فصلنامه چشم انداز حسابداری و مدیریت
دوره ۴، شماره ۴۵، تابستان ۱۴۰۰(جلد چهارم)

www.jamv.ir
ISSN: 2645-4572

چشم انداز حسابداری و مدیریت (دوره چهارم)
شماره ۴۵ / تابستان ۱۴۰۰ (جلد چهارم)/ صص ۱۱۳ -۱۲۳

بهره گیری از رویکرد داده کاوی و شناسایی تقلب در صورت های مالی

دکتر فرهاد دهدار
استادیار گروه حسابداری، واحد شاهرود، دانشگاه آزاد اسلامی، شاهرود، ایران.
dehdar1970@yahoo.com
محمد حسین مرادی
دانشجوی دکتری حسابداری، واحد شاهرود، دانشگاه آزاد اسلامی، شاهرود، ایران. (نویسنده مسئول).
mohamadhosinmoradi66@gmail.com

چکیده
تقلب از پدیده های رایج و متداول در کسب و کار است. پیشگیری با کشف تقلب های با اهمیت در صورت های مالی همواره کانون توجه سرمایه گذاران، قانونگذاران، استاندارد گذاران، مدیران و حسابرسان بوده است. بر عصر حاضر عصر اطلاعات نام نهاده شده و آنچه که منجر به قدرت و موفقیت می شود، اطلاعات است. داده کاوی پل ارتباطی میان علم آمار، علم کامپیوتر، هوش مصنوعی، الگوشناسی، فراگیری ماشین و بار نمایی بصری داده ها می باشد که در چارچوب فرآیندی، استخراج اطلاعات معتبر، از پیش ناشناخته، قابل فهم و قابل اعتماد از پایگاه داده های بزرگ امکان پذیر می گردد و استفاده از آن در تصمیم گیری در فعالیت‌های تجاری مهم نظیر ارتقای کیفیت سودمندی اطلاعات از طریق شناسایی تقلب های مالی میسر می شود. تحقیق حاضر به روش شناخت تاریخی با بهره گیری از اسناد کتابخانه ای و به پشتوانه پیشینه و تحقیقات محققان شواهدی لازم جهت بررسی داده کاوی در تقلب صورت های مالی ارائه می کند.
واژگان کلیدی: تقلب مالی، داده کاوی، درخت تصمیم گیری، شبکه های عصبی، شبکه های گمانه زنی بیزی.

مقدمه
اهداف گزارشگری مالی از نیازها و خواسته های اطلاعاتی استفاده کنندگان برون سازمانی سرچشمه می گیرد. هدف اصلی بیان آثار اقتصادی رویدادها و عملیات مالی بر وضعیت و عملکرد واحد تجاری برای کمک به اشخاص خارجی برای اتخاذ تصمیمات مالی در ارتباط با واحد تجاری است (ثقفی و عرب مازیار یزدی[1]، ۲۰۱۰). اما امروزه ورشکستگی، تقلب و عدم موفقیت شرکت ها همواره مسئله ای پیچیده و درخور توجه بوده است (زارع بهنمیری و ملکیان کله بستی، ۱۳۹۴). شواهد زیادی از حساب سازی در به کارگیری حسابداری وجود دارد و به صورت گسترده از آن استفاده می شود که این امر به عنوان پدیده رایج و مخرب کارایی، نقدینگی و امنیت بازارهای مالی جهانی، بازارهای سرمایه و بدهی به دلیل وجود تقلب در صورت‌های مالی که عامل اساسی در بحران مالی اخیر و رکود اقتصادی جهانی می‌باشد، تهدید می‌کند (اسپاتیس[2] و همکاران، ۲۰۰۷).

[1] Saghafi, A, Arabmazar Yazdi
[2] Spathis

www.jamv.ir
ISSN: 2645-4572

کوسیوارا سال ۲۰۰۲ را از نقطه نظر دفترداری و ادعاها که دارای دستکاری هایی هست که همچنان ادامه دارد سال ناگوار نامید (کوسیوارا[۱] ،۲۰۰۴). بعضی گزارشات هزینه تقلب کسب و کارهای ایالات متحده آمریکا را بیش از ۴۰۰ بیلیون دلار در سال برآورد کرده است (ولز[۲]، ۱۹۹۷).

اسپاتیس[۳] و همکاران (۲۰۰۲) اعتقاد دارند که صورت های مالی متقلبانه به صورت فزاینده ای به طور مکرر در طی سالیان اخیر در حال ازدیاد می باشد.

