن و وجود مقرراتی برای محافظت در برابر سطوح با اهمیت عدم تقارن اطلاعاتی گزارشگری و اعتماد بیش از حد مدیریت به خود هستند که منجر به دست‌کاری ارقام حسابداری و در نهایت به ضرر ذینفعان شرکت، درآینده خواهد بود (بامبر، مک‌کینگ^۸، ۲۰۱۶، ۶۰). از این رو پرسش‌های اصلی تحقیق آن است که اول: آیا الگوریتم یادگیری ماشین آدابوست و الگوریتم طبقه‌بندی احتمالی بیزین در پیش‌بینی بیش‌اطمینانی مدیریت سال جاری و آتی توانایی دارد؟ و دوم: آیا توانایی پیش‌بینی بیش‌اطمینانی مدیریت توسط الگوریتم آدابوست بیشتر از الگوریتم طبقه‌بندی احتمالی بیزین می‌باشد؟ با در نظر داشتن اینکه تاکنون هیچ پژوهشی در داخل کشور به ارائه مدل پیش‌بینی بیش‌اطمینانی مدیریت با بکارگیری الگوریتم‌های هوش مصنوعی انجام نشده است این پژوهش از دو دیدگاه دارای نوآوری می‌باشد. دیدگاه اول: ارائه مدل پیش‌بینی بیش‌اطمینانی مدیریت با روش‌های آدابوست و بیزین برای اولین بار و دیدگاه دوم: استفاده از ۲۲ متغیر مستقل که در روش‌های محاسبه بیش‌اطمینانی مدیریت در مبانی نظری خارج و داخل کشور کاربرد داشته است و اعتبار سنجی آنها جهت ورود به مدل‌های پژوهش توسط الگوریتم متغیرگزینی لاس و گاس (LVF) برای اولین بار انجام گرفته‌است.

بنابراین در این پژوهش تلاش می‌شود در شرکت‌های فعال در بازار سرمایه ایران، برای سال جاری و یک سال مالی بعد، از الگوریتم آدابوست که یکی از روش‌های نوین یادگیری ماشین می‌باشد، برای پیش‌بینی بیش‌اطمینانی مدیریت اعتبارسنجی شود و هدف دوم این است که متغیر بیش‌اطمینانی مدیریت توسط الگوریتم طبقه‌بندی احتمالی بیزین نیز مدل‌سازی و اعتبار سنجی شود تا در نهایت هدف سوم اینکه، قدرت پیش‌بینی اعتماد به نفس کاذب مدیریت توسط این دو روش با یکدیگر مقایسه گردد.

مبانی نظری و مروری بر پیشینه پژوهش

بیش‌اطمینانی یا اعتماد بیش از حد به خود را می‌توان اعتقادی بی‌اساس درباره توانایی‌های شناختی، قضاوت‌ها و استدلال شهودی فرد تعریف کرد. مفهوم بیش‌اطمینانی در مجموعه وسیعی از بررسی‌ها و آزمایش‌های روان‌شناسی از نوع شناختی بررسی شده است که نشان می‌دهد، افراد هم درباره توانایی‌های خود در پیش‌بینی و هم درباره دقت اطلاعات خود، برآوردی بیش از اندازه دارند. (پاکوان^۹، ۲۰۱۳، ۳۷۲).

بیش اطمینانی مدیران

در مبانی نظری تئوری بیش اعتمادی مدیریت نشان داده شده است که این متغیر تأثیر قابل توجهی بر روی تصمیم‌گیری شرکت‌های بزرگ دارد. مدیران بیش اطمینان، احتمال و تأثیر رویدادهای مطلوب را بر جریان‌های نقدی شرکت بیشتر از واقعیت تخمین می‌زنند و آن را برای رویدادهای منفی کمتر از واقعیت ارزیابی می‌کنند، بنابراین در هنگام تصمیم‌گیری در پرداخت بدهی‌های شرکت بدهی کوتاه‌مدت را انتخاب می‌کنند و بنابراین ریسک گزارشگری مالی در شرکت‌هایی که مدیران آن‌ها دارای خوش‌بینی هستند، شدیدتر است (هیتون، ۲۰۱۶، ۳۵). مدیران بیش اطمینان، بازده آتی پروژه‌های سرمایه‌گذاری شرکت را بیشتر برآورد می‌کنند؛ بنابراین ممکن است شناسایی زیان را به تأخیر بیندازند و برآوردهای خوش‌بینانه‌ای در تعیین ارزش دارایی‌های جاری یا بلندمدت داشته باشند (رامشه، ملانظری، ۱۳۹۳، ۵۹).

یشینه پژوهش‌های داخلی و خارجی مرتبط با الگوریتم‌های بیزین و هوش مصنوعی

تاکنون پژوهش‌های داخلی و خارجی در مورد مدل‌سازی در موضوعات مختلف با استفاده از شبکه‌های عصبی بیزین و شبکه‌های عصبی مصنوعی انجام گرفته است، از جمله پژوهش‌های جدید در دست چاپ قابل ذکر را می‌توان، پژوهش جونیور و همکاران^{۱۰} (۲۰۱۹) نام برد که آنان در تحقیق خود با استفاده از یک طبقه‌بندی جدید در الگوریتم بوستینگ توانستند، مدل رقابتی نوینی را پیشنهاد دهند. همچنین (جان و همکاران^{۱۱}، ۲۰۱۸) با استفاده از الگوریتم مدل پیش‌بینی سری زمانی برای صنعت توریسم، نشان دادند: استفاده از مدل‌های ANN و ARIMA با سه سری زمانی (روش ترکیبی)، نتایج بهتری نسبت به سایر روش‌های مرتبط با آن را دارد. در پژوهشی دیگر هورتوگ و زانتن^{۱۲} (۲۰۱۸) با ارائه یک رویکرد بیزین نا پارامتریک برای حل مسائل مربوط به حوزه علوم انسانی از طبقه‌بندی باینری استفاده کردند، این محققان توانستند از سیستم شبکه عصبی مبتنی بر الگوریتم برای پژوهش خود استفاده کنند.

امیدی و همکاران (۱۳۹۶) در پژوهش خود با عنوان «توسعه کاربرد مدل‌های باکس جنکینز، شبکه عصبی مصنوعی و تعدیل نمایی در پیش‌بینی و مدیریت پدیده‌های اجتماعی» از شبکه‌های عصبی در پیش‌بینی‌های خود استفاده نمودند.

در پژوهشی خداکرمی و همکاران (۱۳۹۳) با ارائه مدلی جهت تحلیل پارامترهای هزینه کیفیت با رویکرد شبکه‌های بیزین، پارامترهای مهم COQ را جهت اثربخشی برنامه‌های بهبود شناسایی

مقایسه توانایی الگوریتم‌های یادگیری ماشین آداپوست.../ اعتبار، دارایی، حمیدیان و جعفری

کرده است. شهیکی تاش و همکاران (۱۳۹۳) در پژوهش خود به منظور بررسی ارتباط رشد اقتصادی و ضریب رفاه اجتماعی در ایران از رهیافت بیزین استفاده نمودند. در بررسی علل تورم در اقتصاد ایران مهرآرا و غضنفری (۱۳۹۳) از یک مدل مبتنی بر به‌کارگیری روش میانگین‌گیری مدل بیزینی اثر ۱۳ متغیر اقتصادی را بر رشد تورم مورد مطالعه قرار دادند، متغیرهای درآمد نفتی و واردات نیز تنها از کانال نرخ ارز، رشد تولید و رشد حجم پول بر تورم تأثیرگذارند.

تاکنون تحقیقات داخلی و خارجی در مورد موضوع تحقیق انجام گرفته است که در ادامه تعدادی از پیشینه تجربی پژوهش داخل و خارج از کشور بیان می‌گردد.

