فصلنامه مدلسازی اقتصادی (سال چهاردهم، شماره ۱ «پیاپی ۴۹» بهار ۱۳۹۹، صفحه‌های ۱۵۱-۱۲۷)

شناسایی عوامل موثر بر مطالبات غیرجاری بانک‌ها با استفاده از شبکه‌های عصبی و الگوریتم ماشین بردار پشتیبان[۱]

سجاد کردمنجیری*، ایمان داداشی **، زهرا خوشنود+، حمیدرضا غلام‌نیا روشن×

تاریخ دریافت: ۱۳۹۸/۰۹/۰۸ تاریخ پذیرش: ۱۳۹۹/۰۲/۰۲

چکیده

هدف این مقاله شناسایی عوامل تاثیرگذار بر ایجاد و افزایش مطالبات غیرجاری برای اتخاذ تصمیم مناسب‌تر در اعطای تسهیلات است. بدین منظور برای انتخاب متغیرهای موثر، از الگوریتم‌های تجزیه و تحلیل مولفه‌های هم‌بستگی و لاسو و برای کلاس‌بندی نمونه‌ها، از شبکه‌های عصبی و ماشین بردار پشتیبان استفاده شده است. در این پژوهش، نمونه‌ای از ٦٦٠ مشتری حقوقی بانک سپه برای سال‌های ۱۳۹٦-۱۳۸۵ انتخاب و بر متغیرهای خصوصیتی مستخرج از قراردادهای تسهیلاتی این مشتریان در کنار متغیرهای مالی، غیرمالی، حسابرسی و اقتصادی تمرکز شده است. نتایج نشان داد الگوریتم لاسو با تمرکز بر متغیرهای مالی، اقتصادی و حسابرسی، عملکرد بهتری نسبت به الگوریتم تجزیه و تحلیل مولفه‌های هم‌سایگی داشته و براساس این الگوریتم، ۱۰ متغیر کلیدی تاثیرگذار بر مطالبات غیرجاری شناسایی شدند. با توجه به عملکرد بهتر ماشین‌های بردار پشتیبان با هسته شعاعی، استفاده از آن در مدل‌سازی مطالبات غیرجاری پیشنهاد می‌شود.

طبقه‌بندی JEL: C45, C58, G21.

واژگان کلیدی: تسهیلات بانکی، مطالبات غیرجاری، ماشین بردار پشتیبان، شبکه عصبی.

---

[۱] مقاله مستخرج از رساله دکتری سجاد کردمنجیری به راهنمایی دکتر ایمان داداشی و مشاوره دکتر زهرا خوشنود و دکتر حمیدرضا غلام‌نیا روشن در دانشگاه آزاد اسلامی واحد بابل است.

*دانشجوی دکتری گروه حسابداری، واحد بابل، دانشگاه آزاد اسلامی، بابل، ایران، پست الکترونیکی:
s.kordmanjiri95@gmail.com

**استادیار گروه حسابداری، واحد بابل، دانشگاه آزاد اسلامی، بابل، ایران (نویسنده مسئول)، پست الکترونیکی:
i.dadashi@gmail.com

+استادیار، گروه بانکداری، پژوهشکده پولی و بانکی بانک مرکزی، تهران، ایران، پست الکترونیکی:
rkhoshnoud@yahoo.com

×استادیار گروه حسابداری، واحد بابل، دانشگاه آزاد اسلامی، بابل، ایران، پست الکترونیکی:
Hamid_r_2057@yahoo.com

۱. مقدمه

ارتباط صحیح بین نظام مالی و بخش تولیدی از مهم‌ترین عوامل رشد و توسعه اقتصادی محسوب می‌شود. بانک‌ها به عنوان رکن اصلی نظام مالی، نقش مهمی در تامین وجوه مورد نیاز بخش‌های تولیدی، تجاری، مصرفی و حتی دولتی بر عهده دارند. در ایران نیز با توجه به ساختار اقتصادی و به دلایلی مانند عدم توسعه بازارهای سرمایه و انواع مختلف موسسات و نهادهای غیربانکی، تامین مالی بخش واقعی اقتصاد بیش‌تر بر عهده شبکه بانکی است.

بانک‌ها به دنبال اعطای تسهیلات خود به اشخاصی هستند که ضمن برخورداری از ریسک پایین بتوانند بازدهی متناسب با سود تسهیلات اعطایی داشته باشند. این امر زمانی محقق می‌شود که بانک‌ها قادر به شناسایی مشتریان اعتباری خود اعم از حقیقی و حقوقی بوده و بتوانند آنها را براساس توانایی و تمایل، نسبت به بازپرداخت کامل و به موقع تعهدات، با استفاده از معیارهای مالی و غیرمالی مناسب، طبقه‌بندی کنند. در چنین رویه‌ای، تسهیلات به متقاضیانی پرداخت می‌شود که از ریسک اعتباری کم‌تری برخوردار بوده و احتمال بازپرداخت بدهی آنها در سررسید بیش‌تر است (پورکاظمی، صداقت‌پرست و ده‌پناه، ۱۳۹٦).

در مورد اهمیت مدیریت مطالبات غیرجاری، حتی راینهارت و روگوف[1] (۲۰۱۰) افزایش این نوع مطالبات را به عنوان نشان بالقوه شروع بحران بانکی معرفی کرده‌اند. زیرا اگر بانک‌ها نتوانند مطالبات غیرجاری خود را از حد لازم پایین‌تر نگه دارند، در ارائه تسهیلات جدید بانکی با مشکل روبه‌رو شده و ثبات شبکه بانکی با تهدید مواجه خواهد شد.

در واقع، مطالبات غیرجاری بانک‌ها که «دارایی غیرگردشی[2]» نیز نامیده می‌شوند، نشانگر زیان انباشته‌ای است که در اثر استمرار مجموعه‌ای از فراز و فرودها در سطح کلان و سوء مدیریت‌ها در بخش مالی، در اقتصاد کشور رسوب کرده است (نیلی و محمودزاده، ۱۳۹۳). افزایش این متغیر نشان‌دهنده افزایش ریسک اعتباری بانک‌هاست و نسبت تسهیلات غیرجاری به کل تسهیلات به عنوان شاخصی در شناسایی ریسک اعتباری بانک‌ها به کار می‌رود.

یکی از موضوعات مهم در اعطای تسهیلات بانکی، احتمال غیرجاری شدن مطالبات توسط مشتریان حقیقی و حقوقی بانک‌هاست که با توجه به اطلاعات مندرج در نماگرهای

[1] Reinhart & Rogoff
[2] Non-Return Asset

اقتصادی بانک مرکزی (۱۳۹۷)، نسبت مطالبات غیرجاری به کل تسهیلات اعطایی در بانک‌ها و موسسات اعتباری در کشور طی سال‌های ۱۳۹۵ و ۱۳۹٦ به ترتیب، ۹/۵ و ۱۰/۳ درصد بوده است[۱] که نشان‌دهنده ریسک بالای اعطای تسهیلات در بانک‌های ایران است. از آنجا که این مساله پیامدها و مشکلات زیادی را برای بانک‌ها و در نتیجه، اقتصاد کشور به دنبال دارد؛ بنابراین، شناخت عوامل شکل‌دهنده این نوع از مطالبات برای کاهش و کنترل ریسک اعتباری مشتریان، ارتقای فرایند پرداخت تسهیلات، و بهبود ساختار ترازنامه بانک‌ها ضروری است.

با توجه به اهمیت ریسک اعتباری، تاکنون مدل‌های بسیاری برای پیش‌بینی این ریسک در مشتریان حقیقی و حقوقی بانک‌ها طراحی شده است. حتی مدل‌سازی رتبه‌بندی مشتریان از سال ۱۹٤۰، برای اندازه‌گیری ریسک اعتباری و پشتیبانی تصمیم رد یا قبول درخواست اعتبار به کار گرفته شده است (بلوتی[۲]، ۲۰۱۰).

مدل‌های ابتدایی، عموماً بر روش امتیازدهی آلتمن[۳](۱۹٦۸) و تحلیل ممیزی استوار بوده و در ادامه مدل‌های احتمال خطی، لاجیت و پروبیت مورد استفاده قرار گرفتند. مدل‌های پیچیده‌تری منشعب از هوش مصنوعی نیز، در این حوزه به کار گرفته شده‌اند که از میان آنها می‌توان به سیستم‌های شبکه‌های عصبی و الگوریتم ژنتیک (کیم و آهن[٤]، ۲۰۱۲)، هوش مصنوعی (کارا و کری‌چن[٥]، ۲۰۱۲)، ماشین بردار پشتیبان (وانگ و ما[٦]، ۲۰۱۲) و درخت تصمیم (ادو، گوگان و حسنی[۷]، ۲۰۱۸) اشاره کرد.

در مطالعات اخیر داخلی نیز برای ارزیابی ریسک اعتباری، بر مدل‌های منتج از هوش مصنوعی تمرکز شده است (فلاح‌پور، راعی و هندیجانی‌زاده، ۱۳۹۳؛ داداحمدی و احمدی، ۱۳۹۳؛ محمدیان حاجی کُرد، اصغرزاده زعفرانی و امام‌دوست، ۱۳۹۵؛ حبیبی، کوهی و بعیدی، ۱۳۹۷؛ تاری، ابراهیمی، موسوی و کلانتری، ۱۳۹٦).

