

# طراحی مدلی جهت پیش‌بینی 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)

محمد صادق حری<sup>۱\*</sup>، کاوه مهدوی<sup>۲</sup>

۱. استادیار، عضو هیئت علمی دانشگاه آزاد واحد اراک، اراک، ایران

۲. کارشناس ارشد مدیریت اجرایی، دانشگاه آزاد اسلامی واحد اراک، اراک، ایران

پذیرش: ۱۳۹۴/۰۲/۱۹

دریافت: ۱۳۹۳/۰۵/۲۰

### چکیده

همواره مؤسسات مالی و اعتباری برای آنکه بتوانند حداکثر سود حاصل از سرمایه گذاری‌های خود را دریافت دارند، بدنبال پالایش، جذب و نگهداشت بهترین سرمایه گذاران، مشاوران، مشتریان و قرض‌گیرندگان بوده‌اند. با این وجود، علوم مختلف سعی نموده‌اند روشهای دقیقی برای تفکیک مشتریان ارایه نمایند. از همین رو علمی مانند روانشناسی تا علوم مدیریت، ریاضیات، مالی و ... درصدد تحقق این هدف برآمده‌اند. آنچه که در این پژوهش بدان اشاره خواهد شد ضرورت استفاده از روش‌های نوین داده‌کاوی در ترکیب با روشهای هوش مصنوعی جهت فائق آمدن بر پیچیدگی‌های مسأله است و پاسخ به این سؤال که آیا روش ترکیبی استفاده شده به خوبی رتبه اعتباری مشتریان را پیش‌بینی می‌کند؟ این امر در حالی رخ می‌دهد که نباید بُعد دیگری از مسأله را که همانا انتخاب مهمترین عوامل سنجش (معیارها) هستند را فراموش نمود و در این راستا از قضاوت خبرگان و تحلیل‌های ناپارامتری (آزاد توزیع) به منظور رتبه‌بندی معیارها استفاده گردیده است. که نهایت با انتخاب تعدادی از شاخص‌ها به منظور پیاده‌سازی مدل ترکیبی به این سؤال پاسخ داده خواهد شد که آیا نظر خبرگان در انتخاب معیارها، منتج به پیش‌بینی مناسبی از وضعیت اعتباری مشتریان می‌گردد. سه شاخص "سن"، "سابقه

ارتباط با بانک (مدت حساب) و " میزان اعتبار" برای پیاده‌سازی مدل ترکیبی عصبی فازی انتخاب گردید. و نتایج بیانگر آن می‌باشد که ۸۹,۶۷ درصد از مواقع این سیستم می‌تواند تخمین درستی نسبت به رتبه اعتباری مشتریان ارائه دهد. برای بهینه‌سازی شبکه عصبی- فازی از الگوریتم کلونی مورچگان استفاده گردید که منتج به بهبود عملکرد مدل تا ۹۰,۵ درصد گردید.

**کلیدواژه‌ها:** ریسک اعتباری، رتبه‌بندی اعتباری، داده‌کاوی، هوش مصنوعی، الگوریتم کلونی مورچگان.

## ۱- مقدمه

در بازار پول و سرمایه نقش اصلی را بانک‌ها و مؤسسات اعتبار ایفا می‌کنند و وظیفه کنترل و مدیریت ریسک اعتباری و تخصیص اعتبارات به مشتریان جزء لاینفک از وظایف این سازمان‌ها می‌باشد. بانک‌ها نیازمند روش و ابزاری هستند تا بوسیله آن رتبه اعتباری و میزان ریسک بازگشت اصل و سود اعتبارات تخصیص داده شده به مشتریان را برآورد نموده و از نرخ نکول وام‌های ارائه شده تا حد امکان بکاهند. بهره‌گیری از سیستمی کارآمد در این حوزه مطلوبیت بالایی را برای ذینفعان بانک‌ها و مؤسسات اعتباری فراهم می‌نماید. علی‌رغم اهمیت این موضوع، در نظام بانکی ایران در زمینه تخصیص تسهیلات اعتباری به مشتریان، روند کارآمد و منظمی به منظور تعیین ریسک اعتباری، امتیازدهی، رتبه‌بندی و همچنین تعیین سطح اعتباری بر اساس شاخص‌هایی مؤثر ملاحظه نمی‌شود و عمدتاً بر اساس ممیزی کارشناسان و کمیته‌های اعتباری صورت می‌پذیرد. در این صورت برخورداری از یک مدل ریسک کارآمد، نه تنها تصمیم‌گیری در زمینه اعطای اعتبار و اخذ وثائق را تسهیل می‌کند بلکه باعث خواهد شد که سیستم بانکی و به دنبال آن کشور از الگوی کارآمدی در تخصیص سرمایه به بخش‌های مختلف اقتصادی برخوردار شود. در مطالعات گذشته از روش‌های آماری مانند تحلیل‌های تبعیضی و رگرسیون زبانی، طبقه‌بندی و درخت رگرسیون، رگرسیون لاجیت و پروبیت و روش تحلیل ممیزی برای امتیازدهی و رتبه‌بندی مشتریان استفاده می‌گردید، لیکن اخیراً با توسعه مدل‌های مبتنی بر هوش مصنوعی، منطق فازی و روش‌های فرا ابتکاری از یک سو و توسعه قابلیت‌های نرم‌افزاری و سخت افزای از سوی دیگر مطالعات بسیاری در کاربرد این روشها در مدل‌های امتیازدهی و رتبه‌بندی اعتباری صورت گرفته است. ساختار مقاله حاضر



در ادامه به این صورت است که ابتدا به مروری بر تاریخچه و ادبیات موضوع پرداخته و سپس جامعه آماری و نمونه آماری و متغیرهای مدل مورد بررسی قرار می‌گیرند. در بخش بعدی شاخص‌های اصلی رتبه‌بندی اعتباری مشتریان و نحوه انتخاب مؤثرترین آنها بیان می‌گردد. سپس مفاهیم اصلی سیستم استنتاج فازی عصبی انطباقی تشریح می‌گردد. در بخش بعد به معرفی معیارهای اصلی ارزیابی عملکرد مدل پرداخته می‌شود. روش بهینه‌سازی مدل بوسیله الگوریتم کلونی مورچگان بیان می‌شود. در انتها کاربرد این مدل در رتبه‌بندی اعتباری مشتریان بانک مورد بررسی قرار گرفته و کارایی آن در مورد تشخیص صحیح ارزیابی می‌گردد.

## ۲- مروری بر ادبیات رتبه‌بندی اعتباری

در ارزیابی ریسک اعتباری، درجه‌بندی اعتباری یکی از تکنیک‌های تحلیلی مهم است. درجه‌بندی اعتباری، تکنیکی است که به برخی سازمان‌ها مثل بانک‌های تجاری و شرکت‌های کارت اعتباری در تعیین اینکه آیا اعتبار به مشتریان بدهند یا نه کمک می‌کند. این کار بر اساس مجموعه‌ای از معیارهای تعریف شده انجام می‌شود. معمولاً یک درجه‌ی اعتباری عددی است که میزان اعتبار یک شخص را بر اساس تحلیل‌های کمی تاریخچه‌ی اعتباری و دیگر معیارها تعیین می‌کند و میزان تمایل فرد برای پرداخت قرض یا بدهی خود را تخمین می‌زند. برای اعطای تسهیلات باید درجه اعتبار و قدرت بازپرداخت اصل و سود مبلغ اعتبار را برای دریافت‌کننده تسهیلات تعیین کرد. شانس اینکه وام گیرنده نتواند وام را بازپرداخت کند، ریسک اعتباری یا ریسک عدم بازپرداخت شناخته می‌شود. [۱] در کنار موضوع ریسک اعتباری، مدیریت آن ریسک به امر قابل توجهی تبدیل می‌شود. مدیریت ریسک اعتباری فرآیندی است که در طول آن زیان‌های ناشی از نکول وام طی ادوار تجارت به صورت کمی محاسبه می‌شود و شامل ۳ مرحله، محاسبه احتمال نکول وام‌های مشتریان با استفاده از مدل‌های کمی، محاسبه زیان‌های ناشی از نکول با استفاده از مدل‌های آماری و ارزشیابی مدل با استفاده از داده‌های دوران‌های گذشته می‌باشد. بر اساس نظر کمیته بال سویس [۲] هدف مدیریت ریسک اعتباری بهینه کردن بازده تعدیل شده بانک بر اساس ریسک اعتباری و همچنین کنترل خطر ریسک اعتباری از طریق پارامترهای قابل قبول است.

