



فصلنامه علمی پژوهشی
دانش حسابداری و حسابرسی مدیریت
دوره ۱۵ / شماره ۳ (پیاپی ۵۹) / پائیز ۱۴۰۵
صفحه ۱۷۵ تا ۱۸۶

ارائه الگوی پیش‌بینی تقلب مبتنی بر هوش مصنوعی کاربرد مدل ماشین بردار پشتیبان (SVM)

علی زارعی

دانشجوی دکتری، گروه حسابداری، واحد امارات، دانشگاه آزاد اسلامی
lali5zaree.ir@gmail.com

فریدون رهنمای رودپشتی

استاد تمام، گروه حسابداری، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران
rahnama@iau.ir

حامد خانمحمدی

استادیار گروه حسابداری واحد دماوند، دانشگاه آزاد اسلامی دماوند، ایران
dr.khanmohammadi@yahoo.com

حمیدرضا کردلویی

دانشیار، گروه حسابداری، واحد اسلامشهر، دانشگاه آزاد اسلامی، تهران، ایران
hamidreza.kordlouie@gmail.com

تاریخ دریافت: ۱۴۰۳/۰۵/۰۸ تاریخ پذیرش: ۱۴۰۳/۰۸/۲۹

چکیده

در جهان امروز همزمان با پیشرفت تکنولوژی، کاربرد علم داده و هوش مصنوعی در همه حوزه‌ها از جمله حسابداری امری انکارناپذیر است. هوش مصنوعی در علم حسابداری در مباحث مختلف از جمله انجام فرآیند حسابداری داخلی، گزارشات مرتبط با درآمدها و هزینه‌ها و حتی کشف و پیش‌بینی تقلب کاربرد گسترده‌ای پیدا کرده که در واقع به دلیل حذف خطای انسانی، نتایج بدست آمده از سرعت و قابلیت اتکای بالاتری برخوردار هستند. لذا در این پژوهش هدف بررسی کارایی روش ماشین بردار پشتیبان در جهت پیش‌بینی تقلب است. در این راستا در بازه زمانی ۱۳۹۵ تا ۱۴۰۱، تعداد ۱۳۶ شرکت پذیرفته شده در بورس اوراق بهادار از این طریق مورد بررسی قرار گرفتند. نتایج بیانگر اینست که این مدل با ضریب دقت ۸۵ درصد، ضریب بازخوانی ۹۷ درصد و صحت ۸۳ درصد از کارایی قابل قبولی در جهت پیش‌بینی تقلب برخوردار است. از سویی خروجی حاصل آن از بیانگر اینست که جمع کل دارایی‌های دوره قبل، همچنین جریان نقدی عملیاتی، سود ناخالص دوره قبل و سود خالص و همچنین سرمایه دوره قبل و جاری اثرگذاری را بر ارتکاب تقلب داشته‌اند.

واژه‌های کلیدی: تقلب، شرکت‌های پذیرفته شده در بورس اوراق بهادار، ماشین بردار پشتیبان، هوش مصنوعی.

۱- مقدمه

بر اساس تعریف گزارش انجمن بازرسان رسمی تقلب آمریکا^۱ در سال ۲۰۱۲، تقلب^۲ به معنای سوء استفاده فرد از موقعیت و نفوذ خود به منظور بکارگیری نادرست از دارایی‌ها و منابع یک سازمان است. بر همین اساس تقلب حرفه ای به طور کلی در سه دسته قابل تقسیم‌بندی است:

(۱) فساد مالی^۳: تقلبی است که در آن کارکنان یک واحد از موقعیت و نفوذ خود به منظور دستیابی به منافع مستقیم و یا غیرمستقیم، سوءاستفاده می‌کنند؛ مانند دریافت پورسانت و درگیر شدن در تضاد منافع.

(۲) سوءاستفاده از دارایی‌ها^۴: شامل دزدی یا استفاده نادرست کارکنان از موجودی‌ها و سایر دارایی‌های یک سازمان است. مانند ترفندهای سرقت وجه نقد یک شرکت به صورت مستقیم یا سرقت توسط رسیدهای وجه نقد، پرداخت‌های جعلی، دزدی، سرقت و خرید و فروش غیرقانونی دارایی‌ها و ...

(۳) تقلب در صورت‌های مالی^۵: عبارت است از ارائه نادرست، حذف اقلام و افشا نکردن کافی اطلاعات به منظور فریب استفاده‌کنندگان صورت‌های مالی، به خصوص سرمایه گذاران و اعتباردهندگان که بیشتر اوقات با بیش‌نمایی^۶ دارایی‌ها و درآمدها یا کم‌نمایی بدهی‌ها و هزینه‌ها همراه است و یا برعکس.

در این میان بیش‌نمایی دارایی‌ها و درآمدها یا کم‌نمایی بدهی‌ها و هزینه‌ها بیشترین سهم در وقوع تقلب را دارند از اینرو برای جلوگیری از بروز آن‌ها بایستی توجه خاصی به ابعاد مختلف صورت‌های مالی در شرکت‌های مختلف داشت. لازم به ذکر است کاهش شیوه انجام تقلب به صورت برعکس است یعنی با کم‌نمایی در سطوح درآمدی و یا بیش‌نمایی سطح هزینه‌ها صورت می‌گیرد که در واقع بارزترین شیوه انجام تقلب بشمار می‌روند. بنابراین تقلب گزارشگری مالی عبارت است از تحریف عمدی در صورت‌های مالی برای ارائه یک تصویر نادرست از ساختار و وضعیت مالی یک سازمان. با توجه به گستردگی اقلام تقلب در یک سازمان قاعدتا وقوع آن با آثار منفی و قابل توجهی بر سرمایه‌گذاران و حتی مدیران همراه است. در واقع تقلب با از بین بردن سطح اعتمادپذیری به گزارش‌ها موجب افزایش ریسک و هزینه کسب و کار می‌شود و از سویی صلاحیت حرفه‌ای

حسابداری و حسابرسی نیز تحت‌الشعاع قرار می‌گیرد (خواجوی و منصور ۱۳۹۴). تقلب در صورت‌های مالی گاهی خساراتی فراتر از حد سازمان و یا سهام‌داران خرد می‌تواند در بر داشته باشد، بدین صورت که افشای تقلب خصوصا در شرکت‌های بزرگ و ورشکستگی آن‌ها می‌تواند همانند شوک خارجی، ضربه بزرگی با اقتصاد کشورها وارد نماید که آثار آن تا مدت‌ها تداوم داشته باشد. بدین معنا که افشای تقلب از سوی سازمان‌ها و از سویی ورشکستگی آن‌ها به دلیل عدم اطمینان به فرآیند گزارش دهی، عموماً با از بین رفتن فرصت‌های شغلی و همچنین تضعیف قابلیت اتکا و شفافیت در صنعت مربوطه گردد. بنابراین ارزیابی، شناسایی و حتی پیش‌بینی تقلب یکی از مهم‌ترین وظایف حسابداری در سطح یک سازمان است و استفاده‌کنندگان از صورت‌های مالی انتظار دارند تا برآورد صحیحی از تمامی موارد تقلب در گزارشات حسابرسی دیده شود. در این میان شناسایی و تخمین دقیق تقلب نیز امری بسیار دشوار و حتی در برخی موارد نشدنی است زیرا بسیاری از انواع تقلب به صورت کشف نشده باقی می‌مانند و یا افشا نمی‌شوند (فارل و همکاران ۱۹۹۹). با این حال مطالعات مختلفی به منظور ارزیابی خسارات ناشی از تقلب تا کنون صورت گرفته است. به عنوان مثال بر مطالعات کوتون^۷ نشان می‌دهد که زیان ناشی از عدم برآورد درست تقلب در شرکت‌هایی همچون گلوبال کراسینگ^۸، تایکو^۹، کوئست^{۱۰}، ورلد کام^{۱۱} و انرون به سهام‌داران خود وارد کردند رقمی حدود ۴۶۰ میلیارد دلار بوده است. از سویی در مطالعه موردی کشور ایران، اهمیت شناسایی و پیش‌بینی تقلب تا جاییست که اگر تقلب یکی از شرکت‌های پذیرفته شده در بورس اوراق بهادار فاش گردد موجب زمینه بی‌اعتمادی به کل بازار سرمایه و حتی از بین رفتن اعتماد به سایر شرکت‌ها خواهد شد که نتیجه‌ای جز کاهش ارزش سهام آن‌ها نخواهد داشت و حتی آثار مخرب آن می‌تواند تا مدت‌ها تدام داشته باشد بنابراین استفاده از تکنیک‌های کارآمد برای یک سازمان و در مقیاس وسیع تر یک اقتصاد به منظور پیش‌بینی تقلب به عنوان یک ابزار هشدار پیش از موعد بسیار حائز اهمیت است. در رویکردهای سنتی حسابداری برای شناسایی تقلب عموماً از روش‌های نمونه‌گیری تصادفی یا سیستماتیک استفاده می‌شد که خروجی آن‌ها نمی‌توانست برداشت دقیق و جامعی را ارائه دهد. با پیشرفت تکنولوژی و ارایه متدهای نوین حسابداری مبتنی بر الگوهای

