مدل‌سازی پیش‌بینی کیفیت سود با استفاده از درخت تصمیم و ماشین‌بردار پشتیبان

لقمان حاتمی شیرکوهی[۱]، صغری براری نوکاشتی[۲*]، مریم اوشک سرایی[۳]

[۱]دانشجوی دکتری، گروه حسابداری، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران.

[۲]استادیار، گروه حسابداری، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران(عهده‌دار مکاتبات)

[۳]استادیار، گروه مدیریت صنعتی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران.

تاریخ دریافت: شهریور ۱۴۰۲، اصلاحیه: مهر ۱۴۰۲، پذیرش: آبان ۱۴۰۲

چکیده

سود و کیفیت آن از مهم‌ترین مولفه‌های تصمیم‌گیری برای استفاده‌کنندگان است. از این‌رو پیش‌بینی کیفیت سود برای سرمایه‌گذاران و سایر گروهای ذینفع از اهمیت بسزایی برخوردار است، بدین منظور از روش درخت تصمیم و ماشین‌بردار پشتیبان برای پیش‌بینی کیفیت سود که هدف اصلی تحقیق می‌باشد استفاده شده است. جامعه آماری پژوهش شامل شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران از سال ۱۳۹۰ سال ۱۴۰۰ به مدت ۱۰ سال می‌باشد، بعد از غربالگری تعداد ۱۱۳ شرکت و ۱۱۳۰ مشاهده به عنوان نمونه آماری انتخاب شدند. در این مطالعه به منظور شناسایی و پیش‌بینی کیفیت سود شاخص‌های مرتبط با حاکمیت شرکتی (استقلال هیئت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی)، سیاست تقسیم سود، تامین مالی بدهی و محافظه کاری به عنوان متغیر مستقل وکیفیت اقلام تعهدی اختیاری نماینده شاخص کیفیت سود به عنوان متغیر وابسته مورد بررسی قرار گرفته است. این پژوهش ازنظر هدف کاربردی و به لحاظ روش پژوهشی، توصیفی از نوع همبستگی. تجزیه تحلیل داده‌ها مطابق با استاندارد داده‌کاوی CRISP-DM و اجرای چهار الگوریتم درخت تصمیم شامل (*CHAID, C۵,۰ , C&R, QUEST*) و ماشین‌بردار پشتیبان انجام شده است. نتایج پژوهش نشان می‌دهد که در هر دو روش درخت تصمیم و ماشین‌بردار پشتیبان، استقلال هیئت مدیره بیشترین تاثیر را در پیش‌بینی کیفیت سود دارد. مقایسه نتایج عملکرد دو روش نشان می‌دهد که، تکنیک ماشین‌بردار پشتیبان به جهت دارا بودن صحت و حساسیت بیشتر و خطای کمتر نسبت به تکنیک درخت تصمیم عملکرد بهتری دارد.

واژه‌های اصلی: کیفیت سود، درخت تصمیم، ماشین‌بردار پشتیبان، اقلام تعهدی، محافظه‌کاری.

۱- مقدمه

گزارشگری مالی اطلاعات حسابداری مفیدی را در اختیار بازار سرمایه برای تصمیم‌گیری تجاری قرار می‌دهد[۲۷]، گزارش‌های مالی بطور گسترده توسط مدیران، سرمایه‌گذاران، تحلیلگران، تنظیم کننده‌ها و استانداردگذاران مالی بطورگسترده مورد استفاده قرارمی‌گیرد. شکل سود در صورت مالی اطلاعات حسابداری را خلاصه می‌کند، در حالی که کیفیت سود برای تصمیم‌گیری اقتصادی استفاده می‌شود[۲۱]. سود با کیفیت بالا با ارایه یک معیار دقیق از عملیات یک شرکت مشخص می‌شود. در واقع چنین سودی (سود با کیفیت) پایدار است؛ درآمدهای آتی را بطور قابل اعتماد پیش‌بینی می‌کند، مدیریت نشده است و اطلاعات ارزشمندی را در اختیار کاربران قرار می‌دهد[۲۵].

کیفیت سود معیاری مهم برای سلامت مالی واحد تجاری به شمار می‌رود و مفهومی چند بعدی است و معنای واحدی ندارد. ارزش سهام یک شرکت، تنها به سود هر سهم سال جاری آن بستگی ندارد، بلکه به انتظارات ما از آیندهٔ شرکت و توان سود آوری آتی بستگی دارد. کیفیت سود جنبهٔ بسیار مهمی از سود حسابداری را نشان می‌دهد، زیرا آگاهی سهامداران از کیفیت سود می‌تواند بر قیمت سهام، میزان سود تقسیمی و در نتیجه بازده سهام، تأثیرگذار باشد. به عنوان مثال در مورد تاثیرات کیفیت سود بر بازده سهام تحقیقات مختلفی انجام شده است. دومینگوس[۱۳] نشان دادند افزایش نوسانات خاص با بدتر شدن کیفیت سود مرتبط است؛ اما تحقیقات موگوتو و موزینداتسی[۱][۲۵] نشان داد معیارهای مختلفی برای کیفیت سود وجود دارد و هریک از معیارهای کیفیت سود منحصر به فرد است و جنبه خاصی از سود گزارش شده یک شرکت را نشان می‌دهد؛ بنابراین هریک از معیارهای کیفیت سود می‌تواند به طور متفاوتی بر نوسانات بازده سهام تاثیر بگذارد و مطابق با این نگرش می‌توان گفت که کیفیت سود یک مفهوم چند بعدی است. معیارهای متعدی از جمله کیفیت اقلام تعهدی، محافظه کاری، به موقع بودن، ارتباط ارزشی، تداوم سود، قابلیت پیش‌بینی و یکنواختی سود برای

[۱] Muguto & Muzindutsi

*Brari@iaurasht.ac.ir

اندازه گیری آن استفاده شده است[۱۶]. اگرچه در مجموع هیچ دستورالعمل خاصی در مورد انتخاب معیارهای سود وجود ندارد، اما توصیه می‌شود از معیارهای مختلفی استفاده شود زیرا هر معیار کیفیت سود جنبه خاصی از سود گزارش شده یک شرکت را نشان می‌دهد[۱۱،۱۹].

از آنجایی که واکنش‌های متفاوت سرمایه‌گذاران نسبت به اطلاعات سود، موجب واکنش‌های متفاوت بازار می‌گردد، لذا درک عوامل تأثیرگذار بر کیفیت سود حائز اهمیت به نظر می رسد. فعالان بازار همواره خواهان کیفیت بالای گزارشگری مالی می‌باشند. زیرا کیفیت بالای گزارشگری مالی باعث کاهش عدم تقارن اطلاعاتی و همچنین افزایش شفافیت و بهبود قراردادها می‌گردد. کیفیت بالای صورت‌های مالی به وسیلهٔ کیفیت سود تخمین زده می‌شود. چون سود پایه و اساس تصمیمات سرمایه‌گذاران است[۲۶]. البته همین سود حسابداری به دلایل متنوعی ممکن است کیفیت لازم را نداشته باشد و به جای اینکه به استفاده‌کنندگان از صورت‌های مالی کمک کند زمینه گمراهی آنها را فراهم کند. هدف اصلی این تحقیق بررسی عوامل موثر بر کیفیت سود و کمک به تصمیم‌گیرندگان از جمله سرمایه‌گذاران در شناسایی این موضوع که آیا سود گزارش شده از طرف مدیریت در صورت‌های مالی از کیفیت لازم برخوردار هست یا خیر؟ هدف دیگر تحقیق بیان این موضوع است که تمامی تحقیقات قبلی انجام شده فقط ارتباط یک متغیر مستقل با متغیر کیفیت سود (متغیر وابسته) را مورد مطالعه قرار دادند ولی این تحقیق به طور همزمان تاثیر چندین متغیر موثر بر کیفیت سود را مورد مطالعه قرار می‌دهد تا نتایج آن مورد استفاده تمامی تصمیم‌گیرندگان و استفاده‌کنندگان از صورت‌های مالی قرار گیرد. هدف دیگر این پژوهش این است که با بررسی جامع داده‌های حسابداری شرکت‌های پذیرفته شده در بورس اوراق بهادار برای اولین بار در ایران با استفاده از رویکرد ماشین‌بردار پشتیبان و الگوریتم‌های درخت تصمیم به شناسایی و پیش بینی کیفیت سود اقدام کند و با توجه به کاستی‌های الگوهای طبقه‌بندی شده و پیش‌بینی مرسوم در این پژوهش به منظور شناسایی و پیش‌بینی کیفیت سود از ماشین بردار پشتیبان و چهار الگوریتم درخت تصمیم استفاده نماید.

