



## رویکردهای یادگیری ماشین جهت پیش بینی ریسک اعتباری در حوزه بانکداری

مهرداد رضایی پور<sup>۱\*</sup>

۱- کارشناس ارشد مشاور سرمایه گذاری

### چکیده

مدیریت ریسک نکول اعتباری یکی از چالش‌های اساسی نظام بانکی است که مستقیماً بر ثبات مالی و کارایی تخصیص منابع اثر می‌گذارد. هدف این پژوهش، توسعه و مقایسه مدل‌های پیش‌بینی ریسک نکول و تبیین‌پذیری نتایج آن‌ها با تأکید بر شناسایی محرک‌های اصلی ریسک اعتباری است. داده‌های مورد استفاده شامل اطلاعات مشتریان اعتباری برگرفته از مجموعه داده German Credit بوده که پس از پیش‌پردازش، استخراج متغیرهای مشتق شده و شاخص‌های نوین فشار مالی، به مجموعه‌ای غنی از عوامل اقتصادی و رفتاری تبدیل شده است. در این مطالعه، سه رویکرد مدل‌سازی شامل رگرسیون لجستیک با جریمه LASSO، جنگل تصادفی و XGBoost به کار گرفته شد. ارزیابی عملکرد مدل‌ها با استفاده از معیار AUC نشان داد که مدل LASSO با مقدار AUC برابر با ۰.۷۷۶، جنگل تصادفی با ۰.۷۷۳ و XGBoost با ۰.۷۴۱ عملکرد مناسبی در تفکیک مشتریان خوش حساب و بدحساب دارند. اگرچه اختلاف عملکرد مدل‌ها محدود است، مدل‌های غیرخطی عملکرد بهتری در استخراج الگوهای پیچیده و غیرخطی از داده‌ها از خود نشان دادند. به منظور افزایش شفافیت و تفسیرپذیری، از روش SHAP برای تحلیل مدل جنگل تصادفی استفاده شد. نتایج SHAP نشان داد که شاخص‌هایی نظیر فشار بدهی، شدت تنش مالی، مبلغ اعتبار و مدت بازپرداخت بیشترین نقش را در پیش‌بینی نکول ایفا می‌کنند. همچنین جهت اثر متغیرها حاکی از آن است که افزایش فشار بازپرداخت و بار بدهی احتمال نکول را افزایش داده، در حالی که شاخص‌های نقدینگی و ثبات درآمد نقش محافظتی دارند.

واژگان کلیدی: ریسک اعتباری، یادگیری ماشین، جنگل تصادفی، XGBoost، تبیین‌پذیری مدل.



## مقدمه

در سال های اخیر، با رشد سریع فناوری های اطلاعاتی و افزایش پیچیدگی بازارهای مالی، پیش بینی دقیق ریسک اعتباری به یکی از چالش های اساسی بانک ها و مؤسسات مالی تبدیل شده است. تصمیمات نادرست در زمینه اعطای اعتبار می تواند منجر به افزایش مطالبات غیرجاری، کاهش سودآوری و حتی بروز ریسک های سیستمیک در نظام مالی شود. از این رو، مؤسسات مالی نیازمند ابزارهایی هستند که بتوانند با تحلیل داده های متنوع و پیچیده، ریسک ناشی از وام دهی و سرمایه گذاری را به طور دقیق تری برآورد کنند. در این راستا، مدل های آماری پیشرفته و الگوریتم های یادگیری ماشین به عنوان ابزارهایی کارآمد برای شناسایی الگوهای پنهان در داده های مالی و ارائه پیش بینی هایی با دقت بالا مورد توجه قرار گرفته اند.

مطالعات متعدد نشان می دهند که هوش مصنوعی و فناوری های نوین حسابداری می توانند نقش مهمی در بهبود کیفیت گزارش های مالی و ارتقای تصمیم گیری سرمایه گذاران ایفا کنند. برای مثال، ارائه یک چارچوب برای پذیرش هوش مصنوعی در گزارش دهی مالی نشان می دهد که درک مفید بودن و سهولت استفاده از سیستم های مبتنی بر هوش مصنوعی، تأثیر معناداری بر بهبود کیفیت اطلاعات مالی دارد (حیدریوسف و همکاران ۱۴۰۴). همچنین مطالعات دیگر بیان می کنند که به کارگیری فناوری های نوین حسابداری در بخش دولتی منجر به افزایش دقت و کاهش خطاهای گزارش های مالی می شود (Ahmad, et al. ۲۰۲۴).

افزون بر این، استفاده از الگوریتم های هوش مصنوعی در گزارشگری پایداری موجب افزایش معنادار دقت و کارآمدی تحلیل داده های مرتبط با شاخص های زیست محیطی و اجتماعی می شود (Ahmad et al., ۲۰۲۳). علاوه بر این، استفاده از ابزارهای مبتنی بر هوش مصنوعی در فرآیندهای گزارش دهی مالی، نقش معناداری در ارتقای کیفیت اطلاعات و تسهیل تصمیم گیری آگاهانه سرمایه گذاران ایفا می کند (Anantharaman et al., ۲۰۲۳). افزون بر این، به کارگیری تکنیک های هوش مصنوعی می تواند ضمن افزایش دقت اطلاعات مالی، از طریق کاهش عدم تقارن اطلاعاتی، ریسک تصمیم گیری سرمایه گذاران را نیز کاهش دهد (Hamidi et al., ۲۰۲۳).

هم زمان با این تحولات، ادغام داده های سنتی با داده های نوظهور، ظرفیت های جدیدی را در زمینه پیش بینی ریسک اعتباری فراهم کرده است. داده های نوظهور، از جمله اطلاعات دیجیتال، تراکنش های آنلاین و الگوهای رفتاری مشتریان در بسترهای الکترونیکی، امکان ارائه تصویری جامع تر و پویاتر از وضعیت اعتباری و رفتار مالی مشتریان را فراهم می سازند. شواهد تجربی حاکی از آن است که به کارگیری توأمان این داده ها در کنار مدل های آماری و الگوریتم های یادگیری ماشین، موجب بهبود معنادار دقت پیش بینی ریسک اعتباری و افزایش توان شناسایی مشتریان پرریسک می شود (Agarwal et al., ۲۰۲۰; Berg et al., ۲۰۲۰).



یادگیری ماشین، ماهیت «جعبه سیاه» آن هاست که می تواند پذیرش نتایج مدل را در محیط های نظارتی و تصمیم گیری مالی با محدودیت مواجه سازد.

در این راستا، تبیین پذیری مدل ها و رویکردهای هوش مصنوعی قابل توضیح به عنوان یکی از الزامات اساسی در کاربردهای مالی، به طور فزاینده ای مورد توجه قرار گرفته است. روش هایی مانند SHAP با تخصیص سهم هر متغیر در پیش بینی نهایی مدل، امکان بررسی هم زمان شدت و جهت اثر متغیرها بر خروجی مدل را فراهم می کنند و بدین ترتیب شفافیت تصمیمات مدل را به طور معناداری افزایش می دهند. بهره گیری از چنین چارچوب های تبیین پذیری، افزون بر تقویت اعتماد نهادهای نظارتی و مدیران مالی، زمینه استفاده عملی از نتایج مدل ها را در فرآیندهای تصمیم گیری اعتباری فراهم می سازد. (Lundberg et al. ۲۰۱۷)

بر این اساس، پژوهش حاضر با تمرکز بر ادغام مدل های آماری و الگوریتم های یادگیری ماشین، بهره گیری از داده های نوظهور و به کارگیری چارچوب های تبیین پذیر، تلاش می کند بخشی از خلأ های موجود در ادبیات ریسک اعتباری را پوشش دهد. نتایج این پژوهش می تواند به توسعه مدل های پیش بینی دقیق تر، شفاف تر و کاربردی تر در محیط های مالی کمک کرده و از طریق بهبود تصمیم گیری های اعتباری، کاهش ریسک سیستمیک و افزایش اعتماد سرمایه گذاران، نقش مؤثری در پیشرفت مطالعات حوزه فین تک و مدیریت ریسک ایفا کند.