پیشینه پژوهش‌های داخلی و خارجی بیش‌اطمینانی مدیریت

کانگ و همکارانش^{۱۳} در سال (۲۰۱۷) پژوهشی با عنوان «مهار کوتاه بینی مدیریتی: نقش مدیریتی اعتماد به نفس کاذب در شرکت‌های با مدیریت حرفه‌ای» در دوره زمانی ۲۰۰۹ الی ۲۰۱۵ انجام دادند. آن‌ها در تحقیق از یک مدل سرمایه‌گذاری چند دوره‌ای استفاده کردند و بیش‌اعتمادی به عنوان یک متغیر مستقل و نتیجه تحقیق آن‌ها بیان‌گر این بود که اعتماد به نفس بیش از حد مدیریت تأثیر منفی بر عملکرد مالی شرکت دارند. در پژوهش بهاراتی، دلمان، فو^{۱۴} (۲۰۱۶) با عنوان «بیش‌اطمینانی مدیریت و بازده‌های سهام» نتایج تحقیق نشان دادند که ارتباط معناداری بین اعتماد به نفس مدیرعامل و سیاست‌های شرکت وجود دارد. چن و همکارانش^{۱۵} (۲۰۱۴) در تحقیقی با عنوان «تأثیر اعتماد به نفس بیش از حد مدیران ارشد بر عملکرد بلندمدت شرکت» به این نتیجه دست یافتند که اعتماد به نفس بیش از حد باعث افزایش در مخارج تحقیق و توسعه خواهد شد و به دنبال آن افزایش در مخارج تحقیق و توسعه منجر به بهبود عملکرد بلندمدت شرکت‌ها می‌شود.

بن دیوید، گراهام، هاروی^{۱۶} (۲۰۱۲) نشان دادند که اعتماد به نفس بیش از حد مدیریتی عامل تعیین‌کننده‌ای برای برخی از فعالیت‌های مالی شرکت‌ها هست. این پژوهشگران گزارش می‌دهند که اعتماد به نفس بیش از حد مدیریتی سرمایه‌گذاری را افزایش می‌دهد، استفاده از بدهی را بیشتر و احتمالاً پرداخت سود تقسیم‌شده را کاهش می‌دهد و همچنین احتمال بازخرید سهام را افزایش می‌دهد و باعث می‌شود که از بدهی‌های بلندمدت در مقابل بدهی‌های کوتاه‌مدت بیشتر استفاده شود. همچنین با بررسی تأثیر بیش‌اطمینانی مدیریتی بر کیفیت گزارشگری مالی و محافظه‌کاری شرطی در بورس تهران، توسط عباسی، وکیلی فرد، معروف (۱۳۹۷) نشان دادند که اعتماد بیش از اندازه مدیریتی بر کیفیت گزارشگری مالی تأثیر معکوس و معناداری دارد، اعتماد

بیش از اندازه مدیریتی بر محافظه‌کاری شرطی تأثیر معکوس و معناداری دارد و محافظه‌کاری موجب قوی‌تر شدن تأثیر معکوس بیش‌اطمینانی مدیریتی بر کیفیت گزارشگری مالی می‌شود. نتایج پژوهش چهارم‌حالی، زاهدی (۱۳۹۶) با عنوان «بررسی تأثیر بیش‌اعتمادی مدیران بر چسبندگی هزینه‌ها» حاکی از تأثیر مستقیم معیارهای مختلف بیش‌اعتمادی مدیران بر چسبندگی هزینه‌ها طی دوره پژوهش است و همچنین در پژوهش تلخابی و همکاران (۱۳۹۶) نیز نتایج حاکی از آن است که بیش‌اطمینانی مدیران و پاداش‌های هیات مدیره، انحراف از مدل استاندارد حق‌الزحمه حسابرسی را افزایش می‌دهد و باعث افزایش شدت چسبندگی (رو به بالا) دستمزدهای حسابرسی خواهد شد. همچنین نتایج تحقیق پوردری، بنی‌مهد، احمدزاده (۱۳۹۵) نشان داد، هر چه بیش‌اعتمادی مدیران افزایش یابد، در آن صورت مدیریت سود از نوع حداکثر سازی سود افزایش می‌یابد. نتایج حاصل از پژوهش عربصالحی، امیری، کاظمی نوری (۱۳۹۳) نشان می‌دهد که، در دوره زمانی مورد بررسی اعتماد به نفس بیش از حد مدیران ارشد باعث افزایش حساسیت سرمایه‌گذاری جریان‌های نقدی شده است.

فرضیه‌های پژوهش

در این تحقیق به بررسی و مقایسه توانایی الگوریتم یادگیری ماشین آدابوست و الگوریتم طبقه‌بندی احتمالی بی‌زین در پیش‌بینی بیش‌اطمینانی مدیریت می‌پردازیم. با توجه به موضوع تحقیق و بیان مبانی نظری پژوهش، ۳ فرضیه به شرح زیر در سطح نرخ شناسایی ۹۰ درصد مورد بررسی قرار می‌گیرند:

(۱) الگوریتم (یادگیری ماشین) آدابوست در پیش‌بینی بیش‌اطمینانی مدیریت سال جاری و سال آتی توانایی دارد.

(۲) الگوریتم طبقه‌بندی احتمالی بی‌زین در پیش‌بینی بیش‌اطمینانی مدیریت سال جاری و سال آتی توانایی دارد.

(۳) توانایی پیش‌بینی بیش‌اطمینانی مدیریت توسط الگوریتم یادگیری ماشین آدابوست بیشتر از الگوریتم طبقه‌بندی احتمالی بی‌زین می‌باشد.

روش‌شناسی پژوهش

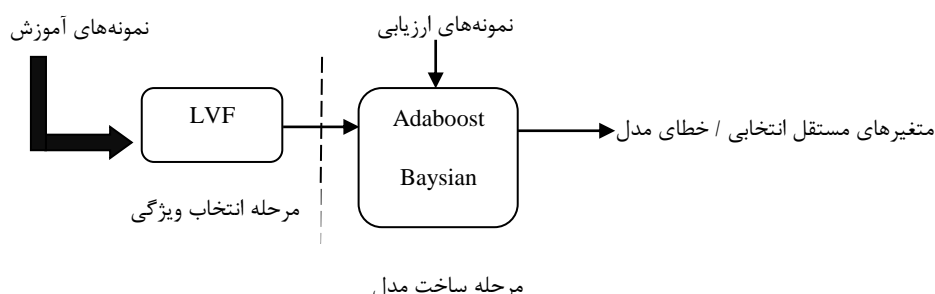
این پژوهش از نظر هدف، تحقیقی کاربردی است و از نظر ماهیت داده‌ها، تحقیق کمی و از نوع پس‌رویدادی می‌باشد که با استفاده از داده‌ها و اطلاعات مالی به روش اسنادکاوی و از طریق

مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست.../ اعتبار، دارایی، حمیدیان و جعفری

مراجعه به صورت‌های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار و بانک‌های اطلاعاتی رهاورد نوین و سراچه‌های مرتبط با سازمان بورس اوراق بهادار، داده‌های گذشته برای پیش‌بینی دوره جاری و آتی جمع‌آوری شده است، همچنین برای جمع‌آوری ادبیات نظری و متون موضوع پژوهش از روش کتابخانه‌ای استفاده شده است. به‌منظور انجام محاسبات و تجزیه و تحلیل‌های آماری، داده‌ها در نرم‌افزارهای Excel طبقه‌بندی شده است و سپس در نرم‌افزار Matlab 2017 جهت آزمون فرضیه‌های پژوهش (اعتبار سنجی قدرت پیش‌بینی بیش‌اعتمادی مدیریت بازار سرمایه ایران) مورد استفاده قرار گرفته است. جامعه‌ی آماری این تحقیق شامل تمامی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۷ الی ۱۳۹۵ هستند که دارای شرایط زیر باشند، داده‌های موردنظر آن‌ها در دسترس باشد و جز شرکت‌های سرمایه‌گذاری، بانک و لیزنگ نباشند. همچنین در دوره زمانی تحقیق، تغییر فعالیت یا تغییر دوره مالی نداده باشند که در مجموع، تعداد اساس اطلاعات مالی تعداد ۳۱۴۵ (شرکت - سال) متشکل از ۳۴ صنعت مختلف با استفاده از نرم‌افزار رهاورد نوین به‌عنوان نمونه آماری انتخاب شده است.

