



روش‌های داده‌کاوی برای کشف تقلب در حسابرسی صورت‌های مالی

آزاده مسلم‌زاده گنتابی*

مقدمه

بر عصر حاضر، «عصر اطلاعات» نام نهاده شده و آنچه که منجر به قدرت و موفقیت می‌شود اطلاعات است. اکنون با کمک فناوری‌های پیچیده‌ای مانند کامپیوتر، ماهواره و وسایلی برای ذخیره‌سازی اطلاعات، می‌توان شروع به جمع‌آوری و ذخیره‌سازی انواع داده‌ها کرد و از توانایی کامپیوترها برای کمک به مرتب‌سازی انبوه اطلاعات بهره برد. متأسفانه، این مجموعه عظیم داده‌های ذخیره شده بر روی کامپیوترهای متفاوت به سرعت دست و پاگیر شدند. این هرج و مرج منجر به ایجاد پایگاه داده‌های ساختار یافته و سیستم‌های مدیریت پایگاه داده^۱ شده است. سیستم‌های مدیریت پایگاه داده کارآمد، سرمایه بسیار

ارزشمندی برای مدیریت حجم انبوهی از داده‌ها و به خصوص برای بازیابی اثربخش و کارآمد اطلاعات از مجموعه‌ای بزرگ در زمان مورد نیاز است. گسترش سیستم‌های مدیریت پایگاه داده نیز به جمع‌آوری انواع اطلاعات کمک کرده است. امروزه، اطلاعاتی به مراتب بیشتر از نیاز در اختیار داریم: ما به اطلاعات معاملات کسب و کار و داده‌های علمی تا تصاویر ماهواره‌ای و اطلاعات نظامی دسترسی داریم. بنابراین دیگر بازیابی اطلاعات به تنهایی برای تصمیم‌گیری کافی نیست. راهکارهای جدیدی برای کمک به تصمیم‌گیری‌های بهتر مدیریتی در مواجهه با مجموعه عظیمی از داده‌ها ایجاد شده است. این راهکارها شامل خلاصه‌سازی خودکار داده‌ها،

استخراج از «اصل» اطلاعات ذخیره شده و کشف الگوهای موجود در داده‌های خام است.

تجزیه و تحلیل داده‌ها از طریق داده‌کاوی

نمایه شماره ۱ نشانگر رابطه مفهومی بین استخراج داده‌ها، تجزیه و تحلیل داده‌ها و داده‌کاوی^۲ است. هنگامی که حسابرس، مراحل کار خود را از استخراج داده‌ها به تجزیه و تحلیل داده‌ها و سپس به داده‌کاوی تغییر می‌دهد، نرم افزارهای مورد استفاده از نظر قابلیت‌ها پیچیده‌تر می‌شود و مقدار بیشتری از قدرت تشخیص و پیش‌بینی را فراهم می‌کند. مثلث نمایه شماره ۱ نشانگر فراوانی نسبی استفاده از

نرم افزارهای مختلف توسط حسابرسان بر مبنای بررسی های بدست آمده از مقالات حرفه ای است. فرآیند داده کاوی به این شرح می باشد:

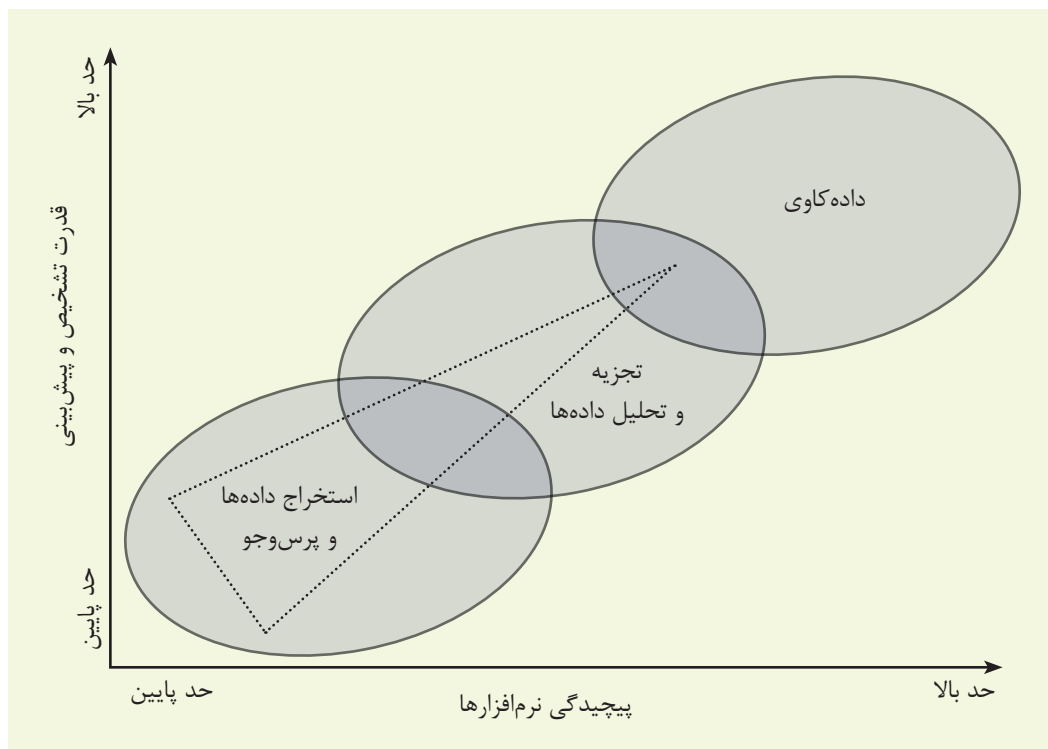
استخراج داده ها

برای حسابرسی صورت های مالی، نرم افزارهای اکسل^۳، ای سی ال^۴، کیس ویر^۵ و آیدیا^۶ جزء کاربردی ترین ابزارهای مورد استفاده برای بررسی اطلاعات صاحبکار می باشند. این ابزارها، در واقع برای جستجو و استخراج داده ها، استفاده می شود و تجزیه و تحلیل داده ها را با انواع آمارهای توصیفی و تعداد محدودی از روش های آماری، انجام می دهد. نرم افزارهای ای سی ال و کیس ویر و آیدیا، شامل ابزارهایی می باشند که برای رویه های حسابرسی بعدی، نمونه سازی می کنند. به این نرم افزارها و نرم افزارهای مشابه مرسوم، «فنون حسابرسی به کمک رایانه»^۷، گفته می شود. حسابرس

پس از انجام هر گونه تجزیه و تحلیل در رسیدگی های مربوط به داده ها، با کمک این نرم افزارها تصمیم می گیرد که می خواهد چه اطلاعاتی را استخراج کند. به عنوان نمونه، اگر حسابرس در کنترل های داخلی شرکتی، نیاز به کنترل امضای چک های بیشتر از ۵۰۰۰ دلار داده شده به فروشندگان را دارد، او یکی از این ابزارها را برای استخراج شماره چک ها و اطلاعات مربوط به تمام چک های بین ۴۹۰۰ و ۴۹۹۹ دلار مورد استفاده قرار می دهد. استخراج چنین اطلاعاتی به این علت صورت می گیرد زیرا حسابرس معتقد است این ریسک وجود دارد که کلاهبردار یا تقلب کننده چک جعلی زیر ۵۰۰۰ دلار بنویسد و به مدیر مالی گزارش ندهد. حسابرس پس از آن، برای تعیین چک های فاقد کنترل های داخلی یا جعلی، به بررسی و تعیین موقعیت چک ها و اسناد پشتوانه آنها می پردازد (Coderre, 2009).

