

## Examining Data Mining in Improving the Audit Process<sup>1</sup>

Mohammad Hossein Setayesh<sup>2</sup>, Mina Sadeghi<sup>3</sup>, Younes  
Masoudi<sup>4</sup>, Elias Dehdari<sup>5</sup>

Received: 2025/02/13

Accepted: 2025/05/17

Research Paper

### Abstract

Auditing has become an essential field worldwide due to the increasing evidence of intentional manipulations in financial reports. This study examined the role of data mining in improving the audit process. The research was conducted using a qualitative descriptive method. Data collection involved reviewing books, theses, journals, and related studies on the topic. Given the breadth and abundance of sources, both printed and digital materials relevant to the subject and accessible were selected for sampling. Additionally, receipts were used to gather information from the sample.

The study's results showed that data mining techniques for fraud detection begin with feature selection. The core principle of data mining is that data can reveal valuable insights. Although data mining is a relatively new field, it has proven to be widely applicable in financial decision-making processes. Applications of data mining methods, based on relevant studies and their nature, span a broad range, including bankruptcy prediction, credit risk assessment, going concern evaluation, financial distress analysis, business unit performance forecasting, and management fraud detection.

**Keyword:** Auditing, Data mining, Process Improvement.

**JEL Classification:** M42, O33.

1. doi: 10.22051/jera.2025.49993.3429

2. Professor, Accounting Department, Faculty of Economics, Management and Social Sciences, Shiraz University, Shiraz, Iran. (setayeh@shirazu.ac.ir).

3. Ph.D. Student, Department of Accounting, Faculty of Economics, Management and Social Sciences, Shiraz University, Shiraz, Iran. (mina\_sadeghi96@yahoo.com).

4. Ph.D. Student, Department of Accounting, Faculty of Economics, Management and Social Sciences, Shiraz University, Shiraz, Iran. (younes.masoudi@mail.um.ac.ir).

5. Ph.D. Student, Department of Accounting, Faculty of Economics, Management and Social Sciences, Shiraz University, Shiraz, Iran. (el.dehdari@gmail.com).

## بررسی داده‌کاوی در بهبود فرآیند حسابداری<sup>۱</sup>

محمدحسین ستایش<sup>۲</sup>، مینا صادقی<sup>۳</sup>، یونس مسعودی<sup>۴</sup>، الیاس دهداری<sup>۵</sup>

تاریخ دریافت: ۱۴۰۳/۱۱/۲۵

تاریخ پذیرش: ۱۴۰۴/۰۲/۲۷

مقاله پژوهشی

### چکیده

حسابرسی تبدیل به یک موضوع ضروری در دنیا شده است زیرا شواهد زیادی در دستکاری‌های عمدی در گزارشات وجود دارد. پژوهش حاضر به بررسی داده‌کاوی در بهبود فرآیند حسابداری پرداخته است. این پژوهش به روش کیفی از نوع پژوهش توصیفی انجام یافته است. جمع‌آوری اطلاعات پژوهش شامل بررسی کتاب‌ها، پایان‌نامه‌ها، مجله‌ها و پژوهش‌های مرتبط با موضوع است. به علت وسعت و کثرت منابع، برای نمونه‌گیری از منابع چاپی و دیجیتالی که مرتبط با بحث و در دسترس بود، استفاده گردید. در این پژوهش از فیش برداری جهت گردآوری اطلاعات از نمونه، بهره‌گرفته شد. نتایج پژوهش نشان داد تکنیک‌های داده‌کاوی برای شناسایی تقلب با انتخاب ویژگی آغاز می‌شود. ایده‌ی اصلی داده‌کاوی این است که داده‌ها می‌توانند سخن بگویند. فناوری داده‌کاوی، با توجه به این که زمینه جدیدی است، اما نشان داده که در فرآیند تصمیم‌گیری امور مالی کاربرد فراوانی دارد کاربرد روش‌های داده‌کاوی با توجه به مطالعات مرتبط و ماهیت آن‌ها، می‌تواند طیف گسترده‌ای شامل پیش‌بینی ورشکستگی، تخمین ریسک اعتباری، وضعیت تداوم فعالیت، درماندگی مالی، پیش‌بینی عملکرد واحد تجاری و تقلب مدیریت را در برگیرد.

واژه‌های کلیدی: بهبود فرآیند، حسابداری و داده‌کاوی.

طبقه‌بندی موضوعی: M42, O33

doi: 10.22051/jera.2025.49993.3429

۱. استاد، گروه حسابداری، دانشکده اقتصاد، مدیریت و علوم اجتماعی، دانشگاه شیراز، شیراز، ایران. (نویسنده مسئول). (setayesh@shirazu.ac.ir).

۲. دانشجوی دکتری، گروه حسابداری، دانشکده اقتصاد، مدیریت و علوم اجتماعی، دانشگاه شیراز، شیراز، ایران. (mina\_sadeghi96@yahoo.com).

۳. دانشجوی دکتری، گروه حسابداری، دانشکده اقتصاد، مدیریت و علوم اجتماعی، دانشگاه شیراز، شیراز، ایران. (younes.masoudi@mail.um.ac.ir).

۴. دانشجوی دکتری، گروه حسابداری، دانشکده اقتصاد، مدیریت و علوم اجتماعی، دانشگاه شیراز، شیراز، ایران. (el.dehdari@gmail.com).

۵. <https://jera.alzahra.ac.ir>

مقدمه<sup>۱</sup>

دوره اقتصاد جهانی به سرعت در حال تحول است و این تغییرات سازمان‌ها و شرکت‌ها را به این وادار می‌کند که برای حفظ رقابت‌پذیری، از مجموعه‌ای متنوع از روش‌های مدیریتی نوآورانه و فناوری‌های پیشرفته اطلاعاتی استفاده کنند. به همین دلیل، داده‌کاوی نقش بسیار مهمی در آینده حرفه حسابداری و حسابداری ایفا خواهد کرد. داده‌کاوی به عنوان یکی از مفاهیم برجسته برای تحلیل هوشمند کسب‌وکار و ابزار حمایتی تصمیم‌گیری شناخته می‌شود. این روش با امکانات و مزایای وسیع خود، از جمله پیش‌بینی دقیق روندهای توسعه آینده سازمان‌ها و تقویت رقابت‌پذیری، به شرکت‌ها کمک می‌کند تا تصمیمات بهتری اتخاذ کنند. همچنین، با توجه به گسترش روزافزون اطلاعات در سیستم‌های یکپارچه و شبکه جهانی وب، نیاز به یک علم مؤثر برای شناسایی و استخراج سریع و صحیح داده‌ها از این حجم عظیم اطلاعات بیش از پیش احساس می‌شود و در این راستا، داده‌کاوی به عنوان فناوری اصلی برای استفاده کارآمد، سریع و دقیق از داده‌های کلان به شمار می‌آید. در حال حاضر، ابزارها و روش‌های داده‌کاوی در حوزه‌های مالی و حسابداری اهمیت بسیار زیادی پیدا کرده‌اند (رضایی، اصغری شلمانی، دلیری دهبه، ۱۴۰۰). در طول سال‌های گذشته، یک انقلاب تکنولوژیک رخ داده است اینترنت که به ویژه به ظهور خدمات مدرن هموار شد در تجارت الکترونیک و انتقال پول تجارت الکترونیک یکی از بسیاری از حوزه‌های اقتصادی در اطلاعات و ارتباطات فن‌آوری که به بهبود کسب و کار کمک کرد، راهی برای مدیریت شرکت‌های متوسط و کوچک، کاهش هزینه‌ها و صرفه‌جویی در زمان و افزایش بهره‌وری هموار شد (خالدقربان و پریتگا، ۲۰۲۲).

داده‌کاوی رویکردی است که برای استخراج داده‌های معنادار از یک مجموعه داده معین با استفاده از یک یا چند رویکرد مانند تکنیک‌های آماری، یادگیری ماشینی، ریاضی یا هوش مصنوعی استفاده می‌شود. در میان این رویکردها، انواع مختلفی از تکنیک‌ها را می‌توان برای کلاهداری مالی به کار برد. مانند بیز ساده<sup>۳</sup>، ماشین بردار پشتیبانی<sup>۴</sup>، رگرسیون لجستیک<sup>۵</sup> و بسیاری دیگر (آلبشراوی و لوول، ۲۰۱۶).

۱. مقاله منتخب بیست و دومین همایش ملی حسابداری ایران: اسفند ۱۴۰۳

2. Khaled Gubran & Peritga  
3. NB  
4. SVM  
5. LR

در دهه اخیر، حجم زیادی از داده‌ها در سیستم‌های اطلاعاتی و بانک‌های داده ذخیره شده است که عمدتاً از طریق نرم‌افزارهای تجاری، تراکنش‌های مالی، منابع و مدیریت کسب‌وکار و همچنین مدیریت ارتباط با مشتری جمع‌آوری شده‌اند (تانگ و مک کلینان<sup>۱</sup>، ۲۰۰۵).

هدف از داده‌کاوی شناسایی الگوها و روابط معتبر، جدید، بالقوه مفید و قابل درک در داده‌های موجود است. در افکار عمومی، کاوش داده‌ها به فرآیند یافتن راهکارهای اطلاعاتی برای سازمان‌ها و شرکت‌ها اشاره دارد. تعریف دقیقی از داده‌کاوی وجود ندارد، اما به طور کلی، داده‌کاوی به استخراج اطلاعات پنهان، ناشناخته و بالقوه سودمند از مجموعه‌های بزرگ پایگاه‌های داده اطلاق می‌شود که قادر است تحلیل‌های ارزشمندی از میان حجم زیاد داده‌ها ارائه دهد (لیو و چن<sup>۲</sup>، ۲۰۰۹). به عبارت دیگر داده‌کاوی علم استخراج اطلاعات مفید از پایگاه‌های داده یا مجموعه داده‌ای می‌باشد (هند، هیکی، پدراایک<sup>۳</sup>، ۲۰۰۱).

