

تشخیص مدیریت سود صنعت خودرو و ساخت قطعات با استفاده از ترکیب شبکه بیزین و درخت تصمیم C5.0

مرضیه سعید مقدم *

داریوش جاوید **

محمود همت فر ***

تاریخ دریافت: ۱۳۹۶/۰۴/۱۱

تاریخ پذیرش: ۱۳۹۶/۱۱/۱۰

چکیده

سرمایه‌گذاران، تحلیل‌گران مالی و سایر ذینفعان به رقم سود گزارش شده در صورت‌های مالی، به عنوان یکی از فاکتورهای مهم تصمیم‌گیری توجه ویژه‌ای دارند، از این رو تشخیص مدیریت سود موجود در صورت‌های مالی از اهمیت خاصی برخوردار است. هدف اصلی این پژوهش بررسی اثربخشی، مدلی برای تشخیص سطح مدیریت سود با استفاده از ترکیب دو تکنیک داده‌کاوی شبکه بیزین و درخت تصمیم C5.0 می‌باشد. برای این منظور از ۲۴ متغیر تأثیرگذار بر مدیریت سود به عنوان متغیرهای مستقل و اقلام تعهدی اختیاری به عنوان متغیر وابسته و جایگزین مدیریت سود، استفاده شده است. به منظور محاسبه اقلام تعهدی اختیاری (مدیریت سود)، از روش توصیه شده توسط دیچو و همکاران (۲۰۱۲)، استفاده گردید. جامعه آماری پژوهش شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، فعال در صنعت خودرو و ساخت قطعات، طی سال‌های ۱۳۸۸ الی ۱۳۹۳ می‌باشد. یافته‌های پژوهش حاکی از آن است که مدل ترکیبی شبکه بیزین با درخت تصمیم C5.0، توانایی بالایی، برای تشخیص سطح مدیریت سود، در صنعت خودرو و ساخت قطعات دارد. بااهمیت‌ترین متغیر شناسایی شده توسط مدل ترکیبی، جهت

*کارشناسی ارشد حسابداری، گروه حسابداری، واحد بروجرد، دانشگاه آزاد اسلامی، بروجرد، ایران

Email: msaeedmoghadam@yahoo.com

**استادیار حسابداری، گروه حسابداری، واحد بروجرد، دانشگاه آزاد اسلامی، بروجرد، ایران (نویسنده مسئول)

Email: darush.javid@yahoo.com

***دانشیار مدیریت مالی، گروه حسابداری، واحد بروجرد، دانشگاه آزاد اسلامی، بروجرد، ایران

Email: dr.hematfar@yahoo.com

تشخیص سطح مدیریت سود در صنعت خودرو و ساخت قطعات، متغیر آستانه عملکرد با نسبت اهمیت ۱۰۰ درصد می‌باشد.

واژه‌های کلیدی: اقلام تعهدی، داده‌کاوی، درخت تصمیم C5.0، شبکه بیزین، مدیریت سود.

۱- مقدمه

اهداف حسابداری برگرفته از نیازها و خواسته‌های اطلاعاتی استفاده‌کنندگان است و از اهداف بنیادین حسابداری بیان وضعیت مالی و عملکرد واحد تجاری برای استفاده‌کنندگان برون سازمانی جهت کمک به آن‌ها در تصمیم‌گیری‌های مالی و سرمایه‌گذاری است، ابزار اصلی انتقال اطلاعات به این اشخاص گزارش‌های مالی می‌باشد (مشایخ، اربابی و رحیمی فر، ۱۳۹۲). صورت سود و زیان منعکس‌کننده عملکرد واحد تجاری در یک دوره مالی است و از برترین شاخص‌های اندازه‌گیری فعالیت‌های یک واحد اقتصادی سود است (غفارپور، ۱۳۹۱) و با توجه به اینکه استفاده‌کنندگان از صورت‌های مالی یک شرکت از جمله سرمایه‌گذاران، بستانکاران، تحلیل‌گران و مشتریان نمی‌توانند به طور مستقیم از عملکرد شرکت آگاه شوند بسیاری از آن‌ها به سود شرکت‌های بزرگ به عنوان یک شاخص بزرگ توجه می‌کنند (چن، چی و وانگ^۱، ۲۰۱۵). در نتیجه مدیران واحد تجاری به دلایل مختلف می‌کوشند تا از طریق کاربرد روش‌های گوناگون حسابداری، سود دوره مالی را در راستای برآورده کردن، سیاست‌ها و اهداف خود تغییر دهند (غفارپور، ۱۳۹۱). لو^۲ (۲۰۰۷) از سرمایه‌گذاران، اعتباردهندگان، قانون‌گذاران، مشتریان و رقبا به عنوان قربانیان مدیریت سود یاد می‌کند، زیرا ممکن است آن‌ها بر پایه اطلاعات دستکاری شده توسط مدیران، اقدام به تصمیم‌گیری کنند و تصمیمات نادرستی را اتخاذ نمایند (کامیابی و نورعلی، ۱۳۹۵). تاریخ حسابداری از وجود وقایع نامطلوبی از دستکاری گزارشات مالی توسط مدیریت خبر می‌دهد. وقوع دو بحران عظیم در بازار سهام آمریکا ناشی از گزارش سودهای ساختگی شرکت انرون و وردکام باعث شد که کیفیت اطلاعات حسابداری مورد تردید واقع شود (رستمی، ۱۳۹۰)؛ بنابراین تشخیص مدیریت سود موجود در صورت‌های مالی اهمیت دارد. مدیریت سود^۳ به برخی از روش‌ها یا روش‌های استفاده شده توسط مدیریت به منظور به دست آوردن سود حسابداری برای رسیدن به اهداف مورد انتظار اشاره دارد (دویل، جنینگ و سلیمان^۴، ۲۰۱۳). مدیریت سود فراتر از یک مرز معین احتمالاً می‌تواند به سوء مدیریت بر سود

1 Chen, Chi and Wang

2 Lo

3 Earnings Management

4 Doyle, Jennings and Soliman

منجر شود. در گذشته اکثر مطالعات مربوط به مدیریت سود به کاوش ارتباط مدیریت سود با موارد دیگر پرداخته‌اند در حال حاضر مطالعات کمتری برای تشخیص سطح مدیریت سود انجام شده است. از این رو به منظور کمک به سهامداران و جلوگیری از متحمل شدن به یک زیان بزرگ در بازار سهام که در نتیجه مدیریت سود، مدیر است. ارائه یک روش جدید که توانایی تشخیص و پیش‌بینی سطح مدیریت سود را داشته باشد ضروری به نظر می‌رسد. علاوه، تعیین راهکار و دستیابی به مدلی برای پیش‌بینی میزان و سطح مدیریت سود، برای حساب‌رسان در کمک به تشخیص سود دستکاری شده، می‌تواند بسیار مفید باشد. از آنجایی که ممکن است شرایط حاکم در صنایع مختلف متفاوت باشد و خصوصیات هر صنعت نیز خود می‌تواند عاملی اثرگذار بر مدیریت سود باشد لذا بررسی مدیریت سود به تفکیک صنعت و تعمیم نتایج آن به صنعت مربوطه، نیز می‌تواند بسیار سودمند باشد. اکثر شرکت‌های فعال در صنعت خودرو و ساخت قطعات جزء شرکت‌های بزرگ بررسی می‌باشد و گروه خودرو و ساخت قطعات یکی از گروه‌های بزرگ و فعال بازار سرمایه است؛ که در جهت‌گیری و حرکت بازار بورس بسیار تأثیرگذار و مهم است (اعطاسی، ۱۳۹۴) و از آنجایی که صنعت خودرو در کشور ما محرک رشد صنایع دیگر است به نحوی که ۶۰ رشته صنعتی دیگر با صنعت خودرو مرتبط هستند و صنعت خودرو سهم ۲۲ درصدی را در صنعت کشور دارا است (دهدشتی نژاد، ۱۳۹۲) و در نتیجه بیشتر مورد توجه سرمایه‌گذاران و سهامداران و... است. اهمیت صنعت خودرو و ساخت قطعات باعث انتخاب آن برای پیاده‌سازی مدل شد. داده‌کاوی و به‌کارگیری آن، یکی از فناوری‌هایی است که در عصر اطلاعات می‌تواند استفاده کنندگان از اطلاعات مالی، اعتباردهندگان، تحلیل‌گران مالی و سایر ذینفعان را در امر تصمیم‌گیری یاری رساند (کیقبادی، فتحی و سیف، ۱۳۹۴). درواقع هدف از داده‌کاوی ایجاد مدل‌هایی برای تصمیم‌گیری است، این مدل‌ها رفتارهای آینده را بر اساس تحلیل‌های گذشته پیش‌بینی می‌کنند (سروش و بحرینی نژاد، ۱۳۸۸). یکی از فرایندهایی که می‌تواند برای پیش‌بینی و تشخیص سطح مدیریت سود مورد استفاده قرار گیرد روش داده‌کاوی است. از این رو، هدف اصلی پژوهش بررسی اثربخشی مدلی، برای تشخیص و پیش‌بینی سطح مدیریت سود با استفاده از تکنیک‌های داده‌کاوی شبکه بیزین^۱ و درخت تصمیم^۲ می‌باشد. با توجه به هدف فوق این پژوهش سعی دارد یک مدل ترکیبی جدید را برای پیش‌بینی و تشخیص سطح مدیریت سود پیشنهاد کند. معرفی مدل‌های ترکیبی، می‌تواند گستره‌ای نوین از تحقیقات داده‌کاوی باشد. مدل ترکیبی این پژوهش، از شبکه بیزین برای

1 Bayesian Network

2 Decision Tree

غربالگری متغیرها، به منظور شناسایی متغیر مستقل بااهمیت استفاده می‌کند که این می‌تواند معرف یکی از کاربردهای شبکه بیزین در حوزه حسابداری باشد و سپس با استفاده از درخت تصمیم C5.0، قواعد معنی‌داری برای مدیریت سود شناسایی می‌گردد. پژوهش حاضر، اولین پژوهش داخلی می‌باشد که در شناسایی و تشخیص سطح مدیریت سود از تکنیک شبکه بیزین استفاده می‌کند. در ادامه مبانی نظری و پیشینه پژوهش، روش‌شناسی و طراحی، اجرا و یافته‌های پژوهش ارائه گردیده و در پایان نیز نتایج پژوهش ارائه شده است.

