Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

سال ششم، شماره‌های ۱۹ و ۲۰، بهار و تابستان ۱۳۹۳
صص: ۱۱- ۲۸

فصلنامهٔ علمی- پژوهشی
فناوری اطلاعات و ارتباطات ایران

ارائهٔ مدل ترکیبی شبکه‌های عصبی با بهره‌گیری از یادگیری جمعی به منظور ارزیابی ریسک اعتباری

[*]شعبان الهی [**]احمد قدس‌الهی [***]حمیدرضا ناجی
*دانشیار، گروه مدیریت، دانشگاه تربیت مدرس، تهران
**کارشناس ارشد، مدیریت فناوری اطلاعات، دانشگاه تربیت مدرس، تهران
***استادیار، دانشگاه شهید بهشتی، تهران

تاریخ دریافت: ۱۳۹۲/۴/۱۳ تاریخ پذیرش: ۱۳۹۲/۱۲/۱۳

چکیده

بانکداری صنعت ویژه‌ایی است که با سرمایه و ریسک برای کسب سود مواجه است. یکی از مهم‌ترین ریسک‌های بانکی، ریسک اعتباری است که حوزهٔ تحقیقاتی پویایی را در مطالعات مدیریت به خود اختصاص داده است. در این پژوهش یک سیستم ترکیبی ارزیابی ریسک اعتباری ارائه می‌شود، که از یادگیری جمعی برای تصمیم‌گیری در مورد اعطای اعتبار به فرد متقاضی استفاده می‌کند. ترکیب تکنیک‌های دسته‌بندی و خوشه‌بندی در این پژوهش، منجر به بهبود عملکرد سیستم می‌شود. برای آموزش شبکه‌های عصبی از مجموعهٔ داده‌های واقعی، از نمونه‌های تقاضای اعتبار در بانکی در آلمان استفاده شده است. مدل پژوهش در قالب یک سیستم چند عاملی ارزیابی ریسک‌اعتباری طراحی شد و نتایج نشان داد که این سیستم صحتّی بالاتر، عملکردی برتر و هزینهٔ کم‌تری، در دسته‌بندی متقاضیان اعتبار نسبت به دیگر روش‌های مشابه حاصل می‌کند.

واژه‌های کلّیدی: شبکهٔ عصبی، یادگیری جمعی، ریسک اعتباری

مقدمه

یکی از خصوصیات بانک‌ها این است که به صورت کوتاه مدت قرض می‌گیرند و به صورت بلند مدت وام می‌دهند و این خصوصیت به این امر منتهی شده که بانک‌ها در معرض ریسک اعتباری قرار بگیرند که به معنی از دست دادن بهره و سود بانک به دلیل متعهد نبودن وام گیرنده به قرار داد خود است. این گونه ریسک دارای تاریخچه‌ایی تقریباً همزمان با پیدایش بانک و سیستم بانکداری است. وجود این ریسک ناشی از این احتمال است که وام گیرنده قادر به پرداخت منظم اصل و فرع بدهی خود نباشد و یا در پرداخت آن کوتاهی کند[1]. منافع قدیمی و سنتی بانکی، در مواجهه با رقابت شدید بین بانک‌ها در حال کاهش است. این امر بانک‌ها را ترغیب می‌کند که با وجود ریسک اعتباری و بحران‌های اعتباری بانک‌ها که در سال‌های اخیر روی داده است، تصمیم به تغییر جهت خود از کسب و کار مداری به مشتری مداری بگیرند، تا اینکه منافع بیش‌تری کسب کنند[39]. با این وجود در سال های اخیر، وام دهی زیاد و بدون ارزیابی و پیش بینی موجب افزایش مشتریان بد حساب و وام‌های معوق شده است. در سال‌های اخیر و قبل از بحران مالی ۲۰۰۷-۲۰۰۸ مدیریت ریسک اعتباری میانه‌تر و ساده‌تری در راستای هم سویی با استانداردهای

نویسندهٔ عهده‌دار مکاتبات: شعبان الهی Elahi@modares.ac.ir

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

اعتباری بکار گرفته می‌شد. اما بحران مالی ۲۰۰۷-۲۰۰۸ این امر را مشهود ساخته است که درک ریسک از طریق بازارهای مالی برای قابلیت و توانایی بانک‌ها به منظور افزایش سرمایه‌های جدیدشان، امری مهم و حیاتی است[4]. اساساً می‌توان بیان کرد که تصمیم در مورد اعطای اعتبار درخواست شده به متقاضی، می‌تواند برای بانک بسیار مهم باشد و کسب و کار و سودآوری بانک را بطور چشمگیری تحت تأثیر قرار دهد. اگر وام گیرنده در پرداخت خود دچار مشکل شود، بانک ضرر می‌کند، از سوی دیگر اگر بانک به افرادی که دارای اعتبار خوبی هستند اعتبار ندهد، مجدداً دچار ضرر و زیان خواهد شد [45]. بطور کل، ریسک اعتباری موجب مشکلات مالی در بانک‌ها خواهد شد و ارزیابی آن نیازمند تکنیک‌های مدل‌سازی پیشرفته است که مرتبط به منبع عدم قطعیت ناشی شده است. بر این اساس مدل‌سازی ریسک اعتباری مؤسسات مالی و بخصوص بانک‌ها، امری مهم و ضروری است[28]. یکی از تکنیک‌های مدیریت ریسک اعتباری که از دیر باز تا کنون برای مدیریت ریسک اعتباری بکار گرفته شده است، تکنیک امتیازدهی اعتباری و رتبه‌بندی مشتریان بانک از جنبهٔ اعتبار آن‌ها است. در عمل اصلی‌ترین روش ارزیابی ریسک اعتباری فرد وام گیرنده، ارزیابی امتیاز اعتباری او است. مبتنی بر یک پایگاه داده قدیمی ، وام گیرندگان به افرادی که وام خود را در موعد مقرّر پرداخته‌اند و یا به پرداخت آن در تاریخ مقرّر متعهّد نبوده‌اند، تقسیم‌بندی می‌شوند[9]. تکنیک امتیازدهی اعتباری به برخی از سازمان‌ها مثل بانک‌ها و شرکت‌های صادر کنندهٔ کارت‌های اعتباری این امکان و توانایی را می‌دهد که بتوانند در مورد اعطای اعتبار به یک فرد، بر اساس یک سری از معیارهای از پیش تعیین شده، تصمیم بگیرند که آیا اعتبار درخواست شده واگذار شود یا خیر. معمولاً امتیاز اعتباری عددی است که ارزش و اعتبار فرد را بر اساس تحلیلات کمّی سابقهٔ اعتباری و دیگر معیارهای مربوط به متقاضی مشخص می‌کند و بیان می‌دارد که تا چه حدّی فرد وام گیرنده به پرداخت قرض خود در موعد مقرّر متعهد خواهد بود[45]. این تکنیک با استفاده از داده‌های قدیمی و تفکیک آماری، سعی دارد که اثرات خصوصیات مختلف متقاضی را بروی نکول وام تفکیک کند[26]. مزایای امتیازدهی اعتباری شامل کاهش هزینهٔ تحلیلات اعتباری، تسریع در امر تصمیم‌گیری مربوط به اعتبارات و الویت‌بندی اعتباری است[21]. مدلی که معمولاً به منظور امتیازدهی اعتباری بکار گرفته می‌شد و رایج بود، تحلیل امتیازدهی اعتباری از روش تحلیل متمایز کنندهٔ چند متغیره بود که برای اولین بار توسط «آلتمن[1]» بکار گرفته شد[3]. این مدل، متغیرهای مالی را که قدرت تشریح آماری دارند، شناسایی می‌کند و از طریق آن‌ها، متقاضیان بد را از خوب متمایز می‌کند. مدل امتیازی می‌دهد که در برخی شرایط این امتیاز به عنوان احتمال نکول شناخته می‌شود، در غیر این صورت، امتیاز را می‌توان برای سیستم دسته‌بندی استفاده کرد، بدین صورت که متقاضی بر اساس امتیاز و آستانهٔ مشخص شده در کلاس خوب یا بد قرار می‌گیرد[5]. روش‌های آماری قدیمی که برای ایجاد مدل‌های امتیازدهی اعتباری بکار گرفته شده‌اند، عبارتند از:

پروبیت[2]، لجیت[3]، رگرسیون خطّی، تحلیل متمایز کنندهٔ خطّی[4]، تحلیل متمایز کنندهٔ نمایی[5]، توبیت[6]، درخت‌های صفر و یکی، روش‌های کمّی، رگرسیون لجستیک[7] [7،42، 45]

با وجود استفاده از این روش‌ها برای ارزیابی ریسک اعتباری، می‌توان بیان کرد که قابلیت آن‌ها در تمایز مشتری خوب از بد هم چنان دارای مشکل است. دو روش رایج که در امتیازدهی اعتباری استفاده شده است، عبارتند از :تحلیل متمایز کننده و رگرسیون لجستیک[8، 24، 32]. ضعف تحلیل متمایز کننده، فرض خطّی بودن رابطه بین متغیر ها است که معمولاً غیر خطّی است و همچنین حساسیت به انحرافات، از فرضیه‌های نرمال بودن چند متغیره‌ها است. رگرسیون لجستیک خروجی‌های دو

1 .Altman
2 .Probit
3 .Logit
4 .Linear Discriminant Analysis (LDA)
5 .Quadratic Discriminant Analysis (QDA)
6 .Tobit
7 .Logistic regression(LR)

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

بخشی می‌دهد و روابط خطّی، بین متغیرها در تابع لجستیک مفروض می‌شود. به دلیل رابطهٔ خطّی، اینگونه بیان می‌شود که هر دو روش دارای عدم صحّت کافی است[38]. در دوره‌های اخیر رویکردهای جدیدی نیز برای توسعهٔ سیستم‌های امتیازدهی اعتباری بکار گرفته شده است. از تکنیک‌های جدید که در حوزهٔ اعتباری مطرح شده و نتایج خوب و موفقی ارائه داده است، کاربرد تکنیک‌های داده کاوی، سیستم‌های فازی، الگوریتم ژنتیک، سیستم‌های عصبی- فازی، و به خصوص شبکه‌های عصبی مصنوعی در حوزهٔ امتیازدهی اعتباری در بخش بانکی و مالی است [5,6,33,34]. تحقیقات اخیر نشان داده است تکنیک‌های هوش مصنوعی مثل شبکه‌های عصبی مصنوعی، محاسبات تکمیلی، الگوریتم ژنتیک و ماشین بردار پشتیبان برای تحلیل‌های آماری، در جهت ارزیابی ریسک اعتباری نتایج خوب و موفقی ارائه داده است [45]. در بین این روش‌ها، شبکه‌های عصبی بسیار مؤثر و کارا، و گزینهٔ خوبی برای جایگزینی با تحلیل متمایز کننده و رگرسیون لجستیک است. در تحقیقات انجام شده، موارد کاربرد شبکه‌های عصبی در امتیازدهی اعتباری نشان داده است که این روش بسیار صحیح‌تر از روش‌های آماری قدیمی است[4، 35،42].

