

مدل سازی پیش بینی کیفیت سود با استفاده از درخت تصمیم و ماشین بردار پشتیبان

لقمان حاتمی شیرکوهی^۱، صغری براری نوکاشتی^{۲*}، مریم اوشک سرایی^۳

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

^۲آستادیار، گروه حسابداری، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران

^۳آستادیار، گروه مدیریت صنعتی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران.

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

چکیده

سود و کیفیت آن از مهم ترین مولفه های تصمیم گیری برای استفاده کنندگان است. از اینرو پیش بینی کیفیت سود برای سرمایه گذاران و سایر گروه های ذینفع از اهمیت بسزایی برخوردار است، بدین منظور از روش درخت تصمیم و ماشین بردار پشتیبان برای پیش بینی کیفیت سود که هدف اصلی تحقیق می باشد استفاده شده است. جامعه آماری پژوهش شامل شرکتهای پذیرفته شده در بورس اوراق بهادار تهران از سال ۱۳۹۰ تا سال ۱۴۰۰ به مدت ۱۰ سال می باشد، بعد از غربالگری تعداد ۱۱۳ شرکت و ۱۱۳۰ مشاهده بعنوان نمونه آماری انتخاب شدند. در این مطالعه به منظور شناسایی و پیش بینی کیفیت سود شاخصهای مرتبط با حاکمیت شرکتی (استقلال هیئت مدیره، استقلال کمیته حسابرسی، مالکیت سازمانی)، سیاست تقسیم سود، تامین مالی بدهی و محافظه کاری بعنوان متغیر مستقل و کیفیت اقلام تعهدی اختیاری نماینده شاخص کیفیت سود بعنوان متغیر وابسته مورد بررسی قرار گرفته است. این پژوهش از نظر هدف کاربردی و به لحاظ روش پژوهشی، توصیفی از نوع همبستگی. تجزیه تحلیل داده ها مطابق با استاندارد داده کاوی CRISP-DM و اجرای چهار الگوریتم درخت تصمیم شامل (CHAID, C5.0, C&R, QUEST) و ماشین بردار پشتیبان، استقلال هیئت مدیره بیشترین تاثیر را در پیش بینی کیفیت سود دارد. مقایسه نتایج عملکرد دو روش نشان می دهد که، تکنیک ماشین بردار پشتیبان به جهت دارا بودن صحت و حساسیت بیشتر و خطای کمتر نسبت به تکنیک درخت تصمیم عملکرد بهتری دارد.

واژه های اصلی: کیفیت سود، درخت تصمیم، ماشین بردار پشتیبان، اقلام تعهدی، محافظه کاری.

۱- مقدمه

سود جنبه بسیار مهمی از سود حسابداری را نشان می دهد، زیرا آگاهی سهامداران از کیفیت سود می تواند بر قیمت سهام، میزان سود تقسیمی و در نتیجه بازده سهام، تأثیرگذار باشد. بعنوان 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 هر یک از معیارهای کیفیت سود منحصر بفرد است و جنبه خاصی از سود گزارش شده یک شرکت را نشان می دهد؛ بنابراین هر یک از معیارهای کیفیت سود می تواند به طور متفاوتی بر نوسانات بازده سهام تاثیر بگذارد و مطابق با این نگرش می توان گفت که کیفیت سود یک مفهوم چند بعدی است. معیارهای متعددی از جمله کیفیت اقلام تعهدی، محافظه کاری، به موقع بودن، ارتباط ارزشی، تداوم سود، قابلیت پیش بینی و یکنواختی سود برای اندازه گیری آن استفاده شده است [۱۶]. اگرچه در مجموع هیچ

گزارشگری مالی اطلاعات حسابداری مفیدی را در اختیار بازار سرمایه برای تصمیم گیری تجاری قرار می دهد [۲۷]، گزارش های مالی بطور گسترده توسط مدیران، سرمایه گذاران، تحلیلگران، تنظیم کننده ها و استاندارد گذاران مالی بطور گسترده مورد استفاده قرار می گیرد. شکل سود در صورت مالی اطلاعات حسابداری را خلاصه می کند، در حالیکه کیفیت سود برای تصمیم گیری اقتصادی استفاده می شود [۲۱]. سود با کیفیت بالا با ارایه یک معیار دقیق از عملیات یک شرکت مشخص می شود. در واقع چنین سودی (سود با کیفیت) پایدار است؛ درآمدهای آتی را بطور قابل اعتماد پیش بینی می کند، مدیریت نشده است و اطلاعات ارزشمندی را در اختیار کاربران قرار می دهد [۲۵].

کیفیت سود معیاری مهم برای سلامت مالی واحد تجاری به شمار می رود 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 آینده شرکت و توان سود آوری آتی بستگی دارد. کیفیت

¹ Muguto & Muzindutsi
*Brari@iaurasht.ac.ir

دستور العمل 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 [۱۹،۱۱].

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 می‌رسد. 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 قراردادهای می‌گردد. کیفیت بالای صورت‌های مالی بوسیله 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 بجای 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 تأثیر 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 ماشین بردار پشتیبان و الگوریتم‌های درخت تصمیم به شناسایی و پیش‌بینی کیفیت سود اقدام کند و با توجه به کاستی‌های الگوهای طبق‌بندی 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.