اعتبارسنجی به مفهوم ارزیابی و سنجش توان بازپرداخت متقاضیان اعتبار و تسهیلات مالی

و احتمال عدم بازپرداخت اعتبارات دریافتی از سوی آنها می‌باشد. امروزه به منظور اعتبارسنجی مشتریان نظام‌هایی نظیر «امتیازدهی اعتباری»<sup>۱</sup> و «رتبه‌بندی مشتریان اعتباری»<sup>۲</sup> تدوین و توسعه یافته‌اند. در کنار روش امتیازدهی اعتباری روش رتبه‌بندی اعتباری قرار دارد. رتبه‌بندی اعتباری در واقع روشی برای شناسایی و موافقت با اعطای وام به متقاضیان باریسک پایین و اجتناب از اعطای وام به متقاضیان با ریسک بالا از طریق طبقه‌بندی آنها می‌باشد. طبقات رتبه‌بندی با نمادهای مختلفی مثل Aaa یا AAA (برای عالی‌ترین کیفیت) یا با اعداد از ۱ تا ۱۰ مشخص می‌شوند. با داشتن امتیازها امکان اندازه‌گیری ریسک اعتباری برای بانکها فراهم می‌شود و می‌توانند آنرا متناسب با پورتفوی اعتباری اداره نمایند. بدین مفهوم که اکسپوژر<sup>۳</sup> (ریسک قابل مشاهده و محاسبه) بانک را در رابطه با انواع ریسک تعدیل و اصلاح نمایند. این مطلب را می‌توان با رابطه زیر توضیح داد:

$$\text{ریسک} = (\text{احتمال عدم بازپرداخت}) \times (\text{نرخ زیان در صورت عدم بازپرداخت}) \times (\text{مبلغ در معرض بازپرداخت})$$

معیارهای امتیازدهی اعتباری، مؤسسه‌های اعتباری و بانکها میبایست با توجه به پیچیدگی فعالیتها و محیط اقتصادی، مدل‌های مناسب جهت ارزیابی امتیازدهی اعتباری مشتریان انتخاب نمایند. در خصوص مشتریان اعتباری حقوقی بزرگ، بانکها از معیار C، برای تصمیم‌گیری استفاده می‌کنند. البته معیار C می‌تواند با تعدیلاتی برای کلیه مشتریان به کار رود؛ ولی برای مشتریان کوچک و متوسط شاخص‌های خاصی در مدل گنجانده می‌شود. [۳، ص ۳۰] این معیار به طور خلاصه عبارتست از - شخصیت، ظرفیت، سرمایه، وثیقه، شرایط.

معیار (LAPP)؛ که شامل: نقدینگی، فعالیت، سودآوری، پتانسیل که برای مشتریان حقوقی و سازمان‌ها و کسب کارها مورد استفاده قرار می‌گیرد. [۳، ص ۳۳] معیار P: اجزای این معیار به طور خلاصه عبارتند از: مردم، محصول، حمایت، پرداختها، شمای کلی آینده. با انجام این بررسی‌ها و ارزشیابی‌ها، تصمیم‌گیرندگان اعتباری قادر خواهند بود تا در خصوص اعطای تسهیلات و چگونگی سقف اعتبار و شیوه‌های کنترل آن و نیز مدت و شیوه بازپرداخت، اظهار نظر نمایند. [۳، ص ۳۶]

از دلایل اهمیت سنجش ریسک اعتباری می‌توان به موارد زیر اشاره نمود:  
الف) اکنون مهمترین عامل ورشکستگی بانکها ریسک اعتباری است. اگر مشتری به موقع



تعهدات خود را بازپرداخت نکنند، این تسهیلات به صورت مطالبات معوق بانکی در می‌آید و این امر موجب اختلال در توزیع اعتبارات بانکی و در نتیجه اختلال در اقتصاد کشور می‌شود. ب) اندازه‌گیری ریسک اعتباری با پیش‌بینی زیانهای عدم بازپرداخت اعتبارات و ایجاد رابطه منطقی بین ریسک و بازده، امکان بهینه‌سازی ترکیب پرتفوی اعتباری، قیمت‌گذاری دارایی‌ها و تعیین سرمایه اقتصادی بانکها را به منظور کاهش هزینه‌های سرمایه‌ای و حفظ توان رقابتی فراهم و نوعی مزیت نسبی برای بانکها و مؤسسات اعتباری ایجاد می‌نماید. ج) در ایران از یک طرف فعالیت بانکها براساس قانون بانکداری بدون ربا و مبتنی بر عقود اسلامی است؛ بنابراین نمیتوان بین بازار پول و سرمایه مرزی قائل شد و از طرف دیگر با توجه به ساختار اقتصادی کشور، عملیات بازار سرمایه (بازار اوراق بهادار و سهام) و سایر شبکه‌های غیربانکی، پیشرفت قابل ملاحظه‌ای نداشته و از این رو سهم قابل توجهی از سرمایه‌گذاری از طریق بازار بانکی انجام می‌گیرد. بنابراین موفقیت بانکها در انجام این امور از اهمیت ویژه‌ای برخوردار است. د) در نظام ربوی پس از پرداخت وام، ارتباط بانک با پول قطع می‌شود و بانک بدون توجه به نوع فعالیت اقتصادی، اصل و فرع پول خود را مطالبه می‌نماید؛ بنابراین با گرفتن ضمانت کافی، لزومی به ارزیابی دقیق از مشتری وجود ندارد (و در صورتیکه ارزیابی انجام شود، در راستای تسهیل مبادلات و انتخاب مشتریان بهتر است) حال آنکه در سیستم بانکداری اسلامی بانک شریک گیرنده تسهیلات در فعالیتهای اقتصادی می‌باشد و به طور عمده سهم آورده فرد به عنوان ضمانت در نظر گرفته می‌شود. بنابراین با توجه به منابع مالکیتی - وکالتی ارزیابی توان بازپرداخت مشتری بسیار اهمیت دارد. [ ۳، ص ۲۱].

### ۳- روش‌های رتبه‌بندی و امتیازدهی اعتباری

روشهای آماری: عموماً تحلیل‌های تبعیضی و رگرسیون زبانی دو روش داده‌کاوی معمول برای ساخت مدل‌های امتیازدهی اعتباری هستند. با این حال، تحلیل تبعیضی خطی<sup>۵</sup> به دلیل فرضیه‌ی آن در مورد طبیعت گروهی داده‌ها و این واقعیت که ماتریس‌های کواریانس کلاس‌های مختلف مساوی نیستند مورد انتقاد واقع شده است.

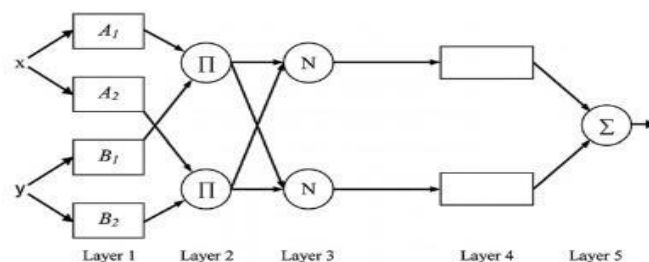
با توجه به مشکلات تحلیل تبعیضی خطی روش‌های رگرسیون منطقی و شبکه‌های عصبی و امتیازدهی اعتباری با استفاده از تکنیک‌های رایج داده‌کاوی، طبقه‌بندی و درخت رگرسیون و