مرور ادبیات پژوهش

در پاسخ به رشد صنعت اعتباری و مدیریت وام‌های بزرگ و کوچک، مدل‌های امتیازدهی رفتاری و اعتباری بصورت فعّال در حال توسعه و ایجاد است و تحقیقات وسیعی در حوزهٔ یادگیری ماشینی و به خصوص شبکه‌های عصبی در جهت ارزیابی ریسک اعتباری انجام شده است و نتایج خوب و مؤثری حاصل شده است، که همگی نمایانگر قابلیت و توانایی این ابزارها در ارزیابی ریسک اعتباری و بطور خاص امتیازدهی اعتباری است. «دزایی[8]» و همکارانش [14] برای تصمیمات امتیازدهی اعتباری سه تکنیک شبکه‌های عصبی، تحلیل متمایز کنندهٔ خطّی و رگرسیون لجستیک را مقایسه کرده‌اند. آن‌ها به این نتیجه رسیدند که شبکه‌های عصبی از دو روش دیگر در طبقه‌بندی درخواست‌های اعتباری به دو دستهٔ خوب و بد بهتر عمل می‌کند. «مالهوترا[9]» و همکارش [29] از مدل‌های عصبی- فازی برای تحلیل درخواست وام مشتریان استفاده کرده‌اند و مزایای سیستم‌های عصبی- فازی را در مقایسه با روش‌های قدیمی آماری در ارزیابی ریسک اعتباری بیان نموده‌اند. «هافمن[10]» [19] یک مدل دسته‌بندی عصبی- فازی و ژنتیک- فازی را برای امتیازدهی اعتباری استفاده کرده است. «بیزین[11]» و همکاران[7] یافتند که بر اساس هشت مجموعهٔ داده امتیازدهی اعتباری واقعی، کوچک‌ترین مربعات ماشین بردار پشتیبان[12] و شبکهٔ عصبی عملکرد خوبی در دسته‌بندی داشته‌اند. «بنسیک[13]» و همکاران [10] به مدل‌سازی امتیازدهی اعتباری کسب و کارهای کوچک با استفاده از رگرسیون لجستیک، شبکه‌های عصبی و درخت تصمیم می‌پردازند. هدف اصلی آن‌ها استخراج خصوصیات مهم برای امتیازدهی اعتباری در کسب و کارهای کوچک است. در اینجا از چهار معماری متفاوت شبکه‌های عصبی با عنوان‌های BP, RBF, PNN, LVQ استفاده شده است. «سیه[14]» [20] از یک رویکرد ترکیبی کاوش در طراحی مدل امتیازدهی اعتباری استفاده کرده است. به دلیل وجود نمونه‌های غیر نمایانگر از یک رویکرد دو مرحله‌ایی استفاده شده است که از نقشه‌های خود سازمانده برای تعیین خوشه‌ها استفاده شده و از الگوریتم کامینز[15] برای یافتن پاسخ نهایی. در اینجا از شبکهٔ عصبی پرسپترون چند لایه‌ای نظارتی پیشرو استفاده شده است. از این طریق صحّت امتیازدهی نسبت به روش‌های دیگر بهبود یافت. «هوانگ[16]» و همکاران[21]، مدل ترکیبی

8 .Desai

9 . Malhotra
10 .Hoffmann
11 .Baesens
12 .Least-Squares Support Vector Machine (LS-SVM)
13 .Bensic
14 .Hsieh
15 .K-means
16 .Huang

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

مبتنی بر ماشین بردار پشتیبان را پیشنهاد داده‌اند که امتیاز اعتباری متقاضی را از طریق خصوصیات ورودی فرد می‌سنجد. «لیم[17]» و «سان[18]»[27] یک مدل امتیازدهی اعتباری، مبتنی بر شبکه‌های عصبی طراحی کردند که بطور پویا تغییرات خصوصیات متقاضی وام را بعد از وام اعمال می‌کند. آن‌ها بیان می‌کنند که با این روش نسبت به روش‌های ایستا، ریسک را کم‌تر خواهند کرد. «عبدو[19]» و همکاران [2] قابلیت و توانایی شبکه‌های عصبی چند لایه‌ایی پیشرو را در مقایسه با روش‌های قدیمی مثل تحلیلات متمایز کننده، پروبیت و رگرسیون لجستیک، در ارزیابی ریسک اعتباری بانک‌ها به منظور امتیازدهی اعتباری می‌سنجند. این تحقیق نمایان ساخت که مدل‌های شبکه عصبی میانگین نرخ دسته‌بندی صحیح‌تری نسبت به دیگر تکنیک‌ها ارائه می‌کنند. «آنجلینی[20]» و همکارانش [5] کاربرد شبکه‌های عصبی را برای ارزیابی ریسک اعتباری مرتبط با کسب و کارهای کوچک ایتالیا مورد بررسی قرار داده‌اند. در این تحقیق دو مدل شبکهٔ عصبی ارائه شده است، یکی با شبکهٔ پیشرو و دیگری با معماری هدف خاص. آن‌ها بیان کرده‌اند که هر دو شبکهٔ عصبی می‌توانند در یادگیری و تخمین احتمال نکول متقاضی بسیار مؤثر و موفق باشند، با این شرط که تحلیلات داده، پیش پردازش داده و آموزش شبکهٔ عصبی، بصورت صحیح و با دقّت انجام شود. «وانگ[21]» و «هوانگ[22]» [40] از شبکهٔ عصبی پس انتشار برای دسته‌بندی متقاضیان اعتبار استفاده کرده بود. «ساسترسیک[23]» و همکارانش [38] یک یک مدل امتیازدهی اعتباری مبتنی بر شبکه‌های عصبی را برای مؤسسات مالی در جایی که داده به میزان کافی موجود نیست، طراحی کرده‌اند. آن‌ها از الگوریتم ژنتیک برای انتخاب متغیرها استفاده کرده‌اند. همچنین از شبکهٔ عصبی کوهنن[24] برای دسته‌بندی مشتریان به دو دستهٔ خوب و بد استفاده کرده‌اند. علاوه بر این، از شبکه‌های عصبی پس انتشار استفاده کرده و خروجی آن را با رگرسیون لجستیک مقایسه کرده‌اند. نتایج بیانگر این بود که شبکه‌های عصبی پس انتشار نتایج بهتری را نسبت به رگرسیون لجستیک نشان می‌دهد. «یو[25]» و همکاران [45] در تحقیق خود یک مدل یادگیرنده ترکیبی چند مرحله‌ایی شبکهٔ عصبی را برای ارزیابی ریسک اعتباری پیشنهاد داده‌اند. مدل پیشنهادی آن‌ها از شش مرحله تشکیل شده بود. «نانی[26]» و «لومینی[27]» [31] اثرات استفاده از یادگیری جمعی را نسبت به مدل‌های دسته‌بندی منفرد می‌سنجند. نتایج بررسی‌های آن‌ها نشان داد که استفاده از رویکرد جمعی موجب افزایش صحّت مدل دسته‌بندی می‌شود. همچنین نتایج نشان داد که برترین روش منفرد، روش شبکهٔ عصبی چند لایه‌ایی است. «کوین[28]» و همکاران [36] در تحقیق خود یک الگوریتم بهبود یافته با نرخ‌های یادگیری متغیر مبتنی بر پس انتشار را ارائه و از آن برای امتیازدهی اعتباری استفاده کرده‌اند. نتایج تحقیق آن‌ها نشان داد که الگوریتم بطور چشمگیری تعداد تکرار شبکه را کاهش داده، زمان آموزش شبکه را کم‌تر کرده و صحّت آموزشی را بهبود بخشیده است، بصورتی که میزان صحّت بالاتری را نسبت به شبکه‌های عصبی پس انتشار ارائه داده است. «التر» و همکاران [15] مدل هوشمند مصنوعی مبتنی بر نرون را برای تصمیمات وام ارائه می‌دهند. در اینجا یک مدل شبکهٔ عصبی چند لایه‌ایی پیشرو با الگوریتم پس انتشار برای تصمیمات وام بانک ارائه می‌شود تا شبکهٔ عصبی را به عنوان یک ابزار خوب برای ارزیابی در خواست اعتباری جهت پشتیبانی تصمیمات بانک‌های اردن در رابطه با وام معرفی کند. نتایج نشان داد که شبکه‌های عصبی تکنیک موفقی است که می‌توان برای ارزیابی درخواست اعتباری استفاده

17 .Lim
18 .Sohn
19 .Aabdou
20 .Angelini
21 .Wang
22 .Huang
23 .Sustersic
24 .Kohenen
25 .Yu
26 .Nani
27 .Lumini
28 .Qin