۱-۱- 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. ضمناً در تئوری حسابداری مفاهیم ساختاری، تفسیری و رفتاری (عملگرایی) 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 جریان نقدی عملیاتی دوره گذشته، حال و آینده تعریف می‌کنند. در این صورت، کیفیت سود ناشی از کیفیت 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 تأثیر دارند در این پژوهش ویژگیهای حاکمیت شرکتی (استقلال هیئت 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 منجر به بروز یک درگیری بالقوه شود که بر کیفیت سود گزارش شده تأثیر بگذارد. مدیریت با سود بهره‌خاص تمایل به تهیه گزارش متناسب با هدف و نه در جهت منافع اصلی دارد. در چنین شرایطی، 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 می‌تواند 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 نموده‌اند" [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] محافظه‌کاری از طریق اقلام تعهدی اختیاری (غیر عملیاتی) اندازه‌گیری می‌کنند، زیرا به نظر آنها حسابداری تعهدی محای برای اجرای محافظه‌کاری است. براساس مفهوم محافظه‌کاری واحدهای تجاری در شرایط مختلف روشهای را انتخاب می‌کنند که حداکثر اثر فزاینده بر سود و جمع داراییها داشته باشد و این نحوه عمل می‌تواند تأثیر 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 بعنوان شاخصی از کیفیت

² Dechow & Dichev

³ Ball & Shivakumar

⁴ Givoly & Hayn

مطالعه صورت گرفته در سال های ۲۰۱۶-۲۰۰۳ مورد بررسی قرار داده اند. نتایج این پژوهش نشان داد که تخصص مالی کمیته حسابرسی رابطه مثبت و معناداری با کیفیت سود دارد و متخصصان مالی حسابداری رابطه قویتری با کیفیت سود نسبت به متخصصان مالی غیر حسابداری دارند. از طرفی سیستم های حاکمیت شرکتی و استانداردهای گزارشگری مالی بین المللی رابطه بین تخصص مالی کمیته حسابرسی و کیفیت سود را تعدیل می کنند.

ستین و مروانیگساری^۸ [۲۹] تأثیر توانایی مدیریت در افزایش کیفیت سود با در نظر گرفتن اثر تعدیلی کمیته حسابرسی در ۱۲ شرکت تولیدی دولتی اندونزی در بین سال های ۲۰۱۸-۲۰۱۶ را مورد بررسی قرار داده اند. نتایج حاصل از برآورد مدل رگرسیون چند متغیره پژوهش نشان داد که بین توانایی های مدیریت و کیفیت سود رابطه مثبت و معناداری وجود دارد. همچنین کمیته حسابرسی تأثیر توانایی مدیریتی را در قبال کیفیت سود تقویت می کند.

استبان و گارسیا^۹ [۲۶] در پژوهش خود به بررسی تأثیر ویژگی های شرکت بر کیفیت سود پرداختند. جامعه آماری این پژوهش هتل های کشور اسپانیا بین سال های ۲۰۰۰ تا ۲۰۱۱ می باشد. ایشان در پژوهش خود برای اندازه گیری کیفیت سود از یک اندازه گیری چند بعدی که شامل پایداری سود و قابلیت پیش بینی و هموارسازی سود می باشد استفاده کرده و متغیرهایی مانند متغیرهای بازار و استراتژی های بازار و ساختار مالکیت و شاخص های حسابرسی به عنوان ویژگی های شرکت مورد استفاده قرار گرفت. نتایج تحقیق نشان داد که متغیرهای بازار و ساختار مالکیت و شاخص های حسابرسی بر کیفیت سود تأثیر گذار 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 می دهد که کیفیت سود با افزایش دقت حسابداری افزایش می یابد. با ملاحظه پژوهشهای بالا می توان گفت که اکثر آنها اثر یک متغیر مستقل وابسته بر یک متغیر وابسته را مورد بررسی قرار داده اند و از روش های ساده رگرسیون و غیر هوشمند برای تجزیه تحلیل داده ها و آزمون فرضیه ها استفاده نموده اند. از آنجایی که همیشه بین متغیر مستقل و متغیر وابسته رابطه خطی وجود ندارد، از این رو در چنین مواقعی باید از روش های دیگری مانند درخت تصمیم و ماشین بردار پشتیبان که جز روشهای داده کاوی هستند و از قدرت پیش بینی کنندگی بالایی برخوردار هستند و میزان صحت و قدرت دقت آن بیشتر از روش های خطی هست استفاده نمود. علاوه بر استفاده از روشهای الگوریتم درخت تصمیم و ماشین بردار پشتیبان استفاده شده در این تحقیق برای پیش بینی کیفیت سود بطور همزمان اثر چندین متغیر بر کیفیت سود مورد بررسی قرار گرفته است تا در یک نگاه همه جانبه اطلاعات جامعی را در اختیار سرمایه گذاران و سایر گروه های ذینفع قرار دهد.