اسپلاین‌های رگرسیون انطباقی چند متغیره ارایه شدند. دلیل استفاده از طبقه‌بندی و درخت رگرسیون و اسپلاین‌های رگرسیون انطباقی چند متغره در امتیازدهی اعتباری، چهار وجهی است. اول اینکه روش‌های طبقه‌بندی و درخت رگرسیون<sup>۶</sup> و اسپلاین‌های رگرسیون انطباقی چند متغره<sup>۷</sup> برخلاف تحلیل تبعیضی خطی و رگرسیون منطقی، قابلیت مدلسازی روابط پیچیده بین متغیرها بدون فرضیات قوی مدل را دارند. به علاوه، برخلاف شبکه‌های عصبی، هر دو قادر به شناسایی متغیرهای مستقل مهم در ساخت درخت و توابع اساسی در مواقعی که متغیرهای زیادی لحاظ شوند هستند. سوم اینکه طبقه‌بندی و درخت رگرسیون و اسپلاین‌های رگرسیون انطباقی چند متغره نیاز به فرآیند طولانی یادگیری ندارند و بنابراین وقتی مجموعه داده‌ها بزرگ است می‌توانند در زمان مدلسازی صرفه‌جویی کنند. در نهایت، مزیت بزرگ طبقه‌بندی و درخت رگرسیون و اسپلاین‌های رگرسیون انطباقی چند متغره نسبت به دیگر روش‌های طبقه‌بندی این است که مدل طبقه‌بندی نتیجه به آسانی تفسیر می‌شود. روش‌های داده‌کاوی، داده کاوی که برخی اوقات به اکتشاف دانش از پایگاه دانش گفته می‌شود، یک روش سیستماتیک برای یافتن الگوها، گرایش و روابط زیرین در داده‌ها است. داده کاوی توجه بسیار از محققان را به دلیل کاربردهای گسترده در تصمیمات تجاری مهم به خود جلب کرده است. اصولاً، محققان داده‌کاوی را به دو دسته تقسیم می‌کنند: متدولوژی‌ها و تکنولوژی‌ها. متدولوژی‌ها از بصری سازی داده‌ها، یادگیری ماشین، روش‌های آماری و پایگاه‌داده‌ی استنتاجی تشکیل شده است. کاربردهای مربوطه با استفاده از این متدولوژی‌ها می‌توانند به صورت طبقه‌بندی، پیش‌بینی، خوشه‌بندی، خلاصه‌سازی، مدلسازی وابستگی، تحلیل ارتباط و تحلیل دنباله‌ای خلاصه شوند. بخش تکنولوژی داده‌کاوی از تکنیک‌هایی مثل روش‌های آماری، شبکه‌های عصبی، درخت‌های تصمیم‌گیری، الگوریتم‌های ژنتیک و روش‌های غیر پارامتری تشکیل شده است. شبکه‌های عصبی مصنوعی<sup>۸</sup>، شبکه‌های عصبی مصنوعی که توسط شبیه‌سازی اصول شبکه‌ی مغز انسانی ارائه شده، یک ابزار مدلسازی غیرخطی منعطف است. توانایی یادگیری انسان با شبکه‌های عصبی مصنوعی به یک محیط کامپیوتری منتقل شده است. به عبارت دیگر، شبکه‌های عصبی مصنوعی توانایی یادگیری از نمونه‌ها را دارد. شبکه‌های عصبی مصنوعی از تعدادی اجزای پردازشی تشکیل شده است که در چارچوب قوانینی که گره یا نرون نامیده می‌شوند گردهم آمده‌اند. از سال ۱۹۹۰، شبکه‌های عصبی مصنوعی به طور



گسترده در مطالعات پیش‌بینی مالی استفاده می‌شود. اکثر این مطالعات نشان داده‌اند که دقت پیش‌بینی شبکه‌های عصبی مصنوعی بهتر از تکنیک‌های آماری سنتی است. با وجود اینکه از شبکه‌های عصبی مصنوعی می‌توان در بسیاری از زمینه‌ها استفاده کرد، معایبی نیز دارد. شبکه‌های عصبی مصنوعی نیاز به یک فرآیند طولانی یادگیری در مدل بهینه دارد. شبکه‌های عصبی مصنوعی بدلیل کمبود تئوری نیز مورد انتقاد قرار گرفته است. فرصتی برای توضیح نتیجه‌ی بدست آمده از شبکه‌های عصبی مصنوعی نیست، به عبارت دیگر، مدل مثل یک جعبه‌ی سیاه کار می‌کند. مدل‌های ترکیبی از حداقل دو تکنولوژی هوش مصنوعی مثل فازی-عصبی استفاده می‌کند و می‌تواند این معایب را حذف و نتایج بهتری تولید کند. سیستم‌های منطق فازی عصبی<sup>۹</sup> بسیار شبیه به استدلال انسانی است. اما این تکنولوژی‌ها معایب خاصی هم دارند. شبکه عصبی مصنوعی قادر به توضیح چگونگی تولید راه‌حل نیست. به عبارت دیگر اطلاعات نگه داری شده در جعبه‌ی سیاه یک اشکال مهم برای شبکه عصبی مصنوعی است. اشکال منطق فازی عدم توانایی در یادگیری است. پس از ترکیبی از این تکنولوژی‌ها استفاده می‌شود. بنابراین معایب این تکنولوژی‌ها با ترکیب آن‌ها از بین می‌رود. برای بهره‌گیری از مزایای شبکه عصبی مصنوعی و مدلسازی منطق فازی، از این تکنولوژی‌ها به صورت همزمان استفاده می‌شود. اخیراً سیستم‌های فازی عصبی توجه زیادی را به خود جلب کرده است زیرا مزیت هر دو روش شبکه عصبی مصنوعی و منطق فازی را دارد. سیستم‌های فازی عصبی توانایی اعمال تجربیات انسانی در مسائل را با استفاده از قوانین فازی دارند. مزیت اصلی مدل‌های فازی عصبی نسبت به تکنولوژی‌های غیر خطی هوش مصنوعی توانایی آن‌ها در یادگیری توسط متغیرهای زبانی است. مزیت دیگر فازی عصبی توانایی اظهار نظر در مورد چگونگی تولید مقادیر خروجی توسط مدل است. به عبارت دیگر، مدل فازی عصبی یک جعبه‌ی سیاه باقی نمی‌ماند. فازی عصبی توسط محققانی برای پیش‌بینی مالی استفاده شده است. سیستم استنتاج فازی عصبی انطباقی<sup>۱۰</sup>، که یک نوع سیستم عصبی فازی است که توسط جانگ در سال ۱۹۹۳ ارائه شد. سیستم استنتاج فازی عصبی انطباقی از تخصص انسانی به شکل فازی و قوانین اگر-آنگاه استفاده می‌کند. سیستم استنتاج فازی عصبی انطباقی توانایی ساخت مدل‌هایی با داده‌های نمونه و تحمل خطا دارد. به عبارت دیگر، سیستم استنتاج فازی عصبی انطباقی پارامترهای مناسب خود را توسط توابع عضویت و ورود داده به سیستم برای یادگیری بهتر تعیین می‌کند.

مهم‌ترین ویژگی که سیستم استنتاج فازی عصبی انطباقی را از شبکه عصبی هوشمند متمایز می‌کند مدلی است که با قوانینی مثل اگر... آنگاه ... ارائه می‌شود. برای ارائه‌ی دو مثال از قانون اگر-آنگاه قوانین زیر را ببینید:



شکل ۲ سیستم استنتاج فازی عصبی انطباقی

لایه‌ی ۱: هر گره‌ی  $i$  در این لایه، یک گره‌ی انطباقی با یک تابع گره توصیف شده توسط عبارات زیر است.

$$O_{1,i} = \mu_{A_i}(x), \text{ For } i = 1, 2$$

$$O_{1,i} = \mu_{B_i}(y), \text{ For } i = 3, 4$$

که  $x$  گره‌ی ورودی  $i$  و  $A_i$  و  $B_i$  برچسب‌های زبانی مرتبط با تابع این گره،  $O_{1,i}$  و  $O_{1,i-2}$  توابع عضویت  $A_i$  و  $B_i$  هستند. این تحقیق از  $\mu_{A_i}(x)$  و  $\mu_{B_i}(y)$  به عنوان توابع عضویت با حداقل و حداکثر ۰ و ۱ بصورت زیر استفاده می‌کنند: (که  $a_i$  و  $b_i$  و  $c_i$  پارامترهای منطقی تابع عضویت هستند).

$$A_i(x) = 1 / (1 + [(x - c_i/a_i)^2]^{b_i})$$

لایه‌ی ۲: هر گره در این لایه یک گره‌ی ثابت برچسب دار  $\Pi$  است که سیگنال‌های ورودی را ضرب و نتیجه را به عنوان خروجی می‌دهد. خروجی این لایه نشان‌دهنده‌ی قدرت درست بودن قانون است که به صورت زیر نشان داده می‌شود:

$$W_{i,1} = A_i(x) \times B_i(y), \quad i = 1, 2$$

لایه‌ی ۳: هر گره در این لایه، یک گره‌ی ثابت است. هر گره با  $N$  برچسب خورده است.  $i$