در واقع با به کارگیری این نرم افزارها، کارایی و بهره وری حسابرسی، کاهش نمی یابد. نرم افزارهای استخراج داده ها، ابزارهای حسابرسی قدرتمندی محسوب می شوند. نرم افزارهای ای سی ال و کیس ویر می تواند در شناسایی الگوهای داده های مشکوک در پایگاه های اطلاعاتی مشتری اثربخش باشد. این ابزارها، اساساً برای جستجوی داده ها و ابزارهای استخراج برای تبدیل حجم انبوه اطلاعات صاحبکار به حجم کمتر و قابل کنترل تر اطلاعات مورد استفاده قرار می گیرد. حسابرس پس از استخراج، تجزیه و تحلیل یا تحقیقاتی (اغلب به طور دستی) در ارتباط با نمونه های استخراج شده از جمعیت، انجام می دهد که این کار به میزان آموزش و تجربه حسابرس برای شناسایی موارد مشکوک استخراج شده بستگی دارد که نیازمند بررسی های اضافی می باشد.

نمایه شماره ۱. روابط بین ابزارهای مختلف در بررسی داده ها



Ref: Gray & Debreceeny: 2014

تجزیه و تحلیل داده‌ها

ابزارهای تجزیه و تحلیل داده‌ها در دسترس حساب‌برسان، طیف وسیعی از فنون تحلیلی از ساده به نسبتاً پیچیده را ارائه می‌دهد. به عنوان نمونه، در سطح پایین، تجزیه و تحلیل داده‌ها شامل آمار توصیفی مقدماتی از قبیل شمارش‌ها، حداقل‌ها، حداکثرها، پراکندگی‌ها و میانگین‌ها می‌باشد. تجزیه و تحلیل‌های سطح پایین‌تر ممکن است شامل تجزیه و تحلیل نسبت‌ها نیز باشد. در سطح بالاتر، تجزیه و تحلیل داده‌ها می‌تواند شامل آمار استنباطی از قبیل رگرسیون چند متغیره و تک متغیره، تجزیه و تحلیل همبستگی و سایر خروجی‌های نرم‌افزارهای آماری باشد. عامل محدودکننده اصلی در استفاده از ابزارهای تجزیه و تحلیل داده‌ها، مانند استخراج داده‌ها، ناشی از اطلاعات انبوه حساب‌برسان است. به عنوان نمونه، حساب‌برسان در سازمان‌های بزرگ، با حجم قابل توجهی از اطلاعات مواجه خواهند بود (Eppler and Mengis, 2004).

داده‌کاوی

داده‌کاوی، کشف دانش در پایگاه داده‌ها^۸ نیز نامیده می‌شود و روشی برای کشف اطلاعات سودمند جدید و بالقوه از بین حجم انبوهی از اطلاعات است. هدف از داده‌کاوی، شناسایی الگوها و ارتباطات قابل فهم، معتبر، جدید و به طور بالقوه سودمند در داده‌های موجود می‌باشد. پیدا کردن الگوهای مفید در داده‌ها با نام‌های مختلف (از جمله داده‌کاوی) در جوامع مختلف (به عنوان نمونه، استخراج دانش، کشف اطلاعات، برداشت اطلاعات، باستان‌شناسی داده‌ها، پردازش الگوی داده‌ها) شناخته شده است (Fayyad et al., 1996).

داده‌کاوی در آغاز توسط آمارگران، محققان پایگاه داده و سیستم اطلاعات مدیریت^۹ و جوامع کسب و کار مورد استفاده قرار می‌گرفت. واژه کشف دانش در پایگاه



مستقیم، متغیر خاصی شناسایی می‌شود که مد نظر قرار گرفته شده است. فناوری داده‌کاوی روابط بین یک متغیر و جمعیت انتخاب شده از دیگر متغیرها را می‌یابد. با داده‌کاوی غیرمستقیم متغیر مشخصی هدف قرار نمی‌گیرد (متغیر وابسته) بلکه هدف یافتن روابط بین متغیرها در جمعی از اطلاعات است. راه دیگر برای توصیف این دو دسته فنون اینطور است که بگوییم رویکرد از بالا به پایین، فرضیه‌های خاص را آزمون می‌کند و رویکرد از پایین به بالا، فرضیه‌های جدید را تولید می‌کند. نمونه‌هایی از داده‌کاوی مستقیم به این شرح است:

طبقه‌بندی: مجموعه‌ای از قوانین براساس داده‌های موجود (مجموعه‌های آموزشی) ایجاد می‌شود تا براساس آن طبقه‌بندی مناسبی برای موضوعی جدید در میان طبقه‌بندی‌های مختلف انجام گیرد. به عنوان نمونه، فنون طبقه‌بندی می‌تواند ریسک‌های صاحبکار حساب‌برس را به ریسک پایین، متوسط یا بالا طبقه‌بندی کند که این طبقه‌بندی براساس مجموعه‌ای از متغیرهای مستقل مانند متغیرهای صورت‌های مالی، تغییرات در تعداد و ارزش سهام انجام می‌گیرد. کیفیت طبقه‌بندی فنون داده‌کاوی می‌تواند از طریق صحت و درستی طبقه‌بندی، تاثیرگذاری و توانایی تفسیر نتایج طبقه‌بندی، سنجیده شود. روش‌های

داده‌ها، به طور کلی برای اشاره به فرایند کلی کشف اطلاعات مفید از داده‌ها، استفاده می‌شود و داده‌کاوی گام خاصی در این فرآیند است (Han and Kamber, 2000). همان‌طور که نمایه شماره ۱ نشان می‌دهد، تجزیه و تحلیل داده‌ها و داده‌کاوی، معانی مشترک دارند. اطلاعاتی که کشف دانش محسوب می‌شود، بیشتر از طریق داده‌کاوی استخراج می‌شود. داده‌کاوی در ارتباط با الگوهای کشف، قواعد یا مدل‌های مبتنی بر مجموعه یا مجموعه‌هایی از اطلاعات می‌باشد. نتایج (الگوها، قواعد و مدل‌ها) برای پیش‌بینی پیامدهای آینده مورد استفاده قرار می‌گیرد. نتایجی نیازمند

بررسی‌های حساب‌برس است که خارج از محدوده پیش‌بینی و الگوها اتفاق می‌افتد (Delen and Al-Hawamdeh, 2009).

فنون داده‌کاوی می‌تواند به دو دسته فنون مستقیم (با رویکرد از بالا به پایین) و غیرمستقیم (با رویکرد از پایین به بالا) تقسیم شود. با استفاده از داده‌کاوی

.....● **عامل محدودکننده اصلی در استفاده از ابزارهای تجزیه و تحلیل داده‌ها، مانند استخراج داده‌ها، ناشی از اطلاعات انبوه حساب‌برسان است. به عنوان نمونه، حساب‌برسان در سازمان‌های بزرگ، با حجم قابل توجهی از اطلاعات مواجه خواهند بود**

طبقه‌بندی شامل درخت تصمیم‌گیری^{۱۰}، شبکه‌های بیزی^{۱۱}، شبکه‌های عصبی^{۱۲}، ماشین‌های بردار پشتیبان^{۱۳} و الگوریتم‌های ژنتیکی^{۱۴} می‌باشد.

برآورد: طبقه‌بندی برای پیامدهای نسبتاً مجزا به کار می‌رود و در آن متغیرها مستقل هستند ولی در فنون برآورد از متغیرهای پیوسته



استفاده می‌شود. به عنوان نمونه، به جای طبقه‌بندی ریسک در سطح پایین، متوسط و بالا، فنون برآورد به سطح ریسک امتیاز می‌دهد و ریسک را از ۰ تا ۱۰ تعیین می‌کنند. سپس این برآوردها می‌توانند از طریق فنی مشابه رگرسیون لجستیک، عمل طبقه‌بندی را انجام دهد و موضوعی را طبقه‌بندی کند. به عنوان نمونه، اگر ریسک بالا برای مشتری، ریسک بالای ۷/۵ می‌باشد. برآوردها می‌توانند برای رتبه‌بندی یک مجموعه و سپس برای تعیین یک آستانه نمره، مورد استفاده قرار گیرد.

