

کاربرد تکنیک های هوش مصنوعی در حوزه مالی و حسابداری

زهره زیودار^۱

تاریخ دریافت: ۱۴۰۱/۰۱/۱۸ تاریخ چاپ: ۱۴۰۱/۰۳/۰۵

چکیده

وجود اطلاعات مالی شفاف و قابل مقایسه رکن اصلی پاسخگویی و تصمیم گیری های اقتصادی آگاهانه است. نقش اصلی گزارشگری مالی، انتقال اثربخش اطلاعات به افراد برون سازمانی به روشی معتبر و به موقع است. مطالعات بسیاری درباره کیفیت گزارشگری مالی در دنیا صورت گرفته است که عمدتاً با تکیه بر گزارش های ساختار یافته و محاسبات کمی و عددی است. این در حالی است که امروزه بخش زیادی از اطلاعات مالی از طریق افشاهای کیفی مانند گزارش های روایتی و اسناد متنی منتشر می شوند. جستجوی ابزارهایی که بتواند با بررسی منابع مالی و تحلیل آن ها از میان داده های خام در گزارش های ناساختاریافته و کیفی، دانش مفید را استخراج نماید، منجر به شکل گیری شاخه ای در هوش مصنوعی گردید که به داده کاوی معروف است. استفاده از داده کاوی در مورد متن خود شاخه ای دیگر را در علوم هوش مصنوعی به نام متن کاوی ایجاد کرد. متن کاوی نوع خاصی از داده کاوی است که بیشتر به دنبال استخراج اطلاعات مفید از داده های متنی ساختار نیافته و نیمه ساختاریافته (مانند اسناد متنی) از طریق تشخیص و نمایش الگوها است. پژوهش های بسیاری نشان داده اند که امروزه جهت سنجش و اندازه گیری و پیش بینی متغیرهای مالی و حسابداری باید از گزارش های کمی فاصله گرفته و بر گزارش های کیفی تکیه بیشتری شود و به جای تکیه صرف بر «اعداد و ارقام» بر تجزیه و تحلیل «متن گزارش ها» نیز اتکا کرد؛ همین امر باعث شده است که پژوهشگران حسابداری و مالی در سال های اخیر تکنیک های هوش مصنوعی را در حوزه های مختلف دانش حسابداری و حسابرسی خصوصاً برای درک رابطه بین متن در افشاهای شرکتی و عناصر بنیادی شرکت بیشتر به کار گیرند. در این مقاله سعی شده است کاربردترین و موثرترین تکنیک های هوش مصنوعی در حیطه مالی، حسابداری و حسابرسی معرفی و تشریح شوند.

واژگان کلیدی

هوش مصنوعی، داده کاوی، متن کاوی، افشاهای روایتی (متنی)

^۱ استادیار، گروه حسابداری، واحد دولت آباد، دانشگاه آزاد اسلامی، اصفهان، ایران (zzivdar55@yahoo.com)

۱. مقدمه

کیفیت گزارشگری مالی یکی از مهمترین مباحثی است که در سال های اخیر پژوهش های بسیاری را در حوزه حسابداری و مدیریت مالی به خود اختصاص داده است. هدف اصلی گزارشگری مالی تامین اطلاعات مورد نیاز برای تصمیم گیری، فراهم ساختن اطلاعات لازم برای ارزیابی عملکرد و توان سود آوری بنگاه اقتصادی است. برای دستیابی به اهداف مذکور گزارشگری مالی باید اطلاعات حسابداری با کیفیت در اختیار استفاده کنندگان قرار دهد (راجگوپال و وینکتچالم، ۲۰۱۱). از سوی دیگر برای سنجش کیفیت گزارشگری مالی مبحث کیفیت افشاء در بسیاری از پژوهش ها مورد توجه اصلی قرار گرفته است. افشاء در لغت به معنای آشکار نمودن اطلاعات است. براساس فرهنگ کehler افشاء عبارت است از نمایش واضح یک واقعیت با یک وضعیت در ترازنامه و یا صورت های مالی دیگر و یادداشت های همراه آن و یا در گزارش حسابرسی (کehler و ارکلویز، ۱۹۷۵). در سطح معنایی، کلمه افشاء به معنی انتشار اطلاعات مالی واحد تجاری، در گزارش های سالانه است. دیدگاه افشای اطلاعات شامل بحث های مدیریت، تجزیه و تحلیل، یادداشت های پیوست صورت های مالی، گزارش هیات مدیره به مجمع و صورت های مالی مکمل می شود (هندریکسن و ون بردا، ۱۹۹۲). انجمن حسابداری آمریکا افشاء را به صورت «جریان اطلاعات از قلمرو خاص به قلمرو عام» تعریف کرده است. افشاء در حسابداری واژه ای فراگیر بوده و تقریباً تمامی فرآیند گزارشگری مالی را در بر می گیرد و به عنوان یکی از اصول اساسی تحت عنوان اصل افشاء شناخته می شود. این اصل ایجاب می کند که صورت های مالی به گونه ای تهیه و ارائه شوند که اهداف گزارشگری مالی شامل ارائه اطلاعات کافی و قابلیت فهم محقق گردد. افشای اطلاعات مربوط در صورت های مالی باید به گونه ای باشد که از یک طرف، امکان اتخاذ تصمیمات آگاهانه را برای استفاده کنندگان فراهم و از طرف دیگر باعث گمراهی آنان نشود. از نظر مخاطبین، افشای اطلاعات مالی برای دریافت کنندگان ثانویه گزارش های مالی نظیر کارکنان، مشتریان عمده، نهادهای دولتی و جامعه نیز مد نظر قرار گرفته است، چنانچه سرمایه گذاران کانون توجه حسابداری باشند، می توان افشاء را ارائه اطلاعات لازم برای عملکرد بهینه بازار کار آمد سرمایه تعریف کرد (هندریکسن، ۱۹۹۲).

حسابداران در ارائه گزارش های مالی هم از داده های کمی و هم از داده های کیفی برای افشای اطلاعات مالی استفاده می کنند. افشای کمی به این معنی است که از داده های عددی برای افشای اطلاعات مالی استفاده می شود؛ افشای کیفی بیشتر از گزارش های متنی برای ارائه اطلاعات به استفاده کنندگان بهره می جوید. در استانداردها و الزامات گزارشگری مالی، هر دوافشای کمی و کیفی مورد توجه قرار گرفته است (بیانیه گزارشگری مالی ۴۸؛ استاندارد بین المللی مالی شماره ۷). اطلاعات کمی گزارش های مالی شرکت ها از اهمیت زیادی برخوردار هستند اما این اطلاعات به تنهایی نمی توانند تصویر کاملی از وضعیت شرکت برای استفاده کنندگان فراهم سازند. جهت استفاده از اطلاعات کمی گزارش های مالی، سرمایه گذاران می بایست ابتدا اقدام به کدگشایی از آن ها کرده، سپس آن ها را پردازش کنند، به همین دلیل تهیه و ارائه اطلاعات نوشتاری در کنار اطلاعات کمی گزارش های مالی، جهت درک بهتر اطلاعات کمی و ارائه اطلاعاتی که در قالب اعداد نمی گنجند از دیرباز مورد توجه تهیه کنندگان، استفاده کنندگان و قانون گذاران اطلاعات حسابداری بوده است، به نحوی که امروزه اطلاعات کیفی نسبت به اطلاعات کمی، حجم بیشتری از گزارش های مالی شرکت ها را به خود اختصاص داده است (هاجک و همکاران، ۲۰۱۴). اصولاً افشاهای کمی دربرگیرنده داده ها و گزارش های ساختاریافته و افشاهای کیفی بیشتر شامل گزارش های ساختار نیافته می باشند. اکثر پژوهش های حسابداری و مالی طی پنجاه سال گذشته بر افشاهای کمی برای سرمایه گذاری، مشارکت و تصمیمات تجاری متمرکز شده است (سیانو و