امین گره میزان قدرت قوانین نام به جمع قدرت دیگر قوانین را مشخص می‌کند. به عبارت دیگر، این لایه قدرت گرهی  $\bar{A}$  را نرمال می‌کند.

$$۰.۳, i = \frac{W_i}{W_1 + W_2}, i = 1, 2$$

لایه ۴: این لایه نتیجه را محاسبه می‌کند. هر گره در این لایه یک گرهی انطباقی با یک تابع گره است:

$$۰.۴, i = W(P_i.x + Q_i.Y + R_i)$$

لایه ۵: گرهی واحد در این لایه، یک گرهی ثابت با برچسب  $\Sigma$  است که خروجی کلی را به عنوان جمع همه‌ی سیگنال‌های ورودی محاسبه می‌کند.

$$۰.۵, i = \frac{\sum_i W_i.F_i}{\sum_i W_i}$$

#### ۴- پیشینه تحقیق

همزمان با رشد میزان اعتبار بخش مالی، بسیاری از مدل‌های امتیازدهی اعتباری نیز توسط بانکها و محققان ارائه شده تا کاربردهای اعتباری خود شامل تحلیل خطی جدا کننده، تحلیل منطقی رگرسیون، اسپلاین رگرسیون انطباقی چند متغیره، درخت رگرسیون و طبقه‌بندی، شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان<sup>۱۱</sup> و الگوریتم ژنتیک<sup>۱۲</sup> را ارزیابی کنند. هر دو روش تحلیل خطی جدا کننده و تحلیل منطقی رگرسیون به طور گسترده در امتیازدهی اعتباری استفاده می‌شوند. با این حال، تحلیل خطی جدا کننده فرضیاتی در نظر می‌گیرد: رابطه‌ی خطی بین متغیرهای مستقل و توزیع نرمال متغیرها. از تحلیل خطی جدا کننده انتقاد شده است زیرا قادر به فراهم آوردن تطابق بین این فرضیه‌ها نیست تحلیل منطقی رگرسیون برای پیش‌بینی روی یک مجموعه داده با نتایج دو بخشی استفاده می‌شود. تحلیل منطقی رگرسیون بر خلاف تحلیل خطی جدا کننده، نیازی به فرضیه‌های نرمالیده ندارد. اما هر دو مدل فرض می‌کنند که یک رابطه‌ی خطی بین متغیرها وجود دارد بنابراین ممکن است همه‌ی این مدل‌ها دقت کافی در امتیازدهی اعتباری را نداشته باشند. کارایی پنج مدل مختلف امتیازدهی اعتباری را با تحلیل خطی جدا کننده، تحلیل منطقی رگرسیون،  $k$  نزدیک‌ترین همسایه و تخمین چگالی کرنل و درخت

رگرسیون و طبقه‌بندی مقایسه کرد. وست بیان کرد که مدل‌های شبکه عصبی مصنوعی به طور موفقیت‌آمیزی در امتیازدهی اعتباری قابل استفاده‌اند و شبکه عصبی مصنوعی می‌تواند جایگزینی برای تحلیل منطقی رگرسیون باشد. [۴] لی و همکاران یک مدل هیبرید امتیازدهی اعتباری ارائه دادند که شبکه عصبی مصنوعی و تحلیل خطی جدا کننده را با هم ترکیب می‌کند. کارایی مدل ترکیبی بسیار موفقیت‌آمیزتر از تحلیل خطی جدا کننده، تحلیل منطقی رگرسیون با شبکه عصبی مصنوعی است. [۵] در تحقیقی دیگر از الگوریتم ژنتیک در فرآیند انتقال سه کاربرد اعتباری رد شده در گروه پذیرش شرطی استفاده کردند و دریافتند که مدل‌هایی که از اسپلاین رگرسیون انطباقی چند متغیره و درخت رگرسیون و طبقه‌بندی روی داده‌های اعتباری کار می‌کنند موفق‌تر از آنهایی هستند که از تحلیل خطی جدا کننده، تحلیل منطقی رگرسیون و شبکه عصبی مصنوعی استفاده می‌کنند. در مطالعه دیگر، ثابت گردید که مدل‌های طبقه‌بند- شبکه عصبی مصنوعی میزان خطای یک طبقه‌بند واحد را به ۳٪ یا ۵٪ کاهش می‌دهند.

به دلیل اهمیت ارزیابی ریسک اعتباری، یک جریان تحقیقی وسیع روی بررسی ریسک اعتباری انجام شده است. ابتدا، بسیاری از تحلیل‌های آماری و روش‌های بهینه‌سازی مثل تحلیل تبعیضی خطی، تحلیل منطقی، تحلیل قیاسی، برنامه‌نویسی خطی، برنامه‌نویسی صحیح،  $k$  نزدیک‌ترین همسایه و درخت طبقه‌بندی به طور گسترده در ارزیابی ریسک اعتباری و مدلسازی وظایف استفاده شده‌اند. با وجود اینکه از این روش‌ها برای ارزیابی ریسک اعتباری استفاده می‌شود، توانایی تمایز مشتریان خوب از مشتریان بد هنوز یک مشکل است و روش‌های موجود محدودیت‌های ذاتی دارند که در آینده می‌تواند بهبود داده شود. مطالعات اخیر نشان داده است که تکنیک‌های هوش مصنوعی مثل شبکه‌های عصبی مصنوعی، محاسبات تکاملی و الگوریتم ژنتیک و ماشین بردار پشتیبان دارای مزایایی بر تحلیل‌های آماری و روش‌های بهینه‌سازی برای ارزیابی ریسک اعتباری در نتایج عملی خود هستند. سبب اعتباری می‌تواند به دو خوشه از اعتبارات تقسیم شده و هر خوشه نیز به تقسیم دو نظام قصور سنگین و قصور سبک بپردازد. برای تعیین ورشکستگی شرکت‌ها، مدل رگرسیون لجستیک به کار برده شد، از این مدل برای اندازه‌گیری ریسک اعتباری اوراق قرضه منتشر شده شرکت‌ها استفاده شد. [۶].

در تحقیقی دیگر، یک متدولوژی امتیازدهی اعتباری براساس تحلیل پوششی داده‌ها پیشنهاد

گردید. آن‌ها داده‌های مالی جاری ۸۲ شرکت تولیدی/صنعتی را که تشکیل دهنده پرتفولیوی اعتباری یکی از بزرگ‌ترین بانک‌های ترکیه بود، برای رتبه‌بندی اعتباری به کار گرفتند. در این پژوهش براساس ادبیات موضوع، ۴۲ نسبت مالی انتخاب شد و از میان آن‌ها ۶ نسبت مهم مالی مورد توجه قرار گرفت. امل و همکارانش پس از اعتبارسنجی مدل با تجزیه و تحلیل رگرسیون دریافته‌اند که روش تحلیل پوششی داده‌ها قادر به تخمین رتبه‌های اعتباری شرکت‌ها بوده و از کارایی لازم برای امتیازدهی اعتباری برخوردار است. [۷] در تحقیقی با عنوان "رویکرد چند گزینه‌ای به رتبه‌بندی اعتباری با به‌کارگیری روش تحلیل پوششی داده‌ها، ارزیابی وام‌گیرندگان با در نظر گرفتن پروژه‌های مالی خصوصی" یک رویکرد چند گزینه‌ای به رتبه‌بندی اعتباری را به وسیله تحلیل پوششی داده‌ها به منظور ارزیابی وام‌گیرندگان برای پروژه‌های مالی خصوصی، پیشنهاد می‌دهند. در این پژوهش تکنیک‌های مختلف رتبه‌بندی اعتباری نظیر تجزیه و تحلیل ممیزی، تحلیل پوششی داده‌ها -تصمیم، شبکه‌های عصبی و ... مقایسه شده‌اند. [۸] دزای و همکاران در سال ۱۹۹۷ به مقایسه شبکه‌های عصبی، تحلیل تمایز خطی و رگرسیون لجستیک پرداختند. آنها در مطالعاتشان به این نتیجه رسیدند که در طبقه‌بندی متقاضیان وام به مشتریان خوش اعتبار و بد اعتبار شبکه‌های عصبی از تحلیل تمایز خطی بهتر عمل می‌کند و نسبتاً عملکرد مشابهی نسبت به رگرسیون لجستیک دارند [۹] چنگ و لی در سال ۱۹۹۹ در یک مدل ترکیبی سیستم استدلال فازی و شبکه‌های عصبی مصنوعی را به منظور انجام یک تحلیل رگرسیون فازی ترکیب نمودند. [۱۰] وست در سال ۲۰۰۰ در تحقیقاتی که داشت به بررسی درستی پنج مدل شبکه عصبی LVQ، RBF، MOE، MP و استدلال تطبیقی فازی در انجام فرایند امتیازدهی اعتباری پرداخت. نتایج این بررسی‌ها با روش‌های آماری سنتی مانند رگرسیون لجستیک خطی و K نزدیکترین همسایه و تحلیل تمایز خطی و درخت تصمیم در امتیازدهی اعتباری بنچ مارک شد. [۱۱] بوسابین و وانوس نیز به برتری روش‌های عصبی - فازی نسبت به روش‌های سنتی موجود، در پیش‌بینی ورشکستگی تجاری شرکت‌ها اشاره می‌کنند. آنها بر این باورند که این روش می‌تواند در تصمیم‌گیری کلیه ذینفعان شرکت‌ها بسیار کارا باشد. [۱۲] کاستیلو و ملین، از یک سیستم خبره با ترکیب شبکه عصبی مصنوعی و منطق فازی به منظور پیش‌بینی قیمت استفاده کردند. آنها مدل‌های مددانی و سوگنو را با یکدیگر مقایسه و کارایی بهتر سیستم‌های استدلال سوگنو را گزارش کردند.

