

مقایسه‌ی مدل‌های مختلف شبکه‌ی عصبی در رتبه‌بندی اعتباری سیستم بانکی و معرفی بهترین مدل (۱۳۸۵-۱۳۹۰)

جعفرزهی (استاد)

دانشکده‌ی مهندسی صنایع، پردیس دانشکده‌های فنی، دانشگاه تهران

محمدرضا شهبازی* (کارشناس ارشد)

دانشکده‌ی مهندسی صنایع، پردیس البرز، دانشگاه تهران

مهندسی صنایع و مدیریت شریف (زمستان ۱۳۹۵)
دوره ۱-۳۲، شماره ۱/۲، ص. ۱۲۷-۱۴۰، (یادداشت فنی)

در نگاهی ساده، حوزه‌ی فعالیت بانک‌ها تجهیز و تخصیص منابع است. لذا، بانک‌ها با در نظر گرفتن ریسک اعتباری مشتریان، به تقاضاهای آن‌ها مبنی بر اخذ تسهیلات جامه‌ی عمل می‌پوشانند. این پژوهش، با هدف انتخاب متغیرهای اثرگذار و مدل بهینه، به منظور رتبه‌بندی اعتباری مشتریان بانکی با استفاده از مدل‌های شبکه‌های عصبی با الگوریتم پس‌انتشار خطا، شبکه‌های عصبی GMDH، شبکه‌های عصبی با الگوریتم شعاع‌محور، مدل‌های لاجیت، پروبیت و تحلیل ممیزی ارائه شده است. لذا ۲۰۰ نفر از مشتریان حقیقی یکی از بانک‌های دولتی در فواصل سال‌های ۱۳۸۵-۱۳۹۰ انتخاب شده‌اند که از این تعداد ۱۰۵ نفر خوش‌حساب و ۹۵ نفر از مشتریان بدحساب بوده‌اند. در مرحله‌ی اول ۹ متغیر به عنوان متغیرهای بی‌اثر در وضعیت اعتباری مشتریان تشخیص داده شد که ۵ متغیر حذف شدند. نهایتاً مقایسه‌ی این مدل‌ها با یکدیگر نشان داد که شبکه‌های عصبی با الگوریتم شعاع‌محور و شبکه‌های عصبی GMDH بالاترین دقت را در پیش‌بینی رفتار اعتباری مشتریان دارد.

واژگان کلیدی: رتبه‌بندی اعتباری، ریسک اعتباری، شبکه‌های عصبی، الگوریتم پس‌انتشار خطا، مدل GMDH، الگوریتم شعاع‌محور، مدل لاجیت، مدل پروبیت و مدل تحلیل ممیزی.

jrazmi@ut.ac.ir
mr.shahbazi@alumni.ut.ac.ir

۱. مقدمه

عدم ثبات سیاسی و اقتصادی در جوامع امروزی از یک‌سو، و الزامات کنترل بازار پول و سرمایه از سوی دیگر بر اهمیت تجربه و تحلیل مدیریت ریسک این بازارها افزوده است. بنابراین، تمام مؤسسات انتفاعی و غیرانتفاعی به‌نوعی با این ریسک مواجه هستند و همین امر باعث دخالت‌های مستقیم قانونی از سوی قانون‌گذاران برای کنترل آن شده است. تعیین «حداقل میزان سپرده‌ی قانونی در بانک‌ها» را می‌توان نمونه‌ی بارز این دخالت از سوی بانک مرکزی برای کنترل ریسک نقدشوندگی دانست. تا جایی که افزایش تقاضای اعتبار، افزایش رقابت و به‌وجود آمدن کانال‌های جدید در فضای اقتصادی، فرصت‌های جدیدی برای مؤسسات اعتباردهنده به‌وجود آورده است. این مسئله، مؤسسات را به سمت تجدید نظر، توان‌مهندسازی و ورود فناوری‌های جدید در فرایند مدیریت اعتبار سوق داده است. مدل‌های رتبه‌بندی اعتباری، یکی از مهم‌ترین و اساسی‌ترین سیستم‌های (ابزارها) تصمیم‌گیری هستند که بخش عمده‌ی اطلاعات

مورد نیاز مؤسسات اعتباردهنده در مدیریت اعتبار را فراهم می‌کنند. هدف مدل‌های رتبه‌بندی اعتباری، پیش‌بینی احتمال عدم بازپرداخت اعتبار از سوی مشتری یا طبقه‌بندی متقاضیان اعتبار به دو گروه خوب و بد (خوش‌حساب و بدحساب) است. به عبارت دیگر، رتبه‌بندی اعتباری مجموعه‌ی بی‌اثر از مدل‌های تصمیم‌گیری و روش‌های مرتبط با آن‌هاست که به اعتباردهندگان در اعطای اعتبار به مشتریان (و تصمیم‌گیری) کمک می‌کند. از آنجا که ارزیابی و سنجش اعتبار مشتریان بر مبنای قضاوت ذهنی کارشناسان بانک صورت می‌گیرد، الگوی امتیازدهی و سنجش ریسک اعتباری مشتریان به بانک‌ها کمک خواهد کرد تا در ارزیابی مشتریان اعتباری ضوابط و معیارهای علمی را به‌کارگیرند و ضمن کاهش ریسک اعتباری و احتمال نکول تسهیلات، اعتبارات را به‌صورت بهینه تخصیص دهند و نیز بتوانند بر سایر ریسک‌ها، نظیر ریسک نقدشوندگی، غلبه کنند. همچنین صرفه‌جویی در وقت و هزینه، حذف قضاوت‌های شخصی و افزایش دقت در ارزیابی متقاضیان انواع تسهیلات، از جمله مزیت‌های این روش است. در نهایت روش‌های آماری مختلفی همچون رگرسیون لجستیک، هموارسازی ناپارامتریک و نیز روش‌هایی نظیر شبکه‌های عصبی در

* نویسنده مسئول

تاریخ دریافت: ۱۳۹۳/۶/۳۰، اصلاحیه ۱۳۹۳/۱۰/۳، پذیرش ۱۳۹۳/۱۰/۱۶.

زمینه‌ی رتبه‌بندی اعتباری مورد استفاده قرار گرفته است. از این میان «شبکه‌های عصبی» با برخورداری از قابلیت طبقه‌بندی، تعمیم و یادگیری الگوها نسبت به سایر روش‌ها از انعطاف‌پذیری بالاتری برخوردار بوده و در سال‌های اخیر مورد توجه بیشتری قرار گرفته است. در این پژوهش نیز، از مدل‌های لاجیت، پروبیت، تحلیل ممیزی چندگانه و شبکه‌های عصبی MLP، RBFN، GMDH استفاده شده تا بتوان با کاربرد روش‌ها به یک ارزیابی صحیح از ریسک اعتباری مشتریان رسید و با کاهش احتمال نکول مطالبات و تسهیلات بانکی به تخصیص بهینه دست یافت.

بانک‌های موفق، به دلایل مختلف، مانند خطر یا هزینه‌های ناشی از نوسان‌های نرخ بهره، تورم، ارز و یا عدم بازپرداخت تسهیلات پرداختی، با بحران‌های متعددی روبرو شده‌اند. بحران‌های اجتماعی و پنهان، مسئولین نهادهای نظارتی و اجرایی سیستم‌های مالی را بر آن داشته تا مدیریت ریسک نهادهای مالی، به‌ویژه بانک‌ها را با جدیت بیش‌تر و کارشناسانه‌تر مورد توجه قرار دهند.

ج) نوآوری در تحقیق

«ریسک اعتباری» یکی از مهم‌ترین ریسک‌های بانکی است که تاکنون اقدامی علمی و عملی برای اندازه‌گیری آن در بانک‌ها صورت نگرفته، و ارزیابی و سنجش اعتبار مشتریان تنها بر مبنای قضاوت ذهنی کارشناسان بانک‌ها صورت می‌گیرد. ارائه‌ی این الگو به بانک‌ها و مؤسسات مالی کمک خواهد کرد تا در ارزیابی مشتریان خود، ضمن به‌کارگیری ضوابط و معیارهای علمی و با کاهش ریسک اعتباری و احتمال نکول تسهیلات، اعتبارات را به‌صورت بهینه تخصیص دهند و از این طریق بر ریسک نقدینگی نیز غلبه یابند.

د) سؤالات پژوهش

-- عوامل مؤثر بر ریسک اعتباری مشتریان حقیقی بانک کدام است؟
 -- کدام یک از خصوصیات فردی مشتریان، اثرگذاری بیشتری در وضعیت اعتباری مشتریان دارد؟
 -- کدام یک از مدل‌های کمی‌سازی ریسک اعتباری، در تعیین ریسک اعتباری مشتریان اطمینان‌بخش‌تر است؟

-- آیا مدل‌های لاجیت، پروبیت، تحلیل ممیزی چندگانه و شبکه‌های عصبی MLP، RBFN و GMDH قابلیت برآورد کمی ریسک اعتباری مشتریان حقیقی متقاضیان تسهیلات را دارد؟

ه) مفروضات پژوهش

-- مدل‌های لاجیت، پروبیت، تحلیل ممیزی چندگانه و شبکه‌های عصبی MLP، RBFN و GMDH از قابلیت برآورد کمی ریسک اعتباری مشتریان حقیقی متقاضیان تسهیلات برخوردارند.

-- قدرت شبکه‌های عصبی GMDH در پیش‌بینی ریسک اعتباری بیشتر از سایر مدل‌هاست.

-- بین مدت بازپرداخت تسهیلات، نرخ سود و نوع تسهیلات (سرمایه در گردش، سرمایه‌ی، مصرفی و غیره) با ریسک اعتباری مشتری رابطه‌ی معناداری وجود دارد.

-- بین خصوصیات فردی مشتری نظیر سن، شغل، وضعیت تأهل و غیره، نسبت معناداری وجود دارد.

