فصلنامهٔ پژوهش‌های پولی ـ بانکی
سال پانزدهم، شمارهٔ ۵۱، بهار ۱۴۰۱
صفحات ۱۲۱ ـ ۱۵۴

شناسایی الگوهای پویایی ریسک اعتباری مشتریان حقیقی بانک‌ها و مؤسسات مالی (مطالعهٔ موردی: سه بانک ایرانی)

صبا مرادی [*] **فریماه مخاطب رفیعی** [†]
عباس سقایی [‡]

تاریخ دریافت: ۱۳۹۸/۱۰/۰۷ تاریخ پذیرش: ۱۴۰۰/۰۶/۲۴

چکیده

بررسی ریسک اعتباری مشتریان از مهم‌ترین دغدغه‌های بانک‌ها است. افزایش مطالبات معوق نشان می‌دهد که مدل‌های مرسوم نتوانسته‌اند ریسک اعتباری مشتریان را به خوبی ارزیابی کنند. عوامل متغیر و غیرقطعی محیطی بر رفتار مشتری اثر می‌گذارد و موجب تغییر ریسک اعتباری مشتری در طول زمان می‌شود، لذا می‌بایست مدلی پویا برای ارزیابی ریسک اعتباری مشتری طراحی شود. پژوهش‌های گذشته از مدل‌های متداولی مانند مدل‌های مالی، شبکه‌های عصبی، و مبتنی بر داده‌های تجربی استفاده کرده‌اند. در این پژوهش با درنظرگرفتن جابه‌جاشدن مشتری در طول زمان در گروه‌های مختلف کم‌ریسک، ریسک متوسط، و ریسک بالا و با به‌کارگیری روش‌های تحلیل داده‌های بزرگ، الگوهای رفتار پویای مشتریان استخراج شده است. برای این منظور، اطلاعات مشتریان حقیقی سه بانک مورد بررسی قرار گرفت و نتایج نشان داد که مدل با درجهٔ حساسیت ۹۲/۱ درصد و درجهٔ تشخیص ۸۹/۱ درصد از کارایی بالایی برخوردار است و در مقایسه با مدل‌های پروبیت، لاجیت، و شبکهٔ عصبی بهتر عمل می‌کند و کارایی بالاتری دارد. دستاورد این پژوهش، تشخیص الگوهای پویای ریسک اعتباری مشتریان است. این روش به این جهت که به آموزش مجدد داده‌ها نیازی ندارد، موجب صرفه‌جویی در زمان و هزینه می‌گردد.

واژه‌های کلیدی: مطالبات معوق، نکول، پویایی رفتار مشتریان، ابزارهای تحلیل دادهٔ‌های بزرگ، الگوهای پویای ریسک اعتباری، خوشه‌بندی پویا.
طبقه‌بندی JEL: C53، C55

---

[*] دکتری، دانشکده مهندسی صنایع، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران؛ saba.moradi@srbiau.ac.ir
[†] دانشیار، دانشکده مهندسی صنایع و سیستم ها، دانشگاه تربیت مدرس،(نویسندهٔ مسئول)؛ f.mokhatab@modares.ac.ir
[‡] دانشیار، دانشکده مهندسی صنایع، دانشگاه علوم و تحقیقات؛ a.saghaei@srbiau.ac.ir

۱ مقدمه

بانک‌ها با عوامل ریسک مختلفی مواجه‌اند که باید آن‌ها را شناسایی، درک، اندازه‌گیری، و مدیریت کنند. مطابق با پیمان بازل ۲[1]، ریسک اعتباری یکی از ریسک‌هایی است که بانک‌ها در تخصیص منابع با آن مواجه‌اند. تامایو[2] (۲۰۰۰) ریسک اعتباری را با توجه به حوزهٔ موردبررسی این تحقیق، احتمال عدم بازپرداخت یا پرداخت با تأخیر تسهیلات بانک‌ها از سوی مشتری تعریف می‌کند، اما در ابعادی وسیع‌تر ون جستل و همکاران[3] (۲۰۰۵) ریسک اعتباری را ریسک نکول قرض‌گیرنده تعریف می‌کنند؛ به این معنی که قرض‌گیرنده نتواند به تعهدات خود برای پس‌دادن قرض عمل کند یا حداقل بموقع آن را پرداخت نکند. تاکنون مدل‌های بسیاری از دسته مدل‌های آماری و تجربی و روش‌هایی با تکیه بر هوش مصنوعی برای ارزیابی ریسک اعتباری ارائه شده است. بررسی این مدل‌ها نشان می‌دهد همهٔ آن‌ها یک ویژگی مشترک دارند یا به‌بیان‌دیگر همهٔ آن‌ها نکته‌ای مهم را از قلم انداخته‌اند. ممکن است مشتری که رتبهٔ اعتباری خوبی کسب کرده است، به‌دلایل مختلف شخصی، اجتماعی مانند رکود اقتصادی یا تحریم، رفتار اعتباری متمایزی از خود نشان دهد. از مصادیق این حالت می‌توان به دورهٔ تحریم در کشورمان اشاره کرد که شرایط غیرنرمال و بسیاری از پیش‌بینی‌ها اشتباه شد. پس این نکته مهم است که اساساً رفتار مشتری ماهیتی پویا دارد؛ به این معنی که در گذر زمان تغییر می‌کند و ایستا نیست و این موضوع باید در ارائهٔ مدل ارزیابی ریسک اعتباری لحاظ شود. مطالعهٔ حاضر در روش پیشنهادی خود به مشتری به‌عنوان عنصری پویا که رفتار او با گذشت زمان تغییر می‌کند، توجه می‌کند و مدلی ارائه می‌دهد که با بروز شرایط مختلف و در شرایط غیرقطعی، ریسک اعتباری مشتری را درست و با خطای بسیار کم پیش‌بینی می‌کند.

در این تحقیق، برخلاف تحقیقات دیگر (سوسا و همکاران[4] (۲۰۱۶)، نظرآقایی و همکاران (۱۳۹۸)، قاسمی و همکاران (۱۳۹۵)، ابراهیمی و همکاران (۱۳۹۱)) که مشتریان به دو گروه خوب و بد تفکیک می‌شوند، مشتریان به سه گروه کم‌ریسک، ریسک متوسط، و ریسک بالا دسته‌بندی می‌شوند که ویژگی‌های هریک از گروه‌ها در ادامهٔ تحقیق آورده شده است و پس از اثبات اینکه مشتریان بانک در طول زمان میان این گروه‌ها جابه‌جا می‌شوند و به‌عبارت‌دیگر،

---

[1] Basel 2 Capital Accord

[2] Tamayo

[3] Van Gestel et al.

[4] Sousa et al.

رفتاری پویا از خود نشان می‌دهند، نتیجه تحقیق ما را به خوشه‌های متنوع‌تری می‌رساند که با ارائه و اتخاذ تصمیمات مدیریتی برای رفتار با هریک از خوشه‌های مشتریان، در واقع، ابزاری برای مدیریت ریسک اعتباری ارائه می‌شود. نتیجهٔ بررسی این تحقیق نشان می‌دهد مشتریان در طول زمان از حیث ریسک اعتباری پویا عمل می‌کنند؛ اما در همین پویایی از قوانینی پیروی می‌کنند که موجب به‌وجودآمدن خوشه‌های متنوع‌تر نسبت‌به تحقیقات پیشین مشابه خود از حیث موضوع می‌شود و نیز قوانین ذکرشده موجب می‌شود بانک بتواند به پیش‌بینی درستی نسبت‌به مشتری دست یابد و چون طیفی از رفتارها را برای مشتری پیش‌بینی می‌کند، می‌تواند نسبت‌به مدل‌هایی که تنها مشتریان را به دو دستهٔ خوب و بد (خوش‌حساب و بدحساب) تفکیک می‌کنند، دقیق‌تر و منعطف‌تر عمل کند. در مقالهٔ حاضر، برای اولین‌بار با استفاده از خوشه‌بندی گراف مدلی پویا برای ارزیابی ریسک اعتباری مشتری ارائه می‌شود.

پرسش این تحقیق این است که آیا مدل ارائه‌شده در این تحقیق نسبت به مدل‌های ایستا ریسک اعتباری را صحیح‌تر شناسایی می‌کند؟ فرضیهٔ تحقیق این است که مدل طراحی‌شده توانایی پیش‌بینی ریسک اعتباری به‌شکل مطلوب را دارد.

روش گردآوری اطلاعات میدانی است و از پایگاه‌دادهٔ سه بانک ایرانی اخذ شده است. در بخش ۲ این مقاله، پیشینهٔ تحقیق‌های انجام‌شده و نقد آن‌ها ارائه می‌شود. در بخش ۳ روش‌شناسی تحقیق و در بخش ۴ جمع‌آوری و پیش‌پردازش داده‌ها توضیح داده می‌شود. در بخش ۵، تحلیل نتایج انجام می‌شود. نهایتاً در بخش ۶، نتیجه‌گیری و بحث در خصوص تحقیق و در بخش ۷، محدودیت‌های تحقیق و پیشنهادهای آتی ارائه می‌شود.

۲ پیشینهٔ تحقیق

تکنیک‌های امتیازدهی اعتباری به‌عنوان ابزار پشتیبان تصمیم یا الگوریتم تصمیم‌گیری خودکار برای طیفی وسیع از مشتریان بانک‌ها استفاده می‌شود (هیات[1]، ۲۰۱۲). اهمیت مدل‌های امتیازدهی اعتباری موجب شده است تا مؤسسات اعتباری مدل‌های امتیازدهی داخلی اعتباری را توسعه دهند (وانگ[2] و همکاران، ۲۰۱۴). به‌دلیل فقدان چهارچوب اعتباری، مدل‌های ریسک اعتباری فعلی بر پایهٔ سعی و خطا هستند و تحلیل ریسک اعتباری

---

[1] Heiat

[2] Wang

مسئلهٔ شناسایی الگوست (کی. راپا[1] و همکاران، ۲۰۱۳). در این زمینه، زانین[2] و همکاران (۲۰۱۶) تلاش کرده‌اند که مسئلهٔ ارزیابی ریسک اعتباری را با رویکردها و تکنیک‌های مختلف حل کنند. در هریک از تحقیقات، آن‌ها تلاش کرده‌اند مدلی با دقت بیشتر از مدل قبلی ارائه کنند. یادگیری ماشینی و تکنیک‌های داده‌کاوی معمولاً در این حوزه استفاده می‌شود. روش‌های متعددی با توانایی‌های متفاوت برای داده‌کاوی وجود دارد؛ مانند درخت تصمیم[3] و شبکه‌های عصبی[4]، ماشین بردار پشتیبان[5]، نزدیک‌ترین همسایه[6]، شبکه‌های بیزین[7]، الگوریتم‌های بر پایهٔ نمونه[8]، و سیستم‌های دسته‌بندی آموزشی[9]. آن‌ها اساساً سه دسته‌اند (جیرجیا و همکاران[10] ، ۲۰۰۶): آماره‌های کلاسیک[11]، هوش مصنوعی[12]، و یادگیری ماشینی[13] (روزن بلات[14]، ۱۹۵۸). ونگ و همکاران (۲۰۱۱) مجموعه داده‌های اعتباری را با استفاده از شبکه‌های عصبی تحلیل کردند. نتایج نشان داد که روش بگینگ[15] بهترین بازدهی را داشته است.

بکحت[16] و همکاران (۲۰۱۴) مدلی با رویکرد شبکهٔ عصبی برای ارزیابی ریسک اعتباری در بانک تجاری در اردن ارائه دادند. نتایج این تحقیق نشان می‌دهد که مدل رگرسیون

---

[1] Kruppa
[2] Zanin
[3] Decision Tree
[4] Artificial Neural Networks
[5] Support Vector Machines
[6] K-nearest Neighbors
[7] Baysian Networks
[8] Instance Base
[9] Learning Classifier Systems
[10] Girjia et al.
[11] Classical Statistics
[12] Artificial Intelligence
[13] Machine Learning
[14] Rosenblatt
[15] Bagging
[16] Hussain Ali Bekhet

لاجیت[1] از مدل رادیال[2] عملکرد بهتری از لحاظ دقت و صحت دارد. ماندالا[3] و همکاران (۲۰۱۲) فاکتورهایی برای ارزیابی ریسک اعتباری در بانکی غیرشهری شناسایی کردند. در این تحقیق با استفاده از روش‌شناسی C5.0، مدل جدید درخت تصمیم ایجاد شده است که این مدل معیاری جدید در تحلیل تقاضای وام پیشنهاد می‌کند.