۲- مبانی نظری و ادبیات پژوهش

در حال حاضر تعریف کیفیت سود و نحوه اندازه‌گیری آن چالش بزرگی در دنیای تئوری حسابداری و تحقیقات تجربی حسابداری محسوب می‌شود. در بعد تئوریک، عده‌ای کیفیت سود را حاصل بیان صادقانه سود معرفی نموده و عده‌ای نیز نزدیکی سود حسابداری به مفهوم سود اقتصادی را ضابطه کیفیت سود قلمداد می‌کنند. ضمنا در تئوری حسابداری مفاهیم ساختاری، تفسیری و رفتاری (عملگرایانه) سود تبیین شده که هر کدام معنای خاص خود را دارد که در صورتی که بتوان این مفاهیم را به تعاریف کیفیت سود مرتبط نمود، شاید قدری ابهام را برطرف نموده و زمینه ساز فراهم شدن اجماع درباره ماهیت کیفیت سود گردد[۹]. تعاریف متعددی برای کیفیت سود مطرح شده است. دیچو و دیچف[۲][۱۰] سود را بر اساس ارتباط اقلام تعهدی با جریانات نقدی عملیاتی دوره گذشته، حال و آینده تعریف می‌کنند. در این صورت، کیفیت سود ناشی ازکیفیت اقلام تعهدی خواهد بود. آنها کیفیت سود را برای شرکت‌هایی که واریانس مجموع خطاهای آنها با هم برابر است را یکسان می‌دانند. بال و شیوا کومار[۳][۴] مفهوم کیفیت را سودمندی صورت‌های مالی برای سرمایه‌گذاران، اعتباردهندگان، مدیران و تمامی بخش‌های مرتبط با شرکت تعریف می‌کنند.

عوامل متعددی بر کیفیت سود تاثیر دارند در این پژوهش ویژگی‌های حاکمیت شرکتی (استقلال هیت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی)، محافظه‌کاری، سیاست تقسیم سود و تامین مالی از طریق بدهی مورد بررسی قرار گرفته‌اند. تئوری نمایندگی رابطه بین مالکان و مدیران شرکت را توصیف می‌کند و در این حالت ممکن است مدیران در جهت افزایش منافع مالکان عمل نکنند و یک نوع تضاد منافع بین آنها ایجاد شود. تئوری نمایندگی یعنی تضاد بالقوهای بین منافع نماینده و کارفرما وجود دارد، که ممکن است منجربه بروز یک درگیری بالقوه شود که بر کیفیت سود گزارش شده تأثیر بگذارد. مدیریت با سود بهره خاص تمایل به تهیه گزارش متناسب با هدف و نه در جهت منافع اصلی دارد. در چنین شرایط، یک سازوکار کنترلی برای همسویی منافع دو طرف لازم است[۷]. حاکمیت شرکتی مفهومی است که برای بهبود عملکرد مدیریت از طریق نظارت به منظور اطمینان از پاسخگویی به سهامداران طراحی شده است. حاکمیت شرکتی علاوه بر اینکه می‌تواند تضاد نمایندگی را کاهش دهد، توانایی ایجاد ارزش افزوده برای سهامداران به صورت حمایت موثرتر، به ویژه برای سرمایه‌گذاران برای بازیابی سرمایه‌های خود به روشی طبیعی و با ارزش بالا را دارد[۲۸]. مکانیزم‌های حاکمیت شرکتی می تواند فرصت‌های مدیریت سود را کاهش و در نتیجه کیفیت سود را افزایش دهد. محافظه‌کاری یکی از عوامل مهم بر کیفیت سود است. حسابداران بطور سنتی محافظه‌کاری را از طریق قاعده"هیچ سودی را پیش‌بینی نکن اما همه زیانها را پیش‌بینی کم بیان نموده اند"[۶]. باسو این قاعده را به معنی تمایل حسابداران به الزام درجه بالاتری از تاییدپذیری در شناسایی اخبار خوب و سودها نسبت به اخبار بد و زیانها تفسیر می‌نماید. گیولی و هاین[۴][۱۸] محافظه کاری از طریق اقلام تعهدی اختیاری (غیر عملیاتی) اندازه‌گیری می کنند، زیرا به نظر آنها حسابداری تعهدی محلی برای اجرای محافظه‌کاری است. براساس مفهوم محافظه‌کاری واحدهای تجاری در شرایط مختلف روش‌های را انتخاب می‌کنند که حداکث اثر فزاینده بر سود و جمع دارایی‌ها داشته باشد و این نحوه عمل می‌تواند تاثیر زیادی بر محتوای اطلاعاتی تهیه شده توسط سیستم داشته باشد و هر چقدر میزان محافظه‌کاری بیشتر باشد کیفیت سود بالاتر است. تحقیقات مختلف نشان می‌دهد که بین کیفیت اقلام تعهدی در شرکت و سیاست تقسیم سود ارتباط وجود دارد. کیفیت اقلام تعهدی به عنوان شاخصی ازکیفیت

---

۲ Dechow & Dichev

۳ Ball & Shivakumar

۴ Givoly & Hayn

گزارش‌های مالی رابطه مستقیمی با سیاست تقسیم سود شرکت دارد. شرکت‌هایی که میزان سود خود را در گزارش‌های مالی اظهار می‌کنند، سیاست‌هایی متناظر با سود اعلام شده، در تقسیم سود شرکت لحاظ می‌کنند[۲۳]. فاما و فرنچ[۵][۱۵] نشان دادند نسبت شرکت‌های آمریکایی که بطور مستمر سود تقسیمی پرداخت می‌کنند در سال ۱۹۹۹ نسبت به سال ۱۹۷۸ روند نزولی داشته است. شواهد حاکی از این است که سود تقسیمی فراهم کننده اطلاعاتی درباره کیفیت سود گزارش شده است[۳۱]. تامین مالی از طریق بدهی بر کیفیت سود تاثیر دارد و در این مورد سه دیدگاه مطرح است. دیدگاه اول تامین مالی از طریق بدهی تاثیر مثبت بر کیفیت سو دارد. در این دیدگاه اعتباردهندگان نسبت به کیفیت اطلاعات حسابداری وام‌گیرندگان واکنش نشان می‌دهند. بدین معنی هرچه کیفیت اطلاعات حسابداری وام‌گیرندگان بیشتر باشد شرایط محدود کننده کمتر و نرخ بهره پایین‌تری را در قراردادهای وام آنها در نظر می‌گیرند و بالعکس. در نتیجه اعتبارگیرندگان سعی می‌کنند با گزارش اطلاعات با کیفیت بالا، هزینه بهره و سایر شروط محدود کننده را کاهش دهند[۲]. دیدگاه دوم تامین مالی از طریق بدهی بر کیفیت سود اثر منفی دارد، بدین معنی که مدیران شرکت‌ها ممکن است برای جلب نظر مثبت اعتباردهندگان، سود را دستکاری کرده تا بتوانند تصویر مطلوبی از توان سودآوری شرکت را به نمایش گذاشته تا سرمایه مورد نیاز را جذب نمایند[۱۲]. در نتیجه می‌توان گفت که بدهی اثر منفی بر کیفیت سود دارد. دیدگاه سوم رابطه بین نامین مالی بدهی و کیفیت سود یک رابطه مستقیم یا معکوس نیست بلکه نوعی رابطه سهمی گون است که در نهایت بوسیله کنش و واکنش بین اثرهای مثبت و منفی تامین مالی از طریق بدهی برکیفیت سود تعیین می‌شود[۱۷].

۳- پیشینه تجربی پژوهش

موگوتو و موزینداتسی[۲۵] در پژوهشی با عنوان معیارهای کیفیت سود و نوسانات بازده سهام در آفریقای نشان دادند که کیفیت اقلام تعهدی و تداوم سود رابطه منفی با نوسانات خاص دارند. علاوه بر این مشخص شد که یکنواختی سود بطور مثبت بر نوسانات خاص تاثیر می‌گذارد و شرکتهایی با سود کمتر یکنواخت، نوسانات بازده سهام بیشتری دارند.

وانگ و توماس[۶][۳۰] به بررسی ارتباط بین مسئولیت اجتماعی شرکت و کیفیت سود در میان شرکت های استخراجی موزامبیک طی سال‌های ۲۰۱۷-۲۰۱۰ پرداخته‌اند. نتایج حاصل از آزمون فرضیه‌های پژوهش با استفاده از مدل رگرسیون داده های تابلویی نشان داد که بین مسئولیت اجتماعی شرکت و کیفیت سود رابطهٔ منفی و معناداری وجود دارد. این یافته‌ها با این ایده سازگار هستند که مدیران فرصت طلب از مسئولیت اجتماعی برای رسیدن به منافع خاص خود استفاده می کنند. چن و کومال[۷][۸] در پژوهشی رابطهٔ بین تخصص مالی کمیتهٔ حسابرسی و کیفیت سود شرکت‌ها را با استفاده از تکنیک فراتحلیل نتایج ۹۰ مطالعهٔ صورت گرفته در سال‌های ۲۰۱۶-۲۰۰۳ مورد بررسی قرار داده‌اند. نتایج این پژوهش شان داد که تخصص مالی کمیتهٔ حسابرسی رابطهٔ مثبت و معناداری با کیفیت سود دارد و متخصصان مالی حسابداری رابطهٔ قویتری با کیفیت سود نسبت به متخصصان مالی غیر حسابداری دارند. از طرفی سیستم‌های حاکمیت شرکتی و استانداردهای گزارشگری مالی بین‌المللی رابطهٔ بین تخصص مالی کمیتهٔ حسابرسی و کیفیت سود را تعدیل می‌کنند.