---

[۱] براساس آمار بانک جهانی، مطالبات معوق باید بین ۲ تا ۵ درصد تسهیلات اعطایی باشد و مقادیر بیش‌تر این نسبت به عنوان ریسک بالا محسوب می‌شود (رستمی، نبی‌زاده و شاهی، ۱۳۹۷).

[2] Bellotti
[3] Altman
[4] Kim & Ahn
[5] Karaa & Krichene
[6] Wang & Ma
[7] Addo, Guegan & Hassani

از این‌رو، در این مقاله نیز با تکیه بر الگوریتم‌های شبکه عصبی یک لایه، شبکه عصبی چند لایه[۱]، شبکه کمّی‌سازی بردار یادگیری[۲] و الگوریتم‌های ماشین بردار پشتیبان[۳] با هسته چند جمله‌ای و شعاعی، از نمونه‌ای از مدل‌های مبتنی بر هوش مصنوعی پرکاربرد در مدل‌سازی ریسک اعتباری استفاده می‌شود. با وجود این، در مقایسه با دیگر مطالعات داخلی، بر مجموعه گسترده‌ای از متغیرها تمرکز شده و در کلاس‌بندی[۴] آنها، از الگوریتم‌های لاسو[۵] و تجزیه و تحلیل مولفه‌های هم‌سایگی[۶] استفاده می‌شود.

در ادامه مقاله بدین شکل سازماندهی می‌شود؛ پس از مقدمه، ادبیات موضوع مرور می‌شود؛ سپس، روش تحقیق بررسی می‌شود؛ در مرحله چهارم، مدل پژوهش برآورد شده و در پایان نیز نتیجه‌گیری و پیشنهادها ارائه می‌شود.

۲. ادبیات موضوع

از آنجا که افزایش مطالبات غیرجاری، منجر به شکل‌گیری ریسک اعتباری و کاهش سوددهی بانک‌ها می‌شود، عموماً در قالب الگوی بانکداری سنتی، شناسایی اهلیت و اخذ وثیقه و تضمین‌های لازم از مشتریان برای پوشش این ریسک مورد توجه قرار می‌گیرد. با وجود این، با افزایش تعداد درخواست‌های اعتباری از سوی مشتریان و در نتیجه، افزایش ریسک اعتباری، ارزیابی میزان شایستگی اعتباری آنها در قالب الگوهای شناسایی اهلیت رایج در «بانکداری رابطه‌ای[۷]» نمی‌تواند از کفایت و پویایی لازم در مدیریت مطالبات برخوردار باشد.

بنابراین، باید قبل از اعطای اعتبار، براساس مدل‌های آماری پیشرفته و مبتنی بر متغیرهای کلیدی، ریسک اعتباری مشتری به خوبی ارزیابی شود تا احتمال غیرجاری شدن مطالبات در

---

[1] Multi-Layer Perceptron (MLP)
[2] Learning Vector Quantization (LVQ)
[10] Support Vector Machine (SVM)
[4] Classification
[5] Least Absolute Shrinkage and Selection Operator (LASSO)
[6] Neighbourhood Component Analysis (NCA)
[۷] منظور از بانکداری رابطه‌ای (Relationship Banking)، توانایی بانک‌ها در به‌دست آوردن اطلاعات گیرندگان تسهیلات از از طریق نظارت مستمر بر تسهیلات موجود؛ الگوی رفتاری آنها در بازپرداخت این تسهیلات، در کنار الگوی درخواست خدمات مختلف بانکی است.

حد امکان کاهش یابد. بدین ترتیب، گذر به روش‌هایی آینده‌نگر و مبتنی بر پیش‌بینی رفتار اعتباری مشتریان پس از اعطای تسهیلات، در کنار شناسایی اهلیت و وضعیت اعتباری فعلی آنها، دریافت وثیقه‌های معتبر ضروری است.

بنابراین، شناسایی عوامل موثر بر شکل‌گیری مطالبات غیرجاری و منتهی به ریسک اعتباری، از اولویت برخوردار است. در واقع، بر پایه این متغیرها می‌توان در گام دوم، مدل‌سازی احتمال نکول و غیرجاری شدن مطالبات براساس رویکردهای نوین را به نحو مناسبی انجام داد.

در این راستا، طی سال‌های اخیر، مطالعات متعددی در مورد ریسک اعتباری و به طور خاص، مطالبات غیرجاری و تاثیر متغیرهای مالی و کلان بر آن در کشور انجام شده است. بررسی این مطالعات در کنار پژوهش‌های خارجی، بیانگر تمرکز بسیاری از آنها بر گروه‌های محدودی از متغیرها بوده است. در این زمینه متغیرهای مالی، از پرکاربردترین متغیرها بوده‌اند (ادو و همکاران، ۲۰۱۸؛ بلانکو، مجیاس، لارا و ریو[1]، ۲۰۱۳؛ کارا و کری‌چن، ۲۰۱۲؛ وانگ و ما، ۲۰۱۲؛ حکیمی‌پور، ۱۳۹۷؛ تاری و همکاران، ۱۳۹٦؛ محمدیان حاجی کُرد و همکاران، ۱۳۹۵؛ میرغفوری و امین‌آشوری، ۱۳۹٤؛ ابراهیمی و دریابر، ۱۳۹۱)[2].

پس از متغیرهای مالی، متغیرهای اقتصادی از قدمت طولانی در حضور در این مدل‌ها برخوردارند (بلانکو و همکاران، ۲۰۱۳؛ سیدشکری و گروسی، ۱۳۹٤)[3]. متغیرهای غیرمالی نیز نیز که بیش‌تر متمرکز بر توصیف کلی اعتبارگیرنده هستند؛ به دلیل نقش مهمی که در شکل‌گیری مطالبات غیرجاری و ریسک اعتباری بانک‌ها ایفا می‌کنند، عموماً در مدل‌سازی

[1] Blanco, Mejias, Lara & Raya

[2] از جمله متغیرهای مالی مورد استفاده می‌توان به موارد زیر اشاره کرد: نسبت کل بدهی به کل دارایی، حقوق صاحبان سهام به کل دارایی، بدهی بلندمدت به کل دارایی، بدهی بلندمدت به حقوق صاحبان سهام، نسبت جاری، حساب دریافتنی به بدهی جاری، حساب دریافتنی به کل فروش، حساب پرداختنی به کل فروش، فروش به دارایی‌ها، فروش به سود قبل از بهره و مالیات، کل دارایی‌ها به دارایی ثابت و اسناد پرداختنی به سرمایه در گردش.

[3] نرخ رشد تولید ناخالص داخلی و تورم از متداول‌ترین متغیرهای اقتصادی مورد استفاده بوده‌اند.

استفاده می‌شوند (بخت و التر[1]، ۲۰۱۲؛ بلانکو و همکاران، ۲۰۱۳؛ تاری و همکاران، ۱۳۹٦؛ پورکاظمی و همکاران، ۱۳۹٦؛ ابراهیمی و دریابر، ۱۳۹۱)[2].

آخرین گروه از متغیرهای مورد استفاده نیز، متغیرهای تصریح‌کننده خصوصیات و ویژگی‌های قرارداد تسهیلات است (بخت و التر، ۲۰۱۲؛ بلانکو و همکاران، ۲۰۱۳؛ پورکاظمی و همکاران، ۱۳۹٦؛ جعفری اسکندری و روحی، ۱۳۹٦)[3].

در ادامه، گفتنی است تاکنون مدل‌های بسیاری برای پیش‌بینی ریسک اعتباری مشتریان حقیقی و حقوقی، بر پایه دامنه متغیرهای توضیحی، طراحی شده و مورد استفاده قرار گرفته است. در این زمینه مدل‌های متنوعی از مجموعه مدل‌های حوزه هوش مصنوعی در مطالعات اخیر استفاده شده‌اند که حتی برخی از آنها عملکرد بهتری در پیش‌بینی این ریسک داشته‌اند.

ادو و همکاران (۲۰۱۸) در تجزیه و تحلیل ریسک اعتباری در قالب پیش‌بینی احتمال ورشکستگی شرکت‌ها با استفاده از اطلاعات متغیرهای مالی مشتریان بانکی در فرانسه به این نتیجه رسیدند که مدل‌های مبتنی بر درخت تصمیم، پایدارتر از مدل‌های مبتنی بر شبکه‌های عصبی چندلایه هستند.

بلانکو و همکاران (۲۰۱۳) امتیازدهی مشتریان یک موسسه تامین مالی خرد در پرو را با استفاده از متغیرهای مالی، غیرمالی، خصوصیتی قرارداد و اقتصادی در قالب سه روش سنتی مبتنی بر مدل‌های تحلیل تمایز خطی، تحلیل تمایز درجه دو و نیز رگرسیون لجستیک در مقایسه با شبکه عصبی چند لایه آزمون کردند. نتایج، نشان‌دهنده عملکرد بهتر مدل هوش مصنوعی در مقایسه با سه مدل سنتی بوده است.

بخت و التر (۲۰۱۲) با در نظر گرفتن دامنه وسیع‌تری از متغیرها شامل سن، جنسیت، درآمد کل، مقدار تسهیلات، هدف از درخواست تسهیلات، ملیت و نرخ بهره، به مقایسه دو

---

[1] Bekhet & Eletter

[2] در تصریح این متغیر در مدل، براساس نوع مشتری و حقیقی یا حقوقی بودن آن، متغیرهای متفاوتی مورد توجه قرار می‌گیرد؛ به عنوان مثال، پورکاظمی و همکاران (۱۳۹٦) در مطالعه خود و متمرکز بر مشتریان حقیقی، درآمد، سن، جنسیت، وضعیت تاهل، سابقه مالی، نوع شغل و افراد تحت تکفل را در تصریح این متغیر منظور کرده‌اند.

