فصلنامه علمی پژوهشی
دانش حسابداری و حسابرسی مدیریت
سال یازدهم / شماره ۴۱/ بهار ۱۴۰۱

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 بهادر تهران

حسین کامرانی
دانشجوی دکتری ،گروه حسابداری،واحد قشم ،دانشگاه آزاد اسلامی،قشم،ایران.
Kamranihossein66@gmail.com

بیژن عابدینی
گروه حسابداری، دانشکده حسابداری و مدیریت، دانشگاه هرمزگان، بندر عباس، ایران
(نویسنده مسئول)
Bizhanabedini.48@gmail.com

تاریخ دریافت: ۹۸/۱۱/۳۰ تاریخ پذیرش: ۹۹/۰۱/۰۵

چکیده

امروزه با توجه به گسترش روزافزون بازارهای مالی و نیاز به جلب سرمایه‌گذاران داخلی و خارجی یکی از مهم‌ترین دغدغه‌های هر شرکت در راستای تأمین منابع مالی افشای صحیح اطلاعات مالی است.هدف از این پژوهش 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 بهادر تهران بود. پژوهش حاضر از نوع پژوهشهای کاربردی و تجربی- همبستگی است. جامعه آماری شامل شرکتهای پذیرفته شده در بورس اوراق بهادر تهران در دوره زمانی ۱۳۹۷-۱۳۹۲است. تحقیقات اخیر مشخص کرده است که سرمایه‌گذاران در فرایند تصمیم‌گیری، شرکت‌هایی را انتخاب می‌کنند که سود آن‌ها از پایداری بالاتر و درواقع از کیفیت بالاتری برخوردار باشد. نتایج نشان داد که در بخش آموزش قدرت پیش بینی الگوریتم ماشین بردار پشتیبان حدود ۸۶ درصد و در آزمون حدود ۸۲ درصد بوده است. همچنین قدرت پیش بینی الگوریتم شبکه عصبی در بخش آموزش ۸۱ درصد و در آزمون ۷۸ درصد بوده است.

واژه‌های کلیدی: تقلب، صورتهای مالی ، شبکه عصبی مصنوعی، ماشین بردار پشتیبان.

۱- مقدمه

تقلب و فساد مالی از دیرباز در فرهنگ‌ها، سامانه‌های قوانین، اصول اخلاقی و ادبیات بشر مذموم و عملی مجرمانه دانسته شده است و انواع رفتارها و مجازات‌های فردی و اجتماعی به صورت مختلف مثل مدارا، طرد، حبس و مجازات های مشابه نتوانسته‌اند این پدیده شوم را، که ناشی از آزمندی انسان‌ها است، ریشه‌کن کند. با این وجود، سازمان‌ها باید در تلاش دائمی برای پیش‌گیری، بازدارندگی و کشف تقلب باشند. سازمانی که از تلاش برای مبارزه با تقلب دست بکشد یا خطر تقلب را بی‌اهمیت بداند بی‌تردید در معرض خطر تقلب بیش تری قرار دارد. کشف‌نکردن تقلب به دلیل پایین‌بودن آن و هم چنین گزارش‌نکردن تقلب و غیره به معنی نبود تقلب نیست بروز بحران‌های مالی متعدد از دهه ۱۹۳۰میلادی تا ورشکستگی شرکت‌هایی چون انرون و وردکام و وقوع تقلب‌های بزرگ در سایر شرکت‌ها در دهه ۲۰۰۰ همگان را برآن داشت تا در ابعاد مختلف مبارزه با تقلب از قبیل پیش‌گیری، کنترل و نظارت تلاش کنند اعمال ووضع قوانینی همچون قانون فساد خارجی۱۹۷۷و قانون ساربنز آکسلیدر آمریکا و به دنبال آن هم سویی برخی کشورها در پیش‌گیری قبل از درمان در زمینه تقلبدر قالب قوانین مشابه، از یک سو مسئولیت مدیران در برابر مدیریت خطر تقلب را افزایش داده و از سوی دیگر موجب ایجاد محیطی شده است که در کنار قوانین و مقررات، مدیران شرکت‌ها و سازمان‌ها خود نیز به دنبال تدوین و اجرای راهبردهایی برای پیش‌گیری و کشف تقلب و سوء استفاده‌های مالی برآیند طبق مثالی رایج در حسابرسی: بیش‌تر تقلب‌ها به صورت تصادفی کشف می‌شود نه با حسابرسی و طراحی سیستم حسابداری. با این حال، کتاب‌های بسیاری درباره حسابرسی تقلب نوشته شده است که به بررسی تقلب پس از وقوع آن پرداخته‌اند. به عبارتی دیگر، حسابرس با استفاده از روش‌های معمول و برحسب تصادف تقلب وقوع یافته‌ای را کشف و وجود آن را تأیید و اثبات می‌کند حسابرسان نقش مهمی در حفاظت از سرمایه‌گذاران در برابر دریافت اطلاعات مالی اشتباه، ناقص و گمراه‌کننده ایفا می‌کنند. هم‌چنین، اگر هوشیاری و شک و تردید حسابرسان بیش‌تر بود، می‌توانست از برخی شکست‌های حسابرسی اخیر جلوگیری کنند (رضایی و کرامبلی، ۲۰۰۷)در گستره اقتصاد جهانی، تقلب تأثیر به سزا و قابل تأکیدی بر گزارش حسابرسان دارد. بنابراین، همواره یکی از بزرگ‌ترین چالش‌های پیش‌روی پژوهشگران فعال در حوزه تقلب این بوده است.

در حال حاضر با توجه به نیاز روز افزون مدیران بهمنظور کسب اطلاعات مالی صحیح، برای اخذ تصمیمات مدیریتی در افق بلند مدت شرکت و نیاز به جلب سرمایه گذاران داخلی و خارجی برای تامین سرمایه و رقابت بر سر این موضوع، اطلاعات ارائهشده درباره وضعیت مالی و عملکرد یک شرکت، اهمیت زیادی برای استفادهکنندگان از صورتهای مالی دارد زیرا همانطور که گفته شد صورتهای مالی برای ارزیابی عملکرد شرکت، سلامت مالی و نظارت بر مدیریت ضروری است و مبنایی برای تصمیمگیری و تخصیص سرمایه است و اطلاعات مالی مفید برای تخصیص منابع کارآمد و تصمیم گیریهای اقتصادی آگاهانه بسیار حائز اهمیت است .بنابراین گزارشات مالی باید صحیح، با دقت و به موقع ارائه شوند تا بتوانند اطلاعات مفیدی را درباره عملکرد مالی و غیر مالی سازمانها در اختیار استفاده کنندگان درون سازمانی و برون سازمانی قرار دهد و به ارائه درست و صادقانه دستاوردهای مالی شرکت منجر خواهد شد و نقش مهمی در پایداری سیستم مالی خواهد داشت.

در سالهای اخیر، موضوع تقلب در صورتهای مالی توجه زیادی را به خود جلب کرده است و تقلب در گزارشات مالی رو به افزایش است. تقلب صورتهای مالی با عناوینی همچون تقلب حسابداری، تقلب مدیریت، و یا گزارش مالی تحریف شده شناخته میشود. این اتفاق زمانی رخ می دهد که صورتهای مالی (گزارشگری) حاوی اطلاعات غلط و یا نادیده گرفتن حقایق مالی (میزان، افشا یا شواهد) برای فریب کاربران باشد. صورتهای مالی متقلبانه اثرات منفی زیادی بر اقتصاد دنیا داشته و به خسارات قابل توجهی برای افراد و شرکتها منجر شده است. شواهد بیانگر این مطب می باشد که در حال حاضر تقلب در صورتهای مالی وجود دارد و این تقلبها برای دنیای کسب و کار بسیار شایع و پرهزینه است

اشکال مختلف تقلب در حوزه های مالی بر اساس گزارش انجمن بازرسان رسمی تقلب عبارتنداز: فساد مالی، سوء استفاده از دارایی ها و تقلب در گزارشگری مالی.

از طرفی تقلب صورتهای مالی تقلب تاثیر قابل ملاحظه ای در مالیات شرکت ها نسبت به سایر تقلبها دارد . بنابراین تقلب پیامدهای ویرانگری برای آینده شرکت، مدیران، کارکنان، حسابرسان، سرمایه گذاران و کل جامعه در پی دارد. تقلب های بزرگ موجب از پا در آمدن واحدهای اقتصادی یا شرکتها شده و زیانهای قابل ملاحظه ای را برای سرمایه گذاران در پی دارد، همچنین هزینه های حقوقی قابل توجهی را تحمیل می نماید، موجب زندانی شدن افراد کلیدی میشود و میتواند باعث از بین رفتن اعتماد در بازار سرمایه میشوند. در نهایت هزینه تقلب در قالب افزایش قیمت کالا و یا ارائه خدمات بر تمام اعضای جامعه تاثیر خواهد گذاشت و از طرفی قیمت سهام شرکت

کاهش مییابد و در نهایت زمانی که تقلب حسابداری آشکار میشود، تاثیر منفی قابل توجهی بر شهرت، نام تجاری و اعتبار شرکت بر جای می گذارد. بدلیل اهمیت موضوع در استانداردهای بین المللی حسابداری و حسابرسی نیز به آن اشاره شده است. بنابراین، بررسی علل ریشه ای عمل تقلب بسیار حایز اهمیت میباشد. از این رو در یک محیط تجاری نیاز مبرمی به روشهای موثر برای پیشگیری و کشف تقلب در صورتهای مالی وجود دارد.به دلیل ضرورت این مساله در پژوهش حاضر به دنبال طراحی الگویی بهمنظور پیش بینی ریسک تقلب صورتهای مالی می باشد.تلاش این پژوهش بر آن است که عواملی را که باعث بروز تقلب در صورتهای مالی میشوند را شناسایی و رتبه بندی کند تا از این طریق کمک شایانی به حسابرسان در کشف تقلب و جلوگیری از وقوع مجدد آن نماید تا به این ترتیب اعتماد را به سرمایه گذاران بازگرداند و موجب افزایش سرمایه گذاری و رونق اقتصادی در جامعه شود.

۲- مبانی نظری و پیشینه پژوهش

۲-۱- تعریف تقلب

در ابتدا نیاز به یک تعریف روشن و دقیق از موضوع تقلب و توضیحی مدل مفهومی که باعث بینش عمیق نسبت به موضوع میشود، لازم به نظر می رسد. به طورکلی تعاریف متعددی در مورد تقلب بیان شده است و اجماع نظری درباره تعریف تقلب وجود ندارد. جوامع حرفهای حسابداران نیز در این مورد، تعاریفی را برای تقلب برشمردند.

استاندارد حسابرسی استرالیا در سال ۲۰۰۴ (AUS، ۲۱۰) تقلب را عمل عمدی یک یا چند نفر از جمله مدیریت، کارکنان و یا اشخاص ثالث، با استفاده از فریب برای بهدست آوردن مزیت غیر عادلانه و غیر قانونی تعریف میکنند

انجمن حسابداران امریکا نیز در بیانیه حسابرسی شماره ۹۹ در اکتبر ۲۰۰۲ سه ویژگی برای تقلب در نظر گرفته و اعتقاد دارد این سه عامل موجب بروز تقلب هستند که به مثلث تقلب معروف هستند Rasha and Andrew۲۰۱۲.

طبق بند ۴ بخش ۲۴۰ استانداردهای حسابرسی ایران، تقلب هرگونه اقدام عمدی و فریب کارانه یک یا چند نفر از مدیران، کارکنان یا اشخاص ثالث، برای برخورداری از یک مزیت غیرقانونی میباشد (کمیته فنی سازمان حسابرسی، ۱۳۸۸). این نکته قابل توجه است که تفاوت بین تقلب و اشتباه در این است که تقلب همراه با قصد و نیت است ولی اشتباهات ناشی از اعمال غیر عمدی است

اکثر مردم از تقلب به عنوان یک عمل بد یاد میکنند. بهطور کلی تقلب نوعی کلاهبرداری و فریبکاری است که باعث ایجاد برتری ناعادلانه متقلب بر دیگران میشود. یعنی یک طرف قرارداد از طرف دیگر اطلاعات بیشتری در مورد موضوع معامله دارد.

انجمن حسابداران خبره نیجریه (۲۰۰۶) تقلب را عمل عمدی یک یا چند نفر شامل مدیریت، کارکنان یا اشخاص ثالث که منجر به ارائه اطلاعات نادرست از صورتهای مالی شود، بیان نموده است.

تعریف تقلب از نگاه انجمن حسابداران رسمی امریکا عبارتند از تحریف عمدی حقیقت به منظور اجبار فردی دیگر در اتکا بر آن، جهت دادن چیز با ارزشی که به وی تعلق دارد یا واگذاری حق قانونی بازنمایی کذب حقیقت امر، خواه با واژگان یا با رفتارها یا ادعایی کذب یا گمراه کننده، یا با پنهان سازی آن چه باید افشا می شده است که گمراه می کند و با این انگیزه صورت می گیرد که فرد دیگری را به نحوی گمراه کند که از نظر حقوقی به ضرر خود عمل کند. در اجرای عملیات حسابرسی گفته می شود که ارائه های نادرست از دو نوع عمل ناشی می شود:

- گزارشگری مالی متقلبانه که ارائه گزارش های مالی نادرست به شکل عمدی به منظور گمراه کردن استفاده کنندگان.
- سوءاستفاده از دارایی ها که به شکل سرقت دارایی های شرکت بوده و اثر آن به خوبی در گزارش های مالی منعکس نشده است (استانداردهای حسابرسی شماره ۹۹ آمریکا،بند ۶)

انجمن حسابرسان داخلی آمریکا، تقلب را به این صورت توصیف می کند: هر گونه عمل غیرقانونی که ویژگی هایی از قبیل فریب، پنهان سازی یا نقض اعتماد دارد. این اعمال به استفاده از تهدید خشونت آمیز یا نیروی فیزیکی وابسته نیست. تقلب به وسیله گروه ها و سازمان ها اجرا می شود، تا پول، اموال، یا خدمات به دست آورند، از پرداخت پول یا از دست دادن خدمات اجتناب کنند یا منافع فردی یا سازمانی را تضمین کنند (استانداردهای حسابرسی داخلی، ۲۰۱۳). همچنین انجمن حسابرسان داخلی، تقلب و رفتار نادرست را به گروه های: ریسک گزارشگری مالی، ریسک عملیاتی و ریسک رعایت تقسیم بندی می کند.

در تعریف دیگری توسط انجمن بازرسان رسمی تقلب (امریکا)، تقلب عبارتند از هر گونه عمل غیرقانونی که دارای یکی از ویژگیهای نیرنگ، خیانت در امانت و پنهانکاری باشد که این اعمال نیازمند کاربرد خشونت و اجبار فیزیکی ندارد. به

عبارتی دیگر تقلب از سوی افراد و سازمان ها برای به دست آوردن پول، دارایی، یا خدمت انجام می شود تا از پرداخت وجوه یا از دست دادن خدمات خودداری کنند؛ یا نفع تجاری یا فردی به دست آورند (گلدمن و کافمن، ۲۰۰۹). بنابراین از دیدگاه این انجمن تمام رویدهای غیر مترقبه، کلک، حیلهگری یا مخفیکاری و... ، راههای غیر منصفانه برای فریب دیگران است (پوریا نسب، ۱۳۸۴).

برینک و ویت (۱۹۸۲) تقلب را تهدیدی که هدف آن سو استفاده از منابع می باشد تعریف کرده اند و آن را به عنوان نگرانی اصلی مدیریت مطرح نمودند.

ادریبجی و دادا (۲۰۰۷) تقلب را فریب عمدی به قصد محروم کردن شخص دیگری از اموال و حقوقش و نیز برنامه ریزی و اجرای آن به صورت مستقیم و غیر مستقیم تعریف نموده است.

آبرشت و همکاران (۱۹۹۵) و هوارد (۲۰۰۵) نیز تقلب را به اختلاس کارمند، تقلب مدیریت، کلاهبرداری، تقلب فروشنده، تقلب مشتری و تقلب متفرقه طبقه بندی کرده است.

میرناد والاس نیز تقلب را برنامه ای از پیش طراحی شده برای فریب دیگران از طریق ارائه اسناد و مدارک ساختگی تعریف نموده است. از آن جایی که تقلب اقدامی عمدی است، تقلب در صورتهای مالی ساده نیست و معلول عوامل متعددی است و به شیوههای متفاوتی نیز اعمال میشود.

رضایی (۲۰۰۵) تقلب در صورتهای مالی را شامل هر گونه اقدام عمدی توسط شرکت برای گمراه کردن استفاده کنندگان از این صورتهای مالی به ویژه سرمایهگذاران و اعتبار دهنگان از طریق تهیه و انتشار صورتهای مالی حاوی اشتباه با اهمیت بیان میکند (رضایی، ۲۰۰۵).

۲-۲- نظریه‌های تقلب

ماهیت و محتوای رفتار کج روانه از گذشته تاکنون و از یک جامعه تا جامعه دیگر بسیار متفاوت است اما این نظریه‌ها از بعضی جنبه‌های نظری با یکدیگر هم پوشانی داشته و از جهاتی می توانند با یکدیگر ترکیب شوند تا درک معقولی از جنبه های عمده رفتار کج روانه فراهم سازند (گیدنز، ۱۳۸۴).

حسابرسان به شناسایی تخلفاتی که منجر به مخدوش شدن گزارشگری مالی شود علاقه زیادی از خود نشان می دهند. نظریه ها و الگوهای کشف تقلب، با وجود این که هر کدام به تنهایی تضمین کننده کشف تقلب در حسابرسی نیستند، اما می توانند مکمل خوبی برای حسابرس در کشف تقلب و علت شناسی آن باشند (مهرانی و حصارزاده، ۱۳۸۷).

نظریه های تقلب را می توان از دو منظر و دیدگاه مورد بررسی قرار داد:

- نظریه های پیش گیری و اصلاح (جرم شناسی)
- نظریه های کشف تقلب (حسابرسی).

نظریه های غالب در مورد پیش گیری و اصلاح (جرم شناسی) عبارتند از:

- نظریه جرم شناسی کلاسیک
- نظریه زیست شناختی جرم
- نظریه های روان شناختی
- نظریه جامعه شناختی

نظریه جرم شناسی کلاسیک: جرم شناسی کلاسیک که مبتنی بر اصول فلسفی سودمندگرایی است، ریشه در این باور دارد که نوع بشر موجوداتی منطقی و حساب گر هستند. در نتیجه، اقداماتی را انجام می دهند تا از درد دوری جویند و خوشنودی را در خود ایجاد کنند. نام دو نظریه پرداز (بکارا از ایتالیا و بنتام از انگلستان) با جرم شناسی کلاسیک درآمیخته است (پوریانسب و احمدی، ۱۳۹۱).

