

## بررسی سودمندی انتخاب متغیرهای پیش‌بین در پیش‌بینی نوع اظهار نظر حساب‌برسان

محمدحسین ستایش<sup>۱</sup>، مصطفی کاظم‌نژاد<sup>۲</sup>، غلامرضا رضایی<sup>۳</sup>، علی اصغر دهقانی سعدی<sup>۴</sup>

**چکیده:** در اغلب پژوهش‌های انجام‌شده، متغیرهای پیش‌بین بدون ضابطه و فقط براساس مطالعات گذشته انتخاب شده‌اند. فرایند انتخاب متغیرها را می‌توان به‌عنوان مرحله پیش‌پردازش برای حذف متغیرهای نامربوط و اضافه و انتخاب متغیرهای بهینه قبل از ایجاد مدل دانست. در این رابطه، پژوهش حاضر به بررسی سودمندی روش انتخاب متغیر مبتنی بر همبستگی برای پیش‌بینی نوع اظهار نظر حساب‌برسان شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران می‌پردازد. طبقه‌بندی‌کننده‌های این پژوهش، شبکه‌های عصبی مصنوعی و رگرسیون لجستیک است. به‌طور کلی، یافته‌های تجربی مربوط به بررسی ۱۲۱۴ مشاهده (سال - شرکت) در بازه زمانی ۱۳۸۶ تا ۱۳۹۳ نشان داد سودمندی استفاده از متغیرهای منتخب روش انتخاب متغیر همبستگی، در عملکرد پیش‌بینی نوع اظهار نظر حساب‌برسان است. به بیان دیگر، در صورت استفاده از متغیرهای منتخب این روش نسبت به استفاده از کلیه متغیرهای اولیه، میانگین دقت افزایش و خطای نوع اول و دوم کاهش خواهد یافت. افزون بر این، یافته‌های پژوهش حاکی از عملکرد مناسب و بهتر شبکه‌های عصبی نسبت به رگرسیون لجستیک است.

**واژه‌های کلیدی:** پیش‌بینی نوع اظهار نظر حساب‌برسان، رگرسیون لجستیک، روش انتخاب متغیر مبتنی بر همبستگی، شبکه‌های عصبی.

۱. استاد حسابداری، دانشگاه شیراز، شیراز، ایران

۲. دکتری حسابداری، دانشگاه شیراز، شیراز، ایران

۳. دانشجوی دکتری حسابداری، دانشگاه شیراز، شیراز، ایران

۴. کارشناس ارشد حسابداری، دانشگاه شیراز، شیراز، ایران

تاریخ دریافت مقاله: ۱۳۹۴/۱۱/۲۶

تاریخ پذیرش نهایی مقاله: ۱۳۹۵/۰۶/۱۵

نویسنده مسئول مقاله: محمدحسین ستایش

E-mail: setayesh@shirazu.ac.ir

### مقدمه

حسابرسی مستقل از طریق اعتباردهی به صورت‌های مالی، تضمین قابلیت اتکا و تأیید کیفیت اطلاعات مالی، موجب حمایت از حقوق تمام ذی‌نفعان شرکت می‌شود (اشبو و وارفیلد، ۲۰۰۳). محصول نهایی حسابرسی هر واحد تجاری، گزارشگری‌ای است که در آن حسابرسان نسبت به صورت‌های مالی صاحبکار اظهارنظر می‌کنند (نیکخواه آزاد، ۱۳۹۳). از سویی، یافته‌های پژوهش‌های گذشته حاکی از آن است که گزارش حسابرس، نوع و محتوای آن، دربردارنده بار اطلاعاتی و پیامدهای اقتصادی با اهمیتی است و می‌تواند بر قیمت یا بازده سهام (چن و زائو، ۲۰۰۰؛ شباهنگ و خاتمی ۱۳۷۷؛ حساس‌یگانه و یعقوبی‌منش، ۱۳۸۲)، پیش‌بینی ورشکستگی (شورورزی، بختیاری، زنده‌دل و اسماعیل‌زاده باغ‌سیاهی، ۱۳۹۰)، پاداش مدیران (دی‌آنجلو، ۱۹۸۱) و تصمیم‌گیری اعتباردهندگان در خصوص اعطای اعتبار (فیرس، ۱۹۸۰) مؤثر باشد. بنابراین، پیش‌بینی دقیق نوع و محتوای گزارش حسابرسی، اطلاعات سودمندی در تصمیم‌گیری استفاده‌کنندگان از صورت‌های مالی محسوب می‌شود.

در سال‌های اخیر، بهبود ظرفیت کسب داده‌ها، هزینه کمتر نگهداری و توسعه فناوری و گستردگی پایگاه‌های اطلاعاتی، به ایجاد مجموعه داده‌هایی با حجم زیاد اضافه‌بار اطلاعاتی منجر شده است. بسیاری از این اطلاعات و متغیرها، نامربوط و اضافه است که اندازه فضای جست‌وجو را افزایش می‌دهد و در نتیجه فرایند پردازش بیشتر اطلاعات را دشوار می‌کند. فضای ذخیره‌سازی و زمان محاسباتی زیادی برای داده‌هایی با حجم زیاد اضافه‌بار اطلاعاتی لازم است که این موضوع، انتخاب متغیرهای بهینه را اجتناب‌ناپذیر می‌کند (ورادراپا و رانگاراجان، ۲۰۱۰). مشکل اصلی در انتخاب مجموعه‌ای از متغیرهای پیش‌بین در پیش‌بینی‌های مالی، وجود تعداد زیاد متغیرهای پیش‌بین بالقوه قابل استفاده است (جاردین، ۲۰۱۰). از سویی، یکی از مشکلات پیش‌بینی نوع اظهارنظر حسابرسان، نداشتن توافق جامع در خصوص متغیرهای پیش‌بین است که ممکن است برخی از متغیرهای حاوی اطلاعات، پارازیت داشته باشند و عملکرد پیش‌بینی را تحت تأثیر قرار دهند (لیانگ، تسای و وو، ۲۰۱۵). از این رو، در اغلب پژوهش‌های انجام‌شده، متغیرهای پیش‌بین بدون ضابطه و فقط براساس مطالعات گذشته انتخاب شده‌اند که این روش معایب زیادی دارد. در این زمینه، استفاده از روش‌های نظام‌مند انتخاب متغیرهای پیش‌بین بهینه، مزیت دارد (جاردین، ۲۰۱۰). از فرایند انتخاب متغیرها می‌توان به‌عنوان مرحله پیش‌پردازش برای کاهش متغیرهای نامربوط و اضافه و انتخاب متغیرهای بهینه قبل از ایجاد مدل استفاده کرد. با توجه به اهمیت پیش‌بینی دقیق نوع اظهارنظر حسابرسان، نقش متغیرهای پیش‌بین بهینه و روش‌های انتخاب آنها در این زمینه حائز اهمیت است؛ زیرا انتخاب متغیرهای بهینه به‌عنوان

