Archive of SID

نشریه علمی ـ پژوهشی
روش های عددی در مهندسی

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

ارائه یک مدل طبقه‌بندی ترکیبی هوشمند مبتنی بر شبکه‌های عصبی پرسپترون چندلایه و رگرسیون فازی به‌منظور تجزیه و تحلیل مسائل امتیازدهی اعتباری

مهدی خاشعی و شیدا تربت*

دانشکده مهندسی صنایع و برنامه‌ریزی سیستم‌ها، دانشگاه صنعتی اصفهان، اصفهان

(دریافت مقاله: ۱۳۹۵/۱۰/۲۹ – دریافت نسخه نهایی: ۱۳۹۶/۰۴/۱۸)

چکیده– بحران‌های مالی موجود در نظام‌های بانکی ناشی از عدم توانایی در مدیریت ریسک‌های اعتباری است. امتیازدهی اعتباری یکی از تکنیک‌های مدیریت ریسک است که ریسک وام‌گیرنده را تحلیل می‌کند. در این مقاله با استفاده از مزایای روش‌های هوش محاسباتی و محاسبات نرم یک روش ترکیبی جدید به‌منظور بهبود مدیریت ریسک‌های اعتباری ارائه شده است. در روش پیشنهادی، به‌منظور مدل‌سازی در شرایط عدم قطعیت، پارامترهای شبکه عصبی، شامل وزن‌ها و خطاها، به‌صورت فازی درنظر گرفته شده‌اند. در این روش، ابتدا سیستم مورد مطالعه با استفاده از شبکه‌های عصبی متامدل‌بندی شده و سپس با به‌کارگیری استنتاجات فازی تصمیم بهینه با بیشترین میزان برتری تعیین خواهد شد. نتایج حاصل از به‌کارگیری روش پیشنهادی بیانگر کارامدی و دقت بالای این روش در تحلیل مسائل امتیازدهی اعتباری است.

واژه‌های کلیدی: امتیازدهی اعتباری، روش‌های طبقه‌بندی، پرسپترون‌های چندلایه، شبکه‌های عصبی مصنوعی، منطق فازی.

A Hybrid Intelligent Classification Model Based on Multilayer Perceptron Neural Networks and Fuzzy Regression for Credit Scoring Problems

M. Khashei and Sh. Torbat*

Department of Industrial Engineering, Isfahan University of Technology, Isfshsn, Iran.

Abstract: ***Financial crises in banking systems are due to inability to manage credit risks. Credit scoring is one of the risk management techniques that analyze the borrower's risk. In this paper, using the advantages of computational intelligence as well as soft computing methods, a new hybrid approach is proposed in order to improve credit risk management. In the proposed method, for modeling in uncertainty conditions, parameters of the neural network, including weights and errors, are considered***

* : مسئول مکاتبات، پست الکترونیکی: s.torbat@in.iut.ac.ir

www.SID.ir

Archive of SID

in the form of fuzzy numbers. In this method, the underlying system is firstly modeled using neural networks and then, using fuzzy inferences, the optimal decision will be determined with the highest degree of superiority. Empirical results of using the proposed method indicate the efficiency and high accuracy of this method in analyzing credit rating problems.

Keywords: *Credit scoring, Classification methods, Multilayer perceptrons (MLPs), Artificial neural networks, Fuzzy logic.*

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

فهرست علائم

b_0	بایاس شبکه عصبی مربوط به نورون خروجی	$\tilde{w}_{i,j}$	وزن فازی شبکه عصبی از نورون i به نورون j
$b_{0,j}$	بایاس شبکه عصبی مربوط به نورون میانی jام	$x_{i,t}$	متغیر توضیح‌دهنده iام در زمان t
$F(x)$	تابع فعال‌سازی شبکه عصبی در لایه خروجی	y_{t-i}	وقفه iام متغیر هدف در زمان t
$f(x)$	تابع چگالی احتمال مربوط به متغیر x	y_t	متغیر هدف در زمان t
$G(x)$	تابع فعال‌سازی شبکه عصبی در لایه میانی	α_i	مرکز عدد فازی مربوط به پارامتر iام
$h(x)$	سطح آستانه مربوط به تابع عضویت متغیر x	β_i	گسترش عدد فازی مربوط به پارامتر iام
$L(x)$	تابع فازی نوع-ال، دابیوس و پرایس برای متغیر x	ε_t	خطای خالص مدل‌سازی در زمان t
P_A	احتمال عضویت یک داده به طبقه Aام	$\mu_{\tilde{w}}(w_{i,j})$	تابع عضویت فازی مربوط به اوزان شبکه عصبی
$w_{i,j}$	وزن اتصالی شبکه عصبی از نورون i به نورون j		

۱- مقدمه

امتیازدهی اعتباری یک تکنیک مدیریت ریسک اعتباری است که ریسک وام‌گیرنده را تحلیل می‌کند. در معنای اولیه آن، امتیازات اعتباری به هر مشتری منتسب می‌شوند تا سطح ریسک آن را نشان دهند. یک مدل مناسب در حوزه امتیازدهی اعتباری باید قدرت تفکیک میان مشتریان را داشته باشد. این‌گونه ازمدل‌ها، به‌طور اساسی متقاضیان را به دو گروه اعتباری خوب و بد تقسیم می‌کند. گروه اعتباری خوب، گروهی هستند که دیون خود را به موقع بازپرداخت می‌کنند و گروه اعتباری بد، گروهی هستند که به احتمال مشخص، دیون آنها نکول خواهد شد [۱]. در گذشته امتیازدهی اعتباری بر اندازه‌گیری این ریسک متمرکز بود که یک مشتری به تعهدات مالی‌اش عمل نکند و پرداخت معوق داشته باشد. به تازگی، امتیازدهی اعتباری، همین‌طور ریسک زیان و اثر کلی ریسک را نیز دربرمی‌گیرد. تکنیک‌های امتیازدهی، امروزه در سراسر چرخه عمر یک اعتبار به‌عنوان ابزار حامی تصمیم‌گیری یا الگوریتم تصمیم‌گیری اتوماتیک برای طیف وسیعی از مشتریان مورد استفاده قرار می‌گیرند. رقابت فزاینده، کانال‌های فروش الکترونیک و مقررات بانکی فعلی کاتالیزورهای مهمی برای کاربرد سیستم‌های امتیازدهی اتوماتیک بوده‌اند.

امتیازدهی اعتباری نخستین بار همچون ابزاری دم‌دستی برای متعهد ساختن اعتبارات خرد مانند وام‌های رهن مسکن، کارت‌های اعتباری، وام‌های قسطی و اعتبارات تجاری خرد معرفی شد. این مدل‌ها امروزه برای نظارت و هدایت ریسک نکول در کل پرتفوی اعتباری یک نهاد مالی شامل بانک‌ها، بنگاه‌ها، حکومت‌ها، تأمین مالی پروژه‌ها، مؤسسات مالی و غیره مورد استفاده قرار می‌گیرد. مدل‌های امتیازدهی اعتباری صرفاً برای تصویب اعتبار به‌کار نمی‌روند، بلکه در حوزه‌های دیگر مانند قیمت‌گذاری، ذخیره‌سازی، محاسبه سرمایه نظارتی و اوراق بهادار نیز مورد استفاده قرار می‌گیرند. با توجه به موفقیت گسترده کاربرد امتیازدهی در محیط اعتباری امروزی، مؤسسات مالی نیز استفاده از تکنیک‌های امتیازدهی را بر مبنای طیفی از سایر اهداف تجاری شروع کرده‌اند. برنامه امتیازدهی، ریسک وام‌گیرنده را در لحظه اعطای اعتبار ارزیابی

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

می‌کند. وقتی وامی نکول می‌شود، سیستم امتیازدهی برای پشتیبانی از امتیازات وصول شده استفاده و پارامترهای ریسک را به‌روزرسانی می‌کند [۲].