نظریه زیست شناختی جرم: نخستین کوشش ها برای تبیین جرم و شکل های دیگر شناسایی جرم اساسا مدعی کشف ویژگی هایی در جمجمه جنایت کاران شد که آنها را از افراد پیرو قانون متمایز می ساخت ، زیر بنای نظریه زیست شناختی به لومبروسو گیدنز پزشک ایتالیایی برمی گردد. این نظریه، جنبه های تشریحی، فیزیولوژی، تشخیص بیماری و بیوشیمی شخص بزهکار را مطالعه می کند (نجفی، ۱۳۸۹).

نظریه های روان شناختی جرم، مانند تبیین های زیست شناختی، تبهکاری را با انواع خاص شخصیت مربوط می دانند. بنا بر نظر فروید، تا اندازه زیادی حس اخلاقی ما ناشی از خودداری هایی است که در کودکی در طی مرحله رشد یاد می گیریم. گفته می شود که روان رنجوران، افرادی کناره گیر و بی احساس هستند که از خشونت به خاطر خود آن لذت می برند (گیدنز، ۱۳۸۴).

نظریه جامعه شناختی: این نظریه بیان می دارد که ماهیت جرم به نهادهای اجتماعی جامعه بستگی دارد. یکی از مهم ترین جنبه های تفکر جامعه شناختی درباره جرم، تأکید بر ارتباطات متقابل میان هم نوایی و کج روی است که از نظر متفکرین این سبک زمینه های اجتماعی مختلف است (گیدنز، ۱۳۸۴). دورکیم، عوامل اجتماعی و تأثیر محیط بر زندگی انسان ها و سوق دادن آنها به سوی بزهکاری را دارای اهمیت فراوانی می دانست (نجفی توانا، ۱۳۸۹). نظریه های غالب در مورد کشف تقلب عبارتند از: نظریه قله کوه یخی و نظریه

سوسک حمام در نظریه قله کوه یخی، تقلب مانند کوه یخی است که روی آب قرار گرفته، وجوه ساختاری بالای سطح آب و در دید همگان قرار دارد اما جنبه های رفتاری زیر سطح آب پنهان است و برای حسابرس آسوده خیال خطر دارد و مانند یخ نهفته در آب در کمین کشتی نشینان است.

عوامل ساختاری شامل: سلسله مراتب، منابع مالی، هدف های سازمان، مهارت ها و توانایی های کارکنان، وضعیت فن آوری، استانداردهای عملکرد و اندازه گیری کارایی.

عوامل رفتاری (جنبه های نهفته) شامل: نگرش ها، احساسات مانند ترس و خشم و...، ارزش ها، هنجارها، هم کنش، پشتیبانی و خرسندی (تیموری، ۱۳۸۶) مطرح شد. این نظریه بر تکرارپذیری تقلب تأکید دارد.

۲-۳- انواع تقلب

اشکال مختلف تقلب در حوزه های مالی بر اساس گزارش انجمن بازرسان رسمی تقلب عبارتند از: فساد مالی، سوء استفاده از دارایی ها و تقلب در گزارشگری مالی (انجمن بازرسان رسمی تقلب، ۲۰۰۸).

- **فساد مالی:** به عنوان تقلبی تعریف میشود که در آن متقلبان از نفوذ خود در یک تراکنش مالی به منظور کارپردازی برای منافع شخصی خود یا دیگری به غلط استفاده میکنند. مانند پذیرش پورسانت و درگیر شدن در تضاد منافع، رشوه خواری، باجگیری و زورگیری اقتصادی.
- **سواستفاده از داراییها:** شامل اختلاس در داراییها که اغلب به عنوان تقلب کارکنان شناخته میشود، میباشد که به وسیله یک شخص و یا گروهی از افراد بهمنظور کسب منافع مالی غیر منصفانه انجام میشود (رامنی و استن بارت، ۲۰۰۶).
- **تقلب گزارشگری مالی:** عبارت از تحریف عمدی در نتایج صورتهای مالی بمنظور ارائه تصویر نادرست از شرکت مانند بیشنمایی داراییها و کمنمایی هزینهها و کمنمایی داراییهای دزدیده شده، می باشد فریبکاری از طریق سند سازی و دستکاری یا تغییر سوابق حسابداری یا مدارک پشتوانه تهیه صورتهای مالی و ارائه نادرست یا حذف عمدی رویدادها، معاملات یا سایر اطلاعات با اهمیت در صورتهای مالی صورت می پذیرد.(وکیلی فرد و همکاران، ۱۳۸۸).

هم چنین، می توان تقلب را بر حسب مفاهیم سازمانی که در آن رخ می دهد، به دو گروه تقلب درون سازمانی و تقلب برون سازمانی طبقه بندی کرد.

تقلب درون سازمانی را نیز می توان به دو گروه تقلب کارکنان و تقلب مدیریت تقسیم بندی کرد. تقلب درون سازمانی کارکنان شامل اختلاس، تقلب در حقوق و دستمزد، تقلب و دستکاری چک، تقلب درصدگیری و پرداختنی، سرقت دانش فنی و اطلاعات محرمانه شرکت می باشد (گلدمن و کافمن، ۲۰۰۹). تقلب درون سازمانی مدیریت شامل سوءاستفاده ماموریت و تفریحات، پنهان کردن، گزارشگری مالی متقلبانه و تضاد منافع، اختلاس/ دزدی، گرفتن رشوه می باشد(گلدمن و کافمن، ۲۰۰۹).

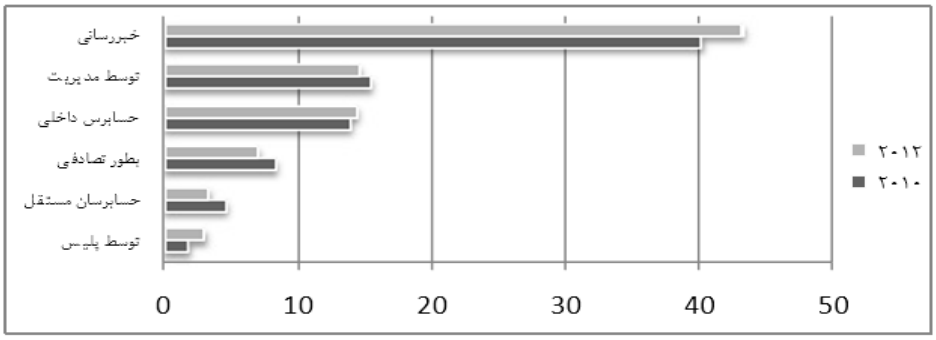
هر چند که میزان و تنوع تقلب مدیریت در مقایسه با تقلب کارکنان کم تر است، ولی تقلب مدیریت، زیان بیش تری را به شرکت تحمیل می کند. زیرا، مدیران اختیار بیش تری دارند و فرصت آنان برای فریب کاری بیش تر از کسانی است که زیردست آنان کار می کنند (گلدمن و کافمن، ۲۰۰۹). تقلب هم چنین ممکن است به دست افراد برون سازمانی مانند فروشندگان، تأمین کنندگان مواد اولیه تولید، پیمانکاران، مشاوران و مشتریان از طریق افزودن بر مبلغ صورتحساب، ارسال صورتحساب مجدد و مواد نامرغوب (بر خلاف قرارداد معامله)، ارائه اطلاعات نادرست درباره کیفیت و ارزش کالاهای خریداری شده یا وضعیت اعتبار مشتریان بر علیه سازمان انجام شود. این گروه برون سازمانی ممکن است به انجام اعمالی مانند رشوه دهی به اعضای سازمان نیز متوسل شوند (تیموری، ۱۳۸۶).

۲-۴- روشهای کشف و پیشگیری از تقلب

به مجموعه عملیات یا اقداماتی که براساس روشها یا متدهایی، سعی در کشف و شناسایی تقلبهای صورتگرفته یا در حال وقوع را دارند شناسایی تقلب گفته میشود. تمرکز بر فعالیت های ضدتقلب از کشف به پیش گیری تکامل یافته است. حرفه ای های ضدتقلب مانند حسابرسان و بازرسان همگی بر این نکته اتفاق نظر دارند که بیش تر قربانیان تقلب به سختی می توانند دارایی های به سرقت رفته در اثر تقلب را پس بگیرند، چون متقلبان هیچ گاه داراییهای بدست آورده از راه تقلب را پس انداز نمی کنند (کرانچر، ۲۰۱۳). پیش گیری از تقلب و کشف آن موضوعاتی به هم وابسته اند اما مفاهیم یکسانی نیستند. پیش گیری از تقلب شامل خط مشی ها، رویکردها، آموزش ها و ارتباطاتی است که از وقوع تقلب پیش گیری می کند در حالی که کشف تقلب بر فعالیت ها و روش

هایی تأکید دارد که بلادرنگ و با حساسیت زمانی، وقوع تقلب را تشخیص می دهد یا در می یابد که تقلبی در شرف وقوع است .

همان طور که نمودارشماره۱ نشان میدهد براساس گزارش سالانه انجمن بازرسان رسمی تقلب یکی از مهمترین راههای کشف تقلب در سازمانها، خبر رسانی از طریق ایجاد و برقراری خطوط تلفنی داغ است.



نمودار شماره۱: راههای اولیه کشف تقلب

۲-۵- توفان فکری (بارش افکار) راهی برای کشف تقلب

توفان فکری، موثرترین و احتمالاً متداولترین فرایند گروهی است. این روش را بیش از شصت سال ٌپیش، اسبورن (۱۹۵۳) در شرکتی تبلیغاتی به منظور افزایش کمیت و کیفیت ایده های تبلیغاتی طراحی کرد و چنان مورد استفاده و استقبال مردم و سازمان ها در غرب قرار گرفت که جزئی از زندگی آنها شد. توفان فکری روشی به منظور برانگیختن خلاقیت در گروه است که برای تولید تعداد زیادی ایده در مورد یک مساله به کار می رود. این واژه به زبان ساده به معنای تشویق افکار است.

در کشور ما، استاندارد حسابرسی شماره ۳۱۵ با عنوان شناخت واحد مورد رسیدگی و محیط آن و برآورد خطرهای تحریف با اهمیت، اعضای گروه حسابرسی را درباره احتمال آسیب پذیری صورت های مالی واحد مورد رسیدگی از تحریف های بااهمیت به بحث و گفت وگو ملزم کرده است. چنین بحث ها، فرصت هایی را برای اعضای با تجربه گروه حسابرسی شامل مدیر حسابرسی، به وجود می آورد تا اطلاعاتی را که بر اساس شناخت از واحد مورد رسیدگی پیدا کرده اند با اعضای گروه در میان گذارند. اعضای گروه نیز فرصت می یابند درباره خطرهای تجاری که واحد مورد رسیدگی در معرض آن قرار دارد و نحوه و احتمال آسیب پذیری صورت های مالی از تحریف های بااهمیت، تبادل اطلاعات کنند. میزان بحث ها به نقش، تجربه و نیازهای اطلاعاتی اعضای گروه حسابرسی بستگی دارد.

۲-۶- روش‌های مورد استفاده

ANN و SVM برای انتخاب متغیرهای مهم مناسب هستند، در حالی که CART، CHAID، C5.0 و QUEST برای طبقه‌بندی، پیش‌بینی و شناسایی متغیرها مناسب هستند [۳-۶]. در مرحله اول، شبکه‌های عصبی مصنوعی (ANN) و ماشین بردار پشتیبانی (SVM) برای نمایش متغیرهای مهم مورد استفاده قرار می‌گیرند. در مرحله دوم، برای ساخت مدل‌های طبقه‌بندی و مقایسه‌ها، چهار تکنیک درخت تصمیم‌گیری (CART، CHAID، C5.0 و QUEST) مورد استفاده قرار می‌گیرند.

شبکه‌های عصبی مصنوعی (ANN)

یک شبکه عصبی مصنوعی (ANN) ساختاری شبیه به نورون در مغز انسان است. این یک سیستم پردازش اطلاعات است که اعصاب بیولوژیکی را تقلید کرده و می‌تواند برای انجام پیش‌بینی‌ها چندین ورودی دریافت و ترکیب کند. شبکه عصبی مصنوعی نوعی از دستگاه هوش مصنوعی است که در آن روش ریاضی برای ایجاد توانایی در رایانه برای نتیجه‌گیری از طریق توانایی محاسبه سریع رایانه استفاده می‌شود. این عمل باید از طریق یک فرآیند یادگیری (یعنی یادگیری ماشین) به طوری که بتواند توانایی استنتاج را داشته باشد - یعنی کسی به آن می‌گوید که چه نوع شرایطی به چه نتیجه‌ای خواهد رسید. اگر نمونه‌های درست را به آن بگوئید، به درستی به شما پاسخ می‌دهد. این حتی می‌تواند نتیجه ممکن را برای نمونه هایی که قبلاً آموخته نشده بیان سازد.

واحدهایی که توسط یک ANN پردازش می‌شوند نورون هستند. آن‌ها دو هدف را دنبال می‌کنند: (۱) عبور، که در آن نقاط داده ناقص برای ورودی‌های گره تأثیر قابل توجهی در شبکه نداشته باشد؛ و (۲) یادگیری سازگار، که به تنظیمات وزن اتصالات بین گره‌ها اشاره دارد. ساختار اصلی ANN شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی است. مقدار خروجی هر عنصر پردازش به واحد پردازش دیگری منتقل و مقدار ورودی آن واحد می‌شود.

شبکه عصبی تکثیر اولیه BPNNبه طور گسترده‌ای، با پارامترهای کلیدی در لایه‌های پنهان، توابع اصلاح خطا و نسبت یادگیری مورد استفاده قرار می‌گیرد.

رفنس و همکاران 24استفاده از معادله زیر را برای تعیین تعداد لایه‌های پنهان بر اساس اصول هم‌گرایی و تعمیم ارائه می دهند:

(۱)

$$\text{No. of layer} = \sqrt{(\text{No. of input} \times \text{No. of output})}$$

تعداد لایه‌های پنهان بعدی لگاریتم طبیعی نورون‌ها در لایه قبلی است. تابع تصحیح خطا از روش تدریجی شیب گرادیان برای به حداقل رساندن تابع خطا استفاده می‌کند.ورودی هر مقدار تمرینی همراه با تعدیل وزن اتصال توسط شبکه همراه است. محدوده تنظیم و تابع خطای در همان جهت حساسیت به مقدار تغییر می‌کند. محاسبات به شرح زیر است:

(۲)

$$\Delta\omega = -\sigma\frac{\partial e}{\partial \omega}$$

نماد σ نشان‌دهنده نسبت یادگیری است که میزان تنظیم برای هر وزن را اندازه‌گیری می‌کند تابع خطا به منظور ارزیابی کیفیت یادگیری به کار می‌رود. هرچه مقدار e بیشتر باشد، خطا بزرگ‌تر و کیفیت یادگیری ضعیف‌تر می‌شود. محاسبه تابع خطا به شرح زیر است:

(۳)

$$e = \frac{1}{2}\sum (O_b - y_b)^2$$

که O_b مقدار خروجی هدف نورون خروجی bام است در لایه خروجی مقدار آموزش، و y_b مقدار خروجی نورون bام خروجی در لایه خروجی مثال آموزش است.وزن اتصال ω_{ab} بین نورون aام و نورون bام در لایه پنهان تابع خطا می‌تواند با معادله زیر از قانون زنجیره بیان شود:

(۴)

$$\frac{\partial e}{\partial \omega} = \frac{\partial e}{\partial yb}\frac{\partial y_b}{\partial net}\frac{\partial \text{net}}{\partial \omega_{ab}} = -(O_b - y_b)\cdot f'(\text{net})\cdot x_a$$

خطای نورون خروجی bام در لایه خروجی به صورت $\forall b$ مشخص می‌شود:

(۵)

$$\forall = (O_b - y_b)\cdot f'(\text{net})$$

مقدار اصلاح وزنی بین لایه خروجی و لایه مخفی به شرح زیر است:

(۶)

$$\Delta\omega_{ab} = -\sigma\frac{\partial \text{e}}{\partial \omega} = \sigma\cdot(O_b - y_b)\cdot f'(\text{net})\cdot x_a = \sigma\cdot\forall\cdot x_a$$

اصلاح ورودی برای نورون خروجی می‌تواند به صورت زیر بیان شود:

(۷)

$$\Delta\tau_b = -\sigma\frac{\partial e}{\partial \tau_b} = -\sigma\cdot\forall_b$$

معادلات (۵) - (۷) شبیه‌سازی‌ها و محاسبات را تکرار می‌کنند تا به تدریج تفاوت بین مقدار مورد نظر و مقدار تخمین زده شده توسط شبکه عصبی را کاهش دهند.

ماشین بردار پشتیبانی (SVM)

ماشین بردار پشتیبانی (SVM)، توسعه یافته توسط وپنیک [۲۵]، یک روش یادگیری هوش مصنوعی است. این تکنیک یادگیری ماشین بر اساس نظریه یادگیری آماری و به حداقل رساندن خطر ساختاری است. هدف این است که صفحه جدایش بهینه برای تقسیم دو یا چند کلاس داده با مکانیسم یادگیری با آموزش داده‌های ورودی را شناسایی کند. این یک نوع یادگیری نظارت شده برای پیش‌بینی و طبقه‌بندی آیتم‌ها در زمینه استخراج داده است.

