فصلنامه علمی پژوهشی دانش سرمایه‌گذاری
دوره ۱۶ / شماره ٤ (پیاپی ٦٤) / زمستان ۱٤۰۶
صفحه ۳۲۱ تا ۳٤٦

بررسی نقش تحلیل گراف و پردازش زبان طبیعی در کشف شبکه‌های تقلب مالی با تعدیل گری مؤلفه‌های حاکمیت شرکتی

امیر انصاری
گروه حسابداری، واحد یاسوج، دانشگاه آزاد اسلامی، یاسوج، ایران
Amir.ansari382@iau.ac.ir

هاشم ولی‌پور
گروه حسابداری، واحد فیروزآباد، دانشگاه آزاد اسلامی، فیروزآباد، ایران(نویسنده مسئول)
Hashem. valipour @iau.ac.ir

حمید صالحی
گروه حسابداری، واحد فیروزآباد، دانشگاه آزاد اسلامی، فیروزآباد، ایران
Salehi.hamid@iau.ac.ir

تاریخ دریافت: ۱۴۰۴/۰۵/۰۹ تاریخ پذیرش: ۱۴۰۴/۰۹/۰۹

چکیده

این پژوهش باهدف بررسی تأثیر تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی (GNN) و پردازش زبان طبیعی (NLP)بر کشف شبکه‌های تقلب مالی انجام‌شده و نقش تعدیل گر مؤلفه‌های حاکمیت شرکتی در افزایش این تأثیرات مورد ارزیابی قرارگرفته است. پژوهش حاضر از نوع کاربردی و توصیفی-پیمایشی بوده و داده‌ها از طریق پرسشنامه‌ای استاندارد و اقتباسی، میان نمونه‌ای ۱۵۰ نفره از متخصصان و خبرگان حوزه‌های مالی، حسابرسی، فناوری اطلاعات، مدیریت ریسک و نظارت مالی در شرکت‌های دولتی و خصوصی ایران جمع‌آوری شده‌است. پرسشنامه‌ها به‌صورت ترکیبی آنلاین و حضوری توزیع شدند. تحلیل داده‌ها با استفاده از نرم‌افزارهای آماری SPSS و PLS-SEM انجام‌شده‌است. نتایج نشان داد که تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی تأثیر مثبت و معناداری بر کشف شبکه‌های تقلب مالی دارد و پردازش زبان طبیعی نیز توانایی استخراج شاخص‌های رفتاری و متنی مؤثر در کشف تقلب مالی را به‌طور معناداری افزایش می‌دهد. نوآوری این پژوهش در ترکیب همزمان رویکردهای GNN و NLP با توجه به مؤلفه‌های حاکمیت شرکتی برای کشف تقلب مالی و ارائه چارچوبی کاربردی برای بهبود شفافیت و کنترل‌های داخلی در سازمان‌ها است.

واژه‌های کلیدی: مؤلفه‌های حاکمیت شرکتی، کشف تقلب مالی، پردازش زبان طبیعی، تحلیل گراف.

۱- مقدمه

تقلب مالی یکی از پیچیده‌ترین و مخرب‌ترین بحران‌های اقتصادی است که پیامدهای گسترده‌ای بر سلامت بازارهای مالی، اعتماد عمومی و شفافیت اطلاعاتی دارد. با توسعه اقتصادی، وقوع تقلب‌های مالی به‌طور مستمر در حال افزایش است و تشخیص و پیشگیری از آن نقش تعیین‌کننده‌ای در حفظ نظم و پایداری نظام مالی ایفا می‌کند (مودپالی،۲۰۲۵). تقلب مالی، به‌ویژه در شرایط رقابتی شدید بنگاه‌ها و بحران‌های اقتصادی فراگیر، نگرانی‌های جدی را در کشورهای درحال‌توسعه و بازارهای نوظهور ایجاد کرده است. تعاریف مختلفی از تقلب مالی ارائه‌شده است که هر یک بر جنبه‌های مشخصی از رفتارهای متقلبانه و ناهنجاری‌های مالی تأکیددارند (کوبیلای و همکاران، ۲۰۲۳). رسوایی‌های مالی شرکت‌های بزرگ جهانی، ازجمله انرون و ورلدکام و آثار زیان‌بار ناشی از آن‌ها طی یک دهه گذشته، منجر به کاهش اعتماد عمومی و به‌ویژه کاهش اعتماد سرمایه‌گذاران نسبت به گزارش‌های مالی و خدمات حسابرسی شده است (خواجوی و ابراهیمی، ۱۳۹۶). در ایران نیز، وقوع تقلب طی سال‌های اخیر به یکی از موضوعات محوری در بازارهای مالی و نهادهای اقتصادی-اجتماعی تبدیل‌شده است. بحران‌های مالی ایران در سال‌های ۱۳۹۰ و ۱۳۹۷ که با کاهش شدید ارزش ریال همراه بود، پیامدهای منفی قابل‌توجهی بر فضای کسب‌وکار و دسترسی سازمان‌ها به منابع مالی ایجاد کرده است. این شرایط اهمیت استفاده از روش‌های علمی و سیستماتیک برای شناسایی و پیش‌بینی تقلب را بیش‌ازپیش آشکار می‌سازد (ملکی کاکلر و همکاران، ۱۴۰۰). در مطالعه موردی ایران، اهمیت شناسایی و پیش‌بینی تقلب تا حدی است که افشای تخلفات مالی یک شرکت پذیرفته‌شده در بورس اوراق بهادار می‌تواند منجر به کاهش اعتماد به کل بازار سرمایه و افت ارزش سهام سایر شرکت‌ها شود و آثار مخرب آن تا مدت‌ها ادامه یابد؛ بنابراین، بهره‌گیری از تکنیک‌های پیشرفته برای پیش‌بینی تقلب، نه‌تنها در سطح سازمانی بلکه در سطح کلان اقتصادی، به‌عنوان یک ابزار هشدار پیش از وقوع از اهمیت حیاتی برخوردار است. درروش‌های سنتی حسابداری، شناسایی تقلب عمدتاً از طریق نمونه‌گیری تصادفی یا سیستماتیک انجام می‌شد که اغلب تصویر جامع و دقیقی از وضعیت مالی سازمان ارائه نمی‌کرد. با پیشرفت فناوری‌های هوشمند، این رویکردها جای خود را به مدل‌های داده محور و الگوریتم‌های پیشرفته داده‌کاوی و هوش مصنوعی داده‌اند. سیستم‌های مبتنی بر هوش مصنوعی قادرند با بهره‌گیری از الگوریتم‌های یادگیری ماشین و شبکه‌های عصبی، رفتارهای انسانی را شبیه‌سازی و تحلیل کنند، الگوهای پنهان تقلب را شناسایی نمایند و پیش‌بینی‌های دقیق‌تری ارائه دهند. بدین ترتیب، ادغام فناوری‌های نوین با متدهای حسابداری امکان شناسایی و پیش‌بینی مؤثر تقلب‌های مالی را فراهم ساخته و سطح اطمینان و دقت تصمیم‌گیری‌های مالی را به‌طور قابل‌توجهی ارتقاء می‌دهد (زارعی و همکاران،۱۴۰۵).

در سال‌های اخیر، گزارش‌های رسمی سازمان حسابرسی و سازمان بورس اوراق بهادار ایران حاکی از افزایش تعداد پرونده‌های تقلب و تخلفات مالی در شرکت‌های بزرگ و متوسط بوده است. به‌عنوان‌مثال، برخی از پرونده‌های مرتبط با تحریف صورت‌های مالی و سوءاستفاده از منابع شرکت‌های دولتی و خصوصی نشان می‌دهد که عدم شفافیت و ضعف سامانه‌های کنترل داخلی، زمینه مناسبی برای وقوع تقلب فراهم کرده است. علاوه بر این، رشد سریع معاملات الکترونیکی و استفاده از ابزارهای مالی پیچیده، شرایط را برای تقلب‌های چندلایه و شبکه‌ای در اقتصاد ایران مهیا کرده است. در این شرایط، تشخیص زودهنگام شاخص‌های تقلب برای جلوگیری از زیان‌های

مالی و کاهش بی‌اعتمادی سرمایه‌گذاران اهمیت ویژه‌ای دارد (جبار زاده و همکاران،۱۳۹۹). روش‌های سنتی کشف تقلب عمدتاً مبتنی بر تحلیل نسبت‌های مالی و گزارش‌های حسابرسی هستند که در مواجهه با پیچیدگی‌های نوین بازار ایران ناکارآمدند و توانایی شناسایی روابط غیرمستقیم و پنهان میان ذی‌نفعان را ندارند. این محدودیت‌ها موجب افزایش خطاهای مثبت و منفی کاذب و کاهش اثربخشی نظارت مالی شده است (فارایولا و همکاران، ۲۰۲۳). تجربه‌های عملی در حسابرسی شرکت‌های ایرانی نشان داده است که روش‌های سنتی، به‌ویژه در مواجهه با معاملات میان شرکت‌های چند شاخه و شبکه‌ای، قادر به تشخیص رفتارهای غیرعادی و تراکنش‌های مشکوک نیستند و این امر موجب افزایش احتمال وقوع زیان‌های مالی گسترده و پیامدهای منفی برای بازار سرمایه می‌شود(احمدی و همکاران،۱۴۰۳).

بنابراین، بهره‌گیری از فناوری‌های نوین هوش مصنوعی و یادگیری ماشین، به‌ویژه تحلیل گراف مبتنی بر شبکه‌های عصبی (GNN) و پردازش زبان طبیعی(NLP) ، برای شناسایی الگوهای پیچیده و چندبعدی تقلب در محیط اقتصادی ایران ضروری است. تحلیل گراف مبتنی بر شبکه‌های عصبی امکان مدل‌سازی دقیق ساختارهای شبکه‌ای پیچیده را فراهم می‌کند؛ در این مدل‌ها، گره‌ها نماینده واحدهای اقتصادی و یال‌ها روابط مالی و تراکنشی میان آن‌ها هستند .این رویکرد، مبتنی بر نظریه شبکه‌های اجتماعی، رفتارها و تصمیمات افراد را در چارچوب ساختار شبکه‌ای تحلیل می‌کند و با استخراج ویژگی‌های نهفته، شبکه‌های همدستی و تراکنش‌های مشکوک را شناسایی می‌نماید (وانگ و همکاران، ۲۰۲۵). همزمان، پردازش زبان طبیعی با تحلیل متون مالی شامل گزارش‌های سالانه، یادداشت‌های توضیحی و افشاهای داوطلبانه، اطلاعات رفتاری و شاخص‌های مخفی مدیران را استخراج می‌کند که می‌تواند نشانه‌های تقلب احتمالی را آشکار سازد (دو و همکاران، ۲۰۲۰). ترکیب این دو رویکرد، تصویری جامع و چندبعدی از فرآیند تقلب ارائه می‌دهد که فراتر از توانایی روش‌های سنتی است.

علاوه بر ضرورت فنی، محیط نهادی و حاکمیتی شرکت‌ها در ایران نیز نقش مهمی در وقوع و پیشگیری تقلب دارد. ضعف در ساختارهای حاکمیتی، مانند ترکیب ناکافی هیئت‌مدیره، محدودیت در استقلال حسابرس، نبود کمیته‌های حسابرسی فعال و کیفیت پایین افشاهای مالی، می‌تواند انگیزه و فرصت ارتکاب تقلب را افزایش دهد. تجربه‌های پژوهشی نشان می‌دهد که در شرکت‌های ایرانی با حاکمیت ضعیف، حتی وجود شاخص‌های هشداردهنده تکنیکی نیز اغلب به‌موقع شناسایی نمی‌شوند و تقلب‌های سازمان‌یافته فرصت بروز می‌یابند. با توجه به این شرایط، ترکیب ابزارهای پیشرفته فنی با تحلیل نهادی و حاکمیتی، می‌تواند اثربخشی کشف تقلب را به‌طور چشمگیری افزایش دهد (کاظمی و پیری،۱۴۰۱). باوجود پیشرفت‌های فنی، بسیاری از پژوهش‌ها نقش عوامل نهادی و ساختارهای حاکمیت شرکتی را در کشف تقلب نادیده گرفته‌اند.تئوری آژانس نشان می‌دهد که ضعف در ساختارهای حاکمیتی، انگیزه و فرصت ارتکاب تقلب را افزایش می‌دهد مؤلفه‌های حاکمیت شرکتی شامل ترکیب هیئت‌مدیره، استقلال حسابرس، فعالیت کمیته حسابرسی و کیفیت افشا می‌توانند رابطه میان شاخص‌های کشف تقلب تکنیکی و وقوع واقعی آن را تعدیل کنند (ویتوساری وهمکاران، ۲۰۲۵). این مسئله در اقتصاد ایران که نظارت و شفافیت گاه محدود است، اهمیت ویژه‌ای پیدا می‌کند.