در ادبیات موضوع مرتبط با روش‌های امتیازدهی اعتباری، روش‌های آماری مختلفی از جمله رگرسیون لجستیک [۳ و ۴]، تحلیل جدایش خطی [۵ و ۶]، تحلیل جدایش درجه دوم [۷] و شبکه‌های بیزین [۸ و ۹] به‌منظور ارزیابی مسائل امتیازدهی اعتباری پیشنهاد شده است. روش‌های هوش مصنوعی مانند شبکه‌های عصبی مصنوعی [۱۰ و ۱۱]، ماشین‌های بردار پشتیبان [۱۲]، درخت تصمیم [۱۳] و الگوریتم ژنتیک [۱۴]، به‌عنوان گروهی دیگر از مدل‌های طبقه‌بندی که عملکرد موفقیت‌آمیزی در تحلیل مسائل امتیازدهی اعتباری داشته‌اند، پیشنهاد شده است. اگرچه روش‌های آماری و هوش مصنوعی، روش‌هایی دقیق و کارا محسوب می‌شوند اما هر دو، مدل‌هایی قطعی به‌شمار می‌آیند که در فرایند مدل‌سازی آنها منطق کلاسیک به‌کار گرفته شده است. بنابراین در شرایطی که با پیچیدگی و عدم‌قطعیت مواجه هستیم مدل‌های کارایی نخواهند بود [۱۵]. از این‌رو در چنین شرایطی مدل‌های پیش‌بینی و طبقه‌بندی فازی از جمله سری‌های زمانی فازی و مدل رگرسیون خطی فازی می‌توانند مناسب باشند، هرچند که کارایی این‌گونه از مدل‌ها نیز همیشه رضایت‌بخش نخواهد بود [۱۶].

بسیاری از محققان بر این باورند که ترکیب مدل‌ها می‌تواند منجر به بهبود کارایی شود. بر این اساس در پژوهش‌های پیشین به‌منظور فائق آمدن بر محدودیت‌های مدل‌های واحد و بهبود کارایی مدل‌های طبقه‌بندی، استفاده از مدل‌های ترکیبی امری رایج و متداول به‌شمار می‌آید. هدف از ترکیب مدل‌ها کاهش میزان ریسک ناشی از استفاده از مدل‌های نامناسب و کسب نتایج کارا است. به‌عبارت دیگر انگیزه اصلی در به‌کارگیری مدل‌های ترکیبی، نتیجه این موضوع است که هر یک از مدل‌هایی که در اینجا گفته شد به‌تنهایی توانایی کافی در تشخیص فرایند تولید داده و تشخیص ویژگی‌های داده‌ها را ندارد [۱۷]. در سال‌های اخیر مدل‌های طبقه‌بندی ترکیبی بسیاری به‌منظور افزایش دقت و کارایی پیشنهاد شده‌اند که در بخش بعدی به‌طور مختصر به برخی از آنها اشاره خواهد شد.

هدف این مقاله استفاده از مزایای اختصاصی شبکه‌های عصبی مصنوعی و استنتاج فازی به‌منظور ارائه روش ترکیبی جدید، جهت بهبود کیفیت تصمیم‌های مالی در راستای اعطای امتیازات اعتباری به مشتریان در بانک‌ها و مؤسسات مالی و اعتباری است. در روش پیشنهادی، ابتدا سیستم مورد مطالعه با استفاده از شبکه‌های عصبی متامدل‌بندی شده و سپس با به‌کارگیری استنتاج فازی تصمیم و سیاست بهینه با بیشترین میزان برتری تعیین خواهد شد. نتیجه حاصل از به‌کارگیری روش پیشنهادی بیانگر کارامدی و دقت بالای این روش در تحلیل مسائل امتیازدهی اعتباری است.

در بخش بعدی به ادبیات موضوع مربوطه به‌طور خلاصه اشاره می‌شود. در قسمت سوم خلاصه‌ای از روش شبکه‌های عصبی مصنوعی و رگرسیون فازی شرح داده شده‌اند. مدل ترکیبی شبکه‌های عصبی مصنوعی و رگرسیون فازی در بخش چهارم توضیح داده شده است. در بخش پنجم در رابطه با اطلاعات اعتباری محک کشورهای ژاپن و استرالیا که در این تحقیق مورد استفاده قرار گرفته است، توضیحاتی ارائه شده است. در بخش ششم استفاده از مدل ترکیبی در امتیازدهی اعتباری و طبقه‌بندی مشتریان به دو گروه خوش‌حساب و بدحساب و درنهایت نتیجه‌گیری آورده شده است.

۲- ادبیات موضوع

ادبیات موضوع مربوط به مدل‌های ترکیبی بسیار گسترده بوده و از زمان اولین تحقیقات انجام شده تاکنون، مطالعات فراوانی در این زمینه صورت گرفته است. در این بخش به تعدادی از مدل‌های طبقه‌بندی ترکیبی که از منطق فازی یا شبکه‌های عصبی مصنوعی در تحلیل و ارزیابی بازارهای مالی، به‌ویژه امتیازدهی اعتباری، استفاده شده است به‌طور مختصر اشاره می‌شود.

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

لـی و همکـاران [۱۸] بـا اسـتفاده از شـبکه‌هـای عصـبی پس‌انتشار و تحلیل جدایش خطی مدلی برای امتیازدهی اعتباری ارائه کردند. هسـیه [۱۹] بـا هـدف طراحـی مـدلی بهینـه بـرای امتیـازدهی اعتبـاری، یـک مـدل ترکیبـی شـامل الگـوریتم‌هـای خوشه‌بندی و شبکه‌های عصبی مصنوعی پیشـنهاد داد. هانـگ و همکاران [۲۰] با پیمودن سه مرحله به‌منظور ساختن مـدل‌هـای امتیازدهی اعتباری مبتنی بر ماشین‌های بردار پشتیبان ترکیبـی بـه ارزیابی امتیازات اعتباری مشـتریان از طریـق ویژگـی‌هـای آنهـا پرداختند. لی و همکاران [۲۱] بـا وارد کـردن ترکیبـی خطـی از توابع کرنل به‌منظور افزایش قابلیت تفسیر مدل‌های طبقـه‌بنـدی اعتباری، مـدلی جـایگزین مبتنـی بـر اسـتراتژی تکـاملی بـرای بهینه‌سازی پارامترها پیشنهاد کردند. کلی‌کیلماز و تورکسن [۲۲] سیسـتم تـابع فـازی نـوع دو را بـرای مـدل‌سـازی در شـرایط عدم‌قطعیت با استفاده از الگوریتم‌های تکاملی (ET2FF) معرفی کرده و آن را در تحلیل مسائل دنیـای واقـع ماننـد مـدل‌سـازی بازارهای مالی به‌کار گرفتند. در این مدل ابتدا، بـه‌منظـور یـافتن ساختارهای پنهان، از خوشه‌بندی فازی بهبود یافته استفاده شـده سپس برای بهینه‌سازی فواصل مجموعـه‌هـای فـازی، الگـوریتم ژنتیک به‌کار گرفته شده است. چن و لی [۲۳] یک مدل ترکیبـی مبتنی بر ماشین‌های بردار پشتیبان پیشنهاد کردنـد. در ایـن مـدل طبقه‌بندی کننده ماشین بردار پشتیبان به‌ترتیب با استفاده از مدل تحلیل جدایش خطی، درخت تصمیم، F-score و مجموعه راف به‌عنوان مراحل پـیش پـردازش بـا هـدف بهینـه‌سـازی وسـعت متغیرها از طریق حذف متغیرهای اضـافی و غیرمـرتبط، ترکیـب شده اسـت. بیـژاک و تومـاس [۲۴] یـک مـدل دو مرحلـه‌ای و همزمان پیشنهاد کردنـد. در ایـن مـدل پیشـنهادی، کـارت‌هـای اعتباری و طبقه‌بندی به‌وسیله درختـان مـدل لجسـتیک بـه‌طـور همزمان بهینه می‌شوند.

کـیم و هـان [۲۵] یـک روش ترکیبـی از نقشـه‌هـای خودسازمان‌ده (SOM) و استنتاج مبتنی بر مورد (CBR) بـرای پیش‌بینی میزان اوراق قرضه شرکتی پیشـنهاد کردنـد. پـارک و هان [۲۶] با استفاده از روش‌هـای تحلیـل سلسـله مراتبـی در استنتاج مبتنی بر مورد به‌عنوان ابزاری برای وزن‌دهی متغیرهـا، باعث بهبود عملکرد قابلیت پیش‌بینی مدل اسـتنتاج مبتنـی بـر مورد در پیش‌بینی نرخ شکست کسب و کار شدند. آهن و کیم [۲۷] با استفاده از الگوریتم ژنتیک در مدل اسـتنتاج مبتنـی بـر مورد، یک رویکرد ترکیبی جدید ارائـه و آن را در پـیش‌بینـی ورشکستگی شرکت‌های بزرگ به‌کار گرفتند. در این رویکـرد، الگوریتم ژنتیک بـه‌طـور همزمـان عملیـات بهینـه‌سـازی وزن متغیرها و انتخاب نمونه برای مدل استنتاج مبتنی بـر مـورد را انجام مـی‌دهـد. کـاپوتورتی و باربـانرا [۲۸] یـک متـدولوژی ترکیبی برای مسائل طبقه‌بندی پیشنهاد کردند. آنها از مدل خود که مبتنی بر نظریه مجموعه‌های راف، ارزیابی جزئـی احتمـال شرطی و مجموعـه‌هـای فـازی بـود بـرای تحلیـل امتیـازدهی اعتباری استفاده کردند. آکوک [۲۹] در مطالعـه‌ای بـه مقایسـه مدل‌های رایج اقتصادسنجی، شـبکه‌هـای عصـبی مصـنوعی و سیستم استنتاج عصبی- فازی تطبیقی (ANFIS) بـرای تحلیـل امتیازدهی اعتباری پرداخت. وی به این نتیجه دست یافت کـه سیستم استنتاج عصبی- فازی تطبیقی به جهت ارائـه تفسـیری دقیق از فرایند تصمیم‌گیری دارای کارایی بـالاتری نسـبت بـه شبکه عصبی مصنوعی و دیگر مدل‌هـای سـنتی اقتصادسـنجی است.