بسیاری از افراد داده‌کاوی را معادل عبارت کشف دانش از داده‌ها می‌دانند، در حالی که برخی دیگر آن را یکی از مراحل اصلی در فرآیند کشف دانش می‌پندارند. فرآیند کشف دانش را می‌توان شامل مراحل زیر دانست:

- ۱- **پالایش داده‌ها:** در این مرحله داده‌های نامناسب و ناسازگار حذف می‌شوند.
- ۲- **یکپارچه‌سازی داده‌ها:** در این مرحله ممکن است منابع مختلف داده با یکدیگر ترکیب شوند.
- ۳- **انتخاب داده‌ها:** در این گام، داده‌های مرتبط با تحلیل از منابع داده‌ای استخراج می‌شوند.
- ۴- **تبدیل داده‌ها:** داده‌ها از طریق تکنیک‌هایی مانند خلاصه‌سازی و تجمع به شکل‌های مناسب‌تری تبدیل می‌شوند.
- ۵- **داده‌کاوی:** این فرآیند کلیدی با بهره‌گیری از روش‌های هوشمند، الگوهای نهفته در داده‌ها را شناسایی می‌کند. *پایگاه علوم انسانی و مطالعات فرهنگی*
- ۶- **ارزیابی الگو:** در این مرحله، الگوهای مهم و جالب از میان تعداد زیادی الگو بر اساس معیارهای مختلف معرفی می‌شوند. *پایگاه علوم انسانی*
- ۷- **ارائه دانش:** مرحله‌ای که در آن با استفاده از تکنیک‌های بصری، دانش استخراج شده به کاربر نمایش داده می‌شود.

مراحل یک تا چهار به عنوان فرآیند پیش‌پردازش داده‌ها در نظر گرفته می‌شوند. طی این مراحل، داده‌ها برای اجرای فرآیند داده‌کاوی آماده‌سازی می‌شوند. همچنین، عملیات داده‌کاوی می‌تواند با مشارکت کاربر یا بهره‌گیری از پایگاه‌های دانش صورت پذیرد (نمازی و مهارلویی، ۱۳۹۷).

در مجموع در پژوهش حاضر به دنبال پاسخ به این سوال هستیم که آیا بررسی داده‌کاوی در فرآیند حسابرسی تاثیر گذار است یا خیر؟

### مبانی نظری

داده‌کاوی از مهمترین ابزارهای در دسترس سازمان‌ها برای ارزیابی، شناسایی و پیش‌بینی ریسک‌های متنوع پیش‌روست است. داده‌کاوی به معنای تجزیه و تحلیل مجموعه‌ای از داده‌های مشاهده‌شده، که معمولاً دارای حجم بالایی هستند، با هدف کشف روابط پنهان و ارائه خلاصه‌ای از داده‌ها به شیوه‌ای نوآورانه است، به گونه‌ای که برای صاحب داده‌ها نیز قابل فهم و بهره‌برداری باشد (هند و همکاران، ۲۰۰۱). مدیریت ابرداده‌ها برای حسابداران و متخصصان حوزه مالی، امری فراتر از فرصتی برای تغییر روش بازی است. آینده تنها محدود به انتقال مستقیم مهارت‌های فعلی نیست؛ بلکه مهارت‌های جدیدی باید توسعه داده شود. با توجه به این نکته، ظهور ابرداده‌ها تغییرهایی را ایجاد می‌کند که برای انطباق با آن‌ها نه تنها باید مهارت‌های جدیدی آموخته شوند، بلکه باید مهارت‌های جدیدی نیز ایجاد گردند تا شامل هر دو جنبه پیش گفته باشند. براساس تعریف انجمن حسابداران مدیریت آمریکا<sup>۱</sup>، ابرداده‌ها بر سه حوزه مهم تأثیر می‌گذارد: ارزش‌گذاری دارایی‌های داده‌ای<sup>۲</sup>، تصمیم‌گیری و مدیریت ریسک (سادگالی و بنابو<sup>۳</sup>، ۲۰۱۹).

رشد در تجارت الکترونیک اکثر شرکت‌ها و سازمان‌ها را قادر به انجام کار می‌کند تراکنش‌های مالی آن‌ها به صورت الکترونیکی از طریق اتخاذ سیستم‌های پرداخت مانند سیستم‌های بیمه مراقبت‌های بهداشتی، مخابرات سیستم‌ها و بخش مالی اخیراً یک افزایش محسوس در تعداد تراکنش‌های مالی به دلیل پذیرش گسترده خدمات بانک‌های اینترنتی و

1. Institute of Management Accountants

۲. اصطلاح دارایی‌های داده‌ای (Data Assets) شامل هر نوع قطعه اطلاعاتی است که ثبت‌های ارزشمندی را در برگیرد.

3. Sadgali & Benabbou

مؤسسات مالی و همچنین تجارت الکترونیک وجود دارد. اینترنت نقش مهمی ایفا کرده است. نقش در پرداخت آنلاین، که بستر مناسبی برای مهاجمان مخرب برای سوء استفاده از این خدمات برای انجام مشاغل متقلبانه منجر به ظهور جرایم سایبری که خدمات بانکداری الکترونیک را هدف قرار می‌دهد واقع شده است (سادگالی و بنابو، ۲۰۱۹).

کلاهبرداری مالی یک مسئله اساسی است که زندگی روزمره نقش مهمی را تحت تأثیر قرار می‌دهد در یکپارچگی و اعتماد در بخش‌های مالی نیز به عنوان هزینه زندگی افراد تأثیر گذار است که به عنوان کلاهبرداری مالی شناخته شده است و باعث ایجاد یک سوء استفاده مالی و که یک نگرانی بزرگ در جامعه اقتصادی است. ایجاد خسارات هنگفت به اقتصاد دولت‌ها، سازمان‌ها، بخش‌های شرکتی یا حتی افراد را باعث می‌شود. می‌توان آن را به صورت تعریف کرد یک عمل نادرست یا غیرقانونی که منجر به سودمندی می‌شود (وانگ، یو، یای، ۲۰۱۸).

داده کاوی یعنی استخراج اطلاعات مفید از مجموعه‌ی داده‌های خام و نامفهوم که این اطلاعات شامل اجزای مختلفی هستند، از جمله کاربردهای دیگر داده کاوی عبارتند از:

- کشف الگوی میان داده‌ها
- پیش‌بینی حدودی نتایج
- به‌دست آوردن اطلاعات کاربردی
- تمرکز بر روی داده‌های بزرگ

به‌طور کلی فرآیند داده کاوی علاوه بر اینکه به ما کمک می‌کند داده‌های نامرتبط و بلااستفاده را از مجموعه‌ی خود حذف کنیم، از طرفی اطلاعات بسیار مفید و کاربردی را در اختیار ما (سازمان) قرار می‌دهد و همچنین به فرآیندهای تصمیم‌گیری سرعت می‌بخشد. به‌طور کلی اکثر سیستم‌های داده کاوی با دو مشکل اساسی دست و پنجه نرم می‌کنند، این مشکلات عبارتند از:

- حجم بالای داده‌های موجود در ورودی
- عدم اطمینان کامل به اطلاعات خروجی

## داده‌های بزرگ<sup>۱</sup>

داده‌های بزرگ در چند سال گذشته علاقه زیادی به حوزه‌های مختلف علمی و مهندسی ایجاد کرده است. علیرغم مزایا و کاربردهای فراوان، چالش‌های زیادی در داده‌های بزرگ وجود دارد که باید برای کیفیت بهتر خدمات با آن‌ها مقابله کرد، به عنوان مثال، تجزیه و تحلیل داده‌های بزرگ، مدیریت داده‌های بزرگ، و حفظ حریم خصوصی و امنیت کلان داده‌ها. کلان داده به معنای میزان عظیمی داده‌های ساختاربندی شده و نشده است که این پتانسیل را دارد که به شرکت‌ها کمک کند تا عملیات‌های خود را، بهبود بخشیده و تصمیمات سریعتر و هوشمندانه‌تری اتخاذ نمایند. تعداد این داده‌ها به قدری است که پردازش آن‌ها به وسیله دیتابیس‌های سنتی و نرم افزارهای موجود، دشوار بود.

در حال حاضر بسیاری از صنایع و دولت‌ها از ارزش داده‌های بزرگ آگاهی دارند. در واقع، استخراج کارآمد داده‌های بزرگ باعث بهتر شدن مزیت رقابتی و ارزش افزوده برای بسیاری از بخش‌ها (اقتصادی، اجتماعی، پزشکی، تحقیقاتی و غیره) می‌شود. داده‌های بزرگ بطور اساسی توسط سه ویژگی اساسی خود مشخص می‌شوند که عبارتند از سرعت (داده‌ها با سرعت بالایی افزایش و تغییر می‌یابند)، گوناگونی (داده‌ها دارای انواع قالب‌های متفاوتی هستند) و حجم (مقدار زیادی از داده‌ها در هر ثانیه تولید می‌شوند). این سه ویژگی باید همزمان وجود داشته باشند تا اثبات شود یک منبع از نوع منبع داده‌های بزرگ است. اگر یکی از این سه ویژگی وجود نداشته باشند نمی‌توان درباره‌ی داده‌های بزرگ بحث کرد. بسیاری از محققان فرصت‌های مرتبط با داده‌های بزرگ را در زمینه‌ی مربوط به خود مشخص کرده‌اند. در این بخش، فرصت‌ها و مزایای داده‌ها بزرگ را در بخش‌های زیر بصورت خلاصه بیان می‌کنیم: سلامت، تحقیقات، تجارت، حمل و نقل، گردشگری، سیاست. هدف ما ارائه‌ی تصویری بزرگ از این موضوع است.

فناوری و روش‌های داده‌های متعارف اکثر اوقات کند و گران‌قیمت هستند و برای کار روی ذخیره‌سازی و پردازش حجم زیاد و رو به رشد داده‌های نامتجانس مناسب نیستند. یکی از چالش‌ها نحوه‌ی فائق آمدن بر ماهیت پیچیده‌ی داده‌های بزرگ (حجم، سرعت و تنوع) است.

چالش دیگر در حال تغییر برای بسیاری از نهادهای مربوطه بسته به نقش آنهاست. علاوه بر این، بازیگران داده‌های بزرگ «بینش بزرگ» نحوه‌ی نمایش موثر هنگام تامین داده‌های بزرگ با چالش‌های دیگری روبرو هستند مثلاً نحوه‌ی تامین مجموعه داده‌های بزرگ که در حال تکامل هستند، نحوه‌ی یکپارچه‌سازی لایه‌های امنیتی بدون تاثیر گذاشتن بر عملکرد سیستم‌ها، نحوه‌ی یکپارچه‌سازی سیاست‌های حریم خصوصی در پلت فرم‌های مشترک داده‌های بزرگ در حین تامین دسترسی سریع و گران‌تری به داده‌ها (چن، دو، ۲۰۰۹).