هدف این پژوهش، ارائه چارچوبی یکپارچه و میان‌رشته‌ای است که ضمن بهره‌گیری همزمان از تحلیل گراف مبتنی بر شبکه‌های عصبی و پردازش زبان طبیعی، نقش تعدیل گر مؤلفه‌های حاکمیت شرکتی در کشف تقلب مالی در محیط اقتصادی ایران را بررسی می‌کند. نوآوری اصلی پژوهش در ترکیب داده‌های ساختاری و متنی با توجه به ساختارهای حاکمیتی است که امکان تحلیل دقیق‌تر و جامع‌تر فرآیند پیچیده تقلب را فراهم می‌آورد. یافته‌های پژوهش می‌تواند نهادهای نظارتی، بورس‌ها و حسابرسان مستقل را در طراحی سامانه‌های پیشگیرانه و هوشمند یاری رسانده و به افزایش شفافیت، کاهش زیان‌های ناشی از تقلب و ارتقای سلامت بازارهای مالی ایران منجر شود. علاوه بر این، مطالعه حاضر با ارائه نمونه‌های تحلیلی از محیط واقعی بازار ایران، به توسعه دانش علمی و کاربردی در حوزه کشف تقلب کمک می‌کند. این پژوهش می‌تواند الگوهای رفتاری و تراکنشی مدیران و شرکت‌ها را شناسایی کند و راهکارهایی برای بهبود سامانه‌های نظارتی و کنترل داخلی ارائه دهد. درنهایت، مدل پیشنهادی می‌تواند به‌عنوان مرجعی برای تحقیقات آینده در حوزه حسابرسی فناوری محور، تحلیل ریسک و حاکمیت شرکتی در ایران مورداستفاده قرار گیرد و به توسعه سیاست‌گذاری‌های مقرراتی و نظارتی کمک نماید.

ساختار این مقاله به ترتیب زیر ارائه‌شده است: در بخش بعدی، به مروری بر ادبیات پژوهش پرداخته می‌شود سپس، روش پژوهش، جامعه آماری، ابزار گردآوری داده‌ها و روش‌های تحلیل داده‌ها بیان می‌شوند. در ادامه، نتایج تحلیل داده‌ها ارائه و نقش تحلیل گراف، پردازش زبان طبیعی و مؤلفه‌های حاکمیت شرکتی در کشف تقلب مالی موردبررسی قرار می‌گیرد. درنهایت، بحث و نتیجه‌گیری ارائه‌شده و توصیه‌های کاربردی برای نهادهای نظارتی و پژوهش‌های آتی مطرح می‌گردد.

۲-مبانی نظری و تئوری‌های مرتبط با فرضیات تحقیق

۲-۱-نقش تحلیل گراف در کشف شبکه‌های تقلب مالی

تحلیل گراف به‌عنوان روشی پیشرفته در شناسایی و مدل‌سازی روابط پیچیده میان عوامل مختلف در داده‌های مالی بزرگ اهمیت فزاینده‌ای یافته است. شبکه‌های تقلب مالی معمولاً شامل مجموعه‌ای از بازیگران و ارتباطات مستقیم و غیرمستقیم میان آن‌ها هستند که ساختار پیچیده‌ای ایجاد می‌کنند در این میان، فناوری‌های نوین تحلیل گراف مانند شبکه‌های عصبی گرافی امکان درک عمیق‌تر روابط غیرخطی و پنهان را فراهم کرده‌اند که روش‌های سنتی قادر به شناسایی آن‌ها نیستند (گوری و همکارن، ۲۰۰۵). مطالعات جدید نشان داده‌اند که استفاده از شبکه‌های عصبی گرافی در تحلیل تراکنش‌های مالی به‌طور قابل‌توجهی دقت کشف تقلب را افزایش داده و همدستی‌های پیچیده میان عوامل تقلب را بهتر شناسایی می‌کند . این فناوری همچنین به شناسایی الگوهای پویای شبکه‌های تقلب کمک می‌کند که درگذر زمان تغییر می‌کنند و در صورت عدم توجه، از چشم ناظران پنهان می‌مانند (وانگ و همکاران، ۲۰۲۵).

علاوه بر این، تحلیل گراف به دلیل قابلیت ذاتی در مدل‌سازی روابط چندوجهی و همزمان میان اجزای شبکه، امکان درک عمیق‌تری از ساختارهای تقلب را فراهم می‌کند. تقلب‌های مالی اغلب به‌صورت شبکه‌هایی با اعضای متعدد، ارتباطات پیچیده و چندلایه پدیدار می‌شوند؛ به‌گونه‌ای که تشخیص بازیگران اصلی یا نهادهای مرکزی،

بدون در نظر گرفتن ارتباطات و مسیرهای میان آن‌ها، بسیار دشوار خواهد بود.در این زمینه، شاخص‌های مختلف تحلیل گراف مانند مرکزیت، درجه، نزدیک‌تر بینابینی ابزارهای کلیدی برای شناسایی نقش و نفوذ بازیگران در شبکه‌های تقلب هستند (هو و همکاران، ۲۰۱۹). به‌عنوان‌مثال، گره‌هایی با مرکزیت بالا معمولاً نقش واسطه‌ای یا رهبری در شبکه را ایفا می‌کنند و تمرکز تحقیقات می‌تواند روی آن‌ها باشد (هارل و همکاران، ۲۰۲۰).

از سوی دیگر، تحلیل گراف قادر است تعاملات زمانی و تغییرات شبکه را نیز به‌صورت پویا مدل‌سازی کند. در بسیاری از موارد، شبکه‌های تقلب به‌مرورزمان ساختار خود را تغییر می‌دهند تا کشف نشوند یا بازیگران جدید وارد شبکه شوند . با استفاده از تکنیک‌های تحلیل گراف پویا می‌توان روند تغییرات شبکه را رصد کرد و الگوهای مشکوک را در مراحل اولیه شناسایی نمود (دِولین و همکاران، ۲۰۱۹). این توانایی تحلیل زمانی به ناظران و حسابرسان امکان می‌دهد پیش‌بینی‌های دقیق‌تری درباره احتمال وقوع تقلب داشته باشند و اقدامات پیشگیرانه مناسبی اتخاذ کنند.

یکی دیگر از مزایای تحلیل گراف، قابلیت ادغام آن با سایر روش‌های هوش مصنوعی و یادگیری ماشین است. به‌خصوص در قالب شبکه‌های عصبی گرافی، این روش می‌تواند ویژگی‌های پیچیده و چندبعدی داده‌های گراف را استخراج کرده و مدل‌هایی با دقت بسیار بالا در تشخیص تقلب بسازد. این شبکه‌ها با یادگیری بازنمایی‌های تعبیه‌شده برای گره‌ها و یال‌ها، می‌توانند روابط غیرخطی و مخفی را شناسایی کنند که سایر الگوریتم‌های سنتی از درک آن‌ها عاجزند (هو و همکاران، ۲۰۲۰).

تحقیقات اخیر نشان می‌دهند که استفاده از شبکه‌های عصبی گرافی در حوزه کشف تقلب مالی نه‌تنها به افزایش دقت تشخیص کمک می‌کند، بلکه زمان پاسخ‌گویی سیستم‌های نظارتی را نیز کاهش می‌دهد و می‌تواند به‌صورت بلادرنگ تراکنش‌ها و تعاملات را تحلیل کند.این موضوع اهمیت فراوانی دارد زیرا شبکه‌های تقلب به‌سرعت تغییر می‌کنند و کشف به‌موقع آن‌ها می‌تواند مانع خسارات مالی گسترده شود (فاچیا و همکاران، ۲۰۲۳).

علاوه بر موارد فوق، تحلیل گراف می‌تواند به شناسایی ساختارهای سلسله‌مراتبی یا چندلایه در شبکه‌های تقلب کمک کند؛ ساختارهایی که معمولاً در سازمان‌های بزرگ یا گروه‌های پیچیده تقلب مشاهده می‌شوند (چنگ و همکاران، ۲۰۱۹). این امکان، راه را برای طراحی مدل‌های دقیق‌تر ریسک و تدوین استراتژی‌های مقابله‌ای بهتر هموار می‌سازد.

درنهایت، تحلیل گراف می‌تواند نقش مهمی در ترکیب داده‌های ساختاری و داده‌های متنی ایفا کند که در بخش پردازش زبان طبیعی به آن اشاره خواهد شد. این ادغام چندمنظوره، قدرت تشخیص و کشف تقلب را به سطح بالاتری می‌برد و امکان تحلیل جامع‌تری از شبکه‌های تقلب فراهم می‌سازد؛ بنابراین، فرضیه اول بر این اساس شکل‌گرفته است که:

فرضیه ۱: تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی تأثیر معناداری در کشف شبکه‌های تقلب مالی دارد.

۲-۲-نقش پردازش زبان طبیعی در کشف شاخص‌های رفتاری و متنی مرتبط با تقلب مالی

پردازش زبان طبیعی به‌عنوان فناوری‌ای پیشرفته، امکان استخراج معانی نهفته و ساختارهای پنهان در داده‌های متنی مالی را فراهم می‌آورد که در تحلیل عددی و ساختاری داده‌ها قابل‌رؤیت نیستند .این داده‌های متنی شامل گزارش‌های سالانه، افشاهای مالی، یادداشت‌های توضیحی، مکاتبات داخلی و حتی پیام‌های رسانه‌های اجتماعی می‌شود که در آن‌ها نشانه‌هایی از رفتارهای غیر شفاف یا غیرقانونی ممکن است وجود داشته باشد (جیانگ، و همکاران، ۲۰۱۹). پیشرفت‌های مدل‌های مبتنی بر مبدل مانند برت (مدل بازنمایی رمزگذار دوطرفه از متون)، رابرتا (برت قوی‌تر شده) و مدل زبانی (GPT) باعث شده است تا سامانه‌های پردازش زبان طبیعی قادر به شناسایی دقیق‌تر لحن، تهدیدات ضمنی، تناقض‌های معنایی و تحریف‌های زبانی باشند که از علائم مهم تقلب مالی به شمار می‌روند به‌طور خاص، تحلیل احساسات، تشخیص ناهماهنگی‌های متنی و استخراج الگوهای زبانی به‌عنوان فنون کلیدی در این حوزه مطرح هستند (دو و همکاران، ۲۰۲۰). همچنین، مطالعات جدید نشان داده‌اند که پردازش زبان طبیعی می‌تواند در کنار تحلیل شبکه‌ای و یادگیری ماشین، الگوهای رفتاری و ارتباطی پنهان را بهتر شناسایی کند و بدین ترتیب، دقت پیش‌بینی تقلب مالی را به‌صورت قابل‌توجهی ارتقاء دهد (ژانگ و همکاران، ۲۰۲۴) برای مثال، تحلیل متون ایمیل‌ها یا پیام‌های داخلی شرکت‌ها می‌تواند نشانه‌هایی از برنامه‌ریزی و هماهنگی برای تقلب را آشکار سازد که در داده‌های تراکنشی دیده نمی‌شود (بولیرس و همکاران، ۲۰۲۴)بنابراین، با توجه به توانمندی‌های پردازش زبان طبیعی در استخراج شاخص‌های رفتاری و متنی با عمق و گستره بالا، فرضیه دوم به شکل زیر مطرح می‌شود:

فرضیه ۲: پردازش زبان طبیعی قابلیت استخراج شاخص‌های رفتاری و متنی مؤثر در کشف شبکه‌های تقلب مالی را داراست.

پردازش زبان طبیعی (Natural Language Processing - NLP) به‌عنوان یکی از فناوری‌های کلیدی در حوزه هوش مصنوعی، در سال‌های اخیر نقش فزاینده‌ای در تحلیل محتوای متنی ایفا کرده است. برخلاف تحلیل‌های عددی و ساختاری که صرفاً بر داده‌های کمی متمرکزند، پردازش زبان طبیعی قادر است معانی ضمنی، مفاهیم پنهان و الگوهای زبانی مشکوک در داده‌های متنی را استخراج کند؛ داده‌هایی که اغلب نادیده گرفته می‌شوند اما می‌توانند سرنخ‌های مهمی درباره رفتارهای غیرقانونی یا غیر شفاف فراهم سازند (لن و همکاران، ۲۰۲۵).

