

مجله پژوهش‌های حسابداری مالی
سال پنجم، شماره سوم، شماره پیاپی (۱۷)، پاییز ۱۳۹۲
تاریخ وصول: ۱۳۹۱/۰۹/۲۱
تاریخ پذیرش: ۱۳۹۲/۰۶/۲۵
صص ۱۳۰-۱۰۵

پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از درخت تصمیم و رگرسیون لجستیک

سید محسن حسینی*، زینب رشیدی^{۱**}

*دانشیار آمار زیستی، دانشگاه علوم پزشکی اصفهان

Hosseini@hlth.mui.ac.ir

^{**}کارشناسی ارشد مدیریت دولتی (گرایش مالی - دولتی)، دانشگاه آزاد خوراسگان

Z_rashidi282@yahoo.com

چکیده

هدف اصلی این مقاله پیش‌بینی ورشکستگی اقتصادی شرکت‌ها در بورس اوراق بهادار تهران با استفاده از درخت تصمیم (CART) و رگرسیون لجستیک است که از تکنیک‌های داده کاوی هستند و می‌توانند به امر پیش‌بینی کمک کنند. نسبت‌های مالی به عنوان متغیرهای مستقل و شرکت‌های سالم و ورشکسته به عنوان متغیر وابسته پژوهش استفاده شده است. جامعه آماری اطلاعات صورت‌های مالی شرکت‌ها در بورس اوراق بهادار تهران طی سال‌های ۱۳۷۸ تا ۱۳۸۹ است. در این پژوهش از نمونه‌گیری استفاده نشده است و شرکت‌ها شامل دو گروه سالم و ورشکسته است گروه ورشکسته بر مبنای ماده ۱۴۱ قانون تجارت و زیان خالص انتخاب شده‌اند و گروه غیر ورشکسته بر اساس معیار سوددهی انتخاب شده‌اند. یافته‌های تحقیق حاکی از آن است که هر دو مدل درخت تصمیم و رگرسیون لجستیک توانایی پیش‌بینی ورشکستگی را با دقت‌های متفاوت دارا هستند و سطح زیر منحنی ROC در مدل رگرسیون لجستیک بیشتر از مدل درخت تصمیم است و از عملکرد بهتری برخوردار است.

واژه‌های کلیدی: پیش‌بینی ورشکستگی، داده کاوی، درخت تصمیم، درخت رگرسیون و طبقه‌بندی، رگرسیون لجستیک

متعددی نیز با توجه به محیط‌های اقتصادی و مالی مختلف، ارائه داده‌اند [۳۷]. در حال حاضر، تکنیک‌های داده کاوی یا کشف دانش (KDD) به سرعت برای کشف الگوهای مفید در مقدار زیادی از داده‌ها، به خصوص در حوزه مسائل مالی توسعه پیدا کرده‌اند. در این پژوهش از مدل‌های درخت تصمیم‌گیری (CART) و رگرسیون لجستیک با روش داده کاوی جهت پیش‌بینی ورشکستگی استفاده شده است به منظور معرفی یک مدل مطلوب که بتواند به عنوان یک سیگنال هشدار دهنده اولیه به مدیران سازمان‌ها؛ مقامات دولتی؛ صاحبان سهام؛ سرمایه‌گذاران، اعتباردهندگان و تحلیل‌گران مالی و کارکنان ... کمک کند.

مروری بر مبانی نظری پژوهش

از دید کلان اقتصادی، بورس اوراق بهادار در فرآیند توسعه اقتصاد ملی از اهمیت زیادی برخوردار است، به طوری که می‌توان گفت میزان فعالیت و اهمیت بورس اوراق بهادار با درجه توسعه یافتگی اقتصاد ملی رابطه‌ای همسو دارد. در واقع، نقش بورس اوراق بهادار در گردآوری نقدینگی راکد و غیر مولد و هدایت منابع پس‌انداز سرگردان به سوی مصارف سرمایه‌گذارای و تأمین منابع مالی فعالیت‌های اقتصادی مولد، جلب مشارکت مردمی، گسترش مبانی مالکیت مردمی، جلوگیری از فرار سرمایه و گسترش اقتصاد زیرزمینی، جذب سرمایه‌های خارجی، کاستن از فشار کسری بودجه دولت و کمک به رشد اقتصادی، انکارناپذیر است [۲۱]. ورشکستگی شرکت‌ها معمولاً بر نقدینگی بازار سرمایه و توسعه اقتصاد مؤثر است. در زمان ورشکستگی، بانک‌ها معمولاً اعتباردهی به شرکت‌های ورشکسته را

پیش‌بینی ورشکستگی به نظر می‌رسد محبوب‌ترین موضوع در استفاده از تکنیک‌های تصمیم‌گیری در داده‌های مالی است. ورشکستگی شرکت‌ها منجر به خسارات اقتصادی برای مدیران، سرمایه‌گذاران، بستانکاران و کارکنان به همراه هزینه‌های اجتماعی می‌شود. به این دلایل پیش‌بینی ورشکستگی یک مسئله مهم در امور مالی است. رقابت روز افزون بنگاه‌های اقتصادی، دستیابی به منابع را محدود کرده است و احتمال ورشکستگی را افزایش داده است. تحقیقات انجام شده قبلی نشان داده است که شرکت‌ها ورشکستگی خود را پنهان می‌کنند و زمانی رسماً اعلام ورشکستگی می‌کنند که در آن زمان تلاش برای جلوگیری از ورشکستگی بیهوده و بسیار دیر است [۴۶]. بر این اساس یکی از اولین تحقیقاتی که در زمینه شرایط مالی شرکت‌ها انجام شده این بوده است که آیا می‌توان زمان ورشکستگی شرکت‌ها را قبل از وقوع و اعلام رسمی تشخیص داد؟

سرمایه‌گذاران و اعتباردهندگان تمایل زیادی برای پیش‌بینی ورشکستگی بنگاه‌ها دارند، زیرا در صورت ورشکستگی هزینه‌های زیادی به آن‌ها تحمیل می‌شود. در ۴۰ سال اخیر موضوع "پیش‌بینی ورشکستگی شرکت‌ها" به عنوان یکی از موضوعات عمده پژوهش در ادبیات مالی تبدیل شده است. تحقیقات آکادمیک زیادی تلاش کرده‌اند بر اساس اطلاعات موجود و روش‌های آماری، بهترین مدل‌های پیش‌بینی ورشکستگی را کشف کنند و در اکثر کشورهای در حال توسعه، پژوهشگران به منظور ساخت مدل‌های جدید، تلاش‌های زیادی نموده و مدل‌های پیش‌بینی

کارگاه‌های کشف دانش و معرفت از پایگاه داده‌ها توسط فیاض، پیاتسکی و همکاران برگزار شد. امروز کنفرانس‌های مختلفی در این زمینه در سراسر دنیا برگزار می‌شود. نگاهی به ترجمه تحت‌اللفظی داده کاوی، به ما در درک بهتر این واژه کمک می‌کند. به معنای استخراج از منابع نهفته و با ارزش زمین اطلاق می‌شود. پیوند این کلمه با کلمه داده، جستجوی عمیق جهت پیدا کردن اطلاعات اضافی مفید که قبلاً نهفته بودند، از داده‌های قابل دسترس حجیم را پیشنهاد می‌کند. داده کاوی یک رشته نسبتاً جدید علمی می‌باشد که از انجام تحقیقات حداقل در رشته‌های آمار، یادگیری ماشین، علوم رایانه خصوصاً مدیریت پایگاه داده‌ها شکل گرفته است. البته مرزهای این رشته‌ها در داده کاوی مبهم می‌باشد [۲۲]. به منظور انتخاب یک مدل، ما باید واقعیت مسأله را مشخص کنیم. واضح است که ما یک مسأله رده‌بندی، که متغیر پاسخ دو دویی و هدف ما پیش‌گویی ورشکستگی شرکت‌ها می‌باشد، داریم. الگوریتم‌های رایج در داده کاوی از این قرارند؛ قواعد وابستگی؛ درخت تصمیم؛ الگوریتم ژنتیک؛ شبکه‌های عصبی؛ K نزدیک‌ترین همسایه؛ رگرسیون خطی / لجستیک [۴۵]. در این پژوهش از مدل‌های درخت تصمیم و رگرسیون لجستیک با روش داده کاوی جهت پیش‌بینی ورشکستگی استفاده شده است.

پیشینه پژوهش

پژوهش‌های داخلی

در ایران پژوهش‌هایی بر روی مدل‌های پیش‌بینی ورشکستگی و درماندگی مالی انجام شده اما پژوهش‌های کمتری در مورد پیش‌بینی

کاهش داده و در ازای وامی که به شرکت‌ها می‌دهند، بهره بالاتری را برای جبران ریسک اضافی درخواست می‌کنند. به صورت مشابهی، مؤسسات سرمایه‌گذاری همچون صندوق‌های بازنشستگی و شرکت‌های بیمه، خرید سهام را کاهش می‌دهند و بیشتر به سراغ سرمایه‌گذاری و خرید اوراق قرضه بانک‌ها یا بازارهای مشابه به آن اقدام می‌کنند. همه این‌ها به کاهش نقدینگی در بازارهای سرمایه، افزایش هزینه سرمایه شرکت‌ها و کاهش رشد اقتصادی منجر خواهد شد [۲۱]. در اینجا مقصود از ورشکستگی، ورشکستگی اقتصادی است و زمانی رخ می‌دهد که بر اثر زیان‌های وارده، حداقل نیمی از سرمایه شرکت از میان برود. یعنی شرکت مشمول ماده ۱۴۱ قانون اصلاحی تجارت شود. در این ماده می‌خوانیم: اگر بر اثر زیان‌های وارده، حداقل نصف سرمایه شرکت از میان برود، هیأت مدیره مکلف است بلافاصله مجمع عمومی فوق‌العاده صاحبان سهام را دعوت نماید تا موضوع انحلال یا بقای شرکت، مورد شور و رأی واقع شود. هرگاه مجمع مزبور رای به انحلال شرکت ندهد، باید در همان جلسه و با رعایت مقررات ماده ۶ این قانون، سرمایه شرکت را به مبلغ سرمایه موجود کاهش دهد [۲۶]. در صورتی که شرکتی مشمول ماده ۱۴۱ قانون اصلاحی تجارت شود، کلیه معاملات سهام آن مسدود می‌شود.

در واقع تکنیک‌های تولید و جمع‌آوری پایگاه داده‌ها بسیار سریع‌تر از توانایی ما در درک و استفاده از آنها رشد کرده است. با توجه به وجود اطلاعات ارزشمند در این پایگاه داده‌ها در اواخر دهه ۸۰ تلاش برای استخراج اطلاعات و استفاده از آن‌ها شروع شد. در فواصل سال‌های ۱۹۹۱ تا ۱۹۹۴

ورشکستگی با استفاده از مدل درخت تصمیم CART توسط محقق یافت شد.

در زیر به ذکر مواردی از پژوهش‌های انجام شده در زمینه پیش‌بینی ورشکستگی پرداخته می‌شود. ابزری، شهشهانی و رسول‌زاده [۱] به بررسی کاربرد مدل آلتمن برای تعیین وضعیت ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در دو صنعت نساجی و ساخت فلزات اساسی پرداختند و نشان دادند که از گروه ورشکسته که شامل ۱۶ شرکت بود مدل آلتمن با ۸۱ درصد اطمینان وضعیت ورشکستگی شرکت‌ها را قبل از ورشکستگی به درستی پیش‌بینی کرده است. گروه غیر ورشکسته نیز شامل ۲۵ شرکت بود که در مدل آلتمن با ۹۲ درصد اطمینان وضعیت عدم ورشکستگی شرکت‌ها را طی دوره ۴ ساله ۷۵ تا ۷۸ به درستی پیش‌بینی کرده است.