زمانی (۲۰۱۱) دربارهٔ الگوی رفتار مشتریان تحقیق کرده است. او شبکه‌های پرسپترون چندلایه[4] و شبکه‌های عصبی با توابع پایهٔ شعاعی[5] را مقایسه کرده است. نتایج تحقیق او نشان می‌دهد، شبکه‌های پرسپترون چندلایه در پیش‌بینی ریسک اعتباری مشتریان بهتر از شبکه‌های با توابع پایه شعاعی است. داننـاس[6] و همکاران (۲۰۱۵) روش ارزیابی ریسک اعتباری ارائه کرده‌اند. این روش متشکل است از ماشین بردار پشتیبان، ارزیابی خارجی، و پنجره‌ای که به بخش‌های مختلف تقسیم شده است. نتایج نشان داده است که مدل پیشنهادی قابل‌مقایسه با طبقه‌بندی‌کننده‌هایی مانند شبکه‌های با توابع پایهٔ شعاعی و رگرسیون لاجیت[7] است.

پینگ[8] و همکاران (۲۰۱۱) مدل هیبریدی ماشین بردار پشتیبان برای ارزیابی ریسک اعتباری مشتری ارائه کردند، نتیجهٔ مدل آن‌ها نشان داد که ماشین بردار پشتیبان بر مبنای طبقه‌بندی هیبریدی و نایبرهود راف[9] بهترین توانایی امتیازدهی اعتباری در مقایسه با سایر طبقه‌بندی کننده‌های هیبریدی را دارند. استفاده از طبقه‌بندی کننده‌های ترکیبی یکی از دغدغه‌های تحقیقات اخیر در یادگیری ماشینی است.

اچ. سیه[10] و همکاران (۲۰۱۰) یک مرحله پیش‌پردازش برای به‌دست‌آوردن طبقه‌بندی‌کنندهٔ گروهی مؤثر معرفی کرده‌اند. آن‌ها این طبقه‌بندی کننده را با استفاده از تکنیک‌های متعدد داده‌کاوی نظیر شبکه‌های عصبی، ماشین بردار پشتیبان و شبکه‌های

---

[1] Logistic
[2] Radial
[3] Narindra Mandalaa
[4] Multi-Layer Perceptron
[5] Radial Bases Function, RBF
[6] Danenas
[7] Logit Regression
[8] Ping
[9] Neighborhood Rough
[10] Hsieh

بیزین[1] ارائه کرده‌اند. نظرآقایی و همکاران (۲۰۱۹) در مدلی عوامل مهم و مؤثر در رفتار اعتباری مشتریان را ارائه داده‌اند. در این مدل از شبکهٔ عصبی و درخت تصمیم فازی استفاده شده است. نتایج تحقیق آن ها نشان می‌دهد که درآمد و تراکنش‌های مالی مشتریان بالاترین اهمیت را در تعیین ریسک اعتباری مشتریان دارد و نیز درخت تصمیم فازی با روش بگینگ از دقت بالاتری در مقایسه با روش شبکهٔ عصبی و درخت تصمیم فازی معمولی برخوردار است.

سوسا و همکاران (۲۰۱۶) چهارچوبی پویا برای ریسک اعتباری مشتریان ارائه کرده‌اند. پردازش داده‌ها در مدل به‌صورت دسته‌ای انجام می‌شد. آن‌ها مدل افزایش تعمیم یافته[2] را به‌عنوان روش طبقه‌بندی از نوع آموزش با ناظر انتخاب کردند. واحدهای آموزشی واحدهای ایستا بودند. در این تحقیق از ضریب Gini برای اندازه‌گیری کارایی استفاده شده است. نمونه‌های آموزش هرماه از داده‌های ماه قبل انتخاب می‌شد. این مدل دو روش برای بررسی نکول دارد: ۱ـ پنجرهٔ حافظه کامل که بر اساس داده‌های قبل است و داده‌های ماه جدید به مجموعهٔ در حال آموزش اضافه می‌شوند. این نما نمی‌تواند با تغییرات بزرگ سازگار شود و ۲ـ پنجرهٔ حافظهٔ کوتاه‌مدت ثابت که نمونه‌های قبلی را فراموش می‌کند و دلیل آن این است که محقق به این نتیجه رسیده است که ارتباط کمّی بین نکول زمان حاضر و نمونه‌های گذشته وجود دارد.

نتیجهٔ بررسی‌های انجام‌شده در تحقیقات ریسک اعتباری و ارزیابی آن نشان می‌دهد تحقیقاتی که در این زمینه انجام شده است در مجموع به سه دسته تقسیم می‌شود. به‌بیان‌دیگر، محققان از ابعاد و زوایای ذیل به این موضوع نگریسته‌اند:

۱ـ بررسی عوامل مؤثر در ریسک اعتباری؛

۲ـ تأثیر ریسک اعتباری در سودآوری و بهره‌وری بانک‌ها و بررسی کیفیت وام‌های بانکی؛ و

۳ـ ارزیابی ریسک اعتباری با استفاده از داده‌کاوی.

از مجموع مدل‌های بررسی‌شده در بخش پیشینهٔ تحقیق، ۹۶ درصد از مدل‌هایی که تاکنون ارائه‌شده‌اند، ایستا هستند. ۴ درصد باقی‌ماندهٔ مدل‌ها که پویا هستند نیز با درنظرگرفتن فاکتورهای قطعی میزان پویایی خود را به چالش کشیده‌اند. این مدل‌ها به‌لحاظ مدل‌سازی و پیاده‌سازی کمبودهایی داشتند و هنوز نتوانستند مدلی ارائه کنند که رفتار مشتری را متأثر از شرایط مختلف محاط بر او و ریسک اعتباری را پیش‌بینی کند، بنابراین

---

[1] Bayesian

[2] Generalised Additive Model (GAM)

استفاده از تکنیک‌های معمول داده‌کاوی با توجه به ماهیت پویای مسئلهٔ ریسک اعتباری کارا نیست و نمی‌تواند پیش‌بینی‌های دقیق انجام دهد. ۸۸ درصد از تحقیقات از مدل‌های هوش مصنوعی و تکنیک‌های داده‌کاوی و ۱۲ درصد از مدل‌های هیبریدی استفاده کرده‌اند.

در این مقاله برای اولین بار مدلی ارائه شده است که به تغییر رفتار مشتری (پویایی) در طول زمان توجه می‌کند. برای این منظور، جابه‌جاشدن مشتری در طول زمان در گروه‌های ریسک اعتباری مختلف کم‌ریسک (متشکل از دو گروه خیلی کم‌ریسک L1 و کم‌ریسک L2)، ریسک متوسط (متشکل از دو گروه با نماد M1 و M2 با تفاوت دو ماه در تأخیر بازپرداخت)، ریسک بالا (متشکل از سه گروه با نمادهای H1، H2، و H3 با تفاوت شش ماه در تأخیر بازپرداخت) با استفاده از خوشه‌بندی پویا، گراف‌کاوی و شناسایی قواعد و الگوهای رفتاری و تکنیک کشف قواعد انجام می‌شود که روشی کاراتر است و به پیش‌بینی‌های دقیق‌تری می‌انجامد؛ بنابراین این تحقیق چه در بیان مسئله چه در دیدگاه پویا به رفتار سیستم و چه در استفاده از تکنیک‌های داده‌کاوی برای مدل‌سازی مسئله نوآوری دارد. قواعد به‌دست‌آمده از مدل‌سازی و داده‌کاوی به‌صورت اگر–آنگاه مطرح می‌شود که برای ارزیابی یا به عبارتی پیش‌بینی ریسک اعتباری استفادهٔ مدیران و کارشناسان اعتباری آسان، کاربردی، و قابل‌فهم است.

به‌طور خلاصه، خلأهای تحقیقاتی که در این مقاله به آن توجه می‌شود، عبارت‌اند از:

۱ــ تعقیب رفتار اعتباری مشتری (نحوهٔ بازپرداخت تسهیلات توسط مشتری) در طول زمان در قالب دنباله‌های انتقال که در تحقیقات پیشین بررسی نشده است؛

۲ــ بررسی الگوهای مشتریان در تغییر وضعیت اعتباری؛

۳ــ تحلیل و بررسی الگوهای مشتریان در بازپرداخت تسهیلات در هر بازه زمانی و تأثیر آن در مدیریت ریسک اعتباری؛ و

۴ــ شناسایی الگوی پویای مشتریان با ریسک اعتباری بالا و آسیب‌رسان به بانک.

۳ روش‌شناسی تحقیق

۳. ۱ رویکرد

این تحقیق از نوع کیفی–کمّی است. راهبرد این تحقیق مطالعهٔ موردی است و هدف آن توصیفی–اکتشافی است. برای پیاده‌سازی طرح پیشنهادی این تحقیق، جدولی از مشتریان سال ۱۳۸۹ تا ۱۳۹۶ در نظر گرفته و به ۳۲ بازه زمانی تقسیم شده است. فاصلهٔ زمانی هر بازه با بازه قبلی سه ماه است. پس‌ازآن، وضعیت اعتباری مشتریان در هریک از بازه‌ها شناسایی

می‌شود. وضعیت اعتباری مشتریان به این صورت تعریف می‌شود که مشتریانی که کمتر از ۲ ماه تأخیر در بازپرداخت دارند، کم‌ریسک هستند و با L نشان داده می‌شوند. مشتریانی که بین ۲ تا ۶ ماه تأخیر بازپرداخت دارند، ریسک متوسط دارند و با M نشان داده می‌شوند و مشتریانی که بیش از ۶ ماه تأخیر بازپرداخت دارند، ریسک بالا دارند و با H نشان داده می‌شوند. در تقسیم جزئی‌تر، مشتریانی که هیچ‌گاه تأخیر نداشته‌اند، با نماد L1، مشتریانی که ۱ یا ۲ ماه تأخیر دارند، با نماد L2، مشتریانی که بین ۲ تا ۴ ماه تأخیر دارند، با نماد M1، مشتریانی که ۵ یا ۶ ماه تأخیر دارند، با نماد M2، مشتریانی که بین ۶ تا ۱۲ ماه تأخیر دارند، با نماد H1، مشتریانی که بین ۱۲ تا ۱۸ ماه تأخیر دارند، با نماد H2، و مشتریانی که بیش از ۱۸ ماه تأخیر دارند، با نماد H3 نمایش داده می‌شوند.

پس از تعیین وضعیت اعتباری مشتری در هریک از بازه‌های زمانی، برای هر مشتری در زمانی که با بانک قرارداد دارد، دنبالهٔ انتقال وضعیت اعتباری تشکیل می‌شود. در مرحلهٔ بعد، دنباله‌های انتقال مشتریان خوشه‌بندی می‌شود. خوشه‌های تولیدشده مشتریانی هستند که در تغییر وضعیت اعتباری رفتار یکسانی داشته‌اند. برای خوشه‌بندی دنباله‌های انتقال از روش خوشه‌بندی پویا استفاده می‌شود. خوشه‌های به‌دست‌آمده با استفاده از روش تحلیل k ـ برترین قواعد و الگوریتم GSP تفسیر و توصیف می‌شوند. در مرحلهٔ آخر، نتایج به خبرگان و کارشناسان ریسک اعتباری ارائه و طرح‌ها و کاربردهای مدیریتی ارائه می‌شود.

۳. ۲ روش پویای ارزیابی ریسک اعتباری

مرور تحقیقات انجام‌شده نشان می‌دهد که از روش‌هایی مانند شبکه‌های عصبی، ماشین بردار پشتیبان، و درخت تصمیم‌گیری برای محاسبه و پیش‌بینی ریسک اعتباری مشتریان استفاده شده است. این مدل‌ها در شرایط قطعیت و ایمن با درصدی از خطا کاربرد داشته‌اند، اما نمی‌توانند در شرایط عدم‌قطعیت درست پیش‌بینی کنند. شرایطی مانند تحریم یا تورم شدید که در آن محاسبات اقتصادی به‌هم می‌خورد عدم‌قطعیت گفته می‌شود. تحقیقات نشان می‌دهد (مرادی، ۱۳۹۷) بیش از ۷۳ درصد مطالبات معوق دارای وثیقهٔ سنگین بوده‌اند. این موضوع از این حیث بانکداران را نگران می‌کند که حتی اخذ وثیقهٔ سنگین که درحال‌حاضر یکی از متداول‌ترین روش‌ها برای مدیریت ریسک اعتباری مشتری است نیز نمی‌تواند تضمینی بر بدون‌ریسک‌بودن مشتری باشد؛ بنابراین، بانک‌ها مشتاقانه از مدلی که بتواند ریسک اعتباری را در همهٔ زمان‌ها درست تخمین بزند و به‌عبارت‌دیگر، تابع زمان باشد و بتواند تغییرات زمانی را در خود لحاظ کند، استقبال می‌کنند. در روشی که در این پژوهش ارائه می‌شود، وضعیت اعتباری مشتریان در بازه‌های زمانی سه‌ماهه بررسی و دنباله‌های انتقال

مشتریان خوشه‌بندی شد. خوشه‌های مشتریان تحلیل و تفسیر شدند و با کمک خبرگان بانکی، خوشه‌ها توصیف شدند و نتایج چشمگیری به‌دست آمد. تشکیل دنباله‌های انتقال مشتریان تاکنون به عنوان بهترین روش برای انعکاس پویایی شناخته شده است. برای به‌دست‌آوردن قواعد دنباله‌ای هریک از خوشه‌ها از روش GSP و k-برترین قواعد استفاده شده است.