#### ۱-۱. ریسک اعتباری و مدل های یادگیری ماشین

ریسک اعتباری<sup>۱</sup> به احتمال عدم ایفای تعهدات مالی توسط وام گیرنده اطلاق می شود و ارزیابی دقیق آن نقش اساسی در کاهش زیان ها و بهبود تصمیم گیری اعتباری دارد (Altman ۱۹۶۸). یکی از مهم ترین شاخص های کمی در این حوزه، احتمال نکول<sup>۲</sup> است که احتمال وقوع نکول در یک افق زمانی مشخص را نشان می دهد و در چارچوب های نظارتی بازل و رویه های مبتنی بر رتبه بندی داخلی (IRB) کاربرد گسترده ای دارد (Crouhy et al. ۲۰۰۶).

مدل های سنتی اعتبارسنجی عمدتاً بر داده های رسمی مالی متکی هستند؛ با این حال، محدودیت این داده ها به ویژه برای مشتریان دارای سابقه اعتباری کوتاه، موجب توجه به داده های نوظهور شده است (Vidal and Barbon ۲۰۱۹). هم زمان، پژوهش های اولیه نشان دادند که الگوریتم های یادگیری ماشین، از جمله شبکه های عصبی و روش های تکاملی، توانایی بالاتری در شناسایی روابط غیرخطی و پیچیده میان متغیرهای مالی و PD دارند (West ۲۰۰۰; Yobas et al. ۲۰۰۰). در سال های اخیر، روش های تجمیعی مانند Random Forest، AdaBoost و Gradient Boosting به طور گسترده برای افزایش دقت پیش بینی ریسک اعتباری استفاده شده اند (Moscato et al. ۲۰۲۱; Tripathi et al. ۲۰۲۲). مطالعه مرجع Lessmann با ارائه یک چارچوب مقایسه ای جامع نشان داد که بسیاری از الگوریتم های یادگیری ماشین عملکردی

<sup>۱</sup> Credit Risk

<sup>۲</sup> Probability of Default – PD



بهرتر از مدل های سنتی نظیر رگرسیون لجستیک دارند و مسیر تحقیقات بعدی را مشخص کرد (Lessmann, Baesens ۲۰۱۵). همچنین، مدل های ترکیبی و چندمرحله ای به عنوان راهکاری برای کاهش محدودیت های الگوریتم های منفرد توسعه یافته اند (Lee and Chen ۲۰۰۵; Dushimimana et al. ۲۰۲۰).

## ۱-۲. داده های نوظهور و فین تک

ظهور فناوری های مالی ۳ و پلتفرم های وام دهی دیجیتال، استفاده از داده های نوظهور مانند ردپاهای دیجیتال، داده های رفتاری موبایل و تعاملات آنلاین را امکان پذیر کرده است. این داده ها می توانند تصویری غنی تر از رفتار اعتباری مشتریان ارائه دهند و دقت پیش بینی ریسک اعتباری را به طور قابل توجهی افزایش دهند (Berg et al. ۲۰۲۰).

مطالعات نشان می دهد که ترکیب داده های نوظهور با الگوریتم های یادگیری ماشین، به ویژه در محیط های وام دهی همتا به هم (Serrano-Cinca and Gutiérrez-Nieto)، منجر به بهبود معنادار عملکرد مدل ها می شود (بنی هاشمی، علی مندگاری ۱۳۹۷). همچنین، شواهد تجربی حاکی از آن است که داده های رفتاری دیجیتال می توانند کمبود داده های مالی سنتی را جبران کرده و شمول مالی را تقویت کنند (Serrano-Cinca and Gutiérrez-Nieto ۲۰۱۶). این تحولات، چارچوب های جدیدی برای ارزیابی ریسک اعتباری در اکوسیستم فین تک فراهم کرده اند (Vallee and Zeng ۲۰۱۹).

## ۱-۳. تبیین پذیری و XAI در مدل های هوش مصنوعی

با افزایش استفاده از مدل های پیچیده یادگیری ماشین در حوزه های حساس مالی، مسئله تبیین پذیری و شفافیت تصمیمات مدل اهمیت ویژه ای یافته است. بسیاری از این مدل ها به عنوان «جعبه سیاه» شناخته می شوند و درک سازوکار تصمیم گیری آن ها برای نهادهای نظارتی و تصمیم گیران دشوار است. در این راستا، هوش مصنوعی تبیین پذیر<sup>۴</sup> به عنوان پاسخی به این چالش مطرح شده است (Adadi and Berrada ۲۰۱۸).

روش های مدل-ناوابسته مانند LIME و به ویژه SHAP، امکان تبیین سهم متغیرها در پیش بینی مدل را فراهم می کنند (Ribeiro, Singh et al. ۲۰۱۶). SHAP با اتکا بر نظریه بازی ها، هم تبیین محلی و هم تبیین جهانی ارائه می دهد و به یکی از پرکاربردترین ابزارهای XAI در مطالعات مالی تبدیل شده است (Lundberg and Lee ۲۰۱۷, Giudici and Raffinetti ۲۰۲۱). پژوهش ها نشان می دهند که استفاده از XAI می تواند ریسک مدل را کاهش داده، قابلیت ممیزی را افزایش دهد و اعتماد نهادهای نظارتی و مشتریان را تقویت کند (Bücker, Bussmann, Giudici et al. ۲۰۲۱, Szepannek et al. ۲۰۲۲).

<sup>۳</sup> FinTech

<sup>۴</sup> Explainable AI – XAI



#### ۴-۱. ریسک سیستمیک و شبکه های مالی

ریسک سیستمیک به خطر سرایت شوک ها در سراسر سیستم مالی اشاره دارد و در دهه های اخیر به یکی از محورهای اصلی تحقیقات مالی تبدیل شده است. شاخص هایی مانند CoVaR برای سنجش اثر متقابل مؤسسات مالی توسعه یافته اند (Acharya, Engle et al. ۲۰۱۲) (Tobias and Brunnermeier ۲۰۱۶). مطالعات نشان می دهد که وابستگی های شبکه ای و غیرخطی نقش مهمی در انتقال ریسک دارند (Adams, Füss et al. ۲۰۱۴, Abbassi, Brownlees et al. ۲۰۱۷).

پژوهش های جدیدتر با بهره گیری از مدل های شبکه ای و الگوریتم های یادگیری ماشین، امکان شناسایی الگوهای پنهان، خوشه های ریسک و نقاط شکننده سیستم مالی را فراهم کرده اند. این رویکردها نشان می دهند که ترکیب یادگیری ماشین و تحلیل شبکه ای می تواند دیدی جامع تر و پیش بینی پذیرتر از ریسک سیستمیک ارائه دهند (Giudici, Hadji-Misheva et al. ۲۰۲۰, Kumar, Rahman et al. ۲۰۲۴).