وایسوکي، ۲۰۱۹). ازسوی دیگر دانشگاهیان و دست اندرکاران معتقدند که افشاهای کمی به تنهایی برای تصمیم گیری های اقتصادی مؤثر نیستند (لو و گو، ۲۰۱۶)؛ لذا اخیراً ادبیات مالی شروع به درک بهتر و توصیف افشاهای کیفی و متنی در حسابداری و اسناد مالی و ارتباطات کرده است. پژوهشگران با استفاده از ابزارهای تحلیل متنی، پیشرفت های چشمگیری در مستند سازی ارتباط بین ویژگی های متنی اسناد حسابداری و مالی و نتایج اقتصادی مختلف دارند. با این حال، تکنیک های تجزیه و تحلیل متنی موجود، اطلاعات کمی را از متن حسابداری و افشای مالی حذف یا نادیده می گیرند. اعداد نقش مهمی در افشاهای حسابداری و مالی دارند و با این که پژوهش هایی نشان می دهد که بین ارقام گزارش های مالی و قابلیت فهم گزارش های کیفی، ارتباط مثبتی وجود دارد با این حال دیده می شود که امروزه پژوهش های تحلیل متنی، سعی دارند نقش ارقام و اعداد را در این افشاها نادیده بگیرند. به همین دلیل ایجاد توازن در به کارگیری هر دو دسته از افشاهای کیفی و کمی در پژوهش های آتی در حوزه مالی و حسابداری اهمیت خواهد داشت (سیانو و وایسوکي، ۲۰۱۹).

افشای ساختاریافته شامل داده های ساختاریافته ای هستند که به طور کاملاً مستقل از همدیگر ولی یکسان از لحاظ ساختاری در یک محل گردآوری شده اند. انواع بانک های اطلاعاتی و صورت های مالی اساسی نمونه هایی از این دسته اطلاعات هستند (داس، ۲۰۱۴). بخش قابل توجهی از دانش بشر، کاملاً ساختار نیافته است؛ و افشای ساختار نیافته حاوی داده های بدون ساختاری هستند که یا دارای یک مدل داده ای از پیش تعیین شده نیستند و یا به صورت پیش فرض سازماندهی نشده اند. داده های ساختار نیافته معمولاً متن سنگینی دارند و اغلب شامل محتوای چندرسانه ای می باشند (فلدمن و سنگر، ۲۰۰۷). اسناد متنی ساختار نیافته، نقش مهمی در افشای اطلاعات مالی شرکت ها دارند، زیرا مقدار زیادی از اطلاعات و استراتژی عملیاتی شرکت ها بین خطوط این اسناد پنهان شده است (یانگ و همکاران، ۲۰۱۷). گزارش های تحلیلی و دوره ای، اعلامیه های سود، کنفرانس های خبری، گزارش های تفسیری مدیریت، گزارش هیات مدیره به مجمع، کتابخانه های دیجیتال، اخبار، کتاب های الکترونیکی، بسیاری از مدارک مالی، مقالات علمی و تقریباً هر چیزی قابل رویت در داخل وب، همگی نمونه هایی از افشاهای ساختار نیافته هستند (نانگ، ۲۰۰۳). ۹۰٪ از اطلاعات جهانی ایجاد شده در دهه گذشته از نوع داده های ساختار نیافته (متن آزاد) بوده است که ۸۰٪ محتوای موجود در شبکه های رایانه ای را به خود اختصاص می دهد. ارزش کامل این اطلاعات قابل بهره برداری توسط دولت ها، بخش عمومی و خصوصی، تجارودانشگاهیان نخواهد بود مگر این که افراد متن گزارش های مالی را بخوانند و یا ابزاری جایگزین برای استخراج اطلاعات ارزشمند از این متون پیدا کنند. رشد چشمگیر داده های ساختار نیافته در بازارهای مالی، کاملاً مشهود است. به عنوان مثال دایرو همکاران (۲۰۱۷)، درمطول شدن گزارش های متنی شرکت های آمریکایی، رشد ۱۱۳ درصدی، مشاهده کردند. این در حالی است که الحاج و همکاران (۲۰۱۹)، گزارش مشابهی از رشد گزارش های تفسیری مدیریت در شرکت های انگلیسی ومطول شدن آن ها بین سال های ۱۹۹۶ تا ۲۰۱۳ ارائه نمودند. پژوهشگران حسابداری و مالی به منظور پردازش و تجزیه و تحلیل افشاهای کیفی و متنی، به زبان شناسی محاسباتی و روش های نوین متن کاوی روی آورده اند (الحاج و همکاران، ۲۰۱۹).

الیوت (۱۹۹۸) پیش بینی کرد که اهمیت افشاهای ساختار نیافته برای تصمیم گیری مالی تا پایان قرن بیستم به مراتب از همتای ساختار یافته خود فراتر می رود. مطابق با این پیش بینی در دو دهه گذشته، مشارکت کنندگان در بازارهای مالی شاهد گسترش سریع حجم داده های ساختار نیافته بوده اند. مثال هایی از این دست عبارتند از: اطلاعاتی های مطبوعاتی شرکت ها؛ محتوای صفحه وب؛ گزارش تحلیلگران؛ گزارش های تحلیلی موجود در رسانه های اجتماعی و مطبوعاتی

ارسال شده توسط انواع ذینفعان شرکت از جمله مدیران درون سازمانی شرکت ها، سرمایه گذاران، مشتریان، تأمین کنندگان، گروه های فشار و سیاستمداران. این گسترش در حجم داده های ساختار نیافته فرصت های مناسبی را برای مشارکت کنندگان در بازارهای مالی و پژوهشگران دانشگاهی ایجاد می کند. تحول در گزارشگری بین المللی شرکت ها خود عامل مهمی در رشد گزارش های ساختار نیافته مالی بود؛ اما از سوی دیگر رشد و گسترش گزارش های ساختار نیافته برای آن دسته از فعالان بازارهای مالی که نگران شیوع اطلاعات نامربوط هستند، به عنوان یک معضل و نگرانی همواره وجود داشته است (لویز و یانگ، ۲۰۱۹).

صورت های مالی شرکت ها دربرگیرنده اطلاعات تاریخی است و حاوی اطلاعات غیرمالی و چشم اندازهای آتی کمتری است. اگر چه گزارش های مالی به درک کاربران از چشم اندازهای آتی شرکت تا حدودی کمک می کنند اما استدلال حاکم معتقد است که برای درک بهتر و کامل تر نیاز به اطلاعات بیشتری حس می گردد. به علاوه این گونه استدلال شده است که به کارگیری استانداردهای گزارشگری بین المللی برای صورت های مالی موجب می شود بسیاری از استفاده کنندگان از این گزارش ها در خواندن و فهم آن ها دچار مشکل شوند. لذا در تلاش برای ایجاد صورت های مالی قابل فهم تر، هیات تدوین استانداردهای حسابداری ممکن است قدم درراهی بگذارد که منجر به مبهم ترشدن گزارش ها گردد؛ بنابراین می توان نتیجه گرفت که شرکت ها نه تنها باید اطلاعاتی راجع به عملکرد خود در صورت های مالی ارائه نمایند بلکه باید برای درک بهتر استفاده کنندگان، اطلاعات غیرمالی و چشم اندازهای آتی را در گزارش های خود اطلاع رسانی کنند (گوش، ۲۰۲۰).

