فصلنامه راهبرد توسعه/ سال بیستم و یکم/ شماره ۳ (پیاپی۸۳)/ پاییز ۱٤۰٤/ ۲۷۰-۲۳۱
Quarterly Journal of Development Strategy, ۲۰۲۵, Vol. ۲۱, No.۳ (۸۳)

رتبه‌بندی اعتباری مبتنی بر یادگیری عمیق حساس به هزینه: ارائه چارچوب ترکیب وزنی بهینه الگوریتم‌های نا همگن

مهدی فرضی[۱]
یعقوب پور کریم[۲]
سید علی پایتخی اسکوئی[۳]
مهدی زینالی[٤]
رسول برادران حسن زاده[٥]

(تاریخ دریافت ۱۴۰۳/۱۰/۲۱ ـ تاریخ تصویب ۱۴۰۴/۱/۲۳)
نوع مقاله: **علمی پژوهشی**

چکیده
رتبه‌بندی اعتباری فرایندی است که توسط تسهیلات‌دهندگان، مانند بانک‌ها یا شرکت‌های کارت اعتباری، برای ارزیابی توانایی بازپرداخت تسهیلات توسط افراد یا کسب‌وکارها به کار

[۱]- گروه حسابداری، واحد تبریز، دانشگاه ازاد اسلامی، تبریز، ایران.
[۲]- گروه حسابداری، واحد تبریز، دانشگاه ازاد اسلامی، تبریز، ایران.(نویسنده مسئول).
pourKarim @iaut.ac.ir
[۳]- گروه اقتصاد، واحد تبریز، دانشگاه ازاد اسلامی، تبریز، ایران.
[۴]- گروه حسابداری، واحد تبریز، دانشگاه ازاد اسلامی، تبریز، ایران.
[۵]- گروه حسابداری، واحد تبریز، دانشگاه ازاد اسلامی، تبریز، ایران.

می‌رود. این ارزیابی با بررسی عواملی نظیر سابقه پرداخت، میزان استفاده از اعتبار، طول سابقه اعتباری و نوع اعتبار مصرفی انجام می‌شود. مدل‌های متنوعی برای رتبه‌بندی اعتباری توسعه یافته‌اند که از الگوریتم‌های پیچیده برای محاسبه امتیاز اعتباری بهره می‌برند. در سال‌های اخیر، استفاده از یادگیری عمیق در این حوزه به دلیل دقت بالا و توانایی تحلیل داده‌های پیچیده، رشد چشمگیری داشته است. شبکه عصبی عمیق **LSTM**، با بهره‌مندی از لایه فراموشی و حافظه بلندمدت، عملکرد برتری نسبت به سایر شبکه‌های عصبی مشابه ارائه می‌دهد. این پژوهش، یک مدل یادگیری عمیق یکپارچه برای رتبه‌بندی اعتباری مشتریان بانکی پیشنهاد می‌دهد که بر پایه یادگیری جمعی و بهینه‌سازی همزمان پارامترها و انتخاب ویژگی‌ها با الگوریتم ژنتیک (**GA**) طراحی شده است. نوآوری کلیدی این مدل، ترکیب مستقل چهار الگوریتم **AdaBoost**، **Bagging**، **Random Forest** و **LSTM** و ادغام خروجی‌های آن‌ها از طریق یک ترکیب وزنی خطی بهینه‌شده است. این رویکرد، با بهره‌گیری از نقاط قوت هر الگوریتم و بهینه‌سازی هایپرپارامترها و ویژگی‌ها، دقت پیش‌بینی را به‌طور قابل توجهی افزایش می‌دهد. ارزیابی مدل پیشنهادی بر روی سه مجموعه داده استاندارد از ایران، استرالیا و آلمان انجام شد. نتایج نشان داد که این روش، بهبود چشمگیری در معیارهای دقت (**ACC**)، صحت (**Precision**)، فراخوانی (**Recall**)، شاخص **F۱** و کاهش خطای طبقه‌بندی (**MC**) ایجاد کرده است. به‌طور خاص، برای داده‌های ایران، دقت به ۰.۹۳۴ رسید که خطای طبقه‌بندی را به ۶.۶٪ کاهش داد. برای داده‌های استرالیا و آلمان نیز دقت به ترتیب به ۰.۹۰۲ و ۰.۸۴۳ دست یافت که نسبت به مدل‌های پیشین در بازه ۲۰۲۳ تا ۲۰۲۵ برتری قابل توجهی دارد. این یافته‌ها، پتانسیل بالای مدل پیشنهادی را در بهبود تصمیم‌گیری اعتباری، کاهش ریسک عدم بازپرداخت و تقویت پایداری مالی بانک‌ها تأیید می‌کند.

واژه‌های کلیدی: رتبه‌بندی اعتباری- یادگیری عمیق- شبکه **LSTM**- یادگیری گروهی- الگوریتم ژنتیک- پایداری مالی

۱. مقدمه

بانک‌ها به عنوان موتور محرکه اقتصاد، با تجهیز پس‌اندازها و تخصیص آنها به فعالیت‌های تولیدی و تجاری، نقشی حیاتی در رشد اقتصادی ایفا می‌کنند. اما ذات این فعالیت‌ها با ریسک اعتباری گره خورده است؛ ریسکی که ناشی از احتمال عدم بازپرداخت تسهیلات توسط وام‌گیرندگان است و در صورت تحقق، نه تنها سودآوری بانکها را کاهش می‌دهد، بلکه ثبات کل نظام مالی را تهدید می‌کند. مطالعات نشان می‌دهد که بخش عمده دارایی‌های بانک‌ها را مطالبات تشکیل می‌دهد و شکست در مدیریت این دارایی‌ها می‌تواند به کاهش حقوق صاحبان سهام و حتی ورشکستگی بینجامد (داس[۱]، ۲۰۱۹). در پاسخ به این چالش، کمیته بازل با ارائه چارچوبهای نظارتی، تحولی اساسی در مدیریت ریسک اعتباری ایجاد کرد. توافقنامه بازل **II** در سال‌های آغازین قرن ۲۱، با تأکید بر «حساسیت به ریسک»، الزامات سرمایه‌ای بانک‌ها را از حالت ثابت و یکسان خارج کرد و آن را به رتبه‌بندی اعتباری وام‌گیرندگان و پرتفوی اعتباری بانک‌ها وابسته ساخت. این توافقنامه، بانک‌ها را ملزم کرد تا از دو روش اصلی برای ارزیابی ریسک استفاده کنند: استفاده از مؤسسات رتبه‌بندی اعتباری یا توسعه مدل‌های داخلی مبتنی بر داده‌های تاریخی. هدف این بود که سرمایه نگهداری شده توسط بانک‌ها، نه تنها متناسب با حجم فعالیت‌ها، بلکه متناسب با سطح ریسک هر وام باشد (لاندو[۲]، ۲۰۰۹).

اما چگونه می‌توان ریسک اعتباری را به صورت عملیاتی کاهش داد؟ یکی از کارآمدترین ابزارها، سیستم‌های امتیازدهی اعتباری است که با تحلیل عوامل کیفی و کمی مانند سابقه مالی مشتری، درآمد، نسبت بدهی به دارایی، و ثبات شغلی، احتمال نکول را پیش‌بینی می‌کند (جیحونی‌پور و همکاران، ۱۴۰۳). این سیستم‌ها با تبدیل ریسک به یک متغیر کمی، به بانک‌ها اجازه می‌دهند تا مشتریان پرریسک را شناسایی کرده و نرخ سود وام‌ها را متناسب با سطح ریسک تعیین کنند. پژوهشها حاکی از آن است که استفاده از این روش‌ها تا ۳۰ درصد از

---

[۱] Das

[۲] Lando

هزینه‌های ناشی از مطالبات معوقه می‌کاهد (توماس[۱] و همکاران، ۲۰۱۳).

با این حال، در جهان امروز که بانک‌ها با چالش‌هایی مانند بین‌المللی شدن فعالیت‌ها، نوآوری‌های مالی پیچیده، و بحران‌های غیرمنتظره روبه‌رو هستند، مدیریت ریسک اعتباری نیازمند رویکردی پویاتر است. بحران مالی ۲۰۰۸ نمونه بارزی بود که نشان داد حتی مدل‌های پیشرفته نیز در صورت بی‌توجهی به عوامل کلان اقتصادی و «ریسک سیستماتیک» می‌توانند شکست بخورند. از این رو، بانک‌ها امروزه نه تنها بر رتبه‌بندی فردی وام‌گیرندگان تمرکز می‌کنند، بلکه با تحلیل روندهای اقتصادی، سیاست‌های پولی، و حتی تحولات ژئوپلیتیک، سعی در پیش‌بینی ریسک‌های نوظهور دارند(ندری و محرابی، ۱۳۹۷).

مدیریت ریسک اعتباری را باید فرایندی مستمر و چندبعدی دانست که ترکیبی از مقررات هوشمند (مانند بازل **II**)، فناوری‌های تحلیلی (مانند هوش مصنوعی در امتیازدهی)، و الگوهای نوین مالی را می‌طلبد. موفقیت در این حوزه، علاوه بر حفظ ثبات و پایداری مالی بانک‌ها، به جذب سرمایه‌گذاران، کاهش هزینه‌های تأمین مالی، و تقویت نقش بانک‌ها به عنوان نهادهای پیشران توسعه پایدار می‌انجامد.

رتبه‌بندی اعتباری مشتریان به عنوان یکی از ابزارهای کلیدی مدیریت ریسک اعتباری، با استفاده از روش‌های متنوعی از جمله الگوریتم‌های ژنتیک، ماشین یادگیری، روش‌های بیزی، تحلیل عاملی، و روش‌های هیبریدی انجام می‌شود. هر یک از این روش‌ها با مزایا و محدودیت‌های خاص خود، نیازمند انتخاب هوشمندانه مبتنی بر شرایط سازمانی و ماهیت داده‌ها هستند. در سال‌های اخیر، یادگیری عمیق به عنوان زیرمجموعه‌ای از هوش مصنوعی، با استفاده از شبکه‌های عصبی مصنوعی عمیق، تحولی در تحلیل داده‌های پیچیده و غیرخطی ایجاد کرده است. این روش با استخراج خودکار ویژگی‌های ارزشمند از داده‌های ورودی، امکان پیش‌بینی دقیق‌تر ریسک اعتباری را فراهم می‌کند و نیاز به تعریف دستی متغیرها را کاهش می‌دهد (گونارسون[۲] و

---

[۱] Thomas

[۲] Gunnarsson

همکاران، ۲۰۲۱). شبکه‌های عصبی عمیق بازگشتی[۱] (**RNN**) و شبکه‌های حافظه کوتاه ـ بلندمدت[۲] (**LSTM**) به عنوان نمونه‌هایی پیشرفته از معماری‌های یادگیری عمیق، توانایی منحصربه‌فردی در پردازش داده‌های دنباله‌ای و زمانی دارند. **RNN** با حفظ وابستگی‌های زمانی بین داده‌ها، امکان مدل‌سازی الگوهای پویا مانند رفتار مالی مشتریان در طول زمان را فراهم می‌کند. با این حال، مشکل فراموشی اطلاعات بلندمدت در **RNN** منجر به توسعه **LSTM** شد که با استفاده از واحدهای دروازه‌ای (گیت[۳]ها)، امکان ذخیره و بازیابی اطلاعات حیاتی در بازه‌های زمانی طولانی‌تر را دارد. این قابلیت، **LSTM** را به ابزاری کارآمد در رتبه‌بندی اعتباری تبدیل کرده است، به ویژه زمانی که وابستگی‌های زمانی پنهان در داده‌ها (مانند سابقه پرداخت‌های مشتری) نقش کلیدی در پیش‌بینی ریسک ایفا می‌کنند (دو و شو[۴]، ۲۰۲۲).

اگرچه مدل‌های یادگیری عمیق به تنهایی می‌توانند دقت بالایی در برخی مجموعه‌داده‌ها نشان دهند، اما وابستگی به ساختار داده و حساسیت به عدم تعادل کلاس‌ها (مانند داده‌های نامتوازن ریسک اعتباری) از چالش‌های اصلی آنهاست. برای مثال، یک کلاسیفایر مبتنی بر **LSTM** ممکن است در داده‌های متوازن عملکردی عالی داشته باشد، اما در داده‌های نامتوازن (که نمونه‌های نکول کمتر از مشتریان سالم هستند) به سمت کلاس اکثریت سوگیری کند. برای حل این مشکل، یادگیری گروهی[۵] به عنوان راهکاری مؤثر پیشنهاد شده است. در این روش، ترکیبی از چندین مدل پایه مانند **LSTM**، شبکه‌های کانولوشنی، یا درخت تصمیم برای بهبود دقت و کاهش واریانس خطا استفاده می‌شوند (آلاراج[۶] و همکاران، ۲۰۲۲).

در این پژوهش، با هدف رتبه‌بندی اعتباری مشتریان یک بانک ایرانی، از یادگیری عمیق گروهی حساس به هزینه استفاده شده است که به صورت یکپارچه با الگوریتم ژنتیک (**GA**) هینه‌سازی

---

[۱] Recurrent Neural Network

[۲] Long Short-Term Memory

[۳] Gate

[۴] Du & Shu

[۵] Ensemble Learning

[۶] Ala’raj

می‌شود. این چارچوب پیشنهادی دو مرحله اصلی دارد:

۱. انتخابی ویژگی بهینه: با استفاده از **GA**، متغیرهای کلیدی از داده‌های خام (مانند سن، درآمد، سابقه اعتباری، و نسبت بدهی به دارایی) استخراج می‌شوند تا ابعاد داده کاهش یابد و نویز حذف شود.