همچنین ثابت نمودند که سیستم هیبرید ساخته شده از مدل‌های آماری مرسوم کارایی بهتری داشته و آن را برای پیش‌بینی نرخ مبادلات ارزی (دلارایسو) پیشنهاد نمودند. [۱۳] در سال ۲۰۰۶ ترینکل در پایان نامه دکترای خود به مقایسه قدرت شبکه عصبی مصنوعی و مدل‌های آماری سنتی در امتیازدهی اعتباری پرداخت. وی دو فرضیه داشت: اول اینکه قدرت طبقه‌بندی شبکه‌های عصبی از روش‌های آماری سنتی بیشتر است و دوم اینکه با تغییر تکنیک‌های تفسیر وزن‌های شبکه عصبی، امتیازدهی اعتبار نتایج متفاوتی خواهد داشت. نتایج پژوهش وی دو فرضیه فوق را اثبات نمود. [۱۴] در سال ۲۰۰۶ بئل و همکاران از شبکه عصبی برای سنجش اعتبار استفاده نمودند. ایشان در این مطالعات با استفاده از داده‌های آژانس‌های اعتباری و کشورهای مختلف از سال ۱۹۸۹ تا سال ۱۹۹۹ به این نتیجه رسیدند که شبکه عصبی مصنوعی نسبت به دیگر روش‌ها دارای عملکرد و کارایی بهتری است. [۱۵] در سال ۲۰۰۷ نیز تسای و وو از شبکه‌های عصبی به طور همزمان برای پیش‌بینی ورشکستگی بانک‌ها و امتیازدهی اعتبار استفاده نمودند. [۱۶] در سال ۲۰۰۲ مالهورتا و مالهورتا مدل شبکه عصبی -فازی آنف یس را برای تشخیص مشتریانی که باید به آنان وام تعلق بگیرد به کار بردند و نتایج آن را با رگرسیون تمایزی خطی مقایسه کردند. برای این ارزیابی و مقایسه آنها داده‌های خود را از ۹ مؤسسه اعتباری گردآوری کردند و از هفت جفت مختلف نمونه‌های آزمایشی و آموزشی‌ای که از این مؤسسه‌ها جمع‌آوری شده بود بهره بردند. آنها از ۵۰۰ مشاهده برای آموزش و ۲۹۰ مشاهده برای آزمایش دو مدل خود استفاده نمودند. نتایج آنان نشان داد که عملکرد مدل شبکه عصبی - فازی آنها بسیار بهتر از مدل دیگر است [۱۷] هوآنگ و همکاران در سال ۲۰۰۴ تکنیک‌های ژنتیک الگوریتم و شبکه‌های عصبی را برای حل مشکلات تحلیل اعتبار به کار بردند [۱۸] همچنین لی و چن در سال ۲۰۰۵ یک فرایند مدلینگ هیبریدی دو مرحله‌ای را با استفاده از شبکه‌های عصبی و اسپلینهای رگرسیون انطباقی چند متغیره برای امتیازدهی اعتبار پیشنهاد داده‌اند. آنها در ابتدا از اسپلینهای رگرسیون انطباقی چند متغیره برای ساخت مدل امتیازدهی اعتبار استفاده نمودند و سپس متغیرهای بدست آمده را به عنوان ورودی وارد سیستم خبره نمودند. داده‌هایی که آنها استفاده نمودند مربوط به یک بانک محلی که وام مسکن به مشتریانش اعطا می‌نمود بود. [۱۹] هسیه در سال ۲۰۰۵ سیستمی هیبریدی را برای امتیازدهی اعتباری بر اساس دسته‌بندی و شبکه‌های عصبی طراحی کرد. وی از دسته‌بندی

برای انجام فرایند آنالیز، طبقه‌بندی و هرس کردن آن‌ها قبل از ورود به شبکه عصبی استفاده نمود و از شبکه عصبی برای ساخت مدل امتیازدهی اعتبار استفاده نمود [۲۰].

لاها در سال ۲۰۰۶ روشی را برای امتیازهای اعتباری توسط کلاسیفایرهای قانون محور فازی معرفی کرد. در این روش ابتدا فرایند یادگیری بر روی قانون محور توسط داده‌های آموزشی انجام می‌شود. سپس قانون  $K-mn$  فازی نیز با آن ترکیب می‌شود. از قابلیت‌های این روش این است که محدودیت‌های تجاری نیز در آن نشان داده می‌شوند [۲۱]. در تحقیقات نیز در سال ۲۰۰۷ جیانو و همکاران با استفاده از شبکه تطبیقی فازی (FAN) به سنجش اعتبار در شرکت‌های مالی کوچک پرداختند. در این مدل ابتدا داده‌های سنجش اعتبار توسط اعداد فازی نمایش داده می‌شوند سپس شبکه فان بر اساس قانونهای استنتاجی که ۲۷ قانون فازی هستند ساخته می‌شود و در نهایت شبکه با استفاده از داده‌های آموزشی اعداد فازی آموزش داده می‌شود و یا فرایند یادگیری بر روی آن انجام می‌شود. مزایای اصلی این سیستم عبارتند از: بازنمایی زبانی، تراکم زبانی و خاصیت یادگیرنده شبکه‌های عصبی [۲۲]. از دیگر کاربردهای سیستم استدلال فازی، مدل رتبه‌بندی مبتنی بر پایگاه قواعد فازی هافمن و همکاران را می‌توان نام برد. در این مدل با استفاده از ترکیب سیستم استدلال فازی تقریبی و توصیفی با الگوریتم ژنتیک، موفق به رتبه‌بندی اعتباری مشتریان حقوقی با دقت بیش از ۸۰٪ شدند [۲۳].

## ۵- روش تحقیق

در این پژوهش، سؤال ابتدایی این است که چه ترکیبی از عوامل می‌توانند در کشف احتمال نکول مؤثر باشد؟ هر یک از این عوامل چه اولویت (وزنی) دارد. از طریق پرسشنامه قضاوت خبرگان حوزه اعتبارات جمع‌آوری و با روش سوروی کدگذاری شد. سپس تحلیل آماری نتایج و در نهایت رتبه‌بندی بوسیله روش تاپسیس و آزمون فریدمن انجام شد. سپس عوامل رد شده و رد نشده در چندین ترکیب مختلف به شبکه عصبی فازی انطباقی سپرده شد و بهترین ترکیب از بین آنها شناسایی و مشخص گردید که تا چه میزان می‌توان به عوامل شناسایی شده به کمک روش تاپسیس و فریدمن خوش‌بین بود. این آزمون، پاسخی مشخص به اعتبار عوامل شناسایی شده در این تحقیق و سؤال اول این تحقیق است. سپس به بررسی کارایی روش شبکه عصبی فازی انطباقی در شناسایی و برقراری ارتباط مناسب برای درک پیش‌بینی نکول (خوش‌حسابی/بد حسابی)

مشتریان پرداخته شد. در انتها با کمک گرفتن از روش بهینه‌سازی کلونی مورچگان مشخص شد که تا چه میزان می‌توان نسبت به بهبود عملکرد شبکه عصبی فازی انطباقی مبادرت نمود.