مدل مفهومی پژوهش

در این بخش رویکرد فیلتر پیشنهادی برای انتخاب متغیرهای مستقل به همراه ارائه مدل پیش‌بینی بیش‌اطمینانی مدیریت یا اعتماد به نفس کاذب بیان می‌گردد. این رویکرد شامل سه مرحله است، مرحله اول انتخاب یک زیرمجموعه از متغیرهای مستقل توسط الگوریتم فیلتر لاس و گاس^{۱۷} (LVF) و مرحله دوم ساخت مدل به منظور پیش‌بینی بیش‌اطمینانی مدیریت مبتنی بر الگوریتم یادگیری ماشین آدابوست^{۱۸} و مرحله سوم نیز ساخت مدل به منظور پیش‌بینی بیش‌اطمینانی مدیریت یا اعتمادبنفس کاذب مدیریت مبتنی بر الگوریتم طبقه‌بندی احتمالی بیزین می‌باشد. الگوریتم طبقه‌بندی احتمالی بیزین، دسته‌بندی‌کننده بیز ساده در یادگیری ماشین به گروهی از دسته‌بندی‌کننده‌های ساده بر پایه احتمالات گفته می‌شود که با متغیرهای تصادفی مستقل ساده مفروض میان حالت‌های مختلف و بر اساس قضیه بیز کاربردی است. به‌طور ساده روش بیز روشی برای دسته‌بندی پدیده‌ها، بر پایه احتمال وقوع یا عدم وقوع یک پدیده است. الگوریتم آدابوست، از الگوریتم بوستینگ انطباقی^{۱۹}، بهینه‌تر شده است و یک الگوریتم یادگیری ماشین با نظارت^{۲۰} و از دسته الگوریتم‌های ترکیبی بهینه‌سازی شده، می‌باشد. الگوریتم آدابوست می‌تواند تعداد زیادی از الگوریتم‌های یادگیری را به منظور ارتقاء عملکرد و بهبود کارایی، با هم ترکیب کند. فرآیند مدل‌سازی پیش‌بینی بیش‌اطمینانی مدیریت در نمودار ۱ ارائه شده است.



شکل (۱) - مدل مفهومی پژوهش

منبع: یافته‌های پژوهشگر

تعریف عملیاتی متغیرهای پژوهش

در تحقیق حاضر بیش‌اطمینانی مدیریت متغیر وابسته است که به معنای تفاوت بین دقت عینی و یقین و اطمینان ذهنی در تصمیم‌گیری‌های مدیریت (فانسکا کاستا^{۲۱} و همکاران، ۲۰۱۷، ۱۷۸۰). مربوط‌ترین تعریف عملیاتی متغیر وابسته تحقیق حاضر، در پژوهش‌های (پایانوتیس و همکاران^{۲۲}، ۲۰۱۷، ۱۱) و (هریبار، یانگ^{۲۳}، ۲۰۱۱، ۳۲)، (کمپل و همکاران^{۲۴}، ۲۰۱۰، ۱۷) و (مالمندیر، تیت، یان^{۲۵}، ۲۰۱۱، ۲۴) و همچنین (شراند و زاچمن^{۲۶}، ۲۰۱۲، ۲۳) بیان شده است. براساس این تحقیقات معیار خطای پیش‌بینی سود هر سهم، اختلاف سود پیش‌بینی شده هر سهم با سود هر سهم واقعی می‌باشد که در صورت وجود اختلاف عدد یک و مدیر بیش اطمینان خواهد بود و بر اساس معادله رگرسیون مخارج سرمایه‌ای زیر اگر باقیمانده رگرسیون بزرگ‌تر از صفر باشد این شاخص برابر یک (وجود بیش‌اطمینانی مدیریت) و در غیر این صورت برابر صفر (عدم وجود بیش‌اطمینانی مدیریت) می‌باشد. همچنین استفاده از این معیار دوم بر این مبنا است که در شرکت‌هایی که دارایی‌ها با نرخ بالاتری نسبت به فروش رشد می‌کنند مدیران این شرکت‌ها نسبت به مدیران سایر شرکت‌ها بیشتر در شرکت سرمایه‌گذاری می‌کنند و بنابراین شاخص معادله رگرسیون زیر برابر یک و حاکی از بیش‌اطمینانی مدیریت خواهد بود.

$$\text{Asset.Gr}_{j,t} = \beta_0 + \beta_1 \times \text{Sale.Gr}_{j,t} + \varepsilon_{j,t}$$

که در آن

$\text{Asset.Gr}_{j,t}$: رشد جمع دارایی‌های سال جاری در مقایسه با سال قبل در هر شرکت

$\text{Sale.Gr}_{j,t}$: رشد فروش سال جاری نسبت به سال قبل در هر شرکت

مقایسه توانایی الگوریتم‌های یادگیری ماشین آداپوست.../ اعتبار، دارایی، حمیدیان و جعفری

$\varepsilon_{j,t}$: خطای مدل که نشان‌دهنده بیش‌اعتمادی مدیریت می‌باشد.

متغیرهای مستقل حسابداری که در این پژوهش به منظور پیش‌بینی بیش اطمینانی مدیریت مورد استفاده قرار گرفته‌اند، شامل: حاشیه سود ناخالص، حاشیه سود عملیاتی، بازده دارایی، سود هر سهم، نسبت جاری، نسبت آنی، نسبت سرمایه در گردش، اهرم مالی، نسبت گردش موجودی‌ها، نسبت گردش دارایی، نسبت گردش دارایی ثابت، گردش حساب‌های دریافتی، کیوتوبین، بازده فروش، بازده سهام، نسبت وجه نقد به دارایی‌ها، نسبت وجه نقد عملیاتی به دارایی‌ها، نسبت دارایی‌های جاری به دارایی‌ها، اندازه شرکت، قیمت سهام، نسبت سود تقسیمی، نسبت P/E هستند که به الگوریتم فیلتر لاس و گاس داده شده‌اند و در نهایت تعداد ۱۱ متغیر مستقل که توسط الگوریتم لاس و گاس فیلتر (LVF) انتخاب شده‌اند، که در جدول (۱) ارائه شده است.

جدول ۱- متغیرهای مستقل مدل پیش‌بینی بیش اطمینانی مدیریت

نام متغیر کمی	تعریف عملیاتی
حاشیه سود ناخالص	سود ناخالص/فروش خالص
حاشیه سود عملیاتی	سود عملیاتی/فروش خالص
بازده دارایی	سود خالص/کل دارایی‌ها
سود هر سهم	سود خالص/تعداد سهام
نسبت گردش دارایی ثابت	فروش/دارایی ثابت
نسبت گردش دارایی	فروش/دارایی
گردش حساب‌های دریافتی	فروش/حساب‌های دریافتی
بازده فروش	سود خالص/فروش خالص
قیمت سهام	قیمت سهام در پایان سال
نسبت سود تقسیمی	سود تقسیمی هر سهم/سود هر سهم
نسبت p/e	قیمت هر سهم/سود هر سهم

منبع: سلیمانی‌رسا، طاهرآبادی، کریمی پویا، (۱۳۹۵)

یافته‌های پژوهش

تحلیل توصیفی متغیرهای پژوهش

به منظور کسب شناخت بیشتر درباره جامعه آماری و متغیرهای مورد مطالعه، خلاصه وضعیت آمار توصیفی متغیرها در جدول ۲ ارائه شده است.

جدول ۲- شاخص‌های توصیفی متغیرهای تحقیق

نام متغیر	میانگین	میانه	کمترین	بیشترین	حرف معیا	چولگی	کشیدگی
حاشیه سود ناخالص	۲/۷۶۷	۰/۲۰۹	۰/۰۰۰	۳۵۷۹/۲۸۶	۹۲/۷۳۵	۳۸/۵۷۸	۱۴۸۸/۸۱۸
حاشیه سود عملیاتی	-۰/۷۹۲	۰/۱۶۵	-۱۴۵۳/۵۷۱	۴/۲۵۶	۳۷/۶۶۷	-۳۸/۵۸۳	۱۴۸۹/۱۰۰
بازده دارایی	۰/۱۲۶	۰/۰۹۹	-۱/۱۰۴	۲/۱۰۱	۰/۱۴۳	۲/۰۷۵	۲۸/۷۷۹
سود هر سهم	۹۲۰/۱۹۱	۵۰۳/۴۸۰	-۸۳۴۷/۱۴۷	۱۵۹۲۴/۱۷۵	۱۴۸۳/۹۳۴	۳/۳۲۰	۲۴/۹۱۰
نسبت گردش دارایی	۰/۹۰۷	۰/۷۶۶	۰/۰۰۰	۲۳/۰۰۹	۰/۸۰۸	۱۴/۴۱۴	۳۷۷/۰۰۷
نسبت گردش دارایی ثابت	۷/۴۹۵	۳/۹۲۲	۰/۰۰۰	۱۹۶/۶۳۱	۱۴/۸۶۳	۷/۲۸۰	۶۷/۲۶۹
گردش حساب‌های دریافتی	۸/۰۲۸	۳/۲۲۴	۰/۰۰۶	۴۶۵/۲۷۶	۲۲/۰۰۷	۱۰/۷۴۱	۱۶۶/۳۷۳
بازده فروش	۲/۵۴۷	۰/۱۳۲	-۳/۳۶۷	۳۴۳۰/۵۷۱	۸۸/۸۸۰	۳۸/۵۸۴	۱۴۸۹/۱۵۲
قیمت سهام	۶۳۳۳/۷۴۰	۳۶۵۳/۵۰۰	۰/۰۰۰	۱۰۵۵۳۱/۰۰۰	۷۹۸۶/۱۱۶	۳/۸۱۳	۲۵/۶۳۵
نسبت سود تقسیمی	۰/۲۰۹	۰/۰۸۴	۰/۰۰۰	۱۶/۰۹۴	۰/۵۳۷	۲۱/۴۹۸	۵۸۵/۱۸۱
اعتماد به نفس کاذب	۰/۳۷۳	۰/۰۰۰	۰/۰۰۰	۱/۰۰۰	۰/۴۸۴	۰/۵۲۵	-۱/۷۲۷
p/e	۱۷۰/۳۲۱	۶/۷۱۷	-۴۳۳/۳۸۲	۱۹۸۳۹۸/۸۳۷	۵۱۴۵/۵۶۱	۳۸/۴۵۲	۱۴۸۲/۲۱۱