منابع متنی در فضای مالی بسیار متنوع هستند. آن‌ها شامل گزارش‌های سالانه شرکت‌ها، صورت‌های مالی حسابرسی شده، یادداشت‌های توضیحی، مکاتبات اداری و ایمیل‌های داخلی، افشاهای رسمی، اظهارات مدیران در کنفرانس‌های خبری و حتی محتوای منتشرشده در رسانه‌های اجتماعی و شبکه‌های آنلاین می‌شوند. در این منابع، ممکن است نشانه‌هایی از قصد پنهان برای دست‌کاری اطلاعات، تحریف واقعیات یا برنامه‌ریزی برای ارتکاب تقلب وجود داشته باشد. به‌ویژه زمانی که زبان به شکل محافظه‌کارانه، ابهام‌آمیز یا احساسی مورداستفاده قرار گیرد، می‌توان احتمال رفتارهای متقلبانه را بالا دانست (کیم، ۲۰۱۴).

پیشرفت‌های اخیر در مدل‌های پردازش زبان طبیعی، به‌ویژه با بهره‌گیری از معماری‌های مبتنی بر مبدل و مدل‌های تولیدی مانند GPT، توانایی تحلیل محتوای زبانی را به سطح جدیدی رسانده‌اند. این مدل‌ها قادرند

لایه‌های معنایی عمیق، لحن، احساسات ضمنی، تهدیدات پنهان و حتی تناقض‌های معنایی در متون را با دقت بسیار بالا شناسایی کنند (لی و همکاران، ۲۰۱۹). این ویژگی‌ها به‌ویژه در افشای گزارش‌های ساختگی یا تحریف‌شده مفیدند، زیرا مدیران یا حسابداران متخلف ممکن است در متون رسمی سعی در پوشاندن واقعیت داشته باشند.

فنون کلیدی مورداستفاده در NLP مانند تحلیل احساسات، تشخیص ناهماهنگی متنی، استخراج نهادها و موجودیت‌ها و طبقه‌بندی موضوعی ابزارهای قدرتمندی برای درک نیت و رفتار افراد در متن محسوب می‌شوند. به‌عنوان‌مثال، یک گزارش مالی که بار احساسی مثبت اغراق‌آمیز دارد، ممکن است تلاشی برای منحرف کردن توجه از زیان‌های پنهان باشد .همچنین، عدم هماهنگی میان متون یادداشت‌های توضیحی و بخش‌های عددی صورت‌های مالی می‌تواند نشانه‌ای از تقلب یا افشای ناقص اطلاعات باشد (ژو و همکاران، ۲۰۲۱).

از دیگر کاربردهای مهمNLP ، تحلیل مکاتبات داخلی شرکت‌ها و ایمیل‌های کارکنان است. در بسیاری از رسوایی‌های مالی بزرگ، مانند موارد شرکت انرون، ردپای تقلب در پیام‌های داخلی کشف‌شده است، جایی که کارکنان درباره برنامه‌ریزی برای گمراه کردن سهام‌داران یا ناظران صحبت کرده‌اند. مدل‌های NLP قادرند الگوهای رفتاری مشکوک، جملات با بار معنایی منفی و اشاره‌های ضمنی به اقدامات غیرقانونی را شناسایی کرده و به‌صورت خودکار هشدار دهند (بولیرس و همکاران، ۲۰۲۴)

افزون بر این، ترکیب NLP با تحلیل شبکه‌ای و الگوریتم‌های یادگیری ماشین موجب ارتقای توانایی سیستم‌های هوشمند در شناسایی الگوهای پنهان ارتباطی و رفتاری شده است. برای نمونه، یک مکاتبه با لحن خاص در متن، وقتی در قالب شبکه ارتباطات بین کارکنان بررسی شود، ممکن است نشان‌دهنده برنامه‌ریزی برای همدستی در تقلب باشد. به‌ویژه در سازمان‌هایی با ساختار پیچیده، چنین تجزیه‌وتحلیل ترکیبی می‌تواند ابعاد پنهان مانده رفتارهای متقلبانه را آشکار سازد (ژانگ و لی، ۲۰۲۰).

درمجموع، پردازش زبان طبیعی باقابلیت تحلیل گسترده، عمیق و هوشمند محتوای متنی، ابزاری قدرتمند برای کشف تقلب محسوب می‌شود. این فناوری نه‌تنها مکمل تحلیل‌های عددی و ساختاری است، بلکه می‌تواند مستقل از آن‌ها نیز شاخص‌های معناداری برای هشدارهای اولیه فراهم کند.

بنابراین، با توجه به توانمندی‌های پردازش زبان طبیعی در استخراج شاخص‌های رفتاری و متنی با عمق و گستره بالا، فرضیه دوم به شکل زیر مطرح می‌شود:

فرضیه ۲ : پردازش زبان طبیعی قابلیت استخراج شاخص‌های رفتاری و متنی مؤثر در کشف شبکه‌های تقلب مالی را داراست.

۲-۳- نقش تعدیل گری مؤلفه‌های حاکمیت شرکتی در رابطه تحلیل گراف و کشف تقلب مالی

نظریه نمایندگی به‌عنوان چارچوبی بنیادی در علوم مالی و حسابداری، بر وجود تضاد منافع ذاتی میان مدیران (نمایندگان) و مالکان یا سهامداران (اصیل‌ها) تأکید دارد (جنسن و مکلینگ، ۱۹۷۶). بر اساس این نظریه، مدیران ممکن است در نبود نظارت مؤثر و ساختارهای کنترل‌کننده، انگیزه‌هایی برای اتخاذ رفتارهای فرصت‌طلبانه مانند

پنهان‌کاری، تحریف اطلاعات مالی و حتی ارتکاب به تقلب داشته باشند. در این میان، حاکمیت شرکتی به‌عنوان مجموعه‌ای از سازوکارهای داخلی و خارجی طراحی‌شده است تا این تضاد را کاهش دهد و رفتار مدیران را هم‌راستا با منافع سهامداران تنظیم کند (بیرول، بورجو. ، ۲۰۱۹).

با گسترش ابزارهای هوشمند مانند تحلیل گراف در کشف روابط پنهان و شبکه‌های پیچیده تقلب، نقش حاکمیت شرکتی نه‌فقط به‌عنوان مانعی برای وقوع تقلب، بلکه به‌عنوان تعدیل گر در اثربخشی این فناوری‌ها موردتوجه قرارگرفته است. تحلیل گراف به‌ویژه در محیط‌هایی اثربخش‌تر عمل می‌کند که داده‌ها شفاف، ساختاریافته و قابل‌ردیابی باشند. حاکمیت شرکتی با تقویت شفافیت اطلاعات، نظارت بر مدیریت و ارتقای پاسخگویی، چنین محیطی را فراهم می‌سازد (مانداجیو همکاران، ۲۰۱۲).

مطالعات تجربی متعددی بر این نکته تأکیددارند که ساختارهای قوی حاکمیت شرکتی، اثربخشی الگوریتم‌های تحلیل گراف را در کشف تقلب افزایش می‌دهند .به‌عنوان‌مثال، استقلال هیئت‌مدیره موجب کاهش نفوذپذیری هیئت‌مدیره از سوی مدیران اجرایی شده و نظارت مؤثرتری را بر فرآیندهای مالی و تصمیم‌گیری‌ها فراهم می‌آورد. این امر منجر به ایجاد داده‌های دقیق‌تر، گزارش دهی شفاف‌تر و کاهش تلاش برای مخفی سازی روابط مشکوک مالی می‌شود (لو و همکاران، ۲۰۲۰).

از سوی دیگر، کیفیت حسابرسی به‌عنوان یکی دیگر از مؤلفه‌های کلیدی حاکمیت شرکتی، نقش مهمی در اعتبار و قابلیت اعتماد داده‌های مورداستفاده در تحلیل گراف دارد. شرکت‌هایی که از خدمات حسابرسی باکیفیت بهره‌مند هستند، معمولاً سطح پایین‌تری از تحریف و دست‌کاری در اطلاعات مالی دارند که این موضوع تحلیل الگوریتم‌های گرافی را تسهیل و تقویت می‌کند (کوتوپیس و همکاران، ۲۰۲۱).

افزون بر آن، افشای داوطلبانه و سطح شفافیت اطلاعات مالی نیز به‌طور مستقیم بر عملکرد تحلیل گراف اثرگذار است. در محیط‌هایی که شرکت‌ها اطلاعات بیشتری در اختیار بازار و نهادهای ناظر قرار می‌دهند، شبکه‌های اطلاعاتی واضح‌تر و ارتباطات مالی شفاف‌تر شده و تشخیص الگوهای تقلب آمیز پنهان از طریق تحلیل گراف آسان‌تر می‌شود (اسلام و رحمان، ۲۰۲۵).

در این میان، تحلیل گراف به‌تنهایی نمی‌تواند تمامی جنبه‌های رفتاری و ساختاری تقلب را آشکار سازد. به‌ویژه در شرکت‌هایی با حاکمیت ضعیف، ممکن است بخشی از روابط مشکوک در قالب ارتباطات غیررسمی، استفاده از شرکت‌های واسطه‌ای، یا معاملات برون دفتری انجام شود که در داده‌های تحلیلی قابل‌شناسایی نباشند. در چنین شرایطی، حتی پیشرفته‌ترین الگوریتم‌های گرافی نیز با محدودیت مواجه می‌شوند (کیم و همکاران، ۲۰۱۴)؛ اما زمانی که مؤلفه‌های حاکمیت شرکتی در سطح بالایی قرار دارند، همان فناوری تحلیل گراف می‌تواند با دسترسی به داده‌های شفاف‌تر، ساختارهای پیچیده تقلب را با دقت بیشتری شناسایی کند.

از منظر ترکیبی، حاکمیت شرکتی نقش هم‌افزا با فناوری تحلیل گراف ایفا می‌کند .این هم‌افزایی باعث می‌شود که ساختارهای سازمانی از یک‌سو احتمال وقوع تقلب را کاهش دهند و از سوی دیگر ابزارهای تحلیل داده را برای شناسایی موارد باقیمانده توانمندتر سازند. چنین ساختاری یک چرخه بازخورد مثبت ایجاد می‌کند که در آن، نظارت قوی و تحلیل فناورانه یکدیگر را تقویت می‌کنند.

شایان‌ذکر است که مطالعات مقطعی در شرکت‌های بورسی در اقتصادهای نوظهور نیز نشان داده‌اند که شرکت‌هایی با هیئت‌مدیره مستقل‌تر، تمرکز مالکیت پایین‌تر و افشای شفاف‌تر، در مواجهه با الگوریتم‌های تحلیل گراف، نرخ کشف تقلب بالاتری را تجربه می‌کنند (بولیرس و همکاران، ۲۰۲۴). این یافته‌ها مؤید آن است که حاکمیت شرکتی نه‌فقط عامل بازدارنده تقلب، بلکه بستری برای ارتقاء کارایی ابزارهای داده محور است.

درنتیجه، تحلیل گراف زمانی می‌تواند به حداکثر کارایی خود برسد که در محیطی با ساختار حاکمیتی قوی پیاده‌سازی شود. در چنین محیطی، داده‌ها کامل، قابل‌اعتماد و قابل‌پیگیری بوده و سیستم نظارتی سازمانی نیز از یافته‌های تحلیل پشتیبانی می‌کند. از این منظر، مؤلفه‌های حاکمیت شرکتی نقش تعدیل گر معنادار و اثرگذار در رابطه بین فناوری تحلیل گراف و شناسایی تقلب مالی ایفا می‌کنند؛ بنابراین، فرضیه سوم به شرح زیر ارائه می‌شود:

فرضیه ۳ : مؤلفه‌های حاکمیت شرکتی نقش تعدیل گر معناداری در رابطه بین تحلیل گراف و کشف شبکه‌های تقلب مالی دارند؛ به‌گونه‌ای که افزایش قدرت حاکمیت شرکتی، اثر تحلیل گراف را افزایش می‌دهد.

۲-۴- نقش تعدیل گری مؤلفه‌های حاکمیت شرکتی در رابطه پردازش زبان طبیعی و کشف تقلب مالی

در سال‌های اخیر، پردازش زبان طبیعی به یکی از ابزارهای قدرتمند برای کشف نشانه‌های زبانی، متنی و رفتاری مرتبط با تقلب مالی تبدیل‌شده است. این فناوری به کمک مدل‌های پیشرفته زبانی مانندBERT ، RoBERTa و GPT، توانایی شناسایی ظرافت‌های معنایی، لحن، ناهماهنگی‌های کلامی و حتی تحریف‌های عمدی در متون مالی را به‌صورت خودکار و با دقت بالا فراهم کرده است (ما، اِکس ، ۲۰۲۱). باوجوداین توانمندی‌ها، باید توجه داشت که میزان اثربخشی شاخص‌های استخراج‌شده توسط پردازش زبان طبیعی تا حد زیادی تحت تأثیر بستر نهادی و سازمانی شرکت‌ها، به‌ویژه ساختار حاکمیت شرکتی، قرار دارد.