فرض کنید که به تعداد n نقطه داده موجود در فضا وجود دارد، $\{(\bar{x}_1, c_1), (\bar{x}_2, c_2), \ldots, (\bar{x}_n, c_n)\}$ و نماد $C_1 \in \{+1, -1\}$ طبقه‌بندی را برای نقطه داده $\bar{x}_1$ نشان می‌دهد. این داده‌ها به عنوان داده‌های آموزشی برای شناسایی صفحه جدایش بهینه به صورت زیر تعریف می‌شود:

(۸)

$$\overline{w}\cdot\overline{x}-\alpha=0$$

نماد $\overline{\omega}$ حاشیه جدایش را مشخص می‌کند و α یک ثابت است. راه‌حل‌های متعددی برای $\overline{\omega}$ وجود دارد، اما $\overline{\omega}$ بهینه مقداری است که دارای حداکثر حاشیه است. معادله زیر راه‌حل مسئله بهینه‌سازی است:

(۹)

$$\begin{aligned}&\text{minimize}\,\tfrac{1}{2}\|\overline{w}\|^2\\&\text{subject to}\,c_1(\overline{w}\cdot\overline{x}-\alpha)\geq 1,1\leq i\leq n\end{aligned}$$

پس از یادگیری شبکه، $\overline{\omega}$ با حداکثر حاشیه، می‌توان با استفاده از معادله (۱۰) بر روی داده‌های آزمون که هنوز طبقه‌بندی نشده است، طبقه‌بندی $\hat{C}$ را ایجاد کند:

(۱۰)

$$\hat{C}=\begin{cases}-1,\ \text{if}\ \overline{w}\cdot\overline{x}-\alpha\ \leq -1\\+1,\ \text{if}\ \overline{w}\cdot\overline{x}-\alpha\ \geq +1\end{cases}$$

درخت تصمیم‌گیری

درخت‌های تصمیم‌گیری یکی از ساده‌ترین روش‌های یادگیری القایی است [۲۶]. آن‌ها می‌توانند متغیرهای مداوم و گسسته را پردازش کنند. ساختار درختی با حقایق و طبقه بندی‌های شناخته شده به منظور تعمیم قواعد مرتبط قضیه به وجود می‌آید. درخت‌های تصمیم‌گیری در این مقاله CART، CHAID، C5.0 و QUEST هستند که در زیر توضیح داده شده اند.

CART (درخت طبقه‌بندی و رگرسیون) یک روش تصمیم گیری دوتایی است که توسط بریمن و همکاران در سال ۱۹۸۴ انجام شده است. برای داده‌های مداوم یا داده‌های غیرپارامتری برای طبقه‌بندی استفاده می‌شود. تصمیم تقسیم شرایط بر اساس مقدار و ویژگی‌های داده‌ها، و هم‌چنین شاخص جینی است. هر بخش اطلاعات را به دو زیرمجموعه تقسیم می‌کند و برای هر زیرمجموعه برای شناسایی شرایط تقسیم بعدی تکرار می‌شود. به منظور ساخت یک ساختار درختی، داده‌ها هم‌چنان به دو زیرمجموعه تقسیم می‌شوند. فرآیند زمانی به پایان می رسد که داده‌ها دیگر قابل تقسیم نیستند.

هدف شاخص جینی جدا کردن بزرگ‌ترین گروه (اندازه گیری شده توسط تعداد مشاهدات) از دیگران در گره است. فرض کنید داده S حاوی N مقدار $C_1, C_2, \ldots, C_N$ باشد. بر اساس شرایط تقسیم v برای ویژگی A، S به $\{S_L، S_R\}$ تقسیم می‌شود. نمادهای l_i و r_i نشان‌دهنده تعداد مشاهدات متعلق به گروه C_i و به ترتیب غیرمتعلق به زیرمجموعه‌های S_L و S_R می باشند (I=1,2,…,N). اگر Cn بزرگ‌ترین رده در S باشد، محاسبه مقدار جینی مطابق معادله (۱۱) به دست می‌آید:

(۱۱)

$$Gini(A,v)=\frac{|S_L|}{|S|}\left[1-\sum_{i=1}^{n}\left(\frac{l_i}{|S_L|}\right)^2\right]+\frac{|S_R|}{|S|}\left[\sum_{i=1}^{n}\left(\frac{r_i}{|S_R|}\right)^2\right]$$

CHAID (آشکارساز تعامل اتوماتیک مربع چی) شاخه‌ای از الگوریتم درخت تصمیم است. الگوریتم CHAID، که توسط کاس در سال ۱۹۸۰ طراحی شده است، به طور عمده بر روی تست‌های مربع چی در فرآیند ساخت درخت تصمیم‌گیری تکیه می‌کند و شاخه تقسیم مطلوب با تکرار فرآیند ترکیب‌ها و تقسیم‌ها مشخص می‌شود. الگوریتم CHAID دارای مزایای خاصی در توسعه درخت‌های تصمیم‌گیری است، چرا که مقادیر انحصاری و مقدار تقسیم بر اساس اهمیت آماری را تأیید می کند. این برای بهینه‌سازی فرآیند شاخه‌ای مفید است.

C5.0 توسط کوینلان [۲۷] به عنوان بهبود ID3 توسعه داده شد. روش ID3 اشاره به کسب اطلاعات به عنوان معیار ساخت درخت‌های تصمیم‌گیری دارد و به طور معمول به علت تعداد زیادی از متغیرهای ورودی به فراگیری می‌انجامد. C5.0 از نسبت سود به جای معیارهای قبلی استفاده می‌کند. با این حال، مفهوم اساسی باقی می‌ماند. توسعه درخت تصمیم‌گیری مبتنی بر آنتروپی است، مهم نیست که چگونه ساختار درخت رشد می‌کند. محاسبات در معادلات (۱۲) و (۱۳) بیان می‌شود. فرض کنید مجموعه‌ای از داده‌ها حاوی دو دسته A و B است که a تعداد مشاهدات برای دسته A و b تعداد مشاهدات برای دسته B است. مقدار مورد انتظار از این مجموعه داده‌ها می تواند با I(A,B) بیان شود:

(۱۲)

$$I(A,B)=-\frac{a}{a+b}log_2\frac{a}{a+b}-\frac{b}{a+b}log_2\frac{b}{a+b}$$

مقدار آنتروپی هر ویژگی به همان شکل برآورد شد. نماد a_i نشان‌دهنده تعداد زیرمجموعه‌های ویژگی C در دسته A است. نماد b_i نشان‌دهنده تعداد زیرمجموعه‌های ویژگی C در دسته B است. مقدار آنتروپی ویژگی C مبتنی بر زیرمجموعه‌های ویژگی C می‌تواند به صورت زیر بیان شود:

(۱۳)

$$E(C)=\sum_{i=1}^{v}\frac{a_i+b_i}{a+b}I(a_i,b_i)$$

معادله (۱۴) برای محاسبه شاخص سود برای ویژگی C و سایر ویژگی‌ها استفاده می‌شود. ساختار درخت با ویژگی‌های با مقادیر شاخص سود بالا شروع به رشد می‌کند:

(۱۴

$$Gain(C) = I(a,b) - E(C)$$

برای حل مسائل مربوط به داشتن بیش از حد شاخه‌های تقسیم شده به علت عدم پردازش ردیف‌های مستمر داده‌ها، C5.0 نسبت سود را طبق معادلات (۱۵) و (۱۶) نرمال و بیان می‌کند:

(۱۵

$$Gain\ ratio = \frac{Gain(C)}{Split(C)}$$

(۱۶

$$Split(C) = -\sum_{i=1}^{n} \frac{S_i}{S} \times log_2\left(\frac{S_i}{S}\right)$$

در معادله (۱۶) نماد S_i تعداد زیرمجموعه‌ها را پس از دسته بندی ویژگی A نشان می‌دهد و S تعداد کل مجموعه‌های داده است. روش C5.0 مقادیر ردیف‌های مستمر داده را به دست می دهد و دوباره نسبت سود را در مقیاس‌های انفرادی محاسبه می کند. مقدار با بزرگ‌ترین سود به عنوان نقطه تقسیم برای ساختار درختی استفاده می‌شود.

QUEST (درخت آماری کارآمد سریع) فرض می‌کند که متغیر هدف برای ایجاد قوانین تقسیم‌بندی مستمر است. این الگوریتم می‌تواند به سرعت محاسبات را انجام داده و می‌تواند از احتمال تبعیض که در سایر روش‌ها دیده می‌شود، اجتناب کند. هم‌چنین برای متغیرهای توضیحی با دسته‌های متعدد مناسب است. با این حال، روش QUEST فقط می‌تواند دسته های دوتایی را مدیریت کند. اگر متغیر هدف برای قانون طبقه بندی مستمر باشد، آزمون ANOVA-F استفاده می‌شود. اگر متغیر هدف قابل دسته‌بندی باشد، آزمون مربع چی استفاده می‌شود. معیار شاخه‌بندی حداقل مقدار p است. متغیر ویژگی کوچک‌تر از سطح اهمیت α به عنوان متغیر شاخه‌بندی بهینه استفاده می‌شود. اگر متغیر کوچک‌تر از α وجود نداشته باشد، تست Levene برای انتخاب متغیر با بیش‌ترین تناقض از لحاظ همگن بودن استفاده می‌شود. در غیر این صورت، بیش‌تر نمی تواند تقسیم شود.

۳- پیشینه پژوهش

تانگوت[1] و کلکارنی (۲۰۱۵)، در پژوهشی به کشف تقلب صورتهای مالی با استفاده از روش دادهکاوی و تحلیل عملکرد پرداختند. برای این منظور از الگوریتم MLFF استقاده شد و به این نتیجه رسیدند که اطلاعات موجود در صورتهای مالی منتشر شده دارای شاخصهای متقلبانه میباشد. علاوه بر این تعداد نسبتا کمی از نسبتهای مالی، به طور گستردهای تعیینکننده نتایج طبقهبندی میباشند. همچنین آنها نشان دادند که شبکه عصبی نسبت به MLFF از دقت بالاتری برخوردار است.

آلدن[2] و همکاران (۲۰۱۲)، با نمونهای متشکل از ۴۵۸ شرکت شامل ۲۲۹ شرکت متقلب و ۲۲۹ شرکت غیرمتقلب، سودمندی طبقهبندیکنندههای مبتنی بر قواعد فازی در کشف الگوهای گزارشگری مالی متقلبانه را مورد بررسی قرار داد. در مجموع، ۳۲متغیر مالی به عنوان پیشبینیکنندههای بالقوه گزارشگری مالی متقلبانه انتخاب شدند. نتایج حاصل از طبقهبندیکنندههای مبتنی بر قواعد فازی با الگوریتم ژنتیک و الگوریتم برآورد توزیع یادگیری مارکو حاکی از سودمندی مدلهای تکاملی برای کشف گزارشگری مالی متقلبانه بود.

لین[3] و همکاران (۲۰۱۶)، با نمونهای متشکل از ۱۲۹ شرکت متقلب و ۴۴۷ شرکت غیرمتقلب به طبقهبندی و رتبهبندی عوامل تقلب طی سالهای ۲۰۱۰-۱۹۹۸ پرداختند. در این پژوهش ۳۲ عامل تقلب که مطابق نظر متخصصان برای کشف تقلب مناسب شناخته شدند بکار رفته است. نتایج حاصل از تکنیکهای دادهکاوی رگرسیون لجستیک، درخت تصمیم و شبکههای عصبی مصنوعی حاکی از دقت بیشتر روشهای درخت تصمیم و شبکههای عصبی مصنوعی نسبت به رگرسیون لجستیک است. در پایان نیز به منظور بهبود دستاوردهای پژوهش مقایسهای بین قضاوت متخصصان و تکنیکهای دادهکاوی صورت گرفته است.

فانینگ و کوگر (۱۹۹۴) از شبکه عصبی برای بسط الگویی به منظور کشف تقلب مدیریت استفاده کردند. آنان، هم چنین، نتایج حاصل از شبکه عصبی را با رگرسیون خطی و رگرسیون لجستیک مقایسه کردند. نتایج این الگو نشان داد که می توان تقلبات احتمالی صورت های مالی را با استفاده از تجزیه و تحلیل مدارک مشتری کشف کرد. هم چنین، نتایج بیانگر این بود که کیفیت الگوی شبکه عصبی از کیفیت الگوهای آماری بهتر است.

گرین و چوی (۱۹۹۷) از روش شبکه عصبی مصنوعی برا کشف تقلب صورتهای مالی استفاده نمودند که نتایج این

مطالعه نشان داد که این روش توانایی کشف تقلب را در صورتهای مالی دارند.

فانینگ وکوگر (۱۹۹۸) از شبکه عصبی مصنوعی (ANN) برای طراحی مدل کشف تقلب مدیریت استفاده نمودند. آنها مدلی را با هشت متغیر با احتمال بـــالای کشف تدوین نمودند. نتایج این پژوهش نشان داد که شبکه عصبی نتایج بهتری در مقایسه با رگرسیون لجستیک و مدلهای آماری استاندارد ارائه داده است.

بوستا و وینبرگ (۱۹۹۸) از شبکه عصبی برای تمایز میان داده های«واقعی» و «دستکاری شده» استفاده کردند. آنان با استفاده از این داده ها شش شبکه عصبی طراحی کرده و از میان شش شبکه عصبی بهترین را انتخاب کردند. شبکه عصبی طراحی شده در ۶۸ درصد موارد توانست از میان ۸۰۰ سری داده های مالی، داده های واقعی را از داده های دستکاری شده متمایز کند.

سرولو و سرولو (۱۹۹۹) از شبکه عصبی برای کشف تقلب استفاده نمودند که نتایج این مطالعه تقلب در گزارشگری مالی را پیشبینی نمود.

فروز (۲۰۰۰) نمونهای از ۴۲ گزارش مالی نادرست و ۹۰ گزارش مالی واقعی استفاده کرد. آزمون براساس روش رگرسیون لجستیک و شبکه عصبی انجام شد. نتایج تجربی نشان داد که روش شبکه عصبی بهتر از روش رگرسیون لجستیک است.

لین و همکاران (۲۰۰۳) از شبکههای مصنوعی، شبکههای مبتنی بر قواعد فازی و رگرسیون لجستیک گام به گام برای پیشبینی کشف تقلب استفاده کردند. نتایج این پژوهش نشان داد که پیشبینی تقلب در گزارشگری مالی به کمک شبکههای عصبی مصنوعی و شبکههای مبتنی بر قواعد منطق فازی در مقایسه با رگرسیون لجستتیک گام به گام نتایج بهتری را به دنبال دارد.

کرکس و همکاران (۲۰۰۷) از درختان تصمیم، شبکههای عصبی و شبکههای مبتنی برا قواعد بیزین برای کشف تقلب استفاده کردند که نتایج نشان داد که شبکههای مبتنی بر قواعد بیزین در مقایسه با سایر روشها توانایی بهتری در کشف تقلب دارند.

هوگس و همکاران (۲۰۰۷) با نمونه ای متشکل از ۳۹۰ شرکت شامل ۵۱ شرکت متقلب و ۳۳۹ شرکت غیر متقلب طی سالهای ۲۰۰۴ -۱۹۹۸ به معرفی رویکرد تکاملی برای کشف گزارشگری مالی متقلبانه پرداختند.در مجموع ۸۵ متغیر مالی و غیر مالی به عنوان پیش بینی کننده بالقوه خطر تقلب در صورتهای مالی انتخاب شدند.نتایج حاصل از الگریتم ژنتیک بیانگر سودمندی این تکنیک و همچنین اهمیت اطلاعات مالی و غیر مالی در کشف گزارشگری مالی متقلبانه است.

کرکوز و همکاران (۲۰۰۷) با نمونه ای متشکل از ۳۸ شرکت متقلب و ۳۸ شرکت غیر متقلب و با استفاده از ۱۰ نسبت مالی به بررسی سودمندی درخت تصمیم ،شبکه های عصبی مصنوعی و شبکه های باور بیزین در کشف صورتهای مالی متقلبانه پرداخت.نتایج پژوهش حاکی از این بود که صورتهای مالی منتشره در بر گیرنده اطلاعات سودمندی برای کشف تقلب در گزارشگری مالی می باشند.همچنین ، مدل شبکه باور بیزین عملکرد بهتری نسبت به شبکه عصبی مصنوعی و درخت تصمیم داشت و خطای نوع اول در هر سه مدل کمی بود.

شیخ (۱۳۸۳) به بررسی اهمیت شاخص های بـالقوه تقلـب مـالی بـا اسـتفاده تحلیل عاملی پرداخت. هدف این تحقیق شناسایی بسترهای خطر از دیدگاه حسابرسان مستقل و مدیران مالی و همچنین ولویِ ا ت بندی آنها بود نسبت به یکدیگربود. در این تحقیق وی به وجود علایم خطری (شرایط و رویداهایی) بیش از نقاط ضعف سیستم های حسابداری و کنترل علامت های داخلی اشاره نمود و نشان داد که تمامی هشدار های ۵ عامل ریسک‌پذیری و خطرجویی، ویژگی های مدیریت، ویژگی‌های اقتصادی، بی‌ثباتی، ناتوانی و وناکافی بودن سودآوری، ویژگی‌های ساختاری اشباع شدهاند.

فرقاندوست و عبدالهی (۱۳۸۷) در تحقیقی به بررسی کابرد سیستم‌های فازی در ارزیابی ریسک تقلب مدیریت پرداختند.

فرقاندوست و براوری (۱۳۸۸) به بررسی کاربرد روشهای تحلیلی در ارزیابی ریسک تحریف صورتهای مالی (تقلب مدیریت) پرداختند. آنها براساس نتیجه نهایی تحلیل آماری در مورد توانایی مدل پیشنهادی در پیشبینی حضور یک شرکت در گروه متقلب و یا غیر متقلب، درصد پیشبینی صحیح توسط مدل نهایی را ۹۰/۵ ٪ محاسبه گردید.

پورحیدری و بذر افشان (۱۳۹۰) به بررسی سودندی استفاده از چک لیست راهنمای کشف تقلب در ارزیابی خطر تقلب مدیریت پرداختند. آنها در ابتدا براساس مرور مطالعات و چک لیست کشف تقلب، پرسشنامه مبتنی بر آن تهیه و تعداد ۶۰۰ پرسشنامه بین حسابرسان توزیع نمودند. نتایج این مطالعه نشان داد که استفاده از چک لیست راهنمای کشف تقلب مدیریت موجب کاهش اشتباهات قضاوت شخصی حسابرس خواهد شد. همچنین افزایش اندازه شرکت مورد حسابرسی و کاهش زمان مورد نیاز برای انجام حسابرسی، احتمال استفاده حسابرس از چک لیست کشف تقلب را افزایش

میدهد. به طور کلی نتایج حاصل از این مطالعه نشاندهنده سودمندی استفاده از چک لیست کشف تقلب در ارزیابی خطر تقلب مدیریت میشود.

محمدی و همکاران (۱۳۹۰) در پژوهشی با عنوان "شناسایی مدیریت آرایشی سود" به بررسی وجود مـدیریت آرایشـی سـود درشرکتهای پذیرفته شده در بورس اوراق بهادار تهران با استفاده از قانون بنفورد و توزیع کـای ـ دو پرداختند. نتایج این پژوهش نشان داد شرکتهای بررسی شده، فاقد مدیریت آرایشی سود بوده است.