روش پیشنهادی این تحقیق از چهار مرحله تشکیل شده است که در ادامه به توضیح آن‌ها پرداخته شده است.

۳. ۲. ۱ گام اول

مرحلهٔ اول شامل جمع‌آوری داده و پیش‌پردازش آن‌هاست. داده‌های اولیهٔ این تحقیق از سه بانک ایرانی به‌دست آمده است. در مرحلهٔ پیش‌پردازش که یکی از مهم‌ترین مراحل پیش از داده‌کاوی است، به مواردی مانند نرمال‌سازی، داده‌های حذف‌شده، یا تکراری پرداخته شده است.

۳. ۲. ۲ گام دوم

مشتریان در طول زمان در وضعیت‌های مختلف ریسک اعتباری قرار می‌گیرند. لذا در مرحلهٔ آغازین تحقیق سعی می‌شود این موضوع در ارائهٔ مدل لحاظ شود. ابتدا، وضعیت ریسک اعتباری مشتریان را در بازه‌های زمانی سه‌ماهه و بر اساس تعریف بازل ۲ تعیین می‌کنیم. مشتری کمتر از ۲ ماه بازپرداخت عقب‌افتاده داشته باشد، کم‌ریسک است و با L (به‌طور جزئی‌تر با L1 و L2) نشان داده می‌شود. مشتریانی که بین ۲ تا ۶ ماه تأخیر بازپرداخت دارند، ریسک متوسط دارند و با M (به‌طور جزئی‌تر بین ۲ تا ۴ ماه تأخیر با M1 و بین ۴ تا ۶ ماه تأخیر با M2) نشان داده می‌شوند، و مشتریانی که بیش از ۶ ماه تأخیر بازپرداخت دارند ریسک بالا دارند و با H (به‌طور جزئی‌تر بین ۶ تا ۱۲ ماه با H1، بین ۱۲ تا ۱۸ ماه با H2، و بیشتر از ۱۸ ماه با H3) نشان داده می‌شوند. چنانچه برای هر مشتری در بازه زمانی سه‌ماهه وضعیت را با L، M، و H نشان دهیم. هر مشتری در طول زمان قرارداد خود با بانک، مالک رشته‌ای از کاراکترهاست مثلاً به این صورت H1-L1-M2-M2-L1-H3. این رشته‌ها در این تحقیق دنبالهٔ وضعیت ریسک اعتباری مشتری یا دنبالهٔ انتقال مشتری نامیده می‌شود. به این صورت توصیف پویایی ریسک اعتباری مشتری در گذر زمان توسط دنباله‌ای یا رشته‌ای از کاراکترها نمایش داده می‌شود.

۳. ۲. ۳ گام سوم

چنانچه در دنبالهٔ وضعیت اعتباری مشتری ملاحظه می‌شود، مشتری از گروه کم‌ریسک به گروه ریسک متوسط یا از گروه ریسک متوسط به گروه ریسک بالا جابه‌جا می‌شود. در این مرحله، به‌دنبال کشف یا استخراج الگوهای جابه‌جاشدن مشتریان بین گروه‌های مختلف ریسک اعتباری هستیم. لذا، برای استخراج الگوهای پویایی و مهاجرت مشتریان بین وضعیت‌های مختلف، دنباله‌های به‌دست‌آمده از وضعیت ریسک اعتباری مشتریان خوشه‌بندی می‌شود. هریک از خوشه‌ها بیانگر مشتریانی است که در رفتار اعتباری خود نسبت‌به بانک مشابه یکدیگر عمل کرده‌اند؛ یعنی به‌طور مثال، مشتریانی که در یک خوشه هستند، در جابه‌جاشدن بین گروه‌های کم‌ریسک، ریسک متوسط، و ریسک بالا مانند یکدیگر عمل کرده‌اند. نکته اینکه رفتار خوشه‌های مختلف با یکدیگر متفاوت است. روش‌های خوشه‌بندی سلسله‌مراتبی، تفکیکی، و خوشه‌بندی مبتنی بر گراف برای خوشه‌بندی داده‌های دنباله‌ای توصیه شده است (دانگ و همکاران، ۲۰۰۷) در این تحقیق از روش خوشه‌بندی تفکیکی استفاده می‌شود.

۳. ۲. ۴ گام چهارم

در این مرحله به‌منظور تفسیر و توصیف خوشه‌های مشتریان به‌دست‌آمده در گام سوم، به بررسی و کاوش قوانین پرداخته می‌شود. این موضوع در حوزهٔ داده‌کاوی به استخراج قوانین انجمنی و کشف الگوهای دنباله‌ای شناخته شده است که در این تحقیق، در بخش مبانی نظری در مورد آن‌ها توضیح داده شد. هرکدام از این قواعد نشان‌دهندهٔ الگوی غالب تغییر وضعیت ریسک اعتباری در سطح کل مشتریان است. برای به‌دست‌آوردن قواعد دنباله‌ای از روش k ـ برترین قواعد استفاده می‌شود. در این روش، در هر خوشه ابتدا قواعد با اطمینان بالای آستانه استخراج می‌شود و سپس k قاعده‌ای که مقدار پشتیبان بالاتری دارد، انتخاب می‌شود؛ k مقداری است که توسط کاربر انتخاب می‌شود. در مورد اصطلاح اطمینان و پشتیبان در بخش مبانی نظری تحقیق شرح مبسوط داده شده است. در استخراج قواعد دنباله‌ای از الگوریتم اپریوری و GSP استفاده می‌شود. در نهایت، قواعد به‌دست‌آمده همراه با مقدار پشتیبان و اطمینان در سطرهای مجزا نمایش داده می‌شود.

۳. ۲. ۵ گام پنجم

پس از انجام‌دادن گام‌های فوق، در گام نهایی نتایج به‌دست‌آمده به خبرگان ریسک اعتباری ارائه می‌شود و خوشه‌های مشتریان تفسیر و توصیه‌های مدیریتی برای هریک از گروه‌های مشتریان اخذ می‌شود.

۴ جمع‌آوری و پیش‌پردازش داده

در این بخش، پیاده‌سازی مدل پیشنهادی با استفاده از داده‌های بانک‌های ایرانی فوق‌الذکر شرح داده می‌شود و نتایج حاصل از بررسی مسئلهٔ این تحقیق در شرایط واقعی ارائه می‌شود. در سطرهای بعدی ابتدا در خصوص جامعهٔ آماری، متغیرهای تحقیق، و آماده‌سازی داده‌ها توضیح داده می‌شود و در ادامه، چگونگی پیاده‌سازی مدل و تحلیل نتایج ارائه می‌شود.

۴. ۱ جامعه و نمونهٔ آماری

جامعهٔ آماری این تحقیق مشتریان حقیقی سه بانک دولتی است که به آن‌ها تسهیلات اعطا شده است.

نمونه‌های (داده‌های) به‌کار گرفته‌شده در این پژوهش، داده‌های مربوط به پروفایل مشتریان و تسهیلات اعطاشده به ایشان در قالب قرارداد منعقدشده با بانک است. نمونهٔ آماری این تحقیق اطلاعات مربوط به ۲,۸۴۰,۰۰۰ مشتری برای ۳۲ دورهٔ سه‌ماهه از سال ۱۳۸۹-۱۳۹۶ بوده است. برای نمونه‌گیری از بین تسهیلات اعطایی، روش نمونه‌گیری تصادفی در فاصلهٔ سال‌های ۱۳۸۹ تا ۱۳۹۶ به‌کار گرفته شده است. برای پیش‌پردازش و آماده‌سازی داده‌ها حدود ۲ ماه زمان صرف شده است. گفتنی است این پژوهش در استفاده از داده‌های بزرگ از میان پژوهش‌هایی ازاین‌دست، منتخب به‌حساب می‌آید.

کیفیت این داده‌ها جز در مواردی محدود مشکلی نداشت. به‌دلیل اینکه داده‌های این پایگاه‌داده حاصل به‌کارگیری سامانهٔ جامع اطلاعاتی تسهیلات است، مشکلات مربوط به داده مانند مفقودشدن یا غیرقابل‌پذیرش‌بودن وجود نداشته است. آنچه موجب بالارفتن زمان آماده‌سازی داده‌ها شد، پراکندگی مشخصه‌ها در جداول مختلف پایگاه‌داده بود که برای تسریع کار، جداول جدید ایجاد و ارتباط منطقی بین آن‌ها برقرار شد. در این پژوهش برای انتخاب فیلدهای موردنظر جهت به‌کارگیری در مدل، با خبرگان و کارشناسان ریسک اعتباری مصاحبه و با توجه به هدف پژوهش، ۳۰ فاکتور انتخاب شد. مشخصه‌ها در جدول ۱ قابل‌مشاهده است.

۴. ۲ روش‌ها و ابزار گردآوری اطلاعات

ابزار گردآوری اطلاعات در این تحقیق پایگاه‌دادهٔ مشتریان و پرداخت‌های آن‌ها در سامانهٔ تسهیلات است. در این تحقیق برای تجزیه‌وتحلیل داده‌ها از زبان برنامه‌نویسی پایتون و محیط ژوپیتر استفاده شده است. اطلاعات از طریق فایل‌های خروجی از اوراکل در اختیار پایتون قرار گرفته است.

۳.۴ متغیرهای تحقیق

مجموعه دادهٔ این مطالعه، شامل ۳۰ مشخصه است که اطلاعات دقیق مربوط به مشتریان و تسهیلات دریافتی را داراست. در این داده‌ها اطلاعات مربوط به مشتری ازجمله میزان درآمد، تعداد درخواست‌های ردشده، تعداد چک‌های برگشتی، و اطلاعات مربوط به قرارداد مانند نرخ سود، نرخ جریمه، هدف از دریافت تسهیلات، مبلغ قسط، تاریخ پرداخت، مبلغ وثیقه، فاصله بین اقساط و تعداد اقساط، و سایر اطلاعات دقیق در خصوص مشتری و تسهیلات موجود است. مشخصه‌ای که بر اساس آن وضعیت اعتباری مشتری در بازه‌های زمانی مختلف تعیین شده است، تعداد ماه‌های ازسررسیدگذشته است.