در دنیای کنونی تمایلات نامحدود بشر در مقابل منابع محدود اقتصادی قرار گرفته است. ظهور و سقوط هر پدیده‌ای ریشه در نیازهای واقعی و منطقی جوامع انسانی دارد؛ ظهور حسابرسی تقلب در عرصه خدمات حرفه ای نیز نمی‌تواند از این قاعده مستثنی باشد. در عصر حاضر پیشرفت فناوری و تغییرات گسترده در محیط تجاری موجب ایجاد شتاب فزاینده در اقتصاد شده است و به علت رقابت روز افزون موسسه ها دستیابی به درآمد مورد انتظار نیز محدود شده است بنابراین زمینه بروز تقلب روز به روز افزایش می‌یابد (جامعی و اصغرزاده[۴]، ۲۰۱۲). تقلب، هر اقدام عمدی یا حذف و از قلم انداختن طرح ریزی شده است که به منظور فریب یا اغفال دیگران صورت می‌گیرد و باعث می‌شود قربانیان آن دچار زیان و عاملان و مرتکبان آن، منافع را به دست آورند. تمام افراد جامعه فارغ از هرگونه فرهنگ و مذهب یا هر ویژگی دیگری در معرض وسوسه ارتکاب به تقلب هستند (رحیمیان و آخوندزاده[۵] ، ۲۰۱۱). تقلب در گزارش مالی می‌تواند تاثیر مخرب بر گزارش شرکت داشته باشد تا نقطه ای که ماهیت شرکت مخاطره افتد. طبق بیانیه تقلب در انگلیس و ولز در سال ۲۰۰۶ تقلب اساساً شامل استفاده از نیرنگ برای نفع شخصی و یا آسیب به دیگری است. تلاش‌های بسیاری برای سنجش میزان واقعی تقلب صورت گرفته است اما جمع‌آوری آمار قابل قبول در مورد تقلب آسان نیست اکثریت تقلب ها آشکار نشده و حتی وقتی که تقلبی آشکار شود، شاید گزارش نشود؛ به این دلیل که شرکتی که قربانی تقلب است نمی‌خواهد چهره ای مخدوش در بین عموم مردم داشته باشد. در تحقیق جرایم اقتصادی جهانی در سال ۲۰۰۷ بیان شده است که بیش از ۴۳ درصد کسب و کارهای بین‌المللی قربانی تقلب در طی دو سال قبل بودند. در انگلیس این رقم بیش از میانگین جهانی است، ٪۴۸ شرکت‌ها در دام تقلب گرفتار شده‌اند (موسسه حسابداران مدیریت[۶] ،۲۰۰۸). افزایش شمار تقلب در گزارشگری مالی و موارد تجدید ارائه صورت های مالی که اغلب با ورشکستگی شرکت‌های بزرگ در آمیخته نگرانی‌هایی را درباره کیفیت گزارشگری صورت های مالی بر انگیخته است. به همین دلیل پیشگیری یا کشف تقلب های با اهمیت در گزارشگری مالی همواره کانون توجه سرمایه گذاران قانونگذاران مدیران و حسابرسان بوده است (خانی[۷]، ۲۰۰۷).

AICPA به صراحت تصدیق کرده که حسابرسان در برابر تقلب های کشف شده پاسخگو هستند (کولینا و سوتو[۸]، ۲۰۰۲). از این گذشته برای توجه به انتظارات از ایشان حسابرسان تکنیک های تحلیلی را برای مرور کارشان به خدمت می‌گیرند چیزی که به آنها اجازه می‌دهد مانده حساب ها را بدون آزمون مرتبط با تراکنش منحصر به فرد آن را برآورد سازند. فراسر[۹] و همکاران (۱۹۹۷) تکنیک های تحلیلی احتمالی اجمالی را به صورت غیر کمی, نمونه کمی و پیشرفته

[1] Koskivaara
[2] Wells
[3] Spathis
[4] Jameie, R. And Asgharzadih, P
[5] Rahimian, N and Akhundzade, M
[6] Chartered Institute of Management Accountants
[7] Khani
[8] Cullinan, P. G., & Sutton, G. S
[9] Fraser

www.jamv.ir
ISSN: 2645-4572

کمی طبقه‌بندی نمودند. تکنیک های پیشرفته کمی دربردارنده روش های پیچیده به کارگیری آمار و هوش مصنوعی مانند شبکه عصبی و تحلیل رگرسیون بود.

داده کاوی[1] یک فرایند تکراری است که در آن، فرآیند به وسیله پوشش آن تعریف می‌شود. همچنین می‌تواند با روش‌های کاملاً خودکار یا دستی انجام شود. داده کاوی می تواند بیشترین سودمندی را در یک زمینه راهنمای طرح تحلیلی اکتشافی داشته باشد در جایی که هیچ تصور قبلی در مورد چیزی که کسب خواهیم کرد تشکل نیافته باشد (کانتارزیک[2]، ۲۰۰۲). بکارگیری تکنیک های داده کاوی برای طبقه بندی های مالی یکی از حوزه‌های پرثمر تحقیقاتی است، بسیاری از قوانین التزامی و واحدهای تحقیقاتی ویژه، افرادی را مامور شناخت فعالیت‌های متقلبانه در اختیار دارند که با موفقیت داده‌کاوی را به کار می‌برند هرچند در مقابل سایر زمینه های آزموده شده مثل پیش بینی ورشکستگی و پریشانی مالی، تحقیقات کاربرد تکنیک های داده کاوی به منظور کشف تقلبات مدیریتی در حداقل قرار دارد (کالدرون و چه[3]، ۲۰۰۲).

در تبیین مبانی داده کاوی یاس (۱۳۹۰) عنوان می‌کند داده کاوی پل ارتباطی میان علم آمار، علم کامپیوتر، هوش مصنوعی، الگوشناسی، فراگیری ماشین و بازنمایی بصری داده می‌باشد و فرآیند استخراج اطلاعات معتبر از پیش ناشناخته، قابل فهم و قابل اعتماد از پایگاه داده های بزرگ و استفاده از آن در تصمیم گیری در فعالیت‌های تجاری مهم است. بنابراین داده کاوی به صورت یک محصول قابل خریداری نمی باشد بلکه یک رشته علمی و فرایندی است که با بایستی به صورت یک پروژه پیاده سازی شود. اصولاً هر جایی که داده وجود داشته باشد داده کاوی نیز معنا می‌یابد این فرآیند دارای مزایای با اهمیت است از جمله:

- ✓ کمک به مدیران در تصمیم‌گیری
- ✓ یافتن مقالات خبری روزآمد
- ✓ کمک به کاربر برای جستجو در وب
- ✓ بهبود امر بازاریابی
- ✓ کشف کلاهبرداری‌ها، تقلب و نابهنجارها

اساس داده کاوی بر مبنای سه فعالیت اصلی است که عبارتند از:

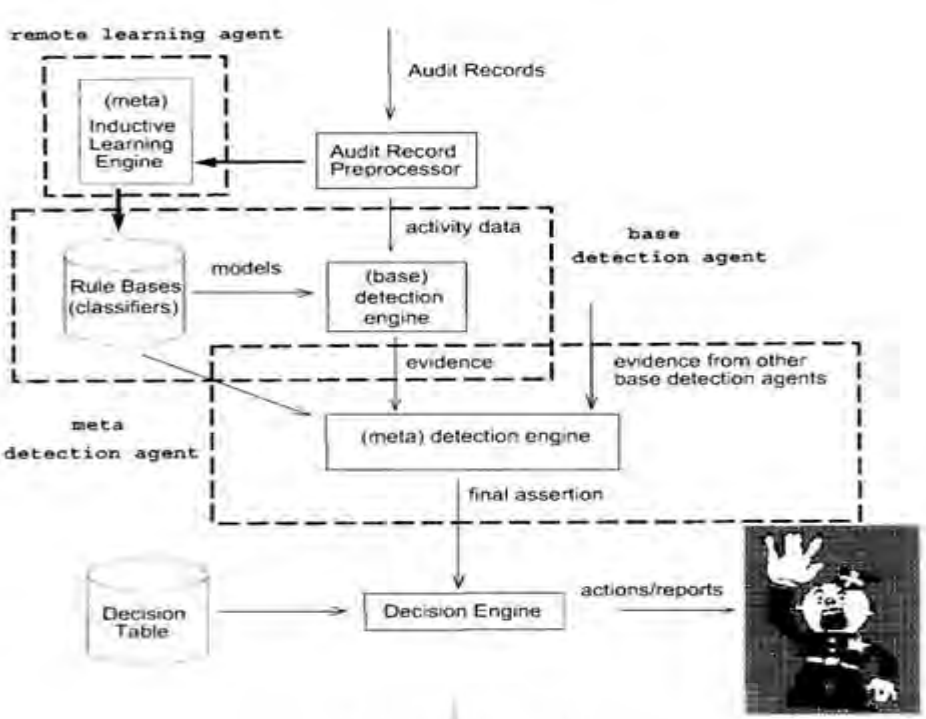
الف) حذف داده ها: داده های بی ارزش و عوامل بیرونی حذف می‌شوند.

ب) فشرده سازی داده ها: این عمل به وسیله کدگذاری داده ها صورت می گیرد.

ج) کشف الگوها: الگوهای موجود در پایگاه داده ها از قبیل طبقه بندی الگوهای زنجیری و غیره کشف می‌شوند.

لی و استلف (۱۹۹۸) چارچوبی تحت عنوان معماری داده کاوی به عنوان نظام پایه کشف تقلب ارائه می‌دهد که در شکل شماره ۱ ارائه می شود (رهنمای رودپشتی،۱۳۹۱).

[1] Data Mining
[2] Kantardzic
[3] Calderon, T. G., & Cheh, J.



شکل ۱

در این مطالعه، سه تکنیک داده کاوی برای کاربرد در کشف تقلبات مدیریتی نظیر درخت تصمیم گیری، شبکه عصبی، شبکه گمانه‌زنی بیزی تبیین می گردد. چن[۱] (۲۰۱۶) علاوه بر سه تکنینک فوق از تکنیک دستگاه بردار پشتیبان[۲] نیز بهره گرفته است.

درخت تصمیم گیری

یک درخت تصمیم گیری[۳](DT) یک ساختار درختی است که هر گره آن یک آزمون برای یک صفت را نشان می دهد و هر انشعاب تشان دهنده یک خروجی از آزمون می باشد. در این طریق، هر شاخه تلاش می کند مشاهدات را در زیر گروه هایی به صورت دو به دو ناسازگار تقسیم کند. خوبی یک انشعاب بر مبنای انتخاب صفاتی است که بهترین جداسازی نمونه را انجام دهد. نمونه ها با موفقیت به دسته هایی تقسیم می شوند، یعنی تا زمانی که هر یک از دو تا دیگر مسیر تقسیم شدن را دیگر طی ننمایند، می توان اختلاف معنی دار را به طور آماری با زیر گروه ها برای ایجاد تقسیم معنی دار مشابه بسیار کوچک تولید کرد. چندین الگوریتم تقسیم پیشنهاده شده است. در ردیابی یا بازیابی متقابل خودکار[۴](AID) بیشترین آماره t معنی دار به عنوان هر تحلیل واریانس استفاده می شود. در مربع AID chi، آمارهٔ chi و طبقه بندی و رگرسیون درختی[۵] (CART) برای استفاده از شاخص تفاوت ها بکار رفته است (کوه و لو[۶]،۲۰۰۴).

تقسیم های پی در پی نمونه ممکن است یک درخت بزرگ تولید کنند. برخی از شاخه های درخت ممکن است بازتاب خلاف قاعده ای از دسته آموزشی باشد مانند ارزش های نادرست با طرح های کلی نگر. به همین دلیل شاخه به میزان

[1] Chen
[2] support vector machine
[3] Decision Tree
[4] Automatic Interaction Detection
[5] Classification and Regression Trees
[6] Koh, H. C., & Low, C

www.jamv.ir
ISSN: 2645-4572

مورد نیاز هرس می شود. هرس شاخه یا حذف نمره های تقسیم شده در راهی است که در نرخ درستی مدل معنی دار نمی باشد. برای تعیین طبقه هدف مشاهده نشده پیشین، وزن صفت هدف دوباره در گره های تصمیم گیری شاخه آزموده می شود. براساس این آزمون یک مسیر پیموده می شود که به هدف طبقه در پیشگویی ختم خواهد شد. مزیت اصلی درخت تصمیم گیری این است که یک راه معنی دار با نمایش دانش کسب شده تدارک می بیند و قواعد طبقه بندی اگر آنگاه را به آسانی استخراج می کند.