[3] از جمله متغیرهای تصریح‌کننده خصوصیات قراردادها می‌توان به موارد زیر اشاره کرد (پورکاظمی و همکاران، ۱۳۹٦): مدت زمان و نحوه بازپرداخت، مبلغ و نوع تسهیلات، نوع وثیقه.

روش رگرسیون لجستیک و شبکه‌های عصبی تابع پایه شعاعی، در پیش‌بینی ریسک اعتباری مشتریان بانکی در اردن پرداختند. نتایج نشان داد مدل رگرسیون لجستیک از نظر میزان دقت کلی، عملکرد کمی بهتر از مدل عملکرد شعاعی است. با این حال، شبکه‌های عصبی تابع پایه شعاعی در شناسایی مشتریانی با پیش فرض قبلی رفتاری، عملکرد بهتری داشته است.

کارا و کری‌چن (۲۰۱۲) در ارزیابی ریسک اعتباری مشتریان بانک، از اطلاعات نسبت‌های مالی ۱٤۳۵ پرونده اعتباری شرکت‌های صنعتی تونس و مدل‌های شبکه‌های عصبی و ماشین بردار پشتیبان استفاده کردند. نتایج بیانگر عملکرد بهتر مجموعه اطلاعات تلفیقی از متغیرهای تعهدی، جریان وجوه نقد و وثیقه، از بین متغیرهای مالی، در پیش‌بینی ریسک اعتباری و کارکرد بهتر مدل شبکه عصبی چند لایه در مقایسه با ماشین بردار پشتیبان است.

وانگ و ما (۲۰۱۲) در ارزیابی ریسک اعتباری مشتریان بر اساس ماشین بردار پشتیبان، از سوابق مالی ۲۳۹ شرکت در بانک صنعتی و تجاری چین استفاده کردند. نتایج نشان می‌دهد که شبکه ماشین بردار پشتیبان می‌تواند به عنوان روشی جایگزین برای ارزیابی ریسک اعتباری شرکت‌ها مورد استفاده قرار گیرد.

تاری و همکاران (۱۳۹٦) با هدف ارزیابی روش‌های رتبه‌بندی اعتباری مشتریان حقیقی بانک‌ها با بررسی سوابق مالی و خصیصه‌های متقاضی و استفاده از سه مدل لاجیت، شبکه عصبی و الگوریتم ژنتیک دریافتند عملکرد شبکه عصبی نسبت به مدل‌های لاجیت و ژنتیک تا حدودی بهتر بوده است.

پورکاظمی و همکاران (۱۳۹٦) در شناسایی عوامل موثر بر احتمال نکول مشتریان حقیقی بانک پاسارگاد، با استفاده از روش شبکه‌های عصبی، اطلاعات تسهیلات ٤۷۰ مشتری را بررسی کردند. نتایج بیانگر قدرت پیش‌بینی مناسب احتمال نکول متقاضیان (۹۲ درصد) توسط روش شبکه‌های عصبی است. طبق این نتایج، متغیرهایی مانند سوءسابقه مالی و نوع وثیقه، تاثیر زیادی بر پیش‌بینی احتمال نکول داشته‌اند.

جعفری اسکندری و روحی (۱۳۹٦) با تحلیل اطلاعات مشتریان و ویژگی‌های قراردادها به ارزیابی ریسک مشتریان با استفاده از روش ماشین بردار پشتیبان بهبود یافته با الگوریتم ژنتیک در قالب رویکرد داده‌کاوی پرداختند. نتایج بیانگر دقت بیش‌تر این روش در مقایسه با مدل استانداردی نظیر رگرسیون لجستیک در پیش‌بینی ریسک اعتباری است.

محمدیان حاجی کُرد و همکاران (۱۳۹۵) در پژوهشی با تمرکز بر نسبت‌های مالی در طراحی مدل ارزیابی ریسک اعتباری مشتریان حقوقی بانک تجارت، از ماشین بردار پشتیبان و الگوریتم ژنتیک استفاده کردند. آنها با استفاده از متغیرهای مالی ۲۸۲ مشتری حقوقی این بانک و بهینه‌سازی ورودی‌های ماشین بردار پشتیبان با استفاده از الگوریتم ژنتیک به نتایجی با عملکرد بهتر در شناسایی مشتریان خوش حساب و بد حساب و پیش‌بینی ریسک اعتباری در مقایسه با عدم استفاده از این الگوریتم در ماشین بردار پشتیبان دست یافتند.

ابراهیمی و دریابر (۱۳۹۱) به پیش‌بینی ریسک اعتباری مشتریان حقوقی با استفاده از رویکرد تحلیل پوششی داده‌ها، رگرسیون لجستیک و شبکه‌های عصبی پرداختند. آنها با توجه به نتایج حاصل از تحلیل عاملی و قضاوت خبرگان، ۸ شاخص مالی و غیرمالی اثرگذار در شناسایی این ریسک را شناسایی کردند. البته نتایج این پژوهش نیز نشان می‌دهند مدل شبکه‌های عصبی نسبت به تحلیل پوششی داده‌ها عملکرد بهتری در پیش‌بینی ریسک اعتباری و گروه‌بندی متقاضیان تسهیلات دارد.

در اکثر مطالعات، تمرکز بر متغیرهای مالی و اقتصادی بوده و علی‌رغم با اهمیت بودن انواع مختلفی از نسبت‌های مالی در شکل‌گیری مطالبات غیرجاری، در مطالعات پیشین به دامنه گسترده‌ای از این متغیرها توجه نشده است. از این‌رو، در این مقاله سعی شده با تمرکز بر دامنه گسترده‌تری از متغیرهای مالی و اقتصادی، در کنار توجه به دیگر متغیرها در شناسایی عوامل موثر بر مطالبات غیرجاری با استفاده از الگوریتم‌های مختلف و روش‌های جدید فناوری در ارزیابی نتایج، گام موثری در این زمینه برداشته شود.

در مجموع، این پژوهش حاوی سه نوآوری و گام موثر نسبت به مطالعات قبلی است. اول، با تمرکز بر متغیر خصوصیتی قراردادها، متغیرهای مالی، غیرمالی، حسابرسی و اقتصادی، بر دامنه گسترده‌تری از انواع متغیرها تمرکز شده است. دوم، در راستای ضرورت به‌کارگیری مدل‌های جدید، در انتخاب متغیرهای تاثیرگذار، از دو نوع الگوریتم استفاده کرده است. سوم، دقت نتایج هر دو الگوریتم انتخاب ویژگی، با سه روش شبکه‌های عصبی و دو نوع الگوریتم ماشین بردار پشتیبان ارزیابی شده است. به این ترتیب، انتظار می‌رود استفاده از نتایج این تحقیق بتواند در فرایند تصمیم‌گیری کارشناسان و کمیته اعتباری در اعطای اعتبار و اخذ

وثیقه‌ها، کمک موثری کرده و در ارتقای الگوی تخصیص منابع بانک‌ها به طرح‌ها و بخش‌های مختلف اقتصادی مفید واقع شود.

۳. روش تحقیق

هدف این مقاله، پوشش شکاف تحقیقاتی موجود در اتخاذ رویکردی جامع‌تر با محدود کردن جامعه مشتریان در کنار دامنه گسترده‌تر متغیرهای کلیدی و به‌کارگیری رویکردهای نوین مبتنی بر هوش مصنوعی در شناسایی عوامل اثرگذار بر مطالبات غیرجاری است. از این‌رو، به دلیل دسترسی به پرونده‌های اعتباری و نظر به حجم زیادتر تسهیلات اعطایی به مشتریان حقوقی در مقایسه با مشتریان حقیقی، با تمرکز بر مشتریان حقوقی بانک سپه و استفاده از رویکردهای مبتنی بر هوش مصنوعی و تمرکز هم‌زمان بر دامنه وسیعی از متغیرها شامل متغیرهای خصوصیتی قراردادها، متغیرهای مالی، غیرمالی، حسابرسی و اقتصادی کشور، در حد لازم شکاف برطرف شده و بر ارزش مطالعه افزوده شده است.

در انتخاب ویژگی‌ها و متغیرهای اثرگذار، از الگوریتم‌های انتخاب ویژگی لاسو و تجزیه و تحلیل مولفه‌های همسایگی استفاده شده و در پایان، براساس روش‌های انتخابی کلاس‌بندی در هوش مصنوعی با تکیه بر عملکرد بهتر، دقت نتایج بررسی می‌شود.

در این پژوهش از مطالعه اطلاعات کمّی و کیفی ٦٦٠ پرونده تسهیلاتی اشخاص حقوقی در قراردادهای بالای ده میلیاردی بانک سپه استفاده شده است[1]. از این تعداد پرونده، ٤١ درصد بدحساب (غیرجاری) و ٥٩ درصد خوش‌حساب و ٦٧ درصد مشتریان فوق را شرکت‌های بورسی و ۳۳ درصد نیز شرکت‌های غیربورسی تشکیل می‌دهند. دوره زمانی این پژوهش با درنظر گرفتن شروع سال طبقه‌بندی دارایی‌ها طبق بخشنامه بانک مرکزی و در دسترس بودن آنها برای دوره ۱۲ ساله از سال ۱۳۹٦-۱۳۸۵ است.