گزارشهای مالی رابطه مستقیمی با سیاست تقسیم سود شرکت دارد. شرکتهایی که میزان سود خود را در گزارشهای مالی اظهار می کنند، سیاست هایی متناظر با سود اعلام شده، در تقسیم سود شرکت لحاظ می کنند [۲۳]. فاما و فرنچ^۵ [۱۵] نشان دادند نسبت شرکتهای آمریکایی که بطور مستمر سود تقسیمی پرداخت می کنند 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 تأثیر دارد و در این مورد سه دیدگاه مطرح است. دیدگاه اول تامین مالی از طریق بدهی تأثیر مثبت بر کیفیت سود دارد. در این دیدگاه اعتبار دهندگان نسبت به کیفیت اطلاعات حسابداری وام گیرندگان واکنش نشان می دهند. بدین معنی هرچه کیفیت اطلاعات حسابداری وام گیرندگان بیشتر باشد شرایط محدود کننده کمتر و نرخ بهره پایین تری را در قراردادهای وام آنها در نظر می گیرند و بالعکس. در نتیجه اعتبارگیرندگان سعی می کنند 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 شرکتهای ممکن است برای جلب نظر مثبت اعتبار دهندگان، 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 می توان گفت که بدهی اثر منفی بر کیفیت سود دارد. دیدگاه سوم رابطه بین تامین 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 بر کیفیت سود تعیین می شود [۱۷].

۲-۲- پیشینه تجربی پژوهش

موگوتو و موزینداتسی [۲۵] در پژوهشی با عنوان معیار های 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 تأثیر می گذارد 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 سال های ۲۰۱۷-۲۰۱۰ پرداخته اند. نتایج حاصل از آزمون فرضیه های پژوهش با استفاده از مدل رگرسیون داده های تابلویی نشان داد که بین مسئولیت اجتماعی شرکت و کیفیت سود رابطه منفی و معناداری وجود دارد. این یافته ها 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 رابطه بین تخصص مالی کمیته حسابرسی و کیفیت سود شرکت ها را با استفاده از تکنیک فراتحلیل نتایج ۹۰

⁸ SeTin & Murwaningsari

⁹ Parte-Esteban & García

¹⁰ Ewert & Wagenhofer

⁵ Fama & French

⁶ Siueia & Wang

⁷ Chen & Komal

۲- روش شناسی پژوهش

این پژوهش یک پژوهش توصیفی، و در حوزه پژوهش های همبستگی است که مبتنی بر اطلاعات واقعی شرکت های پذیرفته شده در بورس اوراق بهادار تهران است. داده های لازم برای اجرای پژوهش از صورت های مالی و گزارشات هیأت مدیره شرکت های پذیرفته شده در بورس اوراق بهادار تهران در شرکت های صنایع تولیدی از سال ۱۳۹۰ تا سال ۱۴۰۰ و همچنین از سایت کدال جمع آوری شده است. تجزیه و تحلیل داده ها با استفاده از نرم افزار اکسل و ایوبوز انجام شده برای استخراج مدل از روش الگوریتم درخت تصمیم و ماشین بردار پشتیبان استفاده شده است.

۱-۲- فرضیه پژوهش

فرضیه اصلی پژوهش در این تحقیق در قالب سوال مطرح می شود به دلیل اینکه روش های روش های داده کاوی مانند روش های آماری از آمار استنباطی برای تحلیل فرضیه ها استفاده نمی کنند و صرفاً روشی هوشمند برای پیش بینی متغیر وابسته هستند، بنابراین فرضیه پژوهش را می توان به صورت زیر بیان کرد. "آیا استفاده از الگوریتم های درخت تصمیم و ماشین بردار پشتیبان بر پیش بینی دقت کیفیت سود موثر

است یا خیر"

۲-۲- جامعه و نمونه آماری پژوهش

جامعه آماری مورد بررسی در این پژوهش شامل شرکت های پذیرفته شده در بورس اوراق بهادار تهران از سالهای ۱۳۹۰ تا سال ۱۴۰۰ به مدت ۱۰ سال می باشد. علت گستردگی حجم جامعه آماری و وجود برخی ناهماهنگی ها میان اعضای جامعه، شرایط زیر برای انتخاب نمونه آماری قرار داده شده و نمونه آماری به روش غربالگری (حذف سیستماتیک) انتخاب شده است. شرکت هایی بعنوان نمونه انتخاب می شوند که دارای شرایط زیر باشند:

۱- سال مالی آنها منتهی به پایان اسفند باشد.

۲- جز شرکت های سرمایه گذاری و واسطه گری مالی نباشد.

۳- طی دوره مورد پژوهش سال مالی خود را تغییر نداده باشد.

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

جدول (۱): صنایع و شرکت های 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 شیمیایی	۸	لاستیک پلاستیک	۳
صنایع غذایی	۷	ماشین الات و تجهیزات	۷
صنعت کاشی و سرامیک	۳	محصولات فلزی	۱
صنعت سیمان	۱۲	دستگاه های برقی	۵
فلزات اساسی	۹	قند و شکر	۲
کانه غیر فلزی	۵	حمل نقل	۱
جمع کل		۱۱۳	

دارد. در این پژوهش برای اندازه گیری آن از مدل تعدیل شده جونز^{۱۱} [۲۰] استفاده شده است. در این پژوهش متغیر وابسته، کیفیت سود می باشد که از طریق کیفیت ارقام تعهدی اختیاری^{۱۲} (DA) اندازه گیری شد. کیفیت ارقام تعهدی نیز از طریق رابطه زیر محاسبه گردید.

۲-۳- نحوه محاسبه متغیرها

این پژوهش یک پژوهش توصیفی و در حوزه پژوهش های همبستگی است که مبتنی بر اطلاعات واقعی شرکت های 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، محافظه کاری.