مرحله پیش‌پردازش (قبل از انجام پیش‌بینی)، یکی از مهم‌ترین مراحل فرایند داده‌کاوی است که افزون‌بر فیلتر کردن متغیرهای نامربوط از داده‌های اولیه، موجب بهبود عملکرد پیش‌بینی می‌شود (تسای، ۲۰۰۹). یافته‌های پژوهش لو (۲۰۱۰) نشان داد انتخاب متغیرهای پیش‌بین و روش‌های آن، نسبت به انتخاب مدل پیش‌بینی، تأثیر بیشتری بر میانگین دقت پیش‌بینی دارد. با این حال، در کنار اهمیت انتخاب متغیرهای پیش‌بین در پیش‌بینی، اغلب پژوهش‌های گذشته با انتخاب بدون ضابطه متغیرهای پیش‌بین اجرا شده‌اند (جاردين، ۲۰۱۰ و لیانگ و همکاران، ۲۰۱۵)؛ بنابراین، می‌توان مشکل موجود را در روش نادرست انتخاب متغیر برای پیش‌بینی نوع اظهارنظر حسابرسان شرکت‌ها دانست. بدین ترتیب، مسئله اصلی این پژوهش بررسی این موضوع است که آیا انتخاب متغیرهای پیش‌بین بهینه از بین متغیرهای اولیه، تأثیر مثبت و معناداری بر عملکرد پیش‌بینی نوع اظهارنظر حسابرسان شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران دارد؟ به بیان دیگر، آیا متغیرهای منتخب روش انتخاب متغیر مبتنی بر همبستگی، بهتر از متغیرهای اولیه نوع اظهار نظر حسابرسان را پیش‌بینی می‌کنند؟ با توجه به اهمیت انتخاب متغیرهای پیش‌بین، پژوهش حاضر به بررسی مزیت روش انتخاب متغیر مبتنی بر همبستگی بر عملکرد پیش‌بینی نوع اظهارنظر حسابرسان شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران می‌پردازد. به بیان دیگر، هدف پژوهش حاضر، مقایسه عملکرد پیش‌بینی نوع اظهارنظر حسابرسان با متغیرهای منتخب این روش انتخاب متغیر و کلیه متغیرهای اولیه است. در این رابطه، پیش‌بینی نوع اظهارنظر حسابرسان با استفاده از دو روش (طبقه‌بندی‌کننده) شبکه‌های عصبی مصنوعی و رگرسیون لجستیک انجام می‌شود.

با توجه به سازوکار پژوهش حاضر و متغیرهای موجود، این مقاله با پیشینه‌های نظری و تجربی مرتبط با موضوع و همچنین روش پژوهش و فرضیه‌های برگرفته از پیشینه نظری و تجربی پژوهش ادامه می‌یابد و پس از آن، نتایج آزمون فرضیه‌ها بیان می‌شود و سرانجام با توجه به نتایج آزمون فرضیه‌ها، پیشینه نظری و تجربی، نتیجه‌گیری شده و با ارائه پیشنهادها به پایان می‌رسد.

### **پیشینه نظری پژوهش**

مشکل اصلی در انتخاب مجموعه‌ای از متغیرهای پیش‌بین در پیش‌بینی‌های مالی، تعداد زیاد متغیرهای پیش‌بین بالقوه قابل استفاده است (جاردين، ۲۰۱۰). از سویی، یکی از مشکلات پیش‌بینی نوع اظهارنظر حسابرسان، نبود توافق جامع در خصوص متغیرهای پیش‌بین است که در نتیجه ممکن است برخی از متغیرها حاوی اطلاعات پارازیت‌دار باشند و عملکرد پیش‌بینی را تحت تأثیر قرار دهند (لیانگ و همکاران، ۲۰۱۵). در اغلب پژوهش‌های انجام‌شده، متغیرهای

پیش بین بدون ضابطه و فقط براساس مطالعات گذشته انتخاب شده‌اند که این روش معایب زیادی دارد؛ برای نمونه، فرض می‌شود متغیری که در پژوهش خاصی مناسب بوده، همواره برای پژوهش‌های دیگر نیز مناسب است. با وجود این، اغلب، زمانی که متغیرها در نمونه دیگر یا با روش متفاوتی استفاده می‌شوند، ویژگی‌های اولیه خود را از دست می‌دهند. در این زمینه، استفاده از روش‌های نظام‌مند انتخاب متغیرهای پیش‌بین بهینه، مزیت دارد (جاردين، ۲۰۱۰). فرایند انتخاب متغیرها را می‌توان به‌عنوان مرحله پیش‌پردازش برای کاهش متغیرهای نامربوط و اضافه و انتخاب متغیرهای بهینه قبل از ایجاد مدل استفاده کرد.

با توجه به اینکه عامل‌های بسیاری از جمله کیفیت داده‌ها، در عملکرد پیش‌بینی مؤثر است، اگر داده‌ها حاوی متغیرها و اطلاعات تکراری و نامربوط باشند یا حاوی اطلاعات پارازیت‌دار و نامطمئن باشند، اخذ دانش از آن داده‌ها مشکل خواهد شد (هال، ۲۰۰۰). افزون بر این، کاهش تعداد متغیرهای پیش‌بین نامربوط یا اضافه (تکراری)، علاوه بر کاهش زمان اجرای الگوریتم‌های آموزشی، به مفهوم عمومی‌تری منجر می‌شود. سایر مزیت‌های بالقوه انتخاب متغیرهای پیش‌بین شامل تسهیل درک و تجسم داده‌ها، کاهش الزامات اندازه‌گیری و ذخیره اطلاعات، مقاومت در برابر اضافه‌بار ابعاد و بهبود عملکرد پیش‌بینی و فراهم کردن بینش بهتری درباره مفهوم زیربنایی از طبقه‌بندی دنیای واقعی است (تسای، ۲۰۰۹).

### پیشینه تجربی پژوهش

اغلب پژوهش‌های اولیه در زمینه پیش‌بینی نوع اظهارنظر حسابرس، به کمک روش آماری رگرسیون لجستیک اجرا شده‌اند. کیسی، واتسون و واینارزج (۱۹۸۸) در پژوهشی به بررسی ارتباط بین ارقام مالی، متغیرهای مؤسسه حسابرسی و گزارش مشروط حسابرسی در شرکت‌های کوچک پرداختند. نتایج آنان نشان داد می‌توان بندهای شرط موجود در گزارش حسابرسی شرکت‌های کوچک را پیش‌بینی کرد. بر اساس نتیجه دیگر این پژوهش، اگر شرکتی گزارش مشروط دریافت کند، احتمال بیشتری وجود دارد که سال‌های بعد نیز گزارش مشروط دریافت کند. مگینا و تساک‌لنگانوس (۲۰۱۱) با استفاده از یازده نسبت مالی به پیش‌بینی نوع اظهارنظر حسابرسان در بورس اوراق بهادار آتن (یونان) پرداختند. یافته‌های این پژوهش حاکی از آن بود که رگرسیون لجستیک، توانایی مناسبی برای پیش‌بینی دارد.

با مرور زمان و با مشخص شدن معایب روش آماری رگرسیون لجستیک (از قبیل ماهیت خطی، وابستگی به توزیع آماری و از پیش مشخص کردن مدل) و مزیت‌های انواع روش‌های هوش مصنوعی (از قبیل شبکه‌های عصبی مصنوعی، درخت تصمیم، ماشین بردار پشتیبان و...)،

پژوهش‌های زیادی به پیش‌بینی نوع اظهارنظر حسابرسان با استفاده از این روش‌ها و مقایسه با رگرسیون لجستیک پرداختند. نتایج پژوهش کرکاس، اسپاتیس، نپولس و منولوپولس (۲۰۰۷) حاکی از عملکرد بهتر شبکه‌های عصبی نسبت به روش خطی بود. فرناندزگومز، گراسیالاگوس و سانچزسرانو (۲۰۱۶) نیز در پژوهشی با استفاده از داده‌های شرکت‌های پذیرفته‌شده در بورس اسپانیا به این نتیجه رسیدند که استفاده از شبکه عصبی بهترین عملکرد (دقت بیش از ۹۸ درصد) را دارد.

پژوهش‌های زیادی به پیش‌بینی نوع اظهارنظر حسابرسان با استفاده از سایر روش‌های هوش مصنوعی از قبیل ماشین بردار پشتیبان (دامپوس، گاگانیس و پاسیوراس، ۲۰۰۵)؛ شبکه‌های عصبی احتمالی (گاگانیس، پاسیوراس و دامپوس، ۲۰۰۷)؛ نزدیک‌ترین همسایگی (گاگانیس، پاسیوراس، اسپاتیس و زپونیدیس، ۲۰۰۷)؛ درخت تصمیم (کرکاس و همکاران، ۲۰۰۷)؛ شبکه‌های بیز (کرکاس و همکاران، ۲۰۰۷) پرداختند. نتایج این پژوهش‌ها نشان‌دهنده برتری روش‌های هوش مصنوعی نسبت به مدل خطی آماری است.