سلیمانی امیری [۱۱] در پژوهش خود تحت عنوان بررسی شاخص‌های پیش‌بینی‌کننده ورشکستگی در شرایط محیطی ایران با مدل رگرسیون چندگانه از اطلاعات ۳۰ شرکت سالم و ۳۰ شرکت ورشکسته استفاده کرده است و مدلی جهت پیش‌بینی بحران مالی در شرکت‌های پذیرفته شده در بورس اوراق بهادار ارائه نموده و آن را تا سه سال قبل از بحران مالی مورد آزمون قرار داده بود که نتایج حاصل از تحقیق حاکی از آن بوده که مدل مزبور در یک، دو و سه سال قبل از بحران مالی، به ترتیب ۹۵٪، ۸۳٪ و ۹۵٪ کل نمونه را به طور صحیحی طبقه‌بندی نموده است.

یاریفرد [۲۸] به بررسی مدل‌های پیش‌بینی ورشکستگی (مدل‌های فولمر و اسپرینگ و زمیجوسکی) در بورس اوراق بهادار تهران پرداخت.

نتایج پژوهش او نشان داد که برای یک سال قبل از ورشکستگی مدل فولمر با اطمینان ۹۰ درصد، مدل اسپرینگ با ۷۶ درصد و مدل زمیجوسکی با اطمینان ۷۱ درصد شرکت‌های ورشکسته را درست پیش‌بینی کرده است.

فلاح‌پور و راعی [۱۷] در پژوهش خود تحت عنوان پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از شبکه‌های عصبی مصنوعی به پیش‌بینی درماندگی در ۸۰ شرکت تولیدی پرداخت و به این نتیجه رسیدند که مدل شبکه عصبی در پیش‌بینی ورشکستگی به طور معنی‌داری نسبت به مدل تفکیکی چند متغیره از دقت پیش‌بینی بیشتری برخوردار است.

مهرانی و همکاران [۲۳] در پژوهش خود به دنبال به دست آوردن مدل‌هایی جدید بر مبنای الگوهای پیش‌بینی ورشکستگی زمیسکی و شیراتا متناسب با شرایط محیطی ایران است. نتایج نشان داد که هر دو الگو توانایی تقسیم شرکت‌ها را به دو گروه ورشکسته و غیر ورشکسته دارند و متغیرهای مستقل الگوها تأثیر یکسانی در پیش‌بینی ورشکستگی شرکت‌ها ندارند.

امینی [۳] به بررسی استفاده از مدل فولمر برای تخمین ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخت. این پژوهش در دوره ۸ ساله از سال ۱۳۷۵ تا ۱۳۸۳ بوده است. نتایج پژوهش نشان داد مدل فولمر با ۷۵ درصد اطمینان یک سال قبل از ورشکستگی و با ۶۸ درصد اطمینان دو سال قبل از ورشکستگی توانسته وضعیت ورشکستگی شرکت‌ها را درست پیش‌بینی کند.

کمیجانی و سعادت‌فر [۲۰] به کاربرد مدل‌های شبکه عصبی در پیش‌بینی ورشکستگی اقتصادی

(توانمندی مالی) بر اساس مدل زاوگین پراختند. نتیجه به دست آمده از تحقیق، حاکی از این امر است که به دلیل شرایط محیطی و اقتصادی حاکم بر بورس ایران، مدل زاوگین برای پیش‌بینی احتمال ورشکستگی کارا نبوده و همچنین مدل مذکور برای چهل شرکت غیر ورشکسته آزمون شد که صد در صد کارایی داشته است.

سعیدی و آقای [۱۰] دو مدل با استفاده از شبکه بیز و یک مدل با استفاده از رگرسیون لجستیک برای نمونه انتخاب شده از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در بازه زمانی ۱۳۷۵-۱۳۸۵ ارائه کردند اولین مدل شبکه ساده بیز که مبتنی بر همبستگی شرطی بوده است و می‌تواند با دقت ۹۰٪ شرکت‌های ورشکسته و غیر ورشکسته را درست پیش‌بینی کند. دومین مدل شبکه بیز که مبتنی بر احتمال شرطی است و با دقت ۹۳٪ شرکت‌های ورشکسته و غیر ورشکسته را درست پیش‌بینی می‌کند و در نهایت، مدل رگرسیون لجستیک که یک مدل خطی است می‌تواند با دقت ۹۰٪ شرکت‌های ورشکسته و غیر ورشکسته را درست پیش‌بینی کند.

رهنمای رودپشتی و همکاران [۱۵] در مطالعه خود به ارائه مبانی تئوریک تحقیق و مقایسه نتایج حاصل از به کارگیری مدل‌های آلتمن و فالمر جهت پیش‌بینی ورشکستگی شرکت‌ها پرداختند. از این رو داده‌های جمع‌آوری شده برای سال‌های ۱۳۷۹ تا ۱۳۸۳ مورد آزمون قرار گرفت. به منظور تجزیه و تحلیل داده‌ها روش‌های آماری باینومیل ناپارامتریک مورد استفاده قرار گرفت. نتایج حاصله حاکی از آن است که در پیش‌بینی یک شرکت، تفاوت معنی‌داری بین نتایج دو مدل وجود دارد. همچنین مدل آلتمن

شرکت‌های بازار بورس پرداختند نتایج پژوهش آن‌ها نشان داد که به کارگیری مدل‌های شبکه عصبی مصنوعی توانایی مدیران را برای مقابله با نوسانات اقتصادی ورشکستگی افزایش می‌دهد. پیش‌بینی ورشکستگی شرکت‌های بازار بورس در سال‌های ۱۳۸۵ و ۱۳۸۶ و ترسیم روند ورشکستگی این شرکت‌ها در دوره ۱۳۶۹-۱۳۸۶ از دیگر بخش‌های این پژوهش است. نتایج نشان می‌دهد که در سال ۱۳۸۵ تحت تأثیر سیاست‌های شفاف‌سازی روند ورشکستگی اقتصادی شرکت‌ها به طور چشمگیری افزایش خواهد یافت که با سازگار شدن شرکت‌ها با شرایط جدید تا حدی این روند در سال ۱۳۸۶ تعدیل می‌شود.

فرج‌زاده دهکردی [۱۸] در پژوهش خود نشان داد که الگوریتم ژنتیک قابلیت پیش‌بینی ورشکستگی تا ۹۰ درصد در نمونه آزمون دارد در حالی که تجزیه و تحلیل تفکیکی چند متغیره تا ۷۳ درصد از ورشکستگی شرکت‌ها را می‌تواند پیش‌بینی نماید.

راعی و فلاح‌پور [۱۴] به بررسی کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از نسبت‌های مالی پراختند و این روش را با رگرسیون لجستیک مقایسه کردند. نتایج این پژوهش نشان داد که در پیش‌بینی درماندگی مالی شرکت‌ها مدل ماشین بردار پشتیبان نسبت به مدل لجستیک به طور معنی‌داری از دقت کلی بیشتری برخوردار است و این مدل توانایی بالاتری نیز در تعمیم‌پذیری دارد.

سلیمانی و نیکومرام [۱۳] به درجه‌بندی و رتبه‌بندی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران بر مبنای توانایی نسبی تداوم فعالیت

در پیش‌بینی ورشکستگی محافظه‌کارانه‌تر از مدل فالمر عمل می‌کند.

پورزمانی و همکاران [۷] در تحقیق خود تلاش کردند وضعیت مالی و اقتصادی شرکت‌ها را با استفاده از نسبت‌های شاخص سودآوری، جریان‌های نقدی و رشد پیش‌بینی کنند، از این رو هدف این پژوهش توسعه و آزمون عملکرد مالی و رویکرد ارزیابی ریسک است و به منظور طبقه‌بندی شرکت‌ها به دو گروه موفق و ناموفق، فرضیه تحقیق برای دو گروه شرکت‌های ورشکسته طبق ماده ۱۴۱ قانون تجارت و شرکت‌های حذف شده از بورس طبق تصمیم هیأت پذیرش بورس با استفاده از تحلیل آماری لاجیت مورد آزمون قرار گرفت، نتایج تحقیق نشان می‌دهد نسبت‌های مالی در هر دو گروه نمونه‌گیری، توانایی طبقه‌بندی شرکت‌ها به دو گروه موفق و ناموفق را دارند اما دقت مدل تدوین شده بر اساس شرکت‌های ورشکسته طبق ماده ۱۴۱ قانون تجارت نسبت به شرکت‌های حذف شده از بورس اوراق بهادار بالاتر است.

سلیمانی [۱۲] به بررسی کارایی الگوهای پیش‌بینی بحران مالی در محیط اقتصادی ایران پرداخته است. نمونه‌های آماری تحقیق شامل ۳۰ شرکت موفق و ۳۰ شرکت ناموفق است که از بین شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران انتخاب شد. فرضیه‌های گروه اول با استفاده از روش تجزیه و تحلیل لوجیب آزمون شد. نتایج نشان داد که الگوهای پیش‌بینی بحران مالی توانایی پیش‌بینی تداوم فعالیت شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران را برای یک و دو سال قبل از توقف فعالیت دارند. برای

آزمون فرضیه‌های گروه دوم از آزمون معنی‌دار بودن تفاوت دو ضریب همبستگی استفاده شد. نتایج نشان داد که این الگوها، تفاوت معنی‌داری در پیش‌بینی تداوم فعالیت شرکت‌ها دارند.

پورحیدری و کوپائی حاجی [۶] در این مطالعه برای شناسایی مهمترین متغیرهای حایز اهمیت در پیش‌بینی بحران مالی و ورشکستگی شرکت‌ها، از مدل تابع تفکیکی خطی استفاده شده و مدلی ۹ متغیره طراحی و ارائه شده است. توانایی پیش‌بینی مدل، با استفاده از اطلاعات شرکت‌های دارای بحران مالی و شرکت‌های فاقد بحران مالی ارزیابی شده است و نتایج بررسی نشان می‌دهد که تا پنج سال قبل از بحران مالی می‌توان با استفاده از مدل با دقت نسبتاً بالا آن را پیش‌بینی نمود.

نیکبخت و شریفی [۲۵] به پیش‌بینی ورشکستگی مالی شرکت‌ها در بورس اوراق بهادار تهران به وسیله شبکه‌های عصبی مصنوعی پرداختند. گروه ورشکسته بر مبنای ماده ۱۴۱ قانون تجارت طی سال‌های ۱۳۷۸ لغایت ۱۳۸۵ انتخاب شده‌اند و گروه غیر ورشکسته نیز به صورت تصادفی انتخاب شده‌اند. صحت پیش‌بینی شبکه‌های عصبی با استفاده از نمودار ROC ارائه شده است. نتایج نشان دادند که تفاوت معناداری بین MDA و ANN وجود دارد. همچنین طبق نتایج کم بودن خطای نوع اول بر خطای نوع دوم پیش‌بینی اولویت دارد.

پورزمانی و همکاران [۸] در این تحقیق به بررسی کارایی دو الگوی تحلیل رگرسیون لوجیست و تحلیل ممیزی چند متغیره جهت پیش‌بینی موفقیت یا عدم موفقیت شرکت‌ها پرداختند با داده‌های مربوط به یک سال قبل از سال مبنای،

درصد، و سه سال قبل از وقوع ۸۲/۵ درصد بود. نتایج تحقیق وضوح برتری مدل شبکه‌های عصبی را بر مدل رگرسیونی نشان می‌دهد.

فیروزیان و همکاران [۱۹] به پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از مدل‌های Z آلتمن و الگوریتم ژنتیک پرداختند. نمونه تحت بررسی شامل ۳۶ شرکت ورشکسته و ۳۶ شرکت غیر ورشکسته طی دوره مالی ۵ ساله ۸۴-۸۸ است. در نهایت نتایج این ۲ مدل با هم مقایسه شده است. مدل الگوریتم ژنتیک توانست به‌طور میانگین در یک سال و دو سال قبل از سال منبأ به ترتیب دقتی معادل ۹۰ و ۹۱/۵ درصد داشته باشد و مدل Z آلتمن دقتی معادل ۸۳/۳۲ و ۸۳/۳۲ درصد داشته باشد. توجه به نتایج مدل الگوریتم ژنتیک دقت بیشتری در پیش‌بینی ورشکستگی دارد.