نیاز به شفافیت در تمامی گزارش های مالی توسط سهامداران، مالکان و استفاده کنندگان از این گزارش ها احساس می شود و الزام برای شفافیت گزارش ها عامل کلیدی در استقرار یک حاکمیت شرکتی موفق است. بسیاری از شرکت ها برای بهبود کیفیت روابط خود با سهامداران الزاماتی برای ارائه گزارش های اختیاری بیشتر تعیین نمودند؛ به همین دلیل افشاهای روایتی و یا به عبارتی گزارشگری روایتی به بخش مهمی از گزارشگری شرکت ها تبدیل شد. این گزارش ها شامل گزارش بررسی عملیاتی و مالی، گزارشگری اجتماعی و زیست محیطی و گزارشگری منابع انسانی هستند. اطلاعات متنی موجود در گزارش های روایتی ماهیتی ساختار نیافته دارد. در سال های اخیر، به ویژه از زمان ظهور بحران مالی جهانی، این بحث افزایش یافته است که آیا گزارش های سالانه به سرمایه گذاران و سایر استفاده کنندگان در اتخاذ تصمیمات درست در راستای چشم انداز شرکت ها کمک می کند. برای پرکردن شکاف اطلاعاتی بین الزامات استانداردهای گزارشگری و نیازهای اطلاعاتی استفاده کنندگان، گزارش های روایتی به وجود آمد. هدف گزارش روایت تهیه اطلاعاتی برای پاسخگویی به نیازهای سهامداران است. گزارش مالی اطلاعات مفیدی در باره وضیت مالی، موقعیت و عملکرد واحد تأمین می کند اما به ندرت همه اطلاعات مورد نیاز برای ارزیابی دقیق یک شرکت را در اختیار سهامداران می گذارد؛ اما رئوس کلی گزارش های روایتی به گونه ای است که چشم اندازی برای بررسی عملیات و توان آتی موسسه را در اختیار استفاده کنندگان قرار می دهد (گوش، ۲۰۲۰). پژوهش های حسابداری تصدیق می کند که ساختار برای درک گزارش های مالی مهم است. مدیران می توانند از ساختار افشاهای روایتی در راستای افزایش درک استفاده کنندگان از ابعاد عملکرد و برنامه های آتی شرکت بهره گیرند (آلی ودی آنجلیس، ۲۰۱۵).

با این که افشاهای روایتی مدتی است که علاقه پژوهشگران را در حوزه مالی و حسابداری به خود جلب نموده است، اما نیاز به جمع آوری و امتیاز دهی دستی به محتوای اطلاعات متنی در افشاهای روایتی باعث شده که کار در این حوزه

همچنان با محدودیت هایی روبه رو باشد. بسیاری از پژوهشگران کماکان به دنبال تقویت تحلیل های دستی به جای روش های نوین محاسباتی هستند (مانند شوماخر ۱۹۹۴، دیویس و برنان ۲۰۰۷، بیتی ۲۰۱۴). از سوی دیگر اتخاذ تکنیک های خودکار نمره دهی توسط پژوهشگران مطرح حسابداری و مالی در گزارش های روایتی باعث شده است که پژوهشگران در سال های اخیر برای تحلیل گزارش های روایتی و متنی به جای استفاده از تحلیل های دستی به محاسبات خودکار و رایانه ای روی آورند که این امر خود موجب به کارگیری بیشتر افشاهای روایتی در تجزیه و تحلیل های مالی شده است (الحاج و همکاران، ۲۰۱۹). به دلیل خصوصیات افشاهای کیفی و روایتی که در بخش مقدمه مقاله به تفصیل مطرح شد؛ نیاز به تکنیک های نوین جهت تحلیل این گزارش ها در حیطه های مالی و حسابداری و حسابرسی به وضوح حس گردیده است؛ لذا هوش مصنوعی و تکنیک های آن به این حوزه ورود نمود که در بخش بعدی مقاله به آن پرداخته خواهد شد. ۲.

۲. داده کاوی^۱، متن کاوی^۲ و جایگاه آن در حوزه حسابداری و مالی

امروزه به دلیل گسترش دانش و پیچیده تر شدن فرآیند تصمیم گیری، استفاده از سیستم های هوش مصنوعی، اهمیت بیشتری پیدا کرده است. هوش مصنوعی رابطه تنگاتنگی با علوم مختلف دارد که از مهم ترین آن ها می توان به منطق ریاضی و مهندسی نام برد. هوش تجاری که مقوله ای از هوش مصنوعی است با جست و جو، استفاده از فناوری ها، ابداع، آزمون و تصمیم گیری های مفید در کسب و کار و همچنین بازار ها و اقتصاد جهانی به پیشرفت های زیادی رسیده است. هوش تجاری مانند چتری دربرگیرنده طراحی ها، پایگاه های داده ای، نرم افزارها و روش شناختی هایی با هدف تحلیل داده ها جهت حمایت از تصمیمات مدیران تجاری و سرمایه گذاران می باشد (توربان و همکاران، ۲۰۱۱).

باپیدایش روزافزون انواع کسب و کارهای نوین که منجر به افزایش پیچیدگی محیط تجاری گردیده است، مدیران سرمایه گذاران و بسیاری از فعالان بازار سرمایه همواره در جستجوی ابزارهایی به منظور پیش بینی دقیق تر عملکرد مالی شرکت ها و حداکثر نمودن سود حاصل از سرمایه گذاری می باشند. داده کاوی فرآیند بررسی و تجزیه و تحلیل مقادیر عظیمی از داده ها جهت استخراج اطلاعات، دانش، کشف الگوهای پنهان، یادگیری الگوها و قوانین معنی دار نهفته در انباره داده است و در حال حاضر یکی از مهم ترین الگوهای موجود در حوزه هوش تجاری و ابزارهای تصمیم گیری است (هان و کامبر، ۲۰۱۲). ویژگی های داده کاوی سبب شده که امروزه در کسب و کار و تجارت الکترونیک از آن استفاده های بسیاری گردد (چاولا و همکاران، ۲۰۰۴). عملیات داده کاوی به یک نوع از داده های محدود نمی شود و معمولاً داده های مختلفی توسط این سیستم ها پذیرفته می شود، لذا تکنیک های گوناگونی برای انواع مختلفی از داده ها مناسب هستند و یافتن یک راه کلی تلاشی بیهوده به نظر می رسد. تکنیک های داده کاوی را می توان بر روی داده های ساختار نیافته (مانند متون) نیمه ساختاریافته (مانند اسناد) و ساختاریافته (مانند جداول و گزارش های عددی) اعمال نمود. عمدتاً داده کاوی از میان داده های ساخت یافته به کسب دانش و استخراج اطلاعات می پردازد.

یکی از منابع داده های انبوهی که توسط انسان تولید می شود «متن است» بر همین اساس یکی از حوزه های داده کاوی، تجزیه و تحلیل متون تهیه شده توسط انسان است که آن را «متن کاوی» می نامند. متن کاوی در واقع یک حوزه ی بین رشته ای است که از مفاهیم داده کاوی، آموزش ماشین، آمار و زبان برای تجزیه و تحلیل متون تهیه شده توسط انسان مانند مقالات، گزارش ها، کتاب ها، وب سایت ها و مواردی از این قبیل استفاده می کند. هدف نهایی از متن کاوی،

1- Data Mining

2- Text mining

استخراج اطلاعات باکیفیت از داخل متن است (رهروی و همکاران، ۱۳۹۷). متن کاوی مشابه داده کاوی است، ابزارهای داده کاوی عمدتاً طراحی شده اند تا داده های ساختاریافته از پایگاه داده را به کار گیرند. می توان گفت، متن کاوی راه حل بهتری برای شرکت ها است. تفاوت متن کاوی و داده کاوی این است که داده کاوی بر روی داده های ساختاریافته پایگاه داده و متن کاوی، بر روی داده های ساختار نیافته و نیمه ساختار یافته مانند رایانامه و مستندات تمام متنی کار می کند. در متن کاوی سعی می گردد از همان تکنیک های داده کاوی استفاده گردد تا به صورت خودکار آمارهایی را جمع آوری نموده و ساختار و معنای مناسبی از متن استخراج گردد. در این موارد، دیدگاه عمومی، استخراج ویژگی های کلیدی از متن است. ویژگیهای استخراج شده به عنوان داده برای تحلیل استفاده می گردد (اسماعیلی، ۱۳۹۱). شناسایی هرزنانه ها، شناخت روابط میان مفاهیم، جستجو و بازیابی، نظارت بر رفتار انسان ها به صورت پنهان، طبقه بندی داده ها، خلاصه سازی متن، از مهمترین کاربردهای متن کاوی است (هان و همکاران، ۲۰۱۲).