۲. ترکیب وزنی مدل‌ها: مدل‌های پایه (شامل **LSTM**، **AdaBoost**، **Bagging** و جنگل تصادفی[۱] (**RF**) با وزن‌های بهینه‌شده توسط **GA** ترکیب می‌شوند تا یک خروجی گروهی با دقت بالاتر تولید شود.

برای ارزیابی مدل، از دو مجموعه داده بنچمارک **UCI** (داده‌های متوازن استرالیا و نامتوازن آلمان) و همچنین داده‌های واقعی مشتریان یک بانک ایرانی استفاده شده است. به منظور پیاده‌سازی مدل، از محیط پایتون و کتابخانه‌های **TensorFlow**، **Keras**، و **Scikit-learn** استفاده می‌شود.

۲. ادبیات و پیشینه تحقیق

در دنیای مدیریت مالی و بانکداری، ارزیابی دقیق اعتبار مشتریان به عنوان یکی از وظایف بنیادین و حیاتی موسسات مالی و اعتباری مطرح است؛ فرآیند رتبه‌بندی اعتباری با تحلیل دقیق عوامل ریسک و احتمال بازپرداخت بدهی‌ها، به بانک‌ها و موسسات اعتباری کمک می‌کند تا با کاهش ریسک‌های ناشی از اعطای تسهیلات، بتوانند استراتژی‌های سودآوری و رشد پایداری را در فعالیت‌های خود به کار گیرند. این ارزیابی، که بر پایه بررسی داده‌های گذشته و شرایط مالی و اقتصادی مشتریان انجام می‌شود، امکان شناسایی الگوهای رفتاری و پیش‌بینی دقیق‌تر نسبت به توانایی بازپرداخت بدهی‌ها را فراهم می‌آورد و به عنوان ابزاری کلیدی در مدیریت ریسک اعتباری مطرح است. ریسک اعتباری، به معنای احتمال عدم توانایی مشتریان در انجام تعهدات مالی خود، یکی از چالش‌های اساسی در حوزه مالی محسوب می‌شود. همانطور که

---

[۱] Random Forest

باتاچاریا[۱] و همکاران (۲۰۲۳) بیان کرده‌اند، این ریسک می‌تواند تأثیرات منفی گسترده‌ای بر پایداری سیستم‌های بانکی داشته باشد؛ به همین دلیل، موسسات مالی از ابزارها و روش‌های متعددی مانند ارزیابی اعتبار، تنظیم سیاست‌های محدودسازی ریسک، تنوع‌بخشی به سبد وام‌ها و استفاده از بیمه‌های اعتباری بهره می‌گیرند. گزارش‌های کمیته بال (بلکی و روتوفسکی[۲]، ۲۰۱۳) نیز به نقش مهم اعطای تسهیلات به عنوان منبع اصلی ریسک اعتباری اشاره دارند؛ این تسهیلات در صورت نوسان شرایط بازپرداخت مشتریان، می‌توانند باعث تغییرات عمده در ارزش اوراق و ابزارهای مشتقه شوند که این امر مستقیماً بر سلامت سیستم بانکی تأثیرگذار است.

با پیشرفت فناوری‌های نوین اطلاعاتی و سیستم‌های رایانه‌ای، روند رتبه‌بندی اعتباری دستخوش تحولاتی چشمگیر شده است. در گذشته که سیستم‌های ارزیابی متنوع و پراکنده موجب ایجاد پیچیدگی در امتیازدهی می‌شد، امروزه با بهره‌گیری از تحلیل‌های پیشرفته، داده‌های بزرگ و الگوریتم‌های هوشمند، امکان ایجاد مدل‌های پیش‌بینی دقیق‌تر فراهم آمده است. این تحولات موجب شده تا علاوه بر ارزیابی وضعیت کنونی مشتریان، توانایی پیش‌بینی رفتارهای آینده و شناسایی زودهنگام ریسک‌های احتمالی نیز بهبود یابد. به عبارتی دیگر، امروزه علاوه بر بررسی معیارهایی نظیر تاریخچه پرداخت، سوابق اعتباری، وضعیت دارایی‌ها و بدهی‌ها و درآمدهای ثابت، عوامل دیگری مانند ثبات شغلی، الگوهای مصرفی و حتی شاخص‌های اقتصادی کلان مورد توجه قرار می‌گیرند (کان[۳] و همکاران، ۲۰۲۴).

مؤسسات مالی و بانک‌ها از مدل‌های امتیازدهی به دو صورت بهره‌مند می‌شوند: از یک سو رتبه‌بندی خارجی که توسط سازمان‌های معتبر بین‌المللی مانند استاندارد اند پورز، فیچ و مودیز ارائه می‌شود و از سوی دیگر رتبه‌بندی داخلی که بر اساس سیستم‌های اختصاصی و سیاست‌های خاص هر بانک تنظیم می‌گردد. در این میان، موسسه استاندارد اند پورز رتبه‌بندی اعتباری را به عنوان اظهار نظر درباره اعتبار یک بدهکار بر مبنای عوامل ریسک تعریف کرده و از دیدگاه

---

[۱] Bhattacharya

[۲] Bielecki & Rutkowski

[۳] Cahn

موسسه مودیز، این امتیازدهی بیانگر توانایی آینده بدهکار در بازپرداخت به موقع اصل و سود است (مون و استوسکی[۱]، ۱۹۹۳؛ اونگ[۲] و همکاران، ۲۰۰۵).

از سویی دیگر، مدیریت ریسک اعتباری نه تنها به عنوان ابزاری جهت پیش‌بینی احتمال نکول مشتریان به کار گرفته می‌شود، بلکه در تعیین سیاست‌های کلان مدیریتی نیز نقش مؤثری دارد. احمد و راجالکسیمی[۳] (۲۰۱۹) سه دسته سیاست کلان در این زمینه معرفی کرده‌اند: سیاست‌های محدودسازی یا کاهش ریسک اعتباری که شامل تمرکز بر اعطای تسهیلات به مشتریان منتخب است؛ سیاست‌های طبقه‌بندی دارایی‌ها و ارزیابی‌های مستمر که به کمک آن‌ها توانایی وصول پرتفوی وام‌ها سنجیده می‌شود؛ و سیاست‌های کاهش زیان که از طریق اقدامات محتاطانه و اعطای اختیارات مدیریتی مناسب، تلاش در حذف یا کاهش خسارات احتمالی دارند. این رویکردها به همراه نظارت دقیق مقامات نظارتی بر سرمایه و اعتبار بانک‌ها، چارچوبی مستحکم برای مقابله با نوسانات و بحران‌های احتمالی فراهم می‌کنند (نائیلی و لاهیرچی[۴]، ۲۰۲۰).

توسعه سیستم‌های رتبه‌بندی اعتباری از دهه ۱۹۵۰ میلادی آغاز شده و با گذر زمان به واسطه پیشرفت‌های فناوری و دسترسی به داده‌های دقیق‌تر، توانسته‌اند نقش اساسی در تصمیم‌گیری‌های اعتباری داشته باشند. در این بین، استفاده از داده‌های بنچمارک و به کارگیری الگوریتم‌های پیشرفته به بهبود دقت پیش‌بینی‌ها کمک شایانی نموده است؛ هرچند که همچنان هیچ مدلی قادر به پیش‌بینی کامل و بدون اشتباه رفتار مشتریان در شرایط متغیر اقتصادی نیست. بنابراین، ترکیب تحلیل‌های آماری با تجربیات مدیریتی و نظارت‌های پیوسته از الزامات اساسی در مدیریت ریسک اعتباری محسوب می‌شود. این رویکرد جامع و چندبعدی در نهایت به حفظ پایداری بانک‌ها و تأمین سلامت اقتصاد کلان کمک می‌کند، چرا که با کاهش ریسک‌های ناشی از عدم بازپرداخت، زمینه ایجاد اعتماد در سیستم مالی و افزایش سرمایه‌گذاری‌های اقتصادی فراهم

---

[۱] Moon & Stotsky

[۲] Ong, Huang & Tzeng

[۳] Ahmed & Rajaleximi

[۴] Naili & Lahrichi

می‌آید (آدی[۱] و همکاران، ۲۰۲۴).

فرآیند رتبه‌بندی اعتباری به عنوان یکی از ارکان اساسی در تحلیل ریسک اعتباری مشتریان، در طول سالیان متمادی دستخوش تحول و تکامل عمیقی شده است. در دوران اولیه، بانک‌ها برای ارزیابی اعتبار مشتریان از رویکردهای سنتی و تجربی بهره می‌گرفتند؛ به عبارت دیگر، تصمیم‌گیری بر اساس روابط شخصی و اعتماد متقابل از اصول بنیادین در تعیین اعتبار محسوب می‌شد. با رشد صنعت بانکی و افزایش پیچیدگی معاملات مالی، ضرورت به کارگیری روش‌های مدلسازی دقیق‌تر و علمی‌تر بیش از پیش احساس گردید؛ به همین منظور، در این دوره برخی مدل‌های آماری مبتنی بر اصول ریاضی و احتمالات معرفی شدند (جوشی[۲]، ۲۰۲۵).

ورود فناوری‌های نوین محاسباتی و افزایش توان پردازشی کامپیوترها موج جدیدی از روش‌های رتبه‌بندی اعتباری را به وجود آورد. در این راستا، ابزارهایی مانند شبکه‌های عصبی مصنوعی و درخت‌های تصمیم‌گیری مطرح شدند که توانستند با ارائه تحلیل‌های دقیق‌تر، جایگزین روش‌های تجربی گردند. از دهه ۱۹۹۰ به بعد، کاربرد روش‌های یادگیری ماشین در رتبه‌بندی اعتباری آغاز شد؛ الگوریتم‌هایی نظیر **SVM** و انواع شبکه‌های عصبی به عنوان گام‌های نخست در این مسیر به کار گرفته شدند. ورود به دهه ۲۰۰۰ و ظهور روش‌های یادگیری عمیق، امکان استفاده از شبکه‌های عصبی عمیق مانند **LSTM** را فراهم ساخت؛ این شبکه‌ها قادر به مدل‌سازی و استخراج الگوهای پیچیده از داده‌های بزرگ و متنوع هستند (موفو و موکوسرا[۳]، ۲۰۱۴).

امروزه، تکنیک‌های یادگیری عمیق به ویژه شبکه‌های عصبی بازگشتی مانند **LSTM** و سایر روش‌های مدلسازی پیشرفته، به عنوان ابزارهای قدرتمندی در رتبه‌بندی اعتباری شناخته می‌شوند. این روش‌ها با دقت بالا در شناسایی الگوهای نهفته و پیچیده، نقش اساسی در بهبود ارزیابی اعتبار مشتریان و مدیریت ریسک اعتباری ایفا می‌کنند. از سوی دیگر، رویکردهای فراابتکاری

---

[۱] Addy

[۲] Joshi

[۳] Mpofu & Mukosera

که مبتنی بر الگوریتم‌های تکاملی، ژنتیک، جستجوی همزمان و الگوریتم‌های شبه‌تصادفی هستند، با ترکیب چندین مدل و الگوریتم در جستجوی پاسخ‌های بهینه، توانسته‌اند دقت و کارایی فرآیند رتبه‌بندی را افزایش دهند. مزایای این روش‌ها شامل دقت بالاتر، تطبیق بهتر با داده‌های پیچیده و متنوع، مقاومت در برابر نویز و بهینه‌سازی مؤثر پارامترها می‌باشد؛ به عبارتی، بهره‌گیری از این رویکردها به بهبود تصمیم‌گیری‌های مرتبط با ریسک اعتباری منجر می‌شود (المفتی[۱] و همکاران، ۲۰۲۳).