رضایی و گل‌دوز [۱۶] به مقایسه قدرت پیش‌بینی الگوهای ورشکستگی زاوگین، زیمسکی و شیراتا در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداختند. در این تحقیق از دو روش آماری تحلیل تمایزی و رگرسیون لجستیک (به سه روش اینتر، پیش‌رونده و پس‌رونده) استفاده گردید؛ نتایج حاکی از دقت ۹۸/۶ درصد الگوی شیراتا، ۸۷ درصد الگوی زاوگین و ۸۹/۶ درصد الگوی زیمسکی در انطباق با شرایط محیطی ایران است.

جنانی و همکاران [۳۸] به مطالعه روی توانایی نسبت‌های مالی در پیش‌بینی ورشکستگی شرکت‌ها (شواهدی از بورس اوراق بهادار تهران) پرداختند. از اطلاعات ۳۶ شرکت ورشکسته و ۳۶ شرکت غیر ورشکسته بر اساس نسبت Q توین در طول دوره ۲۰۰۵-۲۰۱۱ استفاده کردند. نتایج اولیه نشان

الگوی تحلیل ممیزی کارآمدتر عمل می‌کند و با داده‌های دو سال قبل از سال منبأ الگوی لجستیک کارآمدتر عمل می‌کند؛ ولی در مجموع تفاوت معناداری بین این دو الگو وجود ندارد.

نبوی چاشمی و همکاران [۲۷] به پیش‌بینی ورشکستگی شرکت‌ها با استفاده از مدل لجستیک پرداختند برای انجام تحقیق نمونه‌ای شامل ۴۰ شرکت متشکل از ۲۰ شرکت ورشکسته و ۲۰ شرکت غیر ورشکسته انتخاب گردیده است. بر اساس نتایج تحقیق، مدل لجستیک با متغیرهای توضیحی سرمایه در گردش بر کل دارایی، دارایی جاری بر بدهی جاری و سود ناویژه بر فروش، بیشترین قدرت پیش‌بینی را نسبت به ورشکستگی شرکت‌ها در ایران دارند. دقت پیش‌بینی مدل برای سال ورشکستگی ۸۷/۵ درصد و برای یک سال قبل از ورشکستگی ۷۲/۵ درصد و برای دو سال قبل ۵۲/۵ درصد است.

مکیان و کریمی تکلو [۲۴] در پژوهش خود تحت عنوان پیش‌بینی ورشکستگی شرکت‌های تولیدی با استفاده از شبکه‌های عصبی مصنوعی نشان دادند شبکه‌های عصبی در پیش‌بینی ورشکستگی شرکت‌ها از دقت بالا برخوردار است و مشخص نمودند که تنها یکی از شرکت‌های تولیدی استان کرمان با احتمال ۹۷٪ در سال ۱۳۸۷ ورشکسته خواهد شد.

ابزری، منجمی و بخشایشی [۲] مدل شبکه عصبی باز جریانی را برای پیش‌بینی ورشکستگی بررسی کردند و از مدل رگرسیون خطی بعنوان مدل مقایسه‌ای استفاده کردند دقت کلی مدل شبکه عصبی باز جریانی یک سال قبل از وقوع ۹۵ درصد، در پیش‌بینی ورشکستگی دو سال قبل از وقوع ۹۰

شرکت‌ها بررسی کرد، آلمن در این مطالعه از MAD استفاده کرد. مدلی که او به دست آورد و به Z-Score معروف است، هنوز به‌عنوان شاخصی برای سلامت مالی شرکت‌ها و رشکسته از غیر و رشکسته مورد استفاده قرار می‌گیرد. تئوری اصلی آلمن این بود که مدل پیش‌بینی و رشکستگی او که از پنج نسبت مالی تشکیل می‌شود، می‌تواند برای تشخیص شرکت‌های و رشکسته از غیر و رشکسته مورد استفاده قرار گیرد. او پیشنهاد کرد که از مدلش در ارزیابی اعطای وام‌های تجاری فرآیندهای کنترل داخلی و بررسی گزینه‌های سرمایه‌گذاری استفاده شود.

دیکن [۳۳] تحقیقات انجام شده توسط بیور و آلمن را در مدل جدیدی ادغام کرد. دیکن اظهار کرد که روش بیور توان پیش‌بینی‌کنندگی بهتری دارد در حالی که رویکرد آلمن بینش بهتری می‌دهد. دیکن نیز برای پیش‌بینی و رشکستگی از مدل MDA استفاده کرد ولی ۱۴ نسبتی را که در مطالعه بیور مورد استفاده قرار گرفته بودند به کار برد.

اهلسون [۴۲] مدلی را با استفاده از رگرسیون لجستیک توسعه داد. وی در این پژوهش از ۱۰۵ شرکت و رشکسته و ۲۰۵ شرکت غیر و رشکسته بین سال‌های ۱۹۷۰ تا ۱۹۷۶ استفاده کرد. او نه نسبت مالی را به‌عنوان متغیر مستقل بکار برد. از این نسبت مالی، پنج تای آن‌ها در مطالعات قبلی نیز استفاده شده بودند.

اودوم و شاردا [۴۳] نخستین بار از شبکه‌های عصبی در طراحی مدل‌های پیش‌بینی و رشکستگی استفاده کردند. یافته‌های تحقیق نشان داد که نتایج

می‌دهد که هر دو روش توانایی پیش‌بینی و رشکستگی دارند. اما با دقت‌های متفاوت هستند. احتمال و رشکستگی بر اساس روش CART در سال و رشکستگی، یک سال قبل و دو سال قبل از و رشکستگی به ترتیب ۹۷/۲، ۹۷/۲ و ۸۶/۱ بود. بر اساس روش رگرسیون لجستیک به ترتیب ۹۸/۶، ۹۴/۴ و ۸۴/۷ بوده است و میانگین سه سال در مدل CART ۹۳/۵ و در مدل رگرسیون لجستیک ۹۲/۶ بود.

موسوی شیری و همکاران [۴۱] به پیش‌بینی و رشکستگی شرکت‌ها با استفاده از تکنیک‌های داده کاوی پرداختند. نمونه مورد مطالعه شامل ۱۴۴ شرکت فهرست شده در بورس اوراق بهادار تهران از سال ۲۰۰۵ تا ۲۰۰۹ بود و سپس الگوریتم‌های مختلف داده کاوی از قبیل شبکه‌های عصبی، رگرسیون لجستیک، SVM BAYESNET و درخت‌های تصمیم‌گیری (C5, CHAID, QUEST, CART) در سال T-1، T-2 پیاده‌سازی شده است. نتایج نشان می‌دهد که الگوریتم CART کارآمدتر در پیش‌بینی شرکت‌های و رشکسته و غیر و رشکسته در ایران با متوسط دقت ۹۴/۹۳٪ در سه سال است.

پژوهش‌های خارجی

بیور [۳۱] برای بررسی توان نسبت‌های مالی در پیش‌بینی درماندگی مالی، از تجزیه و تحلیل یک متغیری و بیشتر از نسبت‌های مربوط به جریان‌ات نقدی استفاده کرد. در این پژوهش، بیور عدم توانایی شرکت به انجام تعهدات مالی‌اش را به عنوان درماندگی مالی تعریف کرده است.

آلمن [۲۹] برای اولین بار اثر ترکیبات مختلف نسبت‌های مالی را برای پیش‌بینی درماندگی مالی

یکی مدل ساده بیز و دیگری مدل پیچیده بیز استفاده کردند. نمونه انتخابی آنها شامل ۲۲۸ بانک بود. نتایج پژوهش آنها نشان داد مدل ساده بیز با ۸۰ درصد اطمینان وضعیت ورشکستگی و با ۹۳ درصد اطمینان وضعیت عدم ورشکستگی را درست پیش‌بینی کرده‌اند و مدل پیچیده بیز با ۸۸ درصد اطمینان وضعیت ورشکستگی و با ۹۳ درصد اطمینان وضعیت عدم ورشکستگی را درست پیش‌بینی کرده است.

شین و لی [۴۹] از مدل ژنتیک برای پیش‌بینی در ماندگی مالی شرکت‌ها استفاده کردند و نتایج پژوهش آنها نشان داد که مدل ژنتیک علاوه بر مناسب بودن برای پیش‌بینی در ماندگی مالی، درک آن برای استفاده‌کنندگان بسیار آسان است.

سیلن [۳۲] پژوهشی را برای پیش‌بینی ورشکستگی به وسیله سه الگوی حداقل مجموع انحرافات، DEA و C5.0 (یکی از روش‌های درخت تصمیم) انجام دادند. تحقیقات آنها نشان داد که سه روش به ترتیب دقت طبقه‌بندی ۷۸/۹ درصد، ۸۶/۴ درصد و ۸۵/۵ درصد را تأمین می‌کنند.

والاس [۵۳] یک مدل با استفاده از روش شبکه‌های عصبی طراحی کرد. در این مدل از مقادیر نسبت‌های مالی کلیدی که در مطالعات ورشکستگی گذشته به عنوان بهترین نسبت‌ها گزارش شده بودند، استفاده شد. نسبت‌های استفاده شده وی عبارتند از: سرمایه در گردش به کل دارایی‌ها، جریان نقدی به کل دارایی‌ها، سود خالص به کل دارایی‌ها، کل بدهی‌ها به کل دارایی‌ها، دارایی‌های جاری به بدهی‌های جاری، دارایی‌های سریع به بدهی‌های جاری. مدل والاس دارای دقت کلی ۹۴

روش شبکه‌های عصبی صحیح‌تر، دقیق‌تر و معتبرتر از روش تشخیص خطی چند متغیره است.

تام و کیانگ [۵۲] قابلیت پیش‌بینی شبکه‌های عصبی را با پیش‌بینی آنالیز تشخیص خطی، رگرسیون لجستیک، درخت‌های تصمیم و نزدیکترین کای مجاور مقایسه کردند و نشان دادند که نتایج حاصل از روش شبکه‌های عصبی صحیح‌تر، دقیق‌تر و معتبرتر است.

سرانو - سینکا [۴۷] به مطالعه پیش‌بینی ورشکستگی با استفاده از پرسپترون چند لایه MLP پرداخت. در این تحقیق MLP با MDA مورد مقایسه قرار گرفت. دقت کلی پیش‌بینی با استفاده از MDA، ۸۶ درصد و با استفاده از MLP، ۹۴ درصد به دست آمد.

ژانگ، هو، پاتوا و ایندرو [۵۴] بر اساس اطلاعات ۲۲۰ شرکت (متشکل از ۱۱۰ شرکت ورشکسته و ۱۱۰ شرکت غیر ورشکسته) مدلی را با استفاده از ANN ارائه دادند. نتایج این تحقیق نشان داد که دقت کلی پیش‌بینی شبکه‌های عصبی از رگرسیون لجستیک بیشتر است. گفتنی است که در این مطالعه، از MLP به همراه الگوریتم یادگیری GRG2 استفاده شد.

شاه و مرتزا [۴۸] مدلی را با استفاده از ANN برای پیش‌بینی ورشکستگی ارائه دادند. در این مطالعه از اطلاعات ۶۰ شرکت ورشکسته و ۵۴ شرکت غیر ورشکسته بین سال‌های ۱۹۹۲ تا ۱۹۹۴ استفاده شد. دقت پیش‌بینی این مدل ۷۳ درصد به دست آمد.

سارکر و اسریرام [۵۱] برای اولین بار از شبکه‌های بیز برای پیش‌بینی ورشکستگی استفاده کردند. آنها در این پژوهش از دو مدل متفاوت،

درصد است و ۶۵ نسبت مالی مختلف را در مطالعات گذشته بررسی کرده است.