جدول ۱

مشخصات انتخاب‌شدهٔ مشتری

ردیف	گروه مشخصه	تعریف مشخصه
۱	سن	سن مشتری در زمان ارزیابی ریسک اعتباری
۲	آدرس	نشانی مشتری در زمان ارزیابی ریسک اعتباری
۳	تاریخ قرارداد	تاریخ درخواست مشتری و افتتاح موقت قرارداد
۴	طول قرارداد	مدت زمان قرارداد برای تسهیلات مشارکتی از ۱۲ ماه تا ۱۸ ماه قابل‌تمدید تا ۲۴ یا ۳۶ ماه
۵	درصد سهم	درصد سهم مربوط به قراردادهای مشارکتی، مضاربه‌ای، یا فروش اقساطی
۶	کل هزینه	کل هزینهٔ برآوردی برای قرارداد توسط بانک
۷	تعداد واحد	تعداد واحد ملک در خصوص قراردادهای مشارکت
۸	ارزش اعیانی	ارزش اعیانی ملک زیربنا
۹	مبلغ وام	میزان اعتبار تخصیص‌یافته از سوی بانک به قرارداد مشتری که همان مبلغ وام است.
۱۰	نرخ سود	نرخ ماهیانهٔ بازگردانده‌شده توسط مشتری که علاوه بر اصل وام باید به بانک پرداخت شود.
۱۱	نرخ جریمه	نرخ جریمهٔ عدم‌بازپرداخت
۱۲	هدف از دریافت	ایجاد، تأمین سرمایهٔ در گردش، تعمیر، توسعه، و خرید
۱۳	محل مصرف	شهر محل مصرف تسهیلات تخصیص‌یافته از سوی بانک به مشتری
۱۴	تعداد اقساط	تعداد اقساط وام در زمان قرارداد
۱۵	زمان اقساط	معمولاً یک‌ماهه
۱۶	مبلغ قسط	مبلغ قابل‌پرداخت ماهیانه به بانک
۱۷	سود	بر اساس نرخ سود قرارداد در ابتدای هرسال

۱۸	تاریخ	تاریخ افتتاح قرارداد با مشتری
۱۹	متوسط موجودی	متوسط ماهانه موجودی حساب مشتری
۲۰	مبلغ سپرده	حساب ذخیرهٔ مشتری
۲۱	تعداد کل قراردادهای مشتری	تعداد کل قراردادهای منعقدشده بین بانک و مشتری
۲۲	مبلغ وثیقه	مبلغ وثیقهٔ منقول یا غیرمنقول
۲۳	کل مبلغ پرداخت‌نشده	کل میزان بدهی پرداخت‌نشده توسط مشتری
۲۴	تعداد استعلام‌ها	تعداد استعلام‌های سال جاری از مشتری
۲۵	تعداد درخواست‌های ردشده	تعداد درخواست‌های ردشدهٔ مشتری تاکنون
۲۶	میزان درآمد	میزان درآمد مشتری (نکته: اغلب این شاخص بر اساس خوداظهاری فرد است.)
۲۷	روزهای بعد از سررسید	میانگین تعداد روزهایی بعد از موعد مقرر پرداخت‌شده اقساط
۲۸	تاریخ سررسید سپرده	تاریخ سررسید قراردادهای سپرده
۲۹	تعداد چک برگشتی	تعداد چک‌های وصول‌نشدهٔ مشتری
۳۰	نوع مشتری	مشتری حقیقی و مشتری حقوقی

یادداشت. این جدول شامل مشخصه‌های مشتریان و تسهیلات دریافتی آن‌هاست. در ابتدای پژوهش ۸۰ مشخصه در نظرگرفته شد، ولی پس از مذاکره با خبرگان و با توجه به هدف پژوهش، ۳۰ مشخصه انتخاب و استفاده شد.
منبع:محاسبات تحقیق

۴. ۴ تغییر ریسک مشتریان در طول زمان

در این مرحله، وضعیت ریسک اعتباری مشتری در بازه‌های زمانی سه‌ماهه تعیین شد. تعیین وضعیت ریسک اعتباری بر اساس تعریف بازل ۲ از نکول انجام شد. دوره‌های زمانی بر اساس صلاحدید خبره سه‌ماهه در نظرگرفته شده است. مشتریانی که بازه زمانی قراردادشان با بانک پایان یافته و پس‌ازآن تعاملی با بانک نداشته‌اند، در وضعیت پایان‌یافته[1] قرار گرفته‌اند. مشتریانی که در مرحلهٔ تنفس (زمان بین اعطای وام به مشتری و اولین بازپرداخت مشتری است که توسط بانک بر طبق نوع قرارداد با مشتری تغییر می‌کند و می‌تواند تا سه سال نیز طول بکشد) قرار داشته‌اند، در وضعیت کم‌ریسک L1 قرار داده شده‌اند.

در این مرحله، وضعیت‌های مختلف مشتری را به‌ترتیبِ زمان وقوع آن‌ها به‌دنبال یکدیگر قرار می‌دهیم و دنبالهٔ ریسک اعتباری (دنبالهٔ انتقال) مشتری به‌دست می‌آید و ذخیره می‌شود؛ برای نمونه، دنبالهٔ ریسک اعتباری (دنبالهٔ انتقال) مشتری با شناسهٔ ۱۰۰۰ به این

[1] Terminated

شکل است: L1->L1->L1->L1->terminate و دنبالهٔ ریسک اعتباری مشتری با شناسهٔ ۱۰۰۱ به این شکل است: L1 -> L1-> L2->L2->M1->M2->L1->terminate که این‌گونه تفسیر می‌شود: مشتری در مرحلهٔ زمانی اول و دوم جزو مشتریان L1 (با ریسک کم) بوده، تا دو دورهٔ متوالی زمانی بعدازآن نیز جزو مشتریان L2 بوده، پس‌ازآن، به گروه مشتریان M1 پیوسته، سپس در گروه M2 قرار گرفته، در پایان به گروه مشتریان L1 بازگشته، و پس‌ازآن قراردادش با بانک پایان یافته است.

در این مرحله، دنباله‌های ریسک اعتباری را خوشه‌بندی می‌کنیم که موجب می‌شوند مشتریانی که در تغییر وضعیت اعتباری خود و احتمالاً پایدارشدن در وضعیتی مانند یکدیگر رفتارکرده‌اند، در خوشه‌های یکسان قرار داده می‌شوند و به‌تبعِ آن مشتریانی که رفتاری مشابه در طول زمان نداشته‌اند، در خوشه‌های متفاوت قرار می‌گیرند. در زمینهٔ خوشه‌بندی، روش‌های سلسله‌مراتبی شامل اگنس[1] و دیانا[2] توصیه شده است که با توجه به طبیعت کاراکتری دنباله‌های این تحقیق، از روش‌های خوشه‌بندی تفکیکی با الگوریتم CLARA، نرم‌افزار R data، و پکیج[3] و کتابخانه مربوط به خوشه‌بندی ، استفاده شده است.

پیش از اعتمادکردن به پکیج‌های موجود برای خوشه‌بندی، می‌بایست یک‌بار خوشه‌های به‌دست‌آمده چک می‌شد، زیرا خوشه‌بندی عمل بهینه‌سازی مثل رگرسیون یا شبکهٔ عصبی است و درصدی خطا دارد. با توجه به حجم بزرگ دنباله‌ها، برای خوشه‌بندی ابتدا نمونهٔ ۵ درصدی از میان آن‌ها به‌طور تصادف انتخاب و خوشه‌بندی شد. این خوشه‌بندی دستی چک شد و موردتأیید قرار گرفت.

در این گام برای توصیف و تفسیر خوشه‌های به‌دست‌آمدهٔ گام قبل، از قواعد انجمنی و دنباله‌ای استفاده شده است. برای پرهیز از بررسی تمام قواعد خوشه‌ها برای به‌دست‌آوردن قواعد برجسته، هر خوشه از روش k ـ برترین قواعد استفاده شده است. گفتنی است، برای این کار از الگوریتم GSP که بر مبنای اپریوری است و در بخش مبانی نظری تحقیق توضیح داده شد، استفاده شده است. حداقل شرط اطمینان با نظر خبره ۸۰ درصد انتخاب شده است. حداقل و حداکثر فاصلهٔ زمانی بین دو موقعیت متوالی برابر با یک دوره و دو دوره در نظر گرفته شده است. مقدار k بر اساس ۲۰ درصد تعداد قوانین با اطمینان بالای ۷۰ درصد و به صلاحدید خبرگان انتخاب شده است.

---

[1] Agness
[2] Diana
[3] Package

۴. ۵ یافته‌های تحقیق

۴. ۵. ۱ تشخیص الگوهای پویا

گام نهایی ارائهٔ نتایج به خبرگان و توصیف و تفسیر خوشه‌هاست. پس از ارائهٔ نتایج به خبرگان، در پایان هفت دسته از الگوهای پویایی ریسک اعتباری مشتریان شناسایی شد:

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ اول: مشتریان این خوشه در طول دورهٔ زمانی که مشتری بانک هستند، فقط به گروه L1 تعلق دارند که بخش مشتریان کم‌ریسک (بدون ریسک) بانک است و سابقهٔ جابه‌جایی به گروه‌های دیگر ریسک را ندارند، لذا پویایی ندارند. این خوشه با اطمینان ۹۱/۴۷ درصد و پشتیبان ۲۱/۳۲ تا پایان قرارداد خود در گروه L1 باقی می‌مانند. مشتریان این بخش به‌دلیل پرداخت بموقع و در سررسید مقرر تسهیلات اعطایی برای بانک دارای ارزش‌اند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ دوم: مشتریان این خوشه بین گروه L1 و L2 جابه‌جا شده‌اند و چنانچه نتایج نشان می‌دهد، با اطمینان ۷۳/۰۵ درصد و پشتیبان ۵۲/۸۱ درصد در گروه L2 باقی می‌مانند. مشتریان این خوشه با وجود پویایی بین دو گروه در زمرهٔ مشتریان کم‌ریسک بانک قرار دارند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ سوم: مشتریان این خوشه در سه دورهٔ متوالی در گروه L1 قرار دارند، پس‌ازآن به L2 تغییر گروه داده‌اند، و پس‌ازآن به گروه‌های M1، M2، و H1 مهاجرت کرده و دوباره به گروه M1 برگشته‌اند؛ لذا، پویایی زیادی در مشتریان این خوشه به‌چشم می‌خورد. مشتریان این گروه با اطمینان ۸۰/۱۲ درصد و پشتیبان ۶۱ درصد تا پایان قرارداد خود در گروه M2 باقی می‌مانند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ چهارم: مشتریان این خوشه به گروه L1 و L2 برای شش دورهٔ زمانی متوالی تعلق دارند، پس‌ازآن به گروه M1، و در نهایت به گروه M2 مهاجرت می‌کنند. مشتریان این خوشه با اطمینان ۹۱/۷۳ درصد و پشتیبان ۷۰/۳۵ در گروه M2 تا پایان قراردادشان باقی می‌مانند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ پنجم: مشتریان این خوشه به گروه L1 و L2 برای ۹ دورهٔ زمانی متوالی تعلق دارند و بین این دو گروه جابه‌جا می‌شوند، سپس برای پنج دوره زمانی بین گروه M1 و M2 جابه‌جا می‌شوند و در نهایت در گروه H1 قرار می‌گیرند. مشتریان این خوشه با اطمینان ۹۶/۰۷ درصد و پشتیبان ۶۶/۵۱ در گروه H1 تا پایان قراردادشان باقی می‌مانند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ ششم: مشتریان این خوشه به گروه L1 و L2 برای ۹ دورهٔ زمانی متوالی تعلق دارند و بین این دو گروه جابه‌جا می‌شوند، سپس برای

دو دورهٔ زمانی بین گروه M1 و M2 جابه‌جا می‌شوند، پس‌ازآن به گروه H1 مهاجرت می‌کنند و در نهایت در گروه H2 قرار می‌گیرند. مشتریان این خوشه با اطمینان ۹۲/۱ درصد و پشتیبان ۶۷ درصد در گروه H2 تا پایان قراردادشان باقی می‌مانند.

الگوی تغییر ریسک اعتباری مشتریان در خوشهٔ هفتم: مشتریان این خوشه به گروه L1 و L2 برای هشت دورهٔ زمانی متوالی تعلق دارند و بین این دو گروه جابه‌جا می‌شوند، سپس برای دو دورهٔ زمانی بین گروه M1 و M2 جابه‌جا می‌شوند، پس‌ازآن به گروه H1 مهاجرت می‌کنند، و در نهایت در دو دورهٔ زمانی متوالی در گروه H2 قرار می‌گیرند. مشتریان این خوشه با اطمینان ۹۴/۰۳ درصد و پشتیبان ۷۰/۸ درصد در گروه H3 تا پایان قراردادشان باقی می‌مانند.

۴. ۵. ۲ کارایی مدل

در خصوص کارایی مدل‌های ارزیابی ریسک اعتباری، کورشُلم[1] (۲۰۰۴) دو شاخص برای ارزیابی میزان کارایی مدل‌ها معرفی کرده است که در واقع مقایسهٔ پیش‌بینی داده‌های آزمون و مقدار داده‌های واقعی است. این دو شاخص درجهٔ حساسیت و درجهٔ تشخیص است که در ذیل به توضیح آن‌ها می‌پردازیم:

۱ـ درجهٔ حساسیت: نسبتی از مشتریانی است که واقعاً بد (ریسک متوسط و ریسک بالا) هستند و مدل هم آن‌ها را در گروه مشتریان بدحساب (ریسک متوسط و ریسک بالا) قرار می‌دهد.

۲ـ درجهٔ تشخیص: نسبتی از مشتریان واقعاً خوش‌حساب (کم‌ریسک) است که مدل رتبه‌بندی هم آن‌ها را در گروه خوش‌حساب (کم‌ریسک) قرار می‌دهد.

جدول ۲، کارایی مدل را که با استفاده از داده‌های آزمون محاسبه شده است، ارائه می کند. در ردیف درصد درست، درجهٔ حساسیت و تشخیص مدل به ترتیب ۸۹/۱ درصد و ۹۲/۱ درصد است.