۳. ادبیات پژوهش

افزایش توجه صنعت اعتباری به تصمیم‌گیری در مورد اعطاء یا عدم اعطاء اعتبارات به متقاضیان، از سال ۱۹۶۰ باعث رشد دو برابری این صنعت در این مدت کوتاه شده است.^[۱] توجه روزافزون به این موضوع باعث خلق تکنیک‌هایی مفید در این حوزه، تحت عنوان مدل‌های رتبه‌بندی اعتباری توسط بانک‌داران و محققین به‌منظور فرایند ارزیابی متقاضیان شده است؛ هدف این مدل‌ها شناسایی مشتریان خوب و مشتریان بد است. سالیوان (۱۹۸۱) و بیلیر (۲۰۰۴) بیان داشته‌اند که در ارزیابی از طریق تکنیک‌های قضاوتی، هر درخواست به‌صورت مجزا از طریق تصمیم‌گیرنده‌ی

۲. روش‌شناسی پژوهش

۲.۱. ابزار گردآوری داده‌ها و تعریف متغیرها

اطلاعات مورد نیاز در این پژوهش، به‌صورت نمونه‌گیری تصادفی از پرونده‌ی اعتباری ۲۰۰ نفر از مشتریان اعتباری حقیقی یکی از بانک‌های دولتی در فاصله‌ی سال‌های ۱۳۸۵ تا ۱۳۹۰ جمع‌آوری شده، که به‌لحاظ لزوم حفظ امنیت و مسائل امنیتی از افشای نام بانک خودداری می‌کنیم. اطلاعات جمع‌آوری شده با استفاده از نرم‌افزارهایی نظیر Eviwes، SPSS و Multi GMDH مورد تحلیل و بررسی قرار گرفته است. از جمله مشکلات و تنگناهای این پژوهش، می‌توان به محرمانه‌بودن و نقص اطلاعات در پرونده‌های اعتباری مشتریان بانک اشاره کرد.

۲.۲. جامعه‌ی آماری

«مشتریان بدحساب» براساس تعریف ارائه شده توسط کمیته‌ی بال عبارت‌اند از: مشتریانی که یک یا چند مورد از موارد زیر در مورد آنها صدق کند:

- تردید در توانایی بازپرداخت تعهدات، اعم از اصل و بهره‌ی آن؛
- قرارگرفتن در زمره‌ی بدهکارانی که با تشکیل پرونده اعلام ورشکستگی کرده‌اند؛
- قرارگرفتن در زمره‌ی بدهکارانی که بیش از ۹۰ روز از تعهد اعتباری خود را سپری کرده‌اند.

۳.۲. برآورد حجم نمونه و روش نمونه‌گیری

نمونه‌گیری به‌صورت تصادفی و با استفاده از روش نمونه‌گیری خوشه‌یی از پرونده‌ی تسهیلات مشتریان حقیقی یکی از بانک‌های دولتی انجام شده است.

۴.۲. سؤالات و مفروضات پژوهش

چنان که گفته شد «طراحی الگوی ریاضی اندازه‌گیری ریسک اعتباری مشتریان» هدف اصلی این پژوهش است و طرح موضوع آن در قالب سؤالات و مفروضات ذیل قابل بررسی است.

الف) هدف تحقیق

هدف آن است که با استفاده از مدل‌های مذکور، ضمن ارزیابی صحیح از ریسک اعتباری مشتریان و با کاهش احتمال نکول مطالبات، تسهیلات بانکی به‌صورت بهینه تخصیص یابد.

ب) ضرورت تحقیق

در سال‌های اخیر هم‌زمان با تخصیص حجم عظیم منابع مالی توسط بانک‌های مختلف دنیا، شاهد بحران‌ها، زیان‌ها و حتی ورشکستگی‌های متعدد بانک‌ها بوده‌ایم.

را با تحلیل ممیزی رگرسیون لاجیت و پروبیت مقایسه کردند و در پایان، شبکه‌ی عصبی PNN را به‌عنوان دقیق‌ترین روش برگزیدند.

در ایران نیز، مقالات مختلف متعددی در زمینه‌ی کاربرد شبکه‌های عصبی در علوم مختلف -- نظیر فنی و مهندسی، اقتصاد و مدیریت -- به چاپ رسیده است. اما کاربرد شبکه‌های عصبی در رتبه‌بندی اعتباری کم‌تر مورد توجه قرار گرفته است، برای مثال می‌توان به گزارش عملکرد بانک‌های کشور در سال ۱۳۸۹ اشاره کرد.^[۱۵] بیل و جکسون^[۱۶] معیارهای مورد استفاده در رتبه‌بندی اشخاص حقوقی توسط مؤسسه‌ی استاندارد پورز^{۱۲} را با استفاده از شبکه‌های عصبی معرفی می‌کنند؛ از آن جمله می‌توان به ریسک تجارت، ریسک مالی، ساختار بدهی و سرمایه و انعطاف پذیری مالی اشاره کرد. جلیلی^[۱۷-۱۹] مزایای رتبه‌بندی اعتباری اشخاص را با استفاده از مدل‌های شبکه‌ی عصبی، توسعه بازارهای مالی و اعتباری، کاهش تخلفات ناشی از اشخاص حقیقی و حقوقی بانک و سایر مؤسسات اعتباری غیر بانکی در جریان اعطای تسهیلات موجود اعتباری تسریع و افزایش کارایی بانک‌ها و مراکز اعتباری دانسته است. کمیجانی و سعادت‌فر^[۲۰-۲۲] به کاربرد مدل‌های شبکه‌های عصبی در پیش‌بینی ورشکستگی اقتصادی شرکت‌های بازار بورس پرداخته‌اند. کمیجانی^[۲۳] و همکاران به مدل‌سازی ریسک اعتباری بازپرداخت تسهیلات مشتریان اعتباری بانک ملی ایران و کربخانی و فراتی^[۲۴] به بررسی جامع اثر متغیرهای کلان اقتصادی بر منابع و مصارف بانک‌ها پرداخته‌اند. برخی مفاهیم به‌کار رفته در مقاله عبارت‌اند از:

الف) تعریف ریسک اعتباری و شاخص‌ها

ریسک اعتباری از قدیمی‌ترین و بزرگ‌ترین خطرات موجود در دادوستد هاست و خطر سوخت شدن اعتبار اعطایی هم که با کوتاهی در پرداخت بدهی از سوی وام‌گیرندگان پیش می‌آید، یکی از بزرگ‌ترین خطرات مدیریتی تا این زمان است. بدهی است غالبه بر ریسک اعتباری با سازوکارهای اقتصادی همراه است و در نتیجه بیشتر بانک‌ها سیستم رتبه‌بندی داخلی پیشرفت‌ها را برای وام‌گیرندگان در نظر می‌گیرند.

ب) شاخص‌های ریسک اعتباری

رعایت دقت و هوشیاری مبتنی بر ضوابط و معیارهای لازم به‌هنگام اعطای تسهیلات ضرورتی است که لازمه‌ی آن، شناسایی شاخص‌های ریسک اعتباری در اعطای تسهیلات است. این شاخص‌ها، به دو دسته: الف) اطلاعات اعتباری، و ب) صورت‌های مالی تقسیم می‌شوند. از نظر کمیته‌ی بال، عواملی که موجب افزایش ریسک اعتباری بانک‌ها و مؤسسات مالی می‌شود عبارت‌اند از:

- اعطای تسهیلات با حجم بالا به متقاضیان تسهیلات در یک بخش اقتصادی؛
 - اعطای تسهیلات به شرکت‌ها و تشکل‌های صنعتی مادر که به صورت تار عنکبوتی مالک سهم یکدیگرند؛
 - اعطای تسهیلات بدون توجه به درجه اعتباری شرکت مادر و صورت‌های مالی تلفیقی آن‌ها (عدم توجه به ضوابط اعطای تسهیلات به ذی‌نفع واحد)؛
 - اعطای تسهیلات به نهادها و افرادی که توانایی نفوذ در مدیریت بانک‌ها را دارند.
- رعایت معیارها و ضوابط اساسی قبل از اعطای تسهیلات از جمله ابزارهای کاهش ریسک اعتباری است. این معیارها و ضوابط عبارت‌اند از:
- قابلیت اعتماد و اطمینان؛

اعتباری مورد بررسی قرار می‌گیرد.^[۲] مزیت این روش آن است که براساس تجربه دیدگاه تحلیل‌گر اعتباری شکل گرفته است. به عبارت دیگر تکنیک‌های قضای نامتجانس‌اند و از ضعف انگیزشی، کنترل و کمی‌سازی ریسک برخوردارند.

امروزه مطالعات رتبه‌بندی در دو حوزه‌ی متفاوت صورت می‌گیرد: ۱. پیش‌بینی ورشکستگی شرکت‌ها؛ ۲. پیش‌بینی ریسک اعتباری فردی. آلتمن و همکاران (۱۹۹۴) از دو روش تحلیل ممیزی و شبکه‌های عصبی پرسپترون^۱ برای پیش‌بینی ورشکستگی ۱۰۰۰ شرکت ایتالیایی استفاده کردند^[۲] (مدل z-score) و نشان دادند که دقت شبکه‌های عصبی به‌مراتب بالاتر از روش‌های سنتی آماری است. کوانس و فانت (۱۹۹۳) با به‌کارگیری شبکه‌های MLP عصبی و تحلیل ممیزی برای پیش‌بینی ورشکستگی شرکت‌هایی که اطلاعات آن‌ها از کامپوستات^۲ در دوره زمانی ۱۹۷۰-۱۹۸۹ گرفته شده بود^[۳] نشان دادند که شبکه‌های عصبی به‌مراتب دقیق‌تر از روش‌های تحلیل ممیزی‌اند. لاچر و همکاران نیز با بهره‌گیری از اطلاعات به‌دست آمده از کوانتس دقت شبکه‌های عصبی را نشان دادند.^[۵] سالچین بیلی^۳ از شبکه‌های عصبی MLP به‌منظور پیش‌بینی سلامت مالی پس‌انداز و وام‌ها استفاده کرد. وی این روش را با مدل رگرسیونی لاجیت مقایسه کرد.

دیوید وست (۲۰۰۰) ۵ نوع شبکه‌ی عصبی MOE^۴، RBF^۵، FAR^۶، LVQ^۷، MLP را با درخت‌های تصمیم‌گیری، تخمین چگالی کرنل^۸، روش نزدیک‌ترین همسایگی، تحلیل ممیزی و رگرسیون لاجیت در امتیازدهی اعتباری بر روی اطلاعات اعتباری کشور آلمان و استرالیا مقایسه کرد^[۶] و نشان داد که MLP از تمامی آن‌ها دقیق‌تر است.^[۷]

در ادامه، جیسوال به ارتباط بین دارایی‌ها و مسئولیت بانک‌های تجاری،^[۸] و نوره به مباحث ارزیابی ریسک اعتباری^[۹] پرداخته‌اند. نوره در مطالعه‌ی با عنوان «ارزیابی ریسک اعتباری: کاربرد داده‌کاوی در یک بانک روستایی برای ارائه‌ی یک درخت تصمیم‌گیری به‌منظور ارزیابی ریسک اعتباری بانکی»، از تکنیک داده‌کاوی استفاده کرد؛ مدل ارائه‌شده معیاری جدید در تحلیل کاربرد وام فراهم کرد. نتایج ارزیابی‌ها نشان داد که با به‌کارگیری این مدل بانک می‌تواند وام‌های ناکارآمد را تا کم‌تر از ۵ درصد کاهش دهد و در زمره‌ی بانک‌های کارآمد دسته‌بندی شود.