۲- مبانی نظری پژوهش

امروزه، موضوع مدیریت سود، بیشتر مورد توجه پژوهشگران قرار گرفته است. در دهه‌های اخیر پژوهش‌های زیادی در خصوص مدیریت سود انجام گرفته و ادبیات بسیاری از آن‌ها در مورد انگیزه‌های مدیران، اهداف و ابزارهای آن بوده است (حجازی، محمدی، اصلانی و آقاجانی، ۱۳۹۱). مدیریت سود به این معناست که مدیران شرکت با استفاده از انعطاف‌پذیری که در روش‌ها و استانداردهای حسابداری وجود دارد روش‌هایی را انتخاب کنند که سودهای گزارش شده را با سطح مطلوبی که مورد نظر است مطابقت دهند (ولی زاده لاریجانی، رحمانی و ساده، ۱۳۹۷). مدیران برای ارائه کردن تصویری مطلوب از شرکت و کاهش ریسک سرمایه‌گذاری انگیزه‌ای قوی دارند تا سود را مدیریت کنند (دادنوش فر، احمدی و گسگری، ۱۳۹۰). از اواسط دهه ۱۹۸۰ مطالعات مدیریت سود، بر به‌کارگیری مدیران از اقلام تعهدی، برای تأثیر بر سود گزارش شده تمرکز یافت (اعتمادی، مؤمنی و فرج زاده دهکردی، ۱۳۹۱). اقلام تعهدی حسابداری از دو جزء اختیاری و غیر اختیاری تشکیل می‌شوند. جزء غیر اختیاری این اقلام تحت تأثیر شرایط اقتصادی قرار دارد و بخش اختیاری آن توسط مدیران تعیین می‌شود مدیران می‌توانند با اعمال نظر و قضاوت خود روی اقلام تعهدی اختیاری سود شرکت را جهت اهداف خود مدیریت کنند (کردستانی، معصومی و بقایی، ۱۳۹۲). ساده‌ترین و متداول‌ترین شیوه برای مدیریت سود دستکاری اقلام تعهدی، به ویژه اقلام تعهدی اختیاری است و درعین حال تشخیص مدیریت سود صورت گرفته از این طریق، نیز سخت می‌باشد بنابراین اغلب در پژوهش‌های مدیریت سود به‌کارگیری اقلام تعهدی حسابداری برای مدیریت کردن سود مورد توجه قرار گرفته است (عربصالحی، محمودی و محمودی، ۱۳۹۴). مدل مدیریت سود درواقع یک مدل خطی نیست (هاگلوند^۱، ۲۰۱۲). از مزیت‌های داده‌کاوی نسبت به روش‌های آماری سنتی این می‌باشد که در روش‌های آماری سنتی، علیرغم پیش‌بینی‌های قابل قبول مدل‌های آماری، مفروضات صحیح و

محدودکننده‌ی این روش‌ها، نظیر؛ خطی بودن، نرمال بودن، استقلال متغیرهای مستقل از یکدیگر و وجود یک ساختار تابعی از پیش تعریف شده، کاربرد آن‌ها را در دنیای واقعی محدود می‌کند (چن و دیو^۱، ۲۰۰۹). درحالی‌که روش داده‌کاوی عملاً محدودیت‌های، روش‌های آماری سنتی را ندارد. روش‌های داده‌کاوی در داده‌های مالی، می‌تواند در حل مشکلات طبقه‌بندی و پیش‌بینی و آسان‌سازی فرایند تصمیم‌گیری به کار رود (دستگیر و شفیع‌ی سردشت، ۱۳۹۰). اهمیت داده‌کاوی توسط بسیاری از سازمان‌های حرفه‌ای تشخیص داده شده است، انجمن حسابداری رسمی آمریکا داده‌کاوی را به عنوان یکی از ده فناوری برتر برای آینده معرفی کرده است، همچنین انجمن حساب‌رسان داخلی آمریکا این فناوری را در فهرست یکی از چهار اولویت تحقیقاتی خود گنجانده است (قاصدی قزوینی، فائزی رازی و حیدرپور، ۱۳۹۶). بهترین توصیف از داده‌کاوی به وسیله اجتماع آمار و هوش مصنوعی و یادگیری ماشین به دست می‌آید سپس این تکنیک‌ها با کمک یکدیگر برای مطالعه داده و پیدا کردن الگوهای نهفته در آن‌ها استفاده می‌شود (سروش و بحرینی نژاد، ۱۳۸۸). اهداف داده‌کاوی شامل پیش‌بینی، توصیف یا ترکیبی از آن‌هاست، به هر ترتیب اهداف داده‌کاوی با استفاده از روش‌های داده‌کاوی محقق می‌شوند، اصطلاح روش‌های داده‌کاوی در واقع بیانگر جمع‌کثیری از الگوریتم‌ها و فنون است، برخی از این روش‌ها در مورد داده‌های مالی به کار رفته‌اند (دستگیر و شفیع‌ی سردشت، ۱۳۹۰). به عنوان نمونه روش‌های شبکه عصبی، الگوریتم ژنتیک، درخت تصمیم را می‌توان نام برد. در ایران از اواخر دهه‌ی ۸۰ تا به امروز، برخی از پژوهشگران داخلی، روش‌های داده‌کاوی را برای پیش‌بینی و تشخیص مدیریت سود، در پژوهش‌هایشان برگزیدند. طبق بررسی‌های صورت گرفته، از بین روش‌های داده‌کاوی به کار رفته در پژوهش‌های پیش‌بینی مدیریت سود، تکنیک‌های شبکه عصبی و درخت تصمیم از محبوب‌ترین روش‌های داده‌کاوی پژوهشگران داخلی (برای نمونه؛ نجاری، ۱۳۸۹؛ اصلانی، ۱۳۹۰؛ محمودی، ۱۳۹۰؛ حجازی و همکاران، ۱۳۹۱؛ چالاک‌ی و یوسفی، ۱۳۹۱؛ کردستانی و همکاران، ۱۳۹۲؛ صالحی و فرخی پيله رود، ۱۳۹۳؛ عرب‌صالحی و همکاران، ۱۳۹۴؛ صالحی و فرخی پيله رود، ۱۳۹۷) بوده است. مطالعات گذشته (اصلانی ۱۳۹۰؛ حجازی و همکاران، ۱۳۹۱؛ صالحی و فرخی پيله رود، ۱۳۹۳؛ صالحی و فرخی پيله رود، ۱۳۹۷؛ هاگلوند، ۲۰۱۲) نشان داده است که روش‌های داده‌کاوی در پیش‌بینی مدیریت سود نسبت به روش‌های خطی دقیق‌تر است.

۳- پیشینه پژوهش

۳-۱- پیشینه داخلی

صالحی و فرخی پيله رود (۱۳۹۷)، به پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در چهار صنعت کشاورزی، دارویی، نساجی و فرآورده‌ای نفتی، پرداختند. نتایج حاصل از پژوهش آنان نشان داد که مدیریت سود با متغیرهای اقلام تعهدی اختیاری دوره قبل، اقلام تعهدی غیر اختیاری دوره قبل یا آستانه عملکرد و ریسک در چهار روش مدل‌های خطی، شبکه عصبی، درخت‌های C5.0 و Cart دارای بیشترین ارتباط است و دریافتند که روش شبکه عصبی و درخت تصمیم‌گیری در پیش‌بینی مدیریت سود نسبت به روش خطی دقیق‌تر و دارای سطح خطای کمتری است.

فغانی ماکرانی، صالح نژاد و امین (۱۳۹۵)، در پژوهش خود به این نتیجه دست یافتند که شبکه عصبی مصنوعی از توانایی بالایی در پیش‌بینی مدیریت سود نسبت به مدل خطی جونز تعدیل‌شده برخوردار است.

عربصالحی و همکاران (۱۳۹۴)، در مقاله‌ای تحت عنوان پیش‌بینی مدیریت سود با به‌کارگیری شبکه‌های عصبی پرسپترون چند لایه دارای دو لایه پنهان در صنعت خودرو و ساخت قطعات، به این نتایج دست یافتند که شبکه‌های عصبی پرسپترون چند لایه دارای دو لایه پنهان، قادر به پیش‌بینی مدیریت سود در صنعت خودرو و ساخت قطعات در مجموع به میزان ۶۰٫۵٪ و در بهترین طبقه‌بندی کننده تا میزان ۸۱٫۱٪ برای مدیریت سود نزدیک به صفر و ۶۹٪ برای مدیریت سود افزاینده زیاد گردیده است که نشان از قابلیت بالای این مدل دارد و در بین متغیرهای ورودی بیشترین درجه اهمیت مربوط به ریسک شرکت، اندازه شرکت و آستانه عملکرد می‌باشد. صالحی و فرخی پيله رود (۱۳۹۳)، در پژوهشی به عنوان پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در صنایع کشاورزی و نساجی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، به این نتایج رسیدند که روش شبکه عصبی و درخت تصمیم‌گیری در پیش‌بینی مدیریت سود نسبت به روش‌های خطی دقیق‌تر و دارای سطح خطای کمتری است و مدیریت سود با متغیرهای اقلام تعهدی اختیاری دوره قبل (DAI)، اقلام تعهدی غیر اختیاری دوره قبل یا آستانه عملکرد (THOD) و حساسیت در پرداخت بابت عملکرد (PPS)، در روش‌های رگرسیون، شبکه عصبی، درخت کارت دارای بیشترین ارتباط می‌باشد.