مطابق با نظریه نمایندگی، ساختارهای ضعیف حاکمیت شرکتی زمینه‌ساز فرصت‌طلبی‌های مدیریتی و کاهش شفافیت اطلاعاتی هستند (جنسن و مکلینگ، ۱۹۷۶). در چنین محیط‌هایی، حتی در صورت کشف سیگنال‌های هشداردهنده توسط مدل‌هایNLP ، ممکن است این شاخص‌ها به تقلب واقعی منجر شوند، چراکه سازوکارهای کنترلی مؤثر برای پیشگیری از این فرایند وجود ندارد (فاضل و همکاران، ۲۰۲۴). در مقابل، ساختارهای قوی حاکمیت شرکتی با بهره‌گیری از هیئت‌مدیره مستقل، کیفیت بالای حسابرسی، وجود کمیته‌های حسابرسی فعال و افشای شفاف اطلاعات، نه‌تنها بروز تقلب را محدود می‌کنند، بلکه زمینه را برای کارایی بیشتر ابزارهای NLP فراهم می‌سازند (لی و همکاران، ۲۰۲۲).

به‌بیان‌دیگر، حاکمیت شرکتی قوی، نقش تعدیل گر در رابطه میان شاخص‌های رفتاری و زبانی استخراج‌شده از متون مالی و وقوع یا کشف تقلب ایفا می‌کند. برای مثال، زمانی که مدل‌های پردازش زبان طبیعی تناقضات معنایی یا تحریف‌های زبانی را در گزارش‌های سالانه یا مکاتبات داخلی شناسایی می‌کنند، این هشدارها تنها در صورتی منجر به کشف یا جلوگیری از تقلب می‌شوند که نهادهای نظارتی داخلی سازمان دارای قدرت اجرایی کافی باشند (گاوو و همکاران، ۲۰۲۳). در غیر این صورت، شاخص‌ها صرفاً به‌عنوان نشانه‌های بالقوه باقی می‌مانند و منجر به اقدام اصلاحی نخواهند شد.

علاوه بر این، فرهنگ‌سازمانی شکل‌گرفته در چارچوب حاکمیت شرکتی نیز در این میان مؤثر است. ساختارهای حاکمیتی مبتنی بر اخلاق، شفافیت و پاسخگویی، سبب می‌شوند که کارکنان و مدیران سطح پایین‌تر نیز حساسیت بیشتری نسبت به رفتارهای پر ریسک یا تقلب آمیز نشان دهند و اطلاعات مرتبط را راحت‌تر در اختیار سامانه‌های نظارتی قرار دهند (چن و همکاران، ۲۰۲۲). درنتیجه، مدل‌های NLP که با داده‌های غنی‌تری تغذیه می‌شوند، قادر به شناسایی الگوهای پنهان‌تری خواهند بود.

ازنظر فنی، کیفیت داده‌های متنی، ساختاربندی گزارش‌ها و وجود منابع اطلاعاتی متنوع (مانند صورت‌های مالی، گزارش‌های حسابرسی، صورت‌جلسات هیئت‌مدیره، ایمیل‌های داخلی و ...) تأثیر بسزایی در دقت و کارایی مدل‌های زبانی دارد. حاکمیت شرکتی مؤثر با الزام به افشای جامع و قابل‌اتکای اطلاعات، محیط داده‌ای مناسبی برای استقرار موفق پردازش زبان طبیعی ایجاد می‌کند (کامینسکی و همکاران، ۲۰۰۴).

همچنین، مطالعات اخیر نشان می‌دهد که هم‌افزایی میان ابزارهای هوش مصنوعی و سازوکارهای نهادی مانند حاکمیت شرکتی می‌تواند منجر به کشف تقلب‌های پیچیده‌تر شود که در شرایط معمول و بدون تعدیل‌کننده‌های نهادی، از دید تحلیلگران و سیستم‌های خودکار پنهان می‌مانند (میکو و همکاران، ۲۰۱۵).

برای نمونه، در یک مطالعه موردی در شرکت‌های فناوری اطلاعات جنوب شرق آسیا، مشخص شد که در سازمان‌هایی با کمیته حسابرسی فعال و جلسات منظم هیئت‌مدیره، شاخص‌های استخراج‌شده توسط NLP مانند تحلیل لحن منفی در گزارش‌های مالی یا واژگان غیرمعمول در افشاهای مدیریتی) در پیش‌بینی تقلب‌های آتی دقت بسیار بالاتری داشتند نسبت به شرکت‌هایی که فاقد این زیرساخت‌ها بودند (ژانگ و لی، ۲۰۲۰).این یافته‌ها بر نقش دوگانه حاکمیت شرکتی به‌عنوان ابزار پیشگیرانه و تقویت‌کننده ابزارهای تحلیلی تأکید می‌کند.

درنهایت، می‌توان نتیجه گرفت که پردازش زبان طبیعی، هرچند ابزاری قدرتمند در تحلیل متون مالی است، اما برای دستیابی به حداکثر اثربخشی نیازمند محیطی ساختاریافته و مبتنی بر اصول شفافیت، نظارت و پاسخگویی است .مؤلفه‌های حاکمیت شرکتی نه‌تنها از وقوع تقلب پیشگیری می‌کنند، بلکه شرایط لازم برای عملکرد بهتر مدل‌های زبانی را نیز فراهم می‌سازند؛ بنابراین، این ساختارها نقش تعدیل گر مؤثری در رابطه بین شاخص‌های رفتاری و زبانی استخراج‌شده توسط NLP و کشف شبکه‌های تقلب مالی ایفا می‌کنند. بر همین اساس، فرضیه چهارم به‌صورت زیر تعریف می‌شود:

فرضیه ۴: مؤلفه‌های حاکمیت شرکتی نقش تعدیل گر معناداری در رابطه بین پردازش زبان طبیعی و کشف شبکه‌های تقلب مالی دارند؛ به‌گونه‌ای که افزایش قدرت حاکمیت شرکتی، اثر شاخص‌های رفتاری و متنی کشف‌شده توسط پردازش زبان طبیعی را افزایش می‌دهد.

۳- پیشینه تحقیق

کشف تقلب مالی به‌عنوان یکی از چالش‌های اساسی در حوزه حسابرسی و مدیریت مالی، توجه گسترده‌ای را در پژوهش‌های علمی به خود معطوف کرده است. با رشد حجم و پیچیدگی داده‌های مالی، استفاده از فناوری‌های نوین مانند تحلیل گراف و پردازش زبان طبیعی برای شناسایی شبکه‌ها و شاخص‌های پنهان تقلب اهمیت فراوانی

یافته است. در این راستا، مطالعات متعددی در سطح جهانی و داخلی انجام‌شده که هر یک از جنبه‌ای به بررسی این موضوع پرداخته‌اند.

در حوزه تحلیل گراف، تحقیقات متعددی به توسعه مدل‌های یادگیری عمیق مبتنی بر شبکه‌های عصبی گراف پرداخته‌اند. به‌طور مثال، پژوهش تیان و همکاران (۲۰۲۳) مدل ASA-GNN را معرفی کردند که با بهره‌گیری از ساختارهای پیچیده تراکنش‌های مالی، توانسته است دقت شناسایی تقلب را به شکل معناداری افزایش دهد. همچنین، مطالعه کیم و همکاران (۲۰۲۳) با ارائه مدلDRAG ، مکانیزم توجه پویا را در تشخیص تقلب به کار گرفتند که موجب ارتقای عملکرد سیستم‌های کشف تقلب شد. تحقیقاتی مانند دو، کلوین و همکاران (۲۰۲۵) نیز بر اهمیت تحلیل دینامیک شبکه‌های تقلب در طول زمان تأکید داشته‌اند و نشان داده‌اند که تحلیل گراف قادر است تغییرات پیچیده و همدستی‌های پنهان را بهتر آشکار سازد.

در زمینه پردازش زبان طبیعی، تحقیقات گسترده‌ای با استفاده از مدل‌های مبتنی بر ترنسفورمر مانندBERT ، RoBERTaو GPT-4 انجام‌شده است. خوزه و همکاران (۲۰۲۵) با تمرکز بر استخراج شاخص‌های متنی از گزارش‌های مالی و مکاتبات داخلی، نشان دادند که پردازش زبان طبیعی می‌تواند الگوهای رفتاری و نشانه‌های هشداردهنده تقلب را به‌طور دقیق‌تر شناسایی کند. همچنین، هوی و همکاران(۲۰۲۵) در پژوهش با هدف مقابله با محدودیت‌های روش‌های سنتی در تحلیل متون بلند و ترکیب داده‌های چندگانه، یک چارچوب نوین برای تشخیص تقلب مالی شرکت‌ها پیشنهاد می‌کند. این چارچوب با ادغام خلاصه‌سازی معنایی گزارش‌های سالانه توسط مدل‌های زبان بزرگ (LLMs) و مدل‌های یادگیری ماشینی قابل تفسیر (مانند (GBDT با شاخص‌های مالی و حاکمیتی، کارایی و شفافیت بیشتری در شناسایی تقلب مالی فراهم می‌آورد.

در زمینه تلفیق فناوری‌ها، تحقیقات اسلام و همکاران (۲۰۲۵) و وانگ و همکاران (۲۰۲۵) به توسعه مدل‌های ترکیبی پرداخته‌اند که با ادغام یادگیری فدرال، تحلیل گراف و پردازش زبان طبیعی توانسته‌اند دقت پیش‌بینی تقلب را به‌طور چشمگیری ارتقاء دهند. این رویکردها با حفظ حریم خصوصی داده‌ها و تحلیل جامع ساختارهای رفتاری و شبکه‌ای، به‌عنوان روش‌های پیشرو در حوزه کشف تقلب مطرح‌شده‌اند.

زارعی و همکاران (۱۴۰۵) در پژوهشی باهدف ارزیابی کارایی مدل ماشین بردار پشتیبان (SVM) در پیش‌بینی تقلب مالی، داده‌های ۱۳۶ شرکت بورسی را طی سال‌های ۱۳۹۵ تا ۱۴۰۱ تحلیل کردند. نتایج نشان داد مدل SVMبا دقت ۸۵٪، بازخوانی ۹۷٪ و صحت ۸۳٪ عملکرد مطلوبی در تشخیص تقلب دارد. همچنین متغیرهایی مانند جمع کل دارایی‌های دوره قبل، جریان نقد عملیاتی و دارایی‌های ثابت بیشترین تأثیر را در پیش‌بینی تقلب ایفا کرده‌اند. از سوی دیگر، خواجوی و ابراهیمی (۱۳۹۷) به بررسی تأثیر سازوکارهای حاکمیت شرکتی بر تقلب در صورت‌های مالی شرکت‌های بورسی تهران پرداخته است. با استفاده از علائم خطر استاندارد حسابرسی شماره ۲۴۰ و تکنیک‌های داده‌کاوی، شرکت‌های متقلب شناسایی شدند. نتایج پژوهش نشان داد ترکیب هیئت‌مدیره، تمرکز مالکیت و مالکیت نهادی با وقوع تقلب در صورت‌های مالی رابطه معناداری دارد. حاجیان (۱۴۰۴) در پژوهشی باهدف طراحی مدلی مبتنی بر یادگیری ماشین برای کشف تقلب در صورت‌های مالی شرکت‌های پذیرفته‌شده در بورس تهران انجام‌شده است. مدل پیشنهادی با بهره‌گیری از داده‌های تاریخی مالی و الگوریتم‌هایی

مانند جنگل تصادفی، ماشین بردار پشتیبان و شبکه‌های عصبی توسعه یافت .نتایج تحقیق نشان داد مدل طراحی‌شده از دقت بالایی در شناسایی شرکت‌های متقلب برخوردار است و می‌تواند به‌عنوان ابزاری مؤثر برای افزایش شفافیت مالی، بهبود نظارت و ارتقای اعتماد سرمایه‌گذاران به کار گرفته شود. توکلی وهمکاران (۱۴۰۲) در پژوهشی به بررسی نقش پردازش زبان طبیعی (NLP) در تحلیل گزارش‌های مالی می‌پردازد. با توجه به حجم زیاد داده‌های متنی مالی، استفاده از تکنیک‌هایی مانند تحلیل احساسات، طبقه‌بندی و خلاصه‌سازی متون می‌تواند به استخراج سریع‌تر و دقیق‌تر اطلاعات کلیدی کمک کند .یافته‌ها نشان می‌دهد NLP دقت تحلیل‌های مالی را افزایش داده، خطاهای انسانی را کاهش می‌دهد و فرآیند تصمیم‌گیری مالی را بهبود می‌بخشد. در مطالعه‌ای دیگر، ملکی کاکلر و همکاران، (۱۴۰۰) به بررسی کارایی مدل‌های آماری و الگوریتم‌های یادگیری ماشین در پیش‌بینی گزارشگری مالی متقلبانه پرداخته است. با استفاده از ۲۰ متغیر در قالب الگوی پنج‌ضلعی تقلب و داده‌های ۱۶۶ شرکت بورسی طی سال‌های ۱۳۸۸ تا ۱۳۹۷، نتایج نشان داد مدل‌های یادگیری ماشین نسبت به مدل‌های آماری دقت و کارایی بالاتری در شناسایی تقلب دارند. ترکیب الگوریتم‌های درخت تصمیم‌گیریCHAID ، C5و C&R بالاترین دقت را با ۹۲/۶۱ درصد در پیش‌بینی گزارشگری مالی متقلبانه ارائه کرد و نشان داد که روش‌های داده‌کاوی مبتنی بر یادگیری ماشین‌ابزار مؤثری برای پیشگیری و کشف تقلب در صورت‌های مالی هستند.

