فصلنامه اقتصاد مقداری (بررسی‌های اقتصادی سابق)، دوره 7، شماره 4، زمستان 1389، صفحات 172- 151
Archive of SID

کارایی شبکه‌های عصبی، رگرسیون لجستیک و تحلیل تمایزی در پیش بینی نکول

دکتر علی رحمانی و غریبه اسماعیلی *

تاریخ وصول: 89/7/19 تاریخ پذیرش: 89/11/4

چکیده:

مدل‌های آماری مختلفی برای پیش بینی و طبقه بندی در علوم وجود دارد. روش‌های آماری و اقتصادسنجی نظیر رگرسیون، تحلیل تمایزی، سری‌های زمانی، رده بندی و دیگر روش‌ها، بر اساس متغیرها و اطلاعات موجود برای پیش بینی و طبقه بندی یک موضوع خاص به کار می‌روند. مدل‌های آماری متأثر از مفروضات و محدودیت‌های زیادی هستند، بدین لحاظ اخیرا شبکه‌های عصبی به عنوان شیوه‌ی نوین پیش بینی به دلیل عدم نیاز به فروض و محدودیت‌ها در توزیع داده‌ها وکارایی بالاتر آن مورد توجه ویژه قرار گرفته است. هدف از این مقاله مقایسه‌ی توانایی مدل‌های شبکه‌ی عصبی مصنوعی با رگرسیون لجستیک و تحلیل تمایزی برای پیش بینی ریسک نکول است. با استفاده از اطلاعات 23801 قرارداد لیزینگ و انتخاب متغیرهای مدت قرارداد، مبلغ قرارداد، نوع صنعت، نوع قرارداد، نوع تضمین و خط مشی و سیاست اعتباری به عنوان متغیرهای پیش بین مدل‌های رگرسیون لجستیک، تحلیل تمایزی و شبکه‌ی عصبی برازش شد. از تحلیل راک و مقایسه‌ی صحت طبقه بندی برای مقایسه‌ی قدرت پیش بینی مدل‌ها استفاده شد. نتایج حاکی از معنی دار بودن متغیرهای فوق در پیش بینی نکول بوده و مقایسه‌ی قدرت پیش بینی مدل‌ها موید برتری شبکه‌های عصبی نسبت به رگرسیون لجستیک و تحلیل تمایزی است.

طبقه‌بندی JEL: **C53 ،C45 ،C40**

واژه‌های کلیدی: *شبکه‌های عصبی، رگرسیون لجستیک، تحلیل تمایزی، ریسک نکول، لیزینگ*

* به ترتیب، استادیار دانشکده علوم اجتماعی واقتصاد دانشگاه الزهراء، دانشجوی کارشناسی ارشد دانشگاه الزهراء (rahmani@alzahra.ac.ir)

www.SID.ir

Archive of SID

1- مقدمه

یکی از مهمترین کاربردهای روش‌های آماری، شناسایی الگو، طبقه بندی و پیش بینی بر اساس واقعیت‌ها و اطلاعات در دسترس از یک موضوع خاص است. روش‌های آماری که برای مدل سازی روابط بین متغیرها به کار می‌رود دارای تعدادی پیش فرض و محدودیت است. این پیش فرض‌ها شامل نرمال بودن توزیع داده‌ها، عدم هم خطی، یکسان بودن واریانس خطاها و موارد دیگری است و در صورت فقدان این شرایط به کارگیری این مدل‌ها امکان پذیر نبوده و یا با خطای قابل توجهی همراه است. از این رو به روش‌هایی با محدودیت‌ها و فروض کمتر، نیاز می‌باشد و یکی از این روش‌ها شبکه‌های عصبی مصنوعی است. این روش هیچ فرض اولیه‌ای برای توزیع داده‌ها تحمیل نمی‌نماید. در این شبکه‌ها اطلاعات به صورت ضمنی پردازش می‌شود، بر این اساس چنانچه بخشی از سلول‌های شبکه حذف شوند یا عملکرد غلط داشته باشند باز هم احتمال رسیدن به پاسخ صحیح وجود دارد.

یکی از پدیده‌هایی که در دنیای کسب و کار پیش بینی آن اهمیت دارد، ریسک می‌باشد که برای مدلسازی آن تلاش زیادی انجام شده است. (قنبری و دیگران، 1388). ریسک به خصوص برای صنعت لیزینگ و خدمات مالی بسیار مهم است. با توجه به اهمیت لیزینگ در تأمین مالی سرمایه گذاری‌ها، پیش بینی و مستند نمودن ریسک‌ها در این صنعت برای ادامه‌ی حیات و بقای آن امری مسلم و ضروری است. ریسک اعتباری و ریسک نکول یکی از با اهمیت‌ترین ریسک‌ها در این صنعت است. شیوه‌های گوناگونی برای پیش بینی این ریسک استفاده شده است. در این مقاله از شبکه‌های عصبی، رگرسیون لجستیک و تحلیل تمایزی برای پیش بینی این نوع ریسک استفاده شده است و کارایی این سه شیوه در پیش بینی دقیق‌تر نکول ارزیابی شده است.

2- پیشینه‌ی پژوهش

قبل از دهه‌ی 1960 پیش بینی ورشکستگی شرکت‌ها با استفاده از روش‌های آماری و اقتصادی بیشتر از اینکه علم باشد هنر بود. پیشرفت‌ها در روش‌های آماری و اقتصادسنجی فرصت‌های جدیدی را برای پژوهشگران و سرمایه گذاران فراهم

www.SID.ir

Archive of SID

کرد. آلتمن[1] (1968) برای اولین بار از تحلیل تمایزی برای تعیین تابع پیش بینی ورشکستگی شرکت‌ها استفاده کرد و رابطه‌ی بین سوابق نکول و تعدادی از متغیرهای مالی و حسابداری (عمدتاً ترازنامه و سود و زیان) را با استفاده از مدل چند متغیره بررسی نمود. هدف وی شناسایی عوامل مختلف تعیین کننده‌ی نکول بود تا بتواند شرکت‌های در حال نکول را از سایر شرکت‌ها تفکیک نماید. انتقاداتی که از این مدل شد باعث گردید وی اصلاحاتی را در مدل انجام دهد. مدل تعدیل شده در کشورهای گوناگون مورد آزمون قرار گرفت و مشخص شد که پیش بینی‌های این مدل در بیش از 75 درصد موارد درست است (تهرانی و همکاران، 1384، ص 47).