متغیر وابسته پژوهش مبتنی بر مطالبات غیرجاری استخراج شده از پرونده تسهیلاتی مشتریان حقوقی بانک نمونه است. بر اساس دستورالعمل بانک مرکزی جمهوری اسلامی ایران

---

[1] براساس سیاست‌ها و ضوابط اعتباری بانک مرتبط، پرداخت تسهیلات بیش از ده میلیارد ریال به متقاضیان، منوط به ارائه صورت‌های مالی حسابرسی شده است. بنابراین، اطلاعات متغیرهای حسابرسی این مشتریان نیز در دسترس است.

در خصوص طبقه‌بندی دارایی‌های بانک‌ها و موسسات اعتباری، کیفیت تسهیلات براساس سه پارامتر زمان، وضعیت مالی مشتری و رشته فعالیت، در قالب تسهیلات جاری و غیرجاری طبقه‌بندی می‌شود. با تمرکز صرف بر پارامتر زمان، در دارایی‌های جاری، پرداخت اصل و سود تسهیلات یا بازپرداخت اقساط، در سررسید انجام شده و یا حداکثر از سررسید آن دو ماه گذشته است. در طرف مقابل، مطالبات غیرجاری شامل مطالبات سررسید گذشته، معوق و مشکوک‌الوصول بوده و کل مانده تسهیلات از زمان طبقه‌بندی در این گروه‌ها به سرفصل‌های مرتبط منتقل می‌شود[۱].

در ارتباط با متغیرهای توضیحی و اثرگذار بر مطالبات غیرجاری گفتنی است؛ بر اساس ادبیات تحقیق و با تمرکز بر مطالعات داخلی، مجموعه متنوعی از متغیرها در دسترس قرار گرفت. این متغیرها را می‌توان در چهار گروه ویژگی‌های خصوصیتی قراردادها، نسبت‌های مالی، شاخص‌های غیرمالی و اقتصادی طبقه‌بندی کرد. با این حال، در این مقاله در کنار گونه‌های کلی متغیرهای یاد شده، با توجه به اهمیت متغیرهای حسابرسی در پیش‌بینی ورشکستگی[۲] و احتمال قصور مشتری در بازپرداخت تسهیلات، مجموعه‌ای از این نوع متغیرهای مستخرج از گزارش حسابرسی نیز مورد توجه قرار گرفتند.

به این ترتیب، با مطالعه پژوهش‌های انجام شده، پرتکرارترین و مرتبط‌ترین متغیرها در شکل‌گیری مطالبات غیرجاری، جمع‌آوری و با توجه به گزارش‌های مستمر عملکرد نظام بانکی، داده‌های موجود در بانک نمونه، در کنار نظر کارشناسان امور بانکی، متغیرهای پژوهش، نهایی شد.

از این‌رو، از پنج گروه کلی از متغیرها، شامل ویژگی‌های خصوصیتی قراردادها، متغیرهای مالی، غیرمالی، حسابرسی و اقتصادی استفاده شده و متغیرهای با هم‌بستگی بالا حذف شدند.

---

[۲] در دارایی‌های سررسید گذشته، از تاریخ سررسید اصل و سود تسهیلات و یا تاریخ قطع پرداخت اقساط، بیش از ۲ ماه گذشته، ولی تاخیر در بازپرداخت، هنوز از ٦ ماه تجاوز نکرده است. در دارایی‌های معوق، بیش از ٦ ماه و کم‌تر از ۱۸ ماه از تاریخ سررسید و یا قطع پرداخت اقساط آنها سپری شده و مشتری هنوز اقدامی برای بازپرداخت مطالبات موسسه اعتباری ننموده است و در طبقه مشکوک الوصول، بازه زمانی بیش از ۱۸ ماه مورد نظر است (بانک مرکزی، ۱۳۸۵).

[۲] در این زمینه می‌توان به پژوهش اکبری و حاجیها (۱۳۹۵)، مزرعه فراهانی (۱۳۹۵)، منصورفر، پیری، علیخانی و اسدی (۱۳۹۷) رجوع کرد.

بنابراین، توجه به دامنه گسترده‌ای از انواع متغیرها در کنار یکدیگر، از مهم‌ترین ویژگی‌های متمایزکننده این پژوهش نسبت به مطالعات پیشین است.

به این ترتیب، در گروه متغیرهای بیانگر خصوصیات قراردادهای تسهیلاتی منعقد شده با مشتریان حقوقی بانک نمونه، ۱۲ متغیر منظور شدند که عبارتند از مرجع تصویب‌کننده اعتبار ارکان اعتباری (مدیریت منطقه، اداره تسهیلات و هیات مدیره)، نوع عقد قرارداد (مشارکتی، مبادله‌ای)، نوع قرارداد (مشارکت مدنی، فروش اقساطی، سلف، مرابحه، مضاربه و جعاله)، هدف از پرداخت تسهیلات (تامین سرمایه ثابت، تامین سرمایه در گردش)، مبلغ تسهیلات، نحوه بازپرداخت (یک‌جا، اقساطی)، مدت قرارداد (تا یک‌سال، یک تا دو سال، دو تا سه سال، بالای سه سال)، تغییر در وضعیت قرارداد (بدون تغییر، تمدید، تقسیط مجدد)، نرخ سود تسهیلات، نوع وثیقه (اموال غیرمنقول محل فعالیت، اموال غیرمنقول خارج از محل فعالیت، ذمه‌ای، سایر)، مراودات مالی شرکت با بانک (ضعیف، متوسط، خوب، خیلی‌خوب)، وضعیت رفتار مشتری در بازپرداخت تسهیلات قبلی (ضعیف، متوسط، خوب، خیلی‌خوب).

در گروه شاخص‌های مالی که مستخرج از اطلاعات صورت‌های مالی مشتریان حقوقی بانک نمونه هستند؛ شاخص‌های مبتنی بر ایفای تعهدات، پایداری رشد شرکت، تسهیلاتی، سودآوری، جریان نقدی و فعالیت مورد توجه قرار گرفته‌اند که هر یک نیز شامل چندین متغیر متفاوت و در مجموع، مشتمل بر ۲۹ متغیر در جدول (۱) ارائه شده‌اند.

جدول ۱. متغیرهای مالی

گروه شاخص مالی	نام متغیر	اندازه‌گیری متغیر
ایفای تعهدات	نسبت آنی	بدهی‌های جاری / دارایی‌های آنی
	نسبت جاری	بدهی‌های جاری/ دارایی‌های جاری
	نسبت دارایی جاری به دارایی‌ها	دارایی‌ها / دارایی جاری
	اندازه شرکت	لگاریتم جمع دارایی‌های شرکت
	نسبت سرمایه در گردش به دارایی‌ها	جمع دارایی‌ها / سرمایه در گردش
	نسبت سرمایه در گردش به فروش	فروش/ سرمایه درگردش
	نسبت پوشش بهره	هزینه بهره / سود قبل از بهره و مالیات
	نسبت مالکانه	جمع دارایی‌ها / حقوق صاحبان سهام
	نسبت بدهی	جمع دارایی‌ها / جمع بدهی‌ها

پایداری رشد شرکت	نرخ رشد سودخالص	سود خالص دوره قبل /سود خالص دوره قبل- سود خالص دوره جاری
	نرخ رشد دارایی‌ها	دارایی دوره قبل / دارایی‌های دوره قبل- دارایی‌های دوره جاری
	نرخ رشد فروش	فروش دوره قبل /فروش دوره قبل- فروش جاری
	نسبت بقاء	سود خالص / سود انباشته
	نسبت سود انباشته به دارایی‌ها	کل دارایی‌ها / سود انباشته
تسهیلاتی	نسبت تسهیلات به دارایی	کل دارایی‌ها / مانده تسهیلات
	نسبت تسهیلات به بدهی‌ها	کل بدهی‌ها / مانده تسهیلات
	حصه جاری به دارایی جاری	دارایی جاری / حصه جاری تسهیلات
	نسبت تسهیلات به فروش	فروش / مانده تسهیلات
سودآوری	بازده دارایی‌ها	متوسط کل دارایی‌ها / سود خالص دوره
	بازده حقوق صاحبان سهام	متوسط حقوق صاحبان سهام/ سود خالص دوره
	حاشیه سود	کل درآمد عملیاتی / سود خالص
جریان نقدی	گردش وجوه نقد	بدهی های جاری / وجه نقد حاصل از عملیات
	بازده نقدی دارایی‌ها	متوسط دارایی‌ها/ وجه نقد حاصل از عملیات
	کیفیت سود	سود عملیاتی/ وجه نقد حاصل از عملیات
	جریان نقدی به هزینه بهره	هزینه بهره/ جریان وجه نقد
فعالیت	گردش دارایی‌ها	متوسط دارایی‌های ثابت/ فروش
	گردش موجودی	متوسط موجودی کالا/ بهای تمام‌شده کالای فروش رفته
	گردش حساب دریافتنی	متوسط حساب دریافتنی/ فروش
	نسبت هزینه بدهی	متوسط بدهی‌ها/ هزینه مالی

منبع: یافته‌های پژوهش

شاخص‌های غیرمالی نیز با پوشش ۱۱ متغیر، مجموعه‌ای از ویژگی‌های کلی حوزه فعالیت و صنعت مشتری حقوقی بانک نمونه را پوشش می‌دهند که عموماً از اطلاعات صورت‌های مالی آنها یا سوابق اطلاعاتی نزد بانک استخراج شده‌اند. این شاخص‌ها عبارتند از: نوع شرکت

(سهامی عام یا خاص)، سن شرکت (تعداد سال‌های فعالیت از زمان تاسیس)، نوع صنعت (بر اساس طبقه‌بندی بورس)، نوع فعالیت (تولیدی و غیرتولیدی)، سابقه فعالیت در بورس، اندازه هیات مدیره (برحسب تعداد اعضاء)، تنوع محصول (تعداد محصولات و یا خدمات)، تعداد کارکنان، شهر محل فعالیت (تهران، مرکز استان، شهرستان)، سابقه حساب (تعداد سال‌های حساب فعال نزد بانک نمونه) و سابقه تسهیلات (تعداد دوره‌های اخذ تسهیلات از بانک نمونه).