باوجوداین پیشرفت‌ها، در ایران کمتر پژوهشی به‌طور جامع و تلفیقی به بررسی همزمان تحلیل گراف، پردازش زبان طبیعی و نقش تعدیل گر حاکمیت شرکتی در کشف تقلب پرداخته است. علاوه بر این، استفاده از فناوری‌های نوین یادگیری عمیق و مدل‌های ترنسفورمر در این حوزه هنوز به‌طور کامل توسعه‌نیافته است که فرصت مناسبی برای تحقیقات آینده فراهم می‌آورد.

و در پژوهش‌های بین‌المللی نشانگر پیشرفت‌های قابل‌توجه در توسعه مدل‌های پیشرفته یادگیری ماشین و پردازش زبان طبیعی برای کشف تقلب مالی هستند که همزمان نقش تعدیل گر مؤلفه‌های حاکمیت شرکتی را نیز در نظر گرفته‌اند. این رویکردها موجب افزایش دقت و قابلیت تعمیم‌پذیری مدل‌های کشف تقلب شده‌اند. در مقابل، پژوهش‌های داخلی گرچه در سال‌های اخیر به‌صورت محدود و جزئی پیشرفت‌هایی داشته‌اند، اما همچنان به مطالعات تلفیقی و کاربرد فناوری‌های نوین نیازمندند تا بتوانند همگام با روند جهانی حرکت کنند و نیازهای بازار مالی کشور را بهتر پاسخ دهند.

۴-روش‌شناسی تحقیق

این پژوهش از نوع کاربردی محسوب می‌شود؛ زیرا هدف اصلی آن، استفاده از روش‌های نوین داده‌کاوی و هوش مصنوعی—شامل تحلیل گراف و پردازش زبان طبیعی(NLP) در راستای شناسایی و کشف الگوهای پنهان تقلب مالی در بستر واقعی سازمان‌ها و نهادهای اقتصادی است.

از نظر روش گردآوری داده‌ها، این تحقیق در سال ۱۴۰۳ و در زمره مطالعات توصیفی–پیمایشی انجام شد. داده‌های میدانی با استفاده از پرسشنامه جمع‌آوری شده و روابط میان متغیرهای تحقیق از طریق تحلیل‌های آماری بررسی شد.

جامعه آماری تحقیق شامل متخصصان و خبرگان فعال در حوزه‌های مالی، حسابرسی، فناوری اطلاعات، مدیریت ریسک و نظارت مالی در شرکت‌های دولتی و خصوصی ایران است؛ به‌ویژه افرادی که با طراحی، پیاده‌سازی و یا ارزیابی سامانه‌های کشف تقلب، کنترل‌های داخلی و تحلیل ریسک آشنایی دارند. برای دستیابی به نمونه‌ای نماینده از کل جامعه هدف، از روش نمونه‌گیری طبقه‌ای نسبی استفاده شد تا اطمینان حاصل شود که تمامی زیرگروه‌های تخصصی به نسبت مشارکتشان در اکوسیستم کشف تقلب در نمونه نهایی حضورداشته باشند. با در نظر گرفتن وسعت و نامشخص بودن دقیق اندازه جامعه آماری و بر اساس جدول کرجسی و مورگان، حجم نمونه تحقیق ۱۵۰ نفر در نظر گرفته شد. فرایند توزیع پرسشنامه‌ها نیز به دو شیوه حضوری و آنلاین صورت گرفت تا پوشش مناسبی از خبرگان شاغل در نهادهای مختلف حاصل شود.

ابزار گردآوری داده‌ها، یک پرسشنامه اقتباسی نیمه ساختاریافته است که بر پایه ادبیات نظری و تجربی معتبر داخلی و بین‌المللی طراحی‌شده است. این پرسشنامه با بهره‌گیری از مطالعات برجسته نظیرعبدالله و همکاران (۲۰۱۶) و ابری و محمدی (۱۴۰۳) تدوین‌شده و شامل ۴۰ گویه در چهار بعد اصلی است: تحلیل گراف، پردازش زبان طبیعی، کشف تقلب مالی و مؤلفه‌های حاکمیت شرکتی. مقیاس پاسخ‌دهی در قالب طیف پنج‌درجه‌ای لیکرت طراحی‌شده و از «کاملاً مخالفم» تا «کاملاً موافقم» امتداد دارد.

برای بررسی روایی محتوایی ابزار، طراحی گویه‌ها بر اساس چارچوب‌های نظری مورد تأیید نظیر چارچوب COSO (2015)و با نظارت خبرگان دانشگاهی و حرفه‌ای صورت پذیرفته است. به‌منظور اطمینان از پایایی ابزار اندازه‌گیری، از ضریب آلفای کرونباخ برای هریک از ابعاد استفاده‌شده و پایایی ترکیبی نیز محاسبه گردیده است. همچنین، به‌منظور سنجش روایی سازه‌ای، از تحلیل عاملی تأییدی (CFA) بهره گرفته‌شده و شاخص‌هایی نظیر بارهای عاملی، میانگین واریانس استخراج‌شده (AVE) و نسبت واریانس مشترک برای ارزیابی کفایت سازه‌ها به‌کاررفته‌اند.

در بخش تحلیل داده‌ها، از دو نرم‌افزار SPSSبرای آمار توصیفی و بررسی اولیه پایایی و از SmartPLSبرای آزمون فرضیه‌ها با استفاده از مدل‌سازی معادلات ساختاری با رویکرد حداقل مربعات جزئی (PLS-SEM) بهره گرفته‌شده است. ابتدا آمار توصیفی برای تحلیل ویژگی‌های جمعیت شناختی پاسخ‌دهندگان ارائه‌شده و سپس ساختار عاملی تأییدی جهت ارزیابی مناسب بودن مدل اندازه‌گیری اجرا گردیده است. در مرحله بعد، فرضیه‌های اصلی پژوهش از طریق آزمون روابط میان تحلیل گراف، پردازش زبان طبیعی و کشف شبکه‌های تقلب مالی بررسی‌شده‌اند. همچنین، نقش تعدیل گر مؤلفه‌های حاکمیت شرکتی در این روابط از طریق آزمون تعاملات و ضرایب مسیر تعدیل‌شده موردسنجش قرارگرفته است.

درمجموع، این روش‌شناسی با بهره‌گیری از طراحی علمی، نمونه‌گیری نظام‌مند، ابزار معتبر و تحلیل آماری چند سطحی، بستری فراهم می‌سازد برای بررسی دقیق نقش فناوری‌های نوین داده‌کاوی در بهبود نظام‌های کشف تقلب مالی، با توجه به ساختار نهادی و حاکمیتی سازمان‌های ایرانی.

۵-یافته‌های تحقیق

۵-۱-تحلیل و تفسیر اطلاعات جمعیت شناختی

تحلیل داده‌های جمعیت‌شناسی جدول ۱ نشان می‌دهد که بیشترین فراوانی مربوط به گروه سنی ۳۱ تا ۴۰ سال (۳۷/۳٪) و کمترین مربوط به بالای ۵۰ سال (۱۴/۷٪) است که بیانگر حضور غالب افراد میان‌سال و فعال در حوزه‌های مرتبط با پژوهش می‌باشد. ازنظر جنسیت، ۶۱/۳٪ مرد و ۳۸/۷٪ زن هستند که متناسب با ساختار جنسیتی حوزه‌های مالی و فناوری در ایران است. سابقه کاری بیشتر پاسخ‌دهندگان در بازه ۶ تا ۱۰ سال (۳۲٪) متمرکز است که نشان‌دهنده تجربه کاری متوسط و متنوع در نمونه است. در بخش تحصیلات، اکثریت با مدرک کارشناسی ارشد (۵۶٪) حضور دارند که تأییدی بر سطح تخصصی بالای نمونه است. همچنین، توزیع حوزه فعالیت پاسخ‌دهندگان عمدتاً در بخش‌های مالی (۳۴٪) و حسابرسی (۲۸/۷٪) متمرکز بوده و مشارکت مناسبی از حوزه فناوری اطلاعات (۲۱/۳٪) نیز مشاهده می‌شود. این ترکیب جمعیت‌شناسی متنوع و مرتبط با موضوع

جدول ۱: اطلاعات جمعیت‌شناسی

متغیر	دسته‌بندی	فراوانی	درصد(%)
سن	زیر ۳۰ سال	۳۴	۲۲/۷
	۳۱-۴۰ سال	۵۶	۳۷/۳
	۴۱-۵۰ سال	۳۸	۲۵/۳
	بالای ۵۰ سال	۲۲	۱۴/۷
جنسیت	مرد	۹۲	۶۱/۳
	زن	۵۸	۳۸/۷
سابقه کاری	زیر ۵ سال	۲۶	۱۷/۳
	۶-۱۰ سال	۴۸	۳۲/۰
	۱۱-۱۵ سال	۴۲	۲۸/۰
	بالای ۱۵ سال	۳۴	۲۲/۷
تحصیلات	کارشناسی	۳۹	۲۶/۰
	کارشناسی ارشد	۸۴	۵۶/۰
	دکتری	۲۷	۱۸/۰
حوزه فعالیت	مالی	۵۱	۳۴/۰
	حسابرسی	۴۳	۲۸/۷
	فناوری اطلاعات	۳۲	۲۱/۳
	سایر	۲۴	۱۶/۰

پژوهش، به افزایش اعتبار و تعمیم‌پذیری نتایج کمک می‌کند.

۵-۲-آمار توصیفی

نتایج توصیفی نشان می‌دهد که پردازش زبان طبیعی (میانگین = ۳/۹۱، انحراف معیار = ۰/۶۵) و کشف شبکه‌های تقلب مالی (میانگین = ۳/۸۵، انحراف معیار = ۰/۶۸) بالاترین میزان اهمیت را از دیدگاه پاسخ‌دهندگان دارند . تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی با میانگین ۳/۷۲و نوسان بیشتر (انحراف معیار = ۰/۷۴) دیدگاه‌های متنوع‌تری را منعکس می‌کند .مؤلفه‌های حاکمیت شرکتی نیز با میانگین ۳/۶۸ و انحراف معیار ۰/۷۱ نشان‌دهنده ارزیابی متعادل شرکت‌کنندگان نسبت به نقش آن در کشف تقلب است. به طور کلی، داده‌ها حاکی از توجه بالای شرکت‌کنندگان به فناوری‌های نوین در کشف تقلب و اهمیت نسبی حاکمیت شرکتی در این فرآیند است.

جدول (۲) توصیف متغیرهای تحقیق

متغیر	مینیمم	ماکزیمم	میانگین	انحراف معیار
کشف شبکه‌های تقلب مالی	۲	۵	۳/۸۵	۰/۶۸
تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی	۱	۵	۳/۷۲	۰/۷۴
پردازش زبان طبیعی	۲	۵	۳/۹۱	۰/۶۵
مؤلفه‌های حاکمیت شرکتی	۱	۵	۳/۶۸	۰/۷۱

۵-۳- برازش مدل تحقیق و ارزیابی مدل اندازه‌گیری

در این پژوهش، ارزیابی کیفیت مدل اندازه‌گیری با استفاده از روش حداقل مربعات جزئی در چارچوب مدل‌سازی معادلات ساختاری مبتنی بر پی‌ال‌اس (PLS-SEM) و مطابق با پروتکل‌های استاندارد روش‌شناختی (هیر و همکاران، ۲۰۱۷) انجام شد. نتایج تحلیل داده‌ها نشان می‌دهد که تمامی سازه‌های پژوهش شامل «تحلیل گراف»، «پردازش زبان طبیعی»، «مؤلفه‌های حاکمیت شرکتی» و «کشف تقلب مالی» از پایایی و روایی همگرای قابل قبولی برخوردارند.

