

ارزیابی مقایسه‌ای عملکرد توابع شایستگی الگوریتم ژنتیک در رتبه‌بندی مشتریان

علی اقبالی^۱، سید حسین رضوی حاجی آقا^۲، حنان عموزاد مهدیرجی^۳

چکیده: با توجه به اهمیت بحث اعتبارسنجی مشتریان در تصمیم‌گیری مؤسسه‌های اعتباری برای اعطای تسهیلات، تحقیقات گسترده‌ای در خصوص روش‌های رتبه‌بندی اعتباری انجام شده است. الگوریتم‌های ژنتیک به‌عنوان یکی از روش‌های محاسبه تکاملی، از جمله روش‌هایی است که در این زمینه استفاده می‌شود. مقالات بسیاری در مقایسه عملکرد الگوریتم‌های ژنتیک و سایر روش‌های رتبه‌بندی رایج منتشر شده است، اما جزئیات چندانی درباره تابع شایستگی مورد استفاده در الگوریتم ژنتیک ارائه نشده، در حالیکه تابع شایستگی تأثیر شایان توجهی بر عملکرد مدل کلی دارد. برای بررسی بیشتر موضوع، در مقاله حاضر سه تابع شایستگی قابل استفاده در الگوریتم ژنتیک مطرح شده و عملکرد آنها با سایر روش‌های رتبه‌بندی اعتباری شامل رگرسیون لجستیک و تحلیل پوششی داده‌ها مقایسه شده است. نتایج به‌دست آمده حاکی از برتری عملکرد تابع درجه دو الگوریتم ژنتیک از لحاظ مجموع شاخص‌های صحت، تشخیص و حساسیت است.

واژه‌های کلیدی: الگوریتم ژنتیک، تابع شایستگی، تحلیل پوششی داده‌ها، رتبه‌بندی اعتباری، رگرسیون لجستیک، روش‌های ارزشیابی، مدیریت ریسک.

۱. کارشناس ارشد مدیریت صنعتی، دانشکده مدیریت و مالی، دانشگاه خاتم، تهران، ایران

۲. استادیار گروه مدیریت، دانشکده مدیریت و مالی، دانشگاه خاتم، تهران، ایران

۳. استادیار گروه مدیریت صنعتی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۵/۱۱/۱۶

تاریخ پذیرش نهایی مقاله: ۱۳۹۶/۰۲/۳۰

نویسنده مسئول مقاله: حنان عموزاد مهدیرجی

E-mail: h.amoozad@ut.ac.ir

مقدمه

امروزه با توجه به بحران‌های مالی و اقتصادی در سطح جهان و نقش نهادهای مالی در تأمین منابع سرمایه‌ای، تخصیص بهینه منابع و ارتقای کیفیت خدمات اعتبار و تسهیلات توسط بانک‌ها و مؤسسه‌های مالی و اعتباری، ضرورت محسوب می‌شود. گسترش بازارهای مالی، پیچیده‌شدن فرایند تأمین مالی و اخذ تسهیلات در سال‌های گذشته، اهمیت تأیید صلاحیت و اهلیت مشتریان بانک را دوچندان کرده است. در همین رابطه، قانون‌گذاران در برنامه پنج‌ساله بانک‌ها را مکلف کرده‌اند تا پیش از اعطای تسهیلات و اعتبار به مشتریان، نسبت به اعتبارسنجی آنان اقدام کنند. این الزام قانونی از خلأ اطلاعات جامع درباره مشتریان بانک‌ها نشئت می‌گیرد و این خلأ، زمینه‌ساز چالش‌های جدی برای بانک‌های کشور خواهد بود. در حالی که با اعتبارسنجی، جمع‌آوری، ساماندهی، تلفیق، تطبیق و پردازش اطلاعات اعتباری مشتریان، بانک‌ها و مؤسسه‌های مالی و اعتباری می‌توانند با شناخت مناسب‌تر از مشتری، ریسک اعطای تسهیلات را تا حد امکان کاهش دهند و به‌علاوه، نیاز واقعی مشتریان را برآورده و تأمین کنند. ارتباط صحیح بین نظام‌های مالی و تولیدی در هر کشوری از مهم‌ترین عوامل رشد و توسعه اقتصادی به‌شمار می‌رود. در این رابطه، بانک‌ها به‌عنوان بخش‌های اصلی نظام مالی (نظام پایه بانکی)، نقش اصلی را در تأمین مالی بخش‌های مختلف اقتصادی بر عهده دارند.

در ایران نیز به‌دلیل عدم گسترش متناسب بازار سرمایه، عمدتاً بانک‌ها و شبکه بانکی به‌عنوان مهم‌ترین نهاد مالی، وظیفه جمع‌آوری و تخصیص منابع مالی اقتصاد را بر عهده دارند. از این رو، هرگونه تصمیم‌گیری در رابطه با کیفیت جمع‌آوری و تزریق این منابع، آثار بنیادینی بر بخش‌های مختلف اقتصادی می‌گذارد.

توسعه و گسترش فعالیت‌های بانکی از طریق سیستم کارآمد، در پیشرفت و توسعه اقتصاد کشور بسیار مؤثر است و موجب پایداری بانک در محیط رقابتی می‌شود. به بیان دیگر، موفقیت بانک در عملیات بانکی نه تنها بقای بانک را در بلندمدت تضمین می‌کند، بلکه بالندگی، رشد و توسعه اقتصادی را برای کشور به ارمغان می‌آورد. از فعالیت‌های عمده و اساسی که موفقیت بانک را تضمین می‌کند، چگونگی تخصیص منابع است. بدیهی است که بخش عمده منابع بانکی کشور از پس‌اندازهای مردم تشکیل شده و هدایت صحیح این منابع به سمت فعالیت‌های اقتصادی از طریق سیستم کارآمد تخصیص اعتبار، در صنعت بانکداری کشور ضروری است (رجب‌زاده، بهرام‌میرزایی و احمدی، ۱۳۸۸).

در بسیاری از کشورهای جهان، استراتژی‌ها و راهکارهای گوناگونی در این زمینه استفاده شده است (کائو و کالبرگ، ۱۹۹۴). برای ایفای این نقش، بانک‌ها با ریسک‌های متفاوتی روبه‌رو

می‌شوند که یکی از عمده‌ترین آنها ریسک اعتباری است. ریسک اعتباری را به‌طور خلاصه می‌توان «احتمال قصور وام‌گیرنده یا طرف مقابل بانک نسبت به انجام تعهداتش طبق شرایط توافق شده» تعریف کرد (گروه مدیریت ریسک کمیته بازل، ۲۰۰۰).

برای اعطای تسهیلات، باید درجه اعتبار و قدرت بازپرداخت اصل و سود تسهیلات دریافت‌کننده را تعیین کرد تا احتمال عدم برگشت اصل و سود اعطایی، یعنی ریسک درجه اعتبار کاهش یابد. یکی از روش‌های کاهش این ریسک، طراحی نظام سنجش درجه اعتباری برای دریافت‌کنندگان تسهیلات است که مدل رتبه‌بندی یا ارزیابی اعتباری (اعتبارسنجی)، کانون این نظام محسوب می‌شود (شایان آرانی، ۱۳۸۰).

ارائه تسهیلات و اعتبارات، یکی از هسته‌های اصلی کسب‌وکار بانکداری جهانی به‌شمار می‌رود. از جمله تصمیمات اولیه بانک‌ها در این کسب‌وکار، افزایش ارائه وام به مشتریان است. یکی از موارد مهم در مؤسسه‌ها و بانک‌ها، شناخت مشتریان خوب و مشتریان بد از لحاظ پرداخت اقساط است تا بتوان تفاوت‌هایی را میان آنها در نظر گرفت. از روش‌های بسیاری برای رتبه‌بندی اعتبار استفاده شده که از آن جمله می‌توان به رگرسیون خطی تفکیک‌پذیر، درخت تصمیم، مدل لوجیت و پروبیت، روش‌های تصمیم‌گیری چند شاخصه نظیر مجموع ساده وزنی و فرایند تحلیل سلسله‌مراتبی، شبکه‌های عصبی مصنوعی، داده‌کاوی، الگوریتم ژنتیک و... اشاره کرد که در بخش پیشینه پژوهش به شرح کوتاهی از آنها پرداخته شده است.

الگوریتم‌های ژنتیک به‌عنوان یکی از روش‌های بهینه‌سازی، از جمله روش‌هایی است که در اعتبارسنجی مشتریان استفاده می‌شود. یکی از نکات مهم در کاربرد الگوریتم‌های ژنتیک، تعریف توابع شایستگی مناسب به‌منظور سنجش اعتباری مشتریان است. با توجه به روند رو به رشد کاربرد روش‌های هوشمند در حوزه کسب‌وکار، مقاله حاضر به مقایسه عملکرد سه نوع تابع شایستگی در الگوریتم ژنتیک به‌منظور رتبه‌بندی اعتباری مشتریان می‌پردازد و نتایج آن را با سایر روش‌های اعتبارسنجی مقایسه می‌کند.

ساختار مقاله بدین صورت طراحی شده است. در بخش پیشینه پژوهش، انواع روش‌های به‌کاررفته در زمینه اعتبارسنجی مشتریان مرور شده است. در بخش روش‌شناسی پژوهش، انواع روش‌های رایج اعتبارسنجی مشتریان که در تحقیق حاضر با الگوریتم ژنتیک مقایسه شده‌اند، معرفی می‌شود. در ادامه، پس از مرور الگوریتم ژنتیک، انواع توابع شایستگی قابل استفاده در اعتبارسنجی مشتریان تشریح می‌شود. سپس، با توجه به شاخص‌های مد نظر برای اعتبارسنجی مشتریان، آزمایش‌هایی طراحی شده و انواع روش‌های مد نظر روی این داده‌ها به اجرا درمی‌آید و نتایج آنها مقایسه می‌شود.

پیشینه تجربی پژوهش

مروری بر مطالعات مقایسه‌ای در اعتبارسنجی

با توسعه روش‌های اعتبارسنجی در دهه‌های ۱۹۵۰ و ۱۹۶۰، انواع مدل‌های آماری و غیرآماری برای ارزیابی اعتباری مشتریان ارائه شد. امروزه اغلب بانک‌ها و مؤسسه‌های اعتباری از اعتبارسنجی مشتریان به منظور کنترل ریسک اعتباری خود استفاده می‌کنند. روش‌های متنوعی برای ارزیابی اعتباری مشتریان به کار می‌رود که معرفی و مرور تمام آنها خارج از حیطه این مقاله است و علاقه‌مندان می‌توانند به منابع و مقالات موجود در این زمینه مراجعه کنند. با توجه به هدف تحقیق حاضر، این بخش بر مطالعات مقایسه‌ای میان انواع روش‌های اعتبارسنجی تمرکز کرده است. برخی مطالعات نشان داده‌اند که روش‌های هوش مصنوعی و محاسبات تکاملی می‌توانند پیش‌بینی‌های دقیق‌تری را در اختیار استفاده‌کنندگان قرار دهند. در مقابل برخی دیگر، هنوز بر این باورند که روش‌های قدیمی بهتر و مؤثرتر بوده یا استفاده همزمان از هر دو دسته روش‌ها، نتیجه خوبی را به همراه خواهد داشت (دسای، کانوی، کروک و اووراستریت، ۱۹۹۷). با وجود این، لی، چو و لو (۲۰۰۶) معتقدند در سیستم‌های نوین بانکداری، روش‌های سنتی تصمیم‌گیری در خصوص اعتبارسنجی مشتریان جای خود را به سیستم‌های هوشمند داده‌اند. آنگریک رافال (۲۰۰۲) به بررسی پتانسیل روش‌های طبقه‌بندی مبتنی بر هوش مصنوعی، برای بهبود رویه‌های اعطای اعتبار در بانک‌ها پرداخته است. او پس از مقایسه مدل‌های شبکه عصبی کوهنن و سیستم فازی، نتایج بهتری را از کاربرد مدل سیستم‌های فازی گزارش کرد.

کومرود (۲۰۰۲) با بررسی انواع روش‌های اعتبارسنجی، عملکرد این روش‌ها را روی مجموعه‌ای از داده‌های یک بانک فرانسوی بررسی کرده است. نتایج وی نشان داد روش‌های رگرسیون لجستیک و شبکه‌های عصبی MLP و RBF عملکرد مشابهی دارند، هرچند به نظر می‌رسد مدل سنتی تابع لاجیت از بقیه بهتر است.

اونگ، هوانگ و ژنگ (۲۰۰۵) در تحقیقی، برنامه‌ریزی ژنتیکی^۱ را برای ساخت مدل‌های اعتبارسنجی پیشنهاد دادند. برای این منظور دو نمونه را برای مقایسه مدل‌های اعتبارسنجی شامل شبکه‌های عصبی، درخت تصمیم، مجموعه‌های سخت و رگرسیون لجستیک به کار گرفتند. نتایج به دست آمده نشان داد در این خصوص برنامه‌ریزی ژنتیکی عملکرد بهتری دارد.

لی و همکارانش (۲۰۰۶) عملکرد دو روش داده‌کاوی، شامل CART و MARS را در اعتبارسنجی مشتریان بررسی کردند. در این پژوهش، برای سنجش کارایی این دو روش از

۱. یکی از روش‌های هوش مصنوعی مبتنی بر رمزگذاری برنامه‌های رایانه‌ای در قالب مجموعه‌ای از ژن‌ها و تنظیم و تکامل آنها با استفاده از یک الگوریتم تکاملی که توسط کوزو (۱۹۹۲) ارائه شده است.

مجموعه داده‌های کارت اعتباری یک بانک استفاده شده است. نتایج تحقیق نشان داد عملکرد دو روش یادشده نسبت به رویکردهای سنتی تحلیل ممیزی، رگرسیون لجستیک، شبکه‌های عصبی و ماشین بردار پشتیبان بهتر است.

هوانگ، چن و وانگ (۲۰۰۷) یک روش هیبریدی را برای اعتبارسنجی ارائه دادند. به این صورت که ترکیب الگوریتم ژنتیک و ماشین بردار پشتیبان هم برای انتخاب خصیصه‌ها و هم برای بهینه‌سازی پارامترهای مدل اعتبارسنجی به کار گرفته شده است. مقایسه‌هایی نیز بین روش ماشین بردار پشتیبان و شبکه‌های عصبی، برنامه‌ریزی ژنتیکی و درخت تصمیم صورت گرفته که نتایج کمابیش یکسانی داشته است. ولی نتایج تجربی نشان می‌دهد ماشین بردار پشتیبان نسبت به سایر روش‌های داده‌کاوی نتایج بهتری ارائه می‌دهد.

زاگروسکا (۲۰۰۷) تحقیق خود را بر پایه ترکیب طبقه‌بندی با ناظر و بدون ناظر استوار کرده است. به این صورت که در گام اول با استفاده از الگوریتم خوشه‌بندی k-means، مشتریان به گروه‌هایی با ویژگی‌های مشابه دسته‌بندی شدند و در گام بعدی برای هر گروه از مشتریان درخت تصمیم ساخته شده است. روش پیشنهادشده در این تحقیق به بانک اجازه می‌دهد از قوانین متفاوت برای مشتریان مختلف که ریسک‌های بالایی دارند، استفاده کند.

ذکاوت (۱۳۸۲) به بررسی یک نمونه ۱۲۰ تایی از مشتریان حقوقی بانک توسعه صادرات پرداخته است. پس از بررسی صورت‌های مالی مشتریان، پنج نسبت مالی به‌عنوان متغیرهای مؤثر بر رفتار اعتباری انتخاب شد و با استفاده از مدل‌های تحلیل ممیزی گام به گام و رگرسیون لجستیک، مشتریان از نظر ریسک اعتباری تفکیک شدند. یافته‌های تحقیق نشان می‌دهد متغیر نسبت جاری بیشترین قدرت تفکیک مشتریان خوش حساب و بد حساب را در میان سایر متغیرها دارد. از سوی دیگر، روش‌های تحلیل ممیزی و رگرسیون لجستیک در رابطه با دسته‌بندی مشتریان از نظر ریسک اعتباری، نتایج مشابهی ارائه می‌دهند.

ملا ابراهیملو (۱۳۸۴) مدل رتبه‌بندی را به صورت یک مدل پروبیت ترتیبی برآورد کرد. در این مدل از ویژگی‌های کمی و کیفی ۸۲ مشتری حقوقی بانک سامان استفاده شده است. این تحقیق به دنبال آزمون فرضیه‌های آماری در خصوص تأثیر هفت متغیر (ویژگی‌های مشتریان) بر کاهش یا افزایش احتمال نکول وام‌های دریافت شده بود. در نهایت مدل نهایی توانست مشتریان را بر اساس وضعیت اعتباری در سه حالت خوش حساب، سررسید گذشته و معوق، در سطح قابل قبول تفکیک کند.

کشاورز حداد و آیتی‌گازار (۱۳۸۶) در تحقیقی به مقایسه مدل رگرسیون لجستیک (لاجیت) و مدلی با استفاده از روش درخت‌های طبقه‌بندی برآورد شده پرداختند و دقت و کارایی آنها را

مقایسه کردند. نتایج نهایی نشان‌دهنده دقت بیشتر روش درخت‌های طبقه‌بندی از نظر خطای نوع اول و دوم در مقایسه با مدل لاجیت است.

رستمکلائی (۱۳۸۶) با استفاده از ۲۲ متغیر و ۲۳۵ قانون در چارچوب یک سیستم خبره فازی، الگویی برای ارزیابی اعتباری مشتریان تسهیلات اعتباری بانکی طراحی کرده است. استخراج دانش و ارزیابی متغیرها در این سیستم به‌گونه‌ای طراحی شده که به فرایند ذهنی ارزیابان نزدیک باشد. مدل استفاده‌شده، سیستم استدلال میدانی بود و خروجی سیستم، مشتریان را در چهار گروه مختلف بر اساس میزان اعتبارشان طبقه‌بندی کرد.

نیلی و سبزواری (۱۳۸۷) برای ارزیابی ۸۶ مورد از مشتریان حقوقی بانک پارسیان، با استفاده از اطلاعات مالی آنها، مدل‌های امتیازدهی لاجیت و روش غیرپارامتری AHP را بررسی کردند و این دو روش را از نظر میزان دقت در پیش‌بینی مشتریان خوب و بد با یکدیگر مقایسه نمودند. بر این اساس، روش AHP که یک روش کارشناس‌محور است، توانست با دقت نسبتاً بهتری ارزش اعتباری و در نهایت خوب و بد بودن مشتریان را تعیین کند.

البرزی، خان‌بابایی و محمدپورزندی (۱۳۸۹) نیز در پژوهشی برای اعتبارسنجی مشتریان بانک‌ها از الگوریتم ژنتیک در بهینه‌سازی درختان تصمیم‌گیری استفاده کردند که نتایج آنان نشان داد دقت طبقه‌بندی مدل پیشنهادی به‌طور تقریبی از تمام مدل‌های درخت تصمیم‌گیری مقایسه شده در این پژوهش بیشتر است.

همان‌طور که مشاهده می‌شود با توجه به اهمیت مسئله رتبه‌بندی اعتباری در صنعت بانکی، تحقیقات گسترده‌ای در این حوزه انجام شده است. طرح حاضر با در نظر گرفتن دامنه‌ای از توابع برازش، سعی در بسط تحقیقات پیشین با فراهم آوردن انعطاف بیشتر در ارزیابی اعتباری مشتریان دارد.

پیشینه نظری پژوهش

معرفی روش‌های رتبه‌بندی اعتباری

همان‌طور که اشاره شد، امروزه اغلب بانک‌ها و مؤسسه‌های اعتباری که به نوعی با پرداخت تسهیلات و اعطای اعتبار سروکار دارند، نسبت به رتبه‌بندی اعتباری مشتریان اقدام می‌کنند. نگاهی به الگوهای کمی اعتبارسنجی، گویای آن است که اغلب این مدل‌ها چارچوب معنایی و مفهومی مشابهی دارند، اما به شیوه‌های مختلفی نسبت به برآورد پارامترهای مورد نیاز اقدام می‌کنند. به‌طور کلی روش‌های ارزیابی اعتباری مشتریان به دو گروه عمده روش‌های پارامتری و غیر پارامتری دسته‌بندی می‌شوند (کیس، ۲۰۰۳). در خانواده روش‌های پارامتری، می‌توان به مدل‌های احتمال خطی، لاجیت، پروبیت و تحلیل ممیزی اشاره کرد. در روش‌های غیر پارامتری

نیز می توان روش های مبتنی بر برنامه ریزی ریاضی، طبقه بندی درختی، فرایند تحلیل سلسله مراتبی، تحلیل پوششی داده ها، شبکه های عصبی و الگوریتم ژنتیک را نام برد. نکته مهم در روش های اعتبارسنجی، معیارهایی است که به منظور ارزیابی مشتریان استفاده می شود. اگرچه توافق جامعی در این خصوص وجود ندارد، چارچوب های عمومی و کلانی برای راهنمایی کاربران ارائه شده است. از جمله می توان به الگوی 5C شامل پنج معیار شخصیت، ظرفیت، سرمایه، وثیقه، شرایط و ضوابط تسهیلات اشاره کرد. الگوی LAPP نیز از معیارهای نقدینگی، فعالیت، سودآوری و توان بالقوه برای اعتبارسنجی استفاده می کند. همچنین، الگوی 5P معیارهای مردم، محصول، حمایت، پرداخت ها و شمای کلی آینده را برای این منظور پیشنهاد داده است (عرب مازار و روئین تن، ۱۳۸۵). با توجه به اینکه هدف از مقاله حاضر ارائه روش تحلیلی برای اعتبارسنجی مشتریان با استفاده از الگوریتم ژنتیک است، خوانندگان علاقه مند برای آشنایی با سایر روش ها و معیارهای اعتبارسنجی می توانند به مقالات اشاره شده مراجعه کنند.

کاربرد الگوریتم ژنتیک در اعتبارسنجی مشتریان

الگوریتم ژنتیک، الگوریتمی مبتنی بر تکرار است که اصول اولیه آن از علم ژنتیک اقتباس شده و با تقلید از چند فرایند مشاهده شده در تکامل طبیعی، اختراع شده است. ژنتیک علمی درباره چگونگی توارث و انتقال صفات بیولوژیکی در موجودات زنده از طریق کروموزوم ها و ژن ها صحبت می کند. ایده اساسی آن است که در نهایت ژن ها و کروموزوم های برتر باقی مانده و ژن های ضعیف تر از بین می روند. الگوریتم های ژنتیک توسط هالند (۱۹۷۵) ارائه شده و در سطح گسترده ای برای حل انواع مسائل بهینه سازی دارای تابع هدف و بدون تابع هدف استفاده شده است (هائوپت و هائوپت، ۲۰۰۴). با توجه به هدف مقاله حاضر در خصوص مسئله اعتبارسنجی مشتریان، تمرکز این بخش بر کاربرد GA در زمینه فوق خواهد بود.

مسئله اعتبارسنجی مشتریان را می توان در قالب یک مسئله طبقه بندی تصور کرد. به طور کلی مشتریان به دو گروه خوش حساب و بد حساب طبقه بندی می شوند. هر چند امکان طبقه بندی های دیگری نیز وجود دارد، در مقاله حاضر از طبقه بندی کلاسیک استفاده شده است. تعریفی رسمی از مسئله طبقه بندی مشتریان به دو گروه به صورت زیر است:

هر مشتری، مانند x بر اساس D متغیر $x = (x_1, x_2, \dots, x_D)$ دسته بندی می شود که هر متغیر دامنه ای مانند $V_j, j \leq D$ دارد. بر این اساس، فضای مشخصه های ورود به صورت $V = \prod_{j=1}^n V_j = \{x_1, x_2, \dots, x_D | x_j \in V_j\}$ می توان یک تابع امتیازدهی به صورت $\{ \text{خوب}, \text{بد} \}$ تعریف کرد که نوع مشتری اعتباری را پیش بینی می کند. وضعیت واقعی متقاضی اعتباری در نمونه به صورت $\{ \text{خوب}, \text{بد} \}$ γ نشان داده

می‌شود. تابع شایستگی ترکیبی از تابع امتیازدهی و امتیاز متناظر با آن است. برای آموزش الگوریتم ژنتیک نمونه‌ای از مشتریان، S ، با مشخصه‌ها و طبقه اعتباری معلوم استفاده شده است.

$$S = \{(\vec{x}_1, y_1), \dots, (\vec{x}_N, y_N)\} \quad \text{رابطه (۱)}$$

که $\vec{x}_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ بردار مشخصه‌های مشتری و y_i طبقه اعتباری اوست. اعتبار طبقه‌بندی (ϕ) از طریق تعداد پیش‌بینی‌های صحیح تقسیم بر تعداد کل اعضای نمونه (N) محاسبه می‌شود.

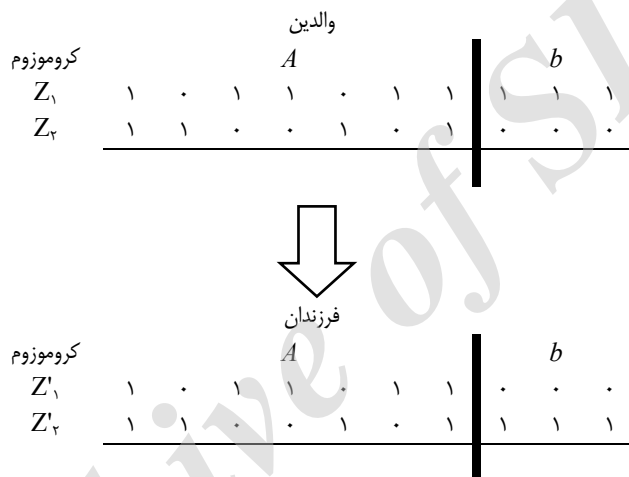
$$\phi(f) = \frac{|\{i \leq N | f(\vec{x}_i) = y_i\}|}{N} \quad \text{رابطه (۲)}$$

در مقاله حاضر سه تابع شایستگی قابل استفاده در الگوریتم ژنتیک برای اعتبارسنجی مشتریان بررسی شده است: ۱. تابع شایستگی خطی؛ ۲. تابع شایستگی نمایی و ۳. تابع شایستگی کوادراتیک. همچنین، به منظور حذف تأثیر پارامترهای الگوریتم ژنتیک بر مسئله، این پارامترها برای هر سه تابع یکسان در نظر گرفته شده‌اند. برای این منظور، از مقادیر پیشنهادشده در مقاله کوزنی (۲۰۱۵) در خصوص پارامترهای الگوریتم ژنتیک استفاده شده است.

هر مدل با جمعیت اولیه از ۴۰۰ کروموزوم آغاز شده است. طول هر کروموزوم بستگی به روش استفاده‌شده دارد که در بخش‌های بعدی توضیح داده می‌شود. پس از تولید جمعیت اولیه، بر اساس مراحل الگوریتم ژنتیک، مجموعه‌ای از گام‌های تکراری دنبال شده است. ابتدا، شایستگی هر کروموزوم در جمعیت اولیه برآورد می‌شود؛ سپس همه کروموزوم‌ها بر اساس امتیاز شایستگی خود مرتب شده و ۵ درصد از جواب‌های عالی بدون تغییر به جمعیت بعد منتقل می‌شود. همچنین، از مهاجرت رو به جلو برای تکرار ۲۰ درصد از بهترین کروموزوم‌ها به نسل بعدی در هر ۲۰ تکرار استفاده می‌شود. گام سوم انتخاب بخشی از جمعیت برای تقاطع است. برای انتخاب کروموزوم‌ها از نمونه‌گیری تصادفی یکنواخت استفاده می‌شود که مشابه با روش چرخ رولت است. چرخ رولت به شیوه‌های گوناگونی تشکیل می‌شود که یکی از پرکاربردترین روش‌ها، متناسب با شایستگی است. در این شیوه چرخ به m بخش تقسیم می‌شود که m تعداد کروموزوم در جمعیت است. اندازه آن نیز برابر با شایستگی کروموزوم است. در این روش کروموزوم‌هایی با شایستگی بیشتر، شانس بیشتری برای انتخاب دارند. احتمال انتخاب p_i هر کروموزوم بستگی به امتیاز شایستگی $\phi(f_i)$ آن دارد و به صورت زیر تعریف می‌شود.

$$p_i = \frac{\phi(f_i)}{\sum_{i=1}^m \phi(f_i)} \quad \text{رابطه (۳)}$$

پس از انتخاب زوج کروموزوم‌های والد، کروموزوم‌های نسل بعدی به سه صورت شکل می‌گیرند: تکرار بدون تغییر والدین در نسل بعدی، تقاطع یا جهش. احتمال تقاطع (p_c) ۸۰ درصد در نظر گرفته شده و برای آن از عملگر تقاطع تک نقطه‌ای استفاده شده است. در این روش تقاطع، نقطه تصادفی یکسانی در دو کروموزوم والد Z_1 و Z_2 انتخاب می‌شود؛ سپس این دو کروموزوم به دو بخش طبقه‌بندی شده که بخش قبل از نقطه انتخاب به ترتیب با Z_{1a} و Z_{2a} و بخش پس از این نقطه به ترتیب با Z_{1b} و Z_{2b} نشان داده می‌شود. کروموزوم‌های فرزند Z'_1 و Z'_2 با جابه‌جایی بخش‌های بعد و قبل نقطه تقاطع به صورت $Z_{1b}Z_{2a}$ و $Z_{1a}Z_{2b}$ تشکیل می‌شوند (شکل ۱).



شکل ۱. تقاطع تک نقطه‌ای

انتخاب نخبگان و مهاجرت رو به جلو، برای اطمینان از عدم تخریب بهترین کروموزوم‌ها به وسیله تقاطع و جهش انجام شده است. با توجه به اینکه اتخاذ دو سیاست فوق موجب کاهش تنوع جمعیت و امکان گیر افتادن در جواب بهینه محلی می‌شود، احتمال نسبتاً بالایی برای جهش ($p_m = 15\%$) در نظر گرفته شده است.

تابع خطی

در این مدل یک تابع چند جمله‌ای به صورت زیر برای طبقه‌بندی متقاضیان وام استفاده شده است:

$$f(x) = a_1x_1 + a_2x_2 + \dots + a_Dx_D + c \quad \text{رابطه ۴}$$

کروموزوم بر اساس پارامترهای a_1, a_2, \dots, a_D و c تشکیل می‌شود که هدف برآورد آنهاست. در این حالت طول هر کروموزوم برابر $D + 1$ خواهد بود. پس از برآورد پارامترها و تشکیل تابع $f(x)$ ، مقدار آن با حد آستانه صفر مقایسه می‌شود. اگر $f(x) > 0$ به دست آمد، متقاضی به عنوان خوب (خوش حساب) شناخته شده و اگر $f(x) \leq 0$ به دست آمد، متقاضی به عنوان بد (بد حساب) معرفی می‌شود. امتیاز شایستگی ϕ بر اساس رابطه ۲ محاسبه می‌شود.

تابع نمایی

در این حالت، از مدل زیر برای برازش داده‌ها استفاده می‌شود (توماس، ادلمن و کروک، ۲۰۰۲):

$$f(x) = a_1 x_1^{b_1} + a_2 x_1^{b_2} + \dots + a_D x_1^{b_D} + c \quad (\text{رابطه ۵})$$

در این حالت ژن‌های هر کروموزوم شامل پارامترهای $a_1, a_2, \dots, a_D, b_1, b_2, \dots, b_D$ و c است. طول هر کروموزوم نیز برابر $D + 1$ خواهد بود. بار دیگر پس از برآورد پارامترهای تابع (رابطه ۵)، اگر $f(x) > 0$ به دست آمد، متقاضی خوش حساب است و اگر $f(x) \leq 0$ به دست آمد متقاضی بد حساب شناخته می‌شود. امتیاز شایستگی ϕ بر اساس رابطه ۲ محاسبه می‌شود.

تابع کوادراتیک

با توجه به اینکه ممکن است اثرات متقابل و درجه دو متغیرها نیز بر برآورد امتیاز هر متقاضی تسهیلات اثرگذار باشد، تابع دیگری که برای برازش امتیاز متقاضیان استفاده شده تابع درجه دو یا کوادراتیک است (رابطه ۶).

$$f(x) = \sum_{j=1}^D a_j x_j + \sum_{i=1}^{D-1} \sum_{j=i+1}^D a_{ij} x_i x_j + \sum_{j=1}^D a_{jj} x_j^2 + c \quad (\text{رابطه ۶})$$

که در این حالت $a_1, a_2, \dots, a_D, a_{12}, a_{13}, \dots, a_{D-1D}, a_{11}, a_{22}, \dots, a_{DD}$ و c پارامترهای مربوط به ژن‌های هر کروموزوم است.

در نتیجه طول هر کروموزوم برابر $(D(D-1)/2) + 1$ خواهد بود. بار دیگر پس از برآورد پارامترهای تابع (رابطه ۶)، اگر $f(x) > 0$ به دست آمد، متقاضی خوش حساب است و اگر $f(x) \leq 0$ به دست آمد متقاضی بد حساب شناخته می‌شود. امتیاز شایستگی ϕ بر اساس رابطه ۲ محاسبه می‌شود.

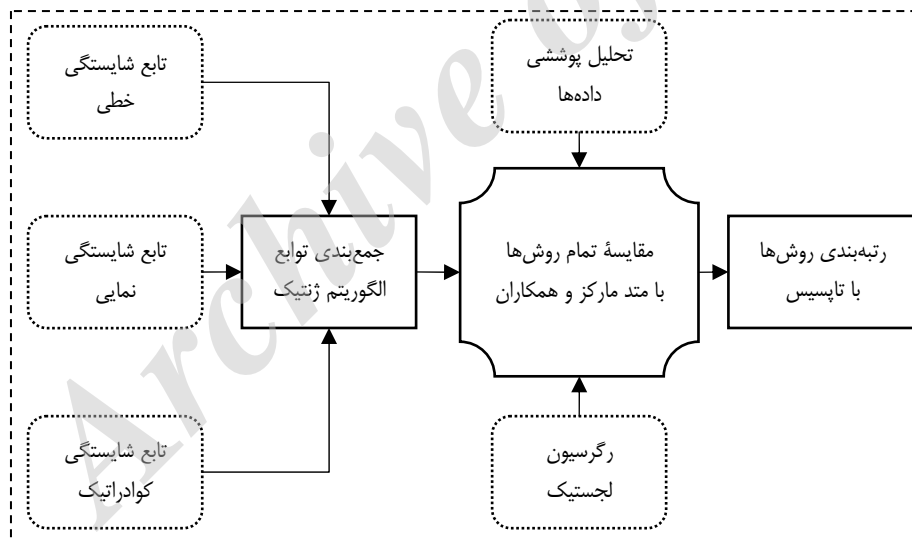
برای برازش توابع فوق با استفاده از الگوریتم ژنتیک، تابع شایستگی به صورت زیر تعریف می‌شود:

$$\text{Min} \sum_{j=1}^n s_j (y_j - f(x_j)) \quad \text{رابطه ۷}$$

که y ارزش اعتباری واقعی متقاضی z ام، n تعداد متقاضیان بررسی شده و $f(x_j)$ امتیاز برآورد شده توسط هر یک از توابع خطی، نمایی و کوادراتیک است.

روش شناسی پژوهش

در تحقیق حاضر، ابتدا وضعیت سه تابع شایستگی الگوریتم ژنتیک روی ۴۰۰ نمونه مشتری متقاضی تسهیلات، برآزش شد؛ سپس نتایج به دست آمده به همراه دو روش تحلیل پوششی داده‌ها و رگرسیون لجستیک بر اساس رویکرد مارکز، گارسیا و سانچز (۲۰۱۳) با یکدیگر مقایسه شدند تا عملکرد برتر شناسایی شود. در پایان نیز بر اساس معیارهای مختلف، روش‌های استفاده شده به کمک روش تاپسیس رتبه‌بندی شدند تا مقایسه پنج رویکرد بررسی شده در تحقیق حاضر مشخص شود. چارچوب اجرایی تحقیق حاضر در شکل ۲ مشاهده می‌شود.



شکل ۲. مراحل اجرای تحقیق

چارچوب گردآوری داده‌ها

برای بررسی و مقایسه عملکرد انواع روش‌های اعتبارسنجی، نمونه‌ای از ۴۰۰ مشتری اعتباری متقاضی تسهیلات یکی از بانک‌های خصوصی کشور به‌عنوان مجموعه داده‌های یادگیری

استفاده شده است. این نمونه شامل ۲۰۵ مشتری خوش حساب و ۱۹۵ مشتری بدحساب است. مشخصه‌های هر مشتری در قالب ۱۰ ویژگی بررسی و ارزیابی شده‌اند که این ویژگی‌ها عبارت‌اند از: سن؛ وضعیت تأهل؛ تحصیلات؛ سطح اشتغال؛ سال‌های اشتغال در کار جاری؛ وضعیت اشتغال همسر متقاضی؛ درآمد متقاضی؛ میزان تسهیلات دریافت‌شده؛ وضعیت مسکن؛ مالکیت خودرو. این شاخص‌ها پس از مرور مقالات و تحقیقات مشابه شناسایی شده‌اند (مارشال، تانگ و میلنه، ۲۰۱۰؛ ایناف، ینکینز و لوین، ۲۰۱۳؛ بلانکو، پینو میاس، رایو، ۲۰۱۳؛ کوزنی، ۲۰۱۵).

نتایج به‌دست آمده از الگوریتم ژنتیک

پس از گردآوری داده‌ها، به تجزیه و تحلیل آنها با روش‌های الگوریتم ژنتیک و سه تابع شایستگی فوق اقدام شده است. برای تنظیم پارامترهای الگوریتم ژنتیک، برخی محققان از روش‌های بهینه‌سازی و تنظیم پارامترها استفاده کرده‌اند (آنگلوا و پنچوا، ۲۰۱۱). از سوی دیگر می‌توان تنظیمات به‌کار رفته و تأییدشده در تحقیقات و مطالعات پیشین را برای پارامترها به‌کار برد. در تحقیق حاضر پارامترهای الگوریتم ژنتیک به‌منظور برازش توابع اعتبارسنجی بر اساس تحقیق کوزنی (۲۰۱۵) تنظیم شده‌اند که شامل احتمال تقاطع ۸۰ درصدی، احتمال جهش ۱۵ درصدی، عملکرد تقاطع تک‌نقطه‌ای و روش انتخاب نمونه‌گیری تصادفی یکنواخت است.

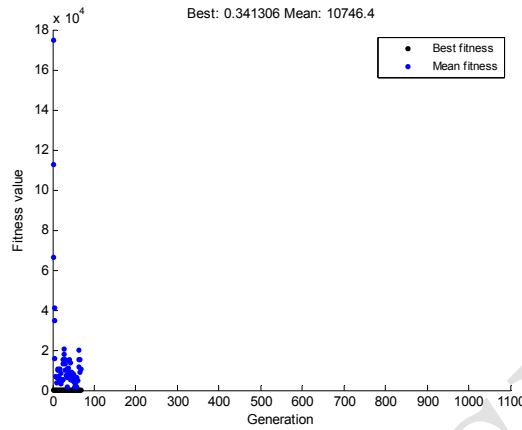
برازش تابع خطی

با توجه به آنچه در بخش‌های قبل دربارهٔ تابع خطی و برازش آن با استفاده از الگوریتم ژنتیک مطرح شد، از optimtool نرم‌افزار MATLAB برای برآورد توابع استفاده شد که تنظیمات تنظیمات الگوریتم به‌صورت زیر به‌دست آمد:

Elticism: 5%
 Forward mitigation 20%
 Mutation 15%
 Selection: stochastic uniform

پس از اجرای کد، نمودار بهترین مقادیر برازنده شده و میانگین مقادیر در شکل ۳ نشان داده شده است. همان‌طور که شکل ۳ نشان می‌دهد، بهترین مقدار برازش ۰/۳۴۱۳۰۸ است. بر اساس نتایج، تابع خطی زیر برای اعتبارسنجی مشتریان برآورده شده است:

$$y = -9.219 + 0.876x_1 + 0.768x_2 - 1.893x_3 - 2.6x_4 + 4.66x_5 - 6.035x_6 - 3.395x_7 - 7.977x_8 + 0.381x_9 - 1.276x_{10}$$



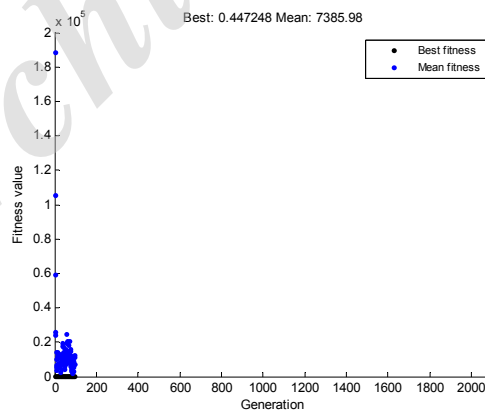
شکل ۳. نمودار میانگین و بهترین مقادیر برازش

برازش تابع نمایی

با تنظیمات مشابه تابع خطی، تابع نمایی به صورت زیر برای ارزیابی امتیاز اعتباری برآورد شده است:

$$y = 1.252 + 1.036x_1^{1.076} - 5.596x_2^{-5.835} + 8.676x_3^{1.736} - 4.176x_4^{7.514} + 8.464x_5^{0.547} - 2.535x_6^{4.209} + 0.082x_7^{-6.708} + 3.882x_8^{-9.263} + 7.917x_9^{3.705} + 8.853x_{10}^{1.518}$$

نمودار بهترین مقادیر برازنده شده و میانگین مقادیر در شکل ۴ نشان داده شده است.



شکل ۴. نمودار میانگین و بهترین مقادیر برازش تابع نمایی

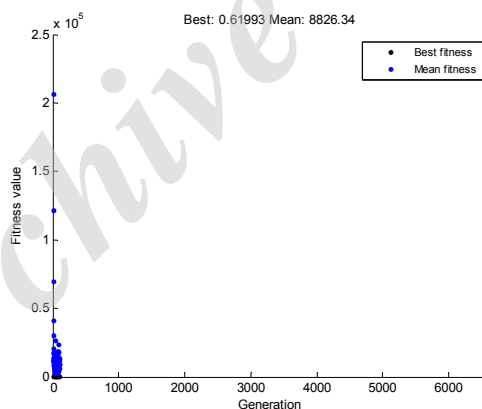
بر اساس شکل ۴، بهترین مقدار برازش ۰/۴۴۷۲۴۸ است.

برازش تابع کوادراتیک

با تنظیمات مشابه، تابع کوادراتیک به صورت زیر برای ارزیابی امتیاز اعتباری برآورد شده است:

$$y = 4.145 + 1.785x_1 + 3.102x_2 + 0.894x_3 + 2.411x_4 - 6.913x_5 - 5.514x_6 + 0.345x_7 - 5.129x_8 - 8.175x_9 + 9.016x_{10} - 1.935x_1^2 + 7.888x_2^2 + 8.780x_3^2 - 4.045x_4^2 - 7.167x_5^2 - 1.122x_6^2 - 3.274x_7^2 - 8.540x_8^2 + 2.345x_9^2 - 8.574x_{10}^2 - 7.154x_1x_2 + 4.209x_1x_3 + 6.803x_1x_4 - 5.385x_1x_5 - 3.999x_1x_6 + 8.875x_1x_7 + 6.838x_1x_8 - 3.078x_1x_9 - 1.399x_1x_{10} + 1.684x_2x_3 - 0.191x_2x_4 - 8.723x_2x_5 + 9.000x_2x_6 + 2.469x_2x_7 + 9.436x_2x_8 - 4.154x_2x_9 - 4.538x_2x_{10} + 0.901x_3x_4 + 2.280x_3x_5 - 3.842x_3x_6 + 6.102x_3x_7 + 0.024x_3x_8 + 1.980x_3x_9 + 5.178x_3x_{10} - 0.575x_4x_5 - 7.227x_4x_6 - 6.989x_4x_7 - 9.650x_4x_8 + 6.872x_4x_9 - 6.677x_4x_{10} + 7.323x_5x_6 - 0.472x_5x_7 - 9.863x_5x_8 + 7.032x_5x_9 + 4.843x_5x_{10} - 6.032x_6x_7 + 6.783x_6x_8 + 8.667x_6x_9 - 2.359x_6x_{10} + 1.700x_7x_8 + 7.501x_7x_9 + 3.013x_7x_{10} + 6.948x_8x_9 + 2.848x_8x_{10} + 3.421x_9x_{10}$$

نمودار بهترین مقادیر برازنده شده و میانگین مقادیر در شکل ۵ نشان داده شده است.



شکل ۵. میانگین و بهترین مقادیر برازش تابع نمایشی

بر اساس شکل ۵، بهترین مقدار برازش ۰/۶۱۹۹۳ است.

تحلیل و اعتبارسنجی نتایج

با توجه به اینکه در مقاله حاضر، الگوریتم ژنتیک به عنوان روشی برای برازش توابع شایستگی خطی و غیرخطی به منظور رتبه‌بندی اعتباری مشتریان استفاده شده است، در این بخش از

تحقیق، عملکرد الگوریتم و توابع پیشنهاد شده نسبت به روش های کلاسیک اعتبارسنجی شامل رگرسیون لجستیک و تحلیل پوششی داده ها، مقایسه می شود. این مقایسه به منظور بررسی عملکرد روش برازش پیشنهاد شده با روش های کلاسیک رتبه بندی اعتباری انجام می گیرد. ارزیابی توابع برازش شده روی مجموعه ای از ۱۰۰ مشتری به عنوان مجموعه داده های آزمایش بررسی شده است. برای ارزیابی نتایج، از سه شاخص صحت (مارکز و همکارانش، ۲۰۱۳)، تشخیص و حساسیت استفاده شده است. جدول ۱ را در نظر بگیرید.

جدول ۱. نتایج دسته بندی

پیش بینی شده	مشاهده شده		کل
	خوش حساب	بد حساب	
خوش حساب	n_{GG}	n_{GB}	$n_{GG} + n_{GB}$
بد حساب	n_{BG}	n_{BB}	$n_{BG} + n_{BB}$
کل	$n_{GG} + n_{BG}$	$n_{GB} + n_{BB}$	

صحت (A) نشان دهنده نسبت متقاضیان خوش حساب و بدحسابی است که به درستی طبقه بندی شده اند.

$$A = \frac{(n_{GG} + n_{BB})}{(n_{GG} + n_{BB} + n_{GB} + n_{BG})} \quad \text{رابطه ۸}$$

تشخیص (Sp) نسبت مشتریان خوش حسابی است که به درستی طبقه بندی شده اند.

$$Sp = \frac{n_{GG}}{(n_{GG} + n_{BG})} \quad \text{رابطه ۹}$$

حساسیت (Se) نیز نسبت مشتریان بدحسابی است که به درستی طبقه بندی شده اند.

$$Se = \frac{n_{BB}}{(n_{BB} + n_{GB})} \quad \text{رابطه ۱۰}$$

در استفاده از رگرسیون لجستیک، وضعیت خوش حساب (۱) یا بدحساب (۰) بودن متقاضی به عنوان متغیر وابسته باینری و مقادیر سایر متغیرها به عنوان متغیر مستقل به کار رفته اند (بولتون، ۲۰۱۰). در استفاده از روش تحلیل پوششی داده ها، از رویکرد چنگ، چیانگ و تانگ (۲۰۰۷) بهره برده شد. برای ارزیابی اعتبار متقاضی، یک واحد ایده آل (DEA_{ideal}) به عنوان ملاک سنجش اعتبار برای مدل مد نظر قرار گرفت. بر این اساس شاخص های از جنس مثبت (شامل وضعیت تأهل، تحصیلات، سابقه کار، اشتغال / عدم اشتغال همسر، درآمد ماهانه، وضعیت مسکن و

مالکیت خودرو) به‌عنوان خروجی و شاخص‌های از جنس منفی (شامل سن، سطح اشتغال و میزان تسهیلات دریافتی) به‌عنوان ورودی در نظر گرفته شدند. جدول ۲ نتایج ارزیابی و جدول ۳ شاخص‌های به‌دست آمده از روش‌های بالا را نشان می‌دهد.

جدول ۲. نتایج رتبه‌بندی با روش‌های مختلف

روش	n_{GG}	n_{GB}	n_{BG}	n_{BB}
تابع خطی	۱۷	۳۹	۱۷	۲۷
تابع نمایی	۸	۳۱	۸	۱۵
تابع درجه دو	۸	۴۸	۸	۳۶
رگرسیون لجستیک	۴۳	۱۳	۴۳	۱
DEA_{ideal}	۱۵	۴۱	۱۵	۲۹

جدول ۳. شاخص‌های اعتبار رتبه‌بندی توسط روش‌های مختلف

روش	صحت	تشخیص	حساسیت
تابع خطی	۰/۴۴	۰/۵۰	۰/۴۰۹
تابع نمایی	۰/۳۷۱	۰/۵۰	۰/۳۳۶
تابع درجه دو	۰/۴۴۰	۰/۵۰	۰/۴۳۹
رگرسیون لجستیک	۰/۴۴۰	۰/۵۰	۰/۰۷۱
DEA_{ideal}	۰/۴۴۰	۰/۵۰	۰/۴۱۴

همان‌طور که در جدول ۳ مشاهده می‌شود، بالاترین درجه صحت مربوط به توابع خطی و درجه دو، رگرسیون لجستیک و روش ایده‌آل تحلیل پوششی داده‌ها است. تمام روش‌ها در شاخص تشخیص، عملکرد مشابهی داشته‌اند. بالاترین درجه حساسیت نیز مربوط به تابع درجه دو است. با در نظر گرفتن وزن یکسان برای این شاخص و کاربرد الگوریتم TOPSIS، رتبه‌بندی عملکرد روش‌های اعتبارسنجی به صورت زیر است:

۱. تابع درجه دو با نزدیکی نسبی ۱؛
 ۲. تحلیل پوششی داده‌های ایده‌آل با نزدیکی نسبی ۰/۹۵۹؛
 ۳. تابع خطی با نزدیکی نسبی ۰/۹۴۵؛
 ۴. تابع نمایی با نزدیکی نسبی ۰/۶۸۴؛
 ۵. رگرسیون لجستیک با نزدیکی نسبی ۰/۱۳۸.
- در واقع بر اساس شاخص‌های فوق، بهترین عملکرد در تشخیص مشتریان مربوط به تابع درجه دو است.

بحث و نتیجه گیری

در تحقیق حاضر سه تابع برازش قابل استفاده در الگوریتم ژنتیک برای اعتبارسنجی مشتریان بررسی پیشنهاد شد: ۱. تابع شایستگی خطی، ۲. تابع شایستگی نمایی و ۳. تابع شایستگی کوادراتیک. همچنین، به منظور حذف تأثیر پارامترهای الگوریتم ژنتیک بر مسئله، این پارامترها برای هر سه تابع یکسان در نظر گرفته شدند. تابع خطی در واقع یک تابع چندجمله‌ای به صورت $f(x) = a_1x_1^{b_1} + a_2x_2^{b_2} + \dots + a_Dx_1^{b_D} + c$ است، تابع نمایی تابعی است که به شکل $f(x) = a_1x_1^{b_1} + a_2x_2^{b_2} + \dots + a_Dx_1^{b_D} + c$ به صورت $f(x) = \sum_{j=1}^D a_jx_j + \sum_{i=1}^{D-1} \sum_{j=i+1}^D a_{ij}x_ix_j + \sum_{j=1}^D a_{jj}x_j^2 + c$ می‌شود. برای برآورد پارامترهای این توابع، از یک تابع شایستگی به صورت $Min \sum_{j=1}^n S_j (y_j - f(x_j))$ الگوریتم ژنتیک، مقادیر پارامترها به دست آمد و توابع مد نظر برازنده شدند.

پس از بررسی و تحلیل‌های انجام شده مشخص شد، تابع درجه دو بهترین عملکرد را در برآورد ارزش اعتباری مشتریان در بانک مورد بررسی داشته است که به حتم این نتیجه برای بانک‌های دیگر متفاوت خواهد بود. بر اساس این یافته‌ها، پیشنهادهای زیر ارائه شده است:

با وجود اهمیت زیاد بحث رتبه‌بندی اعتباری در کاهش احتمال نکول در بازگشت وام، بسیاری از مؤسسه‌های فعال در بازار مالی از روش‌های سنتی برای ارزیابی اعتباری متقاضیان تسهیلات استفاده می‌کنند. الگوریتم ژنتیک به عنوان یکی از الگوریتم‌های تکاملی، ابزار قدرتمندی برای این مؤسسه‌ها در ارزیابی اعتباری مشتریان خود خواهد بود.

در طرح حاضر مجموعه‌ای از یازده متغیر برای ارزیابی اعتباری مشتریان پیشنهاد شد. هر بانک یا مؤسسه اعتباری بر اساس شرایط و مقتضیات خود می‌تواند از شاخص‌های متفاوتی برای این منظور استفاده کند. الگوریتم پیشنهاد شده هیچ محدودیتی برای نوع و تعداد شاخص‌های استفاده شده ندارد.

با توجه به قابلیت‌های الگوریتم ژنتیک، هیچ محدودیتی در شکل و نوع توابع استفاده شده برای برازش ارزش اعتباری مشتریان وجود ندارد. بر این اساس علاوه بر توابع پیشنهاد شده، هر تابع دیگری نظیر تابع bitmask و... قابل استفاده است.

محققان می‌توانند با تعریف طبقات بیشتر، الگوریتم را برای طبقه‌بندی مشتریان در طبقات بیشتر تعمیم دهند. همچنین می‌توان روش‌های دیگری را برای مقایسه توابع شایستگی با مقادیر فازی در الگوریتم ژنتیک به کار برد. علاوه بر این، کاربرد روش‌های دیگری نظیر شبکه‌های عصبی و مقایسه نتایج آنها توصیه می‌شود. در بحث انتخاب متغیرهای اعتبارسنجی نیز،

روش‌هایی نظیر تحلیل پوششی داده‌های گام به گام برای انتخاب متغیرهای مناسب اعتبارسنجی به‌منظور تحقیقات آتی پیشنهاد می‌شود. ضمن آن که تحقیق در تنظیم بهینه‌سازی پارامترهای الگوریتم ژنتیک در این زمینه می‌تواند موضوعی برای مطالعات آتی باشد.