لی [۳۹] مدل درخت تصمیم ۵.۰، GP، CART و شبکه عصبی و رگرسیون لجستیک را در پیش‌بینی ورشکستگی به کار گرفته است و در بازه زمانی ۱۹۹۹ تا ۲۰۰۳ از اطلاعات ۵۵ شرکت ورشکسته و ۱۱۰ شرکت سالم در تایوان برای طراحی مدل‌ها استفاده کرده است، سطح زیر منحنی ROC در مدل درخت تصمیم CART طراحی شده توسط آنها ۸۶/۱۷٪ و مدل درخت تصمیم ۵.۰ طراحی شده توسط آنها ۸۳/۳۶٪، و مدل درخت تصمیم GP طراحی شده توسط آنها ۸۹/۹۵٪ بود و به همین ترتیب مدل رگرسیون لجستیک ۸۲/۴۱٪ و شبکه عصبی ۸۵/۹۴٪ بود.

مین و لی [۴۰] با استفاده از ماشین بردار پشتیبان اقدام به طراحی مدلی برای پیش‌بینی ورشکستگی شرکت‌ها نمودند. پژوهش آن‌ها نشان داد که نسبت به مدل‌های آماری سنتی، از عملکرد بهتری برخوردار است.

آلفار و گارسیا [۳۰] برای پیش‌بینی درماندگی مالی شرکت‌ها به بررسی دو مدل الگوریتم آدابوست و شبکه‌های عصبی مصنوعی پرداختند. نتایج این پژوهش نشان داد الگوریتم آدابوست نسبت به شبکه‌های عصبی مصنوعی عملکردی بهتر دارد و توان پیش‌بینی این روش ۹۱/۱٪ است.

سووشی و گوتو [۵۰] دو روش DEA و DA - DEA را از دیدگاه ارزیابی ورشکستگی مورد بررسی و مقایسه قرار دادند. آن‌ها در کار تحقیقاتی خود نقاط قوت و ضعف متدولوژیکی DEA و DA - DEA را از لحاظ شکست شرکتی بررسی کردند.

چپ و دیگران [۳۵] از الگوی درخت تصمیم برای پیش‌بینی ورشکستگی استفاده کردند و اثبات کردند مدل درخت تصمیم مناسب برای پیش‌بینی است.

اکسی [۳۴] از مدل درخت تصمیم در داده کاوی برای پیش‌بینی ورشکستگی استفاده کرد. ۱۵ متغیر را به کار گرفت. در این مطالعه از اطلاعات ۷۳ شرکت بین سال‌های ۲۰۰۶ تا ۲۰۰۸ استفاده کرد و داده‌ها را به دو مجموعه آموزش و تست تقسیم کرد. مجموعه آموزش شامل ۷۰٪ مشاهدات و مجموعه تست شامل ۳۰٪ مشاهدات است. الگوریتم CART ۸۷/۹۲٪ داده‌ها را در گام آموزش و ۸۸/۵۷٪ داده‌ها را در گام تست به درستی پیش‌بینی کرد در حالیکه ۱۲/۰۸٪ از شرکت‌های ورشکسته در گام آموزش و ۱۱/۴۳٪ را در گام تست با خطا پیش‌بینی کرد.

پرماچاندرا و همکاران [۴۴] با استفاده از الگوی سوپر کارایی جمعی روش DEA بر اساس دو مرز موفقیت و شکست به ارزیابی شرکت‌ها می‌پردازند. آن‌ها نشان دادند که الگوی DEA در پیش‌بینی شکست شرکت‌ها در مقایسه با پیش‌بینی شرکت‌های سالم به نسبت ضعیف عمل می‌کند.

فرضیه‌های پژوهش

با توجه به مبانی نظری پژوهش فرضیه‌های این پژوهش به شرح زیر تدوین شده‌اند:

فرضیه (۱): مدل پیش‌بینی ورشکستگی درخت تصمیم، توانایی پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران را دارد.

فرضیه (۲): مدل پیش‌بینی ورشکستگی رگرسیون لجستیک (LR)، توانایی پیش‌بینی

دلیل انتخاب سود به عنوان معیار انتخاب شرکت‌های سالم به خاطر ارتباط سود و سود انباشته می‌باشد زیرا سود شرکت در پایان سال مالی به حساب سود و زیان انباشته نقل می‌شود. معیارهای ورشکستگی: ۱- شرکت‌ها مشمول ماده ۱۴۱ شده باشند (نسبت سود و زیان انباشته به سرمایه شرکت می‌باشد که زیان انباشته شرکت باید بیش از ۵۰ درصد سرمایه شرکت باشد) ۲- شرکت‌ها در سه سال متوالی دارای زیان خالص باشند. در این پژوهش از همه شماری استفاده شده است بدین معنی که کلیه شرکت‌های جامعه آماری که دارای شرایط ذکر شده هستند به عنوان نمونه انتخاب و بقیه حذف شده‌اند. بدین ترتیب با توجه به معیارهای انتخاب شرکت‌های گروه اول و دوم، ۹۸ شرکت به عنوان شرکت سالم و ۴۶ شرکت به عنوان شرکت ورشکسته انتخاب شدند.

متغیرهای مورد استفاده

متغیرهای پژوهش را می‌توان به دو دسته طبقه‌بندی کرد، که شامل متغیر وابسته و متغیرهای مستقل می‌باشد و به شرح زیر هستند:

متغیر وابسته در این پژوهش شرکت‌های سالم و ورشکسته است. در این پژوهش با توجه به اهداف پژوهش، شرکت‌های ورشکسته مقدار یک و شرکت‌های سالم مقدار صفر را به خود اختصاص داده‌اند.

در این مطالعه متغیرهای مستقل نسبت‌های مالی می‌باشند که بر اساس بهترین نسبت‌ها در نتایج پژوهش‌های قبلی در نظر گرفته شده‌اند و زمان (t)، (t-۱)، (t-۲) نیز به عنوان یک متغیر مستقل وارد مدل شده است که در نگاره (۱) منعکس گردیده است.

ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران را دارد.

فرضیه (۳): توانایی پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران طی مدل پیش‌بینی ورشکستگی بر اساس مدل درخت تصمیم بیش‌تر از مدل رگرسیون لجستیک (LR) است.

جامعه آماری و انتخاب نمونه

جامعه مورد مطالعه در این پژوهش، شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است که در بین سال‌های ۱۳۷۸ تا ۱۳۸۹ صورت‌های مالی خود را به بورس ارائه داده‌اند. در این مطالعه نمونه‌گیری صورت نگرفته بلکه کل جامعه پژوهش مورد بررسی قرار گرفته است و از داخل جامعه مورد مطالعه شرکت‌ها، با در نظر گرفتن معیارهای زیر وارد مطالعه گردیدند:

- ۱- طی دوره تحقیق تغییر سال مالی نداشته باشد.
- ۲- جزء شرکت‌های سرمایه‌گذاری، واسطه‌گری مالی، هلدینگ، لیزینگ و بانک نباشند.
- ۳- صورت‌های مالی شرکت در دسترس باشند.

در مورد هر شرکت از اطلاعات ۳ سال مالی آن‌ها استفاده شده است و این شرکت‌ها به دو گروه شرکت‌های سالم و ورشکسته طبقه‌بندی شده‌اند. سال مبنا (سال t) در مورد شرکت‌های در مانده مالی، بالاترین سالی است که شرکت مشمول ماده ۱۴۱ قانون تجارت شده است و در مورد شرکت‌های سالم، سالی است که اطلاعات یک سال قبل و دو سال قبل، نسبت به آن جمع‌آوری شده است و معیار سالم بودن شرکت‌ها نیز بر اساس سود دهی آن‌ها در سه سال متوالی ارزیابی می‌شود.

نگاره ۱. متغیرهای مستقل

نام متغیر	نماد	نحوه محاسبه (از چپ به راست)
نسبت جاری	CA/CL	دارایی‌های جاری / بدهی‌های جاری
نسبت سرمایه در گردش به کل دارایی‌ها	WC/TA	سرمایه در گردش / کل دارایی‌ها
نسبت بدهی	TD/TA	کل بدهی‌ها / کل دارایی‌ها
نسبت گردش دارائی‌ها	NS/TA	خالص فروش / کل دارایی‌ها
نسبت دارایی‌های جاری به کل دارایی‌ها	CA/TA	دارایی‌های جاری / کل دارایی‌ها
نسبت سود عملیاتی به کل دارایی‌ها	EBIT/TA	سود عملیاتی / کل دارایی‌ها
حاشیه سود عملیاتی	EBIT/NS	سود عملیاتی / خالص فروش
بازده دارایی‌ها	ROA	سود خالص / کل دارایی‌ها
بازده حقوق صاحبان سهام	ROE	سود خالص / حقوق صاحبان سهام
اندازه شرکت	FS	لگاریتم خالص فروش

مدل‌های مورد استفاده

در این پژوهش از روش‌های داده کاوی و مدل‌های درخت تصمیم‌گیری CART و رگرسیون لجستیک استفاده شده است. در ادامه این دو مدل تشریح می‌گردد.

الف) مدل درخت تصمیم‌گیری CART:

درخت تصمیم‌گیری یکی از روش‌های ناپارامتری رده‌بندی کردن است. این روش با به کارگیری تکنیک‌های بسیار ساده، یک الگوی رده‌بندی را برای مشاهدات موجود معرفی می‌نماید. الگوی معرفی شده توسط این روش، از ساختاری بسیار ساده و قابل درک برای تصمیم‌گیری برخوردار است. با اینکه این روش از تکنیک‌های ساده‌ای استفاده می‌نماید ولی در زمینه تشخیص و پیشگویی می‌تواند به خوبی روش‌های پیچیده‌ای نظیر شبکه‌های عصبی عمل نماید [۳۶]. درخت تصمیم یک روش ساده و توانمند برای طبقه‌بندی یک مجموعه به رده‌های متمایز و همگن می‌باشد که

یک گراف غیر چرخشی شبیه درخت دارد که این درخت توسط مجموعه‌ای از سؤالات نشان داده می‌شود معمولاً هر سؤال با توجه به یک متغیر مطرح می‌شود. یک گراف درخت تصمیم از سه جزء اصلی ریشه^۱، گره داخلی^۲ و گره خارجی^۳ (برگ) تشکیل شده است و روند بدین گونه است که ابتدا یک متغیر کمکی به عنوان ریشه انتخاب می‌گردد و با توجه به یک سری از سؤالات و ویژگی‌ها به چندین گره داخلی تقسیم می‌شود. الگوریتم‌های درخت تصمیم انواع مختلفی دارند که مهم‌ترین آن‌ها عبارتند از: Quest، Chaid، C5.0، Cart [۴]. در این پژوهش از درخت تصمیم‌گیری CART به‌عنوان یک الگوریتم مقایسه‌ای استفاده شده است.

CART یکی از انواع درخت رده‌بندی است که توسط بریمن و همکاران در سال ۱۹۸۴ م معرفی

¹ Root

² Internal Node

³ External Node(Leaf)

یک کاهش در تابع ناچوری خواهیم داشت که بر اساس شاخص جینی به صورت زیر تعریف می‌شود:

$$\text{gini gain} = \text{GG}(T, X) \\ = \text{gini}(t) \sum_{j=1}^n p(t_j | t) \cdot \text{gini}(t_j)$$

از بین چندین متغیر، متغیری مناسب است که مقدار بیشتری را برای $\text{GG}(T, X)$ اختیار نماید. این ملاکی برای انتخاب بهترین از بین متغیر است. از این رو با توجه به تابع ناچوری و شاخص جینی، ابتدا مقدار تابع ناچوری در حالت کلی برای متغیر پاسخ محاسبه می‌گردد، در مرحله بعد برای تمام متغیرهای کمکی، با توجه به بهترین تقسیمات دوتایی برای متغیر پاسخ، مقدار تابع ناچوری در هر یک از دو زیر مجموعه ایجاد شده محاسبه و میانگین وزنی آنها از مقدار تابع ناچوری کل کم می‌گردد. از بین متغیرهای کمکی، متغیری که دارای بیشترین مقدار برای این رابطه باشد، در گام اول برای رده‌بندی درختی انتخاب می‌گردد. در برخورد با متغیرهای کمی، از تقسیمات دوتایی استفاده نموده و نقطه‌ای مانند a (نقطه برش) را تعیین می‌کنیم. لازم به ذکر است که نقطه برش در بسیاری از الگوهای رده‌بندی درختی توسط خود شاخص (در اینجا شاخص جینی) مشخص می‌شود. در برخورد با متغیر کیفی، هر سطح متغیر به عنوان یک زیر شاخه درخت رده‌بندی در نظر گرفته می‌شود [۹].