در طول چند دهه گذشته، پژوهش‌های متعددی در حوزه رتبه‌بندی اعتباری با هدف بهبود دقت پیش‌بینی ریسک اعتباری مشتریان انجام شده است. در سال ۱۳۸۹، اخباری و مخاطب‌رفیعی با به کارگیری مدل عصبی ـ فازی **(ANFIS)** توانستند با دستیابی به دقت طبقه‌بندی «۶۹/۳۶ درصد»، گامی ابتدایی در جهت استفاده از روش‌های ترکیبی در تحلیل ریسک بردارند (اخباری و مخاطب‌رفیعی، ۱۳۸۹). به دنبال این مطالعه، در سال ۱۳۹۲، دادمحمدی و احمدی از شبکه عصبی با تابع محرک شعاعی استفاده کردند که نتایج نشان داد این روش قادر به پیش‌بینی دقیق رفتار مشتریان می‌باشد (دادمحمدی و احمدی، ۱۳۹۲). در همان سال، اسراری با مقایسه رگرسیون لاجیت، شبکه عصبی و الگوریتم ژنتیک به دست آورد که این سه روش به‌طور نسبی از دقت مشابهی در پیش‌بینی ریسک اعتباری مشتریان حقیقی برخوردارند (اسراری، ۱۳۹۲). در ادامه، مهدوی و حری در سال ۱۳۹۴ الگوریتم فراابتکاری و هیبریدی چند معیاره مبتنی بر ترکیب شبکه عصبی فازی و کلونی مورچگان را ارائه دادند که سیستم پیشنهادی در «۸۹/۶۷ درصد» از موارد قادر به ارائه تخمین صحیح نسبت به رتبه اعتباری مشتریان بود (مهدوی و حری، ۱۳۹۴). سال ۱۳۹۵ شاهد مطالعه‌ای از گل‌محمدی بود که از ترکیب الگوریتم ژنتیک و رگرسیون لجستیک بهره برد و نتایج نشان‌دهنده برتری الگوریتم ژنتیک در این ترکیب گردید (گل‌محمدی، ۱۳۹۵). همچنین، کوچه‌باغی و رجبی بهجت در سال ۱۳۹۶ الگوریتم فراابتکاری جستجوی گرانشی باینری را معرفی کردند؛ نتایج این مطالعه از منظر معیار**F-Score** ، نشان از

---

[۱] Almufti

توان بالای این الگوریتم در فرایند طبقه‌بندی داشت (کوچه‌باغی و رجبی بهجت، ۱۳۹۶). پاک‌بین و پویانفر در سال ۱۳۹۷ تکنیک طبقه‌بندی ترکیبی پویا و احتمال نرم **(DECSP)** را پیشنهاد کردند؛ یافته‌های این تحقیق حاکی از آن بود که روش **DECSP** نسبت به الگوریتم‌های کلاسیکی مانند **KNN** ، **ANN**، درخت تصمیم، **SVM** و **AdaBoost** عملکرد به مراتب بهتری دارد (پاک‌بین و پویانفر، ۱۳۹۷). در سال ۱۳۹۸، نظرآقایی و همکاران با به‌کارگیری رویکرد یادگیری گروهی از طریق ترکیب شبکه‌های عصبی، درخت تصمیم فازی و بگینگ، به اهمیت عوامل درآمد و تراکنش‌های مالی در تعیین ریسک اعتباری پی بردند و همچنین نشان دادند استفاده از درخت تصمیم فازی همراه با بگینگ دقت طبقه‌بندی را افزایش می‌دهد (نظرآقایی و همکاران، ۱۳۹۸). همان سال، حسامی با استفاده از روش یادگیری عمیق به این نتیجه رسید که این رویکرد می‌تواند به کاهش ریسک عدم پرداخت بدهی مشتریان منجر شود (حسامی، ۱۳۹۸). با گذر زمان، کاربرد روش‌های پیشرفته‌تر همچون یادگیری گروهی و بهبود یافته نیز مورد توجه قرار گرفت؛ ابتیاع و همکاران در سال ۱۴۰۰ با استفاده از یادگیری گروهی مبتنی بر ماشین بردار پشتیبان بر داده‌های بانک پاسارگاد ثابت کردند که این روش از روش‌های معمولی **SVM** و جنگل تصادفی برتری دارد (ابتیاع و همکاران، ۱۴۰۰). در سال ۱۴۰۲، خانجانی و همکاران به کاربرد گسترده هوش مصنوعی، به ویژه روش‌های یادگیری عمیق، در بهبود دقت و کیفیت تصمیم‌گیری مدل‌های رتبه‌بندی اعتباری پرداختند و نشان دادند این روش‌ها نسبت به روش‌های سنتی یادگیری ماشین عملکرد بهتری دارند (خانجانی و همکاران، ۱۴۰۲). از سوی دیگر، در مقطع زمانی بعدی، مطالعاتی نیز به استفاده از تکنیک‌های یادگیری جمعی پرداخته‌اند. به عنوان مثال، آبلان و مانتاس در سال ۲۰۱۴ با به‌کارگیری طرح **Bagging**در چند مدل درخت تصمیم‌گیری به افزایش دقت طبقه‌بندی در رتبه‌بندی اعتباری دست یافتند (آبلان و مانتاس[۱]، ۲۰۱۴). ژوانگ[۲] و همکاران در سال ۲۰۱۵ روش مبتنی بر شبکه بیزی و اطلاعات متقابل **(BNMI)** را پیشنهاد کردند که با اندازه‌گیری وابستگی متغیرهای

---

[۱] Abellán & Mantas

[۲] Zhuang

تصادفی، توانستند کارایی طبقه‌بندی را بهبود بخشند (ژوانگ و همکاران، ۲۰۱۵).
در سال ۲۰۱۶، محمدی و زنگنه با مقایسه چند نوع الگوریتم شبکه‌های عصبی از جمله **BP**، **Resilient**، **Conjugate Gradient**، **Gradient Descent**، **LM** و **BFGS Quasinewton** نشان دادند که شبکه عصبی مبتنی بر الگوریتم **LM** در ارزیابی ریسک اعتباری عملکرد بهتری نسبت به سایر شبکه‌ها داشته و از نظر دقت، نسبت به روش‌های رگرسیون لجستیک و درخت تصمیم برتری دارد (محمدی و زنگنه، ۲۰۱۶). همچنین، گلشنی، باقرزاده و داودی در همان سال با به‌کارگیری ترکیب تحلیل پوششی داده‌ها و تحلیل افتراقی (**DEA-DA**) اثبات کردند که روش پیشنهادی قادر به دسته‌بندی با دقت بالا مشتریان دریافت‌کننده وام در بانک‌های ایرانی می‌باشد (گلشنی، باقرزاده و داودی، ۲۰۱۶). در سال ۲۰۱۷، مطالعاتی نیز به مقایسه میان روش‌های مختلف پرداختند؛ آتیگری، پای و پای نشان دادند که در مقایسه بین رگرسیون لجستیک و شبکه عصبی، روش رگرسیون لجستیک عملکرد بهتری از خود ارائه می‌دهد (آتیگری[۱] و همکاران، ۲۰۱۷). در همین سال، التر و یاسین[۲] با مقایسه تجزیه و تحلیل خطی افتراقی**(LDA)** ، شبکه عصبی چندلایه پرسپترون **(MLP)** و درخت تصمیم‌گیری **CART**به این نتیجه رسیدند که مدل **MLP** بالاترین دقت **(ACC)** و کمترین میزان خطا را دارا می‌باشد (التر و یاسین، ۲۰۱۷). همچنین، ستوریسنو و هالیم[۳] با بهینه‌سازی روش رگرسیون لجستیک با الگوریتم **Nelder-Mead** نشان دادند که این روش نسبت به رگرسیون لجستیک کلاسیک کارایی بهتری دارد (ستوریسنو و هالیم، ۲۰۱۷). دومیتروسکو[۴] در سال ۲۰۱۸ روش ترکیبی رگرسیون لجستیک و درخت تصمیم غیرخطی را معرفی نمود و نتایج حاصل نشان دادند که درخت تصمیم غیرخطی نسبت به رگرسیون لجستیک عملکرد بسیار بهتری در رتبه‌بندی

---

[۱] Attigeri

[۲] Eletter & Yaseen

[۳] Sutrisno & Halim

[۴] Dumitrescu

اعتباری دارد (دومیتروسکو، ۲۰۱۸). در سال ۲۰۱۹، عبدو[۱] و همکاران با مقایسه تحلیل افتراقی، رگرسیون لجستیک، شبکه عصبی چند لایه رو به جلو و شبکه عصبی احتمالی دریافتند که روش‌های شبکه عصبی به‌طور کلی عملکرد بهتری از دو روش دیگر دارند و در این میان شبکه عصبی احتمالی بهترین نتایج را ارائه می‌دهد (عبدو و همکاران، ۲۰۱۹). در سال‌های اخیر، پژوهش‌ها به سمت استفاده از معماری‌های پیشرفته یادگیری عمیق متمایل شده‌اند. ستاری مالکی و همکاران در سال ۲۰۲۱ با بهره‌گیری از **LSTM Autoencoder** نشان دادند که این مدل در انتخاب ویژگی و دسته‌بندی درست مشتریان دقت بالایی از خود دارد (ستاری مالکی و همکاران، ۲۰۲۱). همچنین، دو و شو در سال ۲۰۲۲ با ترکیب یادگیری عمیق **RNN** و الگوریتم بهینه‌سازی **Bionic** ثابت کردند که الگوریتم پیشنهادی نسبت به روش‌های مرسوم یادگیری ماشین و ترکیب‌های مختلف آن عملکرد بهتری دارد (دو و شو، ۲۰۲۲). در سال ۲۰۲۳، وانگ[۲] و همکاران از طریق یادگیری تقویتی عمیق مبتنی بر تکرار تجربه اولویت‌دار طبقه‌ای متوازن **(DQN-BSPER)** نشان دادند که مدل پیشنهادی قادر است از چهار مدل معیار **DRL**و هفت مدل طبقه‌بندی سنتی بهتر عمل کند (وانگ و همکاران، ۲۰۲۳). ژانگ[۳] و همکاران نیز در همان سال یک مدل امتیازدهی اعتباری گروهی بر مبنای رگرسیون لجستیک با اثرات تعادلی و وزن‌دهی ناهمگن معرفی کردند که نتایج آن در مقایسه با ده مدل نماینده در شش مجموعه داده عمومی، قدرت تشخیص نمونه‌های پیش‌فرض و توانایی تعمیم بالاتری را نشان داد (ژانگ و همکاران، ۲۰۲۳). به علاوه، خلیلی و رستگار در سال ۲۰۲۳ مدل بهینه حساس به هزینه ترکیبی را با تلفیق مدل‌های **DT** ، **KNN**، **AdaBoost**، **SVM** و شبکه عصبی احتمالاتی **PNN** ارائه کردند؛ نتایج نشان داد که روش پیشنهادی در هر دو شرایط داده متوازن و نامتوازن دقت بالایی از خود نشان می‌دهد (خلیلی و رستگار، ۲۰۲۳).

این مرور ادبی نشان می‌دهد که طی سال‌ها، پژوهشگران با بهره‌گیری از رویکردهای مختلف از

---

[۱] Abdou

[۲] Wang

[۳] Zhang

مدل‌های سنتی آماری گرفته تا تکنیک‌های پیشرفته یادگیری عمیق و تقویتی، توانسته‌اند به بهبود دقت رتبه‌بندی اعتباری دست یابند. این پیشرفت‌ها نشان‌دهنده تمایل روزافزون به استفاده از هوش مصنوعی و روش‌های فراابتکاری در جهت مدیریت ریسک اعتباری و بهبود تصمیم‌گیری‌های مالی است که می‌تواند زمینه‌ساز توسعه سیستم‌های بانکی پربازده و مقاوم در برابر ریسک‌های ناشی از عدم پرداخت بدهی‌ها گردد.

۳. روش‌شناسی تحقیق

تحقیق حاضر از نظر هدف کاربردی و از نظر روش جز تحقیقات توصیفی است. داده‌های مورد استفاده در تحقیق نیز داده‌های مالی و اطلاعات دموگرافی اشخاص حقیقی یا حقوقی وام‌گیرنده از بانک‌ها می‌باشند. به منظور پیاده‌سازی مدل نیز از نرم‌افزار **Python** استفاده می‌شود.

۳.۱. داده‌ها

داده‌های مورد استفاده در این تحقیق شامل داده‌های مشتریان حقوقی یک بانک ایرانی و دو دیتاست بنچمارک آلمان و استرالیا هست. بررسی داده‌های سه کشور نشان می‌دهد که توزیع مشتریان خوش‌حساب و بدحساب از الگوهای متفاوتی پیروی می‌کند. در ایران، از مجموع ۲,۹۸۷ مشتری حقوقی، ۱,۷۲۱ مشتری (۵۷.۶۲٪) خوش‌حساب و ۱,۲۶۶ مشتری (۴۲.۳۸٪) بدحساب هستند که نسبت خوش‌حساب به بدحساب ۱.۳۶ به ۱ است. در مقابل، آلمان با ۱,۰۰۰ مشتری، نرخ خوش‌حسابی بالاتری دارد: ۷۰۰ مشتری خوش‌حساب (۷۰٪) در برابر ۳۰۰ بدحساب (۳۰٪) که نسبت آن ۲.۳۳ به ۱ می‌باشد. اما استرالیا وضعیت معکوسی دارد: از ۶۹۰ مشتری، ۳۰۷ مورد (۴۴.۴۹٪) خوش‌حساب و ۳۸۳ مورد (۵۵.۵۱٪) بدحساب هستند و نسبت خوش‌حساب به بدحساب ۰.۸۰ به ۱ است (جدول ۱). در جدول ۲ متغیرهای اعتباری مشتریان حقوقی بانک ایرانی و در جدول ۳ متغیرهای کیفی و مقادیر آنها نشان داده شده است.