ستین و مروانیگساری[۸][۲۹] تأثیر توانایی مدیریت در افزایش کیفیت سود با در نظر گرفتن اثر تعدیلی کمیتهٔ حسابرسی در ۱۲ شرکت تولیدی دولتی اندونزی در بین سال‌های ۲۰۱۸-۲۰۱۶ را مورد بررسی قرار داده‌اند. نتایج حاصل از برآورد مدل رگرسیون چند متغیرهٔ پژوهش نشان داد که بین توانایی‌های مدیریت و کیفیت سود رابطهٔ مثبت و معناداری وجود دارد. همچنین کمیتهٔ حسابرسی تأثیر توانایی مدیریتی را در قبال کیفیت سود تقویت می‌کند.

استبان و گارسیا[۹][۲۶] در پژوهش خود به بررسی تأثیر ویژگی‌های شرکت بر کیفیت سود پرداختند. جامعهٔ آماری این پژوهش هتل‌های کشور اسپانیا بین سال‌های ۲۰۰۰ تا ۲۰۱۱ می‌باشد. ایشان در پژوهش خود برای اندازه‌گیری کیفیت سود از یک اندازه‌گیری چند بعدی که شامل پایداری سود و قابلیت پیش‌بینی و هموارسازی سود می‌باشد استفاده کرده و متغیرهایی مانند متغیرهای بازار و استراتژی‌های بازار و ساختار مالکیت و شاخص‌های حسابرسی به عنوان ویژگی‌های شرکت مورد استفاده قرار گرفت. نتایج تحقیق نشان داد که متغیرهای بازار و ساختار مالکیت و شاخص‌های حسابرسی بر کیفیت سود تأثیرگذار بوده ولی سایر متغیرهای ویژگی‌های شرکت تأثیری بر کیفیت سود ندارند.

ایورت و واگن هوفر[۱۰][۱۴] به بررسی روابط میان استانداردهای حسابداری، مدیریت سود و کیفیت سود پرداختند نتایج نشان می‌دهد که کیفیت سود با افزایش دقت حسابداری افزایش می‌یابد. با ملاحظه پژوهش‌های بالا می‌توان گفت که اکثر آنها اثر یک متغیر مستقل وابسته بر یک متغیر وابسته را مورد بررسی قرار داده‌اند و از روش‌های ساده رگرسیون و غیر هوشمند برای تجزیه تحلیل داده‌ها و آزمون فرضیه‌ها استفاده نموده‌اند. از آنجایی که همیشه بین متغیر مستقل و متغیر وابسته رابطه خطی وجود ندارد، از این‌رو در چنین مواقعی باید از روش‌های دیگری مانند درخت تصمیم و ماشین بردار پشتیبان که جز روش‌های داده‌کاوی هستند و از قدرت پیش‌بینی کنندگی بالایی برخوردار هستند و میزان صحت و قدرت دقت آن بیشتر از روش‌های خطی هست استفاده نمود. علاوه بر استفاده از روشهای الگوریتم درخت تصمیم و ماشین‌بردار پشتیبان استفاده شده در این تحقیق برای پیش‌بینی کیفیت سود به طور همزمان اثر چندین متغیر بر کیفیت سود مورد بررسی قرار گرفته است تا در یک نگاه همه جانبه اطلاعات جامعی را در اختیار سرمایه‌گذاران و سایر گروه‌های ذینفع قرار دهد.

---

۵ Fama & French
۶ Siueia & Wang
۷ Chen & Komal
۸ SeTin & Murwaningsari
۹ Parte-Esteban & García
۱۰ Ewert & Wagenhofer

۳- روش‌شناسی پژوهش

این پژوهش یک پژوهش توصیفی، و در حوزه پژوهش‌های همبستگی است که مبتنی بر اطلاعات واقعی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است. داده‌های لازم برای اجرای پژوهش از صورت‌های مالی و گزارشات هیأت مدیره شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در شرکت‌های صنایع تولیدی از سال ۱۳۹۰ تا سال ۱۴۰۰ و همچنین از سایت کدال جمع‌آوری شده است. تجزیه و تحلیل داده‌ها با استفاده از نرم افزار اکسل و ایویوز انجام شده برای استخراج مدل از روش الگوریتم درخت تصمیم و ماشین بردار پشتیبان استفاده شده است.

۴-۱- فرضیه پژوهش

فرضیه اصلی پژوهش در این تحقیق در قالب سوال مطرح می‌شود به دلیل اینکه روش‌های روش‌های داده‌کاوی مانند روش‌های آماری از آمار استنباطی برای تحلیل فرضیه‌ها استفاده نمی‌کنند و صرفا روشی هوشمند برای پیش‌بینی متغیر وابسته هستند، بنابراین فرضیه پژوهش را می‌توان به بصورت زیر بیان کرد. "آیا استفاده از الگوریتم‌های درخت تصمیم و ماشین بردار پشتیبان بر پیش‌بینی دقت کیفیت سود موثر است یا خیر"

۴-۲- جامعه و نمونه آماری پژوهش

جامعه آماری مورد بررسی در این پژوهش شامل شرکتهای پذیرفته شده در بورس اوراق بهادار تهران از سالهای ۱۳۹۰ سال ۱۴۰۰ به مدت ۱۰ سال می‌باشد. به علت گستردگی حجم جامعه آماری و وجود برخی ناهماهنگی‌ها میان اعضای جامعه، شرایط زیر برای انتخاب نمونه آماری قرار داده شده و نمونه آماری به روش غربالگری (حذف سیستماتیک) انتخاب شده است. شرکت‌هایی به عنوان نمونه انتخاب می‌شوند که دارای شرایط زیر باشند:

- سال مالی آنها منتهی به پایان اسفند باشد.
- جز شرکت‌های سرمایه‌گذاری و واسطه‌گری مالی نباشد.
- طی دوره مورد پژوهش سال مالی خود را تغییر نداده باشد.

در نهایت با توجه به موارد اشاره شده در بالا، از مجموع شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، داده‌های مربوط به بازه زمانی ۱۰ ساله از سال ۱۳۹۰ تا ۱۴۰۰ شامل ۱۱۳ شرکت جمعا ۱۱۳۰ سال شرکت به عنوان نمونه در این پژوهش استفاده شده است.

جدول (۱): صنایع و شرکت‌های مورد مطالعه

نوع صنعت	تعداد شرکت در نمونه	نوع صنعت	تعداد شرکت در نمونه
خودرو و قطعات	۱۸	کانه فلزی	۱۰
دارو	۲۰	قند شکر	۲
صنایع شیمایی	۸	لاستیک پلاستیک	۳
صنایع غذایی	۷	ماشین الات و تجهیزات	۷
صنعت کاشی و سرامیک	۳	محصولات فلزی	۱
صنعت سیمان	۱۲	دستگاههای برقی	۵
فلزات اساسی	۹	قند و شکر	۲
کانه غیر فلزی	۵	حمل نقل	۱
جمع کل	۱۱۳		

۴-۳- نحوه محاسبه متغیرها

این پژوهش یک پژوهش توصیفی و در حوزه پژوهش‌های همبستگی است که مبتنی بر اطلاعات واقعی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است. در این پژوهش متغیر وابسته کیفیت سود و متغیرهای مستقل عبارتند: تامین مالی بدهی، استقلال هیئت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی، تقسیم سود، محافظه‌کاری.