## ۶- تعیین وزن و اهمیت شاخص‌ها

متغیرهای مستقل این تحقیق شامل ۱۷ شاخص دریافت شده از بانک است که در سه گروه "شخصی"، "مالی" و "کسب و کار" دسته‌بندی شده است و شامل: وضعیت حساب جاری (حجم گردش مالی در حساب)، سابقه ارتباط با بانک (مدت حساب)، سابقه باز پرداخت وام‌ها (قسط معوقه)، نوع وام، میزان اعتبار، میزان سپرده، وضعیت پرداخت بیمه، نرخ بهره، جنس، تاهل / تجرد، وضعیت ضمانت، محل سکونت، سن، تعداد حساب‌های، شغل، تعداد افراد تحت تکفل، تحصیلات. برای ارزیابی اهمیت نسبی عوامل مذکور پانلی از خبرگان و متخصصان حوزه اعتبارات شامل ۳۲ نفر با ترکیب سطح تحصیلات: ۶ نفر دکترا، ۱۹ نفر کارشناس ارشد، ۷ نفر کارشناس که و ترکیب شغلی: ۴ نفر استاد دانشگاه، ۵ نفر رئیس شعبه بانک، ۶ نفر مدیرعامل سازمان‌ها، ۵ نفر مدیر مالی، ۶ نفر کارشناس اعتبارات، ۶ نفر حسابرس رسمی انتخاب گردید و شاخص‌های فوق‌الذکر در قالب پرسشنامه‌ای با ساختار امتیازدهی لیکرتی ۵ تایی به منظور تعیین امتیاز هر شاخص بین اعضاء پانل توزیع گردید و مطابق روش سوروی نظرات جمع بندی و کد گذاری و آماده تحلیل آماری گردید. ضریب آلفای کرنباخ، شاخصی از هماهنگی درونی در بین گویه‌ها قلمداد می‌شود. ضریب آلفای کرنباخ برای نتایج برابر با ۰/۹۵ محاسبه شد. در اغلب منابع در رابطه با مقدار مطلوب برای ضریب آلفا، به نوشته نونالی استناد شده است. بر اساس استدلال وی، برای پایا قلمداد کردن یک ابزار، حداقل مقدار ۰/۷ برای ضریب آلفا لازم است، لذا میتوان ابزار طرح شده را بر اساس شیوه هماهنگی درونی گویه‌ها، پایا قلمداد کرد. محاسبه ضریب آلفای ۰/۹۵ حاکی از وجود واریانس مشترک (کوواریانس) بالا و در جهت مثبت بین واریانس‌های قضاوت خبرگان می‌باشد. بر این اساس انتظار داریم که همبستگی بین اعضا پانل بالا بوده و انتخاب مناسبی صورت گرفته باشد. برای تعیین میزان اتفاق نظر میان اعضای پانل و روایی ابزار، از ضریب همبستگی کندال استفاده گردید. ضریب هماهنگی کندال مقیاسی است که برای تعیین درجه هماهنگی و موافقت میان چندین دسته رتبه مربوط به N شی یا فرد. ضریب هماهنگی کندال نشان می‌دهد افرادی که چند مقوله را براساس اهمیت آن‌ها امتیازدهی کرده اند، به طور اساسی



معیارهای مشابهی را برای قضاوت درباره اهمیت هر یک از مقوله‌ها به کار برده‌اند و از این لحاظ با یکدیگر اتفاق نظر دارند. ضریب کندال بین ۰ و ۱ متغیر است. اگر ضریب کندال صفر باشد یعنی عدم توافق کامل و اگر یک باشد یعنی توافق کامل وجود دارد. برای پایان تکنیک‌های سوروی و دلفی از ضریب هماهنگی کندال استفاده می‌گردد. نتیجه محاسبات نشان می‌دهد که ضریب کندال تقریباً ۰,۳ می‌باشد که نشان‌دهنده هماهنگی نسبی بین نحوه قضاوت خبرگان در خصوص تعیین اهمیت شاخصهای مؤثر بر نکول اعتبارات توسط مشتریان می‌باشد. پس از اطمینان از صحت ابزار و داده‌های جمع‌آوری شده، داده‌ها توسط روش تاپسیس رتبه‌بندی شدند. بردار وزن مورد نیاز الگوریتم تاپسیس بر اساس دو پارامتر اصلی و تعیین‌کننده در مورد خبرگان یعنی " سطح تحصیلات" و " شغل" آنها محاسبه و تنظیم گردید و همچنین به منظور مقایسه نتایج نیز از روش رتبه‌بندی فریدمن که روشی آماری مبتنی بر میانگین می‌باشد، جهت مقایسه رتبه تأثیرگذاری عوامل به طور همزمان استفاده شده است. آزمون فریدمن یک آزمون ناپارامتری، معادل آنالیز واریانس با اندازه‌های تکراری (درون گروهی) است که از آن برای مقایسه میانگین رتبه‌ها در بین K متغیر (گروه) استفاده گردید. مشخص شد که اختلاف امتیازات شاخص‌ها معنی دار است و همچنین کدام شاخص بیشترین و کدام شاخص کمترین امتیاز را کسب کرده است. نتایج در جدول شماره (۱) ارائه گردیده است. نتایج هر دو روش به وضوح رتبه‌بندی انجام شده توسط روش دیگر را تأیید می‌کنند.

جدول ۱ نتایج مقایسه‌ای رتبه‌بندی شاخص‌ها با روش فریدمن و Topsis

نتایج وزن دهی		عامل	کد	حیطه
TOPSIS	فریدمن			
۰/۳۴۸	۸/۸	وضعیت حساب جاری (حجم گردش مالی در حساب)	۰۱۷	مالی
۰/۸۰۲	۱۵/۸۴	سابقه ارتباط با بانک (مدت حساب)	۰۲۷	مالی
۰/۳۶۲	۹/۵۹	سابقه باز پرداخت وام‌ها (قسط معوقه)	۰۳۷	مالی
۰/۲۹۴	۸/۲۵	نوع وام	۰۴۷	مالی
۰/۵۱۳	۱۱/۱۴	میزان اعتبار	۰۵۷	مالی
۰/۳۰۵	۷/۹۷	میزان سپرده	۰۶۷	مالی
۰/۱۶۴	۵/۳۱	وضعیت پرداخت بیمه	۰۷۷	شغلی
۰/۲۸۶	۷/۹۴	نرخ بهره	۰۸۷	مالی
۰/۲۴۶	۶/۰۲	جنس	۰۹۷	فردی
۰/۱۸۸	۶/۱۹	تأهل / مجرد	۱۰۷	فردی
۰/۳۷۵	۱۰/۰۲	وضعیت ضمانت	۱۱۷	مالی
۰/۲۸۳	۸/۰۵	محل سکونت	۱۲۷	فردی
۰/۲۷۶	۸/۷	سن	۱۳۷	فردی
۰/۵۷۸	۱۳/۱۶	تعداد حساب‌های	۱۴۷	فردی
۰/۲۸۱	۷/۷۵	شغل	۱۵۷	شغلی
۰/۳۱۶	۸/۰۹	تعداد افراد تحت تکفل	۱۶۷	فردی
۰/۴۱۵	۱۰/۱۹	تحصیلات	۱۷۷	فردی

بر اساس وزنهای کسب شده و با انتخاب چندین ترکیب مختلف از جمله ترکیب انتخابی بر اساس قانون پارتو، به بررسی نتایج پرداخته شد. در انتخاب ترکیب‌های مختلف، ابتدا بر مبنای قانون پارتو (قانون ۲۰-۸۰) بیست درصد از عواملی که می‌توانستند بیشترین درصد از وزن را





ایجاد نمایند، استخراج شدند. سپس با کسر تصادفی برخی از عوامل ترکیب‌های شماره ۲، ۳ و ۵ حاصل و به مدد افزودن تصادفی برخی از عوامل به ترکیب پارتو، ترکیب‌های ۱۳، ۲ و ۵ بدست آمدند. کاستن از برخی عوامل مهم سبب افزایش خطا گردید و از سوی دیگر با افزودن به عوامل نیز نسبت به حالت کاستن، میزان کمتری خطا ایجاد شد، اما سبب بروز اشتباه در برآوردهای آتی و افزایش حجم محاسبات نیز مدل در دوره آموزش گردید. در نهایت سه شاخص " سابقه ارتباط با بانک (مدت حساب)"، " میزان اعتبار" و "سن" برای استفاده در مدل شبکه عصبی انتخاب شدند.