⁷ cotton,2002

⁸ Global Crossing

⁹ Tyco

¹⁰ Quest

¹¹ WorldCom

¹ Association of Certified Fraud Examiners (ACFE)

² Fraud

³ Corruption

⁴ Assets Misappropriation

⁵ Fraud in financial statements

⁶ Overstating

مالی متقلبان، تقلب کارکنان، سوءاستفاده از دارایی‌ها و تقلب‌های ورشکستگی در حوزه‌ی مالی قابل مشاهده است (مرادی و سلیمانی مارشک، ۱۳۹۰)

به طور کلی دلایل ارتکاب تقلب بر اساس نظریه‌های مختلفی در طول زمان مطرح شده‌اند. در دهه ۱۹۵۰ نظریه مثلث تقلب که از سوی دونالد کریسی^۳ مطرح گردید به این موضوع اشاره نمود که سه عامل فرصت، توجیه و فشار را از عوامل موثر بر تقلب هستند. در سال ۲۰۰۴ مدل دیامون که توسط ولف هرمانسون مطرح شد به این موضوع اشاره دارد که علاوه بر سه عامل مدل مثلث بعد توانایی نیز بر تقلب تاثیرگذار است. در واقع توانایی به این معناست که متقلب از ضعف موجود در سیستم یا اعتماد موجود است که می‌تواند زمینه ساز تقلب باشد. رامامورتی و همکاران (۲۰۰۹)^۴، با بسط مدل مثلث عوامل محیطی و فرهنگی را نیز اضافه نمودند که مهم ترین نتیجه بدست آمده بر این موضوع دلالت دارد که کمبود اخلاق در یک سازمان بر تقلب اثر معنادار دارد. در سال ۲۰۱۰ نظریه مربع تقلب توسط سایزلویز مطرح گردید این نظریه نیز بر نظریه مثلث تقلب استوار است و به اهمیت تاثیرات اجتماعی و بین المللی بر تقلب اشاره می‌کند. در واقع تاثیر عوامل مذهبی، هنجارهای اجتماعی و یا شرایط اقتصادی در این نظریه مورد بررسی قرار گرفته‌اند. در سال ۲۰۱۱ مدل پنتاگون تقلب مطرح گردید که در واقع مدل بسط داده شده مثلث است و توسط کروهورات معرفی شد. این مدل دو عامل غرور و توانایی را نیز به مدل اضافه نمود که بر این اساس غرور به معنای محق دانستن خود می‌تواند دلیلی بر طمع و ارتکاب تقلب باشد. با بررسی اجمالی نظریه های مطرح در سطح جهان، نظریه پردازان در این حوزه همچون گبگی و آدبیزی (۲۰۱۳)^۴، به این موضوع اشاره می‌کنند که حساب‌رسان و یا حسابداران بایستی بر تمامی نظریه اشراف داشته باشند تا با شناخت عوامل موثر بتوانند اثر هر یک را شناسایی و مورد بررسی قرار دهند. بنابراین در طول زمان همواره تقلب و متغیرهای اثرگذار بر آن بسیار حایز اهمیت بوده‌اند و نظریه پردازان بسیاری به آنها پرداخته‌اند.

ب- هوش مصنوعی (AI)

هوش مصنوعی به عنوان یکی از پیشرفته‌ترین فناوری‌های امروزی در جهان در حال ظهور است. در حقیقت هوش مصنوعی بر مبنای یک سیستم عصبی مصنوعی کار می‌کند که از طریق فناوری‌هایی همچون یادگیری ماشین عصر تازه‌ای از تجزیه و تحلیل داده‌ها پدیدار می‌شود که می‌تواند باعث تقویت و

هوشمند، بسیاری از این روش‌های جایگزین متد سنتی حسابداری شدند. به طوری که در حال حاضر بسیاری از جوامع پیشرفته با هدف کاهش خطای انسانی و همچنین پیش‌بینی تقلب در بکارگیری آن‌ها پیشرو هستند. نرم افزارهای مختلف حسابرسی و به دنبال آن داده کاوی انقلاب بزرگی را در زمینه شناسایی و حتی پیش‌بینی تقلب‌های مالی در جهان به پا کردند که به دنبال آن هوش مصنوعی^۱ مبتنی بر کدها و الگوریتم‌های مختلف به تکامل آن کمک شایانی نمود. هوش مصنوعی شاخه‌ای گسترده از علوم کامپیوتر است که ماشین‌های هوشمندی می‌سازد که از رفتارهای انسانی تقلید می‌کنند در واقع، این ماشین‌ها تصمیماتی می‌گیرند که معمولاً نیازمند سطحی از تجربه‌ی انسانی است. و به انسان‌ها کمک می‌کند تا مشکلات احتمالی را پیش‌بینی کند و بر آن‌ها فائق آید. لذا هدف مطالعه حاضر پیش‌بینی تقلب در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در دوره زمانی ۱۳۹۵ تا ۱۴۰۱ با تکیه بر روش ماشین بردار پشتیبان است. از این روی با استفاده از متغیرهای مالی مندرج در صورت‌های مالی شرکت‌های منتخب و با تعاریف پیش فرض از تقلب، به پیش‌بینی تقلب خواهد پرداخت و به دنبال بررسی ضریب خطا و قابلیت مدل مورد بررسی در زمینه پیش‌بینی تقلب است. بر این اساس پس از بیان مقدمه در بخش دوم مبانی نظری و پیشینه تحقیق مورد بررسی قرار گرفته‌اند. پس از آن مدل و متغیرهای مورد استفاده در تحقیق بررسی شده‌اند و در نهایت نتایج حاصل از مدل آورده شده است.

۲- مبانی نظری و پیشینه تحقیق

۱-۲ مبانی نظری

الف- نظریه های مرتبط با تقلب

در خصوص مبحث تقلب بنا به اهمیت آن در طول زمان نظریه های مختلفی مطرح شده است و عوامل مختلفی به عنوان متغیر اثرگذار بر تقلب مورد بررسی قرار گرفته است.