شاخص‌های توصیفی متغیرهای تحقیق (۱۱ متغیر مستقل انتخاب شده و متغیر وابسته) شامل میانگین، میانه، دامنه تغییرات، انحراف معیار و چولگی و کشیدگی از بین ۲۲ متغیر مستقل حسابداری که در این پژوهش به منظور پیش‌بینی بیش‌اطمینانی مدیریت مورد استفاده قرار گرفته‌اند، در جدول بالا ارائه شده است. به طور مثال، ۰/۳۷۳ از مدیران دارای اعتماد بنفوس کاذب هستند.

الگوریتم‌های پژوهش

انتخاب متغیرهای مستقل توسط الگوریتم فیلتر LVF

لگوریتم لاس وگاس، یک رویکرد فیلتر برای انتخاب متغیرهای مستقل است. این روش فضای جستجو را به صورت تصادفی با استفاده از یک الگوریتم لاس وگاس جستجو می‌کند. یکسری انتخابهای احتمالی انجام می‌دهد تا با استفاده از آن‌ها و به کمک معیار سازگاری، سریع‌تر به جواب بهینه برسد. این الگوریتم به دلیل استفاده از جستجوی تصادفی، محدوده کمتری از فضای کل حالات را جستجو می‌کند، که اندازه این محدوده به حداکثر تعداد تکرار الگوریتم بستگی دارد. پیدا شدن جواب بهینه به اندازه منابع موجود و زمان اجرای الگوریتم بستگی دارد. در هر بار تکرار، تابع

مقایسه توانایی الگوریتم‌های یادگیری ماشین آداپوست... / اعتبار، دارایی، حمیدیان و جعفری

تولید کننده تعدادی از زیرمجموعه‌های ممکن از فضای جستجو را به صورت تصادفی انتخاب می‌کند و در اختیار تابع ارزیابی قرار می‌دهد. تابع تولید کننده تصادفی پارامترهایی دارد که بایستی تنظیم شوند، تنظیم مناسب این پارامترها در سرعت رسیدن به جواب و پیدا شدن جواب‌های بهتر مؤثر است. معیار سازگاری به شدت به داده‌های آموزشی تکیه دارد و به انتخاب یک زیرمجموعه از ویژگی‌ها تمایل دارد. این معیار منجر به انتخاب مجموعه ویژگی‌های کوچک‌تر می‌شود. این معیار زیرمجموعه‌های با کمترین اندازه را بر اساس از دست دادن یک مقدار قابل قبول سازگاری که توسط کاربر تعیین می‌شود، پیدا می‌کند. الگوریتم LVF برای هر زیرمجموعه کاندید، تعداد ناسازگاری را محاسبه می‌کند. ایده آن بر این اساس است که کلاس محتملتر آن است که در میان نمونه‌های این زیرمجموعه متغیرهای مستقل، تعداد بیشتری متعلق به آن کلاس باشند. یک حد آستانه‌ای برای ناسازگاری در نظر گرفته می‌شود، که در ابتدا ثابت و به صورت پیش فرض صفر است و هر زیرمجموعه‌ای که مقدار ناسازگاری آن بیشتر باشد، رد می‌شود. این روش می‌تواند هر زیرمجموعه بهینه را پیدا کند، حتی برای داده‌های دارای نویز، به شرط آنکه سطح نویز درست در ابتدا مشخص شود. یک مزیت این روش این است که احتیاجی نیست که کاربر مدت زیادی برای بدست آوردن زیرمجموعه بهینه منتظر بماند، زیرا الگوریتم هر زیرمجموعه‌ای که بهتر از بهترین جواب قبلی باشد (هم از لحاظ اندازه زیرمجموعه انتخاب شده و هم از لحاظ نرخ سازگاری)، به عنوان جواب باز می‌گرداند. این الگوریتم کارا است، زیرا تنها مجموعه‌هایی برای ناسازگاری تست می‌شوند، که تعداد ویژگی‌های درون آن کمتر یا مساوی بهترین زیرمجموعه‌ای است که تا کنون پیدا شده است. الگوریتم LVF به صورت زیر است. یک زیر مجموعه تصادفی، S ، از N متغیر مستقل در هر تکرار تولید می‌شود. اگر تعداد متغیر مستقل (C) مجموعه S کمتر از بهترین یافته شده تا الان باشد، یعنی $C < C_{best}$ ، برای داده D با متغیرهای مستقل پیشنهاد شده، معیار ناسازگاری بررسی می‌شود. اگر میزان ناسازگاری کمتر از آستانه γ باشد، C_{best} و S_{best} به ترتیب با C و S جایگزین می‌شوند و این مجموعه چاپ می‌شود. اگر $C = C_{best}$ و معیار ناسازگاری ارضا شود، آنگاه یک مجموعه خوب پیدا می‌شود و این مجموعه چاپ می‌شود. این الگوریتم MAX_TRIES مرتبه تکرار می‌شود. کلید موفقیت LVF معیار ناسازگاری آن است. این معیار مشخص می‌کند که چه مقدار داده‌ها می‌تواند کاهش داده شوند تا پذیرفته شود.

الگوریتم ۱- الگوریتم LVF

- ورودی: D داده‌ها، N تعداد متغیرهای مستقل، γ نرخ ناسازگاری مجاز
- خروجی: مجموعه M متغیر مستقل که شرط ناسازگاری را ارضا می‌کنند

$$C_{best} = N \quad (۱)$$

(۲) برای $i = 1$ تا MAX_TRIES تکرار کن

$$S = randomset(seed) \quad (۳)$$

$$C = numOfFeatures(S) \quad (۴)$$

(۵) اگر $C < C_{best}$ آنگاه

(۶) اگر $InconCheck(S, D) < \gamma$ آنگاه

$$S_{best} = S \quad (۷)$$

$$C_{best} = C \quad (۸)$$

(۹) $print_Current_Best(S)$

(۱۰) در غیر این صورت اگر $C = C_{best}$ و $InconCheck(S, D) < \gamma$ آنگاه

(۱۱) $print_Current_Best(S)$

(۱۲) پایان تکرار

داده‌های شرکت-سال برای دو متغیر اعتماد به نفس کاذب سال جاری و سال آتی به الگوریتم LVF داده شدند و در این الگوریتم $\gamma = 0.5$ قرار داده شد و تعداد ۵ متغیر مستقل برای سال جاری (p/e)، بازده سهام، سود هر سهم، نسبت سود تقسیمی، نسبت گردش دارایی) و ۱۱ متغیر مستقل (p/e، بازده دارایی، بازده فروش، حاشیه سود عملیاتی، حاشیه سود ناخالص، سود هر سهم، قیمت سهام، گردش حساب‌های دریافتی، نسبت سود تقسیمی، نسبت گردش دارایی، نسبت گردش دارایی ثابت) برای سال آتی انتخاب شدند. تعداد متغیرهای مستقل سال جاری کمتر از تعداد متغیرهای سال آتی است و قبل از اینکه الگوریتم آدابوست اجرا شود، می‌توان نتیجه گرفت پیش بینی سال آتی سخت‌تر از پیش بینی سال جاری است از آنجا که تعداد متغیرهای مستقل مرتبط سال جاری کمتر از تعداد متغیرهای سال آتی است. همچنین می‌توان مشاهده کرد که ۴ متغیر بین سال جاری و سال آتی مشترک هستند که نشان دهنده اهمیت این متغیرها است.