کردستانی و همکاران (۱۳۹۲)، پیش‌بینی سطح مدیریت سود با استفاده از شبکه‌های عصبی مصنوعی را مورد پژوهش قرار دادند، در این پژوهش با استفاده از متغیرهای موجود در

ادبیات مدیریت سود و به کارگیری مدل شبکه‌های عصبی مصنوعی، سطح مدیریت سود پیش‌بینی شده است.

چالاک و یوسفی (۱۳۹۱)، در مقاله‌ای با عنوان پیش‌بینی مدیریت سود با استفاده از درخت تصمیم‌گیری به این نتیجه رسید که بالاترین میزان دقت پیش‌بینی برای درخت تصمیم‌گیری ۷۴٫۷ درصد می‌باشد.

حجازی و همکاران (۱۳۹۱)، در پژوهش خود به این نتایج دست یافتند که روش شبکه عصبی و درخت تصمیم‌گیری در پیش‌بینی مدیریت سود نسبت به روش‌های خطی دقیق‌تر و دارای سطح خطای کمتری است. ضمناً مدیریت سود با متغیرهای اقلام تعهدی اختیاری (dai) و غیر اختیاری دوره قبل (thod)، عملکرد شرکت، اندازه شرکت، تداوم سود در هر دو روش دارای بیشترین ارتباط می‌باشد.

محمودی (۱۳۹۰) در پایان‌نامه خود به پیش‌بینی مدیریت سود با به کارگیری ترکیب مدل شبکه‌های عصبی مصنوعی و درخت‌های تصمیم پرداخته است. مدل ارائه شده یک نرخ ۱۰۰ درصدی از پیش‌بینی مدیریت سود افزاینده افراطی و یک نرخ ۹۷ درصدی از پیش‌بینی مدیریت سود نزدیک به صفر را تأمین می‌نماید.

۳-۲- پیشینه خارجی

چن و همکاران (۲۰۱۵)، در مقاله‌ای تحت عنوان تشخیص مدیریت سود صنعت بیوتکنولوژی با استفاده از شبکه بیزین، تجزیه و تحلیل مؤلفه‌های اصلی، شبکه عصبی پس انتشار و درخت تصمیم به این نتیجه رسیدند که ترکیب شبکه بیزین با درخت تصمیم‌گیری C5.0 می‌تواند بهترین عملکرد را در کاوش درجه جدی از مدیریت سود تعهدی را داشته باشد که در آن صحت گروه آزمون ۹۸٫۵۱٪ است و اقلام تعهدی اختیاری دوره قبل تأثیر زیادی بر درجه جدی مدیریت سود اقلام تعهدی دارند.

هاگلوند (۲۰۱۲)، در مقاله‌ای تحت عنوان تشخیص مدیریت سود با شبکه عصبی، به این نتیجه رسید که عملکرد مدل برمبنای رگرسیون عمومی شبکه عصبی بهترین عملکرد را داراست، در حالی که مدل خطی مبتنی بر رگرسیون ضعیف‌ترین عملکرد را دارد.

تسای و چیو^۱ (۲۰۰۹)، در مقاله‌ای، به بررسی پیش‌بینی مدیریت سود از طریق شبکه عصبی درخت تصمیم پرداختند و از نتایج پژوهش، مدلی با بالاترین نرخ پیش‌بینی ۸۱٪ در مورد مدیریت سود رو به بالا می‌باشد.

1 Tsai and Chiou

۴- روش‌شناسی پژوهش

۴-۱- روش پژوهش

با توجه به هدف اصلی پژوهش، این پژوهش از لحاظ هدف، از نوع کاربردی و توسعه‌ای است و بر اساس ماهیت و روش از نوع پس‌رویدادی (علی-مقایسه‌ای) می‌باشد. به منظور جمع‌آوری داده‌ها و اطلاعات برای تجزیه و تحلیل، از صورت‌های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار فعال در صنعت خودروسازی و ساخت قطعات در بازه زمانی ۱۳۸۸ الی ۱۳۹۳، استفاده شد. که دستیابی به اطلاعات از طریق پایگاه اینترنتی سازمان بورس اوراق بهادار و نرم‌افزارهای بورس از جمله بانک‌های اطلاعاتی ره‌آورد نوین و تدبیرپرداز تأمین شد. پس از جمع‌آوری اطلاعات مربوط، از نرم‌افزار SPSS Modeler 14.2 استفاده شد.

۴-۲- جامعه و نمونه آماری

در این پژوهش جامعه آماری شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، فعال در صنعت خودرو و ساخت قطعات می‌باشد و نمونه پژوهش بر اساس معیارهای زیر انتخاب شده است؛

۱. سال مالی شرکت منتهی به پایان اسفندماه هر سال باشد.
۲. شرکت در دوره مورد مطالعه تغییر سال مالی نداشته باشد.
۳. شرکت فعال در صنعت خودرو فعالیت هلدینگ نداشته باشد^۱.
۴. شرکت‌هایی که تمامی داده‌های مورد نیاز برای محاسبه ارقام تعهدی اختیاری و متغیرهای مستقل را داشته باشند.

قلمرو زمانی این پژوهش پنج سال در بازه زمانی ۱۳۸۸ لغایت ۱۳۹۳ می‌باشد، البته با توجه به اینکه جهت محاسبه رقم ارقام تعهدی اختیاری که جایگزینی برای مدیریت سود می‌باشد نیازمند اطلاعات دو سال قبل هستیم، داده‌های مربوط به سال ۱۳۸۶ و ۱۳۸۷ نیز به بازه فوق اضافه گردید. طی دوره زمانی پژوهش تعداد ۳۱ شرکت فعال در بورس اوراق بهادار تهران در صنعت خودرو و ساخت قطعات حضور داشته‌اند؛ که ۹ شرکت حائز شرایط بالا بوده‌اند. از آنجا که هر شرکت در طی سال‌های ۱۳۸۸ لغایت ۱۳۹۳ دارای ۶ مجموعه اطلاعات مالی قابل

۱ به طور مثال شرکتی مانند گروه بهمن از جامعه آماری حذف می‌گردد زیرا شرکت گروه بهمن، در سال ۸۷ علاوه بر فعالیت‌های تولیدی خودرو که داشت به «تأسیس هلدینگ گروه بهمن در حوزه‌های تولیدی، بازرگانی و سرمایه‌گذاری پرداخت» و تحول ساختاری با تشکیل هلدینگ بهمن به وجود آورد.

استخراج در صورت‌های مالی است بنابراین تعداد مشاهدات کل شرکت - سال برابر ۵۴ مورد می‌باشد.

۴-۳- متغیرهای پژوهش و شیوه اندازه‌گیری آن‌ها

متغیر وابسته (اقلام تعهدی اختیاری): اقلام تعهدی اختیاری، از کم کردن اقلام تعهدی غیر اختیاری از مجموع اقلام تعهدی، محاسبه می‌شود. مطالعات پیشین، معمولاً از دو روش؛ صورت جریان وجوه نقد و روش ترازنامه برای تخمین مجموع اقلام تعهدی استفاده کرده‌اند. همان‌طور که توسط کالینز و هریر^۱ (۲۰۰۲) نشان داده شده است؛ انحراف مجموع اقلام تعهدی برآورد شده توسط روش صورت جریان وجوه نقد کمتر است. از این رو، برای محاسبه مجموع اقلام تعهدی، روش صورت جریان وجوه نقد، انتخاب گردید. روش صورت جریان وجوه نقد برای محاسبه مجموع اقلام تعهدی: (چن و همکاران، ۲۰۱۵):

$$TAC_{i,t} = EBXI_{i,t} - CFO_{i,t} \quad (۱)$$

متغیرها به شرح زیر تعریف می‌شوند؛

$TAC_{i,t}$: مجموع اقلام تعهدی شرکت i در سال t ; $EBXI_{i,t}$: سود قبل از اقلام غیرمترقبه شرکت i در سال t ; $CFO_{i,t}$: جریان نقدی از فعالیت‌های عملیاتی شرکت i در سال t . به دلیل تفاوت استاندارد حسابداری شماره ۲ ایران و استاندارد ۹۵ FASB آمریکا، تعدیل جریان نقدی از فعالیت‌های عملیاتی به صورت زیر انجام گرفته است. جریان نقدی عملیاتی طبق صورت جریان وجوه نقد+ سود سهام پرداخت شده+ جریان نقد مرتبط با بازده سرمایه‌گذاری‌ها و سود پرداختی بابت تأمین مالی - جریان نقدی مرتبط با مالیات (نوروش و حسینی، ۱۳۸۸). در ادامه از روش اندازه‌گیری مدیریت سود توصیه شده توسط دیچو، اسلون، هاتن و کیم (۲۰۱۲) استفاده می‌شود (چن و همکاران، ۲۰۱۵).

$$WC-ACC = (\Delta CA_{i,t} - \Delta CL_{i,t} - \Delta Cash_{i,t} + \Delta STD_{i,t}) / A_{i,t-1} \quad (۲)$$

متغیرها به شرح زیر تعریف می‌شوند؛

$WC-ACC$: سرمایه در گردش تعهدی غیر نقد شرکت i در سال t ; $\Delta CA_{i,t}$: تغییر در دارایی جاری شرکت i در سال t ; $\Delta CL_{i,t}$: تغییر در بدهی جاری شرکت i در سال t ; $\Delta Cash_{i,t}$: تغییر در وجه نقد شرکت i در سال t ; $\Delta STD_{i,t}$: تغییر در حصة جاری تسهیلات مالی دریافتی شرکت i در سال t ; $A_{i,t-1}$: مجموع دارایی‌های شرکت i در سال $t-1$.

در نهایت اقلام تعهدی اختیاری به شرح زیر محاسبه می‌شود: (چن و همکاران، ۲۰۱۵):

$$DA_{i,t} = TAC_{i,t} / A_{i,t-1} - WC-ACC_{i,t} \quad (۳)$$

متغیرها به شرح زیر تعریف می‌شوند؛

$DA_{i,t}$: اقلام تعهدی اختیاری شرکت i در سال t ، $TAC_{i,t}$: مجموع اقلام تعهدی شرکت i در سال t ، $WC-ACC_{i,t}$: سرمایه در گردش تعهدی غیر نقد شرکت i در سال t .