همچنین سیلوا^[۱۰] به مباحث قانونی و نقش قوانین و مقررات بانکی در اقتصاد تحت عنوان «ریسک اعتباری و اثرات شوک نقدینگی» پرداخته و کاسترو^[۱۱] عوامل اقتصاد کلان ریسک اعتباری در صنعت بانک‌داری را بیان کرده است. کاسترو در مطالعه‌ی با عنوان «عوامل اقتصاد کلان ریسک اعتباری در سیستم بانکی GIPSI»، رابطه‌ی میان پیشرفت‌های اقتصاد کلان و ریسک اعتباری بانکی در کشورهای یونان، ایرلند، پرتغال، اسپانیا و ایتالیا را که اخیراً تحت تأثیر شرایط نامطلوب اقتصادی و مالی قرار گرفته‌اند، مورد بررسی قرار داد. یافته‌های این مقاله نشان می‌دهد که تمام معیارهای سیاسی قابل اجرا برای ترویج رشد، اشتغال، بهره‌وری و رقابت با هدف کاهش بدهی‌های عمومی و خارجی در این کشورها، برای ایجاد ثبات در اقتصاد بسیار ضروری‌اند.

در مقایسه، قدرت شبکه‌های عصبی مصنوعی در مدل‌های آماری سنتی بیشتر است و نیز با تغییر تکنیک‌های تفسیر وزن‌های شبکه‌ی عصبی، امتیازدهی اعتبار نتایج متفاوتی خواهد داشت. نتایج پژوهش نیز این دو فرضیه را اثبات کرد.

موچین چین^[۱۲] و شی هسین هوانگ^[۱۳] با بحث در خصوص برتری روش‌های رده‌بندی الگوریتم ژنتیک و شبکه‌های عصبی در امتیازدهی اعتباری بر دیگر روش‌های آماری، از روش شبکه‌های عصبی در زمینه‌ی پذیرش یا عدم پذیرش درخواست متقاضیان وام استفاده کردند. جان پوینتون و احمد آل مصری^[۱۴] در رتبه‌بندی اعتباری بانک تجاری مصر، شبکه‌های PN^۹، MLFF^{۱۰}، BNS^{۱۱}

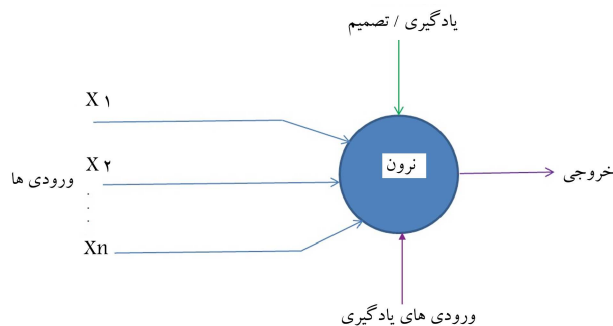
• قابلیت و صلاحیت فنی؛

• ظرفیت مالی و کشش اعتباری (اداره تحقیقات و کنترل ریسک بانک سپه).

۴. مبانی نظری مدل‌های شبکه‌های عصبی

۴.۱. شبکه‌ی عصبی مصنوعی

شبکه‌ی عصبی شامل مجموعه‌ی از عناصر عملیاتی ساده به نام نورون‌هاست که در لایه‌های ورودی، میانی، و خروجی قرار گرفته‌اند و چندین لایه میانی را بین لایه‌ی ورودی و لایه‌ی خروجی در بر می‌گیرند. این شبکه‌ها، به دلیل قابلیت یادگیری و تعمیم در بسیاری از موارد -- نظیر تشخیص الگو، تشخیص هویت، طبقه‌بندی، پردازش صوت و تصویر و سیستم کنترلی -- کاربرد دارند. یک مدل شبکه‌ی عصبی از متغیرهای ورودی در لایه‌ی اول استفاده می‌کند. خروجی شبکه معمولاً راه‌حلی برای یک مسئله است. در این پژوهش، خروجی شبکه می‌تواند نشان‌دهنده‌ی یک متقاضی دارای اعتبار خوب یا بد باشد. بدین ترتیب شبکه پس از آموزش، برای متقاضی دارای اعتبار خوب -- که حائز شرایط دریافت تسهیلات است -- خروجی ۱ و برای متقاضی دارای اعتبار بد -- که فاقد شرایط دریافت تسهیلات است -- خروجی صفر را نشان می‌دهد. مدل‌های شبکه‌ی عصبی برای محاسبه‌ی خروجی‌ها از وزن‌هایی بهره می‌گیرند که به هر یک از ارتباطات نورون‌ها تخصیص می‌یابد. وزن‌ها در قالب ارزش عددی، ارتباط بین دو نورون و اهمیت نسبی هر متغیر ورودی را نشان می‌دهند. فرایند یادگیری نظارتی^{۱۳} در شبکه‌های عصبی، شامل محاسبه‌ی خروجی‌ها و اصلاح مکرر وزن‌ها تا رسیدن به خروجی مناسب، مطابق با بسیار نزدیک (در فاصله‌ی کم‌تر از خطای مورد نظر) به خروجی هدف^{۱۴} است. با تکرار فرایند یادگیری، شبکه قادر به شناسایی مقادیر صحیح وزن‌هاست. در شبکه‌های عصبی مختلف، مقدار خطا براساس الگوریتم‌های یادگیری مورد استفاده محاسبه می‌شود. به دلیل قابلیت الگوریتم یادگیری پس‌انتشار خطا^{۱۵} در شناسایی الگوها در دامنه‌ی وسیعی از داده‌ها، می‌توان از آنها در امور مالی و در پیش‌بینی عملکرد سیستم‌های مالی، بررسی عملکرد سهام، رتبه‌بندی اعتباری و بررسی درخواست‌های وام یا شناسایی مشتریان درمورد کارت‌های اعتباری استفاده کرد. از موارد حائز اهمیت در یادگیری شبکه‌های عصبی که در طبقه‌بندی الگوها و دیگر کاربردها مؤثرند می‌توان به «ساختار شبکه^{۱۶}» اشاره کرد که تعداد نورون‌ها، تعداد لایه‌ها و نحوه‌ی قرارگیری آن‌ها در کنار یکدیگر را شامل می‌شود. از آنجا که مفاهیم بنیادین مطرح شده در مباحث مربوط به آن، از پایه‌ی ریاضی مستحکم و قابل اعتمادی برخوردار بود، بعدها به‌طور گسترده مورد اقبال عمومی محققین در مدل‌سازی تجربی فرایندهای غیرخطی قرار گرفت. شبکه‌های عصبی مصنوعی قابل مقایسه با سیستم عصبی طبیعی نیستند اما، وجود ویژگی‌های مهم -- نظیر قابلیت یادگیری، قابلیت تعمیم، قابلیت پردازش موازی و قابلیت ترمیم خطاها و... -- این شبکه‌ها را در مواردی که نیاز به یک نگاشت خطی یا غیرخطی باشد متمایز می‌سازد.^{۱۵} به‌طور کلی سیستم عصبی انسان‌ها از یک سری عناصر پردازش ساده با نام «نورون^{۱۷}» تشکیل شده که مجموعه‌ی از این نورون‌ها عصب را تشکیل می‌دهند. شیوه‌ی کار نورون‌ها، به‌منظور پردازش اطلاعات چنین است که هرکدام از این نورون‌ها یک سیگنال اولیه، حاوی اطلاعات مهمی از یک نورون دیگر یا محرک خارجی دریافت کرده و آن را به‌کمک توابع فعال‌سازی یا تبدیل، پردازش می‌کنند. این ویژگی (پردازش اطلاعات)، شبکه‌های عصبی را به‌عنوان یک روش محاسباتی قدرتمند به‌منظور یادگیری مثال‌ها و تعمیم این یادگیری به مثال‌هایی که تاکنون مشاهده نشده تبدیل می‌کند.^{۱۶} بنابراین



شکل ۱. نمای کلی شبکه‌ی عصبی مصنوعی.

الگوریتم‌های یادگیری و روش‌های آموزش نیز در دقت یادگیری شبکه‌های عصبی از اهمیت به‌سزایی برخوردارند.

در این پژوهش سعی شده تا با استفاده از چندین ساختار متمایز شبکه‌های عصبی برای رتبه‌بندی اعتباری^{۱۸} مشتریان با اعتبار خوب و بد، نتایج حاصله و نهایتاً مقایسه‌ی میزان خطای هر یک تحلیل و بررسی شود. نمای کلی از یک شبکه‌ی عصبی مصنوعی در شکل ۱ آمده است.

۴.۲. مدل‌های شبکه‌ی عصبی مبتنی بر الگوریتم GMDH

این الگوریتم یکی از الگوریتم‌های مورد استفاده در شبکه‌های عصبی است که جایگزین بسیار مناسبی برای رگرسیون‌های استاندارد با فرم حاصل ضرب است و در سال ۱۹۹۶ توسط ایواخنینکو^{۱۹} برای رهایی از پیچیدگی محاسبات و مشکل وابستگی خطی رگرسیون‌ها ابداع شده^{۲۰} و درحقیقت تکنیکی برای ساخت یک چندجمله‌ی بی‌سبب با مراتب بالاست. این روشی ایده‌آل برای سیستم‌های پیچیده با ساختار نامشخص است که در آن تحلیل‌گر علاقه‌مند به کشف روابط متغیرهای ورودی و خروجی با مراتب بالاست.

الگوریتم ایواخنینکو یک روش اکتشافی است که دانش را از ذات و طبیعت داده‌ها استخراج می‌کند و در آن تحلیل‌های رگرسیونی مبتنی بر یک بنیان ثابت نظری نیست.^{۲۱} بنابراین ایده‌ی اصلی این الگوریتم، طراحی یک مدل بهینه‌ی پیچیده است که فقط مدل را بر پایه‌ی داده‌ها و اطلاعات طراحی کرده و هیچ‌گونه پیش‌زمینه‌ی نظری از نحوه‌ی عملکرد داده‌ها از سوی محقق صورت نمی‌گیرد. این کار صرفاً براساس کشف ارتباط ساده و پیچیده میان داده‌های ورودی و خروجی سیستم صورت می‌گیرد. در نهایت، الگوریتم ایواخنینکو یک مدل خودتنظیم‌گر^{۲۰} است که قابلیت حل مسائل پیش‌بینی، تشخیص، ترکیبات کنترلی و سایر مسائل سیستمی به‌کار برده شده را دارد. اساساً روش‌های تکاملی^{۲۱} مانند الگوریتم ژنتیک، کاربرد وسیعی در مراحل مختلف طراحی شبکه‌های عصبی دارند^{۲۲} به‌طوری که دارای قابلیت‌های منحصر به فردی در یافتن مقادیر بهینه و امکان جست‌وجو در فضاها غیر قابل پیش‌بینی هستند.^{۲۳} از این‌رو در این پژوهش به‌منظور طراحی شکل شبکه‌ی عصبی و تعیین ضرایب آن از الگوریتم ژنتیک استفاده شده است. زیرا شبکه‌ی عصبی GMDH مبتنی بر الگوریتم ژنتیک، ابزاری با قابلیت بالا در مدل‌سازی سیستم‌های غیرخطی پویای پیچیده به‌شمار می‌رود.