پور حیدری و بذرافشان (۱۳۹۱) به بررسی تاثیر ویژگی های فردی (نظیر جنسیت، رشته تحصیلی، سطح تحصیلات و تجربه) در ارزیابی اهمیت بسترهای خطر تقلب پرداختند و نتایج این تحقیق بیانگر این بود که نوع جنسیت تنها در ارزیابی اهمیت بسترهای خطر تقلب مربوط به گروه های "ویژگی مدیریت و نفوذ آن بر محیط کنترلی" و" ویژگی های عملیاتی و ثبات مالی" اثری معنادار دارد اما رشته تحصیلی، سطح تحصیلی و تجربه در ارزیابی اهمیت بسترهای خطر هیچ یک از ۴ گروه دیگر اثری ندارد.

رسـتمی و جعفـری درگیـری (۱۳۹۲) در پژوهشی با عنوان "بررسی قانون بنفورد در بورس اوراق بهادار تهران" به ارزیابی کیفیت دادههـادر بورس اوراق بهادار تهران با استفاده از قانون بنفورد پرداختند. نتایج این تحقیـق حـاکی از آناست که دادههای مورد بررسی از قانون بنفورد تبعیت نمی کنند و نابهنجاریهـای رفتـاری بـازار سرمایه، از دلایل عدم تبعیت از قانون بنفورد است.

۴- تعریف عملیاتی متغیرهای پژوهش

متغیر وابسته

تقلب در صورتهای مالی (Fraud) است که از ماهیت کیفی برخوردار بوده و دارای مقیاس سنجش اسمی است. در اندازهگیری این متغیر به شرکتهای متقلب عدد یک و به شرکتهای غیر متقلب عدد صفر تخصیص داده میشود.

جامعی و همکاران (۱۳۹۵)، معیارهایی را به عنوان شاخص تقلب، به شرح زیر در نظر گرفتند:

۱) وجود اختلافات مالیاتی با حوزه مالیات طبق یادداشت ذخیره مالیات بر درآمد و و بند شرط گزارش حسابرسی پرونده مالیاتی

۲) وجود تعدیلات سنواتی بااهمیت و صورت‌های مالی تجدید ارائه شده.

برای سنجش گزارشگری مالی متقلبانه از متغیر ساختگی استفاده می‌شود، به‌طوری‌که اگر گزارشگری مالی متقلبانه داشته باشند، یک و در غیر این صورت صفر در نظر گرفته می‌شود.

متغیرهای مستقل

متغیرهای مستقل در این تحقیق، از طریق بررسی مقالات معتبر علمی گرد آوری و بعد از روش گلوله برفی و مصاحبه با متخصصان ذیفن و اساتید خبره که تعداد ان ها ۲۰ نفر بود شناسایی و مجموعا 22 متغیر برای بهره وری در تحقیق مورد محاسبه قرار گرفت

متغیرهای مستقل مورد استفاده در تحقیق

طبقه	نسبتهای مورد استفاده	نحوه محاسبه
نسبت های مالی	سود خالص به فروش	NI/SAL
	بازده دارایی (سود خالص به دارایی)	NI/TA
	بازده حقوق صاحبان سهام (سود خالص به حقوق صاحبان سهام)	NI/EQ
	جریان وجه نقد عملیاتی به سود خالص	OCF/NI
	جریان وجه نقد عملیاتی به فروش	OCF/SAL
	جمع بدهی به حقوق صاحبان سهام	TD/EQ
	جمع بدهی به جمع دارایی	TL/TA
	موجودی به جمع دارایی	INV/TA
	حسابهای دریافتنی به جمع دارایی	AR/TA
	داراییهای جاری به کل دارایی	CA/TA
	موجودی به فروش	INV/SAL
	حسابهای دریافتنی به فروش	AR/SAL
	بهای تمام شده کالاهای فروش رفته به فروش	COG/SAL
	گردش داراییها (کل فروش به کل دارایی)	SAL/TA

طبقه	نسبتهای مورد استفاده	نحوه محاسبه
	حسابهای پرداختنی به فروش	AP/SAL
	گردش حسابهای دریافتنی (کل فروش به کل حسابهای دریافتنی)	SAL/AR
	هزینههای عملیاتی به فروش	Cost/SAL
	وجه نقد و معادل وجه به کل دارایی	Cash/TA
	هزینه مالیات بر درآمد به مالیات پرداختی	TAXE/TAXP
نسبت های غیر مالی	کیفیت حسابرسی جزء طبقه اول حسابرس معتمد عدد ۱ در خیر این صورت صفر	Quality Audit
	اظهارنظر حسابرسی مقبول عدد ۱ در خیر این صورت صفر	Opinion
	تجدید ارائه صورت مالی عدد ۱ در خیر این صورت صفر	Restatement

۵- روش شناسی پژوهش

پژوهش حاضر از نوع پژوهشهای کاربردی و تجربی-همبستگی است. جامعه آماری در این تحقیق شامل شرکتهای پذیرفته شده در بورس اوراق بهادر تهران است و قلمرو زمانی تحقیق ۹۷-۹۲میباشد. در نهایت تعداد ۷۹۲ شرکت (۷۹۲ مشاهده) به عنوان نمونه انتخاب شدند که ۱۵۳ مشاهده حاوی تقلب و ۶۳۹ مشاهده فاقد تقلب میباشند و شرکتهای فاقد تقلب انتخابی با توجه به صنعت، حجم فعالیت و ارزش روز شرکت حاوی تقلب (گروه آزمون) در همان سال مالی وقوع تقلب انتخاب شدند. همچنین برای گروه آزمون صورتهای مالی سه سال قبل از تقلب نیز مورد بررسی قرار گرفت. لذا در مجموع صورتهای مالی ۷۹۲ سال- شرکت مورد بررسی قرار گرفتند.. در این پژوهش برای جمعآوری و گرداوری دادهها در مرحله اول از روش نمونه گیری گلوله برفی از بین نخبگان حسابداری استفاده میکنیم بدین صورت که از مقالات معتبر ۵۰ نسبت تاثیر گذار در تقلب شناسایی گردیدکه با استفاده از روش گلوله برفی از بین ۲۰ نفر از متخصصین ودانشگاهیان تعداد ۲۲نسبت در۲دسته مورد تایید قرار گرفت که بعد از آن دوباره از بین این نسبت ها نسبت هایی که تاثیر گذاری بیشتری در تشخیص تقلب صورت های مالی دارندبا مراجعه به روش شبکه عصبی مصنوعی(ANN)انتخاب می شوند و سپس با روش درخت های تصمیم گیری اقدام به مدل سازی مینماییم ومدلی که دارای بالاترین میزان دقت تشخیص FSFو همچنین دارای کمترین میزان خطا باشد بعنوان مدل ایده ال مد نظراست برای جمعآوری اطلاعات مورد نیاز بخش مبانی نظری از مجلات تخصصی لاتین، و برای گردآوری سایر دادهها و اطلاعات مورد نیاز عمدتاً از طریق بانک های اطلاعاتی سازمان بورس اوراق بهادار تهران، گزارش های روزانه و هفتگی سازمان بورس و نرم افزارهای رهاورد نوین و تدبیر پرداز استفاده خواهد شد.روش تحقیق از نوع گردآوری اطلاعات، آمار توصیفی و استنباطی میباشد و هدف از این تحقیق وجود، میزان و نوع رابطه بین متغیرهای مورد آزمون است.. در این پژوهش بهمنظور جمع آوری دادهها و اطلاعات از روش تکنیک های داده کاوی استفاده گردیده است. مبانی نظری پژوهش از کتب، مجلات و سایتهای تخصصی فارسی و لاتین گرد آوری گردیده و دادههای مالی نیز با مراجعه به سایت سازمان بورس و اوراق بهادر تهران و صورتهای مالی شرکتها، گزارشهای هیات مدیره به مجمع عمومی عادی و گزارشهای حسابرسی مستقل و بازرسی قانونی و همچنین با استفاده از نرمافزار رهآورد نوین و تدبیر پرداز گردآوری گردیده است.

دادههای مورد استفاده در این پژوهش از نوع دادههای واقعی است و با روش استقرایی به کل جامعه آماری قابل تعمیم است. از طرفی با توجه به اینکه در این تحقیق از اطلاعات تاریخی استفاده شده است و دادههای مربوط به این پژوهش از طریق اطلاعات مربوط به شرکتهای پذیرفته شده در بورس اوراق بهادر بدست آمده که صورتهای مالی آنها حاوی ریسک تقلب است، بنابراین این پژوهش در دسته تحقیقات شبه آزمایشی قرار میگیرد.بهمنظور تجزیه و تحلیل اطلاعات توسط آزمونهای شبکه عصبی مصنوعی ودرخت های تصمیم گیری استفاده می شود که برای تجزیه تحلیل شبکه عصبی مصنوعی از نرم افزار های spssو نرم افزار matlab R2013و برای درخت های تصمیم گیری از نرم افزار clementin12 استفاده می شود .و برای مدل سازی از مدل های تشخیص fsfاستفاده می کنیم که بالاترین دقت تشخیص و کمترین میزان خطا را برای انتخاب بهترین مدل در نظر می گیریم .

۶-یافته های پژوهش

پژوهش‌های انجام‌شده نشان‌دهنده این موضوع است که در زمینه تشخیص تقلب صورت‌های مالی روش‌های آماری مختلفی استفاده شده است ولی تحقیق حاضر بر ایجاد و به کارگیری هوش مصنوعی و روش‌های یادگیری ماشینی متمرکز شده است تا دقت پیش‌بینی تشخیص تقلب صورت‌های مالی

توسط الگوریتم‌های مانند درخت تصمیم‌گیری، شبکه عصبی مصنوعی و ماشین بردار پشتیبانی در بورس اوراق بهادار تهران جهت تشخیص تقلب صورت‌های مالی مقایسه گردند.

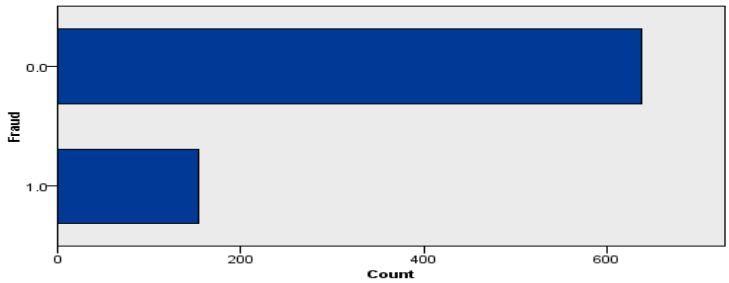
در این بخش الگوریتم‌های پیشنهادی را بر روی مجموعه‌ای از داده‌های واقعی پیاده می‌کنم. مجموعه داده‌های استفاده شده در این مطالعه، مربوط به اطلاعات شرکت پذیرفته‌شده در بورس تهران هستند که از بانک اطلاعات کدال و نرم‌افزار مالی ره‌آورد نوین استخراج شده است.

از این مجموعه عمدتاً برای مقایسه و تجزیه‌وتحلیل روش‌های منتخب طبقه‌بندی و استفاده می‌شود. این داده‌ها در برگیرنده اطلاعات مربوط به ۷۹۲ سال-شرکت (مشاهده) که ، تعداد ۱۵۳ صورت مالی متقلبانه تشخیص داده شدند و مابقی ۶۳۹ عدد غیر متقلبانه تشخیص داده شدند.

توصیف اطلاعات

در این بخش اطلاعات مربوط به داده‌های تحقیق مورد تحلیل توصیفی قرار می‌گیرد.

همانطور که نتایج نشان می دهد از کل ۷۹۲ مشاهده، تعداد ۱۵۳ صورت مالی متقلبانه تشخیص داده شدند و مابقی ۶۳۹ عدد غیر متقلبانه تشخیص داده شدند. درواقع حدود ۸۰ درصد از صورت های مالی، غیر متقلبانه گزارش شده و حدود ۲۰ درصد از آنها متقلبانه شناخته شده اند.



شکل ۱-توصیف اطلاعات متغیر وابسته

جدول توصیف اطلاعات متغیر وابسته

صنعت	تعداد شرکت	تعداد مشاهده (سال-شرکت)	تعداد صورت مالی متقلبانه	درصد
ابزار پزشکی	۱	۶	۶	٪۱۰۰
استخراج زغال سنگ	۱	۶	۰	٪۰
استخراج سایر معادن	۱	۶	۰	٪۰
استخراج کانه های فلزی	۴	۲۴	۱	٪۴
حمل و نقل انبارداری و ارتباطات	۳	۱۸	۱	٪۶
خودرو و قطعات	۱۷	۱۰۲	۳۲	٪۳۱
دارویی	۲۰	۱۲۰	۶	٪۵
دستگاههای برقی	۴	۳۰	۶	٪۲۵
رایانه	۲	۱۲	۰	٪۰
زراعت و خدمات وابسته	۱	۶	۰	٪۰
سیمان آهک گچ	۱۵	۹۰	۱۴	٪۱۶
شیمیایی	۱۰	۶۰	۲۲	٪۳۷
غذایی بجز قند وشکر	۱۰	۶۰	۱۱	٪۱۸
فرآورده های نفتی	۲	۱۲	۶	٪۵۰
فلزات اساسی	۱۳	۷۸	۱۷	٪۲۲
فنی و مهندسی	۲	۱۲	۰	٪۰
قند و شکر	۱	۶	۰	٪۰
کاشی و سرامیک	۵	۳۰	۱۲	٪۴۰

صنعت	تعداد شرکت	تعداد مشاهده (سال- شرکت)	تعداد صورت مالی متقلبانه	درصد
کانی غیر فلزی	۵	۳۰	۳	٪۱۰
لاستیک و پلاستیک	۳	۱۸	۰	٪۰
ماشین آلات و تجهیزات	۶	۳۶	۱۱	٪۳۱
محصولات چوبی	۱	۶	۱	٪۱۷
محصولات فلزی	۳	۱۸	۱	٪۶
منسوجات	۱	۶	۳	٪۵۰

تعداد متغیرهای مستقل یا پیش‌بینی کننده در این تحقیق شامل ۱۹ نسبت مالی و ۳ نسبت غیر مالی است که شرح آنها در جدول ۱ آمده است.

در مرحله اول با استفاده از الگوریتمهای ANN و SVM به منظور انتخاب متغیرهای مهم شناسایی کننده تقلب در صورت های مالی از نرم‌افزار IBM SPSS Modeler 18 استفاده می‌شود. نتایج بکارگیری الگوهای ذکر شده در جدول ۲ شرح داده شده است.

جدول ۱- نسبت های مالی استفاده شده در تحقیق جاری

نام	نماد	نام	نماد
حسابهای دریافتنی به فروش	AR/SAL	سود خالص به فروش	NI/SAL
بهای تمام شده کالاهای فروش رفته به فروش	COG/SAL	بازده دارایی	NI/TA
گردش داراییها	SAL/TA	بازده حقوق صاحبان سهام	NI/EQ
حسابهای پرداختنی به فروش	AP/SAL	بدهی به حقوق صاحبان سهام	TD/EQ
گردش حسابهای دریافتنی	SAL/AR	بدهی به جمع دارایی	TL/TA
هزینههای عملیاتی به فروش	Cost/SAL	جریان وجه نقد عملیاتی به سود خالص	OCF/NI
وجه نقد و معادل وجه به کل دارایی	Cash/TA	جریان وجه نقد عملیاتی به فروش	OCF/SAL
هزینه مالیات بر درآمد به مالیات پرداختی	TAXE/TAXP	موجودی به جمع دارایی	INV/TA
موجودی به فروش	INV/SAL	حسابهای دریافتنی به جمع دارایی	AR/TA
	دارایی‌های جاری به کل دارایی	CA/TA	

جدول ۲- نسبت های غیر مالی استفاده شده در تحقیق جاری

نام	نماد	نام	نماد
کیفیت حسابرسی	Quality Audit	تجدید ارائه صورت مالی	Restatement
	اظهارنظر حسابرسی	Opinion	

ndividual Models

Comparing $S-Fraud with Fraud

'Partition'	Testing		Training	
Correct	192	81.01%	475	85.59%
Wrong	45	18.99%	80	14.41%
Total	237		555	

Comparing $N-Fraud with Fraud

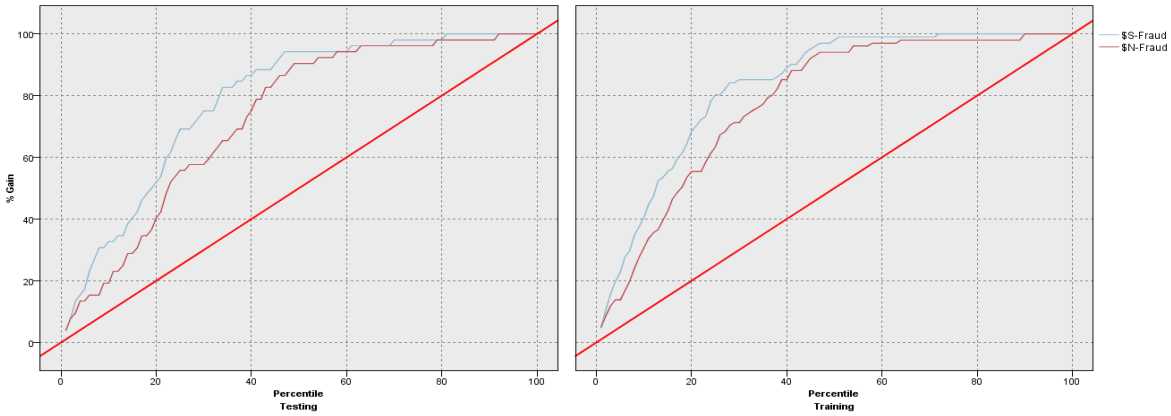
'Partition'	Testing		Training	
Correct	185	78.06%	452	81.44%
Wrong	52	21.94%	103	18.56%
Total	237		555	

مقایسه نتایج مدل های SVM و ANN برای پیش‌بینی تقلب

نتایج استفاده از الگوریتم‌های شبکه عصبیANN و ماشین بردار پشتیبان SVM برای پیش‌بینی تقلب در صورت های مالی شرکت ها در جدول فوق نشان داده شده است. همانطور که این جدول نشان می دهد در بخش آموزش قدرت پیش‌بینی الگوریتم ماشین بردار پشتیبان حدود ۸۶ درصد است. قدرت این الگوریتم در بخش آزمون حدود ۸۲ درصد بوده است. همچنین قدرت پیش‌بینی الگوریتم شبکه عصبی در بخش آموزش ۸۱ درصد و در آزمون ۷۸ درصد بوده است. در بحث الگوریتم هااین اعداد تنها یک پارامتر هستند و چون انحراف معیاری برای آنها گزارش نمی شود نمی توان روی آن آزمون آماری انجام داد که ادعا کرد که این تفاوت معنی دار است یا خیر؟ به عبارت دیگر اعداد مطلقی هستند که صرف بزرگتر یا کوچکتر بودن عدد به معنی تفاوت معنی دار است . و از طرفی دیگر همانطور که جداول دقت و خطا SVMوANN دراخرمقاله نشان داده میشودتفاوت معنی داری بین قدرت تبیین دو مدل SVMوANN وجود دارد.