جدول ۱. جدول مقایسه‌ای توزیع مشتریان در دیتاست‌های مورد مطالعه

کشور	کل مشتریان	خوش‌حساب (تعداد)	خوش‌حساب	بدحساب (تعداد)	بدحساب (٪)	نسبت (خوش‌حساب:بدحساب)
ایران	۲,۹۸۷	۱,۷۲۱	۵۷.۶۲٪	۱,۲۶۶	۴۲.۳۸٪	۱.۳۶:۱
آلمان	۱,۰۰۰	۷۰۰	۷۰٪	۳۰۰	۳۰٪	۲.۳۳:۱
استرالیا	۶۹۰	۳۰۷	۴۴.۴۹٪	۳۸۳	۵۵.۵۱٪	۰.۸۰:۱

منبع: نتایج تحقیق

جدول ۲. متغیرهای اعتباری مشتریان حقوقی بانک ایرانی

ردیف	متغیر	ردیف	متغیر
۱	وضعیت وام	۱۹	نسبت حساب دریافتنی به فروش خالص
۲	میزان وام (ریال)	۲۰	نسبت حساب دریافتنی به بدهی‌ها
۳	موجودی نقد در بانک (ریال)	۲۱	نسبت حساب پرداختنی به فروش خالص
۴	نسبت بدهی به دارایی	۲۲	نسبت فروش به دارایی
۵	نسبت حقوق صاحبان به دارایی	۲۳	نسبت فروش به دارایی ثابت
۶	نسبت بدهی بلند مدت به دارایی	۲۴	نسبت سود خالص به دارایی‌ها
۷	نسبت سود خالص به هزینه مالی	۲۵	نسبت سود خالص به فروش خالص
۸	نسبت دارایی جاری به بدهی جاری	۲۶	نسبت سود خالص به دارایی ثابت
۹	نسبت حساب دریافتی و موجودی نقد به بدهی جاری	۲۷	نسبت سود خالص به حقوق صاحبان
۱۰	نسبت سرمایه در گردش به بدهی‌ها	۲۸	نسبت هزینه فروش به فروش خالص
۱۱	نسبت دارایی جاری به بدهی‌ها	۲۹	اندازه بنگاه
۱۲	نسبت بدهی جاری به دارایی‌ها	۳۰	وضعیت مالکیت ملک (اجاره، سرقفلی، دارای سند مالکیت غیر قابل ترهین)
۱۳	نسبت موجودی نقد به دارایی‌ها	۳۱	ارتباط بین فعالیت و مدرک (۰ یا ۱)
۱۴	نسبت موجودی نقد به فروش خالص	۳۲	حسن شهرت (عددی بین ۰ تا ۱)
۱۵	نسبت سرمایه در گردش به فروش خالص	۳۳	نوع وثیقه (عددی بین ۰ تا ۱ برحسب نوع)
۱۶	نسبت دارایی جاری به فروش خالص	۳۴	مدت بازپرداخت (سال)
۱۷	نسبت موجودی نقد به بدهی جاری	۳۵	نرخ سود
۱۸	نسبت سرمایه در گرش به بدهی‌های جاری		

منبع: نتایج تحقیق

جدول ۳: متغیرهای کیفی و مقادیر آنها

متغیر	مقدار
X_1: وضعیت وام (تسویه شده، معوق، سررسید گذشته، تمدید مدت، استمهال شده)	(۱، ۰٫۸، ۰٫۶، ۰٫۴، ۰٫۲)
X_{30}: وضعیت مالکیت ملک (سرقفلی، دارای سند مالکیت غیر قابل ترهین، اجاره)	(۱، ۰٫۵، ۰)
X_{31}: ارتباط بین فعالیت و مدرک (مرتبط، غیرمرتبط)	(۱، ۰)
X_{32}: نوع وثیقه (ملک (غیرمنقول)، سپرده بلندمدت، اوراق مشارکت، چک و سایر)	(۱، ۰٫۷۵، ۰٫۵، ۰٫۲۵)

منبع: نتایج تحقیق

۲.۳. روش پیشنهادی

در این تحقیق به منظور تجزیه و تحلیل داده‌ها از مدل یادگیری عمیق حساس به هزینه بهینه‌شده با استفاده از الگوریتم فرا ابتکاری **GA** استفاده شده است. همچنین از مدل **LSTM** و چند الگوریتم کلاسیفایر دیگر به منظور رتبه‌بندی اعتباری با مدل یادگیری گروهی استفاده می‌شود. فرایند بهینه‌سازی نیز بصورت یکپارچه و با در نظر گرفتن خطای دسته‌بندی[۱] (**MC**) به عنوان تابع هدف صورت می‌گیرد.

شماتیک روش پیشنهادی بصورت شکل ۱ است و دارای ۴ مرحله است که ۳ مرحله آن بصورت یکپارچه انجام می شود:

۱. مرحله نرمالایز کردن داده‌ها
۲. مرحله انتخاب ویژگی
۳. مرحله کلاسه‌بندی
۴. مرحله بهینه‌سازی با الگوریتم **GA**

---

[۱] Misclassification

۱.۲.۳. مرحله اول: نرمالایز کردن داده‌ها

در روش‌های خوشه‌بندی و دسته‌بندی، مقیاس کردن متغیرها اساسی‌ترین گام است؛ زیرا استفاده از مقادیر متغیرها بدون مقیاس، ممکن است منجر به تأثیر نامتعادل متغیرها با مقادیر بزرگ بر فرایند خوشه‌بندی و دسته‌بندی گردد. یکی از روش‌های رایج مقیاس کردن داده‌ها، نرمال‌سازی آنهاست. با اعمال نرمال‌سازی، مقادیر تمامی متغیرها به بازه‌ای معین مثل ۰ تا ۱ تبدیل می‌شوند. در این تحقیق، فرایند نرمالایز کردن با استفاده از روش **min-max** انجام می‌شود که بصورت زیر است (سوکسامای[۱] و همکاران، ۲۰۲۲):

$$\boldsymbol{X} = \frac{\boldsymbol{X} - \boldsymbol{min}(\boldsymbol{X})}{\boldsymbol{max}(\boldsymbol{X}) - \boldsymbol{min}(\boldsymbol{X})} \quad (۱)$$

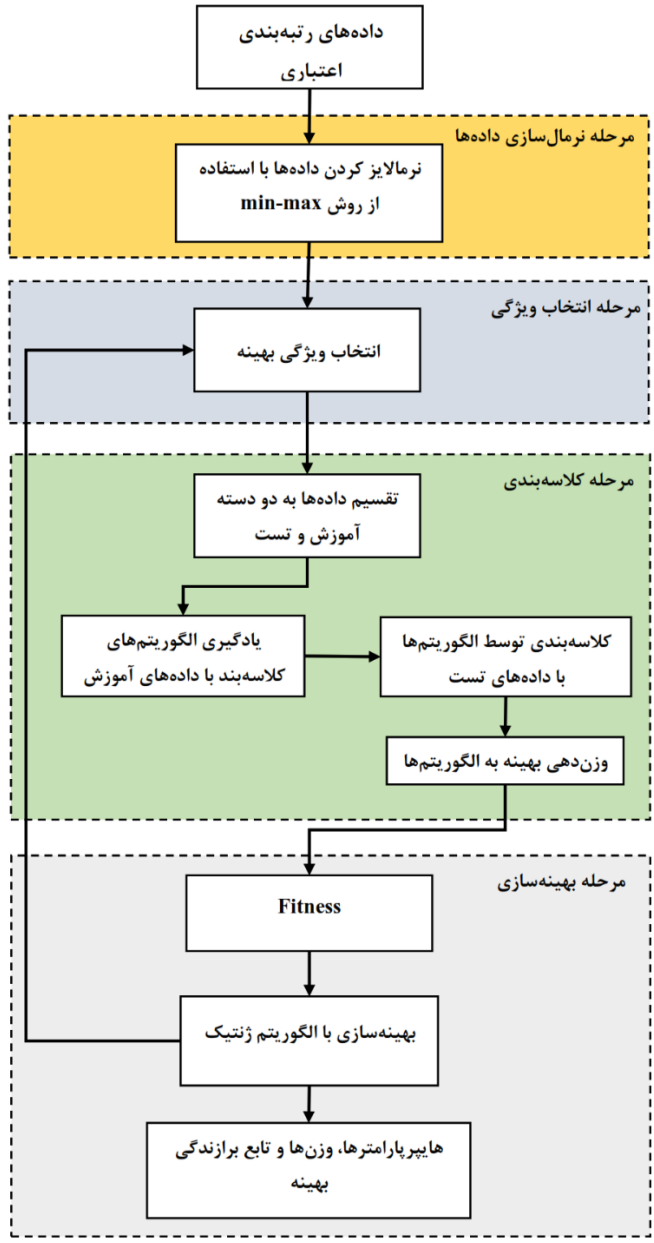
که در آن، $\boldsymbol{X}$ بردار مربوط به داده‌های اعتباری است.

۲.۲.۳. مرحله دوم: انتخاب ویژگی بهینه

در مسائل کلاسه‌بندی، ممکن است تمامی متغیرهای اعتباری در فرایند یادگیری الگوریتم‌های کلاسه‌بندی نقش مؤثری نداشته باشند و یا حتی تأثیر منفی بر این فرایند بگذارند. به منظور اجتناب از این موضوع، بایستی یک فرایند انتخاب ویژگی در نظر گرفته شود که توسط یک الگوریتم بهینه‌ساز مانند **GA** این امر میسر می‌شود. در این فرایند، متغیرهای اعتباری را که تحت عنوان ویژگی در نظر می‌گیریم، به دو حالت انتخاب شدن (۱) یا انتخاب نشدن (۰) توسط الگوریتم **GA** در نظر می‌گیریم. در نهایت آن دسته از ویژگی‌ها که مقدار ۱ برای آنها بدست آمده است، توسط الگوریتم **GA** انتخاب می‌شوند و سایر ویژگی‌ها کنار گذاشته می‌شوند.

---

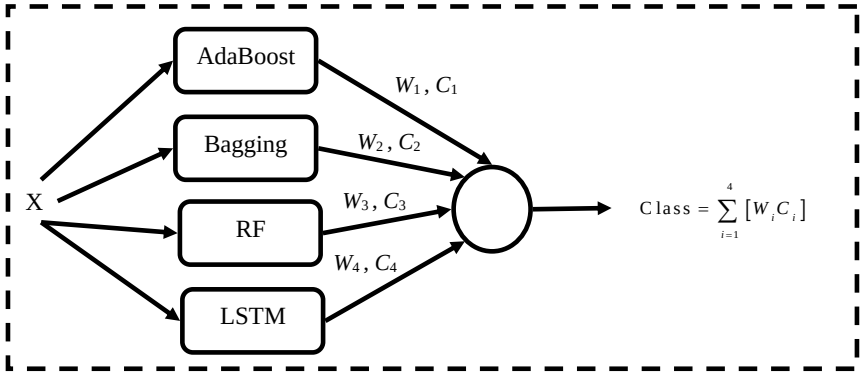
[۱] Suksamai



شکل ۱. شماتیک روش پیشنهادی

۳.۲.۳. مرحله سوم: کلاسه‌بندی بهینه

مرحله کلاسه‌بندی پیشنهادی در این پژوهش، یک کلاسه‌بندی حساس به هزینه و یادگیری جمعی است. حساس به هزینه یعنی الگوریتم‌ها در فرایند یادگیری به آن کلاسی که تعداد سمپل کمتری دارد (دیتای نامتوازن) وزن بیشتری نسبت به کلاس به تعداد سمپل بیشتر می‌دهند تا در فرایند یادگیری توازن برقرار شده و هر دو کلاس بخوبی توسط الگوریتم‌ها آموزش ببینند و نتایج حاصل از تست نیز برای هر دو کلاس مطلوب باشد. چنانچه این اتفاق نیفتد، خطای یادگیری کلاس به تعداد سمپل کمتر خیلی بیشتر از کلاس به تعداد سمپل زیاد می‌شود و در نتیجه الگوریتم بخوبی قادر به کلاسه‌بندی کلاس با تعداد سمپل کمتر نخواهد بود و مقدار **MC** این کلاس بیشتر خواهد شد. از سوی دیگر، یادگیری جمعی یعنی علاوه بر الگوریتم **LSTM** از الگوریتم‌های **Boosting**، **Bagging** و **RF** نیز استفاده خواهد شد و در نهایت بر اساس دقت یادگیری و کلاسه‌بندی الگوریتم‌ها، وزن بهینه برای هر یک از آنها تخصیص می‌یابد. شکل ۲ فرایند یادگیری جمعی پیشنهادی با الگوریتم‌های مذکور را نشان می‌دهد.



شکل ۲. فرایند یادگیری جمعی پیشنهادی

یکی از گام‌های مهم در سنجش یک روش طبقه‌بندی، توزیع مناسب داده‌ها برای مدلسازی می‌باشد. در این پژوهش از روش اعتبارسنجی متقابل[۱] **k-fold** برای تقسیم‌بندی داده‌ها استفاده

[۱] Cross-validation

خواهد شد. بر این اساس، داده‌ها به ۵ بخش مساوی (**k=۵**) بصورت تصادفی تقسیم می‌شوند و مدل طبقه‌بندی بر اساس این داده‌ها ساخته می‌شود. پس از طبقه‌بندی داده‌ها، از روش پیشنهادی برای هر یک از این طبقه‌ها استفاده خواهد شد و مقدار میانگین خطای دسته‌بندی طبقات به عنوان مقدار نهایی **MC** استفاده می‌شود.