جدول (۱). متغیر وابسته

روش محاسبه	نام متغیر وابسته	کد متغیر
طبق روابط (۱ و ۲)، تفسیر شده بالا	اقلام تعهدی اختیاری (DA)	X1

متغیرهای مستقل: متغیرهای مستقل استفاده شده در پژوهش با توجه به پژوهش انجام شده توسط چن و همکاران (۲۰۱۵)، عربصالحی و همکاران (۱۳۹۴)، کردستانی و همکاران (۱۳۹۲)، حجازی و همکاران (۱۳۹۱)، چالاکی و یوسفی (۱۳۹۱)، به شرح جدول ۲ انتخاب شده‌اند.

جدول (۲). متغیرهای مستقل

روش محاسبه	نام متغیر	کد متغیر
اقلام تعهدی اختیاری دوره قبل	اقلام تعهدی اختیاری دوره قبل (DA_{n-1})	X2
اقلام تعهدی غیر اختیاری - اقلام تعهدی غیر اختیاری دوره قبل	آستانه عملکرد (Thod)	X3
لگاریتم طبیعی از مجموع دارایی‌ها	اندازه شرکت (SIZE)	X4
کل بدهی \ ارزش بازار سهام در شروع دوره	نسبت بدهی (LEV)	X5
چنانچه شرکت زیانده باشد در الگوها مقدار زیان ۱ و در غیر این صورت ۰ است.	زیان (LOSS)	X6
سود خالص \ مجموع دارایی‌ها	بازده دارایی (ROA)	X7
جریان‌های نقدی عملیاتی \ کل دارایی‌ها در شروع دوره	جریان‌های نقدی عملیاتی (OCF)	X8
سهام در دست سهامداران نهادی \ کل سهام شرکت * ۱۰۰	نسبت سهام سرمایه‌گذار نهادی (ORGA)	X9
۱ نشان‌دهنده سازمان حسابرسی و ۰ نشان‌دهنده غیر از سازمان حسابرسی	آیا حسابرسی شرکت‌های صنعت خودرو و ساخت قطعات توسط سازمان حسابرسی انجام شده است؟	X10
سود خالص \ جمع حقوق صاحبان سهام	بازده حقوق صاحبان سهام (ROE)	X11
دارایی جاری \ بدهی جاری	نسبت جاری (Curr)	X12
کل بدهی‌ها \ کل دارایی‌ها	نسبت بدهی (DB)	X13
سود خالص \ تعداد سهام در تاریخ ترازنامه	سود هر سهم (EPS)	X14
سود خالص \ فروش خالص	حاشیه سود خالص (PIS)	X15
مبلغ تغییر در درآمد فروش دوره جاری نسبت به دوره گذشته \ درآمد فروش دوره گذشته	فرصت رشد (GROWTH)	X16
تعداد مدیران غیرموظف	تعداد مدیران غیرموظف (IND)	X17
کل بدهی \ جمع حقوق صاحبان سهام	نسبت بدهی به حقوق صاحبان سهام (LA)	X18
بهای تمام شده کالای فروخته \ میانگین موجودی کالا	گردش موجودی کالا (IT)	X19
فروش خالص \ میانگین حساب‌های دریافتنی	گردش وصول مطالبات (RT)	X20
سود قبل از بهره و مالیات \ سود قبل از مالیات	اهرم مالی (FLV)	X21
(دارایی جاری - موجودی کالا - پیش‌پرداخت هزینه) \ بدهی‌های جاری	ضریب قدرت پرداخت (QR)	X22
سود خالص قبل از مالیات \ تعداد کل کارمندان	سودآوری کارمندان (EN)	X23
قیمت سهام \ سود هر سهم	نسبت قیمت به سود (P/E)	X24
جریان‌های نقدی عملیاتی	جریان‌های نقدی عملیاتی (CFO)	X25

۵- پرسش پژوهش

آیا با به کارگیری مدل ترکیبی شبکه بیزین و درخت تصمیم‌گیری می‌توانیم سطح مدیریت سود را در یک شرکت فعال در صنعت خودرو و ساخت قطعات، تشخیص دهیم؟

۶- طراحی؛ اجرا و یافته‌های حاصل از مدل ترکیب شبکه بیزین و درخت

تصمیم C5.0

در این بخش از مقاله ابتدا تکنیک‌های داده‌کاوی استفاده شده در مدل ترکیبی معرفی می‌گردد و سپس مراحل اجرای پژوهش بیان می‌شود. از آنجایی که طبقه‌بندی و تخمین دو شکل از تحلیل داده‌ها محسوب می‌شوند که می‌توان به کمک آن‌ها مدلی جهت توصیف داده‌ها استخراج کرد و روش‌های نظارت شده‌ای مانند طبقه‌بندی و تخمین تلاش می‌کنند تا رابطه‌ی میان صفات خاصه ورودی (که گاه متغیر مستقل نامیده می‌شوند) را با یک یا چندین صفت خاصه هدف (که گاه متغیر وابسته نامیده می‌شوند) کشف کنند. در نهایت این رابطه با یک ساختار به عنوان مدل نمایش داده می‌شود (اسماعیلی، ۱۳۹۲).

لذا در این پژوهش از روش‌های طبقه‌بندی داده‌کاوی، از جمله شبکه بیزین و درخت تصمیم برای طراحی مدل تشخیص مدیریت سود استفاده می‌گردد. یک شبکه بیزین، مدل گرافیکی است که متغیرهای موجود در یک مجموعه داده را که معمولاً گره نامیده می‌شوند همراه با روابط احتمالی و شرطی بین آن‌ها نشان می‌دهند. روابط سببی بین گره‌ها می‌تواند به وسیله شبکه بیزین نمایش داده شود، البته پیوندهای موجود در شبکه (کمان‌ها) لزوماً ارتباط مستقیم علت معلولی را نشان نمی‌دهد. شبکه‌ها در مقابل داده‌های مفقوده بسیار مقاوم‌اند و بهترین پیش‌بینی را ارائه می‌دهند (علیزاده و ملک محمدی، ۱۳۹۳).

یکی از سودمندترین مزایای این شبکه‌ها این است که جهت انجام آنالیز، چیزی به عنوان اندازه حداقل برای داده‌های نمونه، وجود ندارد و این شبکه‌ها در محاسبات خود از تمامی داده‌های موجود استفاده می‌کنند و حتی با وجود نمونه‌هایی با اندازه‌ای نسبتاً کوچک می‌توان به پیش‌بینی‌هایی با دقت نسبتاً خوب دست یافت (خدمت‌گزار سالانقوج، ۱۳۹۲).

ساختار درخت تصمیم یک ساختار درختی، شبیه فلوچارت است. بالاترین گره در درخت، گره ریشه است و گره‌های برگ، دسته‌ها یا توزیع دسته‌ها را نشان می‌دهند. درخت تصمیم یکی از ابزارهای قوی و متداول برای دسته‌بندی و پیش‌بینی می‌باشد. در ساختار درخت تصمیم، پیش‌بینی به دست آمده از درخت در قالب یکسری قواعد توضیح داده می‌شود (غضنفری، علیزاده و تیمورپور، ۱۳۷۸).

یکی از مزایای آن این است که بعد از ساخت درخت به راحتی می‌توان علت استنتاج قواعد بدست آمده را مشاهده نمود به این معنی که درخت تصمیم همانند شبکه عصبی مانند یک جعبه سیاه عمل نمی‌کند و منطق کاری آن روشن و آشکار است (علیزاده و ملک محمدی، ۱۳۹۳). درخت‌های تصمیم به علت دقت مناسبشان و همچنین ساده بودن و فهم راحت آن‌ها، یکی از محبوب‌ترین روش‌های دسته‌بندی هستند (حاجی حیدری و خاکباز، ۱۳۹۲).

در این پژوهش از درخت تصمیم C5.0 استفاده شده است که از جمله مزایای الگوریتم C5.0 درخت تصمیم، سرعت بالا، کارایی حافظه، درخت‌های تصمیم کوچک‌تر، پشتیبانی برای بوستینگ^۱، وزن دهی، غربال‌سازی می‌باشد (سروش و بحرینی نژاد، ۱۳۸۸). با توجه به شکل ۱ که حاوی نمودار جریان پژوهش می‌باشد مراحل اجرای این پژوهش به صورت زیر خلاصه می‌شود:

در مرحله اول: پس از مشخص نمودن جامعه و نمونه آماری پژوهش، داده‌ها جمع‌آوری می‌شوند و متغیرهای پژوهش که شامل:

(۱) متغیر وابسته: مدیریت سود تعهدی (اقدام تعهدی اختیاری)

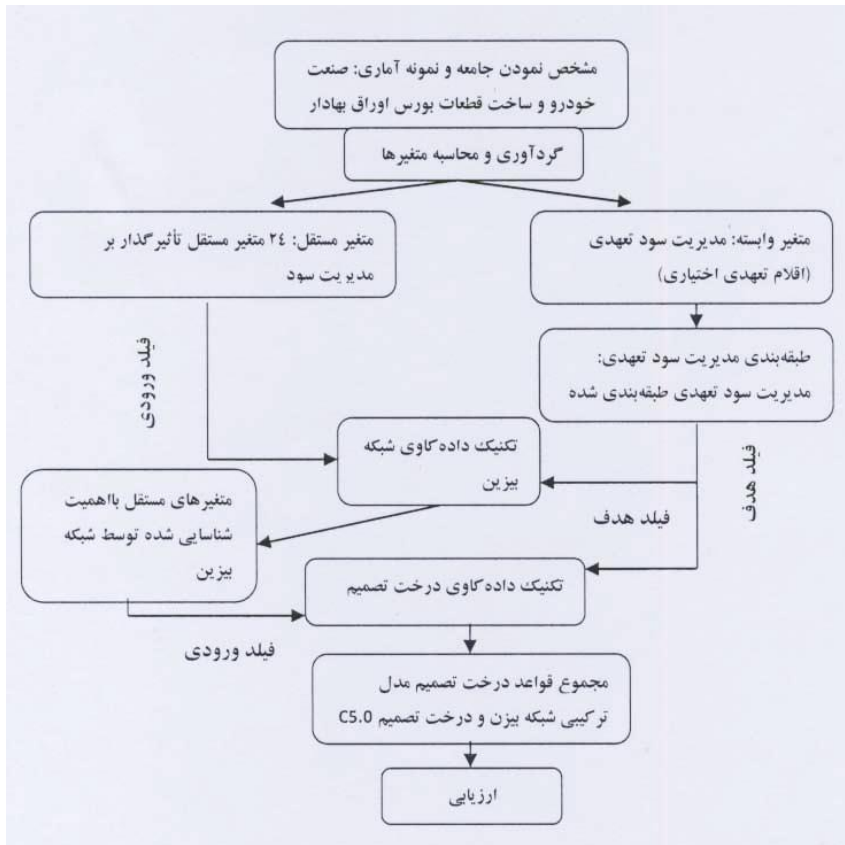
(۲) متغیر مستقل: ۲۴ متغیر تأثیرگذار بر مدیریت سود (مطابق جدول ۲)

می‌باشد گردآوری و محاسبه می‌شوند. در این پژوهش، به منظور تشخیص مدیریت سود، ابتدا بایستی مدیریت سود تعهدی (متغیر وابسته) طبقه‌بندی شود.