پایایی درونی سازه‌ها با دو شاخص آلفای کرونباخ و پایایی ترکیبی موردسنجش قرار گرفت. مقادیر آلفای کرونباخ برای همه سازه‌ها در بازه ۰/۷۰۲ تا ۰/۸۱۹ گزارش شد که بالاتر از آستانه استاندارد ۰/۷ (نانلی و برنشتاین، ۱۹۹۴) است و نشان‌دهنده همسانی درونی مطلوب گویه‌هاست. پایایی ترکیبی نیز به‌عنوان معیاری برتر شناخته می‌شود (هنزلر و همکاران، ۲۰۱۶)، مقادیری بین ۰/۸۰۹ تا ۰/۸۷۳ را نشان داد. این شاخص با در نظر گرفتن همبستگی گویه‌ها و وزن بارهای عاملی، تأیید می‌کند که سازه‌ها از قابلیت اتکای بالایی برخوردارند.

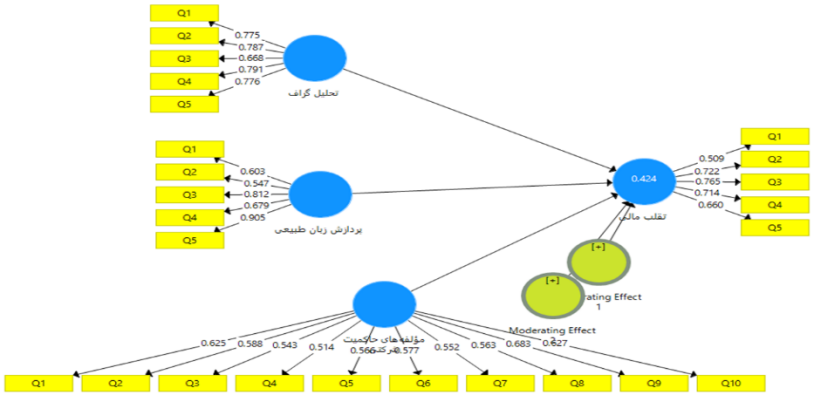
از منظر روایی همگرا، شاخص‌های بارهای عاملی و میانگین واریانس استخراج‌شده (AVE) مورد تحلیل قرار گرفتند. بارهای عاملی تمامی گویه‌ها در محدوده ۰/۵۲۱ تا ۰/۵۷۹ قرار دارد که مطابق با ملاک پذیرش بزرگ‌تر از ۰/۵ است (هیر و همکاران، ۲۰۱۷).

نتایج کلی مطابق با جدول ۲ و شکل ۱ حاکی از آن است که ابزار پژوهش از اعتبار سازه و پایایی درونی کافی برخوردار است. تأیید همزمان شاخص‌های آلفای کرونباخ، پایایی ترکیبی، بارهای عاملی و اِی‌وی‌ای مطابق با

چارچوب‌های روش‌شناختی پیشرفته، گواهی بر قابلیت اعتماد و دقت ابزار پژوهش است (داوری و رضازاده، ۱۳۹۲). اگرچه برخی گویه‌ها نیاز به بازبینی جزئی دارند، اما تأثیر آن‌ها بر کیفیت کلی مدل ناچیز است؛ بنابراین، این مدل قابلیت به‌کارگیری در تحلیل‌های ساختاری پیچیده‌تر و آزمون فرضیه‌های پژوهش را داراست و امکان تعمیم‌پذیری نتایج را در حوزه موضوعی پژوهش فراهم می‌سازد.

جدول (۳) برازش مدل تحقیق و ارزیابی مدل اندازه‌گیری

متغیرها	گویه‌ها	بار عاملی	آلفای کرونباخ	روایی همگرا (AVE)	پایایی ترکیبی
کشف شبکه‌های تقلب مالی	Q1	۰/۵۰۹	۰/۷۰۲	۰/۵۶۳	۰/۸۰۹
	Q2	۰/۷۲۲			
	Q3	۰/۷۶۵			
	Q4	۰/۷۱۴			
	Q5	۰/۶۶۰			
تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی	Q1	۰/۷۷۵	۰/۸۱۹	۰/۵۷۹	۰/۸۷۳
	Q2	۰/۷۸۷			
	Q3	۰/۶۶۸			
	Q4	۰/۷۹۱			
	Q5	۰/۷۷۶			
پردازش زبان طبیعی	Q1	۰/۶۰۳	۰/۸۱۱	۰/۵۲۱	۰/۸۴۰
	Q2	۰/۵۴۷			
	Q3	۰/۸۱۲			
	Q4	۰/۶۷۹			
	Q5	۰/۹۰۵			
مؤلفه‌های حاکمیت شرکتی	Q1	۰/۶۲۵	۰/۷۸۷	۰/۵۴۳	۰/۸۳۸
	Q2	۰/۵۸۸			
	Q3	۰/۵۴۳			
	Q4	۰/۵۱۴			
	Q5	۰/۵۶۶			
	Q6	۰/۵۷۷			
	Q7	۰/۵۵۲			
	Q8	۰/۵۶۳			
	Q9	۰/۶۸۳			
	Q10	۰/۶۲۷			



شکل (۱): مدل تحلیل عاملی تائیدی (با بارهای عاملی مناسب)

۴-۵-روایی واگرا

فورنل و لارکر (۱۹۸۱) روش معتبری را برای ارزیابی روایی واگرا در مدل‌های اندازه‌گیری معرفی نمودند که بر اساس آن میزان همبستگی هر سازه با مؤلفه‌های خود در مقایسه با همبستگی آن سازه با سایر سازه‌های مدل سنجیده می‌شود. روایی واگرا زمانی تأیید می‌گردد که میانگین واریانس استخراج‌شده (AVE) هر سازه، بیش از واریانس مشترک آن سازه با سایر سازه‌ها باشد، به‌عبارت‌دیگر، مربع ضرایب همبستگی بین سازه‌ها باید کمتر از مقدار AVE مربوط به هر سازه باشد. این معیار نشان‌دهنده تفکیک‌پذیری مناسب سازه‌ها و تضمین می‌کند که هر سازه به‌طور معناداری با شاخص‌های خودهمبستگی دارد و از سایر سازه‌ها تفکیک‌پذیر است. جدول (۳) ماتریس روایی واگرای مدل‌های اندازه‌گیری پژوهش حاضر را ارائه می‌دهد که با توجه به موضوع تحقیق—شامل تحلیل گراف، پردازش زبان طبیعی و مؤلفه‌های حاکمیت شرکتی در کشف شبکه‌های تقلب مالی—حاکی از صحت ساختاری و اعتبار تفکیکی بالای سازه‌ها می‌باشد. این نتایج، قابلیت اطمینان و اعتبار مدل اندازه‌گیری را برای تحلیل‌های بعدی و آزمون فرضیات تحقیق تضمین می‌کند و بیانگر استقلال و تمایز هر یک از سازه‌های پژوهش در چارچوب نظری و عملی آن است.

جدول (۴). آزمون فورنل- لارکر

متغیرها	A1	A2	A3	A4	A5	A6
A1	۱/۰۰۰					
A2	۰/۰۱۸	۱/۰۰۰				
A3	۰/۰۰۷	۰/۰۱۹	۱/۰۰۰			
A4	۰/۱۸۴	۰/۰۰۸	۰/۳۳۰	۱/۰۰۰		
A5	۰/۱۶۱	۰/۰۰۷	۰/۲۲۷	۰/۶۱۱	۱/۰۰۰	
A6	۰/۰۱۷	۰/۰۲۶	۰/۰۰۵	۰/۰۹۵	۰/۰۶۹	۱/۰۰۰

که در اینجا:

A1: حاصل‌ضرب تحلیل گراف در مؤلفه‌های حاکمیت شرکتی

A2: حاصل‌ضرب پردازش زبان طبیعی در مؤلفه‌های حاکمیت شرکتی

A3: تحلیل گراف

A4: کشف شبکه‌های تقلب مالی

A5: مؤلفه‌های حاکمیت شرکتی

A6: پردازش زبان طبیعی

۵-۵-شاخص استون-گیسر Q2

شاخص استون-گیسر Q^2 که توسط استون و گیسر (۱۹۷۵) معرفی‌شده است، معیاری کلیدی برای سنجش قدرت پیش‌بینی مدل ساختاری در متغیرهای درون‌زاست. این شاخص بیانگر آن است که مدل‌های دارای برازش ساختاری مطلوب باید قادر به پیش‌بینی دقیق متغیرهای درون‌زا باشند؛ بدین معنا که روابط علی بین سازه‌ها به‌درستی مدل‌سازی شده و اثرات متقابل آن‌ها به‌گونه‌ای است که فرضیه‌های پژوهش را به‌صورت معتبر پشتیبانی نماید. بر اساس توصیه هنسلر و همکاران (۲۰۰۹)، مقادیر Q^2 در سطوح ۰/۰۲، ۰/۱۵ و ۰/۳۵ به ترتیب نشان‌دهنده قدرت پیش‌بینی پایین، متوسط و بالا هستند (هیر و همکاران، ۲۰۱۷).

مطابق جدول (۴)، مقادیر شاخص Q^2 برای تمامی سازه‌های درون‌زای مدل حاضر که شامل مؤلفه‌های تحلیل گراف، پردازش زبان طبیعی و نقش تعدیل گری مؤلفه‌های حاکمیت شرکتی در کشف شبکه‌های تقلب مالی می‌باشد، فراتر از آستانه ۰/۳۵ قرارگرفته‌اند. این امر نشان‌دهنده قابلیت پیش‌بینی قوی مدل ساختاری و صحت روابط فرض شده میان سازه‌ها است؛ بنابراین، نتایج به‌دست‌آمده تأیید می‌کند که مدل پژوهش ضمن برخورداری از قابلیت برازش بالا، از توانمندی پیش‌بینی مناسبی نیز در زمینه شناسایی و تحلیل شبکه‌های تقلب مالی با رویکرد هوش مصنوعی و تأثیر حاکمیت شرکتی بهره‌مند است. این موضوع اعتبار تئوریک و کاربردی مدل را تقویت کرده و امکان تعمیم‌یافته‌ها را در حوزه پژوهشی فراهم می‌سازد.

جدول (۵). شاخص استون-گیسر Q2

متغیر	شاخص Q2
حاصل‌ضرب تحلیل گراف در مؤلفه‌های حاکمیت شرکتی	۰/۴۲۵
حاصل‌ضرب پردازش زبان طبیعی در مؤلفه‌های حاکمیت شرکتی	۰/۵۷۵
تحلیل گراف	۰/۶۲۳
کشف شبکه‌های تقلب مالی	۰/۷۲۱
مؤلفه‌های حاکمیت شرکتی	۰/۴۸۱
پردازش زبان طبیعی	۰/۶۹۷

۶-۵-آزمون فرضیه‌های تحقیق

پس از بررسی برازش مدل‌های اندازه‌گیری، مدل ساختاری، مطابق با الگوریتم تحلیل داده‌ها درروش PLS به بررسی و آزمون فرضیه‌های تحقیق پرداخته می‌شود. جهت بررسی صحت‌وسقم فرضیه‌های تحقیق حاضر از روابط علی مبتنی بر مدل‌سازی معادله‌های ساختاری استفاده شد.

جدول (۶): نتایج فرضیه‌های تحقیق

فرضیه‌ها	ضریب مسیر	آمارهٔ t	ارزش P	نتیجه
تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی -> کشف شبکه‌های تقلب مالی	۰/۴۳۰	۸/۳۲۹	۰/۰۰۰	تأیید
پردازش زبان طبیعی-> کشف شبکه‌های تقلب مالی	۰/۰۹۸	۲/۳۸۷	۰/۰۱۷	تأیید
مؤلفه‌های حاکمیت شرکتی -> کشف شبکه‌های تقلب مالی	۰/۱۴۰	۲/۳۲۸	۰/۰۲۰	تأیید
حاصل‌ضرب تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی در مؤلفه‌های حاکمیت شرکتی -> کشف شبکه‌های تقلب مالی	۰/۰۸۱	۲/۲۱۶	۰/۰۲۷	تأیید
حاصل‌ضرب پردازش زبان طبیعی در مؤلفه‌های حاکمیت شرکتی -> کشف شبکه‌های تقلب مالی	۰/۰۶۶	۳/۸۲۱	۰/۰۰۰	تأیید

نتایج آزمون فرضیات تحقیق در جدول ۶ ملاحظه می‌شود که در ذیل به بررسی نتایج آزمون فرضیات ارائه می‌شود:

فرضیهٔ اول: تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی تأثیر معناداری در کشف شبکه‌های تقلب مالی دارد.