شبکه‌ی عصبی GMDH در برگرفته‌ی مجموعه‌ی از نورون‌هاست که از پیوند جفت‌های مختلف از طریق یک چندجمله‌ی درجه دوم به وجود می‌آید. فرض کنید مجموعه‌ی m از متغیر شامل x_1, x_2, \dots, x_m و یک متغیر y وجود دارد. داده‌های مربوط به هرکدام از x_i ها و متغیر هدف y متغیر خروجی نیز برای یک دوره

در این صورت رابطه‌ی F را می‌توان چنین تجزیه کرد:

$$\begin{aligned} \hat{y}_k &= G(u_i, u_j) \quad i, j = 1, 2 (i \neq j) & k = 1 \\ \hat{u}_k &= G(s_i, s_j) \quad i, j = 1, 2, k, F_1 (i \neq j) \quad F_1 \leq C_{F_1}^1 & k = 2 \\ \hat{s}_k &= G(p_i, p_j) \quad i, j = 1, 2, k, F_2 (i \neq j) \quad F_2 \leq C_{F_2}^1 & k = 3 \\ &\vdots \\ \hat{Z}_k &= G(w_i, w_j) \quad i, j = 1, 2, k, F_l (i \neq j) \quad F_l \leq C_{F_l}^1 & k = F_l + 1 \\ \hat{w}_k &= G(x_i, x_j) \quad i, j = 1, 2, k, F_m (i \neq j) & k = F_m \end{aligned} \quad (6)$$

همان‌گونه که در معادلات بالا مشاهده می‌شود ترتیب روابط فوق از بالا به پایین نمایی از فرایند تجزیه‌ی رابطه‌ی γ به چند جمله‌ی‌های درجه دوم بوده و از طرفی نیز ترتیب این روابط از پایین به بالا بیان‌گر تکمیل رابطه‌ی γ توسط معادلات بازگشتی است. در واقع هدف این الگوریتم یافتن ضرایب مجهول α در سری توابع ولتر است. لازم به ذکر است که تمامی مدل‌های جزئی حاصله از یک ساختار متشابه همانند رابطه‌ی γ برخوردارند:

$$\hat{f}(x_i, x_j) = v_0 + v_1 x_i + v_2 x_j + v_3 x_i^2 + v_4 x_j^2 + v_5 x_i x_j \quad (7)$$

از آنجا که هدفی را که ما در این الگوریتم دنبال می‌کنیم، چیزی جز مدل‌سازی سیستم اولیه نیست، با ترکیب مدل سیستم‌های جزئی و تکرار این عمل می‌توان به مدل اصلی سیستم دست یافت (رابطه‌ی ۸).

$$\hat{y} = v_0 + \sum_{i=1}^m v_i x_i + \sum_{i=1}^m \sum_{j=1}^m v_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m v_{ijk} x_i x_j x_k + \dots \quad (8)$$

بعد از تجزیه‌ی سیستم اصلی به تعداد C_m^1 سیستم جزئی مدلی با دو متغیر ورودی برای هر یک از آن‌ها محاسبه می‌شود. سپس مدل‌های جزئی ایجاد شده دو به دو با هم ترکیب می‌شوند که حاصل این ترکیب، تعداد $\frac{C_m^1(C_m^1-1)}{2}$ سیستم با مدل جزئی جدید با حداقل سه و چهار متغیر ورودی است. البته تعداد متغیرهای وابسته به مدل یا به عبارتی تعدادی ورودی‌های سیستم مهم نیست و تنها دقت تخمین واقعی سیستم اصلی توسط مدل‌های ایجاد شده اهمیت دارد. لذا با در نظر گرفتن این قاعده برای کاهش محاسبات مضاعف و زائد و افزایش راندمان و دقت مدل‌سازی تعدادی از مدل‌های جزئی تشکیل شده را که از دقت و تخمین بالایی نسبت به مدل‌های دیگر برخوردارند انتخاب، و مابقی را حذف می‌کنیم. [۳۹-۳۷]

در دومین مرحله، ترکیب مدل‌های جزئی انتخاب شده یا به عبارتی سیستم‌های ایده‌آل شکل گرفته در مرحله‌ی قبلی مجدداً دو به دو همانند مرحله‌ی قبلی ترکیب شده و سیستم‌های جزئی جدیدی با حداقل پنج و حداکثر شش متغیر ورودی تشکیل می‌شود. به همین طریق در مراحل بعدی نیز با انتخاب و حذف تعدادی از مدل‌های جزئی ایجاد شده عمل ترکیب آن‌ها را ادامه می‌دهیم تا در نهایت به مدلی نسبتاً ایده‌آل دست یابیم.

هدفی که همواره در فرایند ترکیب مدل‌های به‌وجود آمده مورد نظر است دست یافتن به مدلی است که تقریباً تمامی متغیرهای سیستم در آن نمایان و نقش داشته باشند. هدف مورد نظر دیگر انجام ترکیب‌های مکرر، رسیدن به مدلی است که میزان خطای خروجی آن نسبت به سایر مدل‌های محاسبه شده در مراحل قبل کم‌تر باشد. [۴۰]

$$\begin{aligned} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\eta_1} \\ \dots \\ y_n \end{bmatrix} &= \begin{bmatrix} x_{11} & \vdots & x_{1r} & \dots & \vdots & x_{1m} \\ x_{21} & \vdots & x_{2r} & \dots & \vdots & x_{2m} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ x_{\eta_1 1} & \vdots & x_{\eta_1 r} & \dots & \vdots & x_{\eta_1 m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & \vdots & x_{nr} & \dots & \vdots & x_{nm} \end{bmatrix} \end{aligned} \quad \begin{array}{l} \text{مجموعه آموزش} \\ \\ \\ \\ \\ \\ \text{مجموعه آزمون} \end{array}$$

منبع: زمانی (۱۳۸۸)

نمودار ۱. متغیرهای ورودی به الگوریتم GMDH. [۳۲]

زمانی وجود دارد. به عبارتی هر یک از متغیرها به صورت یک بردار شامل اعداد سری زمانی مربوط به آن متغیر است. [۳۲] اطلاعات اولیه‌ی که به منظور ساخت الگوریتم GMDH باید جمع‌آوری کرد مجموعه‌ی n مشاهده است که در نمودار ۱ نشان داده شده است. [۳۲]

برای شروع به‌کار الگوریتم با دو مسئله مواجه هستیم: ۱. تشخیص رابطه‌ی که متغیر خروجی را براساس متغیرهای ورودی x_i ها تولید می‌کند؛ ۲. پیش‌بینی y به‌ازای مقادیر معلوم x_i ها. به عبارتی، تشخیص مدل و رابطه‌ی بین متغیرها (مدل‌سازی) لازم است تا بتوان از روی آن مدل مقادیر آتی متغیر هدف را پیش‌بینی کرد. [۳۲]

مبنای الگوریتم GMDH عبارت است از فرایندی به منظور ساخت یک چندجمله‌ی با مراتب بالا به نام «سری تابع ولتر^۲» است و مطابق رابطه‌ی ۱ ارائه می‌شود: (این چند جمله‌ی را ایواخینسکو نیز می‌نامند). [۳۵]

$$\hat{y} = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \dots \quad (1)$$

برای این منظور در الگوریتم GMDH ابتدا به تجزیه‌ی سری توابع ولتر با چندجمله‌ی‌های دومتغیره‌ی درجه دوم می‌پردازیم. [۳۶]

$$G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \quad (2)$$

در این تجزیه، سری ولتر با مجموعه‌ی از معادلات بازگشتی زنجیره‌ی تبدیل می‌شود، به‌گونه‌ی که مجدداً با جایگذاری جبری هر یک از روابط بازگشتی در این رابطه سری ولتر برقرار می‌شود. رابطه‌ی ۳ توسط تابع f تقریب زده می‌شود:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}) \quad i = 1, 2, 3, \dots, m \quad (3)$$

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{im}) \quad i = 1, 2, 3, \dots, m \quad (4)$$

و در صورتی که تابع f مطابق رابطه‌ی ۵ بیان شود:

$$\hat{y} = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \dots \quad (5)$$

۳.۴. مدل تحلیل ممیزی (DA)

-- توقف، وقتی که خطا در مثال‌های مجموعه‌ی تأیید شده از قاعده‌ی خاصی پیروی کند.

اگر دفعات تکرار کم باشد، خطای زیادی خواهیم داشت و اگر تعداد دفعات زیاد باشد مسئله Over fitting رخ خواهد داد.^[۲۴] با این وجود، موارد زیر را می‌توان به صورت قوانین کلی بیان کرد:

توابع بولی: می‌توان آن را توسط یک شبکه‌ی دولایه پیاده‌سازی کرد.

توابع پیوسته: هر تابع پیوسته‌ی محدود را می‌توان توسط یک شبکه‌ی دولایه تقریب زد. نظریه‌ی مربوط در مورد شبکه‌هایی که از تابع سیگموئید در لایه‌ی پنهان و لایه‌ی خطی در شبکه‌ی خروجی استفاده می‌کنند، صادق است.

توابع دلخواه: هر تابع دلخواه را می‌توان با یک شبکه‌ی سه‌لایه تا حد قابل قبولی تقریب زد.