مرحله دوم: پس از طبقه‌بندی مدیریت سود تعهدی، در مرحله دوم از شبکه بیزین برای غربالگری متغیرهای مستقل و شناسایی متغیرهای مستقل با اهمیت استفاده می‌شود که برای این منظور، مدیریت سود تعهدی طبقه‌بندی شده به عنوان فیلد هدف و ۲۴ متغیر مستقل را به عنوان فیلد ورودی، برای راه‌اندازی شبکه بیزین در نظر می‌گیرد و به این ترتیب شبکه بیزین متغیرهای مستقل با اهمیت را شناسایی می‌کند.

در مرحله سوم: پژوهش، از متغیرهای مستقل با اهمیت شناسایی شده توسط شبکه بیزین، به عنوان فیلد ورودی و از مدیریت سود تعهدی طبقه‌بندی شده به عنوان فیلد هدف برای راه‌اندازی درخت تصمیم C5.0 استفاده می‌کند. در این مرحله نیز با اهمیت‌ترین متغیر مدل ترکیبی شناسایی و مجموعه قواعد درخت تصمیم مدل ترکیبی استخراج می‌شود.

۱ الگوریتم C5.0، روش خاصی برای بهبود نرخ دقت پیش‌بینی دارد که به آن boosting می‌گویند. این روش با ساخت مدل‌های چندگانه به صورت متوالی کار می‌کند. (علیزاده و ملک محمدی، ۱۳۹۳)



شکل (۱). نمودار جریان پژوهش

در مرحله چهارم: مدل ترکیبی شبکه بیزین و درخت تصمیم C5.0، ارزیابی می‌گردد و در ادامه، مراحل ذکر شده به تفسیر بیان می‌شود.

۶-۱- طبقه‌بندی درجه مدیریت سود تعهدی

در مرحله اول؛ به منظور تشخیص سطح مدیریت سود توسط شبکه بیزین و درخت تصمیم، ابتدا بایستی مدیریت سود طبقه‌بندی شود، به عبارت دیگر بایستی سطوح مختلف مدیریت سود تعریف شود برای طبقه‌بندی مدیریت سود (اقدام تعهدی اختیاری) از گره گسسته سازی^۱، نرم‌افزار SPSS Modeler 14.2 استفاده می‌کنیم. این گره مجموعه فیلد جدیدی را بر اساس مقادیر یک یا چند فیلد ایجاد می‌کند (علیزاده و ملک محمدی، ۱۳۹۳). در این گره، از تکنیک

1 Binning

میانگین^۱ و انحراف استاندارد^۲ برای گسسته سازی استفاده می‌کنیم که به این صورت عمل می‌کند از مجموعه نمونه‌ها که در اینجا اقلام تعهدی اختیاری DA است، میانگین و انحراف استاندارد آن‌ها را محاسبه می‌کند و سپس در مرحله اول یک درجه از مقدار انحراف معیار را به مقدار میانگین اضافه می‌کند، مقدار میانگین به علاوه یک درجه انحراف معیار را به عنوان سقف (مدیریت سود بسیار به سمت بالا) در نظر می‌گیریم؛ و در مرحله دوم؛ یک درجه از انحراف معیار، از میانگین کم می‌گردد، مقدار میانگین منهای یک درجه انحراف معیار را، به عنوان کف (مدیریت سود بسیار به سمت پایین) در نظر می‌گیریم. با این حال مقداری از اقلام تعهدی اختیاری که بین سقف و کف باشد، رفتار مدیریت سود جزئی (اندک) تلقی می‌شود.

جدول (۳). آمار توصیفی اقلام تعهدی اختیاری (مدیریت سود تعهدی)

میانگین	انحراف استاندارد	میانگین + انحراف استاندارد	میانگین - انحراف استاندارد
۰,۰۲۴۶۱۹	۰,۱۳۵۲۴۵	۰,۱۵۹۸۶۵	-۰,۱۱۰۶۳

منبع: یافته‌های پژوهش

جدول (۴). فواصل طبقه‌بندی مدیریت سود تعهدی

برچسب	فاصله طبقه‌بندی	نام طبقه‌بندی
۱-	مدیریت سود بسیار رو به پایین (مدیریت سود کاهنده)	$DA < -0,1106361$
۰	مدیریت سود جزئی	$-0,1106361 \leq DA \leq 0,1598667$
۱	مدیریت سود بسیار رو به بالا (مدیریت سود افزاینده)	$DA > 0,1598667$

به این ترتیب سه سطح مدیریت سود (۱، ۰، ۱-) به عنوان برچسب خروجی طبق جدول (۴) تعیین می‌شود. پس از طبقه‌بندی اقلام تعهدی اختیاری DA، به سراغ غربالگری متغیرها پژوهش می‌رویم.

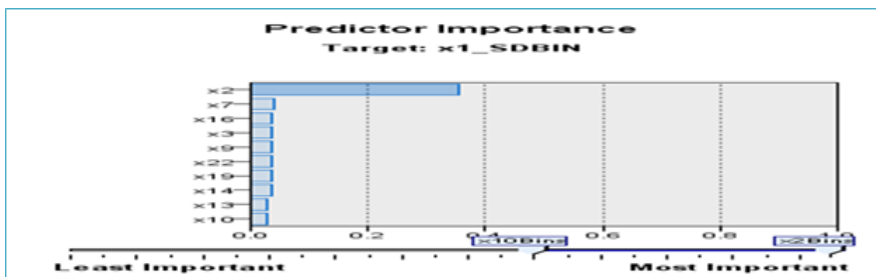
۶-۲- غربالگری متغیرها توسط شبکه بیزین

در این پژوهش در مرحله دوم، از شبکه بیزین برای غربالگری متغیرها، به منظور به دست آوردن متغیرهای مستقل بااهمیت استفاده می‌کنیم. بعد از جمع‌آوری داده‌های مربوط به متغیرها و طبقه‌بندی مدیریت سود (تعیین سه سطح مدیریت سود)، به ساخت مجموعه آموزش و آزمون می‌پردازیم که ۶۰ درصد داده‌ها به عنوان مجموعه آموزش و ۴۰ درصد به عنوان مجموعه آزمون، به صورت تصادفی انتخاب می‌گردد. داده‌های بخش آموزش مدل را می‌سازند و داده‌های بخش آزمون مدل ایجاد شده را مورد ارزیابی قرار می‌دهند. یکی از نیازمندی‌های شبکه بیزین، فیلد هدف طبقه‌ای می‌باشد. فیلد هدف (متغیر وابسته) این پژوهش مدیریت سود

1 Mean

2 Standard Deviation

طبقه‌بندی شده (سه سطح مدیریت سود ۱، ۰، ۱-) می‌باشد و فیلدهای ورودی، ۲۴ متغیر مستقل می‌باشند، برای اینکه بتوانیم شبکه‌ای با کیفیت بالا داشته باشیم از الگوریتم جست‌وجو شبکه بیزین TAN استفاده شده است. عملکرد الگوریتم جست‌وجوی TAN بدین گونه است که یک شبکه بیزین ساده ایجاد می‌کند که در واقع توسعه‌ای از مدل بیز استاندارد می‌باشد این ساختار اجازه می‌دهد که هر پیش‌بینی کننده علاوه بر وابستگی به متغیر هدف بتواند به دیگر پیش‌بینی کننده‌ها نیز وابستگی داشته باشد؛ بنابراین این ساختار دقت دسته‌بندی را افزایش می‌دهد (علیزاده و ملک محمدی، ۱۳۹۳). نمودار شماره ۱، نمودار اهمیت متغیرهای ایجاد شده توسط شبکه بیزین را نمایش می‌دهد؛ که در آن اهمیت متغیرها به شرح جدول ۵ رتبه‌بندی شده است.



نمودار (۱). با اهمیت‌ترین متغیرهای شناسایی شده توسط شبکه بیزین

بنابراین، با استفاده از شبکه بیزین، از ۲۴ متغیر مستقل تأثیرگذار بر مدیریت سود، ۱۰ متغیر (به شرح جدول ۵) به عنوان با اهمیت‌ترین‌ها شناسایی شدند.