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

کرد.«تسایی[۲۹]» و «چن[۳۰]» [39] از یادگیری ماشینی ترکیبی برای امتیازدهی استفاده کرده‌اند. آن‌ها چهار روش ترکیبی را با داده‌های واقعی بانکی در تایوان بررسی کرده‌اند. در اینجا چهار روش دسته‌بندی با عنوان درخت تصمیم، شبکه‌های عصبی، مدل دسته‌بندی بیزین و رگرسیون لجستیک، و دو روش خوشه‌بندی کامینز و بیشینه‌سازی انتظارات بکار گرفته شده است. نتایج تحقیق آن‌ها نشان داد که در روش‌های فردی، رگرسیون لجستیک، در میان روش‌های دسته‌بندی نسبت به دیگر روش‌ها عملکرد بهتری داشته و در میان روش‌های خوشه‌بندی، روش بیشینه سازی انتظارات عملکرد بهتری نسبت به روش دیگر داشته است. نتایج تحقیق نشان داده است که مدل ترکیبی که مبتنی بر رگرسیون لجستیک و شبکۀ عصبی است می‌تواند بیش‌ترین صحّت را به ارمغان بیاورد. «وانگ[۳۱]» و همکاران [41] مقایسه‌ایی بین روش‌های جمعی به منظور امتیازدهی اعتباری انجام داده‌اند. در این تحقیق عملکرد سه روش رایج یادگیری مورد بررسی قرار گرفت. این سه روش عبارتند از بگینگ[۳۲]، بگینگ[۳۲]، بوستینگ[۳۳] و استکینگ[۳۴]. این سه روش یادگیری جمعی برای چهار مدل دسته‌بندی رگرسیون لجستیک، درخت تصمیم، شبکۀ عصبی و ماشین بردار پشتیبان بکار گرفته شد. نتایج حاکی از افزایش صحّت و بهبود عملکرد هر یک از مدل‌های دسته‌بندی از طریق بکارگیری یادگیری جمعی بود. در این تحقیق بگینگ و استکینگ درخت تصمیم بهترین نتایج را ایجاد کرده‌اند.

روش پژوهش

این پژوهش از نقطه نظر هدف، کاربردی و نوع آن، تحلیلی- توصیفی است. از مجموعه داده‌های واقعی بانکی در آلمان[۳۵] برای آموزش و تست شبکه‌های عصبی بکار رفته در مدل، استفاده می‌شود. این پایگاه داده مشتمل بر ۱۰۰۰ نمونه است که هر یک از نمونه‌ها توسط ۲۰ معیار توصیف شده‌اند. این ۲۰ معیار عبارتند از: وضعیت حساب جاری، مدت زمان، سابقۀ اعتباری، هدف، میزان اعتبار، حساب پس‌انداز یا ضمانت، وضعیت شغلی، نرخ درصدی اقساط درآمد از دست رفته، مشخصات فردی و جنسیت، وضعیت ضامن، مدت زمان اقامت در محل کنونی، دارایی، سنّ، دیگر برنامه‌های اقساط، واحد مسکونی، تعداد اعتبارات در این بانک، وضعیت شغلی، تعداد افراد مسئول برای ارائۀ پشتیبانی، دارا بودن تلفن ثابت و وضعیت اقامتی کارگر. از داده‌های این بانک در تحقیقات بسیاری استفاده شده است [38،23،3]

- **مدل پیشنهادی ارزیابی ریسک اعتباری**

همانطور که پیش‌تر ذکر شد، امتیازدهی اعتباری در عمل اصلی‌ترین تکنیک تحلیلی برای مدیریت ریسک اعتباری است. یک مدل صحیح امتیازدهی اعتباری، بایستی امتیازات بالاتر را به شخصی که وام خود را طبق تعهداتش می‌پردازد، تخصیص دهد و امتیاز پایین‌تر را به کسانی تخصیص دهد که بازپرداختشان به تعویق خواهد افتاد. با حجم قابل ملاحظه‌ایی از وام‌ها حتّی بهبود کوچکی هم در صحّت امتیازدهی اعتباری می‌تواند ریسک اعتباری مؤسسه را کاهش دهد و منجر به پس‌اندازهای چشمگیر آتی شود. به دلیل رشد سریع صنعت اعتباری، ساخت یک مدل امتیازدهی مناسب و کارا فعّالیت مهمّی برای کاهش هزینه‌ها و تصمیم‌گیری کارا و مؤثر است. با افزایش مستمر مقیاس و پیچیدگی مؤسسات مالی و بانک‌ها و سرعت تقاضای نقل و انتقالات، نیاز است که تکنیک‌های پیچیده‌تر و دشوارتری برای مدیریت ریسک بکار گرفته شود و میزان در معرض ریسک اعتباری قرار گرفتن، بصورت سریع و در

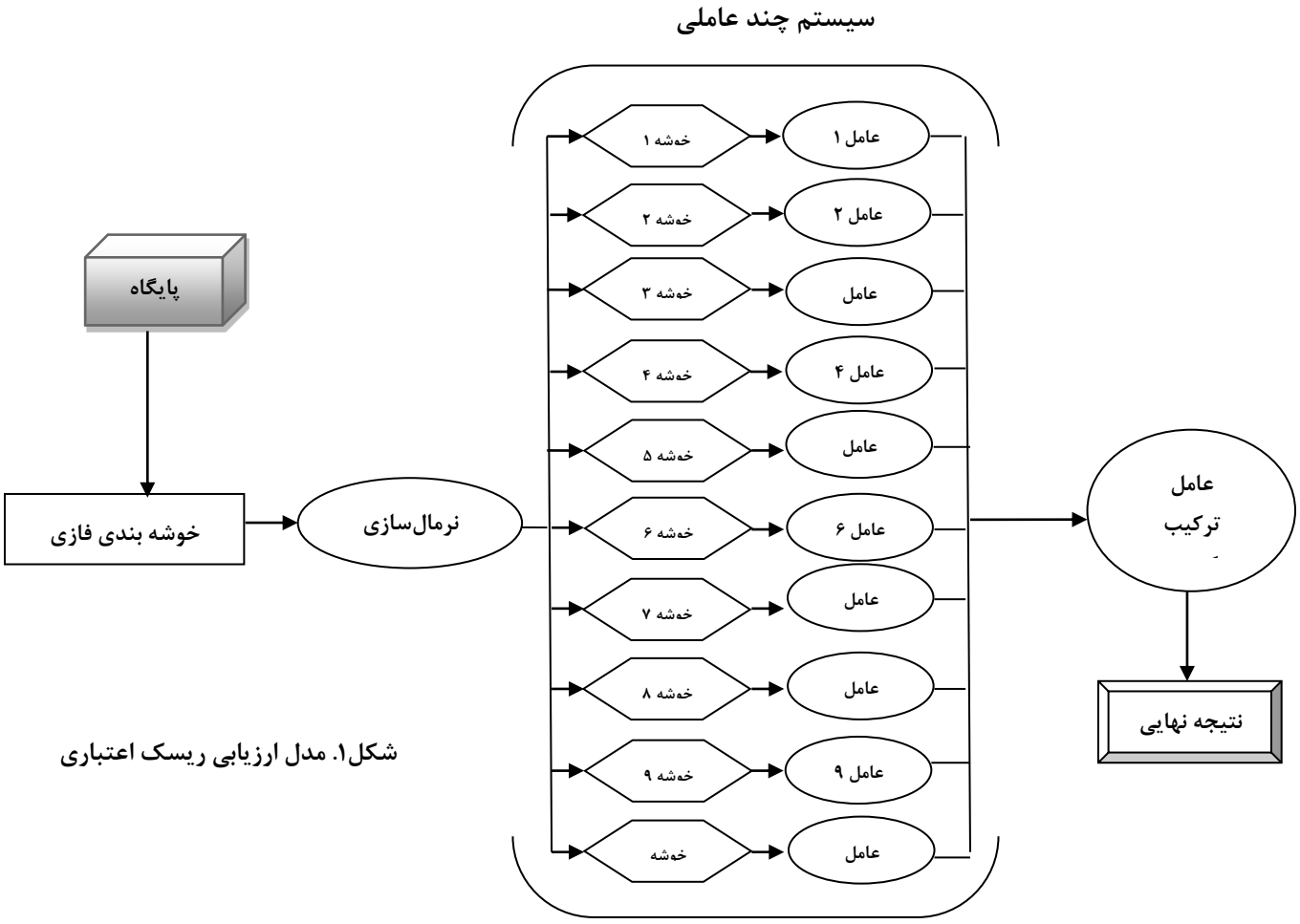
29 .Tsai
30 .Chen
31 .Wang
32 .Bagging
33 .Boosting
34 .Stacking

۳۵.این داده ها به صورت فایل الکترونیک در وب سایت *www.ics.uci.edu* به صورت عمومی قابل دریافت است.

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

زمان کم‌پایش شود. در اینجا به بیان و تشریح مدل پیشنهادی این پژوهش پرداخته می‌شود که درشکل ۱ ارائه شده است. این مدل، مدلی ترکیبی است، به این صورت که با ترکیب تکنیک‌های گوناگون یادگیری ماشینی، اقدام به پیش‌بینی امتیاز فرد خواهد کرد. در یادگیری ماشینی، رویکرد ترکیبی یک حوزهٔ تحقیقاتی فعّال برای بهبود عملکرد دسته‌بندی یا پیش‌بینی نسبت به رویکردهای مستقل یادگیری بوده است.[28،46،39،20،22،25،29، 13].

نتایج تحقیقات حاکی از عملکرد برتر مدل‌های ترکیبی در پیش‌بینی بوده است. در این پژوهش نیز مدلی ترکیبی ارائه شده است، که تکنیک خوشه‌بندی را با تکنیک دسته‌بندی بکار می‌گیرد. برای خوشه‌بندی از تکنیک خوشه‌بندی فازی و از پرسپترون چند لایه‌ایی به عنوان تکنیک دسته‌بندی استفاده شده است. علاوه بر این در این مدل از عامل‌ها بهره گرفته شده است. همانطور که در این شکل دیده می‌شود مجموعهٔ داده یا بانک اطلاعاتی ما به تعدادی خوشه تقسیم می‌شود که جهت انجام سریع عملیات پردازش این داده‌ها، برای هر خوشه یک عامل جداگانه در نظر گرفته می‌شود که این عامل وظیفهٔ پردازش مورد نظر و تعریف شده برای آن عامل را روی خوشه اختصاص یافته به آن انجام می‌دهد. بدین طریق سیستم قادرخواهد بود که بطور همزمان خوشه‌های مختلف را پردازش نموده و سرعت انجام عملیات چندین برابر شود. لازم به ذکر است که علاوه بر پردازش موازی در اینجا عوامل مختلف می‌توانند هر کدام عملیات مختلفی را روی خوشهٔ مربوط به خود انجام دهند. این دو خصوصیت قابلیت بسیار بالایی به مدل پیشنهادی ما، هم در زمینهٔ سرعت عملیاتی و هم انعطاف‌پذیری جهت انجام عملیات مختلف روی یک بانک اطلاعاتی را می‌دهد. در بخش‌های بعدی به تشریح اجزای مدل پژوهش پرداخته می‌شود.