متغیر وابسته: جهت اندازه گیری کیفیت سود معیارهای مختلفی وجود

¹¹ Jones

¹² Discretionary Accruals (DA)

جدول (۳): نحوه محاسبه متغیرهای مستقل پژوهش

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

۲-۴- روش درخت تصمیم

الگوریتم‌های درخت تصمیم یکی از رهیافت‌های ممکن برای تصمیم‌گیری چندمرحله‌ای می‌باشند. ایده اصلی در تصمیم‌گیری چند مرحله‌ای این است که تصمیمات 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 سلسله‌مراتبی می‌باشد (پریجایی مقدم و موسوی، ۲۰۱۲). الگوریتم‌های زیادی در طول سال‌های گذشته برای ساخت درخت تصمیم ایجاد شده‌اند، همانند: C4.5 (1993)، C&RT (1984)، SPRINT (1996)، SLIQ (1996). 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. 4.5 یک ترکیب بسیار خوب از خطا و سرعت می‌باشد. در C4.5 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 دسته بندی نمونه‌ها می‌باشند..

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

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

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

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

TACC_t ارقام تعهدی کل

NDA_t ارقام تعهدی غیراختیاری

DA_t ارقام تعهدی اختیاری

ΔREV_t تغییرات درآمد سال جاری به سال قبل

ΔREC_t تغییرات حسابها و اسناد دریافتنی سال جاری به سال قبل

TA_{t-1} جمع کل دارایی‌های ابتدای دوره

ε_t پسماند مدل رگرسیونی

PPE دارایی‌های ثابت (اموال، ماشین آلات و تجهیزات)

ΔCL تغییرات بدهی‌های جاری

ΔCA تغییرات دارایی‌های جاری

به طور کلی جهت پیش بینی کیفیت سود، میانه توزیع این متغیر در میان سال‌های 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 ارقام 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

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

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

۲-۵- ماشین بردار پشتیبان

۳- نتایج

۳-۱- آمواد سازی و پیش پردازش داده ها

این مرحله یکی از مهم ترین و اثرگذارترین مراحل داده کاوی است و ۸۰ درصد زمان فرایند داده کاوی را به خود اختصاص می دهد. در این پژوهش از متدولوژی CRISP DM به عنوان فرایند استانداردسازی داده کاوی استفاده شد. این روش اولین بار در سال ۱۹۹۶ ارائه شد. این متدولوژی یک چرخه شش مرحله ای دارد و از گام های شناخت کسب و کار، شناخت داده ها، آماده سازی داده ها، مدل سازی، ارزیابی مدل و استقرار تشکیل شده است. داده های مورد استفاده در این پژوهش، داده های شرکت های پذیرفته شده در بورس اوراق بهادار تهران است.

پس از نرمال سازی داده ها در مرحله پیش پردازش داده ها، کل داده ها به دو بخش ۷۰ درصد داده های آموزشی برای ساخت مدل و ۳۰ درصد داده های آزمایشی برای اعتبارسنجی مدل تقسیم گردید. کلیه مراحل در نرم افزار Rapid miner انجام شده است. جدول ۵ اطلاعات آماری پژوهش حاضر را نشان می دهد.

جدول (۵): اطلاعات آماری پژوهش

نام	مقدار
حداقل	-۱
حداکثر	+۱
انحراف معیار	۰.۱۲۵
میانگین	۰.۰۵۱

در مرحله بعدی، با کمک درخت های (CHAID, C5.0, C&RT, QUEST)، قدرت برازندگی شاخص های مورد بررسی بر کیفیت سود تعیین شده و با استفاده از ماتریس انطباق، عملکرد هر یک از درخت های تصمیم مورد بررسی قرار گرفته است.

۳-۲- مدل های درخت تصمیم با کمک تحلیل حساسیت

در نهایت نتایج چهار درخت تصمیم با استفاده از تحلیل حساسیت ترکیب می شود و مدل نهایی شکل می گیرد تا اعتباردهی هر یک از نسبت ها با دقت بیشتری انجام شود. تحلیل حساسیت تکنیکی مبتنی بر ترکیب اطلاعات است که با اجرای آن می توان نتایج چهار درخت تصمیم را ترکیب کرد تا مدل توانمند و صحیح تری ارائه کرد. شش متغیر به عنوان متغیرهای ورودی الگوریتم تعریف شده است. برای تعریف متغیر خروجی از کیفیت سود استفاده شده است. برای پیش بینی کیفیت سود، چهار نوع درخت تصمیم به کار برده شد: CHAID, C&RT, C5.0, QUEST. ابتدا ۷۰٪ داده ها 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 C5.0 و C&RT و QUEST بالاترین اهمیت را به خود اختصاص داده است. با توجه به

ماشین بردار پشتیبان تکنیک های یادگیری آماری است که ابتدا توسط کورتز و واپنیک در سال ۱۹۹۵ معرفی شدند و در بسیاری از مسائل طبقه بندی بسیار موفق بوده اند. ماشین بردار پشتیبان بر اساس مفهوم تصمیم گیری است و مرزهای تصمیم گیری را تعریف می کند. یک سطح تصمیم گیری آن است که بین مجموعه ای از کلاس های مختلف جدا شده است. این الگوریتم ها یک سطح بالاتر را به عنوان یک سطح تصمیم گیری که نمونه ها را به دو کلاس مثبت و منفی طبقه بندی می کند، در نظر می گیرند. این مدل در مقایسه با مدل های دیگر دقت و کارایی بالاتری دارد. به منظور دستیابی به حداکثر دقت پیش بینی، این الگوریتم نیازمند داده های با اندازه بزرگ برای آموزش است [۵].