در بورس اوراق بهادار تهران نیز در چند پژوهش با استفاده از روش‌های آماری به پیش‌بینی نوع اظهارنظر حسابرسان پرداخته شده است. سجادی، فرازند، دستگیر و دهقان‌فر (۱۳۸۷) با بررسی عوامل مؤثر بر گزارش مشروط حسابرسی، به این نتیجه دست یافتند که نسبت جاری و نسبت حساب‌های دریافتی به دارایی، بر گزارش مشروط حسابرسی اثر دارد. افزون بر این، بین گزارش مشروط حسابرسی سال قبل و نوع مؤسسه حسابرسی، با گزارش مشروط سال جاری ارتباط معناداری وجود دارد. ستایش و جمالیان‌پور (۱۳۸۸) به بررسی رابطه بین نسبت‌های مالی و غیرمالی با نوع اظهارنظر حسابرسان پرداختند. نتایج بیان‌کننده آن بود که از میان تمام متغیرهای مالی و غیرمالی، نوع عملکرد بیشترین رابطه را با نوع اظهارنظر حسابرسان دارد.

پورحیدری و اعظمی (۱۳۸۹) به شناسایی نوع اظهارنظر حسابرسان با استفاده از شبکه‌های عصبی پرداختند. بر اساس نتایج آنان، شبکه عصبی عملکرد بهتری در شناسایی نوع گزارش حسابرسان دارد و رگرسیون لجستیک الگوی نامتوازنی در شناسایی انواع اظهارنظر حسابرسان ارائه می‌کند. امینی، محمدی و عباسی (۱۳۹۰) عوامل مؤثر بر صدور گزارش مشروط حسابرسی را با استفاده از شبکه‌های عصبی پرسپترون چندلایه، شناسایی کردند. نتایج پژوهش نشان داد نسبت سود پس از مالیات به فروش، مهم‌ترین عامل مؤثر بر صدور گزارش مشروط حسابرسی است. نتیجه پژوهش‌های ستایش، ابراهیمی، سیف و ساریخانی (۱۳۹۱) و باقرپور و لاشانی، ساعی، مشکانی و باقری (۱۳۹۲) نیز حاکی از عملکرد مناسب شبکه‌های عصبی است.

در پژوهش‌های زیادی با استفاده از سایر روش‌های هوش مصنوعی از قبیل ماشین بردار پشتیبان (ستایش و همکاران، ۱۳۹۱)؛ شبکه‌های عصبی احتمالی (حساس یگانه، تقوی فرد و محمدپور، ۱۳۹۳)؛ درخت تصمیم (ستایش و همکاران، ۱۳۹۱، باقرپور ولاشانی و همکاران، ۱۳۹۲ و ستایش، فتاحی نافچی، عباسپور و روستایی، ۱۳۹۳)؛ شبکه‌های بیز (ستایش و همکاران، ۱۳۹۳) به پیش‌بینی نوع اظهارنظر حسابرسان شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران پرداخته شده است. نتایج این پژوهش‌ها، به استثنای ستایش و همکارانش (۱۳۹۳)، نشان‌دهنده برتری روش‌های هوش مصنوعی نسبت به مدل خطی آماری است.

با بررسی پیشینه پژوهش، مشخص می‌شود هدف بیشتر پژوهش‌های اجراشده، انتخاب مدل‌های بهینه برای پیش‌بینی بوده و در پژوهش‌های اندکی به انتخاب متغیرهای بهینه پرداخته شده است. برای نمونه، کرکاس و همکارانش (۲۰۰۷) از تحلیل واریانس یک‌طرفه و گاگانیس و همکارانش (۲۰۰۷) از تحلیل مؤلفه‌های اصلی برای انتخاب متغیرهای بهینه استفاده کرده‌اند. در ایران نیز باقرپور ولاشانی و همکارانش (۱۳۹۲) و ستایش و همکارانش (۱۳۹۳) از تحلیل واریانس یک‌طرفه و حساس یگانه و همکارانش (۱۳۹۳) از تحلیل مؤلفه‌های اصلی برای انتخاب متغیرهای پیش‌بین بهینه در پیش‌بینی نوع اظهارنظر حسابرسان استفاده کردند. با وجود این، در این پژوهش‌ها عملکرد پیش‌بینی با استفاده از متغیرهای منتخب این روش‌ها و استفاده از کلیه متغیرهای پیش‌بین اولیه مقایسه نشده است. نتایج پژوهش‌های اجراشده در زمینه پیش‌بینی بحران مالی، حاکی از سودمندی روش‌های انتخاب متغیر در پیش‌بینی است (برای نمونه، جاردین، ۲۰۱۰؛ وانگ، ما و یانگ، ۲۰۱۴؛ لیانگ و همکاران، ۲۰۱۵). همچنین، پژوهشی مشاهده نشد که در آن از روش مبتنی بر همبستگی برای انتخاب متغیرهای بهینه در پیش‌بینی نوع اظهارنظر حسابرسان استفاده کرده باشد. این در حالی است که نتایج پژوهش‌های اجراشده در سایر زمینه‌ها (برای نمونه، تسای، ۲۰۰۹ و دلپیان، مارکنتونیو، لاقی و رنزی، ۲۰۱۵) در پیش‌بینی بحران مالی) حاکی از سودمندی روش مبتنی بر همبستگی در انتخاب متغیرهای پیش‌بین بهینه است. با توجه به کاستی مطالعات موجود در بورس اوراق بهادار تهران و تأکید پژوهش‌ها بر الگوهای پیش‌بینی و همچنین انتخاب متغیرها تنها بر اساس پژوهش‌های گذشته، پژوهش حاضر به بررسی سودمندی روش انتخاب متغیر مبتنی بر همبستگی در پیش‌بینی نوع اظهارنظر حسابرسان می‌پردازد.

با توجه به پیشینه‌های نظری و تجربی پژوهش، دو فرضیه زیر در خصوص سودمندی روش انتخاب متغیر مبتنی بر همبستگی، تدوین و آزمون شده است. شایان ذکر است که برتری معیارهای عملکرد پیش‌بینی در زمان استفاده از متغیرهای بهینه انتخاب‌شده روش مبتنی بر

همبستگی، نسبت به استفاده از ۳۵ متغیر اولیه و معنادار بودن این تفاوت از نظر آماری، بیان‌کننده سودمندی روش انتخاب متغیر است؛ زیرا در این صورت، افزون‌بر کاهش تعداد متغیرهای پیش‌بین، عملکرد پیش‌بینی بهبود خواهد یافت.

فرضیه اول: تفاوت معناداری بین عملکرد پیش‌بینی شبکه‌های عصبی در زمان استفاده از متغیرهای منتخب روش مبتنی بر همبستگی و استفاده از کلیه متغیرهای اولیه وجود دارد.

فرضیه دوم: تفاوت معناداری بین عملکرد پیش‌بینی رگرسیون لجستیک در زمان استفاده از متغیرهای منتخب روش مبتنی بر همبستگی و استفاده از کلیه متغیرهای اولیه وجود دارد.

### روش‌شناسی پژوهش

در پژوهش حاضر از روش خطی رگرسیون لجستیک و روش غیرخطی شبکه‌های عصبی برای پیش‌بینی نوع اظهارنظر حسابرسان استفاده شده است؛ همچنین با به‌کارگیری معیارهای ارزیابی (شامل میانگین دقت و خطاهای نوع اول و دوم) عملکرد پیش‌بینی ارزیابی می‌شود. منظور از دقت پیش‌بینی مدل، درصد نوع اظهارنظرهایی است که به‌درستی پیش‌بینی شده‌اند. در خطای نوع اول، ممکن است شرکتی با گزارش تعدیل‌شده، به اشتباه توسط مدل در زیرمجموعه شرکت‌هایی طبقه‌بندی شود که گزارش مقبولی دارند و در خطای نوع دوم، شرکتی که گزارش مقبول دارد به اشتباه در دسته شرکت‌هایی طبقه‌بندی می‌شود که گزارش تعدیل‌شده دارند (چن، ۲۰۱۱؛ ستایش و همکاران، ۱۳۹۱). دقت خطای نوع اول و دوم، براساس ماتریس احتمالات محاسبه می‌شود و این سه معیار به‌طور مستقیم از نرم‌افزار weka به‌دست می‌آید. به‌منظور ارزیابی روش انتخاب متغیر مبتنی بر همبستگی و آزمون فرضیه‌ها، معیارهای ارزیابی به‌دست‌آمده از پیش‌بینی با استفاده از متغیرهای منتخب این روش با معیارهای ارزیابی به‌دست‌آمده از اجرائیدن مرحله انتخاب متغیرهای پیش‌بین (پیش‌بینی با ۳۵ متغیر اولیه) مقایسه می‌شوند. انتخاب متغیرهای پیش‌بین و پیش‌بینی نوع اظهارنظر حسابرسان به‌وسیله نرم‌افزار Weka نسخه ۳-۷ انجام شده است. با اجرای آزمون t زوجی (بر اساس صد دقت در اجرای روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار در هر روش پیش‌بینی) در نسخه ۲۱ نرم‌افزار SPSS، فرضیه‌های پژوهش آزمون می‌شوند.