متغیر وابسته: جهت اندازه‌گیری کیفیت سود معیارهای مختلفی وجود دارد. در این پژوهش برای اندازه‌گیری آن از مدل تعدیل شده جونز[۱۱][۲۰] استفاده شده است. در این پژوهش متغیر وابسته، کیفیت سود می‌باشد که از طریق کیفیت اقلام تعهدی اختیاری[۱۲] (*DA*) اندازه‌گیری شد. کیفیت اقلام تعهدی نیز از طریق رابطه زیر محاسبه

[۱۱] Jones
[۱۲] Discretionary Accruals (DA)

گردید.

$$TACC = \beta_0 + \frac{\beta_1}{TA_{(t-1)}} + \frac{\beta_2 (\Delta REV_t - \Delta REC_t)}{TA_{(t-1)}} + \frac{\beta_3 PPE}{TA_{(t-1)}} + \varepsilon_t$$

$$TACC = (\Delta CA - \Delta Cash) - (\Delta CL - \Delta STD)$$

$$NDA_t = \beta_0 + \frac{\beta_1}{TA_{(t-1)}} + \frac{\beta_2 (\Delta REV_t - \Delta REC_t)}{TA_{(t-1)}} + \frac{\beta_3 PPE}{TA_{(t-1)}}$$

$$DA_t = TACC_t - NDA_t = \varepsilon_t$$

اقلام تعهدی کل $TACC_t$

اقلام تعهدی غیراختیاری NDA_t

اقلام تعهدی اختیاری DA_t

تغییرات در آمد سال جاری به سال قبل ΔREV_t

تغییرات حسابها و اسناد دریافتنی سال جاری به سال قبل ΔREC_t

جمع کل دارایی‌های ابتدای دوره TA_{t-1}

پسماند مدل رگرسیونی ε_t

دارایی‌های ثابت (اموال، ماشین آلات و تجهیزات) PPE

تغییرات بدهی‌های جاری ΔCL

تغییرات دارایی‌های جاری ΔCA

به طور کلی جهت پیش‌بینی کیفیت سود، میانه توزیع این متغیر در میان سال شرکت‌های مختلف استخراج شد و با توجه به منطق زیربنایی آن‌ها به شرکت‌های بالا یا پایین میانه امتیاز (یک و صفر) داده شد. از آنجا که هرچه اقلام تعهدی اختیاری بیشتر باشد، کیفیت سود کمتر است، پس به شرکت‌های بالای میانه عدد صفر و به شرکت‌های زیر میانه عدد یک تعلق گرفت که در جدول ۲ مشاهده می‌شوند.

جدول (۲): چگونگی امتیازدهی و تعیین سطح کیفیت سود

سطوح کیفیت سود	کیفیت اقلام تعهدی	امتیاز
با کیفیت	۱	۱
کیفیت کمتر	۰	۰

متغیرهای مستقل. در این پژوهش با توجه به تحقیقات قبلی متغیرهای مستقل انتخاب شدند که در جدول ۳، علامت اختصاری و نحوه محاسبه هریک از آن‌ها نشان داده شده است.

جدول (۳): نحوه محاسبه متغیرهای مستقل پژوهش

علامت اختصاری	متغیر	نحوه محاسبه
PS	سیاست تقسیم سود	میزان سود تقسیمی توسط شرکت‌ها معیار اندازه گیری سیاست تقسیم سود است که از نسبت سود تقسیمی به سود خالص شرکت بدست می‌آید.
$Dept$	تامین مالی از طریق بدهی	از نسبت مجموع بدهی‌ها (بدهی‌های جاری و غیر جاری) به مجموع دارایی‌ها بدست می‌آید.
$Conser$	محافظه کاری	در این پژوهش با توجه به ارتباط بین کیفیت اقلام تعهدی و کیفیت سود محافظه‌کاری غیرشرطی مد نظر بوده و مطابق مدل گیولی و هاین (۲۰۰۰) از نسبت جمع کل اقلام تعهدی به جمع کل دارایی‌های شرکت بدست می‌آید.
BOD	استقلال هئیت مدیره	استقلال هیت مدیره برابر است با نسبت اعضای هیت مدیره غیر موظف به کل اعضای هیت مدیره
AC	استقلال کمیته حسابرسی	استقلال کمیته حسابرسی برابر است نسبت مدیران غیر اجرایی کمیته حسابرسی به کل مدیران
OOW	مالکیت سازمانی	مالکیت سازمانی برابر است نسبت سهم نگه داری شده توسط مالکان نهادی به کل سهام

۴-۴- روش درخت تصمیم

الگوریتم‌های درخت تصمیم یکی از رهیافت‌های ممکن برای تصمیم‌گیری چندمرحله‌ای می‌باشند. ایده اصلی در تصمیم‌گیری چند مرحله‌ای این است که تصمیمیات پیچیده به چند تصمیم ساده تقسیم شوند و در نهایت با ترکیب این تصمیمات ساده، به تصمیم مورد انتظار دست پیدا کنیم. درخت تصمیم زیرمجموعه‌ای از تصمیم‌گیری‌های سلسله مراتبی می‌باشد (پریجایی مقدم و موسوی، ۲۰۱۲). الگوریتم‌های زیادی در طول سال‌های گذشته برای ساخت درخت تصمیم ایجاد شده‌اند، همانند: C۴,۵ (۱۹۹۳)، C&RT (۱۹۸۴)، SPRINT (۱۹۹۶)، SLIQ (۱۹۹۶). یکی از آخرین مطالعاتی که درخت تصمیم و دیگر الگوریتم‌های یادگیری را مقایسه کرد، به وسیله Tjen-Sien و همکارانش در سال ۲۰۰۰ انجام شد. این مطالعه نشان داد که C. ۴,۵ یک ترکیب بسیار خوب از خطا و سرعت می‌باشد. در C۴,۵ فرض بر این است که کل داده‌های آموزشی در داخل حافظه باشند، بنابراین گرک و همکارانش، چارچوبی برای رشد سریع درخت تصمیم پیشنهاد کردند که به خوبی با اندازه حافظه در دسترس قابل انطباق می‌باشد[۲۲]. درخت تصمیم و شبکه عصبی

متدهای ترتیبی برای دسته‌بندی نمونه‌ها می‌باشند..

۴-۵- ماشین بردار پشتیبان

ماشین بردار پشتیبان تکنیک های یادگیری آماری است که ابتدا توسط کورتز و واپنیک در سال ۱۹۹۵معرفی شدند و در بسیاری از مسائل طبقه بندی بسیار موفق بوده‌اند. ماشین بردار پشتیبان بر اساس مفهوم تصمیم‌گیری است و مرزهای تصمیم‌گیری را تعریف می‌کند. یک سطح تصمیم‌گیری آن است که بین مجموعه‌ای از کلاس‌های مختلف جداشده است. این الگوریتم‌ها یک سطح بالاتر را به عنوان یک سطح تصمیم‌گیری که نمونه‌ها را به دو کلاس مثبت و منفی طبقه‌بندی می‌کند، در نظر می‌گیرند. این مدل در مقایسه با مدل‌های دیگر دقت و کارایی بالاتری دارد. به منظور دستیابی به حداکثر دقت پیش‌بینی، این الگوریتم نیازمند داده‌های با اندازه بزرگ برای آموزش است[۵].

۴-۶- روش اجرایی درخت تصمیم

در این پژوهش از روش درخت تصمیم[۱۳] بر اساس قوانین "اگر-آنگاه" با استفاده از مخزن داده‌های تمام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران که قابل دسترس است، ایجاد شد. مخزن داده پژوهش حاضر شامل ۱۱۳۰ رکورد و ۷ ویژگی است، که مورد تجزیه و تحلیل قرار گرفتند. هدف از انجام این پژوهش پیش‌بینی کیفیت سود با استفاده از معیارهای استقلال هیئت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی، سیاست تقسیم سود، تامین مالی بدهی، و محافظه کاری است. لذا ویژگی کیفیت سود به عنوان برچسب داده ها تعیین شد و مساله از نوع مسائل با ناظر می‌باشد که برای حل آن باید از روش‌های دسته‌بندی[۱۴] استفاده نمود. برای انجام دسته‌بندی از الگوریتم درخت تصمیم با استفاده از نرم‌افزار Rapid miner استفاده شده است.

جدول (۴): اطلاعات مختصر پایگاه داده پژوهش

ویژگی پایگاه‌داده	تعداد ویژگی‌ها	تعداد رکوردها	روش داده‌کاوی	داده‌های گم شده
چندمتغیره	۷	۱۱۳۰	دسته‌بندی	ندارد

بر اساس مطالب گفته شده، فرایند انجام درخت تصمیم به شرح زیر می‌باشد:

- تهیه اطلاعات در سطح کل بورس
- انجام محاسبات نسبت‌های مالی و شاخص‌های پیش‌بینی کننده کیفیت سود
- اجرای مدل‌های درخت تصمیم
- ترکیب نتایج مدل‌های الگوریتم درخت تصمیم با کمک تحلیل حساسیت

[۱۳] Decision tree
[۱۴] Classification

۵- نتایج

۵-۱- آماده‌سازی و پیش پردازش داده‌ها

این مرحله یکی از مهم‌ترین و اثرگذارترین مراحل داده کاوی است و ۸۰ درصد زمان فرایند داده‌کاوی را به خود اختصاص می‌دهد. در این پژوهش از متدولوژی CRISP DM به عنوان فرایند استانداردسازی داده‌کاوی استفاده شد. این روش اولین بار در سال ۱۹۹۶ ارائه شد. این متدولوژی یک چرخه شش مرحله‌ای دارد و از گام های شناخت کسب وکار، شناخت داده‌ها، آماده‌سازی داده‌ها، مدل‌سازی، ارزیابی مدل و استقرار تشکیل شده است. داده‌های مورد استفاده در این پژوهش، داده‌های شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است.