در مدل CART، برای انتخاب اندازه مناسب از درخت رده‌بندی، از روشی تحت عنوان ارزش پیچیدگی (Complexity Cost) استفاده می‌شود. یک الگوی درختی زمانی مناسب است که علاوه بر

گردید. این مدل یک گراف غیر چرخشی شبیه درخت با تقسیمات دوتایی بر اساس متغیرهای کمکی را برای معرفی یک الگوی رده‌بندی و تشخیصی معرفی می‌نماید. درخت تصمیم از سه جز اصلی شامل ریشه، گره داخلی و برگ تشکیل شده و روند بدین گونه است که ابتدا یک متغیر کمکی به عنوان ریشه انتخاب و با توجه به اهداف مطالعه به چندین گره داخلی تقسیم می‌شود. هر گره داخلی نیز مانند ریشه به گره دیگری تقسیم می‌شود تا در نهایت به هر گره یک رده از متغیر پاسخ منتسب گردد این گره‌ها برگ نامیده می‌شود. به منظور انتخاب متغیرهای مهم در الگوی رده‌بندی درختی، در این بررسی از تابعی تحت عنوان تابع ناچوری و شاخصی به نام جینی استفاده گردید. تابع ناچوری برای گره‌ای مانند t و متغیر وابسته با k رده (c_1, c_2, \dots, c_k) به صورت زیر تعریف می‌شود:

$$i(t) = \Phi \left[\begin{array}{l} P(C=C_1 | t), \\ \dots, p(C=C_k | t) \end{array} \right]$$

شاخص جینی اغلب در مدل‌های درختی با تقسیمات دوتایی در هر گره مورد استفاده قرار می‌گیرد و به صورت زیر تعریف می‌شود [۶]:

$$i(t) = \text{gini}(t) \\ = 1 - \sum_{j=1}^k p^2[c=c_j | t] \\ = \sum_{k=1}^k P(c=c_k | t) P(c=c_k | t)$$

رابطه بالا هنگامی که مشاهدات فقط متعلق به یک رده باشند، برابر صفر است و زمانی که احتمال هر رده برابر باشد، بیشترین امکان ممکن را اختیار می‌کند. با در نظر گرفتن متغیر کمکی x که بر اساس آن گره t به n زیر شاخه تقسیم می‌گردد (هر زیر شاخه با T_j برای j از ۱ تا n نشان داده می‌شود)

در رگرسیون لجستیک به جای حداقل کردن مجذور خطاها (که در رگرسیون معمولی صورت می‌گیرد) احتمال را که یک واقع رخ می‌دهد، حداکثر می‌نماید [۵].

روش پژوهش

با توجه به اهداف این پژوهش، نوع پژوهش از نظر هدف کاربردی و از نظر روش انجام پژوهش، توصیفی-تحلیلی از نوع مقطعی است. در این تحقیق در مرحله اول، با استفاده از نسبت‌های مالی شرکت‌های ورشکسته و سالم مندرج در لیست اوراق بهادار تهران، قدرت پیش‌بینی مدل‌های درخت تصمیم و رگرسیون لجستیک برای سال t ، $t-1$ ، $t-2$ در محیط اقتصادی ایران بررسی گردیده است. سپس در مرحله دوم، با استفاده از فنون درخت تصمیم و رگرسیون لجستیک، مناسب‌ترین مدل برای پیش‌بینی ورشکستگی در این سال‌ها برآزش شده است.

ارزیابی اعتبار داده‌های تحقیق

برای امکان ارزیابی اعتبار داده‌های تحقیق از آمار توصیفی استفاده کرده ایم و نتایج حاصل شده نشان از اعتبار داده‌های پژوهش و همچنین قابلیت مقایسه نتایج تحقیق دارد. نتایج در نگاره (۲) ارائه شده است.

اینکه برای مشاهدات موجود (نمونه فراگیری) خوب عمل می‌نمایید، برای مشاهدات جدید (نمونه آزمون) نیز مناسب باشد. با وجود اینکه با بزرگ شدن اندازه درخت رده‌بندی، دقت رده‌بندی در مورد نمونه یادگیری افزایش می‌یابد، این دقت برای نمونه آزمون از اندازه‌ای به بعد کاهش می‌یابد. روش ذکر شده در واقع تعادلی بین دقت درخت رده‌بندی و اندازه آن برقرار می‌نماید و با توجه به اندازه خط و تعداد گره‌های درخت، تصمیم‌گیری می‌نماید که کدام گره از درخت را حذف نماید.

ب) مدل رگرسیون لجستیک:

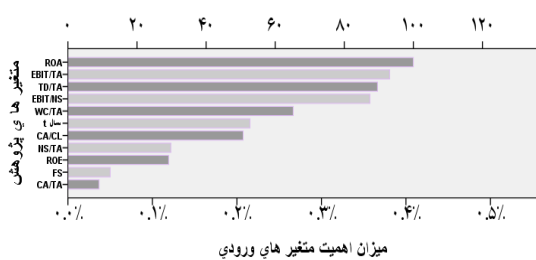
در بسیاری از پژوهش‌ها، متغیر وابسته پیوسته نبوده و ممکن است دو نتیجه داشته باشد. به‌عنوان مثال، فقط یکی از دو ارزش صفر و یک را بپذیرد که ارزش یک به معنای وقوع و ارزش صفر به معنای عدم وقوع حادثه (یا بالعکس) است. برای مثال به کمک میزان تلاش، میزان هوش، موفقیت یا عدم موفقیت فردی را در کنکور تشخیص دهیم و یا درماندگی مالی یک شرکت را با استفاده از چند متغیر تشخیص دهیم. در چنین مواردی از رگرسیون لجستیک استفاده می‌نماییم. رگرسیون لجستیک شبیه رگرسیون معمولی است با این تفاوت که تخمین ضرایب در آن یکسان نیست. رگرسیون لجستیک شبیه رگرسیون معمولی است با این تفاوت که تخمین ضرایب در آن یکسان نمی‌باشد.

نگاره شماره ۲. نتایج آمار توصیفی برای شرکت‌های ورشکسته و شرکت‌های سالم در سال t

FS	CA/TA	ROE	ROA	EBIT/NS	WC/TA	TD/TA	NS/TA	EBIT/TA	CA/CL
شرکت‌های ورشکسته									
۸۷۶۹	۰/۷۷۴	۰/۳۰۶	-۰/۳۲۹	-۰/۴۱۵	-۰/۳۶۲	۲/۳۱۰	۰/۶۲۸	-۰/۱۷۲	۰/۵۰۲
۲/۸۹۴	۰/۲۰۱	۱/۰۶۸	۰/۳۶۷	۰/۶۰۳	۱/۳۷۹	۱/۹	۰/۳۷۰	۰/۲۱۹	۰/۲۹۹
۴/۳۶	۰/۱۸	-۴/۸۳	-۱/۴۳	-۲/۸۷	-۸/۸۱	۰/۱۲	۰/۱۴	-۰/۹۰	۰/۲
۱۲/۳۱	۰/۹۶	۳/۱۸	۰/۰۶	۰/۳۱	۰/۰۵	۹/۶۷	۱/۸۰	۰/۱۱	۱/۱۲
شرکت‌های سالم									
۸/۶۶۱	۰/۶۶۲	۰/۲۷۵	۰/۱۰۶	۰/۱۹۴	۰/۰۹۸	۰/۶۳۳	۰/۹۲۴	۰/۱۴۲	۱/۲۴۹
۲۸/۴۴	۰/۱۸۵	۰/۱۵۷	۰/۰۸۴	۰/۱۶۸	۰/۱۸۸	۰/۱۴۰	۰/۵۷۴	۰/۰۸۹	۰/۶۰۴
۴/۸۸	۰/۱۱	۰/۰۱	۰/۰۰	۰/۰۱	-۰/۳۸	۰/۲۵	۰/۱۶	۰/۰۱	۰/۲۴
۱۲/۳۹	۰/۹۶	۰/۷۲	۰/۳۸	۱/۱۵	۰/۷۹	۰/۹۱	۴/۰۴	۰/۴۳	۵/۸۴

منبع: یافته‌های پژوهش

استفاده در تفکیک داده‌ها به صورت ۵۰ به ۵۰ که ۵۰٪ جزء مجموعه آموزش و ۵۰٪ در مجموعه آزمون قرار می‌گیرند و سپس ۱۰ متغیر وارد مدل گردیده است. نمودار (۱) درجه اهمیت متغیرها را در تفکیک شرکت‌های دو گروه در سال t نشان می‌دهد.



نمودار ۱. میزان تاثیر متغیرهای ورودی بر نتیجه مدل

درخت تصمیم‌گیری در سال t

تأثیرگذارترین متغیرها در سال t-۱ نسبت بدهی، نسبت بازده دارایی‌ها، سود عملیاتی به کل دارایی‌ها، حاشیه سود عملیاتی می‌باشد. تأثیرگذارترین متغیرها در سال t-۲ نسبت بازده دارایی‌ها، سود عملیاتی به کل دارایی‌ها، حاشیه سود عملیاتی، نسبت بدهی است.

پیش‌بینی ورشکستگی شرکت‌ها با استفاده از مدل درخت تصمیم‌گیری CART

در این حالت با استفاده از روش درخت تصمیم، داده‌ها را رده‌بندی کرده و هر کدام از آن‌ها به یک رده تخصیص داده می‌شود. در این روش مشاهده‌ها بر اساس نزدیک‌ترین رده مرتبط رده‌بندی می‌شوند و سعی می‌شود مشاهده‌های موجود در هر رده بیشترین شباهت را به یکدیگر داشته باشند. همچنین از متغیرها به‌عنوان یک رابط استفاده نمی‌شود بلکه متغیرها فقط به‌عنوان تفکیک‌کننده مورد استفاده قرار می‌گیرند. در واقع در درخت تصمیم به وجود آمده تفکیک بر اساس اهمیت متغیرها در به وجود آمدن گره‌های تفکیکی انجام می‌گیرد. در ادامه به اجرای مدل درخت تصمیم در cart در پیش‌بینی ورشکستگی شرکت‌ها می‌پردازیم همچنین جهت آزمون فرضیه‌ها در این مدل از نرم‌افزار Clementine استفاده شده است. جهت پیش‌بینی وضعیت شرکت‌ها، داده‌های پژوهش به دو گروه آموزش و آزمون تفکیک و نسبت مورد

نسبت بازده دارایی‌ها				نسبت بازده دارایی‌ها			
بیشتر از ۰/۰۱۵ -		مساوی و کمتر از ۰/۰۱۵ -		بیشتر از ۰/۰۱۵ -		مساوی و کمتر از ۰/۰۱۵ -	
ورشکسته	سالم	ورشکسته	سالم	ورشکسته	سالم	ورشکسته	سالم
۲	۵۵	۲	۹۶/۵	۲۳	۰	۲	۹۵/۶
%۳/۵	%۹۶/۵	%۳/۵	%۹۶/۵	%۱۰۰	%۰	%۴/۴	%۹۵/۶

نسبت بدهی				نسبت بدهی			
بیشتر از ۰/۸۶۰ -		مساوی و کمتر از ۰/۸۶۰ -		بیشتر از ۰/۸۶۰ -		مساوی و کمتر از ۰/۸۶۰ -	
ورشکسته	سالم	ورشکسته	سالم	ورشکسته	سالم	ورشکسته	سالم
۱	۵۲	۱	۷۵	۰	۴۳	۲	۰
%۱/۹	%۹۸/۱	%۲۵	%۷۵	%۰	%۱۰۰	%۱۰۰	%۰

نمودار ۲. درخت تصمیم ایجاد شده در داده‌های آموزش و آزمون بر اساس متغیرهای پژوهش در سال t

ورشکسته طبقه‌بندی می‌گردد و نسبت بدهی مهم‌ترین متغیر در طبقه‌بندی شرکت‌ها در مرحله بعدی است و زمانی که نسبت بدهی بزرگتر و کوچکتر و مساوی از ۰/۸۶۰ باشد دقت پیش‌بینی در این طبقات به ۱۰۰ درصد می‌رسد. در نهایت نگاره (۳) نتایج مربوط به پیش‌بینی و طبقه‌بندی مدل CART را به تصویر می‌کشد.