یو و همکاران [۳۰] در پژوهش خود با استفاده از یک مـدل هوشمند مبتنی بر تصمیم‌گیری گروهی فازی در حوزه پشـتیبانی تصمیم‌گیری‌های مالی چندمعیاره به ارزیابی امتیازدهی اعتبـاری پرداختند. در مـدل پیشـنهادی تعـدادی از تکنیـک‌هـای هـوش مصنوعی برای تحلیل و ارزیابی سـطح ریسـک مشـتری تحـت معیارهـای از پـیش تعیـین شـده بـه‌کـار گرفتـه شـدند. چـن و همکاران [۳۱] با استفاده از روش ترکیبی ماشین بـردار پشـتیبان به ارزیابی مسائل امتیازدهی اعتباری پرداختند. روش پیشنهادی، مبتنی بر سه استراتژی است: ۱) استفاده از دسته‌بندی و درخـت رگرسیون (CART) برای انتخاب متغیرهای ورودی، ۲) اسـتفاده از رگرسـیون چنـد متغیـره تطبیقـی (MARS) بـرای انتخـاب متغیرهـای ورودی و ۳) اسـتفاده از جسـتجوی شـبکه بـرای

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

بهینه‌سازی پارامترهای مدل. نتایج حاکی از آن است که مدل ترکیبی پیشنهادی نه‌تنها نرخ طبقه‌بندی بهتری دارد، بلکه دارای خطای نوع دوم کمتری نسبت به هر یک از مدل‌های مذکور به‌تنهایی است. لی و چوی [۳۲] در پژوهش خود با استفاده از شبکه‌های عصبی پس‌انتشار و تجزیه و تحلیل چند متغیره، با انتخاب متغیرهای مستقل مناسب به پیش‌بینی ورشکستگی پرداختند. زائو و همکاران [۳۳] در تحقیق خود با استفاده از شبکه عصبی پرسپترون چندلایه و به‌کارگیری روش جدیدی به نام انتخاب تصادفی میانگین، با رویکرد بهینه‌سازی توزیع داده‌ها، به بهبود مدل‌های امتیازدهی اعتباری پرداختند. سون و همکاران [۳۴] با استفاده از متغیرهای بیانی و به‌کارگیری رگرسیون لجستیک فازی به مدل‌سازی مسائل امتیاردهی اعتباری پرداختند. آلاراج و ابود [۳۵] به‌منظور بهبود دقت مدل‌سازی خود، یک مدل امتیازدهی اعتباری ترکیبی ارائه کردند. در این مدل روش‌های ترکیبی هم در فاز پیش‌پردازش داده‌ها و هم در فاز طبقه‌بندی مشتریان به‌کار گرفته شدند. نتو و همکاران [۳۶] با استفاده از ابزارهای هوش محاسباتی و مدل‌های کلاسیک یک چارچوب محاسبه‌ای مبتنی بر رویکردهای مدل‌محور به‌منظور پیش‌پردازش حجم انبوه اطلاعات استفاده کردند.

۳- شبکه‌های عصبی مصنوعی و منطق فازی

همان‌گونه که پیش‌تر نیز اشاره شد، در این مقاله به‌منظور بهبود کیفیت تصمیم‌گیری‌های مالی و مدل‌سازی پیچیدگی‌ها و عدم‌قطعیت‌های موجود در مسئله از ابزارهای هوش محاسباتی و روش‌های محاسبات نرم به‌صورت همزمان استفاده شده است. در ادامه به‌طور مختصر به توضیح هر یک پرداخته شده است.

۳-۱- شبکه‌های عصبی مصنوعی

ابزارهای هوش محاسباتی همچون شبکه‌های عصبی مصنوعی یکی از جمله روش‌های غیرخطی هستند که قادر به تخمین موارد غیرخطی متعدد در داده‌ها بوده و یک چارچوب محاسبه‌ای انعطاف‌پذیر برای دامنه وسیعی از مسائل غیرخطی هستند. شبکه‌های عصبی چندلایه پیشخور با یک لایه مخفی یکی از پرکاربردترین مدل‌های شبکه‌های عصبی هستند [۳۷]. این‌گونه از مدل‌ها از سه لایه پردازش ساده اطلاعات متصل به‌هم تشکیل شده‌اند. رابطه بین خروجی (y_t) و ورودی‌ها ($y_{t-1},...,y_{t-p}$) به‌صورت زیر است:

$$y_t = b_0 + \sum_{j=1}^{Q} w_j \cdot g(b_{0j} + \sum_{i=1}^{P} w_{i,j} \cdot y_{t-i}) + \varepsilon_t \qquad (1)$$

به‌طوری که w_i و w_j پارامترهای مدل بوده که اغلب وزن‌های اتصالی نامیده می‌شوند، P تعداد گره‌های ورودی و Q تعداد گره‌های مخفی هستند.

۳-۲- مدل رگرسیون فازی

مدل‌های کلاسیک از مفهوم عبارت خطا استفاده می‌کنند. به‌عبارت دیگر تخمین‌های این‌گونه از مدل‌ها مقادیر دقیقی بوده و شامل جمله خطا نمی‌شوند. این همان مفهوم پایه‌ای رگرسیون فازی است که توسط تاناکا [۳۸] پیشنهاد شده است. مفهوم اساسی رگرسیون فازی این است که جمله خطا از باقی‌مانده‌های بین مقادیر تخمین زده شده و مقادیر اصلی تولید نمی‌شوند، بلکه در عدم قطعیت پارامترهای مدل و امکان توزیع در ارتباط با مشاهدات حقیقی به‌کار گرفته می‌شوند. یک مدل رگرسیون خطی فازی در حالت کلی به‌صورت زیر است:

$$\tilde{Y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + ... + \tilde{\beta}_n x_n = \sum_{i=1}^{n} \tilde{\beta}_i x_i = [X]'[\tilde{\beta}] \qquad (2)$$

به‌طوری که [X] بردار متغیرهای مستقل، علامت پریم "′" عملگر ترانهاده، n تعداد متغیرها و $\tilde{\beta}_i$ مجموعه‌های فازی بیانگر i–امین پارامتر مدل است. این اعداد فازی (پارامترهای $\tilde{\beta}_i$) به شکل اعداد فازی نوع ال، دوبویس و پراد $(\alpha_i, f_i)_L$ با توزیع احتمال به‌صورت زیر هستند:

$$\mu_{\tilde{\beta}_i}(\beta_i) = L\left\{\left(\alpha_i - \frac{\beta_i}{C}\right)\right\} \qquad (3)$$

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

۴- فرمول‌بندی مدل پیشنهادی

شبکه‌های عصبی پرسپترون چندلایه ابزارهای محاسبه انعطاف‌پذیر با درجه بالایی از کارایی هستند و برای طیف وسیعی از مسائل طبقه‌بندی به‌کار گرفته می‌شوند. شبکه‌های عصبی پرسپترون چندلایه دارای ویژگی‌های متمایزی هستند که آنها را به ابزاری ارزشمند و قوی در حوزه مسائل طبقه‌بندی تبدیل کرده است. از جمله مهم‌ترین این ویژگی‌ها می‌توان به اطلاعات محور بودن و حالت خودتطبیقی این روش‌ها اشاره کرد. با وجود ویژگی‌های گفته شده، این مدل‌ها دارای محدودیت تعداد زیاد داده‌های گذشته برای به‌دست آوردن نتایج دقیق هستند [۳۹]. درصورتی که امروزه به‌علت عدم‌قطعیت محیط و توسعه سریع تکنولوژی، باید موقعیت‌های آینده را با استفاده از داده‌های کم و در بازه‌های زمانی کوتاه‌مدت تحلیل و ارزیابی کرد. بنابراین نیازمند روش‌هایی هستیم که به داده‌های کمتری نیاز داشته و در این‌گونه از شرایط نیز کارا باشند. هدف این روش ترکیبی بهره‌مندی از ویژگی‌های رگرسیون فازی و برطرف کردن محدودیت نیاز به داده‌های زیاد در روش شبکه عصبی به‌منظور ارزیابی ریسک اعتباری است.