٤.٢.٣. مرحله چهارم: بهینه‌سازی یکپارچه با الگوریتم GA

هایپرپارامترهای هر یک از الگوریتم‌ها، در کنار وزن بهینه تخصیصی به الگوریتم‌ها و فرایند انتخاب ویژگی بهینه توسط الگوریتم **GA** و بصورت یکپارچه انجام می‌شود. تابع برازندگی (**Fitness**) نیز میزان خطای کلاسه‌بندی (**MC**) در نظر گرفته می‌شود که برابر است با حاصل جمع تعداد خطای کلاس منفی (مشتری بدحساب) و کلاس مثبت (مشتری خوش حساب) بر کل تعداد سمپل‌ها در فرایند تست:

$$\mathbf{MC} = \frac{\mathbf{FN} + \mathbf{FP}}{\mathbf{TP+TN} + \mathbf{FP} + \mathbf{FN}} \qquad (۲)$$

که تعریف هر یک از متغیرهای فرمول فوق مطابق با جدول ۴ است.

جدول ٤. ماتریس درهم ریختگی

کلاس واقعی	کلاس پیش‌بینی شده	
	مشتریان بد	مشتریان خوب
مشتریان بد	منفی درست[۲] (**TN**)	مثبت کاذب[۱] (**FP**)
مشتریان خوب	منفی کاذب[۴] (**FN**)	مثبت درست[۳] (**TP**)

منبع: نتایج تحقیق

فلوچارت الگوریتم **GA** نیز بصورت شکل ۲ است. بنابراین مسئله بهینه پژوهش حاضر بصورت

---

[۱] False positive

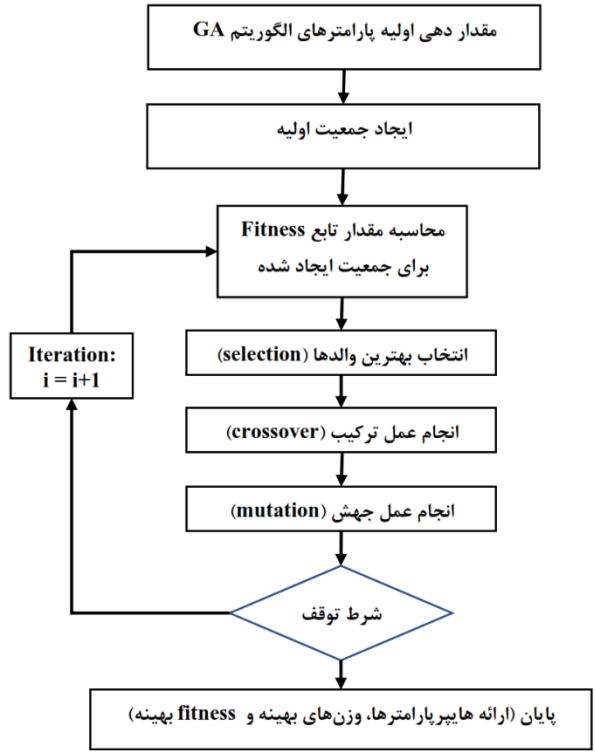
[۲] True negative

[۳] True positive

[۴] False negative

زیر بیان می‌شود:

$$\text{Fitenss} = \min \boldsymbol{MC} \quad (۳)$$



شکل ۲. فلوچارت الگوریتم GA.

٤.۲.۳. ارزیابی روش پیشنهادی

معیارهای مختلفی برای سنجش و ارزیابی روش‌های کلاسه‌بندی وجود دارد که در این پژوهش از هفت معیار مختلف استفاده می‌شود که عبارتند از:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (۴)$$

$$\text{ACC} = \frac{\text{TN} + \text{TP}}{\text{TP+TN} + \text{FP} + \text{FN}} \quad (۵)$$

$$\text{Recall or Sensitivity} = \frac{TP}{TP+FN} \quad (۶)$$

$$\text{F-Score} = (2 \times \text{Recall} \times \text{Precision}) / (\text{Recall} + \text{Precision}) \quad (۷)$$

$$\text{Error}_{\text{Type I}} = \frac{FP}{TN + FP} \quad (۸)$$

$$\text{Error}_{\text{Type II}} = \frac{FN}{FN + TP} \quad (۹)$$

$$MC = \frac{FN + FP}{TP+TN + FP + FN} \quad (۱۰)$$

٤. یافته‌های تحقیق

در این بخش به بررسی نتایج حاصل از روش پیشنهادی پرداخته شده است. همانطور که در بخش پیشین بدان اشاره شد، روش پیشنهادی یک روش بهینه یکپارچه هست که با استفاده از تکنیک‌های مختلف و انتخاب ویژگی به اعتبارسنجی مشتریان می‌پردازد. چهار روش مورد استفاده در این تحقیق عبارتند از روش‌های **Bagging**، **AbaBoost**، **RF** و **LSTM**. بنابراین هایپرپارامترهای این الگوریتم‌ها از طریق الگوریتم ژنتیک بهینه خواهد شد. جدول ۵ هایپرپارامترهای مدل را بهمراه فضای جستجو (کران بالا (**UB**) و کران پایین (**LB**)) نشان داده است. همچنین در جدول ۶ پارامترهای الگوریتم ژنتیک نشان داده شده است. علاوه بر این، برای انتخاب ویژگی، مقدار ۰ به منظور انتخاب نشدن ویژگی و مقدار ۱ به معنای انتخاب شدن ویژگی در نظر گرفته شده است. با در نظر گرفتن وزن الگوریتم‌ها، سه هایپرپارامتر وزنی نیز که مقدار آنها بین ۰ و ۱ بوده و مجموع آنها برابر ۱ هست، لحاظ شده است. بنابراین، در مجموع با در نظر گرفتن ۱۱ هایپرپارامتر نشان داده شده در جدول ۵ و سه پارامتر وزنی الگوریتم‌ها و **n** تعداد ویژگی دیتاست‌ها، تعداد هایپرپارامترهای مسئله که از راه بهینه‌سازی با الگوریتم **GA** حاصل می‌شود، برابر با ۱۴ + **n** است.

جدول ۵. هایپرپارامترهای الگوریتم‌ها

الگوریتم	پارامتر	LB	UB
AdaBoost	نرخ یادگیری	۰.۰۱	۱.۰
	تعداد تخمین گرها	۵۰	۵۰۰
	ماکزیمم عمق	۱	۱۰
Bagging	تعداد تخمین گرها	۵۰	۵۰۰
	ماکزیمم عمق	۱	۱۰
	نرخ ماکزیمم تعداد سمپل	۰.۰۱	۱.۰
	نرخ ماکزیمم تعداد ویژگی	۰.۰۱	۱.۰
RF	تعداد تخمین گرها	۵۰	۵۰۰
	ماکزیمم عمق	۱	۱۰
	نرخ ماکزیمم تعداد سمپل	۰.۰۱	۱.۰
	نرخ ماکزیمم تعداد ویژگی	۰.۰۱	۱.۰
LSTM	تعداد لایه‌های پنهان	۱۰	۵۰
	نرخ **dropout**	۰.۰۱	۰.۵
	نرخ یادگیری	۰.۰۱	۱.۰
	تعداد **epoch**	۱۰	۱۰۰

منبع: نتایج تحقیق

جدول ٦. پارامترهای الگوریتم ژنتیک

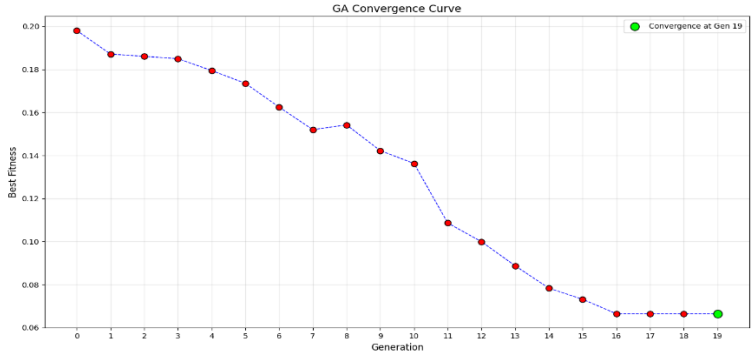
پارامتر	مقدار
اندازه جمعیت	۳۰
نرخ جهش	۰.۲
نرخ ترکیب	۰.۸
عملگر انتخاب	**Tournament selection**
حداکثر تعداد نسل	۳۰
تلورانس	۱e-۵
معیار توقف	رسیدن به تلورانس مدنظر یا حداکثر تعداد نسل

منبع: نتایج تحقیق

همانطور که مشخص شد، دیتای مشتریان حقوقی بانک ایرانی به نسبت ۵۷٪ مشتری خوش‌حساب به ۴۲٪ مشتری بدحساب است و در نتیجه یک دیتای تقریباً متوازن محسوب می‌شود. فرآیندی همگرایی مدل پیشنهادی برای دیتای ایران در نسل ۱۹ به دست آمده (شکل ۳) و نشان‌دهنده فضای جستجوی نسبتاً پیچیده برای پارامترهای مدل می‌باشد. نتایج به‌دست آمده از دیتاست ایران نشان می‌دهد که الگوریتم‌های منفرد در این چارچوب دارای عملکرد قابل قبولی از نظر معیارهایی مانند دقت (**ACC**)، میزان صحت (**Precision**)، فراخوانی (**Recall**) و شاخص **F۱** هستند (جدول ۷ تا ۱۰). بر اساس نتایج به‌دست‌آمده از داده‌های ایران، عملکرد الگوریتم‌های منفرد در این چارچوب از نظر معیارهای ارزیابی نظیر دقت (**Accuracy**)، میزان صحت (**Precision**)، فراخوانی (**Recall**) و شاخص **F۱** قابل قبول بوده است. به‌طور مشخص، مدل **AdaBoost** دقتی در حدود ۰.۹۱۴، مدل **Bagging** دقتی برابر ۰.۹۲۷، و مدل‌های **Random Forest** و **LSTM** دقت‌هایی در بازه ۰.۹۱۴ تا ۰.۹۳۱ را به ثبت رسانده‌اند. با ترکیب خروجی‌های این الگوریتم‌ها و استفاده از وزن‌های بهینه‌شده شامل **AdaBoost:** ۰.۲۹۵، **Bagging:** ۰.۲۱۹، **Random Forest:** ۰.۱۸۸ و **LSTM:** ۰.۲۹۷، عملکرد کلی مدل ترکیبی اندکی بهبود یافته است. این مدل ترکیبی به میانگین دقتی معادل ۰.۹۳۴، میزان صحت ۰.۸۷۹، فراخوانی ۰.۸۹۸ و شاخص **F۱** برابر ۰.۸۸۸ دست یافته است. علاوه بر این، کاهش قابل توجهی در خطاهای نوع اول (**False Positive**) با میانگین

۰.۰۶۳ و خطاهای نوع دوم (**False Negative**) با میانگین ۰.۱۱۶ مشاهده شده و میزان کل خطای طبقه‌بندی به ۰.۰۶۶ رسیده است (جدول ۱۱).

مقادیر بهینه به‌دست آمده برای هر الگوریتم در جدول ۱۲ نشان داده شده است. هایپرپارامترهای بهینه برای هر الگوریتم نیز به‌خوبی با ویژگی‌های داده‌ها و نیازهای مسئله تطبیق داده شده‌اند. برای **AdaBoost**، نرخ یادگیری ۰.۱۴۹، تعداد تخمین‌گرها ۱۲۷ و حداکثر عمق ۶؛ برای **Bagging**، تعداد تخمین‌گرها ۲۷۴ و حداکثر عمق ۵؛ برای **Random Forest**، تعداد تخمین‌گرها ۲۳۸ و حداکثر عمق ۶؛ و برای **LSTM**، تعداد لایه‌های پنهان ۳۹، نرخ **dropout** برابر ۰.۱۸۸، نرخ یادگیری ۰.۲۸۹ و تعداد **epoch** برابر ۱۰۹ به‌عنوان مقادیر بهینه تعیین شده‌اند. این تنظیمات نشان‌دهنده انطباق مناسب مدل‌ها با داده‌ها و پیچیدگی‌های مسئله طبقه‌بندی است.