پس از نرمال سازی داده‌ها در مرحله پیش پردازش داده‌ها، کل داده‌ها به دو بخش ۷۰ درصد داده‌های آموزشی برای ساخت مدل و ۳۰ درصد داده‌های آزمایشی برای اعتبارسنجی مدل تقسیم گردید. کلیه مراحل در نرم‌افزار Rapid miner انجام شده است. جدول ۵ اطلاعات آماری پژوهش حاضر را نشان می‌دهد.

جدول (۵): اطلاعات آماری پژوهش

نام	مقدار
حداقل	-۱
حداکثر	+۱
انحراف معیار	۰,۱۲۵
میانگین	۰,۰۵۱

در مرحله بعدی، با کمک درخت‌های (*CHAID, C۵,۰ , C&RT , QUEST*)، قدرت برازندگی شاخص‌های مورد بررسی بر کیفیت سود تعیین شده و با استفاده از ماتریس انطباق، عملکرد هر یک از درخت‌های تصمیم مورد بررسی قرار گرفته است.

۵-۲- مدل‌های درخت تصمیم با کمک تحلیل حساسیت

در نهایت نتایج چهار درخت تصمیم با استفاده از تحلیل حساسیت ترکیب می‌شود و مدل نهایی شکل می‌گیرد تا اعتباردهی هر یک از نسبت ها با دقت بیشتری انجام شود. تحلیل حساسیت تکنیکی مبتنی بر ترکیب اطلاعات است که با اجرای آن می‌توان نتایج چهار درخت تصمیم را ترکیب کرد تا مدل توانمند و صحیح‌تری ارائه کرد. شش متغیر به عنوان متغیرهای ورودی الگوریتم تعریف شده است. برای تعریف متغیر خروجی از کیفیت سود استفاده شده است. برای پیش‌بینی کیفیت سود، چهار نوع درخت تصمیم به کار برده شد: CHAID,C&RT, C۵,۰ , QUEST. ابتدا ۷۰٪ داده‌ها برای آموزش و ۳۰٪ آن ها برای تست در نظر گرفته شدند. همان طور که در جدول ۶ مشخص است شاخص استقلال هیئت مدیره در چهار الگوریتم CHAID و C۵,۰ و C&RT و QUEST بالاترین اهمیت را به خود اختصاص داده است. با توجه به

تشریح معیارهای عملکرد درخت تصمیم، ابتدا با پیاده‌سازی ماتریس انطباق چهار درخت تصمیم (جدول ۷) به ارزیابی عملکرد هر یک از مدل های درخت تصمیم در جدول ۸ پرداخته شد.

جدول (۶): میزان اهمیت هریک از متغیرهای مستقل

مدل	CHAID	C۵,۰	C&RT	QUEST
BOD	۰,۶۳۶	۰,۴۸۷	۰,۵۳۹	۰,۴۷۴
AC	۰,۲۱۵	۰,۱۴۲	۰,۰۸۷	۰,۹۶
OOW	۰,۱۷۸	۰,۱۸۹	۰,۲۱۴	۰,۱۸۸
PS	۰,۱۸۷	۰,۲۶۳	۰,۲۱۸	۰,۲۶۹
Dept	۰,۱۶۶	۰,۲۲۹	۰,۰۱۹	۰,۲۴۸
Conser	۰,۰۴۱	۰,۰۴۹	۰,۱۳۲	۰,۰۴۹

جدول (۷): ماتریس اختلال (انطباق) مدل‌های درخت تصمیم

نوع مدل		ناموفق (۰)	موفق (۱)			صحت کلی (٪)	صحت هر گروه (٪)
CHAID	ناموفق (۰)	۶۰۹	۳۵	صحیح	۱۰۲۷	۹۳,۷۵	۹۴,۴۵
	موفق (۱)	۳۸	۴۴۸	غلط	۷۳	۶,۲۵	۹۲,۷۵
	جمع	۶۴۷	۴۸۳		۱۱۳۰		
C&RT	ناموفق (۰)	۶۳۶	۲۳	صحیح	۱۰۹۶	۹۷,۰۹	۹۸,۳۹
	موفق (۱)	۱۱	۴۶۰	غلط	۳۴	۲,۹۱	۹۵,۲۴
	جمع	۶۴۷	۴۸۳		۱۱۳۰		
C۵,۰	ناموفق (۰)	۶۱۰	۲۵	صحیح	۱۰۶۸	۹۴,۶۹	۹۴,۶۰
	موفق (۱)	۳۷	۴۵۸	غلط	۶۲	۵,۳۱	۹۴,۸۲
	جمع	۶۴۷	۴۸۳		۱۱۳۰		
QUEST	ناموفق (۰)	۵۹۳	۱۰۲	صحیح	۹۷۴	۸۶,۶۴	۹۲,۱۲
	موفق (۱)	۵۴	۳۸۱	غلط	۱۵۶	۱۳,۳۶	۷۸,۸۸
	جمع	۶۴۷	۴۸۳		۱۱۳۰		

جدول (۸): ارزیابی عملکرد مدل‌های درخت تصمیم

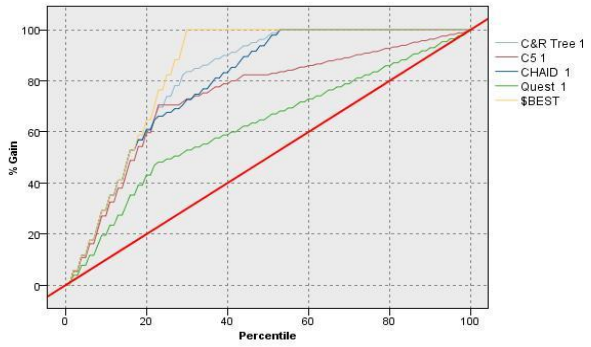
نوع مدل	صحت کلی	دقت	حساسیت	ویژگی	نسبت مثبت کاذب	نسبت منفی کاذب	معیار f	منطقه زیر منحنی
CHAID	۰,۹۳۴	۰,۹۲۵	۰,۹۲۹	۰,۹۴۲	۰,۰۵۲	۰,۰۷۴	۰,۹۲۱	۰,۹۸۴
C&RT	۰,۹۷۱	۰,۹۵۸	۰,۹۷۳	۰,۹۶۸	۰,۰۳۹	۰,۰۲۹	۰,۹۶۸	۰,۹۷۴
C۵,۰	۰,۹۴۸	۰,۹۴۹	۰,۹۲۷	۰,۹۶۹	۰,۰۳۸	۰,۰۷۸	۰,۹۳۳	۰,۹۷۳
QUEST	۰,۸۶۳	۰,۷۸۴	۰,۸۷۲	۰,۸۶۱	۰,۱۳۶	۰,۱۲۶	۰,۸۳۸	۰,۹۳۹

همانطور که در جدول ۵ مشخص است با توجه به تمام معیارهای ارزیابی مدل به جز معیار منطقه زیر منحنی، درخت C&RT بهترین سطح عملکرد را داشته است و با توجه به معیار منطقه زیر منحنی درخت CHAID عملکرد نسبتا بهتری داشته است. درخت C۵,۰ دومین سطح عملکرد را بر اساس تمام معیارها به خود اختصاص داده است. با این حال درخت CHAID با صحت کلی ۹۳/۷۵ درصد و درخت QUEST با صحت کلی ۸۶/۶۴ درصد عملکرد خوبی را داشته اند. با توجه به جدول ۵ می توان دید که صحت گروه ناموفق در سه درخت CHAID, C&RT و QUEST بالاتر از صحت گروه موفق بوده است ولی در درخت C۵,۰ با تفاوت کمی، صحت گروه موفق بالاتر از صحت گروه ناموفق بوده است.

نمایش نموداری معیارهای عملکرد با استفاده از نمودار کیفیت در شکل ۱ به تصویر کشیده شده است. کیفیت به صورت نسبتی از کل موفقیت‌ها است که در هر چارک اتفاق می‌افتد و به صورت (تعداد موفقیت‌ها در چارک/کل تعداد موفقیت‌ها) * ۱۰۰٪ تعریف می‌شود. مدل C&RT عملکرد بسیار خوبی را در بسیاری از چارک‌ها نشان می‌دهد. منحنی‌ها از نقطه ۰٪ شروع می‌شود و به به سمت ۱۰۰٪ از چپ به راست حرکت می‌کنند.

تحلیل حساسیت مدل‌های درخت تصمیم : تحلیل حساسیت تکنیکی مبتنی بر ترکیب اطلاعات است که با اجرای آن می‌توان نتایج چهار درخت تصمیم را ترکیب کرد تا مدل توانمند و صحیح‌تری ارائه کرد. همانطور که اشاره شد، درجدول ۹ میزان اهمیت هر یک از شاخص‌ها در مدل‌های درخت تصمیم بیان شده است.