۲-۶- روش اجرایی درخت تصمیم

در این پژوهش از روش درخت تصمیم^{۱۳} بر اساس قوانین "اگر-آنگاه" با استفاده از مخزن داده های تمام شرکت های 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 پیش بینی 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 برجسته داده ها تعیین شد و مساله از نوع مسائل با ناظر می باشد که برای حل آن باید از روش های دسته بندی^{۱۴} استفاده نمود. برای انجام دسته بندی از الگوریتم درخت تصمیم با استفاده از نرم افزار Rapid miner استفاده شده است.

جدول (۶): اطلاعات مختصر پایگاه داده پژوهش

ویژگی پایگاه داده	تعداد ویژگی ها	تعداد رکوردها	روش داده کاوی	داده های گم شده
چندمتغیره	۷	۱۱۳۰	دسته بندی	ندارد

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

- تهیه اطلاعات در سطح کل بورس
- انجام محاسبات نسبت های مالی و شاخص های پیش بینی کننده کیفیت سود
- اجرای مدل های درخت تصمیم
- ترکیب نتایج مدل های الگوریتم درخت تصمیم با کمک تحلیل حساسیت

¹³ Decision tree
¹⁴ Classification

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

جدول (۶): میزان اهمیت هریک از متغیرهای مستقل

مدل	CHAID	C5.0	C&RT	QUEST
<i>BOD</i>	۰.۶۳۶	۰.۴۸۷	۰.۵۳۹	۰.۴۷۴
<i>AC</i>	۰.۲۱۵	۰.۱۴۲	۰.۰۸۷	۰.۰۹۶
<i>OOW</i>	۰.۱۷۸	۰.۱۸۹	۰.۲۱۴	۰.۱۸۸
<i>PS</i>	۰.۱۸۷	۰.۲۶۳	۰.۲۱۸	۰.۲۶۹
<i>Dept</i>	۰.۱۶۶	۰.۲۲۹	۰.۰۱۹	۰.۲۴۸
<i>Conser</i>	۰.۰۴۱	۰.۰۴۹	۰.۱۳۲	۰.۰۴۹

جدول (۷): ماتریس اختلال (انطباق) مدل های 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		
C5.0	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 مدل های 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
C5.0	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

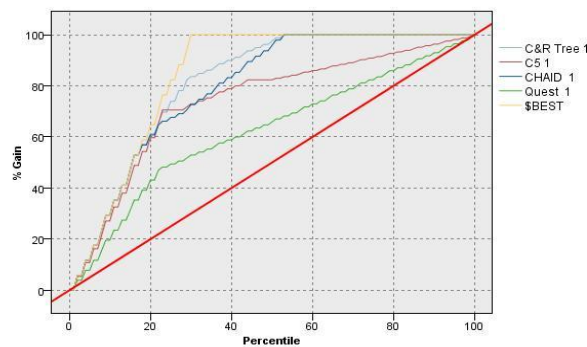
درخت CHAID با صحت کلی ۹۳/۷۵ درصد و درخت QUEST با صحت کلی ۸۶/۶۴ درصد عملکرد خوبی را داشته اند. با توجه به جدول ۵ می توان دید که صحت گروه ناموفق در سه درخت CHAID, C&RT و QUEST بالاتر از صحت گروه موفق بوده است ولی در درخت C5.0 با تفاوت کمی، صحت گروه موفق بالاتر از صحت گروه ناموفق بوده است.

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 نسبتاً بهتری داشته است. درخت C5.0 دومین سطح عملکرد را بر اساس تمام معیارها به خود اختصاص داده است. با این حال

همانطور که در جدول ۹ مشخص شده است سه متغیر استقلال هیئت مدیره، تقسیم سود و استقلال کمیته حسابرسی به ترتیب با میزان $0.356/6\%$ ، $0.204/4\%$ و $0.142/2\%$ درصد مهم ترین شاخص ها 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 ریپید مایندر نیز مورد بررسی قرار گرفت که در ادامه نتایج آن آورده شده است. در این قسمت ابتدا داده های آموزشی در اختیار درخت تصمیم قرار داده شد تا دانش حاکم بر داده های مدل را سازی کند و در ادامه برای پیش بینی کلاس از داده های آزمایشی استفاده شد. جدول ۱۰ نمایانگر متغیرهای مورد استفاده در پژوهش حاضر می باشد. شکل ۲ درخت تصمیم حاصل از داده های پژوهش را نشان می دهد.

نمایش نموداری معیارهای عملکرد با استفاده از نمودار کیفیت در شکل ۱ به تصویر کشیده شده است. کیفیت به صورت نسبی از کل موفقیت ها است که در هر چارک اتفاق می افتد و به صورت (تعداد موفقیت ها در چارک/کل تعداد موفقیت ها) * ۱۰۰٪ تعریف می شود. مدل C&RT عملکرد بسیار خوبی را در بسیاری از چارک ها نشان می دهد. منحنی ها از نقطه ۰٪ شروع می شود و به سمت ۱۰۰٪ از چپ به راست حرکت می کنند.