#### ۲-روش تحقیق

در این پژوهش، چارچوبی ترکیبی از آمار کلاسیک و یادگیری ماشین برای پیش بینی ریسک اعتباری و شناسایی ریسک های نوظهور بانکی ارائه شده است. این چارچوب شامل پنج لایه تحلیلی مکمل است که کل چرخه داده، مدل سازی، ارزیابی، تبیین و تصمیم گیری را پوشش می دهد:

۱. لایه تولید و آماده سازی داده ها<sup>۵</sup>
۲. لایه انتخاب ویژگی ها<sup>۶</sup>
۳. لایه مدل سازی آماری و یادگیری ماشین<sup>۷</sup>
۴. لایه ارزیابی آگاه از ریسک<sup>۸</sup>
۵. لایه تبیین پذیری و تحلیل ریسک<sup>۹</sup>

هدف اصلی این ساختار، ارائه پیش بینی دقیق احتمال نکول (Habibi tabar) همراه با شفافیت و قابلیت تبیین برای مدیران ریسک و نهادهای نظارتی است. استفاده از روش های تبیین پذیر مانند مقادیر SHAP امکان شناسایی محرک های اصلی ریسک و تحلیل منطق تصمیم گیری مدل را فراهم می سازد.

<sup>۵</sup> Data Layer

<sup>۶</sup> Feature Selection Layer

<sup>۷</sup> Modeling Layer

<sup>۸</sup> Risk-aware Evaluation Layer

<sup>۹</sup> Explainability & Risk Analysis Layer



## ۲-۱. داده‌ها و انتخاب ویژگی‌ها

داده‌های مورد استفاده شامل مجموعه German Credit Data منتشر شده توسط مخزن UCI است که شامل ۱۰۰۰ مشاهده، ۲۰ متغیر توضیحی و یک متغیر هدف (نکول/عدم نکول) ۱۰ می‌باشد. این داده‌ها شامل اطلاعات اعتباری، مالی و فردی وام‌گیرندگان بوده و با چارچوب‌های استاندارد ارزیابی احتمال نکول در رویکردهای مبتنی بر رتبه‌بندی داخلی (IRB) هم‌راستا هستند.

انتخاب ویژگی‌ها در دو سطح انجام شد: نخست، متغیرهای اصلی اعتباری شامل مشخصات وام، سابقه اعتباری، توان بازپرداخت ۱۱ و ویژگی‌های فردی؛ که انتخاب آن‌ها مبتنی بر ادبیات کلاسیک ریسک اعتباری و مطالعاتی نظیر Thomas et al. (Kleinbaum, Klein et al.) و Lessmann et al. (Ahmad, Shamsuddin et al.) است.

### جدول ۱. متغیرهای ورودی مدل ریسک اعتباری

نقش در ریسک	تعریف عملیاتی	متغیر	دسته
شدت تعهد مالی	کل مبلغ تسهیلات	مبلغ وام	مشخصات وام
افق ریسک	تعداد ماه	مدت بازپرداخت	مشخصات وام
ریسک رفتاری	رفتار گذشته	سابقه اعتباری	سابقه اعتباری
توان بازپرداخت	فشار نقدینگی	نسبت قسط به درآمد	توان بازپرداخت
ضربه‌گیر شوک	سطح ذخایر مالی	پس‌انداز	وضعیت مالی
ناهمگنی ریسک	چرخه عمر	سن	ویژگی فردی
متغیر وابسته	نکول/عدم نکول	نکول	برچسب

جدول ۱ متغیرهای اصلی ورودی و نیز شاخص‌های مشتق شده منتخب را نشان می‌دهد.

<sup>۱۰</sup> Modeling Layer Default (Target)

<sup>۱۱</sup> Installment Rate



در سطح دوم، مجموعه‌ای از ویژگی‌های مشتق شده به منظور شناسایی ریسک‌های نوظهور طراحی شد؛ از جمله شاخص فشار بدهی<sup>۱۲</sup>، شاخص شکنندگی مالی<sup>۱۳</sup>، تراکم اعتباری<sup>۱۴</sup> و شاخص تنش کوتاه‌مدت<sup>۱۵</sup>. این شاخص‌ها با هدف استخراج سیگنال‌های زودهنگام فشار مالی و الگوهای پرریسک ساخته شده‌اند.

## جدول ۲. شاخص‌های مشتق شده منتخب برای شناسایی ریسک‌های نوظهور

نوع ریسک	منطق محاسبه	شاخص
فشار بدهی	مبلغ وام / مدت	شاخص فشار بدهی
ریسک زودهنگام	قسط × مبلغ	شاخص شکنندگی مالی
بیش‌اعتباری	اعتبارات فعال / سابقه شغلی	شدت اعتباری
ریسک نقدینگی	وضعیت پس‌انداز + حساب	شاخص نقدینگی
هشدار زودهنگام	قسط / مدت	شاخص تنش کوتاه‌مدت

## ۲-۲. لایه مدل‌سازی آماری و یادگیری ماشین

در این لایه، مدل‌های مختلف با هدف پیش‌بینی PD و مقایسه عملکرد توسعه داده شدند. از یک سو، مدل‌های آماری کلاسیک شامل رگرسیون لجستیک و لاسو<sup>۱۶</sup> به‌عنوان مدل‌های پایه و قابل تفسیر مورد استفاده قرار گرفتند. بطوریکه لاسو با منظم‌سازی، مسئله هم‌خطی و تفکیک کامل<sup>۱۷</sup> داده‌ها را کنترل می‌کند. از سوی دیگر، الگوریتم‌های یادگیری ماشین شامل درخت تصمیم، جنگل تصادفی و XGBoost برای مدل‌سازی روابط غیرخطی و تعاملات پیچیده میان متغیرها به کار گرفته شدند.

فرآیند مدل‌سازی شامل تقسیم داده‌ها به مجموعه‌های آموزش و آزمون، پیش‌پردازش داده‌ها، آموزش مدل‌ها و تنظیم پارامترها است. ارزیابی عملکرد مدل‌ها بر اساس معیارهایی نظیر AUC-ROC، F1-Score، Accuracy و آماره KS انجام شد و مدل نهایی با توجه به دقت، پایداری و قابلیت تبیین انتخاب گردید.

<sup>۱۲</sup> Debt Burden Index

<sup>۱۳</sup> Financial Stress Indicator

<sup>۱۴</sup> Credit Intensity

<sup>۱۵</sup> Short-Term Stress Index

<sup>۱۶</sup> LASSO

<sup>۱۷</sup> Complete Separation



### ۳-۲. لایه ارزیابی آگاه از ریسک:

ارزیابی مدل‌ها صرفاً به معیارهای متداول یادگیری ماشین محدود نشد، بلکه شاخص‌های ریسک‌محور نیز مدنظر قرار گرفت. این شاخص‌ها شامل توزیع PD، تحلیل زیان مورد انتظار و بررسی رفتار مدل در نواحی پرریسک است. نخست، توزیع احتمال نکول (Habibi tabar) برای بررسی نحوه تفکیک وام‌گیرندگان در سطوح مختلف ریسک تحلیل شد؛ به‌گونه‌ای که توان مدل در شناسایی مشتریان پرریسک و خوشه‌بندی سطوح ریسک به‌صورت دقیق ارزیابی گردد. سپس، زیان مورد انتظار به‌عنوان یکی از شاخص‌های کلیدی مدیریت ریسک اعتباری، با ترکیب مؤلفه‌های PD، نرخ زیان در صورت نکول و میزان قرارگیری در معرض ریسک مورد توجه قرار گرفت تا پیامدهای مالی پیش‌بینی‌های مدل به‌صورت کمی قابل ارزیابی باشد.