شکل۱. مدل ارزیابی ریسک اعتباری

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

۱. خوشه‌بندی فازی

اولین فاز مدل، خوشه‌بندی فازی داده‌های پژوهش است. علت بکارگیری خوشه‌بندی فازی فراهم نمودن داده‌های آموزشی مختلف بر اساس میزان خصوصیات مشترکشان و برطرف نمودن عدم قطعیت موجود در داده‌های پژوهش است. خوشه‌بندی یک تکنیک دسته‌بندی غیر نظارتی است که از ویژگی‌های نمونه‌ها برای دسته‌بندی آن‌ها به گروه‌هایی با خصوصیات مشترک استفاده می‌کند. خوشه‌بندی داده‌های پژوهش به ما کمک خواهد کرد که خصوصیات مشابه نمونه‌ها را پیدا کنیم و نمونه‌ها را به گروه‌های همگن و مشابه گروه‌بندی کنیم و در نتیجه پیش‌بینی و دسته‌بندی‌های ناصحیح را که به دلیل نمونه‌های آموزشی نامناسب رخ می‌دهد، کم کنیم. گاهی اوقات حتّی با یک مدل دسته‌بندی بسیار صحیح، قابلیت مدل برای پیش‌بینی یک نمونهٔ جدید محدود می‌شود. این محدودیت به دلیل الگوهای نامناسب دسته‌بندی است که در داده‌های آموزشی بروز می‌کند و عدم قطعیت موجود در این داده‌ها، پیچیدگی یادگیری را افزایش می‌دهد. اگر کیفیت نمونه‌ها بهبود داده شود، توانایی مدل نیز افزایش می‌یابد. از خوشه‌بندی به عنوان فرایند پیش پردازش داده‌ها استفاده شده است تا کیفیت داده‌های ورودی شبکه‌های عصبی بهبود داده شود و در نتیجه از بکارگیری مدل جمعی شبکه‌های عصبی عملکرد بالاتری حاصل شود. در این مدل از خوشه‌بندی فازی استفاده می‌شود، به طوریکه این امکان وجود دارد یک نمونه به یک یا چند خوشه با درجه‌های عضویت متفاوت بین صفر و یک تعلق داشته باشد. یکی از تکنیک‌های خوشه‌بندی فازی روش FCM[۳۶] است. FCM یک تکنیک خوشه‌بندی است، بصورتیکه هر داده یا نمونه می‌تواند به بیش از یک خوشه تعلق داشته باشد و یا ممکن است هر داده یا نمونه به هر کدام از خوشه‌های تعیین شده تا حدودی تعلق داشته باشد، که این امر از طریق درجهٔ عضویت تبیین می‌شود.

۲. نرمال‌سازی

یکی از مشکلاتی که در رابطه با استفاده از شبکه‌های عصبی در حوزهٔ امتیازدهی اعتباری در تحقیقات قبلی مشاهده می‌شود، بحث نرمال‌سازی داده‌ها است. ارزش‌های عددی که برای گره‌های ورودی شبکه تعیین می‌شوند معمولاً بین صفر و یک هستند و برای ارزیابی اعتباری، این ارزش‌ها نمایانگر خصوصیات متقاضی اعتبار هستند که هر یک، در ارزش‌های عددیشان متفاوت هستند. در صورت اجرای یک نرمال‌سازی ساده، به عنوان مثال تقسیم هر داده بر بزرگ‌ترین دادهٔ مجموعهٔ کلّی داده‌ها، اطلاعات مؤثر و مفیدی را از دست خواهیم داد. در مجموعه داده‌های بیشینهٔ عددی برابر با ۱۸۴۲۴ است. در صورتیکه کلّیه داده‌ها را بر این عدد به منظور نرمال‌سازی تقسیم کنیم، ممکن است اطلاعات مفیدی را از دست بدهیم، به دلیل اینکه با اینکار اکثر داده‌ها به سمت صفر میل می‌کند. این کار منتهی به آموزش ناکافی شبکه می‌شود. بنابراین نرمال‌سازی داده‌های ورودی متقاضی اعتبار بایستی با دقت انجام گیرد، در حالی که مفاهیم هر کدام از خصایص حفظ می‌شود. در اینجا نرمال‌سازی نیز به عنوان یک پیش پردازش برای شبکه‌های عصبی مورد استفاده، محسوب می‌شود. این پیش پردازش به منظور استفاده از داده‌ها در آموزش و تست شبکه‌های عصبی انجام می‌گیرد. در این پژوهش، نرمال‌سازی از طریق یافتن بیشهٔ عددی در هر ستون (معیار) از مجموعه داده‌ها و تقسیم باقی ورودی‌های هر ستون به آن مقدار بیشینه انجام می‌گیرد.

۳. شبکه‌های عصبی مورد استفاده در پژوهش

برای طراحی شبکهٔ عصبی و به منظور نزدیک‌تر کردن پیش‌بینی سیستم به پیش‌بینی‌های صحیح‌تر و با دقت‌تر از شبکهٔ عصبی پرسپترون چند لایه‌ایی استفاده شده است. تکنیک‌های زیادی در زمینهٔ شبکهٔ عصبی وجود دارد که پرسپترون چند لایه‌ایی یکی از گزینه‌های انتخابی برای

36. Fuzzy C-Means clustering

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

تصمیمات امتیازدهی اعتباری است. پرسپترون چند لایه‌ایی یکی از پر طرفدارترین الگوریتم‌های شبکهٔ عصبی است و بیش‌ترین استفاده را در امور تجاری و امتیازدهی اعتباری داشته است که این امر به دلیل توانایی زیاد شبکه‌های عصبی در مدل‌سازی مسائل است[42]. شبکهٔ عصبی پرسپترون چند لایه‌ایی از لایه‌های ورودی، میانی و خروجی، و از نرون‌های به هم مرتبط تشکیل شده است. نرون‌های یک لایه مطابق با یک‌سری از اوزان، با یکدیگر ترکیب شده و به لایهٔ بعدی منتقل می‌شوند. در طول مرحلهٔ یادگیری، نمونه‌ها به شبکه ارائه می‌شود و اوزان و بایاس[37] های شبکه بر اساس نرخ خطای حاصل از خروجی‌ها تنظیم و اصلاح می‌شوند. بر این اساس دوره‌های زیادی به منظور آموزش شبکه مورد نیاز است. معمولاً زمان آموزش شبکه بالاست و فرایند آموزش تا وقتی که خطا به قدر کافی کاهش نیابد و یا معیارهای دیگر برآورده نشود، متوقف نخواهد شد. برای ساخت یک مدل شبکهٔ عصبی پیش‌بینی کننده، داده‌های آموزش و تست شبکه مورد نیاز است. داده‌های آموزش برای توسعهٔ مدل شبکهٔ عصبی مصنوعی استفاده می‌شوند و داده‌های تست برای ارزیابی توانایی مدل مورد استفاده قرار می‌گیرند. تکنیک‌های دسته‌بندی بر اساس نمونه‌هایی هستند که بخش‌های ورودی را به یکی از خروجی‌های مطلوب و مورد نظر آدرس‌دهی می‌کنند. به عبارتی، یک الگوی دسته‌بندی کننده را می‌توان از طریق فرایند آموزش و یادگیری ایجاد کرد. فرایند آموزش مدل دسته‌بندی، محاسبهٔ فاصلهٔ تقریبی بین ورودی و خروجی نمونه و ایجاد برچسب خروجی‌های مورد نظر و صحیح از داده‌های آموزشی است. این فرایند، فاز ایجاد مدل نامیده می‌شود. زمانیکه مدل ایجاد شد، شبکه می‌تواند یک نمونهٔ ناشناخته را به یکی از دسته‌های شناخته شده در داده‌های آموزشی اختصاص دهد. نکتهٔ حائز اهمیت در اینجا، چگونگی تقسیم داده‌های پژوهش به داده‌های آموزش و آزمایش است. اگر چه یک راه حل کلّی در این مورد وجود ندارد، اما فاکتورهایی از قبیل خصوصیات مسئله مورد بحث، نوع داده‌ها و اندازهٔ داده‌های در دسترس بایستی در تصمیم‌گیری مدّ نظر قرار گیرند. با توجّه به این امر، یک مسئله در هنگام استفاده از شبکه‌های عصبی، بکارگیری نسبت بالایی از داده‌های آموزشی نسبت به داده‌های تست است. متناسب با داده‌هایی که بکار گرفته می‌شود، نسبت بالای داده‌های آموزشی به تست، یادگیری معناداری را در پی نخواهد داشت. نسبت‌های بکار گرفته شده در تحقیقات انجام شده قبلی عبارتند از: ۸۰٪-۲۰٪ [۲] – ۷۱٪-۲۹٪ [۱۲] – ۶۸٪-۳۲٪ [۲۰] – ۷۰٪-۳۰٪ [۷] – ۶۹٪-۳۱٪ [۳۸] – ۶۷٪-۳۳٪ [۳۷] و ۶۲٪-۳۸٪ [۶]

برای رفع این مشکل در این پژوهش نسبت‌های مختلف داده‌های آموزش و تست شبکه‌های عصبی مورد آزمایش قرار گرفت و در نهایت نسبتی برای هر شبکه انتخاب شد که بالاترین میزان صحّت را برای شبکه در طی فرایند آموزش و تست فراهم کرد. بر این اساس برای هر شبکه تعداد نه طرح آموزشی پیاده‌سازی شد که در این نه طرح، برای هر شبکه یک نسبت آموزش به تست مورد استفاده قرار گرفت.