۱-۲. طبقه بندی پژوهش های مبتنی بر داده کاوی در حوزه مالی و حسابداری

نسبت به روش های کمی که به صورت سنتی در حسابداری و مالی استفاده می شوند، تجزیه و تحلیل های متنی، اصولاً از دقت کمتری نسبت به تحلیل های کمی برخوردارند، اما شناخت این روش به همان اندازه روش های کمی حایز اهمیت است زیرا در گزارش معمولی یک شرکت، گزارش های روایتی متنی به طور متوسط هشتاد درصد گزارش سالانه را در برمی گیرد و بقیه گزارش حاوی اعداد و داده های کمی است. چنین حجم بالایی از افشای اجباری، فهم و تفسیر اطلاعات موجود در این نوع از گزارش ها را ضروری می سازد (لو و همکاران، ۲۰۱۷). اخیراً با افزایش قدرت محاسبه و اندازه گیری در طی نیم قرن گذشته و تمرکز جدی تر بر روش های متنی که از زیرمجموعه های هوش مصنوعی و مهندسی اینترنت می باشد، کاربرد این روش در پژوهش های علمی بیشتر گردیده و در شاخه های مختلفی از علوم و دانش به کار گرفته می شود. در حسابداری و مالی، دسترسی آنلاین به مقالات به روز و جدید، کنفرانس های خبری، اسناد و مدارک منتشر شده از سوی شرکت ها و نهادهای قانون گذار و متون موجود در شبکه های اجتماعی، خوراک زیادی را برای استفاده از این فن آوری فراهم می کند (لوگران و مکدونالد، ۲۰۱۶).

اولین بار در حوزه حسابداری اینگرام و فریزر^۳ (۱۹۸۰) مقاله ای را برای معرفی تحلیل محتوا برای ارزیابی کمی داده های روایتی بر اساس تکرار کلمات منحصر به فرد ارائه کردند. بعد از پژوهش فوق، مقالات حسابداری و مالی برای کمی کردن اطلاعات متنی، فعال شدند. از دیگر مقالات و محققان پیشگام در زمینه تحلیل متنی فریزر و همکاران (۱۹۸۴)، آنتویلر و فرانک^۶ (۲۰۰۴)، داس و چن (۲۰۰۷)، تتلاک^۷ (۲۰۰۷)، لی^۸ (۲۰۰۸)، لی (۲۰۱۰)، کرنی و لیو^۹ (۲۰۱۴)، داس (۲۰۱۴) اشاره نمود. بیشتر این محققان در حیطه حسابداری و حسابرسی، عمدتاً تاثیر اطلاعات کیفی را بر ارزشگذاری حقوق صاحبان سهام بررسی کرده اند. پژوهش های این محققان نشان داده است که واژه های انتخاب شده توسط مدیران برای توصیف عملکردشان و زبان به کار گرفته شده توسط رسانه برای گزارش عملکرد شرکت ها و بازارها، ارتباط موثری با بازده آتی سهام، سود و حتی فعالیت های متقلبان مدیران در آینده داشته است. مشخصاً سرمایه گذاران بازار سهام، برای

3- Lo & et al

4- Loughran & Mcdonald

5- Ingram & Frazier

6- Antweiler & Frank

7- Tetlock

8- Li

9- Kearney&Liu

ارزشیابی های خود بیشتر از داده های کمی استفاده می کنند اما هم چنان که نظام حسابداری و مالی این روش و فن آوری جدید را می پذیرد، باید محققان و پژوهشگران مطمئن باشند که آیا این روش معیار مناسبی برای اندازه گیری حقیقت هست یا خیر (لوگران و مکدونالد، ۲۰۱۶). مقالات بسیاری در حوزه حسابداری از روش متن کاوی برای تحلیل های متن خود استفاده کرده اند (لی ۲۰۱۰؛ لئو و کرنی ۲۰۱۴؛ داس ۲۰۱۳؛ لوگران و مکدونالد ۲۰۱۶؛ دایر و همکاران ۲۰۱۷؛ الحاج و همکاران ۲۰۱۹). بیشترین توجه امروزه به این موضوع معطوف شده است که آیا گزارش های روایتی همچون گزارش تفسیری مدیریت و یا گزارش هیات مدیره به مجمع، شامل اطلاعات افزایشی فراتر از اطلاعات مندرج در صورت های مالی و یادداشت های همراه است (لوئیس و یانگ^{۱۰}، ۲۰۱۹). لی (۲۰۰۸) یکی از پیشگامان پژوهش در حیطه متن کاوی در حسابداری، ناخوانایی در گزارش های کیفی متن را عاملی بر عدم پایداری سود در شرکت ها معرفی کرد. لی (۲۰۰۸) یک رابطه واضح در خصوصیات زبان شناسی درافشاهای روایتی مشخص نمود و نشان داد که عملکرد شرکت رابطه مثبتی با درجه خوانایی گزارش های متن آن دارد. وی در ادامه پژوهش های خود تاثیر به کارگیری واژگان و لحن مثبت و منفی آن ها را در پیش بینی های عملکرد آتی بررسی کرد.

از آن جایی که تحلیل متن هنوز یکی از حوزه های نوظهور در حسابداری و مالی است، الگوریتم طبقه بندی آن هنوز به خوبی بسط و توسعه نیافته است و دیدگاه های مختلفی در این حوزه وجود دارد. دیدگاه اول متعلق به امانی و فادللا^{۱۱} (۲۰۱۷) است که به منظور ارائه چارچوبی جامع در زمینه کاربرد داده کاوی در حسابداری، زیرساخت های «حسابداری» و «داده کاوی» را به طور همزمان بررسی کردند. آن ها با توجه به وظایف اصلی داده کاوی که شامل توصیف^{۱۲}، پیش بینی^{۱۳} و طبقه بندی^{۱۴} می باشد به معرفی چارچوبی پرداختند که در آن، دو دیدگاه شناخته شده برای گزارشگری در «حسابداری» شامل دیدگاه گذشته نگر^{۱۵} و دیدگاه آینده نگر^{۱۶} را با سه هدف شناخته شده برای «داده کاوی» شامل توصیف، پیش بینی و طبقه بندی، ترکیب کردند و شش حالت ترکیبی که فقط چهار حالت آن به طور منطقی امکان پذیر است شناسایی نمودند: ۱) داده کاوی توصیفی در گزارشگری گذشته نگر، ۲) داده کاوی طبقه بندی در گزارشگری گذشته نگر، ۳) داده کاوی طبقه بندی در گزارشگری آینده نگر، ۴) داده کاوی پیش بینی کننده در گزارشگری آینده نگر (رهروی و همکاران، ۱۳۹۷).

دیدگاه دوم متعلق به فرانکل و همکاران^{۱۷} (۲۰۱۶) است. آن ها پژوهش هایی را که با استفاده از اطلاعات کیفی و تکنیک متن کاوی به تحلیل متون مالی و حسابداری می پردازند به سه گروه تقسیم کردند: گروه اول مطالعات، از فرهنگ واژگان برای ارزیابی خصوصیات افشاء، مانند عدم اطمینان، لحن^{۱۸} و رقابت، استفاده کرده اند: (آلی و دی آنجلیس^{۱۹} ۲۰۱۵؛ موسلو و همکاران^{۲۰} ۲۰۱۴؛ لی و همکاران^{۲۱} ۲۰۱۳؛ کراوت و موسلو^{۲۱} ۲۰۱۳؛ روجرز و همکاران^{۲۲} ۲۰۱۱). گروه دوم پژوهش ها به