پیش‌بینی: طبقه‌بندی و برآورد از جمله فنون داده‌کاوی هستند که برای افشای اطلاعات ویژگی‌های متغیرها، مورد استفاده قرار می‌گیرند که قبلاً در مجموعه‌ای از داده‌ها، شناسایی شده است. پژوهشگران داده‌کاوی اغلب بین فنون "پیش‌بینی" با طبقه‌بندی و برآورد تفاوت قائل می‌شوند. هدف از پیش‌بینی، آزمون سیستماتیک فرضیه و همچنین یافتن وضعیت نامطلوبی است که خارج از محدوده مورد انتظار اتفاق می‌افتد. این فنون پیش‌بینی مشابه فنون طبقه‌بندی، در قالب مجموعه‌های آموزشی برای مدل‌سازی اولیه مانند نمونه‌های آزمایشی یا روش‌های خود راه‌اندازی در مجموعه داده‌های نهایی استفاده می‌شود (Han and Kamber, 2006).

نمونه‌هایی از داده‌کاوی غیرمستقیم به این شرح است:

گروه‌بندی ترکیبی:

هدف از گروه‌بندی ترکیبی، یافتن روابط بین متغیرها در مجموعه داده‌ها می‌باشد. در این گروه‌بندی هیچ متغیر مستقل و وابسته خاصی وجود ندارد. تحلیلگر داده‌ها به دنبال جستجوی روابط بین متغیرها می‌باشد. در

نسبت به زیرگروه‌ها دسته‌بندی شود. هدف از طبقه‌بندی روش مستقیم و خوشه‌بندی غیر مستقیم، مشابه می‌باشد. با این حال شیوه عمل و نتایج کاملاً متفاوت است. در طبقه‌بندی مستقیم، زیرگروه‌ها دارای دسته‌های از پیش تعریف شده برحسب متغیرهای از پیش انتخاب شده می‌باشند. ولی در خوشه‌بندی غیرمستقیم، زیرگروه‌ها از پیش انتخاب شده نیستند. حسابرس باید با استفاده از دانش خود در ارتباط با مدل حسابداری و مدل‌های کسب و کاری که توسط مشتری به کار گرفته شده، مفاهیم اصلی زیرگروه‌ها را در خوشه‌بندی تفسیر کند. در مجموعه‌ای از ثبت‌های دفتر روزنامه، خوشه‌بندی گروه‌های مورد انتظار ثبت‌های دفتر روزنامه شناسایی خواهد شد. نمونه‌ای از چنین گروهی، ثبت‌های روزنامه مربوط به درآمد فروش، مطالبات، هزینه‌های محصول و موجودی، می‌باشد. همچنین ممکن است گروه‌هایی از ثبت‌های دفتر روزنامه شناسایی شود که براساس الگوی شناخته شده و معینی ایجاد نشده باشد. فنون خوشه‌بندی شامل روش سلسله مراتبی^{۱۵}، نقشه‌های خودسازمان‌دهی براساس شبکه‌های عصبی و فنون مبتنی بر تراکم^{۱۶}، می‌باشد. جستجوی داده‌های پرت، زیرمجموعه مهمی از خوشه‌بندی برای کشف تقلب محسوب می‌شود. فنون مورد استفاده برای شناسایی داده‌ها، شامل الگوریتم‌های مبتنی بر تراکم و فاصله^{۱۷} است (Han and Kamber, 2006).

توضیحات و تصورات: هدف اصلی از داده‌کاوی کشف آمارهای توصیفی جالب توجه در پایگاه داده هاست. بیان توضیحات و تصورات مانند گروه‌بندی ترکیبی و خوشه‌بندی، جزئی از اولویت‌های کار حسابرس برای تفسیر نتایج خواهد بود. همچنین نقطه شروع خوبی در نتایج حاصل از داده‌کاوی است که موجب ترغیب حسابرس برای تأکید بیشتر بر داده‌کاوی

حسابرسی، فنون داده‌کاوی می‌تواند برای یافتن روابط در میان ثبت‌های روزنامه مورد استفاده حسابرس قرار گیرد. نمونه‌ای از گروه‌بندی ترکیبی می‌تواند شامل ثبت‌های دفتر روزنامه اموال، ماشین‌آلات و تجهیزات، تعمیر و نگهداری ماشین‌آلات باشد. حسابرس ثبت‌های روزنامه‌ای را مورد بررسی قرار می‌دهد که خارج از گروه‌بندی ترکیبی مورد انتظار روی می‌دهد. همچنین روابط می‌توانند براساس داده‌هایی باشند که از ترتیبی مانند الگوی زمانی پیروی می‌کنند. به نظر می‌رسد الگوی استخراج متوالی برای روابط بر حسب حوادث مربوط رخ می‌دهد. این دسته از داده‌کاوی الزامات روشنی برای شناسایی

الگوها دارد. در این مورد می‌توان به عنوان نمونه، به مجموعه‌ای نامناسب اما نسبتاً کوچک از نقل و انتقالات بین حساب‌ها، اشاره کرد.

خوشه‌بندی: هدف از خوشه‌بندی این است که متغیرها را براساس روابط موجود در مجموعه داده‌ها،

حسابرس باید با استفاده از دانش خود در ارتباط با مدل حسابداری و مدل‌های کسب و کاری که توسط مشتری به کار گرفته شده، مفاهیم اصلی زیرگروه‌ها را در خوشه‌بندی تفسیر کند

(حتی داده‌کاوی مستقیم) برای به اثبات رسانیدن یا رد تفسیرهای اولیه خود از نتایج توصیفی می‌شود. داده‌کاوی به طور معمول بیشتر برای داده‌های بسیار ساختارمند مورد استفاده قرار می‌گیرد، درحالی که متن‌کاوی، روشی در حال توسعه از داده‌کاوی است که الگوهای آماری و زبانی را در متن شناسایی می‌کند. با افزایش دسترسی به حجم انبوه متون از طریق اینترنت و جاهای دیگر، متن‌کاوی به طوری فزاینده و گسترده به روشی مورد استفاده در بسیاری از فنون تبدیل شده است. وقتی متن برای جمعی از ذینفعان فرستاده می‌شود به طور نمونه از طریق پست الکترونیکی، تجزیه و تحلیل شبکه‌های اجتماعی می‌تواند برای مجموعه‌ای از متون به کار برده شود. فرآیند داده‌کاوی و نقش کاربر (برای جداسازی تحلیل وظایف)، حوزه‌های جدیدتری در داده‌کاوی محسوب می‌شوند که دارای توان بسیار زیادی در کشف تقلب هستند (Berry and Kogan, 2010).

استفاده از داده‌کاوی برای حسابرسی و تقلب

در سال‌های اخیر، استفاده از فنون داده‌کاوی در حسابرسی‌های صورت‌های مالی افزایش یافته است، در این بخش به بررسی پژوهشی در زمینه داده‌کاوی می‌پردازیم که برای درک الگوهای تقلب در افشای اطلاعات صورت‌های مالی، فرآیندهای کسب و کار در شرکت، ثبت‌های روزنامه و همچنین اطلاعات تهیه شده توسط شرکت برای استفاده‌های داخلی و خارجی می‌باشد. برنامه طولانی مدت بسیاری از داده‌کاوی‌ها در راستای کاوش افشای اطلاعات صورت‌های مالی بوده است. پس از وقفه‌ای به مدت یک دهه، مجدداً توجه‌ای افزون بر استخراج داده‌های صورت‌های مالی برای شناسایی گزارش‌های جعلی بالقوه شرکت، صورت

گرفت. به طور نمونه، داده‌های کلیدی اصلی صورت‌های مالی (مانند سود ناخالص و سود خالص) و نسبت‌ها (مانند بدهی‌های طولانی مدت/سرمایه‌ها و ذخایر) برای شناسایی گزارش‌های جعلی بررسی و کاوش می‌شوند. نمونه‌ای از نتایج این کاوش، مجموعه داده‌هایی مرتبط با ۱۰۱ شرکت چینی با تقلب‌های شناخته شده و به همان میزان، شرکت‌های عاری از تقلب بود. مجموعه‌ای از فنون داده‌کاوی به کار گرفته شده شامل ماشین‌های بردار پشتیبان، شبکه‌های اجتماعی، برنامه پیدایشی، شبکه‌های عصبی می‌باشد. از فنون شبکه‌های عصبی، شبکه احتمالاتی است که با توانایی شگفت‌انگیزی برای شناسایی صحیح بیش از ۹۰ درصد مجموعه داده‌های سازمان‌ها در انجام داده‌کاوی مستقیم، سودمندترین روش شبکه عصبی محسوب می‌شود. الگوریتم‌های تکاملی با قابلیت تمایز نسبتاً بالایی برای شناسایی تقلب در صورت‌های مالی مورد استفاده قرار می‌گیرند (Alden et al., 2012).