همان‌طور که با توجه به جدول ۶ مشاهده می‌شود تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی به میزان ۰/۴۳۰ بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازهٔ ۱/۹۶± قرار دارد (t= ۸/۳۲۹)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ بنابراین با افزایش تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی میزان کشف شبکه‌های تقلب مالی افزایش می‌یابد.

فرضیهٔ دوم: پردازش زبان طبیعی قابلیت استخراج شاخص‌های رفتاری و متنی مؤثر در کشف شبکه‌های تقلب مالی را داراست.

همان‌طور که با توجه به جدول ۶ مشاهده می‌شود پردازش زبان طبیعی به میزان ۰/۰۹۸ بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازهٔ ۱/۹۶± قرار دارد (t= ۲/۳۸۷)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ بنابراین با افزایش پردازش زبان طبیعی میزان کشف شبکه‌های تقلب مالی افزایش می‌یابد.

فرضیهٔ سوم: مؤلفه‌های حاکمیت شرکتی نقش تعدیل گر معناداری در رابطه بین تحلیل گراف و کشف شبکه‌های تقلب مالی دارند؛ به‌گونه‌ای که افزایش قدرت حاکمیت شرکتی، اثر تحلیل گراف را افزایش می‌دهد.

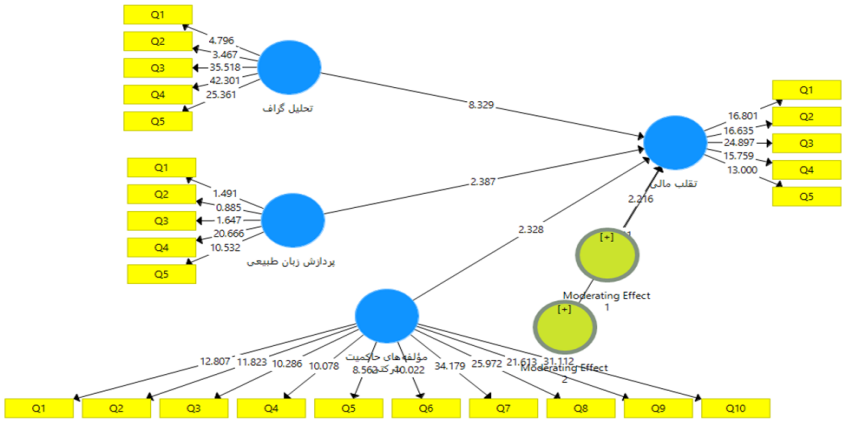
همان‌طور که با توجه به جدول ۶ مشاهده می‌شود تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی به میزان ۰/۴۳۰

بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t= ۸/۳۲۹)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود. مؤلفه‌های حاکمیت شرکتی به میزان ۰/۱۴۰ بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t= ۲/۳۲۸)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ و همچنین، حاصل‌ضرب تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی در مؤلفه‌های حاکمیت شرکتی به میزان ۰/۰۸۱ بر شناسایی الگوهای پیچیده و نامعمول مرتبط با تقلب‌های مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t= ۲/۲۱۶)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ بنابراین با افزایش قوت مؤلفه‌های حاکمیت شرکتی میزان تأثیر به‌کارگیری تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی بر شناسایی الگوهای پیچیده و نامعمول مرتبط با تقلب‌های مالی افزایش می‌یابد.

فرضیۀ چهارم: مؤلفه‌های حاکمیت شرکتی نقش تعدیل گر معناداری در رابطه بین پردازش زبان طبیعی و کشف شبکه‌های تقلب مالی دارند؛ به‌گونه‌ای که افزایش قدرت حاکمیت شرکتی، اثر شاخص‌های رفتاری و متنی کشف‌شده توسط پردازش زبان طبیعی را افزایش می‌دهد.

همان‌طور که با توجه به جدول ۶ مشاهده می‌شود پردازش زبان طبیعی به میزان ۰/۰۹۸ بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t= ۲/۳۸۷)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ بنابراین با افزایش پردازش زبان طبیعی میزان کشف شبکه‌های تقلب مالی افزایش می‌یابد. مؤلفه‌های حاکمیت شرکتی به میزان ۰/۱۴۰ بر کشف شبکه‌های تقلب مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t= ۲/۳۲۸)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ و همچنین، حاصل‌ضرب پردازش زبان طبیعی در مؤلفه‌های حاکمیت شرکتی به میزان ۰/۰۶۶ بر شناسایی الگوهای پیچیده و نامعمول مرتبط با تقلب‌های مالی تأثیر مثبت (مستقیم) دارد و با توجه به این‌که مقدار t خارج از بازۀ ۱/۹۶± قرار دارد (t=۳/۸۲۱)، این ضریب مسیر در سطح ۰/۰۵ معنی‌دار می‌باشد درنتیجه این فرضیه با اطمینان ۹۵٪ تأیید می‌شود؛ بنابراین با افزایش قوت مؤلفه‌های حاکمیت شرکتی میزان تأثیر پردازش زبان طبیعی بر شناسایی الگوهای پیچیده و نامعمول مرتبط با تقلب‌های مالی افزایش می‌یابد.

شکل ۲ ضرایب مسیر و مقادیر t مدل آزمون شده و جدول ۵ نتایج آزمون فرضیه‌ها را به‌طور خلاصه نشان می‌دهند. همان‌طور که قبلاً اشاره شد، به‌منظور بررسی معنادار بودن ضرایب مسیرها در مدل از آمارۀ t استفاده‌شده است. اگر t خارج از بازه ۱/۹۶± باشد، ضریب مسیر در سطح ۰/۰۵ و اگر t خارج از بازه ۲/۵۸± باشد، ضریب مسیر در سطح ۰/۰۱ معنی‌دار است.



شکل (۲): ضرایب مسیر و مقادیر t مدل آزمون شده

۶-نتیجه‌گیری و پیشنهاد

با گسترش روزافزون پیچیدگی‌های تقلب‌های مالی، سازمان‌ها و نهادهای مالی بیش‌ازپیش نیازمند بهره‌گیری از فناوری‌های پیشرفته هوش مصنوعی برای شناسایی ناهنجاری‌ها و پیشگیری از خسارات گسترده هستند. پژوهش حاضر باهدف بررسی تأثیر تحلیل گراف مبتنی بر شبکه‌های عصبی گرافی و پردازش زبان طبیعی (NLP) در کشف تقلب‌های مالی، همراه با نقش تعدیل گر مؤلفه‌های حاکمیت شرکتی انجام شد تا نشان دهد چگونه فناوری‌های هوشمند می‌توانند با ساختارهای حاکمیتی هم‌افزایی داشته و اثربخشی سیستم‌های کشف تقلب را افزایش دهند. یافته‌های پژوهش نشان داد که تحلیل گراف با استخراج روابط چندلایه و شناسایی الگوهای پنهان میان ذی‌نفعان، امکان کشف ساختارهای پیچیده و سازمان‌یافته تقلب را فراهم می‌آورد و داده‌های متنی نیز از طریق پردازش زبان طبیعی به منبع اصلی و مکمل در تحلیل‌های کشف تقلب تبدیل می‌شوند.

نتایج تحلیل نشان داد که مؤلفه‌های حاکمیت شرکتی، شامل شفافیت، پاسخگویی و کنترل‌های داخلی، نقش تعدیل گر مهمی در افزایش اثر فناوری‌های هوش مصنوعی ایفا می‌کنند؛ به‌طوری‌که در سازمان‌هایی با ساختار حاکمیتی قوی، دقت و قابلیت اعتماد خروجی الگوریتم‌ها افزایش‌یافته و در سازمان‌هایی با حاکمیت ضعیف، فناوری‌های هوشمند نقش جبرانی دارند. این یافته نشان می‌دهد که هم‌افزایی فناوری و ساختارهای حاکمیتی می‌تواند به بهبود کیفیت اطلاعات مالی، کاهش نویز داده‌ها و افزایش تمکین کارکنان به مقررات منجر شود. همچنین مشاهده شد که ترکیب الگوریتم‌های تقویتی، شبکه‌های عصبی گرافی، یادگیری ماشین وNLP ، بهترین عملکرد را در محیط‌های داده‌ای پیچیده و غیر ساختاریافته نشان داد و توانست الگوهای تقلب زنجیره‌ای و چندلایه را با دقت بالاتری شناسایی کند. تفسیر نتایج حاکی از آن است که فناوری‌های هوش مصنوعی به‌تنهایی نمی‌توانند جایگزین سیستم‌های کنترل و حاکمیت شرکتی شوند؛ بلکه کارکرد مؤثر آن‌ها در هماهنگی با ساختارهای

حاکمیتی و سیاست‌های نظارتی محقق می‌شود. به‌طور خاص، تحلیل گراف با توانایی کشف روابط پنهان و NLP با شناسایی الگوهای زبانی و رفتاری غیرعادی، می‌توانند داده‌های عددی و متنی را به‌صورت یکپارچه تحلیل کرده و امکان پیش‌بینی وقوع تقلب را افزایش دهند. یافته‌ها همچنین نشان دادند که در سازمان‌هایی با حاکمیت ضعیف، فناوری‌های هوشمند نقش جبرانی ایفا می‌کنند، اما در محیط‌هایی با نظارت فعال و سیستم‌های حسابرسی قوی، سهم فناوری‌های هوشمند در کشف تقلب کاهش می‌یابد که این امر اهمیت تطبیق فناوری با ویژگی‌های نهادی هر سازمان را برجسته می‌کند. مقایسه نتایج تحقیق حاضر با پژوهش‌های پیشین نشان می‌دهد که یافته‌ها با مطالعات تیان و همکاران (۲۰۲۳) و کیم و همکاران (۲۰۲۳) هم‌راستا هستند. همچنین، نتایج تحقیق حاضر با مطالعات وانگ و همکاران (۲۰۲۵) مطابقت دارد که بر اهمیت تحلیل دینامیک شبکه‌های تقلب در طول زمان تأکید کرده‌اند و نشان داده‌اند تحلیل گراف قادر است همدستی‌های پنهان و تغییرات پیچیده را بهتر آشکار سازد. در حوزه پردازش زبان طبیعی، یافته‌های حاضر با مطالعات چن و همکاران (۲۰۲۳)، خوزه و همکاران (۲۰۲۴)، سینگ و همکاران (۲۰۲۵) و توکلی و همکاران (۱۴۰۲) هم‌راستا هستند که نشان داده‌اند مدل‌های مبتنی بر ترنسفورمر و تکنیک‌های تحلیل احساسات می‌توانند الگوهای رفتاری و نشانه‌های هشداردهنده تقلب را به شکل دقیق‌تر شناسایی کنند.

بر اساس این نتایج، پیشنهاد می‌شود سازمان‌ها به‌صورت راهبردی بر تقویت مؤلفه‌های حاکمیت شرکتی شامل استقلال واحدهای نظارتی، تفکیک وظایف، ممیزی مستمر و طراحی رویه‌های دقیق کنترل تمرکز کنند تا زیرساخت‌های کنترلی مستحکمی فراهم شود که عملکرد بهینه فناوری‌های هوشمند را تضمین کند. بهره‌گیری از رویکردهای چندمدلی و ترکیبی در فناوری‌های هوش مصنوعی، به‌ویژه تلفیق تحلیل گراف و پردازش زبان طبیعی، برای پوشش یکپارچه ابعاد عددی، متنی و ساختاری داده‌ها توصیه می‌شود. ایجاد نظام داده محور مستحکم با استانداردسازی، اعتبارسنجی داده‌ها و آموزش تخصصی منابع انسانی، تدوین چارچوب‌های اخلاقی و نظارتی برای ارزیابی مستمر عملکرد الگوریتم‌ها و حمایت نهادهای ناظر از طریق مشوق‌ها و سیاست‌های حمایتی، از دیگر توصیه‌های کلیدی است. علاوه بر این، پژوهش‌های آتی می‌توانند به بررسی اثر فناوری‌های هوش مصنوعی در سازمان‌ها با ویژگی‌های نهادی و فرهنگی متفاوت، توسعه مدل‌های پیش‌بینی تقلب مبتنی بر داده‌های بین‌المللی، تحلیل نقش عوامل رفتاری و فرهنگی سازمان در تقویت عملکرد الگوریتم‌ها و کیفیت داده‌های ورودی و همچنین بررسی ترکیب‌های نوین الگوریتم‌های هوش مصنوعی و یادگیری ماشین برای بهبود دقت و قابلیت اعتماد مدل‌ها بپردازند.