جامعه آماری پژوهش شامل تمام شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران است که از ابتدای سال ۱۳۸۶ تا پایان ۱۳۹۳ در بورس فعال بوده‌اند. در این پژوهش از نمونه‌گیری استفاده نمی‌شود، اما نمونه پژوهش براساس معیارهای زیر انتخاب شده است:

۱. به‌منظور افزایش قابلیت مقایسه، سال مالی شرکت‌ها منتهی به پایان اسفند باشد؛

۲. به دلیل ماهیت متفاوت فعالیت، جزء شرکت های سرمایه گذاری، بانک ها و مؤسسه های مالی، نباشد؛
۳. اطلاعات مالی شرکت برای اجرای این پژوهش، در دوره زمانی مد نظر به طور کامل ارائه شده باشد.

با توجه به بررسی های انجام شده، ۱۲۱۴ مشاهده (سال - شرکت)، حائز شرایط بالا بودند و برای اجرای پژوهش از آنها استفاده شد. از ۱۲۱۴ نمونه بررسی، گزارش حسابرسی ۳۲۷ مورد مقبول و مابقی تعدیل شده بود.

### روش انتخاب متغیر و طبقه بندی کننده های استفاده شده

به طور کلی متغیری مناسب است که به متغیر وابسته (طبقه، در طبقه بندی) ارتباط داشته باشد، اما نسبت به سایر متغیرهای پیش بین مربوط، اضافه یا زائد نباشد. اگر همبستگی بین دو متغیر، معیاری برای مناسب بودن در نظر گرفته شود، تعریف یاد شده به این صورت خواهد شد؛ متغیری مناسب است در صورتی که همبستگی بالایی با متغیر وابسته (طبقه) و همبستگی پایینی با سایر متغیرهای پیش بین داشته باشد. به بیان دیگر، اگر همبستگی بین یک متغیر پیش بین و متغیر وابسته (طبقه) به اندازه کافی بالا باشد تا برای پیش بینی متغیر وابسته (طبقه) مربوط باشد و همبستگی بین آن و سایر متغیرهای پیش بین مربوط، به سطح معینی نرسد، به طوری که نتوان به وسیله سایر متغیرهای مربوط متغیر یاد شده را پیش بینی کرد، آن متغیر، متغیر مناسبی برای پیش بینی (طبقه بندی) قلمداد خواهد شد. در این حالت، مسئله اصلی انتخاب متغیر، جست و جوی معیار مناسب همبستگی بین متغیرها و رویه منطقی برای انتخاب متغیرهای مناسب بر اساس این معیار خواهد بود (یو و لیو، ۲۰۰۳).

روش مبتنی بر همبستگی، همبستگی های بین متغیرهای پیش بین و همچنین متغیرهای پیش بین با متغیر وابسته را محاسبه می کند، سپس به جست و جوی فضای زیرمجموعه متغیرها می پردازد. زیرمجموعه یافت شده در جست و جو که بالاترین سودمندی (بر اساس رابطه ۱) را داشته باشد، برای کاهش ابعاد داده های آموزشی اولیه و داده های آزمایشی استفاده می شود (هال و اسمیت، ۱۹۹۹):

$$Merit_s = \frac{k \bar{r}_{cf}}{\sqrt{k + k(k-1) \bar{r}_{ff}}} \quad \text{رابطه ۱}$$

در رابطه ۱،  $Merit_s$  سودمندی اکتشافی زیرمجموعه متغیر  $S$  شامل  $K$  متغیر؛  $\bar{r}_{cf}$  میانگین همبستگی متغیر مستقل - وابسته و  $\bar{r}_{ff}$  میانگین همبستگی بین متغیر مستقل - متغیر مستقل



است. رابطه یادشده ضریب همبستگی پیرسون است که در آن تمام متغیرها استاندارد شده‌اند. صورت کسر، بیان‌کننده کیفیت (حدود) پیش‌بینی‌کنندگی گروهی از متغیرهای مستقل است و مخرج آن، میزان حشو (اضافه‌بودن) بین آنها را نشان می‌دهد.

طبقه‌بندی‌کننده‌ها، روش‌های پیش‌بینی را نشان می‌دهند که برای پیش‌بینی متغیرهای وابسته دووجهی (از قبیل ورشکسته درمقابل غیرورشکسته یا گزارش مقبول درمقابل تعدیل‌شده) استفاده می‌شوند. در این پژوهش از طبقه‌بندی‌کننده‌های شبکه‌های عصبی و رگرسیون لجستیک استفاده شده است. دلیل استفاده از دو طبقه‌بندی‌کننده، کسب نتایج تجربی برای فراهم‌آمدن امکان مقایسه است. طی دهه‌های اخیر، از شبکه‌های عصبی مصنوعی برای پیش‌بینی نوع اظهارنظر حساب‌رسان استفاده زیادی شده است. هر شبکه عصبی مصنوعی، شامل مجموعه‌ای از نرون‌های به هم متصل است که به هر مجموعه از این نرون‌ها لایه گفته می‌شود. نقش نرون‌ها در شبکه‌های عصبی، پردازش اطلاعات است. این کار در شبکه‌های عصبی مصنوعی به‌وسیله یک پردازشگر ریاضی که همان تابع فعال‌سازی است، انجام می‌شود. تابع فعال‌سازی، بر اساس نیاز خاص مسئله‌ای که قرار است به‌وسیله شبکه عصبی حل شود، از سوی طراح انتخاب می‌شود. ساده‌ترین شکل شبکه، فقط دو لایه دارد. لایه ورودی و لایه خروجی که عملکرد آن شبیه سیستم ورودی - خروجی است و از ارزش نرون‌های ورودی برای محاسبه ارزش نرون خروجی استفاده می‌کند (منهاج، ۱۳۷۷). در این پژوهش از شبکه‌های عصبی پرسپترون چندلایه با الگوریتم آموزشی پس‌انتشار خطا استفاده شده است. دلایل استفاده از شبکه‌های عصبی پرسپترون چندلایه با الگوریتم آموزشی پس‌انتشار خطا به شرح زیر است:

۱. در دهه‌های اخیر، شبکه‌های عصبی مصنوعی به‌طور وسیعی برای پیش‌بینی متغیرهای مالی استفاده شده‌اند (آل‌فارو، گارسیا، گامز و الیزاندو، ۲۰۰۸ و راوی کومار و راوی، ۲۰۰۷)؛ به‌طوری که در اغلب مطالعات تجاری کاربردی، از پرسپترون چندلایه استفاده شده و مشهورترین روش آموزشی نیز پس‌انتشار خطاست (تسای، ۲۰۰۹ و کرمی و بیک‌بشرویه، ۱۳۹۰).

۲. دستاورد پژوهش‌های تجربی بیان‌کننده دقت بالای این مدل هوش مصنوعی است. این مدل‌ها می‌توانند روابط غیرخطی را شناسایی کنند و عملکرد خوبی در صورت وجود اطلاعات پارازیت‌دار که اغلب در پیش‌بینی‌های مالی اتفاق می‌افتد، نشان می‌دهند (راوی کومار و راوی، ۲۰۰۷؛ آل‌فارو و همکاران، ۲۰۰۸ و پورحیدری و اعظمی، ۱۳۸۹).