شبکه های عصبی

شبکه های عصبی[1](NN) یک فن آوری کامل در اجرای یک تـئوری و شناخت حوزه های کاربردی آن است. یک NN متشکل از تعداد سلول عصبی و امثال آن بصورت واحدهای پردازش به هم پیوسته است. اشتراک در هر اتصال با یک ارزش عددی است که وزن نامیده می شود. هر سلول عصبی علامتی از اتصال سلول عصبی می گیرد و ترکیب علامت های ورودی محاسبه می شود. مجموع علامت ورودی برای سلول عصبی j به صورت

$$u_j = \sum w_{ij} \times x_i$$

است. چنانکه X علامت ورودی از سلول عصبی i و W وزن اتصال بین سلول عصبی i و j می باشد (رهنمای رودپشتی، ۱۳۹۱).

اگر قدرت علامت ورودی ترکیب شده بیش از یک آستانه ای باشد در این صورت ارزش ورودی به وسیله بخش ویژه انتقال شبکه عصبی انتقال داده می شود و در پایان سلول عصبی برانگیخته می گردد (هان و کمبر[2]، ۲۰۰۰).

سلول های عصبی در لایه هایی چیده می شوند. هر لایه شبکه متشکل از یک ورودی (اولیه) و یک خروجی (ثانویه) به عنوان کوچکترین جزء هر لایه است. در بین ورودی و خروجی لایه ممکن است یک لایه مخفی یا بیشتر وجود داشته باشد. اختلاف انواع NNها در اختلاف تعداد لایه های آنها است. طرح های سازمان یافته طبیعی[3](SOM) تنها یک لایه ورودی و خروجی دارند. در حالی که یک NN پس انتشار یک یا چندین لایه مخفی بیشتر دارد.

بعد از اینکه معماری شبکه مشخص شد، شبکه باید آموزش ببیند، در شبکه های پس انتشار یک الگو برای لایه ورودی و محاسبه خروجی نهایی در لایه خروجی بکار می رود. خروجی با نتیجه مورد درخواست مقایسه شده و خطاها به عقب NN انتشار می یابد تا وسیله تنظیم وزن اتصال ها قرار گیرد. این فرایند تا زمانی تکرار می شود که به یک نرخ خطای قابل پذیرش رسید. شبکه های عصبی که هیچ فرضی در مورد صفات مستقل ندارند، قابلیت رسیدگی به داده های مخدوش یا متناقض را ندارند و پیشنهاد انعطاف پذیر مناسب بر مشکلاتی هستند که یک الگوریتم چاره ساز در آن کارایی ندارد. از آنجایی که NNهای پس انتشار بیشترین عمومیت را به عنوان یک پیش بینی کننده و طبقه بندی کننده مشکلات دارد (سهل و ون کاتاچالام[4]، ۱۹۹۵).

شبکه های گمانه زنی بیزی[5] (شبکه های احتمالات شرطی)

[1] Neural Networks
[2] Han, J., & Camber, M
[3] Self-organizing maps
[4] Sohl, J. E., & Venkatachalam, A.
[5] Bayesian Belief Networks

www.jamv.ir
ISSN: 2645-4572

طبقه بندی بیزی بر مبنای قضیه آماری بیز است. قضیه بیز یک خوشه بندی از احتمالات پس گرا تدارک می بیند. بر اساس قضیه بیز، اگر H یک فرضیه باشد– مانند اینکه شی (یا موضوع) X به طبقه C تعلق داشته باشد آنگاه احتمال جا افتادن (تحقق) فرضیه به صورت

$$P(H|X) = (P(X|H) \times P(H))|P(X)$$

می باشد.

اگر یک شئ (یا موضوع) X به یک گروهی از طبقات متناوب i تعلق داشته باشد برای تعیین طبقه بندی شی در یک خوشه بندی طبقه بندی شده بیزی، احتمالات ممکن به صورت $P(C_i|X)$ برای تمام طبقات Ci ممکن بوده و بیشترین احتمال تخصیص شی (یا موضوع) به طبقه ای خاص می باشد (رهنمای رودپشتی، ۱۳۹۱).

طبقه بندی های بیزی ساده با فرض حالت مستقل ساخته می شود که این در علم آمار از یک ارزش صفت داده شده در یک طبقه که مستقل از ارزش سایر صفات است نتیجه می شود. فرض ساده خوشه بندی به صورت $P(C_i|X)$ است. اگر این فرض درست انگاشته شود طبقه بندی ساده بیزی بهترین برآورد صحیح را در مقایسه با بقیه طبقه بندی ها خواهد داشت. هر چند در بسیاری از موارد این فرض تایید نمی شود چون که وابستگی بین صفات می تواند وجود داشته باشد.

شبکه های گمانه زنی بیزی اجازه گزارش مقادیر وابسته زیر مجموعه ای از صفات را می دهد. یک BNN نمودار ناچرخه ای (مارپیچی) مستقیم است که هر گاه یک صفت را نشان می دهد و هر فلش نشان دهنده وابستگی احتمالی است. اگر یک فلش به صورت اگر گره A به گرۀ B رسم شود، آنگاه A پدر B و B زاده A است. در یک شبکه گمانه زنی هر متغییر شرطی مستقل از نازاده آن که توسط پدرش داده شده، می باشد (هان و کمبر، ۲۰۰۰).

برای هر گره X موجود یک جدول احتمالات شرطی نیز موجود است، هم چنانکه احتمالات شرطی ویژه برای هر ارزش X برای هر ترکیب ارزش های ممکن پدرانشان نیز موجود است: توزیع شرط .P(X|Parents(X))

احتمال یک مجموعه ,n(X1, X2, X3, … , Xn) توزیع دارند که به این شکل است.