جدول (۵). اهمیت نسبی با اهمیت‌ترین متغیرهای شبکه بیزین

متغیر	متغیر	اهمیت نسبی
X ₂	اقدام تعهدی اختیاری دوره قبل (DA _{n-1})	۰,۳۵
X ₇	بازده دارایی (ROA)	۰,۰۴
X ₁₆	رشد فروش (GROWTH)	۰,۰۴
X ₃	آستانه عملکرد (THOD)	۰,۰۴
X ₉	درصد مالکیت سهام سرمایه‌گذار نهادی (ORGA)	۰,۰۴
X ₂₂	ضریب قدرت پرداخت (QR)	۰,۰۴
X ₁₉	نسبت گردش کالا (IT)	۰,۰۴
X ₁₄	سود هر سهم (EPS)	۰,۰۴
X ₁₃	نسبت بدهی (DB)	۰,۰۳
X ₁₀	آیا حسابرسی شرکت توسط سازمان حسابرسی صورت گرفته است؟	۰,۰۳

۶-۳- ترکیب شبکه بیزین و درخت تصمیم

و در مرحله سوم؛ متغیرهای بااهمیت شناسایی شده توسط شبکه بیزین، در ایجاد درخت تصمیم استفاده می‌شوند و داده‌ها به دو بخش آموزش (۶۰ درصد) و آزمون (۴۰ درصد) تقسیم می‌گردند، الگوریتم درخت تصمیم انتخابی در این مرحله از مدل‌سازی C5.0 می‌باشد. متغیر مدیریت سود (اقدام تعهدی اختیاری) طبقه‌بندی شده به عنوان متغیر هدف (متغیر وابسته) و متغیرهای نشان داده شده در جدول ۵؛ که توسط شبکه بیزین به عنوان متغیرهای بااهمیت شناسایی شده‌اند به عنوان متغیرهای ورودی (متغیر مستقل)، درخت تصمیم C5.0 استفاده می‌شود. درخت تصمیم C5.0، تمامی فیلدهای پایگاه داده را بررسی می‌کنند تا به فیلدی برسند که بهترین دسته‌بندی و پیش‌بینی را با تقسیم داده‌ها به زیرگروه‌ها انجام دهد. این پروسه به صورت بازگشتی تکرار می‌شود تا بازمی‌تقسیم داده‌ها به زیرگروه‌های دیگری شکسته شوند. زیر نمونه‌ها دوباره شاخه زده می‌شوند و این فرایند تا زمانی که نتوان زیر نمونه‌ها را شاخه زد ادامه می‌یابد. در پایان، پایین‌ترین سطح شاخه‌ها آزموده می‌شوند و شاخه‌هایی که نقش مهمی در مدل ندارند حذف یا هرس می‌شوند (علیزاده و ملک محمدی، ۱۳۹۳).

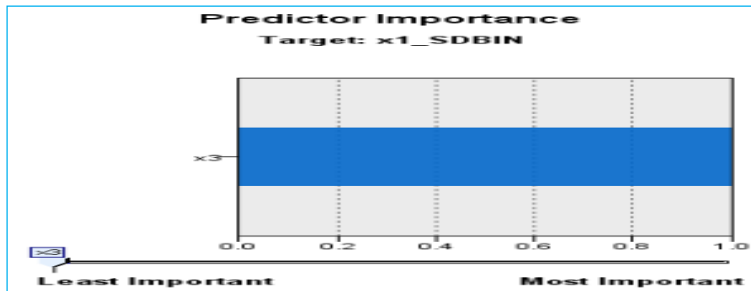
با توجه به جدول ۶، درخت تصمیم روی ۶۰ درصد نمونه‌ها یعنی در ۳۳ نمونه یادگیری صورت گرفته است و روی ۴۰ درصد باقی‌مانده یعنی ۲۱ نمونه، آزمون شده است. نتایج آزمون نشان می‌دهد که ۹۰٫۴۸٪ درصد نمونه‌های آزمون یعنی ۱۹ نمونه درست دسته‌بندی شده است و ۹٫۵۲٪ درصد نمونه‌ها، ۲ نمونه، اشتباه دسته‌بندی شده‌اند. پس مدل ترکیبی حاصل دارای صحت ۹۰٫۴۸٪ می‌باشد.

جدول (۶). داده‌های آموزش و آزمون مدل ترکیب شبکه بیزین و

درخت تصمیم C5.0 و صحت آن

تقسیم‌بندی	آموزش	آزمون
صحیح	۳۱	۱۹
اشتباه	۲	۲
جمع	۳۳	۲۱

معمولاً در مدل‌سازی بر متغیرهایی که اهمیت بیشتری دارند تمرکز می‌شود و متغیرهای کم‌اهمیت‌تر حذف می‌شوند. نمودار ۲ بااهمیت‌ترین متغیرشناسایی شده درخت تصمیم C5.0، مدل ترکیبی شبکه بیزین و درخت تصمیم را نمایش می‌دهد. طبق نمودار، متغیر x_3 (آستانه عملکرد) بااهمیت‌ترین متغیر با نسبت اهمیت ۱ یا ۱۰۰ درصد می‌باشد.



نمودار (۲). با اهمیت ترین متغیر شناسایی شده توسط درخت تصمیم C5.0 مدل ترکیبی شبکه بیزین و درخت تصمیم

در بخش آخر مدل سازی، به تولید مجموعه قواعد از درخت تصمیم مدل ترکیبی می پردازیم و با آزمون اضافی درخت تصمیم C5.0، قواعد برگرفته از آن استخراج گردید، می توان یک مدل به شکل قوانین تعریف کرد که البته شاخه های پایانی درخت تا ریشه درخت همگی در ایجاد این قوانین نقش دارند. معمولاً مجموعه های قوانین، مهم ترین اطلاعات مرتبط با درخت تصمیم را در بر دارند و نسخه ساده شده یا استخراج شده از اطلاعات یافته شده در درخت تصمیم را ارائه می کند و از پیچیدگی کمتری برخوردار هستند (علیزاده و ملک محمدی، ۱۳۹۳). مجموعه قوانین استخراج شده از مدل ترکیبی، در جدول ۷ نشان داده شده است.

جدول (۷). قواعد استخراج شده مدل ترکیبی شبکه بیزین و درخت تصمیم C5.0

قانون ۱ برای طبقه مدیریت سود بسیار رو به پایین (مدیریت سود کاهنده)	اگر $x3 \leq -0,361$ باشد، مدیریت سود بسیار رو به پایین (مدیریت سود کاهنده) خواهد بود.
قانون ۱ برای طبقه مدیریت سود جزئی	اگر $-0,361 < x3 \leq 0,247$ باشد، مدیریت سود جزئی خواهد بود.
قانون ۱ برای طبقه مدیریت سود بسیار رو به بالا (مدیریت سود افزاینده)	اگر $x3 > 0,247$ باشد مدیریت سود بسیار رو به بالا (مدیریت سود افزاینده) خواهد بود.

۴-۶- ارزیابی مدل

پس از مدل سازی بایستی به ارزیابی مدل ایجاد شده پرداخت. نتایج ارزیابی باعث بهبود مدل شده و مدل را قابل استفاده می نماید. برای بررسی صحت مدل، داده ها به دو دسته آموزش (۶۰ درصد) و آزمون (۴۰ درصد) تقسیم شدند. داده های بخش آموزش مدل را می سازند و داده های بخش آزمون مدل ایجاد شده را تست و مورد ارزیابی قرار می دهند. شاخص های مختلفی برای ارزیابی عملکرد روش های دسته بندی وجود دارد، معیارهای مختلفی همچون حساسیت، دقت، صحت و ... که در ادامه برخی از این موارد توضیح داده می شود.

در این پژوهش از سه معیار ارزیابی؛ دقت، حساسیت و صحت، جهت ارزیابی مدل پیشنهادی «تشخیص سطح مدیریت سود»^۱ استفاده می شود. مقادیر بالای ۸۰ در صد معیارهای مذکور، بیانگر توان بالای الگوریتم در رده بندی متغیرها و مقادیر بین ۷۰ تا ۸۰ در صد بیانگر قابل قبول بودن نتایج الگوریتم است (افلاطونی، ۱۳۹۴). برای محاسبه مقادیر شاخص های دقت و حساسیت از پارامترهای جدول ماتریس پیرشانی (اغتشاش)^۱ ایجاد شده در محیط نرم افزار، استفاده می گردد؛ که در رابطه های ۱ و ۲، دو معیار حساسیت و دقت معرفی شده اند. در رابطه های مزبور، مثبت در ست (TP) نشان دهنده تعداد نمونه هایی که جز دسته مثبت بوده و در ست پیش بینی شده اند. مثبت نادر ست (FP) نشان دهنده تعداد نمونه هایی که نادر ست به عنوان دسته مثبت پیش بینی شده اند.

$$(۱) \quad \text{حساسیت} = \frac{TP}{\text{POS}} = \frac{\text{تعداد داده های بر چسب مثبتی که درست دسته بندی شده اند}}{\text{کل تعداد داده های مثبت}}$$

$$(۲) \quad \text{دقت} = \frac{TP}{TP + FP}$$

رابطه (۱) و (۲)؛ (غضنفری و همکاران، ۱۳۸۷).

طبق ماتریس پیرشانی داده های آزمون، نشان داده شده در جدول ۸، ۱ نمونه که متعلق به دسته ۱- (مدیریت سود کاهنده) بودند به دسته های ۰ تعلق گرفته است. ۱ نمونه که متعلق به دسته ۰ (مدیریت سود جزئی) بودند به عنوان دسته ۱ پیش بینی شده اند. طبق ماتریس هیچ یک از نمونه های متعلق به دسته ۱ (مدیریت سود افزایشنده) به اشتباه پیش بینی نشده اند.

جدول (۸). ماتریس پیرشانی داده های آزمون مدل ترکیبی

دسته بندی	(-۱)	(۰)	(۱)
(-۱) مدیریت سود کاهنده	۲	۱	۰
(۰) مدیریت سود جزئی	۰	۱۵	۱
(۱) مدیریت سود افزایشنده	۰	۰	۲

حساسیت و دقت محاسبه شده از ماتریس پیرشانی داده های آزمون، مدل ترکیبی مطابق جدول ۹ می باشد. همان طور که مشاهده می شود دقت طبقه ۱- (مدیریت سود کاهنده)، ۱ می باشد به عبارت دیگر هیچ یک از طبقه های دیگر، به عنوان طبقه ۱- به اشتباه پیش بینی نشده اند. همچنین در مقابل یک نمونه از طبقه ۱- به اشتباه پیش بینی شده است و دارای حساسیت ۶۶,۶۶ درصد می باشد. از این رو می توان گفت که دقت طبقه ۱- (مدیریت سود کاهنده) ۱۰۰ درصد می باشد.