این نتایج در نمودار زیر، که بیانگر شاخص Gain برای پیش‌بینی الگوریتم ها است، نیز نشان داده می‌شود.

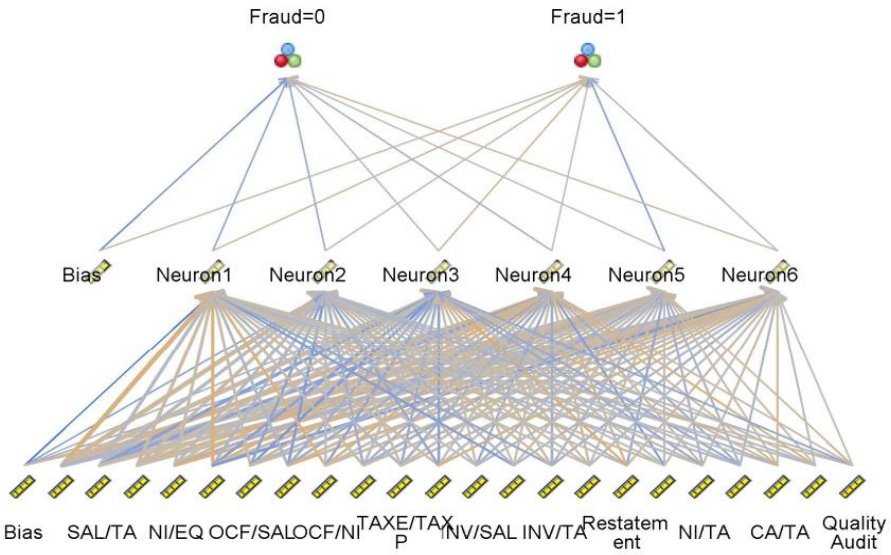
در بخش بعدی رتبه بندی متغیرهای پیش‌بینی کننده تقلب در صورت های مالی توسط دو الگوریتم ذکر شده می آید.



نمودار ۱-شاخص GAIN برای مقایسه پیش‌بینی قدرت الگوریتم ANN و SVM

الگوریتم ANN

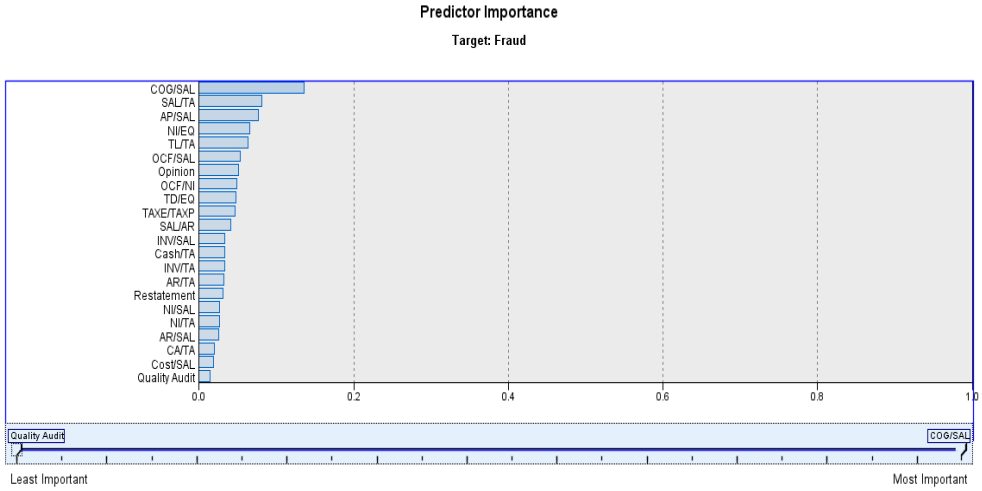
شبکه عصبی طراحی شده به شکل زیر می باشد.



نمودار ۲- شبکه عصبی طراحی شده برای پیش‌بینی تقلب در صورت های مالی

جدول ۳- تعیین میزان اهمیت متغیرها در پیش‌بینی توسط الگوریتم شبکه عصبی

Rank	Nodes	Importance	Rank	Nodes	Importance
1	COG/SAL	0.1356	**12**	INV/SAL	0.0334
2	SAL/TA	0.0813	**13**	Cash/TA	0.0332
3	AP/SAL	0.077	**14**	INV/TA	0.0328
4	NI/EQ	0.065	**15**	AR/TA	0.0317
5	TL/TA	0.0629	**16**	Restatement	0.0305
6	OCF/SAL	0.0532	**17**	NI/SAL	0.0268
7	Opinion	0.0507	**18**	NI/TA	0.026
8	OCF/NI	0.0485	**19**	AR/SAL	0.0247
9	TD/EQ	0.0475	**20**	CA/TA	0.0192
10	TAXE/TAXP	0.0467	**21**	Cost/SAL	0.0189
11	SAL/AR	0.0409	**22**	Quality Audit	0.0135



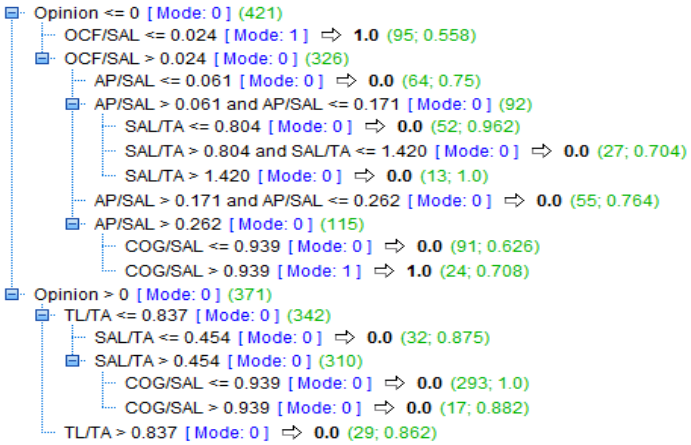
نمودار ۳- تعیین میزان اهمیت متغیرها در پیش‌بینی توسط الگوریتم شبکه عصبی

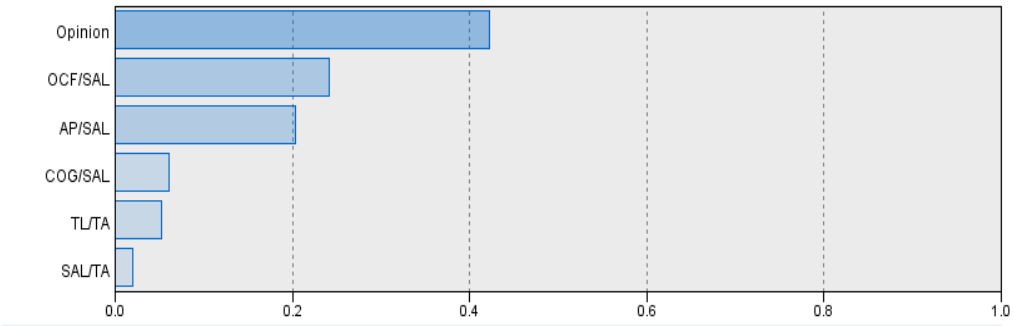
بر اساس استاندارد روش شناختی جان[۴] (۲۰۱۸)، متغیرهایی که بیش از ۵ درصد قدرت تبیین دارند در مرحله دوم پیش‌بینی بکار گرفته می شوند. بر این اساس در الگوریتم شبکه عصبی متغیرهای، بهای تمام شده کالاهای فروش رفته به فروشCOG/SAL ،گردش دارایی-هاSAL/TA ، حساب-های پرداختنی به فروشAP/SAL ، بازده حقوق صاحبان سهام NI/EQ هستند.

پیش‌بینی با متغیرهای منتخب از طریق الگوریتم ANN

الگوریتم ANN-CHAID

نتایج برارزش الگوریتم CAHID برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس روش





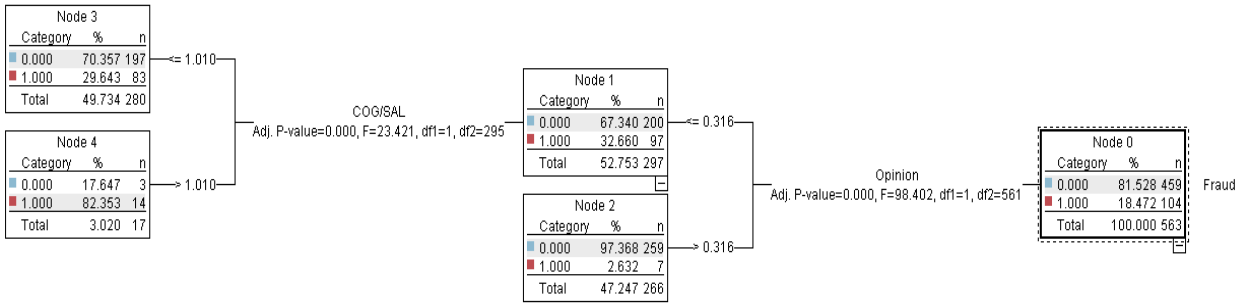
نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم ANN-CHAID

در مرحله قبلی ۷ متغیر انتخاب شدند اما قرار نیست در این مرحله همه ۷ متغیر قدرت تبیین داشته باشند. در این مدلسازی یکی از متغیرها بطور کلی نتوانسته در پیش بینی حضور داشته باشد. درواقع وارد مدل شده اما چون قدرت تبیین نداشته نرم افزار آن را در مدلسازی کلا در نظر نگرفته است. همانطور که نمودار نشان می دهد در بین متغیرهای پیش‌بینی کننده اصلی، تنها متغیر اظهار نظر حسابرس جزو متغیرهای غیر مالی پیش‌بینی کننده مهم اما مهمترین متغیر با حدود ۴۲ درصد قدرت پیش‌بینی قرار دارد. قدرت پیش‌بینی متغیرهای بعدی به ترتیب، ۲۴، ۲۰، ۶، ۵ و ۲ درصد است.

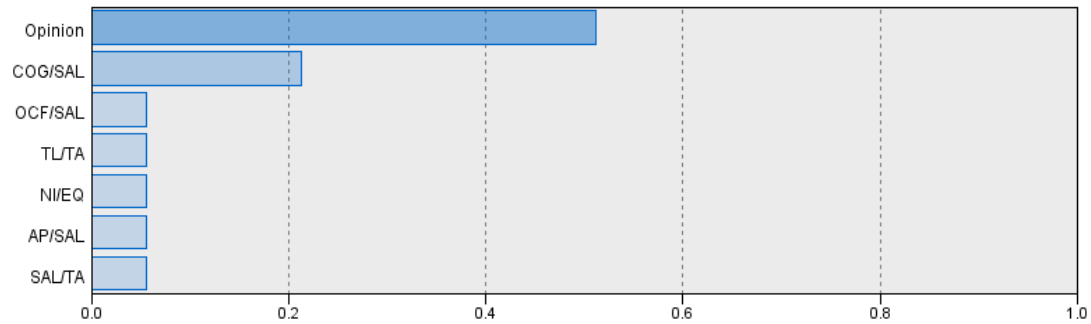
الگوریتم ANN-QUEST

نتایج برارزش الگوریتم QUEST برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس روش شبکه مصنوعی عصبی به صورت زیر است.

مشابه با مدل قبلی همانطور که نمودار نشان می دهد در بین متغیرهای پیش‌بینی کننده اصلی، تنها متغیر اظهار نظر حسابرس جزو متغیرهای غیر مالی پیش‌بینی کننده مهم اما مهمترین متغیر با حدود ۵۱ درصد قدرت پیش‌بینی قرار دارد. قدرت پیش‌بینی متغیرهای بعدی به ترتیب، ۲۱، ۶، ۶، ۶، ۶ و ۶ درصد است.



نمودار نتایج بکارگیری الگوریتم ANN-QUEST برای پیش‌بینی تقبل در صورتهای مالی



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم ANN-QUEST

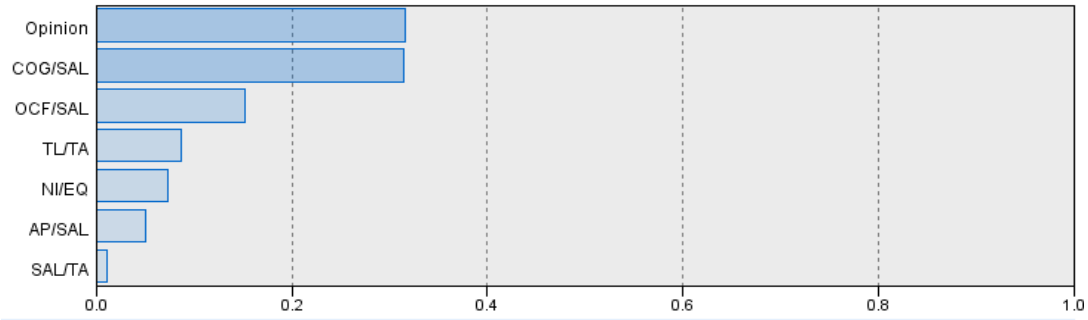
الگوریتم ANN-C5.0

نتایج برارزش الگوریتم C5.0 برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس روش شبکه مصنوعی عصبی به صورت زیر است.

در تایید نتایج دو مدل قبلی و همانطور که نمودار نشان می دهد در بین متغیرهای پیش‌بینی کننده اصلی، تنها متغیر اظهار نظر حسابرس جزو متغیرهای غیر مالی پیش‌بینی کننده مهم ومهمترین متغیر با حدود ۳۲ درصد قدرت پیش‌بینی قرار دارد. قدرت پیش‌بینی متغیرهای بعدی به ترتیب، ۳۱، ۱۵، ۹، ۷، ۵ و ۱ درصد است.

جدول نتایج بکارگیری الگوریتم ANN-C5.0 برای پیش‌بینی تقبل در صورتهای مالی

- Opinion <= 0 [Mode: 0] (421)
  - COG/SAL <= 1.024 [Mode: 0] (399)
    - OCF/SAL <= 0.021 [Mode: 1] (82)
      - COG/SAL <= 0.533 [Mode: 0] ⇨ **0.0** (5; 1.0)
      - COG/SAL > 0.533 [Mode: 1] (77)
        - AP/SAL <= 1.037 [Mode: 1] ⇨ **1.0** (73; 0.603)
        - AP/SAL > 1.037 [Mode: 0] ⇨ **0.0** (4; 1.0)
    - OCF/SAL > 0.021 [Mode: 0] (317)
      - SAL/TA <= 1.384 [Mode: 0] (290)
        - AP/SAL <= 0.273 [Mode: 0] (191)
          - SAL/TA <= 1.196 [Mode: 0] ⇨ **0.0** (181; 0.823
          - SAL/TA > 1.196 [Mode: 1] (10)
            - AP/SAL <= 0.144 [Mode: 0] ⇨ **0.0** (3; 1.0)
            - AP/SAL > 0.144 [Mode: 1] ⇨ **1.0** (7; 0.857
        - AP/SAL > 0.273 [Mode: 0] (99)
          - AP/SAL <= 0.476 [Mode: 1] ⇨ **1.0** (47; 0.596)
          - AP/SAL > 0.476 [Mode: 0] ⇨ **0.0** (52; 0.75)
      - SAL/TA > 1.384 [Mode: 0] ⇨ **0.0** (27; 1.0)
    - COG/SAL > 1.024 [Mode: 1] ⇨ **1.0** (22; 0.909)
- Opinion > 0 [Mode: 0] (371)
  - NI/EQ <= 0.244 [Mode: 0] (169)
    - TL/TA <= 0.735 [Mode: 0] ⇨ **0.0** (132; 0.977)
    - TL/TA > 0.735 [Mode: 0] (37)
      - COG/SAL <= 1.309 [Mode: 0] ⇨ **0.0** (34; 0.882)
      - COG/SAL > 1.309 [Mode: 1] ⇨ **1.0** (3; 1.0)
  - NI/EQ > 0.244 [Mode: 0] ⇨ **0.0** (202; 1.0)



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم ANN-C5.0

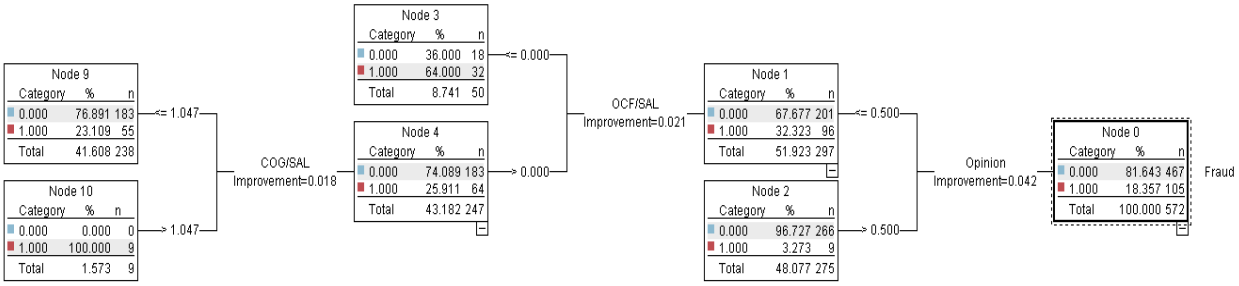
الگوریتم ANN-CART

نتایج برارزش الگوریتم CART برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس روش شبکه مصنوعی عصبی به صورت زیر است.

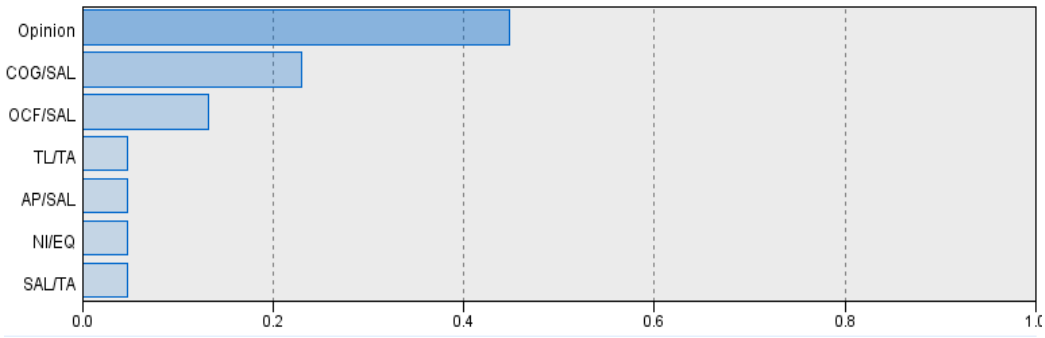
در تایید نتایج دو مدل قبلی و همانطور که نمودار نشان می دهد در بین متغیرهای پیش‌بینی کننده اصلی، تنها متغیر اظهار نظر حسابرس جزو متغیرهای غیر مالی پیش‌بینی کننده مهم و مهمترین متغیر با حدود ۴۵ درصد قدرت پیش‌بینی قرار

دارد. قدرت پیش‌بینی متغیرهای بعدی به ترتیب، ۲۳، ۱۳، ۵، ۵، ۵ و ۵ درصد است.