ساختار یک شبکه را می توان از پیش تعریف نمود یا از داده ها استنباط کرد به منظور طبقه بندی یک گرۀ از میان گره ها می توان طبقه گره را تعریف نمود. شبکه توانایی خوشه بندی احتمال هر طبقه متناوبی را دارد.

قبل از پیاده سازی تکنیک های معرفی شده بایستی متغیرهای موردنیاز با استفاده از تحقیقات پیشین و مبانی نظری انتخاب شوند. به طور مثال در تحقیق اسپاتیس و همکاران (۲۰۰۷)، گزینش متغییرها با استفاده از معرفی داده های خط دهنده با مشارکت مبتنی بر کارهای تحقیقاتی پیشین مرتبط با موضوعات صورت های مالی تقلبی[1](FFS) بوده است. مانند کار انجام شده به وسیله اسپادیس (۲۰۰۲)، اسپادیس و سایر همکاران (۲۰۰۲)؛ فانینگ و کوگر (۱۹۹۸)؛ پرسونس (۱۹۹۵)؛ استیک (۱۹۹۵)؛ فروز، پارک و پاستنا (۱۹۹۱)؛ لئوبیک، اینینگ، ویتینگ هام (۱۹۸۹) و کینی و مک دانل (۱۹۸۹) که در بردارنده پیشنهادهایی با اشاره به FFS می باشد. پریشانی مالی ممکن است یک مشوق برای تقلبات مدیریتی باشد (فانینگ و گوگر، ۱۹۹۸؛ اسنیک، ۱۹۹۱؛ لیوبیک و سایر همکاران، ۱۹۹۸: کینی و مک دانل، ۱۹۸۵). به منظور تطبیق نمودن یک نسبت در باب پریشانی مالی امتیاز Z التمن که مشهور است را برای برآورد پریشانی مالی بکار گرفتند.

پیرسوس (۱۹۹۵) معتقد بود که یک پرسش مطرح این است که آیا یک ساختار بدهی سنگین در گزارش مالی متقلبانه دخالت دارد. یک ساختار بدهی سنگین ممکن است احتمال FFS را افزایش دهد بدین صورت که این موضوع با ریسک حقوق مالکان و مدیریت بدهی توسط مالکان مرتبط است. مدیران ممکن است بدهی هایی را که براساس قراردادهای بدهی در صورتهای مالی در حالی سررسید شدن می باشد را دستکاری کنند. این موضوع اشاره به این دارد که سطوح

[1] fraudulent financial statements

www.jamv.ir
ISSN: 2645-4572

بالاتری از بدهی ممکن است احتمال FFS را افزایش دهد. اسپاتیس و همکاران (۲۰۰۷) آن را به وسیله استفاده از لگاریتم کل بدهی (LOGDEBT)، نسبت بدهی به ارزش ویژه (DEBTEQ) و نسبت کل بدهی به کلی دارایی ها (TDTA) اندازه گرفتند.

انگیزه دیگر تقلب مدیریتی نیاز به رشد پایدار است. شرکت های ناتوان در رسیدن به نتایج مشابه عملکرد گذشته ممکن است درگیر فعالیت های متقلبانه برای حفظ روند پیشین خود شوند (استیک، آلبرچت و براون، ۱۹۹۱). شرکت هایی که رشد سریعی داشته اند ممکن است فرایندی را تحت نظر قرار دهند که قابلیت نظارت مناسب را تدارک می بیند و آن را پشت سر گذاشته باشند (فانینگ و کوگر، ۱۹۹۸). برای اندازه گیری رشد، از نسبت رشد فروش (SALGRTH) استفاده کردند.

حسابرسی معدودی از حساب هایی که استفاده از برآورد در آنها مجاز است، مشکل می باشد در نتیجه آنها مستعد تحریف های متقلباته هستند. حساب های دریافتنی موجودی و فروش ها در این گروه قرار می گیرند (پیرسونس، ۱۹۹۵)؛ استیک (۱۹۹۱) و فروز و همکاران (۱۹۹۱) معتقدند که مدیریت ممکن است در حسابهای دریافتنی دستکاری کنند. فعالیت های متقلبانه ثبت فروش قبل از اینکه تحقق یابد ممکن است حساب های دریافتنی را اضافه تر نشان دهد (فانینگ و کوگر، ۱۹۹۸). در تحقیق مذکور حساب های دریافتنی را با استفاده از نسبت حساب های دریافتنی / فروش (RECSAL) و نسبت حساب های دریافتنی به حساب های دریافتنی برای دو سال پیاپی (RETREND) و متغییر رقومی ۱۱REC که اشاره به تغییرات ۱۰٪ دارد، مدنظر قرار گرفت. بسیاری از تحقیقات به این موضوع اشاره دارند که ممکن است مدیریت موجودی ها را دستکاری کند (استیک ۱۹۹۱؛ پیرسون ۱۹۹۵؛ اس چلیت، ۲۰۰۲). گزارش موجودی ها با بهای تمام شده کمتر و ثبت موجودی های راکد بعضی از تدابیر شناخته شده است. موجودی ها به وسیله استفاده از نسبت موجودی به فروش (INVSAL) و موجودی به مجموع دارایی ها (INVTA) بررسی کردند. حاشیه سود ناخالص نیز مستعد دستکاری است. شرکتها ممکن است تطابق فروشی هایشان را با بهای تمام شده کالای فروش رفته مربوط انجام ندهند. در نتیجه حاشیه سود ناخالص، خالص درآمد و استحکام ترازنامه افزایش یابد(فانینگ و کوگر،۱۹۹۸). حاشیه سود ناخالص را با استفاده از نسبت فروش منهای حاشیه سود ناخالص (COSAL) و نسبت سود ناخالص / مجموع دارایی ها (GPTA) و نسبت حاشیه سود ناخالص به حاشیه سود ناخالص دو سال پیاپی (GMTREND) و متغییر دو تایی باینری ۱۱GM که اشاره به ۱۰٪ افزایش نسبت به ارزش سال پیش دارد، آزمون شد.