یکی از خواص الگوریتم پس‌انتشار خطا این است که در لایه‌های پنهان شبکه، ویژگی‌های ناشکاری از داده‌ی ورودی نشان می‌دهد و این موضوع با افزایش تعداد تکرار بیشتر حاصل خواهد شد.^[۲۵]

تحلیل ممیزی یک روش قدرتمند چندمتغیره‌ی آماری است که توانایی گروه‌بندی داده‌ها را در دو گروه یا بیشتر دارد. در این مدل، با استفاده از یک یا چند معادله‌ی مبتنی بر رابطه‌ی خطی بین متغیرهای مستقل، متغیرهای وابسته به یک یا چند گروه تقسیم می‌شود. توابع ممیزی عبارت‌اند از:^[۲۱]

$$Z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (9)$$

که در آن، Z نمایاننده‌ی امتیاز ممیزی، α عرض از مبدأ، و β_i ضریب متناظر با متغیر X است. بنابراین، فرض‌های اصلی آن عبارت‌اند از:

۱. متغیرهای مستقل در مقیاس فاصله‌ی محاسبه می‌شوند.

۲. ماتریس کواریانس متغیرهای مستقل یکسان می‌دهد.

۳. متغیرهای مستقل دارای توزیع نرمال چندمتغیره‌اند.^[۲۲]

۴.۴. شبکه‌های عصبی شعاع‌محور (RBN)

شبکه‌های عصبی شعاع‌محور ساختاری نسبتاً متفاوتی دارند. بدین ترتیب که لایه‌های پنهانی به توابع متقارن شعاعی محدود شده‌اند. این شبکه‌ها عکس‌العمل موضعی به اطلاعات ورودی نشان می‌دهند. خروجی واحد گاوسی z_m عبارت است از:

$$y(k) = \sum_{j=1}^m w_{kj} \exp(-\|x - \mu_j\|^2 / 2\sigma_j^2) + w_k \quad \forall j = 1, 2, \dots, m \quad (10)$$

که در آن x تابع ورودی -1 بعدی است، μ_j بردار وزن‌های -1 ورودی واحد گاوسی z_m است. پارامتر σ_j^2 فاصله‌ی بیرونی که تابع گاوسی عکس‌العمل نشان می‌دهد و در نتیجه تصفیه‌ی تابع فرضی را تحت تأثیر قرار می‌دهد. $y(k)$ به عنوان خروجی RBN ترکیب خطی وزنی m تابع گاوسی محور است. w_{kj} تورش متناظر با $y(k)$ و μ_j ها عناصر ماتریس وزنی w هستند. شبکه‌های RBN به عنوان شبکه‌های موضعی شناخته شده‌اند، زیرا خروجی $y(k)$ تنها براساس توابع با مرکزیت نزدیک به ماتریس ورودی X محاسبه می‌شود.

۵.۴. الگوریتم پس‌انتشار خطا

این الگوریتم یکی از معمول‌ترین الگوریتم‌های یادگیری نظارت شده در شبکه‌های عصبی چندلایه‌ی پیشرو^{۲۲} محسوب می‌شود و اساس آن، تصحیح پارامترهای آزاد شبکه با توجه به خطای شبکه یا تفاوت پاسخ واقعی شبکه و پاسخ مطلوب است. برای تعیین مقدار تصحیح از گرادیان^{۲۳} است.^[۲۳] در این روش با استفاده از گرادیان سعی می‌شود تا مربع خطای بین خروجی‌های شبکه و تابع هدف کمینه شود.

$$\Delta w_{ji} = \eta \sigma_j x_{ji} \quad (11)$$

معمولاً الگوریتم پس‌انتشار خطا پیش از خاتمه، هزاران بار با استفاده از همان داده‌های آموزشی تکرار می‌شود. شروط مختلفی را می‌توان برای خاتمه‌ی الگوریتم به کار برد:

-- توقف، بعد از تکرار به دفعات معین.

-- توقف، وقتی که خطا از یک مقدار تعیین شده کم‌تر شود.

۶.۴. مدل تابع لاجبیت

این مدل یکی از مدل‌های رگرسیونی است که در آن وقتی متغیر وابسته‌ی ما باینری یا موهومی باشد مورد استفاده قرار می‌گیرد. در این حالت Y به عنوان متغیر وابسته تنها می‌تواند مقادیر 0 و 1 را داشته باشد. اگر X_i متغیر مستقل باشد احتمال این که متغیر Y مقدار 1 را بپذیرد برابر است با:

$$P_i = E(y = 1 | x_i, i = 1, 2, \dots, n) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni})}} \quad (12)$$

این تابع را تابع توزیع تجمعی لجستیک می‌نامند و دوباره چنین بازنویسی می‌شود:

$$P_i = \frac{1}{1 + e^{-Z_i}} \quad -\infty < Z_i < +\infty, \quad 0 < P_i < 1 \quad Z_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} = \ln(p / (1 - p)) \quad (13)$$

یعنی L (معروف به لاجبیت) که لگاریتم نسبت برتری یا مزیت است نه تنها برحسب X ، بلکه (نکته‌ی مهم از نظر تخمین) برحسب پارامترها نیز خطی است. ویژگی‌های مدل لاجبیت عبارت است از:

۱. همچنان که P بین 0 و 1 (یعنی Z بین $-\infty$ تا $+\infty$) نوسان می‌کند لاجبیت از $-\infty$ تا $+\infty$ تغییر می‌کند. یعنی مدل‌های لاجبیت اگر چه برحسب ضرورت بین 0 و 1 قرار می‌گیرند، در این فاصله محدود نیستند.

۲. اگر چه L بر حسب X خطی است؛ خود احتمال‌ها چنین نیستند.

۳. تعبیر مدل لاجبیت: β ضریب زاویه‌ی میزان تغییر در L را به ازای یک واحد تغییر در X یا به عبارت دیگر میزان تغییر در لگاریتم نسبت برتری یا مزیت را به ازای یک واحد تغییر اندازه می‌گیرد.

تخمین مدل لاجبیت در تحقیق، با استفاده از نرم‌افزار Eviews انجام شد. به منظور تخمین مدل لاجبیت در مرحله‌ی اول، وجود رابطه‌ی هم‌خطی بین متغیرهای اولیه با درجه اطمینان ۹۵ درصد بررسی شد. در مرحله‌ی بعد به منظور دست‌یابی به مدل بهینه‌ی اعتبارسنجی، کلیه متغیرها را وارد مدل می‌کنیم و سپس با استفاده از

۶. مدل لاجیت

مدل‌های لاجیت و پروبیت در مواردی کاربرد دارند که متغیر وابسته قابل مشاهده نباشد.

۱.۶. تخمین مدل

به منظور تخمین مدل لاجیت، در مرحله اول وجود رابطه‌ی هم خطی بین متغیرهای اولیه را با درجه اطمینان ۹۵ درصد بررسی می‌کنیم. در جدول ۱ نتایج حاصل از این آزمون نشان داده شده است.

در مرحله‌ی بعد، به منظور دست‌یابی به مدل بهینه‌ی اعتبارسنجی، تمامی متغیرها در مدل وارد شده و در گام بعد، با استفاده از معناداری ضرایب رگرسیونی، متغیرهای نهایی اثرگذار بر وضعیت اعتباری مشتریان (۵ متغیر) انتخاب شد. در جدول ۲ شکل نهایی تخمین ارائه شده است.

در این تحلیل، متغیرهای وضعیت اقامت، تعداد اقساط، وضعیت اشتغال و نسبت وثایق به مقدار اعتبارات و وضعیت‌های تأهل به عنوان متغیرهای اثرگذار شناسایی شده‌اند. شکل کلی تابع لاجیت یا لگاریتم احتمال قصور در بازپرداخت عبارت است از:

$$\ln(P/1 - P) = 1,961302349 \times HOUSE + 0,04922019868 \times INSALTMENT - NO - 1,04439367 \times JOB - 1,9651643455 \times LOAN - COLLATERAL - 1,252849156 \times MARRIAGE + 1,03882212 \quad (15)$$

از آنجا که علامت نسبت وثایق به مقدار وام منفی است، هر قدر این نسبت برای یک مشتری بیشتر باشد احتمال قصور وی کم‌تر است. این امر باعث می‌شود که احتمال حضور وی در گروه مشتریان خوش حساب بیشتر شود. از همین منطبق برای تفسیر همه‌ی متغیرها استفاده می‌شود.

۲.۶. تحلیل اثر نهایی

تفسیر مقادیر ضرایب مدل لاجیت پیچیده است، زیرا ضرایب برآورد شده حاصل یک مدل دوگزینه‌ی است که نمی‌تواند به عنوان اثر نهایی روی متغیر وابسته شود. اثر نهایی x_i روی احتمال شرطی به وسیله‌ی رابطه‌ی ۱۶ تعیین می‌شود. [۳۸]

$$E(y/(x), \beta) / \delta x_j \delta = F(-x\beta) \beta_j \quad (16)$$

مقادیر محاسبه شده‌ی اثر نهایی هر یک از متغیرهای نهایی در جدول ۳ آمده است.

۳.۶. آزمون نیکویی برازش

در این پژوهش، سعی شده با استفاده از آزمون هاسمر - لموشوف^{۲۵} مقدار برازش شده‌ی مورد انتظار با مقادیر واقعی هر گروه ارزیابی و مقایسه شود. این آزمون بر پایه‌ی پیش‌بینی احتمال $Y = 1$ ، مقدار برازش شده را گروه‌بندی می‌کند. آماره‌ی این آزمون مطابق رابطه‌ی ۱۷ محاسبه می‌شود:

$$\sum_{j=1}^n (O_j - E_j)^2 \sim x^2 n - 2 \quad (17)$$

معناداری ضرایب رگرسیونی متغیرهای نهایی اثرگذار بر وضعیت اعتباری، مشتریان انتخاب می‌شوند. در نهایت، پنج متغیر اثرگذار بر وضعیت اعتباری مشتریان انتخاب شد که در این تحلیل، متغیرهای وضعیت اقامت، تعداد اقساط، وضعیت اشتغال، نسبت وثیقه‌ها به مقدار اعتبارات، و وضعیت تأهل به عنوان متغیرهای اثرگذار شناسایی شد.

۷.۴. مدل پروبیت

این مدل نیز مانند مدل لاجیت یک تابع CDF^{۲۴} است با این تفاوت که در مدل لاجیت تابع لجستیک، اما در مدل پروبیت نرمال است. اگر I_i شاخص وقوع حادثه باشد، چنین بیان می‌شود:

$$I_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} \quad (14)$$

از سوی دیگر I_i^* را مقدار بحرانی شاخص فوق می‌نامیم. $I_i \geq I_i^*$ حادثه وقتی به وقوع می‌پیوندد که احتمال وقوع حادثه عبارت باشد از:

$$P_i = \Pr(y = 1) = t$$

t متغیر نرمال استاندارد است؛ یعنی: $t \sim N(0, 1)$. برای تخمین پارامترهای این مدل، از روش بیشترین راست‌نمایی استفاده می‌شود.