در ادامه، رفتار مدل در نواحی پرریسک توزیع با استفاده از شاخص‌های Tail Risk (شامل Bastani, Value at Risk) و (Asgari et al.) Conditional Value at Risk (CVaR) بررسی شد. این تحلیل امکان ارزیابی پایداری مدل در شرایط بحرانی و سناریوهای بدبینانه را فراهم می‌سازد؛ شرایطی که در آن‌ها خطای پیش‌بینی می‌تواند منجر به زیان‌های قابل توجه بانکی شود. در نهایت، آزمون‌های تنش رای شبیه‌سازی شوک‌های اعتباری و ارزیابی واکنش مدل نسبت به تغییرات شدید متغیرهای کلیدی انجام شد. این رویکرد ارزیابی چندبعدی، امکان مقایسه مدل‌ها را نه تنها از منظر دقت آماری، بلکه از دیدگاه کارایی عملی، پایداری ریسک و قابلیت اتکای تصمیم‌گیری در محیط‌های واقعی بانکی فراهم می‌کند.

### ۴-۲. لایه تبیین‌پذیری و تحلیل ریسک

به‌منظور افزایش شفافیت تصمیمات مدل‌های غیرخطی و اطمینان‌پذیری کاربردهای بانکی از روش‌های تبیین‌پذیر شامل مقادیر SHAP و نمودارهای وابستگی جزئی برای تحلیل اثر واقعی متغیرها و تعاملات آن‌ها بر احتمال نکول استفاده شد. این رویکرد امکان شناسایی متغیرهای کلیدی، جهت و شدت اثر آن‌ها بر ریسک اعتباری، خوشه‌بندی وام‌گیرندگان بر اساس پروفایل ریسک و شناسایی زود هنگام ریسک‌های نوظهور را فراهم می‌سازد. بدین ترتیب، قدرت پیش‌بینی مدل‌های یادگیری ماشین با قابلیت تبیین و الزامات نظارتی ترکیب شده و تصمیم‌گیری اعتباری به سطحی شفاف و قابل اتکا ارتقا می‌یابد.

### ۳. یافته‌ها:

#### ۳-۱. تحلیل توصیفی داده‌ها:

در این بخش، آمار توصیفی مجموعه‌ای از متغیرهای کلیدی مورد استفاده در مدل‌های پیش‌بینی احتمال نکول ارائه می‌شود تا تصویری کلی از ساختار داده‌ها و ناهمگنی ریسک در نمونه پژوهش فراهم گردد. تمرکز بر متغیرهایی است



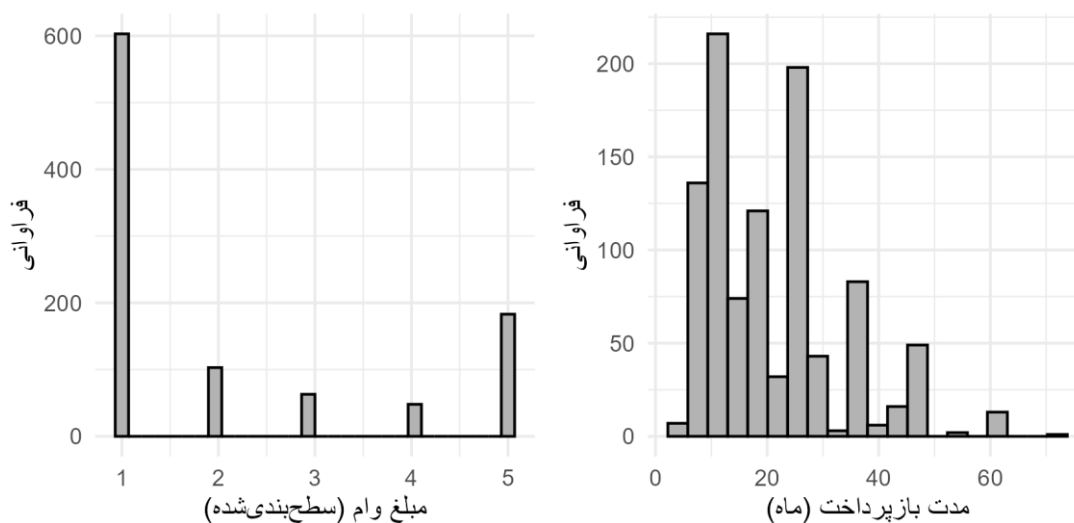
که در ادبیات ریسک اعتباری به عنوان محرک های اصلی نکول شناخته می شوند (Brown & Thomas et al., ۲۰۰۲; Mues, ۲۰۱۲).

با توجه به ماهیت کدگذاری شده و دسته ای بخش قابل توجهی از متغیرهای مجموعه داده German Credit، آمار توصیفی صرفاً برای متغیرهای پیوسته و ترتیبی با تفسیر اقتصادی مستقیم شامل مدت بازپرداخت، مبلغ وام، نرخ تعهد اقساط و سن وام گیرنده گزارش شده است.

جدول ۲. آمار توصیفی متغیرهای منتخب

متغیر	تعداد	میانگین	انحراف معیار	مینیمم	ماکسیمم
مبلغ وام (کدگذاری شده)	۱۰۰۰	۲.۱۰	۱.۵۸	۱	۵
مدت بازپرداخت (ماه)	۱۰۰۰	۲۰.۹۰	۱۲.۰۶	۴	۷۲
نرخ تعهد اقساط	۱۰۰۰	۲.۸۵	۱.۱۰	۱	۴
سن (کدگذاری شده)	۱۰۰۰	۱.۱۶	۰.۳۶	۱	۲

جدول ۲ نشان می دهد که متغیر مدت بازپرداخت از پراکندگی قابل توجهی برخوردار است که بیانگر ناهمگنی افقی های زمانی بازپرداخت در میان وام گیرندگان می باشد. همچنین، توزیع متغیرهای ترتیبی مرتبط با مبلغ وام، نرخ اقساط و سن، وجود سطوح متفاوتی از فشار مالی، توان بازپرداخت و ریسک رفتاری را در نمونه مورد بررسی منعکس می کند. این ناهمگنی، ضرورت استفاده از مدل های انعطاف پذیر و غیرخطی برای تحلیل ریسک اعتباری و ریسک های نوظهور را تقویت می کند.





### شکل ۱. توزیع مبلغ وام و مدت بازپرداخت در داده‌های اعتباری

شکل ۱ توزیع متغیرهای کلیدی مبلغ وام و مدت بازپرداخت را نشان می‌دهد. پراکندگی نسبتاً بالای متغیر مدت بازپرداخت بیانگر ناهمگنی قابل توجه در افق‌های زمانی بازپرداخت وام‌گیرندگان است که می‌تواند منجر به تفاوت در تجمع ریسک نکول در دوره‌های مختلف شود. از سوی دیگر، متغیر مبلغ وام که به صورت ترتیبی کدگذاری شده است، وجود سطوح متفاوتی از ریسک اعتباری را در میان مشتریان منعکس می‌کند و نشان‌دهنده تمایز میان وام‌های با فشار بدهی پایین تا بالا می‌باشد. این الگوها مؤید ضرورت استفاده از مدل‌های انعطاف‌پذیر و غیرخطی برای شناسایی ناهمگنی ریسک و تحلیل ریسک‌های نوظهور بانکی هستند.

### ۲-۳. مقایسه عملکرد مدل‌های پیش‌بینی ریسک اعتباری

در این بخش، عملکرد مدل‌های آماری کلاسیک و الگوریتم‌های یادگیری ماشین در پیش‌بینی احتمال نکول مورد مقایسه قرار می‌گیرد. هدف اصلی، ارزیابی توان تمایز مدل‌ها، پایداری پیش‌بینی‌ها و قابلیت کاربرد آن‌ها در چارچوب مدیریت ریسک اعتباری است. منحنی گیرنده-عملیاتی (Breiman) بیانگر رابطه بین نرخ کشف صحیح نکول<sup>۱۸</sup> و نرخ خطای هشدار کاذب<sup>۱۹</sup> در سطوح مختلف آستانه تصمیم‌گیری است. مقدار سطح زیر منحنی ROC به عنوان یک معیار مستقل از آستانه، توان مدل را در رتبه‌بندی وام‌گیرندگان بر اساس ریسک نکول نشان می‌دهد؛ به گونه‌ای که مقادیر بالاتر AUC بیانگر قدرت تفکیک بهتر بین مشتریان نکول‌کننده و غیرنکول‌کننده هستند. در ادبیات مدیریت ریسک، مقادیر AUC بالاتر از ۰.۷۵ به عنوان عملکرد مناسب برای کاربردهای عملی نظیر رتبه‌بندی اعتباری، تخصیص سرمایه و سیستم‌های هشدار زودهنگام تلقی می‌شوند.