در نمودار (۲) با توجه به الگوی رده‌بندی درختی cart، نسبت بازده دارایی‌ها مهم‌ترین نقش را جهت پیش‌بینی وضعیت شرکت‌ها ایفا می‌کند. در تفسیر این نمودار می‌توان گفت به‌عنوان مثال در گام آموزش، برای طبقه‌ای که نسبت بازده دارایی‌های آن بزرگتر از ۰/۰۱۵ - باشد در داده‌های آموزش ۹۵/۶٪ از شرکت‌ها را به عنوان شرکت سالم و ۳/۳٪ از شرکت‌ها به‌عنوان شرکت

نگاره ۳. نتایج دقت پیش‌بینی درخت تصمیم

میانگین سه سال	درصد پیش‌بینی درست			موارد تشخیص به‌عنوان شرکت ورشکسته			موارد تشخیص به‌عنوان شرکت سالم			دسته	داده‌های آموزش	داده‌های آزمون
	t-2	t-1	t	t-2	t-1	t	t-2	t-1	t			
	%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	۰	۰	۰	۵۱	۴۲			
%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	۲۳	۲۵	۲۵	۰	۰	۰	دقت تفکیک کلی مدل		
%۱۰۰	%۱۰۰	%۱۰۰	%۱۰۰	۱	۲	۳	۴۶	۵۴	۵۲	سالم	ورشکسته	
%۹۶/۹۷	%۹۵/۷	%۱۰۰	%۹۵/۲	۲۲	۲۱	۲۰	۱	۰	۱	دقت تفکیک کلی مدل		
%۹۶/۴	%۹۷/۱	%۹۷/۴	%۹۴/۷									

منبع: یافته‌های پژوهش

پیش‌بینی ورشکستگی شرکت‌ها با استفاده از

مدل رگرسیون لجستیک

به منظور ساخت مدل رگرسیون لجستیک از نرم‌افزار spss نسخه ۲۰ استفاده شده است. در این بخش جهت پیش‌بینی ورشکستگی شرکت‌ها، ۱۰ نسبت به طور همزمان وارد مدل رگرسیون لجستیک گردید. در این مدل جهت بررسی معنی‌داری کلی مدل از آماره کای-دو استفاده می‌گردد و گام مناسب برای مدل از طریق معیار دو برابر لگاریتم تابع درستنمایی (2LL-) انتخاب می‌گردد. نگاره (۴) آماره کای-دو مربوط به برآزش و معنی‌داری

کلی مدل را به تصویر می‌کشد. با توجه به اطلاعات منعکس در نگاره (۴) سطح معنی‌داری مربوط به آزمون کای-دو (۰/۰۰۱) کمتر از میزان خطای (۰/۰۵) می‌باشد و این نشان از معنی‌داری کلی مدل می‌باشد و گامی که مقدار 2LL- آن در بین گام‌های انجام شده حداقل باشد انتخاب می‌گردد. پس از بررسی معنی‌داری کلی مدل از آزمون والد جهت بررسی معنی‌داری متغیرهای مستقل استفاده گردیده است. همه متغیرها وارد مدل می‌شوند متغیرهایی که معنی‌دار نیستند از مدل خارج می‌شوند و متغیرهای معنی‌دار باقی می‌مانند.

نگاره ۴. آماره کای-دو مربوط به برآزش و معنی‌داری کلی مدل رگرسیون لجستیک

سال	کای-دو	درجه آزادی	سطح معنی‌داری	معیار منفی دو برابر لگاریتم تابع درستنمایی
سال ۱	۱۶۱/۸۱۷	۲	۰/۰۰۱	۱۸/۶۰۰
سال ۱-۲	۱۶۶/۷۸۷	۲	۰/۰۰۱	۱۳/۶۳۰
سال ۲-۳	۱۷۱/۶۱۶	۲	۰/۰۰۱	۸/۸۰۲

منبع: یافته‌های پژوهش

نگاره ۵. نتایج مربوط به آزمون والد و ضرایب متغیرها

سال	نام متغیر مستقل استخراج شده	ضریب	انحراف استاندارد	آماره والد	درجه آزادی	سطح معنی‌داری
	بازده دارایی	-۷۸/۶۷۸	۲۶/۳۸۳	۸/۸۹۳	۱	۰/۰۰۳
سال ۱	بازده حقوق صاحبان سهام	-۱/۵۶۷	۰/۷۶۴	۴/۲۱۰	۱	۰/۰۴
	مقدار ثابت	-۰/۵۵۸	۰/۶۴۵	۰/۷۵۰	۱	۰/۳۸۶
	بازده دارایی	-۹۲/۹۸۵	۳۷/۱۱۰	۶/۲۷۸	۱	۰/۰۱۲
سال ۱-۲	بازده حقوق صاحبان سهام	-۱/۴۳۴	۰/۷۳۰	۳/۸۶۱	۱	۰/۰۴۹
	مقدار ثابت	-۰/۵۴۹	۰/۷۹۴	۰/۴۷۸	۱	۰/۴۸۹
	نسبت بدهی	۲۰/۵۸۵	۱۳/۷۷۸	۲/۲۳۲	۱	۰/۱۳۵
سال ۲-۳	بازده حقوق صاحبان سهام	-۱۴۲/۴۸۲	۸۱/۰۲۸	۳/۰۹۲	۱	۰/۰۷۹
	مقدار ثابت	-۱۷/۲۰۴	۱۱/۹۰۰	۲/۰۹۰	۱	۰/۱۴۸

منبع: یافته‌های پژوهش

توجه به اطلاعات منعکس در نگاره (۵) متغیرهای مستقل باقی مانده در سطح خطای ۰/۰۵ معنی‌دار

نگاره (۵) نتایج مربوط به آماره والد، ضریب ثابت و ضرایب متغیرها را به تصویر می‌کشد. با

بوده و با متغیر وابسته رابطه معنی‌دار دارند. در این نمودار می‌توان مشاهده کرد تأثیرگذارترین عوامل در سال t نسبت بازده دارایی و بازده حقوق صاحبان سهام می‌باشد و رابطه بازده حقوق صاحبان سهام و بازده دارایی‌ها با ورشکستگی، منفی است و با کاهش نسبت بازده حقوق صاحبان سهام و بازده دارایی‌ها احتمال ورشکستگی افزایش می‌یابد. تأثیرگذارترین عوامل در سال $t-1$ نسبت بازده دارایی و بازده حقوق صاحبان سهام است. در سال $t-2$ نسبت بدهی و نسبت بازده حقوق صاحبان سهام معنادار نمی‌باشند.

نگاره ۶. آماره هاسمر و لمشو مربوط به اعتبار مدل

کای - دودرجه آزادی سطح معنی‌داری			
سال t	۸/۸۵۲	۸	۰/۳۵۵
سال $t-1$	۸/۴۱۴	۸	۰/۳۹۴
سال $t-2$	۰/۰۴۵	۸	۱

منبع: یافته‌های پژوهش

با توجه به اطلاعات منعکس شده در نگاره (۶) آماره هاسمر - لمشو دارای توزیع X^2 با $J-2$ درجه آزادی می‌باشد (J تعداد گروه‌ها است) که نشان می‌دهد توزیع نسبت به نمونه حساس است. لذا زمانی که تعداد نمونه خیلی بزرگ باشد، تفاوت‌های کمی بین مقادیر مشاهده شده و پیش‌بینی شده از

طریق این روش نشان داده می‌شود و آزمون معتبر است. مقدار این آماره در مدل برآورد شده در سال t دارای توزیع X^2 با ۸ درجه آزادی برابر ۸/۸۵۲ می‌باشد و احتمال آن بزرگ‌تر از ۰/۰۵ و برابر ۰/۳۵۵ بدست آمده است و مقدار این آماره در مدل برآورد شده در سال $t-1$ دارای توزیع X^2 با ۸ درجه آزادی برابر ۸/۴۱۴ می‌باشد و احتمال آن بزرگ‌تر از ۰/۰۵ و برابر ۰/۳۹۴ بدست آمده است و مقدار این آماره در مدل برآورد شده در سال $t-2$ دارای توزیع X^2 با ۸ درجه آزادی برابر ۰/۰۴۵ است و احتمال آن بزرگ‌تر از ۰/۰۵ و برابر ۱ بدست آمده است. بنابراین فرض صفر که بیانگر نکویی برآزش است، پذیرفته می‌شود (رد نمی‌گردد). پس متغیرهای مستقل مدل قدرت پیش‌بینی را دارند.

پس از بررسی معنی‌دار کلی مدل و آماره‌های مربوط به آن، جهت آزمون فرضیه‌ها، باید دقت دسته‌بندی مدل در پیش‌بینی ورشکستگی شرکت‌ها مورد بررسی قرار گیرد. نگاره (۷) نتایج مربوط به دقت پیش‌بینی مدل را در سه سال به تصویر می‌کشد.

نگاره ۷. دقت پیش‌بینی در مدل رگرسیون لجستیک

میانگین سه سال	درصد پیش‌بینی درست			موارد تشخیص به‌عنوان شرکت ورشکسته			موارد تشخیص به‌عنوان شرکت سالم			دسته
	$t-2$	$t-1$	t	$t-2$	$t-1$	t	$t-2$	$t-1$	t	
	٪۹۹/۳۳	٪۹۸	٪۱۰۰	٪۱۰۰	۲	۰	۰	۹۶	۹۸	
٪۹۷/۸	٪۹۷/۸	٪۹۷/۸	٪۹۷/۸	۴۵	۴۵	۴۵	۱	۱	۱	ورشکسته
٪۹۸/۸۳	٪۹۷/۹	٪۹۹/۳	٪۹۹/۳							دقت تفکیک کلی مدل

منبع: یافته‌های پژوهش

آزمون فرضیه‌های پژوهش

دقت دسته‌بندی از اهمیت خاصی برخوردار است. با توجه به نتایج به دست آمده در بخش‌های قبل، دقت دسته‌بندی بر اساس هریک از مدل‌های فوق در نگاره (۸) خلاصه شده است.