همانطور که در جدول ۹ مشخص شده است سه متغیر استقلال هیئت مدیره، تقسیم سود و استقلال کمیته حسابرسی به ترتیب با میزان ۳۵/۶٪ ، ۲۰/۴ ٪ و۱۴/۲ ٪ درصد مهم‌ترین شاخص‌ها در معادله مدل تحلیل حساسیت هستند. با توجه به اینکه الگوریتم C&RT کاراترین درخت تصمیم شناخته شده مجدد جهت افزایش اعتبار نتیجه، این الگوریتم با نرم افزار رپید ماینر نیز مورد بررسی قرار گرفت که در ادامه نتایج آن آورده شده است. در این قسمت ابتدا داده‌های آموزشی در اختیار درخت تصمیم قرار داده شد تا دانش حاکم بر داده‌های مدل‌سازی کند و در ادامه برای پیش‌بینی کلاس از داده‌های آزمایشی استفاده شد. جدول ۱۰ نمایانگر متغیرهای مورد استفاده در پژوهش حاضر می‌باشد. شکل ۲ درخت تصمیم حاصل از داده‌های پژوهش را نشان می‌دهد.



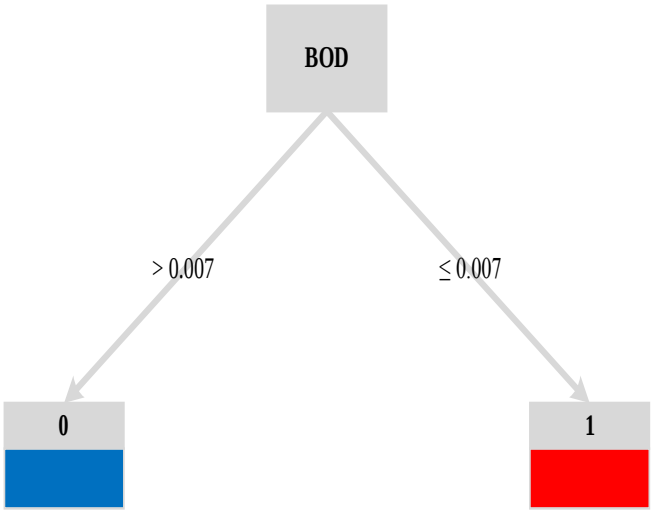
شکل (۱): نمودار سود مبتنی بر الگوریتم‌های درخت تصمیم

جدول (۹): میزان اهمیت هریک از متغیرهای مستقل در تحلیل حساسیت

نسبت های مالی	تحلیل حساسیت
استقلال هئیت مدیره	۰,۳۵۶
استقلال کمیته حسابرسی	۰,۱۴۲
مالکیت سازمانی	۰,۰۹۸
تقسیم سود	۰,۲۰۴
تامین مالی از طریق بدهی	۰,۱۰۲
محافظه کاری	۰,۰۹۸

جدول (۱۰): متغیرهای پژوهش در الگوریتم C&RT

علامت اختصاری	متغیر	وزن با همبستگی
BOD	استقلال هئیت مدیره	۰,۳۵۶
AC	استقلال کمیته حسابرسی	۰,۱۴۲
OOW	مالکیت سازمانی	۰,۰۹۸
PS	تقسیم سود	۰,۲۰۴
Dept	تامین مالی از طریق بدهی	۰,۱۰۲
Conser	محافظه کاری	۰,۰۹۸
EQ	کیفیت سود	



شکل (۲): درخت تصمیم با الگوریتم C&RT

با توجه به درخت ایجاد شده معیاراستقلال هیئت مدیره (BOD) به عنوان ریشه درخت انتخاب شد که نشان دهنده بالابودن اهمیت و تاثیرگذار بودن آن است. اگر مقدار ترکیب هیئت مدیره بالاتر از ۰،۰۰۷ باشد پیش‌بینی به درستی صورت نمی‌گیرد.

الف) ماتریس در هم ریختگی

ماتریس در هم ریختگی به ماتریسی گفته می‌شود که در آن عملکرد الگوریتم‌های مربوطه را نشان می‌دهد. این ماتریس دو قطر دارد؛ قطر اصلی نمایانگر داده‌هایی است که به درستی دسته‌بندی شده‌اند و قطر فرعی نمایانگر داده‌هایی است که به اشتباه دسته‌بندی شده‌اند. عناصر این ماتریس به شرح زیر است[۳]:

TP تعداد نمونه‌های مثبت که به درستی مثبت طبقه‌بندی شده‌اند.

FP تعداد نمونه‌های منفی که به اشتباه مثبت طبقه‌بندی شده‌اند.

TN تعداد نمونه‌های منفی که به درستی منفی طبقه‌بندی شده‌اند.

FN تعداد نمونه های مثبت که به اشتباه منفی طبقه‌بندی شده‌اند.

حال با استفاده از اپراتورهای این الگوریتم دقت مدل اندازه‌گیری شد (جدول ۱۱ و ۱۲):

جدول (۱۱): ماتریس درهم ریختگی مدل درخت تصمیم

	درست: ۰	درست: ۱	دقت
پیش بینی:۰	۲۳	۶۳۶	۹۷،۰۹٪
پیش بینی:۱	۴۶۰	۱۱	۱۰۰٪
حساسیت	۱۰۰٪	۲،۹۱٪	

جدول (۱۲): ارزیابی عملکرد مدل درخت تصمیم

معیار	مقدار(٪)	انحراف معیار(٪)
صحت	۷۹	۲،۸±
خطا	۲۱	± ۲،۸
AUC	۸۲،۹	± ۳،۲
دقت	۱۰۰	± ۰
حساسیت	۵۵،۷	± ۴
میانگین هارمونیک	۷۳،۴	± ۳،۲
وضوح	۱۰۰	± ۰

براساس معیارهای اندازه‌گیری، معیار دقت که تعداد پیشی‌بینی کیفیت سود را که به درستی دسته‌بندی شده‌اند را نشان می‌دهد ۹۷،۰۹ ٪ می‌باشد که به این معناست ۲،۹۱٪ از نمونه ها به اشتباه در کلاس دیگر طبقه‌بندی شده‌اند.

ب) ماتریس همبستگی

این ماتریس مشخص می‌کند در یک پژوهش دو ویژگی باهم ارتباط معنادار دارند یا خیر. میزان همبستگی عددی بین ۱- و ۱+ است. مقدار ۱+ نشان دهنده رابطه مثبت و کامل بین دو ویژگی است و معیار ۱- نشان دهنده ارتباط منفی و صفر نشان دهنده عدم ارتباط است[۲۴].

جدول (۱۳): ماتریس همبستگی مدل درخت تصمیم

attributes	BOD	AC	OOW	Coser	Dept	EQ=۱	PS
BOD	۱	۰،۰۳۶	۰،۰۲۸	-۰،۰۱۶	۰،۰۲۳	-۰،۰۲۶	-۰،۰۴۱
AC	۰،۰۳۶	۱	۰،۰۲۶	-۰،۰۰۳	-۰،۰۰۲۴	-۰،۰۲۶	۰،۰۳۶
OOW	۰،۰۲۸	۰،۰۲۶	۱	۰،۲۱۴	۰،۰۴۷	۰،۰۳۹	۰،۰۴۸
Conser	-۰،۰۱۶	-۰،۰۰۳	۰،۲۱۴	۱	-۰،۰۴۳	-۰،۱۱۶	-۰،۰۵۵
Dept	۰،۰۲۳	-۰،۰۰۲۴	۰،۰۴۷	-۰،۰۴۳	۱	۰،۰۶۶	۰،۰۳۷
EQ=۱	-۰،۰۲۶	-۰،۰۲۶	۰،۰۳۹	-۰،۱۱۶	۰،۴۶۶	۱	۰،۰۴۵
PS	-۰،۰۴۱	۰،۰۳۶	۰،۰۴۸	-۰،۰۵۵	۰،۰۳۷	۰،۰۴۵	۱

ج) نمودارهای ROC و AUC

ROC یک ابزار مدل‌سازی قوی برای تصمیم‌گیری‌ها در زمانی که ارزش‌های آستانه‌ای مدنظر است می‌باشد. این منحنی نمودار پراکندگی از حساسیت برای یک سیستم طبقه‌بندی کننده دوبه دویی است.

با استفاده از این نمودار می‌توان سه معیار اساسی در بررسی صحت یک مدل شامل دقت، ویژگی و حساسیت را استخراج کرد. AUC سطح زیر نمودار ROC است که یک معیار استاندارد از میزان تشخیص مربوط به روش شناسایی است. این مقدار نشان دهنده کارایی مدل مورد استفاده است و هرچه این مقدار بیشتر باشد کارایی آن مدل بیشتر است.

۵-۳- روش اجرایی ماشین بردار پشتیبان

بر اساس توضیحات گفته شده، در این قسمت نتایج روش ماشین بردار پشتیبان آورده شده است. جدول ۱۴ مدل Kernel تکنیک ماشین بردار پشتیبان را نشان می‌دهد.