جدول ۳. عملکرد مدل‌ها بر اساس معیار AUC-ROC

مدل	مساحت زیر منحنی ROC
-----	---------------------

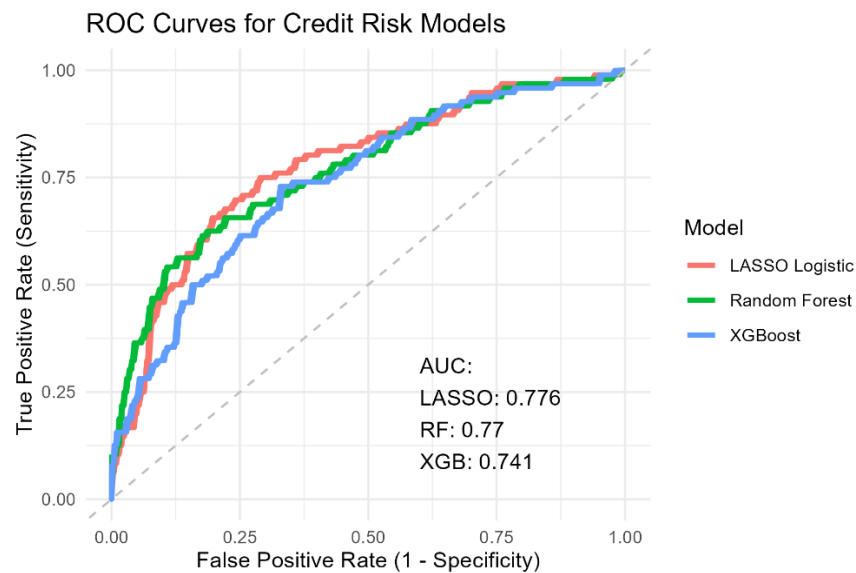
<sup>۱۸</sup> True Positive Rate

<sup>۱۹</sup> False Positive Rate



۰,۷۷۶	رگرسیون لجستیک لاسو
۰,۷۷۳	جنگل تصادفی
۰,۷۴۱	<b>XGBoost</b>

جدول ۳ عملکرد سه مدل اصلی شامل رگرسیون لجستیک لاسو، جنگل تصادفی و XGBoost را بر اساس معیار-AUC ROC گزارش می کند. نتایج بیانگر آن است که مدل رگرسیون لجستیک لاسو با مقدار AUC برابر با ۰.۷۷۶، بالاترین توان تفکیک بین وام گیرندگان نکول کننده و غیرنکول کننده را ارائه می دهد. این یافته نشان می دهد که در حضور ویژگی های مشتق شده و کنترل هم خطی متغیرها، مدل رگرسیون لاسو همچنان می تواند عملکردی رقابتی و قابل اتکا داشته باشد؛ موضوعی که با نتایج مطالعات پیشین هم راستا است. (Lessmann et al., ۲۰۱۵)



شکل ۲. منحنی ROC برای رتبه بندی وام گیرندگان

شکل ۲ منحنی های ROC مربوط به سه مدل رگرسیون لجستیک لاسو، جنگل تصادفی و XGBoost را نشان می دهد. مقادیر AUC بیانگر توان تمایز مدل ها در تفکیک نکول کنندگان از غیرنکول کنندگان هستند و نشان می دهند که هر سه مدل از قدرت پیش بینی قابل قبولی برخوردارند. نمودار ROC امکان ارزیابی عملکرد مدل ها را از سه منظر فراهم می کند: توان تفکیک کلی از طریق AUC، رفتار مدل ها در آستانه های مختلف تصمیم گیری، و تعیین نقطه بهینه تصمیم گیری عملی با استفاده از شاخص Youden که موازنه مناسبی میان حساسیت و ویژگی ایجاد می کند.



هرچند ROC صرفاً عملکرد کلی مدل‌ها را نشان می‌دهد، تفاوت میان مدل‌های آماری و یادگیری ماشین را می‌توان با توجه به ویژگی‌های ساختاری الگوریتم‌ها تبیین کرد. به‌طور خاص، جنگل تصادفی و XGBoost به دلیل توانایی در مدل‌سازی روابط غیرخطی و تعاملات پیچیده میان متغیرها، ظرفیت بالاتری برای شناسایی الگوهای رفتاری و ریسک‌های نوظهور دارند (Chen, ۲۰۱۶; Breiman, ۲۰۰۱).

### ۳-۳. تبیین‌پذیری مدل و تحلیل ریسک‌های نوظهور

#### ۳-۳-۱. تحلیل تبیینی مبتنی بر SHAP

به‌منظور تبیین تصمیمات مدل جنگل تصادفی و شناسایی عوامل کلیدی مؤثر بر احتمال نکول، از چارچوب SHAP استفاده شد. تحلیل SHAP در این پژوهش به‌صورت تبیین‌پذیری کلی و مبتنی بر نمونه‌گیری از کل داده‌ها انجام شده است؛ به‌طوری که میانگین قدر مطلق اثر SHAP<sup>۲۰</sup> به‌عنوان شاخص شدت اثر متغیرها و میانگین SHAP به‌عنوان شاخص جهت اثر آن‌ها بر ریسک نکول محاسبه شد. (Lundberg, Lee, ۲۰۱۷)

#### جدول ۴. مهم‌ترین متغیرهای اثرگذار بر خروجی مدل جنگل تصادفی بر اساس مقادیر SHAP

متغیر	جهت (اثر میانگین SHAP)	شدت اثر (قدر مطلق میانگین SHAP)
فشار بدهی بلند مدت	-۰.۰۵۳	۰.۰۵۳
شاخص فشار مالی	-۰.۰۳۶	۰.۰۳۶
دوره باز پرداخت وام	-۰.۰۳۶	۰.۰۳۶
شاخص استرس کوتاه مدت	-۰.۰۳۰	۰.۰۳۰
بدهکاران کوتاه مدت	۰.۰۳۰	۰.۰۳۰
سابقه اعتباری	-۰.۰۲۶	۰.۰۲۶
وضعیت شخصی/جنسیت	-۰.۰۲۳	۰.۰۲۳
وضعیت حساب جاری	۰.۰۲۳	۰.۰۲۳
مبلغ اعطا وام	-۰.۰۲۲	۰.۰۲۲

<sup>۲۰</sup> SHAP (Mean |SHAP|)



۰.۰۱۹	۰.۰۱۹	وضعیت حساب جاری (عددی)
-------	-------	------------------------

نتایج جدول ۴ نشان می‌دهد که متغیرهای وضعیت حساب جاری، مدت بازپرداخت و سابقه اعتباری، بالاترین میانگین قدر مطلق اثر SHAP را داشته و به‌عنوان محرک‌های اصلی ریسک نکول شناسایی می‌شوند. همچنین، علامت میانگین جهت اثر SHAP بیانگر نقش افزایشی یا کاهش‌ی این متغیرها بر احتمال نکول در سطح کل داده‌ها است.