در این پژوهش برای پیش‌بینی ورشکستگی شرکت‌ها از مدل درخت تصمیم و رگرسیون لجستیک استفاده شده است. در مدل‌های داده کاوی

نگاره ۸. مقایسه دقت پیش‌بینی مدل درخت تصمیم‌گیری CART و رگرسیون لجستیک

مدل		دقت پیش‌بینی در سال t		دقت پیش‌بینی در سال t-۱		دقت پیش‌بینی در سال t-۲	
		ورشکسته	سالم	ورشکسته	سالم	ورشکسته	سالم
درخت	داده‌های آموزش	٪۱۰۰	٪۱۰۰	٪۱۰۰	٪۱۰۰	٪۱۰۰	٪۱۰۰
تصمیم‌گیری	داده‌های آزمون	٪۹۴/۵	٪۹۵/۲	٪۹۶/۴	٪۱۰۰	٪۹۷/۹	٪۹۵/۷
رگرسیون لجستیک		٪۱۰۰	٪۹۷/۸	٪۱۰۰	٪۹۷/۸	٪۹۸	٪۹۷/۸

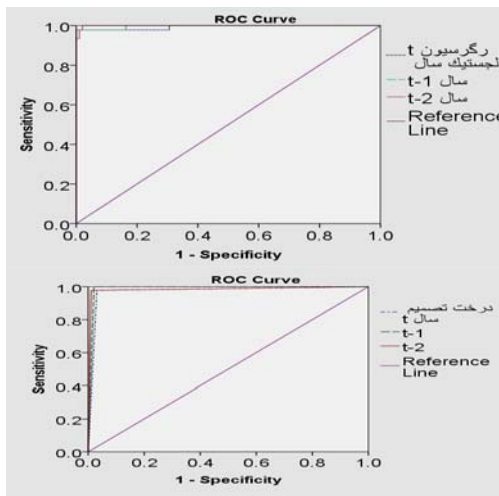
منبع: یافته‌های پژوهش

نگاره ۹. سطح زیر منحنی ROC

سال t	سال t-۱	سال t-۲
۰/۹۷۴	۰/۹۹۰	۰/۹۸۴
درخت تصمیم‌گیری CART		
۰/۹۹۳	۰/۹۹۶	۰/۹۹۹
رگرسیون لجستیک		

منبع: یافته‌های پژوهش

این نتیجه بیانگر آن است که از نظر سطح زیر منحنی ROC مدل رگرسیون لجستیک در سال t، t-۱، t-۲ توانایی بالاتری در پیش‌بینی ورشکستگی دارد و به این مدل از عملکرد بهتری برخوردار است و فرضیه سوم پژوهش رد می‌شود.



نمودار ۳. سطح زیر منحنی ROC در مدل درخت تصمیم و رگرسیون لجستیک

به‌طورکلی و بر مبنای نتایج حاصل از دو مدل فوق می‌توان گفت جهت بررسی فرضیه اول و دوم فرضیه صفر رد می‌شود و فرضیه یک مبنی بر اینکه الگوها توانایی درست پیش‌بینی کردن را دارند پذیرفته می‌شود. جهت آزمون فرضیه سوم از سطح زیر منحنی ROC استفاده شده است. در مطالعات مدل‌سازی جهت سنجش میزان دقت مدل و پیش‌بینی صورت گرفته از منحنی ROC و سطح زیر منحنی نیز استفاده می‌کنند. محدوده زیر منحنی، میزان توانایی مدل در تفاوت قائل شدن بین دو نتیجه است این میزان تمایز نام دارد. هر چه سطح زیر منحنی ROC به عدد ۱ نزدیک‌تر باشد دقت مدل در معیار خوب و هر چه این عدد به ۰/۵ نزدیک‌تر باشد، نشان از دقت پایین مدل و پیش‌بینی نامناسب مدل است. برای مقایسه عملکرد دو مدل از سطح زیر منحنی ROC که با استفاده از روش شناسی‌های داده کاوی به دست می‌آید استفاده می‌شود. نتایج به شرح نگاره (۹) است.

نتیجه

پیش‌بینی ورشکستگی شرکت‌ها یکی از مطالعه‌های جالب و با اهمیت در حوزه مالی است. با پیش‌بینی ورشکستگی و پس از آن ریشه‌یابی مسأله و حل آن، می‌توان به نتایج بسیار رضایت‌بخشی دست یافت. در این پژوهش از مدل درخت تصمیم و رگرسیون لجستیک که از روش‌های داده کاوی هستند، جهت پیش‌بینی ورشکستگی شرکت‌های داخلی، استفاده شده است. نتایج به دست آمده از پژوهش نشان داد، اطلاعات صورت‌های مالی دارای قدرت پیش‌بینی‌کننده بالا می‌باشند و شرکت‌های گروه ورشکسته در مراحل ورشکستگی، فروش، سود ویژه و دارایی‌های کمتری نسبت به گروه دوم دارند که در نهایت به تفاوت معنادار نسبت‌های مالی دو گروه منجر خواهد شد. مدل ایجاد شده توسط الگوریتم CART توانست به‌طور میانگین در دو سال قبل از ورشکستگی در داده‌های آزمون دقتی معادل ۹۷/۱ درصد داشته باشد؛ در حالی که مدل رگرسیون لجستیک دقتی معادل ۹۷/۹ درصد داشته، در یکسال قبل از سال مبنا مدل الگوریتم CART به‌طور میانگین دقتی معادل ۹۷/۴ درصد داشته در حالی که مدل رگرسیون لجستیک دقتی معادل ۹۹/۳ درصد داشته، در سال مبنا مدل الگوریتم CART به‌طور میانگین دقتی معادل ۹۴/۷ درصد داشته ضمن اینکه در این سال مدل رگرسیون لجستیک دقتی معادل ۹۹/۳ درصد داشته است. بنابراین با توجه به نتایج حاصل دقت کلی و سطح زیر منحنی ROC در مدل رگرسیون لجستیک نسبت به مدل CART بیشتر است؛ پس به‌طور کلی مدل رگرسیون لجستیک ابزار مناسب‌تری برای پیش‌بینی ورشکستگی شرکت‌ها

است. همچنین با استفاده از نتایج حاصل از آزمون فرضیه‌های پژوهش می‌توان نتیجه گرفت، مدل‌های الگوریتم CART و مدل رگرسیون لجستیک ابزارهای مناسبی برای پیش‌بینی ورشکستگی محسوب می‌شوند و با آزمون فرضیه سوم مشاهده شد که مدل رگرسیون لجستیک محافظه‌کارتر از مدل CART عمل می‌نماید. نتایج توان پیش‌بینی دو مدل با نتایج پژوهش‌های صورت گرفته توسط جنانی و همکاران (۲۰۱۳) و موسوی شیرینی و همکاران (۲۰۱۲) و لی (۲۰۰۴) و راعی و فلاح‌پور (۱۳۸۷) و سعیدی و آقایی (۱۳۸۸) و اهلسون (۱۹۸۰) و تام و کیانگ (۱۹۹۲) و ژانگ و همکاران (۱۹۹۹) و جپ و همکاران (۲۰۱۰) و اکسی (۲۰۱۱) سازگار است و نتایج مقایسه‌ای دو مدل با نتایج پژوهش‌های صورت گرفته توسط جنانی و همکاران (۲۰۱۳) و موسوی شیرینی و همکاران (۲۰۱۲) و لی (۲۰۰۴) ناسازگار است.

پیشنهادهای کاربردی حاصل از نتایج پژوهش

با توجه به نتایج حاصل از پژوهش و اهمیت موضوع می‌توان پیشنهادهای زیر را ارائه نمود:

الف) سازمان بورس و اوراق بهادار تهران برای اطمینان نسبی از وضعیت مناسب مالی آتی شرکت‌ها می‌تواند از مدل‌های این پژوهش در پذیرش شرکت جدید به سازمان بورس استفاده نماید.

ب) بانک‌ها و مؤسسات مالی و اعتباری از مدل‌های این پژوهش در اعطا وام‌ها با مبالغ بالا به متقاضیان صاحبان صنایع می‌توانند استفاده نمایند.

ج) کارگزاران بورس، تحلیل‌گران و مشاوران مالی که وظیفه آن‌ها تجزیه و تحلیل وضعیت مالی شرکت‌های داخل بورس و تشریح وضعیت مالی آینده شرکت‌ها برای متقاضیان خرید سهام شرکت‌ها

منابع

- می‌باشند نیز می‌توانند از مدل‌های این پژوهش استفاده نمایند.
- د) مدیران شرکت‌ها همواره مایل به ارزیابی نقاط ضعف و پیش‌بینی تهدیدهای آتی می‌باشند. یکی از روش‌های ارزیابی نقاط ضعف مالی در آینده و در نهایت درماندگی مالی استفاده از مدل‌های مبتنی بر داده‌های مالی می‌باشد. مدیران از مدل‌های ارائه شده در این پژوهش برای تشخیصی درماندگی مالی می‌توانند استفاده نمایند.
- ه) شرکت‌های سرمایه‌گذاری همواره در پی کسب وضعیت مناسب برای سرمایه‌گذاری هستند. اصولاً آنها قبل از اقدام به سرمایه‌گذاری در صنایع مختلف، وضعیت این صنایع را با روش‌های گوناگون مورد ارزیابی و تجزیه و تحلیل قرار می‌دهند تا نسبت به سودآوری آنها مطمئن شوند. یکی از انواع روش‌های ارزیابی مورد استفاده این شرکت‌ها، ارزیابی وضعیت ورشکستگی است. بنابراین با توجه به نتایج حاصل از این تحقیق، شرکت‌های سرمایه‌گذاری می‌توانند از مدل‌های استفاده شده در این پژوهش برای ارزیابی ورشکستگی شرکت‌ها استفاده نمایند.
- و) صنایع همواره برای مقاصد توسعه محصول و کنترل بر مواد اولیه مورد نیاز خود، تصمیم به ادغام و خرید شرکت‌های دیگر می‌گیرند. این صنایع در زمان تصمیم به خرید، آنها را مورد ارزیابی قرار داده و وضعیت مالی آنها را بررسی کامل می‌نمایند، تا پس از ادغام دچار مشکلات مالی نشود. با توجه به نتایج ارائه شده در این تحقیق، این صنایع می‌توانند وضعیت فعالیت این شرکت‌ها را در آینده با استفاده از مدل‌های بکار گرفته شده در این پژوهش پیش‌بینی نمایند.
- ۱- ابزری، مهدی، میرشمس شهشانی، مرتضی و مهدی رسول‌زاده. (۱۳۸۰) بررسی کاربرد مدل آلتمن برای تعیین وضعیت ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران (صنایع نساجی و ساخت فلزات اساسی). پایان‌نامه کارشناسی ارشد، دانشکده مدیریت دانشگاه اصفهان، اصفهان.
- ۲- ابزری، مهدی، منجمی، امیرحسن و مروه بخشایش. (۱۳۹۰) پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از شبکه‌های عصبی باز جریانی. پایان‌نامه کارشناسی ارشد، دانشکده مدیریت دانشگاه اصفهان، اصفهان.
- ۳- امینی، پیمان. (۱۳۸۵) بررسی امکان سنجی استفاده از مدل فولمر برای تخمین ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. پایان‌نامه کارشناسی ارشد، رشته حسابداری دانشگاه تربیت مدرس، تهران.
- ۴- آذر، عادل، احمدی، پرویز و محمد وحید بسط. (۱۳۸۹) طراحی مدل انتخاب نیروی انسانی با رویکرد داده کاوی، نشریه مدیریت فناوری و اطلاعات، ش ۴، صص ۲۲-۳.
- ۵- پیرایش، رضا، منصور، علی و صابر امجدیان. (۱۳۸۸) طراحی مدل‌های ریاضی مبتنی بر جریان‌های نقدی برای پیش‌بینی ورشکستگی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، مجله توسعه سرمایه، ش ۴، صص ۹۴-۷۳.
- ۶- پورحیدری، امید و مهدی کوپائی حاجی. (۱۳۸۹) پیش‌بینی بحران مالی شرکت‌ها با استفاده از مدل مبتنی بر تابع تفکیکی خطی. مجله پژوهش‌های حسابداری مالی، دوره ۲، ش ۱، صص ۴۶-۳۳.