پارامترهای شبکه‌های عصبی مقادیر قطعی هستند ($w_{i,j}\left(i=0,1,2,...,p \quad j=1,2,...,q\right)$ و $w_j\left(j=0,1,2,...,q\right)$) درصورتی که در روش جدید به‌جای به‌کارگیری این مقادیر قطعی، پارامترهای شبکه به‌صورت اعداد فازی درنظر گرفته شده‌اند. ($\tilde{w}_{i,j}\left(i=0,1,2,...,p \quad j=1,2,...,q\right)$ و $\tilde{w}_j\left(j=0,1,2,...,q\right)$) مدل ترکیبی با پارامترهای فازی به‌صورت زیر است:

$$\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^{q} \tilde{w}_j \cdot g(\tilde{w}_{0,j} + \sum_{i=1}^{p} \tilde{w}_{i,j} \cdot y_{t-i})), \qquad (۴)$$

به‌قسمی که y_t مشاهدات، $\tilde{w}_{i,j}\left(i=0,1,2,...,p \quad j=1,2,...,q\right)$ و $\tilde{w}_j\left(j=0,1,2,...,q\right)$ اعداد فازی و g و f به‌ترتیب توابع فعال‌سازی لایه میانی و لایه خروجی هستند. با جای‌گذاری، معادله فوق به‌صورت زیر تبدیل می‌شود:

$$\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^{q} \tilde{w}_j \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^{q} \tilde{w}_j \cdot \tilde{X}_{t,j}), \qquad (۵)$$

به قسمی که $\tilde{X}_{t,j} = g(\tilde{w}_{0,j} + \sum_{i=1}^{p} \tilde{w}_{i,j} \cdot y_{t-i})$ است. وزن‌های اتصالی بین لایه ورودی و لایه میانی در این معادله به‌صورت اعداد فازی مثلثی $\tilde{w}_{i,j} = \left(a_{i,j}, b_{i,j}, c_{i,j}\right)$ مطابق زیر است:

$$\mu_{\tilde{w}_{i,j}}\left(w_{i,j}\right) = \begin{cases} \frac{1}{b_{i,j}-a_{i,j}}\left(w_{i,j}-a_{i,j}\right) & a_{i,j} \le w_{i,j} \le b_{i,j}, \\ \frac{1}{b_{i,j}-c_{i,j}}\left(w_{i,j}-c_{i,j}\right) & b_{i,j} \le w_{i,j} \le c_{i,j}, \\ 0 & \text{otherwise}, \end{cases}$$

(۶)

به‌طوری که $\mu_{\tilde{w}}\left(w_{i,j}\right)$ تابع عضویت یک مجموعه فازی است که با پارامترهای $w_{i,j}$ مشخص می‌شود. حال با استفاده از اصل گسترش، تابع عضویت $\tilde{X}_{t,j} = g(\tilde{w}_{0,j} + \sum_{i=1}^{p} \tilde{w}_{i,j} \cdot y_{t-i})$ مطابق زیر خواهد بود [۳۹]:

$$\mu_{\tilde{X}_{t,j}}\left(X_{t,j}\right) = \begin{cases} \frac{\left(X_{t,j} - g\left(\sum_{i=0}^{p} a_{i,j} y_{t,i}\right)\right)}{g\left(\sum_{i=0}^{p} b_{i,j} y_{t,i}\right) - g\left(\sum_{i=0}^{p} a_{i,j} y_{t,i}\right)} \\ \text{if} \quad g\left(\sum_{i=0}^{p} a_{i,j} y_{t,i}\right) \le X_{t,j} \le g\left(\sum_{i=0}^{p} b_{i,j} y_{t,i}\right), \\ \frac{\left(X_{t,j} - g\left(\sum_{i=0}^{p} c_{i,j} y_{t,i}\right)\right)}{g\left(\sum_{i=0}^{p} b_{i,j} y_{t,i}\right) - g\left(\sum_{i=0}^{p} c_{i,j} y_{t,i}\right)} \\ \text{if} \quad g\left(\sum_{i=0}^{p} b_{i,j} y_{t,i}\right) \le X_{t,j} \le g\left(\sum_{i=0}^{p} c_{i,j} y_{t,i}\right), \\ 0 \qquad \text{otherwise}, \end{cases}$$

(۷)

به‌طوری که $y_{t,i} = y_{t-i} \quad \left(t=1,2,...,k, \quad i=1,2,...,p\right)$ و $y_{t,i} = 1 \quad \left(t=1,2,...,k, \quad i=0\right)$. با درنظر گرفتن داده‌های فازی مثلثی $\tilde{X}_{t,j}$ با تابع عضویتی مطابق رابطه (۷) و پارامترهای فازی مثلثی $\tilde{w}_j$ تابع عضویت $\tilde{y}_t = f(\tilde{w}_0 + \sum_{j=1}^{q} \tilde{w}_j \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^{q} \tilde{w}_j \cdot \tilde{X}_{t,j})$ مطابق با رابطه زیر خواهد بود:

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

$$\mu_{\tilde{Y}}(y_t) \cong \begin{cases} \dfrac{-B_1}{2A_1} + \left[\left(\dfrac{B_1}{2A_1}\right)^2 - \dfrac{C_1 - f^{-1}(y_t)}{A_1}\right]^{\frac{1}{2}} \\ \quad \text{if } C_1 \le f^{-1}(y_t) \le C_3, \\ \dfrac{B_2}{2A_2} + \left[\left(\dfrac{B_2}{2A_2}\right)^2 - \dfrac{C_2 - f^{-1}(y_t)}{A_2}\right]^{\frac{1}{2}} \\ \quad \text{if } C_3 \le f^{-1}(y_t) \le C_2, \\ 0 \qquad \text{otherwise,} \end{cases} \tag{۸}$$

به‌قسمی که:

$$A_1 = \sum_{j=0}^{q}(e_j - d_j)\cdot\begin{pmatrix} g\left(\sum_{i=0}^{p} b_{i,j}\cdot y_{t,i}\right) \\ -g\left(\sum_{i=0}^{p} a_{i,j}\cdot y_{t,i}\right)\end{pmatrix},$$

$$B_1 = \sum_{j=0}^{q}\left(d_j\cdot\begin{pmatrix} g\left(\sum_{i=0}^{p} b_{i,j}\cdot y_{t,i}\right) \\ -g\left(\sum_{i=0}^{p} a_{i,j}\cdot y_{t,i}\right)\end{pmatrix} + g\left(\sum_{i=0}^{p} a_{i,j}\cdot y_{t,i}\right)\cdot(e_j - d_j)\right),$$

$$A_2 = \sum_{j=0}^{q}(f_j - e_j)\cdot\begin{pmatrix} g\left(\sum_{i=0}^{p} c_{i,j}\cdot y_{t,i}\right) \\ -g\left(\sum_{i=0}^{p} b_{i,j}\cdot y_{t,i}\right)\end{pmatrix},$$

$$B_2 = \sum_{j=0}^{q}\left(f_j\cdot\begin{pmatrix} g\left(\sum_{i=0}^{p} c_{i,j}\cdot y_{t,i}\right) \\ -g\left(\sum_{i=0}^{p} b_{i,j}\cdot y_{t,i}\right)\end{pmatrix} + g\left(\sum_{i=0}^{p} c_{i,j}\cdot y_{t,i}\right)\cdot(f_j - e_j)\right),$$

$$C_1 = \sum_{j=0}^{q}\left(d_j\cdot g\left(\sum_{i=0}^{p} a_{i,j}\cdot y_{t,i}\right)\right)$$

$$C_2 = \sum_{j=0}^{q}\left(f_j\cdot g\left(\sum_{i=0}^{p} c_{i,j}\cdot y_{t,i}\right)\right),$$

$$C_3 = \sum_{j=0}^{q}\left(e_j\cdot g\left(\sum_{i=0}^{p} b_{i,j}\cdot y_{t,i}\right)\right),$$