تقلب شامل اقدامات عمدی است و کسی مرتکب آن می‌شود که از فریب، نیرنگ و حیله‌گری استفاده کرده است. اینگونه اقدامات را در دوگونه از ظاهر سازی می‌توان طبقه‌بندی کرد: اظهار کذب و پوشیده داشتن حقیقت. تقلب شامل فریب، آز، ریسک دستگیری، تخطی از اعتماد، توجیه عقلی و مواردی از این دست می‌شود. تقلب اقدامات نامناسبی است که به اظهار نظر نادرست اما با اهمیت در صورت‌های مالی منجر می‌شود. اشکال مختلفی از تقلب از قبیل صورت‌های

³ Ramamoorti

⁴ Gbegi & Adebisi

¹ artificial intelligence

² Cressey

علیزادگان و همکاران (۱۴۰۱)، تأثیر تیپ شخصیتی و اخلاق حرفه‌ای بر توانایی حسابرسان در کشف تقلب با استفاده از تئوری رفتار برنامه ریزی شده با نقش تردید حرفه‌ای را مورد مطالعه قرار دادند. بر این اساس از جامعه آماری حسابرسان شاغل در سازمان حسابرسی و مؤسسات حسابرسی عضو جامعه حسابداران رسمی ایران استفاده شده که بر اساس جدول مورگان ۳۰۲ نفر به عنوان نمونه تحقیق انتخاب شده‌اند. در این تحقیق برای ارزیابی ارتباط بین متغیرهای هدف از رویکرد حداقل مربعات جزئی بهره گرفته شده است. نتایج بیانگر اینست که انواع تیپ‌های شخصیتی حسابرسان و اخلاق حرفه‌ای آنها و همچنین تردید حرفه‌ای بر تقلب رابطه مثبت و معناداری برجای دارد.

کامرانی و عابدینی (۱۴۰۱)، به تدوین مدل کشف تقلب صورتهای مالی با استفاده از روش های شبکه عصبی مصنوعی و ماشین بردار پشتیبانی در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداختند. این مطالعه در دوره زمانی ۱۳۹۷-۱۳۹۲ با تکیه بر روش‌های الگوریتم ماشین بردار پشتیبانی نتایج تحقیق نشان میدهد که در پیش بینی الگوریتم ماشین بردار پشتیبان حدود ۶۸ درصد و در آزمون حدود ۶۱ درصد بوده است. همچنین قدرت پیش بینی الگوریتم شبکه عصبی در بخش آموزش ۶۷ درصد و در آزمون ۱۶ درصد بوده است.

نجفی و همکاران (۱۴۰۱) به ارائه الگویی برای کشف تقلب مالیاتی بر اساس تیپ‌های شخصیتی مدیران مالی شرکت‌ها با استفاده از رویکرد شبکه‌های عصبی پرداختند. جامعه آماری این تحقیق شامل کلیه مدیران مالی شرکت‌های بورسی و غیربورسی در سال ۱۳۹۹ که معاف از مالیات نیستند، می‌باشد. همچنین این پژوهش از طریق پرسشنامه به جمع‌آوری داده پرداخته و با استفاده از نرم افزار spss به تخمین داده‌ها پرداخته است. نتایج گویای اینست که الگوی شبکه عصبی طراحی شده با ۱۰ نرون در لایه پنهان دارای دقت ۷۹.۵ در کشف تقلب مدیران می‌باشد. همچنین طبق نتایج حاصل از آزمون مدل رگرسیونی، تیپ‌های شخصیتی روان رنجوری، برون‌گرایی، انعطاف پذیری و توافق پذیری بر تقلب مالیاتی تأثیر مثبت و معنی‌دار و تیپ شخصیتی باوجدان بودن تأثیر منفی و معنی‌دار بر تقلب مالیاتی دارد.

دهدار و مرادی (۱۴۰۰)، با بهره‌گیری از رویکرد داده کاوی به شناسایی تقلب در صورتهای مالی پرداختند. بر این اساس به معرفی روش‌های درخت تصمیم‌گیری، شبکه‌های عصبی و شبکه‌های گمانه زنی بیزی پرداختند. نتایج نشان می‌دهد که در راستای تحقیقات پیشین این روش‌ها قابلیت شناسایی تحریف‌های موجود در صورتهای مالی را دارا می‌باشند و کاربرد آن‌ها توصیه می‌شود.

بازتعریف رویکرد تمامی متخصصان از جمله در حوزه در حسابداری، امور مالی و حسابرسی شود. در واقع هوش مصنوعی بر پایه داده‌های موجود رفتارهای موجود در آن حوزه را مورد بررسی قرار می‌دهد و قاعدتا رفتاری که خارج از حیطه موجود باشد را شناسایی می‌نماید. به بیانی دیگر با ادغام انبوهی از اطلاعات، رفتار متقلب را شناسایی می‌نماید که بر همین اساس ابزاری قدرتمند در راستای پیش‌بینی تقلب نیز خواهد بود. هوش مصنوعی دارای کاربردهای زیادی در حسابداری و حسابرسی است که از مهم‌ترین آنها می‌توان به کاهش خطای انسانی، ساده سازی روش‌ها، امنیت و دقت حسابرسی پیشرفته و همچنین حسابرسی پیوسته اشاره نمود که مجموعه این ویژگی‌ها منجر به استفاده روز افزون هوش مصنوعی در این حوزه شده و مطالعات بسیاری نیز در سال‌های به بررسی آن در حوزه‌های مختلف علم حسابداری اختصاص داده شده است.

۲-۲ مطالعات داخلی

برنا و همکاران (۱۴۰۱)، به ارائه الگوی عوامل موثر بر وقوع تقلب در صورتهای مالی با رویکرد حسابداری دادگاهی: بر اساس روش تحلیل مضمون پرداختند. در این پژوهش در دوره زمانی ۱۴۰۰-۱۳۹۹ تقلب در برخی از شرکت‌های متقلب ایرانی مورد بررسی قرار گرفته اند که این امر با تکیه بر روش پرسشنامه‌ای صورت گرفته است. نتایج گویای اینست که علاوه بر عوامل فردی، تقلب معلول عوامل ساختاری و کلان برون سازی است پرشکوه و همکاران (۱۴۰۱)، رابطه کیفیت حسابرسی، ثبات و هدف مالی با امکان تقلب در صورتهای مالی را مورد بررسی قرار دادند. این تحقیق در جامعه آماری شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران و در دوره زمانی ۱۳۹۷-۱۳۹۲ و با تکیه بر رویکرد رگرسیون لاجیت انجام شده است. نتایج نشان می‌دهد که بین کیفیت حسابرسی و امکان تقلب رابطه معنی و معناداری وجود دارد. بر این اساس با افزایش کیفیت حسابرسی می‌توان امکان ارتکاب تقلب کاهش می‌یابد.

عربیان اصل (۱۴۰۱) تأثیر تکنولوژی هوش مصنوعی بر حسابداری پالایشگاه‌های کشور با تأکید بر مدیریت کاهش هزینه‌های حسابداری و تقلب مطالعه موردی: پالایش گاز فجر جم را مورد بررسی قرار دادند. این تحقیق از نوع پرسشنامه‌ای بوده که در جامعه آماری کلیه کارکنان شرکت پالایش گاز فجر توزیع شده است. نتایج با تکیه بر نرم افزار SPSS مورد تجزیه و تحلیل قرار گرفتند و گویای این هستند که تکنولوژی هوش مصنوعی بر کاهش هزینه‌های حسابداری پالایشگاهی تأثیر مثبت و معناداری داشته و در کاهش تقلب، نیز موثر است.