- ۷- پورزمانی، زهرا، جهان‌شاد، آزیتا و شهرام عین‌قلایی. (۱۳۸۸) پیش‌بینی وضعیت مالی و اقتصادی شرکت‌ها با استفاده از نسبت‌های مالی مبتنی بر سودآوری، جریان‌های نقدی و رشد. *پژوهشنامه حسابداری مالی و حسابرسی*، دوره ۱، ش ۳، صص ۹۳-۱۱۵.
- ۸- پورزمانی، زهرا، توانگر حمزه کلایی، افسانه و آوا کیارسی. (۱۳۸۹) بررسی کارایی الگوی لجیت و تحلیل تمایزی چند متغیره در پیش‌بینی وضعیت مالی شرکت‌های بورس اوراق بهادار تهران. *پژوهشنامه حسابداری مالی و حسابرسی*، دوره ۲، ش ۵: ۱۲۲-۹۴.
- ۹- حسینی، محسن، تذهیبی، مهدی، امینی، مسعود، زارع، اصغر و حسن جهانی هاشمی. (۱۳۸۹) کاربرد درخت رده‌بندی برای پیشگویی رتیوپانی دیابتیک و مقایسه با مدل تابع تشخیص در بیماران دیابتی نوع ۲، *مجله دانشکده پزشکی اصفهان*، ش ۱۰۴، صص ۱۷-۲۷.
- ۱۰- سعیدی، علی و آرزو آقایی. (۱۳۸۸) پیش‌بینی درماندگی مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از شبکه‌های بیز، *فصلنامه بررسی‌های حسابداری و حسابرسی*، دوره ۱۶، ش ۵۶، صص ۷۸-۵۹.
- ۱۱- سلیمانی امیری، غلامرضا. (۱۳۸۲) نسبت‌های مالی و پیش‌بینی بحران مالی شرکت‌ها در بورس اوراق بهادار تهران. *تحقیقات مالی*، دوره ۵، ش ۱۵، صص ۱۳۶-۱۲۱.
- ۱۲- سلیمانی، غلامرضا. (۱۳۸۹) ارزیابی کارایی الگوهای پیش‌بینی بحران مالی برای شرکت‌های ایرانی. *مجله دانش حسابداری*، دوره ۱، ش ۲، صص ۱۵۸-۱۳۹.
- ۱۳- سلیمانی، اعظم و هاشم نیکومرام. (۱۳۸۷) ارزیابی توانمندی مالی شرکت‌های بهادار تهران و ارائه مدل مناسب ارزیابی. *پژوهشنامه اقتصادی*، دوره ۲۹، ش ۲، صص ۲۵۳-۲۷۹.
- ۱۴- راعی، رضا و سعید فلاح‌پور. (۱۳۸۷) کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از نسبت‌های مالی، *فصلنامه بررسی‌های حسابداری و حسابرسی*، دوره ۱۵، ش ۵۳، صص ۳۴-۱۷.
- ۱۵- رهنمای رودپشتی، فریدون، علی‌خانی، راضیه و مهدی مران‌جوری. (۱۳۸۸) بررسی کاربرد مدل‌های پیش‌بینی ورشکستگی آلتمن و فالمر در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. *بررسی‌های حسابداری و حسابرسی*، دوره ۱۶، ش ۵۵، صص ۳۴-۱۹.
- ۱۶- رضایی، فرزین و مهدی گل‌دوز (۱۳۹۰). مقایسه قدرت پیش‌بینی الگوهای ورشکستگی زاوگین، زیمسکی و شیراتا در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. *مجله مدیریت توسعه و تحول*، ش ۶، صص ۸۱-۶۹.
- ۱۷- فلاح‌پور، سعید و رضا راعی. (۱۳۸۳) پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از شبکه‌های عصبی مصنوعی، *تحقیقات مالی*، ش ۱۷، صص ۶۹-۳۹.
- ۱۸- فرج‌زاده دهکردی، حسن. (۱۳۸۶) کاربرد الگوریتم ژنتیک در پیش‌بینی ورشکستگی. پایان‌نامه کارشناسی ارشد، دانشگاه تربیت مدرس، تهران.
- ۱۹- فیروزیان، محمود؛ جاوید، داریوش و نرگس نجم‌الدینی. (۱۳۹۰) کاربرد الگوریتم ژنتیک در پیش‌بینی ورشکستگی و مقایسه آن با مدل Z آلتمن در شرکت‌های پذیرفته شده در بورس اوراق بهادار

- تهران. بررسی‌های حسابداری و حسابرسی، دوره ۱۸، ش ۶۵، صص ۹۹-۱۱۴.
- ۲۰- کمیجانی، اکبر و جواد سعادت‌فر. (۱۳۸۵) کاربرد مدل‌های شبکه عصبی در پیش‌بینی ورشکستگی اقتصادی شرکت‌های بازار بورس. *جستارهای اقتصادی*، دوره ۳، ش ۶، صص ۴۳-۱۱.
- ۲۱- گلریز، حسن. (۱۳۷۴) *بورس اوراق بهادار*. تهران: انتشارات امیرکبیر.
- ۲۲- مشکانی، علی و عبدالرضا ناظمی (۱۳۸۸). *مقدمه‌ای بر داده کاوی*. مشهد: موسسه چاپ و انتشارات دانشگاه فردوسی مشهد.
- ۲۳- مهرانی، ساسان، مهرانی، کاوه، منصفی، یاشار و غلامرضا کرمی. (۱۳۸۴) بررسی کاربردی الگوهای پیش‌بینی ورشکستگی زیمنسکی و شیراتا در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. *بررسی‌های حسابداری و حسابرسی*، دوره ۱۲، ش ۴۱، صص ۱۳۱-۱۰۵.
- ۲۴- مکیان، نظام‌الدین، سلیم کریمی تکلو. (۱۳۸۹) پیش‌بینی ورشکستگی شرکت‌های تولیدی با استفاده از شبکه‌های عصبی مصنوعی. *فصلنامه اقتصاد مقداری*، دوره ۶، ش ۱، ۱۴۴-۱۲۹.
- ۲۵- نیکبخت، محمدرضا و مریم شریفی. (۱۳۸۹) پیش‌بینی ورشکستگی مالی شرکت‌های بورس اوراق بهادار تهران با استفاده از شبکه‌های عصبی مصنوعی. *مجله مدیریت صنعتی*، دوره ۲، ش ۴، صص ۱۸۰-۱۶۳.
- ۲۶- ناصرزاده، هوشنگ. (۱۳۷۴) *قانون تجارت*. تهران: نشر دیدار.
- ۲۷- نبوی چاشمی، سیدعلی، احمدی، موسی و صادق مهدوی فرح‌آبادی. (۱۳۸۹) پیش‌بینی ورشکستگی شرکت‌ها با استفاده از مدل لاجیت.
- مجله مهندسی مالی و مدیریت پرتفوی، ش ۵، صص ۵۵-۸۱.
- ۲۸- یاری‌فرد، رضا. (۱۳۸۲) *بررسی مدل‌های پیش‌بینی ورشکستگی (مدل‌های فولمر و اسپرینگ)* در بورس اوراق بهادار تهران. پایان‌نامه کارشناسی ارشد حسابداری، دانشگاه مازندران، مازندران.
- 29- Altman, E.I. (1968). Financial ratios, disarmament analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23 (4): 589-609.
- 30- Alforo, E, Garcia, N, Gámez, M, Elizondo, D. (2008). Bankruptcy forecasting: An empirical comparison of adaboost and neural networks. *Decision Support Systems*, 45(1): 110-122.
- 31- Beaver, W.H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, Empirical Research in Accounting: Selected studies, 4: 71-111.
- 32- Cielen, A.P. (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research*, 154: 526-532.
- 33- Deakin, E.B. (1972). "A discriminant analysis of predictors of Business failure". *Journal of Accounting Research*, 10: 167-179.
- 34- Eksi, I. (2011). Classification of firm failure with classification and regression trees. *Journal of Finance and Economics*, 76: 113-120.
- 35- Gepp, A, Kumar, K, Bhattacharya, S. (2010). Business failure prediction using decision tree. *Journal of Forecasting*, 29: 536-555.
- 36- Kiers H.A.L, Rasson J.P, Groenen P.J.F, Schader M. (2000). *Data analysis classification and related methods*. Namur, Belgium: International Federation of Classification Societies (IFCS), 428p.
- 37- Jones, F.L. (1978). Current techniques in bankruptcy prediction. *Journal of Accounting Literature*, 6: 131-164.
- 38- Janani GH, Nabavi Chashmi A, Faghani Makrani KH. (2013). A study on capability of financial ratios in predicting bankruptcy of firms: Evidence from Tehran Stock Exchange. *Management Science Letters*. 3: 2119-2124.

- 47- Serrano- Cinaco C. (1997). Feed forward neural networks in the classification of Financial information. *The European Journal of Financial*, 3: 183-202.
- 48- Shah J.R, Murtza M.B. (2000). A network based clustering procedure for bankruptcy prediction. *American Business Review*, 18 (20): 80- 86.
- 49- Shin K, Lee Y. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Application*, 23 (3): 321- 328.
- 50- Sueyoshi T, Goto M. (2009). Methodological comparison between DEA (data envelopment analysis) and DEA- DA (discriminant analysis) form the perspective of bankruptcy assessment. *European Journal of Operational Research*, 199: 561-575.
- 51- Sarkar S, Sriram R.S. (2001). Bayesian Models for Early Warning of Bank Failures. *Management Science* , 47(11): 1457-1475.
- 52- Tam K.Y, Kiang M. (1992). Managerial Applications of Neural Networks: the Case of Bank Failure Predictions. *Management Science*, 38 (7): 926-947.
- 53- Wallace A. (2004). Risk Assessment by Internal Auditors Using Past Research on Bankruptcy Applying Bankruptcy Models. *J. Account. Res*, 56(16): 109-131.
- 54- Zhang G.H, Hu M, Eddy Patuwo B.C, Indro D. (1999). Artificial neural network in bankruptcy prediction: General framework and cross- validation analysis *European Journal of Operational Research*, 116 (1): 16-32.
- 39- Lee, W.C. (2004). An empirical comparison of bankruptcy models - evidence from Taiwan. *Sin Pac Financial Journal*, 149-168. Available: http://ir.lib.au.edu.tw/bitstream/987654321/812/1/CM05-pa_lee_01-0001-u
- 40- Min H, Jae & Lee C,Young. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28: 603- 614.
- 41- Mousavi Shiri M, Ahangary M, Vaghfi H, Kholousi A. (2012). Corporate bankruptcy prediction using data mining techniques: evidence from iran. *African Journal of Scientific Research*. 8 (1): 403-426.
- 42- Ohlson J.A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18 (1): 109-131.
- 43- Odom M.D, Sharda R. (1990). A neural network model for bankruptcy prediction. *IJCNN International Joint Conference on Neural Networks*, 2: 163-168.
- 44- Premachandra 1M, Chen Y,Watson J. (2011). Dea as a tool for predicting corporate failure and success: A case of bankruptcy assessment. *Omega*, 39(6): 620-626.
- 45- Rygielski C, Wang J.C, Yen D.C. (2002). Data mining techniques for customer relationship management. *Technology in Society*, 24(4): 483-502.
- 46- Ravi Kumar P, Ravi V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques– A review. *European Journal of Operational Research*, 180 (1): 1-28

Bankruptcy Prediction of Companies listed Corporations in Tehran Stock Exchange by Using Decision Tree and Logistic Regression

S. M. Hosseini

Associate Professor of Biological Statistics, University of Isfahan, Iran

Z. Rashidi*

MA student of Public- financial Management at Azad University of Khorasgan, Iran

Abstracts:

The main goal of this study is the prediction of companies listed in Tehran Stock Exchange using decision tree and logistic regression, which are data mining methods and can help facilitate prediction. Financial ratios are independent variables and healthy and bankrupt companies are dependent variables. Statistical population of the research is information about financial statements of companies listed in Tehran Stock Exchange during 1999 – 2010. No sampling was used in this study; we have two groups of healthy and bankrupt companies. Bankrupt companies group was selected based on article 141 of Commercial law and the healthy group was selected based on profitability criteria. Results suggest that both decision tree and logistic regression methods predict bankruptcy with different accuracy. Area under ROC curve in logistic regression model is more than in decision tree model.

Keywords: prediction bankruptcy, data mining, decision tree, cart, logistic regression

* Z_rashidi282@yahoo.com