### روش‌های داده‌کاوی

اهداف داده‌کاوی را می‌توان به دو دسته پیش‌بینی و توصیف یا ترکیبی از هر دو تقسیم کرد. پیش‌بینی بر دقت و توانایی در پیش‌بینی نتایج تمرکز دارد، در حالی که توصیف به فرآیند استخراج و ارائه داده‌ها مرتبط است. در پیش‌بینی، مادامی که مدل توانایی پیش‌بینی دارد، برای کاربر اهمیت چندانی ندارد که مدل تا چه حد بازتاب‌دهنده واقعیت باشد؛ به‌عنوان نمونه، مدلی که شاخص‌های مالی را به‌صورت غیرخطی ترکیب می‌کند تا نرخ ارز را تخمین بزند. از سوی دیگر، مدل توصیفی بر تفسیر واقعیت تأکید دارد. به‌عنوان مثال، مدلی که متغیرهای اقتصادی و جمعیتی را با سطح پیشرفت آموزشی مرتبط می‌کند، می‌تواند مبنایی برای ارائه پیشنهادی سیاست‌گذاری اجتماعی باشد. در بسیاری از موارد، کاربردهای کشف دانش به‌صورت ترکیبی از مدل‌های پیش‌بینی و توصیفی هستند و به هر دو نوع مدل نیاز دارند (دستگیر و شفیع‌ی سردشت، ۱۳۹۰).

اهداف داده‌کاوی از طریق استفاده از روش‌ها و تکنیک‌های مختلف داده‌کاوی به دست می‌آید. این اصطلاحات و روش‌ها در حقیقت نمایانگر مجموعه‌ای از الگوریتم‌ها و فنون هستند که از علوم مختلفی مانند آمار، یادگیری ماشین، پایگاه‌های داده و تحلیل‌های استنتاجی نشأت می‌گیرند. برخی از این تکنیک‌ها در زمینه داده‌های مالی نیز به کار گرفته شده‌اند (علیخاندان، ۱۳۸۵).

### شبکه‌های عصبی

شبکه‌های عصبی مشابه مغز انسان عمل کرده و از نظر ساختاری به آن شباهت دارند. مغز به عنوان یک سیستم پردازش اطلاعات از اجزای بنیادی به نام نورون‌ها تشکیل شده است. در شبکه‌های عصبی مصنوعی نیز مجموعه‌ای از نورون‌های متصل به یکدیگر وجود دارد که به هر

کدام از این مجموعه‌ها لایه گفته می‌شود. نورون یا گره کوچک‌ترین واحد پردازش اطلاعات است که پایه‌گذار عملکرد شبکه‌های عصبی می‌باشد. شبکه‌های عصبی با وجود تنوع زیاد، ساختار مشابهی دارند. به طور معمول، یک شبکه عصبی از سه لایه ورودی، پنهان و خروجی تشکیل می‌شود. لایه ورودی تنها اطلاعات را دریافت کرده و مانند متغیر مستقل عمل می‌کند، بنابراین تعداد نورون‌های آن به تعداد متغیرهای وابسته بستگی دارد. در حالی که لایه پنهان هیچ معنای مستقیمی ندارد و تنها به عنوان نتیجه‌میان‌ی در محاسبات خروجی عمل می‌کند، این لایه در فرآیند آموزش از اهمیت ویژه‌ای برخوردار است (پورحیدری و اعظمی، ۱۳۸۹).

یکی از معایب شبکه‌های عصبی این است که تفسیر چگونگی دستیابی این شبکه‌ها به تصمیمات خود، امری پیچیده و دشوار است. از سوی دیگر، نقدهایی که به این شبکه‌ها وارد شده، اشاره به تعدادی از عواملی دارد که باید به صورت تجربی تعریف شوند، مانند توپولوژی شبکه. برخلاف روش‌های آماری سنتی، شبکه‌های عصبی مصنوعی نیازی به فرضیات خاصی درباره ویژگی‌های توزیع داده‌ها و متغیرهای مستقل ورودی ندارند (دستگیر و شفیع‌سردشت، ۱۳۹۰).

شبکه‌های عصبی عمدتاً در شناسایی تقلب‌های مربوط به کارت‌های اعتباری، بیمه خودرو و سوءاستفاده‌های شرکتی مورد استفاده قرار می‌گیرند (نگایی، دینه، سویتا، پروان، ۲۰۲۲). چن و دو (۲۰۰۹) در پژوهشی با بهره‌گیری از شبکه‌های عصبی مصنوعی، ۶۸ شرکت فعال در بورس تایوان را بررسی کردند. آن‌ها با ترکیب داده‌های مالی و غیرمالی، مدلی برای پیش‌بینی بحران مالی ایجاد کردند. یافته‌های این پژوهش نشان داد که شبکه‌های عصبی مصنوعی در مقایسه با روش‌های آماری سنتی، عملکرد دقیق‌تری در پیش‌بینی بحران‌های مالی دارند.

### درخت تصمیم

درخت تصمیم در واقع یک نمودار گرافی است که ساختاری مشابه درخت دارد و هر گره آن نمایانگر یک آزمون یا مقایسه برای یک متغیر خاص است. یال‌هایی که از گره‌ها منشعب می‌شوند، نشان‌دهنده تصمیماتی هستند که بر اساس هر نتیجه‌گیری از آزمون‌های انجام‌شده به دست می‌آید. در این شیوه، تلاش بر این است که مشاهدات به گروه‌های مشخص تقسیم شوند. یکی از مزایای مهم این روش، عدم وابستگی به نوع توزیع داده‌ها و روابط میان متغیرهای ورودی

است. در واقع، درخت تصمیم مدلی مفهومی و ساده ارائه می‌دهد که بر اساس مجموعه‌ای از تصمیمات پایه‌ای عمل می‌کند. همچنین، الگوریتم‌های یادگیری مورد استفاده در این روش از سرعت بالایی برخوردار هستند. یکی از ویژگی‌های کلیدی درخت‌های تصمیم، توانایی آن‌ها در ساده‌سازی فرایندهای پیچیده تصمیم‌گیری به مجموعه‌ای از تصمیمات ساده‌تر است که به راحتی قابل تفسیر و درک می‌باشند (یزدانی و کنگاوری، ۱۳۸۳).

### رگرسیون لجستیک

امروزه در بسیاری از پژوهش‌ها، از ترکیب چندین عامل برای دستیابی به یک هدف خاص استفاده می‌شود، به گونه‌ای که نتیجه‌ای بهینه حاصل گردد. در علم آمار، این نوع تحلیل‌ها اغلب با بهره‌گیری از روش‌های گوناگون رگرسیون انجام شده و نتایج آن بررسی می‌شود. در روش رگرسیون، متغیرهای مستقل به منظور برآورد متغیر پاسخ (یا همان متغیر هدف) مورد استفاده قرار می‌گیرند. رگرسیون لجستیک نوع خاصی از رگرسیون محسوب می‌شود که زمانی به کار می‌رود که متغیر پاسخ دارای دو یا چند دسته مجزا باشد. این روش یک مدل ریاضی ارائه می‌دهد که می‌تواند ارتباط بین چندین متغیر مستقل  $X$  و یک متغیر وابسته دوحالتی یا چندحالتی  $Y$  را توصیف کند. متغیر دوحالتی به متغیری اطلاق می‌شود که تنها دو وضعیت ممکن دارد، مانند زنده یا مرده بودن، حضور یا غیبت، و داشتن یا نداشتن یک ویژگی خاص. معمولاً برای این نوع متغیرها از کدهای صفر و یک استفاده می‌شود، به این صورت که مقدار یک برای نشان دادن وجود آن ویژگی و مقدار صفر برای عدم وجود آن در نظر گرفته می‌شود (ستایش، فتاحی نافچی، عباس پور، روستایی، ۱۳۹۳).

### الگوریتم ژنتیک

این الگوریتم بر پایه نظریه تکامل داروین بنا شده و کاربرد آن بر اصول ژنتیک طبیعی متمرکز است. مفاهیم ابتدایی این الگوریتم نخستین بار در سال ۱۹۶۲ توسط هلند و همکارانش مطرح شد. مطابق این نظریه، موجوداتی که سازگاری بیشتری با محیط خود دارند، از شانس بالاتری برای بقا و تولیدمثل برخوردارند، و در نتیجه احتمال انتقال ویژگی‌های آن‌ها به نسل‌های آینده افزایش می‌یابد (موسی زادگان و ذکری، ۱۳۸۷). در الگوریتم ژنتیک، هر پاسخ به یک مسئله خاص در قالب یک مجموعه از مقادیر عددی بیان می‌شود. این مجموعه‌ها که به آن‌ها رشته گفته می‌شود، نشان‌دهنده یک جواب از مجموعه پاسخ‌های مسئله هستند. الگوریتم ژنتیک فرآیند خود را با

مجموعه‌ای از پاسخ‌های اولیه که نسل نخست را تشکیل می‌دهند، آغاز می‌کند. در طول تکرارها و نسل‌های متوالی، تعداد این پاسخ‌ها ثابت می‌ماند، اما کیفیت آن‌ها به مرور ارتقا پیدا می‌کند. در هر مرحله، پاسخ‌ها از طریق یک تابع هدف ارزیابی شده و نسل جدید بر مبنای این ارزیابی شکل می‌گیرد. به این صورت که برخی پاسخ‌ها مستقیماً به نسل بعد منتقل می‌شوند، در حالی که برخی دیگر از طریق اعمال تغییرات در کروموزوم‌هایشان دچار جهش شده و به پاسخ‌های جدیدی تبدیل می‌شوند. در اصل، این الگوریتم با بهره‌گیری از مفاهیم تکامل، به دنبال یافتن راه‌حلی بهینه برای مسائل مورد نظر است (دستگیر و شفیع سردهشت، ۱۳۹۰).

### نظریه مجموعه اولیه

نظریه مجموعه ابتدایی در سال ۱۹۹۱ توسط پالاک معرفی شد. این نظریه به عنوان تعمیمی از مفهوم دوتایی در نظر گرفته می‌شود که از نظریات شناخته‌شده در ریاضیات مربوط به بازه‌ها است (عربانی و کلانتری، ۱۳۸۱).

اگرچه مسئله اندازه‌گیری کیفیت، بهبود کیفیت یادگیری و تکمیل دانش ناظران در مواجهه با اطلاعات ناقص هنوز به طور کامل حل نشده است، اما با این حال نظریه مجموعه ابتدایی، ابزاری قدرتمند برای استخراج دانش از اطلاعات پراکنده و ناقص فراهم می‌آورد. این نظریه می‌تواند برای توضیح و ارزیابی وابستگی بین ویژگی‌ها و برای سنجش اهمیت ویژگی‌ها و ارتباط آنها با داده‌های متناقض نیز مورد استفاده قرار گیرد.