المر وپروفسکی[2] (1998) برای پیش بینی توانایی پرداخت وام‌ها از مدل شبکه‌ی عصبی چند لایه پرسپترون استفاده نموده‌اند. متغیرهای استفاده شده در این مدل همان متغیرهای به کار گرفته شده توسط آلتمن بوده و نتایج حاکی از این بود که قدرت پیش بینی مدل پرسپترون بیشتر از مدل‌های نمره دهی اعتباری بوده است.

دلار فتیس گرینو[3] (2001) مورد نرخ‌های بازیافت در صنعت لیزینگ را بررسی کرد و به این نتیجه رسید که در صنعت لیزینگ ریسک اعتباری پایین است.

اسمیت و استوک[4] (2003) تحقیقی را با عنوان نرخ‌های بازیافت در صنعت لیزینگ با استفاده از پایگاه داده 32259 قراردادی برای لیزوروپ (فدراسیون انجمن لیزینگ‌های اروپا) انجام داد. نتایج این تحقیق علاوه تأیید یافته‌های دلار فتیس وگرینو (2001) با دیدگاه‌های رایج در مورد لیزینگ سازگار بود، یعنی زیان ناشی از نکول قراردادها کلاً پایین بوده و لیزینگ اتومبیل نرخ‌های بازیافت بهتری را نسبت به بخش تجهیزات نشان می‌دهد.

گوکاسیان و سیمان[5] (2007) در "استراتژی‌هایی برای پیش بینی نکول در قرارداد اجاره‌ی تجهیزات" با استفاده از 250/000 قرارداد اجاره در طی دوره‌ی

[1] Altman
[2] Elmer and Borowski
[3] De Laurentis and Geriano
[4] Schmit and Stuyck
[5] Goukasian and Seaman

www.SID.ir

Archive of SID

زمانی 2002 تا 2005 و به کارگیری سه روش رگرسیون لجستیک، تحلیل تمایزی و شبکه‌ی عصبی به نتایج زیر رسیدند: درجه‌ی رتبه بندی ترکیبی[6] پی نت (یک سیستم رتبه بندی اعتباری)، متغیرهای جمعیت شناسی سنتی،[7] عقود اجاره‌ی قبلی شرکت[8] و سابقه‌ی استقراض[9] پیش بینی کننده‌های برجسته‌ی ریسک اعتباری در هر سه مدل طبقه بندی یاد شده بوده‌اند. نتیجه بیانگر این بود که بر خلاف انتظارات، تحلیل تمایزی پیش بینی دقیق‌تری نسبت به رگرسیون لجستیک و شبکه‌های عصبی ارائه داده است.

مدرس و ذکاوت (1382) برای طراحی مدل‌های ریسک اعتباری برای بانک توسعه‌ی صادرات ایران یک نمونه 120 تایی از مشتریان حقوقی را انتخاب و از روش‌های تحلیل تمایزی و رگرسیون لجستیک استفاده کردند. متغیرهای نسبت جاری، نسبت بدهی به مجموع دارایی‌ها، نسبت حقوق صاحبان سهام به مجموع دارایی‌ها، نسبت سود قبل از کسر مالیات به حقوق صاحبان سرمایه، نسبت سود قبل از کسر مالیات به خالص فروش به عنوان پیش بینی کننده‌های نکول در مدل‌های تحلیل تمایزی و رگرسیون لجستیک استفاده شد. نتایج نشان داد که بر اساس متغیرهای مالی می‌توان مشتریان حقوقی بانک توسعه را از نظر ریسک اعتباری به خوش حساب و بد حساب دسته بندی کرد. از بین متغیرهای مالی استفاده شده در این تحقیق متغیر نسبت جاری بیشترین سهم را در تفکیک مشتریان داشت.

فروش و سپهری[10] (۱۳۸۴) در مقاله‌ی "مقایسه‌ی روش‌های خطی سنتی با تکنیک شبکه‌های عصبی مصنوعی در پیش بینی قیمت سهام بورس تهران" مدل‌های قیمت گذاری دارایی‌های سرمایه‌ای (*CAPM*) و سه عاملی فاما و فرنچ به عنوان مدل‌های خطی را با *CAPM* شبکه‌های عصبی مصنوعی به عنوان مدل‌های غیرخطی مقایسه کردند. نتایج آماری نشان داد که مدل *CAPM* دقت بهتری از مدل سه عاملی در بورس اوراق بهادار تهران دارد. همچنین نتایج نشان داد

[6] Composite paynet Rating Score
[7] Conventional demographic variables
[8] Companys prior leasing
[9] Borrowing history
[10] Forosh and Sepehri

www.SID.ir

Archive of SID

شبکه‌های عصبی مصنوعی دقت بیشتری نسبت به هر دو مدل دیگر در پیش بینی قیمت سهام دارد.

تهرانی و فلاح شمس[11] (1385) در "طراحی وتبیین مدل ریسک اعتباری در نظام بانکی کشور" با استفاده از داده‌های اعتباری 316 مشتری حقوقی بانک‌های کشور و با استفاده از مدل‌های احتمال خطی، لجستیک و شبکه‌های عصبی مصنوعی اقدام به طراحی و آزمون کارایی مدل ریسک اعتباری نموده است. نتایج حاکی از این بود که ارتباط بین متغیرها در مدل پیش بینی ریسک اعتباری به صورت خطی نبوده و تابع‌های نمایی و سیگموئید مناسب‌ترین مدل‌های پیش بینی ریسک اعتباری است و بیشترین کارایی برای پیش بینی ریسک اعتباری به ترتیب مربوط به شبکه‌های عصبی مصنوعی و مدل لجستیک می‌باشد.

کمیجانی و سعادت فر[12] (1385) در کاربرد مدل‌های شبکه‌ی عصبی در پیش بینی ورشکستگی اقتصادی شرکت‌های بورس اوراق بهادار تهران به این نتیجه رسیدند که ساختار اصلی پرسپترون سه و چهار لایه برای پیش بینی ورشکستگی شرکت‌ها به مدل‌هایی شبیه یکدیگر منتهی می‌شود که در این میان شبکه‌ی سه لایه از قدرت پیش بینی بیشتری نسبت به شبکه‌ی چهار لایه برخوردار است. همچنین به کارگیری مدل‌های مبتنی بر شبکه‌ی عصبی توانایی مدیریت را برای مقابله با نوسان‌های اقتصادی و ورشکستگی نسبت به مدل‌های رقیب افزایش می‌دهد. پیش بینی ورشکستگی اقتصادی شرکت‌های بورس در سال‌های 1385 و 1386 و ترسیم روند ورشکستگی این شرکت‌ها در دوره‌ی 86-1369 از دیگر موضوعات این مقاله بوده است. آنها نتیجه گرفتند که در سال 1385 تحت تأثیر سیاست‌های شفاف سازی روند ورشکستگی اقتصادی شرکت‌ها به طور چشمگیری افزایش خواهد یافت که با سازگار شدن شرکت‌ها با شرایط جدید، تا حدی این روند در سال 1386 تعدیل می‌شود.