۵. تحلیل خروجی‌ها و نتایج پژوهش

در این پژوهش ۲۰۰ نفر از مشتریان حقیقی یکی از بانک‌های دولتی به صورت کاملاً تصادفی انتخاب شده‌اند. در بین این مشتریان ۱۵۰ نفر در گروه مشتریان خوش حساب و ۹۵ نفر در گروه مشتریان بدحساب قرار گرفتند. «مشتریان بدحساب» بر اساس تعریف کمیته‌ی بال عبارت‌اند از:

۱. مشتریانی که در توانایی بازپرداخت تعهدات‌شان، اعم از اصل و سود آن تردید وجود داشته باشد؛

۲. بدهکارانی که بیش از ۹۰ روز از تعهد اعتباری خود را سپری کرده‌اند؛

۳. بدهکارانی که با تشکیل پرونده اعلام ورشکستگی کرده باشند.

این حجم نمونه از بین مشتریان سال‌های ۱۳۸۵-۱۳۹۰ انتخاب شده است. اطلاعات جمع‌آوری شده عبارت‌اند از:

۱. جنسیت (صفر به‌ازای زن و ۱ به‌ازای مرد)

۲. سن

۳. وضعیت تأهل (صفر مجرد و ۱ متأهل)

۴. نرخ بهره

۵. نوع وثیقه (صفر به‌ازای وثیقه‌ی غیرملکی و ۱ به‌ازای وثیقه‌ی ملکی)

۶. تعداد اقساط

۷. نسبت مقدار وثیقه‌ها به مبلغ وام

۸. وضعیت سکونت (صفر ملکی و ۱ استیجاری)

۹. وضعیت اشتغال (صفر فاقد شغل و ۱ اشتغال)

زیر می‌باشد:

$$l = 1,166,539 \times HOUSE + 0,29529 \times INSALTMENT - NO - 0,621309 \times JOB - 0,986014 \times LOAN - COLLATERAL - 0,736157 \times MARRIAGE + 0,578431 \quad (18)$$

۱.۷. تحلیل اثر نهایی

مقادیر محاسبه شده اثر نهایی هر یک از متغیرهای نهایی مدل پروبیت در جدول ۷ آمده است.

۲.۷. آزمون نیکویی برازش

در این آماره نیز تعداد ۱۰ گروه در نظر گرفته شده است. مقدار آماره هاسمر - لموشوف طبق جدول ۸ معادل ۵,۳۷ بوده و از ۱۶ کم تر است. این نشان می‌دهد که نیکویی برازش با سطح اطمینان ۹۵ درصد تأیید شده است.

۳.۷. جدول پیش‌بینی مورد انتظار

این جدول پیش‌بینی مورد انتظار مدل پروبیت را نشان می‌دهد. در این تحقیق مقدار برش ۰/۵ در نظر گرفته شده است. نتایج این تحلیل در جدول ۹ قابل مشاهده است.

۸. مدل تحلیل ممیزی

۱.۸. تخمین مدل

کلیدی متغیرها را وارد مدل می‌کنیم و سپس با استفاده از معناداری ضرایب رگرسیونی، متغیرهای نهایی اثرگذار بر وضعیت اعتباری مشتریان انتخاب می‌شود. در جدول ۱۰ شکل نهایی تخمین و در جدول ۱۱ آزمون معناداری ضرایب متغیرهای نهایی ارائه شده است. بدین ترتیب خط ممیزی ایجاد شده توسط این تحلیل عبارت است از:

$$De\ default = 2,046 \times HOUSE - 1,059 \times MARRIAGE + 0,040 \times COLLATERAL - TYPE - 1,011 \quad (19)$$

با وارد کردن مقادیر وضعیت تأهل، وضعیت اقامت و نوع قرارداد در این تابع، اگر مقادیر به دست آمده به‌ازای پیش‌فرض منفی بود، به گروه مشتریان خوش حساب و در صورت مثبت بودن به گروه مشتریان بدحساب تعلق می‌گیرد.

۲.۸. جدول پیش‌بینی مورد انتظار

در جدول ۱۲ پیش‌بینی مورد انتظار مدل تحلیل ممیزی ارائه شده است.

۹. تخمین شبکه‌های عصبی با الگوریتم پس‌انتشار خطا

در این پژوهش، شبکه‌ی MLP با استفاده از نرم‌افزار Clementine تخمین زده شده است. معماری شبکه‌ی تخمین زده شده توسط این نرم‌افزار دارای ۹ نورون

جدول ۳. اثر نهایی متغیرهای نهایی.

متغیر	اثر نهایی	ضرایب
وضعیت سکونت	۱,۰۱۷۹۸۳	۱,۹۶۱۳۰۲
تعداد اقساط	۰,۰۲۵۵۴۷	۰,۰۴۹۲۲
وضعیت اشتغال	-۰,۰۵۴۲۰۸	-۱,۰۴۴۳۹۴
وام/وثیقه	-۰,۰۸۵۷۲۶	-۱,۰۶۵۱۶۴۳
وضعیت تأهل	-۰,۰۶۵۰۲۷	-۱,۰۲۵۲۸۴۹

جدول ۴. خروجی آزمون هاسمر - لموشوف.

آزمون نیکویی برازش	
تعداد مشاهدات	۲۰۰
تعداد گروه‌ها	۱۰
مقدار آماره هاسمر - لموشوف	۵,۹۲۷۷۳۴۳۸۸۶

جدول ۵. خروجی پیش‌بینی مورد انتظار.

کل	مشتریان		
	بد	خوب	
۱۰۵	۲۳	۸۲	$P(DEP = 1) \leq C$
۹۵	۷۲	۲۳	$P(DEP = 1) > C$
۲۰۰	۹۵	۱۰۵	کل
۱۵۴	۷۲	۸۲	صحیح
۷۵,۴۹	۷۲,۷۳	۷۸,۱۰	درصد صحیح
۲۴,۵۱	۲۷,۲۷	۲۱,۹۰	درصد ناصحیح

در O_z ، Z اعداد مشاهده شده در گروه است؛ در N_z ، Z تعداد کل مشاهدات در گروه است؛ در E_z ، Z تعداد پیش‌بینی شده در گروه است.

در این آماره، تعداد ۱۰ گروه در نظر گرفته شده است. مقدار آماره هاسمر - لموشوف طبق جدول ۴ معادل ۵,۹۳ و از ۱۶ کم تر است. این نشان می‌دهد که نیکویی برازش با سطح اطمینان ۹۵ درصد تأیید شده است.

۴.۶. جدول پیش‌بینی مورد انتظار

این جدول طبقه‌بندی، درست و نادرست را براساس یک قاعده پیش‌بینی که کاربر تعیین می‌کند، نشان می‌دهد. در این تحقیق، مقدار برش ۰/۵ در نظر گرفته شده است. نتایج این تحلیل در جدول ۵ قابل مشاهده است.

۷. مدل پروبیت

همانند خروجی آزمون هم‌خطی متغیرها (جدول ۱)، در مرحله‌ی بعدی به مدل بهینه‌ی اعتبارسنجی کلیدی متغیرها را وارد می‌کنیم و سپس با استفاده از معناداری ضرایب رگرسیونی، متغیرهای نهایی اثرگذار بر وضعیت اعتباری مشتریان انتخاب می‌شود. در نهایت، ۵ متغیر به‌عنوان متغیرهای اثرگذار بر وضعیت اعتباری مشتریان انتخاب شدند. جدول ۶ شکل نهایی تخمین را نشان می‌دهد.

رابطه تخمین زده به‌ازای شاخص وقوع قصور با استفاده از مدل پروبیت به‌شکل

جدول ۶. تخمین نهایی.

آزمون ضریب	آماره Z	متغیر	خطای انحراف معیار	
۰٫۰۰۰	۵٫۵۱۶۶۴۰	۰٫۲۱۱۴۵۸	-۰٫۹۸۶۰۱۴	وضعیت سکونت
۰٫۰۰۰۸۲	۲٫۶۴۵۶۰۴	۰٫۰۱۱۱۶۱	۰٫۷۳۶۱۵۷	تعداد اقساط
۰٫۰۰۰۲۱	-۳٫۰۷۰۶۰۵	۰٫۲۰۲۳۴۱	۰٫۵۷۸۴۳۱	وضعیت اشتغال
۰٫۰۰۰۹۹	-۲٫۵۷۹۴۹۰	۰٫۳۸۲۲۵۱	-۱٫۶۵۱۶۴۳	وام/وثیقه
۰٫۰۰۰۰۵	-۳٫۴۹۷۲۱۵	۰٫۲۱۰۴۹۸	-۱٫۲۵۲۸۴۹	وضعیت تأهل
۰٫۱۴۱۳	۱٫۴۷۱۰۰۹	۰٫۳۹۳۲۲۱	۱٫۰۰۳۸۸۲	ضریب ثابت
۰٫۴۸۵۲۹۴	میانگین متغیر وابسته خودرگرسیون برداری	۰٫۳۰۶۷۷۴		میانگین وابسته
۰٫۴۰۵۶۴۴	انحراف معیار یا خطای استاندارد میانگین	۰٫۵۰۱۰۱۳		انحراف معیار یا خطای استاندارد میانگین
۳۲٫۵۸۰۳۲	مجموع مازاد مجذور	۱٫۰۱۹۲۳۹		مجموع مازاد مجذور
-۹۷٫۹۶۲۴۲	لگاریتم درست‌نمایی	۱٫۱۱۶۸۳۱		لگاریتم درست‌نمایی
-۱۴۱٫۳۱۳۸	بهینه‌سازی لگاریتم تابع درست‌نمایی	۱٫۰۵۸۷۱۷		بهینه‌سازی لگاریتم تابع درست‌نمایی
۰٫۴۸۰۲۰۸	میانگین لگاریتم درست‌نمایی	۸۶٫۷۰۲۷۲		آماره (۵ درجه آزادی) ضرایب متغیرها
۲۰۰		۰٫۰۰۰۰۰۰		قانون نسبت احتمال
		Total obs=۱۰۵		Obs With Dep=۰
		۹۵		Obs With Dep=۱

متغیر وابسته: پیش فرض،

روش به‌کارگرفته شده: مدل لاجیت،

تاریخ ۱۳۹۱/۰۷/۱۱ ساعت: ۱۵:۱۹

تعداد نمونه: ۲۰۰ و تعداد مشاهدات: ۲۰۰ مورد

هم‌گرایی پس از ۵ تکرار به‌دست آمده و ماتریس کواریانس با استفاده از مشتقات دوم محاسبه شده است.

جدول ۹. خروجی پیش‌بینی مورد انتظار.