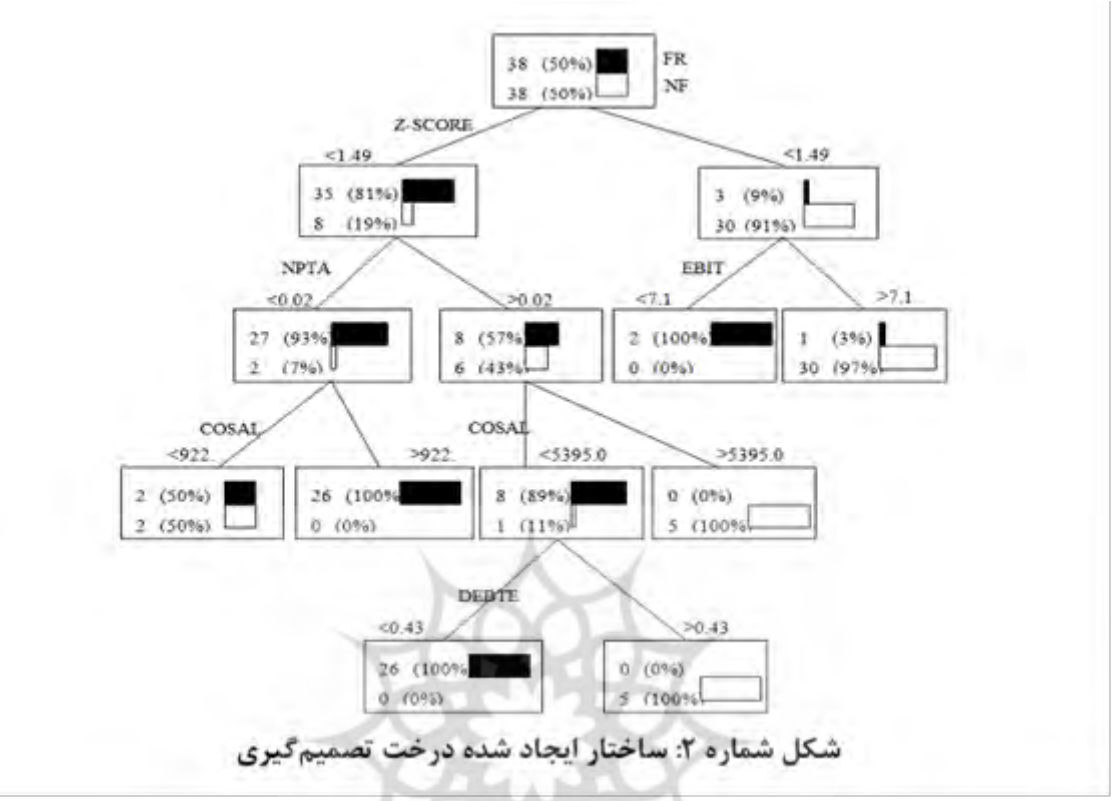
اسپاتیس(۲۰۰۲) در یک مطالعه رگرسیون لاجستیک برای پیش بینی FFS معتقد بود سود خالص / مجموع دارایی ها (NPTA) و سرمایه در گردش/ مجموع دارایی ها (WCTA) نسبت های معنی داری هستند. به علاوه اسپاتیس و همکارانش(۲۰۰۲) به این موضوع اشاره داشتند که همچنین نسبت سود خالص فروش (NPSAL) بیشترین معنی داری را در خود دارد. در این مطالعه، ارتباط بعضی از پرچم های قرمز صورتهای مالی با قابلیت پیش بینی FFS آنها آزموده شد. این نسبت ها عبارتند از: لگاریتم مجموع دارایی ها (LTA)، سرمایه در گردش (WCAP)، نسبت تاسیسات و تجهیزات به مجموع دارایی ها (NFATA) و فروش به مجموع دارایی ها (SALTA) و دارایی های جاری / بدهی های جاری (CACL) و خالص درآمد / دارایی های ثابت (NIFA) و موجودی نقد / مجموع دارایی ها (CASHTA) و دارایی های آتی / بدهی های جاری (QACL)، سود قبل از بهره و مالیات (EBIT) و بدهی های بلند مدت / مجمع دارایی ها (LTDTA)، (رهنمای رودپشتی، ۱۳۹۱).

حال بعد از معرفی متغیرهای تحقیق به طور مثال و تبیین هر کدام از تکنیک های داده کاوی، بایستی آنها پیاده سازی گردند؛ به عنوان مثال مدل درخت تصمیم گیری با استفاده از نرم افزار ویرایش تحقیقاتی سیپنیا[1] می تواند اجرا شود.

[1] Sipina Research Edition

www.jamv.ir
ISSN: 2645-4572

اسپاتیس و همکاران (۲۰۰۷) در تحقیق خود از این نرم افزار جهت درخت تصمیم گیری شرکت های مورد تقلب بهره گرفتند که نمودار درخت تصمیم گیری تحقیق آنها در شکل شماره ۲ نشان داده شده است.



شکل شماره ۲: ساختار ایجاد شده درخت تصمیم‌گیری

آنها نشان دادند در این مورد درخت تصمیم گیری تمام موارد فاقد تقلب (۱۰۰٪) و ۳۵ مورد از ۳۸ مورد تقلب را به طور صحیح طبقه بندی شد که در جدول شماره (۱) متغیر های جداساز نمود یافته در درخت تصمیم گیری را شرح می دهد.