الگوریتم آدابوست

آدابوست، مختصر شده از بوستینگ انطباقی یک الگوریتم یادگیری ماشین با نظارت می‌باشد. آدابوست می‌تواند تعداد زیادی از الگوریتم‌های یادگیری را به منظور ارتقاء عملکرد و بهبود کارایی، با هم ترکیب کند. طبقه‌بند پایه‌ای که برای الگوریتم آدابوست استفاده می‌شود، فقط کافی است از طبقه‌بند تصادفی بهتر باشد و به این ترتیب بهبود عملکرد الگوریتم با تکرارهای بیشتر افزایش می‌یابد. حتی طبقه‌بندهای با خطای بالاتر از تصادفی با گرفتن ضریب منفی، عملکرد کلی را بهبود می‌بخشند. آدابوست به داده‌های نویزدار و دارای بخش‌های مجزا حساس می‌باشد و در مسائل بیش برآزش حساسیت کمتری نسبت به الگوریتم‌های یادگیری دیگر دارد. ایده اساسی این الگوریتم به این صورت است که به هر یک از نمونه‌های آموزشی وزن خاصی نسبت داده می‌شود. در ابتدا وزن تمامی نمونه‌ها یکسان می‌باشد ولی در هر تکرار، ساختار آموزش یافته ضعیف، طبقه‌بندی را ارائه می‌دهد و وزن نمونه‌هایی که توسط آن طبقه‌بند به اشتباه طبقه‌بندی شده‌اند افزایش می‌یابد. بدین ترتیب تمرکز الگوریتم بر روی نمونه‌هایی قرار می‌گیرد که به سختی طبقه‌بندی شده‌اند. طبقه‌بند نهایی به صورت رأی‌گیری اکثریت بر روی طبقه‌بندها صورت می‌پذیرد، که در آن طبقه‌بندهایی که خطای کمتر دارند دارای وزن بیشتری می‌شوند. الگوریتم آدابوست انواع مختلفی دارد از جمله نوع آلفا و بتا، که بهبودهای متنوعی بر روی آن صورت گرفته است. الگوریتم آلفا به صورت زیر قابل بیان است:

الگوریتم آدابوست نوع آلفا

-
- **ورودی:** X مجموعه داده‌های آموزش، T تعداد تکرار الگوریتم، T وزن اولیه نمونه‌ها
 - **خروجی:** ω وزن نهایی طبقه‌بندها، h طبقه‌بندهای نهایی
-

الف) مقداردهی اولیه:

توزیع ابتدایی یکنواخت برای نمونه‌های آزمایشی در شروع الگوریتم: $D_1(i) = 1/n$

تعیین تعداد تکرار: T

انتخاب وزن یا معیار اطمینان $\omega_t \in R$

ب) تکرار به ازای $t=1, \dots, T$:

اعمال طبقه‌بند h_t به نمونه‌ها

محاسبه خطای طبقه‌بندی $(\varepsilon_t)h_t$

$$\omega_t = \frac{1}{2} \ln\left(\frac{1 - \varepsilon_t}{\varepsilon_t}\right)$$

تعیین مقدار وزن (ω_t) هر طبقه‌بندی:

$$D_{t+1}(i) = \frac{D_t(i)e^{-\omega_t y_i h_t(x_i)}}{Z_t}$$

بهینه نمودن توزیع مجموعه آموزشی:

وقتی Z_t فاکتور نرمال‌سازی وزن‌ها می‌باشد.

$$\omega = \omega_t, h = h_t, D = D_t, Z = Z_t$$

و y_i خروجی هدف است.

پایان تکرار

دو الگوریتم آلفا و بتا در طبقه‌بندی باینری به یکدیگر شبیه می‌باشند، ولی در طبقه‌بندی‌های بیش از دو کلاس از یکدیگر متمایزند. به همین علت نیازی به معرفی الگوریتم بتا نمی‌باشد. عمل بوستینگ (تقویت) را می‌توان به صورت حداقل کردن یک تابع محدب روی یک مجموعه محدب از توابع در نظر گرفت. به طور خاص تابعی که حداقل می‌شود نمایشی است:

Error!)

$$G = \sum_i e^{y_i f(x_i)}$$

No text of specified style in (\document.

فرم نهایی الگوریتم آدابوست یافتن مدل زیر است.

Error!)

$$f(x) = \sum_t \omega_t h_t(x)$$

No text of specified style in (document.

الگوریتم طبقه‌بندی احتمالی بی‌زین NBC

مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست... / اعتبار، دارایی، حمیدیان و جعفری

در این روش فرض بر این است که کلاس‌ها مستقل از یکدیگر بوده و احتمال بر چسب زدن داده‌ی x با s_j با $P(s_j)$ مشخص می‌شود که $s_j \in \Omega = \{\omega_1 = correct, \omega_2 = bankrupt\}$. احتمال اینکه داده‌ی x با s برچسب زده شود اگر داده‌های کلاس ω_k به عنوان ورودی داده شود در صورت مستقل بودن طبقه‌بندها با رابطه‌ی زیر مشخص می‌شود.

$$P(S | \omega_k) = P(s_1, s_2 | \omega_k) = \prod_{i=1}^2 P(s_i | \omega_k) \quad (1)$$

در این صورت احتمال پسین برای برچسب زدن به X با رابطه ۲ (قانون بیز) محاسبه می‌گردد:

$$P(\omega_k | S) = \frac{P(\omega_k) P(S | \omega_k)}{P(S)} = \frac{P(\omega_k) \prod_{i=1}^2 P(s_i | \omega_k)}{P(S)}, k = 1, 2 \quad (2)$$

مخرج کسر فوق به ω_k بستگی ندارد و قابل صرف نظر کردن است. در نتیجه میزان تعلق X به ω_k از رابطه زیر بدست می‌آید:

$$\mu_k(x) \propto P(\omega_k) \prod_{i=1}^2 P(s_i | \omega_k) \quad (3)$$

آزمون فرضیه پژوهش

مدل‌های پیش‌بینی غیرخطی الگوریتم آدابوست و الگوریتم طبقه‌بندی احتمالی بیزین برای ارزیابی مدل‌های طبقه‌بندی الگوریتم آدابوست و الگوریتم احتمالی بیزین از نرخ شناسایی استفاده شده است. همچنین به منظور اجرای عادلانه و بررسی پدیده بیش‌بزارش از ده روش اعتبار سنجی متقابل^{۲۷} استفاده شده است. شرکت‌سال‌ها به دو دسته آموزشی و ارزیابی توسط ده روش اعتبار سنجی متقابل تقسیم می‌گردند. داده‌های آموزش به الگوریتم‌های آدابوست و بیزین با طبقه‌بندهای پایه درخت تصمیم مبتنی بر ضریب جینی داده می‌شوند. پس از اجرای فرآیند یادگیری الگوریتم‌ها، به منظور اینکه بررسی شود چقدر مدل‌های آدابوست و بیزین، فرآیند یادگیری را با موفقیت سپری کرده‌اند، ابتدا همان داده‌های آموزش را که قبلاً به الگوریتم داده شده است، مجدداً به مدل داده می‌شود، با این تفاوت که این بار مدل‌های آدابوست و بیزین مقدار متغیر وابسته را پیش‌بینی می‌کند، سپس میانگین ۱۰ نرخ شناسایی به ۱۰ روش متقابل محاسبه شده و در جدول ۳ گزارش شده است. میانگین نرخ شناسایی مدل الگوریتم آدابوست و بیزین، برای ارزیابی درصد

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره سی و نهم / تابستان ۱۳۹۸

یادگیری مدل‌ها جهت پیش‌بینی بیش‌اطمینانی یا اعتمادبه‌نفس کاذب مدیریت است، هر چقدر خطاهای یادگیری به صفر نزدیک‌تر باشند، یا میزان یادگیری به صد درصد نزدیک‌تر باشد، نشان‌دهنده یادگیری بهتر مدل هستند. با توجه به نتایج ارائه‌شده، در مدل آدابوست و مدل احتمالی بیزین، نرخ شناسایی سال جاری بیشتر از نرخ شناسایی سال آتی است، از این رو، الگوریتم آدابوست و الگوریتم احتمالی بیزین مدل سال جاری را دقیق‌تر از سال آتی یاد گرفته است. با بررسی میانگین درصد یادگیری مدل پیش‌بینی الگوریتم آدابوست در سال جاری و سال آتی نسبت به مدل پیش‌بینی الگوریتم احتمالی بیزین سال جاری و سال آتی، مدل پیش‌بینی آدابوست قدرت یادگیری بیشتری برای بیش‌اطمینانی مدیریت را دارد.