پرولز (۲۰۱۱) اعمال به وجود آمده برای شناسایی تقلب در صورت‌های مالی را نظارت می‌کند. این اعمال شامل



سطوح پایین تقلب مشاهده شده در برخی از مؤسسات حسابرسی و نوسان در نرخ قیمت بین آنها می‌باشد، به طور نمونه، تشخیص‌های غلط و عدم تشخیص (عدم تشخیص نسبت به تشخیص غلط برای حسابرس و بازار سرمایه به طور چشمگیری خطرناکتر است) و ماهیت داده‌های نامرتب صورت‌های مالی. پس از حسابداری این موارد در طراحی مجموعه داده‌هایشان، پرولز، رگرسیون لجستیک را می‌یابد و ماشین بردار پشتیبان^{۱۸} به طور گسترده‌ای مورد استفاده قرار می‌گیرد. روش طبقه‌بندی هر دو در تشخیص شرکت‌های متقلب به خوبی عمل می‌کند و همچنین نسبت به فنون دیگر مانند انواع مختلف شبکه‌های عصبی، بهتر عمل می‌کند. براساس نوشته‌های پرولز، رگرسیون لجستیک و ماشین بردار پشتیبان برای استقرار در محیط‌های تولید نسبتاً کارآمد هستند (Perols, 2011).

فرآیندکاوی، یکی دیگر از رشته‌های در حال رشد در استفاده از داده‌کاوی برای حسابرسی می‌باشد. این روش، اطلاعات مرتبط با فرآیند کسب و کار را از وقایع ثبت شده توسط سیستم‌های اطلاعاتی سازمان‌ها استخراج می‌کند که به طور معمول سیستم‌های برنامه‌ریزی منابع انسانی^{۱۹} می‌باشند. تحقیقات فرآیندکاوی در زمینه فرآیندهای تصمیم‌گیری سازمان، رعایت مقررات و مدیریت ریسک و ساختارهای کنترل انجام شده است. فرآیندکاوی می‌تواند شامل انواع منابع داده، مانند گردش کار و تجزیه و تحلیل وظایف در سیستم‌های برنامه‌ریزی منابع انسانی، باشد. بیشتر تحقیقات اخیر بر استخراج فرآیند دانش ذاتی در فعالیت‌ها متمرکز بوده است. آلس^{۲۰} و همکاران و جنز^{۲۱} (Jans et al., 2010) و همکاران هر دو سیستم‌هایی برای جمع‌آوری و استخراج فرآیندهای حاصل از پرونده‌های حجیم را ایجاد نمودند. در محدوده تقلب کارمند،

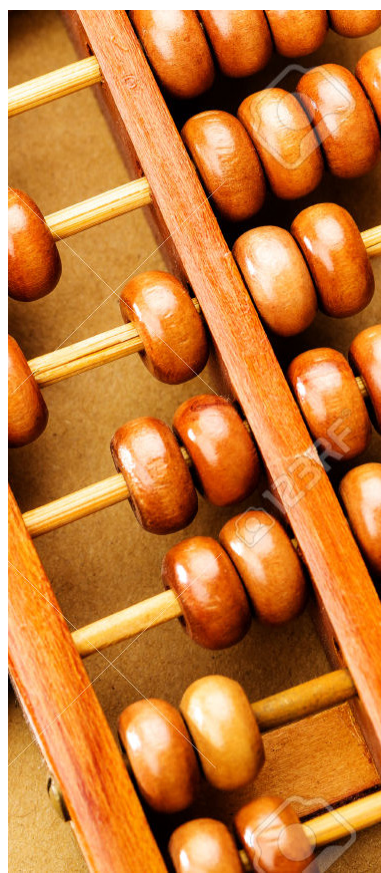
فنون خوشه‌بندی تک متغیری و چند متغیری را به کار گرفتند و برای شناسایی معاملات بالقوه جعلی در مجموعه داده‌های انبوه مربوط به پیش نیازهای خرید یک شرکت، این فنون را مورد استفاده قرار دادند. همان طور که در این دسته از داده‌کاوی معمول است، تجزیه و تحلیل مطلقاً نمی‌تواند تقلبات فعلی را شناسایی کند، بلکه تجزیه و تحلیل، داده‌های پرت و سایر معاملاتی را شناسایی می‌کند که متناسب با الگوها و روندهای پیش‌بینی شده نمی‌باشند. علاوه بر این، تعیین تعداد سفارشات خرید مشکوک، می‌تواند حجم زیادی را برای رسیدگی در برگرد. به این ترتیب، یک نمونه انتخاب می‌شود و مورد بررسی‌های بیشتر قرار می‌گیرد.

در ایجاد فرصت برای ارتکاب تقلب، درک نقش‌هایی که کاربران می‌توانند در سیستم‌های سازمانی ایفا کنند، ورودی مهمی برای حسابرسی کنترل‌های داخلی و معاملات خاص می‌باشد. درک نقش‌ها، هم راستا با تجزیه و تحلیل تفکیک وظایف انجام می‌شود. از اولین مقالات در زمینه نقش کاوی، متعلق به (Colantonio et al., 2011) می‌باشد که در آن روشی برای مدل سازی پیشنهاد شده و این روش را برای مجموعه داده‌های انبوه در شرکت بزرگ، کاربرد دارد.

مسلماً حوزه‌ای که حسابرسان بیشترین تجزیه و تحلیل را از مجموعه داده‌های انبوه انجام می‌دهند، تجزیه و تحلیل ثبت‌های روزنامه صاحبکار براساس بیانیه‌های استناداردهای حسابرسی است. دبرسنسی و گری (۲۰۱۰) اولین گام را در داده‌کاوی ثبت‌های دفتر روزنامه برداشتند. ثبت‌های دفتر روزنامه مشخصات غیرعادی دارند که ناشی از اصول فرآیندهای کسب و کاری است که موجب ایجاد این ثبت‌ها شده است. آنها به طور نمونه، مجموعه‌ای از ۲۹ گروه از ثبت‌های دفتر روزنامه را استخراج

کردند و سپس روش‌هایی را چون قانون بنفورد^{۲۲} ارایه دادند که برای این مباحث کاربرد ندارد. آرگیرو (۲۰۱۲) از شبکه‌های خودسازمانده^{۲۳} برای شناسایی رویدادها و معاملات مشکوک در مجموعه‌ای از ثبت‌های دفتر روزنامه برای سازمان واحد، استفاده کرده و اشتباهات را در یک مجموعه ریشه‌یابی نموده. روش شبکه‌های خودسازمانده، اشتباهات را با درجه بالایی از دقت و همچنین تجزیه و تحلیل هزینه طبقه‌بندی نادرست ثبت‌ها، بین خطای نوع اول و نوع دوم، شناسایی می‌کند. (Argyrou, 2012)

حوزه قابل توجه برای حسابرسی، کاربرد فنون داده‌کاوی متن است. از مهم‌ترین منابع ارتباطات شرکت‌های داخلی، پست الکترونیکی است که روشی فراگیر برای ارتباط با شرکت‌ها و بین کارمندان در شرکت‌ها و ارتباط با فروشندگان، مشتریان، معاونان و مشاوران است. متن همان چیزی هست که در پست الکترونیکی آورده می‌شود. ماهیت پست الکترونیکی با تاریخ، دریافت کننده، و عنوان ایجاد می‌شود که در محل موضوع، حجم شبکه اجتماعی و زمان مورد نیاز برای داده‌کاوی پست الکترونیکی



افزوده می‌شود. علاوه بر این، با توجه به اهمیت حفظ پست الکترونیکی برای کشف الزامات دعوی قضایی و نظارتی در آینده، سازمان‌ها آرشیو پست‌های الکترونیکی را نگهداری می‌کنند.