در مورد شاخص‌های حسابرسی نیز بر ٤ متغیر کلیدی تاکید می‌شود که از صورت‌های مالی حسابرسی شده آنها استخراج شده‌اند (نوع اظهارنظر حسابرسی[۱]، اظهارنظر حسابرس بابت ابهام در تداوم فعالیت[۲]، تعداد بندهای اظهارنظر در تعدیل گزارش حسابرسی و اندازه موسسه حسابرسی[۳]).

در گروه شاخص‌های اقتصادی نیز ۵ متغیر شامل: نرخ تورم سالانه، شاخص بهای کالای مصرفی، نرخ رشد تولید ناخالص داخلی، شاخص قیمت مسکن و نرخ بهره استفاده شدند.

٤. برآورد مدل و تجزیه و تحلیل آن

٤-۱. انتخاب ویژگی‌ها

انتخاب متغیرهای کلیدی، یک مرحله مهم و چالش برانگیز در مدل‌سازی آماری است. در این پژوهش، به منظور تدوین بهترین مدل با کم‌ترین متغیرهای ممکن، از دو الگوریتم کاهش ویژگی لاسو و تجزیه و تحلیل مولفه‌های همسایگی استفاده شده است. الگوریتم لاسو به عنوان عملگر تنظیم و انتخاب حداقل مقدار مطلق، اولین‌بار توسط رابرت تیب شیرانی[٤] در سال ۱۹۹٦ طراحی شد. روش لاسو محدودیتی برای مجموع ارزش‌های مطلق پارامترهای مدل ایجاد می‌کند که مقدار آن باید کم‌تر از یک حد بالایی باشد. در این روش، یک فرایند

---

[۱] این متغیر به صورت اظهارنظر مقبول، مشروط، مردود و یا عدم اظهارنظر است.

[۲] این متغیر می‌تواند بیانگر ابهام یا عدم ابهام در تداوم فعالیت باشد.

[۳] در پژوهش رضایی، ویسی‌حصار و قندچی (۱۳۹۸) سازمان حسابرسی و موسسه حسابرسی مفید راهبر به عنوان موسسات حسابرسی بزرگ و دیگر موسسات حسابرسی به عنوان موسسات حسابرسی کوچک در نظر گرفته شده‌اند.

[4] Robert Tibshirani

کوچک‌سازی اعمال می‌شود که در آن، ضرایب متغیرهای رگرسیون، کاهش و برخی از آنها، حتی به صفر می‌رسند.

در این فرایند، متغیرهایی که هنوز یک هسته خنثی صفر ندارند، پس از فرایند کوچک شدن، به عنوان بخشی از مدل انتخاب می‌شوند. هدف این فرایند، به حداقل رساندن خطای پیش‌بینی است (فونتی و بلیستر[1]، ۲۰۱۷).

با استفاده از این الگوریتم، ۱۰ متغیر که تاثیرگذاری آنها نسبت به ٦۱ متغیر اصلی بیش‌تر بود، انتخاب شدند (جدول ۲) که به ترتیب اندازه تاثیرگذاری عبارتند از: نرخ رشد سود خالص، کیفیت سود، مانده تسهیلات به کل بدهی‌ها، نرخ تورم سالانه، نسبت جاری، بازده حقوق صاحبان سهام، سود انباشته به دارایی، اظهارنظر حسابرس بابت ابهام در تداوم فعالیت، بندهای تعدیلی اظهارنظر حسابرس مستقل و نرخ بهره.

الگوریتم لاسو[2] برای هر یک از متغیرهای انتخاب شده، یک ضریب تاثیرپذیر تعیین می‌کند که میزان اهمیت آن متغیر را، در دسترس قرار می‌دهد. همان طور که مشاهده می‌شود، بیش‌ترین متغیرهای تاثیرگذار از مجموعه متغیرهای مالی و سپس، متغیرهای اقتصادی و حسابرسی انتخاب شده‌اند.

دیگر الگوریتم مورد استفاده در انتخاب ویژگی‌ها، تجزیه و تحلیل مؤلفه‌های همسایگی است. این الگوریتم به عنوان روشی غیرپارامتری با هدف حداکثرسازی دقت پیش‌بینی در الگوریتم‌های رگرسیون و کلاس‌بندی استفاده می‌شود. اساس این الگوریتم، مبتنی بر K رکورد نزدیک‌تر به مجموعه رکوردهای آموزشی است. با استفاده از این الگوریتم، ۱۰ متغیر که تاثیرگذاری آنها نسبت به ٦۱ متغیر اصلی بیش‌تر بود، انتخاب شدند (جدول ۲) که به ترتیب تاثیرگذاری عبارتند از: دارایی‌های جاری به کل دارایی‌ها، مانده تسهیلات به کل دارایی‌ها، نرخ رشد دارایی‌ها، سن شرکت، مراودات مالی شرکت با بانک، نرخ بهره، نسبت آنی، بازده دارایی‌ها، نوع صنعت، اندازه موسسه حسابرسی.

---

[1] Fonti & Belister

[2] برای الگوریتم تجزیه و تحلیل مولفه‌های همسایگی، چنین وزن‌هایی مقدور نیست و این الگوریتم به صورتی است که تنها به ترتیب تاثیرپذیری، تعداد مشخصی از ویژگی‌ها را انتخاب می‌کند.

نتایج بیانگر تاثیرگذاری بیش‌تر متغیرهای مالی در این الگوریتم در مقایسه با سایر متغیرهاست. با استفاده از این الگوریتم از هر گروه کلی متغیرهای مورد بررسی، حداقل یک متغیر به عنوان متغیر تاثیرگذار شناسایی شده و بیش‌تر متغیرهای انتخاب شده، مربوط به گروه متغیرهای مالی هستند.

به این ترتیب، با حضور متغیرهای اقتصادی در قالب هر یک از الگوریتم‌های مرتبط، در مجموع، نقش کلیدی متغیرهای اقتصادی در کنار دیگر متغیرها، به ویژه متغیرهای مالی، در رفتار مطالبات غیرجاری تائید می‌شود.

جدول ۲. انتخاب ویژگی‌های تاثیرگذار

متغیرهای انتخابی با الگوریتم لاسو			متغیرهای انتخابی با الگوریتم تجزیه و تحلیل مولفه‌های همسایگی	
گروه متغیر	متغیر انتخاب شده	اندازه تاثیرپذیری	گروه متغیر	متغیر انتخاب شده
مالی	رشد سود خالص	۱/۰۷	مالی	دارایی جاری به کل دارایی‌ها
مالی	کیفیت سود	۱/۰۲	مالی	تسهیلات به کل دارایی‌ها
مالی	تسهیلات به کل بدهی‌ها	۰/۷۳	مالی	نرخ رشد دارایی‌ها
اقتصادی	نرخ تورم سالانه	۰/٦۸	غیر مالی	سن شرکت
مالی	نسبت جاری	۰/٤۹	قراردادها	مراودات مالی شرکت با بانک
مالی	بازده حقوق صاحبان سهام	۰/۳۵	اقتصادی	نرخ بهره
مالی	سود انباشته به دارایی	۰/۳٤	مالی	نسبت آنی
حسابرسی	اظهارنظر تعدیلی بابت تداوم فعالیت	۰/۳۲	مالی	بازده دارایی‌ها
حسابرسی	بندهای تعدیلی اظهارنظر	۰/۲۹	غیر مالی	نوع صنعت
اقتصادی	نرخ بهره	۰/۲٦	حسابرسی	اندازه موسسه حسابرسی
تعداد گروه‌های پوشش داده شده = ۳			تعداد گروه‌های پوشش داده شده = ۵	
حضور دو متغیر از گروه متغیرهای اقتصادی			حضور یک متغیر از گروه متغیرهای اقتصادی	

منبع: یافته‌های پژوهش

٤-۲. کلاس‌بندی

برای کلاس‌بندی نمونه‌ها و ارزیابی توانایی ویژگی‌ها در شناسایی عوامل موثر مطالبات غیرجاری در دو الگوریتم کلیدی مورد استفاده، روش‌های گوناگونی مانند شبکه کمّی‌سازی بردار یادگیری، شبکه‌های عصبی تک لایه، چند لایه، و ماشین بردار پشتیبان با هسته‌های چند جمله‌ای و شعاعی آزمایش شدند که در ادامه این شبکه‌ها و چگونگی پیاده‌سازی آنها توضیح داده شده است.