انصاری[13] (1388) برای اعتبارسنجی مشتریان بانک پارسیان از شبکه‌ی عصبی استفاده نمود. نتایج بیانگر این بوده که متغیرهای نسبت دارایی جاری به کل دارایی، سود قبل از کسر مالیات به کل بدهی، فروش به کل بدهی، سود خالص

[11] Tehrani and Falah Shams
[12] Komijani and Saadatfar
[13] Ansari

www.SID.ir

Archive of SID

به حقوق صاحبان سهام وکل بدهی به حقوق صاحبان سهام اهمیت بالاتری داشته و وزن‌های بالاتری را به خود اختصاص داده است. شبکه‌ی عصبی 5 متغیره در داده‌های آموزش کارایی بالاتری نسبت به رگرسیون لوجیت داشت اما در داده‌های آزمایش لاجیت کارایی بیشتری نسبت به شبکه‌ی عصبی داشت. شبکه‌ی عصبی 17 متغیره بر رگرسیون لوجیت ارجحیت داشت وکارایی آن در هر دسته از داده‌ها بالاتر بود.

3- پرسش و فرضیه‌ی پژوهش

در این پژوهش در پی یافتن پاسخی به پرسش زیر هستیم:

آیا می‌توان بر اساس شبکه‌ی عصبی، رگرسیون لجستیک و تحلیل تمایزی مدل معنی داری برای تعیین ریسک نکول قراردادهای لیزینگ طراحی نمود؟

منظور از نکول قرارداد در سوال فوق این است که حداقل 3 قسط قرارداد پرداخت نشده و معوق باشد. این تعریف کمیته بال (2002) از نکول و تعریف بانک مرکزی جمهوری اسلامی (1388) از مطالبات سررسید گذشته است.

بر اساس پرسش بالا، فرضیه‌ی این پژوهش به شرح زیر در نظر گرفته شده است:

بر اساس شبکه‌های عصبی، رگرسیون لجستیک و تحلیل تمایزی می‌توان مدل معنی داری برای تعیین ریسک نکول قراردادهای لیزینگ طراحی نمود.

4- روش پژوهش

این پژوهش از لحاظ روش یک پژوهش از نوع همبستگی است. در برازش رگرسیون لجستیک از شیوه‌ی گام به گام پیشرو[14] استفاده شد و معنی داری ضرایب رگرسیون از طریق تابع درستنمایی و آزمون والد سنجیده شد. تابع درستنمایی یک روش برای آزمون معنی‌داری ضرایب مدل رگرسیون است. مقایسه‌ی مقدار مشاهده شده با مقادیر پیش بینی شده با استفاده از تابع درستنمایی صورت می‌گیرد و آزمون والد یکی از راه‌هایی است که معمولاً برای آزمون معنی داری ضرایب در رگرسیون لجستیک به کار برده می‌شود.

[14] Forward

www.SID.ir

Archive of SID

در برازش تحلیل تمایزی از روش لاندای ویلکس استفاده شده است. آماره‌ی لاندای ویلکس برای بررسی قدرت همبستگی درونی بین متغیرهای مورد استفاده در تحلیل تمایزی به کار رفته است.

برای مدلسازی شبکه‌ی عصبی از پرسپترون چند لایه‌ای با یک لایه‌ی پنهان استفاده شده است. با توجه به دو گانه بودن متغیر پاسخ، یک ملاک مناسب برای سنجش کیفیت مدل‌های برازش شده و تعیین توان پیش بینی آنها استفاده از مساحت زیر منحنی مشخصه عملکرد (ROC) است. این منحنی عبارت است از حساسیت برحسب یک منهای ویژگی و مساحت زیر آن عددی بین 0 و1 است و به عنوان ملاکی برای سنجش توانایی پیش بینی مدل‌ها مورد استفاده قرار می‌گیرد. هرچه این عدد نزدیک به یک باشد نشان دهنده‌ی توانایی بیشتر مدل پیش بینی است. برای برازش مدل‌ها از نرم افزار *spss* نسخه 18 استفاده شده است.

با توجه به اینکه هدف اصلی در این پژوهش مقایسه‌ی دقت مدل‌های مختلف در پیش بینی ریسک نکول است، با استفاده از معیارهای آماری مناسب مقایسه بین مدل‌ها انجام می‌شود تا بتوان یکی از آنها را به عنوان مدل بهینه انتخاب و استفاده نمود.

4-1- رگرسیون لجستیک

رگرسیون لجستیک نوعی از رگرسیون است که متغیرهای پیش بین (مستقل) می‌تواند هم در مقیاس کمی و هم در مقیاس مقوله‌ای باشد، ولی متغیر وابسته مقوله‌ای دو سطحی است. این دو مقوله به گونه‌ای معمول به عضویت یا عدم عضویت در یک گروه اشاره دارد.

این مدل رگرسیون، شبیه رگرسیون معمولی است با این تفاوت که روش تخمین ضرایب یکسان نمی باشد و به جای حداقل کردن مجذور خطاها (که در رگرسیون معمولی انجام می‌شود)، احتمالی که یک واقعه رخ می‌دهد را حداکثر می‌نماید. در رگرسیون لجستیک از مفهومی به نام بخت برای مقدار متغیر وابسته استفاده می‌شود در اصطلاح آماری بخت به معنی احتمال رخداد یک پیشامد (p_i) بر احتمال عدم رخداد ($1-p_i$) آن می‌باشد. احتمال بین 0 و 1تغییر می‌کند. در حالی که بخت ممکن است بیش از یک باشد. واژه‌ی کلیدی در تحلیل رگرسیون

www.SID.ir

Archive of SID

لجستیک سازه‌ای به نام لوجیت[15] است که لگاریتم طبیعی بخت می‌باشد. رگرسیون لجستیک به صورت زیر تعریف می‌شود:

$$Zi = In\left(\frac{Pi}{1-pi}\right) = b\circ + \sum biXi$$

در معادله‌ی پیشگفته In بیانگر لگاریتم طبیعی است. در مدل رگرسیون لجستیک احتمال رخداد پیشامد مورد نظر (عدم بازپرداخت اقساط قرارداد لیزینگ از سوی مشتری) بر اساس رابطه‌ی زیر محاسبه می‌شود:

$$Pi = \prod i(x1, x2, \ldots, xk) = e^{b\circ} + \frac{\sum bixi}{1+e^{b\circ+\sum bixi}}$$

یکی از منافع رگرسیون لجستیک بی‌نیازی آن به مفروضات محدود کننده‌ی آماری در رابطه با متغیرهاست. مسأله‌ی اصلی که در این مدل وجود دارد این است که تغییرات یکسان وضعیت اقتصادی شرکت همیشه تغییرات یکسانی را در احتمال به دست آمده به دنبال ندارد. این موضوع زمانی که احتمال نزدیک مقادیر 1 و یا 0/5 باشد اهمیت بیشتری پیدا می‌کند (تهرانی و همکاران، 1384، ص 49).