اینست که الگوی XGBoost در این مطالعه خروجی بهتری را نسبت به سایرین داشته است. که در واقع الگوریتم‌های بهینه XGBoost بیانگر دقت ۹۶.۵٪ در پیش‌بینی سطح تقلب هستند. کورلیوسیک و کارگر^۲ (۲۰۲۳)، به مرور مطالعات مرتبط با پیش‌بینی متغیرهای مختلف در حسابداری مالی با استفاده از هوش مصنوعی پرداخت. بر این اساس نویسندگان، به بررسی ۴۷ مقاله مرتبط را بر اساس جستجوی وب را انتخاب و مورد مطالعه قرار دادند. مطالعات صورت گرفته بر مبنای حجم نمونه، روش مورد استفاده و دوره زمانی طبقه‌بندی شدند. خروجی تحقیق معرفی چندین الگوریتم مبتنی بر یادگیری ماشین است که به نتایج دقیق تری (دقت ضریب پیش‌بینی^۳) منتج می‌شوند و همچنین تایید کاربرد و کارایی این الگوها در علم حسابداری است. ژائو و همکاران^۴ (۲۰۲۲)، به شناسایی و پیش‌بینی تقلب با بکارگیری روش SMOTE و الگوریتم‌های یادگیری ماشین در شرکت‌های منتخب پرداختند. در این مطالعه ۳۶۲ متغیر مالی با استفاده از انواع روش‌های مبتنی بر هوش مصنوعی در شرکت‌های منتخب مورد بررسی قرار گرفتند. نتایج این تحقیق نشان می‌دهد الگوریتم XGBOOST از قدرت پیش‌بینی و شناسایی بالایی برخوردار با ضریب دقت ۹۹ درصد برخوردار است. خاکساری و همکاران^۵ (۲۰۲۲)، در مطالعه‌ای تحت عنوان توسعه مدلی برای پیش‌بینی گزارشگری مالی متقلبان، با استفاده از رهیافت بنیاد و اسپاتیش و رگرسیون لجستیک به پیش‌بینی تقلب در صورت‌های مالی پرداختند. بر این اساس ۹۹ شرکت فعال در بورس اوراق بهادار تهران در بازه ۲۰۰۹-۲۰۱۹ مورد بررسی قرار گرفتند. نتایج نشان می‌دهد که نتایج مدل‌های بکار رفته در تحقیق برای پیش‌بینی تقلب کارا در نمونه آماری مناسب نیستند اما مدل تعدیل شده آنها دارای دقت ۷۲ درصد و ۶۴ درصد هستند.

کازمی و پیری^۶ (۲۰۲۲)، به پیش‌بینی طرح تقلب گزارشگری مالی از طریق رویکرد یادگیری ماشین - طبقه بندی چند کلاسه پرداختند. بر این اساس داده‌های ۱۳۴ شرکت پذیرفته شده در بورس اوراق بهادار تهران در دوره زمانی ۲۰۰۹ تا ۲۰۲۱ بر اساس مدل لجستیک و با بکارگیری یادگیری ماشین مورد بررسی قرار گرفته است. نتایج بیانگر کارایی این روش در پیش‌بینی تقلب است.

دهقان و همکاران^۷ (۲۰۲۱)، به پیش‌بینی تقلب در صورت‌های مالی با استفاده از تکنیک یادگیری ماشین در کشور امارات

ابراهیمیان و همکاران^۱ (۱۳۹۹) در مطالعه‌ای تحت عنوان ارائه مدلی به منظور پیش‌بینی احتمال تقلب در صورت‌های مالی با استفاده از رگرسیون لجستیک ۸۳ شرکت را در دوره زمانی ۱۳۹۶-۱۳۹۰ مورد بررسی قرار دادند. در این تحقیق پس از طبقه‌بندی شرکت‌های متقلبان و غیر متقلبان به تخمین مدل بر اساس رگرسیون لجستیک پرداختند. نتایج نشان می‌دهد که با استفاده از نسبت‌های سود ناخالص، تقسیم برفروش، سود ناخالص تقسیم برکل دارایی‌ها، سود قبل از مالیات دارایی‌های ثابت تقسیم بر بدهی‌های جاری، سرمایه در گردش تقسیم برکل دارایی‌ها و برکل دارایی‌ها می‌توان احتمال تقلب در صورت‌های مالی را برآورد نمود.

رضایی و همکاران^۱ (۱۳۹۹)، در مطالعه ای تحت عنوان کشف تقلب صورت‌های مالی با توجه به گزارش حسابرسی صورت‌های مالی، در دوره زمانی ۱۳۹۳ تا ۱۳۹۶، در نمونه ای به حجم ۱۶۴ شرکت و از طریق روش نمونه‌گیری حذف سیستماتیک، به این نتیجه‌گیری دست یافتند که تمامی تکنیک‌ها قابلیت کشف تقلب صورت‌های مالی را در سطح نسبتا بالایی دارند و تکنیک پیشنهادی ترکیبی با میزان نرخ پیش‌بینی ۹۶.۲٪ دارای دقت و توان ارزیابی بالاتری نسبت به سایر تکنیک‌ها است.

تاراسی و همکاران^۱ (۱۳۹۸)، با استفاده از تکنیک شبکه‌های عصبی به کشف تقلب صورت‌های مالی پرداختند. نتایج این پژوهش نشان داد که تکنیک شبکه عصبی و مدل سازی مبتنی بر شبکه عصبی از صحت ۹۷.۴ درصد برخوردار است و با طراحی و آموزش دقیق، می‌توان شبکه عصبی طراحی نمود که با دقت قابل قبولی بتواند احتمال گزارشگری مالی متقلبان شرکت را کشف و پیش‌بینی نمایند.

۲-۳ مطالعات خارجی

الی و همکاران^۱ (۲۰۲۳)، به ارائه یک مدل قدرتمند برای پیش‌بینی تقلب در صورت‌های مالی شرکت‌ها بر اساس تکنیک یادگیری بهینه شده XGBoost پرداختند. در این مطالعه چندین روش مبتنی بر هوش مصنوعی با هدف پیش‌بینی تقلب در صورت‌های مالی شرکت‌های عمومی منتخب در منطقه منا (خاورمیانه و شمال آفریقا) مورد آزمون قرار گرفت. برای بررسی داده‌ها نامتقارن نیز از الگوریتم‌های SMOTE استفاده شده همچنین با استفاده از زبان برنامه نویسی پایتون روش‌های متعدد دیگری به منظور پیش‌بینی تقلب استفاده شده است. نتایج گویای

⁵ Khaksari and et al

⁶ Kazemi.T and Piri

⁷ Dehghan.A and Khedr.A

¹ Al Ali. H and et al

² Kureljusic and Karger

³ prediction accuracy

⁴ Zhao and et al

۳- مدل تحقیق

ماشین بردار پشتیبان

یکی از روش‌های طبقه بندی نظارت شده که در سال‌های اخیر مطرح شده است روش SVM است. اساس این دیدگاه بر مبنای شناخت الگوی آماری است. SVM دارای قابلیت‌های زیر است:

- ✓ ریسک ساختاری را کاهش می‌دهد.
 - ✓ بردارهای نزدیک به مرزها را مشخص می‌کند.
 - ✓ از فوق صفحه با ابعاد زیاد برای طبقه بندی غیرخطی استفاده می‌کند
 - ✓ از کرنل‌های خطی و غیرخطی استفاده می‌کند.
- این روش برای فشرده‌سازی داده‌ها بسیار مناسب است و قابلیت انعطاف زیادی متناسب با مسئله دارد؛ اما از معایب این روش می‌توان به این نکته اشاره کرد که داده‌ها دور افتاده می‌تواند مرزها را خراب کنند. بنابراین داده‌های یادگیری باید تمیز باشند تا نقاط مرزی براساس همسایگی‌های بین کلاس‌ها باشد نه اندازه‌گیری‌های اشتباه و نویزی. همچنین انتخاب کرنل‌ها در SVM قانون خاصی ندارد و تابع قانون کلی است.