فهرست منابع

- البرزی، م.، محمدپورزرندی، م.، خان‌بابایی، م. (۱۳۸۹). به‌کارگیری الگوریتم ژنتیک در بهینه‌سازی درختان تصمیم‌گیری برای اعتبارسنجی مشتریان بانک‌ها، فصلنامه مدیریت فناوری اطلاعات، ۲(۴)، ۳۸-۲۳.
- ذکاوت، س. (۱۳۸۲). مدل‌های ریسک اعتباری مشتریان بانک توسعه صادرات ایران، پایان‌نامه کارشناسی ارشد رشته بانکداری، مؤسسه عالی بانکداری.
- رجب‌زاده قطری، ع.، بهرام‌میرزایی، آ.، احمدی، پ. (۱۳۸۸). طراحی سیستم هوشمند ترکیبی رتبه‌بندی اعتباری مشتریان بانک‌ها با استفاده از مدل‌های استدلالی فازی ترکیبی. پژوهشنامه بازرگانی، ۱۴(۵۳)، ۲۰۱-۱۵۹.
- رستم کلایی، س. (۱۳۸۶). ارائه مدل تصمیم‌یار فازی برای ارزیابی وام بانکی در مدیریت بانکداری الکترونیک، پایان‌نامه کارشناسی ارشد مدیریت فناوری اطلاعات، تهران: دانشگاه شهید بهشتی.
- شایان آرانی، ش. (۱۳۸۰). نوآوری در ابزارهای مالی بانکداری اسلامی. مجله پیام بانک، (۲۹۲)، ۵-۵.
- عرب مازار، ع.، رویین‌تن، پ. (۱۳۸۵). عوامل مؤثر بر ریسک اعتباری مشتریان بانکی (مطالعه موردی: بانک کشاورزی). جستارهای اقتصادی، ۳(۶)، ۸۰-۴۵.
- کشاورز حداد، غ.، آیتی گازار، ح. (۱۳۸۶). مقایسه کارکرد مدل لاجیت و روش درخت‌های طبقه‌بندی و رگرسیون در فرایند اعتبارسنجی متقاضیان حقیقی برای استفاده از تسهیلات بانکی، پژوهش‌های اقتصادی (رشد و توسعه پایدار)، ۷(۴)، ۹۷-۷۱.
- ملاابراهیم‌لو، م. (۱۳۸۴). تدوین یک مدل رتبه‌بندی اعتباری برای مشتریان حقوقی بانک سامان، پایان‌نامه کارشناسی ارشد اقتصاد، تهران: دانشگاه شریف.
- نیلی، م.، سبزواری، ح. (۱۳۸۷). برآورد و مقایسه مدل درجه‌بندی اعتباری لاجیت با روش تجزیه و تحلیل سلسله‌مراتبی (AHP). ثریف ویژه علوم مهندسی (ویژه مهندسی صنایع، مدیریت و اقتصاد)، ۲۴(۴۳)، ۱۱۷-۱۰۵.
- Alborzi, M., Mohammad Pourzarandi, M. E. & Khanbabaie, M. (2011). Using Genetic Algorithm in Optimizing Decision Trees for Credit Scoring of Banks Customers. *Journal of Information Technology Management*, 2(4), 23-38. (in Persian)

- Angelova, M., Pencheva, T. (2011). Tuning genetic algorithm parameters to improve convergence time. *International Journal of Chemical Engineering*, 2011, 1-7.
- Angryk Rafal, A. (2002). *Credit risk assessment for a small business, using multidimensional membership functions: design and applications*. Master thesis, Poland: Technical University of Szczecin.
- Arabmazar, A. & Ruintan, P. (2006). Determinants of Credit Risk among Bank Clients A Case Study; Agricultural Bank of Iran. *Journal of Iran's Economic Essays*, 3(6), 45-80. (in Persian)
- Blanco, A., Pino-Mejías, R. & Rayo, S. (2013). Credit scoring models for the microfinance industry using neural networks: evidence from Peru. *Expert Systems with Applications*, 40(1), 356-364.
- Bolton, C. (2010). *Logistic regression and its application in credit scoring*. Pertoria: University of Pertoria.
- Cheng, E. W. L., Chiang, Y. H., Tang, B. S. (2007). Alternative approach to credit scoring by DEA: Evaluating borrowers with respect to PFI projects. *Building and Environment*, 42(4), 1752-1760.
- Desai, V.S., Conway, D.G., Crook, J.N. & Overstreet, G.A. (1997). Credit-scoring models in the credit-union environment using neural networks and genetic algorithms. *IMA Journal of Management Mathematics*, 8(4), 323-346.
- Einav, L., Jenkins, M., Levin, J. (2013). The impact of credit scoring on consumer lending. *RAND Journal of Economics*, 44(2), 249-274.
- Haupt, R. L. & Haupt, S. E. (2004). *Practical genetic algorithms*. 2nd ed., New Jersey: John Wiley & Sons.
- Holland, J. H. (1975). *Adaption in natural and artificial systems*. Michigan: University of Michigan Press.
- Huang, C. L., Chen, M. C. & Wang, C. J. (2007). Credit scoring with data mining approach based on support vector machines. *Expert Systems with Applications*, 33(4), 847-856.
- Kao, D. L. & Kallberg, J. (1994). Strategies for measuring and managing risk concentration in loan portfolios. *Journal of Commercial Bank Lending*, 76(5), 18-27.
- Keshavarz Haddad, G. R. & Ayati Gazar, H. (2008). A Comparison between Logit Model and Classification Regression Trees (CART) in Customer Credit Scoring Systems. *The Economic Research (Scientific Research Quarterly)*, 7(4), 71-97. (in Persian)
- Kiss, F. (2003). Credit Scoring Processes from a Knowledge Management. *Periodica Polytechnica*, 11(1), 95-110.

- Komorad, K. (2002). *On credit scoring estimation*. Berlin: Masters thesis Humboldt University.
- Kozeny, V. (2015). Genetic algorithms for credit scoring: Alternative fitness function performance comparison. *Expert Systems with Applications*, 42(6), 2998-3004.
- Lee, T.S., Chiu, C.C., Chou, Y.C. & Lu, C.J. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Systems*, 50(4), 1113-1130.
- Marques, A. I., Garcia, V. & Sanchez, J. S. (2013). A literature review on the application of evolutionary computing to credit scoring. *Journal of the Operational Research Society*, 64(9), 1384-1399.
- Marshall, A., Tang, L. & Milne, A. (2010). Variable reduction, sample selection bias and bank retail credit scoring. *Journal of Empirical Finance*, 17(3), 501-512.
- Mollaebrahimlo, M. H. (2005). *A model of corporate customer credit rating for Saman Bank*. Master Thesis, Tehran: Sharif University of Technology. (in Persian)
- Nili, M. & Sabzevari, H. (2008). Estimation and comparison of Logit credit rating model with analytic hierarchy process (AHP) method. *Sharif Journal*, 43, 105-117. (in Persian)
- Ong, C. S., Huang, J. J. & Tzeng, G.H. (2005). Building credit scoring models using genetic programming. *Expert Systems with Applications*, 29(1), 41-47.
- Rajabzadeh Ghatari, A., Bahram Mirzaei, A. & Ahmadi, P. (2010). Hybrid Intelligent Credit Ranking System Using Fuzzy Hybrid-Reasoning Models. *Iranian Journal of Trade Studies Quarterly*, 14(53), 159-201. (in Persian)
- Risk Management Group of the Basel Committee (2000). *Principles for the Management of Credit Risk*. Basel Committee on Banking Supervision, Basel.
- Rostamkolaei, S.A. (2007). *Fuzzy decision support model for loaning in electronic banking management*. Master Thesis, Tehran: Shahid Beheshti University. (in Persian)
- Shayanarani, S. (2001). Innovation in Islamic banking financial tools. *Bank Message*, 292, 5-5. (in Persian)
- Thomas, L. C., Edelman, D. B. & Crook, J. N. (2002). *Credit scoring and its applications*. Philadelphia: SIAM.
- Zakrzewska, D. (2007). On integrating unsupervised and supervised classification for credit risk evaluation. *Information Technology and Control*, 36(1), 98-102.
- Zekavat, S.M. (2007). *Iran Export Development Bank's Customers credit rating models*. Master Thesis, Tehran: Iran Banking Institute. (in Persian)