شکل ۳. نمودار همگرایی مدل پیشنهادی برای دیتای ایران(خروجی نرم افزار)

جدول ۷. عملکرد الگوریتم AdaBoost برای دیتای ایران

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۱۴	۰.۹۳۰	۰.۹۳۶	۰.۹۳۳	۰.۰۹۴	۰.۰۹۰	۰.۰۸۶
۲	۰.۹۲۶	۰.۹۲۳	۰.۹۲۴	۰.۹۲۴	۰.۰۹۵	۰.۰۹۵	۰.۰۷۴
۳	۰.۹۱۴	۰.۹۰۴	۰.۹۲۶	۰.۹۱۵	۰.۱۲۴	۰.۰۹۵	۰.۰۸۶
۴	۰.۹۰۲	۰.۹۳۴	۰.۹۳۶	۰.۹۳۵	۰.۰۹۵	۰.۰۹۰	۰.۰۹۸
۵	۰.۹۱۶	۰.۹۲۵	۰.۸۹۸	۰.۹۱۱	۰.۰۹۵	۰.۱۵۴	۰.۰۸۴
Mean	**۰.۹۱٤**	**۰.۹۲۳**	**۰.۹۲٤**	**۰.۹۲٤**	**۰.۱۰۱**	**۰.۱۰۵**	**۰.۰۸٦**

منبع: نتایج تحقیق

جدول ۸. عملکرد الگوریتم Bagging برای دیتای ایران

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۴۰	۰.۹۵۴	۰.۹۷۷	۰.۹۶۶	۰.۰۳۰	۰.۰۲۱	۰.۰۶۰
۲	۰.۹۰۷	۰.۹۰۵	۰.۸۹۱	۰.۸۹۸	۰.۱۲۳	۰.۱۳۶	۰.۰۹۳
۳	۰.۹۳۵	۰.۹۶۶	۰.۹۰۷	۰.۹۳۵	۰.۰۲۵	۰.۰۵۶	۰.۰۶۵
۴	۰.۹۳۴	۰.۹۰۶	۰.۹۵۸	۰.۹۳۱	۰.۰۵۶	۰.۰۷۹	۰.۰۶۶
۵	۰.۹۲۰	۰.۹۲۴	۰.۹۳۶	۰.۹۳۰	۰.۰۶۴	۰.۰۵۴	۰.۰۸۰
Mean	**۰.۹۲۷**	**۰.۹۳۱**	**۰.۹۳٤**	**۰.۹۳۲**	**۰.۰٦۰**	**۰.۰٦۹**	**۰.۰۷۳**

منبع: نتایج تحقیق

جدول ۹. عملکرد الگوریتم RF برای دیتای ایران

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۳۳	۰.۹۶۹	۰.۹۷۰	۰.۹۷۰	۰.۰۲۹	۰.۰۲۲	۰.۰۶۷
۲	۰.۹۴۵	۰.۹۲۴	۰.۹۹۰	۰.۹۵۶	۰.۰۶۱	۰.۰۱۳	۰.۰۵۵
۳	۰.۹۳۰	۰.۹۳۵	۰.۹۵۷	۰.۹۴۶	۰.۰۳۸	۰.۱۳۴	۰.۰۷۰
۴	۰.۹۲۵	۰.۹۱۵	۰.۹۷۷	۰.۹۴۵	۰.۰۸۱	۰.۰۲۱	۰.۰۷۵
۵	۰.۹۱۹	۰.۹۲۵	۰.۹۶۱	۰.۹۴۳	۰.۰۶۸	۰.۰۳۰	۰.۰۸۱
Mean	**۰.۹۳۱**	**۰.۹۳٤**	**۰.۹۷۱**	**۰.۹۵۲**	**۰.۰۵۵**	**۰.۰٤٤**	**۰.۰٦۹**

منبع: نتایج تحقیق

جدول ۱۰. عملکرد الگوریتم LSTM برای دیتای ایران

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۲۵	۰.۹۱۵	۰.۹۴۴	۰.۹۲۹	۰.۰۸۸	۰.۱۰۲	۰.۰۷۵
۲	۰.۹۰۷	۰.۹۴۳	۰.۹۶۸	۰.۹۵۶	۰.۰۶۹	۰.۰۴۴	۰.۰۹۳
۳	۰.۹۲۵	۰.۸۹۷	۰.۹۷۹	۰.۹۳۷	۰.۱۵۳	۰.۰۲۰	۰.۰۷۵
۴	۰.۹۰۷	۰.۸۸۵	۰.۹۱۵	۰.۹۰۰	۰.۱۷۴	۰.۰۸۷	۰.۰۹۳
۵	۰.۹۰۶	۰.۹۳۶	۰.۹۴۴	۰.۹۴۰	۰.۰۷۱	۰.۱۰۲	۰.۰۹۴
Mean	**۰.۹۱٤**	**۰.۹۱۵**	**۰.۹۵۰**	**۰.۹۳۲**	**۰.۱۱۱**	**۰.۰۷۱**	**۰.۰۸٦**

منبع: نتایج تحقیق

جدول ۱۱. عملکرد مدل ترکیبی برای دیتای ایران

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۲۳	۰.۸۸۶	۰.۹۵۲	۰.۹۱۸	۰.۰۵۲	۰.۰۷۸	۰.۰۷۷
۲	۰.۹۲۴	۰.۸۳۱	۰.۸۹۹	۰.۸۶۴	۰.۱۰۱	۰.۱۱۱	۰.۰۷۶
۳	۰.۹۳۱	۰.۹۳۹	۰.۹۱۵	۰.۹۲۷	۰.۰۳۰	۰.۰۹۵	۰.۰۶۹
۴	۰.۹۲۹	۰.۸۳۳	۰.۷۹۸	۰.۸۱۵	۰.۰۹۸	۰.۲۰۲	۰.۰۷۱
۵	۰.۹۶۳	۰.۹۰۳	۰.۹۲۷	۰.۹۱۵	۰.۰۳۵	۰.۰۹۳	۰.۰۳۷
Mean	**۰.۹۳٤**	**۰.۸۷۹**	**۰.۸۹۸**	**۰.۸۸۸**	**۰.۰٦۳**	**۰.۱۱٦**	**۰.۰٦٦**

منبع: نتایج تحقیق

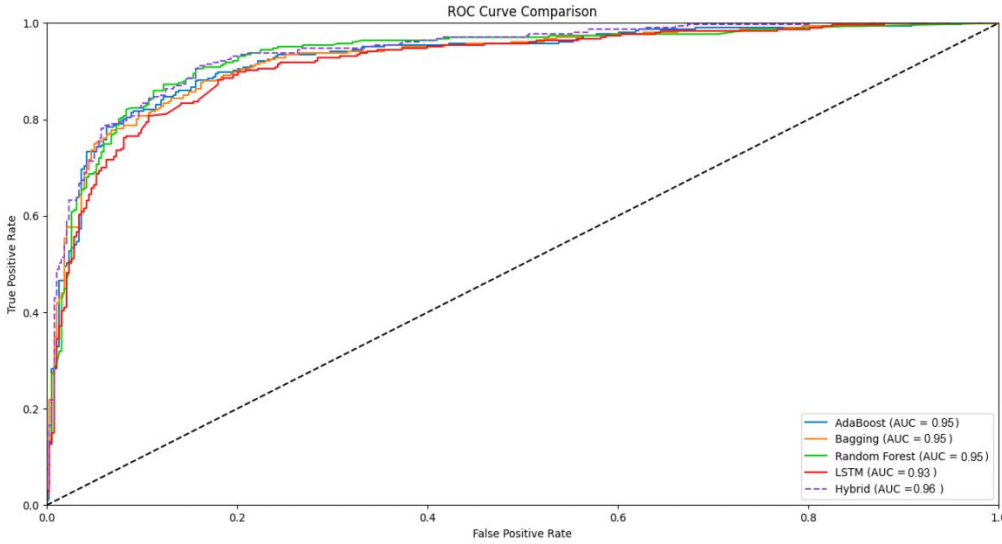
جدول ۱۲. مقادیر بهینه هایپرپارامترهای مدل برای دیتای ایران

الگوریتم	پارامتر	LB	UB	مقدار بهینه
AdaBoost	نرخ یادگیری	۰.۰۱	۱.۰	۰.۱۴۹
	تعداد تخمین‌گرها	۵۰	۵۰۰	۱۲۷
	ماکزیمم عمق	۱	۱۰	۶
Bagging	تعداد تخمین‌گرها	۵۰	۵۰۰	۲۷۴
	ماکزیمم عمق	۱	۱۰	۵
	نرخ ماکزیمم تعداد سمپل	۰.۰۱	۱.۰	۰.۱۹۶
	نرخ ماکزیمم تعداد ویژگی	۰.۰۱	۱.۰	۰.۲۴۵
RF	تعداد تخمین‌گرها	۵۰	۵۰۰	۲۳۸
	ماکزیمم عمق	۱	۱۰	۶
	نرخ ماکزیمم تعداد سمپل	۰.۰۱	۱.۰	۰.۴۳۲
	نرخ ماکزیمم تعداد ویژگی	۰.۰۱	۱.۰	۰.۷۹۶
LSTM	تعداد لایه‌های پنهان	۱۰	۵۰	۳۹
	نرخ **dropout**	۰.۰۱	۰.۵	۰.۱۸۸
	نرخ یادگیری	۰.۰۱	۱.۰	۰.۲۸۹
	تعداد **epoch**	۱۰	۱۰۰	۱۰۹
وزن الگوریتم‌ها	**AbaBoost**	۰	۱	۰.۲۹۵
	Bagging	۰	۱	۰.۲۱۹
	RF	۰	۱	۰.۱۸۸
	LSTM	۰	۱	۰.۲۹۷

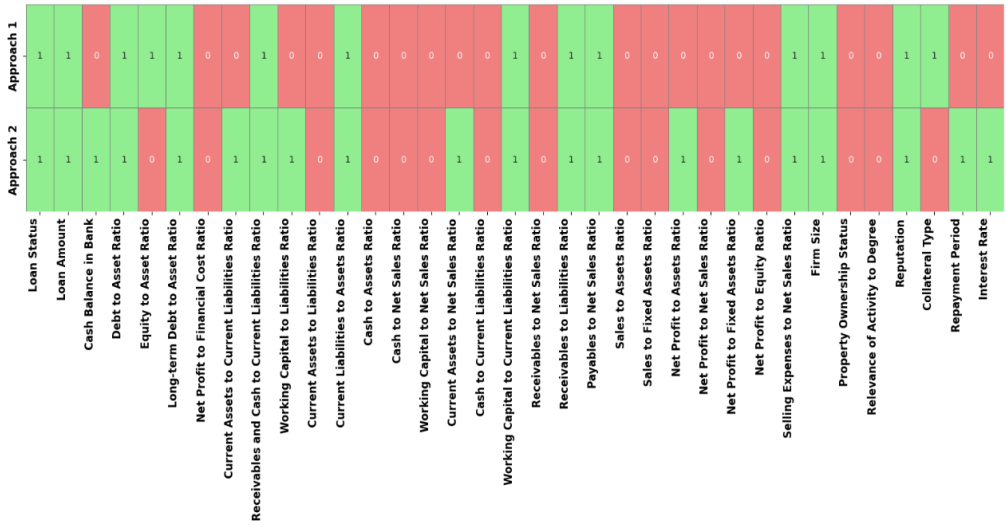
منبع: نتایج تحقیق

نمودار **ROC** الگوریتم‌های مدل رویکرد اول نیز در شکل ۴ نشان داده شده است. همانطور که در این شکل مشاهده می‌شود، مقدار **ROC** الگوریتم‌ها مقدار قابل توجه برابر ۰.۹۵ هست که نشان از دقت بالای آنها برای دیتای ایران دارد. همچنین مقدار **ROC** مدل ترکیبی بیشتر از مدل‌های مذکور هست و به مقدار ۰.۹۶ رسیده است که نشان می‌دهد، رویکرد ترکیبی پیشنهادی

باعث بهبود دقت نهایی می‌شود.



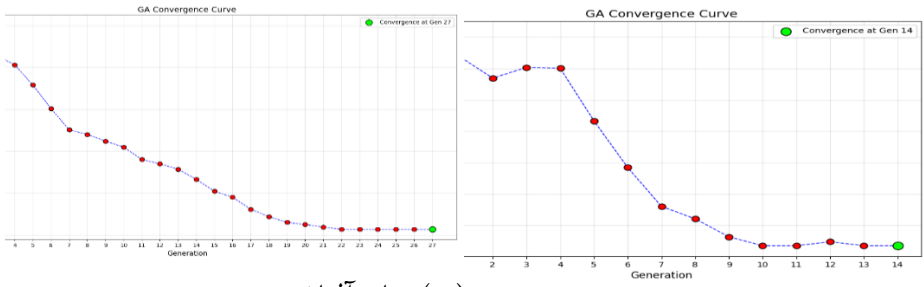
شکل ٤. نمودار ROC مدل پیشنهادی برای دیتای ایران(خروجی نرم افزار)



شکل ٥. مقایسه ویژگی‌های بهینه انتخاب شده دو رویکرد پیشنهادی برای دیتای ایران(خروجی نرم افزار)

در انتخاب ویژگی‌های بهینه که در شکل ۵ نشان داده شده است، مشاهده می‌شود که ویژگی‌های ۱ (وضعیت وام)، ۲ (میزان وام)، ۴ (نسبت بدهی به دارایی)، ۵ (نسبت حقوق صاحبان به دارایی)، ۶ (نسبت بدهی بلند مدت به دارایی)، ۹ (نسبت حساب دریافتی و موجودی نقد به بدهی جاری)، ۱۲ (نسبت بدهی جاری به دارایی‌ها)، ۱۸ (نسبت سرمایه در گردش به بدهی‌های جاری)، ۲۰ (نسبت حساب دریافتنی به بدهی‌ها)، ۲۱ (نسبت حساب پرداختنی به فروش خالص)، ۲۸ (نسبت هزینه فروش به فروش خالص)، ۲۹ (اندازه بنگاه)، ۳۲ (حسن شهرت) و ۳۳ (نوع وثیقه) به عنوان ویژگی‌های بهینه انتخاب شده‌اند.