تحلیل حساسیت مدل های درخت تصمیم: تحلیل حساسیت تکنیکی مبتنی بر ترکیب اطلاعات است که با اجرای آن می توان نتایج چهار درخت تصمیم را ترکیب کرد تا مدل توانمند و صحیح تری ارائه کرد. همانطور که اشاره شد، در جدول ۹ میزان اهمیت هر یک از شاخص ها در مدل های درخت تصمیم بیان شده است.



شکل (۱): نمودار سود مبتنی بر الگوریتم های درخت تصمیم

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

نسبت های مالی	تحلیل حساسیت
استقلال هیئت 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

وزن با همبستگی	متغیر	علامت اختصاری
۰.۳۵۶	استقلال هیئت مدیره	BOD
۰.۱۴۲	استقلال کمیته حسابرسی	AC
۰.۰۹۸	مالکیت سازمانی	OOW
۰.۲۰۴	تقسیم سود	PS
۰.۱۰۲	تامین مالی از طریق بدهی	Dept
۰.۰۹۸	محافظه کاری	Conser
	کیفیت سود	EQ

(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

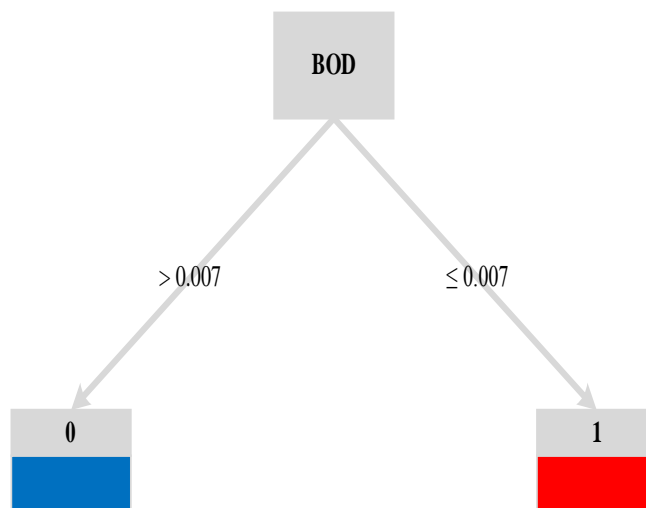
براساس معیارهای اندازه گیری، معیار دقت که تعداد پیش‌بینی کیفیت سود را که به درستی دسته بندی شده اند را نشان می دهد ۹۷.۰۹٪ می باشد که به این معناست ۲.۹۱٪ از نمونه ها به اشتباه در کلاس دیگر طبقه بندی شده اند.

ب) ماتریس همبستگی

این ماتریس مشخص می کند 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	<i>BOD</i>	<i>AC</i>	<i>OOW</i>	Coser	Dept	EQ=1	PS
<i>BOD</i>	۱	۰.۰۳۶	۰.۰۲۸	-۰.۰۱۶	۰.۰۲۳	-۰.۰۲۶	-۰.۰۴۱
<i>AC</i>	۰.۰۳۶	۱	۰.۰۲۶	-۰.۰۰۳	-۰.۰۰۲۴	-۰.۰۲۶	۰.۰۳۶
<i>OOW</i>	۰.۰۲۸	۰.۰۲۶	۱	۰.۲۱۴	۰.۰۴۷	۰.۰۳۹	۰.۰۴۸
Coser	-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=1	-۰.۰۲۶	-۰.۰۲۶	۰.۰۳۹	-۰.۱۱۶	۰.۰۶۶	۱	۰.۰۴۵
PS	-۰.۰۴۱	۰.۰۳۶	۰.۰۴۸	-۰.۰۵۵	۰.۰۳۷	۰.۰۴۵	۱



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 پیش بینی به درستی صورت نمی گیرد.

الف) ماتریس درهم ریختگی

ماتریس درهم ریختگی به ماتریسی گفته می شود که در آن عملکرد الگوریتم های مربوطه را نشان می دهد. این ماتریس دو قطر دارد؛ قطر اصلی نمایانگر داده هایی است که به درستی دسته بندی شده اند و قطر فرعی نمایانگر داده هایی است که به اشتباه دسته بندی شده اند. این ماتریس به شرح زیر است [۳]:

TP تعداد نمونه های مثبت که به درستی مثبت طبقه بندی شده اند.

FP تعداد نمونه های منفی که به اشتباه مثبت طبقه بندی شده اند.

TN تعداد نمونه های منفی که به درستی منفی طبقه بندی شده اند.

FN تعداد نمونه های مثبت که به اشتباه منفی طبقه بندی شده اند.

حال با استفاده از اپراتورهای این الگوریتم دقت مدل اندازه گیری شد

	بدهی
-۳۰۷۹.۹۷۷	محافظه کاری
۲	تعداد کلاس
۵۳	تعداد بردار پشتیبان برای کلاس صفر
۵۶	تعداد بردار پشتیبان برای کلاس یک

ج) نمودارهای ROC و AUC

ROC یک ابزار مدل سازی قوی برای تصمیم گیری ها در زمانی که ارزش های آستانه ای مد نظر است می باشد. این منحنی نمودار پراکندگی از حساسیت برای یک سیستم طبقه بندی کننده دوجه دویی

است. با استفاده از این نمودار می توان 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.