در ادامه به بررسی معیارهای ارزیابی و اعتبارسنجی ماشین بردار پشتیبان طراحی شده پرداخته شد. نتیجه حاصل از اجرای عملگر performance classification در جدول ۱۵ و ۱۶ نمایش داده شده است. با توجه به مقدار صحت برای ماشین بردار پشتیبان ایجاد شده که برابر با ۹۸،۵ درصد می باشد عملکرد خوب مدل را نشان می‌دهد.

دقت در این مدل ۱۰۰٪ ، میزان حساسیت ۵۶،۶٪ است. در ادامه به مقایسه نتایج دو روش مورد استفاده در پیش بینی کیفیت سود پرداخته شد.

جدول (۱۴): نتایج تکنیک ماشین بردار پشتیبان

عنوان		مقدار
مجموع بردارهای پشتیبان		۱۰۹
Bias (Offset)		-۱۲،۸۵۴
درجه معیارها	استقلال هئیت مدیره	-۲۴۱۵۸۶،۳
	استقلال کمیته حسابرسی	-۳۲-۹۶۵۴
	مالکیت سازمانی	-۹۱۴۱،۱۴۱
	تقسیم سود	-۳۶۴۹۷،۷۸۵
	تامین مالی از طریق بدهی	-۲۲۵۷۶۰،۶۷۷
	محافظه کاری	-۳۰۷۹،۹۷۷
تعداد کلاس		۲
تعداد بردار پشتیبان برای کلاس صفر		۵۳
تعداد بردار پشتیبان برای کلاس یک		۵۶

با توجه به مقایسه نتایج عملکرد دو تکنیک (جدول ۱۷)، تکنیک ماشین بردار پشتیبان به جهت دارا بودن صحت و حساسیت بیشتر و خطای کمتر نسبت به تکنیک درخت تصمیم عملکرد بهتری دارد. همچنین در هر دو روش متغیر استقلال هیئت مدیره (BOD) بالاترین وزن و اهمیت را نسبت به سایر متغیر ها دارد و با توجه منحنی ROC، مدل ماشین بردار پشتیبان از کارایی بالاتری نسبت به مدل درخت تصمیم برخوردار است. با توجه به اینکه سطح زیر منحنی مدل ماشین بردار پشتیبان از مدل درخت تصمیم بیشتر است (شکل ۳) کارایی آن نیز بیشتر از مدل درخت تصمیم است.

جدول (۱۵): ماتریس درهم ریختگی مدل ماشین بردار پشتیبان

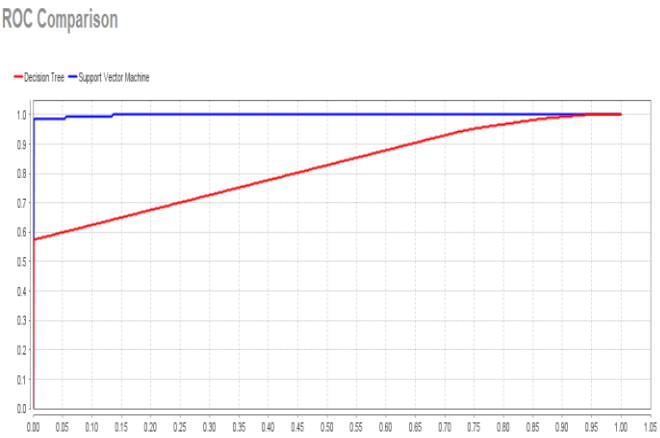
	درست: ۰	درست: ۱	دقت
پیش بینی:۰	۴۵۸	۳۷	۹۸،۵۱٪
پیش بینی:۱	۴۸۳	۶۴۷	۱۰۰٪
حساسیت	۱۰۰٪	۱،۴۹٪	

جدول (۱۶): ارزیابی عملکرد مدل ماشین بردار پشتیبان

معیار	مقدار(٪)	انحراف معیار(٪)
صحت	۷۸،۳	± ۳،۸
خطا	۲۱،۲	± ۳،۸
AUC	۹۹،۹	± ۰،۲
دقت	۱۰۰	± ۰
حساسیت	۵۶،۶	± ۴،۶
میانگین هارمونیک	۷۲،۲	± ۳،۸
وضوح	۱۰۰	± ۰

جدول (۱۷): مقایسه نتایج دو روش درخت تصمیم و ماشین بردار پشتیبان

عنوان	دقت(٪)	صحت(٪)	حساسیت(٪)	خطا(٪)	انحراف معیار(٪)	میانگین هارمونیک(٪)
درخت تصمیم	۱۰۰	۷۸	۵۵,۷	۲۲	± ۳,۲	۷۱,۲
ماشین بردار پشتیبان	۱۰۰	۷۸,۶	۵۶,۶	۲۱,۴	± ۳,۸	۷۲,۴



شکل (۳): منحنی ROC دو روش درخت تصمیم و ماشین بردار پشتیبان

۶- بحث

نتیجه تحقیق نشان می‌دهد که در هر دو روش الگوریتم درخت تصمیم و ماشین بردار پشتیبان استقلال هیت مدیره و سیاست تقسیم سود به ترتیب به ترتیب با ۳۵/۶٪و ۲۰/۰۴٪ بیشترین تاثیر را جهت پیش‌بینی کیفیت سود نسبت به سایر متغیرها داشته‌اند .و ماشین بردار پشتیبان با دقت ۹۸/۵٪ نسبت به روش درخت تصمیم عملکرد بهتری جهت پیش‌بینی کیفیت سود داشته و کارایی آن بالاتر است. نتیجه این تحقیق تایید کننده نتایج تحقیق حبیب‌زاده و همکاران در سال ۹۷ می‌باشد که تاثیر شبکه عصبی و مقایسه آن با درخت تصمیم و ماشین بردار پشتیبان، بر سودآوری شرکت‌ها را مورد مطالعه قرار دادند. نتایج تحقیق آنها نشان داد که درخت تصمیم با دقت ۹۳/۵درصد بیشترین در پیش بینی روند سودآوری شرکتهای پذیرفته شده در بورس اوراق بهادار تهران دارد.یافته‌های این پژوهش بیان می‌کند که ماشین بردار پشتیبان نسبت به درخت تصمیم دقت، صحت بیشترو خطای کمتری در پیش‌بینی کیفیت سود دارد با نتایج پژوهش گنگ و همکاران (۲۰۱۵) و کیم و همکاران (۲۰۱۳) تطابق دارد. با توجه به نتیجه این تحقیق استفاده از روش‌های داده کاوی جهت افزایش دقت پیش‌بینی‌های مبتنی بر داده‌های مالی برای افراد مشارکت کننده در بازار سرمایه سودمند است، تا ازتحمیل زیان‌های عمده مرتبط با خطای داده‌های مالی به سرمایه‌گذاران جلوگیری شود. باتوجه به یافته‌های این تحقیق به سرمایه‌گذاران بایستی هنگام تصمیم‌گیری درباره خرید یا فروش سهام خود به ویژگی‌های حاکمیتت شرکتی و مالکیت شرکت‌ها توجه کنند زیرا بر اساس نتایج این تحقیق ویژگی‌های حاکمیت شرکتی بر کیفیت سود آنها تاثیر دارد و توصیه می‌شود در شرکت‌هایی سرمایه‌گذاری کنند که دارای حاکمیت شرکتی قوی‌تر و ساختار مالکیت متمرکزتری هستند زیرا هر چه تمرکز مالکیت شرکت متمرکزتر باشد موجب بهبود عملکرد و سودآوری شرکت می‌شود.

از طرفی سهامداران بایستی از مزایای وجود مدیران غیرموظف و مستقل در هیئت مدیره و مالکیت سازمانی که از معیارهای حاکمیت شرکتی هستند، آگاه شوند. چون در صورتی که اکثریت اعضای هیت مدیره را مدیران غیرموظف تشکیل دهند هیت مدیره از کارایی بیشتری برخوردار خواهد بود.

۷- نتیجه‌گیری

هدف این پژوهش، پیش‌بینی کیفیت سود با شاخص‌های مرتبط با تامین مالی بدهی، استقلال هیئت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی، محافظه کاری و سیاست تقسیم سود می‌باشد. برای این منظور داده‌های لازم جهت محاسبه متغیرهای مستقل از صورت‌های مالی شرکتهای پذیرفته شده در بورس اوراق بهادار برای دوره زمانی ۱۳۹۰ تا ۱۴۰۰ به تعداد ۱۱۳ شرکت و ۱۱۳۰ مشاهده استخراج شد. بعد از شناسایی عوامل موثر، پیش‌بینی کیفیت سود با استفاده از الگوریتم‌های درخت تصمصم شامل شامل (CHAID, C۵,۰, C&RT, QUEST) و ماشین بردار پشتیبان انجام شد. نتایج الگوریتم درخت تصمیم و ماشین بردار پشتیبان نشان می‌دهد در هر دو مورد استقلال هیئت مدیره بیشترین تاثیر را درپیش بینی کیفیت سود دارد. مقایسه نتایج تکنیک درخت تصمیم و ماشین بردار نشان می‌دهد که ماشین بردار پشتیبان با دقت ۹۸/۵٪ ، حساسیت ۵۶/۶٪ و خطای ۲۱/۴٪ نسبت به روش درخت تصمیم عملکرد بهتری جهت پیش‌بینی کیفیت سود داشته و در هر دو روش استقلال هیئت مدیره بالاترین وزن و اهمیت نسبت به سایر متغیرهای مربوط به پیش‌بینی کیفیت سود دارد. باتوجه به اینکه پژوهش حاضر تأثیر تعدادی از متغیرها را برکیفیت سود بررسی کرده و به پیش بینی کیفیت سود به وسیله تکنیک درخت تصمیم و ماشین بردار پشتیبان پرداخته است، انجام پژوهش مشابه برای تک تک شرکت‌ها بااستفاده از گزارش‌های مالی میان دوره‌ای و یا استفاده از

مدل‌های دیگر و ترکیب الگوریتم‌های مختلف، انجام پژوهش درخصوص اثرات روانی پیش‌بینی و همچنین استفاده از سایر متغیرهای موثر بر کیفیت سود موضوعی است که می‌تواند در پژوهش‌های آتی مد نظر قرار گیرد.