به منظور اعتبارسنجی مدل پیشنهادی، از داده‌های بنچمارک استرالیا و آلمان استفاده شده است. نمودار همگرایی مدل پیشنهادی برای دیتای استرالیا و آلمان در شکل ۶ نشان داده شده است. همانطور که از این شکل مشخص است، الگوریتم برای دیتای استرالیا در نسل ۱۴ و برای دیتای آلمان در تکرار ۲۷ به همگرایی رسیده است که نشان‌دهنده فضای جستجوی پیچیده برای تنظیم پارامترهای مدل می‌باشد. نتایج برآورد مدل برای این دو دیتاست نیز به ترتیب در جدول ۱۳ و ۱۴ آورده شده است. مدل ترکیبی برای دیتای استرالیا دارای میانگین دقتی برابر ۰.۹۰۲، **Precision** برابر ۰.۸۷۹، **Recall** برابر ۰.۸۹۸ و **F1-Score** برابر ۰.۸۸۸ است. همچنین، کاهش قابل توجهی در خطاهای نوع **I (False Positive)** به میانگین ۰.۰۶۳ و خطاهای نوع **II (False Negative)** به میانگین ۰.۱۱۶ به همراه میزان کل خطای طبقه‌بندی (**MC**) برابر با ۰.۰۹۸ مشاهده شده است. برای دیتای آلمان نیز منجر به دستیابی به میانگین دقت ۰.۸۴۳، **Precision** ۰.۹۰۲، **Recall** ۰.۹۱۸ و **F1-Score** ۰.۹۱۰ شده است؛ همچنین، نرخ خطای نوع **I** و **II** به ترتیب به مقادیر ۰.۱۲۶ و ۰.۱۱۴ و **MC** برابر ۰.۱۵۷ کاهش یافته است.



(ب) دیتای آلمان

(الف) دیتای استرالیا

شکل ٦. نمودار همگرایی مدل پیشنهادی برای دیتای استرالیا و آلمان(خروجی نرم افزار).

جدول ۱۳. عملکرد مدل ترکیبی برای دیتای استرالیا

Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۹۱۴	۰.۸۸۶	۰.۹۵۲	۰.۹۱۸	۰.۰۵۲	۰.۰۷۸	۰.۰۸۶
۲	۰.۸۸۱	۰.۸۳۱	۰.۸۹۹	۰.۸۶۴	۰.۱۰۱	۰.۱۱۱	۰.۱۱۹
۳	۰.۹۴۶	۰.۹۳۹	۰.۹۱۵	۰.۹۲۷	۰.۰۳۰	۰.۰۹۵	۰.۰۵۴
۴	۰.۸۳۷	۰.۸۳۳	۰.۷۹۸	۰.۸۱۵	۰.۰۹۸	۰.۲۰۲	۰.۱۶۳
۵	۰.۹۳۱	۰.۹۰۳	۰.۹۲۷	۰.۹۱۵	۰.۰۳۵	۰.۰۹۳	۰.۰۶۹
Mean	**۰.۹۰۲**	**۰.۸۷۹**	**۰.۸۹۸**	**۰.۸۸۸**	**۰.۰٦۳**	**۰.۱۱٦**	**۰.۰۹۸**

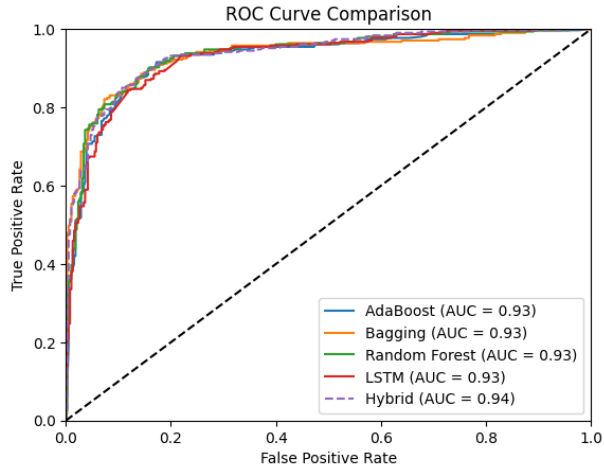
منبع: نتایج تحقیق

جدول ١٤. عملکرد مدل ترکیبی برای دیتای آلمان

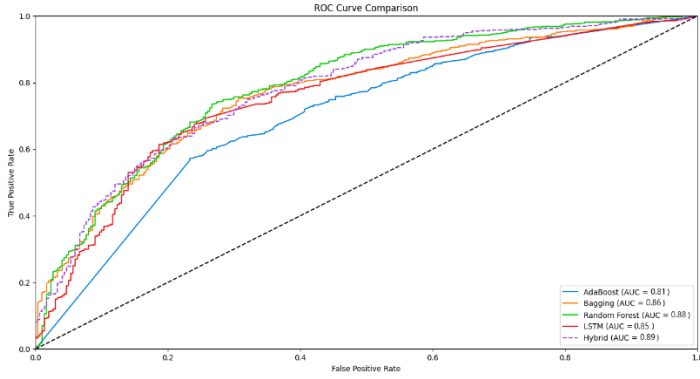
Fold	ACC	Precision	Recall	F1-Score	E-Type I	E-Type II	MC
۱	۰.۸۳۶	۰.۸۸۰	۰.۸۹۹	۰.۸۸۹	۰.۱۶۷	۰.۱۷۱	۰.۱۶۴
۲	۰.۸۰۵	۰.۸۹۷	۰.۸۸۷	۰.۸۹۲	۰.۱۲۱	۰.۱۴۳	۰.۱۹۵
۳	۰.۸۴۵	۰.۸۹۵	۰.۹۱۶	۰.۹۰۵	۰.۱۳۱	۰.۱۰۴	۰.۱۵۵
۴	۰.۸۶۰	۰.۸۹۳	۰.۹۳۱	۰.۹۱۲	۰.۱۲۷	۰.۰۸۹	۰.۱۴۰
۵	۰.۸۷۰	۰.۹۴۵	۰.۹۵۶	۰.۹۵۰	۰.۰۸۵	۰.۰۶۴	۰.۱۳۰
Mean	**۰.۸٤۳**	**۰.۹۰۲**	**۰.۹۱۸**	**۰.۹۱۰**	**۰.۱۲٦**	**۰.۱۱٤**	**۰.۱۵۷**

منبع: نتایج تحقیق

نمودار **ROC** الگوریتم‌های مدل برای دیتاست استرالیا نیز در شکل ۷ نشان داده شده است. همانطور که در این شکل مشاهده می‌شود، مقدار **ROC** الگوریتم‌ها مقدار قابل توجه برابر ۰.۹۳ هست که نشان از دقت بالای آنها برای دیتای استرالیا دارد. همچنین مقدار **ROC** مدل ترکیبی بیشتر از مدل‌های مذکور هست و به مقدار ۰.۹۴ رسیده است که نشان می‌دهد، رویکرد ترکیبی پیشنهادی باعث بهبود دقت نهایی می‌شود. همچنین نمودار **ROC** دیتای آلمان (شکل ۸) نشان می‌دهد که الگوریتم‌های منفرد مقدار **ROC** بین ۰.۸۳ و ۰.۸۶ دارند و مدل ترکیبی به **ROC** برابر ۰.۸۹ دست یافته که دلالت بر دقت بسیار بالا و توان تفکیک قوی مدل‌ها در دیتای آلمان دارد.



شکل ۷. نمودار ROC مدل پیشنهادی برای دیتای استرالیا(خروجی نرم افزار)



شکل ۸. نمودار ROC مدل پیشنهادی با رویکرد اول برای دیتای آلمان(خروجی نرم افزار)

مقایسه نتایج مدل پیشنهادی با نتایج سایر محققین در جدول ۱۵ آورده شده است. همانطور که از این جدول مشاهده می‌شود، دقت مدل پیشنهادی به ویژه برای دیتای متوازن آلمان بیشتر از مدل‌های مشابه معرفی شده توسط سایر محققیق برای حالت بدون دستکاری داده با استفاده از روش‌های نمونه‌برداری است. این امر نشان می‌دهد که مدل پیشنهادی توانایی بالایی برای حصول نتایج بسیار مناب و با قوام برای دیتای پیچیده است که ذاتاً در دنیای واقعی، بانک‌ها با آن روبرو هستند.

جدول ۱۵. مقایسه نتایج مدل پیشنهاد یا سایر منابع

رفرنس	روش	sampling	دیتاست آلمان	دیتاست استرالیا
تریپاتی[۱] و همکاران(۲۰۲۸)	**NB-DT-QDA-TDNN-PNN (Majority Voting)**	×	۰.۷۹۴	۰.۹۰۹
	NB-DT-QDA-TDNN-PNN (Weighted Voting)	×	۰.۷۹۴	۰.۹۰۹
	NB-DT-QDA-TDNN-PNN (Layered Majority Voting)	×	۰.۸۰۵	۰.۹۰۸
	NB-DT-QDA-TDNN-PNN (Layered Weighted Voting)	×	۰.۸۴۲	۰.۹۳
خلیلی و رستگار(۲۰۲۳)	**Stacking**	×	۰.۸۱۶	۰.۹۳۸
آرولبا و سان[۲](۲۰۲۴)	**XGBoost + SMOTE-ENN**	✓	***۰.۸۹۵***	۰.۹۱۰
روفیک[۳] و همکاران(۲۰۲۴)	**Stacking + SMOTE**	✓	۰.۸۳۲	-
روش پیشنهادی	**Stacking**	×	**۰.۸۴۳**	**۰.۹۰۲**

منبع: نتایج تحقیق

[۱] Tripathi

[۲] Aruleba& Sun

[۳] Rofik

۵. بحث و نتیجه‌گیری

پژوهش حاضر با ارائه چارچوب ترکیبی مبتنی بر یادگیری عمیق حساس به هزینه و بهینه‌سازی همزمان با الگوریتم ژنتیک، گامی اساسی در ارتقای دقت پیش‌بینی ریسک اعتباری مشتریان حقوقی بانک‌ها برداشت. نتایج حاصل از تست‌های گسترده روی داده‌های ایران، استرالیا و آلمان نشان‌دهنده توانایی مدل در کاهش خطای طبقه‌بندی به ترتیب به مقادیر ۶.۶٪ در داده‌های ایران، ۹.۸٪ در داده‌های استرالیا و ۱۵.۷٪ در داده‌های آلمان است. این بهبود عملکرد ناشی از ترکیب هوشمندانه الگوریتم‌های مختلف، بهینه‌سازی دقیق هایپرپارامترها و انتخاب هدفمند ویژگی‌های کلیدی (مانند «وضعیت وام»، «نسبت بدهی به دارایی» و «حسن شهرت») می‌باشد. از منظر تطبیق مدل با ویژگی‌های اقتصادی ـ مالی هر منطقه و نحوه برخورد با داده‌های نامتوازن، یافته‌های پژوهش موید این نکته است که مدل‌های عمومی نمی‌توانند به‌طور یکسان در تمامی بازارها عملکرد مطلوبی داشته باشند؛ لذا بومی‌سازی پارامترها و به‌کارگیری تنظیمات خاص برای هر بازار از ضرورت‌های مهم محسوب می‌شود. به‌علاوه، کاهش قابل توجه خطاهای نوع اول (شناسایی نادرست مشتریان خوب به‌عنوان بدحساب) و نوع دوم (شناسایی نادرست مشتریان بد به‌عنوان خوش‌حساب) نشان‌دهنده کارایی بالای رویکرد ترکیبی پیشنهادی در مدیریت ریسک اعتباری است.

با توجه به یافته‌های این پژوهش، پیشنهادات زیر مطرح می‌شود:

پژوهش حاضر با ارائه چارچوب ترکیبی مبتنی بر یادگیری عمیق حساس به هزینه و بهینه‌سازی همزمان با الگوریتم ژنتیک، گامی اساسی در ارتقای دقت پیش‌بینی ریسک اعتباری مشتریان حقوقی بانک‌ها برداشت. نتایج حاصل از تست‌های گسترده روی داده‌های ایران، استرالیا و آلمان نشان‌دهنده توانایی مدل در کاهش خطای طبقه‌بندی به ترتیب به مقادیر ۶.۶٪ در داده‌های ایران، ۹.۸٪ در داده‌های استرالیا و ۱۵.۷٪ در داده‌های آلمان است. این بهبود عملکرد ناشی از ترکیب هوشمندانه الگوریتم‌های مختلف، بهینه‌سازی دقیق هایپرپارامترها و انتخاب هدفمند ویژگی‌های کلیدی (مانند «وضعیت وام»، «نسبت بدهی به دارایی» و «حسن شهرت») می‌باشد. از منظر تطبیق مدل با ویژگی‌های اقتصادی ـ مالی هر منطقه و نحوه برخورد با داده‌های نامتوازن، یافته‌های پژوهش موید این نکته است که مدل‌های عمومی نمی‌توانند به‌طور یکسان در تمامی بازارها

عملکرد مطلوبی داشته باشند؛ لذا بومی‌سازی پارامترها و به‌کارگیری تنظیمات خاص برای هر بازار از ضرورت‌های مهم محسوب می‌شود. به‌علاوه، کاهش قابل توجه خطاهای نوع اول (شناسایی نادرست مشتریان خوب به‌عنوان بدحساب) و نوع دوم (شناسایی نادرست مشتریان بد به‌عنوان خوش‌حساب) نشان‌دهنده کارایی بالای رویکرد ترکیبی پیشنهادی در مدیریت ریسک اعتباری است.