کل	مشتریان		
	خوب	بد	
۱۱۰	۸۲	۲۸	$P(DEP = 1) \leq C$
۹۰	۲۱	۶۹	$P(DEP = 1) > C$
۲۰۰	۱۰۵	۹۵	کل
۱۵۴	۸۲	۷۲	صحیح
۷۵٫۴۹	۷۹٫۰۵	۷۱٫۷۲	درصد صحیح
۲۴٫۵۱	۲۰٫۹۵	۲۸٫۲۸	درصد ناصحیح

جدول ۱۰. ضرایب متغیرهای اثرگذار میانگین‌های گروهی آزمون‌های برابری.

تابع	
۱	
وام/وثیقه	۰٫۴۰
وضعیت تأهل	-۱٫۰۵۹
وضعیت سکونت	۲٫۰۴۶
مقدار ثابت	-۱٫۰۱۱

جدول ۷. اثر نهایی متغیرهای نهایی.

متغیر	اثر نهایی	ضرایب
وضعیت سکونت	۰٫۴۶۴۶۲۱	۱٫۹۶۱۳۰۲
تعداد اقساط	۰٫۰۱۱۷۶۱	۰٫۰۴۹۲۲
وضعیت اشتغال	-۰٫۲۴۷۴۶	-۱٫۰۴۴۳۹۴
وام/وثیقه	-۰٫۳۹۲۷۲	-۱٫۶۵۱۶۴۳
وضعیت تأهل	-۰٫۲۹۳۲	-۱٫۲۵۲۸۴۹

جدول ۸. آزمون نیکویی برازش.

آزمون نیکویی برازش	
تعداد مشاهدات	۲۰۰
تعداد گروه‌ها	۱۰
مقدار آماره‌ها سمر - لموشوف	۵٫۳۷۳۲۹۲۰ ۱۷۹۹

جدول ۱۱. آزمون معنی داری ضرایب متغیرهای نهایی.

ویلیکس لامبادا	فیشر	یک درجه آزادی	دو درجه آزادی	Σ
وام/وثیقه	۰/۹۶۱	۸/۱۰۹	۲۰۲	۰/۰۰۵
وضعیت تأهل	۰/۹۵۰	۱۰/۶۳۹	۲۰۲	۰/۰۰۱
وضعیت سکونت	۰/۷۷۹	۵۷/۲۲۸۳	۲۰۲	۰/۰۰۰

جدول ۱۲. خروجی پیش بینی مورد انتظار.

گروه‌های پیش‌بینی شده				
مقادیر پیش‌بینی شده	مشتریان			کل
	خوب	بد	مشتریان خوب	
مشتریان بد	۲۵	۷۰	۹۵	۱۰۵
مشتریان خوب	۸۱	۲۴	۱۰۵	۱۰۰
مشتریان بد	۲۵	۷۰	۹۵	۱۰۰
مشتریان خوب	۷۷/۱	۲۲/۹	۱۰۰	۱۰۰
مشتریان بد	۳۰/۳	۶۹/۷	۱۰۰	۱۰۰

جدول ۱۳. ارزش نسبی شبکه MLP.

متغیر	ارزش نسبی
وضعیت مسکونی	۰/۳۴۷۹
وضعیت اشتغال	۰/۳۱۳۶۹۱
وضعیت تأهل	۰/۲۰۵۴۴۹
نرخ بهره	۰/۱۰۹۸۶۵
نوع وثایق	۰/۱۰۱۹۹۹
جنسیت	۰/۰۹۶۹۹۴۳
سن	۰/۰۴۵۲۵۶
تعداد اقساط	۰/۰۴۱۴۷۵۵
نسبت وام به وثایق	۰/۰۲۳۹۵۵۷

ورودی و ۱ نورون خروجی است. دو لایه میانی در این شبکه وجود دارد. در لایه اول ۲۰ نورون، و در لایه دوم ۱۵ نورون وجود دارد. تابع انتقال در لایه اول لوجستیک، در لایه میانی تابع تانژانت هیپربولیک و در لایه خروجی خطی است. ارزش نسبی هر یک از متغیرهای این شبکه در جدول ۱۳ آمده است. میزان صحت پیش‌بینی‌های این شبکه در جدول ۱۴ ارائه شده است.

۱۰. تخمین شبکه‌های عصبی با الگوریتم تابع شعاع محور (RBN)

ارزش نسبی هر یک از متغیرها در این شبکه در جدول ۱۵ ارائه شده است. میزان صحت پیش‌بینی‌های این شبکه نیز در جدول ۱۶ آمده است.

۱۱. تخمین شبکه‌های عصبی با الگوریتم GMDH

شبکه‌ی عصبی تخمین زده شده دارای ۲ لایه میانی است. در این تخمین، میزان خطای مشاهده شده برای ۳۰ درصد مشاهدات ۱۵/۹۶ درصد و میزان خطای مشاهده شده برای ۷۰ درصد مشاهدات ۸/۵۰ درصد است. وضعیت ترتیب

جدول ۱۴. صحت پیش‌بینی.

MLP		
MDA	تجزیه و تحلیل چندگانه	۰/۳۰۶۵۳۳
RMSE	میانگین مجذور مربعات خطا	۰/۳۶۱۴۲۷
MSE	میانگین مربعات خطا	۰/۱۳۰۶۲۹
Accuracy	اصل صحت و دقت	۰/۸۳۳۳۳۳
MDA	تجزیه و تحلیل چندگانه	۰/۳۵۸۲۵۵
RMSE	میانگین مجذور مربعات خطا	۰/۴۲۸۷۷۷
MSE	میانگین مربعات خطا	۰/۱۸۳۸۵
Accuracy	اصل صحت و دقت	۰/۷۲۹۱۶۷
MDA	تجزیه و تحلیل چندگانه	۰/۳۴۳۰۴۲
RMSE	میانگین مجذور مربعات خطا	۰/۴۱۰۱۱۸
MSE	میانگین مربعات خطا	۰/۱۶۸۱۹۷
Accuracy	اصل صحت و دقت	۰/۷۵۹۸۰۴

جدول ۱۵. ارزش نسبی شبکه RBN.

متغیر	ارزش نسبی
وضعیت مسکونی	۰/۳۵۰۸۷۱
وضعیت تأهل	۰/۳۳۷۱۸۷
وضعیت اشتغال	۰/۱۱۹۱۳۲
نوع وثایق	۰/۱۰۸۵۲۳
نسبت وام به وثایق	۰/۰۹۲۰۲۷۸
جنسیت	۰/۰۶۵۰۳۱۵
نرخ بهره	۰/۰۶۴۴۸۶۳
تعداد اقساط	۰/۰۶۰۲۲۵۹
سن	۰/۰۴۸۷۸۶۲

جدول ۱۶. صحت پیش‌بینی.

RBN		
MDA	تجزیه و تحلیل چندگانه	۰/۳۸۸۳۳۲
RMSE	میانگین مجذور مربعات خطا	۰/۴۶۴۳۹
MSE	میانگین مربعات خطا	۰/۲۱۵۶۵۸
Accuracy	اصل صحت و دقت	٪۰/۶۲۰۵
MDA	تجزیه و تحلیل چندگانه	۰/۲۴۳۱۹۲
RMSE	میانگین مجذور مربعات خطا	۰/۳۳۴۸۰۸
MSE	میانگین مربعات خطا	۰/۱۱۲۰۹۶
Accuracy	اصل صحت و دقت	٪۰/۹۰
MDA	تجزیه و تحلیل چندگانه	۰/۳۴۴۴۵۶
RMSE	میانگین مجذور مربعات خطا	۰/۴۳۰۳۴۷
MSE	میانگین مربعات خطا	۰/۱۸۵۱۹۸
Accuracy	اصل صحت و دقت	٪۷۰/۵۹

جدول ۱۷. اهمیت نسبی متغیرها.

GMDH
وضعیت مسکونی
نوع وثایق
تعداد اقساط
وضعیت اشتغال
نرخ بهره

جدول ۱۹. مقایسه‌ی دقت مدل‌های مورد استفاده در پژوهش.

روش مورد استفاده	دقت پیش‌بینی (درصد)
GMDH	۹۰
RBF	۹۰
MLP	۷۲٫۹۲
MDA	۷۳٫۵
Probit	۶۵
Logit	۶۳٫۳۳

جدول ۱۸. خطای مقادیر پیش‌بینی.

GMDH		
MDA	تجزیه و تحلیل چندگانه	۰٫۳۲۹۳۴۵
RMSE	میانگین مجذور مربعات خطا	۰٫۴۱۲۳۴۸
MSE	میانگین مربعات خطا	۰٫۱۰۰۳۱
Accuracy	اصل صحت و دقت	٪۷۳٫۶۱
MDA	تجزیه و تحلیل چندگانه	۰٫۲۰۸۳۵۹
RMSE	میانگین مجذور مربعات خطا	۰٫۲۸۴۱۰۷
MSE	میانگین مربعات خطا	۰٫۰۸۰۷۱۷
Accuracy	اصل صحت و دقت	٪۵۰٫۹۰
MDA	تجزیه و تحلیل چندگانه	۰٫۲۹۳۷۶۱
RMSE	میانگین مجذور مربعات خطا	۰٫۳۷۹۱۶
MSE	میانگین مربعات خطا	۰٫۱۴۳۷۶۲
Accuracy	اصل صحت و دقت	٪۷۸٫۴۳

اقتصادسنجی مورد استفاده شامل مدل‌های لاجیت و پروبیت است و از سوی دیگر، از شبکه‌های عصبی GMDH، MLP و RBFN به‌منظور امتیازدهی و رتبه‌بندی مشتریان استفاده شده است. نتایج به‌دست آمده از این پژوهش در جدول ۱۹ قابل مشاهده است. باید توجه داشت این مقادیر بر اساس ۶۰ مشاهده‌ی انتهایی، به‌عنوان نمونه‌ی آزمایشی ارائه شده است.

چنان‌که مشاهده می‌شود روش GMDH و RBN دارای بالاترین سطح دقت، و مدل لاجیت دارای پایین‌ترین سطح دقت است. همچنین داده‌های این جدول نشان‌دهنده‌ی کارایی بالای شبکه‌های عصبی GMDH و RBN در مقایسه با دیگر روش‌هاست. از سوی دیگر، شبکه‌های عصبی MLP به‌لحاظ دقت در پیش‌بینی دارای رتبه‌ی دوم هستند. البته از محدودیت‌های پژوهش می‌توان به عدم وجود اکثر متغیرهای شناسایی شده، به‌عنوان متغیرهای اثرگذار بر رفتار اعتباری مشتریان در پایگاه داده‌های اطلاعات مشتریان اشاره کرد.