جدول ۳- نتایج درصد یادگیری مدل‌های پیش‌بینی آدابوست و مدل بیزین

مدل پیش‌بینی الگوریتم بیزین		مدل پیش‌بینی الگوریتم آدابوست		ردیف
سال جاری	سال آتی	سال جاری	سال آتی	
۸۷/۵۱	۹۳/۱۰	۸۹/۷۰	۹۵/۱۷	۱
۹۰/۳۰	۹۳/۸۵	۹۰/۵۳	۹۵/۶۴	۲
۸۶/۲۱	۹۲/۷۳	۹۰/۷۴	۹۴/۹۴	۳
۸۷/۹۵	۹۳/۶۰	۹۰/۳۹	۹۵/۲۹	۴
۸۵/۳۱	۹۳/۶۰	۹۰/۸۷	۹۹/۹۴	۵
۸۶/۱۹	۹۴/۴۷	۹۱/۹۰	۹۴/۷۰	۶
۸۶/۲۱	۹۲/۹۷	۹۱/۲۸	۹۵/۰۵	۷
۸۴/۸۷	۹۲/۱۰	۹۱/۲۱	۹۵/۴۱	۸
۸۷/۰۹	۹۳/۰۹	۹۰/۸۷	۹۴/۸۱	۹
۸۷/۵۱	۹۴/۲۲	۹۱/۶۳	۹۴/۴۰	۱۰
۸۶/۹۲	۹۳/۳۷	۹۰/۹۱	۹۵/۰۴	میانگین

به‌منظور بررسی عدم رخداد پدیده‌ای به نام بیش‌برازش، داده‌های مربوط به شرکت-سال‌های ارزیابی مدل‌های پیش‌بینی که تابع‌حال الگوریتم آن‌ها را مشاهده نکرده است، به مدل‌های به دست آمده الگوریتم‌های آدابوست و بیزین داده می‌شوند. الگوریتم آدابوست و الگوریتم احتمالی بیزین، اعتماد به نفس کاذب مدیریت را برای تمامی این شرکت-سالها (داده‌های ارزیابی) پیش‌بینی

مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست.../ اعتبار، دارایی، حمیدیان و جعفری

می‌کنند، از مقایسه مقادیر پیش‌بینی‌شده با مقادیر واقعی میزان کارایی مدل‌های پیش‌بینی، برای شرکت-سال‌هایی که تاکنون الگوریتم‌ها آن‌ها را مشاهده نکرده‌اند، به دست می‌آید. این فرآیند توسط ده روش اعتبار سنجی متقابل انجام می‌شود تا تمامی شرکت‌ها حداقل یک بار به‌عنوان داده‌های ارزیابی مورد آزمون قرار بگیرند تا نتایج حاصل از مدل‌های پیش‌بینی قابلیت اعتماد بیشتری داشته باشند. در جدول ۴ میانگین این نرخ‌های شناسایی برای ارزیابی عدم وجود پدیده بیش‌برازش (ارزیابی کارایی مدل‌های پیش‌بینی آدابوست و بیزین)، نشان داده شده است. مدل‌های الگوریتم آدابوست و الگوریتم احتمالی بیزین برای شرکت-سال‌هایی که تا به حال ندیده است (داده‌های ارزیابی مدل‌های پیش‌بینی)، دارای دقت پیش‌بینی نزدیک به شرکت-سال‌های آموزش (داده‌های آموزشی مدل‌های پیش‌بینی) است و فقط مقداری خطا به آن اضافه شده است. از آنجا که اختلاف نرخ شناسایی داده‌های آموزشی و داده‌های ارزیابی مدل‌های پیش‌بینی کم است، پس بیش‌برازش برای پیش‌بینی آدابوست و بیزین اتفاق نیافتاده است و با در نظر گرفتن میانگین نتایج ارزیابی به دست آمده برای دو مدل پیش‌بینی الگوریتم آدابوست و الگوریتم بیزین در کل کارایی مدل پیش‌بینی آدابوست در مقایسه با مدل پیش‌بینی بیزین بیشتر است یا به عبارت دیگر با توجه به میانگین کل نتایج به دست آمده مدل پیش‌بینی آدابوست نسبت به مدل پیش‌بینی بیزین برای پیش‌بینی بیش‌اطمینانی مدیریت، قابلیت کاربرد بیشتری دارد.

جدول ۴- نتایج ارزیابی کارایی (کاربردی بودن) مدل‌های پیش‌بینی آدابوست و بیزین

ردیف	مدل پیش‌بینی الگوریتم آدابوست سال جاری سال آتی	مدل پیش‌بینی الگوریتم بیزین سال جاری سال آتی
۱	۹۳/۶۲	۹۴/۳۱
۲	۹۵/۲۴	۸۵/۵۲
۳	۹۵/۲۱	۸۶/۶۶
۴	۹۴/۷۱	۹۴/۷۳
۵	۹۵/۲۱	۸۴/۲۱
۶	۹۵/۲۴	۸۸/۱۵
۷	۹۵/۲۴	۸۱/۳۳
۸	۹۴/۱۵	۸۵/۵۲
۹	۹۵/۲۴	۹۰/۶۶
۱۰	۹۵/۷۷	۹۰/۷۸
میانگین	۹۴/۹۶	۸۷/۸۴

بحث و نتیجه گیری

یکی از مهم ترین خصوصیات شخصیتی مدیران، اعتماد به نفس بیش از حد است. با در نظر گرفتن اهداف پژوهش حاضر، الگوریتم های آدابوست و بیزین به منظور پیش بینی بیش اطمینانی مدیریت در شرکت های پذیرفته شده بورس اوراق بهادار تهران اعتبار سنجی شد و با توجه به نتایج تجزیه و تحلیل داده های پژوهش که پیش از این توضیح داده شد، مدل های پیش بینی بیش اطمینانی مدیریت برای دو سال یعنی سال جاری و آتی (دوره زمانی تحقیق ۸۷ تا ۹۵) ارائه شد که با توجه به نتایج ارائه شده، مدل های پیش بینی الگوریتم آدابوست و الگوریتم احتمالی بیزین، هر دو توانایی پیش بینی بیش اطمینانی مدیریت را دارند؛ اما مدل پیش بینی آدابوست در مقایسه با مدل پیش بینی بیزین نتایج بهتری برای پیش بینی اعتماد به نفس کاذب مدیریت دارد که نشان دهنده قدرت بیشتر در یادگیری و کارایی بهتر این مدل به منظور پیش بینی اعتماد به نفس کاذب مدیریت می باشد؛ بنابراین فرضیه اول پژوهش مبنی بر اینکه مدل پیش بینی الگوریتم (یادگیری ماشین) آدابوست، در پیش بینی بیش اطمینانی مدیریت سال جاری و سال آتی توانایی دارد، با در نظر داشتن نرخ شناسایی ۹۰٪ به عنوان آستانه شناسایی قابلیت کارایی و توانایی بالای مدل را در سال جاری و آتی نشان می دهد و بنابراین فرضیه اول پژوهش پذیرفته می شود.

فرضیه دوم پژوهش نیز که بیان می دارد، در سطح نرخ شناسایی ۹۰ درصد، مدل پیش بینی الگوریتم طبقه بندی احتمالی بیزین، در پیش بینی بیش اطمینانی مدیریت سال جاری و سال آتی توانایی دارد، با توجه به میانگین نرخ های شناسایی یا نتایج درصد یادگیری مدل پیش بینی الگوریتم بیزین (عدد سال آتی = ۸۶/۹۲) و میانگین نرخ شناسایی ده روش اعتبارسنجی متقابل با داده های تست برای ارزیابی کارایی مدل پیش بینی الگوریتم بیزین (عدد سال آتی = ۸۷/۸۴) رد می شود. همچنین با توجه به تحلیل نتایج کسب شده، می توان بیان نمود که توانایی پیش بینی بیش اطمینانی مدیریت توسط الگوریتم یادگیری ماشین آدابوست بیشتر از الگوریتم طبقه بندی احتمالی بیزین می باشد؛ بنابراین فرضیه سوم پژوهش نیز پذیرفته می شود و در نهایت می توان بیان نمود که، مدل پیش بینی آدابوست نسبت به مدل پیش بینی بیزین برای پیش بینی بیش اطمینانی مدیریت، قابلیت کاربرد بیشتری دارد. نکته قابل توجه این است که در پژوهش حاضر ۲۲ متغیر مستقل حسابداری بررسی شدند و متغیرهای p/e ، بازده دارایی، بازده فروش، حاشیه سود عملیاتی، حاشیه سود ناخالص، سود هر سهم، قیمت سهم، گردش حساب های دریافتی، نسبت سود تقسیمی، نسبت گردش دارایی، نسبت گردش دارایی ثابت انتخاب شدند، این در حالی است که با