4-2- تحلیل تمایزی

تحلیل تمایزی تکنیکی آماری است که در مواردی که دو یا چند گروه شناخته شده وجود داشته باشد و روشی است که هم به رگرسیون و هم به واریانس بسیار نزدیک است. درجه‌ی تمایزی به صورت ترکیب خطی از متغیرهای پیش بینی محاسبه می‌شود. تحلیل تمایزی با به کارگیری برخی از ابزارهای طبقه بندی، برای حداقل سازی واریانس درون هر گروه و حداکثر کردن واریانس در گروه‌های مختلف عمل می‌نماید (گوکاسیان[16] و همکاران، 2007، ص 9).

تحلیل تمایزی دو محدودیت عمده دارد: محدودیت اول ضرورت نرمال بودن توزیع متغیرهای پیش‌بینی کننده است و محدودیت دیگر ضرورت برابر بودن ماتریس واریانس/کوواریانس متغیرهای مزبور در بین دو گروه است. معمول‌ترین شکل آن تحلیل تمایزی خطی است که مدل‌های احتمال خطی را به کار می‌برد. این مدل‌ها در مواردی که گروه‌های نادری را پیش‌بینی می‌نماید خطاهای عدم

[15] Logit
[16] Goukasian

www.SID.ir

Archive of SID

طبقه بندی بالایی دارند و از این رو معمولاً نمونه‌های مساوی برای هر گروه به کار می‌رود (مدرس، ذکاوت، 1381، ص 4).

4-3- شبکه‌های عصبی

در اواخر هزاره دوم در هوش مصنوعی (*AI*)[17] پیشرفت عمده‌ای صورت گرفت. در حقیقت کاربرد هر مدل پیش بینی به عنوان بخشی از فرایند تصمیم گیری می‌تواند به صورت *AI* تفسیر شود.

اما یک تکنیک وجود دارد که *AI* را در بطن خود دارد، این تکنیک شبکه‌های عصبی (*NN*) هستند که به عنوان شبکه‌ای از اجزای محاسباتی که می‌توانند به خروجی‌ها پاسخ دهند و یاد بگیرند که با محیط سازگار شوند توصیف می‌شوند.

نیورال تی (2002) اظهار می‌دارد: برای ایجاد *NN* چندین روش وجود دارد، که برای اهداف امتیازدهی اعتباری، پرسپترون چند لایه‌ای بهترین روش[18] است، زیرا که این شیوه دارای هر دو مزیت غیر خطی بودن و در نظر گرفتن اثرات متقابل متغیرها است. همچنین این روش به عنوان "طبقه بند جامع[19]"شناخته می‌شود به این دلیل که از لحاظ تئوریکی قادر به طبقه بندی هر فرآیند تصمیم گیری است. سایر روش‌ها در ایجاد *NN* عبارتند از تابع اصلی محوری (*RBF*)،[20] نقشه‌های خود سازمان (*SOM*)[21] و شبکه‌های کوهنی[22] (اندرسون،[23] 2007، ص 174).

روش شبکه‌ی عصبی همانند روش تجزیه و تحلیل تمایزی غیر خطی است که فرض غیر خطی بودن و استقلال متغیرهای توضیحی حذف شده است و در آن روابط پنهان بین متغیرهای توضیحی به عنوان یک متغیر اضافی وارد تابع می‌شود. شبکه‌ی عصبی از کارهای بیولوژیکی مربوط به مغز و سیستم نرون‌های انسان تقلید می‌نماید. شبکه‌ی عصبی همانند مغز انسان از گره‌هایی تشکیل شده است که عمل پردازش عناصر را انجام می‌دهند، همانند نرون‌های واقعی این گره‌ها از طریق

---

[17] Artificial intelligence
[18] MLP or back propagation
[19] Universal classifier
[20] Radial basis function
[21] Self-organising maps
[22] Kohnen network
[23] Anderson

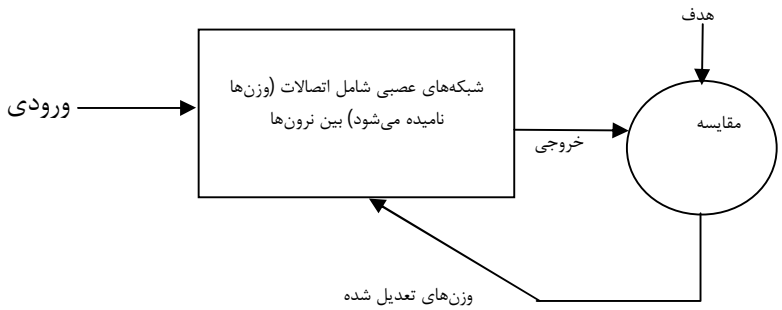
www.SID.ir

Archive of SID

اتصالات داخلی موزونی به همدیگر مرتبط شده‌اند. هر گره علایم ورودی را دریافت و به خروجی تبدیل می‌نماید که اگر مفهوم و علامت قانع کننده‌ای نباشد مجدداً به عنوان یک علامت ورودی به گره‌های دیگر وارد و ارسال می‌شود (ابراهیم لو 1384).

در طبیعت ساختار شبکه‌های عصبی از طریق نحوه‌ی اتصال بین اجزا تعیین می‌شود. بنابراین می‌توان یک ساختار مصنوعی به تبعیت از شبکه‌های طبیعی ساخت و با تنظیم مقادیر هر اتصال تحت عنوان وزن اتصال نحوه‌ی ارتباط بین اجزای آن را تعیین نمود. پس از تنظیم یا همان آموزش شبکه‌ی عصبی اعمال یک ورودی خاص به آن منجر به دریافت پاسخ خاص می‌شود. همان طور که در شکل زیر مشاهده می‌شود شبکه بر مبنای تطابق و همسنجی بین ورودی و هدف سازگار می‌شود تا اینکه خروجی شبکه و هدف بر هم منطبق گردند. عموماً تعداد زیادی از این زوج‌های ورودی و خروجی به کار گرفته می‌شوند تا در این روند که از آن تحت عنوان یادگیری نظارت شده یاد می‌شود شبکه آموزش داده شود.