رگرسیون لجستیک شبیه به رگرسیون معمولی است با این تفاوت که روش تخمین ضرایب در آنها یکسان نیست. رگرسیون لجستیک، به‌جای حداقل‌کردن مجذور خطاها - که در رگرسیون معمولی انجام می‌شود - احتمال رخداد واقعه را حداکثر می‌کند (مؤمنی و فعال‌قیومی، ۱۳۸۶). در

بیشتر پژوهش‌های اجرا شده در زمینه پیش‌بینی نوع اظهار نظر حساب‌برسان، برای ارزیابی مدل‌های مبتنی بر هوش مصنوعی، از نوعی مدل خطی استفاده می‌شود. با توجه به دلایل زیر، در این پژوهش از مدل خطی رگرسیون لجستیک برای مدل مقایسه‌ای استفاده شده است:

۱. بین مدل‌های آماری شهرت بیشتری دارد و در پژوهش‌های اجرا شده از همه بیشتر استفاده شده است (سعیدی و آقای، ۱۳۸۸)؛

۲. رگرسیون لجستیک نسبت به تحلیل ممیزی (مدل خطی دیگری که در پیش‌بینی استفاده زیادی از آن می‌شود) ابزار قوی‌تری است (مؤمنی و فعال قیومی، ۱۳۸۶).

### متغیرهای پژوهش

متغیر وابسته پژوهش حاضر، نوع اظهار نظر حساب‌برسان است. بر اساس استاندارد حسابرسی بخش ۷۰۰ (گزارشگری نسبت به صورت‌های مالی)، اظهار نظر حساب‌برسان به دو دسته تعدیل نشده (مقبول) و تعدیل شده (شامل بدون اظهار نظر، اظهار نظرهای مشروط و مردود) تقسیم می‌شود. بنابراین، در خصوص سنجش متغیر وابسته از متغیر دو وجهی استفاده شده است که در صورت دریافت گزارش تعدیل شده، عدد یک و در غیر این صورت عدد صفر به هر شرکت اختصاص داده می‌شود. این متغیر در پژوهش‌های ستایش و جمالیان پور (۱۳۸۸)، پورحیدری و اعظمی (۱۳۸۹) و عباس‌زاده و منظرزاده (۱۳۹۰) نیز با همین تعریف به کار رفته است.

برای انتخاب متغیرهای پیش‌بین (مستقل) در مرحله اول با بررسی ادبیات و پیشینه پژوهش (برای نمونه، کرکاس و همکاران، ۲۰۰۷؛ گانگیس و همکاران، ۲۰۰۷؛ ستایش و جمالیان پور، ۱۳۸۸؛ پورحیدری و اعظمی، ۱۳۸۹؛ باقرپور و لاشانی و همکاران، ۱۳۹۲ و ستایش و همکاران، ۱۳۹۳)، حدود ۶۰ متغیر پیش‌بین شناسایی شد. از بین متغیرهای شناسایی شده، ۳۵ متغیری که در ادبیات پیش‌بینی گزارش حسابرسی بیشتر استفاده شده و داده‌های لازم برای سنجش آنها از طریق پایگاه‌های اطلاعاتی سازمان بورس اوراق بهادار و نرم‌افزارهای تدبیر پرداز و ره‌آورد نوین در دسترس بود، انتخاب شد. جدول ۱، میانگین این متغیرها و نحوه سنجش هر یک را نمایش می‌دهد. متغیرهای نوع گزارش سال قبل، نوع عملکرد، دعاوی حقوقی و اندازه مؤسسه حسابرسی با استفاده از متغیرهای مجازی اندازه‌گیری شده‌اند؛ به این ترتیب که اگر گزارش حسابرسی سال قبل شرکت مقبول باشد عدد صفر و در غیر این صورت عدد یک به آن اختصاص یافته است. اگر شرکت سود ده باشد عدد صفر و در غیر این صورت عدد یک برای آن لحاظ شده است. در صورت وجود دعاوی حقوقی علیه شرکت عدد یک و در غیر این صورت عدد صفر به آن اختصاص یافته است. در نهایت، اگر حسابرسی شرکت را سازمان حسابرسی انجام داده باشد عدد یک و در صورتی که حسابرسی شرکت را سایر مؤسسه‌های عضو جامعه حسابداران رسمی انجام

داده باشند، عدد صفر اختصاص یافته است. اجرای روش انتخاب متغیر مبتنی بر همبستگی، به انتخاب شش متغیر بهینه (متغیرهای ردیف‌های ۱، ۳، ۶، ۲۰، ۲۹ و ۳۲ در جدول ۱) از بین ۳۵ متغیر اولیه منجر شد. شایان ذکر است که در این پژوهش، از داده‌های یک سال قبل شرکت‌ها برای پیش‌بینی نوع اظهار نظر حسابرس استفاده شده است.

جدول ۱. متغیرهای پیش‌بین، نحوه سنجش و میانگین آن

ردیف	متغیرهای پیش‌بین	نحوه سنجش	میانگین گزارش‌های مقبول	میانگین گزارش‌های تعدیل‌شده
۱	نوع گزارش حسابرسی سال قبل			
۲	نوع عملکرد			
۳	دعای حقوقی			
۴	اندازه مؤسسه حسابرسی			
۵	سن (عمر) شرکت برحسب سال	LD	۲۱/۴۲	۱۹/۳۴
۶	حاشیه سود ناخالص	GP / S	۰/۳۲	۰/۲۳
۷	حاشیه سود خالص	NI / S	۰/۲۷	۰/۱۰
۸	بازده حقوق صاحبان سهام	NI / SE	۰/۵۱	۰/۳۹
۹	بازده کل دارایی‌ها	NI / TA	۰/۱۵	۰/۱۱
۱۰	سود انباشته به مجموع دارایی‌ها	RE / TA	۰/۱۹	۰/۱۰
۱۱	نسبت آنی	(CA-Inv-PP) / CL	۰/۷۶	۰/۶۲
۱۲	نسبت جاری	CA / CL	۱/۱۸	۱/۰۷
۱۳	نسبت سرمایه در گردش به دارایی‌ها	(CA-CL) / TA	۰/۱۴	۰/۱۱
۱۴	نسبت حساب‌های دریافتی به کل دارایی‌ها	AR / TA	۰/۲۴	۰/۲۷
۱۵	لگاریتم طبیعی کل دارایی‌ها (اندازه شرکت)	Ln (TA)	۱۳/۲	۱۲/۸
۱۶	لگاریتم طبیعی فروش خالص	Ln (S)	۱۲/۹	۱۲/۷

## ادامه جدول ۱

ردیف	متغیرهای پیش‌بین	نحوه سنجش	میانگین گزارش‌های مقبول	میانگین گزارش‌های تعدیل‌شده
۱۷	لگاریتم طبیعی تعداد کارکنان	Ln (PN)	۶/۱۲	۵/۹۷
۱۸	نسبت بدهی (اهرم مالی)	TL / TA	-/۶۱	-/۶۹
۱۹	نسبت بدهی بلندمدت به حقوق صاحبان سهام	LTD / SE	-/۲۷	-/۳۳
۲۰	پوشش هزینه‌های مالی	EBIT / IE	۹۸/۵۴	۹۳/۳۱
۲۱	سرانه ذخیره مزایای پایان خدمت کارکنان	RP / PN	۲۷/۳۴	۲۴/۴۵
۲۲	سرانه دارایی‌ها	TA / PN	۹۲۱/۲۳	۸۷۵/۴۵
۲۳	سرانه فروش	S / PN	۷۷۴/۸۵	۶۵۷/۷۶
۲۴	سرانه سود خالص	NI / PN	۱۵۶/۳۷	۷۵/۶۴
۲۵	دوره وصول مطالبات	ART / 365	۱۳۱/۴۶	۱۴۲/۳۷
۲۶	گردش متوسط مجموع دارایی‌ها	S / TA <sub>(ave)</sub>	-/۷۵	-/۸۱
۲۷	گردش متوسط دارایی‌های ثابت	S / FA <sub>(ave)</sub>	۶/۵۸	۷/۶۲
۲۸	گردش متوسط حساب‌های دریافتی	S / AR <sub>(ave)</sub>	۴/۳۸	۴/۱۴
۲۹	گردش متوسط موجودی کالا	COGS/Inv <sub>(ave)</sub>	۱۰/۴۶	۴/۹۳
۳۰	وجه نقد به‌دست‌آمده از فعالیت‌های عملیاتی به فروش	OCF / S	-/۱۷	-/۱۲
۳۱	وجه نقد فعالیت‌های سرمایه‌گذاری به فروش	ICF / S	-/۱۲	-/۰۷
۳۲	ذخیره مالیات به فروش	TR / S	-/۰۳	-/۰۴
۳۳	سود (زیان) انباشته به فروش	RE / S	-/۱۷	-/۰۷
۳۴	رشد سود	%Δ (NI)	-/۲۶	-/۱۸
۳۵	رشد شرکت	%Δ (TA)	-/۲۸	-/۲۰