۴. رویکرد یادگیری جمعی[38]

به منظور بهبود عملکرد یک مدل دسته‌بندی، ترکیب و بکارگیری چند مدل دسته‌بندی، در حوزهٔ یادگیری ماشینی پیشنهاد شده است[31،43،45،47]. سیستم‌های دسته‌بندی کنندهٔ چندگانه، مبتنی بر ترکیب مجموعه‌ایی از مدل‌های دسته‌بندی است که ترکیب آن‌ها صحّت عملکرد بیش‌تری را نسبت به استفادهٔ منفرد و مستقل از آن‌ها در اختیار کاربر خواهد گذاشت. مدل جمعی دسته‌بندی کننده، مجموعه‌ایی از مدل‌های دسته‌بندی است که تصمیمات دسته‌بندی منفرد آن‌ها با یک روش‌شناسی خاص ترکیب می‌شود[31]. بطور کلّی یک مدل دسته‌بندی منفرد، به عنوان یک عامل با پارامترها و متغیرهای ثابت، نوعی انحراف استنتاجی را خواهد داشت. به منظور کاهش خطای انحراف توسط یک عامل، از چند

37. Bias

38. Ensemble

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

عامل استفاده می‌کنیم. تصمیمات چند عاملی‌ها در نهایت توسط روشی مناسب ترکیب خواهد شد. هدف اصلی روش جمعی، ترکیب مجموعه‌ایی از مدل‌هاست که هر کدام وظیفهٔ مشابهی را انجام می‌دهند تا یک مدل ترکیبی کلّی‌تر و بهتر به دست آید، که هم صحّت بیش‌تری دارد و هم تصمیمات پایدارتری نسبت به مدل‌های فردی ارائه می‌دهد. اساس ترکیب مدل‌های دسته‌بندی شبکهٔ عصبی دستیابی به صحّت و عملکرد بالاتر از یک شبکهٔ مستقل است. عملکرد مدل جمعی وابسته به این است که هر یک از مدل‌های دسته‌بندی بکار گرفته شده در مدل جمعی به خوبی آموزش ببینند. بنابراین تنوع و گوناگونی بین مدل‌های یادگیرنده بایستی حفظ شود. ترکیب مدل‌های دسته‌بندی در صورتی مؤثر و کارا عمل خواهد کرد که این مدل‌ها با یکدیگر متفاوت باشند. این تنوع را می‌توان از راه‌های مختلفی ایجاد کرد، مثل معماری‌های گوناگون، مجموعهٔ داده‌های آموزشی مختلف یا اختصاص تصادفی آن‌ها و یا بر اساس تحلیلات خوشه‌بندی. ایجاد واحد ترکیب کننده و حفظ تنوع میان دسته‌بندی کننده‌ها، دو عامل اصلی برای ساخت مدل جمعی است. نکتهٔ کلّیدی رویکرد یادگیری جمعی مبتنی بر گرفتن نظر خبرگان قبل از تصمیم‌گیری نهایی است. این کار نشأت گرفته شده از فرایند تصمیم‌گیری در جامعه است. اثبات شده است که، گرفتن نظر و ایدهٔ گروهی از خبرگان منجر به تصمیمات بهتری نسبت به تصمیمات یک فرد خبره می‌شود. اگر چه این امر مفهومی ساده می‌نمایاند، اما مزایای حاصل شده از آن توسط محققان بسیاری، مطابق با ادبیات پژوهش، در برنامه‌های خودکار تصمیم‌گیری، به اثبات رسیده است.

۵. ترکیب نتایج

روش‌های اکثریت آراء[۳۹]، رتبه‌بندی[۴۰] و میانگین وزنی[۴۱]، سه روش معمول برای ترکیب نتایج در مدل‌های جمعی است. روش اکثریت آراء پرکاربردترین استراتژی ترکیب نتایج در رویکرد جمعی است که برای مشکلات مربوط به دسته‌بندی انجام می‌گیرد [44]. در تحقیقات مربوط به امتیازدهی اعتباری از رویکردهای اکثریت آراء و میانگین وزنی استفاده شده است. در این پژوهش از روشی جدید استفاده شد که منجر به افزایش صحّت خروجی مدل گردید. در این رویکرد از مجموع درجات عضویت هر خوشه به عنوان وزن عامل مربوط به آن خوشهٔ خاص استفاده شده است، به گونه‌ایی که مجموع درجه‌های عضویت خوشهٔ یک، به عنوان وزن شبکهٔ عصبی یک یا عامل یک در نظر گرفته شده است. در این پژوهش میزان صحّت کلّی هر روش به عنوان معیار برتری آن مطرح خواهد شد. علاوه بر صحّت کلّی، صحّت نوع اول و صحّت نوع دوم نیز برای هر روش محاسبه شده است. علاوه بر میزان صحّت هر یک از روش‌های امتیازدهی اعتباری، به عنوان معیار ارزیابی دیگری که در رابطه با عملکرد مدل‌های دسته‌بندی در تحقیقات مختلف بکار گرفته می‌شود، از معیار محدودهٔ زیر نمودار برای مقایسهٔ عملکرد روش‌های ذکر شده در این پژوهش استفاده شده است. این معیار محدودهٔ زیرگراف خصوصیات عملیاتی دریافتی است. گراف ROC[۴۲] تکنیک مؤثری برای رتبه‌بندی مدل‌ها و بصری‌سازی عملکرد آن‌ها است[16].

معمولاً ROC یک گراف دو بعدی است. که محور عمودی آن میزان حساسیت (صحّت نوع دوم) است و محور افقی آن میزان اختصاصی بودن (صحّت نوع اول) منهای یک است. برای انجام عملیات رتبه‌بندی مدل‌های دسته‌بندی، یک راه معمول و رایج محاسبه محدودهٔ زیر نمودار ROC است که به اختصار محدودهٔ زیر نمودار[۴۳] نامیده می‌شود. به دلیل اینکه محدودهٔ زیر نمودار بخشی از یک مربع واحد است، میزان آن برای هر مدل همیشه بین صفر و یک خواهد بود.

39 .Majority voting
40 .Ranking
41 .Weighted averaging

42 .Area under the Receiver Operating Characteristic(ROC)
43 .Area Under Curve (AUC)

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

نتایج پژوهش

در گام ابتدایی از رویکرد خوشه‌بندی فازی برای تقسیم داده‌های این پژوهش به تعداد خوشه‌های معین استفاده شده است. اساس تقسیم‌بندی خوشه‌ها مبتنی بر بیش‌ترین تشابه در درون خوشه‌ها و کم‌ترین تشابه بین مرکز هر خوشه است. تعداد ۱۰ خوشه از مجموعهٔ داده‌های پژوهش استخراج شد. با توجّه به فازی بودن خوشه‌بندی، ممکن است که یک داده حداقل به یک خوشه و حداکثر به ۱۰ خوشه تعلق داشته باشد.

جدول۱. تعداد اعضای خوشه‌های فازی

خوشه	تعداد اعضا	خوشه	تعداد اعضا
۱	**۳۲۰**	**۶**	**۲۳۶**
۲	**۱۶۷**	**۷**	**۱۸**
۳	**۷۱**	**۸**	**۹۷**
۴	**۴۰**	**۹**	**۱۵۷**
۵	**۱۰۰**	**۱۰**	**۱۹۶**

جدول۲. مشخصات ساختاری عامل‌های پژوهش

صحّت	طرح	نسبت	تعداد گره‌های لایهٔ پنهان	تعداد لایهٔ پنهان	تعداد گره خروجی	تعداد گره ورودی	شبکهٔ عصبی
٪۸۰	**۵**	**% - ۵۰ %۵۰**	**۷،۴**	**۲**	**۱**	**۲۰**	**عامل ۱**
٪۸۲	**۳**	**% - ۳۰ %۷۰**	**۲،۴**	**۲**	**۱**	**۲۰**	**عامل ۲**
٪۸۶	**۷**	**% - ۷۰ %۳۰**	**۱۶،۹**	**۲**	**۱**	**۲۰**	**عامل ۳**
٪۸۱	**۸**	**% - ۸۰ %۲۰**	**۲،۲**	**۲**	**۱**	**۲۰**	**عامل ۴**
٪۸۴	**۹**	**% - ۹۰ %۱۰**	**۳،۸**	**۲**	**۱**	**۲۰**	**عامل ۵**
٪۷۶	**۷**	**% - ۷۰ %۳۰**	**۱۵،۱۰**	**۲**	**۱**	**۲۰**	**عامل ۶**
٪۸۰	**۸**	**% - ۸۰ %۲۰**	**۱۳۹،۱۸**	**۲**	**۱**	**۲۰**	**عامل ۷**
٪۷۰	**۹**	**% - ۹۰ %۱۰**	**۲۷،۶**	**۲**	**۱**	**۲۰**	**عامل ۸**
٪۷۵	**۴**	**% - ۴۰ %۶۰**	**۲،۲**	**۲**	**۱**	**۲۰**	**عامل ۹**
٪۸۱	**۵**	**% - ۵۰ %۵۰**	**۴،۵**	**۲**	**۱**	**۲۰**	**عامل ۱۰**

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

همانطور که ذکر شد برای هر شبکه نسبت‌هایی مختلف از داده‌های آموزش و تست در قالب نه طرح آموزشی بکار گرفته شد. همچنین برای هر شبکه برترین معماری از جنبهٔ میزان صحّت کلّی هر شبکه طراحی شد. نتایج مربوط به نسبت و معماری بهینه برای هر شبکه در جدول۲ گردآوری شده است. معماری و ساختار بهینه طی فرایندهای مختلف آموزش و تست با معماری‌ها و ساختارهای مختلف برای هر شبکه به دست آمده است. سیستم ارزیابی ریسک اعتباری بر اساس ده عامل (شبکهٔ عصبی) با ساختار تعیین شده طراحی شد. موضوع دیگری که در این پژوهش به آن پرداخته شده است، تعیین معیار جداسازی[۴۴] مناسب و بهینه است. در بسیاری از تحقیقات این معیار جداسازی برابر با ۰.۵ در نظر گرفته شده است[23]. همچنین در تحقیقی دیگر مقدار ۰.۶ به عنوان معیار جداسازی شبکه در نظر گرفته شده بود[45]. با توجّه به اینکه خروجی شبکه‌های عصبی عددی بین صفر و یک است، تعیین استراتژی تصمیم‌گیری، می‌تواند در صحّت خروجی شبکه تأثیرگذار باشد. برای رفع این مشکل، در این پژوهش معیارهای متفاوت جداسازی برای خروجی شبکه‌های عصبی در نظر گرفته شد و نتایج مربوط به این معیارها ثبت شد. نه طرح جداسازی برای مدل پیشنهادی و با هر سه روش ترکیب نتایج، پیاده‌سازی شد. نتایج مربوط به نه طرح جداسازی در جدول۳ ارائه شده است. در این جدول سه معیار ارزیابی صحّت نوع اول و دوم و صحّت کلّی محاسبه شده است.