استدلال مبتنی بر مورد:

استدلال مبتنی بر مثال، روشی است که برای حل مسائل جدید، از پاسخ‌های مسائل مشابه قبلی استفاده می‌کند. این تکنیک به گونه‌ای طراحی شده است که با استفاده از تجربیات گذشته و نتایج به دست آمده، برای مواجهه با مسائل مشابه در آینده به کار می‌رود. استدلال مبتنی بر مثال به روشی اشاره دارد که در آن افراد از راه‌حلی که برای مسائل پیشین استفاده کرده‌اند، برای مواجهه با مسائل جدید بهره می‌برند (فانز، قدسی پور و غضنفری، ۱۳۸۵). به این ترتیب، از دانش به دست آمده از حل مسائل قبلی به عنوان مرجعی برای رسیدن به راه‌حلی برای مسائل جدید استفاده می‌شود (دستگیر و شفیع سردهشت، ۱۳۹۰).

فرآیند حل مسئله با استفاده از روش CBR شامل چهار مرحله اساسی است:

۱. یافتن مورد مشابه به مسئله جدید؛

۲. به کارگیری پاسخ مسئله مشابه برای ارائه یک راه‌حل پیشنهادی برای مسئله جدید؛
۳. بررسی و اصلاح پاسخ پیشنهادی در صورتی که تفاوت‌هایی بین شرایط مسئله جدید و مسائل مشابه وجود داشته باشد؛
۴. ذخیره‌سازی مسئله جدید و راه‌حل آن برای استفاده در مواقع بعدی.

### نیویز

یکی از روش‌های یادگیری کاربردی و مؤثر، روش یادگیری بیزی است. این روش راه‌حل‌های مفیدی را برای مسائل عملی ارائه می‌دهد. استدلال بیزی بر پایه احتمالات برای استنتاج و نتیجه‌گیری استوار است. اصل اصلی این روش بر این فرضیه بنا شده که برای هر متغیر، یک توزیع احتمالی وجود دارد که با مشاهده داده‌های جدید و تحلیل توزیع احتمال آن، می‌توان تصمیمات بهینه‌ای اتخاذ کرد. این روش در مقایسه با شبکه‌های عصبی و درخت تصمیم‌ها از نظر عملکرد در مسائل دسته‌بندی کارایی مشابهی دارد (ستایش، فتاحی نافچی، عباس پور، روستایی، ۱۳۹۳).

### پارت

پارت یک کلاس برای تولید یک لیست تصمیم است. از الگوریتم پارت برای شناسایی دانش، الگو و همچنین برای شناسایی قوانین مختلف استفاده می‌شود (ستایش، فتاحی نافچی، سمیره عباسپور، روستایی، ۱۳۹۳).

### اهداف داده کاوی

اصلی‌ترین اهداف مرتبط با داده کاوی را می‌توان در دو هدف اصلی به شرح زیر جمع‌بندی کرد (سرکالواندا، ۲۰۰۲).

۱. تایید و رسیدگی: این روش برای تایید یا رد فرضیات مطرح شده، یا برای توضیح شرایط و رویدادهای مشاهده شده به کار می‌رود. محدودیت این هدف آن است که فرضیات، شرایط یا رویدادهای مشاهده شده ممکن است تحت تأثیر آگاهی و دانش تحلیل‌گر قرار گیرند.
۲. کشف: در این رویکرد، فرد سعی دارد به طور خودکار الگوهای ناشناخته گذشته را شناسایی کند. به عبارت دیگر، با بررسی داده‌های گذشته، الگوها و روندهای موجود

را کشف می‌کند تا بتواند از طریق آن‌ها اطلاعات مفیدی در مورد نحوه رفتار و ساختار داده‌ها در آینده کسب نماید.

هر دو هدف فوق در نهایت به یک نتیجه مشترک منتهی می‌شوند که به آن خروجی نهایی داده‌کاوی اطلاق می‌شود و آن شناسایی الگویی است که می‌تواند برای پیش‌بینی روندهای آینده مورد استفاده قرار گیرد.

در سال‌های گذشته حساب‌برسان پی به افزایش شدید حجم معاملات و پیچیدگی داده‌های مالی و غیر مالی شرکت‌های صاحبکار برده‌اند. به منظور حسابرسی این شرکت‌ها حساب‌برسان پی در پی باید با داده‌های حجیم و ساختار پیچیده این داده‌ها سرو کار داشته باشند. چن و دوو (۲۰۰۰) در پژوهشی با عنوان «استفاده از روش‌های داده‌کاوی و شبکه‌های مصنوعی برای پیش‌بینی مدل اضطراب مالی» از شبکه‌های عصبی مصنوعی برای پیش‌بینی تقلب استفاده کردند و نشان دادند که شبکه‌های عصبی مصنوعی از دقت بالایی در کشف تقلب برخوردارند. بهیمنی (۲۰۰۹) به بررسی تداوم فعالیت از دیدگاه حساب‌برس پرداختند. در این پژوهش با توجه به متغیرهای مالی و غیر مالی، احتمال دریافت اظهار نظر تعدیل شده را پیش‌بینی کردند. ابهام در تداوم فعالیت گزارش حساب‌برسان بررسی شد و بیان کردند که شرکت‌های دریافت‌کننده این گزارشها، مدت زمان بیشتری برای دریافت وام باید سپری کنند. بنابراین، حساب‌برسان دیگر نمی‌توانند تنها به ابزارهای گزارش‌دهی یا ساده‌سازی فرآیند حسابرسی تکیه کنند. ابزارهایی مانند داده‌کاوی که به صورت خودکار اطلاعات را از بین حجم وسیعی از داده‌ها استخراج می‌کنند، می‌توانند بسیار کارآمد و مفید باشند. اگرچه استفاده از فنون داده‌کاوی در فرآیند حسابرسی در واقع زمینه نسبتاً جدیدی می‌باشد. اما با توجه به رویداد اخیر از جمله سقوط انرون و آرتور اندرسن نشان‌دهنده مشکلات قابل توجهی در فرآیند حسابرسی است. از طرفی دیگر داده‌کاوی نشان داده است که در بسیاری از کاربردهای تجاری مرتبط با حسابرسی از جمله پیش‌بینی ورشکستگی، تداوم فعالیت و مضیقه مالی، کشف تقلب، پیش‌بینی عملکرد واحد تجاری، ارزیابی ریسک اعتباری دارای مزیت‌های فراوانی خواهد بود (علیخانداده، ۱۳۸۵).

### پیش‌بینی ورشکستگی

یکی از عواملی که می‌تواند به سرمایه‌گذاران در اتخاذ تصمیمات کمک کند، داشتن ابزارها و مدل‌های مناسب برای تحلیل وضعیت مالی و شرایط شرکت‌ها است. یکی از این ابزارها، مدل

پیش‌بینی ورشکستگی شرکت هاست. سرمایه‌گذاران همواره در پی این هستند تا با امکان پیش‌بینی ورشکستگی شرکت‌ها از سوخت‌اصل و فرع پول خود جلوگیری کنند. بنابراین پیش‌بینی ورشکستگی شرکت‌ها یکی از موضوعات اساسی برای کشورها، سرمایه‌گذاران، سهامداران، مشتریان و ... است. روش‌های مختلفی اعم از روش‌های آماری و داده‌کاوی را می‌توان برای پیش‌بینی ورشکستگی به کار گرفت. روش‌های آماری و کلاسیک سابقه طولانی در پیش‌بینی ورشکستگی شرکت‌ها دارند در حالی که روش‌های داده‌کاوی و هوش مصنوعی به تازگی در این زمینه مورد استفاده قرار می‌گیرند.

کیسی (۱۹۹۸) در پژوهشی ارتباط میان ارقام مالی، متغیرهای مربوط به سازمان حسابرسی و گزارش‌های مشروط حسابرسی در شرکت‌های کوچک را بررسی کردند. یافته‌های این مطالعه نشان می‌دهد که با به کارگیری تحلیل لجستیک چندگانه و استفاده از اطلاعات موجود در صورت‌های مالی، امکان پیش‌بینی بندهای مشروط در گزارش‌های حسابرسی شرکت‌های کوچک فراهم می‌شود. علاوه بر این، نتایج نشان‌دهنده این است که در صورتی که یک شرکت گزارش مشروط دریافت کند، احتمال دریافت گزارش مشروط در سال‌های آینده نیز بیشتر خواهد بود.

پیش‌بینی ورشکستگی مهم‌ترین و مشهورترین کاربرد روش‌های داده‌کاوی در داده‌های مالی می‌باشد. ورشکستگی باعث خسارت‌های اقتصادی برای مدیران و سرمایه‌گذاران و اعتبار دهندگان و تمامی ذی‌نفعان می‌شود. و هزینه‌های اجتماعی زیادی هم در بر خواهد داشت (ومتر<sup>۱</sup>، ۲۰۰۷). پیش‌بینی ورشکستگی توسط داده‌های صورت‌های مالی نخستین بار از آثار التمن سرچشمه گرفته است. التمن استدلال می‌کند ورشکستگی فرآیندی بلندمدت است، از این رو صورت‌های مالی، می‌تواند حاوی هشدارهایی از رخداد یک ورشکستگی قریب‌الوقوع باشند. وی با به کارگیری تحلیل تشخیصی چندگانه مدلی را برای پیش‌بینی ورشکستگی شرکت‌ها ایجاد کرد. در همین حال بسیاری از محققان نیز مدل‌های جایگزین را با استفاده از فنون آماری گسترش می‌دادند.

واگل (۲۰۱۷) ۱۲۰ شرکت را با داده‌کاوی مورد بررسی و تجربه و تحلیل قرار دادند و نشان دادند که شبکه‌های عصبی به عنوان بهترین مدل برای پیش‌بینی ورشکستگی شرکت‌ها توجیه می‌شود و آنها یافتند که استفاده مناسب و انتخاب تکنیک‌های داده‌کاوی به افزایش دقت پیش

بینی مدل کمک می کند. سازمان های مالی می توانند از این مدل های پیش بینی بهره مند شوند زیرا به آن ها اجازه می دهد که موفقیت کسب و کار را در آینده پیش بینی و بر این اساس تصمیم بگیرند.

### کاربرد داده کاوی در کشف تقلب مدیران

تحلیل، ارزیابی و شناسایی گزارش های جعلی که گاهی به نام گزارش های دستکاری شده شناخته می شوند، تاریخچه ای طولانی در عرصه حسابداری و امور مالی دارند. سهامداران مدیران را به عنوان نمایندگان خود در شرکت ها منصوب می کنند و به همین دلیل طبیعی است که نسبت به دقت و صحت گزارش های مالی مدیران حساسیت نشان دهند. حسابرسان نیز در جریان بررسی های خود مسئول ارزیابی و شبیه سازی ریسک تقلب در روش های به کار رفته توسط مدیریت سازمان هستند تا بر اساس این ارزیابی ها، میزان آزمون های حسابرسی و در نتیجه هزینه های پیشنهادی خود را پیش بینی کنند. تقلب از سوی مدیریت شرکت می تواند باعث گمراه کردن ذینفعان از طریق گزارش های مالی نادرست و گمراه کننده شود که این موضوع تأثیر منفی بر قیمت سهام خواهد گذاشت (موهاناد<sup>۱</sup>، ۲۰۲۲).