دبروسینی و گری (۲۰۱۱) به معرفی و تجزیه و تحلیل متن‌کاوی و تحلیل شبکه عصبی می‌پردازد که در داده‌کاوی پست الکترونیکی از منظر تقلب و حسابرسی کاربرد دارد. آنها نمونه‌ای عملی از تجزیه و تحلیل شبکه اجتماعی پست الکترونیکی به وجود آوردند که براساس تحلیل محور حسابرسی از طریق دسترسی به آرشیو پست‌های الکترونیکی انرون می‌باشد. آنها با انجام پژوهشی بالقوه در زمینه استفاده از داده‌کاوی پست الکترونیکی، برای حسابرسی و کشف تقلب، ارزیابی ساختارمندی ارائه نمودند (Debreceeny & Gray, 2011).

تجزیه و تحلیل تقلب، موضوع مهمی در کاربرد متن برای کشف تقلب است. این رویکرد در بررسی اهرم‌ها روشی شناخته شده است که به موجب آن اطلاعات نادرست ارائه شده به طور عمدی با تقلب در کلامی ساختگی توسط فرستنده منتقل می‌شود که در نهایت قصد فریب دریافت‌کنندگان اطلاعات را دارند. الگوهای شناخته شده در کلام متقلب شامل سطوح بالایی از کلام فعال و احساسات منفی در نوشتن است. نظریه‌های بنیادی شامل نظریه دستکاری اطلاعات و نظریه تقلب بین فردی می‌باشد. همفری و همکاران (۲۰۱۱) نشانه‌های زبانی را در بحث تجزیه و تحلیل مدیریت،^{۲۴} بکار بردند و از این روش برای شناسایی تقلبات احتمالی استفاده نمودند که می‌تواند حاکی از ارائه نادرست و جعلی اطلاعات باشد. ۱۰۱ مورد توسط کمیسیون بورس اوراق بهادار در بیانیه‌های راهبردی حسابداری و حسابرسی^{۲۵} به عنوان نمونه‌هایی از تقلب صورت‌های مالی برگزیدند. آنها ۱۰۱ نمونه متقلب با تعداد

مشابهی از نمونه‌های غیر متقلب مطابقت دادند. آنها با استفاده از روشی، مشخصات زبانی مانند زبان فعال، تاثیرگذاری، طول کلمه و پیچیدگی جمله را شناسایی کردند و روشی را ایجاد کردند که ۶۷ درصد از نمونه‌های تقلب و غیر تقلب را به درستی شناسایی می‌کرد و مدل جدیدی برای تشخیص تقلب پیشنهاد دادند که مدل تشخیص تقلب محاسباتی^{۲۶} نامیده

می‌شود. همچنین با ایجاد بردارهای تجزیه ارزش‌های منفرد که نوعی از داده‌کاوی است، روش خود را پایه‌گذاری کردند. آنها یکی از نرم افزارهای تجاری داده‌کاوی به نام ساس^{۲۷} را برای انجام استخراج اطلاعات به کار بردند. گلانسی و یولدف (۲۰۱۱) از بیانیه‌های راهبردی حسابداری و حسابرسی منتشر شده توسط کمیسیون بورس اوراق بهادار، برای شناسایی تقلب استفاده کردند. با استفاده از افشای اطلاعات متنی که بیانیه‌های راهبردی حسابداری و حسابرسی بیان کرده بود، قادر به شناسایی شرکت‌های

حوزه قابل توجه برای حسابرسی، کاربرد فنون داده‌کاوی متن است. از مهم‌ترین منابع ارتباطات شرکت‌های داخلی، پست الکترونیکی است که روشی فراگیر برای ارتباط با شرکت‌ها و بین کارمندان در شرکت‌ها و ارتباط با فروشندگان، مشتریان، معاونان و مشاوران است

بالقوه متقلب شد (Glancy & Yadav, 2011).

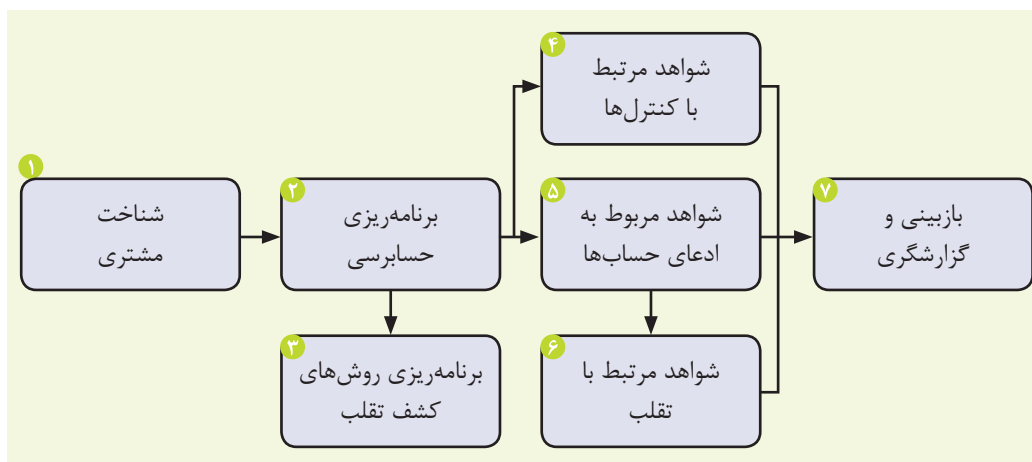
داده‌کاوی در هر مرحله از کار حسابرسی

در این بخش با اهداف حسابرسی و چشم‌اندازی از مراحل مختلف آن برای تشخیص تقلب آشنا خواهیم شد و به بررسی گسترش حوزه اطلاعاتی حسابرسان به منظور بهبود عملکرد حسابرس در تشخیص تقلب می‌پردازیم. حسابرسان ممکن است

احساس کنند داده‌های مورد بررسی در حسابرسی کافی نیست، مانند اطلاعاتی که شامل تنها بخش کوچکی از اطلاعات مرتبط با تقلب مشتری در درون سازمان می‌باشد. علاوه بر این، حجم وسیعی از اطلاعات از خارج سازمان در اختیار صاحبکار قرار می‌گیرد که به طور بالقوه با ارزیابی موارد مربوط به اظهارات اشتباه ناشی از تقلب مرتبط است. بیانیه شماره ۱ حسابرسی اظهار می‌دارد که حسابرس «مسئولیت دارد که برنامه‌ریزی کند و حسابرسی را به گونه‌ای

انجام دهد تا اطمینان معقولی حاصل کند که آیا صورت‌های مالی عاری از تحریف با اهمیت است یا اینکه بر مبنای اشتباه یا تقلب ایجاد شده است» (PCAOB, 2003). تأکید این بیانیه بر کشف تقلب است. بیانیه استانداردهای حسابرسی شماره ۹۹ ایالات متحده، تقلب را به گزارشگری مالی متقلبانه و ادعاهای اشتباه ناشی از سوء استفاده از دارایی‌ها، تجزیه کرده است که هر دو آنها به یک میزان قابل توجه و مهم هستند. طبق بند ۶ بیانیه استانداردهای حسابرسی شماره ۹۹ گزارشگری مالی متقلبانه عبارت از «ادعاهای اشتباه عمدی یا حذف مقادیر یا افشای اطلاعات در صورت‌های مالی طراحی شده به قصد فریب استفاده‌کنندگان صورت‌های مالی است که موجب می‌شود صورت‌های مالی از کلیه جهات با اهمیت در انطباق با اصول پذیرفته شده حسابداری^{۲۸}، ارائه نشود». بیانیه استانداردهای حسابرسی شماره ۹۹ و سایر استانداردهای حسابرسی، حسابرس را ملزم می‌کند که انواع روش‌های تحلیلی و برنامه‌ریزی و روش‌های محتوایی حسابرسی را برای پشتیبانی از تشخیص اشتباهات ناشی از گزارشگری متقلبانه انجام دهد. نگرش خاصی از مراحل مختلف حسابرسی در نمایه شماره ۲ نشان داده شده است (PCAOB, 2002).