هر سه روش در معیارهای مختلف جداسازی دارای درجهٔ صحّت متفاوتی هستند. در اینجا برای هر سه روش، هر سه معیار صحّت نوع اول، صحّت نوع دوم و صحّت کلّی مورد محاسبه قرار گرفته است. با توجّه به روند درجهٔ صحّت کلّی، قابل ذکر است که بیش‌ترین میزان صحّت برای روش اکثریت آراء در معیار ۰.۳ و ۰.۴ واقع شده است (۷۷.۸۵). برای روش میانگین وزنی برترین میزان صحّت در معیار ۰.۴ و ۰.۵ حاصل شده است (۷۶.۰۷). برای روش درجهٔ عضویت بیش‌ترین میزان صحّت در معیارهای ۰.۳ و ۰.۴ روی داده است (۷۷.۸۵). با بررسی نتایج، برتری روش درجهٔ عضویت در معیارهای مختلف با توجّه به اینکه میانگین صحّت بالاتری نسبت به دو روش دیگر دارد، مشهود است. میانگین این سه روش عبارتند از :

اکثریت آراء : ۷۵.۲۸ ، میانگین وزنی : ۷۴.۳۶ ، درجهٔ عضویت : ۷۵.۳۹

با بررسی روند تغییرات میزان صحّت هر یک از روش‌ها، مشهود است که در معیار جداسازی ۰.۴ هر سه روش دارای برترین عملکرد هستند. همچنین با میل معیار جداسازی به سمت صفر و به سمت یک میزان صحّت کلّی کاهش می‌یابد. در این میان روش اکثریت آراء در معیارهای ۰.۱ و ۰.۲ عملکرد بهتری نسبت به دو روش دیگر داراست. از معیار ۰.۴ الی ۰.۹ برتری نسبی در ایجاد نتایج صحیح‌تر، به روش درجهٔ عضویت اختصاص دارد. علاوه بر این یافته‌ها، نتایج روش‌های دیگر که در ادبیات پژوهش مرور شد، برای مقایسه با نتیجه سیستم پیشنهادی، مورد استفاده قرار گرفت. به منظور دستیابی به یک رویکرد مناسب جهت ترکیب نتایج خروجی مجموعهٔ عامل‌های مدل پیشنهادی این پژوهش، اقدام به تست سیستم با هر سه روش اکثریت آراء، میانگین وزنی و درجهٔ عضویت شده است.

برای مدل پیشنهادی برترین عملکرد در هر سه روش ترکیب نتایج مّد نظر قرار گرفته شده است. علاوه بر این برای مقایسهٔ روش‌های قدیمی تحلیل متمایز کننده و رگرسیون لجستیک استفاده شده است. از روش‌های جدید نیز روش‌های نزدیک‌ترین همسایگی، درخت تصمیم، شبکهٔ RBF، ماشین بردار پشتیبان و شبکهٔ عصبی به صورت منفرد استفاده شده است. نتایج روش‌های مختلف در جدول۴ ارائه شده است. با توجّه به نتایج ارائه شده در جدول۴ ، آشکار است که رویکرد ترکیبی جمعی پیشنهاد شده در این پژوهش بالاترین میزان صحّت کلّی را در بین روش‌های دیگر دارا است. بگینگ، بوستینگ و استکینگ نیز رویکردهایی هستند که در حوزهٔ یادگیری جمعی مطرح می‌شوند.

44 .Cut off

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

جدول۳. نتایج طرح‌های معیار جداسازی

معیار جداسازی	روش ترکیب نتایج								
	اکثریت آراء			میانگین وزنی			درجهٔ عضویت		
	صحّت(%)			صحّت(%)			صحّت(%)		
	نوع اول	*نوع دوم*	*کلّی*	*نوع اول*	*نوع دوم*	*کلّی*	*نوع اول*	*نوع دوم*	*کلّی*
۰.۱	۷۰.۱۱	۸۰.۸۲	۷۷.۵	۷۵.۸۶	۷۴.۶۱	۷۵	۷۵.۸۶	۷۳.۵۷	۷۴.۲۸
۰.۲	۶۳.۲۱	۸۳.۹۳	۷۷.۵	۷۰.۱۱	۷۷.۲	۷۵	۶۸.۹۶	۷۸.۲۳	۷۵.۳۵
۰.۳	۴۴.۸۲	۹۲.۷۴	۷۷.۸۵	۵۶.۳۲	۸۳.۹۳	۷۵.۳۵	۵۱.۷۲	۸۹.۶۳	۷۷.۸۵
۰.۴	۴۱.۳۷	۹۴.۳۰	۷۷.۸۵	۵۴.۰۲	۸۶.۰۱	۷۶.۰۷	۴۸.۲۷	۹۱.۱۹	۷۷.۸۵
۰.۵	۲۶.۴۳	۹۶.۳۷	۷۴.۶۷	۳۹.۰۸	۹۲.۷۴	۷۶.۰۷	۴۳.۶۷	۹۲.۲۲	۷۷.۱۴
۰.۶	۲۲.۹۸	۹۷.۴۰	۷۴.۲۸	۳۲.۱۸	۹۲.۷۴	۷۳.۹۲	۳۷.۹۳	۹۷.۲۲	۷۵.۷۱
۰.۷	۱۷.۲۴	۹۸.۹۶	۷۳.۵۷	۲۲.۹۸	۹۶.۸۹	۷۳.۹۲	۲۶.۴۳	۹۶.۷۳	۷۴.۶۴
۰.۸	۱۲.۶۴	۹۹.۴۸	۷۲.۵	۱۷.۲۴	۹۶.۸۹	۷۲.۱۴	۲۱.۸۳	۹۷.۴۰	۷۳.۹۲
۰.۹	۹.۱۹	۱۰۰	۷۱.۷۸	۱۱.۴۹	۹۸.۹۶	۷۱.۷۸	۱۰.۳۴	۹۹.۴۸	۷۱.۷۸

جدول۴. نتایج روش‌های مختلف امتیازدهی اعتباری بروی داده‌های بانک آلمان

روش	صحّت(%)			AUC	توضیحات
	نوع اول	*نوع دوم*	*کلّی*		
DA	67.49	64.91	65.91	66.2	
Logistic Regression	48.68	84.67	71	66.675	
KNN	12.3	97	72.2	54.65	
Decision tree	49.93	78.55	70.35	64.24	
RBFN individual	39.47	86.29	68.5	62.88	
SVM individual	27.63	97.58	71	62.605	
MLP individual	55.26	85.48	74	70.37	
Bagging DT	48.28	86.34	74.92	67.31	[۴۱]
Bagging NN	48.40	87.20	75.56	67.8	
Bagging SVM	43.42	89.86	75.93	66.64	
Boosting DT	40.89	82.96	72.77	61.925	
Boosting NN	49.20	83.63	73.3	66.415	
Boosting SVM	45.62	89.44	76.3	67.53	
Stacking	45	89.24	75.97	67.12	
Ensemble MLP(Majority vote)	44.82	92.74	77.85	68.78	معیار جداسازی ۰.۳
Ensemble MLP(Weighted averaging)	54.02	86.01	76.07	70.015	معیار جداسازی ۰.۴
Ensemble MLP(Membership degree)	51.72	89.63	77.85	70.675	معیار جداسازی ۰.۳

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

در بسیاری از تحقیقات از رویکردهای جمعی برای مسائل مربوط به دسته‌بندی استفاده شده و نتایج خوبی ارائه شده است. با این وجود، نتایج حاکی از این است که مدل پیشنهادی این پژوهش می‌تواند ص حتّی بالاتر از این رویکردها ارائه دهد. بالاترین صحّت در مدل این پژوهش، در دو روش ترکیب نتایج اکثریت آراء و روش ترکیب نتایج پیشنهادی پژوهش، درجهٔ عضویت، حاصل شده است. همچنین بیش‌ترین صحّت نوع دوم مربوط به روش ماشین بردار پشتیبان است که بیش‌ترین تعداد مشتریان خوب را به درستی دسته‌بندی کرده است. در این میان برترین صحّت نوع اول به روش تحلیل متمایزکننده اختصاص دارد. بطور کلّی، کلّیه روش‌های تست شده در این پژوهش از صحّت نوع اول خوبی برخوردار نیستند. با تحلیل داده‌ها و نتایج روش‌ها می‌توان بیان کرد که دلیل پایین بودن صحّت نوع اول، کم بودن داده‌های مربوط به متقاضیان بد در مجموعهٔ داده‌های واقعی این پژوهش است، بطوریکه تعداد داده‌های مربوط به متقاضیان بد در مجموعهٔ دادهٔ اولیه ۳۰۰ نمونه بوده و مابقی ۷۰۰ نمونه مربوط به متقاضیان خوب بانک است. بنابراین می‌توان نتیجه گرفت که مدل ترکیبی این پژوهش، صحّت بالاتری را نسبت به دیگر روش‌ها دارد. میزان صحّت برتر ذکر شده (۷۷.۸۵)، در دو روش ترکیب نتایج اکثریت آراء و روش درجهٔ عضویت، و در معیارهای جداسازی ۰.۳ و ۰.۴ حاصل شده است، اما این چهار روش دارای محدودهٔ زیر نمودار متفاوت هستند که نمایانگر عملکرد هر یک از این روش‌ها است. برترین محدودهٔ زیر نمودار متعلق به مدل ترکیبی جمعی با روش درجهٔ عضویت و معیار ۰.۳ است. علاوه بر محدودهٔ زیر نمودار، به منظور نمایان ساختن برتری روش پیشنهادی، میزان هزینهٔ تصمیم‌گیری نادرست برای هر چهار روش برتر از لحاظ میزان صحّت کلّی محاسبه شده است. محاسبهٔ این هزینه بر اساس ماتریس هزینه‌ایی است که در جـدول۵ نمایان شده. با توجّه به این ماتـریس،(هزینه[۴۵])