شکل 1: نحوه‌ی عمل شبکه‌های عصبی در پردازش داده‌ها



مأخذ: کیا، ص 29

آموزش دسته‌ای شبکه به معنای تغییر وزن‌ها یا بایاس‌های شبکه بر مبنای مجموعه‌ی همه‌ی بردارهای ورودی می‌باشد. در حالی که آموزش گام به گام وزن‌ها و بایاس‌های شبکه را پس از اعمال تک تک بردارهای ورودی تنظیم می‌کند،

www.SID.ir

Archive of SID

از آموزش گام به گام گاهی تحت عنوان آموزش "برخط[24]" یا انطباقی یاد مـی‌شـود (کیا، 1389، ص 29).

5- جامعه و نمونه‌ی آماری

داده‌های مورد استفاده در این پژوهش شامل اطلاعات قراردادهای لیزینگ منعقد شده در طی سال‌های 82 الی 87 است که از سه شرکت لیزینگ بزرگ که سهم عمده‌ای را در بازار لیزینگ ایران بر عهده دارند، اخذ شده است. کلیه‌ی قراردادهای مذکور مورد بررسی قرار گرفته است و لذا نمونه‌ی تحقیق شامل 23801 فقره قرارداد لیزینگ است.

6-یافته‌های پژوهش

قبل از ارائه‌ی نتایج حاصل از برازش مدل، معنی دار بودن متغیرها بر اساس آزمون t برای متغیرهای کمی و آزمون من ویتنی[25] برای متغیرهای رده‌ای بررسی شده است. نتایج آزمون‌ها به شرح زیر است:

جدول 1: مقایسه‌ی مقادیر میانگین متغیرهای کمی در قراردادهای نکول شده و نشده

	آماره t	درجه‌ی آزادی	سطح معنی داری	فاصله اطمینان 95%	
مدت قرارداد	-6/647	12840/744	0/000	40276405	4339022
مبلغ قرارداد	-0/564	19039	0/573	16/792	25/000/000

بر اساس آزمون فوق متغیر مدت قرارداد معنی دار ولی متغیر مبلغ قرارداد با توجه به اینکه سطح معنی داری آن بیشتر از 5% است معنی دار نیست.

جدول 2: مقایسه‌ی مقادیر میانگین متغیرهای رده‌ای در قراردادهای نکول شده و نشده

	نوع صنعت	نوع قرارداد	نوع تضمین
من وایتنی	3721/3094	40276405	4339022
Z	-22/668	-16/792	-0/733
سطح معنی داری	0/000	0/000	0/464

[24] Online
[25] Man-whitini

www.SID.ir

Archive of SID

بر اساس آزمون من وایتنی صورت گرفته که نتایج آن در جدول 2 فوق آمده متغیرهای نوع قرارداد و نوع صنعت معنی دارند، اما متغیر نوع تضمین در سطح معنی داری 5% متغیر معنی داری نیست.

6-1- رگرسیون لجستیک

در این مرحله از کل قراردادها به میزان ۲۰% به طور تصادفی یک نمونه خارج نموده که تعداد آن ۴۶۷۲ قرارداد می‌باشد و براساس ۸۰% قرارداد باقیمانده‌ی مدل رگرسیون لجستیک به روش گام به گام پیشرو برازش گردید:

جدول 3: خلاصه‌ی نتایج برازش مدل رگرسیون لجستیک

متغیر	ضرایب (β_i)	انحراف معیار (SE)	آماره والد ($wald$)	سطح معنی داری
مقدار ثابت	-2/917	0/104	788/371	0/000
مدت قرارداد	0/006	0/000	142/159	0/001
نوع صنعت	0/053	0/016	11/044	0/000
نوع قرارداد	0/508	0/033	242/762	0/000
نوع تضمین	0/143	0/008	298/723	0/000

مدل لجستیک طراحی شده برای پیش بینی ریسک نکول براساس ضرایب به صورت زیر می‌باشد:

$$Zi = In\left(\frac{Pi}{1-Pi}\right)$$

$$= -2/917 + 0/006X1 + 0/053X2 + 0/508X3$$

آماره‌ی والد محاسبه شده برای هر یک از متغیرها و سطح خطای محاسبه شده بیانگر معنی داری ضرایب در مدل است.

جدول 4: شاخص‌های نیکویی برازش مدل رگرسیون

شاخص	میزان
کای دو	663/460
مقدار تابع درستنمایی	24995/519
Cox & snell R-squared	0/034
Negelkerk R-squared	0/046

www.SID.ir

Archive of SID

آزمون هوسمر لیموشاو هم در سطح خطای ۵% معنی دار است. مقدار درستنمایی برای این مدل برابراست با ۲۴۹۸۱/۶۷۳.

6-2- تحلیل تمایزی

تحلیل تمایزی نیز با 80 درصد از داده‌ها برازش شده است و20 درصد باقیمانده از داده‌ها را به عنوان داده خارج از برازش برای ارزیابی و آزمون بررسی اعتبار الگوهای برازش شده به کار می‌بریم. نتایج آزمون برابری میانگین در جدول 5 ارائه شده است:

جدول 5: آزمون F برابری میانگین در برازش تحلیل تمایزی

	لاندای ویلکس	F	Df_1	Df_2	سطح معنی داری
مدت قرارداد	0/997	47/921	۱	19039	0/000
نوع صنعت	0/981	372/374	۱	19039	0/000
نوع قرارداد	0/984	313/289	۱	19039	0/000
نوع تضمین	0/998	32/466	۱	19039	0/000

جدول فوق حاکی از معنی داری کلیه‌ی متغیرها در سطح 5% است.