نمایه شماره ۲. مراحل کار حسابرسی



Ref: Gray & Debreceeny (2014)



انتظارات از مانده حساب‌های واقعی اعمال می‌کند. اغلب حساب‌رسان تجزیه و تحلیل نسبتاً ساده‌ای در ارتباط با نسبت‌های اصلی اطلاعات صورت‌های مالی انجام می‌دهند تا وقوع این انتظارات را نشان بدهند. داده‌کاوی ترکیبی از تجزیه و تحلیل داده‌های خارج سازمانی (اطلاعاتی در ارتباط با روند ملی و صنعت) با اطلاعات صاحبکار در داخل سازمان، می‌تواند این بخش از حسابرسی را موثرتر و کارآمدتر سازد. در سطح تولید ملی و صنعت، درآمد و سودآوری میانگین هستند و مشتریان لزوماً در سطوح مختلفی نسبت به این مقدار میانگین عمل خواهند کرد، با این حال، داده‌کاوی می‌تواند از ترکیب این مجموعه داده‌ها به ارائه نتایج برتری در توسعه برنامه‌های حسابرسی بپردازد.

برنامه‌ریزی روش‌های مربوط به تقلب:
بیانیه استانداردهای حسابرسی شماره ۹۹ حساب‌رسان را ملزم کرده است که به طور سیستماتیک و منظم، ریسک اشتباهات بالقوه ناشی از تقلب را به عنوان بخشی از فرآیند حسابرسی، رسیدگی کند. حساب‌رسان باید طی فرآیند برنامه‌ریزی حسابرسی، ریسک‌های تقلب را در مدلی روشمند مانند روش طوفان فکری^{۲۹} مطرح کنند که توسط تیم حسابرسی انجام می‌شود. اجرای انواع روش‌های تحلیلی نیز بخش مهمی از این مرحله حسابرسی محسوب می‌شود. بیانیه استانداردهای حسابرسی شماره ۹۹ اشاره می‌کند که حساب‌رسان باید

در برنامه‌ریزی حسابرسی، روش‌های تحلیلی مربوط به درآمد را با هدف شناسایی روابط غیرعادی یا غیرمنتظره، انجام دهد. از جمله حساب‌های درآمدی که ممکن است حاکی از تحریف با اهمیت ناشی از گزارشگری مالی متقلبانه

..... ●
در ایجاد فرصت برای ارتکاب تقلب، درک نقش‌هایی که کاربران می‌توانند در سیستم‌های سازمانی ایفا کنند، ورودی مهمی برای حسابرسی کنترل‌های داخلی می‌باشد
●

حسابرسی بر در دسترس بودن داده‌های برون سازمانی برای مشتری مانند انتشار صورت‌های مالی، انتشار مطبوعات و تجزیه و تحلیل‌ها، گزارش تحلیلگر، قیمت سهام، نظارت بایگانی‌ها، متمرکز است. داده‌های با کیفیت عملکرد بالا برای داده‌کاوی توسط صاحبکار در اختیار حساب‌رسان قرار می‌گیرد. از آنجایی که هر صاحبکاری، ساختار داده‌های منحصر به فردی دارد، هزینه کسب و تفسیر این اطلاعات، نسبتاً بالا است و همچنین، توجه زیادی نیاز است تا اطمینان حاصل شود که از فنون داده‌کاوی مناسبی برای بدست آوردن نتایجی با ارزش بالا، استفاده شده است.

برنامه‌ریزی حسابرسی مبتنی بر ریسک:
برنامه‌ریزی حسابرسی مبتنی بر ریسک در پی ارزیابی ریسک سطح بالای مشتری انجام می‌شود. حساب‌رسان به عنوان بخشی از فرآیند برنامه‌ریزی، به طور عادی، انواع روش‌های تحلیلی را برای وقوع

مراحل حسابرسی

با اشاره به نقش بالقوه‌ای داده‌کاوی در هر مرحله از حسابرسی، ماهیت هریک از مراحل حسابرسی به شرح زیر ارائه می‌شود:
شناخت مشتری: خطر تحریف با اهمیت ناشی از اشتباه یا تقلب در طی حسابرسی ارزیابی می‌شود. افزایش شناخت نسبت به مشتری، جنبه کلیدی از سطح بالای برنامه‌ریزی حسابرسی است. حساب‌رسان باید به عنوان بخشی از مراحل اولیه فرایند برنامه‌ریزی، انواع عوامل ریسک مرتبط با مشتری را درک کند. این شناخت شامل مالکیت و ساختارهای سازمانی، ماهیت فرآیندهای ارزش افزوده، شریک تجاری و روابط با اشخاص وابسته و محیطی است که مشتری در آن فعالیت می‌کند. با توجه به پیچیدگی در نوع مشتری‌های حسابرسی از نظر فرآیندهای کسب و کار، ساختار سازمانی، و روابط شریک تجاری، عوامل بالقوه‌ای در بهره‌برداری از ابزارها و فنون داده‌کاوی برای افزایش قدرت تحلیلی حساب‌رسان بر عملکرد داده‌های مشتری، روابط خارجی و شبکه‌ها وجود دارد. به طور معمول بیشتر داده‌کاوی‌های قبل از

باشد. نمونه‌ای از چنین روش‌های تحلیلی، مقایسه حجم فروش است که از مقادیر درآمد ثبت شده با ظرفیت تولید مشخص می‌شود. مازاد مقادیر فروش بر ظرفیت تولید می‌تواند نشان‌دهنده ثبت فروش ساختگی و جعلی

باشد. به عنوان نمونه‌ای دیگر، تجزیه و تحلیل روند درآمد ماهانه و برگشت از فروش در طی ماه و مدت کوتاهی پس از مدت گزارشگری ممکن است حاکی از وجود موافقت نامه‌ای مخفیانه با مشتری برای برگرداندن محصولات باشد که این کار مانع از شناسایی درآمد می‌شود. درحالی که این نمونه‌ها به طور موقت، نسبت‌های روند کار را به طور نسبتاً ساده مورد بررسی قرار می‌دهد، زمینه‌ای را برای اجرای روش‌های ارزیابی ریسک حسابرسی نیز فراهم می‌کند که می‌تواند با ابزار داده‌کاوی انجام شود. به عنوان نمونه، به جای تجزیه و تحلیل نسبت‌ها از معیارهای شناخته شده (مانند گردش موجودی مواد و کالا و سودآوری خط تولید)، داده‌کاوی با استفاده از معیارهای متعدد می‌تواند منجر به تدوین مدل پیچیده‌تری از فرآیند ارزش افزوده مشتری و مشخصات ریسک شود. از بخش‌های مهم در مرحله برنامه‌ریزی، رسیدگی به خطرات ناشی از ضعف در کنترل داخلی است. (بندهای ۴۴ و ۴۵) بیانیه استانداردهای شماره ۹۹ ایالت متحده حسابرسان را ملزم کرده است که طراحی و اجرای کنترل‌های در نظر گرفته شده را برای جلوگیری از تقلب مورد ارزیابی قرار دهند. مشتریان اغلب به عنوان بخشی از تعهد خود برای نظارت بر اثربخشی کنترل داخلی، ماتریس گسترده‌ای از ادعاهای حساب‌ها و طراحی و وجود کنترل‌ها، فراهم می‌کنند. این ماتریس‌ها می‌تواند

داده‌کاوی، کشف دانش در پایگاه داده‌ها نامیده می‌شود و روشی برای کشف اطلاعات سودمند جدید و بالقوه از بین حجم انبوهی از اطلاعات است

تجزیه و تحلیل داده‌های نسبتاً مشخص را انجام دهد و به ارزیابی ریسک‌ها و کنترل‌های کلیدی برای آزمون‌های بعدی بپردازد. **شواهد حسابرسی برای کنترل‌ها:** پس از مرحله برنامه‌ریزی، جمع‌آوری شواهد

حسابرسی آغاز می‌شود. ارزیابی اثربخشی عملیاتی کنترل داخلی از جنبه‌های کلیدی این فرآیند است. کاربرد بالقوه داده‌کاوی در زمینه کنترل داخلی، فرآیند کاوی در عملیات کنترل‌های داخلی در سراسر چرخه عملیات است. فرآیند کاوی به طور معمول در سیستم ثبت مربوط اعمال می‌شود، مانند سیستم‌های مدیریت منابع سازمان که دفتر کل را پشتیبانی می‌کند. نمونه‌هایی از جمع‌آوری شواهد در عملیات کنترل‌های داخلی، انجام فرآیند کاوی در فرآیندهای کنترل و زیرپاگذاری کنترل‌های فعلی است.