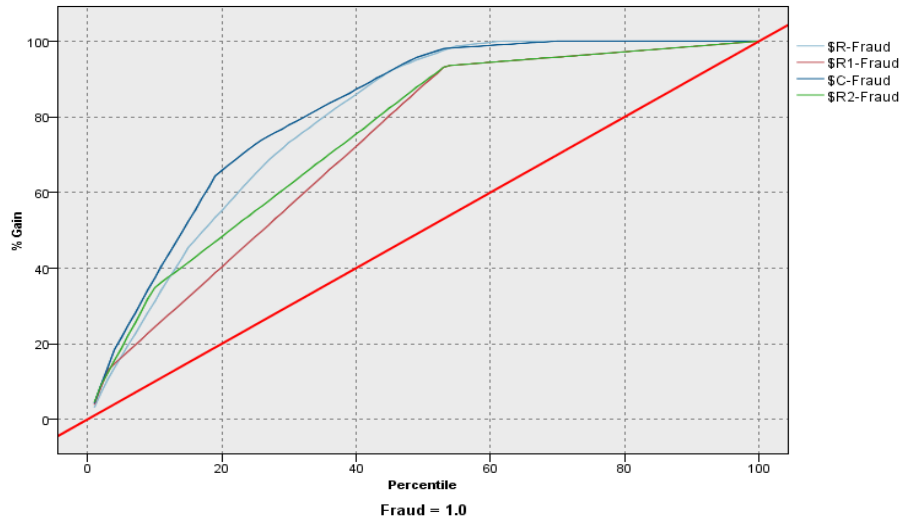
نتایج نشان می دهد از میان متغیرهای غیر مالی و غیر مالی، متغیر اظهار نظر حسابرس در تشخیص تقلب (در هر چهار الگوریتم) عامل موثر شناخته شده است که نتیجه میگیریم در الگوریتم‌های مبتنی بر شبکه عصبی، عامل غیر مالی اظهار نظر حسابرس بعنوان مهمترین متغیر برای تشخیص تقلب در صورتهای مالی بوده و بین۳۲ تا ۵۱ درصد قدرت پیش‌بینی دارد.



نمودارنتایج بکارگیری الگوریتم ANN-CART برای پیش‌بینی تقلب در صورتهای مالی



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم ANN-CART



نمودار شاخص GAIN الگوریتم ۴ تحقیق با متغیرهای منتخب از طریق ANN برای مقایسه نتایج

Individual Models

Comparing $R-Fraud with Fraud

Correct	660	83.33%
Wrong	132	16.67%
Total	792	

Comparing $R1-Fraud with Fraud

Correct	656	82.83%
Wrong	136	17.17%
Total	792	

Comparing $C-Fraud with Fraud

Correct	686	86.62%
Wrong	106	13.38%
Total	792	

Comparing $R2-Fraud with Fraud

Correct	667	84.22%
Wrong	125	15.78%
Total	792	

جدول مقایسه توان پیش‌بینی نتایج ۴ الگوریتم تحقیق با متغیرهای منتخب از طریق ANN

مقایسه چهار الگوریتم

نتایج تبیین شده در نمودار و جدول بالا نشان می دهد که از بین ۴ الگوریتم بکار گرفته شده الگوریتم C5 بالاترین قـدرت تبیین، معادل ۸۶ درصد را دارد. (اولین جدول از بالا به پـایین مربوط به CHAID دومین جدول مربوط بـه QUEST سـومین جدول، C5.0 و آخری مربوط به CART است)

در این قسمت با استفاده از الگوریتم شبکه عصبی مصنوعی بمنظور انتخاب متغیرهای پیش‌بینی کننده تقلب در صورت های مالی و با بکارگیری درخت های تصمیم گیریCHAID، QUEST، C5.0 و CART به انتخاب مدل ایده ال پرداخته است. نتایج مقایسه ای این الگوریتم ها که در جداول زیر آمده است، C5.0 قویترین است.

همانطور که نتایج معیارهای دقت[5] و خطا[6] نشان می دهد بطور کلی الگوریتم C5.0 و پس از آن الگوریتمCART بالاترین قدرت تبیین را با استفاده از متغیرهای انتخاب شده توسط شبکه عصبی را دارد. پس از آنها الگوریتم CAHID و در نهایت ضعیف ترین الگوریتم برای تبین متغیر تقلب در صورت های مالی الگوریتم QUEST است.

جدول مقایسه الگوریتم‌های بکاررفته پس از انتخاب متغیرها با روش شبکه عصبی با پارامتر FSF

MODEL	FSF Detection Accuracy	Non-FSF Detection Accuracy	Overall Accuracy
ANN + QUEST	0.664	0.552	0.776
ANN + CHAID	0.7125	0.617	0.808
ANN + CART	0.799	0.732	0.866
ANN + C5.0	0.841	0.788	0.894

جدول مقایسه قدرت تبیین الگوریتم‌های بکاررفته پس از انتخاب متغیرها با روش شبکه عصبی با معیار خطا

MODEL	Type I Error Rate	Type II Error Rate	Overall Error Rate
ANN + QUEST	0.2335	0.334	0.144
ANN + CHAID	0.209	0.281	0.133
ANN + CART	0.1935	0.274	0.127
ANN + C5.0	0.184	0.241	0.106

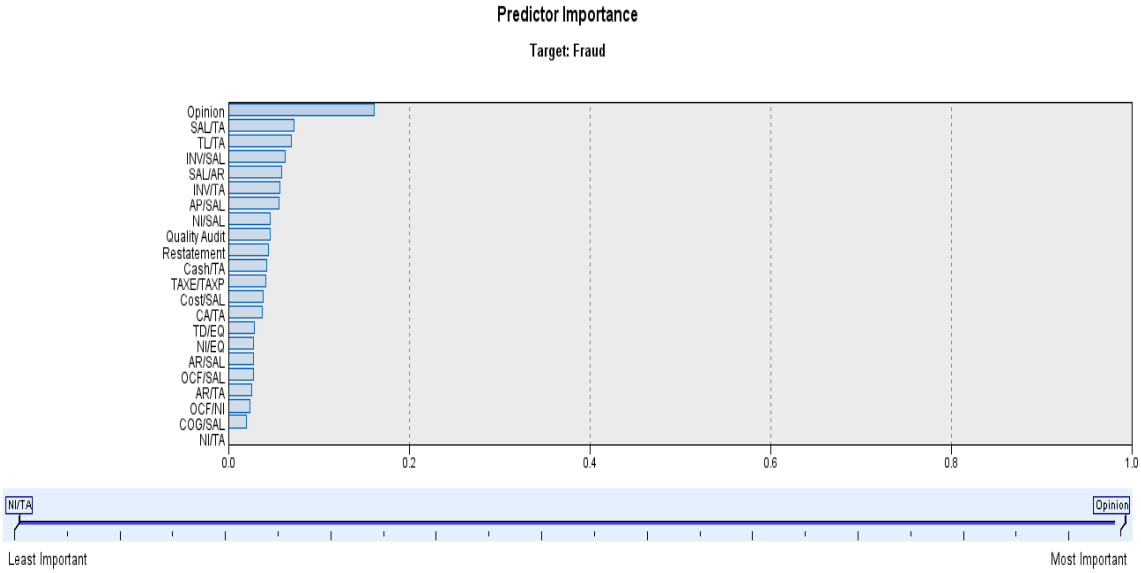
الگوریتم SVMهمانطور که نتایج بصورت زیر نشان می‌دهد:

بر اساس معیار قدرت تبیین بیش از ۵ درصد در الگوی ماشین بردار پشتیبان نیز متغیرهای، اظهارنظر حسابرسیOpinion ، گردش دارایی-هاSAL/TA ، بدهی به جمع داراییTL/TA ، موجودی به فروشINV/SAL ، گردش حساب-های دریافتنی SAL/AR ، موجودی به جمع داراییINV/TA ، حساب-های پرداختنی به فروشAP/SAL ، سود خالص به فروشNI/SAL ، کیفیت حسابرسی Quality Audit متغیرهای منتخب هستند.

در مرحله دوم تحقیق از چهار الگوریتم CART و CHAID و C5.0 و QUEST برای مدلسازی استفاده می شوند.

جدول تعیین میزان اهمیت متغیرها در پیش‌بینی توسط الگوریتم ماشین بردار پشتیبان (SVM)

Rank	Nodes	Importance	Rank	Nodes	Importance
1	Opinion	0.1619	12	TAXE/TAXP	0.0429
2	SAL/TA	0.0751	13	Cost/SAL	0.0417
3	TL/TA	0.0732	14	CA/TA	0.0400
4	INV/SAL	0.0645	15	TD/EQ	0.0370
5	SAL/AR	0.0622	16	NI/EQ	0.0349
6	INV/TA	0.0601	17	AR/SAL	0.0335
7	AP/SAL	0.0598	18	OCF/SAL	0.0331
8	NI/SAL	0.0532	19	AR/TA	0.0302
9	Quality Audit	0.0574	20	OCF/NI	0.0269
10	Restatement	0.0468	21	COG/SAL	0.0223
11	Cash/TA	0.0441	22	NI/TA	0.0113



نمودار تعیین میزان اهمیت متغیرها در پیش‌بینی توسط الگوریتم ماشین بردار پشتیبان

پیش‌بینی با متغیرهای منتخب از طریق الگوریتم SVM

- الگوریتم SVM-CHAID

نتایج برارزش الگوریتم CAHID برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس الگوریتم ماشین بردار پشتیبان به صورت زیر است.

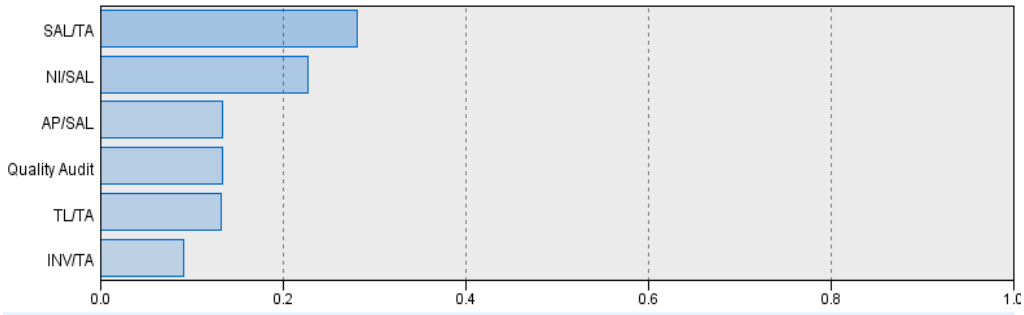
در مرحله قبلی۹ متغیر انتخاب شدند اما قرار نیست در این مرحله همه ۹متغیر قدرت تبیین داشته باشند. در این مدلسازی سه مورد از متغیرها بطور کلی نتوانسته در پیش بینی حضور داشته باشد. درواقع وارد مدل شده اما چون قدرت تبیین نداشته نرم افزار آنها را در مدلسازی کلا در نظر نگرفته است.همانطور که نتایج نشان می دهد، مهمترین متغیر پیش‌بینی کننده متغیر گردش دارایی ها با ۲۸ درصد قدرت پیش‌بینی بعنوان یک متغیر مالی قرار دارد. تنها متغیر غیر مالی در این الگو متغیر کیفیت حسابرسی بوده و توانسته ۱۳ درصد از تغییرات متغیر تقلب در صورتهای مالی را پیش‌بینی کند. سایر متغیرها به ترتیب توانسته اند ۲۳، ۱۳، ۱۳ و ۹ درصد از تغییرات متغیر وابسته را پیش‌بینی کنند.

AP/SAL <= 0.171 [Mode: 0] (396)
- SAL/TA <= 0.454 [Mode: 0] ⇨ **0.0** (46; 0.739)
- SAL/TA > 0.454 and SAL/TA <= 1.091 [Mode: 0] (265)
  - NI/SAL <= 0.040 [Mode: 0] ⇨ **0.0** (33; 0.697)
  - NI/SAL > 0.040 [Mode: 0] (232)
    - INV/TA <= 0.418 [Mode: 0] ⇨ **0.0** (207; 0.937)
    - INV/TA > 0.418 [Mode: 0] (25)
      - NI/SAL <= 0.126 [Mode: 1] ⇨ **1.0** (8; 0.625)
      - NI/SAL > 0.126 [Mode: 0] ⇨ **0.0** (17; 0.882)
- SAL/TA > 1.091 [Mode: 0] ⇨ **0.0** (85; 0.988)

AP/SAL > 0.171 and AP/SAL <= 0.262 [Mode: 0] (159)
- TL/TA <= 0.316 [Mode: 0] ⇨ **0.0** (15; 1.0)
- TL/TA > 0.316 and TL/TA <= 0.573 [Mode: 0] (59)
  - INV/TA <= 0.076 [Mode: 1] ⇨ **1.0** (9; 0.667)
  - INV/TA > 0.076 [Mode: 0] ⇨ **0.0** (50; 0.84)
- TL/TA > 0.573 and TL/TA <= 0.837 [Mode: 0] (72)
  - Quality Audit <= 0 [Mode: 0] ⇨ **0.0** (27; 0.852)
  - Quality Audit > 0 [Mode: 0] ⇨ **0.0** (45; 0.978)
- TL/TA > 0.837 [Mode: 1] ⇨ **1.0** (13; 0.615)

AP/SAL > 0.262 [Mode: 0] (237)
- NI/SAL <= -0.068 [Mode: 1] ⇨ **1.0** (62; 0.532)
- NI/SAL > -0.068 [Mode: 0] ⇨ **0.0** (175; 0.714)

جدول نتایج بکارگیری الگوریتم SVM-CHAID برای پیش‌بینی تقبل در صورتهای مالی

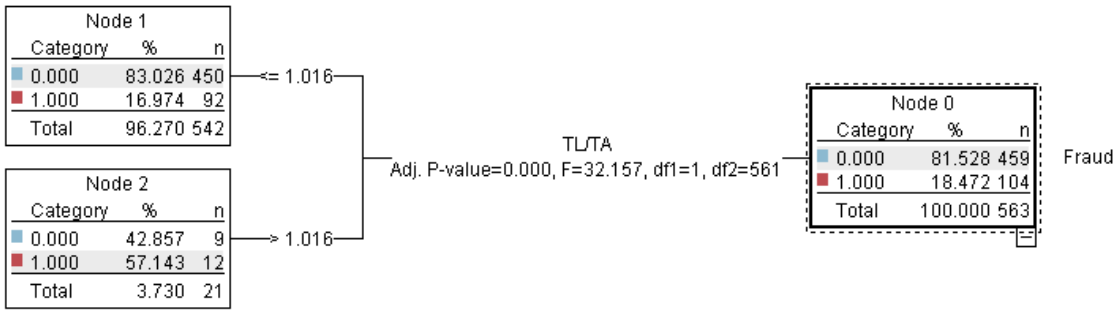


نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم SVM-CHAID

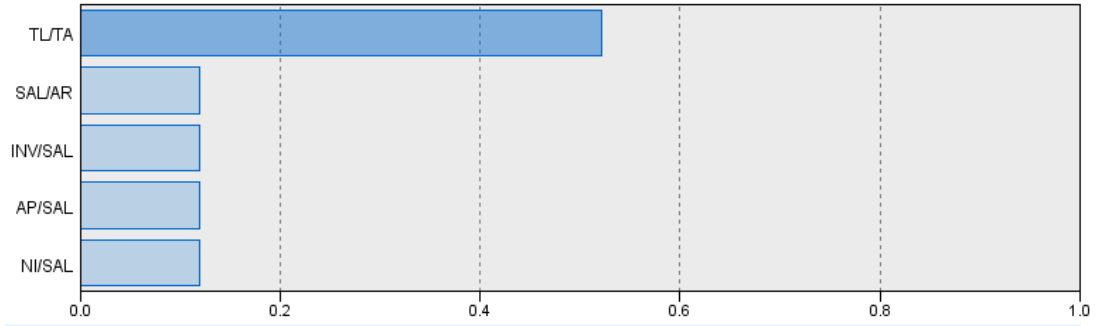
الگوریتم SVM-QUEST

نتایج برارزش الگوریتم QUEST برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس الگوریتم ماشین بردار پشتیبان به صورت زیر است.

در مرحله قبلی۹ متغیر انتخاب شدند اما قرار نیست در این مرحله همه ۹متغیر قدرت تبیین داشته باشند. در این مدلسازی ۴مورد از متغیرها بطور کلی نتوانسته در پیش بینی حضور داشته باشد. . درواقع وارد مدل شده اما چون قدرت تبیین نداشته نرم افزار آنها را در مدلسازی کلا در نظر نگرفته است. همانطور که نتایج نشان می دهد، مهمترین متغیر پیش‌بینی کننده متغیر نسبیت بدهی به دارایی با ۵۲ درصد قدرت پیش‌بینی بعنوان یک متغیر مالی قرار دارد. در این الگو هیچ متغیر غیر مالی حضور ندارد. سایر متغیرها هرکدام به تنهایی توانسته اند ۱۲ درصد از تغییرات متغیر وابسته را پیش‌بینی کنند.



نمودار نتایج بکارگیری الگوریتم SVM-QUEST برای پیش‌بینی تقبل در صورتهای مالی



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم SVM-QUEST

- الگوریتم SVM-C5.0

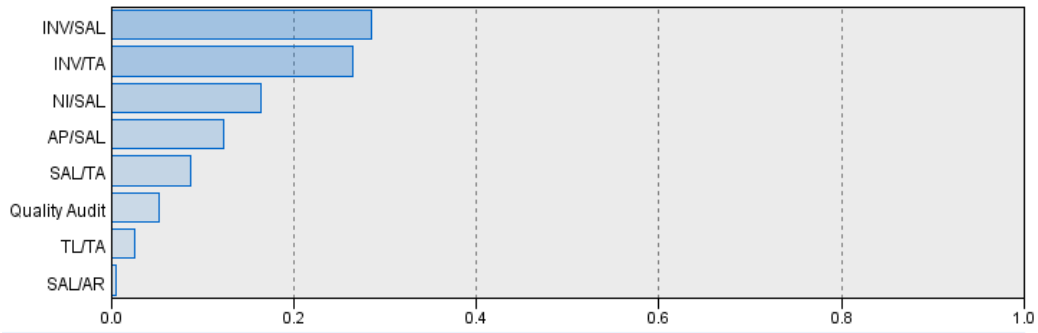
نتایج برارزش الگوریتم C5.0 برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس الگوریتم ماشین بردار پشتیبان به صورت زیر است.