حال با درنظر گرفتن سطح آستانه h برای میزان توابع عضویت تمامی مشاهدات یک مدل برنامه‌ریزی غیرخطی مطابق با رابطه زیر به‌دست می‌آید:

$$\text{Min} \sum_{t=1}^{k}\sum_{j=0}^{q}\left(f_j g\left(\sum_{i=0}^{p} c_{i,j}y_{t,i}\right)\right) - \left(d_j g\left(\sum_{i=0}^{p} a_{i,j}y_{t,i}\right)\right)$$

$$\text{Subject.to} \begin{cases} \dfrac{-B_1}{2A_1} + \left[\left(\dfrac{B_1}{2A_1}\right)^2 - \dfrac{C_1 - f^{-1}(y_t)}{A_1}\right]^{\frac{1}{2}} \le h \\ \quad \text{if } C_1 \le f^{-1}(y_t) \le C_3, \text{ for } t = 1,2,....,k, \\ \dfrac{B_2}{2A_2} + \left[\left(\dfrac{B_2}{2A_2}\right)^2 - \dfrac{C_2 - f^{-1}(y_t)}{A_2}\right]^{\frac{1}{2}} \le h \\ \quad \text{if } C_3 \le f^{-1}(y_t) \le C_2, \text{ for } t = 1,2,....,k, \end{cases} \tag{۹}$$

درنهایت خروجی مدل به‌صورت فازی و پیوسته است در حالی که در مسائل طبقه‌بندی خروجی به‌صورت غیرفازی و گسسته است. بنابراین برای به‌کارگیری مدل پیشنهادی در مسائل طبقه‌بندی نیازمند بهبود و تعدیل مدل هستیم. برای این منظور، در مرحله اول به هر طبقه یک مقدار عددی اطلاق می‌شود سپس احتمال عضویت خروجی در هر طبقه مطابق زیر محاسبه می‌شود:

$$P_A = 1 - P_B = \frac{\int_{-\infty}^{m} f(x)\,dx}{\int_{-\infty}^{+\infty} f(x)\,dx} = 1 - \frac{\int_{m}^{+\infty} f(x)\,dx}{\int_{-\infty}^{+\infty} f(x)\,dx} \tag{۱۰}$$

به قسمی که P_A و P_B به‌ترتیب احتمال عضویت در هر یک از طبقه‌های A و B است. در نهایت یک نمونه به طبقه‌ای منسوب می‌شود که خروجی آن دارای احتمال بزرگ‌تری باشد. همان‌طور که قبلاً بیان شد، در مدل پیشنهادی خروجی به‌صورت فازی است بنابراین بهتر است برای هر طبقه مقادیر عددی بزرگ درنظر گرفته شود تا مدل نسبت به تغییرات در ورودی حساسیت بیشتری داشته باشد.

۵- مجموعه اطلاعات اعتباری محک

در این بخش به‌منظور نمایش تناسب و اثربخشی مدل پیشنهادی برای طبقه‌بندی داده‌های مالی از مجموعه‌ای از اطلاعات اعتباری مربوط به کشورهای ژاپن و استرالیا استفاده شده است. توضیح مختصری در رابطه با هر یک از این مجموعه اطلاعات در قسمت

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

جدول ۱- مجموعه اطلاعات اعتباری ژاپن

ویژگی‌ها	نوع	مقادیر	مقادیر (پیشن)
۱	دسته‌ای	۰ و ۱	a,b
۲	عددی	۱۳/۷۵–۸۰/۲۵	۱۳/۷۵–۸۰/۲۵
۳	عددی	۰–۲۸	۰–۲۸
۴	دسته‌ای	۱ و ۲ و ۳	u,y,i
۵	دسته‌ای	۱ و۲ و ۳	g,p,gg
۶	دسته‌ای	۱ و ۲و ۳و...و ۱۴	ff,d,i,k,j,aa,m,c,w,,q,r,cc,x
۷	دسته‌ای	۱ و ۲و ۳ و...و ۹	ff,dd,j,bb,v,n,o,h,z
۸	عددی	۰–۲۸/۵	۰–۲۸/۵
۹	دسته‌ای	۰و ۱	t,f
۱۰	دسته‌ای	۰ و ۱	t,f
۱۱	عددی	۰–۶۷	۰–۶۷
۱۲	دسته‌ای	۰ و ۱	t,f
۱۳	دسته‌ای	۱ و۲ و ۳	s,g,p
۱۴	عددی	۰–۲۰۰۰	۰–۲۰۰۰
۱۵	عددی	۰–۱۰۰۰۰۰	۰–۱۰۰۰۰۰
کلاس	گسسته	۰ و ۱	+ و –

بعدی ذکر شده است. این مجموعه اطلاعات، به‌علت دارا بودن ترکیب مناسبی از ویژگی‌ها که همزمان شامل ارزش‌های عددی کوچک و ارزش‌های عددی بزرگ هستند مورد توجه بسیاری از محققان در حوزه مدیریت مالی قرار گرفته‌اند.

۵-۱- مجموعه اطلاعات اعتباری ژاپن

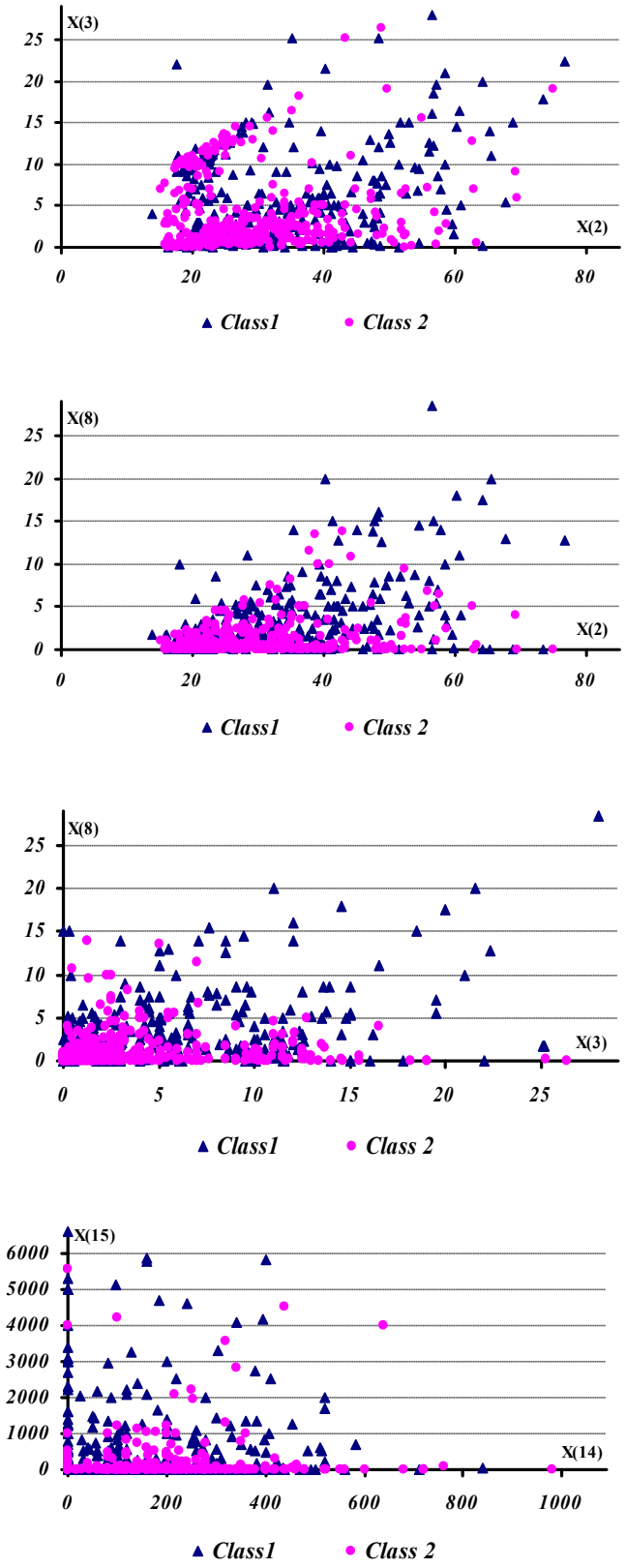
به‌طور کلی این مجموعه اطلاعات شامل ۶۹۰ مشاهده است. در این بین ۳۰۷ نمونه (حدود ۴۴/۵ درصد) از افراد مورد مطالعه در زمره مشتریان خوش‌حساب و ۳۸۳ نمونه (حدود ۵۵/۵ درصد) در زمره مشتریان بدحساب قرار دارند. این مجموعه داده در مجموع شامل ۳۷ مقدار گم شده است. در این مقاله تمامی این ۳۷ مشاهده از مجموعه داده‌های اصلی حذف شده‌اند. مجموعه داده جدید شامل ۶۵۳ مشاهده است که ۳۵۷ مورد از آنها در زمره مشتریان خوش‌حساب و ۲۹۶ مورد در زمره مشتریان بدحساب قرار می‌گیرند. همچنین این مجموعه اطلاعات شامل ۱۵ ویژگی است که ۹ مورد از این ویژگی‌ها از نوع دسته‌ای و ۶ مورد از آنها به‌صورت عددی هستند. لازم به‌ذکر است به‌دلیل محرمانه بودن و محافظت از اطلاعات، از بیان دقیق نام ویژگی‌ها و ارزش آنها به‌صورت مستقیم خودداری شده است. خلاصه‌ای از اطلاعات مورد استفاده در جدول (۱) قابل مشاهده است.