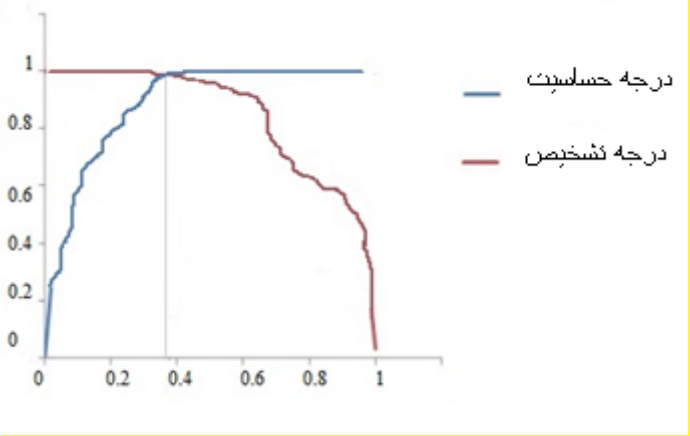
برای قضاوت در خصوص مشتریان و گروه‌بندی آن‌ها به مشتریان کم‌ریسک، ریسک متوسط، و ریسک بالا می‌بایست احتمال نکول در بازه [0,1] مشخص شود که احتمال نکول بیش از یک نقطه یعنی مشتری ریسک متوسط و ریسک بالا و کمتر از آن یعنی مشتری کم‌ریسک است.

حد آستانهٔ بهینهٔ مدل با توجه به معیارهای ارزیابی تعیین می‌شود. کورشُلم توابع هدف مختلفی برای به‌دست‌آوردن بهینه ارائه می‌کند. در این مقاله، حد آستانهٔ بهینه، برابر مقداری

---

[1] Korsholm

است که در آن مجموع درجهٔ حساسیت و درجهٔ تشخیص مدل ماکزیمم شود. در شکل ۵ دو نمودار درجهٔ حساسیت و درجهٔ تشخیص نمایش داده شده است. طبق شکل، حد آستانهٔ بهینه ۰/۳۷ است. در جدول ۲ کارایی مدل با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۵۶۰,۰۰۰ است. درجهٔ حساسیت و تشخیص مدل به‌ترتیب ۸۹/۱ درصد و ۹۲/۱ درصد است.



شکل ۵. حد آستانهٔ بهینه. این شکل حد آستانهٔ بهینه را که از تلاقی دو نمودار درجهٔ حساسیت و درجهٔ تشخیص به‌دست می‌آید، نشان می‌دهد.

جدول ۲

کارایی مدل

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	Y=1 ریسک متوسط و بالا	کل
P(y)<=C	۳۳۰,۸۱۵	۳۰,۰۷۴	۳۶۰,۸۸۹
P(y)> C	۳۹,۹۲۶	۱۵۹,۱۸۵	۱۹۹,۱۱۱
کل	۳۷۰,۷۴۱	۱۸۹,۲۵۹	۵۶۰,۰۰۰
درست	۳۴۱,۰۸۲	۱۶۸,۴۴۱	۵۰۹,۵۲۳
درصد درست	٪۹۲/۱	٪۸۹/۱	٪۹۰/۹۸
درصد نادرست	٪۷/۹	٪۱۰/۹	٪۹/۰۱

منبع: محاسبات تحقیق

۴. ۵. ۳ مقایسهٔ کارایی مدل با مدل‌های لاجیت، پروبیت، و شبکهٔ عصبی

برای مقایسهٔ کارایی مدل تحقیق، مدل‌های لاجیت، پروبیت، و شبکهٔ عصبی تهیه شد و معناداری مدل‌ها تایید شد. چون بررسی این مدل‌ها رسالت این تحقیق نیست، برای پرهیز از طولانی‌شدن مقاله، در این بخش از توضیح مدل‌سازی‌های فوق صرف‌نظر می‌کنیم و فقط به داده‌های مربوط به بررسی مقایسهٔ کارایی مدل که درجهٔ تشخیص و حساسیت است می‌پردازیم. در مورد تمام مدل‌ها مانند مدل این تحقیق، از داده‌های آزمون که ۵۶۰/۰۰۰ است، استفاده می‌کنیم. در ذیل به‌ترتیب مشاهدهٔ تعداد مشتریان هرگروه و نتیجهٔ محاسبات مربوط به کارایی مدل‌های لاجیت، پروبیت، و شبکهٔ عصبی به‌ترتیب در جدول ۳، جدول ۴، و جدول ۵ آمده است.

درجهٔ حساسیت و تشخیص مدل لاجیت به‌ترتیب ۶۱ درصد و ۶۵ درصد است. محاسبات در جدول ۳ آمده است.

جدول ۳

کارایی مدل لاجیت

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	=1Y ریسک متوسط و بالا	کل
P(y)<=C	۲۸۵,۳۹۵	۵۶,۷۰۸	۳۴۲,۱۰۳
P(y)> C	۷۸,۶۰۵	۱۳۹,۲۹۲	۲۱۷,۸۹۷
کل	۳۶۴,۰۰۰	۱۹۶,۰۰۰	۵۶۰,۰۰۰
درست	۲۳۶,۶۰۰	۱۱۹,۵۶۰	۳۵۶,۱۶۰
درصد درست	٪۶۵	٪۶۱	٪۶۳
درصد نادرست	٪۳۵	٪۳۹	٪۳۷

یادداشت. در این جدول، کارایی مدل لاجیت با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۵۶۰,۰۰۰ است. درجهٔ حساسیت و تشخیص مدل به‌ترتیب ۶۱ درصد و ۶۵ درصد است.

مقایسهٔ کارایی مدل با مدل پروبیت

درجهٔ حساسیت و تشخیص مدل پروبیت به‌ترتیب ۵۸ درصد و ۶۲ درصد است. محاسبات در جدول ۴ آمده است.

جدول ۴

کارایی مدل پروبیت

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	Y=1 ریسک متوسط و بالا	کل
P(y)<=C	۲۷۶,۰۲۰	۸۴,۹۸۱	۳۶۱,۰۰۱
P(y)> C	۱۰۳,۹۸۰	۹۵,۰۱۹	۱۹۸,۹۹۹
کل	۳۸۰,۰۰۰	۱۸۰,۰۰۰	۵۶۰,۰۰۰
درست	۲۳۵,۶۰۰	۱۰۴,۴۰۰	۳۴۰,۰۰۰
درصد درست	٪۶۲	٪۵۸	٪۶۰/۷
درصد نادرست	٪۳۸	٪۴۲	٪۳۹/۳

یادداشت. در این جدول، کارایی مدل پروبیت با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۵۶۰,۰۰۰ است. درجهٔ حساسیت و تشخیص مدل به‌ترتیب ۵۸ درصد و ۶۲ درصد است.

مقایسهٔ کارایی مدل با مدل شبکهٔ عصبی

درجهٔ حساسیت و تشخیص مدل شبکهٔ عصبی به‌ترتیب ۶۸/۹ درصد و ۷۰/۹ درصد است. محاسبات در جدول ۵ آمده است.

جدول ۵

کارایی مدل شبکهٔ عصبی

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	Y=1 ریسک متوسط و ریسک بالا	کل
P(y)<=C	۲۲۶,۱۸۰	۹۵,۰۵۵	۳۲۱,۲۳۵
P(y)> C	۲۳,۵۶۶	۲۱۹,۱۹۹	۲۳۸,۷۶۵
کل	۲۴۹,۷۴۶	۳۱۰,۲۵۴	۵۶۰,۰۰۰
درست	۱۷۷,۳۱۹	۲۱۴,۰۷۵	۳۹۱,۳۹۴
درصد درست	٪۷۰/۹	٪۶۸/۹	٪۶۹/۸
درصد نادرست	٪۲۹/۱	٪۳۱/۱	٪۳۰/۲

یادداشت. در این جدول، کارایی مدل شبکهٔ عصبی با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۵۶۰,۰۰۰ است. درجهٔ حساسیت و تشخیص مدل به‌ترتیب ۶۸/۹ درصد و ۷۰/۹ درصد است.

۴. ۵. ۴ بررسی حساسیت مدل پیشنهادی به مبلغ وام

برای بررسی حساسیت مدل پیشنهادی به مبلغ وام، داده‌های آزمون را به دو بخش[1] تقسیم می‌کنیم که عبارت‌اند از بخش مبلغ بالا و بخش مبلغ پایین[2].

تعداد داده‌های آزمون برای بخش مبلغ پایین ۴۵۹,۰۰۰ مشتری است. برای محاسبهٔ میزان حساسیت مدل به مبالغ پایین وام، جدول کارایی مدل را تشکیل می‌دهیم و محاسبات فوق را یک‌بار برای این نمونه ۴۵۹,۰۰۰تایی انجام می‌دهیم. چنانکه در جدول ۶ مشاهده می‌کنید، درجهٔ تشخیص و حساسیت مدل به وام‌های مبلغ پایین به‌ترتیب ۹۲ درصد و ۸۱ درصد است. نتایج محاسبات گویای این است که درصد مشتریان واقعاً کم‌ریسک در وام‌های با مبلغ پایین که مدل نیز آن‌ها را کم‌ریسک تشخیص داده، ۹۲ درصد است و درصد مشتریان واقعاً با ریسک بالا در وام‌های با مبلغ پایین که مدل نیز آن را با ریسک بالا شناسایی کرده است، ۸۱ درصد است.

جدول ۶

کارایی مدل پیشنهادی برای بخش مبلغ پایین

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	Y=1 ریسک متوسط و ریسک بالا	کل
P(y)<=C	۴۱۸,۱۹۰	۴۰۹۰	۴۲۲,۲۸۰
P(y)> C	۱۰,۰۱۱	۲۶,۷۰۹	۳۶,۷۲۰
کل	۴۲۸,۲۰۱	۳۰,۷۹۹	۴۵۹,۰۰۰
درست	۳۹۸,۱۲۴	۲۴,۹۴۸	۴۲۳,۰۷۲
درصد درست	٪۹۲	٪۸۱	٪۹۲
درصد نادرست	٪۸	٪۱۹	٪۸

یادداشت. در این جدول، کارایی مدل پیشنهادی برای وام‌ها با مبلغ پایین با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۴۵۹,۰۰۰ است. درجهٔ تشخیص و حساسیت مدل به‌ترتیب ۹۲ درصد و ۸۱ درصد است.

تعداد داده‌های آزمون برای بخش مبلغ بالا ۲۲۱,۰۰۰ مشتری است. برای محاسبهٔ میزان حساسیت مدل به مبالغ بالای وام، جدول کارایی مدل را تشکیل می‌دهیم و محاسبات فوق را یک‌بار برای این نمونه ۲۲۱,۰۰۰تایی انجام می‌دهیم. چنان‌که در جدول ۷ مشاهده

---

[1] Segment

[2] برای حفظ محرمانگی اطلاعات، از ذکر اعداد ردهٔ مبالغ بالا و پایین وام پرهیز می‌شود.

می‌کنید، درجهٔ تشخیص و حساسیت مدل به وام‌های مبلغ بالا به‌ترتیب ۹۰/۳ درصد و ۹۳ درصد است. نتایج محاسبات گویای این است که درصد مشتریان واقعاً کم‌ریسک در وام‌های با مبلغ بالا که مدل نیز آن‌ها را کم‌ریسک تشخیص داده، ۹۰/۳ درصد است و درصد مشتریان واقعاً با ریسک بالا در وام‌های با مبلغ بالا که مدل نیز آن را با ریسک بالا شناسایی کرده، ۹۳ درصد است.

جدول ۷

کارایی مدل پیشنهادی برای بخش مبلغ بالا

مشاهدات واقعی برآورد	Y=0 کم‌ریسک	Y=1 ریسک متوسط و ریسک بالا	کل
P(y)<=C	۶۴۱۲	۶۰۳۱۸	۶۶۷۳۰
P(y)> C	۴۸۳	۱۵۳۷۸۷	۱۵۴۲۷۰
کل	۶۸۹۵	۲۱۴۱۰۵	۲۲۱۰۰۰
درست	۶۲۲۷	۱۹۹۱۱۸	۲۰۵۳۴۵
درصد درست	٪۹۰/۳	٪۹۳	٪۹۲/۹
درصد نادرست	٪۹/۷	٪۷	٪۷/۱

یادداشت. در این جدول، کارایی مدل پیشنهادی برای وام‌ها با مبلغ بالا با استفاده از داده‌های آزمون که در آموزش مدل استفاده نشده بودند، محاسبه شده است. تعداد رکوردها ۲۲۱,۰۰۰ است. درجهٔ تشخیص و حساسیت مدل به‌ترتیب ۹۰/۳ درصد و ۹۳ درصد است

چنان‌که گفته شد، نتایج جداول ۶ و ۷ نشان می‌دهد که درجهٔ حساسیت مدل به وام‌های با مبلغ بالا ۹۳ درصد و به وام‌های با مبلغ پایین ۸۱ درصد است. همچنین درجهٔ تشخیص مدل برای وام‌های با مبلغ بالا ۹۰/۳ درصد و به وام‌های با مبلغ پایین ۹۲ درصد است. با توجه به تعاریف مربوط به درجهٔ حساسیت و تشخیص که در ابتدای این بخش آورده شد، اعداد فوق این‌گونه تحلیل می‌شود: حساسیت مدل به وام‌های مبلغ بالا بیشتر است، لذا مدل در شناسایی مشتریان واقعاً ریسک بالا در وام‌های با مبلغ بالا به میزان ۱۲ درصد موفق‌تر از وام‌های با مبلغ پایین است. اعداد به‌دست‌آمده برای درجهٔ تشخیص نشان می‌دهد مدل در شناسایی مشتریانی که واقعاً ریسک پایین دارند، در وام‌های با مبالغ بالا و پایین تقریباً یکسان و با تفاوتی اندک به میزان ۱/۷ درصد عمل می‌کند.