LD تاریخ درج؛ GP سود ناخالص؛ S فروش؛ NI سود خالص؛ SE حقوق صاحبان سهام؛ TA مجموع دارایی‌ها؛ RE سود انباشته؛ CA دارایی‌های جاری؛ Inv موجودی کالا؛ PP پیش‌پرداخت‌ها؛ CL بدهی‌های جاری؛ AR حساب‌ها و اسناد دریافتی؛ PN تعداد کارمند؛ TL مجموع بدهی‌ها؛ LTD بدهی‌های بلندمدت؛ EBIT سود قبل بهره و مالیات؛ IE هزینه‌های مالی؛ RP ذخیره مزایای پایان خدمت؛ ART گردش مطالبات؛ FA دارایی‌های ثابت؛ COGS بهای تمام‌شده کالای فروش‌رفته؛ OCF خالص جریان نقد عملیاتی؛ ICF خالص جریان نقدی فعالیت‌های سرمایه‌گذاری؛ TR ذخیره مالیات؛ %Δ درصد رشد سالانه.

### روایی متقابل

در روش holdout که در اغلب پژوهش‌های حسابداری و مالی (به‌ویژه در ایران) استفاده شده است، داده‌ها به دو دستهٔ مجموعهٔ آموزشی و مجموعهٔ آزمایشی تقسیم می‌شوند. این روش‌ها محدودیت‌های بارزی دارند. روش holdout تخمین‌گر بدبینانه‌ای است؛ زیرا تنها بخشی از داده‌ها برای آموزش به روش پیش‌بینی ارائه می‌شود. هر چه تعداد نمونهٔ بیشتری برای مجموعهٔ آزمایشی خارج شود، تعصب برآورد بیشتر می‌شود. از سویی، نمونه‌های آزمایشی کوچک‌تر (با تعداد کمتر) به معنای این است که فاصلهٔ اطمینان دقت بیشتر خواهد بود. بنابراین، روش یادشده روش مناسبی نیست (کوهاوی، ۱۹۹۵). در مقابل، روش روایی متقابل به دلیل سادگی، شفافیت و جامعیت، راهبرد متداول و مناسبی است و نتایج بسیاری از مطالعات پیشین، حاکی از عملکرد بهتر این روش است. در این رابطه، در پژوهش حاضر به منظور بررسی تعمیم‌پذیری پیش‌بینی‌های انجام‌شده، از روایی متقابل ۱۰ بخشی استفاده می‌شود. روایی متقابل ۱۰ بخشی برای برآورد درصد خطای واقعی کاملاً قابل اتکا و کافی است (هو، ۲۰۱۰). در این روش، نمونهٔ اصلی به ۱۰ نمونهٔ فرعی مختلف دسته‌بندی می‌شود. ۹ نمونهٔ فرعی برای نمونه‌های آموزشی استفاده می‌شود و نمونهٔ فرعی باقی‌مانده برای نمونهٔ آزمایشی، تحت آزمون قرار می‌گیرد. این شیوه تا جایی تکرار می‌شود که هر یک از ۱۰ نمونهٔ فرعی برای نمونهٔ آزمایشی تحت آزمون قرار گیرد. در این پژوهش، روایی متقابل ۱۰ بخشی با استفاده از اجزای مختلف مجموعهٔ داده‌ها، به طور مستقل ۱۰ بار انجام شده است (روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار). یافته‌های اغلب پژوهش‌ها (برای مثال کوهاوی، ۱۹۹۵) حاکی از آن است که در مسائل دنیای واقعی، روایی متقابل ۱۰ بخشی، بهترین روش انتخاب مدل است. مزیت روش یادشده این است که تمام نمونه‌ها در نهایت هم به عنوان داده‌های آموزشی و هم به عنوان داده‌های آزمایشی استفاده می‌شوند. افزون بر این، استفاده از روایی متقابل، از بروز مشکل بیش‌برازش و مشکلات مربوط به نتایج برون‌نمونه‌ای جلوگیری می‌کند.

### یافته‌های پژوهش

جدول ۲، میانگین دقت و خطای نوع اول و دوم پیش‌بینی نوع اظهارنظر حسابرسان را با استفاده از شبکه‌های عصبی در دو حالت (بر اساس ۶ متغیر منتخب روش مبتنی بر همبستگی و ۳۵ متغیر اولیه) نشان می‌دهد. در مرحلهٔ آموزش، ماتریسی که تعداد ستون‌های آن ۳۵ (به تعداد متغیرهای پیش‌بین) و تعداد سطرهای آن به اندازهٔ نمونهٔ پژوهش است، شکل می‌گیرد. این ماتریس مانند ورودی در پیش‌بینی عمل می‌کند. خروجی شبکهٔ عصبی نیز گره‌ای است که

بیان کننده نوع اظهار نظر حسابرس (به صورت یک متغیر دودویی) است. در مرحله آزمایش نیز با توجه به اجرای روایی متقابل، پس از آن که داده‌ها به ۱۰ قسمت دسته‌بندی شدند، از ۹۰ درصد داده‌ها برای آموزش استفاده می‌شود که در آن نوع اظهار نظر حسابرس نیز برای یادگیری به در اختیار شبکه عصبی قرار می‌گیرد؛ از ۱۰ درصد باقی داده‌ها نیز برای آزمایش استفاده می‌شود که در آن نوع اظهار نظر حسابرس توسط شبکه پیش‌بینی می‌شود. این شیوه در نرم‌افزار Weka تا جایی ادامه می‌یابد که هر یک از ۱۰ نمونه فرعی به عنوان نمونه آزمایشی تحت آزمون قرار گیرد. در این پژوهش، روایی متقابل ۱۰ بخشی با استفاده از اجزای مختلف مجموعه داده‌ها، به طور مستقل ۱۰ بار اجرا شده است. نرون‌های لایه اول به تعداد متغیرهای پیش‌بین است و لایه آخر فقط یک نرون (نوع اظهار نظر) دارد.

برای آزمون فرضیه اول و بررسی تفاوت معنادار بین میانگین معیارهای ارزیابی عملکرد پیش‌بینی در این دو حالت، آزمون  $t$  زوجی اجرا شده است. برای این کار، از دقت‌های روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار استفاده شده است که به ایجاد ۱۰۰ دقت در هر پیش‌بینی منجر می‌شود. با توجه به آماره  $t$  و مقدار احتمال مربوط به آن، تفاوت معناداری بین میانگین معیارهای عملکرد پیش‌بینی شبکه‌های عصبی در زمان استفاده از متغیرهای منتخب روش مبتنی بر همبستگی و استفاده از کلیه متغیرهای اولیه وجود دارد. بنابراین، با توجه به بهتر بودن معیارهای عملکرد در حالت استفاده از متغیرهای منتخب روش مبتنی بر همبستگی نسبت به استفاده از کل متغیرها و معنادار بودن آن از نظر آماری، می‌توان استنباط کرد که روش انتخاب متغیر مبتنی بر همبستگی، تأثیر مثبت و معناداری بر عملکرد پیش‌بینی شبکه‌های عصبی می‌گذارد؛ زیرا با اینکه تعداد متغیرهای پیش‌بین از ۳۵ به ۶ متغیر کاهش یافته، عملکرد پیش‌بینی بهتر شده است.