داده‌ها به‌صورت تصادفی به دو دسته، داده‌های آموزش و داده‌های تست، تقسیم‌بندی شده‌اند. از بین آنها ۵۰ درصد داده‌ها مربوط به آموزش و ۵۰ درصد مابقی مربوط به تست مدل هستند. نمودارهایی از توزیع دوبعدی از تقابل هر دو ویژگی رسم شده است که در شکل (۱) نمودارهای مربوط به تقابل ویژگی‌های $(A_۲,A_۳)$، $(A_۲,A_۸)$، $(A_۳,A_۸)$ و $(A_{۱۴},A_{۱۵})$ به‌عنوان نمونه قابل مشاهده است.

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019



شکل ۱- توزیع دو بعدی از ویژگی‌های کارت‌های اعتباری ژاپن

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

جدول ۲- مجموعه اطلاعات اعتباری استرالیا

ویژگی‌ها	نوع	مقادیر	مقادیر (پیشن)
۱	گسسته	۰ و ۱	a,b
۲	پیوسته	۱۳/۷۵–۸۰/۲۵	۱۳/۷۵–۸۰/۲۵
۳	پیوسته	۰–۲۸	۰–۲۸
۴	گسسته	۱ و ۲و ۳	p,g,gg
۵	گسسته	۱ و ۲ و ۳ و...و ۱۴	ff,d,i,k,j,aa,m,c,w,e,q,r,cc,x
۶	گسسته	۱ و ۲ و ۳ و...و ۹	ff,dd,j,bb,v,n,o,h,z
۷	پیوسته	۰–۲۸/۵	۰–۲۸/۵
۸	گسسته	۰ و ۱	t,f
۹	گسسته	۰و۱	t,f
۱۰	پیوسته	۰–۶۷	۰–۶۷
۱۱	گسسته	۰ و ۱	t,f
۱۲	گسسته	۱ و ۲ و ۳	s,g,p
۱۳	پیوسته	۰–۲۰۰۰	۰–۲۰۰۰
۱۴	پیوسته	۰–۱۰۰۰۰۰	۰–۱۰۰۰۰۰
کلاس	گسسته	۰ و ۱	- و +

۵-۲- مجموعه اطلاعات اعتباری استرالیا

به‌طور کلی این مجموعه اطلاعات شامل ۶۹۰ مشاهده و ۱۴ ویژگی است. ۸ مورد از این ویژگی‌ها دارای مقادیر گسسته و شش مورد دارای مقادیر پیوسته هستند. در این بین ۳۰۷ نمونه (حدود ۴۴/۵ درصد) از افراد مورد مطالعه در زمره مشتریان خوش‌حساب و ۳۸۳ مورد (حدود ۵۵/۵ درصد) در زمره مشتریان بدحساب قرار دارند. خلاصه‌ای از اطلاعات مورد استفاده در جدول (۲) قابل مشاهده است.

داده‌ها به‌صورت تصادفی به دو دسته، داده‌های آموزش و داده‌های تست، تقسیم‌بندی شده‌اند. از بین آنها ۵۰ درصد داده‌ها مربوط به آموزش و ۵۰ درصد مابقی مربوط به تست مدل هستند. نمودارهایی از توزیع دوبعدی از تقابل هر دو ویژگی رسم شده است که در شکل (۲) نمودارهای مربوط به تقابل ویژگی‌های (A_2,A_3)، (A_2,A_7)، (A_3,A_{10}) و (A_2,A_{10}) به‌عنوان نمونه قابل مشاهده است.

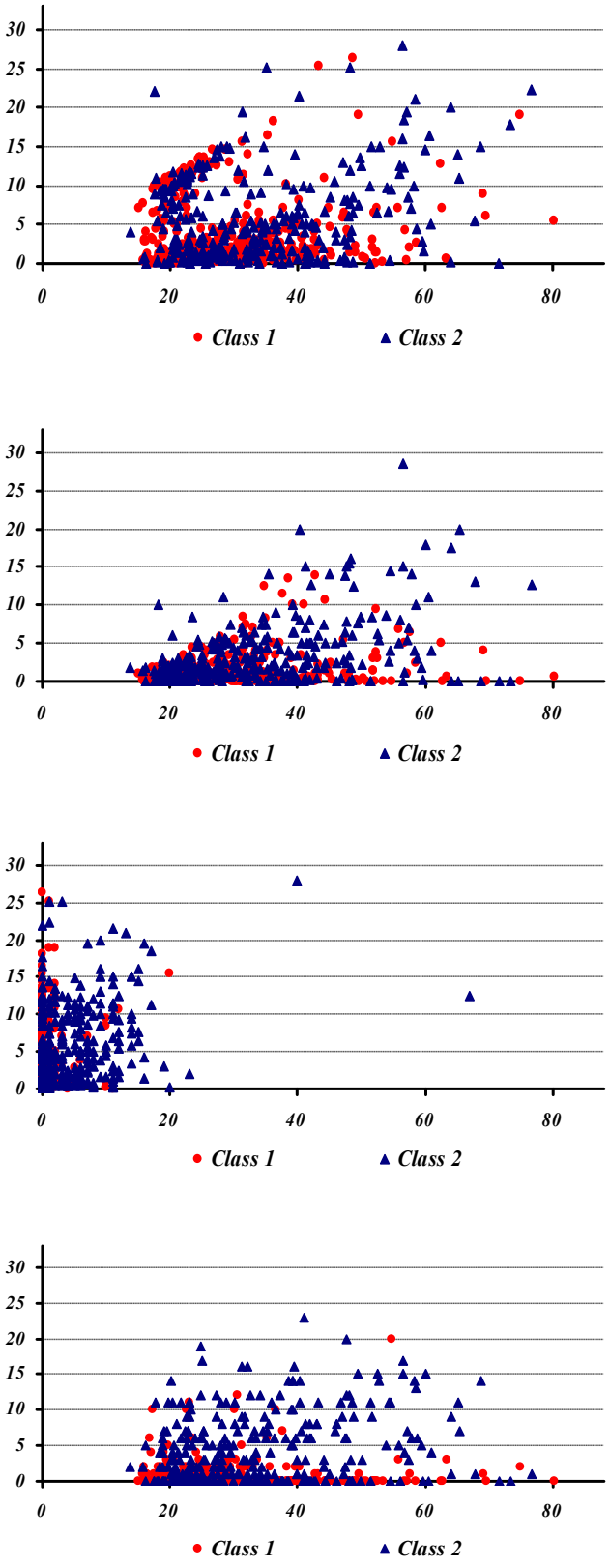
۶- به‌کارگیری مدل پیشنهادی در تحلیل امتیازدهی اعتباری

در این قسمت کارایی مدل ترکیبی پیشنهادی، با استفاده از اطلاعات اعتباری کشورهای ژاپن و استرالیا، در تحلیل امتیازدهی اعتباری نشان داده شده است. در مرحله اول به‌منظور دستیابی به بهترین ساختار شبکه عصبی مبتنی بر مفهوم پرسپترون‌های چند لایه [۳۶]، با استفاده از نرم‌افزار MATLAB 7، شبکه‌های عصبی با ساختارهای مختلف از لحاظ کارایی با یکدیگر مقایسه شده‌اند. درنهایت شبکه ایده‌آل مربوط به اطلاعات اعتباری کشور ژاپن، شبکه‌ای متشکل از ۱۵ نورون ورودی، ۱۶ نورون میانی و یک نورون خروجی (به‌صورت $N^{(15-16-1)}$) و شبکه ایده‌آل مربوط به اطلاعات اعتباری کشور استرالیا، شبکه‌ای متشکل از ۱۴ نورون ورودی، ۱۴ نورون میانی و یک نورون خروجی (به‌صورت $N^{(14-14-1)}$) بوده است. هنگامی که ساختار بهینه تعیین شد، شبکه به‌وسیله الگوریتم پس‌انتشار در ۱۰۰ تکرار آموزش دیده، سپس در مرحله