۵ تحلیل نتایج

در این تحقیق ریسک اعتباری مشتریان با استفاده از مدل پیشنهادی ارزیابی پویای ریسک اعتباری ارزیابی شد و محاسبات نشان می‌دهد، الگوهای تغییر وضعیت ریسک اعتباری مشتریان شامل هفت الگوست؛ به‌بیان‌دیگر، هفت خوشهٔ مختلف از مشتریان که حاکی از هفت روش رفتاری انتقالی بین وضعیت‌های مختلف ریسک اعتباری است، شناسایی شد. بر اساس k ـ برترین قواعد، خوشه‌های مختلف مشتریان به‌صورت ذیل توصیف می‌شوند:

خوشهٔ ۱: این خوشه شامل ۱۹۳/۶۴۷ مشتری است. قواعد این خوشه نشان می‌دهد که مشتری در طول بازه زمانی هشت‌ساله رفتار پایداری در بخش کم‌ریسک داشته است. این بخش از مشتریان برای بانک بی‌ضرر هستند و جزو مشتریان بسیار کم‌ریسک یا بسیار قوی بانک است که باید راهبرد وفاداری بیشتر برای آن‌ها طراحی شود.

خوشهٔ ۲: این خوشه شامل ۱۳۰,۲۸۷ مشتری است. قوانین به‌دست‌آمده نشان می‌دهد که این مشتریان اغلب در وضعیت کم‌ریسک بوده‌اند و در برخی بازه‌های زمانی، در وضعیت با ریسک متوسط قرار گرفته‌اند، ولی باز به وضعیت کم‌ریسک رفته و در این وضعیت پایدار شده‌اند. این خوشه، خوشهٔ مشتریان با ریسک کم نام‌گذاری می‌شود.

خوشهٔ ۳: این خوشه شامل ۷۹۴,۵۱۸ مشتری است. قوانین به‌دست‌آمده نشان می‌دهد که این مشتریان اغلب در وضعیت کم‌ریسک بوده‌اند و در برخی بازه‌های زمانی، در وضعیت با ریسک متوسط قرار گرفته، پس‌ازآن هرگز به بخش کم‌ریسک بازنگشته، در وضعیت ریسک بالا قرار گرفته، و دوباره با پرداخت مبلغی از بدهی به بخش ریسک متوسط بازگشته‌اند و تا زمان تسویه‌حساب قرارداد در آن بخش مانده‌اند. این خوشه، خوشهٔ مشتریان با ریسک متوسط نامیده می‌شود.

خوشهٔ ۴: این خوشه شامل ۳۲۳,۳۴۰ مشتری است که در بازه‌های اول تا چهارم در بخش L بوده‌اند و سپس به بخش M مهاجرت کرده‌اند. قوانین به‌دست‌آمده نشان می‌دهد اغلب این مشتریان از بازه ششم به بعد به M رفته‌اند و پس‌ازآن هرگز به L بازنگشته‌اند، ولی زمان قرارداد آن‌ها تمام شده و هرگز به H هم نرفته‌اند. این خوشه مشتریان با ریسک متوسط آسیب‌گرایانه نام‌گذاری می‌شود. چنانچه اگر زمان قرارداد آن‌ها به پایان نمی‌رسید، در وضعیت ریسک بالا قرار می‌گرفتند.

خوشهٔ ۵: تعداد مشتریان این خوشه ۳۷۳,۶۱۰ است. مشتریان این خوشه پس از ۹ دورهٔ زمانی از وضعیت L به M رفته و پس‌ازآن به وضعیت H رفته‌اند، ولی به‌دلیل اینکه دورهٔ قراردادشان به پایان رسیده است، پیش از یک سال از عدم بازپرداخت اقساط مجبور به تسویهٔ کامل شده‌اند و از طریق صدور حکم اجرائیه تسویه با بانک انجام شده است.

خوشهٔ ۶: این خوشه متشکل از ۶۲۷,۳۴۵ مشتری است. مشتریان این خوشه با نوسان بین وضعیت‌های L و M در نهایت به وضعیت اعتباری H1 رفته و بعد از دو دوره ماندن در وضعیت H1، به وضعیت H2 رفته‌اند، پس‌ازآن قرارداد اتمام یافته است. اغلب این مشتریان مشمول شرایط امهال بدهی شده‌اند؛ بااین‌حال، این دسته از مشتریان به بانک آسیب می‌زنند و برای دوره‌های بعدی به ایشان تخصیص اعتبار نخواهد شد.

خوشهٔ ۷: این خوشه شامل ۳۵۷,۲۵۳ مشتری است که با توجه به موضوع قرارداد، قراردادهای تمدیدشده دارند. این دسته از مشتریان در دورهٔ اول قرارداد در وضعیت‌های L یا M به‌سر می‌بردند، ولی در دورهٔ تمدید قرارداد، وضعیت اعتباری‌شان تغییر یافت، به‌طوری‌که از دورهٔ دوم به بعد از وضعیت اعتباری M به H رفتند و هرگز به وضعیت‌های ریسک متوسط بازنگشتند و تا بیش از ۱۸ ماه در این وضعیت اعتباری ماندند. از این گروه، تنها ۷۲۰۵ مشتری با استفاده از روش‌های تقسیط یا ارائهٔ اموال منقول و غیرمنقول قادر به بازگرداندن مطالبات بانک شدند. طلب بانک از سایر بدهکاران در دستهٔ مطالبات سوخت‌شده قرار گرفت.

با توجه به اینکه مشتریانی که تسهیلات به آن‌ها اعطا شده است، در هنگام افتتاح حساب و درخواست تسهیلات، از سوی بانک کم‌ریسک شناسایی شده‌اند، پس از به‌دست‌آمدن خوشه‌های فوق و با توجه به تعداد نمونهٔ اولیه که ۲,۸۴۰,۰۰۰ مشتری است، می‌توان نتیجه گرفت تنها مشتریان خوشه‌های ۱ و ۲ که در مجموع تعداد ۳۲۳,۹۳۴ نفر هستند، جابه‌جا نشده و در خوشهٔ خود پایدار بودند. مابقی مشتریان از خوشهٔ اولیه جابه‌جا شده‌اند که تعداد آن‌ها با توجه به نمونهٔ اولیه ۲,۵۱۶,۰۶۶ نفر است و معادل ۸۸/۵ درصد از مشتریان هستند. ذکر این عدد در گزارش به مدیران بانکی برای تصمیم‌گیری بسیار مهم است که ۸۸/۵ درصد از مشتریان در آغاز اشتباه شناسایی شده‌اند یا براثرِ وجود مشکلات غیرقابل‌پیش‌بینی رفتار غیرنرمالی از خود نشان داده‌اند. این گزارش اهمیت مدل پیشنهادی را بیش‌ازپیش آشکار می‌کند.

از نتایج به‌دست‌آمده می‌توان برای انتخاب و تدوین راهبرد، پرتفوی بانک، و توسعهٔ محصولات استفاده کرد. پس از به‌دست‌آمدن نتایج، تحلیل‌های استخراج‌شده به خبرگان و کارشناسان ریسک اعتباری ارائه و توصیه‌هایی اخذ شد که عبارت‌اند از:

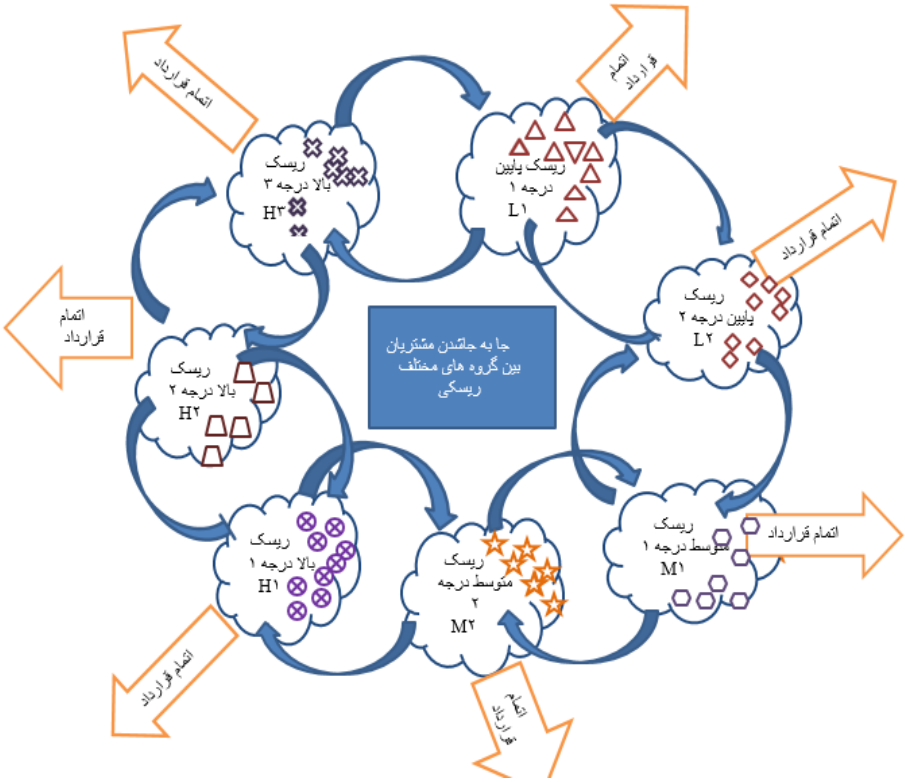
- مشتریان با ریسک بسیار کم و خوش‌حساب: این دسته از مشتریان محبوب بانک هستند و مشتریان هدف برای اعطای تسهیلات‌اند؛ لذا می‌بایست در تخصیص منابع در اولویت قرار بگیرند. برای ایشان برنامه‌های خاصی برای تشویق در تعریف محصولات در نظر

گرفته می‌شود، مثل تعیین مبلغ اقساط طی مذاکره با مشتری، تعیین زمان‌های پرداخت با بررسی ترازنامهٔ بانک و پرتفو به‌نحوی که بانک متضرر نشود.

- مشتریان با ریسک کم: این گروه از مشتریان نیز از مشتریان ارزشمند بانک هستند و در دستهٔ مشتریان هدف بانک قرار دارند. این مشتریان در تخصیص منابع در اولویت قرار می‌گیرند و بعد از مشتریان گروه با ریسک بسیارکم و بااخلاق قرار دارند. برای این دسته از مشتریان نیز برنامه‌های خاصی برای تشویق و جایزهٔ خوش‌حسابی در نظر گرفته می‌شود.
- مشتریان با ریسک متوسط کم‌ضرر: این مشتریان برای بانک اهمیت زیادی دارند، چرا که بخشی از مشتریان هدف بانک هستند و شایستگی اعطای تسهیلات را دارند؛ ولی با توجه به اینکه گاهی در پرداخت خود تعلل می‌ورزند، ضرر کمی را به بانک وارد می‌آورند. با بررسی ویژگی‌های این مشتریان، باید دید مشکل عدم بازپرداخت اقساطشان چیست و چگونه می‌توان آن را اصلاح کرد. به‌طور تجربی، برخی از کارشناسان ریسک که با مشتریان در ارتباط بوده‌اند، اذعان داشته‌اند که برخی از مشتریان نه به‌جهتِ مشکل مالی، بلکه به این دلیل که فراموش کرده‌اند پرداخت اقساط را به تعویق انداخته‌اند. برای این مشتریان می‌توان از سیستم یادآوری استفاده کرد. در مواردی که این راهکار هم کارآمد نباشد، می‌توان از سیستم دستور پرداخت استفاده کرد به این صورت که مشتری شماره‌حسابی را معرفی می‌کند و مطمئن است که در آن همیشه موجودی به‌اندازهٔ اقساط تسهیلات خود دارد و طی دستور برداشت، به بانک مجوز می‌دهد که در تاریخ معین ماهانه مبلغ قسط از حساب او کسر شود.
- مشتریان با ریسک متوسط: این دسته از مشتریان جزو مشتریان هستند که به‌طور عمد خود را در وضعیت مخاطره‌آمیز نسبت‌به بانک قرار می‌دهند و پیش از اینکه برای آن‌ها از سوی بانک نامه‌ای ارسال شود یا به وضعیت اجرائیه بروند، اقساط خود را پرداخت می‌کنند. راهکاری که می‌توان از آن برای این دسته از مشتریان استفاده کرد، اخذ جریمه‌ای است که بیش از مبلغ سود حاصل از دیرپرداخت کردن اقساطشان باشد.
- مشتریان با ریسک بالا و آسیب‌گرایانه: این دسته از مشتریان جزو مشتریانی هستند که ریسک اعتباری قابل‌توجهی دارند. شرایط بد کسب‌وکار، مالی، و اقتصادی حتماً به توانایی بازپرداخت اقساطشان صدمه می‌زند و آن‌ها را در آستانهٔ عدم بازپرداخت یا نکول قرار می‌دهد. تنها حاشیهٔ ایمنی که می‌تواند برای این دسته از مشتریان وجود داشته باشد، ارزش وثیقه منقول یا غیرمنقول آن‌هاست که از فروش آن‌ها بانک می‌تواند مطالبات خود را به‌دست آورد.

- مشتریان با ریسک بالا: این دسته از مشتریان احتمال نکول بسیار بالایی دارند و فاقد درصدی از حاشیهٔ ایمنی هستند. هرگونه شرایط محیطی بحران‌زا آن‌ها را در آستانهٔ نکول و حتی ازدست‌دادن کسب‌وکار و سرمایه‌شان قرار می‌دهد. در این‌صورت، احتمال بازپرداخت مطالبات بانک و تسویه از طریق فروش وثایق نیز ضعیف است. در این‌گونه موارد، اوراق قرضهٔ اخذشده احتمال بازگشت متوسطی دارد.
- مشتریان با ریسک بسیار بالا و آسیب‌رسان: با توجه به محاسبات انجام‌شده و تعداد وام‌های بازپرداخت‌شده توسط این گروه و اینکه به‌جز در موارد معدود مابقی جزو مطالبات سوخت‌شدهٔ بانک قرار می‌گیرد، در خصوص این دسته از مشتریان توصیه مؤکد است که تخصیص منابع بانک صورت نگیرد.

در این تحقیق سعی شد با دیدی خاکستری به مشتری نگاه شود، نه کاملاً خوش‌حساب و نه بدحساب؛ به‌عبارت‌دیگر، رفتار مشتری معلول فاکتورهایی است که در طول زمان تغییر می‌کند و در واقع رابطهٔ مشتری با بانک محاط در شرایط خانوادگی، اقتصادی، اجتماعی، و سیاسی اطراف اوست؛ لذا همان‌طور که شکل ذیل نشان می‌دهد، مشتری در طول زمان بین گروه‌های مختلف ریسک اعتباری جابه‌جا می‌شود؛ اما تحلیل نتایج این تحقیق نشان داد که به‌طورکلی رفتارهای مشتریان از هفت مدل اصلی تبعیت می‌کند که در واقع بانک با اجرای مدل این تحقیق و پیش‌بینی خوشه‌ای که مشتری به آن تعلق دارد، می‌تواند رفتار مشتری (ریسک اعتباری مشتری) را در آینده پیش‌بینی کند. شکل‌های ۶ و ۷ به‌خوبی مفاهیم فوق را تبیین می‌کند.



شکل ۶. رفتارشناسی مشتریان. این شکل جابه‌جاشدن مشتری بین گروه‌های مختلف ریسک اعتباری را نشان می‌دهد. در این شکل، گروه‌های مختلف ریسک اعتباری از خیلی کم‌ریسک با نماد L1 تا بالاترین درجهٔ ریسک با نماد H3 نشان داده شده است و با فلش‌هایی مشخص‌شده که مشتری از یک گروه در برههٔ زمانی دیگر که در این تحقیق سه ماه است، به گروه دیگری جابه‌جا می‌شود. همچنین مشتریان در هریک از گروه‌ها که باشند، می‌توانند براثرِ تسویه‌حساب یا خاتمهٔ قرارداد از گروه خارج شود که با فلش اتمام قرارداد نشان داده شده است.
منبع:یافته‌های تحقیق

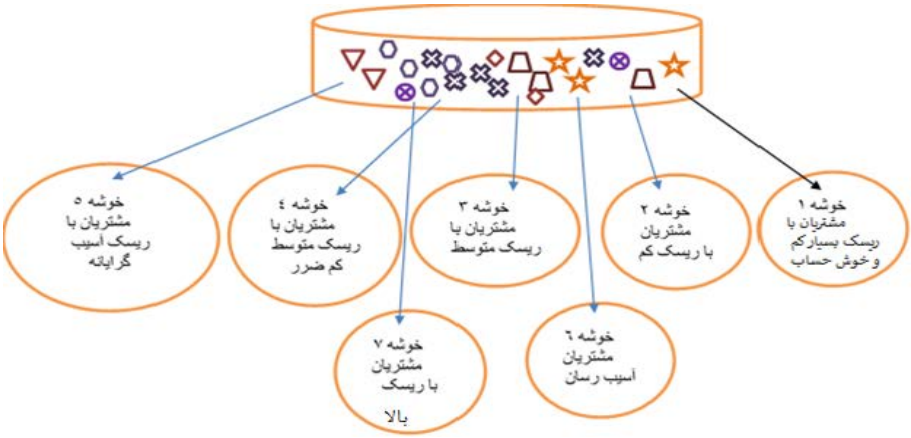
جدول ۸، گویای مطالب موجود در شکل ۶ است به این صورت که وضعیت آغازین ریسک اعتباری مشتری، وضعیت‌های مشتری در طول زمانی که مشتری بانک بوده است و وضعیت پایانی مشتری را نشان می‌دهد.

جدول ۸

وضعیت ابتدایی و انتهایی ریسک مشتریان در الگوهای اصلی تغییر ریسک

شمارهٔ الگو	وضعیت آغازین	وضعیت میانی	وضعیت نهایی
۱	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک‌ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک	خیلی کم‌ریسک
۲	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک – کم‌ریسک – کم‌ریسک ـ خیلی کم‌ریسک – کم‌ریسک	کم‌ریسک
۳	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ کم‌ریسک – ریسک متوسط ۱ـ ریسم متوسط ۲ـ ریسک بالا ۱ـ ریسک بالا ۱ـ ریسک متوسط ۲ـ ریسک متوسط ۱	ریسک متوسط – درجهٔ ۲
۴	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ کم‌ریسک – کم‌ریسک ـ ریسک متوسط ۱ـ ریسک متوسط ۱ ـ ریسک متوسط ۲	ریسک متوسط – درجهٔ ۲
۵	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ کم‌ریسک ـ کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ کم‌ریسک ـ ریسک متوسط ۱ـ ریسک متوسط ۱ ـ ریسک متوسط ۲ ـ ریسک متوسط ۱ ـ ریسک متوسط ۲– ریسک بالا ۱	ریسک بالا از درجهٔ ۱
۶	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک – خیلی کم‌ریسک – کم‌ریسک – ریسک متوسط ۱ـ ریسک متوسط ۲ـ ریسک بالا ۱ـ ریسک بالا ۱ـ ریسک بالا ۲	ریسک بالا از درجهٔ ۲
۷	خیلی کم‌ریسک	خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک ـ خیلی کم‌ریسک – کم‌ریسک– ریسک متوسط ۱ـ ریسک متوسط ۲ـ ریسک بالا ۱ـ ریسک بالا ۱ـ ریسک بالا ۱ـ ریسک بالا ۲ـ ریسک بالا ۲	ریسک بالا از درجهٔ ۳

منبع: یافته‌های تحقیق



شکل ۷. خوشه‌های شناسایی‌شدهٔ مشتریان در رفتار پویا. مشتریان در رفتار ریسک اعتباری ایستا عمل نمی‌کنند و در مقاطع زمانی مختلف رفتار مختلفی از ایشان بروز می‌کند؛ ولی در عین پویا رفتارکردن، در این پویایی هفت روش مختلف عمل می‌کند. هفت خوشه که از نتیجهٔ تحقیق به‌دست آمده و در این شکل نشان داده شده است، حاصل هفت روش مختلف رفتاری است.
منبع: یافته‌های تحقیق

۶ نتیجه‌گیری و بحث

تجربه صنعت بانکی نشان می‌دهد ارزیابی ریسک اعتباری با دیدگاه ایستا نسبت به عملکرد و رفتار مشتریان خطای زیادی دارد که منجر به بروز ضررهای جبران ناپذیری برای بانک‌ها می‌شود. بانک ها برای مدیریت ریسک اعتباری ناچار به تجهیز ارزیابی مشتریان به دیدی نو می‌باشند. دیدگاهی که در آن مشتری و رویکرد و رفتار او با گذشت زمان در حال تغییر است و عملکرد او در شرایط پیرامون او محاط است. در این تحقیق یک روش پویا براي ارزیابی ریسک اعتباری مشتریان با استفاده از دنباله‌های به‌دست‌آمده از تغییر وضعیت ریسک اعتباري مشتري در گذر زمان ارائه شد، که نوآوری این مدل است و در هیچ یک از مقالات مربوط به ارزیابی ریسک اعتباری از چنین روشی استفاده نشده است.

برای بررسی کارایی روش‌های ارزیابی ریسک اعتباری دو شاخص وجود دارد که عبارتند از درجهٔ حساسیت و درجهٔ تشخیص. در بخش یافته‌های تحقیق، محاسبات انجام‌شده نشان داد که مدل با درجهٔ حساسیت ۸۹/۱% و درجهٔ تشخیص ۹۲/۱% از کارایی بالایی برخوردار است. که این نشان می‌دهد فرضیهٔ این تحقیق رد نشد و مدل پیشنهادی توانایی پیش‌بینی ریسک اعتباری را دارد. برای بررسی پرسش این تحقیق در پاسخ به اینکه آیا این مدل از

مدل‌های پیشین نتایج بهتری دارد، به محاسبهٔ کارایی مدل‌های لاجیت، پروبیت، و شبکهٔ عصبی پرداختیم و نتایج را با کارایی مدل این تحقیق بررسی کردیم. نتایج نشان می‌دهد مدل این تحقیق از درجهٔ حساسیت و تشخیص بالاتر از تمام مدل‌های ذکرشده برخوردار است و به‌بیان‌دیگر، کارایی بالاتری از مدل‌های فوق دارد. پس‌ازاین مدل، کارایی مدل‌های شبکهٔ عصبی، مدل لاجیت، و پروبیت به‌ترتیب در رتبهٔ دوم، سوم، و چهارم قرار دارند.

تفاوت ذکرشده بین مدل‌های با رویکرد ایستا (به‌عنوان مثال، در مدل‌های شبکهٔ عصبی، لاجیت، و پروبیت) و پویا در مطالبات معوق مشتریان و آنچه در واقعیت رخ می‌دهد، در بخش پیوست نمودار۱ نشان داده شده است.

در روش پیشنهادی، پس از ایجاد دنباله‌های ریسک اعتباری مشتریان، خوشه‌بندی دنباله‌های ریسک اعتباری مشتریان انجام شد و پس‌ازآن الگوهای تغییر وضعیت مشتریان بین بخش‌های مختلف ریسک اعتباری استخراج شد. نتایج نشان داد که همهٔ مشتریان از هفت الگوی رفتاری در تغییر رفتار اعتباری خود تبعیت می‌کنند. هریک از خوشه‌ها مجموعه‌ای از مشتریان را نشان می‌دهد که تغییر وضعیت اعتباری یکسانی دارند. پس از تحلیل و توصیف رفتار هریک از خوشه‌ها توسط گروه خبرهٔ بانکی برای هریک از مجموعه‌ها نام‌گذاری انجام شد. نام‌ها عبارت‌اند از مشتریان با ریسک بسیارکم و خوش‌حساب، مشتریان با ریسک کم، مشتریان با ریسک متوسط و کم‌ضرر، مشتریان با ریسک متوسط، مشتریان با ریسک بالای آسیب‌گرایانه، مشتریان با ریسک بالا، و مشتریان با ریسک بالای آسیب‌رسان. هریک از این مجموعه‌ها نوعی الگوی غالب وضعیت اعتباری را نشان می‌دهند. برای هریک از این خوشه‌ها، قواعد دنباله‌ای مربوط به آن‌ها به‌دست آورده شد که الگوهای رفتار اعتباری آن‌ها را نشان می‌دهد.

بانک‌ها می‌توانند ویژگی‌های مربوط به مشتریان خود را به این مدل به‌عنوان ورودی تحویل دهند و خوشهٔ مشتری را شناسایی و ریسک اعتباری او را پیش‌بینی کنند و در خصوص اعطای تسهیلات یا اعتبار به مشتری تصمیم‌گیری کنند.

۷ محدودیت‌های تحقیق و پیشنهادهای آتی

متغیرهای دیگری نیز برای ارزیابی ریسک اعتباری مشتری کاربرد دارد که در پایگاه‌دادهٔ بانک‌های موردمطالعه موجود نبود. چنانچه پایگاه اطلاعاتی بر اساس استانداردهای جهانی غنی‌تر شود، امکان بهینه‌سازی مدل وجود دارد. همچنین توصیف و تحلیل رفتار مشتریان با استفاده از دانش خبره انجام شد؛ چنانچه بتوان این بخش را نیز به هوش مصنوعی محول کرد، مدل به‌طور کامل از مداخله و اتکای به کارشناسی انسان بی‌نیاز می‌شود.

باتوجه به نتایج تحقیق، پیشنهاد می‌شود مدیریت ریسک اعتباری نیز به‌صورت پویا انجام پذیرد. توجه به روابط مشتریان ذی‌نفع واحد و ارائهٔ مدلی پویا به تفکیک اعتبار خرد و کلان زمینهٔ تحقیقاتی دیگر است که می‌تواند موردتوجه قرار گیرد.

فهرست منابع

نظرآقایی م.؛ غیاثی ح.؛ و خواه چافی، م. (۱۳۹۸). دسته‌بندی ریسک اعتباری مشتریان حقیقی با استفاده از یادگیری جمعی (مطالعهٔ موردی بانک سپه). *فصلنامهٔ پژوهش‌های پولی-بانکی*، ۱۲(۳۹)، ۱۶۶-۱۲۹

قاسمی ا. و دنیایی هریس، ط. (۱۳۹۵). اندازه‌گیری ریسک اعتباری مشتریان با رویکرد شبکهٔ عصبی در یکی از بانک‌های دولتی. *مجلهٔ مهندسی مالی و مدیریت اوراق بهادار*، ۲۷.

ابراهیمی، م. و دریابر، ع. (۱۳۹۱). مدیریت ریسک اعتباری در نظام بانکی- رویکرد تحلیل پوششی داده‌ها و رگرسیون لجستیک و شبکهٔ عصبی. *فصلنامه دانش سرمایه‌گذاری*، ۲.

تاری، ف.؛ ابراهیمی، ا.؛ موسوی، ج.؛ و کلانتری، م. (۱۳۹۵). مقایسهٔ مدل‌های شبکهٔ عصبی، الگوریتم ژنتیک و لاجیت در ارزیابی ریسک اعتباری مشتریان. *فصلنامهٔ پژوهش‌های پولی-بانکی*، ۳۴، ۶۸۰-۶۵۷.

مرادی، ص. (۱۳۹۷). مدل‌سازی پویای ریسک اعتباری مشتریان. *دانشگاه علوم و تحقیقات*، تهران

Abdou, H. A., & Pointon, J. (2009). Credit scoring and decision making in Egyptian public sector banks. *International Journal of Managerial Finance*, 5(4), 391-406. Doi: 10.1108/17439130910987549

Agrawal R, & Srikant R. (1993). Mining sequential patterns. Proceeding of 11th Int.Conf. of Data Engineering. *IEEE computer society*, pp. 3-14.

Bekhet H., & Eletter S. (2014). Credit risk assessment model for Jordanian commercial banks: Neural scoring approach. *Review of Development Finance*, 4(1), 20-28. Doi: 10.1016/j.rdf.2014.03.002

Bensic M., Sarlija N., & Zekic-Susac M. (2005). Modeling small-business credit scoring by using logistic regression, neural networks and decision trees. *Intelligent System In Account. Fin. Manage*: 133-150. doi:10.1002/isaf.261.

Chan, C. C. H. (2008). Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer. *Expert Systems with Applications*, 34, 2754-2762. doi:10.1016/j.eswa.2007.05.043.

Chen W, Xiang G, Liu Y, & Wang K. (2012). Credit risk evaluation by hybrid data mining technique. *Systems Engineering Procedia*, 3, 194-200. Doi: 10.1016/j.sepro.2011.10.029

Danenas P, Garsva G. (2015). Selection of support vector machines based classifiers for credit risk domain. *Expert Systems with application*, 42, 3194-3204. Doi: 10.1016/j.eswa.2014.12.001.

Dong, G. & Pei, J. (2007). *Sequence Data Mining*. Heidelberg: Springer.

Fournier-Viger P, Faghihi, U., Nkambou, R., Nguifo, E. M. (2012). CM Rules: Mining sequential rules common to several sequences. *Knowledge-Based Systems*, 25(1), 63-76.

Fournier-Viger, P. & Tseng, V.S. (2011). Mining top-k sequential rules. Advanced Data Mining and Applications, *Proceedings of 7th International Conference,* ADMA 2011, Beijing, China, December 17-19, 2011,Springer, Berlin and Heidelberg, pp. 180-194. Doi: 10.1007/978-3-642-25856-5_14.

Girjia N, & Sirvatsa S.K. (2006). A research study: Using data mining in knowledge base business strategies. *Information Technology Journal*, 5(3), 590-600. Doi: 10.3923/itj.2006.590.600.

Heiat A. (2012). Comparing performance of data mining models for computer credit scoring. *International Journal of Finance and Economics*, 12(1), 78-83.

Hsieh N.C, & Hung L.P. (2010). A data driven ensemble classifier for credit scoring analysis. *Expert Systems with Applications*, 37, 534-545. doi:10.1016/j.eswa.2009.05.059

Jarrow, R. & Turnbull, S. (1995). Pricing derivatives on financial securities subject to default risk. *Journal of finance*, 50(1), 53-86. Doi: 10.1111/j.1540-6261.1995.tb05167.x

Jewell, J. & Livingston, M. (1998). Split rating, bond yields and underwriter spreads for industrial bonds. *Journal of Financial Research*, 1(2), 185-204. doi:10.1111/j.1475-6803.1998.tb00679.x

Kaufman L, & Rousseeuw P.J. (2009). *Finding groups in data: An introduction to cluster analysis*. New Jersey: John Wiley & Sons.

Korsholm L. (2004). Analysis of diagnostic studies, sensitivity and specificity positive predicted values ROC curves tests based on logistic regression.

Department of statistics and demography, University of Southern Denmark.

Kruppa J, Schwarz, Armonger G, & Ziegler A. (2013). Customer credit risk: Individual probability estimates using machine learning. *Expert Systems with Applications*, 40(13), 5125-5131. Doi: 10.1016/j.eswa.2013.03.019.

Moradi S, & Mokhat Rafieie F. (2019). A dynamic credit risk assessment model with data mining techniques: Evidence from Iranian banks. *Financial Innovation*, 5, 15, doi: 10.1186/s40854-019-0121-9

Narindra Mandalaa G, & Nawangpalupia C. Fransiscus. (2012). Assessing credit risk: An application of data mining in a rural bank. *Procedia Economics and Finance,* 4, 406-412. Doi: 10.1016/S2212-5671(12)00355-3

Ping Y. & Yongheng, L. (2011). Neighborhood rough set and SVM based hybrid credit scoring classifier. *Expert Systems with Applications*, 38(9), 11300-11304. doi:10.1016/j.eswa.2011.02.179

Rosenblatt F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev*, 65, 386-408.

Shahari, F., Zakaria, R., & Rahman, S. (2015). Investigation of the expected loss of sharia credit instruments in global Islamic banks. *International Journal of Managerial Finance*, 11(4), 503-512. Doi: 10.1108/IJMF-12-2014-0196

Sousa M, Gama J, Brandao. (2016). A new dynamic modeling framework for credit risk assessment. *Expert Systems with Applications*, 45, 341-351. Doi: 10.1016/j.eswa.2015.09.055

Srikant R, & Agrawal R. (1996). Mining sequential patterns: Generalizations and performance improvements, *EDBT 96*, pp. 1-17. Doi: 10.1007/BFb0014140

Tamayo, M. (2000). Managerial applications of neural networks: The case of bank failure predictions. *Management science*, 38(7), 926-947. Doi: 10.100711060-017-2657-8

Twala B. (2010). Multiple classifier application to credit risk assessment. *Expert Systems with Applications*, 37(4), 3326-3336. Doi: 10.1016/j.eswa.2009.10.018

Van Gestel T., Baesense, B., Van Dijcke, P., Suykens, J., & Garsia, J. (2005). Linear and nonlinear credit scoring by combining logistic regression and Support Vector Machines. *Journal of Credit risk*, 1, (4), 31-60. Doi: 10.21314/JCR.2005.025

Wang G., Hao, J., Ma, J., & Jiang, H. (2011). A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, 38, 223-230. Doi: 10.1016/j.eswa.2010.06.048

Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature sekection for corporate bankruptcy prediction. *Expert Systems with Applications*, 41, 2353- 2361. Doi: 10.1016/j.eswa.2013.09.033

West, D. (2000). Neural network credit scoring models. *Computers and Operations Research*, 27, 1131-1152. Doi: 10.1016/S0305-0548(99)00149-5

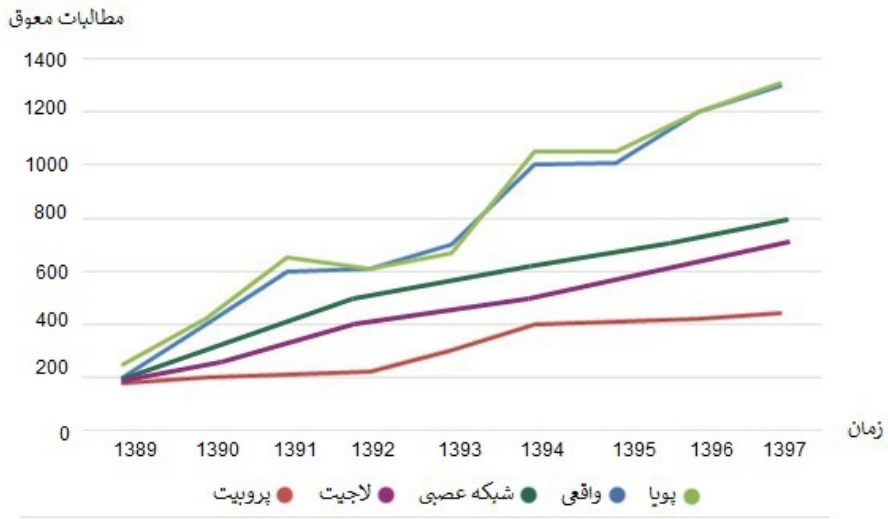
Yeh I.C, & Lien Ch. (2009). The comparisons of datamining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2), 2473-2480. doi:10.1016/j.eswa.2007.12.020

Zamani, S. (2011). Evaluating of predicting power of ANN in order to predict customer's credit risk. *Thesis*.Master of science, Accounting, Sistan University, Zahedan

Zanin, M. et al. (2016). Combining complex networks and data mining: Why and how. *Physics reports*, 635, 1-44. Doi: 10.1101/054064.

Zhang, Z., Gao, b. G., & Shi, Y. (2014). Credit risk evaluation using multi-criteria optimization classifier with kernel, fuzzification and penalty factors. *European Journal of Operational Research*, 237(1),335-348. doi:10.1016/j.ejor.2014.01.044.

پیوست



شکل ۱. نمودار مطالبات معوق بر اساس واقعیت، مدل‌های لاجیت، پروبیت، شبکهٔ عصبی، و مدل پویای پیشنهادی این تحقیق

یادداشت. در این شکل، محور عمودی مقدار مطالبات معوق است و فاصله زمانی بین سال‌های ۱۳۸۹ تا ۱۳۹۷ محور افقی است و نمودار به ترتیب از پایین به بالا (که با رنگ متفاوت نشان داده شده است و در راهنمای نمودار توضیح داده شده است) محاسبهٔ میزان پیش‌بینی مطالبات معوق بر اساس مدل پروبیت، مدل لاجیت، مدل شبکهٔ عصبی است. نمودار چهارم که با رنگ آبی نمایش داده شده است نشان‌دهندهٔ میزان واقعی مطالبات معوق و نمودار آخر که با رنگ سبز نشان داده شده است نمودار مطالبات معوق بر اساس مدل پیشنهادی این تحقیق است. قرابت مقدار پیشنهادی مدل با واقعیت به‌وضوح قابل‌مشاهده است که این نشان کارایی و قابلیت پیش‌بینی بالای مدل است.