تفاوت این روش با شبکه عصبی در این است که شبکه عصبی با خط داده‌ها را از هم جدا می‌کند در صورتیکه در اینجا می‌توان مرزها را با توابع غیر خطی بیان کرد. ایده اصلی در SVM کاهش ریسک است. اگر داده‌های ورودی به ماشین x ها باشند ماشین وظیفه دارد که به هر داده یک خروجی y نسبت دهد. در واقعیت مسئله y خروجی مطلوب است و ماشین از روی ورودی $f(x, \alpha)$ را می‌سازد که این تابع با تغییر متغیر α قابل تنظیم است. برای مثال در شبکه عصبی با ساختار مشخص، α وزن‌ها و بایاس‌های نورون‌ها می‌باشد که باید تنظیم شوند. خطای یادگیری ماشین یا ریسک به صورت رابطه زیر است:

$$R(\alpha) = \int 1/2 |y - f(x, \alpha)| p(x, y) dx dy \quad (1)$$

که $p(x, y)$ تابع چگالی احتمال توام ورودی و خروجی است. در حالت عملی فقط تعداد محدودی از ورودی‌ها و خروجی‌ها برای یادگیری استفاده می‌شوند. پس:

$$R(\alpha)_{emp} = \int 1/2l \sum_{i=1}^l |y_i - f(x_i, \alpha)| \quad (2)$$

که l تعداد داده‌ها هستند. نامساوی زیر را با احتمال $0 \leq \eta < 1$ برای ریسک ارائه شده است:

$$R(\alpha) \leq R_{emp}(\alpha) + \sqrt{\frac{h \left(\log \left(\frac{2l}{h} \right) - \log \left(\frac{\eta}{4} \right) \right)}{l}} \quad (3)$$

پرداختند. در این پژوهش سعی شده از روش‌های متفاوتی همچون رگرسیون لجستیک، درخت تصمیم‌گیری، شبکه‌های عصبی و... به بررسی متغیر هدف در دوره زمانی ۲۰۱۰ تا ۲۰۱۸ بپردازد. نتایج این تحقیق بیانگر اینست که رهیافت SVM با ضریب دقت ۸۹.۵۴ درصد بیشترین توان پیش‌بینی تقلب را دارا می‌باشد. علی و همکاران^۱ (۲۰۲۲)، در پژوهشی به بررسی مطالعات صورت گرفته شده مبنی بر شناسایی تقلب بر اساس یادگیری ماشین پرداختند. در این تحقیق ۹۳ مقاله مورد بررسی قرار گرفته است. نتایج نشان می‌دهد که یادگیری ماشین و شبکه عصبی دارای کاربرد گسترده‌ای در مطالعات انجام شده دارند و همچنین دارای خروجی قابل‌اتکایی هستند که می‌تواند از آنها برای شناسایی تقلب در صورت‌های مالی بهره‌جست.

بنانی و همکاران^۲ (۲۰۲۱)، با استفاده از تکنیک‌های یادگیری ماشین به پیش‌بینی تقلب در صورت‌های مالی مطالعه موردی امارات متحده عربی پرداختند. در این پژوهش با استفاده از دایتا بیس بلومبرگ و Osiris و بکارگیری چند روش هوش مصنوعی و شبکه‌های عصبی به پیش‌بینی تقلب در شرکت‌های منتخب در دوره زمانی ۲۰۱۰ تا ۲۰۱۸ پرداختند. نتایج این تحقیق گویای اینست که روش ML از قدرت پیش‌بینی مناسبی برخوردار است.

جان^۳ (۲۰۲۱)، با تکیه بر روش یادگیری عمیق^۴ به شناسایی تقلب در صورت‌های مالی برای توسعه پایدار بازار سرمایه تحت اطلاعات نامتقارن پرداختند. این تحقیق در دوره زمانی ۲۰۱۹-۲۰۰۱ اطلاعات مالی و غیر مالی ۱۵۳ شرکت فعال در بورس اوراق بهادار با استفاده از الگوریتم‌های شبکه‌های عصبی همچون RNN و LSTM مورد بررسی قرار گرفتند. نتایج نشان می‌دهد که مدل LSTM دارای خروجی بهتری با ضریب دقت ۹۴.۸۸ درصد است.

چی‌چن و همکاران (۲۰۱۵)، به تجزیه و تحلیل ۴۲ تفاوت‌های میان تکنیک‌های داده‌کاوی و خبره‌سنجی در کشف تقلب صورت‌های مالی پرداختند. روش‌های تجزیه و تحلیل داده‌ها به کاربرده شده در این پژوهش شامل رگرسیون منطقی، CART، شبکه‌های ANNs بوده است. نتایج این پژوهش بیانگر کارایی و توانایی بالتر تکنیک‌های داده‌کاوی نسبت به خبره‌سنجی است.

³ Jan⁴ Deep Learning¹ Ali and et al² Bannany and et al

i از ۱ تا تعداد داده و v از ۱ تا بعد داده تغییر می‌کند.

$$\frac{\partial L}{\partial b} = -\sum_i \alpha_i y_i = 0 \quad (3-9)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha - \mu_i$$

$$\mu_i(w \cdot \Phi(x_i) + b) - 1 + \xi_i \geq 0, \xi_i \geq 0, \alpha_i \geq 0$$

$$\alpha_i \{y_i (w \cdot \Phi(x_i) + b) - 1 + \xi_i\} = 0,$$

شکل قبل حالتی که کلاس‌ها از هم صورت خطی تفکیک نشوند را نشان داده است. در این حالت بردارهایی به عنوان خطا وجود دارد. در حالتی Φ تابعی غیر خطی باشد آنگاه svm غیر خطی خواهد بود. از جمله کرنل‌هایی که در SVM به کار برده می‌شوند، می‌توان به کرنل خطی، کرنل چندجمله‌ای، RBF و سیگموئید اشاره کرد. هر یک از این کرنل‌ها بسته به مسئله تفکیک‌پذیری مورد نظر می‌توانند نقاط قوت و ضعف خود را داشته باشند که می‌بایست آن‌ها را در آزمایش بر روی داده‌ها تعیین کرد. این کرنل‌های متداول عبارتند از:

$$K(x, y) = (x \cdot y + 1)^p \quad (3-2)$$

$$K(x, y) = \exp(-\|x - y\|^2 / 2\sigma^2)$$

$$K(x, y) = \tanh(\beta x \cdot y - \lambda)$$

۴- متغیرهای تحقیق

متغیرهای که در این تحقیق مورد استفاده قرار گرفتند به شرح زیر هستند:

نرخ رشد دارایی‌های ثابت، نرخ رشد کل دارایی‌ها، نرخ رشد حساب‌های دریافتی، رشد بهای تمام شده کالای فروش رفته، رشد فروش، رشد استهلاک، رشد بدهی‌های بلندمدت، رشد موجودی وجوه نقد، رشد سود خالص، رشد جریان نقدی عملیاتی، رشد کل بدهی‌ها، رشد دارایی‌های جاری، رشد بدهی‌های جاری، رشد سود ناخالص، رشد سود (زیان) عملیاتی، رشد حقوق صاحبان سهام، رشد موجودی مواد و کالا، رشد سرمایه، رشد سود و زیان انباشته. همچنین برچسب‌گذاری به منظور تعلیم مدل از گزارش حساب‌برسان مبنی بر مقبول، مردود، مشروط و عدم اظهارنظر استفاده شده است. لازم به ذکر است که تمامی متغیرهای بکار گرفته در مدل از سایت کدال^۱ و همچنین بورس اوراق بهادار ایران^۲ استخراج شده‌اند.

که h عدد صحیح غیرمنفی است و بعد (VC) نامیده می‌شود. این بعد ظرفیت تفکیک‌کنندگی تابع f را نشان می‌دهد. برای مثال برای یک خط مستقیم حداکثر سه نقطه را در صفحه می‌توان از هم جدا کرد (VC=3). هدف ماشین کاهش سمت راست نامعادله است.