شواهد حسابرسی در مورد ملاحظات مربوط به خطر تقلب بالقوه: حسابرس پس از اجرای عملیات رسیدگی حسابرسی، حسابرس روش‌های محتوایی را اجرا می‌کند که برخی از آنها به طور خاص مربوط به کشف تقلب است. به طور نمونه، بیانیه استانداردهای حسابرسی شماره ۹۹ ایالت متحده اشاره می‌کند که روش‌های تحلیلی محتوایی درآمد، داده‌های جمع‌آوری شده مانند مقایسه درآمد کسب شده ماهانه، توسط خط تولید یا بخش کسب و کار را استخراج می‌کند.

داده‌کاوی بر روی داده‌های شرکت و بانک اطلاعاتی درآمدها و دیگر پایگاه‌های داده مشتری، به عنوان نمونه‌ای از روش‌های تحلیلی مرتبط با تقلب، حائز اهمیت خواهد بود. به غیر از روش‌های تحلیلی محتوایی، بیانیه استانداردهای حسابرسی شماره ۹۹

اشاره می‌کند که تغییر در ماهیت، زمانبندی و وسعت روش‌های حسابرسی برای پاسخگویی حسابرس به ارزیابی ریسک‌های خاص تقلب مورد نیاز است.

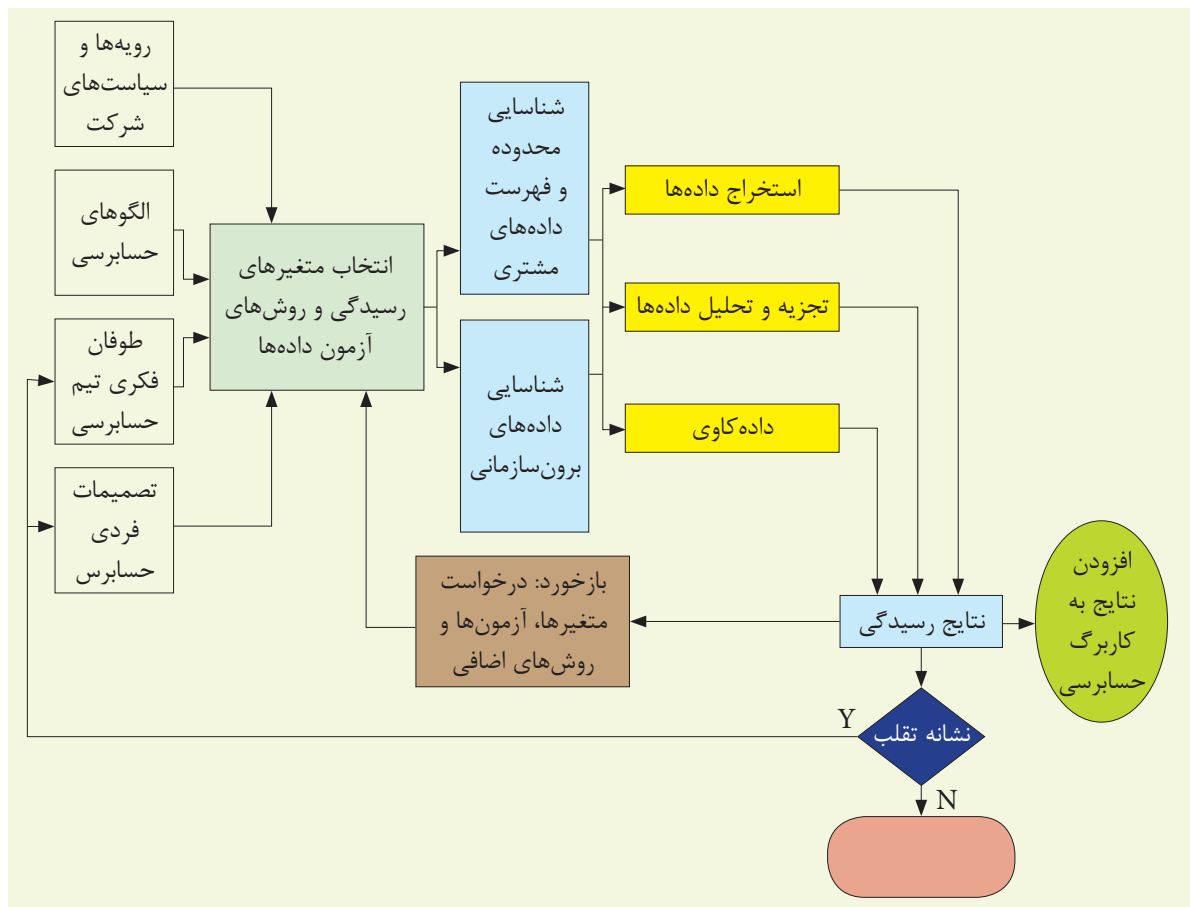
نتیجه‌گیری حسابرسی، بررسی و گزارشگری: طبق بیانیه استانداردهای حسابرسی شماره ۹۹، حسابرس باید "در زمان نزدیک به اتمام اجرای کار حسابرسی"، احتمال وقوع تحریف ناشی از تقلب را ارزیابی کند این کار می‌تواند شامل ارزیابی مجدد استفاده از روش‌هایی از داده‌کاوی باشد که قبل از شروع کار به دلیل بی ثمر دانستن آن، در فرآیند برنامه‌ریزی کنار گذاشته شده بودند یا اجرای روش‌های جدید داده‌کاوی به دلیل بی نتیجه بودن روش‌های دیگر در طی اجرای عملیات حسابرسی باشد. به طور نمونه، روش‌های تحلیلی محتوایی یا سایر آزمون‌ها بر درآمد می‌تواند حاکی از نیاز به کاوش باشد. پس از این فرآیند بازبینی نهایی از شواهد حسابرسی انجام می‌گیرد که منجر به تصمیم‌گیری در ارتباط با ماهیت گزارش نهایی حسابرسی می‌شود.

برنامه‌ریزی با استفاده از داده‌کاوی

مباحث قبلی نشان داد که داده‌کاوی چگونه می‌تواند در هر مرحله از حسابرسی مورد استفاده قرار گیرد. حسابرس باید با توجه به استفاده از داده‌کاوی، فرآیند آزمون داده‌ها را نیز انجام دهد، این فرآیند در نمایه شماره ۳ نشان داده شده است.

در گام اول، تیم حسابرسی باید درباره متغیرهایی (و منابع داده‌ای که شامل این متغیرها می‌شود) تصمیم‌گیری کند که برای دستیابی به اهداف حسابرسی خود (بخشی از آن شامل انواع تقلبی است که از جلسه طوفان فکری حاصل می‌شود) نیاز دارد و اینکه چه چیزی وضعیت نامطلوب را در کار ایجاد می‌کند. این متغیرهای داخلی و خارجی مرتبط با مشتری و منابع داده

نمایه شماره ۳. فرآیند بررسی داده در حسابرسی
Ref: Gray & Debreceeny (2014)



از ابزار داده‌کاوی، در این مرحله او باید تعیین کند که کدامیک از فنون داده‌کاوی برای رسیدگی به مشخصه‌های داده مناسب‌ترین روش می‌باشد.