در مرحله قبلی۹ متغیر انتخاب شدند اما قرار نیست در این مرحله همه ۹متغیر قدرت تبیین داشته باشند. در این مدلسازی ۱مورد از متغیرها بطور کلی نتوانسته در پیش بینی حضور داشته باشد. . درواقع وارد مدل شده اما چون قدرت تبیین نداشته نرم افزار آن را در مدلسازی کلا در نظر نگرفته است. در تایید نتایج مدل SVM-CHAID و همانطور که نتایج نشان می دهد، مهمترین متغیر پیش‌بینی کننده یک متغیر مالی بوده است. متغیر موجودی به فروش توانسته با ۲۸ درصد قدرت پیش‌بینی در صدر پیش‌بینی کننده های تقلب قرار گیرد. تنها متغیر غیر مالی در این الگو متغیر کیفیت حسابرسی بوده و توانسته ۵ درصد از تغییرات متغیر تقلب در صورتهای مالی را پیش‌بینی کند. سایر متغیرها به ترتیب توانسته اند ۲۶، ۱۶، ۱۲، ۹، ۵ و ۰۰۱ درصد از تغییرات متغیر وابسته را پیش‌بینی کنند.

جدول نتایج بکارگیری الگوریتم SVM-C5.0 برای پیش‌بینی تقبل در صورتهای مالی

INV/SAL <= 0.744 [Mode: 0] (726)
- AP/SAL <= 0.268 [Mode: 0] (534)
  - SAL/TA <= 1.373 [Mode: 0] (450)
    - NI/SAL <= 0.043 [Mode: 0] (69)
      - INV/SAL <= 0.074 [Mode: 1] ⇨ **1.0** (4; 1.0)
      - INV/SAL > 0.074 [Mode: 0] (65)
        - TL/TA <= 0.371 [Mode: 1] ⇨ **1.0** (3; 1.0)
        - TL/TA > 0.371 [Mode: 0] ⇨ **0.0** (62; 0.758)
    - NI/SAL > 0.043 [Mode: 0] (381)
      - NI/SAL <= 0.217 [Mode: 0] ⇨ **0.0** (186; 0.957)
      - NI/SAL > 0.217 [Mode: 0] (195)
        - TL/TA <= 0.537 [Mode: 0] ⇨ **0.0** (127; 0.921)
        - TL/TA > 0.537 [Mode: 0] (68)
          - INV/SAL <= 0.089 [Mode: 1] ⇨ **1.0** (5; 1.0)
          - INV/SAL > 0.089 [Mode: 0] (63)
            - AP/SAL <= 0.027 [Mode: 1] (11)
              - Quality Audit <= 0 [Mode: 0] ⇨ **0.0** (6; 0.667)
              - Quality Audit > 0 [Mode: 1] ⇨ **1.0** (5; 1.0)
            - AP/SAL > 0.027 [Mode: 0] ⇨ **0.0** (52; 0.885)
  - SAL/TA > 1.373 [Mode: 0] ⇨ **0.0** (84; 0.988)
- AP/SAL > 0.268 [Mode: 0] (192)
  - Quality Audit <= 0 [Mode: 0] (68)
    - INV/TA <= 0.215 [Mode: 0] (50)
      - SAL/AR <= 2.967 [Mode: 0] ⇨ **0.0** (44; 0.705)
      - SAL/AR > 2.967 [Mode: 1] ⇨ **1.0** (6; 1.0)
    - INV/TA > 0.215 [Mode: 0] ⇨ **0.0** (18; 0.944)
  - Quality Audit > 0 [Mode: 0] ⇨ **0.0** (124; 0.685)

INV/SAL > 0.744 [Mode: 1] (66)
- TL/TA <= 0.736 [Mode: 0] (40)
  - Quality Audit <= 0 [Mode: 0] ⇨ **0.0** (22; 0.864)
  - Quality Audit > 0 [Mode: 1] (18)
    - INV/TA <= 0.333 [Mode: 1] ⇨ **1.0** (4; 1.0)
    - INV/TA > 0.333 [Mode: 0] (14)
      - SAL/TA <= 0.482 [Mode: 0] ⇨ **0.0** (5; 1.0)
      - SAL/TA > 0.482 [Mode: 1] ⇨ **1.0** (9; 0.667)
- TL/TA > 0.736 [Mode: 1] (26)
  - NI/SAL <= 0.147 [Mode: 1] ⇨ **1.0** (24; 0.917)
  - NI/SAL > 0.147 [Mode: 0] ⇨ **0.0** (2; 1.0)



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم SVM-C5.0

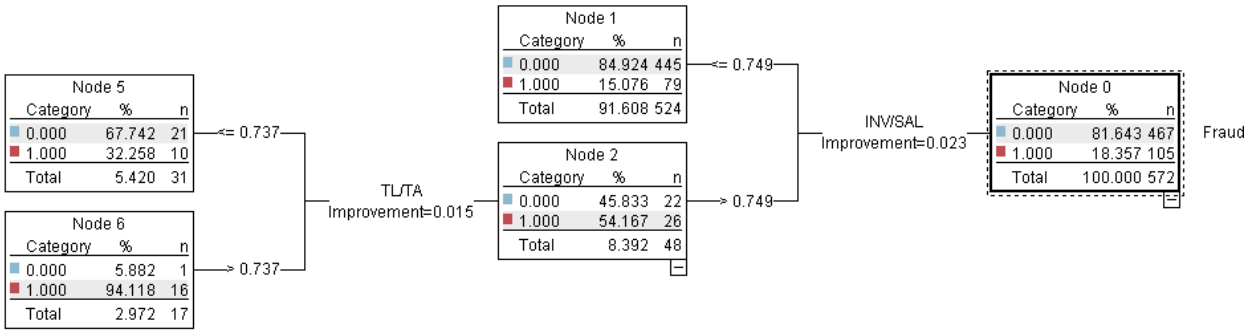
الگوریتم SVM-CART

نتایج برارزش الگوریتم CART برای شناسایی تقلب در صورت های مالی پس از انتخاب متغیرهای تاثیر گذار بر اساس الگوریتم ماشین بردار پشتیبان به صورت زیر است.

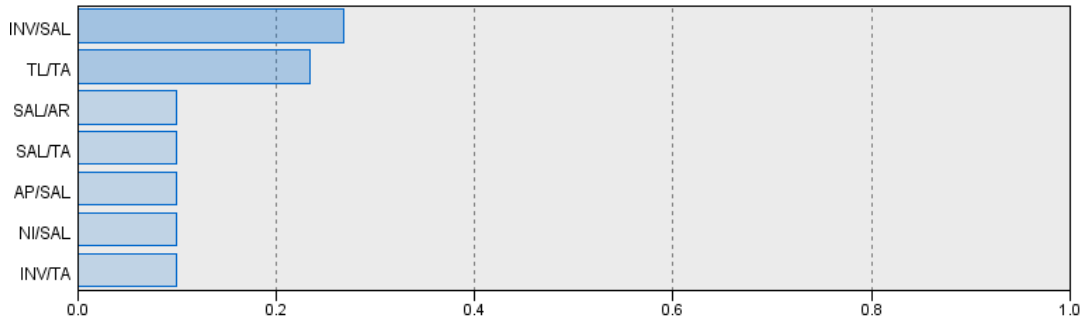
در مرحله قبلی۹ متغیر انتخاب شدند اما قرار نیست در این مرحله همه ۹متغیر قدرت تبیین داشته باشند. در این مدلسازی ۳مورد از متغیرها بطور کلی نتوانسته در پیش بینی حضور داشته باشد. . درواقع وارد مدل شده اما چون قدرت تبیین نداشته نرم افزار آنها را در مدلسازی کلا در نظر نگرفته

است. مشابه با نتایج مدل قبلی و همانطور که نتایج نشان می دهد، در این مدل نیز مهمترین متغیر پیش‌بینی کننده یک متغیر مالی بوده است. متغیر موجودی به فروش توانسته با ۲۷ درصد قدرت پیش‌بینی در صدر پیش‌بینی کننده های تقلب قرار گیرد. در این الگو متغیر غیر مالی حضور ندارد. سایر متغیرها به ترتیب توانسته اند ۲۷، ۲۳، ۱۰، ۱۰، ۱۰، ۱۰ و ۱۰ درصد از تغییرات متغیر وابسته را پیش‌بینی کنند.

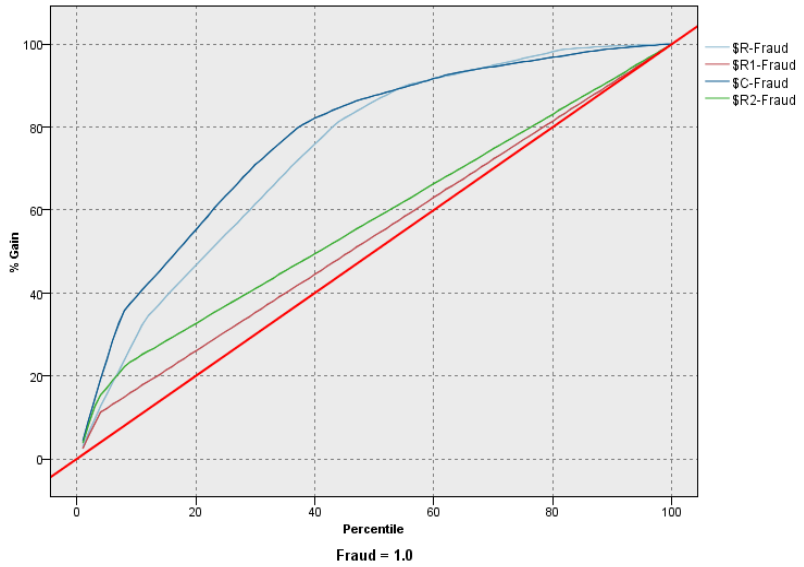
از میان متغیرهای مالی وغیر مالی، متغیرهای مالی نسبت موجودی به فروش (در دو الگوریتم از ۴ الگوریتم)و نسبت بدهی به دارایی ونسبت گردش دارایی در تشخیص تقلب عامل موثری شناخته شده است.



نمودار نتایج بکارگیری الگوریتم SVM-CART برای پیش‌بینی تقبل در صورتهای مالی



نمودار اهمیت متغیرهای مالی و غیر مالی در پیش‌بینی با الگوریتم SVM-CART



نمودار:شاخص GAIN برای مقایسه نتایج ۴ الگوریتم تحقیق با متغیرهای منتخب از طریق SVM

Individual Models

Comparing $R-Fraud with Fraud

Correct	651	82.2%
Wrong	141	17.8%
Total	792	

Comparing $R1-Fraud with Fraud

Correct	643	81.19%
Wrong	149	18.81%
Total	792	

Comparing $C-Fraud with Fraud

Correct	686	86.62%
Wrong	106	13.38%
Total	792	

Comparing $R2-Fraud with Fraud

Correct	657	82.95%
Wrong	135	17.05%
Total	792	

جدول مقایسه توان پیش‌بینی نتایج ۴ الگوریتم تحقیق با متغیرهای منتخب از طریق SVM

مقایسه چهار الگوریتم

نتایج تبیین شده در نمودار و جدول بالا نشان می دهد که از بین ۴ الگوریتم بکار گرفته شده الگوریتم C5.0 بالاترین قدرت تبیین، معادل ۸۶ درصد را دارد. این نتیجه نتایج بخش قبلی را تایید می کند. (از بالا به پایین اولین جدول مربوط به CHAID دومین جدول مربوط به QUEST سومین جدول، C5.0 و آخری مربوط به CART است).

جمع بندی

در این قسمت استفاده از الگوریتم ماشین بردار پشتیبان به منظور انتخاب متغیرهای پیش‌بینی کننده تقلب در صورت های مالی و بکارگیری الگوریتم‌های CHAID، QUEST، C5.0 و CART پرداخته است. نتایج مقایسه ای این الگوریتم ها که در جداول زیر آمده است C5.0 قویترین است

جدول FSF جدول مقایسه الگوریتم‌های بکار رفته پس از انتخاب متغیرها با روش ماشین بردار پشتیبان با پارامتر FSF

MODEL	FSF Detection Accuracy	Non-FSF Detection Accuracy	Overall Accuracy
SVM + C5.0	0.837	0.674	0.7555
SVM + CART	0.794	0.589	0.6915
SVM + CHAID	0.593	0.186	0.3895
SVM + QUEST	0.545	0.091	0.318

جدول مقایسه قدرت تبیین الگوریتم‌های بکار رفته پس از انتخاب متغیرها با روش شبکه عصبی مصنوعی

MODEL	Type I Error Rate	Type II Error Rate	Overall Error Rate
SVM + C5.0	0.165	0.115	0.16
SVM + CART	0.192	0.154	0.1705
SVM + CHAID	0.226	0.155	0.1845
SVM + QUEST	0.291	0.177	0.2225

در این بخش نتایج جدول معیار FSF نشان می دهد که قوی ترین الگو C5.0 است. پس از آن الگوی CART قرار دارد. سپس الگوریتم CHAID و الگوریتم QUEST ضعیف ترین الگوریتم است. نتایج مقایسه الگوریتم ها بر اساس معیار خطا در جدول آمده است. این نتایج تایید کننده نتایج جدول قبلی است. نتایج نشان می دهد که در الگوهای مبتنی بر ماشین بردار پشتیبان، متغیرهای غیر مالی نقش عمده ای در پیش‌بینی ندارند.

۷- بحث و نتیجه گیری : بحث و نتیجه گیری

چهار متغییر دربین متغیر های انتخاب شده توسط هر دو الگوریتم(SVM و ANN) در مرحله اول یکسان است که خود عوامل قدرتمندی در کشف تقلب است.در کل ۱۲ متغیر مهم هستند که از بین این ۱۲متغیر، چهارعامل قدرتمندتراز دیگرعوامل میباشند.که این ۴ متغیر با رنگ آبی در جداول زیرمشخص شده است

جدول متغییر های تاثیرگذار درتشخیص تقلب بوسیله ANN,SVM

الگوریتمANN	**SAL/TA** AP/SAL NI/EQ TL/TA OCF/SAL Opinion **COG/SAL**	**COG/SAL** **SAL/TA** **AP/SAL** **NI/EQ** **TL/TA** **OCF/SAL** **Opinion** **INV/SAL** **SAL/AR** **INV/TA** **NI/SAL** **Quality Audit**
الگوریتمSVM	**Opinion** **SAL/TA** **TL/TA** **INV/SAL** **SAL/AR** **AP/SAL** **INV/TA** **NI/SAL** **Quality Audit**	

بر اساس نتایج تجربی در مرحله دوم، مدل ANN+ C5.0 بالاترین میزان دقت تشخیص FSF و دقت کلی را در میان هشت مدل ساخته شده در این مطالعه برای تشخیص تقلب صورت‌های مالی گزارش می‌دهد. ANN+ C5.0 هم‌چنین دارای کم‌ترین میزان خطای نوع I و میزان خطای نوع II است. بنابراین، این مدل ، بهترین مدل ساخته شده برای تشخیص تقلب صورت‌های مالی در این مطالعه است.

این مطالعه با استفاده از تکنیک‌های استخراج داده برای شناسایی تقلب صورت‌های مالی شرکت‌ها بسیار متفاوت از مطالعات قبلی است. در مرحله اول، این مطالعه یک ANN و SVM را برای نمایش متغیرهای مهم به کار می‌برد. در مرحله دوم، برای ساخت مدل‌های تشخیص، چهار روش تکنیک تصمیم‌گیری (CART، CHAID، C5.0، و QUEST) مورد استفاده قرار می‌گیرند. برای تکنیک‌های استخراج داده مورد استفاده در این مطالعه، ANN و SVM برای انتخاب متغیرهای مهم مناسب هستند؛ CART، CHAID، C5.0، و QUEST برای طبقه‌بندی، پیش‌بینی و شناسایی متغیرها مناسب هستند. به منظور بهبود دقت تشخیص مدل‌ها، مجموع ۲۲ متغیر که ممکن است بر تقلب صورت‌های مالی تأثیر بگذارد انتخاب می شوند (از جمله متغیرهای مالی و غیرمالی؛). ۲۲ متغیر توسط ANN یا SVM نمایش داده می‌شوند و آن‌هایی که دارای اثر بیش‌تری هستند (مقدار اهمیت متغیر ≤ ۰/۰۵) انتخاب می شوند. ۲۲ متغیر به طور مستقیم از طریق CART، CHAID، C5.0، و QUEST مدل‌سازی می‌شوند. به عبارت دیگر، مدلی که تنها در یک مرحله ساخته شده بر مبنای مفاهیم اساسی آمار، دقت مدل بسیار کاهش یافته و سخت‌تر می‌شود. بنابراین، این مطالعه، این مدل را در دو مرحله ایجاد می‌کند..