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019



شکل ۲- توزیع دو بعدی از ویژگی‌های کارت‌های اعتباری استرالیا

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

جدول ۳- نرخ طبقه‌بندی نادرست در هر مدل

مدل	نرخ طبقه‌بندی نادرست (٪)	
	اطلاعات اعتباری ژاپن	اطلاعات اعتباری استرالیا
تحلیل جدایش خطی (LDA)[۱]	۱۳/۸	۱۴/۰
تحلیل جدایش درجه دوم (QDA)[۲]	۱۵/۶	۱۹/۹
Kامین نزدیک‌ترین همسایه (KNN)[۳]	۱۳/۸	۱۴/۲
ماشین‌های بردار پشتیباند (SVM)[۴]	۱۳/۸	۲۲/۵
شبکه‌های عصبی مصنوعی (ANN)[۵]	۱۱/۴	۱۲/۳
سیستم استنتاج فازی- عصبی تطبیقی (ANFIS)	۱۰/۸	۱۱/۷
مدل ترکیبی پیشنهادی	۹/۷	۱۰/۹

جدول ۴- میزان بهبود مدل پیشنهادی در مقایسه با سایر مدل‌ها

مدل	میزان بهبود (٪)	
	اطلاعات اعتباری ژاپن	اطلاعات اعتباری استرالیا
تحلیل جدایش خطی (LDA)	۲۹/۷۱	۲۲/۱۴
تحلیل جدایش درجه دوم (QDA)	۳۷/۸۲	۴۵/۲۳
Kامین نزدیک‌ترین همسایه (KNN)	۲۹/۷۱	۲۳/۲۴
ماشین‌های بردار پشتیبان (SVM)	۲۹/۷۱	۵۱/۵۶
شبکه‌های عصبی مصنوعی (ANN)	۱۴/۹۱	۱۱/۳۸
سیستم استنتاج فازی- عصبی تطبیقی (ANFIS)	۱۰/۱۸	۶/۸۳

دوم حداقل ابهام در پارامترهای فازی با درنظر گرفتن سطح آستانه h=۰، برای به‌دست آوردن پارامترهای فازی با حداقل پهنا با استفاده از معادله (۹)، تعیین شده است. نرخ طبقه‌بندی نادرست در هر مدل و درصد بهبود مدل پیشنهادی (با درنظر گرفتن h=۰)، در مقایسه با سایر مدل‌های طبقه‌بندی در ارزیابی اطلاعات، به‌ترتیب در جدول‌های (۳) و (۴) آورده شده است.

همان‌طور که نتایج به‌دست آمده از جدول‌های بالا نشان می‌دهند، در تحلیل اعتباری اطلاعات مربوط به کشور ژاپن مدل پیشنهادی ما، با ۹/۷ درصد خطا در نرخ طبقه‌بندی نادرست و ایجاد ۱۴/۹۱ درصد بهبود در ارزیابی اطلاعات نسبت به مدل پرسپترون چندلایه سنتی، عملکرد بالای خود را نمایان می‌سازد. به‌علاوه مدل پیشنهادی به‌ترتیب منجر به ۲۹/۷۱، ۳۷/۸۲، ۲۹/۷۱، ۲۹/۷۱ و ۱۰/۱۸ درصد بهبود نسبت به مدل‌های تحلیل جدایش خطی (c=۰)، تحلیل جدایش درجه دوم (c=۰)، Kامین نزدیک‌ترین همسایه (K=۵)، ماشین‌های بردار پشتیبان (c=۰) و سیستم استنتاج فازی- عصبی تطبیقی در ارزیابی اطلاعات شده است. به‌طور مشابه، مدل پیشنهادی در تحلیل اطلاعات اعتباری مربوط به کشور استرالیا با ۱۰/۹ درصد، دارای کمترین میزان خطا در مقایسه با سایر مدل‌ها است. همچنین این مدل ترکیبی به‌ترتیب باعث ۲۲/۱۴، ۴۵/۲۳، ۲۳/۲۴، ۵۱/۵۶، ۱۱/۳۸ و ۶/۸۳ درصد بهبود نسبت به مدل‌های تحلیل جدایش خطی (c=۰)، تحلیل جدایش درجه دوم (c=۰)، Kامین نزدیک‌ترین همسایه (K=۱۳)، ماشین‌های بردار پشتیبان (c=۱)، شبکه‌های عصبی مصنوعی و سیستم استنتاج فازی- عصبی

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

تطبیقی در ارزیابی اطلاعات شده است.

۷- نتیجه‌گیری

در سال‌های اخیر تجزیه و تحلیل ریسک اعتباری به‌عنوان یکی از حوزه‌های مهم تحقیقاتی در زمینه مدیریت مالی شناخته شده است. مدیریت ریسک و امتیازدهی اعتباری یکی از تکنیک‌های مهم و تأثیرگذار در ارزیابی ریسک اعتباری و یکی از حوزه‌های کاربردی اصلی مسائل طبقه‌بندی، مبتنی بر جداسازی کارت‌های اعتباری مشتریان و تفکیک آنها به دو گروه اعتباری خوب و بد است. اعطای اعتبار به مشتریان نیازمند تجزیه و تحلیل همه جانبه عوامل و متغیرهای مؤثر بر آنها به‌منظور تصمیم‌گیری صحیح است. اما وجود ابهام و پیچیدگی بین این متغیرها، تصمیم‌گیری در این موارد را بسیار دشوار ساخته است. یافته‌های نظری و عملی حاکی از آن است که استفاده از مدل‌های ترکیبی، به‌ویژه زمانی که مدل‌ها دارای ساختاری کاملاً متفاوت باشند، امری رایج به‌منظور کاهش نرخ طبقه‌بندی نادرست است.

در این مقاله یک مدل پرسپترون چندلایه، که در آن از مزایای خاص محاسبات هوشمند نرم و منطق فازی استفاده شده، به‌عنوان مدلی جایگزین برای طبقه‌بندی صحیح مشتریان پیشنهاد شده است. در روش پیشنهادی، به‌منظور مدل‌سازی بهتر بازارهای مالی در شرایط عدم قطعیت، پارامترهای شبکه عصبی پرسپترون چندلایه، شامل وزن‌ها و خطاها، به‌صورت پارامترهای فازی درنظر گرفته شده‌اند. در این روش ابتدا سیستم مورد مطالعه با استفاده از شبکه‌های عصبی متامدل‌بندی شده و سپس با به‌کارگیری استنتاج فازی، تصمیم و سیاست بهینه با بیشترین میزان برتری تعیین خواهد شد. مدل پیشنهادی بر خلاف مدل‌های تحلیل جدایش خطی و درجه دوم، فاقد هرگونه فرض شده‌هایی از جمله وجود ارتباط بین متغیرهای مستقل و وابسته است. همچنین این مدل بر خلاف طبقه‌بندی کننده Kامین نزدیک‌ترین همسایه نیازمند ذخیره‌سازی داده‌های آموزش نیست. ماشین‌های بردار پشتیبان نیازمند تعیین پارامتر جریمه و تابع کرنل هستند درصورتی که در مدل پیشنهادی بدون تنظیم این پارامترها، تنها با تکیه بر روند آموزش، طبقه‌بندی کننده نهایی تعیین می‌شود و در نهایت بر خلاف پرسپترون‌های چندلایه، برای دستیابی به نتایج کارا نیازمند حجم زیادی از اطلاعات نیست. نتیجه حاصل از به‌کارگیری روش پیشنهادی بیانگر کارامدی و دقت بالای این روش نسبت به سایر مدل‌های طبقه‌بندی آماری و هوشمند در تحلیل مسائل امتیازدهی اعتباری است.