٤-۲-۱. شبکه کمّی‌سازی بردار یادگیری

این شبکه یک روش کلاس‌بندی الگو است که هر کدام از خروجی‌ها، نمایش‌دهنده یک کلاس است و هر یک توسط بردار وزن آن کلاس مشخص می‌شوند. بردارهای وزن توسط مجموعه‌های آموزشی، مقداردهی اولیه شده و سپس توسط الگوریتم‌های یادگیری نظارت شده[1] بهینه می‌شوند. بعد از یادگیری، شبکه چندی‌سازی، بردار یادگیری ورودی را به کلاسی با نزدیک‌ترین فاصله برداری به آن نسبت می‌دهد.

شبکه از دو لایه تشکیل شده است. در لایه اول، بردار ورودی به خوشه‌هایی نگاشت می‌شود که توسط شبکه در حین آموزش پیدا شده است. سپس در لایه دوم، ترکیبی از گروه‌های خوشه‌های لایه اول به کلاس‌های هدف نگاشت می‌شوند. مجموع تعداد خوشه‌های لایه اول، توسط تعداد نرون‌های مخفی[2] تعیین می‌شوند. تعداد نرون‌های لایه مخفی به تعداد واحدهای ورودی و خروجی، تعداد نمونه‌های آموزش، مقدار نویز در کلاس‌های هدف و پیچیدگی تابعی که باید یاد گرفته شود، بستگی دارد. اگر تعداد نرون‌های مخفی لایه اول زیاد باشد، تعداد خوشه‌های بیش‌تری یاد گرفته می‌شوند و نگاشت پیچیده‌تری از ورودی‌ها به خروجی‌ها ایجاد می‌شود. در این مدل، ۸ نرون درنظر گرفته شد (کوهنن[3]، ۱۹۹۸).

---

[1] Supervised Learning
[2] Hidden Neuron
[3] Kohonen

٤-۲-۲. شبکه‌های عصبی

شبکه‌های عصبی از عناصر پردازشی ساده با اتصالات داخلی بسیار زیاد تشکیل شده‌اند. وزن اتصالات داخلی توسط داده‌های آموزشی به دست می‌آیند. تعداد ویژگی‌ها، تعداد نرون‌های لایه ورودی و تعداد کلاس‌ها، تعداد نرون‌های لایه خروجی را مشخص می‌کنند و شبکه ممکن است دارای لایه‌های میانی باشد (اسوزیل، کواسنیکا و پُسپیچال[1]، ۱۹۹۷).

در این مقاله شبکه عصبی تک لایه و چند لایه با تابع سیگموید[2] غیرخطی آزمایش شد. در بین شبکه‌های عصبی تک لایه، شبکه با ده نرون ورودی (تعداد ویژگی‌ها)، هشت نرون میانی و یک نرون پایانی (تعداد کلاس‌ها) به بهترین نتیجه رسید. تعداد نرون‌های لایه میانی به عوامل زیادی بستگی دارد و معمولاً با یک قانون سرانگشتی، عددی بین تعداد نرون‌های ورودی و خروجی است. در این مقاله تعداد نرون‌های لایه میانی به صورت صعودی و یکی یکی آزمایش شد و برای بیش‌تر از هشت نرون در لایه میانی، نتیجه بهتری به دست نیامد. پیاده‌سازی شبکه عصبی در نرم‌افزار متلب انجام شده است که در آزمایش‌های انجام شده، تابع آموزشی پس‌انتشار انعطاف‌پذیر[3]، نسبت به دیگر تابع‌های آموزشی، از لحاظ سرعت و حافظه مصرفی عملکرد بهتری داشت.

٤-۲-۳. ماشین‌های بردار پشتیبان

در الگوریتم ماشین بردار پشتیبان به عنوان الگوریتم تشخیص الگو یا دسته‌بندی اشیاء در کلاس‌های خاص، کلاس‌بندی‌ها براساس قانون بیش‌ترین حاشیه انجام می‌شود. در مسائل غیرخطی، ماشین‌های بردار پشتیبان از حلقه هسته برای به دست آوردن مرزهای غیرخطی استفاده می‌کنند. به این ترتیب، نمونه‌های آموزش که ۸۰ درصد کل نمونه هستند و به روش تصادفی انتخاب شده‌اند؛ به صورت غیرخطی به فضای بردار ویژگی با ابعاد بالاتر نگاشت شده و ابرصفحه‌ای[4] با حاشیه ایجاد می‌شود. در فضای ویژگی، تنها نیاز به محاسبه تابع کرنل

---

[1] Svozil, Kvasnicka & Pospichal
[2] Sigmoid
[3] Resilient Backpropagation
[4] Hyper Plane

(هسته)[1] است که نوع هسته می‌تواند خطی، چند جمله‌ای، شعاعی و سیگموئد باشد (آبه[2]، ۲۰۱۰).

در هسته چندجمله‌ای، تعداد جملات هسته با p نشان داده می‌شود. انتخاب ضریب p به صورت صعودی از ۲ شروع می‌شود و تا زمان عدم بهبود در دقت یادگیری، ادامه می‌یابد. در این مقاله برای $p = 6$ بهترین نتیجه حاصل شد. ماشین بردار پشتیبان با کرنل شعاعی، دارای دو پارامتر مهم c و γ است که تأثیری زیادی در نتیجه دارند و باید در ابتدا به الگوریتم آموزش داده شوند.

ضریب c پارامتر سختی یا نرمی ماشین بردار پشتیبان است. مثلاً اگر پارامتر c کم انتخاب شود؛ بردار تصمیم، نرم انتخاب می‌شود و در صورتی که پارامتر c بسیار زیاد انتخاب شود، بردار تصمیم بسیار سخت‌گیر می‌شود. پارامتر γ مقدار ضریب هسته برای تابع هسته است. هر چه مقدار این ضریب بیش‌تر باشد، الگوریتم به سوی برازش دقیق‌تر براساس مجموعه داده‌های آموزشی می‌رود. مقدار زیاد ضریب γ ممکن است باعث تعمیم خطا و مشکل بیش برازش شود.

در این مقاله برای ویژگی‌های منتخب از الگوریتم لاسو به منظور به دست آوردن پارامترهای هسته ماشین بردار پشتیبان، از جست و جوی جامع دومرحله‌ای استفاده شده است. در مرحله اول، تخمین پارامتر جست و جو دشوار بوده است؛ زیرا یافتن پارامترهای جست جوی اول، صرفاً برای پیدا کردن محدوده و تسریع بخشیدن پارامترهای جستجوی دوم است.

در جست و جوی دوم، پارامترc و γ به ترتیب، برابر با ٤۰ و ۰/۰۰۷ به دست آمد. به همین ترتیب، برای ویژگی‌های انتخاب شده الگوریتم تجزیه و تحلیل مؤلفه‌های همسایگی، این روند تکرار شده که در جست جوی دوم، مقدار پارامتر c و γ به ترتیب، برابر با ۱۰۰ و ۰/۰۲۸ به دست آمد. سپس این پارامترها در الگوریتم ماشین بردار پشتیبان جایگزین شد. برای پیاده‌سازی الگوریتم ماشین بردار پشتیبان، از کتابخانه ماشین بردار پشتیبان[3] استفاده شده است. این کتابخانه شامل انواع الگوریتم‌های ماشین بردار پشتیبان و رگرسیون است.

---

[1] Kernel Function
[2] Abe
[3] Libraray for Support Vector Machines (LIBSVM)

٤-۳. ارزیابی نتایج

برای بررسی عملکرد الگوریتم‌ها در کلاس‌بندی صحیح مطالبات غیرجاری از مطالبات جاری، معیار دقت[1] به کار گرفته شده است. این معیار به صورت رابطه (۱) تعریف می‌شود (السون و دلن[2]، ۲۰۰۸):

$$\text{دقت} = \frac{TP + TN}{P + N} \qquad (۱)$$

در این رابطه، TP تعداد تشخیص‌های مثبت درست و یا به عبارت دیگر، تعداد نمونه‌هایی که مطالبات غیرجاری بوده و صحیح تشخیص داده شده است. TN تعداد تشخیص‌های منفی درست و یا به عبارت دیگر، تعداد نمونه‌های مطالبات جاری است که صحیح تشخیص داده شده است. P بیانگر تعداد کل مثال‌های مثبت یا تعداد نمونه‌های مطالبات غیرجاری است. N تعداد کل مثال‌های منفی یا تعداد نمونه‌های مطالبات جاری است.

با توجه به اینکه در هر مرحله از اجرای الگوریتم‌ها، نمونه‌های آزمون به صورت تصادفی (۲۰ درصد از کل نمونه‌ها) و متفاوت انتخاب می‌شوند؛ بنابراین، مقدار دقت بر داده‌های آموزش و آزمون متفاوت است. بدین ترتیب، در صورتی که بخواهیم براساس داده‌های آموزش و آزمون، به دقت هر یک از الگوریتم‌ها دست یابیم؛ برای هر یک از روش‌های کلاس‌بندی، ده مرتبه الگوریتم تکرار می‌شود که مقدار میانگین به دست آمده، در جدول (۳) برای داده‌های آموزش و آزمون گزارش شده است.