مقایسه توانایی الگوریتم‌های یادگیری ماشین آدابوست.../ اعتبار، دارایی، حمیدیان و جعفری

تحلیل محتوای مبانی نظری داخلی انجام شده و متغیرهای بکار رفته در مدل‌های پژوهش محققانی مانند یحیی‌زاده‌فر و شمس (۱۳۹۶) متغیرهای اندازه شرکت، حاشیه سود عملیاتی، نسبت دارایی مشهود به کل دارایی‌ها، مهرگان و دلشاد (۱۳۹۶) بازده عادی و غیر عادی سهام، عبدلی، امیری اسفندقه (۱۳۹۵)، رفیع نژاد، لازمی (۱۳۹۵) قیمت سهام و ریسک آن، فروغی و قاسم‌زاد (۱۳۹۴) قیمت سهام، اندازه شرکت، اهرم مالی، بازده دارایی، انحراف معیار بازده ماهیانه شرکت، عربصالحی، امیری، کاظمی نوری (۱۳۹۳) معیار کیوتوبین، اهرم مالی، اندازه شرکت، بازده سهام، و فروغی و نخبه فلاح (۱۳۹۳) اندازه شرکت، نسبت وجوه نقد عملیاتی، با متغیرهای پژوهش حاضر مشترک هستند و همچنین پژوهش حاضر تمامی متغیرهای حسابداری مدل‌های به کار رفته در مبانی نظری داخل کشور را تا کنون پوشش می‌دهد و تعداد متغیر بیشتری را نیز مورد تحقیق و بررسی قرار داده است.

سایر پژوهشگران نیز مانند کانگ و همکاران (۲۰۱۷) که با استفاده از یک مدل سرمایه‌گذاری چند دوره‌ای تأثیر بیش‌اطمینانی مدیریت بر عملکرد مالی را سنجیدند و هیتون (۲۰۱۶) نیز با ارائه مدلی برای بررسی فرایند تصمیم‌گیری مدیران با خوش‌بینی بیش از حد بر تأثیر بیش‌اطمینانی مدیریت بر بیش ارزش‌گذاری فرصت‌های سرمایه‌گذاری شرکت تأکید داشت. (چن و همکاران، ۲۰۱۴) در تحقیق خود تنها به بررسی عملکرد بلندمدت شرکت تحت تأثیر بیش‌اطمینانی مدیریت پرداختند، (فانسکااکاستا و همکاران، ۲۰۱۷) نیز از اقلام تعهدی و بیش‌اطمینانی مدیریت، در پژوهش خود استفاده کردند. در پژوهش‌های قبلی، محققان از تعداد محدودی متغیرهای حسابداری استفاده نموده و مدل‌های بررسی‌شده نیز رگرسیونی بودند که تحقیق حاضر در سطح گسترده‌تری با ارائه مدل‌های غیرخطی پیش‌بینی توسط الگوریتم‌های یادگیری ماشین و طبقه‌بندی احتمالی بیزین صورت گرفته است. علاوه بر مطالب ذکرشده، می‌توان بیان نمود که نتایج پژوهش حاضر در استفاده از الگوریتم‌های ترکیبی جهت مدل‌سازی و پیش‌بینی متغیر بیش‌اطمینانی با نتایج تحقیق (جان و همکاران، ۲۰۱۸) هماهنگی دارد، همچنین از منظر به‌کارگیری الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی و بیزین با پژوهش‌های (هورتوگ، زانتن، ۲۰۱۸) و (جونپور و همکاران، ۲۰۱۹) سازگاری دارد با توجه به نتایج پژوهش می‌توان پیشنهاد‌های زیر را برای پژوهش‌های آینده ارائه داد:

سرمایه‌گذاران و سایر استفاده‌کنندگان از اطلاعات مالی شرکت‌ها، اثرات مخرب بیش‌اعتمادی مدیران را مدنظر قرار دهند. تحلیل‌گران بازارهای مالی و کارگزاران بورسی تأثیرات بلندمدت

فصلنامه مهندسی مالی و مدیریت اوراق بهادار / شماره سی و نهم / تابستان ۱۳۹۸

بیش اطمینانی مدیران شرکت‌ها در بازار سرمایه را به‌عنوان متغیر مربوط به تصمیمات سرمایه‌گذاری در نظر گیرند. توصیه می‌شود از سایر الگوریتم‌های مرتبط با هوش مصنوعی و شبکه‌های عصبی بیزین و مصنوعی برای پیش‌بینی بیش اطمینانی مدیریت استفاده شود و سپس با نتایج این پژوهش مطابقت داده شود. همچنین پیشنهاد می‌شود در پژوهش‌های آینده، از روش متغیر گزینی متفاوتی استفاده شود تا بتوان با سایر متغیرهایی که می‌تواند بر بیش اطمینانی مدیریت تأثیرگذار و تعیین‌کننده باشد، مدل‌های پیش‌بینی اعتماد به نفس کاذب مدیران را بهبود بخشید.

فهرست منابع

- ۱) امید، محمدرضا، امیدی، نبی، شیری، اردشیر، محمدی‌پور، رحمت‌اله، (۱۳۹۶)، توسعه کاربرد مدل‌های باکس جنکینز، شبکه عصبی مصنوعی و تعدیل نمایی در پیش‌بینی و مدیریت پدیده‌های اجتماعی (مطالعه موردی: پیش‌بینی روند ازدواج و طلاق در استان ایلام)، آینده‌پژوهی مدیریت، ۲۸(۱۱۰): ۴۱-۵۴.
- ۲) تلخایی، فاطمه، نیکومرام، هاشم، طالب‌نیا، قدرت‌اله، رویایی، رمضانعلی، (۱۳۹۶)، بررسی چسبندگی حق‌الزحمه‌های حسابرسی و عوامل ساختاری مؤثر بر انحراف حق‌الزحمه‌های حسابرسی از مدل استاندارد، فصلنامه دانش حسابداری و حسابرسی مدیریت، ۶(۲۴): ۹۵-۱۱۲.
- ۳) چهارمحالی، علی اکبر، زاهدی، (۱۳۹۶)، جواد، بررسی تأثیر بیش‌اعتمادی مدیران بر چسبندگی هزینه‌ها، فصلنامه دانش حسابداری و حسابرسی مدیریت، ۶(۲۳): ۳۱-۴۵.
- ۴) خداکرمی، وحید، حقی، فرشاد، امیری، سمیه، آقابابایی، سلمان، (۱۳۹۳)، تحلیل پارامترهای هزینه‌های کیفیت با رویکرد شبکه‌های بیزین، مدیریت صنعتی، سال نهم، ویژه‌نامه مدیریت و حسابداری، دانشگاه آزاد اسلامی واحد سنندج
- ۵) رامشه، منیژه، ملانظری، مهناز، (۱۳۹۳)، بیش اطمینانی مدیریت و محافظه‌کاری حسابداری، دانش حسابداری، ۵(۱۶): ۵۵ - ۷۹.
- ۶) رفیع‌نژاد، عزیز، لازمی، سعید، (۱۳۹۵)، بررسی رابطه اعتماد بنفس بیش از حد مدیران و ریسک سقوط قیمت سهام در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، مطالعات اقتصاد مدیریت مالی و حسابداری، ۲(۲): ۱-۱۰.
- ۷) شهیکی‌تاش، محمدنبی، مولایی، صابر، دینارزهی، خدیجه، (۱۳۹۳)، بررسی ارتباط رشد اقتصادی و ضریب رفاه اجتماعی در ایران بر اساس رهیافت بیزین، ۴(۱۶): ۴۱-۵۲.
- ۸) عباسی، ابراهیم، وکیلی‌فرد، حمیدرضا، معروف، مرتضی، (۱۳۹۷)، بررسی تأثیر بیش‌اطمینانی مدیریتی بر کیفیت گزارشگری مالی و محافظه‌کاری شرطی در بورس تهران، فصلنامه دانش حسابداری و حسابرسی مدیریت، ۷(۲۵): ۱۹۳-۲۰۵.
- ۹) عبدلی، محمدرضا، امیری‌اسفندقه، ملیحه، (۱۳۹۵)، بررسی ارتباط بین اطمینان بیش از حد مدیران و پیش‌بینی مدیریت سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، فصلنامه بورس اوراق بهادار، ۹(۳۵): ۵-۲۷.
- ۱۰) عربصالحی، هدی، امیری، هادی، کاظمی نوری، سپیده، (۱۳۹۳)، بررسی تأثیر اعتماد به نفس بیش از حد مدیران ارشد بر حساسیت سرمایه‌گذاری جریان‌های نقدی، پژوهش‌های حسابداری مالی، ۶(۲۰): ۱۲۸-۱۱۵.