دسته‌بندی درست متقاضی خوب و بد هیچ هزینه‌ایی در بر نخواهد داشت. در صورتیکه مشتری خوب به عنوان مشتری بد معرفی شود، مقدار هزینهٔ ۱ واحد جریمه در نظر گرفته می‌شود و در صورتیکه مشتری بد به عنوان مشتری خوب معرفی شود، میزان هزینهٔ ۵ برای روش بکار گرفته شده منظور می‌شود. اساساً معرفی متقاضی بد به عنوان متقاضی خوب هزینه‌ایی بیش‌تر در بر خواهد داشت و بانک را متحمّل ضرر و زیان بیش‌تری خواهد کرد. با محاسبهٔ میزان هزینهٔ چهار روش برتر این پژوهش نتایج جدول۶ حاصل شد. در جدول۶ مشاهده می‌شود که روش شبکهٔ عصبی چندلایه با روش ترکیب نتایج درجهٔ عضویت و معیار جداسازی ۰.۳ برترین عملکرد را برای بانک خواهد داشت، با توجّه به اینکه بیش‌ترین صحّت، برترین عملکرد و کم‌ترین هزینه را برای بانک فراهم خواهد کرد. بطور موردی می‌توان به نتایج حاصل شده از این پژوهش مطابق با موارد زیر اشاره کرد:

۱. استفاده از روش یادگیری جمعی و روش ترکیبی می‌تواند عملکرد بهتری را نسبت به روش‌های منفرد و مستقل ارائه دهد.
۲. شبکهٔ عصبی، هم در قالب مدل منفرد و هم در قالب مدل جمعی، قابلیت مناسبی را نسبت به تکنیک‌های دیگر برای ارزیابی ریسک اعتباری دارد.
۳. روش درجهٔ عضویت، عملکرد برتری نسبت به روش‌های میانگین وزنی و اکثریت آراء در ترکیب نتایج اعضای مدل جمعی ارائه می‌کند.

نتیجه‌گیری

بانک‌ها با ریسک‌های متفاوتی روبرو هستند، اما خطرناک‌ترین آن‌ها، ریسک اعتباری است، که به معنی از دست دادن بهره و سود بانک به دلیل متعهد نبودن وام گیرنده به قرارداد خود است. در عمل اصلی‌ترین روش ارزیابی ریسک اعتباری، ارزیابی امتیاز اعتباری متقاضی است.

۴۵. این ماتریس در وب سایت داده‌های پژوهش موجود است

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

جدول۵. ماتریس هزینه

وضعیت حقیقی / نتیجه مدل	خوب	بد
خوب	۰	۵
بد	۱	۰

جدول۶. هزینهٔ دسته‌بندی ناصحیح روش‌های برتر پژوهش

هزینهٔ دسته‌بندی ناصحیح	صحّت	معیار جداسازی	روش ترکیب نتایج	روش امتیازدهی اعتباری
۲۵۴	۷۷.۸۵	۰.۳	اکثریت آراء	پرسپترون چندلایه‌ایی جمعی
۲۶۶	۷۷.۸۵	۰.۴	اکثریت آراء	پرسپترون چندلایه‌ایی جمعی
۲۳۰	۷۷.۸۵	۰.۳	درجهٔ عضویت	پرسپترون چندلایه‌ایی جمعی
۲۴۲	۷۷.۸۵	۰.۴	درجهٔ عضویت	پرسپترون چندلایه‌ایی جمعی

روش‌های مختلفی برای امتیازدهی اعتباری تاکنون بکار گرفته شده است. با این وجود یک سیستم پایاتر و دقیق‌تر می‌تواند صحّت پیش‌بینی را افزایش دهد و افزایش صحّت پیش‌بینی حتّی به مقدار کم، می‌تواند بطور چشمگیری هزینه‌های آتی بانک را کاهش دهد. در این پژوهش مدلی ترکیبی مبتنی بر رویکرد جمعی برای امتیازدهی اعتباری در حوزهٔ ریسک اعتباری ارائه شد. این مدل از ترکیب تکنیک‌های دسته‌بندی و خوشه‌بندی حاصل شد. برای خوشه‌بندی از خوشه‌بندی فازی و از شبکه‌های عصبی مصنوعی به عنوان مدل دسته‌بندی استفاده شد. همچنین برای بهبود صحّت و عملکرد مدل از رویکرد یادگیری جمعی بهره گرفته شده است، که این رویکرد یک پارادایم قوی یادگیری ماشینی است که مزیت آشکاری را در بسیاری از حوزه‌ها از خود ارائه داده است. بر اساس رویکرد جمعی از ده شبکهٔ عصبی به عنوان داور و در قالب عامل نرم‌افزاری بهره گرفته شده است. متفاوت با روش‌های ترکیب نتایج بکار گرفته شده در تحقیقات قبلی در حوزهٔ یادگیری جمعی، از جمله اکثریت آراء و میانگین وزنی، در این پژوهش روش ترکیب جدیدی بکار گرفته شد که منجر به بهبود صحّت سیستم ارزیابی ریسک اعتباری در مقایسه با روش‌های ذکر شده گردید. هدف این مدل، دسته‌بندی متقاضیان اعتبار به دو دسته متقاضیان "خوب" و "بد" بر اساس یک سری از معیارها و خصوصیات فرد متقاضی است. شبکه‌های عصبی از طریق مجموعهٔ داده‌های واقعی بانکی در آلمان مورد آموزش و تست قرار گرفتند. معماری شبکه‌های بکار گرفته شده مبتنی بر شبکهٔ عصبی چند لایه‌ایی است. علاوه بر ارائهٔ مدل ترکیبی و ارائهٔ روش جدیدی برای ترکیب نتایج، در این پژوهش به چند مسئله مهم و تأثیرگذار در صحّت و عملکرد سیستم ارزیابی پرداخته شد، که در تحقیقات قبلی مورد توجّه قرار نگرفته است. اولین مسئله، نرمال‌سازی داده‌های ورودی شبکه است. با توجّه به اینکه نرمال‌سازی داده‌ها نقش مهمّی در عملکرد نهایی مدل خواهد داشت، داده‌های ورودی شبکه‌ها مطابق با روشی ساده اما کارا نرمال شد. دومین مسئله انتخاب یک نسبت مناسب برای آموزش و تست شبکه‌های عصبی بکار گرفته شده در این پژوهش است. همانطور که در بخش‌های قبلی ذکر شد در تحقیقات اخیر نرخ‌های متفاوتی برای آموزش و تست شبکه‌های عصبی بکار گرفته شده است. نرخ داده‌های آموزش و تست متناسب با نوع مسئله و داده‌های مسئله می‌تواند متفاوت باشد و بر میزان کارآیی و اثربخشی شبکهٔ عصبی تأثیرگذار است و منجر به درجه‌های صحّت متفاوتی می‌شود. در اینجا نه طرح آموزش و تست برای هر شبکه پیاده‌سازی شد که در نهایت، طرحی برای سیستم انتخاب شد که

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

دارای بالاترین میزان صحّت بود. علاوه بر دو مورد ذکر شده، مسئلهٔ تعیین یک معیار جداسازی مناسب و بهینه، نیز در این پژوهش مطرح شد. در بسیاری از تحقیقات این معیار معادل ۰.۵ در نظر گرفته شده است. اما در این پژوهش اقدام به تست سیستم در نه طرح با معیارهای مختلف جداسازی شد و برای سیستم ارزیابی ریسک اعتباری میزان بهینهٔ این معیار انتخاب شد که بالاترین صحّت را در دسته‌بندی متقاضیان ایجاد کرد. برای اعتبار سنجی مدل پیشنهادی، از چند روش قدیمی و جدید استفاده شد. همچنین خروجی هر یک از عامل‌های سیستم با روش‌های ترکیب نتایج اکثریت آراء و میانگین وزنی، و روش جدید درجهٔ عضویت که در این پژوهش معرفی شد، ترکیب شد تا عملکرد این سه روش نیز با یکدیگر مقایسه شود. معیار ارزیابی روش‌های مختلف، محدودهٔ زیر نمودار و میزان دسته‌بندی درست متقاضیان و یا به عبارتی میزان صحّت کلّی هر یک از روش‌ها است. نتایج نشان داد که مدل پیشنهادی بسیار مؤثر و کارا است و منجر به سیستمی شد که بصورت مناسب قادر به دسته‌بندی ورودی‌ها با صحّت بالا است. در این پژوهش نشان داده شد که استفاده از رویکرد جمعی یا به عبارتی استفاده از چندین مدل دسته‌بندی به جای یک مدل دسته‌بندی، می‌تواند منجر به بهبود عملکرد سیستم دسته‌بندی شود. همچنین در میان روش‌های ذکر شده، برترین صحّت متعلق به مدل ترکیبی مبتنی بر رویکرد جمعی است که با روش درجهٔ عضویت، نظر هر یک از داوران را با یکدیگر ترکیب می‌کند. صحّت این مدل از نتیجهٔ رویکردهای دیگر یادگیری جمعی که در تحقیقات اخیر مورد استفاده قرار گرفته بود، بالاتر بود. با توجّه به نتایج، قابل بیان است که عملکرد شبکه‌های عصبی در حوزهٔ ریسک اعتباری بخصوص به عنوان سیستم غیر خطّی پنهان در مقایسه با روش‌های امتیازدهی قدیمی بهتر است و در میان روش‌های منفرد جدید نیز بهترین میزان صحّت مربوط به این مدل دسته‌بندی با معماری پرسپترون چندلایه‌ایی است. با انجام ارزیابی‌های مختلف از نظر صحّت، عملکرد و هزینهٔ پیش‌بینی نادرست، مشخص شد که مدل ترکیبی با رویکرد جمعی و روش ترکیب نتایج درجهٔ عضویت با معیار جداسازی ۰.۳ بیش‌ترین صحّت، برترین عملکرد و کم‌ترین هزینه را در پیش‌بینی متقاضیان اعتبار از بانک حاصل خواهد کرد.