جدول ۲. عملکرد شبکه‌های عصبی بر اساس متغیرهای منتخب و متغیرهای اولیه

عملکرد	متغیرهای منتخب	متغیرهای اولیه	آماره $t$	مقدار احتمال
میانگین دقت	۰/۹۲	۰/۸۶	۵/۳۲۴	۰/۰۰۰
میانگین خطای نوع اول	۰/۰۵	۰/۱۲	۵/۶۵۷	۰/۰۰۰
میانگین خطای نوع دوم	۰/۱۰	۰/۱۶	۴/۵۶۳	۰/۰۰۰

جدول ۳ میانگین دقت، خطای نوع اول و دوم پیش‌بینی نوع اظهار نظر حسابرس را با استفاده از رگرسیون لجستیک در دو حالت (بر اساس ۶ متغیر منتخب روش مبتنی بر همبستگی و ۳۵ متغیر اولیه) نشان می‌دهد. برای آزمون فرضیه دوم و بررسی وجود تفاوت معنادار بین

میانگین معیارهای ارزیابی عملکرد پیش‌بینی در این دو حالت، آزمون  $t$  زوجی اجرا شده است. برای این کار از دقت‌های به‌دست‌آمده از روایی متقابل ۱۰ بخشی با ۱۰ بار تکرار استفاده شد که به ایجاد ۱۰۰ دقت در هر پیش‌بینی انجامید. با توجه به آماره  $t$  و مقدار احتمال مربوط به آن، تفاوت معناداری بین میانگین معیارهای عملکرد پیش‌بینی رگرسیون لجستیک در زمان استفاده از متغیرهای منتخب روش مبتنی بر همبستگی و استفاده از کلیه متغیرهای اولیه وجود دارد. بنابراین، با توجه به بهتر بودن معیارهای عملکرد در حالت استفاده از متغیرهای منتخب روش مبتنی بر همبستگی نسبت به استفاده از کل متغیرها و معنادار بودن آن از نظر آماری، می‌توان استنباط کرد که روش انتخاب متغیر مبتنی بر همبستگی، تأثیر مثبت و معناداری بر عملکرد پیش‌بینی رگرسیون لجستیک می‌گذارد؛ زیرا با اینکه تعداد متغیرهای پیش‌بین از ۳۵ به ۶ متغیر کاهش یافته، عملکرد پیش‌بینی بهتر شده است.

همچنین، با توجه به نتایج معیارهای عملکرد در جدول‌های ۲ و ۳، می‌توان گفت عملکرد پیش‌بینی شبکه‌های عصبی در کلیه موارد از رگرسیون لجستیک بهتر است.

جدول ۳. عملکرد رگرسیون لجستیک بر اساس متغیرهای منتخب و متغیرهای اولیه

عملکرد	متغیرهای منتخب	متغیرهای اولیه	آماره $t$	مقدار احتمال
میانگین دقت	۰/۸۵	۰/۸۱	۴/۴۵۱	۰/۰۰۰
میانگین خطای نوع اول	۰/۱۴	۰/۱۷	۳/۸۲۴	۰/۰۰۰
میانگین خطای نوع دوم	۰/۱۷	۰/۲۱	۴/۳۹۷	۰/۰۰۰

### نتیجه‌گیری

با توجه به محتوای اطلاعاتی گزارش حسابرس مستقل، پیش‌بینی نوع اظهارنظر در تصمیم‌گیری‌های مالی، اهمیت به‌سزایی دارد. مرحله انتخاب متغیرهای پیش‌بین یکی از مراحل است که باید قبل از پیش‌بینی انجام شود و گام مؤثری در انتخاب اطلاعات ارزشمندتر از بین اطلاعات وسیع به‌شمار می‌رود. به سخنی دیگر، هدف این مرحله، فیلتر کردن اطلاعات نامربوط یا اضافه است که می‌تواند عملکرد الگو را با کاهش تلاش برای آموزش بهبود بخشد. علی‌رغم اهمیت مرحله انتخاب متغیرهای پیش‌بین قبل از پیش‌بینی، این مرحله کمتر در ادبیات پیش‌بینی نوع اظهارنظر حسابرسان بررسی شده است و بیشتر مطالعات بدون توجه به انتخاب متغیرهای پیش‌بین، بر ایجاد مدل‌های پیش‌بینی اثربخش‌تر با قابلیت‌های پیش‌بینی بهتر تأکید داشته‌اند. از این رو، در پژوهش حاضر به بررسی سودمندی روش انتخاب متغیر مبتنی بر همبستگی پرداخته

شد. به منظور کسب نتایج تجربی برای فراهم شدن امکان مقایسه، از دو طبقه بندی کننده شبکه های عصبی و رگرسیون لجستیک استفاده شد.

به طور کلی، یافته های پژوهش حاکی از تأثیر مثبت استفاده از روش انتخاب متغیر مبتنی بر همبستگی (نسبت به استفاده نکردن از این روش ها و پیش بینی به کمک ۳۵ متغیر اولیه) بر عملکرد پیش بینی نوع اظهار نظر حسابرسان در هر دو طبقه کننده یاد شده است. به بیانی، در صورت استفاده از روش انتخاب متغیر مبتنی بر همبستگی، میانگین دقت افزایش یافته و خطای نوع اول و دوم کاهش می یابد. دلیل برتری معیارهای عملکرد در حالت انجام مرحله انتخاب متغیرها نسبت به انجام ندادن آن، مسئله اضافه بار ابعاد است. به نظر می رسد اضافه کردن متغیرهای بیشتر، پارازیت و در نتیجه خطا را افزایش می دهد؛ اما اضافه کردن متغیرها تا یک حد معین می تواند به بهبود پیش بینی کمک کند و اضافه کردن متغیرهای بیشتر، مسئله اضافه بار ابعاد را به وجود می آورد. افزون بر این، یافته های پژوهش حاکی از برتری طبقه بندی کننده شبکه های عصبی نسبت به رگرسیون لجستیک است؛ زیرا در زمان پیش بینی با این طبقه بندی کننده، ضمن آن که دقت پیش بینی افزایش یافت، خطای نوع اول و دوم کمتر از زمان پیش بینی با رگرسیون لجستیک بود. نتیجه این پژوهش مبنی بر عملکرد بهتر شبکه های عصبی مصنوعی نسبت به رگرسیون لجستیک با یافته های گاگانیس و همکاران (۲۰۰۷) و پورحیدری و اعظمی (۱۳۸۹)، سازگار است، اما با پژوهش ستایش و همکارانش (۱۳۹۳) همخوانی ندارد. افزون بر این، برتری شبکه های عصبی نسبت به روش خطی با یافته های حجازی، محمدی، اصلانی و آقاجانی (۱۳۹۱) در پیش بینی مدیریت سود، همسو است. نتیجه این پژوهش مبنی بر سودمندی روش های انتخاب متغیرهای پیش بین نیز با یافته های جاردین (۲۰۱۰) و لیانگ و همکارانش (۲۰۱۵) همخوانی دارد.

## References

- Abbaszadeh, M. R. & Manzarzadeh, H. (2011). The effect of Characteristics of Board of Director on Reporting of Independent Auditors and Accepting Companies which Was Accepted in Tehran Stock Exchange, *The Iranian Accounting and Auditing Review*, 18 (63), 95-112. (in Persian)
- Alfaro, E., García, N., Gámez, M. & Elizondo, D. (2008). Bankruptcy Forecasting: An Empirical Comparison of Adobos and Neural Networks, *Decision Support Systems*, 45 (1), 110-122.
- Amini, P., Mohammadi, K. & Abbasi, SH. (2011). The investigation of effecting Factors on the issue of qualified auditing report: Application of neural