همبستگی کانونی در این برازش برابر است با ۰/۱۸۶ ضرایب تابع تمایزی به شرح جدول زیر است:

جدول 6: ضرایب تمایزی مدل برازش شده

	تابع
	1
مدت قرارداد	-0/10
نوع صنعت	0/917
نوع قرارداد	1/065
نوع تضمین	0/058
مقدار ثابت	-3/868

در نتیجه بر اساس جدول فوق می‌توان تابع تمایزی را به صورت زیر بیان نمود:

$$D = -3/868 - 0/1F1 + 0/917F3 + 1/065F4 + 0/058F5$$

براساس ضریب‌های تابع ممیز استاندارد متغیر $F_۴$ (نوع قرارداد) حداکثر سهم را در تابع ممیز و در نتیجه در تفکیک دو گروه دارد. در رتبه‌های بعد به

www.SID.ir

Archive of SID

ترتیب F_3 (نوع صنعت)، (F_5 نوع تضمین) وF_1 (مدت قرارداد) قرار دارند و متغیر F_2 (مبلغ قرارداد) اصلاً در تابع ممیز قرار ندارد و کمترین اثر را داشته و به دلیل معنی دار نبودن اثر آن حذف شده است.

3-6- شبکه‌ی عصبی

در شبکه‌ی طراحی شده تعداد کل قراردادها 15201 (79/8%) به نمونه‌ی آموزش و 3840 (20/2%) به نمونه‌ی خارج از برازش تعلق دارد. اطلاعات شبکه به شرح ذیل است:

جدول 7: خلاصه‌ی اطلاعات شبکه‌ی طراحی شده

لایه‌ی ورودی	عامل	۱	نوع تضمین
	متغیرهای کمکی	1	مدت قرارداد
		2	مبلغ قرارداد
		3	نوع صنعت
		4	نوع قرارداد
		تعداد گره ها	9
		روش مقیاس بندی مجدد متغیرهای کمکی	استاندارد نمودن
لایه‌ی نهان		تعداد لایه‌ای نهان	1
		تعداد گره در لایه‌ی نهان	6
		تابع محرک	تانژانت هایپربولیک
لایه‌ی خروجی	متغیر وابسته	1	وضعیت قرارداد- نکول یا عدم نکول
	تعداد گره ها		2
	تابع محرک		سافت مکس
	تابع خطا		کراس -اینتروپی

اهمیت متغیرها با استفاده از تحلیل حساسیت در جدول زیر ارائه گردیده:

جدول 8: اهمیت هر یک از متغیرها در شبکه‌ی طراحی شده

	اهمیت	اهمیت نرمال بودن
نوع تضمین	0/308	%91/3
مدت قرارداد	0/079	%23/5
مبلغ قرارداد	0/179	%53/2
نوع صنعت	0/96	%28/4
نوع قرارداد	0/337	%100/0

www.SID.ir

Archive of SID

جدول طبقه بندی زیر نتایج عملی حاصل از به کارگیری شبکه را نشان می‌دهد. در این جدول برای هر قرارداد در صورتی که احتمال پیش بینی شده بیشتر از 0/5 باشد متغیر پاسخ در طبقه نکول شده قرار می‌گیرد. در هر نمونه سلول‌های واقع در قطر اصلی پیش بینی درست است و خارج از قطر پیش بینی نادرست است.

جدول 9: نتایج طبقه بندی قراردادها بر اساس شبکه‌ی عصبی طراحی شده

مشاهده شده		پیش بینی شده		
		نکول شده	نکول نشده	درصد صحت
آموزش	نکول شده	6669	2030	76/7%
	نکول نشده	3819	2683	41/3%
	درصد کلی	69/0%	33/0%	61/5%
خارج از برازش	نکول شده	1644	505	76/5%
	نکول نشده	975	716	42,3%
	درصد کلی	68/2%	31/8%	61/5%

از داده‌هایی که برای ایجاد مدل استفاده شده اند 6669 از 8699 قراردادی که نکول شده‌اند به درستی طبقه بندی شده‌اند و به همین ترتیب 3/ 23 درصد، نادرست طبقه بندی شده‌اند. نمونه‌ی خارج از برازش به تأیید اعتبار مدل کمک می‌نماید که در آن 61/5% از قراردادها به درستی پیش بینی شده‌اند. با توجه به اینکه درصد صحت پیش بینی در داده‌های خارج از برازش بیشتر از داده‌های آموزش است و از سویی دیگر قاعده‌ی توقف گزارش شده در خلاصه‌ی مدل این است که تغییر تقریب در معیار خطای آموزش (0/0001) به دست آمده در نتیجه وجود آموزش بیش از حد مورد انتظار نیست و از این رو نیاز به ایجاد نمونه‌ی آزمایش نیست.

6-4- مقایسه‌ی کارایی سه مدل در پیش بینی

www.SID.ir

Archive of SID

جدول 10: طبقه بندی قراردادهای مورد بررسی بر اساس پیش بینی مدل رگرسیون لجستیک، تحلیل تمایزی و شبکه‌ی عصبی

وضعیت واقعی				
مدل	پیش بینی مدل	نکول نشده	نکول شده	کارایی
رگرسیون لجستیک	نکول نشده	*(50/7)۹۸۴۱	(7/8)۱۵۱۵	٪86/2
	نکول شده	(30/5)۵۹۱۰	(۱۱)۲۱۳۵	٪26/5
	صحت کلی			٪61/7
تحلیل تمایزی	نکول نشده	3127(16/4)	(23/2)۴۴۲۴	٪41/4
	نکول شده	(15/4)۲۹۲۵	(۴۵)۸۵۶۵	٪74/5
	صحت کلی			٪61/4
شبکه عصبی	نکول نشده	(42/8)۸۱۵۳	(14/8)۲۸۲۰	٪74/3
	نکول شده	(22/2)۴۲۳۷	(20/2)۳۸۴۴	٪47/6
	صحت کلی			٪63/0

*درصد از کل قراردادها

بر اساس جدول 10 مدل رگرسیون لجستیک 61/7%، مدل تحلیل تمایزی 61/4% و مدل شبکه‌های عصبی 63% قراردادهای نکول شده و نشده را به درستی پیش بینی کرده‌اند.

منحنی راک برای مدل‌های مورد بررسی در نمودارهای 1 و 2 نشان داده شده است. بر اساس این نمودارها سطح زیر منحنی راک برای مدل‌های رگرسیون لجستیک، تحلیل تمایزی و شبکه‌ی عصبی به ترتیب 0/623، 0/595 و 0/659 است که در جدول 11 ارائه شده است.