از منظر تفسیر بانکی، وضعیت حساب جاری به‌عنوان شاخص نقدینگی کوتاه‌مدت و توان بازپرداخت، بیشترین سهم را در تغییرات پیش‌بینی احتمال نکول دارند. همچنین، دوره‌های بازپرداخت طولانی و سابقه اعتباری ضعیف، اثر معناداری بر افزایش احتمال نکول دارند؛ یافته‌ای که با ادبیات کلاسیک ریسک اعتباری هم‌خوانی دارد (Altman, ۱۹۶۸; Thomas et al., ۲۰۰۲).

## ۵. بحث و تفسیر نتایج

### ۵-۱. چرایی تمرکز تحلیل SHAP بر مدل جنگل تصادفی

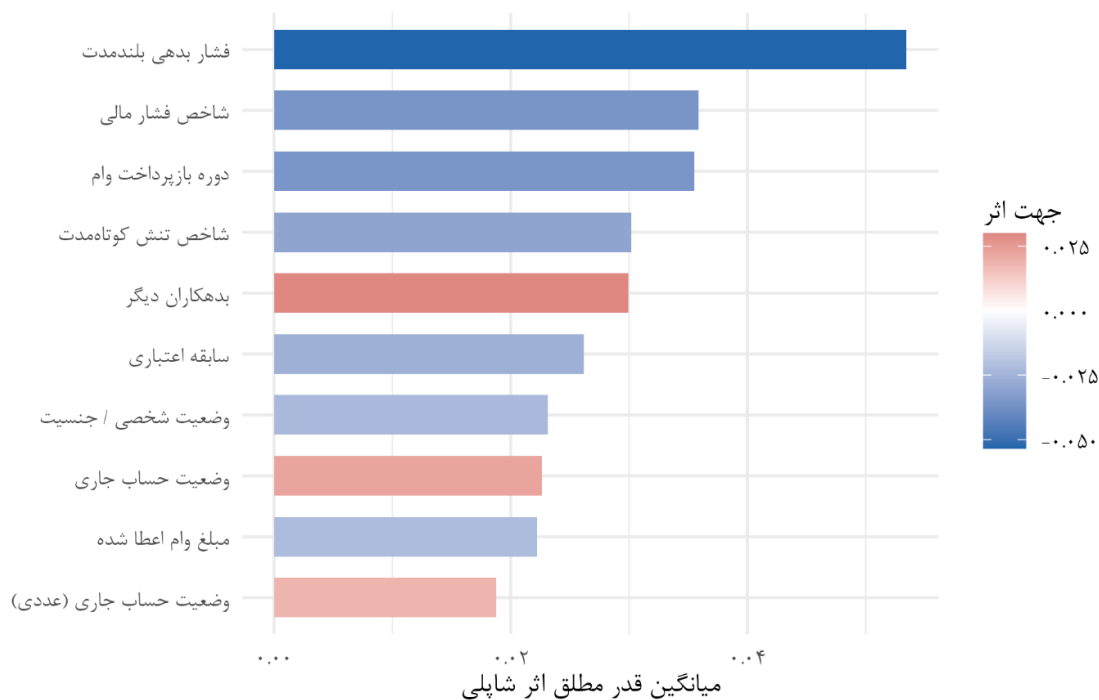
اگرچه در این پژوهش سه مدل شامل رگرسیون لجستیک لاسو، جنگل تصادفی و XGBoost توسعه داده شد، تحلیل تبیین‌پذیری مبتنی بر SHAP صرفاً برای مدل جنگل تصادفی ارائه گردید. دلیل این انتخاب، ماهیت تعادلی جنگل تصادفی میان دقت پیش‌بینی و پایداری تبیین‌پذیری است. برخلاف مدل لاسو که ساختاری خطی و ذاتاً قابل تفسیر دارد، مدل‌های درختی غیرخطی نیازمند ابزارهای XAI برای استخراج منطق تصمیم‌گیری هستند. از سوی دیگر، اگرچه XGBoost نیز مدلی قدرتمند است، اما ساختار تقویتی و حساسیت بالاتر آن به تنظیمات پارامترها، موجب نوسان بیشتر مقادیر SHAP و کاهش پایداری تفسیر در سطح کل داده‌ها می‌شود. مطالعات پیشین نشان داده‌اند که جنگل تصادفی در مقایسه با مدل‌های Boosting، تفاسیر SHAP پایدارتر و قابل‌اعتمادتری در کاربردهای بانکی و نظارتی ارائه می‌دهد (Lundberg et al., ۲۰۱۷; Bussmann et al., ۲۰۲۱; Bucker et al., ۲۰۲۲). بنابراین، تمرکز بر جنگل تصادفی امکان ارائه تبیینی شفاف، قابل اتکا و سازگار با الزامات نظارتی را فراهم می‌سازد.

### ۵-۲. محرک‌های اصلی ریسک نکول (تفسیر اقتصادی متغیرهای کلیدی)

نتایج تحلیل SHAP نشان می‌دهد که شاخص‌های مشتق‌شده فشار مالی مهم‌ترین نقش را در پیش‌بینی ریسک نکول ایفا می‌کنند. بر اساس میانگین قدر مطلق مقادیر SHAP، شاخص فشار بدهی بلندمدت قوی‌ترین محرک نکول شناسایی شد و پس از آن، شاخص فشار کوتاه‌مدت، مدت بازپرداخت وام و شاخص تنش مالی قرار دارند. این یافته نشان می‌دهد که ابعاد پنهان فشار مالی و شکنندگی بازپرداخت، نقش تعیین‌کننده‌تری نسبت به متغیرهای کلاسیک در رفتار نکول دارند.



شکل ۲ نمودار توزیع اثر این متغیرها را بر خروجی مدل نشان می‌دهد. تمرکز جرم نقاط و شدت رنگ‌ها بیانگر آن است که با افزایش فشار بدهی و تنش مالی، احتمال نکول به‌طور غیرخطی افزایش می‌یابد. طول میله‌ها نشان‌دهنده اهمیت متغیرها و رنگ آن‌ها بیانگر جهت اثر متغیر بر ریسک نکول است (قرمز: اثر افزایشی، آبی: اثر کاهش). این حساسیت بالا نسبت به متغیرهای مشتق شده نشان می‌دهد که الگوریتم جنگل تصادفی، در مقایسه با مدل‌های آماری کلاسیک، ظرفیت تحلیلی بالاتری برای کشف ریسک‌های نوظهور و رفتارهای اعتباری پیچیده دارد. (Giudici, Raffinetti, ۲۰۲۱)



شکل ۳. نمودار توزیع اثر برای ۱۰ متغیر برتر در مدل جنگل تصادفی

در کنار این متغیرها، مدت بازپرداخت وام به‌عنوان شاخص فشار مالی بلندمدت، و وضعیت حساب جاری به‌عنوان نمایانگر نقدینگی کوتاه‌مدت، نقش مکملی در پیش‌بینی نکول دارند. دوره‌های بازپرداخت طولانی‌تر و وضعیت نامطلوب حساب جاری، انعطاف‌پذیری مالی مشتری را کاهش داده و احتمال نکول را افزایش می‌دهند.