برای تحلیل متن، محققان معمولاً از روش های مبتنی بر "داده کاوی" بهره می برند. در حال حاضر، داده کاوی به یکی از ابزارهای حیاتی در زمینه هوش تجاری تبدیل شده است. حسابداری برای بیش از سه دهه از این تکنیک ها استفاده کرده است. در حقیقت، حسابداری یکی از اولین حوزه های تجاری است که به منظور بهبود فرآیند تصمیم گیری از روش های داده کاوی بهره برده است. داده کاوی به عنوان یک فرآیند برای شناسایی الگوهای نوین، مفید و پیچیده که برای انسان قابل درک نیستند، از حجم زیادی از داده ها و همچنین به عنوان روشی برای استخراج دانش از میان انبوهی از اطلاعات شناخته می شود (امانی و فادللا<sup>۲</sup>، ۲۰۱۷).

امانی و فادللا (۲۰۱۷) تحقیقات خود را که در آن ها از داده کاوی به منظور ارزیابی، شناسایی، محاسبه یا پیش بینی ریسک تقلب در حسابداری استفاده شده، به دو دسته "گذشته نگر-توصیفی" و "آینده نگر-پیش بینی" تقسیم کرده اند. پژوهش حاضر با تحلیل محتوای گزارش های هیئت مدیره به مجامع عمومی شرکت ها در سال های گذشته، کلمات کلیدی با

1. Mohand

2. Amani and Fadlalla

بیشترین تأثیر در توصیف شاخص ریسک تقلب در این شرکت‌ها را شناسایی کرده است. سپس با بهره‌گیری از این واژه‌ها، شرکت‌هایی با ریسک تقلب بالا از سایر شرکت‌ها تفکیک شده‌اند. به این ترتیب، این پژوهش را می‌توان در گروه "گذشته‌نگر-توصیفی" قرار داد. تحقیقات این دسته عمدتاً با مرور وضعیت مالی گذشته یک واحد تجاری، به تحلیلگران این فرصت را می‌دهند تا الگوها و روندهای موجود در داده‌ها را شناسایی کرده و مشکلات و راه‌حل‌های آن‌ها را بررسی نمایند.

### داده‌کاوی و کاربرد آن در حسابداری

امروزه رسوایی جهانی شرکت‌ها و هم‌چنین ورشکستگی موسسات این سوال را مطرح که آیا تکنیک‌های حسابداری انجام شده بروز بوده و با شرایط کنونی انطباق دارد و یا این که استفاده از تکنیک‌های حسابداری بروز به نحو مطلوب انجام نمی‌شود (کووشیارا<sup>۱</sup>، ۲۰۰۷).

### کاربرد ابزار تحلیل داده در حسابداری داخلی

ابزارهای تجزیه و تحلیل داده، فرصت‌های عظیمی را برای حوزه حسابداری داخلی ایجاد کرده‌اند. با ظهور داده‌های کلان، واحدهای حسابداری داخلی در سازمان‌ها به استفاده از این داده‌ها برای شناسایی ریسک‌ها و درک بهتر کسب و کار پی برده‌اند. این امر به حساب‌برسان داخلی کمک می‌کند تا بر حوزه‌ها یا تراکنش‌هایی تمرکز کنند که در آن‌ها کنترلی وجود ندارد یا کنترل‌ها به درستی اجرا نمی‌شوند. حال اگر حساب‌برسان داخلی دارای مجموعه مهارت‌هایی در حوزه سامانه‌های اطلاعاتی باشند، امکان خواهند داشت تا با داده‌کاوی اطلاعات مربوط به صدور صورت حساب مشتریان، روندهای عملیاتی متعددی را شناسایی کنند. این روندها دربرگیرنده مواردی چون اعتبارهای اعطایی، صورت حساب‌های باطل شده یا دوباره محاسبه شده، خطاهای رخداده در فرآیند صدور صورت حساب‌ها و معیارهای کلیدی مورد استفاده در طرح‌های انگیزشی مشتریان، هستند. بدین ترتیب از آنجایی که واحدهای حسابداری به طور مستمر از روش ابزار تحلیل داده استفاده می‌کنند، امکان کشف فرصت‌های جدیدی وجود دارد که می‌توان روی آن‌ها سرمایه‌گذاری کرد. همچنین به دلیل رشد و بلوغ روزافزون این حوزه، ارزیابی ریسک‌های موجود در طول فرآیند برنامه‌ریزی حسابداری نیز بیش از پیش

تبدیل به فرآیندی تحلیل محور شده است. بنابراین، نه تنها ابزار تحلیل داده موجب ایجاد چنین تغییرهایی می‌شود، بلکه می‌توان از آن به منظور تفکیک و بررسی معادله‌ها، مجموعه‌های داده‌ها و فرآیندهایی که در آن‌ها ریسکی مورد شناسایی قرار می‌گیرد، استفاده کرد. از سوی دیگر، کاربرد ابزار تحلیل داده در حوزه حسابرسی مستقل در مقایسه با حسابرسی داخلی، با همان سرعت توسعه نیافته است؛ هرچند قابلیت‌های خوب و مناسبی برای ایجاد تغییرهایی چشمگیر در این حوزه نیز وجود دارد (ACCA، ۲۰۱۳).

### پیشینه پژوهش

در تحقیقات بین‌المللی، کاربرد روش‌های مبتنی بر داده‌کاوی برای شناسایی تقلب‌های مدیریتی، نمونه‌های فراوانی وجود دارد. در ادامه، به برخی از جدیدترین مطالعات مرتبط با این موضوع اشاره خواهد شد.

چن و وو (۲۰۰۹) در پژوهشی با عنوان «استفاده از روش‌های داده‌کاوی و شبکه‌های مصنوعی برای پیش‌بینی مدل اضطراب مالی» از شبکه‌های عصبی مصنوعی برای پیش‌بینی تقلب استفاده کردند و نشان دادند که شبکه‌های عصبی مصنوعی از دقت بالایی در کشف تقلب برخوردارند.

راویسانکار (۲۰۱۱) در پژوهشی با عنوان «تشخیص تقلب در صورت‌های مالی و متغیرهای انتخاب شده با استفاده از روش‌های داده‌کاوی» به بررسی شرکت‌هایی که در گزارش‌های مالی خود دست به تقلب می‌زنند، پرداختند. آن‌ها روش‌های داده‌کاوی را روی مجموعه‌ای از داده‌های شرکت‌ها در دو مرحله مورد آزمایش قرار دادند. در مرحله نخست، با انتخاب ۱۸ ویژگی، شبکه‌های عصبی احتمالاتی عملکرد بهتری نسبت به سایر روش‌ها داشت. در مرحله دوم نیز، با انتخاب ۱۰ ویژگی، شبکه‌های عصبی احتمالاتی و برنامه‌نویسی حاشیه‌ای از نظر دقت در سطحی برابر و برتر از سایر روش‌ها قرار گرفتند.

اولسون (۲۰۱۲) در پژوهشی از روش‌های داده‌کاوی برای پیش‌بینی ورشکستگی شرکت‌ها بهره گرفتند. یافته‌های این تحقیق نشان داد که اگرچه هزینه استفاده از داده‌کاوی در مقایسه با مدل‌هایی مانند شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان، درخت تصمیم و رگرسیون لجستیک بالاتر است، اما استفاده از این روش می‌تواند میزان خطای احتمالی در فرآیند پیش‌بینی را کاهش دهد.

هوبرگ و لوس<sup>۱</sup> (۲۰۱۷) در پژوهش خود، بخش "تجزیه و تحلیل‌های مدیریت" در گزارش‌های شرکتی را بررسی کرده و دریافتند که لحن گزارش‌های مرتبط با ثقل، به طور معناداری از استانداردهای معمول فاصله گرفته و به سمت بیان متقلبانه متمایل است. به عنوان نمونه، در گزارش‌های مشکوک به ثقل، مدیران کمتر به جزئیات عملکرد شرکت پرداخته و بیشتر بر نکات مثبت عملکرد خود تأکید می‌کنند، در حالی که اطلاعات دقیق‌تری درباره تیم مدیریتی ارائه نمی‌دهند. همچنین، وانگ (۲۰۱۷) نشان دادند که هم داده‌های مالی و هم اطلاعات متنی استخراج‌شده از گزارش‌های مدیریتی، می‌توانند در شناسایی شرکت‌های متقلب مؤثر باشند. در این میان، روش‌های مبتنی بر تحلیل بیزی عملکرد بهتری در تفکیک شرکت‌های غیرمتقلب از خود نشان می‌دهند.

کاوام و آموکو، لورد<sup>۲</sup> (۲۰۲۱) این مقاله به جنبه‌هایی از چگونگی شکل‌گیری حسابداری و حسابداران در مسائل اقتصادی و سیاست‌گذاری مربوط به استخراج، تولید و استفاده از سوخت‌های معدنی و فسیلی می‌پردازد. با ارزیابی اینکه آیا حسابداری روایی توسط یک شرکت معدنی فراملی با افرادی که در منطقه‌ای تحت تأثیر عملیات معدن کار می‌کنند یا زندگی می‌کنند هماهنگ است یا خیر، به بررسی نحوه اعمال داده‌ها، محاسبات و ارتباطات مربوط به پایداری می‌پردازد. این شامل ارتباطاتی است که افراد درگیر بین حسابداری و پایداری درک می‌کنند. داده‌ها از طریق کار میدانی کیفی در داخل و اطراف معدن دامنگ در غنا، شامل مصاحبه با کارکنان و در جامعه، و تجزیه و تحلیل اسناد به دست آمد. مدیران شرکت‌ها از حسابداری روایتی برای حمایت از این ادعاها استفاده می‌کنند که سرمایه‌گذاری هنگامی روی پایداری می‌کنند، بنابراین به قول آن‌ها مجوز اجتماعی برای فعالیت از طرف سهامداران جامعه میزبان دارند. این نشان دهنده شکل‌گیری حسابداری در انتخاب‌های تخصیص منابع، از جمله نحوه مدیریت سرمایه سهامداران برای ایجاد ارزش اجتماعی بیشتر و عملکرد پایدار است. با این حال، اگرچه بسیاری از مردم محلی خود را از فعالیت‌های اقتصادی-اجتماعی شرکت معدنی بهره‌مند می‌دانند، اما حسابداری را کسب و کار خود نمی‌دانند، بلکه بیشتر اقتصادی است تا محیطی یا اجتماعی. استنباط این است که حسابداری همچنان در خدمت اهداف کنترل مدیریت تولید، توزیع ارزش افزوده به نفع تامین‌کنندگان سرمایه و مدیریت تصویر منعکس شده در مفهوم

---

1. Hoberg and Lewis

2. Kwame, Amoako & Lord

داشتن مجوز اجتماعی برای فعالیت است. علیرغم اعتقاد آن‌ها به اینکه حسابداری و حسابداران نقشی در پایداری دارند، عموماً نمی‌توانند این نقش‌ها را شناسایی کنند. این یافته‌ها حاکی از آن است که اگر ارائه‌دهندگان حساب در مورد مسئولیت‌پذیری شرکتی در قبال افراد آسیب‌دیده جدی هستند، باید بیشتر تلاش کنند تا اطمینان حاصل شود که جنبه‌های محیطی و اجتماعی به اندازه کافی مورد توجه قرار می‌گیرند تا آن افراد را متقاعد کنند که واقعاً با شرایط مساوی با آن‌ها در ارتباط هستند. متقاعد کردن مخاطبان گسترده‌تر مبنی بر اینکه گزارش‌هایی که تولید می‌کنند قابل اعتماد و مرتبط با پایداری در عمل هستند.