۳-۳- 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 می دهد.

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 می دهد.

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 (۱۵): ماتریس درهم ریختگی مدل ماشین بردار پشتیبان

	دقت	درست: ۱	درست: ۰
پیش بینی: ۰	۹۸.۵۱٪	۳۷	۴۵۸
پیش بینی: ۱	۱۰۰٪	۶۴۷	۴۸۳
حساسیت		۱.۴۹٪	۱۰۰٪

جدول (۱۴): نتایج تکنیک ماشین بردار پشتیبان

عنوان	مقدار
مجموع بردارهای پشتیبان	۱۰۹
Bias (Offset)	-۱۲۸۵۴
استقلال هیئت مدیره	-۲۴۱۵۸۶.۳
استقلال کمیته حسابرسی	-۳۲-۹۶۵۴
مالکیت سازمانی	-۹۱۴۱.۱۴۱
تقسیم سود	-۳۶۴۹۷.۷۸۵
تامین مالی از طریق	-۲۲۵۷۶۰.۶۷۷

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

در هیئت مدیره و مالکیت سازمانی که از معیارهای حاکمیت شرکتی هستند، آگاه شوند. چون در صورتی که اکثریت اعضای هیئت مدیره را مدیران غیر موظف تشکیل دهند هیئت مدیره از کارایی بیشتری برخوردار خواهد بود.

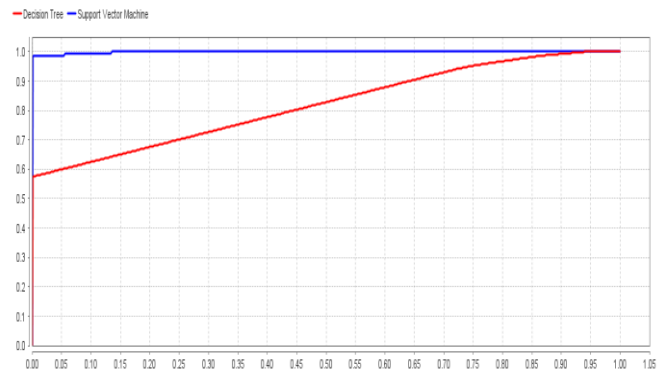
۵- نتیجه گیری

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 الگوریتم های درخت تصمیم شامل شامل (CHAID, C5.0, C&RT,) و ماشین بردار پشتیبان (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 در پیش‌بینی 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 بر کیفیت سود بررسی کرده و به پیش‌بینی کیفیت سود به وسیله تکنیک درخت تصمیم و ماشین بردار پشتیبان پرداخته است، انجام پژوهش مشابه برای تک شرکت‌ها با استفاده از گزارش‌های مالی میان دوره ای و یا استفاده از مدل‌های دیگر و ترکیب الگوریتم‌های مختلف، انجام پژوهش در خصوص 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 و سجاد موسوی. (۲۰۱۲). یادگیری درخت تصمیم با استفاده از شبکه عصبی برای افزایش پایداری وانعطاف پذیری، مجله ایرانی انفورماتیک پزشکی، ش ۳.
- [2] Armstrong, C. S., Guay, W. R., & Weber, J. P. (2010). **The role of information and financial reporting in corporate governance and debt contracting.** *Journal of accounting and economics*, 50(2-3), 179-234.
- [3] Azizi, Z., Abdolvand, N., Ghalibaf Asl, H., & Rajaei Harandi, S. (2021). **The Impact of Persian News on Stock Returns Through Text Mining Techniques.** *Iranian Journal of Management Studies*, 14(4), 799-816.
- [4] Ball, R., & Shivakumar, L. (2005). **Earnings quality in UK private firms: comparative loss recognition timeliness.** *Journal of Accounting And Economics*, 39(1), 83-128.

ROC Comparison



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

۴- بحث

نتیجه تحقیق نشان می‌دهد که در هر دو روش الگوریتم درخت تصمیم و ماشین بردار پشتیبان استقلال هیئت 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 داشته اند. 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 حبیب زاده و همکاران در سال ۹۷ می‌باشد که تاثیر شبکه عصبی و مقایسه آن با درخت تصمیم و ماشین بردار پشتیبان، بر سود آوری شرکتها را مورد مطالعه قرار دادند. نتایج تحقیق آنها نشان داد که درخت تصمیم با دقت ۹۳/۵ درصد بیشترین تاثیر را در پیش‌بینی روند سود آوری شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران دارد. یافته‌های این پژوهش بیان می‌کند که ماشین بردار پشتیبان نسبت به درخت تصمیم دقت، صحت بیشتر و 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 پیش‌بینی های مبتنی بر داده های مالی برای افراد مشارکت کننده در بازار سرمایه سودمند است، تا از تحمیل زیان های عمده مرتبط با خطای داده های مالی به سرمایه گذاران جلوگیری شود. باتوجه به یافته های این تحقیق به سرمایه گذاران بایستی هنگام تصمیم گیری درباره خرید یا فروش سهام خود به ویژگی های حاکمیت شرکتی و مالکیت شرکت ها توجه کنند زیرا بر اساس نتایج این تحقیق ویژگی های حاکمیت شرکتی بر کیفیت سود آنها تاثیر دارد و توصیه می‌شود در شرکت هایی سرمایه گذاری کنند که دارای حاکمیت شرکتی قوی تر و ساختار مالکیت متمرکزتری هستند زیرا هر چه تمرکز مالکیت شرکت متمرکزتر باشد موجب بهبود عملکرد و سود آوری شرکت می‌شود.