سابقه اعتباری و مبلغ وام اگرچه اثر معناداری دارند، اما شدت تأثیر آن‌ها نسبت به شاخص‌های مشتق شده فشار مالی کمتر است. این نتیجه نشان می‌دهد که مدل جنگل تصادفی، فراتر از شاخص‌های سنتی، قادر به شناسایی الگوهای



پیچیده و غیرخطی ریسک نکول و ریسک های نوظهور اعتباری است؛ یافته ای که با ادبیات نوین ریسک اعتباری هم خوانی دارد. (Altman, ۱۹۶۸; Thomas, ۲۰۱۷)

از منظر اقتصادی و بانکی، وضعیت حساب جاری نمایانگر نقدینگی کوتاه مدت و جریان وجوه در دسترس وام گیرنده است. مشتریانی با وضعیت نامطلوب حساب جاری معمولاً توان کمتری در مدیریت شوک های نقدینگی دارند و در مواجهه با نوسانات درآمدی، سریع تر به نکول سوق داده می شوند. مدت بازپرداخت وام به عنوان شاخصی از فشار مالی بلندمدت عمل می کند؛ دوره های بازپرداخت طولانی تر می توانند منجر به انباشت بدهی، افزایش حساسیت به تغییرات اقتصادی و کاهش انعطاف پذیری مالی شوند. سابقه اعتباری نیز پروکسی مستقیم رفتار مالی گذشته وام گیرنده است و نتایج این پژوهش هم راستا با ادبیات کلاسیک نشان می دهد که سابقه ضعیف بازپرداخت، احتمال نکول آینده را به طور معناداری افزایش می دهد. (Altman, ۱۹۶۸, Thomas, ۲۰۱۷)

## ۶. نتیجه گیری

نتایج حاصل از تحلیل مدل های پیش بینی ریسک اعتباری نشان می دهد که همه مدل ها، شامل رگرسیون لجستیک لاسو، جنگل تصادفی و XGBoost، توانایی قابل قبولی در تفکیک وام گیرندگان نکول کننده و غیر نکول کننده دارند، اما تفاوت های ساختاری و عملکردی آنها قابل توجه است. مدل رگرسیون لجستیک لاسو با AUC برابر با ۰.۷۷۶ بالاترین توان تفکیک را ارائه داد، که نشان دهنده مزیت استفاده از مدل های آماری منظم شده در حضور ویژگی های مشتق شده و کنترل هم خطی متغیرها است. این یافته با نتایج مطالعات پیشین در زمینه رتبه بندی اعتباری و پیش بینی نکول هم راستا است. (Lessmann et al., ۲۰۱۵)

از سوی دیگر، مدل های درختی مانند جنگل تصادفی و XGBoost به دلیل توانایی مدل سازی روابط غیرخطی و تعاملات پیچیده میان متغیرها، ظرفیت بالاتری برای شناسایی الگوهای رفتاری نوظهور و ریسک های مخفی دارند. تحلیل های مبتنی بر SHAP نشان داد که جنگل تصادفی به طور خاص توانایی بالایی در شناسایی محرک های اصلی ریسک و الگوهای نوظهور دارد و نسبت به XGBoost پایداری تبیین پذیری بالاتری ارائه می دهد. حساسیت بالای XGBoost به تنظیمات پارامترها و ساختار تقویتی آن موجب نوسان بیشتر مقادیر SHAP و کاهش پایداری تفسیر در سطح کل داده ها می شود، که محدودیت مهمی در کاربرد عملی و گزارش به مدیران ریسک به شمار می رود.

از منظر اقتصادی و عملیاتی، نتایج تحلیل SHAP نشان می دهد که متغیرهایی مانند وضعیت حساب جاری، مدت بازپرداخت و سابقه اعتباری بیشترین نقش را در پیش بینی ریسک نکول دارند. این متغیرها نمایانگر نقدینگی کوتاه مدت، فشار مالی بلندمدت و رفتار مالی گذشته وام گیرندگان هستند و تأثیر معناداری بر احتمال نکول دارند. شاخص های مشتق شده مانند فشار بدهی و تنش مالی نیز به شناسایی زود هنگام ریسک های نوظهور کمک می کنند و نشان می دهند



که مدل های یادگیری ماشین قادر به کشف روابط غیر خطی و پیچیده در داده های اعتباری هستند که مدل های آماری کلاسیک کمتر قادر به استخراج آنها هستند.

با توجه به یافته ها، ترکیب مدل های آماری تبیین پذیر با الگوریتم های یادگیری ماشین، تعادلی مؤثر میان دقت پیش بینی، شفافیت و قابلیت کاربرد عملی در محیط بانکی ایجاد می کند. این ترکیب به مدیران ریسک امکان می دهد که نه تنها تصمیمات مبتنی بر داده اتخاذ کنند، بلکه منطق پشت پیش بینی های مدل را نیز درک کنند و ریسک های نوظهور را به موقع شناسایی نمایند.

در نهایت این پژوهش نشان داد که جنگل تصادفی تعادل مطلوبی میان دقت پیش بینی و تبیین پذیری ارائه می دهد و توانایی شناسایی محرک های اصلی و ریسک های نوظهور را دارد. متغیرهایی مانند وضعیت حساب جاری، مدت بازپرداخت و سابقه اعتباری بیشترین نقش را در پیش بینی نکول دارند و شاخص های مشتق شده مانند فشار بدهی و تنش مالی به شناسایی زود هنگام ریسک ها کمک می کنند.

کاربردهای عملی شامل رتبه بندی اعتباری مشتریان، طراحی سیستم های هشدار زود هنگام و تخصیص بهینه سرمایه است.

پیشنهادات برای تحقیقات آینده:

۱. استفاده از داده های گسترده تر و متنوع شامل تراکنش های مشتریان،
۲. توسعه مدل های ترکیبی برای افزایش دقت و پایداری،
۳. بررسی شاخص ها و ویژگی های مشتق شده جدید برای شناسایی ریسک های نوظهور،
۴. تحلیل عملکرد مدل ها در شرایط بحرانی یا شوک های اقتصادی.

## منابع

- Abbassi, P., C. Brownlees, C. Hans and N. Podlich (۲۰۱۷). "Credit risk interconnectedness: What does the market really know?" *Journal of Financial Stability* ۲۹: ۱-۱۲
- Acharya, V., R. Engle and M. Richardson (۲۰۱۲). "Capital shortfall: A new approach to ranking and regulating systemic risks." *American Economic Review* ۱۰۲(۳): ۵۹-۶۴.
- Adadi, A. and M. Berrada (۲۰۱۸). "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." *IEEE access* ۶: ۵۲۱۳۸-۵۲۱۶۰.
- Adams, Z., R. Füss and R. Gropp (۲۰۱۴). "Spillover effects among financial institutions: A state-dependent sensitivity value-at-risk approach." *Journal of Financial and Quantitative Analysis* ۴۹(۳): ۵۷۵-۵۹۸.
- enAgarwal, S., S. Alok, P. Ghosh and S. Gupta (۲۰۲۰). "Financial inclusion and alternate credit scoring for the millennials: role of big data and machine learning in fintech." *Business School,*