ژنگ، عبدالحمید و هو (۲۰۲۴) به بررسی الگوریتم داده کاوی در شناسایی تقلب حسابداری با استفاده از فناوری اطلاعات شهر هوشمند در بورس شانگهای برای ۶۴۱ گزارش سالانه پرداختند. نتایج نشان می‌دهد که الگوریتم‌های داده کاوی مورد استفاده در فناوری اطلاعات شهر هوشمند برای شناسایی تقلب در حسابداری می‌توانند به بهبود دقت تقلب در حسابداری، بهبود بی‌طرفی و اثربخشی حسابرسی کمک کنند.

سجادی، فرازمنند، دستگیر و دهقان (۱۳۸۷) در پژوهشی عوامل تأثیرگذار بر گزارش مشروط حسابرسی را مورد بررسی قرار دادند. نتایج این پژوهش حاکی از آن است که نسبت جاری و نسبت حساب‌های دریافتی به دارایی بر احتمال ارائه گزارش مشروط حسابرسی تأثیر دارند. همچنین، رابطه معناداری بین گزارش مشروط حسابرسی در سال گذشته و نوع مؤسسه حسابرسی با گزارش مشروط در سال جاری مشاهده گردیده است.

ستایش و جمالیانپور (۱۳۸۸) به بررسی رابطه بین نسبت‌های مالی و غیر مالی با نوع اظهارنظر حسابرسی پرداختند. نتایج حاصل از این پژوهش نشان می‌دهد که توزیع متغیرهای مستقل شرکت‌ها با اظهارنظرهای متفاوت، در اکثر موارد یکسان نیست. همچنین نتایج بیانگر این است که از میان تمام متغیرهای مالی و غیر مالی، نوع عملکرد بیشترین رابطه را با نوع اظهار نظر حسابرسی دارد.

احمدپور، طاهرآبادی، عباسی (۱۳۸۹) در پژوهشی به بررسی تأثیر متغیرهای مالی و غیر مالی بر صدور اظهارنظر مشروط حسابرسی پرداختند. نتایج این پژوهش نشان می‌دهد که متغیرهای مالی بیشترین تأثیر را بر صدور اظهارنظر مشروط دارد.

پورحیدری و اعظمی (۱۳۸۹) در پژوهشی به بررسی نحوه شناسایی نوع اظهارنظر حسابرسان با بهره‌گیری از شبکه‌های عصبی پرداختند. یافته‌های تحقیق نشان می‌دهد که شبکه عصبی در

تشخیص نوع گزارش حسابرسی عملکرد بهتری از خود نشان داده است، در حالی که رگرسیون لجستیک در تفکیک انواع اظهار نظر حسابرس، الگوی نامتوازی ارائه می‌دهد.

کمالیان، نیک، نفس، افشاری زاده، پوررضا (۱۳۸۹) در پژوهشی به شناسایی عوامل کلیدی تأثیرگذار بر رتبه افشای اطلاعات شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران از منظر داده کاوی پرداختند. بر اساس بررسی‌های صورت گرفته در سایر کشورها و با در نظر گرفتن شرایط حاکم بر بازار سرمایه ایران، ۱۹ عامل مورد بررسی قرار گرفت و از طریق روش‌های داده کاوی، عواملی که بیشترین فراوانی را در مجموعه داده‌ها داشتند، شناسایی شدند.

امینی، محمدی، عباسی (۱۳۹۰) با بهره‌گیری از شبکه‌های عصبی پرسپترون چندلایه، به شناسایی عوامل تأثیرگذار بر صدور گزارش مشروط حسابرسی پرداختند. یافته‌های این پژوهش نشان می‌دهد که نسبت سود پس از کسر مالیات به فروش، به‌عنوان مهم‌ترین عامل در صدور گزارش مشروط حسابرسی شناخته می‌شود.

انواری رستمی، خادمی زارع، علی حیدری بیوکی، نشاط (۱۳۹۰) در پژوهشی برای تعیین محرک هزینه در سیستم‌های هزینه‌یابی بر مبنای فعالیت، از روش داده کاوی و تحلیل عاملی استفاده کردند. نتایج بیانگر آن است که به کارگیری موفق روش ارائه شده، تأییدی بر سودمندی و اثربخشی نتایج به کارگیری روش پیشنهادی برای سیستم بانکی است.

عباسی، قنبری، نجاری، شیپوریان (۱۳۹۱) در پژوهشی با عنوان «ضرورت گسترش استفاده از روش‌های داده کاوی در حسابداری و حسابرسی» مطرح کردند که با توجه به قابلیت‌های پیشرفته فناوری داده کاوی و هوش مصنوعی در پردازش پایگاه‌های داده گسترده و شناسایی الگوهای پیچیده و غیرخطی، پژوهشگران به تدریج از روش‌های آماری سنتی به سمت استفاده از هوش مصنوعی گرایش پیدا کرده‌اند. یافته‌های این تحقیق حاکی از آن است که استفاده از تکنیک‌های داده کاوی به حسابرسان این فرصت را می‌دهد که در زمانی کوتاه و با هزینه کمتر، گزارش‌هایی بر پایه روش‌های علمی تهیه کنند که دارای دقت و اعتماد بالایی باشند. بنابراین، گسترش و به کارگیری این روش‌ها در زمینه‌های مالی امری ضروری و غیرقابل اجتناب به شمار می‌آید.

باقرپور، ساعی، مشکاتی (۱۳۹۱) در پژوهشی با عنوان «پیش‌بینی گزارش حسابرس مستقل در ایران با رویکرد داده کاوی» از سه روش داده کاوی، شامل درخت تصمیم، شبکه‌های عصبی مصنوعی و رگرسیون لجستیک، برای پیش‌بینی گزارش حسابرس مستقل بهره گرفتند. نتایج این تحقیق نشان داد که میانگین دقت مدل ایجاد شده از طریق تکنیک درخت تصمیم در مقایسه با

دو روش دیگر بالاتر بوده و این روش توانسته است گزارش حسابرسی را با دقت میانگین ۸۸٫۶۴٪ پیش‌بینی نماید.

نمازی و مهارلوئی (۱۳۹۷) در پژوهشی با عنوان «پیش‌بینی فرار مالیاتی با بهره‌گیری از الگوریتم‌های داده‌کاوی مبتنی بر درخت تصمیم» به بررسی این موضوع پرداختند. این مطالعه به ارزیابی قابلیت پیش‌بینی فرار مالیاتی شرکت‌های حاضر در بورس اوراق بهادار تهران با استفاده از مدل‌های درخت تصمیم پرداخته است. جامعه آماری تحقیق شامل تمامی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران در دوره زمانی ۱۳۸۴ تا ۱۳۹۴ می‌باشد و نمونه مورد بررسی نیز شامل ۱۰۸۱ سال-شرکت است. برای تحلیل داده‌ها، از روش‌های آماری تحلیل واریانس یک‌طرفه و الگوریتم‌های داده‌کاوی درخت تصمیم استفاده شده است. داده‌ها نیز با استفاده از نرم‌افزارهای SPSS و Weka پردازش و تحلیل گردیده‌اند. نتایج این تحقیق نشان می‌دهد که الگوریتم‌های جنگل تصادفی، کاهش خطای هرس، J48، LMT، ریشه تصمیم و درخت تصادفی به ترتیب از دقت و کارایی بیشتری در پیش‌بینی فرار مالیاتی برخوردار هستند. علاوه بر این، نتایج تحلیل واریانس یک‌طرفه نشان‌دهنده این است که تفاوت‌های موجود در کارایی روش‌های مختلف درخت تصمیم از نظر آماری معنادار می‌باشد.

قرائی (۱۴۰۰) در تحقیقی تحت عنوان «تاثیر داده‌کاوی در ارتقای سطح مدیریت دانش در بانک‌ها» به بررسی پرداخت. با پیشرفت فناوری اطلاعات و اهمیت داده‌ها در عصر طلایی اطلاعات و لزوم تحلیل داده‌ها در بازار کسب و کار، داده‌کاوی در حال تبدیل شدن به یک حوزه استراتژیک مهم برای بسیاری از سازمان‌های تجاری از جمله بانک‌ها شده است. فرآیند تجزیه و تحلیل داده‌ها و استخراج اطلاعات ارزشمند در اثر ارتباط با مشتریان بانکی باعث به وجود آمدن الگویی جهت رفع نواقص خدمت رسانی است. امروزه مشتریان نظرات خود را از طریق سامانه‌های مختلف با واحدهای نظارتی بانک‌ها در میان گذاشته تا به وسیله آن بتوانند ضمن حل مشکلات بوجود آمده بانک‌ها را در ارتقای سطح خدمت رسانی کمک نمایند که این امر منجر به پررنگ شدن اثر داده‌کاوی در سازمان شده است. در این مقاله ضمن بررسی سامانه ارتباطات مردمی در بانک‌ها به تاثیر داده‌کاوی در ارتقای مدیریت دانش سازمانی خواهیم پرداخت.