- network method, *Journal of Management Accounting*, 4(11), 25-39. (in Persian)
- Ashbaugh, H. & Warfield, T.D. (2003). Audits as A Corporate Governance Mechanism: Evidence from The German Market. *Journal of International Accounting Research*, 2 (1), 1-21.
- Bagherpoor Valashani, M., Saei, M., Meshkani, A. & Bagheri, M. (2013). Prediction of Independent Auditor Opinion in Iran: Data Mining Approach, *Accounting and Auditing Research*, 5 (19), 134-150. (in Persian)
- Chen, C. P. & Zhao, R. (2000). An emerging market's reaction to initial modified audit opinions: Evidence for the shanghai stock exchange. *Contemporary Accounting Research*, 17 (3), 429-555.
- Chen, M.Y. (2011). Predicting Corporate Financial Distress based on Integration of Decision Tree Classification and Logistic Regression. *Expert Systems with Applications*, 38(9), 11261-11272.
- DeAngelo, L. (1981). Auditor Size and Auditor Quality. *Journal of Accounting and Economics*, 3 (3), 183-199.
- Dellepiane, U., Marcantonio, M., Laghi, E. & Renzi, S. (2015). Bankruptcy Prediction Using Support Vector Machines and Feature Selection during the Recent Financial Crisis. *International Journal of Economics and Finance*, 7 (8), 182-195.
- Doumpos, M., Gaganis, C. & Pasiouras, F. (2005). Explaining qualifications in audit reports using a support vector machine methodology. *Intelligent Systems in Accounting, Finance and Management*, 13 (4), 197-215.
- Fernandez-GaMez, M.A., Garcia-Lagos, F. & Sanchez-Serrano, J.R. (2016). Integrating Corporate Governance and Financial Variables for the Identification of Qualified Audit Opinions with Neural Networks. *Neural Computing and Applications*, 27(5), 1-18.
- Firth, M. (1980). A note on the impact of audit qualifications on lending and credit decisions. *Journal of Banking & Finance*, 4 (3), 257-267.
- Gaganis, Ch., Pasiouras, F. & Doumpos, M. (2007). Probabilistic Neural Networks for the Identification of Qualified Audit Opinions, *Expert Systems with Applications*, 32 (1), 114-124.
- Gaganis, C., Pasiouras, F., Spathis, C. & Zopounidis, C. (2007). A Comparison of Nearest Neighbours, Discriminant and Logit Models for Auditing Decisions, *Intelligent Systems in Accounting, Finance and Management*, 15 (1-2), 23-40.

- Hall, M.A. & Smith, L.A. (1999). Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper. *Proceedings of the Twelfth International Florida Artificial Intelligence Research Society Conference*, May 01 – 05, 235-239.
- Hall, M. A. (2000). Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning. *In Proceedings of the Seventeenth International Conference on Machine Learning*, (June 29 - July 02), 359-366.
- Hasas Yeghane, Y. & Yaghobi Manesh, S. (2003). The impact of audit reports on stock price. *Accounting Empirical Studies*, 1(3), 27-58. (in Persian)
- Hasas Yeghane, Y., Taghavifard, M. & Mohammadpoor, F. (2014). Using Probabilistic Neural Networks to Identifying. *Auditing: Theory and Practice*, 1 (1), 131-159. (in Persian)
- Hejazi, R., Mohamadi, Sh., Aslani, Z. & Aghajan, M. (2012). Earnings Management Prediction Using Neural Networks and Decision Tree in TSE. *The Iranian Accounting and Auditing Review*, 19 (68), 31-46. (in Persian)
- Hu, Y. C. (2010). Analytic Network Process for Pattern Classification Problems Using Genetic Algorithms. *Information Sciences*, 180(13), 2528–2539.
- Jardin, P. (2010). Predicting Bankruptcy Using Neural Networks and Other Classification Methods: The Influence of Variable Selection Techniques on Model Accuracy. *Neurocomputing*, 73 (10-12), 2047-2060.
- Karami, Gh. & Beik Boshroye, S. (2011). Corporate Governance and Equity Valuation: The Model by Using Artificial Neural Network. *The Iranian Accounting and Auditing Review*, 18 (64), 129-150. (in Persian)
- Keasey, K., Watson, R. & Wynarczyk, P. (1988). The small company audit qualification: a preliminary investigation. *Accounting and Business Research*, 18 (72), 323-333.
- Kirkos, E., Spathis, C., Nanopoulos, A. & Manolopoulos, Y. (2007). Identifying Qualified Auditors opinion: A Data Mining Approach. *Journal of Emerging technologies in Accounting*, 4(1), 183-197.
- Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence*, 1137-1143.
- Liang, D., Tsai, C.H. & Wu, H.T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems*, 73(1), 289-297.
- Lo, S.C. (2010). The Effects of Feature Selection and Model Selection on the Correctness of Classification. *Proceedings of the 2010 IEEE IEEM*, 989-993.

- Maggina, A. & Tsaklanganos, A. A. (2011). Predicting Audit Opinions Evidence from the Athens Stock Exchange. *The Journal of Applied Business Research*, 27 (4), 53-68.
- Mahdavi, Gh. & Ghayori Moghadam, A. (2010). Investigating the information content of qualified audit reports in companies listed on Tehran Stock Exchange. *Accounting and Auditing Research*, 2(6), 65-85. (in Persian)
- Menhaj, M. B. (1998). *Foundations of Neural Networks*. Tehran: Publication Center of Professor Hesabi. (in Persian)
- Momeni, M. & Faal Ghayomi, A. (2007). *Statistical Analysis Using SPSS* (1<sup>th</sup> ed). Tehran: New Book Publication. (in Persian)
- Moradi, M. & Fakhrabadi, A. (2010). Evaluation of Cultural Factors Effect on Auditors Evaluation of Internal Control and Assess Control Risk, *Financial Accounting Research*, 1 (1-2), 89-102. (in Persian)
- Nikkhah Azad, A. (2014). *Audit Fundamental Concepts Statement* (3<sup>th</sup> ed), Tehran: Audit Organization. (in Persian)
- Pourheidari, O. & Aazami, Z. (2010). Predicting Auditor's Opinions: A Neural Networks Approach, *Accounting Knowledge*, 1(3), 77-97. (in Persian)
- Ravi Kumar, P. & Ravi, V. (2007). Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques - A Review, *European Journal of Operational Research*, 180 (1), 1-28.
- Saeedi, A. & Aghaie, A. (2010). Predicting Financial Distress of firms Listed in Tehran Stock Exchange Using Bayesian networks. *The Iranian Accounting and Auditing Review*, 16 (3), 59-78. (in Persian)
- Sajadi, H., Farazmand, H., Dastgir, M. & Dehghanfar, D. (2008). Factors affecting on qualified audit reports, *Accounting Empirical Studies*, (18), 123-145. (in Persian)
- Setayesh, M. & Jamalianpoor, M. (2009). The Investigating relationship between financial ratios and non-financial variables with the auditor's opinion. *Accounting and Auditing Research*, 1(2), 130-157. (in Persian)
- Setayesh, M., Ebrahimi, F., Seif, M. & Sarikhani, M. (2013). Forecasting the Type of Audit Opinions: A Data Mining Approach. *Journal of Management Accounting*, 5 (4), 69-82. (in Persian)
- Setayesh, M., Fatahi Nafchi, H., Abbaspoor, S. & Roustaei, M. (2014). Offering a new approach to the issuance of audit reports by using data mining (Case Study: Companies Listed in Tehran Stock Exchange). *Journal of Audit Science*, 14 (57), 5-26. (in Persian)

- Shabahang, R. & Khatami, M. (1999). The impact of audit requirements on share price and the analysis of annual financial statements by officials in Tehran Stock Exchange. *Management and Economics*, 3 (40), 23-50. (in Persian)
- Shourvarzi, M. R., Bakhtiyari, M., ZendeDel, A. & Esmailzadeh, M. (2011). Comparison of independent auditors and financial variables in predicting bankruptcy. *The Iranian Accounting and Auditing Review*, 18(65), 63-78. (in Persian)
- Tsai, C. (2009). Feature Selection in Bankruptcy Prediction. *Knowledge Based Systems*, 22 (2), 120-127.
- Veerabhadrapa, J. & Rangarajan, L. (2010). Multi-Level Dimensionality Reduction Methods Using Feature Selection and Feature Extraction. *International Journal of Artificial Intelligence and Applications*, 1(4), 54-68.
- Wang, G. Ma, J. & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41 (5), 2353-2361.
- Yu, L. & Liu, H. (2003). Feature Selection for High-Dimensional Data: A Fast Correlation-based Filter Solution. *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC, 856-863.

Archive of SID