- (۱۱) عظیمی، عابد، مشایخی، بیتا، (۱۳۹۵)، تأثیر توانایی‌های مدیریت بر رابطه بین مدیریت سود واقعی و عملکرد آتی شرکت، بررسی‌های حسابداری و حسابرسی، ۲۳(۲): ۲۵۳-۲۶۷.
- (۱۲) فروغی، داریوش، قاسم‌زاد، پیمان، (۱۳۹۴)، تاثیر اطمینان بیش از حد مدیریت بر ریسک سقوط آتی قیمت سهام، ۲(۲): ۷۲-۵۵.
- (۱۳) فروغی، داریوش، نخبه‌فلاح، زهرا، (۱۳۹۳)، تاثیر اطمینان بیش از حد مدیریتی بر محافظه‌کاری شرطی و غیرشرطی، پژوهش‌های حسابداری مالی، ۶(۱): ۲۷-۴۴.
- (۱۴) مهرآرا، محسن، غضنفری، (۱۳۹۳)، بررسی علل تورم در اقتصاد ایران مبتنی بر رویکرد متوسط‌گیری بیزین، راهبرد اقتصادی، ۳(۱۰): ۸-۳۷.
- (۱۵) مهرگان، محمدرضا، دلشاد، افسانه، (۱۳۹۶)، بررسی تاثیر تغییر نگرش مدیریت از بیش اطمینانی به مدیریت منطقی بازده سهام، دانش حسابداری مالی، ۴(۲): ۱-۲۶.
- (۱۶) نصیری پوردی، سلاله، بنی مهد بهمن، احمدزاده حمید، (۱۳۹۵)، بیش‌اعتمادی مدیر و بیش‌نمایی سود، فصلنامه علمی پژوهشی حسابداری مدیریت، ۹(۳۰): ۵۵-۶۷.
- (۱۷) یحیی‌زاده‌فر، محمود، شمس، شهاب‌الدین، (۱۳۹۶)، بررسی رابطه خوش‌بینی مدیریتی و هموارسازی سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، مدیریت دارایی و تامین مالی، ۵(۳): ۷۷-۸۸.

18) Abhinav, Rishabh, Pindoriya, Naran M, Wu, Jianzhong, Long, Chao, (2017). "Short-term wind power forecasting using wavelet-based neural network, Energy Procedia, 142(3): 455-460, available at: <https://doi.org/10.1016/j.egypro.2017.12.071>.

19) Aroenkamol, Prasart Jongj, and Laux, Volker, (2017). "Insider versus Outsider CEOs, Executive Compensation, and Accounting Manipulation," Journal of Accounting and Economics, available at: <http://dx.doi.org/10.1016/j.jacceco.2017.01.002>.

20) Baig, Mirza M. and Mian.M., and Awais, El-Sayed and M. El-Alfy (2017), "AdaBoost-based artificial neural network learning, Neuro computing". 248: 120-126, available at: <https://doi.org/10.1016/j.neucom.2017.02.077>

21) Bamber, Matthew, and McMeeking, Kevin, (2016), "An examination of international accounting standard-setting due process and the implications for legitimacy". The British Accounting Review, 48(1):59-73.

22) Berk, J. and Stanton, R. (2007). "Managerial ability, compensation, and the closed-end fund discount". The Journal of Finance, 62 (2): 529-556.

- 23) Ben-David, I., and Graham, J. and Harvey, C. (2012). “Managerial overconfidence and corporate policies”. Working Paper, Available at: <http://www.nber.org/papers/w13711>.
- 24) Bharati, Rakesh. and Doellman, Thomas. and Xudong, Fua. (2016). “CEO confidence and stock returns”. <https://doi.org/10.1016/j.jcae.2016.02.006>.
- 25) Campbell, T.C. and Gallmeyer, M.F. and Johnson, S. A. and Rutherford, J. Stanley, B. (2010). “CEO Optimism and Forced Turnover”. *Journal of Financial Economics*, 101(3): 695-712.
- 26) Chen, S. and Lai, S. and Liu, C. and McVay, S. (2014). “Overconfident managers and internal controls”. Working paper, National Taiwan University and University of Washington.
- 27) Cravens, K. S. and Oliver, E. G. and Stewart, J. S. (2010). “Can a positive approach to performance evaluation accomplish your goals?”. *Kelley School of Business*, No53: 269-279.
- 28) Demerjian, P. Lev, B. and McVay, S. (2012). “Quantifying managerial ability: A new measure and validity tests”. *Management Science*, 58(7): 1229-1248.
- 29) Elshandidya, Tamer, Philip J. Shrivess, Matt Bamberg, Santhosh Abraham, (2018). “Risk reporting: A review of the literature and implications for future research”. *Journal of Accounting Literature* 40 (3): 54-82.
- 30) Fonseca Costa, Daniel, and de Melo Carvalho, and de Melo F., and Moreira, B.C. et al. (2017). “Scientometrics behavioral finance and decision making with cognitive biases such as overconfidence”, anchoring effect and confirmation bias. 111(3): 1775-1799.
- 31) Heaton, J. B. (2016). “Managerial optimism and corporate finance”, *Financial Management*, 31: 33-45.
- 32) Hartog, Jarno, and Harry van Zanten, (2018). “Nonparametric Bayesian label prediction on a graph”. *Computational Statistics & Data Analysis*, 120: 111-131, available at: <https://doi.org/10.1016/j.csda.2017.11.008>
- 33) Hribar, P. and Yang, H. (2011). “CEO overconfidence and management forecasting CEO overconfidence and management forecasting”. Working Paper, University of Iowa. <http://www.daedalustrust.org.uk/2011/11/ceooverconfidence-and-management-forecasting-2011>.

- 34) Junior, Bertini, and Joao, Roberto, and Nicolettib, Maria.and do Carmo, (2019). “An iterative boosting-based ensemble for streaming data classification”. *Information Fusion* 45 (2019) 66–78, available at: <https://doi.org/10.1016/j.inffus.2018.01.003>.
- 35) Jun, Wang.and Yuyan, Luo.and Lingyu, Tang, and Ge Peng, (2018). “Modeling a combined forecast algorithm based on sequence patterns and near characteristics: An application for tourism demand forecasting”. Online <https://doi.org/10.1016/j.chaos.2018.01.028>.
- 36) Malmendier, U. G. Tate, and J. Yan. (2011). “Overconfidence and Early. Life experiences: The effect of managerial traits on corporate financial policies”. *The Journal of Finance* 66 (5): 1687–733.
- 37) Marucci-Wellman, Helen R, and Corns, Helen L, and Lehto, Mark R, (2017). “Classifying injury narratives of large administrative databases for surveillanceacA practical approach combining machine learning ensembles and human review”. *Accident Analysis & Prevention*, Volume 98: 359-371, available at: <https://doi.org/10.1016/j.aap.2016.10.014>.
- 38) Pakwan, G. (2013). “Industry Determinants of Organizational Culture”, *Academy of Management Review*, 369- 415.
- 39) Panayiotis C. Andreou, and John A. Doukas, and Koursaros, Demetris, and Louca, Christodoulos,. (2017), “CEO Overconfidence and the Valuation Effects of Corporate Diversification & Refocusing Decisions”. Electronic copy available at: <https://ssrn.com/abstract=2898469>.
- 40) Scherand, C.M. and Zechman, S.L. (2012). “Executive Overconfidence and the Slippery Slope to Financial Misreporting”. *Journal of Accounting and Economics*, 53(1): 86–107.

- 1 Aroenkamol & Laux
- 2 Berk & Stanton
- 3 Elshandidya et al
- 4 Heaton
- 5 Marucci-Wellman & Corns & Lehto
- 6 Abhinav et al
- 7 Baig et al
- 8 Bamber & McMeeking
- 9 Pakwan
- 10 Junior et al
- 11 Jun et al
- 12 Hartog & Zanten
- 13 Kang et al
- 14 Bharati & Doellman & Thomas & Xudong
- 15 Chen et al
- 16 Ben-David & Graham & Harvey
- 17 Las Vegas Filter
- 18 Adaboost
- 19 Adaptive Boosting
- 20 Supervise
- 21 Fonseca Costa et al
- 22 Panayiotis et al
- 23 Hribar & Yang
- 24 Campbell et al
- 25 Malmendier & Tate & Yan
- 26 Scherand & Zechman
- 27 10-Fold Cross-Validation