- National University of Singapore Working Paper, SSRN ۳۰۷۸۲۷.
- Aghaei, A. M., J. Gholami, A. Sangchooli, Y. Rostam-Abadi, S. Olamazadeh, M. Ardeshir, S. Baheshmat, B. Shadloo, M. Taj, K. Saeed and A. Rahimi-Movaghar (۲۰۲۳). "Prevalence of injecting drug use and HIV, hepatitis B, and hepatitis C in people who inject drugs in the Eastern Mediterranean region: a systematic review and meta-analysis". *Lancet Glob Health* ۱۱(۸): e۱۲۲۵-e۱۲۳۷.
  - Ahmad, A., H. Abusaimah, A. Rababah, M. Alqsass, N. Al-Olima and M. Hamdan (۲۰۲۴). "Assessment of effects in advances of accounting technologies on quality financial reports in Jordanian public sector." *Uncertain Supply Chain Management* ۱۲(Ratkowsky, Lowry, McMeekin, Stokes, & Chandler): ۱۳۳-۱۴۲.
  - Ahmad, S., N. R. Shamsuddin and H. Farizad (۲۰۱۵). "Risks of divorce: comparison between Cox and Parametric Models".
  - Altman, E. I. (۱۹۶۸). "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy." *The journal of finance* ۲۳(Ratkowsky et al.): ۵۸۹-۶۰۹.
  - Amin, M. S. M. (۲۰۰۲). *Synthesis and Characterization of New Chelating Resins Containing Thiourea as the Functional Group*, Universiti Putra Malaysia.
  - Anantharaman, D., Rozario, A., & Parker, C. A. (۲۰۲۳). Artificial intelligence and financial reporting quality. *Available at SSRN 4625279*.
  - Bastani, K., E. Asgari and H. Namavari (۲۰۱۹). "Wide and deep learning for peer-to-peer lending." *Expert Systems with Applications* ۱۳۴: ۲۰۹-۲۲۴.
  - Berg, T., V. Burg, A. Gombović and M. Puri (۲۰۲۰). "On the rise of fintechs: Credit scoring using digital footprints." *The Review of Financial Studies* ۳۳(Ratkowsky et al.): ۲۸۴۵-۲۸۹۷.
  - Bracke, P., A. Datta, C. Jung and S. Sen (۲۰۱۹). "Machine learning explainability in finance: an application to default risk analysis".
  - Breiman, L. (۲۰۰۱). "Random Forests." *Machine Learning* ۴۵(Ratkowsky et al.): ۵-۳۲.
  - Bücker, M., G. Szepannek, A. Gosiewska and P. Biecek (۲۰۲۲). "Transparency, auditability, and explainability of machine learning models in credit scoring." *Journal of the Operational Research Society* ۷۳(Ratkowsky et al.): ۷۰-۹۰.
  - Bussmann, N., P. Giudici, D. Marinelli and J. Papenbrock (۲۰۲۱). "Explainable machine learning in credit risk management." *Computational Economics* ۵۷(Ratkowsky et al.): ۲۰۳-۲۱۶.
  - Crouhy, M., D. Galai and R. Mark (۲۰۰۶). *The essentials of risk management*, McGraw-Hill New York.
  - Dushimimana, B., Y. Wambui, T. Lubega and P. E. McSharry (۲۰۲۰). "Use of machine learning techniques to create a credit score model for airtime loans." *Journal of Risk and Financial Management* ۱۳(۸): ۱۸۰.
  - Giudici, P., B. Hadji-Misheva and A. Spelta (۲۰۲۰). "Network based credit risk models." *Quality Engineering* ۳۲(Ratkowsky et al.): ۱۹۹-۲۱۱.
  - Giudici, P. and E. Raffinetti (۲۰۲۱). "Shapley-Lorenz eXplainable artificial intelligence." *Expert systems with applications* ۱۶۷: ۱۱۴۱۰۴.
  - Hamidi, K. S., Al-Saadi, M. N., & Al-Shibli, G. I. (۲۰۲۳). Using artificial intelligence techniques to improve financial reporting and its impact on investor decisions. *Tikrit Journal of Administrative and Economic Sciences*, 19(Ratkowsky et al.), ۵۸-۷۵.
  - Habibi tabar, J. (۲۰۰۹). "Divorce and the Family Support Bill: With an emphasis on the prevention of divorce." *Women's Strategic Studies* ۱۱(۴۳): ۳۵-۶۴.
  - Khajeh-Kazemi, R., B. Golestan, K. Mohammad, M. Mahmoudi, S. Nedjat and M. Pakravan (۲۰۱۱). "Comparison of generalized estimating equations and quadratic inference functions in superior versus inferior Ahmed glaucoma valve implantation." *Journal of research in medical*



- sciences: the official journal of Isfahan University of Medical Sciences ۱۶(۳): ۲۳۵.
- Kleinbaum, D. G., M. Klein, D. G. Kleinbaum and M. Klein (۲۰۱۲). "Competing risks survival analysis." *Survival Analysis: A self-learning text*: ۴۲۵-۴۹۵.
  - Kumar, G., M. R. Rahman, A. Rajverma and A. K. Misra (۲۰۲۴). "Predicting systemic risk of banks: a machine learning approach." *Journal of Modelling in Management* ۱۹(Ratkowsky et al.): ۴۴۱-۴۶۹.
  - Lee, T.-S. and I.-F. (Chen, ۲۰۱۶) (۲۰۰۵). "A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines." *Expert Systems with applications* ۲۸(Ratkowsky et al.): ۷۴۳-۷۵۲.
  - Lessmann, S., B. Baesens, H.-V. Seow and L. C. Thomas (۲۰۱۵). "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research." *European journal of operational research* ۲۴۷(Ratkowsky et al.): ۱۲۴-۱۳۶.
  - Lundberg, S. M. and S.-I. Lee (۲۰۱۷). "A unified approach to interpreting model predictions." *Advances in neural information processing systems* ۳۰.
  - Moscato, V., A. Picariello and G. Sperlí (۲۰۲۱). "A benchmark of machine learning approaches for credit score prediction." *Expert Systems with Applications* ۱۶۵: ۱۱۳۹۸۶.
  - Ribeiro, M. T., S. Singh and C. Guestrin (۲۰۱۶). "Model-agnostic interpretability of machine learning." *arXiv preprint arXiv:۱۶۰۶.۰۵۳۸۶*.
  - Ribeiro, M. T., S. Singh and C. Guestrin (۲۰۱۶). "Why should i trust you?" Explaining the predictions of any classifier. *Proceedings of the ۲۲nd ACM SIGKDD international conference on knowledge discovery and data mining*.
  - Serrano-Cinca, C. and B. Gutiérrez-Nieto (۲۰۱۶). "The use of profit scoring as an alternative to credit scoring systems in peer-to-peer (P2P) lending." *Decision Support Systems* ۸۹: ۱۱۳-۱۲۲.
  - Tobias, A. and M. K. Brunnermeier (۲۰۱۶). "(CoVaR." *The American Economic Review* ۱۰۶(Ratkowsky et al.): ۱۷۰۵.
  - Tripathi, D., A. K. Shukla, B. R. Reddy, G. S. Bopche and D. Chandramohan (۲۰۲۲). "Credit scoring models using ensemble learning and classification approaches: a comprehensive survey." *Wireless Personal Communications* ۱۲۳(Ratkowsky et al.): ۷۸۵-۸۱۲.
  - Vallee, B. and Y. Zeng (۲۰۱۹). "Marketplace lending: A new banking paradigm?" *The Review of Financial Studies* ۳۲(Ratkowsky et al.): ۱۹۳۹-۱۹۸۲.
  - Vidal, M. F. and F. Barbon (۲۰۱۹). "Credit scoring in financial inclusion." *Washington: CGAP*.
  - West, D. (۲۰۰۰). "Neural network credit scoring models." *Computers & operations research* ۲۷(۱۱-۱۲): ۱۱۳۱-۱۱۵۲.
  - Yobas, M. B., J. N. Crook and P. Ross (۲۰۰۰). "Credit scoring using neural and evolutionary techniques." *IMA Journal of Management Mathematics*) ۱۱ Ratkowsky et al. ۱۲۵-۱۱۱): (.

• حیدریوسف، ه. پ. پیری و پ. چالاک (۱۴۰۴). "مدل آینده‌نگری به‌کارگیری هوش مصنوعی در گزارشگری مالی با تأکید بر سودمندی درک‌شده و سهولت استفاده‌شده." *تحقیقات مالی* ۷۹(۲۷): ۶۶۲-۶۸۴.

- Chen, T. (۲۰۱۶). *XGBoost: A Scalable Tree Boosting System. Cornell University*.
- Ratkowsky, D., Lowry, R., McMeekin, T., Stokes, A., & Chandler, R. (۱۹۸۳). Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *Journal of bacteriology*, 154(۳), ۱۲۲۲-۱۲۲۶.