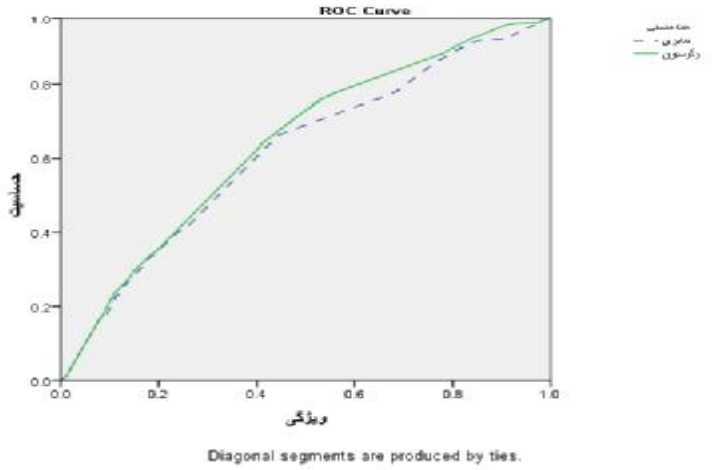
جدول 11: مساحت زیر منحنی راک در سه مدل برازش شده

مدل	مساحت زیر منحنی راک
رگرسیون لجستیک	0/623
تحلیل تمایزی	0/595
شبکه‌ی عصبی	0/659

www.SID.ir

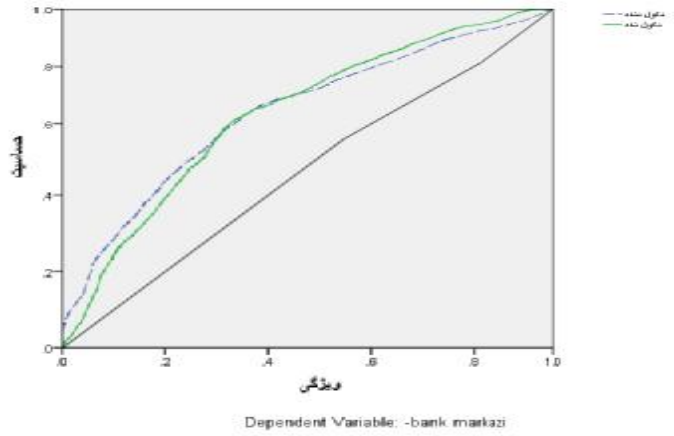
Archive of SID

شکل 2: منحنی *ROC* در مدل‌های رگرسیون لجستیک و تحلیل تمایزی برازش شده



مأخذ: نتایج پژوهش

شکل 3: منحنی *ROC* در شبکه‌ی عصبی طراحی شده برای قراردادهای نکول شده و نشده



مأخذ: نتایج پژوهش

نکته‌ی مهم دیگر مطرح در این پژوهش مقایسه وضعیت متغیر مبلغ قرارداد در سه روش می‌باشد. این متغیر در روش رگرسیون لجستیک و تحلیل تمایزی متغیر معنی داری نیست، اما در شبکه‌ی عصبی متغیری معنی دار با سطح اهمیت 0/179 و دارای رتبه‌ی سوم در بین پنج متغیر پیش بین استفاده شده است. این تفاوت را می‌توان با توجه به یکی از ویژگی‌های اصلی شبکه‌ی عصبی توجیه کرد. این ویژگی توانایی شبکه‌ی عصبی در کشف (شناسایی الگو) و ردیابی ارتباطات بین

www.SID.ir

Archive of SID

داده‌ها و به ویژه اثرات مقابل است (اندرسون، 2007، ص 175). می‌توان معنی‌دار بودن متغیر مبلغ قرارداد را با توجه به اثرات متقابل آن با سایر متغیرها در شبکه طراحی شده توجیه نمود و در واقع این ویژگی یکی از مزیت‌های شبکه‌ی عصبی نسبت به دو روش دیگر دانست.

به دلیل اینکه هر شرکت لیزینگ خط مشی اعتباری و سیاست وصول خاص خود را دارد نوع و چگونگی این خط مشی می‌تواند بر میزان نکول یا عدم نکول قراردادهای آن تأثیر گذار باشد. از آن جایی که نمونه‌ی مورد مطالعه در این تحقیق مربوط به قراردادهای سه شرکت است، به منظور بررسی تأثیر خط مشی اعتباری و سیاست وصول، مدل‌ها به تفکیک اطلاعات هر یک از شرکت‌ها برازش گردید. نتایج مربوط به صحت طبقه بندی و مساحت زیر منحنی *ROC* به منظور مقایسه‌ی قدرت پیش‌بینی مدل‌ها به تفکیک هر یک از شرکت‌ها در جداول زیر ارائه گردیده است:

جدول 12: مقایسه‌ی صحت طبقه بندی کلی سه مدل برازش شده در اطلاعات قراردادهای هر یک از سه شرکت

مدل / شرکت	رگرسیون لجستیک	تحلیل تمایزی	شبکه های عصبی
شرکت اول	%59/2	%57/1	%60/5
شرکت دوم	%83/5	%77/4	%86/9
شرکت سوم	%61/2	%61/1	%62/0

در هر سه شرکت مدل شبکه‌های عصبی عملکرد برتری را نسبت به دو مدل دیگر داشته و سپس رگرسیون لجستیک و تحلیل تمایزی در رده‌های بعدی قرار دارند. اما در مقایسه‌ی سه شرکت، صحت کلی مدل‌ها با هم متفاوت هستند. به طور کلی مدل‌ها در شرکت دوم عملکرد برتری داشته‌اند و این موید این مسأله است که نوع خط مشی شرکت‌ها در زمینه‌ی نکول مؤثر است. برای بررسی دقیق‌تر کارایی سه مدل از مساحت زیر منحنی *ROC* استفاده می‌نماییم

www.SID.ir

Archive of SID

جدول 13: مساحت زیر منحنی *ROC* در سه مدل برازش شده

مدل / شرکت	رگرسیون لجستیک	تحلیل تمایزی	شبکه های عصبی
شرکت اول	0/593	0/593	0/620
شرکت دوم	0/721	0/721	0/905
شرکت سوم	0/613	0/612	0/627

جدول فوق عملکرد برتر شبکه‌های عصبی در سه شرکت را تأیید می‌نماید و مدل‌ها در شرکت دوم عملکرد بهتری را نسبت به دو شرکت داشته است، نتایج مساحت زیر منحنی در شرکت اول و سوم نتایج مشابه و نزدیک به هم داشته است.