با توجه به یافته‌های این پژوهش، پیشنهادات زیر مطرح می‌شود:

- **ادغام داده‌های کیفی و رفتاری:** توصیه می‌شود مطالعات آینده بر روی ادغام داده‌های غیر عددی نظیر سابقه تعاملات مشتریان با بانک، نظرات مشتریان و نوسانات بازار سرمایه متمرکز شوند تا توانایی مدل در تحلیل شرایط پیچیده و زودگذر افزایش یابد.
- **توسعه مدل‌های پویا:** توسعه نسخه‌های به‌روز و خودکار مدل با قابلیت تنظیم مجدد پارامترها در مواجهه با تغییرات اقتصادی (مانند تورم یا تحریم‌ها) می‌تواند افزایش قابلیت اطمینان پیش‌بینی‌ها در بلندمدت را تضمین کند.
- **رویکردهای تفسیرپذیر (XAI):** استفاده از ابزارهای تفسیرپذیری مدل به منظور تحلیل دقیق‌تر نقش هر ویژگی در تصمیم‌گیری، علاوه بر افزایش شفافیت سیستم، می‌تواند به تدوین سیاست‌های اعتباری مبتنی بر شواهد منجر شود.
- **سرمایه‌گذاری در فناوری‌های هوش مصنوعی:** نهادهای مالی و بانکی می‌توانند با سرمایه‌گذاری در فناوری‌های پیشرفته و بهبود سیستم‌های مدیریتی، از مدل‌های رتبه‌بندی اعتباری مبتنی بر یادگیری عمیق و جمعی بهره‌مند شده و در کاهش ریسک‌های اعتباری و خسارات مالی گام‌های موثری بردارند.
- **تبیین شفاف نتایج:** ارائه توضیحات دقیق و شفاف درباره انتخاب ویژگی‌ها و تفسیر نتایج مدل، می‌تواند به افزایش اعتماد نهادهای نظارتی و ارتقای تعامل میان بانک‌ها و مراجع نظارتی بیانجامد.

منابع

- ابتیاع، مجید و همکاران. (۱۴۰۰). رتبه بندی اعتباری مشتریان بانک به کمک روش جدید گروهی بر پایه ماشین بردار پشتیبان: مطالعه موردی بانک پاسارگاد، ***محاسبات نرم***، دوره ۱۰(۲)، ۲-۱۵.
- اخباری، مهدیه؛ مخاطب‌رفیعی، فریماه. (۱۳۸۹). کاربرد سیستم‌های استدلال عصبی ـ فازی در رتبه‌بندی اعتباری مشتریان حقوقی بانک‌ها، ***مجله تحقیقات اقتصادی***، شماره ۹۲، پاییز ۸۹، صص. ۲۱-۱.
- اسراری، مصطفی. (۱۳۹۲). مقایسه کارایی مدل‌های شبکه عصبی، الگوریتم ژنتیک و لاجیت در ارزیابی ریسک اعتباری مشتریان حقیقی: مطالعه موردی در بانک کشاورزی، **پایان‌نامه کارشناسی ارشد، رشته MBA، گرایش مالی، دانشگاه علوم اقتصادی**.
- پاک‌بین، پرستو؛ پویانفر، احمد. (۱۳۹۷). رتبه‌بندی اعتباری مشتریان بانک با تکنیک طبقه‌بندی ترکیبی پویا و احتمال نرم (مطالعه موردی: بانک پاسارگاد)، **پایان‌نامه کارشناسی ارشد، رشته مهندسی مالی و مدیریت ریسک، دانشگاه خاتم.**
- جیحونی‌پور، مهرداد؛ اعظمی، سمیه؛ دل‌انگیزان، سهراب. (۱۴۰۳). مدلسازی و شناسایی روابط علی بین عوامل اصلی ریسک اعتباری در سیستم بانکی با استفاده از تکنیک تصمیم گیری دیمتل، ***نشریه: سیاست گذاری اقتصادی***، ۱۷(۳۳)، ۱۸۰-۲۱۱.
- حسامی، علی. (۱۳۹۸). امتیازدهی اعتباری مشتریان حقیقی بانکی با استفاده از روش یادگیری عمیق مورد مطالعه: مشتریان اعتباری یکی از بانک‌های ایرانی، **پایان‌نامه کارشناسی ارشد، مهندسی صنایع ـ سیستم‌های مالی، دانشگاه اراک**.
- خانجانی، محسن و همکاران. (۱۴۰۲). رتبه بندی اعتباری مشتریان حقوقی با استفاده از روشهای یادگیری ماشین و یادگیری عمیق، **اولین کنفرانس بین المللی توانمندی مدیریت، مهندسی صنایع، حسابداری و اقتصاد**.
- دادمحمدی، دانیال؛ احمدی، عباس. (۱۳۹۳). رتبه‌بندی اعتباری مشتریان بانک با استفاده از شبکه عصبی با اتصالات جانبی، **فصلنامه توسعه مدیریت پولی و بانکی**، سال دوم، شماره ۳، تابستان ۱۳۹۳.
- کوچه‌باغی، سمیرا؛ رجبی‌بهجت، امیر. (۱۳۹۶). انتخاب زیرمجموعه‌ای از ویژگی‌ها مبتنی بر الگوریتم جستجوی گرانشی باینری برای رتبه‌بندی اعتباری مشتریان بانکی با

استفاده از ترکیب الگوریتم‌های طبقه‌بندی، **پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی، واحد بندرعباس.**

- گل‌محمدی، نسرین. (۱۳۹۵). رتبه‌بندی اعتباری مشتریان بانک با استفاده از الگوریتم فراابتکاری(ژنتیک) (مورد مطالعه: مشتریان حقیقی بانک مسکن ـ شعب منطقه غرب تهران)، **پایان‌نامه کارشناسی ارشد، مرکز آموزش عالی رجاء، دانشکده مهندسی صنایع**.
- مهدوی، کاوه؛ حری، محمدصادق. (۱۳۹۴). طراحی مدلی جهت پیش بینی رتبه اعتباری مشتریان بانکها با استفاده از الگوریتم فراابتکاری و هیبریدی چند معیاره شبکه عصبی فازی – کلونی مورچگان (مطالعه موردی شعب پست بانک استان تهران)، **پژوهش‌های مدیریت در ایران**، دوره ۱۹، شماره ۱، ۹۱-۱۱۶.
- ندری، کامران؛ محرابی، لیلا. (۱۳۹۷). بررسی انواع ریسک و مدیریت ریسک در نظام بانکداری اسلامی، **توسعه راهبرد**، ۵۴(۱۴)، ۱۶۰-۱۸۱.
- نظرآقایی، مهدی. (۱۳۹۸). دسته بندی ریسک اعتباری مشتریان حقیقی با استفاده از یادگیری جمعی (مطالعه موردی بانک سپه)، **پژوهش‌های پولی ـ بانکی**، شماره ۳۹، ۱۲۹-۱۶۶.

- Abdou, H. A., Mitra, S., Fry, J., & Elamer, A. A. (۲۰۱۹). Would two-stage scoring models alleviate bank exposure to bad debt?. ***Expert Systems with Applications***, ۱۲۸, ۱-۱۳.
- Abellán, J., & Mantas, C. J. (۲۰۱٤). Improving experimental studies about ensembles of classifiers for bankruptcy
- Addy, W. A., Ugochukwu, C. E., Oyewole, A. T., & Chrisanctus, O. (۲۰۲٤). Predictive analytics in credit risk management for banks: A comprehensive review.
- Ala'raj, M., Abbod, M. F., Majdalawieh, M., & Jum'a, L. (۲۰۲۲). A deep learning model for behavioural credit scoring in banks. ***Neural Computing and Applications***, ۱-۲۸.
- Almufti, S. M., Shaban, A. A., Ali, Z. A., Ali, R. I., & Fuente, J. A. D. (۲۰۲۳). Overview of Metaheuristic Algorithms. ***Polaris Global Journal of Scholarly Research and Trends***, ۲(۲), ۱۰-۳۲.
- Aruleba, I., & Sun, Y. (۲۰۲٤). Effective credit risk prediction using ensemble classifiers with model explanation. ***IEEE Access***.

- Attigeri, G. V., Pai, M. M., & Pai, R. M. (۲۰۱۷). Credit risk assessment using machine learning algorithms. ***Advanced Science Letters***, ۲۳(٤), ۳٦٤۹-۳٦٥۳.
- Bhattacharya, A., Biswas, S. K., & Mandal, A. (۲۰۲۳). Credit risk evaluation: a comprehensive study. **Multimedia Tools and Applications**, ۸۲(۱۲), ۱۸۲۱۷-۱۸۲٦۷.
- Bielecki, T. R., & Rutkowski, M. (۲۰۱۳). *Credit risk: modeling, valuation and hedging*. **Springer Science & Business Media**.
- Cahn, C., Girotti, M., & Salvade, F. (۲۰۲٤). Credit ratings and the hold-up problem in the loan market. ***Management Science***, ۷۰(۳), ۱۸۱۰-۱۸۳۱.
- Das, S. R. (۲۰۱۹). Credit risk derivatives. In *World Scientific Reference on Contingent Claims Analysis in Corporate Finance: Volume ۳: **Empirical Testing and Applications of CCA*** (pp. ۳۷۳-٤۰۰).
- Du, P., & Shu, H. (۲۰۲۲). Exploration of financial market credit scoring and risk management and prediction using deep learning and bionic algorithm. ***Journal of Global Information Management (JGIM)***, ۳۰(۹), ۱-۲۹.
- Dumitrescu, E., Hue, S., Hurlin, C., & Tokpavi, S. (۲۰۱۸). *Machine Learning for Credit Scoring: Improving Logistic Regression with Non Linear Decision Tree Effects* (**Doctoral dissertation, PhD thesis. Paris Nanterre University, University of Orleans**).
- Eletter, S. F., & Yaseen, S. G. (۲۰۱۷). Loan decision models for the Jordanian commercial banks. ***Global Business and Economics Review***, ۱۹(۳), ۳۲۳-۳۳۸.
- Gunnarsson, B. R., Vanden Broucke, S., Baesens, B., Óskarsdóttir, M., & Lemahieu, W. (۲۰۲۱). Deep learning for credit scoring: Do or don't?. ***European Journal of Operational Research***, ۲۹٥(۱), ۲۹۲-۳۰٥.
- Joshi, S. (۲۰۲٥). Review of Gen AI Models for Financial Risk Management. ***International Journal of Scientific Research in Computer Science, Engineering and Information Technology***, ۱۱(۱), ۷۰۹-۷۲۳.
- Khalili, N., & Rastegar, M. A. (۲۰۲۳). Optimal cost-sensitive credit scoring using a new hybrid performance metric. ***Expert Systems with Applications***, ۲۱۳, ۱۱۹۲۳۲.
- Lando, D. (۲۰۰۹). Credit risk modeling. In *Credit Risk Modeling*. **Princeton University Press**.
- Maleki, M. S., Motevallian, S. N., Hosseini, F., Sabokrou, M., & Maleki, H. S. (۲۰۲۱, October). Improvement of credit scoring by lstm autoencoder model. In ***۲۰۲۱ ۱۱th International Conference on Computer Engineering and Knowledge (ICCKE)*** (pp. ۱۸۲-۱۸۷). IEEE.
- Mpofu, T. P., & Mukosera, M. (۲۰۱٤). Credit scoring techniques: a survey. ***International Journal of Science and Research (IJSR)***, ۲۳۱۹-۷۰٦٤.

- Naili, M., & Lahrichi, Y. (۲۰۲۲). Banks' credit risk, systematic determinants and specific factors: recent evidence from emerging markets. **Heliyon**, ۸(۲).
- Rofik, R., Aulia, R., Musaadah, K., Ardyani, S. S. F., & Hakim, A. A. (۲۰۲٤). The optimization of credit scoring model using stacking ensemble learning and oversampling techniques. ***Journal of Information System Exploration and Research***, ۲(۱).
- Suksamai, C., Hengpraprohm, K., & Silachan, K. (۲۰۲۲, July). A Comparison of Normalization Data Transformation Efficiency Affecting with Bank Customer Credit Data Classification using Data Mining Techniques. **The ۱٤th NPRU National Academic Conference Nakhon Pathom Rajabhat University**.
- Sutrisno, H., & Halim, S. (۲۰۱۷, September). Credit Scoring Refinement Using Optimized Logistic Regression. **In** ***۲۰۱۷ International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)*** (pp. ۲٦-۳۱). IEEE.
- Thomas, L., Crook, J., & Edelman, D. (۲۰۱۷). *Credit scoring and its applications*. **Society for industrial and Applied Mathematics**.
- Tripathi, D., Edla, D. R., & Cheruku, R. (۲۰۱۸). Hybrid credit scoring model using neighborhood rough set and multi-layer ensemble classification. ***Journal of Intelligent & Fuzzy Systems***, ۳٤(۳), ۱۵٤۳-۱۵٤۹.
- Wang, Y., Jia, Y., Fan, S., & Xiao, J. (۲۰۲۳). Deep Reinforcement Learning Based on Balanced Stratified Prioritized Experience Replay for Customer Credit Scoring in Peer-to-Peer Lending.
- Zhang, R., Liguo, X., & Qin, W. An Ensemble Credit Scoring Model Based on Logistic Regression with Heterogeneous Balancing and Weighting Effects. ***Available at SSRN*** ٤۱٦۷۸۲۱.
- Zhuang, Y., Xu, Z., & Tang, Y. (۲۰۱۵, September). A credit scoring model based on bayesian network and mutual information. **In** ***۲۰۱۵ ۱۲th Web Information System and Application Conference (WISA)*** (pp. ۲۸۱-۲۸٦). IEEE.