هاشم زاده و پهلوان (۱۴۰۰) در تحقیقی تحت عنوان «مروری بر سیستم‌های مدیریت ریسک در گمرک با استفاده از تکنیک‌های داده‌کاوی» به بررسی پرداختند. سیاست اقتصادی و

گمرکی خارجی کشورها و اتحادیه‌ها، تحت تاثیر جهانی سازی، فرآیندهای یکپارچگی و چالش‌های مربوط به زمان حال، بصورت پویا در حال تغییر است که از طریق تحول و انطباق مداوم خود منجر به شرایط اقتصادی و تجاری جدیدی می‌شود. همراه با توسعه تکنولوژی‌های کنترل گمرک جهت پشتیبانی از سیاست، طرح‌های طراحی شده برای دور زدن چنین کنترل‌هایی نیز تکامل یافته‌اند. بنابراین، مسئولان گمرک و سایر آژانس‌های مرزی به دنبال مدیریت مخاطرات مربوطه از طریق سیستم‌های مدیریت ریسک پیچیده (RMS) هستند. تکنیک‌های داده‌کاوی همانند دسته‌بندی، رگرسیون و خوشه‌بندی برای پیش‌بینی ریسک و مدیریت آن در گمرک استفاده می‌شود. استفاده از الگوریتم‌های دسته‌بندی با کارایی بالا همانند شبکه‌های عصبی، روش فازی و رگرسیون لجستیک، بهبود الگوریتم‌های دسته‌بندی، پاک‌سازی داده‌ها و همچنین یادگیری تجمعی در بهبود مدیریت ریسک گمرک به کار گرفته می‌شود. به دلیل کاربردی بودن مسائل مربوط به مدیریت ریسک گمرک، در این مقاله یک ارزیابی بر روی مسائل مربوط به مدیریت ریسک گمرک و در این راستا شناسایی برخی از تکنیک‌های داده‌کاوی خواهیم داشت. هدف این مقاله، بررسی و مطالعه کارهای انجام شده در راستای سیستم‌های مدیریت ریسک در گمرک با تکنیک‌های داده‌کاوی می‌باشد. زیودار (۱۴۰۱) در پژوهشی به بررسی کاربرد تکنیک‌های هوش مصنوعی در حوزه مالی و حسابداری پرداخت. در این پژوهش سعی شده است پرکاربردترین و موثرترین تکنیک‌های هوش مصنوعی در حیطه مالی، حسابداری و حسابرسی معرفی و تشریح شوند. برزگری دهج، یعقوب نژاد، کیقبادی و جهان‌شاد (۱۴۰۲) در پژوهشی به بررسی انتخاب برای حسابرسی مالیاتی با استفاده از الگوریتم‌های داده‌کاوی پرداخته‌اند. در این پژوهش، اطلاعات اظهارنامه‌های مالیاتی اشخاص حقوقی از سال ۱۳۹۳ تا ۱۳۹۵ برای ارزیابی ریسک مورد استفاده قرار گرفته‌است. الگوریتم‌های مورد استفاده در این پژوهش، روش‌های دسته‌بندی ماشین بردار پشتیبان، شبکه عصبی، درخت تصمیم و نزدیک‌ترین همسایه بوده است. نتایج پژوهش مؤید آن است که الگوریتم شبکه عصبی به عنوان بهترین الگوریتم برای برآورد ریسک اظهارنامه، معرفی می‌شود.

### یافته‌های پژوهش

نتایج پژوهش‌های بررسی شده به شرح جدول ۱ می‌باشد.

جدول ۱. یافته‌های پژوهش

۲	هوش مصنوعی
۱	رگرسیون لجستیک
۱	الگوریتم‌های داده‌کاوی درخت تصمیم
۱۵	داده‌کاوی

### بحث و نتیجه‌گیری

در عرصه تجارت، اطلاعات مالی به‌عنوان یک دارایی استراتژیک به حساب می‌آید. این اطلاعات از سوی نهادهایی مانند بانک‌ها، بازارهای بورس، سازمان‌های مالیاتی، پایگاه‌های داده مخصوص حسابرسان و حسابداران و دیگر موسسات جمع‌آوری و نگهداری می‌شود. تکنیک‌های داده‌کاوی در داده‌های مالی قادرند برای حل مشکلات مربوط به دسته‌بندی، پیش‌بینی و تسهیل فرآیندهای تصمیم‌گیری مورد استفاده قرار گیرند. برخی از مسائلی که در این زمینه به طبقه‌بندی مالی مرتبط هستند شامل پیش‌بینی ورشکستگی شرکت‌ها، ارزیابی ریسک اعتباری، گزارش تداوم فعالیت، تحلیل در ماندگی مالی و پیش‌بینی عملکرد واحدهای تجاری می‌باشد. توربان (۲۰۰۹) داده‌کاوی را به‌عنوان فرآیندی تعریف کرده‌اند که در آن از تکنیک‌های آماری، ریاضی، هوش مصنوعی و یادگیری ماشین برای استخراج و شناسایی اطلاعات مفید و سپس کسب دانش از مجموعه‌های بزرگ داده استفاده می‌شود. بریسون (۲۰۱۷) نیز تعریف مشابهی ارائه دادند و آن را فرآیند استخراج و کشف الگوهای پنهان از مجموعه‌های وسیع داده دانستند. از این رو، اهمیت داده‌کاوی در زمینه‌های مالی و حسابداری می‌تواند دامنه وسیعی را شامل شود، از شناسایی تقلب گرفته تا بهبود سودآوری در واحدهای تجاری. بنابراین، دسته‌بندی مسائل مالی مانند پیش‌بینی ورشکستگی شرکت‌ها، ارزیابی ریسک اعتباری، گزارش تداوم فعالیت، بحران مالی و پیش‌بینی عملکرد شرکت‌ها اهمیت زیادی دارد؛ چرا که داده‌کاوی فرآیند استخراج دانش‌های پنهان از داده‌های عظیم است و به‌عنوان یکی از ابزارهای کلیدی در تجزیه و تحلیل پیشرفته تجارت و پشتیبانی از تصمیم‌گیری شناخته می‌شود. در سال‌های اخیر، استفاده از روش‌های داده‌کاوی در حسابرسی صورت‌های مالی به‌طور چشمگیری افزایش یافته است. این مقاله به بررسی کاربرد داده‌کاوی در فرآیند حسابرسی پرداخته و تکنیک‌های رایج آن را معرفی کرده است که می‌توانند به بهبود فرآیند حسابرسی کمک کنند. همچنین، روش‌های داده‌کاوی مورد استفاده در تحقیقات مختلف شامل شبکه‌های

عصبی مصنوعی، الگوریتم‌های ژنتیک، درخت تصمیم‌گیری، نظریه مجموعه‌های فازی، استدلال مبتنی بر مورد و برنامه‌نویسی ریاضی هستند. طبق بررسی‌های انجام شده، شبکه‌های عصبی مصنوعی به‌عنوان یکی از پرطرفدارترین روش‌های داده‌کاوی شناخته شده‌اند. امید است با پیشرفت روش‌ها و مدل‌های کاربردی در داده‌کاوی، این تکنیک به یکی از ابزارهای مهم و پرکاربرد در تصمیم‌گیری‌های مالی تبدیل شود.

به منظور ارتقای دانش در زمینه تأثیر بررسی داده‌کاوی بر فرآیند حسابرسی توصیه می‌شود تحقیقات آینده بر توسعه مدل‌های ترکیبی داده‌کاوی و یادگیری ماشین بر تحلیل رفتار حسابرسان تمرکز کنند. بهره‌گیری از داده‌کاوی در حوزه حسابرسی می‌تواند دامنه‌ای گسترده از کشف تقلب تا کیفیت حسابرسی را پوشش دهد. همچنین، برگزاری کارگاه‌های آموزشی برای حسابرسان به منظور آشنایی با تکنیک‌های نوین جمع‌آوری شواهد، مستندسازی، بررسی‌های تحلیلی و در نهایت تصمیم‌گیری، می‌تواند درک عمیق‌تری از این روابط ارائه دهد. علاوه بر این، همکاری میان دانشگاه و حرفه در طراحی ابزارهای تحلیلی مبتنی بر داده‌های واقعی، نقش مهمی در کاربردی‌سازی نتایج پژوهش‌ها خواهد داشت. با این حال، یکی از محدودیت‌های موجود در این زمینه، عدم رغبت حسابرسان برای حرکت به سوی داده‌کاوی است که لزوم تحقیقات بیشتر در این زمینه را برجسته می‌سازد.

### ملاحظات اخلاقی

حامی مالی: مقاله حامی مالی ندارد.

مشارکت نویسندگان: تمام نویسندگان در آماده‌سازی مقاله مشارکت داشته‌اند.

تعارض منافع: بنا بر اظهار نویسندگان در این مقاله هیچ‌گونه تعارض منافی وجود ندارد.

تعهد کپی‌رایت: طبق تعهد نویسندگان حق کپی‌رایت رعایت شده‌است.

## منابع

- امینی، پیمان؛ کامران، محمدی؛ شعیب، عباسی. (۱۳۹۰). بررسی عوامل موثر بر صدور گزارش مشروط حسابرسی: کاربرد روش شبکه عصبی. *حسابداری مدیریت*، ۴(۱۱): ۲۵-۳۹.
- انواری رستمی، علی اصغر؛ خادمی زاده، حسن؛ علی حیدری بیوکی، طاهره؛ نشاط، نجمه. (۱۳۹۰). تعیین محرک هزینه در سیستم های هزینه یابی بر مبنای فعالیت با استفاده از روش های داده کاوی و تحلیل عاملی. *بررسی های حسابداری و حسابرسی*، ۱۸(۶۳): ۲۱-۳۸.
- احمدپور، احمد؛ طاهر آبادی، علی اصغر؛ عباسی، شعیب. (۱۳۸۹). تأثیر متغیرهای مالی و غیر مالی بر صدور اظهار نظر مشروط حسابرسی (شرکت های پذیرفته شده در بورس اوراق بهادار تهران). *نشریه علمی (وزارت علوم)*، ۹.
- باقرپور ولاشانی، محمدعلی؛ ساعی، محمدجواد؛ مشکاتی، علی؛ باقری، مصطفی. (۱۳۹۱). پیش بینی گزارش حسابررس مستقل در ایران: رویکرد داده کاوی. *نشریه علمی (وزارت علوم)*، ۱۹.
- برزگری دهج، محمد؛ یعقوب نژاد، احمد؛ کیقبادی، امیررضا؛ جهانشاد، آرزنا. (۱۴۰۲). انتخاب برای حسابرسی مالیاتی با استفاده از الگوریتم های داده کاوی. *پژوهشنامه مالیات*، ۳۱(۵۹): ۸-۲۵.
- پورحیدری، امید؛ اعظمی، زینب. (۱۳۸۹). شناسایی نوع اظهار نظر حسابرسان با استفاده از شبکه های عصبی. *دانش حسابداری*، ۱(۳): ۷۷-۹۷.
- رضایی، یاسر؛ اصغری شلمانی، مصطفی؛ دلیری دهنه، حسین. (۱۴۰۰). معرفی یک چارچوب مناسب سازماندهی برای کاربردهای داده کاوی در حسابداری و حسابرسی: مروری بر تکنیک های پر کاربرد طبقه بندی داده های مالی. *فصلنامه رویکردهای پژوهشی نوین در مدیریت و حسابداری*، ۵: ۱۵۰۷-۱۵۲۵.
- زیودار، زهره. (۱۴۰۱). کاربرد تکنیک های هوش مصنوعی در حوزه مالی و حسابداری. *رویکردهای پژوهشی نوین مدیریت و حسابداری*، ۶(۲۰): ۱۵۵۷-۱۵۷۲.
- سجادی، سید حسین؛ فرازمنند، حسن؛ دستگیر، محسن؛ دهقان فر، دلشاد. (۱۳۸۶). عوامل موثر بر گزارش مشروط حسابرسی، *مطالعات تجربی حسابداری مالی*، ۵(۱۸): ۱۲۳-۱۴۶.
- ستایش، محمدحسین؛ مظفر، جمالیاپور. (۱۳۸۸). بررسی رابطه نسبت های مالی و متغیرهای غیرمالی با اظهار نظر حسابررس. *تحقیقات حسابداری و حسابرسی*، ۱(۲): ۱۳۰-۱۵۷.