7- نتیجه گیری

با توجه به اهمیت شناسایی ریسک نکول و یافتن راهکارهایی برای مدیریت و کنترل آن، دسترسی به مدل‌هایی که با دقت بالا بتواند این رخداد را پیش بینی نماید، نقش مهمی به لحاظ اقتصادی و عملی در بقا و سلامت صنعت لیزینگ دارد. در این مطالعه سه مدل مختلف برای پیش بینی ریسک نکول برآورد گردید و یافته‌های حاصل مقایسه شدند. برای برازش مدل از اطلاعات 19041 فقره قرارداد و برای تأیید اعتبار مدل‌ها از اطلاعات 4670 فقره قرارداد استفاده شد. متغیرهای معنی‌دار در هر دو مدل رگرسیون لجستیک و تحلیل تمایزی شامل مدت قرارداد، نوع صنعت، نوع قرارداد و نوع تضمین بود و متغیر مبلغ قرارداد در هیچ یک از دو مدل معنی دار نبود. اما همین متغیر (مبلغ قرارداد) در شبکه‌ی عصبی معنی‌دار بود. این تفاوت را می‌توان به ویژگی مربوط به شبکه‌ی عصبی در شناخت اثرات متقابل متغیرها و شناسایی الگوهای موجود در داده‌ها مرتبط نمود و توجیه کرد.

برای بررسی کارایی مدل‌ها در تفکیک قراردادها به نکول شده و نکول نشده از جدول طبقه بندی با استفاده از یک نقطه‌ی برش، استفاده شد. علاوه بر آن از تحلیل منحنی راک استفاده شد. نتایج بیانگر این است که شبکه‌ی عصبی از رگرسیون لجستیک و تحلیل تمایزی در پیش بینی نکول کاراتر است. علاوه بر پنج متغیری که در برازش‌های فوق مدنظر قرار گرفت، خط مشی اعتباری و سیاست وصول می‌تواند نرخ نکول را در شرکت‌های لیزینگ تحت تأثیر قرار دهد. لذا مدل‌ها با داده‌های هر یک از شرکت‌ها به صورت مجزا برازش شد و در این

www.SID.ir

Archive of SID

حالت هم عملکرد برتر شبکه‌های عصبی تأیید شد. از سویی متفاوت بودن صحت طبقه بندی و مساحت زیر منحنی *ROC* در سه شرکت مؤید تأثیر گذار بودن خط مشی اعتباری سیاست وصول در نرخ نکول است.

نتایج حاصل از این تحقیق مبنی بر عملکرد برتر شبکه‌های عصبی نسبت به رگرسیون لجستیک و تحلیل تمایزی با نتایج تحقیقات قبلی شامل المر وپروفسکی (1998)، فلاح شمس (1385)، انصاری (1388) منطبق است، اما با نتایج گوکاسیان و همکاران (2007) هماهنگ نیست.

www.SID.ir

Archive of SID

فهرست منابع:

ابراهیم لو، محمد حسین. (1384). بررسی مدل ریسک اعتباری مشتریان حقوقی بانک سامان. پایان نامه کارشناسی ارشد.

امین ناصری، محمدرضا و سیدحسام الدین انوار. (1383). کاربرد شبکه عصبی در پیش بینی نرخ ارز. دومین کنفرانس بین المللی مدیریت.

انصاری، سارا. (1388). اعتبار سنجی مشتریان حقوقی بانک پارسیان با استفاده از شبکه عصبی مصنوعی. پایان نامه کارشناسی ارشد.

بانک مرکزی جمهوری اسلامی. (1388). آیین نامه وصول مطالبات سررسید گذشته، معوق و مشکوک الوصول موسسات اعتباری مصوب هیئت وزیران.

تهرانی، رضا و میرفیض فلاح شمس. (1384). طراحی و تبیین مدل ریسک اعتباری در نظام بانکی کشور. مجله علوم انسانی و اجتماعی دانشگاه شیراز، 2(43):60-45.

سپهری، محمدمهدی و حمید فروش. (1384). مقایسه روشهای خطی سنتی با تکنیک شبکه‌های عصبی مصنوعی در پیش بینی قیمت سهام بورس تهران. دومین کنفرانس بین المللی فناوری اطلاعات و دانش.

سدهی، مرتضی، یداله محرابی، انوشیروان کاظم نژاد و فرزاد حدائق. (1388). مقاسیه مدل‌های شبکه عصبی مصنوعی با رگرسیون لجستیک وتحلیل ممیزی در پیش بینی سندرم متابولیک. مجله غدد درون ریز ومتابولیسم ایران، 6 (11): 646-638.

فروش، حمید و محمدمهدی سپهری. (1384). مقایسه روش‌های خطی سنتی با تکنیک شبکه‌های عصبی مصنوعی در پیش بینی قیمت سهام بورس تهران. مجموعه مقالات دومین کنفرانس بین المللی فناوری اطلاعات و دانش: 36-24.

قنبری، علی، محسن خضری و رقیه ترکی سمایی. (1388). تخمین ریسک سیستماتیک در مقیاس‌های زمانی مختلف با استفاده از آنالیز موجک برای بررسی اوراق بهادار تهران. اقتصاد مقداری، 6(4): 50- 29.

کمیجانی، اکبر و جواد سعادت‌فر. (1385). کاربرد مدل‌های شبکه عصبی در پیش بینی ورشکستگی اقتصادی شرکت‌های بازار بورس. دو فصلنامه جستارهای اقتصادی، 6 (3): 44-11.

کیا، مصطفی. (1389). شبکه‌های عصبی در *MATLAB*. انتشارات کیان رایانه سبز، چاپ سوم.

مدرس، احمد و سید مرتضی ذکاوت. (1381). مدل‌های ریسک اعتباری مشتریان بانک (مطالعه موردی)، فصلنامه حسابرس، 19(14): 58-54.

www.SID.ir

Archive of SID

Altman. E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, 23(4): 589-609.

Anderson, R. (2007). Theory and Practice for Retail Credit Risk, Management and Decision. Automation Published in United States, Oxford University, NewYorK,.

Basel Committeeon Banking Supervision. (2002). Principles for Management of Credit Risk, BIS, Basel, Switzerland

De Laurentis G. & M. Geriano. (2001). Leasing Recovery Rates Leaseurope –Bocconi University Business School Research, 14 pages.

Elmer, P. J. & D. M. Borowski. (1988). An Expert System and Neural Networks Approach To Financial Analysis, Financial Management, 12(5): 66-76.

Goukasian, L. & S. Seaman. (2009). Sttrategies for Predicting Equipment Lease Default, Journal of Equipment Lease Financing, 27(1):1-7.

Schmit, M. & J. Stuyck. (2003). recovery rates in the leasing industry. Working paper, Leaseurope, 1-37.

Murtagh, J. P. (2005). Credit Risk: Contract Characteristics for Success, Journal of Equipment Lease Financing, 23(1): 9-14.

www.SID.ir