جدول (۹). حساسیت و دقت محاسبه شده از ماتریس پیشانی داده‌های آزمون مدل ترکیبی

پاریشن / معیار	حساسیت	دقت
-۱	۰,۶۶۶۶	۱
۰	۰,۹۳۷۵	۰,۹۳۷۵
۱	۱	۰,۶۶۶۶
میانگین	۰,۸۶۸۰	۰,۸۶۸۰

میانگین حساسیت و دقت مدل ترکیبی شبکه بیزین و درخت تصمیم، ۸۶,۸۰ درصد می‌باشد. صحت به دست آمده از مدل ترکیبی ۹۰,۴۸ درصد می‌باشد و میزان خطای مدل ۹,۵۲ درصد است، در نتیجه می‌توان ادعا کرد که مدل از صحت خوبی برخوردار است و توانایی بالایی در تشخیص مدیریت سود دارد.

۷- نتیجه‌گیری و بحث

در این پژوهش، به بررسی اثربخشی، مدل ترکیبی شبکه بیزین و درخت تصمیم برای پیش‌بینی و تشخیص سطح مدیریت سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، فعال در صنعت خودرو و ساخت قطعات، پرداخته شد. برای این منظور، مدیریت سود (اقدام تعهدی اختیاری) به سه طبقه (مدیریت سود افزاینده، مدیریت سود جزئی، مدیریت سود کاهنده) دسته بندی گردید و سپس با اهمیت‌ترین متغیرهای مستقل، برای راه‌اندازی درخت تصمیم C5.0 توسط شبکه بیزین شناسایی شد و در نهایت، قواعد معنی‌داری توسط درخت تصمیم C5.0 مدل ترکیبی استخراج گردید. با توجه به پژوهش حاضر نتیجه می‌گیریم که به کارگیری شبکه بیزین می‌تواند به سهامداران، سرمایه‌گذاران، تحلیل‌گران مالی و سایر استفاده‌کنندگان صورت‌های مالی، در شناسایی تأثیرگذارترین (بااهمیت‌ترین) متغیرها بر مدیریت سود کمک کند. با توجه به نتایج به دست آمده از پژوهش، از ترکیب شبکه بیزین با درخت تصمیم C5.0، قواعد معنی‌داری با صحت ۹۰,۴۸ درصد، استخراج شده است و این بیانگر توان بالای مدل ترکیبی شبکه بیزین و درخت تصمیم‌گیری C5.0 در تشخیص سطح مدیریت سود می‌باشد. با به کارگیری آن قواعد، می‌توان وضعیت شرکت‌های فعال در صنعت خودرو و ساخت قطعات بورس اوراق بهادار تهران را از لحاظ سطوح مختلف مدیریت سود (مدیریت سود افزاینده، مدیریت سود جزئی، مدیریت سود کاهنده)، بررسی کرد. نتایج حاصل از درخت تصمیم C5.0، مدل ترکیبی نشان می‌دهند که X3 (آستانه عملکرد) تأثیر زیادی بر مدیریت سود دارد و درجه اهمیت آن ۱۰۰ درصد می‌باشد و از آنجایی که اقدام تعهدی غیر اختیاری تحت تأثیر

عوامل اقتصادی خارجی قرار دارد و در کشور، بیشتر شرکت‌های فعال در صنعت خودرو و ساخت قطعات از حمایت‌های مستقیم و غیرمستقیم دولت استفاده می‌کنند و در نتیجه اکثراً تحت تأثیر عوامل اقتصادی خارجی می‌باشند؛ بنابراین شناسایی متغیر آستانه عملکرد، توسط مدل به عنوان تأثیرگذارترین متغیر بر مدیریت سود در صنعت خودرو و ساخت قطعات، مناسب به نظر می‌رسد.

حسابرسی صورت‌های مالی اغلب محدود به محدودیت زمانی حسابرسی، هزینه و منابع انسانی است؛ بنابراین برای حسابرس سخت است که سود دستکاری شده را از میان داده‌های عظیم و پیچیده صورت‌های مالی پیدا کند، ساختار قوانین و قواعد تصمیم‌گیری فوق می‌تواند به کمک برخی از کارکنان حسابرسی برآید، هنگامی که هزینه و زمان حسابرسی محدود است. با توجه به هدف پژوهش و اینکه میانگین حساسیت و دقت مدل ترکیبی شبکه بیزین و درخت تصمیم، ۸۶٫۸۰ درصد می‌باشد. صحت به دست آمده از مدل ترکیبی ۹۰٫۴۸ درصد می‌باشد و میزان خطای مدل ۹٫۵۲ درصد است می‌توان گفت که مدل توانایی بالایی را در تشخیص سطح مدیریت سود دارد.

در ادامه نتایج پژوهش با پژوهش‌های مشابه بررسی می‌شود. پژوهش عربصالحی و همکاران (۱۳۹۴)، نشان می‌دهد که شبکه‌های عصبی پرسپترون چند لایه دارای دو لایه پنهان قادر به پیش‌بینی مدیریت سود در صنعت خودرو و ساخت قطعات در مجموع به میزان ۶۰٫۵٪ می‌باشد. درحالی‌که نتایج پژوهش حاضر نشان می‌دهد که مدل ترکیب شبکه بیزین و درخت تصمیم C5.0 دارای صحت ۹۰٫۴۸ درصد می‌باشد در نتیجه می‌توان ادعا کرد که مدل ترکیب شبکه بیزین و درخت تصمیم C5.0، از نتایج بهتری برای پیش‌بینی و تشخیص مدیریت سود برخوردار است. البته لازم به ذکر است که در پژوهش صورت گرفته توسط عربصالحی و همکاران (۱۳۹۴)، از ۱۲ عامل مؤثر بر مدیریت سود برای سنجش توانایی شبکه عصبی در پیش‌بینی مدیریت سود و سطوح مختلف آن استفاده گردیده است و همچنین اطلاعات شش ماهه هر شرکت نمونه جمع‌آوری شده است، درحالی‌که در پژوهش حاضر از ۲۴ متغیر مؤثر بر مدیریت سود و اطلاعات سالانه شرکت استفاده شده است و همچنین چن و همکاران (۲۰۱۵)، در پژوهش‌شان به این نتیجه رسیدند که مدل ترکیب شبکه بیزین و درخت تصمیم C5.0، با صحت ۹۸٫۵۱ درصد بهترین مدل برای تشخیص مدیریت سود در صنعت بیوتکنولوژی می‌باشد. که تقریباً با نتایج به دست آمده با پژوهش حاضر هماهنگ است.

بااهمیت‌ترین متغیر شناسایی شده در مدل ترکیبی این پژوهش آستانه عملکرد THOD می‌باشد. این با پژوهش‌های انجام گرفته توسط عربصالحی و همکاران (۱۳۹۴)، صالحی و فرخی

پيله رود (۱۳۹۳) و حجازی و همکاران (۱۳۹۱) که به این نتیجه رسیدند که آستانه عملکرد نیز جزء متغیرهای بااهمیت می‌باشد، مطابقت دارد. نتایج حاصل از این پژوهش و مطالعات گذشته (صالحی و فرخی پيله رود، ۱۳۹۷؛ فغانی ماکرانی و همکاران، ۱۳۹۵؛ عربصالحی و همکاران، ۱۳۹۴؛ صالحی و فرخی پيله رود، ۱۳۹۳؛ کردستانی و همکاران، ۱۳۹۲؛ چالاکی و یوسفی، ۱۳۹۱؛ محمودی، ۱۳۹۰؛ چن و همکاران، ۲۰۱۵؛ هاگلوند، ۲۰۱۲؛ تسایی و چیو، ۲۰۰۹) نشان داده است که به‌کارگیری فنون داده‌کاوی در پیش‌بینی مدیریت سود، تأثیرگذار است و این می‌تواند کاربرد جدیدی از کارکردهای داده‌کاوی، در حوزه‌ی مالی و حسابداری باشد و از محدودیت‌های پژوهش فعلی این است که با توجه به وجود عامل تورم و تورم قابل‌ملاحظه در محیط اقتصادی ایران و این که بخشی از روند افزایشی اقلام تعهدی در طی زمان به تأثیرپذیری طبیعی این اقلام از تورم مربوط می‌شود، در مدل محاسبه اقلام تعهدی اختیاری و سایر محاسبات مربوط به متغیرهای پژوهش این عامل تعدیل نشده است.

۸- پیشنهادها

به سرمایه‌گذاران و تحلیل‌گران مالی پیشنهاد می‌شود که با استفاده از قواعد ایجاد شده در این پژوهش، سطح مدیریت سود در شرکت‌های فعال در صنعت خودرو و ساخت قطعات را پیش‌بینی و تشخیص دهند و از این طریق، استفاده‌کنندگان از صورت‌های مالی، از ابتلا به زیان ناشی از مدیریت سود، پیشگیری کنند.

به سازمان بورس و اوراق بهادار و سازمان حسابرسی پیشنهاد می‌شود که در تدوین استانداردهای حسابداری و قوانین مالی موضوع مدیریت سود را مدنظر داشته باشند و رهنمودهای لازم را برای محدود نمودن مدیریت سود ارائه دهند و با استفاده از مدل این پژوهش، نرم‌افزاری کاربردی برای تشخیص مدیریت سود ایجاد کنند.

۹- منابع

اسماعیلی، مهدی. (۱۳۹۲). مفاهیم و تکنیک‌های داده‌کاوی. کاشان: ناشر دانشگاه آزاد اسلامی واحد کاشان.