واژه‌نامه

1. linear discriminant analysis
2. quadratic discriminant analysis
3. K-nearest neighbor
4. Support Vector Machines
5. artificial neural networks

مراجع

1. Lee , T. S., Chiu, C. C., Lu, C. J., and Chen, I. F. "Credit Scoring using the Hybrid Neural Discriminant Technique", *Journal of Expert Systems with Applications*, Vol. 23, pp. 245-254, 2002.
2. Gestel, T. V., and Baesens, B., *Credit Risk Management*, Published in the United States by Oxford University Press Inc., New York, 2009.
3. Westgaard, S., and Van der Wijst, N., "Default Probabilities in a Corporate Bank Portfolio: A Logistic Model Approach", *European Journal of Operational Research*, Vol. 135, No. 2, pp. 338-349, 2001.
4. Charitou, A., Neophytou, E., and Charalambous, C., "Predicting Corporate Failure: Empirical Evidence for the UK", *European Accounting Review*, Vol. 13, No. 3, pp. 465-497, 2004.
5. Lee, G., Sung, T. K., and Chang, N., "Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction", *Journal of Management Information Systems*, Vol. 16, pp. 63-85, 1999.
6. Bellotti, T., and Crook, J., "Support Vector Machines for Credit Scoring and Discovery of Significant

www.SID.ir

Archive of SID

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

Features", *Expert Systems with Applications*, Vol. 36, pp. 3302-3308, 2008.
7. Sharma, S., *Applied Multivariate Techniques*, New York, NY: Wiley, SPSS 1997-Statistic Modeling for Windows 95/98/NT, SPSS Inc, 1998.
8. Baesens, B., Verstraeten, G., Poel, D. V. d., Egmont-Petersen, M., Kenhove, P. V., and Vanthienen, J., "Bayesian Network Classifiers for Identifying the Slope of the Customer -lifecycle of Long-life Customers", *European Journal of Operational Research*, Vol. 156, pp. 508-523, 2004.
9. Sun, L., and Shenoy, P. P., "Using Bayesian Networks for Bankruptcy Prediction: Some Methodological Issues", *European Journal of Operational Research*, Vol. 180, pp. 738-753, 2007.
10. Lee, T. S., and Chen, I. F., "A Two-stage Hybrid Credit Scoring Model using Artificial Neural Networks and Multivariate Adaptive Regression Splines", *Expert Systems with Applications*, Vol. 28, pp. 743-752. 2005.
11. Kline, D. M., "Two-group Classification using the Bayesian Data Reduction Algorithm", *Complexity*, Vol. 15, pp. 43-49, 2010.
12. Huang, C. L., Chen, M. C., and Wang, C. J., "Credit Scoring with a Data Mining Approach Based on Support Vector Machines", *Expert Systems with Applications*, Vol. 33, pp. 847-856, 2007.
13. Huang, J. -J., Tzeng, G. -H., and Ong, C. -S., "Two-stage Genetic Programming (2SGP) for the Credit Scoring Model", *Applied Mathematics and Computation*, Vol. 174, No. 2, pp. 1039-1053, 2006.
14. Chen, M. C., and Huang, S. H., "Credit Scoring and Rejected Instances Reassigning Through Evolutionary Computation Techniques", *Expert Systems with Applications*, Vol. 24, pp. 433-441, 2003.
15. Khashei, M., Bijari, M., and Raissi, Gh. A., "Improvement of Auto-regressive Integrated Moving Average Models using Fuzzy Logic and Artificial Neural Networks (ANNs)", *Neurocomputing*, Vol. 72, pp. 956-967, 2009.
16. Celikyilmaz, A., and Turksen, I. B., "Enhanced Fuzzy System Models with Improved Fuzzy Clustering Algorithm", *IEEE Transactions on Fuzzy Systems*, Vol. 16, pp. 779-794, 2008.
17. Khashei, M., Bijari, M., and Hejazi, S. R., "Combining Seasonal ARIMA Models with Computational Intelligence Techniques for Time Series Forecasting", *Soft Computing*, Vol. 16, pp. 1091-1105, 2012.
18. Lee, T. S., Chiu, C. C., Lu, C. J., and Chen, I. F., "Credit Scoring using the Hybrid Neural Discriminate Technique", *Expert Systems with Applications*, Vol. 23, pp. 245-254, 2002.
19. Hsieh, N -Ch., "Hybrid Mining Approach in the Design of Credit Scoring Models", *Expert Systems with Applications*, Vol. 28, pp. 655-665, 2005.
20. Hung, C. L., Chen, M. C., and Wang, C. J., "Credit Scoring with a Data Mining Approach Based on Support Vector Machines", *Expert Systems with Applications*, Vol. 33, pp. 847-856, 2007.
21. Li, J., Wei, L., Li, G., and Xu, W., "An Evolution Strategy-based Multiple Kernels Multi-criteria Programming Approach: The Case of Credit Decision Making", *Decision Support Systems*, Vol. 51, pp. 292-298, 2011.
22. Celikyilmaz, A., and Turksen, I. B., "Uncertainty Modeling of Improved Fuzzy Functions with Evolutionary Systems", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 38, pp. 1098-1110, 2008.
23. Chen, F. -L., and Li, F. -Ch., "Combination of Feature Selection Approaches with SVM in Credit Scoring", *Expert Systems with Applications*, Vol. 37 pp. 4902-4909, 2010.
24. Bijak, K., and Thomas, L. C., "Does Segmentation Always Improve Model Performance in Credit Scoring", *Expert Systems with Applications*, Vol. 39, No. 3, pp. 2433-2442, 2012.
25. Kim, K. S., and Han, I., "The Cluster-indexing Method for Case based Reasoning using Self-organizing Maps and Learning Vector Quantization for Bond Rating Cases", *Expert Systems with Applications*, Vol. 21, pp. 147-156, 2001.
26. Park, C., and Han, I., "A Case-based Reasoning with the Feature Weights Derived by Analytic Hierarchy Process for Bankruptcy Prediction", *Expert Systems with Applications*, Vol. 23, No. 3, pp. 255-264, 2002.
27. Ahn, H., and Kim, K. -J., "Bankruptcy Prediction Modeling with Hybrid Case-based Reasoning and Genetic Algorithms Approach", *Applied Soft Computing*, Vol. 9, pp. 599-607, 2009.
28. Capotorti, A., and Barbanera, E., "Credit Scoring Analysis using a Fuzzy Probabilistic Rough Set Model", *Computational Statistics and Data Analysis*, Vol. 56, No. 4, pp. 981-994, 2012.
29. Akkoc, S., "An Empirical Comparison of Conventional Techniques, Neural Networks and the Three Stage Hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) Model for Credit Scoring Analysis: The Case of Turkish Credit Card Data", *European Journal of Operational Research*, Vol. 222, pp. 168-178, 2012.
30. Yu, L., Wang, S., and Lai, K., "An Intelligent-agent-based Fuzzy Group Decision Making Model for Financial Multicriteria Decision Support: The Case of Credit Scoring", *European Journal of Operational Research*, Vol. 195, pp. 942-959, 2009.
31. Chen, W., Ma, C., and Ma, L., "Mining the Customer Credit using Hybrid Support Vector Machine Technique", *Expert Systems with Applications*, Vol. 36, pp. 7611-7616, 2009.
32. Lee, S., and Choi, W. S., "A Multi-industry Bankruptcy Prediction Model using Back-

www.SID.ir

Archive of SID

propagation Neural Network and Multivariate Discriminant Analysis", *Expert Systems with Applications*, Vol. 40, pp. 2941-2946, 2013.
33. Zhao, Z., Xu, S., Kong, B. H., Kabir, M. M. J., Liu, Y., and Wasinger, R., "Investigation and Improvement of Multi-layer Perceptron Neural Networks for Credit Scoring", *Expert Systems with Applications*, Vol. 42, pp. 3508-3516, 2015.
34. Sohn, S. Y., Kim, H. D., and Yoon, H. J., "Technology Credit Scoring Model with Fuzzy Logistic Regression", *Applied Soft Computing*, Vol. 43, pp. 150-158, 2016.
35. Alaraj, M., and Abbod, F., "A New Hybrid Ensemble Credit Scoring Model Based on Classifiers Consensus System Approach", *Expert Systems With Applications*, Vol. 64, pp. 36-55, 2016.
36. Neto, R., Adeodato, P. J., and Salgado, C. A., "A Framework for Data Transformation in Credit Behavior Scoring Applications Based on Model Driven Development", *Expert Systems With Applications*, Vol. 72, pp. 1-13, 2016.
37. Khashei, M., and Bijari, M., "Hybridization of the Probabilistic Neural Networks with Feed-forward Neural Networks for Forecasting", *Engineering Applications of Artificial Intelligence*, Vol. 25, pp. 1277-1288, 2012.
38. Tanaka, H., "Fuzzy Data Analysis by Possibility Linear Models", *Fuzzy Sets and Systems*, Vol. 24, pp. 363-375, 1987.
39. Khashei, M., Hejazi, S. R., and Bijari, M., "A New Hybrid Artificial Neural Networks and Fuzzy Regression Model for Time Series Forecasting", *Fuzzy Set and Systems*, Vol. 159, pp. 769-786, 2008.

Downloaded from jcme.iut.ac.ir at 14:27 IRDT on Saturday April 27th 2019

www.SID.ir