از این‌رو، مشاهده می‌شود که بهترین عملکرد در بین کلاس‌بندها، متعلق به ماشین بردار پشتیبان با هسته شعاعی و با مقادیر دقت ۹۸/٦ و ۹۱/۸ درصد به ترتیب، بر داده‌های آموزش و آزمون در حالت به‌کارگیری الگوریتم لاسو در انتخاب ویژگی‌ها حاصل شده است.

به این ترتیب، نتایج این مطالعه با وجود استفاده از دامنه گسترده‌تری از متغیرها در شناسایی عوامل موثر بر مطالبات غیرجاری در نمونه‌ای از مشتریان حقوقی بانک سپه با استفاده از شبکه‌های عصبی که براساس مطالعات بلانکو و همکاران (۲۰۱۳)، راعی و فلاح‌پور (۱۳۸۷)، ابراهیمی و دریابر (۱۳۹۱)، میرغفوری و امین‌آشوری (۱۳۹۳)، تاری و همکاران (۱۳۹٦) و جعفری اسکندری و روحی (۱۳۹٦)، استفاده از آن در تدوین ریسک اعتباری در

[1] Accuracy
[2] Olson & Delen

مقایسه با مدل‌های سنتی توصیه شده بود؛ در مجموع، بیانگر عملکرد بهتر الگوریتم انتخاب ویژگی لاسو در مقایسه با عملکرد الگوریتم انتخاب ویژگی تجزیه و تحلیل مولفه‌های همسایگی تحت شبکه ماشین بردار پشتیبان با هسته شعاعی است.

جدول ۳. نتایج میانگین کلاس‌بندی‌ها (طبقه‌بندی‌ها) برای ویژگی‌های حاصل از دو الگوریتم

کلاس‌بندی	دقت روی داده‌های				پارامترها	
	آموزش		آزمون			
	لاسو	تجزیه و تحلیل مولفه‌های هم‌سایگی	لاسو	تجزیه و تحلیل مولفه‌های هم‌سایگی	لاسو	تجزیه و تحلیل مولفه‌های هم‌سایگی
شبکه چندی‌سازی بردار یادگیری	۷۸/۱	۷۳/۲	۷۳/۳	۷۰/۹	HN = 8	HN = 8
شبکه عصبی تک لایه	۸۱/۸	۷٤/٤	۷۸/۲	۷۳/٦	H1=8	H1=8
شبکه عصبی چندلایه	۸٤/۱	۷۵/۵	۷۹/۲	۷۳/٤	H1 = H2 = 8	H1 = H2 = 8
شبکه ماشین بردار با هسته چندجمله‌ای	۹٦/۱	۷۸/۲	۸۸/۱	۷۵/۲	P = 6	P = 6
شبکه ماشین بردار با هسته شعاعی	۹۸/٦	۷۸/۹	۹۱/۸	۷۵/۱	c =40 $\gamma = 0/007$	c = 100 $\gamma = 0/028$

HN بیانگر تعداد نرون‌های مخفی؛ H1، لایه مخفی اول؛ H2، لایه مخفی دوم؛ c، ضریب سختی یا نرمی P بیانگر تعداد جملات هسته، و γ بیانگر ضریب هسته‌اند.

منبع: یافته‌های پژوهش

از این‌رو، با توجه به مجموعه کلاس‌بندی‌های مختلف مورد استفاده در این مقاله، استفاده از کلاس‌بندی شبکه ماشین بردار پشتیبان با هسته شعاعی پیشنهاد می‌شود و حتی نتایج، تا حدود زیادی تاییدکننده پژوهش راعی و فلاح‌پور (۱۳۷۸) و جعفری اسکندری و روحی (۱۳۹٦) در خصوص عملکرد بهتر شبکه ماشین بردار پشتیبان است.

بنابراین، براساس نتایج، گسترش دامنه متغیرهای مورد بررسی به متغیرهای حسابرسی در کنار مجموعه متداول متغیرهای مالی (میرغفوری و امین آشوری، ۱۳۹٤ و محمدیان حاجی کُرد و همکاران، ۱۳۹۵) و اقتصادی (حکیمی‌پور، ۱۳۹۷) می‌تواند برای بانک‌ها و موسسات اعتباری کمک‌کننده باشد.

اهمیت این مساله در با مقایسه نتایج مطالعه محمدیان حاجی کُرد و همکاران (۱۳۹۵) در بررسی ریسک اعتباری مشتریان حقوقی با تمرکز صرف بر متغیرهای مالی مشتریان حقوقی نسبت به این پژوهش با تمرکز بر گسترش دامنه متغیرهای مشتریان حقوقی و درصد پیش‌بینی دقت داده‌های آزمون در مدل ماشین بردار پشتیبان (٦۹ درصد در مطالعه محمدیان حاجی کُرد و همکاران، در مقایسه با مقدار ۹۸/٦ درصد نتیجه این پژوهش) نیز به خوبی تایید می‌شود.

۵. نتیجه‌گیری و پیشنهادها

در این مقاله با تمرکز بر ٦٦۰ پرونده از مشتریان حقوقی بانک سپه برای دوره زمانی ۱۳۹٦-۱۳۸۵ جهت شناسایی عوامل موثر بر مطالبات غیرجاری، ویژگی‌های پرونده‌های تسهیلاتی شامل متغیرهای خصوصیتی قراردادها در کنار متغیرهای مالی، غیرمالی و حسابرسی شرکت‌های حقوقی اعتبارگیرنده، به همراه متغیرهای کلان اقتصادی کشور مورد توجه قرار گرفت که منجر به انتخاب اولیه ٦۱ متغیر در این زمینه شد.

در ادامه با توجه به پیشینه پژوهش مبتنی بر عملکرد بهتر مدل‌های هوش مصنوعی در مقایسه با مدل‌های رگرسیونی تحلیل ممیزی، لاجیت و پروبیت در پیش‌بینی درماندگی مالی و ارزیابی ریسک اعتباری، بر گروه منتخبی از این مدل‌ها در پیش‌بینی مطالبات غیرجاری به عنوان ریسک اعتباری تمرکز شد. در این راستا، در گام اول از الگوریتم‌های انتخاب متغیرهای تاثیرگذار لاسو و تجزیه و تحلیل مولفه‌های همسایگی استفاده شد و ۱۰ متغیر تاثیرگذار متفاوت بر مطالبات غیرجاری در قالب هر یک از این الگوریتم‌ها شناسایی شد. هر چند در

قالب هر یک از این الگوریتم‌ها مجموعه متغیرهای متفاوتی از پنج گروه اصلی انتخاب شدند؛ اما در کل، نتایج بیانگر ضرورت گسترش دامنه متغیرهای مورد بررسی از دامنه محدود و متعارف متغیرهای مالی و اقتصادی در گام اول به گروه متغیرهای حسابرسی و در گام بعد به گروه متغیرهای غیرمالی و شاخص قرارداد بوده است.

در ارزیابی دقت نتایج نیز از دامنه گسترده‌ای از الگوریتم‌ها در مقایسه با مطالعات قبلی استفاده شده است که عبارتند از: الگوریتم‌های شبکه کمّی‌سازی بردار یادگیری، شبکه عصبی یک لایه، چند لایه، و ماشین بردار پشتیبان با هسته چند جمله‌ای و شعاعی.

نتایج نشان‌دهنده عملکرد بهتر الگوریتم لاسو در انتخاب متغیرهای تاثیرگذار در فاز کلاس‌بندی با استفاده از الگوریتم ماشین بردار پشتیبان با هسته شعاعی بوده است. به این ترتیب، نتایج با یافته‌های راعی و فلاح‌پور (۱۳۸۷) و محمدیان حاجی کُرد و همکاران (۱۳۹۵) در عملکرد بهتر الگوریتم ماشین بردار پشتیبان برای حل مسائل طبقه‌بندی و شناسایی ریسک اعتباری، در مقایسه با سایر روش‌ها، تطابق دارد.

با توجه به دامنه متفاوت متغیرهای انتخاب‌شده موثر بر مطالبات غیرجاری در قالب الگوریتم لاسو در مقایسه با الگوریتم تجزیه و تحلیل مولفه‌های همسایگی و عملکرد بهتر الگوریتم لاسو متمرکز بر مجموعه متغیرهای مالی، اقتصادی و حسابرسی، توصیه می‌شود. در مدل‌سازی مطالبات غیرجاری، اهمیت ویژه‌ای برای متغیرهای مالی، اقتصادی و حسابرسی در مقایسه با مدل‌های مبتنی بر متغیرهای غیرمالی و خصوصیتی قراردادها در نظر گرفته شود.

از این‌رو، پیشنهاد می‌شود کارشناسان اعتباری، در تهیه و تنظیم گزارش اعتباری، ارکان اعتباری تسهیلات، جهت تصمیم‌گیری پرداخت تسهیلات، کارشناسان اداره وصول مطالبات، در تهیه و تنظیم گزارش کارشناسی و ارکان وصول مطالبات، در ارزیابی توانایی مالی شرکت‌ها جهت تقسیط مجدد مانده بدهی قراردادها، ضمن مطالعه صورت‌های مالی حسابرسی شده و متغیرهای کلیدی آن، در تجزیه و تحلیل مالی شرکت‌ها، از نسبت‌های مالی منتخب در این مقاله در کنار متغیرهای اقتصادی برای پیش‌بینی دقیق‌تر توانایی مالی شرکت‌ها و احتمال غیرجاری شدن مطالبات مشتری استفاده کنند. البته استفاده از الگوریتم ماشین‌های بردار

پشتیبان با هسته شعاعی در کلاس‌بندی با توجه عملکرد بهتر آن در این مطالعه نیز پیشنهاد می‌شود.