جدول شماره (۱): تفکیک متغیر ها در درخت تصمیم گیری

متغیرهای منتخب	The spilling variables
متغیرها	Variables
Z آلتمن	Z-SCORE
سود خالص به کل دارایی ها	NPTA
سود قبل از بهره و مالیات	EBIT
فروش منهای حاشیه سود ناخالص	COSAL
بدهی به ارزش ویژه	DERTEQ

جهت پیاده سازی تکنیک شبکه های عصبی می توان از نرم افزار ناکلاس۷[1] با شبکه غیرخطی برای طبقه بندی به منظور ساخت یک شبکه پیش خور مفهومی چند لایه بکار برد. اسپاتیس و همکاران (۲۰۰۷) در این باب در تحقیقشان عنوان کردند، بعد از آزمون تعدادی از طرح های متنوع و انجام آموزش های مقدماتی، وضعیتی با یک لایه مخفی محتوی پنج گرهٔ مخفی را برگزیدیم. شبکه انتخاب شده دوم، با استفاده از تمامی نمونه ها آموزش دید و در مقابل دسته

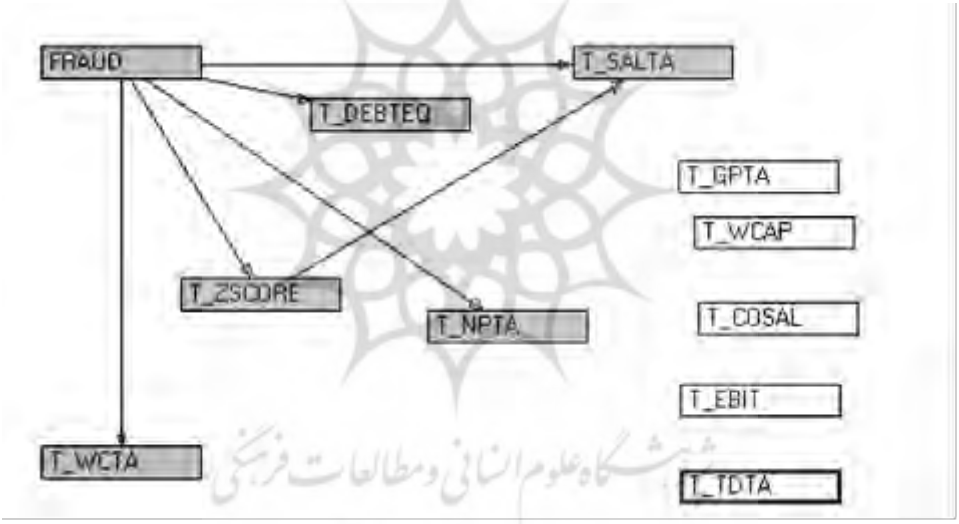
[1] Nuclass7

www.jamv.ir
ISSN: 2645-4572

آموزشی آزمون شد. شبکه موفق شد تمام موارد را به درستی طبقه بندی کند؛ بدین گونه که به موفقیت ۱۰۰ درصدی در عمل دست یافت.

در نهایت جهت پیاده سازی تکنیک شبکه های گمانه زنی بیزی می توان از نرم افزار BN بهره گرفت.

اسپاتیس و همکاران (۲۰۰۷) در آزمایش شبکه گمانه زنی بیزی از نرم افزاری که استفاده کردند، BN با قدرت پیش بینی بود که عنوان شد این نرم افزار ظرفیت آموزش یک طبقه بندی کننده را با داده ها دارد و متعلقات الگوریتمی که بر پاشده دسته ای از الگوریتمهای مبتنی بر آزمون مستقل شرطی است و نیاز به تنظیم گره نیست (جنگ و گرینر[1]، ۲۰۰۱). به اقتضای محدودیت های نرم افزار باید مقادیر را به صورت نظری اعمال کردند. پس از آزمون روش های نظری مختلف (با پهنا و ژرفای مساوی)، روش نظری بر نگری شده را انتخاب کردند. بر خلاف سایر روش های نظری، مبنای نظری در گاشت محور[2] در رده بندی اطلاعاتی بکار رفت. این ساختار شباهتی را که فاصله تعریف شده ممکن است برای کمک به بهبود طبقه بندی صحیح ایجاد کنند را می سازد (هان و کمبر، ۲۰۰۰). از این رو برای آموزش شبکه گمانه زنی، تمام نمونه ها را به عنوان دسته آموزشی استفاده و بعد از آموزش شبکه، آن را در مقابل دسته آموزشی آزمون نمودند. شبکه به درستی ۷۲ مورد را طبقه بندی نمود. در این مورد مدل ۳۷ مورد تقلب (۹۷٪) و ۳۵ مورد فاقد تقلب (۹۲٪) را به درستی طبقه بندی نمود. شکل شماره (۳) ساختار شبکه گمانه زنی را نشان می دهد.



شکل شماره (۳): ساختار شبکه گمانه زنی بیزی

همانطور که در شکل شماره (۳) می توان دید، شبکه گمانه زنی به منظور وفق دادن تعمیم یافتگی بیشتر، ظاهری با ملاحظه به انگیزه های تحریف صورت های مالی مدنظر داشته است. بر اساس شبکه تقلبات گزارش داده شده به شدت وابستگی به ورودی متغیرهای امتیاز Z، NPTA، DEBTEQ, SALTA و WCTA دارند. هر کدام از این متغیرها ادعای اختلاف یک جنبه از وضعیت مالی شرکت را دارد. امتیاز Z باز تایی از پریشانی مالی است؛ DEBTEQ به اهرم اشاره دارد؛ NPTA نشان از قابلیت سودآوری، SALTA به عملکرد فروش اشاره دارد و WCTA نشان از نقدشوندگی است.