تحقیقات آتی

با انجام این پژوهش و نتایجی که حاصل شد، زمینه‌هایی برای تحقیقات آتی فراهم آمد. از جمله این زمینه‌ها می‌توان به این موارد اشاره کرد: استفاده از یک پایگاه داده بزرگ‌تر برای سیستم پیشنهادی، استفاده از مدل جمعی ماشین بردار پشتیبان به عنوان مدل دسته‌بندی، بررسی تأثیر افزایش داوران بروی مدل‌های دیگر دسته‌بندی جمعی و همچنین مدل بکار گرفته شده در این پژوهش.

منابع

۱.اداره مطالعات و کنترل ریسک بانک تجارت .۱۳۸۶ . *مدیریت ریسک در بانکداری* .تهران : بانک تجارت با همکاری نشر جهانگیر.

2.Abdou, H., J. Pointon, & A. Elmasry. 2008. Neural nets versus conventional techniques in credit scoring in Egyptian banking. *Expert systems with applications* 35 (3): 1275-1292.

3.Altman, I. E. 1968. Financial ratios, discriminant analysis and the prediction of corporate banlruptcy. *The journal of finance* 23 (4): 589-611.

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

4.Altunbas, Y., L. Gambacorta, & D. M. Ibanez. 2009. Bank risk and monetary policy. *Journal of financial stability* 5 (3): 224-255.

5.Angelini, E., G. D. Tollo, & A. Roil. 2008. A neural network approach for credit risk evaluation. *The quarterly review of economics and finance* 48 (4): 733-755.

6.Atiya, A. F. 2001. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE transactions on neural networks* 12 (4): 929-935.

7.Baesens, B., R. Setiono, C. Mues, & J. Vanthienen. 2003. Using neural network rule extraction and decision tables for credit-risk evaluation. *Management science* 49 (3): 312-329.

8.Baesens, B., T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, & J. Vanthienen. 2003. Benchmarking state-of-art classification algorithm for credit scoring. *Journal of operational research society* 54 (6): 627-635.

9.Bellotti, T. 2010. A simulation study of Basel II expected loss distributions for a portfolio of credit card. *Journal of financial services marketing* 14(4): 268-277.

10.Bensic, M., N. Sarlija, & M. Zekic-Susac. 2005. Modelling small-business credit scoring by using logistic regression, neural networks and decision trees. *Intelligent systems in accounting, finance and management* 13 (3): 133-150.

11.Bologna, G., & R. D. Appel. 2002. A comparison study on protein fold recognition. In Proceeding of *Ninth international conference on neural information processing*, (pp. 2492-2496). Singapore.

12.Boros, E., Hammer, P. L., Ibaraki, T., Kogan, A., Mayoraz, E., & Muchnik, I. (2000). An implementation of logical analysis of data. *IEEE transactions on knowledge and data engineering* 12(2): 292-306.

13.Chen, W., C. Ma, & L. Ma. 2009. Mining the customer credit using hybrid support vector machine technique. *Expert systems with applications* 36(4), 7611-7616.

14.Desai, V. S., J. N. Crook, & G. A. Overstreet. 1996. A comparison of neural networks and linear scoring models in the credit union environment. *European journal of operational rsearch* 95(1): 24-37.

15.Eletter, S. F., S. G. Yaseen, & G. A. Elrefae. 2010. Neuro-based artificial intelligence model for loan decision. *American journal of economics and business administration* 2(1): 27-34.

16.Fawcett, T. 2004. *ROC Graphs: Notes and practical considerations for researchers.* HP laboratories Palo Alto: Intelligent enterprise technologies laboratory.

17.Fierrez-Aguilar, J., L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia, & D. Maltoni. 2005. An on-line signature verification system based on fusion of local and global information. In Proceeding of *audio- and video- based biometric person authentication*, (pp. 523-532). New york.

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

18.Gath, I., & A. B. Geva. 1989. Unsupervised optimal fuzzy clustering. *IEEE transactions on pattern analysis and machine intelligence* 11(7): 773-781.

19.Hoffmann, F., B. Baesens, J. Martens, F. Put, & J. Vanthienen. 2002. Comparing a genetic fuzzy and a neurofuzzy classifier for credit scoring. *International journal of intelligent systems* 17(11): 1067-1083.

20.Hsieh, N. C. 2005. Hybrid mining approach in design of credit scoring model. *Expert systems with applications* 28(4): 655-665.

21.Huang, C. L., M. C. Chen, & C. J. Wang. 2007. Credit scoring with a data mining approach based on support vector machines. *Expert systems with applications* 33(4): 847-856.

22.Jain, A., & A. M. Kumar. 2007. Hybrid neural network models for hydrologic time series forecasting. *Applied soft computing* 7(2): 585-592.

23.Khashman, A. 2010. Neural Networks for Credit Risk Evaluation: Investigation of Different Neural Models and Learning Schemes. *Expert System with Application* 37(9): 6233-6239.

24.Lee, T. S., & I. F. Chen. 2005. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert sysytems with application* 28(4): 743-752.

25.Lee, T. S., C. C. Chiu, C. J. Lu, & I. F. Chen. 2002. Credit scoring using the hybrid neural discriminant technique. *Expert system with application* 23(3): 245-254.

26.Li, X., W. Ying, J. Tuo, B. Li, & W. Liu. 2004. Application of classification trees to consumer credit scoring methods in commercial banks. In Proceeding of *IEEE international conference on systems, Man and Cybernetics* (pp. 4112-4117). Netherlands.

27.Lim, M. K., & S. Y. Sohn. 2007. Cluster-Based dynamic scoring model. *Expert system with applications* 32(2): 427-431.

28.Lin, S. L. 2009. A new two-stage hybrid approach of credit risk in banking industry. *Expert system with application* 36(4): 8333-8341.

29.Malhotra, R., & D. K. Malhotra. 2002. Diferentiating between good credits and bad credits using neuro-fuzzy systems. *European journal of operational research* 136(1): 190-211.

30.Melville, P., & R. J. Mooney. 2005. Creating diversity in ensembles using artificial. *Information fusion: Special issue on diversity in multiclassifier systems* 6(1): 99-111.

31.Nanni, L., & A. Lumini. 2009. An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. *Expert systems with applications* 36(2): 3028-3033.

32.Noh, H. J., T. H. Roh, & I. Han. 2005. Prognostic personal credit risk model considering censored information. *Expert systems with application* 28(4): 753-762.

33.Pang, S., Y. Wang, & Y. Bai. 2002. Credit scoring model based on neural network. In Proceeding of *International Conference on*

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025

Machine Learning and Cybernetics (pp. 1742-1746). Beijing.

34.Piramuthu, S. 1999. Financial credit-risk evaluation with neural and neurofuzzy systems. *European journal of operational research* 112(2): 310-321.

35.Piramuthu, S., M. J. Shaw, & J. A. Gentry. 1994. A classification approach using multi-layer neural networks. *Decision support system* 11(5): 509-525.

36.Qin, R., L. I. Liu, & J. Xie. 2010. An application of improved BP neural network in personal credit scoring. In Proceeding of *The second international conference on computer modeling and simulation (*pp. 238-241). Sanya : IEEE.

37.Setiono, R., B. Baesens, & C. Mues. 2008. Recursive neural network rule extraction for data with mixed attributes. *IEEE transactional on neural networks* 19(2): 299-307.

38.Sustersic, M., D. Mramor, & J. Zupan. 2009. Consumer credit scoring models with limited data. *Expert system with application* 36(3): 4736-4744.

39.Tsai, C.-F., & M. L. Chen. 2010. Credit rating by hybrid machine learning techniques. *Applied soft computing* 10(2): 374-380.

40.Wang, C. M., & Y. F. Huang. 2009. Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data. *Expert systems with applications* 36(3): 5900-5908.

41.Wang, G., J. Hao, J. Ma, & H. Jiang. 2010. A comparative assessment of ensemble learning for credit scoring. *Expert systems with applications . In Press.*

42.West, D. 2000. Neural network credit scoring models. *Computers and operations research* 27(11): 1131-1152.

43.Xie, H., S. Han, X. Shu, X. Yang, X. Qu, & S. Zheng. 2009. Solving credit scoring problem with ensemble learning: A case study. In Proceeding of *The second international symposium on knowledge acquisition and modeling* (pp. 51 - 54). Wuhan : IEEE.

44.Yu, L., S. A. Wang, & K. K. Lai. 2008. Credit risk assessment with a multistage neural network ensemble learning approach. *Expert systems with applications* 34(2): 1434-1444.

45.Yu, L., S. Wang, & K. K. Lai. 2009. An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: the case of credit scoring. *European journal of operational research* 195(3): 942-959.

46.Zhang, D., M. Hifi, Q. Chen, & W. Ye. 2008. A hybrid credit scoring model based on genetic programming and support vector machines. In Proceeding of *The fourth international conference on natural computation* (pp. 8-12). Jinan: IEEE.

47.Zhou, L., & K. K. Lai. 2009. Multi-agent ensemble models based on weighted least square SVM for credit risk assessment. In Proceeding of *Global Congress on Intelligent Systems* (pp. 559 – 563). Xiamen: IEEE.

Downloaded from jor.iranaict.ir at 3:14 +0430 on Saturday July 12th 2025