## ۷- ایجاد سیستم استنتاج فازی (FIS)

سیستمی که یک نداشت از ورودی به خروجی را با استفاده از منطق فاز فرموله میکند به نام سیستم استنتاج فازی شناخته می‌شود. سیستم استنتاج فازی همچنین سیستم مبتنی بر قواعد نیز نامیده می‌شود. زیرا این سیستم‌ها از تعدادی عبارت « اگر - آنگاه » ساخته شده است. با توجه به اینکه ورودی‌ها و خروجی‌ها هر دو فازی می‌باشند سیستم استنتاج فازی مددانی را انتخاب نموده و پیاده‌سازی شد. پس از ایجاد سیستم استنتاج فازی، ابتدا متغیرها را تعریف می‌کنیم که شامل ۳ ورودی و یک خروجی می‌شود. ورودی‌ها شامل: مدت حساب (V02)، میزان اعتبار (V05)، سن (V13) و خروجی خوشه اعتباری (Y2) می‌باشد.

بعد تابع عضویت برای هر یک از آنها تعیین شد. توابع عضویت هر متغیر بر اساس ساختار فازی که برای آنها تعریف شده است از میان مجموع توابع عضویت انتخاب گردیده و پارامترهای آنها نیز محاسبه و در مدل لحاظ شدند. به علت اینکه تغییر نرم عوامل در این تحقیق مد نظر قرار گیرد در مدل‌سازی توابع عضویت ورودی‌ها از نوع گوسی و برای خروجی به علت اینکه هدف کلاس‌بندی مشتریان است از تابع مثلثی استفاده گردیده است. سن با ۳ تابع عضویت، مدت ارتباط با بانک با ۴ تابع عضویت، میزان اعتبار با ۵ تابع عضویت و خروجی یا همان خوشه اعتباری با ۸ تابع عضویت در مدل تعریف شدند.

پایگاه داده قوانین استنتاج فضایی است از نگاشتهای گزاره‌های متشکل از ترکیبات مختلف متغیرهای ورودی به پارامتر خروجی. به شکل عام کلیه ترکیبات ممکن از ۳ حالت سن و ۴ حالت مدت حساب و ۵ حالت میزان اعتبار کلیه گزاره‌های ضربی ۳ مورد فوق را شامل می‌گردد

که مجموعاً ۶۰ قانون مختلف را ایجاد می‌کند. به منظور بهینه نمودن و کاهش زمان آموزش شبکه عصبی با بررسی و تحلیل آماره موجود و همچنین پیمایش درخت تصمیم ۱۶ قانون استنتاج فازی تعریف گردید. سپس با بررسی و تحلیل نمودارهای پوششی و عملگرهای فازی مدل مشخص گردید که نمودارهای سطح و آنالیز تجمعی قوانین فضای پاسخ را به خوبی پوشش می‌دهند.

سیستم استنتاج فازی تعریف شده را به منظور بهینه‌سازی با شبکه عصبی تلفیق نموده و سیستم استنتاج فازی-عصبی انطباقی را ایجاد نموده. نکته مهم در این بخش این است که نوع سیستم استنتاج فازی لزوماً می‌بایست سوگنو باشد. در این تحقیق با توجه به توضیحات بخش قبل و الزامات مسأله سیستم به صورت مددانی طراحی گردیده است. لذا در مرحله اول و پیش بارگذاری سیستم استنتاج فازی، ابتدا با استفاده از تابع (mam2sug) در خط فرمان متلب، سیستم استنتاج فازی که به صورت مددانی طراحی شده است به نوع سوگنو تبدیل کرده ایم.

سپس با استفاده از بخشی آماره تحقیق، سیستم تعریف شده را آموزش دادیم تا پایگاه داده قوانین استنتاج موجود در آن منطبق بر ساختار آماری موجود در داده‌های تحقیق بهینه شود. به این منظور ابتدا داده‌های آماری پژوهش شامل ۱۰۰۰ رکورد از مشتریان پست بانک استان تهران که اعتبار دریافت کرده‌اند را به سه مجموع به صورت تصادفی تقسیم می‌کنیم. ۷۰٪ داده‌ها برای آموزش، ۱۵٪ برای اعتبار سنجی، ۱۵٪ برای آزمایش و راستی سنجی مدل. پارامترهای اصلی آموزش یعنی روش بهینه‌سازی، این پارامتر روش بهینه‌سازی توابع عضویت را مشخص می‌نماید که در این پژوهش از روش هیبرید که از روشهای حداقل مربعات و روش کاهش شیب پس انتشار می‌باشد استفاده شده است. تعداد دوره‌های آموزش، حداکثر تعداد دوره‌های آموزش را مشخص می‌کند که در این تحقیق پس از تکرارهای متعدد، عدد ۲۰۰۰ برای این پارامتر تعریف گردید و خطای مطلوب: این پارامتر میزان اختلاف بین خروجی مدل و داده‌های واقعی را مشخص می‌کند. در این پژوهش خطای مطلوب کمتر از ۰/۵ تعریف شده است به این دلیل که فضای بین هر گروه از رتبه‌های اعتباری یک واحد می‌باشد و با احتساب خطای کمتر از ۰/۵ می‌توان با گرد کردن نتایج ورودی را رتبه‌بندی کنیم. در مرحله اول آموزش با پارامترهای فوق با میزان خطای ریشه میانگین مربعات ۰/۶۲۰۸۶ به پایان رسید. به منظور تقدیق نتایج و افزایش دقت مدل، فرآیند آموزش را بر روی مدل موجود مجدد



تکرار می‌نماییم تا به کمینه خطای ممکن دست پیدا کنیم و به  $0/61908$  تقلیل یافت. از داده‌های معتبرسازی برای بررسی میزان عمومیت سیستم استنتاج فازی در هر تکرار استفاده می‌شود. داده‌های واری دارای ساختار مشابه با داده‌های آموزش هستند. در صورت وقوع بیش برآزش، سیستم به مجموعه داده‌های مستقل دیگر پاسخ مناسبی نمی‌دهد. بخصوص در صورتی که داده‌ها دارای اختلال باشند. در این وضعیت استفاده از داده‌های واری و معتبرسازی بسیار مفید واقع می‌شود. طی روال آموزش، داده‌های واری نیز به سیستم اعمال می‌شود و با بررسی خروجی، از وقوع بیش برآزش جلوگیری می‌شود. در این مدل با استفاده از داده‌های معتبر سازی خطای RMSE را که در مرحله قبل  $0/61908$  بود تا  $0/57482$  کاهش دادیم. با بررسی توزیع داده‌های معتبرسازی پژوهش و نتایج خروجی مدل، مقدار خطا را به میزان تقریبی ۴ درصد کاهش داده و مدل را در سطح معقولی بهینه نموده ایم. پس از آموزش نهائی مدل با استفاده از داده‌های آزمایش، عملکرد نهائی مدل بررسی گردید و خطای پیش‌بینی آن  $0/57468$  محاسبه گردید. یکی از پارامترهای بررسی بهینه بودن مدل این است که خطای محاسبه شده برای داده‌های آموزش و آزمایش به هم نزدیک باشند. چنانچه مشاهده می‌گردد در مدل فوق این شرط محقق گردیده است. نمودارهای سطح به شکل قابل توجهی بهبود یافته و پوشش مناسب‌تری از فضای مسأله ایجاد نموده اند. در طراحی یک سیستم استنتاج فازی تمام تلاشها در راستای بهینه نمودن مدل و کاهش خطای پیش‌بینی می‌باشد.