گام چهارم، سپس حسابرس باید نتایج هشدارها و وضعیت‌های نامطلوب را مورد بررسی قرار دهد و براساس نتایج بررسی‌ها تعیین کند که چه اقداماتی باید انجام بگیرد. در تحقیقات فعلی، اگر حسابرس جستجوهای اطلاعاتی ساده‌ای به کار برد که دارای ارزش شناختی نباشد یا ارزش شناختی کمی داشته باشد، در این صورت، او ملزم به انجام رسیدگی کامل است. زیرا به کمک ابزارهای تجزیه و تحلیل داده‌ها، می‌توان به معاملات مشکوک‌تری اشاره کرد و در نتیجه، موجب افزایش بهره‌وری حسابرس

تا این متغیرها در تجزیه و تحلیل داده‌ها و نرم افزار کاوش مورد استفاده حسابرسان قرار گیرد. به طور نمونه، اگر حسابرس بخواهد سوالاتی مطرح کند که شامل شماره شناسایی فروشنده، شماره چک، مبلغ چک و تاریخ چک‌های موجود در پایگاه اطلاعاتی مشتری است، باید نام حوزه‌های مناسب و جداول مربوط در پایگاه اطلاعاتی مشتری را تعیین کند.

گام سوم، حسابرس باید تعیین کند که چه روش‌ها و ابزارهایی را برای بررسی داده‌های جمع‌آوری شده به کار می‌گیرد. در صورت تمایل حسابرس به استفاده

براساس رویه‌ها و سیاست‌های مؤسسات حسابرسی انتخاب خواهند شد. الگوهای حسابرسی توسط موسسه حسابرسی و برای اهداف حسابرسی خاص ایجاد می‌شوند. طوفان فکری مربوط به تقلب در برگرفته جلسه ای متشکل از تیم حسابرسی است و متغیرهایی ممکن است توسط حسابرسان

مستقل پیشنهاد شود که بر اساس تجربیات شخصی خود یا مواردی باشد که در طی حسابرسی فعلی و یا حسابرسی‌های گذشته در مشتری دیده‌اند.

گام دوم، متغیرهای داخلی و خارجی در منابع داده مناسبی ترسیم می‌شود

اگر حسابرس جستجوهای اطلاعاتی ساده‌ای به کار برد که ارزش شناختی کمی داشته باشد، در این صورت ملزم به انجام رسیدگی کامل است

a financial statement audit—Interim statements on auditing standards (SAS) no. 99. Washington: DC: Public Company Accounting Oversight Board; 2003.

- Perols J. Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Auditing 2011; 30 (2): 19–50.
- Van der Aalst WMP. Process mining: Discovery, conformance and enhancement of business processes. 1st. Ed. New York: Springer; 2011.

پی‌نوشت‌ها:

1. Database Management System (DBMS)
2. Data Mining
3. Excel
4. Association for Computational Linguistics (ACL)
5. Case Ware
6. International Debate Education Association (IDEA)
7. Computer Assisted Audit Techniques (CAATs)
8. Knowledge Discovery in Databases (KDD)
9. Management Information System (MIS)
10. Decision tree
11. Bayesian belief networks
12. Neural networks
13. Support vector machines
14. Genetic algorithms
15. Hierarchical methods
16. Density-based techniques
17. Distance-based algorithms
18. Support vector machines (SVM)
19. Enterprise Resource Planning (ERP)
20. Alles
21. Jans
22. Benford's law
23. Self Organizing Map (SOM)
24. Management's Discussion & Analysis (MD&A)
25. Accounting and Auditing Enforcement Release (AAERs)
26. Computational Fraud Detection Model (CFDM)
27. SAS's Enterprise Miner
28. Generally accepted accounting principles (GAAP)
29. Brainstorming

* دانشجوی دوره کارشناسی ارشد دوره حسابرسی موسسه آموزش عالی خاتم

بالاترین سطح از توانایی تشخیص می‌باشد و همچنین موجب کاهش رسیدگی‌های بعدی می‌شود که تیم حسابرسی باید انجام دهد.

منابع و مأخذ:

این مقاله، ترجمه آزادی از مقاله زیر است:

Debreceeny R S, Gray G L. A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits: 2014

سایر منابع و مأخذ:

- Alles M, Brennan G, Kogan A, Vasarhelyi M A. Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at Siemens. Int. J. Account. Inf. Syst. 2006; 7 (2): 137–61.
- Argyrou A. Auditing Journal entries using Self-Organizing Map. 18th Americas Conference on Information Systems. Seattle: Association for Information Systems; 2012.
- Berry M W, Kogan J. Text mining: Applications and theory. Chichester: U.K: Wiley; 2010.
- Coderre D G. Computer-aided fraud prevention and detection: A step-by-step guide. Hoboken, NJ: John Wiley & Sons; 2009.
- Colantonio A, Di Pietro R, Ocello A, Verde NV. A new role mining framework to elicit business roles and to mitigate enterprise risk. Decis. Support. Syst. 2011; 50 (4): 715–31.
- Delen D, Al-Hawamdeh S. A holistic framework for knowledge discovery and management. Commun. ACM 2009; 52 (6): 141–5.
- Eppler M J, Mengis J. The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. Inf. Soc. 2004; 20 (5): 325–44.
- Glancy F H, Yadav S B. A computational model for financial reporting fraud detection. Decis. Support. Syst. 2011; 50 (3): 595–601.
- Han J, Kamber M. Data mining: Concepts and techniques. 2nd ed. San Francisco: Morgan Kaufmann; 2006.
- PCAOB. Consideration of fraud in

شد. ابزار داده‌کاوی دارای بالاترین سطح از توانایی تشخیص می‌باشد و همچنین موجب کاهش رسیدگی‌های بعدی می‌شود که تیم حسابرسی باید انجام دهد.

خلاصه و نتیجه‌گیری

داده‌کاوی، کشف دانش در پایگاه داده‌ها نامیده می‌شود و روشی برای کشف اطلاعات سودمند جدید و بالقوه از بین حجم انبوهی از اطلاعات است. هدف از داده‌کاوی، شناسایی الگوها و ارتباطات قابل فهم، معتبر، جدید و به طور بالقوه سودمند در داده‌های موجود می‌باشد. فناوری داده‌کاوی روابط بین یک متغیر و جمعیت انتخاب شده از دیگر متغیرها را می‌یابد. حسابرسان ممکن است احساس کنند داده‌های مورد بررسی در حسابرسی کافی نیست، مانند اطلاعاتی که تنها شامل بخش کوچکی از اطلاعات مرتبط با قلب مشتری در درون سازمان می‌باشد. علاوه بر این، حجم وسیعی از اطلاعات از خارج سازمان در اختیار صاحبکار قرار می‌گیرد که به طور بالقوه با ارزیابی موارد مربوط به اظهارات اشتباه ناشی از قلب مرتبط است. بیانیه استاندارد حسابرسی شماره ۹۹ حسابرسان را ملزم کرده است که به طور سیستماتیک و منظم، ریسک اشتباهات بالقوه ناشی از قلب را به عنوان بخشی از فرآیند حسابرسی، رسیدگی کند. حسابرسان باید طی فرآیند برنامه‌ریزی حسابرسی ریسک‌های قلب را در مدلی روشمند مانند روش طوفان فکری مطرح کنند که توسط تیم حسابرسی انجام می‌شود. اگر حسابرسان جستجوهای اطلاعاتی ساده‌ای به کار ببرند که دارای ارزش شناختی نباشد یا ارزش شناختی کمی داشته باشد، حسابرسان در این صورت ملزم به انجام رسیدگی کامل است. به کمک ابزارهای تجزیه و تحلیل داده‌ها، می‌توان به معاملات مشکوک اشاره کرد و در نتیجه موجب افزایش بهره‌وری حسابرسی شد. ابزار داده‌کاوی دارای