این مطالعه هم‌چنین اعتبارسنجی متقابل ده برابری را که در محافل علمی به عنوان یک روش دقیق‌تر برای افزایش دقت تشخیص شناخته شده است می‌پذیرد. در میان هشت گزارش مالی، مدل‌های تشخیص تقلب ایجاد شده توسط این مطالعه، چند مورد از آن‌ها دارای دقت بیش از ۸۰ درصد می‌باشند. دقت تشخیص مدل‌ها در زمینه تحقیقات علوم اجتماعی خوب است همان‌طور که برای مدل ANN+ C5.0 رتبه اول در دقت کلی رتبه‌بندی شد)، مدل ANN+CART (رتبه دوم در دقت کلی)، CART (درخت طبقه‌بندی و رگرسیون) و C5.0 (توسعه درخت تصمیم‌گیری مبتنی بر آنتروپی است) دو تکنیک مهم الگوریتم درخت تصمیم‌گیری هستند. تحقیقات تجربی نیز دریافتند که توانایی طبقه‌بندی، پیش‌بینی و شناسایی هر دو تکنیک (CART و C5.0) بسیار خوب است. از سوی دیگر، در مورد نتایج تجربی این مطالعه، نمی‌توان به طور خودسرانه استدلال کرد که توانایی یا اثر استفاده از ANN برای فیلتر کردن متغیرهای مهم الزاماً بهتر از مدل SVM است، چرا که بستگی به وضعیت کلی ساخت مدل دارد. به عنوان مثال، در میان مدل‌های ایجاد شده توسط این مطالعه، مرحله I از ANN و مرحله دوم از چهار تکنیک الگوریتم درخت تصمیم‌گیری با دقت کلی استفاده می‌کند. با این حال، این مطالعه هم‌چنین درمی‌یابد که در میان ۸ مدل رتبه‌های توانایی طبقه‌بندی،

پیش‌بینی و شناسایی CART، CHAID، C5.0 و QUEST در مرحله دوم، عبارتند از: C5.0> CART > CHAID > QUEST براساس بحث فوق، این مطالعه نشان می‌دهد که مدل‌های تشخیص تقلب در حساب‌های مالی دقیق و مؤثر است. این هم چنین برای دیگر محققان یا متخصصان آموزنده است. هم‌چنین می‌تواند برای مطالعات آینده با استفاده از سایر تکنیک‌های استخراج داده برای تشخیص تقلب صورت‌های مالی شرکت‌ها مورد توجه قرار گیرد.گزارش‌های مالی، همه ذی‌نفعان در بازار مالی را با اطلاعات مفید در مورد وضعیت فعلی و چشم‌انداز شرکت ارائه می‌دهند. تقلب در صورت‌های مالی، تحریف تعمدی اظهارات یا نادیده گرفتن در ارقام یا پاورقی‌ها به منظور گمراه کردن کاربران است. هیئت مدیره می‌تواند مدیریت را برای رسیدن به اهداف درآمدی تحت فشار قرار دهد، یا مدیریت ممکن است اگر پاداش آن‌ها با درآمد گزارش شده مرتبط باشد در درآمدها به منظور اطمینان از سود خود تقلب کند. تقلب کردن در تهیه صورت‌های مالی ممکن است به مدیریت در منافع شخصی، مانند تبلیغات، افزایش حقوق و یا بالاتر بردن ارزش سهام شرکت کمک کند. شرکت‌ها گاهی اوقات اظهارنامه‌های مالی دروغین را برای دسترسی به تأمین مالی بدهی‌های بلند مدت یا حمایت از قیمت سهام عرضه می کنند. هر مورد تقلب در شرکت‌ها یک ضربه سنگین به سرمایه‌گذاران، طلب‌کاران و سهام‌داران است که بازار سرمایه را به خطر انداخته و بازارهای مالی را خراب می‌کند. تقلب در اظهارنامه‌های مالی به طور جدی به توسعه پایدار شرکت‌ها و بازارهای مالی آسیب می‌رساند. بنابراین، ایجاد یک مدل مؤثر برای تشخیص تقلب صورت‌حساب‌های مالی شرکت‌ها، موضوع مهم و فوری است.

اظهارنامه‌های مالی تقلبی شرکت‌ها، به ویژه شرکت‌های ذکر شده، نه تنها از ذی‌نفعان مانند سهام‌داران، سرمایه‌گذاران، سرمایه‌گذاران بالقوه، تأمین‌کنندگان، مشتریان و خریداران کلاه برداری می‌کنند، بلکه باعث خسارت به بقا و پایداری کاری شرکت‌ها و نیز سلامت بازارهای مالی و توسعه پایدار می‌شوند. به عنوان مثال، در سال ۲۰۰۱، انرون باعث آسیب جدی به اقتصاد آمریکا شد و به طور قابل توجهی بر توسعه پایدار بازارهای مالی آمریکا تأثیر گذاشت. حتی اگر حسابداران و حسابرسان با پیروی از قوانین و مقررات مربوطه برای شناسایی سوء رفتارهای جعلی از سوی مدیریت، صورت‌های مالی را بازبینی کنند، همیشه کوتاهی‌هایی خواهد شد. ایجاد یک مدل دقیق و مؤثر برای تشخیص تقلب‌های صورت‌های مالی شرکت ها، که می‌تواند توسط CPAها و حسابرسان اعمال شود، می تواند کلاه‌برداری‌های مالی شرکت‌ها و خطرات حسابرسی را برای ایجاد تأثیر بازدارنده بر مدیریت شرکت‌ها و کمک به شرکت‌ها و بازارهای مالی به شدت کاهش دهد. توسعه پایدار این مطالعه دقیق و مؤثر دو مرحله‌ای برای تشخیص صورت های مالی تقلبی می‌تواند به عنوان مرجع برای حسابداران، حسابرسان، تحلیل‌گران اوراق بهادار، تحلیل‌گران مالی، متخصصان دانشگاه و شرکت‌ها و ناظران بازارهای مالی استفاده شود. این اهمیت عملی و علمی این پژوهش است.

فهرست منابع

* پور حیدری ، بذر افشان ، ۱۳۹۱ .بررسی اثر ویژگی های فردی حسابرسان درتعیین اهمیت بسترهای خطر تقلب. فصلنامه . پژوهش های تجربی حسابداری مالی، ص ۵۲-۲۸.
* رستمی، محمد رضا، جعفری درگیری، اعظم، ۱۳۹۲. بررسی قانون بن فورد در بورس اوراق بهادر، پژوهشهای مدیریت در ایران. ۱۷ (۱): ۱۱۰-۹۵.
* شیخ، محمد جواد ۱۳۸۳، بررسی اهمیت شاخص‌های تقلب بالقوه مـالی بـا اسـتفاده ازتکنیک تحلیل عاملی در شرکتهای ایرانی، رساله دکتری حسابداری،دانشگاه تهران،دانشـکده مدیریت.
* فرقاندوست حقیقی، برواری، فرید.(۱۳۸۸). بررسی کاربرد روشهای تحلیلی در ارزیابی ریسک تحریف صورتهای مالی تقلب مدیریت،فصلنامه دانش وپژوهش حسابداری، شماره شانزدهم.
* Association of Certified Fraud Examiners (ACFE). (2014). Report to the nations on occupational fraud and abuse. Austin, TX: The association of certified fraud examiners, Inc .
* Association of Certified Fraud Examiners (ACFE). (2016). Report to the nations on occupational fraud and abuse. Austin, TX: The association of certified fraud examiners, Inc .
* Association of Certified Fraud Examiners. (2014). Fraud examiners manual. Austin, Texas. ACFE
* Ata A., Ibrahim H. Seyrek(2009). "the Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms, Suleyman Demirel University". The Journal of Faculty of Economics and Administrative Sciences, Vol. 14(2), pp. 157-170.
* Ata A., Ibrahim H. Seyrek, the Use of Data Mining Techniques in Detecting Fraudulent Financial Statements: An Application on Manufacturing Firms, Suleyman Demirel University, The Journal of Faculty of Economics and Administrative Sciences, Vol. 14(2), 2009, pp. 157-170
* Autore, D. M., Hutton, I., Peterson, d. R., & Smith, I. H. (2014). The effect of securities litigation on

* Carpenter, T. D., & Reimers, J. L. (2005). Unethical and fraudulent financial reporting: Applying the theory of planned behavior. Journal of Business Ethics, 60(2), 115129. doi: 10.1007/s10551-004-7370-9
* Cecchini, M., Aytug, H., Koehler, G. J., and Pathak, P. (2010). Making words work: using financial text as a predictor of financial events. Decision Support Systems, 50(1), 164-175 .
* Chen W.S. and Y.K. Du(2010). "Using Neural Networks and Data Mining Techniques for The Financial Distress Prediction Model". Expert Systems with Applications, Vol. 36, pp.4075–4086. Elliott R.K., and J.J. Willingham(1980). Management Fraud: Detection and Deterrence, Petro celli Books, NewYork, p.4.
* Chen W.S. and Y.K. Du, Using Neural Networks and Data Mining Techniques for The Financial Distress Prediction Model, Expert Systems with Applications, Vol. 36, pp.4875– 4886, 2889.
* Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. SpringerPlus, 5(1), 89.
* Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. Springerplus, 589. doi:10.1186/s40064-016-1707-6
* Clikeman, P. M., & Diaz, J. (2014). ABC Electronics: An Instructional case illustrating auditors' use of preliminary analytical procedures. Current Issues In Auditing, 8(1), I1-I10. doi:10.2308/ciia-50775
* Cohen, J., Ding, Y., Lesage, C., & Stolowy, H. (2010). Corporate fraud and managers' behavior: Evidence from the press. J Bus Ethics Journal of Business Ethics, 95(2), 271-315. doi: 10.1007/s10551-011-0857-2
* Peytcheva, M. (2014). Professional skepticism and auditor cognitive performance in a hypothesis-testing task. Managerial Auditing Journal, 29(1), 27-49. doi: http://dx.doi.org/10.1108/MAJ-04-2013-0852
* Peytcheva, M., & Warren, D. E. (2013). How auditors perceive sanction severity and the detection of violations: Insights into professional vulnerabilities. Accounting & The Public Interest, 13(1), 1-13. doi:10.2308/apin-10343
* Plumlee, R. D., Rixom, B. A., & Rosman, A. J. (2015). Training auditors to perform analytical procedures using metacognitive skills. Accounting Review, 90(1), 351369. doi:10.2308/accr-50856
* Polit-O'Hara D, Beck CT. Essentials of nursing research: Methods, appraisal, and utilization. Philadelphia: Lippincott Williams & Wilkins; 2006.
* Power, M. (2013). The apparatus of fraud risk. Accounting, Organizations & Society, 38(6/7), 525-543. doi:10.1016/j.aos.2012.07.004
* Public Company Accounting Oversight Board (PCAOB). (2016). PCAOB inspections. Retrieved from: https://pcaobus.org/inspections/Pages/default.aspx
* Purda, L., & Skillicorn, D. (2015). Accounting variables, deception, and a bag of words: assessing the tools of fraud detection. Contemporary external financing, Journal of Corporate Finance, 27, 231-250. Doi:10.1016/j.jcorpfin.2014.05.007 .
* Azrina, M. Y., & Ling Lai, M. (2014). An integrative model in predicting corporate tax fraud. Journal of Financial Crime, 21(4), 424 .
* based on data-mining techniques. IJCA International Journal of Computer Applications, 39(1), 37-47 .
* Beasley, M.S., Carcello, J.V., Hermanson, D.R., and Neal, T.L. (2010). Fraudulent financial reporting 1998-2007: An Analysis of U.S. Public Companies. Retrieved from http://www.coso.org/documents/COSOFRAUDSTUDY2010_001.pdf
* Bei, Z., & Wijewardana, W. P. (2012). Working capital policy practice: Evidence from Sri Lankan companies. Procedia-Social and Behavioral Sciences, 40, 695-700 .
* Beneish, M. D. (1997). Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance. Journal of Accounting and Public Policy Journal of Accounting and Public Policy, 16(3), 271309 .
* Bhattacharjee, S., Maletta, M. J., & Moreno, K. K. (2016). The Role of Account Subjectivity and Risk of Material Misstatement on Auditors' Internal Audit Reliance Judgments. Accounting Horizons, 30(2), 225-238. doi:10.2308/acch51363 Bishop, C. C., Hermanson, D. R., & Houston, R. W. (2013). PCAOB Inspections of International Audit Firms: Initial Evidence. International Journal Of Auditing, 17(1), 1-18. doi:10.1111/j.1099-1123.2012.00453.x
* Blay, A. D., Sneathen Jr, L. D., & Kizirian, T. (2007). The effects of fraud and going-concern risk on auditors' assessments of the risk of material misstatement and resulting audit procedures. International Journal of Auditing, 11(3), 149-163.
* Bonner, S.E., Palmrose, Z.V. and Young, S.M. (۱۹۹۸) , ``Fraud type and auditor litigation. An analysis of SEC accounting and auditing enforcement releases'', The Accounting Review, Vol. ۷۳ No. ٤, pp. ٥٠٣-٣٢.
* Boritz, J. E., Kochetova-Kozloski, N., & Robinson, L. (2015). Are fraud specialists relatively more effective than auditors at modifying audit programs in the presence of fraud risk? Accounting Review, 90(3), 881-915. doi:10.2308/accr- 50911
* Boučková, M. (2015). Management Accounting and Agency Theory, Procedia Economics and Finance, 25, 5-13. doi: 10.1016/S2212-5671(15)00707-8 .
* Boyle, D. N., Carpenter, B. W., & Hermanson, D. R. (2012). CEOs, CFOs, and Accounting Fraud. CPA Journal, 82(1), 62-65.
* Bui, T., & Amaria, P. (2014). How do financial analysts in Vietnam perceive the relationship between the corporate board structure and financial statement fraud? International Journal of Business, Accounting, and Finance, 8 (1), 1-23 .
* Burton, F. G., Wilks, T. J., & Zimbelman, M. F. (2013). How auditor legal liability influences the detection and frequency of fraudulent financial reporting. Current Issues in Auditing, 7 (2), 9-15. doi:10.2308/ciia-50566

prevention, roles of technology, and white-collars getting splattered with red! Journal Of Accounting & Finance (2158-3625), 16(1), 26-44 .
* Simon N,(1881), Note on the Frequency of Use of the Different Digit in Natural numbers, American Journal of Mathematics, No. 4, pp. 39-40
* Simon, C. A. (2012). Individual auditors' identification of relevant fraud schemes. Auditing: A Journal Of Practice & Theory, 31(1), 1-16. doi:10.2308/ajpt-10169
* Sirikulvadhana. S (2002). Data Mining As A Financial Auditing Tool. Available on line At: Pafis.Shh.Fi/Graduates/Supsir01.Pdf.
* Smith, G. S. (2015). The Past, present, and future of forensic accounting. CPA Journal, 85(3), 16-21 .
* Solomon, S. J. (2015). Let the numbers talk: A look into the use of the Benford's analysis as a method of detecting data fraud. Academy Of Management Annual Meeting Proceedings, 1. doi:10.5465/AMBPP.2015.15089
* Song, X., Hu, Z., Du, J., & Sheng, Z. (2014). Application of machine learning methods to risk assessment of financial statement fraud: Evidence from china. Journal of Forecasting, 33(8), 611-626. doi:10.1002/for.2294
* Stančić, P., Dimitrijević, D., & Stančić, V. (2013). Forensic accounting - the response of the profession to financial statement frauds. TEME: Casopis Za Društvene Nauke, 37(4), 1879-1897 .
* Summers, S.L. and Sweeney, J.T. (۱۹۹۸) , ``Fraudulently misstated financial statements and insider trading: an empirical analysis'', The Accounting Review, Vol. ۷۳ No. ۱, pp. ۱۳۱-٤٦.
* Sun, J., & Liu, G. (2016). Does analyst coverage constrain real earnings management? The Quarterly Review of Economics and Finance, 59, 131-140. doi:10.1016/j.qref.2015.03.009
* Syriopoulos, T., & Tsatsaronis, M. (2012). Corporate governance mechanisms and financial performance: CEO duality in shipping firms. Eurasian Business Review, 2(1), 1-30 .
* Tabuena, J. (2014). Just what is "reasonable assurance"? Compliance Week, 11(121), 38-39 .
* Tamari, M. (1966). Financial ratios as a means of forecasting bankruptcy. Management International Review, 4 .
* Whitney, M. P. (2015). Review of Neurobiology and the development of human morality: Evolution, culture, and wisdom. Journal Of Moral Education, 44(4), 537-539. doi:10.1080/03057240.2015.1069479

یادداشت‌ها

[1] Tangod
[2] Alden
[3] Lin
[4] Jan, C. L. (2018). An effective financial statements fraud detection model for the sustainable development of financial markets: Evidence from Taiwan. Sustainability, 10(2), 513.
[5] Accuracy
[6] Error

Accounting `Research, 32(3), 1193-1223. doi:10.1111/1911-3846.12089
* Rauch, B., Göttsche, M., Brähler, G., & Kronfeld, T. (2014). Deficit versus social statistics: Empirical evidence for the effectiveness of the Benford's law. Applied Economics Letters, 21(3), 147-151. doi:10.1080/13504851.2013.844319
* Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data-mining techniques. Decision Support Systems, 50(2), 491-500. doi:10.1016/j.dss.2010.11.006
* Reffett, A. B. (2010). Can identifying and investigating fraud risks increase auditors' liability? The Accounting Review, 85(6), 2145-2167 .
* Robu, I., & Robu, M. (2013). Proceduri de audit pentru estimarea riscului de frauda bazate pe indici de detectare a manipularilor contabile/Audit procedures for estimating the fraud risk based on indexes for detection of accounting manipulation. Audit Financiar, 11(10), 3-16 .
* Roxas, M. L. (2011). Financial statement fraud detection using ratio and digital analysis. Journal of Leadership, Accountability & Ethics, 8(4), 56-66 .
* Ruankaew, T. (2016). Beyond the fraud diamond. International Journal Of Busines Management & Economic Research, 7(1), 474-476 .
* Rui. Hort, Beatriz De S. L. Pires De Lima And Carlos Cristiano H. Borges (2009).Data Preprocessing Of Bankruptcy Prediction Models Using Data Mining Techniques. Available on line At: Www.Blog.Campe.Com.Br/Wp-Content/Uploads/2009/03/Witpress_Conf-2.Pdf.
* Sabau, A. S. (2012). Survey of clustering based financial fraud detection research. Informatica Economica, 16(1), 110-122 .
* Salisbury, L. T. (2013). Lack of professional skepticism reason audits fail to pass muster, says tempro. Accounting Policy & Practice Report, 9(20), 817-818 .
* Salkind, N. J. (2012). Exploring research. Boston: Prentice Hall
* Santos, J. B., & Ledur Brito, L. A. (2012). Toward a subjective measurement model for firm performance. BAR - Brazilian Administration Review, 9(2), 95-117 .
* Santoso, N. T. (2018). Predicting Financial Statement Fraud with Fraud Diamond Model of Manufacturing Companies Listed in Indonesia. In State-of-the-Art Theories and Empirical Evidence (pp. 151-163). Springer, Singapore.
* Schuchter, A., & Levi, M. (2015). Beyond the fraud triangle: Swiss and Austrian elite fraudsters. Accounting Forum, 39(3), 176-187. doi:10.1016/j.accfor.2014.12.001
* Shamki, Dhia'a, (2009), "Internal Audit Responsibilities in Auditing Financial Systems Fraud", Business e-Bulletin, Vol. 1, Issue 1, PP. 25-32
* Sharma, A. K., & Panigrahi, P. (2012). A review of financial accounting fraud detection
* Simha, A., & Satyanarayan, S. (2016). Straight from the horse's mouth: Auditors' on fraud detection and