پیشنهادات برای پژوهش‌های آتی عبارت است از:

-- استفاده از شبکه‌های عصبی فازی.

-- استخراج مدل برای مشتریان حقوقی.

-- استفاده از الگوریتم ژنتیک در ارزیابی مشتریان.

-- استخراج مدل برای مشتریان فعال در یک صنعت خاص.

-- استفاده از مدل شبکه‌های عصبی در رتبه‌بندی اعتباری دیگر بانک‌ها، تا به توانایی روش‌های مذکور اضافه شود و شبکه‌ی عصبی برتر قابلیت اعتماد بیشتری به دست آورد.

اهمیت نسبی این متغیرها در جدول ۱۷ ارائه شده است. مقادیر خطای پیش‌بینی در جدول ۱۸ آمده است.

۱۲. نتیجه‌گیری و پیشنهادات

در این پژوهش با توجه به اهمیت رتبه‌بندی اعتباری مشتریان بانکی، سعی شده تا روش‌های آماری و اقتصادسنجی با روش‌های داده‌کاوی مقایسه شود. روش‌های

پانوشته‌ها

1. Perceptron
2. Compusstat
3. Salcheen Berger
4. mixture of experts
5. radial basis function
6. fuzzy adaptive resonance
7. learning vector quantization
8. Kernal density estimation
9. probabilistic neural nets
10. multi-layer feed-forward nets
11. best net search
12. standard & poors (S&P)
13. supervised learning
14. target output
15. back-propagation

16. topology of network
17. Neuron
18. credit scoring
19. A.G. Ivakhnenko
20. self-organizing
21. evolutionary
22. volterra functional series
22. multi-layer feed forward (MLFF)
23. gradient descent
24. cumulative distribution function
25. Hosmer-Lemeshow

منابع (References)

1. Leea, T.-S., Chiu, C.-C., Lu, C.-J. and Chen, I-F. "Credit scoring using the hybrid neural discriminant

- technique”, *Expert System Applications*, **23**(3), pp. 245-254 (2002).
2. Sullivan, A.C., *Consumer Finance*, In E.I.Altman (Ed), Financial Handbook. New York: John Wiley & Sons (1981).
 3. Bailey, M. *Costumer Credit Quality, Underwriting, Scoring, Fraud Prevention and Collections*, Kingswood, Bristol: White Box Publishing (2004).
 4. Coats, P.K. and Fant, L.F. “Recognizing financial distress patterns using a neural network tool”, *Financial Management*, **22**(3), pp. 142-155 (Autumn 1993).
 5. Lacher, R.C., Coast, P.K., Sharma, S.C. and Fant, L.F. “A neural network for classifying the financial health of a firm”, *European Journal of Operational Research*, **85**(1), pp. 53-65 (1995).
 6. West, D. “Neural network credit scoring models”, *Computer and Operation Research*, **27**(11-12), pp. 1131-1152 (2000).
 7. Howland, J.C. and Voss, M.S. “Natural gas prediction using the group method of data handling”, ASC (2003).
 8. Jaiswal, S. “Relationship between asset and liability of commercial banks in India, 1997-2008”, *International Research Journal of Finance and Economics*, **49** pp. 43-58 (2010).
 9. Mandala, I.G.N.N, Nawangpalupi, C.B. and Praktikto, F.R. “Assessing credit risk: An application of data mining in a rural bank”, *Procedia Economics and Finance*, **4**, pp. 406-412 (2012).
 10. Castro, V. “Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI”, *Economic Modeling*, **31**, 672-683 (2014).
 11. Da Silva, M.S. and Divino, J.A. “The role of banking regulation in an economy under credit risk and liquidity shock”, *North American Journal of Economics and Finance*, **26**, pp. 266-281 (2013)
 12. Chen, M.-C. and Huang, S.-H. “Credit scoring and rejected instances reassigning through evolutionary computation techniques”, *Expert System with Applications*, **24**(4), pp. 433-441 (2003).
 13. Shu-Heng, C., *Genetic Algorithm and Genetic Programming in Computational Finance*, Springer Kluwer Academic Publications, New York, USA: pp. 90-102 (2002).
 14. Pointon, J., El-Mesry, A. and Cramer, J.S., *The LOGIT Model: An Introduction for Economists*, Edward Arnold (1991).
 15. Institute of Iran Banking “Report on the performance of banks in 2010”, *the Twenty-Second International Conference on Islamic Banking* (2011).
 16. Bill, R. and Jackson, T., *Introduction to Neural Networks*, Translated by Doctor Alborzi.M, Academic Press and Sharif University, Second edition (2004).
 17. Mahdavi, G. and Behmanesh, M. “The investment company’s stock price prediction model using artificial neural networks; Case study: Investment company”, *Alborz, Economic Journal*, pp. 211-233 (2005).
 18. Jalili, M., Moshiri, S. and Morovvat, H. “Prediction of Tehran stock index returns using linear and nonlinear models”, *Journal of Business Research*, **41**, (2006).
 19. Panya, A.S., Kondo, T., Shah, T.U. and Gandhi, V.R. “Prediction of stock market characteristics uses neural networks in proceeding of the SPIE”, *The International Society for Optical Engineering*, **3722**, pp. 189-197 (1999).
 20. Saadat, A., Alborzi, M. and Yaghoub Nehad, A. and Maghsoud, H. “Application of artificial neural networks in predicting the financial performance and stock prices”, *Journal of Accounting Research*, **22**, pp. 119-137 (2009).
 21. Huang, J.-J., Tzeng, G.-H., Ong, C.-S. “Two-stage genetic programming (2SGP) for the credit scoring model”, *Applied Mathematics and Computation*, **174**(2), pp. 1039-1053 (2006).
 22. Lemke, F. and Muller, J.A. “Self-organizing modeling in financial risk control”, *Proceeding of the 15th IMACS Word Congress on Scientific Computation, Modeling and Application Mathematics*, **6**, pp. 723-738 (1997).
 23. Kimiagari, M.A and Colleagues, “Repayment of credit facilities to customers’ credit risk model (scope shahreza Iran Melli Bank)”, *The Ninth International Conference on Industrial Engineering* (2012).
 24. Karimkhani, A. and Forati, M. “Effect of macroeconomic variables on the sources and uses of banks, research management and risk control”, *Iran Sepah Bank* (2012).
 25. Menjomee, A., Abzari, M. and Rayati Shoazi, A. “Prediction of stock prices in the stock market, using fuzzy neural networks and genetic algorithms compared with artificial neural network”, *Journal of Quantitative Economics*, **8**(III), pp. 1-26 (2009).
 26. Zarenejad, M., Feghhe Majidi, A. and Rezaee, R. “Exchange rate forecasting model using artificial neural networks and ARIMA Model”, *Quarterly Journal of Economics*, **5**(4), pp. 107-130 (2008).
 27. Ivakhnenko, G.A. and Muller, J.A. “Recent development of self-organizing modeling in prediction and analysis of stock market”, *Microelectron. Reliab Journal*, **37**, pp. 1053-1072 (1996)
 28. Ivakhnenko, G.A. “The review of problems solvable by algorithms of the method of data handling (GMDH)”, *Pattern Recognition and Image Analysis*, **5**(4), pp. 527-535 (1995).
 29. Mehrara, M., Behradmehr, N., Ahrari, M. and Moshagheh, M. “Prediction of volatility of crude oil prices using neural networks GMDH”, *Energy Economics Studies*, **25**, pp. 89-112 (2010).
 30. Mehrara, M., Moeeni, A., Ahrari, M. and Hamooni, A. “Modeling and forecasting Tehran stock exchange and determine variables”, *Quarterly Economic Bulletin*, **50**, pp. 31-51 (2009).
 31. Jamali, A., Nariman-zadeh, N. and Atashkari, K. “Inverse modeling of multi-objective thermodynamically optimized and GAI 14th annual (international)”, *Mechanical Engineering Conference*, Isfahan University of Technology, Isfahan, Iran (2006).
 32. Zamani, N. “Analysis, comparison and forecasts crude oil sales revenue using neural networks”, MS Thesis, Tehran University (2009).

33. Scott, D.E. and Hutchinson, C.E. "The GMDH algorithm – a technique for economic modeling", *Modeling and Simulation*, **7**, pp. 729-733 (1999).
34. Nofarasti, M., *Unit Root and Co-Integration in Econometrics*, Published by Institute Rasa Cultural Services (2008).
35. Water, P.R., Wibier, E.J., Kerckhoffs, H. and Koppelaar, H. "GMDH-based stock price prediction", *Neural Network Word*, **7**(5), pp. 552-563 (1997).
36. Scott, D.E. and Hutchinson, C.E. "The GMDH algorithm- a technique for economic modeling", *Proc. 7th Pittsburgh Conf. Modeling Simul.*, University of Massachusetts (1976).
37. Abbasinejad, H., *Advanced Econometrics*, Tehran University Publications (2005).
38. Abrishami, H., Mehrara, M., Ahrari, M. and Mirghasemi, S. "Modeling and forecasting economic growth of the neural network approach GMDH", *Economic Research Journal*, **88**, pp. 1-24 (2009).
39. Abrishmi, H., Moeeni, A., Mehrara, M., Ahrari, M. and Soleimanikia, F. "Modeling and forecasting gasoline prices using neural network GMDH", *Journal of Economic Studies*, **XII**(36), pp. 37-58 (2008).
40. Sharzeie, G., Ahrari, M. and Fakhraee, H. "Water demand prediction in Tehran by using structural patterns, time series and neural network type of GMDH", *Economic Research Journal*, **84**, pp. 151-175 (2008).
41. Salchen Berger, L.M., Cinar, E.M. and Lash, N.A. "Neural network: A new tool for predicting thrift failures", *Decision Sciences*, **23**(4), pp. 899-916 (1992).
42. Goleusov, I.V. and Kondrasheva, S.A. "Comparative analysis of the interdependence structure if macroeconomic indices of COMECON member countries by the group method of data handling", *Soviet Journal of Automation and Information Sciences c/c of Automatika*, **20**(3), pp. 39-43 (1997).
43. Srinivasan, N., Ravichandran, V., Chan, K.L., Vidhya, J.R., Ramakirishnan, S. and Krishnan, S.M. "Exponentiated back propagation algorithm for multilayer feedforward neural networks", *Neural Information Processing. ICONIP apos, 02. Proceedings of the 9th International Conference on* (2002).
44. Lippmann, R. "An introduction to computing with neural nets", *IEEE Assp Magazine*, **4**(2), pp. 4-22 (1987).
45. Grene, W. "Marginal effects in the censored regression model", *Economic letters*, **64**, pp. 43-49 (1998).