منابع و مأخذ

[۱] پریجایی مقدم، احمد. موسوی، سجاد. (۲۰۱۲). **یادگیری درخت تصمیم بااستفاده ازشبکه عصبی برای افزایش پایداری وانعطاف‌پذیری**،مجله ایرانی انفورماتیک پزشکی، ش ۳.

[۲] Armstrong, C.S., Guay, W. R., Weber, J.P. (۲۰۱۰). **The Role of Information and Financial Reporting in Corporate Governance and Debt Contracting.** Journal of accounting and economics, ۵۰(۲-۳), ۱۷۹-۲۳٤.

[۳] Azizi, Z., Abdolvand, N., Ghalibaf Asl, H., Rajaee Harandi, S. (۲۰۲۱). **The Impact of Persian News on Stock Returns through Text Mining Techniques**. Iranian Journal of Management Studies, ۱٤(٤), ۷۹۹-۸۱٦.

[٤] Ball, R., Shivakumar, L. (۲۰۰۵). **Earnings Quality in UK Private Firms: Comparative Loss Recognition Timeliness**. Journal of Accounting and Economics, ۳۹(۱), ۸۳-۱۲۸.

[۵] Bhatia, S., Bajaj, R., Hazari, S. (۲۰۱٦). **Analysis of Credit Card Fraud Detection Techniques.** International Journal of Science and Research, ۵(۳), ۱۳۰۲-۱۳۰۷.

[٦] Bliss, J. (۱۹۲٤). **Management Through Accounts**. The Ronald Press Co., New York .

[۷] Boediono, G.S.B. (۲۰۰۵). **Kualitas Laba: Studi Pengaruh Mekanisme Corporate Governance dan Dampak Manajemen Laba dengan Menggunakan Analisis Jalur**. Simposium Nasional Akuntansi ۸. Solo.

[۸] Chen, S., Komal, B. (۲۰۱۸). **Audit Committee Financial Expertise and Earnings Quality: A Meta-Analysis**. Journal of Business Research, ۸٤, ۲۵۳-۲۷۰.

[۹] Clyde, P.S. (۲۰۰۷). **Financial Accounting an Introduction to Concepts, Methods, and Uses**. Canada: South Cangage Learning.

[۱۰] Dechow, P. M., Dichev, I. D. (۲۰۰۲). **The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors**. The accounting review, ۷۷(s-۱), ۳۵-۵۹.

[۱۱] Dechow, P., Ge, W., Schrand, C. (۲۰۱۰). **Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences.** Journal of accounting and economics, ۵۰(۲-۳), ۳٤٤-٤۰۱.

[۱۲] DeFond, M.L., Jiambalvo, J. (۱۹۹٤). **Debt Covenant Violation and Manipulation of Accruals.** Journal of accounting and economics, ۱۷(۱-۲), ۱٤۵-۱۷٦.

[۱۳] Domingues, A. (۲۰۱٦). **Idiosyncratic Volatility and Earnings Quality: Evidence from United Kingdom**. FEP Working Papers, ۵۷۹: ۳–٤۲.

[۱٤] Ewert, R., Wagenhofer, A. (۲۰۱۳). **Accounting Standards, Earnings Management, and Earnings Quality**. Earnings Management, and Earnings Quality (April ۳۰, ۲۰۱۳). papers.ssrn.com.

[۱۵] Fama, E.F., French, K.R. (۲۰۰۱). **Disappearing Dividends: Changing firm Characteristics or Lower Propensity to Pay?**. Journal of Financial economics, ٦۰(۱), ۳-٤۳.

[۱٦] Francis, J., LaFond, R., Olsson, P.M., Schipper, K. (۲۰۰٤). **Costs of Equity and Earnings Attributes**. The accounting review, ۷۹(٤), ۹٦۷-۱۰۱۰.

[۱۷] Ghosh, A., Moon, D. (۲۰۱۰). **Corporate Debt Financing and Earnings Quality**. Journal of Business Finance & Accounting, ۳۷(۵-٦), ۵۳۸-۵۵۹.

[۱۸] Givoly, D., Hayn, C. (۲۰۰۰). **The Changing Time-Series Properties of Earnings, Cash Flows and Accruals: Has Financial Reporting Become More Conservative?.** Journal of accounting and economics, ۲۹(۳), ۲۸۷-۳۲۰.

[۱۹] Gutierez A.L., Rodriguez, M.C. (۲۰۲۰). **Measuring Earnings Quality: a Proposal for a Theoretical Framework and an Empirical Method.** https://aeca.es method. https://aeca.es/ wp-content/uploads/۲۰۱٤/۰۵/۱۰a.pdf. Accessed ۵ June ۲۰۲۰.

[۲۰] Jones, J. J. (۱۹۹۱). **Earnings Management During Import Relief Investigations.** Journal of accounting research, ۲۹(۲), ۱۹۳-۲۲۸.

[۲۱] Khuong, N. V., Abdul Rahman, A. A., Thuan, P. Q., Liem, N. T., Anh, L. H. T., Thuy, C. T. M., & Ly, H. T. N. (۲۰۲۲). **Earnings Management, Board Composition and Earnings Persistence in Emerging Market.** Sustainability, ۱٤(۳), ۱-۱۵.

[۲۲] Kotsiantis, S.B. (۲۰۱۳). **Decision Trees: a Recent Overview**. Artificial Intelligence Review, ۳۹(٤), ۲٦۱-۲۸۳.

[۲۳] Lennox, C., Lisowsky, P., Pittman, J. (۲۰۱۳). **Tax Aggressiveness and Accounting Fraud.** Journal of accounting research, ۵۱(٤), ۷۳۹-۷۷۸.

[۲٤] Mathew, R.J., Largen, J., Claghorn, J.L. (۱۹۷۹). **Biological Symptoms of Depression.** Psychosomatic Medicine, ٤۱(٦), ٤۳۹-٤٤۳.

[۲۵] Muguto, L., Muzindutsi, P.F. (۲۰۲۲). **A Comparative Analysis of the Nature of Stock Return Volatility in BRICS and G۷ Markets**. Journal of Risk and Financial Management, ۱۵(۲), ۱-۲۷.

[۲٦] Parte-Esteban, L., García, C.F. (۲۰۱٤). **The Influence of Firm Characteristics on Earnings Quality.** International Journal of Hospitality Management, ٤۲, ۵۰-٦۰.

[۲۷] Perotti, P., Wagenhofer, A. (۲۰۱٤). **Earnings Quality Measures and Excess Returns.** Journal of business finance & accounting, ٤۱(۵-٦), ۵٤۵-۵۷۱.

[۲۸] Pratiwi, W., Pralita, T. (۲۰۲۱). **The Influence of Corporate Governance Mechanism, Accounting Conservatism, And Company Size on Earnings Quality.** Journal of Economics, Finance and Management Studies, ٤(۰۱), ۲٦-۳٤.

[۲۹] SeTin, S., Murwaningsari, E. (۲۰۱۸). **The Effect of Managerial Ability towards Earning Quality with Audit Committee as Moderating Variable.** Journal of Business and Retail Management Research, ۱۲(۳), ۱۷۸-۱۸۹.

[۳۰] Siueia, T.T., Wang, J. (۲۰۱۹). **The Association between Corporate Social Responsibility and Earnings Quality: Evidence from Extractive Industry: La Asociación Entre Las Actividades de Responsabilidad Social Corporativa y la Calidad de Los ingresos: Evidencia de la Industria Extractiva**. Revista de Contabilidad-Spanish Accounting Review, ۲۲(۱), ۱۱۲-۱۲۱.

[۳۱] Skinner, D. J., Soltes, E. (۲۰۱۱). **What Do Dividends Tell us about Earnings Quality**?. Review of Accounting Studies, ۱٦(۱), ۱-۲۸.