اصلانی، زهرا. (۱۳۹۰). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. پایان‌نامه کارشناسی ارشد حسابداری، دانشگاه الزهرا (س).

اعتمادی، حسین؛ منصور مؤمنی و حسن فرج زاده دهکردی. (۱۳۹۱). مدیریت سود چگونه کیفیت سود شرکت‌ها را تحت تأثیر قرار می‌دهد؟. پژوهش‌های حسابداری مالی ۲: ۱۰۱-۱۲۲.

اعطاسی، امیررضا. (۱۳۹۴). بررسی گروه خودرو سازی در گذار از دوران رکود. سایت روزنامه جهان اقتصاد (۱۳۹۴/۱۰/۱۸)، <http://jahaneghtesad.com>.
افلاطونی، عباس. (۱۳۹۴). بررسی عوامل تعیین‌کننده سطح نگهداشت موجودی‌های مواد و کالا و رتبه‌بندی آن‌ها با استفاده از الگوریتم‌های درخت تصمیم و شبکه عصبی. فصلنامه حسابداری مالی ۲۷: ۵۲-۳۰.

چالاکی، پری و مرتضی یوسفی. (۱۳۹۱)، پیش‌بینی مدیریت سود با استفاده از درخت تصمیم‌گیری. نشریه مطالعات حسابداری و حسابرسی ۱: ۱۲۳-۱۱۰.

حجازی، رضوان؛ شاپور محمدی؛ زهرا اصلانی و مجید آقاجانی. (۱۳۹۱). پیش‌بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران. بررسی‌های حسابداری و حسابرسی ۲: ۴۶-۳۱.

خدمتگزار سالانقوچ، یاسر. (۱۳۹۲). بررسی خوشه‌بندی در شبکه‌های حسگر بی سیم و ارائه یک متد جدید در این رابطه. پروژه کارشناسی پیوسته کامپیوتر گرایش نرم‌افزار، دانشگاه مهندسی فناوری‌های نوین قوچان.

دادنوش فر، علی؛ محمد رمضان احمدی و ریحانه گسگری. (۱۳۹۰). بررسی رابطه بین مدیریت سود و بازده واقعی و غیرعادی سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. فصلنامه حسابداری مالی ۱۱: ۱۴۳-۱۱۹.

دستگیر، محسن و سید احسان حسینی. (۱۳۹۲). مروری جامع بر مدیریت سود (بخش دوم). نشریه مطالعات حسابداری و حسابرسی ۸: ۱۹-۴.

دستگیر، محسن و مرتضی شفیع سردشت. (۱۳۹۰). فناوری داده‌کاوی؛ رویکردی نوین در حوزه مالی. دانش حسابرسی ۵: ۲۷-۶.

دهدشتی نژاد، علی. (۱۳۹۲). چرا معامله گران سهام خودرویی‌ها را انتخاب می‌کنند؟ تجزیه و تحلیل صنعت خودرو در بورس (۱۳۹۲/۰۶/۱۰). انتشار آنلاین <http://epsnews.blogfa.com/post/11870>

رستمی، مریم. (۱۳۹۰). تأثیر مدیریت سود بر کیفیت گزارشگری مالی در شرکتهای پذیرفته شده در بورس اوراق بهادار تهران. رساله کارشناسی ارشد حسابداری، دانشگاه آزاد اسلامی واحد تهران مرکز.

سروش، علیرضا و ارد شیر بحرینی نژاد. (۱۳۸۸). هو شمندی کسب و کار و داده کاوی: یک استراتژی برای به کارگیری داده ها و برگشت سرمایه. تهران: انتشارات ناقوس. صالحی، مهدی و لاله فرخی پيله رود. (۱۳۹۳). پیش بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم در صنایع کشاورزی و نساجی شرکتهای پذیرفته شده در بورس اوراق بهادار تهران. *مجله اقتصاد کشاورزی* ۴: ۸۹-۱۰۵.

صالحی، مهدی و لاله فرخی پيله رود. (۱۳۹۷). پیش بینی مدیریت سود با استفاده از شبکه عصبی و درخت تصمیم. *پژوهش های حسابداری مالی و حسابرسی* ۳۷: ۲۴-۱. عربصالحی مهدی؛ سیاوش محمودی و علی محمودی. (۱۳۹۴). پیش بینی مدیریت سود با به کارگیری شبکه های عصبی پرسپترون چند لایه دارای دو لایه پنهان در صنعت خودرو و ساخت قطعات. *چهارمین کنفرانس ملی مدیریت و حسابداری*، تهران.

علیزاده، سمیه و سمیرا ملک محمدی. (۱۳۹۳). داده کاوی و کشف دانش گام به گام با نرم افزار Clementine، تهران: انتشارات دانشگاه صنعتی خواجه نصیرالدین طوسی. غضنفری، مهدی؛ سمیه علیزاده و بابک تیمورپور. (۱۳۸۷). داده کاوی و کشف دانش، تهران: انتشارات دانشگاه علم و صنعت ایران.

غفارپور، فرشید. (۱۳۹۱). مقایسه مدل های مختلف پیش بینی اقلام تعهدی اختیاری به منظور شناسایی مدیریت سود. پایان نامه کارشناسی ارشد حسابداری، دانشگاه علامه طباطبائی.

فغانی ماکرانی، خسرو؛ سید حسن صالح نژاد و وحید امین. (۱۳۹۵). پیش بینی مدیریت سود مبتنی بر مدل جونز تعدیل شده با استفاده از مدل شبکه عصبی مصنوعی و الگوریتم ژنتیک. *مجله مهندسی مالی و مدیریت اوراق بهادار* ۲۸: ۱۳۶-۱۱۷.

قاصدی قزوینی، فرشید؛ فرشاد فائزی رازی و فرزانه حیدرپور. (۱۳۹۶). استفاده از روش های داده کاوی در پیش بینی و پاسخ به نیاز در حوزه سرمایه گذاری جسورانه. *دانش مالی تحلیل اوراق بهادار* ۳۵: ۱۱۹-۹۹.

کامیابی، یحیی و مهدیه نورعلی. (۱۳۹۵). بررسی اثر میانجی محافظه‌کاری حسابداری بر رابطه بین عدم تقارن اطلاعات و مدیریت سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. فصلنامه حسابداری مالی ۲۹: ۲۶-۱.

کردستانی، غلامرضا؛ جواد معصومی و وحید بقایی. (۱۳۹۲). پیش‌بینی سطح مدیریت سود با استفاده از شبکه‌های عصبی مصنوعی. مجله پیشرفت‌های حسابداری دانشگاه شیراز ۱: ۱۶۹-۱۹۰.

کیقبادی، امیررضا؛ سمیه فتحی و سمیرا سیف. (۱۳۹۴). رتبه‌بندی میزان تأثیر اقلام کلیدی ترازنامه‌ای و نسبت‌های سودآوری در انتخاب پرتفوی بهینه (با استفاده از تکنیک‌های داده‌کاوی). پژوهش‌های حسابداری مالی و حسابرسی ۲۸: ۸۶-۷۵.

محمودی، علی. (۱۳۹۰). پیش‌بینی مدیریت سود با به‌کارگیری ترکیب مدل شبکه‌های عصبی مصنوعی و درخت‌های تصمیم. پایان‌نامه کارشناسی ارشد حسابداری، دانشگاه علامه طباطبائی.

مشایخ، شهناز؛ زهرا اربابی و معصومه رحیمی‌فر. (۱۳۹۲). بررسی انگیزه‌های مدیریت سود. فصلنامه پژوهش حسابداری ۱: ۷۰-۵۳.

نجاری، معصومه. (۱۳۸۹). پیش‌بینی مدیریت سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار ایران با استفاده از تکنیک‌های داده‌کاوی. پایان‌نامه کارشناسی ارشد حسابداری، دانشگاه فردوسی مشهد.

نوروش، ایرج و سید علی حسینی. (۱۳۸۸). بررسی رابطه بین کیفیت افشاء (قابلیت اتکا و به موقع بودن) و مدیریت سود. بررسی‌های حسابداری و حسابرسی شماره ۲: ۱۳۴-۱۱۷.

ولی‌زاده لاریجانی، اعظم؛ علی رحمانی و شقایق ساده. (۱۳۹۷). رابطه‌ی بین افشای گزارش کنترل‌های داخلی، هزینه‌های نمایندگی و مدیریت سود. فصلنامه دانش حسابداری و حسابداری مدیریت ۲۵: ۴۰-۲۹.

هان، ژیاوی؛ میشلین کامبر و ژان پی. (۱۳۹۲). داده‌کاوی - مفاهیم و روش‌ها، جلد دوم. ترجمه‌نسترن حاجی‌حیدری و سید بهنام خاکباز. (۱۳۹۲). تهران، انتشارات دانشگاه تهران.

Chen, F.H., D.J. Chi, and Y.C. Wang. (2015). Detecting biotechnology industry's earnings management using Bayesian network, principal component analysis, back propagation neural network, and decision tree. *Economic Modelling* 46(40):1-10.

- Chen, W., and Y. Du. (2009). Using Neural Networks And Data Mining Techniques For The Financial Distress Prediction Model. **Expert Systems With Applications** 36: 4075–4086.
- Collins, D.W., and P. Hribar. (2002). Errors in estimating accruals: implications for empirical research. **J. Account. Res** 40 (1): 105–134.
- Dechow, P.M., R.G. Sloan, A.P. Hutton, and J.H. Kim. (2012). Detecting earnings management: a new approach. **J. Account. Res** 50(2): 275–334
- Doyle, J.T., J.N. Jennings, and M.T. Soliman. (2013). Do managers define non-GAAP earnings to meet or beat analyst forecasts?. **Journal of Accounting and Economics** 56(1): 40–56.
- Höglund, H. (2012). Detecting earning management with neural networks. **Expert Systems with Applications** 39(10): 9564–9570.
- Tsai, C.F., and Y.J. Chiou. (2009). Earnings Management Prediction: A pilot Study of combining Neural Networks and Decision Trees. **Expert systems with Application** 36(3): 7183-7191.