10- Lewis & Young

11- Amani&Fadlalla

12- Description

13- Prediction

14- Classification

15- Retrospection

16- Prospection

17- Frankel et al

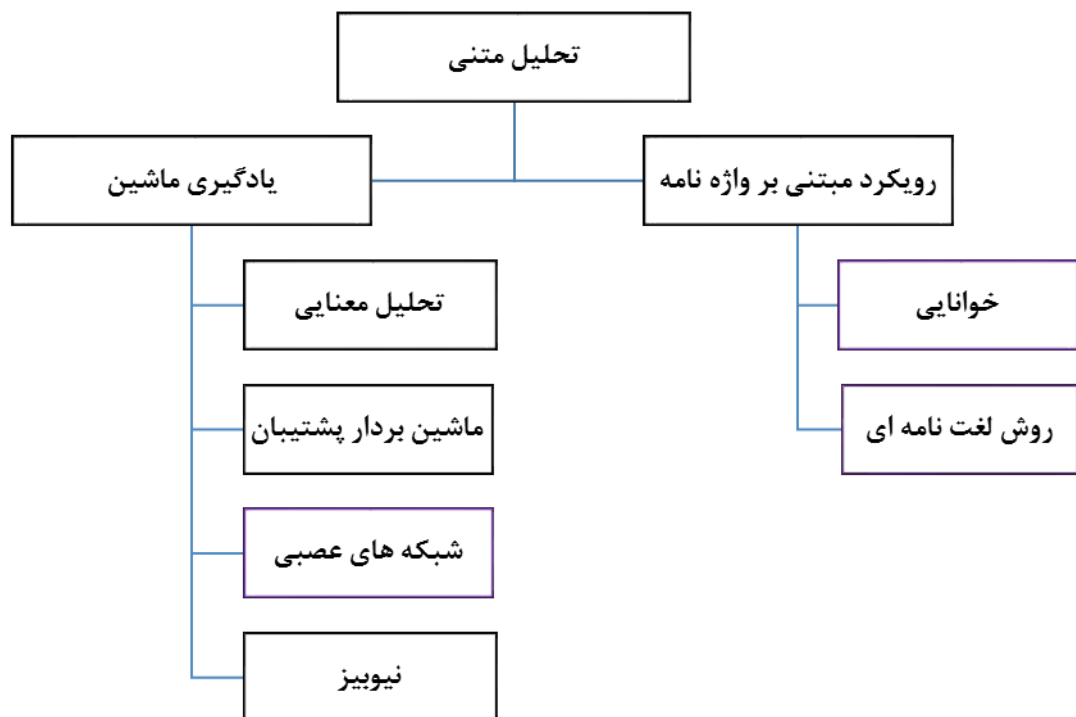
18- Tone

19- Allee&DeAngelis

20- Muslu et al

21- Kravet&Muslu

بررسی و سنجش ارتباط بین خوانایی^{۲۳} (قابلیت فهم)، میزان شباهت افشاء و طول متن افشاء شده با عناصر بنیادی شرکت پرداخته اند (لوگران و مکد و نالد ۲۰۱۶؛ بنسال و همکاران^{۲۴} ۲۰۱۵؛ پترسون و همکاران ۲۰۱۵؛ لی و همکاران ۲۰۱۴ و براون و تاکر ۲۰۱۱). گروه سوم از محققان معتقد هستند که پژوهشگران می توانند برای کاهش شدت کار در حیطه های کاری و وظیفه ای خود، از تکنیک های پردازش زبان طبیعی که از سایر حوزه های دیگر وارد شده است استفاده کنند (بائووداتا ۲۰۱۴؛ لانگ و لارنس ۲۰۱۵؛ کمبل و همکاران ۲۰۱۴؛ کراوت و موسلو ۲۰۱۳؛ لی ۲۰۱۰). عمومی ترین روش شناختی هایی که به وسیله محققان حسابداری و مالی در تحلیل متنی به کار می رود در شکل (۱) نشان داده شده است (لی گو و همکاران ۲۰۱۶). این طبقه بندی براساس تکنیک های کاربردی در داده کاوی، به ویژه متن کاوی شکل گرفته است. پژوهش ها و مطالعات بسیاری از طریق تحلیل متنی و تکنیک های ذکر شده آن در شکل (۱) به بررسی گزارش های روایتی (متنی) پرداخته اند. دو تکنیک ماشین بردار پشتیبان و روش لغت نامه ای از پرکاربردترین و موثرترین روشهای متن کاوی در حوزه مالی و حسابداری هستند که در بخش های بعدی به آن ها به تفصیل اشاره خواهد شد.



شکل (۱): روش های کلی برای تحلیل متنی (لی گو و همکاران ۲۰۱۶، ص ۱۵۵).

۲-۲. تکنیک ماشین های بردار پشتیبان

الگوریتم ماشین های بردار پشتیبان اولیه که در ۱۹۶۳ توسط ولادیمیر وپنیک ابداع شد و در سال ۱۹۹۵ توسط وپنیک و کورتس برای حالت غیر خطی تعمیم داده شد، یک ابر صفحه یا مجموعه ای از ابر صفحه ها در فضاهای با ابعاد محدود یا چندبعدی ایجاد می کند. این ابر صفحه ها می توانند به منظور طبقه بندی و تفکیک بین دو گروه مورد استفاده قرار بگیرند. بهترین تفکیک بین دو گروه از داده ها توسط ابر صفحه ای ایجاد می شود که بیشترین فاصله را از اولین مشاهده از هر

22- Rogers et al

23- Readability

24- Bonsall et al

گروه، داشته باشد. به این فاصله اصطلاحاً حاشیه عملیاتی گفته می شود. می توان گفت هرچه حاشیه عملیاتی بزرگتر باشد خطای کمتری در تعمیم نتایج حاصل از تفکیک، وجود خواهد داشت. ممکن است ابر صفحه های بسیاری وجود داشته باشند که بتوانند تفکیک مذکور را انجام دهند؛ اما روش ماشین بردار پشتیبان منطقه ابر صفحه ای را انتخاب می کند که بیشترین فاصله یا حاشیه را بین دو گروه ایجاد کند. به زبان ریاضی، روش ماشین بردار پشتیبان ابر صفحه ای را انتخاب می کند که فاصله آن از اولین نقطه (مشاهده) از هر دو گروه از داده ها بیشینه باشد که به آن اصطلاحاً ابر صفحه با بیشترین حاشیه گفته می شود.

اخیراً رگرسیون بردار پشتیبان به علت توانایی پیش بینی و دسته بندی به صورت همزمان در ادبیات مالی در سال های اخیر، توجه زیادی را بین پژوهشگران به خود جلب نموده است. از جمله کاربردین تکنیک در حوزه مالی ارزشیابی مدل های پیش بینی ورشکستگی شرکت ها مین و لی (۲۰۰۵)، پیش بینی روند بازار سرمایه (یو و همکاران، ۲۰۰۵) پیش بینی قیمت سهام (بائو و همکاران، ۲۰۰۵)، بوده است. اولین بارونپالانیسوامی (۲۰۰۱) در بورس استرالیا اقدام به گارگیری این تکنیک نمود. او با استفاده از داده های ده ساله و بر اساس ماشین بردار پشتیبان دو کلاسه سهام ها را دسته بندی نمود و با استفاده از ۲۵٪ سهام برتر در طی پنج سال اخیر پرتفویی با بازده خیره کننده ۵ برابر شاخص بازار بدست آورد. از جمله کاربردهای این تکنیک در حوزه متن کاوی پژوهش مانلا و موریرا (۲۰۱۶) است که معیاری از عدم اطمینان را بر مبنای سنجش های مبتنی بر متن بر روی مقالات مجلات وال استریت ژورنال از سال ۱۸۹۰ بنا کردند. بر اساس همین معیار آنها دریافتند که پوشش خبری مربوط به جنگ ها و سیاستهای دولتی، اکثر نوسانات مربوط به ریسک را توصیف می کند.