[1] Cheng, J., & Greiner, R
[2] Entropy-based

www.jamv.ir
ISSN: 2645-4572

نتیجه گیری

داده کاوی به عنوان یکی از ده دانش در حال توسعه، پل ارتباطی میان علم آمار، علم کامپیوتر، هوش مصنوعی، الگوشناسی، فراگیری ماشین و بازنمایی بصری داده می باشد که کاربرد گسترده ایی در زمینه های پزشکی، صنعت، مهاجرت، هتلداری، مدیریت ریسک و بالاخص کشف تقلب دارد. به طوری که امروزه فعالیت حسابرسی باید چیرگی بر موارد رو به فزونی تقلب مدیران باشد، نیازمند ابزارهای موثر می باشد. تکنیک های داده کاوی که ادعا می شود ظرفیت های پیشرفته ای در طبقه بندی و پیش بینی دارند، توان تسهیل کنندگی برای حسابرسان در انجام وظیفه کشف تقلب مدیریتی را دارند. ادعای این مطالعه بررسی سودمندی عملکرد سه تکنیک داده کاوی بکار رفته از جمله درخت تصمیم گیری، شبکه های عصبی و شبکه های گمانه زنی بیزی بود. نتایج نشان داد که موید نتایج تحقیقات پیشین، به این مطلب اشاره دارد که داده های صورت های مالی، محتوی اطلاعاتی جهت کشف تحریف است. این مطالعه همچنین متغیرهای معینی از اطلاعات عمومی آماده که حسابرسان باید زمان بیشتری به آن اختصاص دهند را توصیه کرده است. چارچوب روش شناسی پیشنهادی مورداستفاده می تواند کمکی برای حسابرسان داخلی و خارجی و مالیاتی و سایر اعتباردهندگان گزارشات، اشخاص حقیقی و حقوقی، سرمایه گذاران، معامله گران سهام، شرکت های وکالت، تحلیلگران اقتصادی، موسسات رتبه بندی اعتباری و سامانه بانکی باشد. برای حسابرسی حرفه ای این مطالعه می تواند در مساعدت برای شناسایی روش های پاسخگویی در برابر کشف FFS سودمند باشد.

منابع

- ✓ رهنمای رودپشتی، فریدون، (۱۳۹۱)، داده کاوی و کشف تقلب های مالی، دانش حسابداری و حسابرسی مدیریت، سال اول، شماره سوم، صص ۱۱۷-۳۳.
- ✓ زارع بهنمیری، محمدجواد، ملکیان کله بستی، اسفندیار، (۱۳۹۴)، پیش بینی تقلب در صورت های مالی با استفاده از نسبت های مالی شرکت های پذیرفته شده در بورس اوراق بهادار تهران، چشم انداز مدیریت مالی، شماره ۱۲، صص ۸۰-۶۵.
- ✓ Calderon, T. G., & Cheh, J. J. (2002). A roadmap for future neural networks research in auditing and risk assessment. International Journal of Accounting Information Systems, 3(4), 203–236.
- ✓ Chartered Institute of Management Accountants (2008). Fraud Risk Management: A Guide to Good Practice, First Edition, Uk.
- ✓ Chen, Suduan.(2016). Detection of fraudulent financial statements using the hybrid data mining approach, Chen SpringerPlus (2016) 5:89 DOI 10.1186/s40064-016-1707-6.
- ✓ Cheng, J., & Greiner, R. (2001). Learning Bayeaaa eeiie eekkkkkkciiiii i e aggrtt add sy PPPoceennng ffeeeeeeeeeeeeeeee
- ✓ conference of the Canadian society on computational studies of intelligence: advances in eiiificttttttt tlligecce((... 1111–151).
- ✓ Cullinan, P. G., & Sutton, G. S. (2002). Defrauding the public interest: a critical examination of reengineered audit processes and the likelihood of detecting fraud. Critical Perspectives on Accounting, 13(3), 297–310.
- ✓ Fraser, I. A. M., Hatherly, D. J., & Lin, K. Z. (1997). An empirical investigation of the use of analytical review by external auditors. The British Accounting Review, 29(1), 35–47.
- ✓ Jameie, R. And Asgharzadih, P, (2012). Reviews The Performance Gap Fraud Field Of View And Use Independent Auditors Auditing Services, The First Regional Conference On New Approaches Accounting And Auditing, Bandar Gaz, Azad University Of Bandar Gaz Branch, Http: // Www.Civilica.Com / Paper-Rcnaaa01Rcnaaa01_015.Html.

- ✓ aa rrrr zz 2222222aa ii gggg cnnce aggrtt iley— IEEE Press.
- ✓ Koh, H. C., & Low, C. K. (2004). Going concern prediction using data mining techniques. Managerial Auditing Journal, 19(3), 462– 476.
- ✓ Rahimian, N and Akhundzade, M. (2011). Role of Internal Audit in Preventing and Detecting Fraud, Journal of Certified Public Accountants, 15, 40-44
- ✓ Saghafi, A, Arabmazar Yazdi, Ali (2010). Financial Reporting Quality and Investment Inefficiency, Journal of Financial Accounting Research, 2(4), Serial (6).
- ✓ Sohl, J. E., & Venkatachalam, A. R. (1995). A neural network approach to forecasting model selection. Information & Management, 29(6), 297–303.
- ✓ Spathis, C, Kirkos, E,.and Manolopoulos, Y(2007). Datamining techniques for the detection of fraudulent financial statement. Expert systems with applications, vol 32, pp 995 – 1003.
- ✓ Saaiii C & Ziiiiiiii iii C 22.... .ee eeciigg aaiii fied fiaaccia eeeeeeeeeeee a comparative study using multicriteria analysis and multivariate statistical techniques. The European Accounting Review, 11(3), 509–535.
- ✓ Wells, J. T. (1997). Occupational Fraud and Abuse. Austin, TX: Obsidian Publishing. aaaaaaaa aa, 44444444Afifififi nerr a nekkkkkkkk aaalyiica eeeee rr oceeeeee nnn ageiial Auditing Journal, 19(2), 191–223.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی