فصلنامه علمی پژوهشی دانش سرمایه‌گذاری
دوره ۱۶ / شماره ۱ (پیاپی ۶۱) / بهار ۱۴۰۶
صفحه ۲۱۵ تا ۲۲۷

ارزیابی ریسک اعتباری در وام‌دهی نفر به نفر با استفاده از روش‌های رگرسیون هسته‌ای وزن‌دار شده و رتبه‌بندی ساده

هابیل خاوری
گروه حسابداری و مالی، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران
habilkhavari@yahoo.com

فاطمه احمدی
گروه حسابداری و مالی، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران(نویسنده مسئول).
fatemehahmady۶۰@gmail.com

مجتبی مرادپور
گروه حسابداری و مالی، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران.
Moradpour.mo@gmail.com

رحمت اله محمدی پور
گروه حسابداری و مالی، واحد ایلام، دانشگاه آزاد اسلامی، ایلام، ایران.
mohammadipourrahmatollah@gmail.com

رضا مشهدی زاده
گروه مدیریت، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران.
r.mashhadizadeh@zarholding.com

تاریخ دریافت: ۱۴۰۲/۰۳/۱۰ تاریخ پذیرش: ۱۴۰۲/۰۸/۲۷

چکیده

وام‌دهی نفر به نفر روش دیگری از سرمایه‌گذاری است که موجب حذف دخالت موسسات مالی سنتی شده و در سال های اخیر در دیگر کشورها رشد روزافزونی داشته است. هدف این مقاله ارزیابی ریسک اعتباری مبتنی بر نمونه با قابلیت ارزیابی ریسک و بازده هر وام با استفاده از رگرسیون لاجستیک با تابع گوسی که فاقد رویکرد رتبه بندی است در وام دهی نفر به نفر می باشد. برای تایید و اعتبار سنجی تأثیر روش پیشنهادی از پلتفرم وام دهی لندینگ کلاب استفاده شده است. که هر یک از وام گیرندگان دارای مشخصات اعتباری اند. روش اعتبار سنجی k فولد برای تعیین نمونه های اموزشی و تست استفاده شده است. از اینرو با توجه به ده هزار نمونه داده، هفت هزار نمونه وام به تصادف به عنوان نمونه وام آموزش و مابقی سه هزار نمونه وام به عنوان نمونه وام تست در نظر گرفته شده است. نتایج حاصل از آزمایش ها نشان دهنده عملکرد بهتر روش سرمایه گذاری داده محور مبتنی بر رگرسیون هسته‌ای در مقایسه با روش‌های رتبه بندی در کاربرد وام دهی نفر به نفر است.

واژه‌های کلیدی: ریسک اعتباری، روش داده محور، رگرسیون، رتبه‌بندی، وام‌دهی نفر به نفر.

۱- مقدمه

ظهور شیوه های نوین تأمین مالی، چالش های بزرگی در نحو عملکرد صنعت مالی سنتی به همراه داشته و فرصت های جدیدی را نیز برای تحول و توسعه پایدار اقتصاد واقعی فراهم کرده است (لئو و همکاران، ۲۰۲۲). نتایج بررسی های متعدد نشان دهنده آن ا ست یکی از ا صلی ترین موانع رشد و توسعه و حتی ادامه حیات بنگاه های کوچک و متوسط، مشکل تأمین مالی است که این امر، رشد و توسعه این بنگاه ها را با مشکل روبه رو کرده و باعث شده است فقط بنگاه های بزرگ در بازار باقی بماند و زمینه حذف سایر بنگاه ها فراهم شود(صافی دستجردی و همکاران، ۲۰۲۱). این امر نشان دهنده آن است که تأمین مالی، مقدم بر کارآفرینی با فناوری پیشرفته و رشد زیاد است از این رو، به نظر می رسد که تأمین مالی، ماهیت کارآفرینی را به سمت شرکت های با فناوری و رشد زیادتر تغییر می دهد(یو و فلمینگ، ۲۰۲۱). وام دهی همتا به همتا نوعی جدید از وام دهی است که به وام دهی فرد به فرد معروف بوده و به اشخاص امکان وام دهی و قرض گرفتن مستقیم از یک دیگر در پلتفرم‌های اینترنتی و بدون دخالت واسطه های مالی سنتی را می دهد. در این نوع بازار، وام گیرندگان بر اساس نیازشان اقدام به ثبت برگه هایی که به لیست یا فهرست وام می گویند، کرده و جزئیات وام درخواستی شان از جمله مقدار وام و شرایط دیگر را اعلام می کنند سپس وام دهندگان بسته به بودجه ای که در اختیار دارند می‌توانند بخشی از وام‌های درخواستی موجود در لیست را تامین کنند. اگر مقدار کل وام درخواست شده در لیست، در بازه زمانی از پیش مشخص شده ای توسط وام دهندگان پرداخت و برآورد شود، آنگاه لیست درخواست وام تبدیل به لیست وام گیرندگان می‌شود. بنابراین وام دهی فرد به فرد بخاطر حذف واسطه های مالی عرفی و ارائه بستر پویاتر وام دهی (دسترسی آسان و سریع وام دهندگان و وام گیرندگان به سرمایه گذاری) قدرت بالایی در کاهش هزینه های سرمایه گذاری و افزایش راندمان بازار مالی دارد. پلتفرم‌های وام دهی فرد به فرد، واسطه میان وام دهنده و وام گیرنده می شوند و با تسهیل ارتباط میان این دو طرف و اجرای فرآیندهای اعتبارسنجی کمک می کنند تا فردی که توانایی ارائه تسهیلات در محدوده وام‌های خرد دارد، با فردی که نیاز به وام خرد دارد به هم متصل شده و بر اساس شرایطی مشخص با هم توافق کرده و تسهیلات به فرد متقاضی وام اعطا گردد. ابرداده، بررسی گزاراها را به منظور شناسایی پلتفرم های مشکل دار وام دهی نفر به نفر برای دولت تسهیل می کند. البته یک سیستم نظارتی مبتنی بر شکایات کاربران از سوی بخش های مختلف با پیروی از شرایط شیوه های وام دهی نفر به نفر و رگوالتورهای موجود نیز استفاده می شود(سوریونو و همکاران، ۲۰۲۱). طبعا با توجه به خرد بودن مبالغ وام در این پلتفرم‌ها، مشتریان این پلتفرم‌ها، کسب وکارهای کوچک و اشخاص حقیقی هستند. پلتفرم‌های وام دهی جمعی و وام دهی اجتماعی، پلتفرم های مشابهی هستند که همگی در زمره استارت آپ های فین تک قرار می گیرند. بنابراین یکی از روش های سرمایه گذاری پر طرفدار که به‌دلیل حذف واسطه‌گری بانک‌ها علاقه افراد و فعالان اقتصادی را به خود جذب کرده است، روش وام دهی نفر به نفر می‌باشد. افزایش استفاده از این نوع وام دهی، اعتبار دهندگان را ملزم به استفاده از تکنیک هایی به منظور جلوگیری از ورشکستگی می‌کند. ارزیابی ریسک اعتباری تسهیلات گیرندگان می‌تواند کمک شایانی به کاهش ریسک ورشکستگی اعتبار دهندگان در این نوع وام

دهی کند(کلافت[1]، ۲۰۰۸). تأمین مالی با چالش اعتماد رو به رو است. فقدان اعتماد، مانع از بقای آن به عنوان جایگزین تأمین مالی برای کارآفرینان و سرمایهگ ذاران خرد بالقوه خواهد شد بنابراین بهتر ا ست پلتفرم های جمع سپاری تأمین مالی برای ایجاد اعتماد در ۴ سطح مختلف کار کند: اعتماد به پلتفرم، کاربران، پروژه و تأمین مالی. برخی پلتفرم ها به ویژه پلتفرم های وام دهی نفر به نفر فقط به این موضوع توجه می کند که چگونه برای سرمایه گذاران خرد و نه برای کارآفرینان، معتمد به نظر برسند(فریرا، ۲۰۲۲).

این نوع وام دهی، با ارائه کانال سرمایه گذاری جدید، بودجه جایگزین و مهمی را فراهم می کند. کانال سرمایه گذاری رفاه وام گیرندگان و وام دهندگان را بهبود می بخشد(جیانگ، ۲۰۲۱).

یکی از چالش های مهم در بحث سپرده گذاری و وام دهی در بانک داری نفر به نفر، عدم بازپرداخت وام می‌باشد که می‌تواند منجر به ورشکستگی اعتبار دهندگان گردد بطوریکه در برخی موارد حتی اگر درصد کمی از وام‌های اعطایی قابل وصول نباشد، اعتبار دهنده با خطر ورشکستگی روبه رو خواهد شد. در این راستا پلتفرم‌های معروف وام دهی نفر به نفر از جمله لندینگ کلاب به منظور کاهش ریسک مواجه با مشکل عدم بازپرداخت وام‌های اعطایی و به عبارتی دیگر پیش‌بینی ریسک اعتباری تسهیلات گیرندگان، شاخص‌هایی را پیشنهاد داده اند. شاخص‌های پیشنهادی پلتفرم لندینگ کلاب که به شاخص‌های ارزیابی ریسک معروف هستند برای تصمیم گیری در جهت اعطای هر وام[2] بکار می روند. به عبارتی دیگر برای ارزیابی ارزشمندی اعتباری هر وام می توان به شاخص‌های ارزیابی اعتباری که لندینگ کلاب در اختیار گذارده است، استفاده نمود. از جمله این شاخص‌ها، میزان وام درخواستی، امتیاز فایکو[3] (رتبه ای که اعتبار اسناد اتکایی[4] به متقاضیان وام می‌دهد از A1 تا G5)، نسبت بدهی به درآمد، بدهکاری های وام گیرنده در ۲ سال اخیر، میزان بدهکاری فعلی وام گیرنده، وضعیت مالکیت منزل وام گیرنده نرخ بهره، مبلغ سرمایه گذاری شده، درآمد سالانه، تعداد کارت های اعتباری فعال، میزان بدهی خطوط اعتباری، میزان استفاده از خطوط اعتباری، تعداد حساب های جاری فعال، تعداد حساب های فعال بدهی و.. می‌باشد. بنابراین بکارگیری روشهای آماری مفیدی که بتواند به تحلیل و بررسی روابط میان شاخص‌های اعتباری متقاضیان وام به منظور تصمیم گیری ارزش اعتباری متقاضی بپردازد، امری مهم بوده و نیاز به مطالعه و بررسی عمیقی دارد(امکتر، ۲۰۱۴).

از آنجایی‌که هدف این مقاله بررسی دو روش ارزیابی ریسک اعتباری مبتنی بر رگرسیون و بررسی مسئله تصمیمات بهینه سرمایه گذاری در جهت کمینه سازی ریسک می باشد، در این مقاله مروری بر تعاریف و مفاهیم مقدماتی از جمله ریسک، وام دهی نفر به نفر، رگرسیون خطی و لاجستیک خواهیم داشت و در انتها نیز به بررسی مسائل مطرح در زمینه روشهای آماری از جمله رگرسیون و کاربردشان در خصوص شناسایی ریسک اعتباری و مرور برخی از تحقیقات انجام شده در راستای ارزیابی ریسک اعتباری به صورت جزئی‌تری پرداخته شد. با وجود پیشرفت های بسیار زیاد تکنولوژی و وجود داده های فراوان اگر این داده ها به درستی مورد پردازش قرار نگیرند

[1] Klafft, 2008
[2] Loan
[3] FICO
[4] Letter of Credit

می‌تواند مشکلات زیادی را ایجاد کنند. همچنین از آنجایی که بکارگیری ضرایب وزنی بهینه در روش رگرسیون منجر به کاهش زمان اجرای روش و افزایش دقت پیش‌بینی می‌گردد، نیاز به روش‌های بهینه‌تر و موثرتری در جهت استخراج وزن های مفید در فرآیند رگرسیون می‌باشد (توماس و همکاران[1]، ۲۰۱۷). در این مقاله به ترتیب به بررسی روشهای موثر از جمله رگرسیون هسته ای و روشهای مبنی بر رتبه بندی با پایه رگرسیون در این زمینه پرداخته می‌شود.

مبانی نظری و پیشینه پژوهش

وام دهی فرد به فرد پدیده تجارت الکترونیکی جدید در حوزه مالی با مزیت هایی از جمله سرمایه گذاری مقرون به صرفه است که متکی بر سیستم اطلاعات مبتنی بر اینترنت می‌باشد. در مقاله های مختلف از وام دهی فرد به فرد به عنوان وام دهی آنلاین اجتماعی، تامین اعتبار خرد (ارائه اعتبار مالی به اقشار ضعیف)، تامین مالی انبوه یا جذب سرمایه مردمی، یاد شده است بطوریکه هر یک از آنها تاکید بر مشخصه‌های منحصر به فرد این نوع وام دهی دارند (دیو و همکاران[2]، ۲۰۲۰). یعنی تعاملات اجتماعی آنلاین میان شرکت کنندگان، معاملات با مبلغ اندک و تعداد بسیاری سرمایه گذار که ممکن است در سرمایه گذاری وام متخصص نباشند. شرکت کنندگان در وام دهی فرد به فرد را تقریبا می توان به دو گروه دسته بندی کرد : وام گیرندگان و وام دهندگان. همانند موسسات اعتبار دهنده سنتی در وام دهی فرد به فرد نیز هدف از تصمیم گیری و ارزیابی ریسک از منظر وام دهنده و وام گیرنده متفاوت می‌باشد. درک فرایند تصمیم گیری در وام دهی نفر به نفر، مستلزم درک نحو انتقال اطلاعات بین بنیانگذاران و سرمایه گذاران است به ویژه زمانی که چالشه ای اجتماعی با پروژه هایی که درک آنها ازنظر فنی آسان نیست، مورد توجه قرار می گیرد. در اینجاست که نبودِ تقارن اطلاعات بین بنیانگذار و سرمایه گذار حائز اهمیت می شود(پابست و همکاران، ۲۰۲۱).

بنابراین وام دهندگان یا سرمایه گذاران قبل از وام دهی باید امکان سننجی وام انجام دهند. وام دهنده با تجزیه وتحلیل اعتبار، به این اطمینان می رسند که آیا مشنتری با توجه به قرارداد، توانایی انجام تعهدات و پرداخت منظم اصل و سود را دارد یا خیر (فارادهارمستوتی و لائورنتکسیوس، ۲۰۲۱).

سرایت ریسک اعتباری میان پلتفرم های وام دهی نفر به نفر، پنج عامل مختلف دارد: همبستگی بین پلتفرم های وام دهی نفر به نفر، میزان ایمن سازی پلتفرم های حساس، نرخ حذف نهایی پلتفرم های وام دهی نفر به نفر از سوی رگولاتورها، میزان حذف پلتفرم های پرریسک و نرخ تبدیل پلتفرم های پرریسک به پلتفرم های ایمن. این عوامل، بر سرایت ریسک اعتباری پلتفرم های وام دهی نفر به نفر اثر می گذارد. بنابراین اداره های دولتی باید اقدامات نظارتی خود را برای بهبود مدیریت ریسک اعتباری، مقاومت در برابر ریسک و تحمل ریسک این پلتفرم ها تقویت کند(ژائو و همکاران، ۲۰۲۱)

[1] Thomas, et al.2017
[2] Du, et al. 2020

موجاهیدین و همکاران (۲۰۲۲) نشان داده اند که اعتماد، تأثیر چشم گیری بر رفتار وام گیرندگان دارد. اعتماد، عاملی اساسی در شکل دادن به درک و رفتار کاربر با استفاده از این فناوری است. این متغیر بر محیطی اثر می گذارد که پلتفرم وام دهی در آن وجود دارد و همه آنها سیاست های یکسانی ندارد درنتیجه بسنیاری از کاربران در انتخاب پلتفرم مطمئن سردرگم می شوند (موجاهدین و همکاران، ۲۰۲۲).

وانگ و همکاران (۲۰۱۹) در تحقیق خود با عنوان " اطلاعات نرم در وام دهی همتا به همتا آنلاین: شواهد از یک پلتفرم پیشرو در چین. تحقیقات و کاربردهای تجارت الکترونیک" به مواردی مانند تقسیم بندی سطوح اعتبار، تعیین نماینده وصول معوقات، اشتراک اطلاعات با سایر مؤسسات مالی، مراحل حراج، بررسی و بازپرداخت پس از صدور فهرست وام رسیدند (وانگ و همکاران، ۲۰۱۹)

سانگ و همکاران (۲۰۱۸) نیز در تحقیق خود با موضوع " تجزیه و تحلیل عملکرد سیستم عامل های وام دهی آنلاین همتا به همتا در چین " به این اشاره کردند که مرحله گسترش بازار که هدف آن بیشینه سازی سود وام گیرنده و مرحله مدیریت ریسک که هدفش حصول اطمینان از بازپرداخت به موقع پول وام دهندگان است(سانگ و همکاران، ۲۰۱۸).

هاروی (۲۰۱۸) نیز در کتاب خود تحت عنوان " راهنمای سرمایه گذاری همتا به همتا برای مبتدیان "ریسک های وام دهی را به چهار دسته ریسک عملکرد، پلتفرم، بازار و نقدینگی تقسیم می کند(هاروی ۲۰۱۸).

محقق گوئو و همکاران در سال ۲۰۱۶ روشهای مناقصه برای مزایده های اندک را مورد بررسی و تحلیل قرار دادند. مطالعه آنها نشان داد که رفتار مناقصه در بین سرمایه گذاران مشابه بوده و سرمایه گذاران از استراتژی های مناقصه بسیار متفاوتی برای دنبال کردن رفتار دیگری و به فروش رساندن معاملات خود استفاده می کنند و در نتیجه مشخص است که با کاهش هزینه معاملات، به پایین ترین سطح خود موجب بالا رفتن مزایده ها و انجام حجم زیادی از معاملات در زمان محدود می‌شود و این امر نوسانات قیمتی را نیز تشدید نموده و به نوعی پدیده رفتارهای گله ای[1] (یعنی زمانیکه هر کس رفتارهای دیگری را دنبال می‌کند) شکل میگیرد، بنابراین پدیده رفتار گله ای بیشتر در مزایده های وام اجتماعی آنلاین دیده می‌شود و رخ دادن این پدیده نیز موجب عدم تعادل در سیستم معامله می گردد(گوئو و همکاران، ۲۰۱۶).

محقق وو و همکاران در سال ۲۰۱۱ روش‌های مناقصه برای مزایده های اندک را مورد بررسی و تحلیل قرار دادند. مطالعه آن‌ها نشان داد که رفتار مناقصه در بین سرمایه گذاران مشابه بوده و سرمایه گذاران از استراتژی های مناقصه بسیار متفاوتی برای دنبال کردن رفتار دیگری و به فروش رساندن معاملات خود استفاده می کنند و در نتیجه مشخص است که با کاهش هزینه معاملات، به پایین ترین سطح خود موجب بالا رفتن مزایده ها و انجام حجم زیادی از معاملات در زمان محدود می‌شود و این امر نوسانات قیمتی را نیز تشدید نموده و به نوعی پدیده رفتارهای گله‌ای (یعنی زمانیکه هر کس رفتارهای دیگری را دنبال می‌کند) شکل می‌گیرد، بنابراین پدیده رفتار گله ای بیشتر در مزایده های وام اجتماعی آنلاین دیده می‌شود و رخ دادن این پدیده نیز موجب عدم تعادل در سیستم معامله می‌گردد (وو و همکاران، ۲۰۱۱).

[1] Herding behavior

محقق پورو در سال ۲۰۱۰ به طراحی سیستم کمک یار وام گیرنده پرداختند که سیستم پیشنهادی آنها به منظور یاری رساندن به وام گیرندگان در جهت تعیین میزان و نوع استراتژی های مهم از جمله نرخ بهره، مبلغ وام درخواستی به نوعی که تاثیر موفقی در موافقت با درخواست وامشان داشته باشد را ارائه نمودند. روش پیشنهادی آنها تاثیر بسزایی در موافقت با وام درخواستی وام گیرندگان در لیست درخواست وام داشت (پورو و همکاران[1]، ۲۰۱۰).

برخی محققان نیز از منظر وام دهندگان به بررسی تصمیم گیری و ارزیابی ریسک اعتباری و رفتار مناقصه پرداخته اند، به عنوان مثال محقق سان شین و همکاران در سال ۲۰۱۰ نقش حسابهای اجتماعی در تأثیرگذاری بر تصمیمات وام دهندگان را مورد بررسی قرار دادند. کلافت و همکاران در سال ۲۰۰۸ به بررسی میزان سودجویی و منفعت وام دهندگان در وام دهی فرد به فرد پرداختند. براساس آزمایشاتی که بر روی داده های پلتفرم پراسپر انجام دادند، آنها به این نکته مهم پی بردند که موفقیت وام دهندگان و سطح رضایت آنها از بازپرداخت وام‌ها بستگی به استفاده از استراتژی صحیح و دقیق دارد. محقق کروم در سال ۲۰۰۹ دریافتند که بسیاری از معاملات براساس تصمیمات زیر بهینه انجام می‌شود (کلافت ، ۲۰۰۸).

مطالعات دیگر روش های ارزشمندی در مورد نحوه نمایش وام برای وام دهنده و نحوه انتخاب وام‌ها توسط آن‌ها و موفقیت در مناقصه ارائه می دهند اما مجهز به پشتیبانی از تصمیم گیری مؤثر برای راهنمایی در جهت انتخاب سرمایه گذارهای شخصی اعتبارات و تعیین مبلغ مناسب برای پیش‌بینی در هریک از آنها وجود ندارد. از آنجایی‌که تصمیم گیری در مورد تأمین یا عدم تأمین اعتبار وام‌های خاص یکی از وظایف اصلی سرمایه گذاران در بازار وام دهی فرد به فرد است، می‌توان ادعا کرد که روشهای ارزیابی وام‌های سنتی نیز به وام‌های فرد به فرد مرتبط می‌شود. در برخی از تحقیقات به منظور تمایز وام خوب (وام‌هایی که انتظار می رود با موفقیت بازپرداخت شوند) از موارد بد (وام‌هایی که پیش‌بینی می‌شود با تاخیر بازپرداخت شوند)، اخیرا از تکنیک های مختلف داده کاوی استفاده شده است که در ادامه به هر یک از آن ها اشاره خواهیم کرد: پائولو و همکاران روش رگرسیون لجستیک (پائلو و همکاران، ۲۰۱۹) هاباچی و بِنباچِر از روش تحلیل ممیزی خطی، ژانگ از روش نزدیک ترین kهمسایگی، آبهیجیت گوها و سیدهالدینگ از روش درخت تصمیم گیری، لیندا و همکاران از زنجیره های مارکوف، دی جوندجی و همکاران از روش تحلیل بقا، رابون و همکاران از روش برنامه ریزی خطی و غیر خطی، وِلَک اِسوِن از روش شبکه های عصبی، تائو از روش ماشین های بردار پشتیان، شِن و همکاران از روش الگوریتم بهینه سازی ژنتیک بهره گرفته اند. مطالعات مذکور طبقه بندی هر وام را مورد بررسی قرار می‌دهند اما به طور کلی بهینه سازی سرمایه گذاری وام را به دقت مورد بررسی قرار نمی‌دهند. در سال های اخیر روشهای مبتنی بر هسته بسیاری در کابردهای مختلف از جمله امتیازدهی اعتباری و مسائل بهینه سازی مورد استفاده قرار گرفته اند(لئونارد و همکاران، ۲۰۲۰). در این مقاله از ضرایب رگرسیون به عنوان وزن های بهینه برای ارزیابی ریسک اعتباری استفاده نموده ایم. روش رگرسیون هسته ای یکی از تکنیک های انعطاف پذیر آماری برای بررسی روابط غیر خطی می‌باشد که در کاربردهای مالی و اقتصادی همانند اندازه گیری های غیرپارامتری، روابط غیر خطی بین نرخ بازده

[1] Puro, et al. 2010

بین المللی واقعی و فرآیندهای انتشار متغیر با زمان برای پیش‌بینی خسارات مالی رایج شده است. بنابراین روش رگرسیون هسته ای و روش‌های رتبه بندی در یاری رساندن به سرمایه گذاران شخصی در وام دهی فرد به فرد برای تصمیم گیری مورد بررسی قرار گرفته است.

وام‌دهی نفر به نفر

وام دهی همتا به همتا نوعی جدید از وام دهی است که به وام دهی نفر به نفر معروف بوده و به اشخاص امکان وام دهی و قرض گرفتن مستقیم از یک دیگر در پلتفرم اینترنتی و بدون دخالت واسطه های مالی سنتی را می دهد. در این نوع بازار، وام گیرندگان براساس نیازشان اقدام به ثبت برگه هایی که به لیست یا فهرست وام معروف است، کرده و جزئیات وام درخواستی شان از جمله مقدار وام و شرایط دیگر را اعلام میکنند سپس وام دهندگان بسته به بودجه ای که در اختیار دارند میتوانند بخشی از وام های درخواستی موجود در لیست را تامین میکنند. از جمله پلتفرم های وام دهی نفر به نفر می‌توان به لندنیگ کلاب اشاره نمود (لاریمور و همکاران، ۲۰۱۱).

برخلاف موسسات مالی سنتی، در وام دهی فرد به فرد، هر سرمایه گذار شخصی، اعتبار بسیار کمتری برای سرمایه گذاری دارد اما از آنجا که امکان تامین بخشی از بودجه وام وجود دارد یا به عبارت دیگر هر سرمایه گذار می‌تواند بخشی از هر وام را تامین کند، پویا سازی به شکل مؤثری در این نوع وام دهی امکان پذیر می‌شود. بنابراین سرمایه گذاران در بازار وام دهی فرد به فرد نه تنها نیاز به تصمیم گیری برای تامین اعتبار وام دارند، (یعنی اینکه بودجه خود را صرف تامین بودجه کدام متقاضی وام کنند)، بلکه باید در مورد میزان مبلغی که بهتر است به هر وام درخواستی اختصاص دهند نیز تصمیم گیری کنند (یعنی چه مبلغی به هر متقاضی وام اختصاص دهند) تا نرخ بازده در بازه مورد انتظار کمینه گردد. بنابراین اگرچه این مسئله را می‌توان به عنوان یک مسئله بهینه سازی سبد سرمایه گذاری معمولی ارائه و حل نمود اما ارزیابی دقیق ریسک اعتباری هر وام جداگانه در پلتفرم وام دهی فرد به فرد که از جمله ورودی لازم و مهم در مسئله بهینه سازی پرتفوی می‌باشد، خود مسئله ای بسیار چالش برانگیز و مورد بحث است. بعلاوه در اکثر بازارهای وام دهی فرد به فرد واقعی، محدودیاتی از جمله حداقل مبلغ سرمایه گذاری وجود دارد که باید مورد توجه قرار گیرد. چنین محدودیتهایی برای سرمایه گذاران شخصی که بودجه کمی دارند، مهم و تاثیر گذار است. بنابراین راه حل وتصمیم گیری نوینی که موجب سودآور و پر منفعت شدن فرآیند وام دهی فرد به فرد گردد، لازم است. در وام دهی نفر به نفر هر فردی قادر است براساس ریسک سرمایه گذاری و نرخ بازده برای پرداخت یا دریافت وام اقدام کند. آمارها نشان دهنده پرداختی وام موسساتی نظیر لندینگ کلاب[1] بزرگ ترین موسسه وام دهنده نفر به نفر در کشور ایالات متحده است که در حال حاضر از مرز سه میلیارد دلار نیز گذشته است که با توجه به مبلغ پانصد میلیون دلار در سال ۲۰۱۲ رشد فوق العاده ای را تجربه نموده است. این رقم در انگلستان نزدیک به هشتصد میلیون پوند است که توسط شرکت هایی نظیر فاندینگ سیرکل[2] پرداخت شده است. نکته قابل ذکر نیز این است که نرخ نکول این وام‌ها بسیار کمتر از وام‌های پرداختی

[1] Lending club (LC)
[2] Funding circle

توسط بانک‌ها می‌باشد. از مزیت های این نوع بانکداری نرخ بازده سرمایه بالاتر برای سپرده گذار، نرخ بهره پایین برای وام گیرنده، نوسان کمتر نرخ بهره نسبت به سایر روشها و امکان نرخ بهره[1] متناسب با ریسک می‌باشد. این پلتفرم‌ها حجم زیادی از نقدینگی و نرخ پایینی از نکول را حتی در مقایسه با کارتهای اعتباری به خود تعلق داده اند. در اکثر کشورها مقررات نظارتی بر این پلتفرم‌ها به تدریج و به آهستگی در حال شکل گیری است(گوئو و همکاران، ۲۰۱۶).

روش پژوهش

تحلیل ریسک اعتباری در فرایند اعطای وام به متقاضیان اعتبار روز به روز از اهمیت بیشتری برخوردار می‌شود. یکی از مهمترین جنبه ها در تحلیل ریسک، تشخیص امتیاز هر متقاضی وام در جهت تعیین متقاضیان خوب از بد از لحاظ بازپرداخت وام می‌باشد. شناسایی افراد بدحساب در میان متقاضیان وام نقش کلیدی در اعطای وام با ریسک کم در حوزه وام دهی دارد. یکی از روشهای رایج ارزیابی ریسک اعتباری، روش رتبه بندی ساده[2] و رتبه بندی توسعه یافته[3]می باشد. روش رتبه بندی ساده برخلاف روش رگرسیون هسته ای وزن دار شده فاقد ماتریس فاصله می باشد و صرفا از احتمال نکول هر یک از وام های آموزش و تست به شیوه اختصاص رتبه به هر وام تست براساس شاخص اعتباری درجه بهره می برند. روش رتبه بندی توسعه یافته نیز برخلاف روش رگرسیون هسته ای وزن دار شده فاقد ماتریس فاصله می باشد و صرفا از احتمال نکول هر یک از وام های آموزش و تست به شیوه اختصاص رتبه به هر وام تست براساس شاخص اعتباری زیردرجه بهره می برد. روش رتبه بندی توسعه یافته، نسخه توسعه یافته روش رتبه بندی ساده می باشد و تفاوتش با روش رتبه بندی ساده در استفاده از شاخص اعتباری زیردرجه خلاصه می شود بطوریکه در رتبه بندی توسعه یافته، هر درجه به زیر درجه هایی تفکیک شده و براساس زیردرجه های مستخرج به محاسبه احتمال نکول و سپس نرخ بازده و ریسک برای هر وام تست پرداخته می شود. ساده ترین روشهای ارزیابی ریسک اعتباری که به دلیل سادگی در فهم و پیاده سازیش محبوبیت بسیاری یافته است، روشهای مبتنی برنمونه با بکارگیری رگرسیون لجستیک می‌باشد. اخیرا در جدیدترین تحقیقات، تکنیکی با استفاده از وزن دهی به نمونه های داده برای بهبود روش رگرسیون لجستیک پیشنهاد شده است که از شاخص‌های اعتباری کمی برای ارزیابی ریسک اعتباری استفاده می‌کند. افزایش شاخصه های اعتباری به عنوان متغیرهای مستقل در رگرسیون وزن دار شده می‌تواند منجر به سنجش دقیقتر ریسک اعتباری گردد، از اینرو در این مقاله از مشخصه های اعتباری بیشتری به عنوان متغیر مستقل و عوامل پیش‌بینی کننده ریسک در روش رگرسیون وزن دار شده استفاده می‌شود بطوریکه مشخصه های اعتباری در نظر گرفته شده برای متقاضی وام عبارت اند از: میزان وام درخواستی، امتیاز فایکو، بدهکاریهای دو سال اخیر، نسبت بدهی به درآمد، نرخ بهره، وضعیت مالکیت مسکن، میزان مانده تمامی حسابها، درآمد سالانه و بدهی در ۶ ماه اخیر. در این بخش به ارائه

[1]. Interest Rate
[2] Rating-based Method
[3] Refined Rating-based Method

روشی برای ارزیابی ریسک اعتباری و سپس رتبه بندی اعتباری متقاضیان وام با اسنفاده از روش رگرسیون هسته ای وزن دار شده و روش رتبه بندی ساده پرداخته می شود.

رگرسیون هسته‌ای گوسی

در این بخش ابتدا مدل مبتنی بر نمونه ارزیابی ریسک اعتباری در جهت تعیین ارزیابی کمی نرخ بازده و ریسک سرمایه گذاری بر روی وام‌ها شرح داده شده سپس با بکارگیری رگرسیون هسته ای، فواصل احتمال نکول هر وام به وزنهای بهینه ای تغییر داده می‌شود و در ادامه آن به بررسی روشهای رتبه بندی بر پایه مدل رگرسیون برای تخمین ریسک سرمایه گذاری و نرخ بازده پرداخته می شود (گوئو و همکاران، ۲۰۱۶).

مدل مبتنی بر نمونه

پیش‌بینی نمونه داده های آینده بر مبنای نمونه داده های گذشته (سوابق داده ها) روشی رایج و معمول است. به عبارتی دیگر یک روش مفید برای ارزیابی ریسک اعتباری وام‌های درخواستی جدید، بکارگیری مشخصات اعتباری دریافت کنندگان وام در گذشته است، برای بهینه سازی پرتفوی سرمایه می‌توان از پروسه پیش‌بینی نرخ بازده و ریسک سرمایه گذاری بر روی هر وام بهره گرفت. بدین صورت که برای پیش‌بینی نرخ بازده وام درخواستی i (یعنی نمونه تست (نمونه جدید یا وام جدید)) می‌توان براساس مشخصات اعتباری دریافت کنندگان تسهیلات اعطایی در گذشته (یعنی اعضای رای دهنده در روش رگرسیون) و نرخ بازده مشاهده شده در وام های اعطایی در گذشته $R_j(j = 1,2,..,n)$ به پیش‌بینی نرخ بازده وام های جدید μ_i پرداخت. برای این منظور می‌توان از میانگین وزن دار شده عملکرد تسهیلات اعطایی در گذشته با استفاده از معادله (۱.۴) استفاده نمود(گوئو و همکاران، ۲۰۱۶).

$$\mu_i = \sum_{j=1}^{n} w_{ij} R_j \,. \qquad (1.4)$$

به‌طوری‌که w_{ij} وزن وام گذشته (آموزشی) j ام برای پیش‌بینی نرخ بازده وام جدید (تست) i می‌باشد. این روش مشابه با روش رای گیری است که درآن، هر رای با وزنهای مختلف محاسبه می‌شود. برای پیش‌بینی ریسک سرمایه گذاری نمی‌توان مستقیما از میانگین وزن دار شده استفاده نمود و برای ارزیابی ریسک سرمایه گذاری بر روی وام درخواستی جدید (i)، از واریانس وزن دار شده یعنی σ_i^2 در میان اعضای رای دهنده مطابق با (۲.۴) استفاده می شود.

$$\sigma_i^2 = \sum_{j=1}^{n} w_{ij}\left(R_j - \mu_i\right)^2 . \quad (2.4)$$

به‌طوری‌که μ_i میانگین وزن دار شده در (۱.۴) و w_{ij} وزن وام j ام (وزن وام‌های با مشخصات اعتباری مشابه در گذشته) می‌باشد. وزن در مدلهای مبتنی بر نمونه براساس مجاورت وام جدید (نمونه تست) با اعضای رای دهنده

(وام‌های اعطایی در گذشته) تعیین می‌شود. از این اصل واضح است که نمونه های مشابه تر (نزدیک تر) دارای وزن بیشتر درحالیکه نمونه های کم شباهت تر، وزن کمتری را دارا هستند. در این خصوص، فاصله احتمال نکول بین وام قدیم و جدید (آموزشی و تست) i و j به صورت زیر تعریف می‌شود :

$$d_{ij} = |p_i - p_j|. \quad (3.4)$$

به‌طوری‌که p_i و p_j به ترتیب احتمال نکول وام i و j می‌باشد. فاصله بیشتر میان احتمالهای نکول موجب وزن کمتر (w_{ij}) برای وام درخواستی جدید خواهد شد.

وزن هسته‌ای

یکی از روش‌های آماری برای یافتن روابط غیر خطی میان جفتی از متغیرهای تصادفی، رگرسیون هسته ای می‌باشد. اگر هر مشاهده (وام یا نمونه) در بعد X و Y بررسی گردد، بطوریکه X متغیر پیش‌بینی و Y متغیر پاسخ یا وابسته باشد و n نمونه $\{(xj, yj)| j = 1, 2,..., n\}$ موجود باشد آنگاه می‌توان نتیجه y را با استفاده از مشاهدات پیش‌بینی کننده x براساس معادله رگرسیون هسته ای تخمین زد:

$$y = \hat{f}(x) = \frac{\sum_{j=1}^{n} \mathrm{K}\left(\frac{x - x_j}{h}\right) y_j}{\sum_{j=1}^{n} \mathrm{K}\left(\frac{x - x_j}{h}\right)}. \quad (4.4)$$

به‌طوری‌که $\mathrm{K}(.)$ تابع هسته از نوع گوسی می‌باشد. معادله هسته گوسی در (۵.۴) آمده است:

$$\mathrm{K}(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (5.4)$$

در روش هسته گوسی به مشاهدات (وام‌های) نزدیک تر به وام x، وزنهای بیشتر و به نمونه های دورتر از نمونه وام x، وزنهای کمتری اختصاص می یابد. پارامتر h>0 پهنای باند نامیده می‌شود که نسبت اطلاعات محلی دربرابر اطلاعات جزئی مورد استفاده در مجموع را تعیین می‌کند.

در روش مدل‌سازی ریسک مبتنی بر نمونه، احتمال نکول وام جدید i، p_i و احتمال نکول وام های اعطایی در گذشته j، p_j به عنوان متغیر پیش بینی در نظر گرفته می‌شود. متغیر پاسخ، نرخ بازده در وام های آموزشی (R_j) است. بنابراین با جایگذاری معادله تابع گوسی و متغیر پیش‌بینی و پاسخ در معادله۴ می‌توان معادله تخمین نرخ بازده وام در (۵.۴) را به صورت زیر بازنویسی کرد :

$$\mu_i = \frac{\sum_{j=1}^{n} \mathrm{K}\left(\frac{x_i - x_j}{h}\right) y_j}{\sum_{j=1}^{n} \mathrm{K}\left(\frac{x_i - x_j}{h}\right)} = \frac{\sum_{j=1}^{n} \mathrm{K}\left(\frac{p_i - p_j}{h}\right) R_j}{\sum_{j=1}^{n} \mathrm{K}\left(\frac{p_i - p_j}{h}\right)}. \quad (6.4)$$

به‌طوری‌که μ_i، R_j، p_i و p_j به ترتیب بیانگر میانگین وزن دار شده نرخ بازده در نمونه وام های مشاهده شده (آموزشی)، نرخ بازده وام های آموزشی، احتمال نکول وام جدید i و احتمال نکول وام اعطایی در گذشته j می‌باشد.

با مقایسه نرخ بازده پیش‌بینی شده در (۱.۴) با معادله رگرسیون هسته ای در (۶.۴) می‌توان مستقیما وزن بهینه را از (۷.۴) استخراج کرد :

$$w_{ij} = \frac{\mathrm{K}\left(\frac{p_i - p_j}{h}\right)}{\sum_{j=1}^{n} \mathrm{K}\left(\frac{p_i - p_j}{h}\right)} = \frac{\mathrm{K}\left(\frac{d_{ij}}{h}\right)}{\sum_{j=1}^{n} \mathrm{K}\left(\frac{d_{ij}}{h}\right)} \quad . \qquad (7.4)$$

به‌طوری‌که d_{ij} فاصله احتمال نکول بین وام i و j می‌باشد که در معادله (۳.۴) تعریف شد. به عبارت دیگر وزن وام (w_{ij}) براساس فاصله احتمال نکول (عدم باز پرداخت وام) وام های گذشته از وام های جدید یعنی فاصله d_{ij} تعیین می گردد و وام با فاصله احتمال نکول کمتر دارای وزن بیشتری خواهد بود.

فواصل و وزن هسته ای

یک روش مستقیم و آسان برای تعیین شباهت بین وام‌های مختلف، محاسبه فواصل احتمال نکول آن‌ها می‌باشد که در معادله ۳ تعریف شد. با محاسبه فواصل احتمال نکول میان وام‌ها برای نمونه وام تست (یعنی وام درخواستی که میخواهیم احتمال نکولش را براساس مشخصات دریافت کنندگان وام‌های اعطایی در گذشته یعنی وام l1 و l2 محاسبه کنیم) داریم : $d_{13} = |p_1 - p_3|$ و $d_{23} = |p_2 - p_3|$ بطوریکه d_{13} فاصله بین احتمال نکول وام آموزشی ۱ تا وام تست ۳ و d_{23} فاصله بین احتمال نکول وام آموزشی ۲ تا وام تست ۳ می‌باشد. فاصله احتمال نکول میان وام ۱ (وام آموزشی) و وام ۳ (وام تست) بیشتر از فاصله احتمال نکول میان وام ۲ (آموزشی) و وام ۳ (وام تست) می‌باشد یعنی $d_{13} > d_{23}$. بنابراین وزن رای دهنده وام آموزشی L1 کمتر از وام آموزشی L2 می‌باشد. می‌توان از روش رگرسیون هسته ای در (۷.۳)، میزان اهمیت وزنهای رای دهنده را سنجید، از اینرو با استفاده از (۷.۴)، داریم: $w_{13} < w_{23}$. هسته گوسی در معادله ۵.۳، محدوده مقادیر خروجی بیشتری را برای ورودی های با مقادیر کمتر فراهم می‌کند (گیو، ۲۰۱۶).

مدل تصمیم گیری سرمایه گذاری

در این بخش تصمیمات سرمایه گذاری در وام دهی نفر به نفر به عنوان مسئله انتخاب پرتفوی مجهز به برخی قیود در نظر گرفته می‌شود. سپس مدل ارزیابی ریسک اعتباری مبتنی بر نمونه در بخش قبل با روش انتخاب پرتفوی سرمایه ترکیب می‌شود و فرآیند تصمیم گیری برای سرمایه گذاری انجام می‌گیرد. هدف از فرآیند تصمیم گیری سرمایه گذاری، تخصیص بهینه اعتبار به فهرست وام‌های درخواستی می‌باشد.

مسئله پرتفوی سرمایه برای کمینه کردن ریسک سرمایه گذاری (یا تخصیص اعتبار) به صورت زیر مدل می‌شود:

$$\sum_{i=1}^{l} \lambda_i^2 \sigma_i^2. \quad (8.4)$$

به شرطی که $\lambda_i \geq 0$ و $\sum_i \lambda_i = 1$ برقرار باشد و سطح بازده مورد انتظار معادل با $R^* = \sum_i \mu_i \lambda_i$ باشد. متغیر بهینه سازی λ_i نشان دهنده نسبت بهینه اعتبار سرمایه گذاری شده بر روی وام i ام می‌باشد به عبارتی دیگر نسبت بهینه اعتبار تخصیص یافته به وام iام است.

فرآیند ارزیابی ریسک اعتباری و تصمیم گیری برای سرمایه گذاری بر روی وام ها در شکل۲ نشان داده شده است.

متغیرهای ورودی برنامه

DataH: مجموعه داده آموزشی: n مجموعه وام که در گذشته تخصیص اعتبار داده شده است

Datal: مجموعه داده تست، l مجموعه وام که در لیست تخصیص اعتبار قرار دارد و قرار است به آن اعتبار تخصیص یابد.

M: مینیمم سرمایه مورد نیاز برای تخصیص اعتبار (سرمایه گذاری) به یک وام

R^* : نرخ کلی بازده مورد انتظار

متغیر خروجی برنامه:

OM_i : درصد اعتبار تخصیص یافته به هر وام درخواستی i در نمونه داده تست (i=1,2,...,l)

// مقدار دهی اولیه و مرحله آموزش

پیش پردازش $\left(DataH و Datal\right) \rightarrow \left(n, \{R_j\}_{j=1}^{n}, l, \{e_i\}_{i=1}^{l}\right)$

رگرسیون لاجستیک $\rightarrow \left(\{\hat{\beta}_k\}_{k=0}^{p}, l, \{\hat{p}_j\}_{j=1}^{n}\right)$

تنظیم پهنای باند $(DataH) \rightarrow h$

// شروع روش ارزیابی ریسک اعتباری مبتنی بر نمونه

For $i \leftarrow 1\ to\ l\ do$

$\hat{p}_i \leftarrow \frac{1}{1+e^{\hat{\beta}_0+\hat{\beta}_1 x_{1i}+\cdots+\hat{\beta}_p x_{pi}}}$

For $j \leftarrow 1\ to\ n\ do$

$d_{ij} \leftarrow |\hat{p}_i - \hat{p}_j|;$

$w_{ij} \leftarrow kernelweight(d_{ij}, h);$

End

$\mu_i \leftarrow \sum_{j=1}^{n} w_{ij} R_j;$

$\sigma_i^2 \leftarrow \sum_{j=1}^{n} w_{ij}(R_j - \mu_i)^2;$

End

// بهینه سازی پرتفوی

$$Minimize \quad \sum_{i=1}^{l} \lambda_i^2 \sigma_i^2;$$

$$Subject\ to$$

$$\sum_{i=1}^{l} \lambda_i \mu_i = R^*$$

$$\sum_{i=1}^{l} \lambda_i = 1$$

$$\begin{cases} m \leq \lambda_i M \leq e_i & if\ \ loan\ \ i\ \ is\ invested \\ \lambda_i = 0 & otherwise \end{cases}$$

$$\{\lambda_i\} \leftarrow Prog\left(m, M, R^*, \{\mu_i, \sigma_i^2, e_i\}, \quad i = 1, \ldots l\right)$$

Output $OM_i = \lambda_i M$

در مرحله اولیه و فرآیند آموزش روش (خط۱ − ۳)، ابتدا تعداد کل وام‌های گذشته در مجموعه آموزشی $DataH(n)$، عملکرد یا نرخ بازده مورد انتظار هر وام در مجموعه داده آموزشی (R_j)، تعداد وام‌های قرار گرفته در لیست درخواست که وضعیت بازپرداخت وام و نرخ بازده شان نامشخص می‌باشد (l)، میزان وام درخواست شده در لیست درخواست (e_i) استخراج می‌گردد. سپس رگرسیون لاجستیک بر روی وام‌های موجود در بخش آموزشی $DataH$ اعمال شده و احتمال نکول (p_j) هر وام آموزشی مشخص می‌شود. سرانجام از روش اعتبار سنجی متقابل یک طرفه برای تنظیم بهینه پهنای باند (h) استفاده می‌شود.

در خط (۴-۱۲) فرآیند ارزیابی مبتنی بر نمونه ریسک اعتباری به روش رگرسیون هسته ای انجام میگیرد، برای این منظور ابتدا با استفاده از روش رگرسیون لاجستیک، احتمال نکول (p_i) هر وام در مجموعه تست Datal در خط ۴ و ۵ محاسبه شده سپس برای هر وام i در مجموعه تست Datal، فاصله احتمال نکولش با هر وام j در مجموعه آموزشی محاسبه می‌شود که نام این فاصله d_{ij} نامیده می‌شود، یعنی فاصله میان احتمال نکول هر وام i در مجموعه تست با هر وام j در مجموعه آموزشی. بازه i و j نیز به ترتیب به تعداد نمونه داده های تست (وام‌های بخش تست) و نمونه داده های آموزشی (وام‌های بخش آموزش) یعنی در محدوده $i = 1,..,l, j = 1,..,n$ می‌باشد. سپس با استفاده از ماتریس فاصله d_{ij} بدست آمده و بهینه سازی پارامتر h و بدست آوردن مقدار بهینه ای برای پارامتر پهنای باند h، در معادله خط ۸ به محاسبه وزنهای هسته ای w_{ij} پرداخته می‌شود. بدین صورت که ماتریس فاصلهd_{ij} و پارامتر h در معادله هسته گوسی در (۵.۴) الی (۷.۴)، جایگذاری شده سپس وزنهای مبتنی بر هسته گوسی w_{ij} حاصل می‌گردد و به عبارتی معادلات ۵ الی ۷ پروسه نگاشت تابع گوسی بر روی ماتریس فاصله را نشان می دهد تا فاصله هسته ای شده به دست آید. سرانجام با قرار دادن وزنهای هسته ای حاصل شده w_{ij} در معادله میانگین وزن دار شده نرخ بازده یعنی (۱.۴)، به ارزیابی نرخ بازده (μ_i) پرداخته می‌شود و برای محاسبه ریسک سرمایه گذاری از واریانس وزن دار شده در (۲.۴) استفاده می‌شود بدین ترتیب که با جایگذاری متغیرهایی از جمله وزن های مبتنی بر هسته گوسی w_{ij}، نرخ بازده وام های آموزشی (R_j) و میانگین وزن دار شده μ_i در (۱.۴)، نرخ بازده برای هر وام در مجموعه تست (هر وام در فهرست درخواست) پیش‌بینی می‌گردد.

در خط ۱۳ الی ۱۹، پروسه بهینه سازی پرتفوی و نحوه تخصیص اعتبار به وام های تست انجام می شود و به عبارتی دیگر درصد سرمایه ای از سرمایه کلی که بهتر است برای هر وام در لیست درخواست (وام‌های تست) به سرمایه گذار پیشنهاد شود را تخمین میزند.

روش رتبه بندی ساده

روش مبتنی بر رتبه بندی۱ روشی است که تخمین احتمال نکول وام های آموزشی را براساس مراجعه به شاخص اعتباری درجه هر وام آموزشی انجام می دهد. مدل مبتنی بر رتبه بندی اصلاح شده۲، نسخه پیشرفته از مدل مبتنی بر رتبه بندی است که برای تخمین بتا و سپس احتمال نکول وام های آموزشی از درجه های با سطوح جزئی تر با مراجعه به شاخص اعتباری زیر درجه وام های آموزشی بهره می برد.

در روش رگرسیون هسته ای که پیشتر به آن اشاره شد، معادله میانگین μ_i برای محاسبه نرخ بازده و معادله واریانس σ_i^2 برای محاسبه ریسک سرمایه گذاری بکار می رفت، روش RBM نیز برای محاسبه نرخ بازده و ارزیابی ریسک سرمایه گذاری از روش میانگین و واریانس گیری بهره می برد اما نحوه محاسبه میانگین و واریانس در جهت محاسبه نرخ بازده و ریسک سرمایه گذاری در روش RBM متفاوت از روش رگرسیون هسته ای است، به عبارتی دیگر در روش رتبه بندی با پایه رگرسیون، نرخ بازده و ریسک سرمایه گذاری (μ_i و σ_i^2) صرفا براساس نمونه وام‌های آموزشی که از لحاظ مشخصه اعتباری درجه۳ با وام درخواستی مورد بررسی (وام بخش تست) مشابهت دارند، محاسبه می‌شود و برای این منظور ضرایب بتا حاصل از رگرسیون لاجستیک به ازای هر یک از درجات به طور جداگانه محاسبه می‌شود. بنابراین مشخصات داده مورد استفاده برای هر دو روش رگرسیون و RBM یکسان است، پس از آن روند محاسبه روش RBM متفاوت می‌شود و این تفاوت در نحوه محاسبه احتمال نکول، میانگین μ_i و واریانس σ_i^2 برای پیش‌بینی نرخ بازده و ریسک سرمایه گذاری می‌باشد. برای مثال در پایگاه داده لندینگ کلاب، شاخصی با نام درجه موجود است که روش رتبه بندی، با استفاده از شاخص درجه، اقدام به تفکیک نمونه وام های آموزشی و دسته بندیشان براساس مقدار شاخص درجه می کند بطوریکه وام های آموزشی با درجه مشابه در دسته ای جداگانه قرار می گیرند، از اینرو ابتدا با توجه به مقادیر شاخص مربوط به درجه هر نمونه وام آموزشی که با ستون درجه در پایگاه داده لندینگ کلاب مشخص شده است و معمولا دارای ۷ نوع مقدار (درجه) است (با مقادیر A تا G مشخص شده است که در فاز نرمالسازی، مقادیر اسمی به عددی ۱ الی ۷ تبدیل شده اند بطوریکه که بجای متغیر A، عدد ۱ در نظر گرفته شده است و به همین ترتیب برای درجه های دیگر عدد ۲ الی ۷ در نظر گرفته شده است)، نمونه وام‌های آموزشی که دارای درجات مشابهی هستند در مجموعه ای جداگانه قرار می‌گیرند بطوریکه هر مجموعه شامل وام های آموزشی با درجه همسان است. پس از تفکیک نمونه وام‌های آموزشی برحسب درجه شان، پردازش محاسبه نرخ میانگین و وایانس (برای حصول نرخ بازده و ریسک سرمایه گذاری) برای هر مجموعه تفکیک شده به طور جداگانه انجام می‌شود. به طور مثال فرض کنید در حال بررسی یک نمونه وام درخواستی جدید و درصدد پیش‌بینی نرخ بازده و ریسک سرمایه گذاری بر روی وام درخواستی (وام تست) هستیم، از آنجایی که روش پیش‌بینی مبتنی بر رتبه بندی نیازمند آموزش اولیه و مراجعه به نمونه های وام در بخش آموزش است، ابتدا به نمونه وام‌های موجود در بخش آموزش مراجعه می‌شود و نمونه وام های آموزشی براساس مشخصه اعتباری درجه به مجموعه های مجزایی تفکیک می شوند بطوریکه چندین

[1] Rating-based method: RBM
[2] Refined Rating- based method: RBM+
[3] Grade

ماتریس (مجموعه) که هر مجموعه حاوی نمونه وام‌های آموزشی متعلق به درجه همسانی اند، حاصل خواهد شد، سپس رگرسیون لجستیک بر روی هر ماتریس به طور جداگانه اعمال شده و ضرایب بتا مربوطه شان استخراج می‌شود. پس از آن، احتمال نکول هر وام آموزشی براساس ضرایب بتای استخراج شده مربوط به درجه شان انجام می گیرد. در مرحله بعد نوبت به محاسبه میانگین و واریانس نرخ بازده و ریسک سرمایه گذاری وام های آموزشی پرداخته می‌شود. در واقع یکی دیگر از تفاوت های دیگر رتبه بندی با روش رگرسیون هسته ای در محاسبه احتمال نکول وام های تست و ماتریس فاصله می باشد بطوریکه در روش رتبه بندی نیاز به محاسبه احتمال نکول وام تست و سپس محاسبه ماتریس فاصله براساس ارتباط میان احتمال نکول وام آموزشی و تست نیست و صرفا از میانگین و واریانس هر دسته وام آموزشی دارای درجه همسان با درجه وام تست برای استخراج نرخ بازده و ریسک (میانگین و واریانس) وام تست استفاده می شود.

نمونه پژوهش

در کار محقق گوئو از روش k فولد[1] برای تفکیک نمونه وام های آموزشی و تست به منظور آزمایش روش ها استفاده شد اما در این مقاله از دو روش از جمله روش ۷۰ به ۳۰ درصد و روش k فولد استفاده شد و مقایسه ای نیز با تنظیمات لحاظ شده در کار محقق گوئو (یعنی با شش شاخصه اعتباری) صورت گرفت. تعداد کل وام‌های مورد بررسی در این مقاله ۱۰ هزار نمونه وام می‌باشد، در حالت تفکیک نمونه های آموزشی و تست به روش ۷۰ به ۳۰ درصد، ۷۰ درصد (۷۰۰۰ وام) به عنوان نمونه وام‌های آموزش دهنده در بخش مجموعه آموزشی و مابقی ۳۰ درصد (۳۰۰۰ نمونه وام) به عنوان نمونه وام‌های تست در بخش مجموعه تست مورد استفاده قرار گردید. در روش اعتبارسنجی k فولد، نمونه وام‌ها به k زیرمجموعه افراز می شوند. از این k زیرمجموعه، هر بار یک مجموعه از وام (بخش) برای تست و k-1 تای دیگر مجموعه وام برای آموزش بکار میروند. این روال k بار تکرار می‌شود و همه نمونه وام‌ها دقیقا یکبار برای آموزش و یکبار برای اعتبارسنجی بکار می روند و در نهایت میانگین نتیجه این k بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود. در این پایان نامه از k=10 یعنی روش 10-Fold استفاده می‌شود. همچنین لازم به ذکر است که در این نوع تفکیک، نحوه انتخاب وام‌ها از میان پایگاه داده برای بخش آموزش و تست به طور تصادفی انجام می‌شود.

یافته های پژوهش

برای تایید و اعتبار سنجی تأثیر روش پیشنهادی از پلتفرم وام دهی لندینگ کلاب استفاده شده است (کلاب، ۲۰۱۶). هر یک از وام گیرندگان دارای مشخصات اعتباری اند که جزئیات آنها در جدول۱ درج شده است. روش اعتبار سنجی k فولد برای تعیین نمونه های اموزشی و تست استفاده شده است. از اینرو با توجه به ده هزار نمونه داده، ۷ هزار نمونه وام به تصادف به عنوان نمونه وام آموزش و مابقی ۳ هزار نمونه وام به عنوان نمونه وام تست در نظر گرفته شده است. با توجه به جدول۱، هر یک از نمونه وام های بخش اموزش و تست دارای ۶ متغیر مستقل

[1] K-Fold

و ۱ متغیر وابسته میباشد. شش متغیر مستقل شامل اطلاعات اعتباری هر وام گیرنده و یک متغیر وابسته نشان دهنده نرخ بازگشت وام توسط هر وام گیرنده میباشد. متغیر وابسته دارای چند برچسب کلاس در رنج ۰ تا ۱ میباشد بطوریکه ۰، نرخ بازگشت با ریسک بسیار بالا و ۱ نرخ بازگشت با ریسک بسیار پایین میباشد و طبعا عددی که به مراتب به ۱ نزدیک تر باشد، شانس بیشتری برای وام گرفتن دارد چون به مراتب ریسک پایینی خواهد داشت. نوع نمونه های این مجموعه داده از نوع real و integer می باشند. این مجموعه داده شامل چندین هزار نمونه از افراد وام گیرنده است که از این چندین هزارنمونه نمونه نیمی از نمونه ها متعلق به کلاس ۱ و ۰ و از نیم دیگر در حدود ۷۰ درصد متعلق به کلاس ۰.۵ و ۳۰ درصد مابقی متعلق به برچسب کلاس‌های دیگر می‌باشد.

جدول ۱. مشخصات اعتباری (متغیرهای مستقل)

نام ویژگی	معادل فارسی
loan amount	مبلغ وام درخواستی
Debt to income ratio	نسبت بدهی به درآمد
interest rate	نرخ بهره
borrower's home ownership status	وضعیت مالکیت منزل وام گیرنده
Average current balance of all accounts	متوسط موجودی حساب ها
Annual income	درآمد سالانه

منبع: (کلاب، ۲۰۱۶).

جدول۲. نمونه وام های آموزشی و تست مورد استفاده در روشهای مورد بررسی. الف. نمونه وام های آموزشی

شماره وام	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*	*R*	*Y*
1	20000	1	12.45	110000	694	0	2	15	1	0
2	20000	1	22.21	70000	684	0	3	16	1	0
3	10000	0	34.7	55000	689	0	3	19	0	1
4	20000	0	17.61	85000	709	0	1	6	1	0
5	10000	1	13.07	85000	689	1	0	2	1	0
6	22400	0	22.98	95000	714	0	2	12	0.5	0
7	8000	0	34.8	42000	704	0	1	10	1	0
8	28000	0	21.6	92000	724	0	0	2	1	0
9	23000	1	18.28	64000	704	0	1	6	0.5	0
10	18000	0	8.68	112000	804	0	2	11	1	0
11	9600	0	22.44	60000	699	0	0	4	1	0
12	25000	0	26.02	109000	749	1	0	4	1	0

شماره وام	X1	X2	X3	X4	X5	X6	X7	X8	R	Y
13	28000	2	7.49	125000	699	0	2	13	1	0
14	15000	0	14.34	44000	689	1	0	1	1	0
15	16000	0	23.35	109000	684	1	3	19	1	0
16	17600	0	26.59	88000	784	0	0	1	1	0
17	6000	0	13.28	100000	699	1	0	4	1	0
18	20000	2	18.83	65000	839	1	0	1	1	0
19	16000	0	18.96	65000	679	0	2	12	0	1
20	12000	1	24.04	98000	694	0	1	9	1	0
21	25000	0	34.53	79000	734	1	2	14	0	1
22	7200	0	8.47	72500	729	1	1	10	1	0
23	7000	0	13.06	59000	734	1	0	4	1	0
24	27500	0	6.79	195000	704	0	2	15	1	0
25	8800	0	4.18	92000	669	1	1	8	1	0
26	14650	1	21.44	43160	694	0	4	24	1	0
27	11000	0	21.34	69000	684	1	1	7	1	0
28	28000	0	31.88	75000	689	0	3	20	1	0
29	10000	0	9.44	91392	714	0	2	12	1	0
30	24000	2	10.24	77900	744	1	1	8	1	0
31	35000	0	17.36	106000	734	0	2	12	0	1
32	14000	0	24.29	60000	694	0	2	15	0	1
33	21875	1	17.76	47590	704	0	2	11	1	0
34	16000	0	28.92	62000	724	0	1	6	0	1
35	14025	0	18	39000	679	1	3	20	1	0
36	20000	2	16.9	70000	684	0	2	14	0	1
37	19000	2	13.09	75000	669	2	2	15	1	0
38	15850	2	34.85	45000	759	1	5	27	1	0
39	25000	0	9.54	150000	799	0	0	1	1	0
40	4200	0	26.21	102500	704	0	2	13	1	0
41	15000	1	10.17	40000	674	0	1	9	1	0
42	12000	0	14.3	53750	704	0	0	1	1	0
43	11000	0	15	85000	759	0	0	2	1	0
44	15000	0	13.24	110000	674	0	1	6	1	0
45	18500	0	13.2	70000	679	1	2	12	1	0

شماره وام	X1	X2	X3	X4	X5	X6	X7	X8	R	Y
46	8000	0	21.23	45000	674	0	1	9	1	0
47	5000	1	13.73	32000	689	0	0	5	1	0
48	25000	0	4.65	135000	709	0	1	10	1	0
49	20000	0	29.15	55000	754	0	0	2	1	0
50	15000	1	18.17	77213	674	2	2	13	1	0
51	17925	2	27.78	39000	724	1	3	18	1	0
52	15000	0	13.89	92000	684	1	1	8	1	0
53	14000	0	14.97	100000	689	0	1	8	1	0
54	10000	1	21.75	33000	714	0	1	10	1	0
55	10400	0	24.37	104433	699	3	5	26	1	0
56	20000	0	10.78	63000	699	0	1	9	1	0
57	3600	0	4.91	55000	679	1	2	14	1	0
58	35000	0	17.06	110000	789	0	2	15	0.5	0
59	24700	0	16.06	65000	719	4	2	11	1	0
60	16000	0	26.4	70000	724	0	2	12	1	0
61	11950	1	10.2	34000	694	0	2	13	1	0
62	18000	1	9.39	150000	669	1	4	22	0	1
63	1400	0	34.95	64000	704	0	2	12	1	0
64	20000	0	14.67	180000	684	0	1	7	1	0
65	16000	0	33.18	48000	689	0	2	15	1	0
66	29900	0	21.77	65000	719	0	2	12	1	0
67	8650	1	24.49	55000	679	4	4	23	1	0
68	5000	1	12.7	79000	669	0	2	13	1	0
69	4225	1	14.22	35000	729	0	2	15	0	1
70	24250	0	20.84	75000	664	0	5	28	0	1
71	21000	1	14.47	52000	684	1	2	14	1	0
72	16800	0	34.29	118000	684	0	2	12	0.5	0
73	8650	0	7.28	100000	754	0	0	1	1	0
74	20000	1	23.45	75000	729	0	1	8	1	0
75	16000	0	17.74	142000	684	2	2	14	1	0
76	2500	0	17.6	50000	674	4	3	18	1	0
77	7200	1	17.35	55000	684	0	3	16	1	0
78	21000	1	12.14	54649	674	0	2	13	1	0

شماره وام	X1	X2	X3	X4	X5	X6	X7	X8	R	Y
79	23850	2	24.71	68046	669	0	3	18	0.5	0
80	15700	0	29.13	48000	689	0	3	17	0.5	0
81	28800	2	18.01	60000	804	0	2	11	0.5	0
82	35000	0	6.46	128000	689	0	2	11	1	0
83	20000	0	12.28	145000	699	0	2	12	1	0
84	28000	0	28.52	85000	674	0	2	15	0.5	0
85	10000	1	10.81	40100	684	1	0	5	1	0
86	7200	0	19.25	50000	679	0	1	7	1	0
87	16000	1	9.6	42000	674	0	2	13	0	1
88	1500	0	16.11	45000	749	0	0	2	1	0
89	10000	1	31.04	42500	709	1	2	11	1	0
90	17600	1	17.56	44000	684	0	4	23	0.5	0
91	16250	0	34.95	50000	694	0	2	13	1	0
92	23975	1	33.2	70000	689	1	4	23	0	1
93	30000	0	18.5	175000	729	1	3	16	1	0
94	20200	1	34.84	60000	724	0	3	20	0	1
95	15000	1	20.18	80000	779	1	0	1	1	0
96	9000	2	11.98	50000	804	2	1	7	1	0
97	27300	0	24.63	65000	664	0	3	18	0	1
98	23100	2	20.43	110000	664	0	4	24	0	1
99	15000	0	29.28	65000	724	0	1	9	1	0
100	10000	0	14.78	41600	699	0	1	9	1	0
101	10000	1	30.84	30000	719	2	3	17	1	0
102	18000	0	24.65	75000	719	1	0	4	1	0
103	3600	0	17.44	85000	679	2	1	10	1	0
104	5000	2	6.52	72000	714	0	1	8	1	0
105	7200	2	13.21	20000	709	1	3	18	0	1
106	12000	1	31.95	40000	669	0	1	9	1	0
107	17475	1	23.58	39750	699	0	2	14	0	1
108	9000	0	8.43	47000	679	0	1	6	1	0
109	11550	1	21.07	38000	704	0	3	17	0	1
110	33600	1	14.01	104000	694	2	1	9	1	0
111	12000	0	23.84	65000	789	0	1	8	1	0

شماره وام	X1	X2	X3	X4	X5	X6	X7	X8	R	Y
112	12000	0	26.32	39400	709	1	1	7	1	0

ب. نمونه وام های تست

شماره وام	X1	X2	X3	X4	X5	X6	X7	X8
113	6000	0	14.45	105000	689	1	0	5
114	16000	2	6.39	90000	704	0	2	12
115	10500	1	8.51	70100	664	0	3	16
116	4000	1	14.83	44500	674	0	1	7
117	4200	1	31.02	32000	709	2	2	14
118	16000	0	14.45	96000	729	1	0	1
119	3500	0	17.97	122400	664	2	1	9
120	12000	2	14.53	68000	694	0	1	7
121	24000	2	26.89	95000	679	1	1	10
122	23525	0	13.33	70000	679	0	3	19
123	2000	0	6.78	140000	850	1	1	6
124	15000	0	14.86	93000	759	1	0	3
125	14875	1	36.3	37231	669	3	3	20
126	21600	0	17.64	54000	694	1	2	13
127	8000	1	23.26	42000	709	1	2	14
128	24000	1	29.92	82193	699	0	2	14
129	35000	0	34.69	120000	704	0	4	24
130	5475	0	33.47	108000	674	2	4	21
131	22400	0	28.48	48000	709	0	1	8
132	7000	1	21.96	83000	684	0	2	14
133	20000	2	36.94	55000	744	1	1	9
134	3200	1	20.61	30407	714	0	1	6
135	25825	1	31.68	68000	679	1	5	30
136	35000	0	24.19	125000	714	0	2	13
137	8000	1	14.86	50000	719	0	0	1
138	2500	1	38.06	16488	774	1	1	10
139	16000	1	13.74	87500	699	0	1	9
140	20400	0	38.65	75000	729	0	3	20
141	5000	0	21.21	78000	704	1	1	10
142	8000	0	11.52	33130	669	1	3	19
143	7500	2	37.2	15000	689	0	3	19
144	16000	0	8.78	70000	709	1	3	16
145	9500	1	27.95	28000	699	0	3	16
146	15600	0	14.37	52000	724	1	2	13
147	21000	0	14.07	82000	734	0	2	11
148	18850	0	14.56	55000	704	2	4	25
149	20000	1	22.62	60000	689	2	1	10
150	5000	1	27.9	17000	694	1	1	10
151	8000	0	14.15	45000	794	0	0	3
152	4575	1	10.03	56000	669	1	3	18
153	5600	1	19.85	48000	664	0	1	7
154	13150	1	20.25	40000	769	0	0	1

شماره وام	*X1*	*X2*	*X3*	*X4*	*X5*	*X6*	*X7*	*X8*
155	9075	0	8.39	40000	819	0	0	1
156	19600	0	18.99	80000	714	1	3	16
157	14000	1	19.24	45500	694	0	3	16
158	28000	0	24.83	75000	689	0	1	6
159	24000	0	31.57	75181.21	724	1	2	12
160	5000	1	9.16	30000	679	0	2	11

جدول۳. نمونه احتمال نکول وام های آموزشی و تست

الف. احتمال نکول وام های آموزشی

شماره وام	p_j
1	0.15538
2	0.368678
3	0.280899
4	0.1168
5	0.064673
6	0.150362
7	0.22132
8	0.168382
9	0.295212
10	0.009773
11	0.120725
12	0.055516
13	0.252733
14	0.080887
15	0.080245
16	0.048195
17	0.020536
18	0.02977
19	0.185634
20	0.198136
21	0.146175
22	0.013003
23	0.017609
24	0.03978
25	0.027523
26	0.301315
27	0.076642
28	0.45017
29	0.035172
30	0.099428
31	0.153027
32	0.182136
33	0.307387
34	0.162496
35	0.108171
36	0.442059
37	0.152308
38	0.232555
39	0.012122
40	0.068093
41	0.24735
42	0.092446
43	0.02562
44	0.095961
45	0.082163
46	0.171361
47	0.15045
48	0.048816
49	0.139067
50	0.082861
51	0.307607
52	0.051986
53	0.09046
54	0.202011
55	0.013357
56	0.115018
57	0.024891
58	0.06113
59	0.009982
60	0.13356
61	0.168026
62	0.065073
63	0.130814
64	0.063032
65	0.346777
66	0.247901
67	0.025361
68	0.129389
69	0.081508
70	0.339111
71	0.188227
72	0.246384
73	0.013201
74	0.214011
75	0.021817
76	0.006645
77	0.180819
78	0.320737
79	0.682873
80	0.289965
81	0.168156
82	0.142612
83	0.059544
84	0.440908
85	0.093091
86	0.131988
87	0.249382
88	0.026143
89	0.17876
90	0.328335
91	0.350194
92	0.398738
93	0.0369
94	0.408704
95	0.030832
96	0.010471
97	0.479092
98	0.533777
99	0.152828
100	0.10179
101	0.091915
102	0.079518
103	0.018441
104	0.085299
105	0.126758
106	0.521609
107	0.358116
108	0.084238
109	0.233558
110	0.112766
111	0.033197
112	0.09354

ب. احتمال نکول وام های تست

شماره وام	p_i
1	0.024871
2	0.158439
3	0.1644
4	0.177905
5	0.075848
6	0.025768
7	0.01703
8	0.259466
9	0.441381
10	0.201338
11	0.000524
12	0.01524
13	0.175839
14	0.113955
15	0.107551
16	0.450336
17	0.444085
18	0.046039
19	0.299308
20	0.178932
21	0.346227
22	0.135028
23	0.453351
24	0.242558
25	0.102304
26	0.077886
27	0.149261
28	0.271509
29	0.034739
30	0.065647
31	0.652418
32	0.033383
33	0.326876
34	0.043875
35	0.07455
36	0.037658
37	0.129425
38	0.179052
39	0.017258
40	0.07116
41	0.260137
42	0.091801
43	0.00844
44	0.062596
45	0.2615
46	0.349152
47	0.142531
48	0.140604

محاسبه ماتریس فاصله

ماتریس فاصله با اختلاف احتمال نکول وام های آموزشی از تست تشکیل می شود. یعنی شباهت میان هر جفت نمونه وام آموزشی و تست (هر جفت وام نظیر به نظیر در بخش آموزش و تست) از طریق تفاضل میان احتمال نکول هر نمونه وام آموزشی با احتمال نکول نمونه وام تست حاصل می‌شود. نتایج حاصل از پیاده سازی این بخش در متلب در جدول زیر مشاهده می‌شود. بطوریکه در ماتریس بدست آمده، سطرهای نشان دهنده نمونه وام‌های آموزشی (وام‌های بخش آموزش) و ستونها نشان دهنده نمونه وام‌های تست (وام‌های بخش تست) هستند.

جدول۴: ماتریس فاصله میان نمونه وام آموزشی و نمونه تست اول

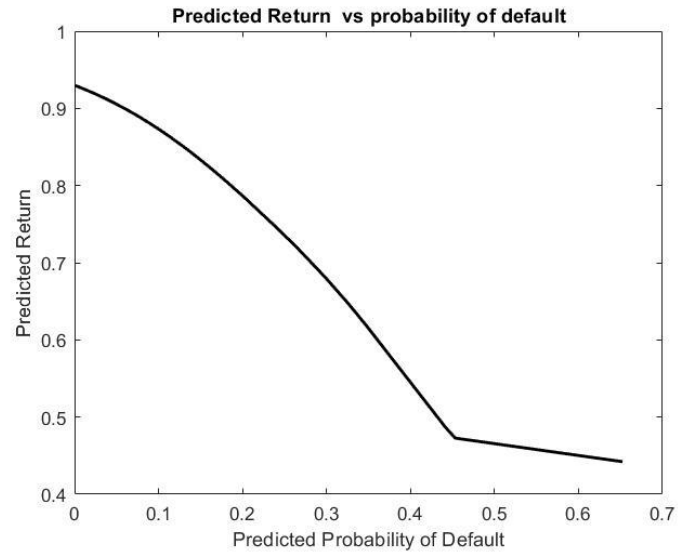
نمونه تست/ نمونه آموزشی	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
۱	*0.130509*	*0.003059*	*0.00902*	*0.022526*	*0.079532*	*0.129612*	*0.13835*	*0.104086*
۲	*0.343807*	*0.210239*	*0.204278*	*0.190772*	*0.29283*	*0.34291*	*0.351648*	*0.109212*
۳	*0.256028*	*0.12246*	*0.116499*	*0.102994*	*0.205051*	*0.255131*	*0.263869*	*0.021433*
٤	*0.091929*	*0.041639*	*0.0476*	*0.061105*	*0.040952*	*0.091032*	*0.09977*	*0.142666*
٥	*0.039802*	*0.093765*	*0.099727*	*0.113232*	*0.011174*	*0.038905*	*0.047643*	*0.194793*
٦	*0.125492*	*0.008076*	*0.014038*	*0.027543*	*0.074515*	*0.124594*	*0.133333*	*0.109104*
۷	*0.19645*	*0.062882*	*0.05692*	*0.043415*	*0.145473*	*0.195552*	*0.204291*	*0.038146*
۸	*0.143511*	*0.009943*	*0.003982*	*0.009524*	*0.092534*	*0.142614*	*0.151352*	*0.091084*
۹	*0.270342*	*0.136774*	*0.130812*	*0.117307*	*0.219365*	*0.269445*	*0.278183*	*0.035747*

نمونه تست/ نمونه آموزشی	1	2	3	4	5	6	7	8
۱۰	0.015098	0.148666	0.154627	0.168132	0.066075	0.015995	0.007257	0.249693
۱۱	0.095854	0.037713	0.043675	0.05718	0.044877	0.094957	0.103695	0.138741
۱۲	0.030645	0.102923	0.108884	0.122389	0.020332	0.029748	0.038486	0.20395
۱۳	0.227862	0.094294	0.088333	0.074827	0.176885	0.226965	0.235703	0.006733
۱٤	0.056016	0.077552	0.083513	0.097018	0.005039	0.055119	0.063857	0.178579
۱٥	0.055374	0.078194	0.084155	0.09766	0.004397	0.054477	0.063215	0.179221
۱٦	0.023324	0.110243	0.116205	0.12971	0.027653	0.022427	0.031165	0.211271
۱۷	0.004335	0.137903	0.143864	0.157369	0.055312	0.005232	0.003506	0.23893
۱۸	0.004899	0.128669	0.13463	0.148135	0.046078	0.004002	0.01274	0.229696
۱۹	0.160764	0.027196	0.021234	0.007729	0.109787	0.159866	0.168605	0.073832
۲۰	0.173265	0.039698	0.033736	0.020231	0.122288	0.172368	0.181106	0.06133
۲۱	0.121304	0.012263	0.018225	0.03173	0.070328	0.120407	0.129145	0.113291
۲۲	0.011868	0.145436	0.151397	0.164903	0.062845	0.012765	0.004027	0.246463
۲۳	0.007262	0.140829	0.146791	0.160296	0.058239	0.008159	0.000579	0.241857
۲٤	0.014909	0.118659	0.12462	0.138125	0.036068	0.014012	0.02275	0.219686
۲٥	0.002652	0.130916	0.136877	0.150382	0.048325	0.001755	0.010493	0.231943
۲٦	0.276444	0.142876	0.136915	0.12341	0.225467	0.275547	0.284285	0.041849
۲۷	0.051771	0.081797	0.087758	0.101264	0.000794	0.050874	0.059612	0.182824
۲۸	0.425299	0.291731	0.28577	0.272264	0.374322	0.424402	0.43314	0.190704
۲۹	0.010301	0.123267	0.129228	0.142733	0.040676	0.009404	0.018142	0.224294
۳۰	0.074557	0.059011	0.064972	0.078478	0.02358	0.07366	0.082398	0.160038
۳۱	0.128156	0.005412	0.011373	0.024879	0.077179	0.127259	0.135997	0.106439
۳۲	0.157265	0.023697	0.017736	0.004231	0.106288	0.156368	0.165106	0.07733
۳۳	0.282516	0.148948	0.142987	0.129482	0.231539	0.281619	0.290357	0.047921
۳٤	0.137625	0.004057	0.001904	0.01541	0.086648	0.136728	0.145466	0.09697
۳٥	0.0833	0.050268	0.056229	0.069734	0.032323	0.082403	0.091141	0.151295
۳٦	0.417188	0.283621	0.277659	0.264154	0.366211	0.416291	0.425029	0.182593
۳۷	0.127437	0.006131	0.012092	0.025597	0.07646	0.12654	0.135278	0.107158
۳۸	0.207684	0.074116	0.068155	0.05465	0.156707	0.206787	0.215525	0.026911
۳۹	0.012748	0.146316	0.152278	0.165783	0.063725	0.013646	0.004907	0.247344
٤۰	0.043222	0.090346	0.096307	0.109812	0.007755	0.042325	0.051063	0.191373
٤۱	0.222479	0.088912	0.08295	0.069445	0.171502	0.221582	0.23032	0.012116
٤۲	0.067576	0.065992	0.071954	0.085459	0.016599	0.066679	0.075417	0.167019
٤۳	0.000749	0.132819	0.13878	0.152286	0.050228	0.000148	0.00859	0.233846
٤٤	0.07109	0.062478	0.068439	0.081945	0.020113	0.070193	0.078931	0.163505
٤٥	0.057292	0.076276	0.082237	0.095742	0.006315	0.056395	0.065133	0.177303
٤٦	0.14649	0.012922	0.006961	0.006544	0.095513	0.145593	0.154331	0.088105
٤۷	0.125579	0.007989	0.01395	0.027455	0.074602	0.124682	0.13342	0.109016
٤۸	0.023945	0.109623	0.115584	0.129089	0.027032	0.023048	0.031786	0.21065
٤۹	0.114196	0.019372	0.025333	0.038838	0.063219	0.113299	0.122037	0.120399
٥۰	0.05799	0.075578	0.081539	0.095045	0.007013	0.057093	0.065831	0.176605
٥۱	0.282736	0.149169	0.143207	0.129702	0.231759	0.281839	0.290577	0.048141
٥۲	0.027115	0.106452	0.112414	0.125919	0.023862	0.026218	0.034956	0.20748
٥۳	0.065589	0.067978	0.07394	0.087445	0.014612	0.064692	0.07343	0.169006

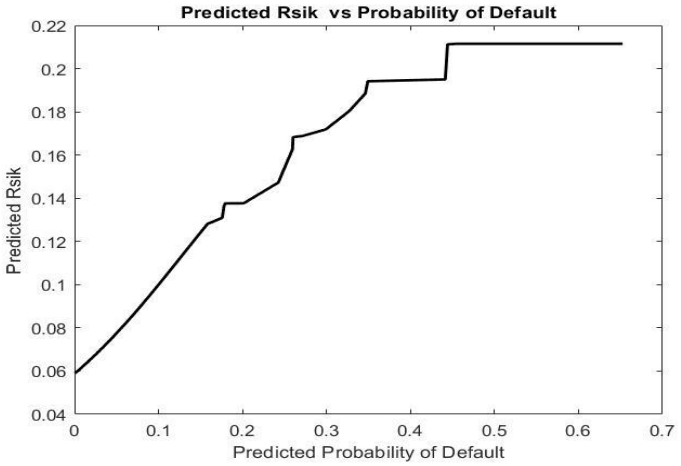
نمونه تست/ نمونه آموزشی	1	2	3	4	5	6	7	8
٥٤	0.17714	0.043572	0.037611	0.024105	0.126163	0.176243	0.184981	0.057455
٥٥	0.011514	0.145082	0.151043	0.164548	0.062491	0.012411	0.003673	0.246109
٥٦	0.090147	0.04342	0.049382	0.062887	0.039171	0.08925	0.097988	0.144448
٥٧	2.01E-05	0.133548	0.139509	0.153014	0.050957	0.000877	0.007861	0.234575
٥٨	0.036259	0.097309	0.10327	0.116776	0.014718	0.035362	0.0441	0.198336
٥٩	0.014889	0.148457	0.154418	0.167923	0.065866	0.015786	0.007048	0.249484
٦٠	0.108689	0.024878	0.03084	0.044345	0.057713	0.107792	0.11653	0.125906
٦١	0.143155	0.009587	0.003626	0.009879	0.092178	0.142258	0.150996	0.09144
٦٢	0.040202	0.093365	0.099327	0.112832	0.010775	0.039305	0.048043	0.194393
٦٣	0.105944	0.027624	0.033586	0.047091	0.054967	0.105046	0.113785	0.128652
٦٤	0.038161	0.095407	0.101368	0.114873	0.012816	0.037264	0.046002	0.196434
٦٥	0.321906	0.188338	0.182377	0.168871	0.270929	0.321009	0.329747	0.087311
٦٦	0.22303	0.089462	0.083501	0.069996	0.172053	0.222133	0.230871	0.011565
٦٧	0.00049	0.133078	0.139039	0.152544	0.050487	0.000407	0.008331	0.234105
٦٨	0.104518	0.02905	0.035011	0.048516	0.053541	0.103621	0.112359	0.130077
٦٩	0.056637	0.07693	0.082892	0.096397	0.00566	0.05574	0.064478	0.177958
٧٠	0.314241	0.180673	0.174711	0.161206	0.263264	0.313343	0.322082	0.079645
٧١	0.163357	0.029789	0.023828	0.010322	0.11238	0.16246	0.171198	0.071238
٧٢	0.221513	0.087945	0.081984	0.068478	0.170536	0.220616	0.229354	0.013082
٧٣	0.011669	0.145237	0.151199	0.164704	0.062646	0.012567	0.003828	0.246265
٧٤	0.189141	0.055573	0.049612	0.036106	0.138164	0.188244	0.196982	0.045454
٧٥	0.003054	0.136622	0.142583	0.156088	0.054031	0.003951	0.004787	0.237649
٧٦	0.018226	0.151793	0.157755	0.17126	0.069203	0.019123	0.010385	0.252821
٧٧	0.155948	0.02238	0.016419	0.002913	0.104971	0.155051	0.163789	0.078647
٧٨	0.295866	0.162299	0.156337	0.142832	0.244889	0.294969	0.303707	0.061271
٧٩	0.658002	0.524434	0.518473	0.504968	0.607025	0.657105	0.665843	0.423407
٨٠	0.265094	0.131526	0.125565	0.112059	0.214117	0.264197	0.272935	0.030499
٨١	0.143285	0.009717	0.003756	0.009749	0.092308	0.142388	0.151126	0.09131
٨٢	0.117741	0.015827	0.021788	0.035293	0.066764	0.116844	0.125582	0.116854
٨٣	0.034673	0.098895	0.104856	0.118362	0.016304	0.033776	0.042514	0.199922
٨٤	0.416038	0.28247	0.276509	0.263003	0.365061	0.415141	0.423879	0.181443
٨٥	0.06822	0.065348	0.071309	0.084814	0.017243	0.067323	0.076061	0.166375
٨٦	0.107117	0.026451	0.032412	0.045917	0.05614	0.10622	0.114958	0.127478
٨٧	0.224511	0.090944	0.084982	0.071477	0.173534	0.223614	0.232352	0.010084
٨٨	0.001272	0.132296	0.138257	0.151762	0.049705	0.000375	0.009113	0.233323
٨٩	0.153889	0.020322	0.01436	0.000855	0.102912	0.152992	0.16173	0.080706
٩٠	0.303464	0.169897	0.163935	0.15043	0.252487	0.302567	0.311305	0.068869
٩١	0.325323	0.191755	0.185794	0.172289	0.274346	0.324426	0.333164	0.090728
٩٢	0.373867	0.240299	0.234338	0.220832	0.32289	0.37297	0.381708	0.139272
٩٣	0.01203	0.121538	0.127499	0.141005	0.038947	0.011133	0.019871	0.222565
٩٤	0.383833	0.250265	0.244304	0.230798	0.332856	0.382936	0.391674	0.149238
٩٥	0.005961	0.127607	0.133568	0.147074	0.045016	0.005064	0.013802	0.228634
٩٦	0.0144	0.147968	0.153929	0.167435	0.065377	0.015297	0.006559	0.248995
٩٧	0.454222	0.320654	0.314692	0.301187	0.403245	0.453325	0.462063	0.219627

نمونه تست/ نمونه آموزشی	1	2	3	4	5	6	7	8
۹۸	0.508906	0.375338	0.369377	0.355872	0.457929	0.508009	0.516747	0.274311
۹۹	0.127957	0.005611	0.011572	0.025078	0.07698	0.12706	0.135798	0.106638
۱۰۰	0.076919	0.056649	0.06261	0.076115	0.025942	0.076022	0.08476	0.157676
۱۰۱	0.067044	0.066523	0.072485	0.08599	0.016067	0.066147	0.074885	0.167551
۱۰۲	0.054647	0.078921	0.084882	0.098388	0.00367	0.05375	0.062488	0.179948
۱۰۳	0.00643	0.139997	0.145959	0.159464	0.057407	0.007327	0.001411	0.241025
۱۰٤	0.060428	0.07314	0.079101	0.092606	0.009451	0.059531	0.068269	0.174167
۱۰٥	0.101887	0.031681	0.037642	0.051147	0.05091	0.10099	0.109728	0.132708
۱۰٦	0.496738	0.363171	0.357209	0.343704	0.445761	0.495841	0.504579	0.262143
۱۰۷	0.333245	0.199677	0.193716	0.18021	0.282268	0.332348	0.341086	0.09865
۱۰۸	0.059367	0.074201	0.080162	0.093668	0.00839	0.05847	0.067208	0.175228
۱۰۹	0.208688	0.07512	0.069159	0.055653	0.157711	0.207791	0.216529	0.025907
۱۱۰	0.087895	0.045673	0.051634	0.06514	0.036918	0.086998	0.095736	0.1467
۱۱۱	0.008327	0.125241	0.131203	0.144708	0.04265	0.00743	0.016168	0.226268
۱۱۲	0.06867	0.064898	0.070859	0.084365	0.017693	0.067773	0.076511	0.165925

در ادامه نتایج حاصل از هر یک از دو روش مورد بررسی بر روی داده های مذکور در شکل ۱و۲ نشان داده شده است. همچنین لازم به ذکر است که پس از مرتب کردن احتمال های نکول به صورت صعودی، نمودار مربوط به رابطه نرخ بازده و ریسک سرمایه گذاری با هر یک از احتمال نکول های مورد نظر نشان داده شده است. با مقایسه نتایج حاصل از رابطه احتمال نکول با نرخ بازده در هر دو روش، مشاهده می شود که نمودار تخمین زده شده در روش رگرسیون هسته ای رابطه جزئی تر و ریز تری را میان احتمال نکول و نرخ بازده نشان می دهد که برای نمودار مربوط به رابطه احتمال نکول و ریسک سرمایه گذاری هم این مهم صادق است، به عبارتی دیگر در روش رگرسیون هسته ای رابطه احتمال نکول - نرخ بازده، رابطه احتمال نکول – سرمایه گذاری دارای ساختاری دانه ای است و از نظریه دانه بندی ریز تبعیت میکنید و روابط را جزئی تر نشان می دهد اما در روش رتبه بندی، نمودار حاصل از روابط احتمال نکول- نرخ بازده و احتمال نکول- ریسک سرمایه گذاری در بازه های طولانی ثابت و مشابه است ارتباط جزئی و دقیق میان احتمال نکول با تخمین های پیش بینی شده را آنطور که باید به دقت و مطابق با واقعیت نشان نمی دهد. از اینرو روش رگرسیون هسته ای در مقایسه با روشهای رتبه بندی قادر به تخمین دقیق تر نرخ بازده و ریسک سرمایه گذاری و ارتباط این دو تخمین (نرخ بازده و ریسک سرمایه گذاری) با احتمال نکول وام می باشد. ازاینرو بنابر منحنی های تخمین در روش مبتنی بر رگرسیون هسته ای در مقایسه با روش مبتنی بر رتبه بندی هموارتر می‌باشد بطوریکه ظرفیت بیشتری برای پیش‌بینی نرخ بازده و ریسک سرمایه گذاری داشته و در مقایسه با روشهای مبتنی بر مدل رتبه بندی ساده و نسخه توسعه یافته اش در کل پیش‌بینی دقیقتری در احتمال نکول وام و نرخ بازده دارد.

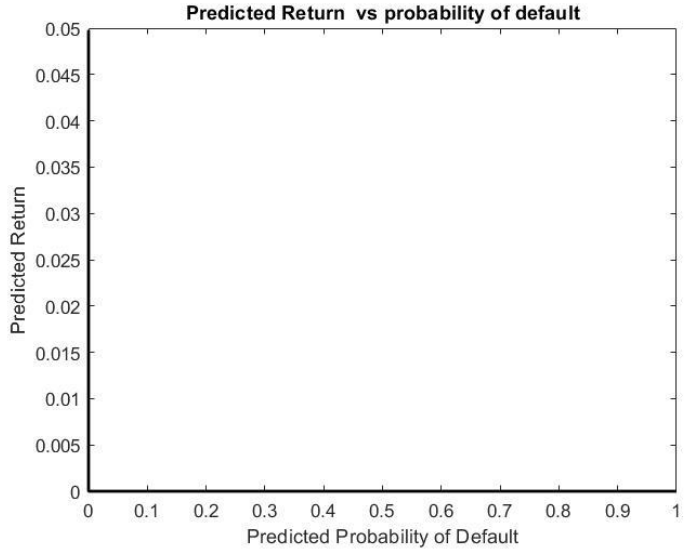


الف. احتمال نکول وام در برابر بازده وام

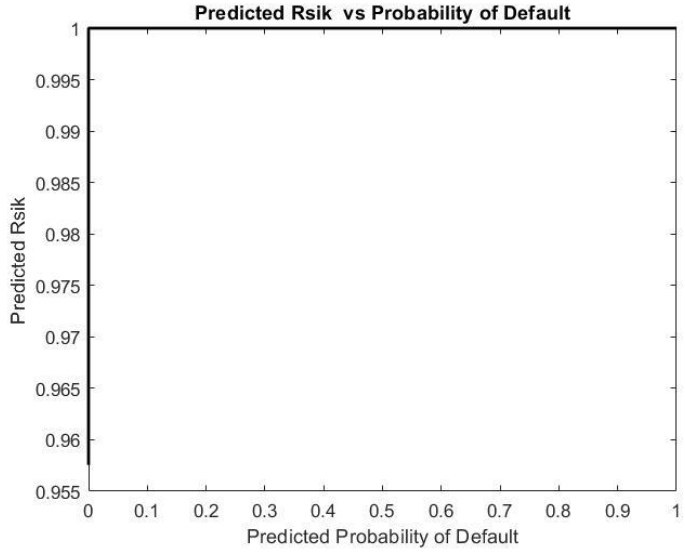


ب. احتمال نکول وام در برابر ریسک سرمایه گذاری

شکل ۱. روش رگرسیون هسته ای



الف. احتمال نکول وام در برابر نرخ بازده



ب. احتمال نکول وام در برابر ریسک سرمایه گذاری

شکل ۲. روش رتبه بندی

نتایج حاصل از میزان و درصد اعتبار اختصاص یافته به وام های تست در هر دو روش مورد بررسی در جدول۵ بخش الف الی پ نشان داده شده است. در جدول۵مشاهده می شود که به کدامیک از وام های تست، تخصیص اعتبار داده شده است و همچنین میزان و درصد اعتبار اختصاص یافته را نشان می دهد. لازم به ذکر است که در بهینه سازی سبد، سرمایه کل، ۵ هزار دلار در نظر گرفته شد. مجموع مقادیر موجود در ستون درصد اعتبارتخصیص یافته معادل با ۱ یا همان ۱۰۰ درصد و مجموع مقادیر موجود در ستون میزان اعتبارتخصیصی از سرمایه کل به هر وام تست معادل با ۵ هزار دلار می باشد.

جدول۵- میزان و درصد اعتبار تخصیص یافته و به هر وام (الف. رگرسیون هسته ای)

شماره وام تست	درصد اعتبار	میزان اعتبار	شماره وام تست	درصد اعتبار	میزان اعتبار
1	0.043754	$218.77	25	0.019562	$97.81
2	0.012164	$60.82	26	0.024857	$124.29
3	0.011649	$58.25	27	0.013038	$64.19
4	0.010622	$53.11	28	0.006769	$33.85
5	0.025377	$126.89	29	0.039247	$196.24
6	0.043321	$216.61	30	0.02819	$140.95
7	0.047738	$238.69	31	0.011992	$59.96
8	0.007072	$34.36	32	0.039835	$199.17
9	0.004545	$22.72	33	0.005651	$28.25
10	0.009216	$46.08	34	0.035531	$177.65
11	0.057456	$287.28	35	0.025716	$128.58
12	0.048702	$243.51	36	0.038014	$190.07
13	0.010768	$53.84	37	0.015323	$76.62
14	0.01756	$87.80	38	0.010543	$52.71
15	0.018623	$93.12	39	0.047617	$238.08
16	0.004536	$22.68	40	0.026626	$133.13
17	0.004541	$22.71	41	0.007054	$34.27
18	0.03471	$173.55	42	0.021641	$108.20
19	0.006165	$30.82	43	0.052556	$262.78
20	0.010551	$52.75	44	0.029104	$144.52
21	0.005335	$26.68	45	0.007018	$34.09
22	0.014617	$73.09	46	0.005291	$26.46
23	0.004535	$22.67	47	0.013749	$68.74
24	0.007557	$37.79	48	0.013965	$69.82

ب. رتبه بندی ساده

شماره وام تست	درصد اعتبار	میزان اعتبار
2	0.090774	$453.87
5	0.090774	$453.87
14	0.090774	$453.87
15	0.090774	$453.87
16	0.090774	$453.87

میزان اعتبار	درصد اعتبار	شماره وام تست
$453.87	0.090774	20
$453.87	0.090774	24
$453.87	0.090774	34
$453.87	0.090774	35
$453.87	0.090774	47
$461.29	0.092257	48

جدول۶ نتایج نرخ بازده حاصل از روش بهینه سازی سبد سهام به واسطه ارزیابی ریسک اعتباری را در هر فولد از اعتباری سنجی k فولد به ازای سه روش مورد بررسی نشان می دهد. در روش اعتباری سنجی k فولد، براساس معیار k ، کل نمونه ها به k بخش مساوی تجزیه می شوند و هر بار تا رسیدن به $k=10$، یک بخش از تمام بخش ها به عنوان نمونه های تست در نظر گرفته می شود، بنابراین اگر $k=10$ انتخاب شود، کل نمونه وام ها به ۱۰ بخش مساوی تجزیه می شوند و هر بار یک بخش به عنوان نمونه وام های تست و مابقی به عنوان نمونه های آموزشی در نظر گرفته می شوند، به طور مثال بار اول، بخش اول به عنوان نمونه وام های تست انتخاب می شود و مابقی به عنوان نمونه وام های آموزشی و بار دوم یعنی به ازای $k=2$، بخش دوم به عنوان نمونه وام های تست انتخاب می شود و مابقی بخش ها به عنوان نمونه های آموزشی و بهمین منوال تا رسیدن به $k=10$، بخش دهم به عنوان نمونه وام های تست و مابقی بخشها یعنی ۹ بخش دیگر به عنوان نمونه وام آموزشی انتخاب می گردند، از اینرو در جدول۳، نتایج حاصل از متوسط نرخ بازده به ازای هر فولد موردی بررسی یعنی از $k=1$ الی $k=10$ نشان داده می شود و در انتها متوسط، انحراف معیار نرخ های بازده در فولدهای مختلف یا همان زیر مجموعه تست های مختلف محاسبه می شود.

جدول۶- میانگین، انحراف معیار نرخ بازده در زیرمجموعه تست های مختلف

رتبه بندی ساده	رگرسیون هسته ای	شماره فولد	رتبه بندی ساده	رگرسیون هسته ای	شماره فولد
0.1252	0.161742	7	0	0.166382	1
0	0.173526	8	0	0.176213	2
0.10214	0.132975	9	0.00032	0.188248	3
0.022415	0.18132	10	0	0.172842	4
0.03624	0.173012	میانگین	0	0.199514	5
0.051083	0.01766	انحراف معیار	0.11233	0.177356	6

نتیجه گیری

در این مقاله به بررسی روش های رگرسیون هسته ای وزن دار شده و مقایسه اش با روش رتبه بندی ساده در جهت ارزیابی ریسک اعتباری و سپس تصمیم گیری برای ارائه پرتفویی بهینه پرداخته شد. لازم به ذکر است با مقایسه روش های مورد بررسی با شاخص های در نظر گرفته شده در کار محقق گوئو و شاخص های انتخابی

توسط ما نشان داده شد که شاخص هایی که در نظر گرفته ایم منجر به پیش بینی نرخ بازده و ریسک دقیق تری خواهد شد، بنابراین با مقایسه نتایج میان کار محقق گوئو و روش مورد بررسی ما نشان داده شد که شاخص های در نظر گرفته شده در کار محقق گوئو کم بوده و با افزایش میزان شاخص های اعتباری می توان پیش بینی دقیقتری را برای نرخ بازده و ریسک داشت.

درخصوص مقایسه میان روش اعتباری سنجی k فولد و تفکیک نمونه های داده به روش معمول ۷۰ به ۳۰درصد، واضح است که با مقایسه نتایج حاصل از سه روش مورد بررسی، روش k فولد به دلیل انتخاب تصادفی نمونه های داده هم در بخش آموزشی و هم در بخش تست موجب انتخاب پویا و متغیر نمونه های وام از نواحی مختلف پایگاه داده و قرار گیری آن در دو بخش آموزشی و تست می گردد و از همین رو به دلیل مورد انتخاب قرار دادن نمونه های داده با شاخص ها و نرخ بازده های آموزشی متغیری در بخش تست، شاهد پیش بینی دقیقتری هستیم و روش اعتبار سنجی k فولد و مقایسه با روش معمول ۷۰ به ۳۰ درصد قدرت پیش بینی بهتری داشته است. با مشاهده نتایج دقت پیش‌بینی حاصل از نرخ بازده وام با روش های داده محور مبتنی بر رگرسیون هسته ای و روشهای مبتنی بر رتبه بندی با حالت تغییر یافته در روش پیشنهادی از جمله تغییر و افزایش در مشخصه های اعتباری روشها واضح است که روش داده محور مبتنی بر رگرسیون هسته ای شده به روش گوسی، عملکرد بهتری نسبت به روش مبتنی بر رتبه بندی دارد و علت اصلی آن در این است که روش رگرسیون هسته ای از تمام وام های گذشته برای تخمین ریسک وام های تست استفاده می کند و دارای آموزش طولانی اما دقیقتر است اما مدلهای رتبه بندی صرفا از وام های آموزشی دارای درجه مشابه با درجه وام های تست برای محاسبه ریسک و بازده وام های تست بهره می برد و دارای پروسه آموزش کوتاه اما غیر دقیق است. همچنین دقت پیش‌بینی مشخص شده از روش رگرسیون هسته ای گوسی به نمونه یادگیری شبیه تر است. در نتایج حاصل شده از روش مبتنی بر رتبه بندی، مقادیر نرخ بازپرداخت وام به ازای هر نمونه داده در نمونه یادگیری مشخص نیست اما در روش رگرسیون هسته ای گوسی واضح می‌باشد. بدین معنا که روش بهبود یافته، قابلیت شبیه سازی روش رگرسیون معمول و مبنی بر رتبه بندی را بهبود می‌دهد. نمونه داده هر رکورد نباید توسط روش رگرسیون یکسانی پردازش شوند. همچنین افزایش شاخص‌های اعتباری موجب بهتر شدن دقت پیش‌بینی شد. میزان دقت نهایی روش رگرسیون با هسته گوسی بر روی پایگاه داده لندینگ کلاب همگرایی مناسبی داشته است و نسبت به روشهای مبتنی بر رتبه بندی نتیجه به مراتب بهتری دارد و با مقایسه میزان دقت های نهایی پیش‌بینی در میان روش رگرسیون هسته ای گوسی و روشهای مورد مقایسه دیگر می توان نتیجه گرفت که علی رغم شرایط یکسان در نظر گرفته شده برای تمام روشها، روش رگرسیون هسته ای گوسی نتیجه بهتری را در حل مسئله پیش‌بینی ریسک بازپرداخت وام داشته است. اگرچه زمان اجرای روشهای مورد مقایسه کمتر از این مقدار بوده است اما از نظر کیفی میزان دقت حاصل از پیش‌بینی ریسک سرمایه گذاری و ریسک بازپرداخت وام در آنها کمتر از روش رگرسیون هسته ای شده با تابع گوسی می‌باشد. مزیت روش مورد بررسی مبتنی بر رگرسیون هسته ای در مقایسه با روشهای دیگر در زیر بخش سازی خودکار وزن دار نمودن نمونه وام‌ها براساس شاخص‌های اعتباریشان می‌باشد. همچنین این زیر بخش سازی، امکان می‌دهد که راه حلهای مسئله قادر به یافتن تمام نقاط بهینه به طور همزمان

باشند. از اینرو روش پیشنهادی به دلیل وزن دار نمودن نمونه های داده، با دقت بیشتری به پیش‌بینی می‌پردازد، بعلاوه روش رگرسیون هسته ای در مقایسه با روش معمول مبتنی بر رتبه بندی به دلیل پردازش اطلاعات در دو مرحله کلی و اعمال وزنهای آگاهانه به هر متغیر مستقل (شاخص‌های اعتباری)، در هر یک از دو مراحل با دقت بیشتری به تحلیل روابط میان متغیر های مستقل (شاخص‌های اعتباری) در جهت تصمیم گیری مقدار متغیر وابسته (ریسک سرمایه گذاری و نرخ بازده وام) می‌پردازد.