از طرفی سهامداران بایستی از مزایای وجود مدیران غیرموظف و مستقل

- [20] Jones, J. J. (1991). **Earnings management during import relief investigations.** *Journal of accounting research*, 29(2), 193-228.
- [21] 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. (2022). **Earnings Management, Board Composition and Earnings Persistence in Emerging Market.** *Sustainability*, 14(3), 1-15.
- [22] Kotsiantis, S. B. (2013). **Decision trees: a recent overview.** *Artificial Intelligence Review*, 39(4), 261-283.
- [23] Lennox, C., Lisowsky, P., & Pittman, J. (2013). **Tax aggressiveness and accounting fraud.** *Journal of accounting research*, 51(4), 739-778.
- [24] Mathew, R. J., Largen, J., & Claghorn, J. L. (1979). **Biological symptoms of depression.** *Psychosomatic Medicine*, 41(6), 439-443.
- [25] Muguto, L., & Muzindutsi, P. F. (2022). **A Comparative Analysis of the Nature of Stock Return Volatility in BRICS and G7 Markets.** *Journal of Risk and Financial Management*, 15(2), 1-27.
- [26] Parte-Esteban, L., & García, C. F. (2014). **The influence of firm characteristics on earnings quality.** *International Journal of Hospitality Management*, 42, 50-60.
- [27] Perotti, P., & Wagenhofer, A. (2014). **Earnings quality measures and excess returns.** *Journal of business finance & accounting*, 41(5-6), 545-571.
- [28] Pratiwi, W., & Pralita, T. (2021). **The Influence of Corporate Governance Mechanism, Accounting Conservatism, And Company Size on Earnings Quality.** *Journal of Economics, Finance and Management Studies*, 4(01), 26-34.
- [29] SeTin, S., & Murwaningsari, E. (2018). **The effect of managerial ability towards earning quality with audit committee as moderating variable.** *Journal of Business and Retail Management Research*, 12(3), 178-189.
- [30] Siueia, T. T., & Wang, J. (2019). **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*, 22(1), 112-121.
- [31] Skinner, D. J., & Soltes, E. (2011). **What do dividends tell us about earnings quality?.** *Review of Accounting Studies*, 16(1), 1-28.
- [5] Bhatia, S., Bajaj, R., & Hazari, S. (2016). **Analysis of credit card fraud detection techniques.** *International Journal of Science and Research*, 5(3), 1302-1307.
- [6] Bliss, J. (1924). **Management Through Accounts,** The Ronald Press Co., New York .
- [7] Boediono, G. S.B. (2005). **Kualitas Laba: Studi Pengaruh Mekanisme Corporate Governance dan Dampak Manajemen Laba dengan Menggunakan Analisis Jalur.** *Simposium Nasional Akuntansi 8.* Solo.
- [8] Chen, S., & Komal, B. (2018). **Audit committee financial expertise and earnings quality: A meta-analysis.** *Journal of Business Research*, 84, 253-270.
- [9] Clyde, P. S. (2007). **Financial accounting an introduction to concepts, methods, and uses.** Canada: South Cengage Learning.
- [10] Dechow, P. M., & Dichev, I. D. (2002). **The quality of accruals and earnings: The role of accrual estimation errors.** *The accounting review*, 77(s-1), 35-59.
- [11] Dechow, P., Ge, W., & Schrand, C. (2010). **Understanding earnings quality: A review of the proxies, their determinants and their consequences.** *Journal of accounting and economics*, 50(2-3), 344-401.
- [12] DeFond, M. L., & Jiambalvo, J. (1994). **Debt covenant violation and manipulation of accruals.** *Journal of accounting and economics*, 17(1-2), 145-176.
- [13] Domingues, A. (2016). **Idiosyncratic Volatility and Earnings Quality: Evidence from United Kingdom.** *FEP Working Papers*, 579: 3-42.
- [14] Ewert, R., & Wagenhofer, A. (2013). **Accounting standards, earnings management, and earnings quality.** *Earnings Management, and Earnings Quality* (April 30, 2013). papers.ssrn.com.
- [15] Fama, E. F., & French, K. R. (2001). **Disappearing dividends: changing firm characteristics or lower propensity to pay?.** *Journal of Financial economics*, 60(1), 3-43.
- [16] Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2004). **Costs of equity and earnings attributes.** *The accounting review*, 79(4), 967-1010.
- [17] Ghosh, A., & Moon, D. (2010). **Corporate debt financing and earnings quality.** *Journal of Business Finance & Accounting*, 37(5-6), 538-559.
- [18] Givoly, D., & Hayn, C. (2000). **The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative?.** *Journal of accounting and economics*, 29(3), 287-320.
- [19] Gutierrez A.L., Rodriguez, M.C. (2020). **Measuring earnings quality: a proposal for a theoretical framework and an empirical method.** <https://aeca.es/method>. <https://aeca.es/wp-content/uploads/2014/05/10a.pdf>. Accessed 5 June 2020.