در SVM خط مرزهای بین کلاس‌ها با خط ایجاد می‌شود. فرض کنید دو کلاس وجود دارد $(y_i \in \{+1, -1\})$ آنگاه باید همه داده‌ها در شرایط زیر صدق کنند:

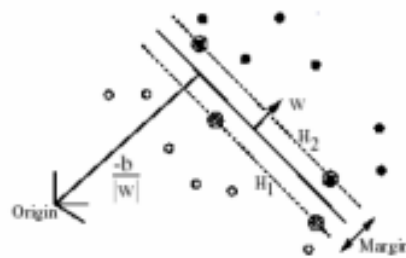
$$\begin{aligned} x_i \cdot w + b &\geq +1 \text{ for } y_i = +1 \\ x_i \cdot w + b &\leq -1 \text{ for } y_i = -1 \end{aligned} \quad (3-4)$$

یا به عبارت کلی تر:

$$y_i(x_i \cdot w + b) - 1 \geq 0 \quad \forall i \quad (5)$$

که در شکل ۳ - ۱ حالت تفکیک‌پذیری کلاس‌ها نشان داده شده‌اند. در حالت کلی مسئله، ممکن است کلاس‌ها از هم جدا نباشند بنابراین:

$$\begin{aligned} y_i \left((w \cdot \Phi(x_i)) + b \right) \\ \geq 1 - \xi_i \quad \xi_i \geq 0 \quad i \\ = 1, \dots, l \end{aligned} \quad (6)$$



شکل ۳ - ۱ جداسازی خطی بین کلاس‌ها

که Φ نگاشتی است که بر ورودی اعمال می‌شود. حال باید رابطه زیر حداقل شود. بنابراین از لاگرانژ استفاده می‌کنیم.

$$L(w, b, \alpha, \mu) = 1/2 \|w\|^2 + C \sum_i \xi_i - \sum_{i=1}^l \alpha_i (y_i (w \cdot \Phi(x_i) + b) - 1 + \xi_i) - \sum_i \mu_i \xi_i$$

که C ضریبی است که به استفاده‌کننده داده می‌شود و مقدار بزرگ آن نشان‌دهنده جریمه بزرگ خطا است. یکی از روش‌های بهینه‌کردن رابطه بالا استفاده از شرایط Karush-Kuhn-Tucker است.

$$\frac{\partial L}{\partial w_v} = w_v - \sum_i \alpha_i y_i x_{iv} = 0 \quad (8)$$

² <https://www.tsetmc.com/>

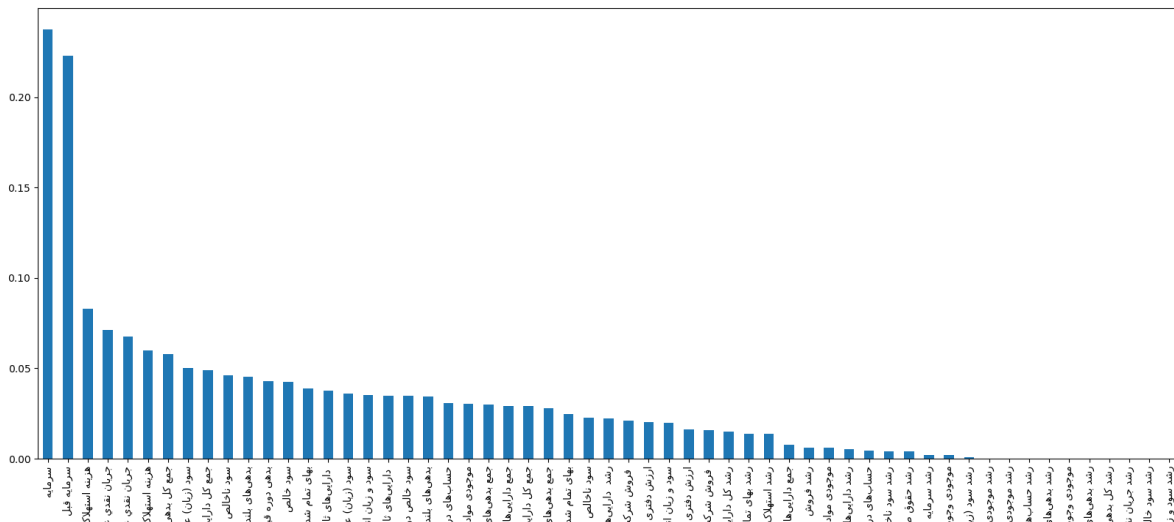
¹ <https://codal.ir/>

۵- یافته های تحقیق

عملکرد بهتر، جلوگیری از بیش برآزش و نیز سرعت اجرای برنامه ، از روش برآورد اطلاعات متقابل^۱ استفاده شده است. اطلاعات متقابل (MI) بین دو متغیر تصادفی یک مقدار غیر منفی است که وابستگی بین متغیرها را اندازه گیری می کند. اگر و فقط اگر دو متغیر تصادفی مستقل باشند برابر با صفر است و مقادیر بالاتر به معنای وابستگی بیشتر است. این تابع به روش های ناپارامتری مبتنی بر تخمین آنتروپی از فاصله k نزدیک ترین همسایه ها متکی است. شکل زیر بیانگر میزان اهمیت متغیرها براساس روش انتخابی و اطلاعات به دست آمده از داده ها بوده که به صورت نزولی مرتب شده اند.

پس از مشخص شدن میزان اهمیت داده ها و به منظور کاهش ابعاد و در نتیجه سرعت بیشتر و کارایی روش انتخابی ۱۵ مورد از بهترین ویژگی ها مورد استفاده قرار گرفته اند. در نهایت ویژگی های مورد استفاده عبارتند از: دارایی های ثابت دوره قبل، جمع کل دارایی های دوره قبل، بهای تمام شده کالای فروش رفته دوره قبل، هزینه استهلاک، هزینه استهلاک دوره قبل، بدهی های بلندمدت دوره قبل، سودخالص، جریان نقدی عملیاتی، جریان نقدی عملیاتی دوره قبل، جمع کل بدهی ها، بدهی دوره قبل، سود ناخالص قبل، سود(زیان) عملیاتی قبل، سرمایه، سرمایه قبل.

این مطالعه به دنبال پیش بینی تقلب در صورت های مالی شرکت های پذیرفته شده در بورس اوراق بهادار تهران می باشد. لذا برای انتخاب نمونه آماری از میان شرکت های پذیرفته شده در بورس اوراق بهادار تهران شرکتهایی انتخاب شده اند که از یکسو دارای فعالیت های تولیدی داشته باشند و از سوی دیگر شرکت های منتخب بایستی دارای سال مالی منتهی به اسفند ماه باشند که در نهایت ۱۳۶ شرکت مورد بررسی قرار گرفتند. در این راستا از داده های سالانه بورس اوراق بهادار تهران موجود در سامانه کدال در بازه زمانی ۱۳۹۵ - ۱۴۰۱ استفاده شده است. به منظور استفاده از روش ماشین بردار با هدف ارایه الگوی پیش بینی تقلب از متغیر های مربوط به صورت های مالی و همچنین نسبت های مالی شرکت ها به صورت نرخ رشد استفاده شده است. از سویی بر اساس گزارش حسابرسان مبنی بر مقبول، مشروط، عدم تایید و عدم اظهار نظر، داده ها در مدل برچسب گذاری شده اند. با توجه به پراکندگی داده های موجود و در راستای دستیابی به نتایج دقیقتر در ابتدا متغیرهای با اهمیت انتخاب شده اند. در واقع انتخاب ویژگی های مناسب نقش بسیار مهمی در عملکرد و کارایی روش انتخابی دارد. داده های مورد استفاده در این تحقیق دارای ۵۷ ویژگی هستند که به منظور



نمودار ۱- ضریب اهمیت ویژگی

منبع: یافته های تحقیق

جدول ۱- نتایج خروجی مدل ماشین بردار پشتیبان

صحت	دقت	بازخوانی	امتیاز - F
۸۳٪	۸۵٪	۹۷٪	۹۰٪

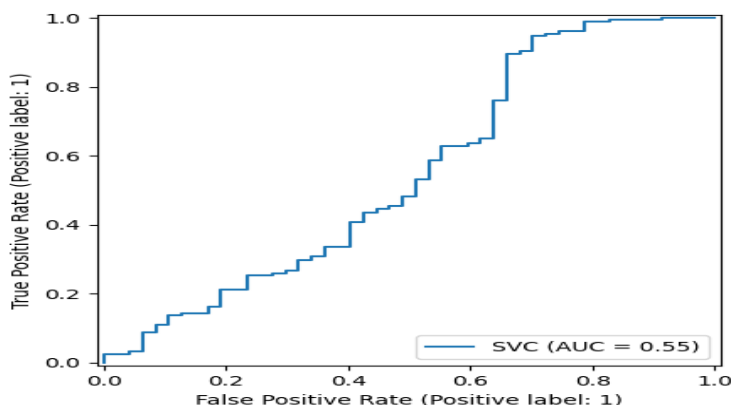
بر این اساس نتایج حاصل از مدل ماشین بردار پشتیبان (SVM) به شرح زیر است:

^۱ Mutual information

ان است که در صورت نامتقلب بودن، این خروجی به میزان ۹۷ درصد به درستی در دسته غیرمقلبانه قرار گرفته است. امتیاز F نیز بیانگر آن است که این آزمایش به میزان ۹۰ درصد قابل اتکا بوده است. منحنی راک ROC Curve که از آن با نام منحنی مشخصه عملکرد سیستم Receiver Operating Characteristics نام برده می‌شود، از ابزارهای تشخیصی عملکرد یک سیستم به حساب می‌آید که براساس دسته بندی مثبت صحیح و مثبت کاذب عمل کرده و بیانگر قرار تشخیص مدل در نواحی مطلوب یا نامطلوب می‌باشد. نمودار ROC زیر بیانگر آن است که مدل پس از آموزش به خوبی توانسته است داده‌ها را در دسته بندی متقلبانه قرار دهد.

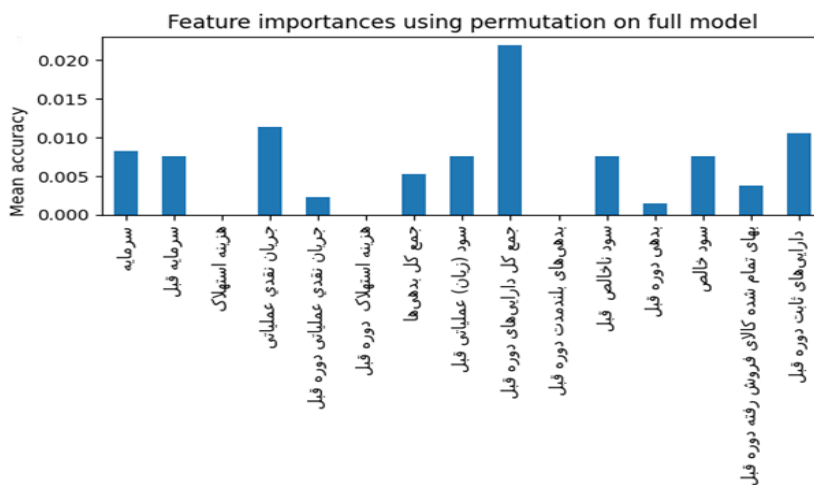
نمودار عملکرد مدل روی داده های آموزشی و تست به صورت زیر است:

جدول فوق بیانگر این است که مدل آموزش دیده با صحت ۸۳ درصد می‌تواند روی داده های واقعی عملکرد صحیح داشته و داده ها را دسته بندی کند. در اینجا این بدان معناست که مدل توانسته است با صحت ۸۳ داده را در صورت تقلب در دسته تقلب و در صورت متقلبانه بودن در دسته نامتقلبانه قرار دهد. همچنین دارای ضریب دقت ۸۵ درصد می‌باشد که این بدان معناست که مدل توانسته در صورت تشخیص درست متقلبانه بودن، این کار با دقت ۸۵ درصد انجام شده است. چنانچه میزان دقت پایین باشد بیانگر آن است که مدل نسبت زیادی از تشخیص های متقلب بودن را در دسته اشتباه قرار داده است. که در اینجا به این معنی است که در صورت تقلب، به میزان ۸۵ درصد به درستی در دسته خود قرار گرفته است. همچنین بازخوانی به مفهوم این است که مدل با نرخ بازخوانی ۹۷ درصد توانسته است نتیجه درست را تشخیص دهد. این میزان به معنای



نمودار ۲- عملکرد مدل

منبع: یافته های تحقیق



نمودار ۳- ضریب اهمیت ویژگی ها

منبع: یافته های تحقیق

همان طور که نمودار نشان میدهد متغیرهای جمع کل دارایی‌های دوره قبل، همچنین جریان نقدی عملیاتی، سود ناخالص دوره قبل و سود خالص و همچنین سرمایه دوره قبل و جاری بیشترین میزان اثرگذاری بر تقلب را دارا هستند.

۶- نتیجه گیری

در این مطالعه هدف ارائه الگویی مبتنی بر هوش مصنوعی به منظور کشف و پیش بینی تقلب در صورت‌های مالی است. بدین منظور، جامعه آماری شرکت‌های پذیرفته شده منتخب در بورس اوراق بهادار در دوره زمانی ۱۳۹۵ تا ۱۴۰۱ مورد بررسی قرار گرفته است. نتایج بیانگر اینست که مدل ماشین بردار پشتیبان (SVM) با ضریب دقت ۸۵ درصد، ضریب بازخوانی ۹۷ درصد و صحت ۸۳ درصد از کارایی قابل قبولی در جهت پیش‌بینی تقلب در شرکت‌های مورد بررسی برخوردار است. مطالعاتی همچون های چی چن و همکاران (۲۰۱۵)، تاراسی و همکاران (۲۰۱۷) و رضایی و همکاران (۱۳۹۹)، کامرانی و عابدینی (۱۴۰۱) نیز کارایی ماشین بردار پشتیبان را مورد آزمون و تایید نموده اند و نتایج این تحقیق نیز با این مطالعات تطابق دارد. از سویی نتایج بیانگر اینست که متغیرهای مالی تاثیر بالایی بر ارتکاب تقلب دارند و معناداری آنها تایید شده که این نکته با مطالعه خواجهی و ابراهیمی (۱۳۹۶) نیز مطابقت دارد. از میان متغیرهای بررسی شده در مدل جمع کل دارایی‌های دوره قبل، همچنین جریان نقدی عملیاتی، سود ناخالص دوره قبل و سود خالص و همچنین سرمایه دوره قبل و جاری از میان متغیرهای مورد بررسی در مدل بیشترین تاثیر را بر تقلب داشته‌اند که این نتایج تایید دیگری بر صحت خروجی بدست آمده از طریق روش ماشین بردار پشتیبان است. در واقع بر اساس تعاریف مرتبط با تعریف انجمن بازرسان رسمی، بیش‌نمایی دارایی‌ها و درآمدها یا کم‌نمایی بدهی‌ها سهم بالایی از وقوع تقلب در شرکت‌های مختلف را دارا هستند بنابراین تغییر در سطح دارایی‌ها نسبت به دوره قبل و یا جریان نقدی می‌تواند سیگنال مهمی مبنی بر ارتکاب تقلب تعبیر شود.

فهرست منابع

ابراهیمیان محمد رضا، ایزدی نیا ناصر و امیری هادی (۱۳۹۹) در مطالعه‌ای تحت عنوان ارائه مدلی به منظور پیش‌بینی احتمال تقلب در صورت‌های مالی با استفاده از رگرسیون لجستیک، دانش حسابرسی، سال بیستم، شماره ۷۹.

برنا محمدرضا، برادران حسن زاده رسول، فضل زاده علیرضا و بادآور نهندی یونس (۱۴۰۱)، الگوی عوامل موثر بر وقوع

تقلب در صورت‌های مالی با رویکرد حسابداری دادگاهی: بر اساس روش تحلیل مضمون (تم)، پژوهش‌های کاربردی در گزارشگری مالی، دوره یازدهم، شماره ۲۱.

تاراسی، مجتبی، بنی طالبی دهکردی، بهاره، زمانی، بهزاد، (۱۳۹۸)، "پیش‌بینی گزارشگری مالی متقلبان از طریق شبکه عصبی مصنوعی (ANN)، فصلنامه علمی پژوهشی حسابداری مدیریت، ۱۲، ۶۳-۷۹.