نقطه قوت رگرسیون بردار پشتیبان این است که پژوهشگران را ملزم به تعیین شکل واژه نامه های از پیش تعیین شده، موضوعات و افشاهای طبقه بندی شده دستی نمی کند. این روش به قضاوت پژوهشگران برای ارزیابی محتوای افشاهای متنی برای توصیف عناصر بنیادی شرکت استناد نمی کند. رگرسیون بردار پشتیبان به پژوهشگران این اجازه را می دهد که لغت نامه ای پویا از لغات کلیدی و وزن آنها تعیین نمایند که می توانند با تغییر اوضاع اقتصادی در نوسان و تغییر باشند. در نتیجه رگرسیون بردار پشتیبان دو نتیجه مهم برای تمام روش های آماری که به طور وسیع در تحقیقات حسابداری انجام می شوند به همراه دارد. اول این که رگرسیون بردار پشتیبان می تواند برای انواعی از افشاها، زبان ها و متون محتوایی با هزینه نسبی پایین به کار برده شود. دوم این که رگرسیون بردار پشتیبان می تواند در شرایطی به کار گرفته شود که پژوهشگران از قبل لغاتی را که با عناصر بنیادی قابل سنجش در ارتباط هستند، مشخص کرده باشند. تکنیک های تجزیه و تحلیل متنی کار فشرده کمتری را لازم دارد و به قضاوت پژوهشگران نیاز کمتری دارد و می تواند به آسانی برای نمونه های بزرگ به کار گرفته شود (لی گو و همکاران، ۲۰۱۶).

۳-۲. تکنیک مبتنی بر فرهنگ لغت

رویکرد مبتنی برواژه نامه کاربردهای فراوانی در حوزه حسابداری و مالی دارد. کیف لغات که یکی از ساده ترین و عمومی ترین روش های این رویکرد است، وسیع ترین و عمومی ترین روشی است که در مطالعات مالی و حسابداری استفاده می شود. از نظر فنی، یک لغت نامه مجموعه ای از اقلام جدول بندی شده ای است که با ویژگی ها و معیارهای مشخصی مرتبط است؛ بنابراین متن باید به فهرستی از لغات و کلماتی محدود شود که به واسطه آن ها ویژگی های خاص و معینی از اسناد نشان داده شود. بر اساس همین فهرست، محقق می تواند لغات مرتبط با هر ویژگی و خاصه را شمارش کند

و یک معیار مقایسه ای از احساسات را فراهم نماید. از آنجایی که رویکرد مبتنی بر واژه نامه برای برنامه نویسی و حتی تکرار بسیار آسان است، لذا از محبوبیت بیشتری در بین پژوهشگران حوزه متن کاوی برخوردار است. به طور عمومی دو لغت نامه بیشترین کاربرد را در پژوهش های متن کاوی در حوزه مالی داشته است: هاوارد و لوگران و مک دونالد (لی گووهمکاران ۲۰۱۶). شیوه تهیه لغت نامه های فوق بدین گونه است که پژوهشگر در ابتدای کار پژوهشی خود تعداد زیادی از واژگان برگرفته شده از یک فرهنگ لغات حسابداری معتبرا (مانند آکسفورد) به عنوان نماینده و پروکسی متغیری که قصد سنجش آن را دارد در نظر می گیرد. بسیاری از این لغات که در گزارش های کیفی و متنی تحت بررسی (مثل گزارش تفسیری مدیریت) کم تکرار هستند و یا به کار نرفته اند، از تجزیه و تحلیل ها حذف می شوند. فهرست به دست آمده شامل لغات و اصطلاحاتی است برای سنجش متغیر حسابداری و مالی مورد نظر استفاده خواهد شد. همان طور که در بالا اشاره شد ویژگی هایی چون لحن، ابهام، قدرت تبیین و تفسیر معیارهای متنی و کیفی مالی و حسابداری توسط واژگان برگرفته از فرهنگ لغات مورد سنجش قرار می گیرد فهرست لغات لوگران و مک دونالد همواره به روزرسانی شده و در وبسایت شخصی ایشان موجود است.

۳. نتیجه گیری

در هر سازمانی تصمیم گیری بسیار مهم است و برای تصمیم گیری به اطلاعات نیاز است. انبوه داده ها حاوی اطلاعات ارزشمندی است که معمولا انسان به سادگی قادر به تشخیص آن ها نیست و تشخیص این اطلاعات ارزشمند نیاز به الگوریتم هایی دارد که داده ها را واکاوی کرده و الگوها، روندها و اطلاعات موجود در دل این داده ها را مشخص کند. اهم اهتمام داده کاوی نیز، کشف دانش نهان نهفته در داده هاست. یکی از منابع داده های حسابداری که جزء داده های ساختار نیافته طبقه بندی شده است گزارش های روایتی هستند که امروزه با کمک تکنیک های هوش مصنوعی مانند رگرسیون بردار پشتیبان و روش لغت نامه ای مورد بررسی و تجزیه و تحلیل قرار می گیرند تا بتوانند برخی عناصر مهم صورت های مالی که جزء اطلاعات ساختاریافته هستند را ارزیابی، توصیف و سنجش نمایند. گزارش های متنی (روایتی) در سال های اخیر مورد توجه پژوهشگران زیادی در حوزه مالی و حسابداری قرار گرفته است. گزارش های توصیفی هیات مدیره، اسناد توصیفی با اهمیتی هستند که از طریق آن مدیریت فرصت می یابد باورها و ارزش های خود را به استفاده کنندگان گزارش های مالی، منتقل کند. گزارش های روایتی برای درک اطلاعات کمی با اهمیت هستند؛ البته زمانی مفید خواهند بودند که واژگان به کاررفته در آن ها منعکس کننده اطلاعات زیربنایی شرکت باشد؛ به عبارت دیگر گزارش های روایتی تحت تاثیر مدیریت نوشتار قرار نگیرند. زمانی که مدیران از مدیریت نوشتار در گزارش های خود استفاده می کنند، عدم تقارن اطلاعاتی افزایش پیدا می کند. این مقاله سعی کرد تا با معرفی موثرترین و پرکاربردترین تکنیک های هوش مصنوعی در حیطه مالی و حسابداری به نقش و جایگاه هوش مصنوعی در پژوهش های نوین حسابداری اشاره کند. استفاده از تکنیک های متن کاوی (به ویژه رگرسیون بردار پشتیبان و لغت نامه ای) قادر است افشاهای کیفی و ناساختار یافته شرکت ها را به تخمینی از عناصر بنیادی در سطح شرکت تبدیل کند.

۴. منابع

اسماعیلی، مهدی. (۱۳۹۱). داده کاوی مفاهیم و تکنیک ها. انتشارات: نیاز دانش، چاپ سوم، ص ۳۸
 رهروی دستجردی، علیرضا، فروغی، دکتر داریوش، کیانی، دکتر غلامحسین. (۱۳۹۷). ارزیابی خطر تقلب مدیران با استفاده از روش داده کاوی. دانش حسابداری. ۹(۱)، ۹۱-۱۱۴. doi: 10.22103/jak.2018.10913.2495