باوجوداین، محدودیت‌هایی نیز وجود دارد؛ ازجمله محدودیت در نمونه و داده‌های مالی موردمطالعه که ممکن است تعمیم نتایج به سایر صنایع و کشورها را محدود کند، کیفیت و دسترسی محدود به داده‌های متنی و ساختاریافته که ممکن است بر دقت الگوریتم‌ها اثرگذار باشد و عدم بررسی کامل عوامل انسانی و رفتاری که می‌توانند بر عملکرد مدل‌های هوش مصنوعی تأثیرگذار باشند. درمجموع، یافته‌های این پژوهش تأکید دارد که هم‌افزایی بین فناوری‌های نوین و ساختارهای حاکمیتی قوی کلید موفقیت در کشف و پیشگیری از تقلب‌های

مالی پیچیده است و بهره‌گیری مستقل از فناوری‌های هوش مصنوعی بدون حمایت ساختارهای کنترلی، پاسخگوی چالش‌های این حوزه نخواهد بود.

فهرست منابع

ابری، ف.، و محمدی، س. (۱۴۰۳). *نقش هوش مصنوعی در شناسایی تقلب‌های مالی*. سومین کنفرانس ملی رویکردهای نوین در حسابداری، حسابرسی و مالی، علی‌آباد. بازیابی از https://civilica.com/doc/2284911

احمدی، س. ج.، فغانی ماکرانی، خ.، و فاضلی، ن. (۱۴۰۳). *تکنیک‌های داده‌کاوی و پیش‌بینی تقلب صورت‌های مالی*. فصلنامه علمی–پژوهشی دانش حسابداری و حسابرسی مدیریت، ۱۳(۴۲)، ۱۵–۲۸.

جبارزاده، س.، حیدری، ی.، عبدی، م.، و کاظمی علوم، م. (۱۳۹۹). *ضعف در کنترل‌های داخلی و احتمال تقلب در گزارشگری مالی*. فصلنامه بورس اوراق بهادار، ۵۲، ۱۶۲–۱۸۷.

حاجیان، ا. (۱۴۰۴). *طراحی مدلی مبتنی بر یادگیری ماشین برای کشف تقلب در صورت‌های مالی شرکت‌های پذیرفته‌شده در بورس تهران*. بیست‌وچهارمین کنفرانس بین‌المللی پژوهش در مدیریت، اقتصاد و توسعه. بازیابی از https://civilica.com/doc/2339633

خواجوی، ش.، و ابراهیمی، م. (۱۳۹۷). *بررسی تأثیر سازوکارهای حاکمیت شرکتی در تقلب در صورت‌های مالی شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران*. مجله مدیریت دارایی و تأمین مالی، ۶(۲)، ۷۱–۸۴. https://doi.org/10.22108/amf.2017.21385

داوری، ف.، و رضازاده، م. (۱۳۹۲). *مدیریت داده‌ها و تحلیل آماری با نرم‌افزار Smart PLS*. تهران: نگاه دانش.

زارعی، ع.، رهنمای رودپشتی، ف.، خان‌محمدی، م.، و کردلوئی، ح. (۱۴۰۵). *ارائه الگوی پیش‌بینی تقلب مبتنی‌بر هوش مصنوعی (SVM)*. دانش حسابداری و حسابرسی مدیریت، ۱۵(۵۹)، ۱۷۵–۱۸۶.

کاظمی، ت.، و پیری، پ. (۱۴۰۱). *پیش‌بینی طرح تقلب در گزارشگری مالی با استفاده از یادگیری ماشین در فضای چند کلاسه*. پژوهش‌های تجربی حسابداری، ۱۲(۴)، ۲۵۵–۲۸۰. https://doi.org/10.22051/jera.2022.41290.3040

ملکی کاکلر، ح.، بحری ثالث، ج.، جبارزاده کنگرلویی، س.، و آشتاب، ع. (۱۴۰۰). *کارایی مدل‌های آماری و الگوهای یادگیری ماشین در پیش‌بینی گزارشگری مالی متقلبانه*. اقتصاد مالی، ۱۵(۵۴)، ۲۶۷–۲۹۲.

هاشمی، س. س.، جعفری، س. م.، و نورالهزاده، ن. (۱۴۰۱). *ارائه الگوی اثرگذاری فرهنگ‌سازمانی بر نقش حسابرس در مبارزه با فساد*. دانش حسابداری و حسابرسی مدیریت، ۱۱(۲)، ۳۳۱–۳۵۵.

توکلی، م.، رحیمی، ا.، و نظری، ف. (۱۴۰۲). *نقش پردازش زبان طبیعی در استخراج بینش‌های کلیدی از گزارش‌های مالی*. فصلنامه حسابداری، امور مالی و هوش محاسباتی، ۱(۲)، ۱–۲۰. https://doi.org/10.12345/jaf.2025.123456

Abdullah, M., Ismail, L. H., & Smith, M. (2016). Fraud detection and prevention in the Malaysian public sector: Accounting practitioners perceptions. *Procedia Economics and Finance, 35*, 138–145.

Boulieris, P., Pavlopoulos, J., Xenos, A., & Vassalos, V. (2024). Fraud detection with natural language processing. *Machine Learning, 113*(5), 5087□5108. https://doi.org/10.1007/s10994-023-06354-5

Birol, B. (2019). Corporate governance and fraud detection: A study from Borsa Istanbul. *Eurasian Journal of Business and Management, 7*(1), 44□64. https://doi.org/10.15604/ejbm.2019.07.01.005

Chen, B., Zhang, J., Zhang, X., Dong, Y., Song, J., Zhang, P., ... Tang, J. (2022). GCCAD: Graph contrastive learning for anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*.

Cheng, D., Tu, Y., Ma, Z., Niu, Z., & Zhang, L. (2019). Risk assessment for networked-guarantee loans using high-order graph attention. *IJCAI-19*, 5822□5828.

Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. *Proceedings of the 29th ACM CIKM*, 315□324.

Faccia, A., McDonald, J., & George, B. (2023). NLP sentiment analysis and accounting transparency. *Computers, 13*(1), 5. https://doi.org/10.3390/computers13010005

Farayola, E. D., Olatoye, M., Mhlongo, T. O., & Oke, T. (2023). Forensic accounting in the digital age. *Finance & Accounting Research Journal, 5*(11), 342□360. https://doi.org/10.51594/farj.v5i11.614

Gao, Y., Wang, X., He, X., Liu, Z., Feng, H., & Zhang, Y. (2023). Alleviating structural distribution shift in graph anomaly detection. *WSDM 2023*, 357□365.

Gori, M., Monfardini, G., & Scarselli, F. (2005). A new model for learning in graph domains. *IJCNN 2005*, 729□734.

Hair, J. F., Hult, G. T., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Sage Publications.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using PLS. *International Marketing Review, 33*(3), 405□431.

Harl, M., Weinzierl, S., Stierle, M., & Matzner, M. (2020). Explainable predictive business process monitoring using gated graph neural networks. *Journal of Decision Systems*, 1□16.

Hu, B., Zhang, Z., Shi, C., Zhou, J., Li, X., & Qi, Y. (2019). Cash-out user detection with hierarchical attention. *AAAI Conference on Artificial Intelligence*, 946□953.

Hu, B., Zhang, Z., Zhou, J., Fang, J., Jia, Q., & Fang, Y., et al. (2020). Loan default analysis with multiplex graph learning. *CIKM 2020*, 2525□2532.

Islam, M. S., & Rahman, N. (2025). AI-driven fraud detection in financial institutions. *Journal of Computer Science and Technology Studies, 7*(1), 100□112.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305□360.

Jiang, J., Chen, J., Gu, T., Choo, K. K. R., Liu, C., Yu, M., et al. (2019). Anomaly detection with GCN. *MILCOM 2019*, 109□114.

Kim, Y. (2014). Convolutional neural networks for sentence classification. *EMNLP 2014*, 1746□1751.

Kim, H., Choi, J., & Whang, J. J. (2023). Dynamic relation-attentive GNNs for fraud detection. *ICDM MLoG Workshop*.

Koutoupis, A., Kyriakogkonas, P., Pazarskis, M., & Davidopoulos, L. (2021). Corporate governance and COVID-19. *Corporate Governance, 21*, 969□982.

Kubilay, E., Raiber, E., Spantig, L., Cahl□kov□, J., & Kaaria, L. (2023). Can you spot a scam? *Journal of Development Economics, 165*, 103147.

Lan, C., Mohammed, K., & Billy, E. (2025, March). *The role of NLP in identifying fraudulent activities in financial communication and documentation.*

Li, Y., Yang, T., & Liu, X. (2019). Detecting financial fraud using NER. *Expert Systems with Applications, 127*, 181□192.
Li, Q., He, Y., Xu, C., Wu, F., Gao, J., & Li, Z. (2022). Dual-augment GNN for fraud detection. *CIKM 2022*, 4188□4192.
Luo, J., Peng, C., & Zhang, X. (2020). Impact of CFO gender on corporate fraud. *Pacific-Basin Finance Journal, 63*, 101404.
Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q. Z., Xiong, H., & Akoglu, L. (2021). Graph anomaly detection with deep learning. *IEEE TKDE*.
Mandac□, P. E., & Kahyaoglu, S. B. (2012). Internal auditing and corporate governance in ERM. *M□DAV Journal*, 43□66.
Miko, N. U., & Kamardin, H. (2015). Impact of audit committee and audit quality on preventing earnings management. *Social and Behavioral Sciences, 172*, 651□657.
Modepalli, U. P. (2025). Network analytics for identifying fraud rings and systemic risk. *Journal of Information Systems Engineering and Management, 10*(58s), 616.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
Saias, J. (2025). Advances in NLP techniques for detecting message-based threats. *Electronics, 14*(13), 2551.
Wang, J., Zhang, S., & Xiao, Y. (2022). A review on GNN methods in financial applications. *Journal of Data Science, 20*(2), 111□134.
Witosari, D., & Bandi. (2025). Corporate governance and financial fraud in ASEAN. *Quarterly Reviews, 8*(4), 1□22.
Zhang, X., & Li, Y. (2020). Fraud detection in financial reports via deep text classification. *Information Processing & Management, 57*(4), 102207.
Zhang, Y., Liu, T., & Li, W. (2024). Corporate fraud detection using linguistic readability vectors. *International Review of Financial Analysis, 95*, 103405.
Zhou, Y., & Kapoor, G. (2021). Enhancing financial fraud detection with ML and NLP. *Computers & Security, 101*, 102123.
Du, K., Zhao, Y., Mao, R., Xing, F., & Cambria, E. (2025). Natural language processing in finance: A survey. *Information Fusion, 115*, 102755.

Financial Economics
Vol. (16) Issue (64) Winter 2027

Investigating the Role of Graph Analysis and Natural Language Processing in Detecting Financial Fraud Networks: The Moderating Role of Corporate Governance Mechanisms

Amir.Ansari
Department of Accounting, Yas.C., Islamic Azad University, Yasuj, Iran
Amir.ansari382@iau.ac.ir

Hashem Valipour
Department of Accounting, Fir.C., Islamic Azad University, Firuzabad, Iran
(Corresponding Author)
Email: Hashem. valipour @iau.ac.ir

Hamid Salehi
Department of Accounting, Fir.C., Islamic Azad University, Firuzabad, Iran.
Salehi.hamid@iau.ac.ir

Abstract

This study was undertaken to examine the influence of Graph Neural Network (GNN)-based graph analytics and Natural Language Processing (NLP) techniques on the detection of financial fraud networks, with specific attention to the moderating effect of corporate governance mechanisms. Employing an applied and descriptive-survey research design, data were acquired via a standardized and adapted questionnaire administered to a sample of 150 professionals and experts in finance, auditing, information technology, risk management, and financial oversight within Iranian public and private sector organizations. The questionnaire was distributed through a hybrid online and in-person approach. The research investigates the extent to which GNN-based graph analytics and NLP impact the identification of financial fraud networks, and further analyzes the moderating role of corporate governance elements in augmenting these effects. The findings reveal that GNN-based graph analytics exerts a statistically significant and positive influence on the detection of financial fraud networks. Furthermore, NLP demonstrates a significant enhancement in the extraction of salient behavioral and textual indicators pertinent to financial fraud detection.

Additionally, corporate governance mechanisms are shown to function as significant moderators in the relationships under investigation. The reinforcement of corporate governance structures amplifies the effect of GNN-based graph analytics on the identification of intricate financial fraud patterns. Similarly, corporate governance mechanisms moderate the impact of NLP on the detection of fraud-related patterns, such that the enhanced robustness of corporate governance correlates with an increased effect of NLP on financial fraud detection.

Keywords: Corporate Governance Mechanisms, Financial Fraud Detection, Natural Language Processing, Graph Analytics