خواجهی شکراله و ابراهیمی مهرداد (۱۳۹۶)، بررسی تأثیر متغیرهای حسابرسی بر احتمال وقوع تقلب در صورت‌های مالی: شواهدی از شرکتهای پذیرفته شده در بورس اوراق بهادار تهران، دانش حسابرسی، سال هفدهم، شماره ۶۸.

دهدار فرهاد و مرادی محمد حسین (۱۴۰۰)، بهره‌گیری از رویکرد داده کاوی و شناسایی تقلب در صورت‌های مالی، فصلنامه چشم انداز حسابداری و مدیریت، دوره ۴، شماره ۴۵.

رسولی پر شکوه سجاد، حاجی هاشمی ورنوسفادانی منصوره و شهری مریم (۱۴۰۱)، رابطه کیفیت حسابرسی، ثبات و هدف مالی با امکان تقلب در صورتهای مالی، فصلنامه رویکردهای پژوهشی نوین در مدیریت و حسابداری، سال ششم، شماره ۸۴.

رضایی مهدی، ناظمی اردکانی مهدی و ناصر صدر آبادی علیرضا (۱۳۹۹)، پیش‌بینی تقلب صورتهای مالی با استفاده از رویکرد کریسپ (CRISP)، فصلنامه علمی پژوهشی دانش حسابداری و حسابرسی مدیریت سال دهم / شماره ۰۴ / زمستان ۰۰۴۴.

عربیان اصل (۱۴۰۱) تاثیر تکنولوژی هوش مصنوعی بر حسابداری پالایشگاه‌های کشور با تاکید بر مدیریت کاهش هزینه‌های حسابداری و تقلب مطالعه موردی: پالایش گاز فجر جم، پژوهش‌های کاربردی در مدیریت و حسابداری، شماره ۲۵، بهار ۱۴۰۱، ص ۵۰-۴۰.

علیزادگان لیلا، صمدی لرگانی محمود و ایمنی محسن (۱۴۰۱)، تأثیر تیپ شخصیتی و اخلاق حرفه‌ای بر توانایی حسابرسان در کشف تقلب با استفاده از تئوری رفتار برنامه ریزی شده با نقش تردید حرفه‌ای، پژوهش‌های حسابداری مالی و حسابرسی، دوره ۱۴، شماره ۲.

کامرانی حسین، بیژن عابدینی (۱۴۰۱)، تدوین مدل کشف تقلب صورتهای مالی با استفاده از روش‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبانی در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، فصلنامه

- Economics Empirical Research in Accounting, Winter 2022, V. 12, No. 46, pp.255-276.
- Khaksari.I, Shoorvarzi .M, Mehrazeen.A and Massihabadi.A(2022), Developing a model to predict fraudulent financial reporting, Int. J. Nonlinear Anal. Appl. In Press, 1–13 ISSN: 2008-6822 (electronic)
<http://dx.doi.org/10.22075/ijnaa.2022.27975.377>
3.
- Kureljusic.M and Karger.E (2023), Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda, Journal of Applied Accounting Research, 10.1108/JAAR-06-2022-0146.
- Ramamoorti, S., Morrison, D., & Koletare, J. W. (2009). Bringing fraud to fraud: understanding the state of- mind of the C-Level suite/white collar offender through “A-B-C” analysis. Working Paper, Institute for Fraud Prevention
- Zhao, Z.; Bai, T. Financial Fraud Detection and Prediction in Listed Companies Using SMOTE and Machine Learning Algorithms. Entropy 2022, 24, 1157. <https://doi.org/10.3390/e24081157>.
- علمی پژوهشی دانش حسابداری و حسابرسی مدیریت، سال یازدهم، شماره ۴۱.
- مرادی، مهدی و مجتبی سلیمانی مارشک (۱۳۹۰). روانشناسی و جامعه شناسی تقلب. فصلنامه دانش حسابرسی، شماره ۴۲، بهار، ۱۳۹۰ صص ۵۶-۷۱
- نجفی سمیه، صالحی اله کرم و امیری هوشنگ(۱۴۰۱)، ارائه الگویی برای کشف تقلب مالیاتی بر اساس تیپ‌های شخصیتی مدیران مالی شرکتها با استفاده از رویکرد شبکه‌های عصبی، پژوهشنامه مالیات، شماره پنجاه و سه.
- Al Ali. A, M. Khedr. A, El-Bannany. M, Sakeena Kanakkayil. S (2023), A Powerful Predicting Model for Financial Statement Fraud Based on Optimized XGBoost Ensemble Learning Technique, Learning Technique. Appl. Sci. 2023, 13, 2272. <https://doi.org/10.3390/app13042272>.
- Ali, A.; Abd Razak, S.; Othman, S.H.; Eisa, T.A.E.; Al-Dhaqm, A.; Nasser, M.; Elhassan, T.; Elshafie, H.; Saif, A. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. Appl. Sci. 2022, 12, 9637. <https://doi.org/10.3390/app12199637>.
- Bannany.M, Dehghan.A, Khedr.A, Prediction of Financial Statement Fraud using Machine Learning Techniques in UAE, 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD).
- Chi-Chen, Lin, An-An Chiu, Shaio Yan Huang, David C. Yen, (2015),” Detecting The Financial Statement Fraud: The Analysis Of The Differences Between Data Mining Techniques And Experts Judgments”, Knowledge-Based Systems, 89(15), Pp.459-470.
- Cressey, D. R. (1950). The criminal violation of financial trust. American Sociological Review. 15(6): 738-743.
- Dehghan.A and Khedr.A, Prediction of Financial Statement Fraud using Machine Learning Techniques in UAE, [2021 18th International Multi-Conference on Systems, Signals & Devices \(SSD\)](https://doi.org/10.3390/app12199637).
- Gbegi, D. O., & Adebisi, J. F. (2013). The new fraud diamond model—how can it help forensic accountants in fraud investigation in Nigeria. European Journal of Accounting Auditing and Fiancé Research, 1(4), 129-138.
- Jan.C-L (2021), Detection of Financial Statement Fraud Using Deep Learning for Sustainable Development of Capital Markets under Information Asymmetry, Sustainability 2021, 13, 9879. <https://doi.org/10.3390/su13179879>.
- Kazemi.T and Piri.P(2022), Financial Reporting Fraud Scheme Prediction via Machine Learning Approach – Multiclass Classification, Alzahra University, Faculty of Social Sciences and



Accounting Knowledge & Management Auditing
Vol. 15/ No. 59/ Autumn 2026

Presenting a fraud prediction model based on artificial intelligence (support vector machine (SVM) model)

Ali Zarei

PhD student, Accounting Department, Emirates Branch, Islamic Azad University

Fraydoon Rahnamay Roodposhti

Full Professor, Accounting Department, Science and Research Unit, Islamic Azad University, Tehran, Iran

Hamed Khanmohammadi

Assistant Professor of Accounting Department, Damavand Branch, Islamic Azad University of Damavand, Iran

Hamidreza Kordaloi

Associate Professor, Department of Accounting, Islamshahr Branch, Islamic Azad University, Tehran, Iran

Abstract

In today's world, along with the advancement of technology, the application of data science and artificial intelligence in all fields, including accounting, is undeniable. Artificial intelligence has been widely used in accounting science in various topics, including internal accounting process, reports related to income and expenses, and even detection and prediction of fraud, which actually results in higher speed and reliability as a result of eliminating human error are. Therefore, in this research, the aim is to investigate the effectiveness of the support vector machine method in predicting fraud. In this regard, in the period from 2015 to 2015, the number of 136 companies admitted to the stock exchange was examined in this way. The results show that this model with an accuracy factor of 85%, a readability factor of 97% and an accuracy of 83% has an acceptable efficiency in predicting fraud. On the other hand, its output shows that the total assets of the previous period, as well as operating cash flow, gross profit of the previous period and net profit, as well as the capital of the previous and current period, had an effect on committing fraud.

Keywords: Fraud, stock exchange companies, support vector machine, artificial intelligence