- Agana, J. A., Mireku, K., & Appiah, K. O. (2015). Comparative Predictive Abilities of Earnings and Operating Cash Flows on Future Cash Flows: Empirical Evidence from Ghana. *Accounting and Finance Research*, 4(3), p40.
- Ajina, A., M Laouiti, B Msolli. (2016). Guiding through the Fog: Does annual report readability reveal earnings management? *Research in International Business and Finance* 38, 509-516,
- Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction, Techniques, Cornell University, arxiv.org/abs/1707.02919
- Al- Attar, A. and Hussein, S. (2004). Corporate Data and Future Cash Flows, *Journal of Business, Finance & Accounting*, 31(7 & 8): 861- 903.
- Al-Attar, A., Hussein, S., Zuo, L. (2008). Earnings quality, bankruptcy risk and future cash flows. *Accounting and Business Research*, 38(1): 5-20.
- Allee, K., DeAngelis, M., (2015). The structure of voluntary disclosure narratives: evidence from tone dispersion. *Journal of Accounting Research*, 53(2), 241–274.
- 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
- Antweiler W, Frank M. (2004). Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* 59, (3): 1259-1294.
- Aoun, D. J. (2008). "The Effect of Cash Flow and Size on the Investment Decisions of ICT Firm: A Dynamic Approach", *Information Economics and Policy*, Vol.2, pp.120-134.
- Aureli, Selena. (2017). A comparison of content analysis usage and text mining in corporate disclosure. *The International Journal of Digital Accounting Research*, 17: 1-32
- Bagaeva, A. (2010), "The quality of published accounting information in Russia", *International Research Journal of Finance and Economics*, 63, 71-77.
- Bajari, P., Nekipelov, D., Ryan, S., Yang, M., (2015). Machine learning methods for demand estimation. *American Economic Review*, 105(5), 481–485.
- Bao, Y., Datta, A., (2014). Simultaneously discovering and quantifying risk types from textual risk disclosures *Management Science*, 60, 1371–1391.
- Beattie, V. (2014). Accounting narratives and the narrative turn in accounting research: issues, theory, methodology, methods and a research framework. *British Accounting Review*, 46(2), 111-134. DOI: 10.1016/j.bar.2014.05.001
- Beattie, Vivien, McInnes, Bill, Fearnley, Stella, (2004). A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes. *Accounting Forum* 28 (3), 205–236
- Bhatia. M., & Dhamija, S. (2015). Voluntary Disclosure of Financial Ratios in India. *South Asian Journal of Management*, 22(1), 29-51.
- Bochkay, K., and C.B. Levine, (2017). Using MD&A to Improve Earnings Forecasts, *Journal of Accounting, Auditing & Finance*, 1-27, DOI: 10.1177/0148558X17722919
- Bonsall, S., A. Leone, and B. Miller. (2015). A Plain English Measure of Financial Reporting Readability. Working paper.
- Boubakri, F. (2012). "The Relationship Between Accruals Quality Earnings Persistence And Accrual Anomaly In The Canadian Context". *International Journal of Economics And Finance*. Vol 4, No 6, pp 51-62.
- Brown, S., and J. Tucker. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, 49: 303–346.

- Buehlmaier, M., and T. Whited. (2014). Looking for Risk in Words: A Narrative Approach to Measuring the Pricing Implications of Finance Constraints." Working paper, University of Rochester.
- Cao, V. N., Chau, F., & Paudyal, K. (2013). "The Accruals Anomaly: An investigation from firm growth perspective". Monash University, 900 Dandenong Road, Caulfield East, VIC 3145, Australia. Email: Viet. C
- Campbell, J., Chen, H., Dhaliwal, D., Lu, H., Steele, L., (2014). The disclosure content of mandatory risk factor disclosures incorporate filings. Review of Accounting Study, 19, 396-455.
- Celik, S., Ozkan, N., & Deniz Akarim, Y. (2013). "An Investigation of the Accrual Anomaly in the Turkish Stock Market". Middle Eastern Finance and Economics -No 19, Pp 8-16.
- Chan, K., Chan, L., Jegadeesh, N. and J. Lakonishok (2006). Earnings quality and stock returns. Journal of Business, 79(3):1041-1082.
- Chen, J. V., and F. Li. (2017). Estimating the Amount of Estimation in Accruals. Available at SSRN: <https://ssrn.com/abstract=2738842>
- Chen, Y.; Lingling, Zh.; and L. Zhang (2013). "Financial Distress Prediction for Chinese Listed Manufacturing Companies", Procedia Computer Science. Vol. 17, pp. 678-686.
- Chen, X., Sun, Y., and Xu, X. (2016), "Free Cash Flow, Over-Investment and Corporate Governance in China", Pacific-Basin Finance Journal, 37 (2), PP. 81-108.
- Cheng, C. S. and D. Hollie (2008); "The Usefulness of Core and Non-Core Cash Flows in predicting Future Cash Flows", working paper, University of Houston
- Cheng CSA, Liu CS, Schaefer T (1996), Earnings permanence and the incremental information content of cash flows from operations. Journal of Accounting Research, 34(Spring):173-181.
- Cheng, C.S., Liu, C. Z. and Thomas, W. (2012), Abnormal accrual estimates and evidence of mispricing, Journal of Business Finance & Accounting, January/ March: pp23-57
- Cherkassky, V., Ma, Y., (2004). Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw. 17, 113-126.
- Cochran, J.; Darrat, A. F.; and K. Elkhail (2006). "On Bankruptcy of Internet Companies: An Empirical Inquiry", Journal of Business Research, Vol. 59, pp. 1193-1200
- Collins, D. W. and P. Hribar (2000). "Errors in Estimating Accruals: Implications for Empirical Research", Journal of Accounting Research, Vol. 40, No. 1, pp. 105-134.
- Copland, T., Koller, T., & Murria, J. (2000). Valuation: Measuring and Managing The Value of Companies. New York: John Wiley and Sons.
- Das SR. Text and context: language analytics in finance (2014). Found Trends Finance, 8(3):145-261
- Das SR, Chen MY (2007). Yahoo! for amazon: sentiment extraction from small talk on the web. Management Science, 53(9):1375-1388.
- DeAngelo, L. (1986). Accounting numbers as market valuation substitutes: A study of management buyout of public stockholders. The Accounting Review, Vol. 61, pp. 400-420.
- Dechow, P. (1994). Accounting earnings and cash flows as measures of firm performance: the role of accounting accruals, Journal of Accounting and Economics, 18:3-42.
- Dechow, P.M., W. Ge, and K.M. Schrand (2010): Understanding earnings quality: A review of the proxies, their determinants and their consequences, Journal of Accounting and Economics, 50: 344-
- Dechow, P. M., and I. D. Dichev. (2002). The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review, 77 (Supplement): 35-5.
- De Franco, G.; O. Hope; D. Vyas; and Y. Zhou (2015). Analyst Report Readability. Contemporary Accounting Research, 32: 76-104.
- Degeorge, F.; Patel, J.; and R. Zeckhauser (1999). "Earning Management to Exceed Thresholds" Journal of Business, Vol. 72, No. 1, pp. 1-33.

- Dey, R. Mithu., and Lucy Lim, (2015). Accrual reliability, earnings persistence, and stock prices: revisited, *American Journal of Business*, 30 (1): 22-48
- Dharan, B. G. (2003). "Accruals Management with Financing and Investing Transactions", Available Online - At: www.ruf.rice.edu [6 October 2012].
- Dvise, A. K., and I. Tama-Sweet. (2012). Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases Versus MD&A. *Contemporary Accounting Research*, 29: 804-37.
- Dyer, T., Lang M., and Stice-Lawrence, L. (2016). The Ever-Expanding 10-K: Why are 10-Ks Getting so Much Longer (and Does it Matter)? Working paper.
- Fairfield, P. S. Whisenant, and T. Yohn, (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review* 78, 353-371.
- Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, 49, 76-82.
- Farshadfar, S. Ng, C. Brimble, M. (2008); "The relative ability of earnings and cash flow data in forecasting future cash flows", *Pacific Accounting Review* Vol. 20 No. 3, pp. 254-268.
- Fedorova, E.; Gilenko, E.; and S. Dovzhenko (2013). "Bankruptcy Prediction for Russian Companies: Application of Combined", *Expert Systems with Applications*, Vol. 40, Issue. 18, pp. 7285-7293.
- Feldman, R.; S. Govindaraj; J. Livnat; and B. Segal. (2010). Management's Tone Change, Post Earnings Announcement Drift and Accruals. *Review of Accounting Studies*, 15: 915-53.
- Fisher, I. E., M. R. Garnsey, & M. E. Hughes. (2016). Natural Language Processing in Accounting, Auditing and Finance: A Synthesis of the Literature with a Roadmap for Future Research, *Intelligent Systems in Accounting, Finance and Management*, Vol. 23, Issue. 3: 157-214
- Fisher, Margaret R. Garnsey, Sunita Goel, and Kinsun Tam. (2010). The Role of Text Analytics and Information Retrieval in the Accounting Domain. *Journal of Emerging Technologies in Accounting*, 7(1): 1-24
- Frankel, R. (1992). 'Accounting information and firms with low-grade bonds'. Working paper, Stanford University
- Frankel, R., Jennings, J. and J. Lee. (2016). Using Unstructured and Qualitative Disclosure to Explain Accruals. *Journal of Accounting and Economics*, 62, (2-3): 209-227
- Francis, J. LaFond, R. Olsson, P. Schipper, K. (2004). Cost of equity and earning attributes. *The Accounting Review*, 79: 137-164.
- Francis, J., LaFond, R., Olsson, P., Schipper, K. (2005). The market pricing of accruals quality, *The journal of Accounting and economics*, 39: 295-327
- Frazier KB, R. W. Ingram; and B. M. Tenyson. (1984). A Methodology for the Analysis of Narrative Accounting Disclosures. *Journal of Accounting Research*, 318-331
- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani. (2013). *An introduction to statistical learning: with applications in R*. New York: Springer,
- Garcia J. L., & Sogorb, F. (2014). Sensitivity of External Resource to Cash Flow under Financial Constraints. *International Business Review*, Vol. 23, Pp. 920-930.
- Gong, G.; Yue-Li, L., and H. Xie (2009). "The Association between Management Earning Forecast Errors and Accruals", *The Accounting Review*, Vol. 84, No. 2, pp. 497-530.
- Greenberg, R. R., G. L. Johnson and K. Ramesh. (1986). "Earnings Versus Cash Flow as Predictor of Future Cash Flow Measures". *Journal of Accounting, Auditing and Finance*, Vol. 1, Pp. 266-277.
- Guay, W. (2006). Discussion of role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research*, Vol. 44, pp. 112-125.
- Han, I., Kim, B. Y., Lee, J., & Park, S. H. (2013). "Information Asymmetry and the Accrual Anomaly". KAIST College of Business Working Paper Series, (2013)034