ستایش، محمد حسین؛ ابراهیمی، فهیمه؛ سیف، سید مجتبی؛ ساریخانی، مهدی. (۱۳۹۱). پیش بینی نوع اظهار نظر حسابرسان با استفاده از داده کاوی. *حسابداری مدیریت*، ۵(۱۵): ۶۹-۸۲.

ستایش، محمد حسین؛ فتاحی نافچی، حسن؛ عباس پور، سمیره؛ روستایی، میثم. (۱۳۹۳). ارزیابی رویکرد نوین بر صدور گزارش حسابرسی با استفاده از داده کاوی. *دانش حسابرسی*، ۱۴(۵۷): ۵-۲۶.

قرائی، علی اصغر. (۱۴۰۰). تاثیر داده کاوی در ارتقای سطح مدیریت دانش در بانک‌ها؛ دومین کنفرانس بین‌المللی چالش‌ها و راهکارهای نوین در مهندسی صنایع و مدیریت و حسابداری.

کمالیان، امین رضا؛ نیک‌نفس، علی اکبر؛ افشاری زاده، امید؛ پوررضا، غلامعلی. (۱۳۸۹). تبیین مهمترین عوامل موثر بر رتبه افشای اطلاعات شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با رویکرد داده کاوی، *بورس اوراق بهادار*، ۳(۱۱): ۱۲۵-۱۴۴.

عباسی، ابراهیم؛ قنبری، یوسف؛ نجاری، معصومه؛ شیپوریان، سعید. (۱۳۹۱). لزوم توسعه به کارگیری روش‌های داده کاوی در حسابداری و حسابرسی؛ دهمین همایش ملی حسابداری ایران.

نمازی، محمد؛ صادق‌زاده مهارلوئی، محمد. (۱۳۹۷). پیش‌بینی فرار مالیاتی با استفاده از الگوریتم‌های داده کاوی درخت تصمیم. *فصلنامه حسابداری مالی*، ۹(۶۳): ۷۶-۱۰۰.

هاشم زاده، بدرالدین؛ پهلوان، اسماء. (۱۴۰۰). مروری بر سیستم‌های مدیریت ریسک در گمرک با استفاده از تکنیک‌های داده کاوی؛ چهارمین همایش بین‌المللی مهندسی فناوری اطلاعات، کامپیوتر و مخابرات ایران.

## References

- Abbasi, E; Ghanbari, Y; Najari, M; & Shipourian, S. (2012). The necessity of developing the application of data mining methods in accounting and auditing. *10th National Accounting Conference of Iran*. (In Persian)
- Al Shbail, M. O; Alshurafat, H; Ananzeh, H; & Al-Msiedeem, J. M. (2022). Dataset of Factors affecting online cheating by accounting students: The relevance of social factors and the fraud triangle model factors. *Data in Brief*, 40, 107732.
- Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: A decade review from 2004 to 2015. *Journal of Data Science*, 14(3), 553-569.
- Amani, F.A; Fadlalla, & A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. *International Journal of Accounting Information Systems*, 24, 32-58.

- Amini, P; Kamran, M; & Shoaib, A. (2018). Investigating the factors affecting the issuance of a conditional audit report: the application of the neural network method. *Management Accounting*, 4(11), 25-39. (In Persian)
- Anwari Rostami, A; Khademi Zare, H; Ali Heydari Beyuki, I; & Neshat, N. (2018). Determining cost drivers in activity-based costing systems using data mining and factor analysis methods. *Accounting and Auditing Surveys*, 18(63), 21-38. (In Persian)
- Ahmadpour, A; Tahir Abadi, A; & Abbasi, S. (2009). The effect of financial and non-financial variables on the issuance of conditional audit opinion (companies admitted to the Tehran Stock Exchange). *Scientific publication (Ministry of Science)*, 9. (In Persian)
- Abbasi, A; Ghanbari, Y; Najari, M; & Shipourian, S. (2011). The need to develop the use of data mining methods in accounting and auditing; *The 10th National Accounting Conference of Iran*. (In Persian)
- Bagharpour Velashani, M. A; Sai, M. J; Mashkati, A; & Bagheri, M. (2017). Predicting the independent auditor's report in Iran: data mining approach. *Scientific Journal (Ministry of Science)*, 19. (In Persian)
- Barzegari Dehaj, M; Yaghoub Nejad, A; Keyghabadi, A. R; & Jahanshad, A. (2023). Selection for tax audit using data mining algorithms. *Journal of Taxation Research*, 31(59): 8-25. (In Persian)
- Carta, S. Fenu, G. Recupero, D. R. & Saia, R. (2019). Fraud detection for E-commerce transactions by employing a prudential multiple consensus model, *Journal of Information Security and Applications*, 46, 13–22.
- Chen, W. & DU, Y. (2009). Using neural networks and data mining techniques for the financial dis tress prediction model. *Expert systems with applications*. 36(2), 4075-4086.
- Che, D; Safran, M; & Peng, Z. (2013, April). From big data to big data mining: challenges, issues, and opportunities. In *International conference on database systems for advanced applications* (pp. 1-15). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Deepa, N; Pham, Q. V; Nguyen, D. C; Bhattacharya, S; Prabadevi, B; Gadekallu, T. R; ... & Pathirana, P. N. (2022). A survey on blockchain for big data: Approaches, opportunities, and future directions. *Future Generation Computer Systems*, 131, 209-226.
- Gharaei, A. A. (2021). The impact of data mining on enhancing knowledge management in banks; *The 2nd International Conference on New Challenges and Solutions in Industrial Engineering, Management, and Accounting*. (In Persian)

- Hand, D; Heikki, M; & Padhraic. S, (2001), *Principles of Data Mining*, MIT Press, Cambridge: MA.
- Hashemzadeh, B; & Pahlavan, A. (2021). A review of risk management systems in customs using data mining techniques; *The 4th International Conference on Information Technology, Computer, and Telecommunications Engineering of Iran*. (In Persian)
- Hoberg, G; & Lewis, C. (2017). Do fraudulent firms produce abnormal disclosure? *Journal of Corporate Finance*, 43, 58-85.
- Koskivaara, E. & Barbro B. (2007). Artificial neural network Assistant ANNA for continuous Auditing and monitoring of Financial Data. *journal of Emerging Technologies in Accounting*. 4(1), 29-45.
- Kwame, O., Beverley, R; & Lord, K. (2021). Narrative accounting for mining in Ghana: An old defence against a new threat. *Resources Policy*, 74, 102439.
- Kamalian, A; Nik Nafs, A; Afsharizadeh, A; & Pourreza, G. (2008). Explaining the most important factors influencing the information disclosure rating of companies admitted to the Tehran Stock Exchange with a data mining approach. *Stock Exchange*, 3(11), 125-144.(In Persian)
- Namazi, M; & Sadeghzadeh Maharlouei, M. (2018). Predicting tax evasion using decision tree data mining algorithms. *Financial Accounting Quarterly*, 9(63), 76-100. (In Persian)
- Pourhydari, A; & Azami, Z. (1389). Identifying the type of auditors' comments using neural networks. *Accounting Knowledge*, 1(3), 77-97. (In Persian)
- Praise, M; & Jamalianpour, M. (2008). Examining the relationship between financial ratios and non-financial variables with the auditor's opinion. *Accounting Research*, 1(2), 13-157. (In Persian)
- Rezaei, Y; Asghari Shalmani, M; & Deliri Dehbeneh, H. (2021). Introducing a suitable organizational framework for data mining applications in accounting and auditing: a review of popular classification techniques for financial data. *Quarterly Journal of Modern Research Approaches in Management and Accounting*, 5(83), 1507-1525. (In Persian)
- Sadgali, I; Sael, N; & Benabbou. F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Procedia Comput*, 148, 45-54.
- Sajjadi, H; Farazmand, H; Dastgir, M; & Dehghanfar, D. (2007). Factors affecting the qualified audit report. *Empirical Studies in Financial Accounting*, 5(18), 123-146. (In Persian)
- Setayesh, M. H; & Mozaffar, J. (2009). Investigating the relationship between financial ratios and non-financial variables with auditor's opinion. *Accounting and Auditing Research*, 1(2), 130-157. (In Persian)

- Setayesh, M. H; Ebrahimi, F; Seif, M; & Sarikhani, M. (2012). Predicting the type of auditors' opinion using data mining. *Management Accounting*, 5(15), 69-82. (In Persian)
- Setayesh, M. H; Fattahi Nafchi, H; Abbaspour, S; & Roustaei, M. (2014). Presenting a new approach to issuing audit reports using data mining. *Auditing Science*, 14(57): 5-26. (In Persian)
- Zheng, X; Hamid, M. A. A; & Hou, Y. (2024). Data mining algorithm in the identification of accounting fraud by smart city information technology. *Heliyon*, 10(9).
- Zivodar, Z. (2022). Application of artificial intelligence techniques in finance and accounting. *Modern Research Approaches in Management and Accounting*, 6(20): 1557-1572. (In Persian)

#### COPYRIGHTS



This is an open access article under the CC BY-NC-ND 4.0 license.

پژوهشگاه علوم انسانی و مطالعات فرهنگی  
پرتال جامع علوم انسانی