منابع

- ابراهیمی، مرضیه، دریابر، عبداله (۱۳۹۱). مدیریت ریسک اعتباری در نظام بانکی- رویکرد تحلیل پوششی داده‌ها و رگرسیون لجستیک و شبکه عصبی. *فصلنامه دانش سرمایه‌گذاری*، ۱ (۲): ۶۲-۳۵.
- اداره بررسی‌ها و سیاست‌های اقتصادی (۱۳۸۵-۱۳۹۷). نماگرهای اقتصادی، بانک مرکزی جمهوری اسلامی ایران.
- اداره مطالعات و مقررات بانکی (۱۳۸۵). دستورالعمل طبقه‌بندی دارایی‌های موسسات اعتباری، بانک مرکزی جمهوری اسلامی ایران.
- اداره مطالعات و مقررات بانکی (۱۳۹٤). آیین‌نامه وصول مطالبات غیرجاری موسسات اعتباری (ریالی و ارزی)، بانک مرکزی جمهوری اسلامی ایران.
- اکبری، سپیده، حاجیها، زهره (۱۳۹۵). رابطه ریسک ورشکستگی مالی و اظهارنظر حسابرس در مورد ابهام در تداوم فعالیت. ششمین کنفرانس ملی و چهارمین کنفرانس بین‌المللی حسابداری و مدیریت با رویکردهای کاربردی و پژوهش نوین، بهمن ۱۳۹۵، تهران.
- تاری، فتح‌اله، ابراهیمی، سید احمد، موسوی، سید جعفر، کلانتری، محمود (۱۳۹٦). مقایسه مدل‌های شبکه عصبی، الگوریتم ژنتیک و لاجیت در ارزیابی ریسک اعتباری مشتریان. *فصلنامه پژوهش‌های پولی- بانکی*، ۱۰ (۳٤) : ٦۸۰-٦۵۷.
- پورکاظمی، محمدحسین، صداقت پرست، الدار، ده پناه، رضا (۱۳۹٦). برآورد احتمال نکول مشتریان حقیقی بانک با استفاده از روش شبکه‌های عصبی (مطالعه موردی: بانک پاسارگاد). *فصلنامه مطالعات مالی و بانکداری اسلامی*، ۲ (٦۷): ۲٤-۱.
- جعفری اسکندری، میثم، روحی، میلاد (۱۳۹٦). مدیریت ریسک اعتباری مشتریان بانکی با استفاده از روش ماشین بردار تصمیم بهبود یافته با الگوریتم ژنتیک با رویکرد داده کاوی. *فصلنامه مدیریت دارایی و تامین مالی*، ۵ (٤): ۳۲-۱۷.
- حبیبی، رضا، کوهی، حسن، بعیدی، حسین (۱۳۹۷). تصمیمات تسهیلات‌دهی بانک با استفاده از روش الگوریتم ژنتیک (مطالعه موردی: مشتریان حقیقی بانک سپه). *فصلنامه مطالعات مالی و بانکداری اسلامی*، ٤ (۹): ۷۱-۳۳.

– حکیمی‌پور، نادر (۱۳۹۷). ارزیابی عوامل تاثیرگذار بانکی بر مطالبات غیرجاری بانک‌های ایران (رویکرد مدل پانل پویا GMM). *فصلنامه اقتصاد مالی*، ۱۲ (٤۲): ۱۱۹–۹۹.

– داداحمدی، دانیال، احمدی، عباس (۱۳۹۳). رتبه بندی مشتریان بانک با استفاده از شبکه عصبی با اتصالات جانبی. *فصلنامه توسعه مدیریت پولی و بانکی*، ۲ (۳): ۲۸–۱.

– راعی، رضا، فلاح‌پور، سعید (۱۳۸۷). کاربرد ماشین بردار پشتیبان در پیش‌بینی درماندگی مالی شرکت‌ها با استفاده از نسبت‌های مالی، *بررسی‌های حسابداری و حسابرسی*، ۱۵ (۵۳): ۳٤–۱۷.

– سیدشکری، خشایار، گروسی، سمیه (۱۳۹٤). بررسی عوامل موثر بر مطالبه‌های غیرجاری در نظام بانکی کشور، *فصلنامه اقتصاد مالی*، ۹ (۳۱): ۱۱۸–۹۵.

– رستمی، محمدرضا، نبی‌زاده، احمد، شاهی، زهرا (۱۳۹۷). بررسی عوامل موثر بر ریسک اعتباری بانک‌های تجاری ایران با تاکید بر عوامل خاص بانکی و کلان اقتصادی. *فصلنامه مدیریت دارایی و تامین مالی*، ٦ (٤): ۹۲–۷۹.

– رضایی، فرزین، ویسی‌حصار، ثریا، قندچی، فرشین (۱۳۹۸). رابطه حق‌الزحمه حسابرسی با فرصت‌های رشد و بازده غیرعادی شرکت‌ها. *فصلنامه دانش حسابرسی*، ۱۹ (۷٦): ۱۲۵–۸۷.

– فلاح‌پور ، سعید، راعی، رضا، هندیجانی‌زاده، محمد (۱۳۹۳). رویکرد شبکه عصبی مبتنی بر کلونی زنبور عسل مصنوعی جهت تخمین رتبه اعتباری مشتریان بانک‌ها. *فصلنامه مهندسی مالی و مدیریت اوراق بهادر*، ۵ (۲۱): ۵۳–۳۳.

– محمدیان حاجی کرد، امین، اصغرزاده زعفرانی، ملیحه، امام دوست، مصطفی (۱۳۹۵). بررسی ریسک اعتباری مشتریان حقوقی با استفاده از مدل ماشین بردار پشتیبان و مدل هیبریدی الگوریتم ژنتیک مطالعه موردی بانک تجارت. *فصلنامه مهندسی مالی و مدیریت اوراق بهادار*، ۷ (۲۷): ۳۲–۱۷.

– مزرعه فراهانی، نقی (۱۳۹۵). بحران مالی، ورشکستگی و بند ابهام درباره تداوم فعالیت شرکت در اظهارنظر حسابرس. پایان‌نامه کارشناسی ارشد، دانشکده حسابداری، دانشگاه آزاد اسلامی، واحد شهر قدس.

– منصورفر، غلامرضا، پیری، پرویز، علیخانی، زهرا، اسدی، مریم (۱۳۹۷). پیش‌بینی درماندگی مالی با توجه به اثرات بندهای تعدیلی گزارش حسابرس مستقل. شانزدهمین همایش ملی حسابداری ایران، اصفهان.

– میرغفوری، سید حبیب‌اله، امین‌آشوری، زهره (۱۳۹٤). ارزیابی ریسک اعتباری مشتریان بانک‌ها. *فصلنامه کاوش‌های مدیریت بازرگانی*، ۷ (۱۳): ۲٤۷–۲٦٦.

- نیلی، فرهاد و محمودزاده، امینه (۱۳۹۳). مطالبات غیرجاری یا دارایی‌های مسموم بانک‌ها، پژوهشکده پولی و بانکی، بانک مرکزی جمهوری اسلامی ایران، ۹۳۰۲۵.

- Abe, S. (2010). Support vector machines for pattern classification, second edition. Springer-Verlag, London.

- Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks*, 6(2), 1-38.

- Altman, E. I., (1968). Financial ratios discriminate analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609.

- Bekhet, H., Eletter, S. (2012). Credit risk management for the jordanian commercial banks: A business intelligence approach, *Australian Journal of Basic and Applied Sciences*, 6(9), 188–195.

- Bellotti, T. (2010). A simulation study of basel ii expected loss distributions for a portfolio of credit cards, *Journal of Financial Service Marketing*, 14(4), 268-277.

- Blanco, A., Mejias, R., Lara, J., and Rayo, S. (2013). Credit Scoring Models for the Microfinance Industry using Neural Networks: Evidence from Peru, *Expert Systems with Applications*, 40 (1), 356–364.

- Fonti, V., & Belitser, E. (2017). Feature selection using LASSO, research paper in business analytics, Vrije Universiteit Amsterdam.

- Karaa, A., Krichene, A., (2012). Crediti–risk assessment using support vector machine and multilayer neural network models: A comparative study - case of Tunisian Bank. *Journal of Accounting and Management Information Systems*, 11(4), 587-620.

- Kim, K. J., & Ahn, H. (2012). A corporate credit rating model using multiclass support vector machines with an ordinal pairwise partitioning approach, *Computers & Operations Research*, 39(8), 1800-1811.

- Kohonen, T. (1998). Learning vector quantization, the handbook of brain theory and neural networks, Cambridge, MIT Press, Cambridge, MA, USA, pp, 37–540.

- Reinhart, M. C., & Rogoff, K. S. (2010), From financial crash to debt crisis, NBER Working Paper. 15795.

- Olson, D. L., & Delen, D. (2008). Advanced data mining techniques, Berlin Heidelberg, Springer.

- Svozil, D., Kvasnicka, V., & Pospichal, J., (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, 39(1), 43-62.

- Wang, G., & Ma, J. (2012). A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine. *Expert Systems with Applications*, 39(5), 5325-5331.