پیشنهادهای تحقیق

پیشنهاداتی برای محققان در جهت ادامه کار فعلی و بهبود آن ارائه می شود.

۱. بکارگیری شبکه عصبی با استفاده از تکنیک های بهینه سازی برای امتیازدهی به افراد وام گیرنده در جهت ارزیابی خوش حساب و بدحساب بودن مشتریان.

در این نوع پیشنهاد می‌توان شاخص‌های اعتباری موجود در لندینگ کلاب را با متغیر وابسته دو حالته از نوع خوش حساب و بدحساب در نظر گرفت. بدین صورت که متغیرهای مستقل شاخص‌های اعتباری برای هر نمونه داده وام گیرنده می‌توانند از طریق شبکه عصبی صرفا متغیر وابسته با دو برچسب کلاس خوش حساب و بدحسابی را پیش‌بینی کنند.

۲- اضافه نمودن گام پیش پردازش انتخاب ویژگی به روش شبکه عصبی در جهت انتخاب شاخص‌های اعتباری مفید در پایگاه داده لندینگ کلاب در جهت افزایش دقت پیش‌بینی بازپرداخت وام و حذف آزمون های سعی و خطا.

پایگاه داده لندینگ کلاب حاوی شاخص‌های اعتباری(ویژگیهای داده) بسیاری می‌باشد که برخی از آنها نامرتبط و زائد می‌باشند. شاخص‌های زائد و نامفید فاقد اهمیت برای پیش‌بینی بوده و تاثیر چندانی در پیش‌بینی نخواهند گذاشت بعلاوه به دلیل ایجاد مشکل نفرین ابعاد می‌تواند موجب عملکرد پایین دقت پیش‌بینی گردد. روش انتخاب ویژگی منجر به انتخاب زیر مجموعه ای مفید از کل شاخص‌های اعتباری در پایگاه داده لندینگ کلاب و در نتیجه کاهش بعد داده و بهتر شدن عملکرد پیش‌بینی می‌گردد. به عبارت دیگر بکارگیری تمام شاخص‌های اعتباری موجود در پایگاه داده لندینگ کلاب نه تنها موجب افزایش پیچیدگی محاسباتی و متعاقبا افزایش زمان اجرا بلکه گاهی موجب کاهش دقت پیش‌بینی نیز می‌گردد. در این مواقع تکنیک انتخاب ویژگی برای کاهش بعد داده می‌تواند خیلی موثر واقع گردد. روشهای انتخاب ویژگی زیر مجموعه ای از ویژگیهای شاخص‌های اعتباری با دقت بالا را انتخاب می‌کنند. تاکنون روشهای مختلفی ارائه شده اند که از روشهای بهینه سازی به همراه معیارهای ارتباط سنجی از جمله مجموعه های راف، اطلاعات متقابل نرمال شده، شاخص رند، معیار رگرسیون و... در جهت انتخاب ویژگیهای مفید بهره میبرند اما برخی از آنها یا زمان اجرای بسیار بالایی را برای یافتن زیر مجموعه ویژگیهای بهینه صرف می‌کنند یا دچار بهینگی محلی و همگرایی زودرس و ارائه پاسخ نامناسبی می‌شوند. از اینرو یک پیشنهاد می‌تواند استفاده از تکنیک بهینه سازی جایا به همراه یک معیار سنجش ارتباط میان متغیرهای

مستقل شاخص‌های اعتباری(ویژگی های داده مستقل) باشد. الگوریتم بهینه سازی جایا به دلیل مزیتهایش می‌تواند نقش بارزی در کاهش زمان اجرا و افزایش دقت در انتخاب ویژگی های مفید داده داشته باشد

محدودیت های جمع آوری داده ها

بدون رضایت قبلی از لندینگ کلاب، نمی توان موارد زیر را انجام داد:

۱) برای دسترسی به این سایت یا جمع آوری هرگونه اطلاعات از سایت (از جمله، بدون محدودیت، ربات ها، اسکریپت ها یا سایر دستگاه ها یا برنامه های خودکار) از هر وسیله خودکار استفاده کنید.

۲) چارچوب سایت، استفاده از تکنیک های قاب بندی برای محصور کردن هرگونه محتوا یا سایر اطلاعات اختصاصی، قرار دادن پنجره های بازشو بر روی صفحات این سایت یا تأثیر دیگر در نمایش صفحات این سایت.

۳) هرگونه تبلیغات ناخواسته یا غیرمجاز، مواد تبلیغاتی، "نامه ناخواسته"، "هرزنامه"، "نامه های زنجیره ای"، "طرح های هرمی" یا هر نوع درخواست دیگر تجاری یا غیر تجاری را بارگذاری، ارسال، ایمیل، ارسال یا در غیر این صورت در دسترس قرار دهید یا ارتباطات انبوه از هر نوع با هر عضو لندینگ کلاب، یا شخص ثالث.

۴) از داده های ارائه شده توسط لندینگ کلاب، که به هر طریقی ارائه می شود، برای هرگونه استفاده یا اهداف رقابتی استفاده کنید. شما همچنین موافقت می کنید که هرگز از داده های ارائه شده لندینگ کلاب، در گذشته برای رقابت با محصولات یا خدمات لندینگ کلاب، استفاده نکنید.

۵) از این سایت به هر روشی که قانون قابل اجرا را نقض کند یا باعث آسیب، غیرفعال کردن، فشار بیش از حد یا آسیب رساندن به این سایت شود، استفاده شود یا از سایت در جهت تفریح یا مواردی دیگر که جنبه تفریحی و تبلیغات داشته باشد، استفاده شود

فهرست منابع

Club, L., Lending club. 2016, Recuperado el.

de Paula, D.A.V., et al., Estimating credit and profit scoring of a Brazilian credit union with logistic regression and machine-learning techniques. 2019.

Djeundje, V.B. and J.J.E.J.o.O.R. Crook, Dynamic survival models with varying coefficients for credit risks. 2019. 275(1): p. 319-333.

Du, N., et al., Prosocial compliance in P2P lending: A natural field experiment. 2020. 66(1): p. 315-333.

Emekter, R., et al., Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending. 2015. 47(1): p. 54-70.

Faradharmastuti, C., &Laurentxius, J. (2021). Factors and benefits that affect lender's interest in giving loans in peer to peer (P2P) lending platform. Binus Business Review, 12(2), 121-130. https://doi.org/10.21512/bbr.v12i2.6359

Ferreira, V., Papaoikonomou, E., &Terceño, A. (2022). Unpeel the layers of trust! A comparative analysis of crowdfunding platforms and what they do to generate trust. Business Horizons, 65(1), 7-19. https://doi.org/10.1016/j.bushor.2021.08.004.

Guo, Y., et al., Instance-based credit risk assessment for investment decisions in P2P lending. 2016. 249(2): p. 417-426.

Habachi, M., S.J.C.B. Benbachir, and Management, Combination of linear discriminant analysis and expert opinion for the construction of credit rating models: The case of SMEs. 2019. 6(1): p. 1685926.

Harvey, J. (2018). Peer to peer investing guide for beginners. https://www.amazon.com/Peer -Peer - Investing-Guide-Beginnersebook/dp/B079J6JX62/ref.

Jiang, J., Liao, L., Wang, Z., & Zhang, X. (2021). Government affiliation and peer-to-peer lending platforms in China. Journal of Empirical Finance, 62(June), 87-106. https://doi.org/10.1016/j.jempfin.2021.02.004.

Klafft, M. Online peer-to-peer lending: a lenders' perspective. in Proceedings of the international conference on E-learning, E-business, enterprise information systems, and E-government, EEE. 2008.

Larrimore, L., et al., Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. 2011. 39(1): p. 19-37.

Leonard, T., Logistic Regression, in A Course in Categorical Data Analysis. 2020, Chapman and Hall/CRC. p. 139-152.

Luo, S., Sun, Y., Yang, F., & Zhou, G. (2022). Does fintech innovation promote enterprise transformation? Evidence from China. Technology in Society, 68, 1-13. https://doi.org/10.1016/j.techsoc.2021.101821.

Möstel, L., M. Pfeuffer, and M.J.C.S. Fischer, Statistical inference for Markov chains with applications to credit risk. 2020: p. 1-26.

Mudjahidin, Hidayat, A. A., &Aristio, A. P. (2022). Conceptual model of use behavior for peer-to-peer lending in Indonesia. Procedia Computer Science, 197, 215-222. https://doi.org/10.1016/j.procs.2021.12.134.

Pabst, S., Wayand, M., & Mohnen, A. (2021). Coordinating contributions in crowdfunding for sustainable entrepreneurship. Journal of Cleaner Production, 319, 1-8. https://doi.org/10.1016/j.jclepro.2021.128677.

Puro, L., et al., Borrower decision aid for people-to-people lending. 2010. 49(1): p. 52-60.

R. Emekter, Y. Tu, B. Jirasakuldech, and M. J. A. E. Lu, "Evaluating credit risk and loan performance in online Peer-to-Peer (P2P) lending," vol. 47, no. 1, pp. 54-70, 2014.

Raboun, O., E. Chojnacki, and A. Tsoukias. Dynamic-R: A New “Convincing” Multiple Criteria Method for Rating Problem Statements. in International Conference on Decision Support System Technology. 2019. Springer.

Safidastjerdi, D., Tayebi, K., &Elahi, N. (2021). Loan interest rate uncertainty and financing SMEs listed in Tehran stock exchange.Journal of Asset Management and Financing, 9(2), 1-20. https://doi.org/10.22108/AMF.2020.124509.1578. (In Persian)

Song, P., Chen, Y., Zhou, Z., & Wu, H. (2018). Performance analysis of peer-to-peer online lending platforms in China, Sustainability,10(9), 1-15. https://doi.org/10.3390/su10092987

Suryono, R. R., Budi, I., & Purwandari, B. (2021). Detection of fintech P2P lending issues in Indonesia. Heliyon, 7(4), 1-10.https://doi.org/10.1016/j.heliyon.2021.e06782

Tao, W. and D. Chang. Credit Risk Assessment of P2P Lending Borrowers based on SVM. in 1st International Conference on Business, Economics, Management Science (BEMS 2019). 2019. Atlantis Press.

Thomas, L., J. Crook, and D. Edelman, Credit scoring and its applications. 2017

Wang, C., Zhang, W., Zhao, X., & Wang, J. (2019). Soft information in online peer-to-peer lending: Evidence from a leading platformin China. Electronic Commerce Research and Applications, 36, 1-15, https://doi.org/10.1016/j.elerap.2019.100873.

Welack, S.J.A.a.S., Artificial Neural Network Approach to Counterparty Credit Risk and XVA. 2019.

Wu, J. and Y.J.J. Xu, A Decision Support System for Borrower's Loan in P2P Lending. 2011. 6(6): p. 1183-1190.

Ye, X., et al., Loan evaluation in P2P lending based on random forest optimized by genetic algorithm with profit score. 2018. 32: p. 23-36.

Yu, S., & Fleming, L. (2021). Regional crowdfunding and high tech entrepreneurship. Research Policy, 1-18.https://doi.org/10.1016/j.respol.2021.104348

Zhang, H., The impact of distance, feature weighting and selection for KNN in credit default prediction. 2020.

Zhao, C., Li, M., Wang, J., & Ma, S. (2021). The mechanism of credit risk contagion among internet P2P lending platforms based on a SEIR model with time-lag. Research in International Business and Finance, 1-10. https://doi.org/10.1016/j.ribaf.2021.101407.

Financial Economics
Vol. (16) Issue (61) Spring 2027

Credit risk assessment in person-to-person lending using weighted nuclear regression methods and simple ranking

Habil Khavari
Accounting and finance Department, Ilam Branch, Islamic Azad University, Ilam, Iran (corresponding author).
habilkhavari@yahoo.com

Fatemeh Ahmadi
Accounting and finance Department, Ilam Branch, Islamic Azad University, Ilam, Iran.
fatemehahmady60@gmail.com

Mojtaba Moradpour
Accounting and finance Department, Ilam Branch, Islamic Azad University, Ilam, Iran.
Moradpour.mo@gmail.com

Rahmatollah Mohamadipour
Accounting and finance Department, Ilam Branch, Islamic Azad University, Ilam, Iran
mohammadipourrahmatollah@gmail.com

Reza Mashhadizadeh
Management Department, Karaj Branch, Islamic Azad University, Karaj, Iran.
r.mashhadizadeh@zarholding.com

Abstract
Person-to-person lending is another method of investment that eliminates the involvement of traditional financial institutions and has been growing more and more in recent years in other countries. The purpose of this article is to evaluate the credit risk based on the sample with the ability to evaluate the risk and return of each loan using logistic regression with Gaussian function that does not have a ranking approach in person-to-person lending. Lending Club's lending platform has been used to confirm and validate the effectiveness of the proposed method. that each borrower has a credit profile. The k-fold validation method has been used to determine the training and test samples. Therefore, according to ten thousand data samples, seven thousand random loan samples have been considered as training loan samples and the remaining three thousand loan samples have been considered as test loan samples. The results of the experiments show the better performance of the data-driven investment method based on nuclear regression compared to the ranking methods in the application of person-to-person lending.
Key words: credit risk, data-driven method, regression, rating, person-to-person lending.