- Hanna, D.J. (1995). 'Financial distress and unexpected cash flows'. Working paper, University of Chicago.
- Healy, P. & Wahlen, J. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13 (4): 365-383.
- Hoglund, H. (2012). "Detecting Earnings Management with Neural Networks", *Expert Systems with Applications*, Vol. 39, pp. 9564-9570.
- Hribar, P., Samuel J. Melessa, R. Christopher Small and Jaron H. Wilde. (2017). Does Managerial Sentiment Affect Accrual Estimates? Evidence from the Banking Industry, *Journal of Accounting and Economics*, 63(1): 26-50
- Hsieh Chia-Chun, Kai Wai Hui and Yao Zhang. (2016). Analyst Report Readability and Stock Returns, *Journal of Business Finance & Accounting*, 43(1) & (2), 98-130
- Huang, X. (2013). Managerial ability and real earnings management. Unpublished Doctoral Dissertation, Oklahoma University.
- Ingram RW, Frazier KB. (1980). Environmental performance and corporate disclosure. *Journal of Accounting Research*, 18(2):614-622
- Lang, M., and Stice-Lawrence, L., (2015). Textual Analysis and International Financial Reporting: Large Sample Evidence, *Journal of Accounting and Economics*, 110-135
- Leuz, C., and, P. Wysocki. (2016). The Economics of Disclosure and Financial Reporting Regulation: Evidence and Suggestions for Future Research. *Journal of Accounting Research*, 54:525-622.
- Lev., B. (2013). Earning Quality: Hand book of key Global Financial Markets, Institutions and Infrastructure, Chapter 36: 391-400.
- Lev, B., Li, S. and T. Sougiannis. (2010). "The Usefulness of Accounting Estimates for Predicting Cash Flows and Earnings". *Review of Accounting Study*, Vol. 15, No. 4, PP. 779-807.
- Li F. (2008). Annual report readability, current earnings, and earnings persistence, *Journal of Accounting and Economics*, 45(2):221-247.
- Li, F. (2010a) Textual Analysis of Corporate Disclosures: A Survey of the Literature. *Journal of Accounting Literature*, 29: 143-65.
- Li F. (2010b). The information content of forward looking statements in corporate filings a naive Bayesian machine learning approach. *Journal of Accounting Research*, 48(5): 1049-1102.
- Li Guo, Feng Shi, Jun Tu., (2016). Textual analysis and machine learning: Crack unstructured data in finance and accounting. *The Journal of Finance and Data Science*, 2: 153-170
- Lo., K., Ramos., F. and R. Rogo., (2017). Earnings management and annual report readability, *Journal of Accounting and Economics*, (Article in Press.)
- Loughran T, McDonald B., (2016). Textual Analysis in Accounting and Finance: A Survey., *Journal of Accounting Research*, 54, (4): 1187-1230.
- Lundholm, R. J.; R. Rogo; and J. Zhang. (2014). Restoring the Tower of Babel: How Foreign Firms Communicate with US Investors. *The Accounting Review*, 89:1453-85.
- Manela, A., Moreira, A., (2017). News implied volatility and disaster concerns. *Journal of Financial Economics*, 123(1): 137-162
- Rajgopal, Shiva and Venkatachalam, Mohan. (2011). Financial Reporting Quality and Idiosyncratic Return Volatility, *The Journal of Accounting and Economics*, 51: 1-20
- Riley, T.J., Gu'n R. Semin and Alex C. Yen. (2014). Patterns of Language Use in Accounting Narratives and Their Impact on Investment- Related Judgments and Decisions. *Behavioral Research in Accounting*, 26(1): 59-84.
- Rogers, J., VanBuskirk, A., Zechman, S., (2011). Disclosure tone and shareholder litigation. *The Accounting Review*, 86:2155-2183.
- Schipper, K. (1989). Commentary on Earnings Management. *Accounting Horizons*, 3(4):91-102
- Subramanyam A., Titman, S. (1996). The pricing of discretionary accruals. *Journal of Accounting and Economics* 22:93-115

- Tetlock PC. (2007). Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance*, 62(3):1139-1168.
- Tsai, C.F. and Y.J. Chiou (2009). "Earnings Management Prediction: A Pilot Study of Combining Neural Networks and Decision Trees", *Expert Systems with Applications*, Vol. 36, pp. 7183-7191.
- Turban, E., Sharda, R., & Delen, D. (2011). *Decision support and business intelligence systems*. USA: Prentice Hall Press.
- Vapnik, V., (1998), *Statistical Learning Theory*. Wiley, New York
- Data mining: Practical machine learning tools and techniques. Morgan Kaufmann. 2nd

Application of artificial intelligence techniques in the field of finance and accounting

Zohreh Zivdar¹

Date of Receipt: 2022/04/07 Date of Issue: 2022/05/26

Abstract

Transparent and comparable financial information is a key element of accountability and informed economic decisions. The main role of financial reporting is the effective transfer of information to outsiders. the world, which are mainly based on structured reports and quantitative and numerical calculations. Today, however, much of the financial information is disseminated through qualitative disclosures such as narrative reports and textual documents. The search for tools that can extract useful knowledge by examining financial resources and analyzing them from raw data in unstructured and qualitative reports, led to the formation of a branch in artificial intelligence known as data mining. The use of data mining on its own text created another branch of artificial intelligence called text mining. Text mining is a special type of data mining that seeks to extract useful information from unstructured and semi-structured text data (such as text documents) by recognizing and displaying patterns. Many studies have shown that today, to measure and predict financial and accounting variables, we must move away from quantitative reports and rely more on qualitative reports, and instead of relying solely on "numbers" on "analysis" of the text. Reports "have also relied on; this has led accounting and financial researchers in recent years to use artificial intelligence techniques in various areas of accounting and auditing knowledge, especially to understand the relationship between text in corporate disclosures and corporate fundamentals.". In this article, we have tried to introduce and explain the most widely used and effective artificial intelligence techniques in the field of finance, accounting and auditing

Keywords

Artificial Intelligence, Data Mining, Text Mining, Narrative Disclosures (Textual)

1. Assistant Professor, Department of Accounting, Dolatabad Branch, Islamic Azad University, Isfahan, Iran (zzivdar55@yahoo.com).