در یک مدل یا برآورد گر آماری به صورت کلی عملکرد به این صورت می‌باشد که داده‌ها به صورت موازی به مدل داده می‌شود و نتایج با نتایج واقعی در هر دور مقایسه گردیده و اختلاف بین خروجی مدل و مقادیر واقعی محاسبه گردیده و سعی می‌گردد خطای تجمعی مدل با تغییر در پارامترهای مدل کاهش یافته و به صفر میل کند. یک راه برای بهینه و میزان کردن مدل با داده‌های واقعی این است که پارامترهای مدل را آنقدر تغییر دهیم تا به بهترین حالت ممکن دست یابیم. بیشتر این پارامترها در دوره‌های متوالی آموزش مدل توسط شبکه عصبی تا حد امکان اصلاح و بهینه گردیده است. لیکن در این تحقیق برای افزایش دقت و جلوگیری از افتادن الگوریتم در نقاط کمینه نسبی از یک سو و از سوی دیگر جلوگیری از بیش برآزش مدل از الگوریتم بهینه‌سازی کلونی مورچگان به عنوان یک تابع در مدل استفاده نموده ایم. در این پژوهش به منظور اصلاح ارزش وزنی قوانین در سیستم استنتاج فازی، ارزش وزنی قوانین

بوسیله الگوریتم کلونی مورچه اصلاح گردیده و سپس مدل با داده‌های آزمایش مجدداً تست شده است. الگوریتم بهینه‌سازی کلونی مورچگان در واقع یک الگوریتم بهینه‌سازی با استفاده از هوش ازدحامی می‌باشد. بدین معنا که در این نوع از الگوریتم‌ها و مسائلی که به این روش قابل حل می‌باشند، در صورتی نتایج بهینه حاصل می‌گردد که از تجربیات تعداد زیادی از گره‌های فعال استفاده نماییم.

سیستم استنتاج فازی- عصبی انطباقی نهایی با استفاده از داده‌های آزمایش مدل اولیه سنجیده شده و نتایج حاصل نشان‌دهنده بهبود نسبی نتایج می‌باشد. میزان خطای آزمایش متوسط از سطح  $0/57468$  به  $0/55055$  کاهش یافته است. نتایج نشان می‌دهد استفاده از این روش منجر به کاهش  $4/19\%$  خطای تجمعی برازش مدل گردیده است و مدل سیستم استنتاج فازی- عصبی انطباقی را به مقدار قابل قبولی بهبود داده است. لازم به توضیح است که این نتیجه برای یک ابر بهینه ساز که نتایج یک الگوریتم بهینه‌سازی دیگر را بهبود می‌بخشد و با توجه به نتایج مراحل پایانی سیستم استنتاج فازی- عصبی انطباقی که در آن شیب نمودار کاهش خطا به شدت به صفر میل می‌نمود، نتیجه رضایت بخشی می‌باشد.

## ۸- فرآیند اعتبارسنجی سیستم استنتاج فازی- عصبی انطباقی

برای اعتبارسنجی سیستم، شاخص‌های متعددی وجود دارند که مهمترین آنها در جدول شماره (۲) تشریح شدند.

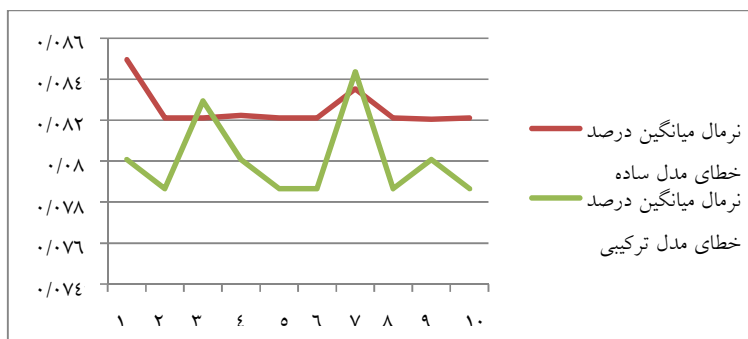
جدول ۲ معیارهای اعتبارسنجی سیستم استنتاج فازی- عصبی انطباقی

مقدار بهینه	بازه مقادیر	معادل انگلیسی	معیار اعتبارسنجی
هر چه کمتر باشد بهتر است.	از صفر تا بی نهایت	SSE	جمع مجزورات خطا
هر چه کمتر باشد بهتر است	از صفر تا بی نهایت	RMSD یا RMSE	ریشه میانگین مربعات خطا
بین $0/6$ تا $0/8$ مناسب / بین $0/8$ تا $1$ عالی	از صفر تا یک	$R^2$	ضریب تشخیص



ادامه جدول ۲

مقدار بهینه	بازه مقادیر	معادل انگلیسی	معیار اعتبارسنجی
بین ۰/۶ تا ۰/۸ مناسب / بین ۰/۸ تا ۱ عالی	از صفر تا یک	Adjusted-R <sup>2</sup>	ضریب تشخیص تصحیح شده
در مقایسه دو سیستم، هر چه کمتر باشد بهتر است	از صفر تا بی نهایت	MSE	میانگین مربعات خطا
در مقایسه دو سیستم، هر چه کمتر باشد بهتر است	از صفر تا بی نهایت	MAE	میانگین خطای مطلق
هر چه کمتر باشد بهتر است	از صفر تا بی نهایت	NRMSD	نرمال انحراف جذر میانگین مربعات
هر چه کمتر باشد بهتر است	از صفر تا بی نهایت	CV(NRMSD)	ضریب تغییرات انحراف جذر میانگین مربعات



نمودار ۱ مقایسه استفاده از شبکه عصبی فازی انطباقی به تنهایی و در ترکیب با کلونی مورچگان

با بهینه شدن مقادیر برای سه شاخص ریشه میانگین مربعات خطا و ضریب تشخیص و ضریب تشخیص تصحیح شده، مقادیر شاخصهای دیگر نیز بطور معمول بهینه خواهند بود. ولی در صورت وجود اختلاف ناچیز، اولویت با این سه شاخص می‌باشد. نکته پایانی اینکه به جز استفاده از الگوریتم کلونی مورچگان، سایر مقادیر تنظیم شده برای آموزش سیستم‌ها،

یکسان می‌باشد. در جدول (۳) نتایج اعتبار سنجی بهترین عملکرد مدل با شاخص‌های فوق‌الذکر ارایه گردیده است. در نمودار (۱) نتایج مقایسه‌ای ۱۰ اجرای متفاوت مدل به صورت مقایسه‌ای ارایه گردیده است. نمودارهای مربوط به سایر خطاها نیز به مقیاس روند و شکلی مشابه این نمودار داشته و فاصله عملکردی بین دو مدل را در مقیاس نمایش می‌دهند.

جدول ۳ شاخص‌های اعتبار سنجی شبکه عصبی فازی انطباقی به تنهایی و در ترکیب با کلونی مورچگان

اعتبارسنجی سیستم استنتاج فازی-عصبی انطباقی			شاخص اعتبار سنجی	
% بهبود نتایج	مدل ترکیبی	مدل ساده		
%/۸۲۲	۲۱۲/۱۷۳۷۱۲	۲۳۱/۱۷۹۹۷۱۷	SSE	جمع مجزورات خطا
%/۴۲۰	۰/۵۵۰۵۵	۰/۵۷۴۶۸	RMSD و RMSE	ریشه میانگین مربعات خطا
%/۹۵	۰/۹۰۵۲۳۷۲۹	۰/۸۹۶۷۴۸۵۶۱	R <sup>2</sup>	ضریب تشخیص
%/۰,۹۴	۰/۹۰۲۲۳۳۴۳	۰/۸۹۳۸۶۱۱۷۳۸	Adjusted-R <sup>2</sup>	ضریب تشخیص تصحیح شده
%/۸۲۲	۰/۳۰۳۱۰۵۳	۰/۳۳۰۲۵۷۱۰۲	MSE	میانگین مربعات خطا
%/۳۳۷	۶۲۰۳۵۰۹۰	۰/۶۴۲۰۰۸۶۳۱	MAE	میانگین خطای مطلق
%/۴۷/۲۰	۰/۰۷۸۶۵	۰/۰۸۲۰۹۷۱۴۳	NRMSD	نرمال انحراف جذر میانگین مربعات
%/۴۲۰	۰/۱۲۳۷۱۹۱	۰/۱۲۹۱۴۱۵۷۳	CV(NRMSD)	ضریب تغییرات انحراف جذر میانگین مربعات

## ۹- نتیجه گیری کلی

در این تحقیق، به بررسی استفاده از تکنیک سیستم استنتاج فازی عصبی انطباقی در ترکیب با الگوریتم بهینه‌سازی کلونی مورچگان جهت پیش‌بینی میزان نکول وام‌ها در شعب پست بانک استان تهران پرداخته شده است. از آنجایی که بسیاری از تکنیک‌های قدیمی و مبتنی بر



احتمالات و فرآیندهای تصادفی نمی‌توانند در تعامل با داده‌های با ماهیت ابهام‌آلود و فازی گونه، مناسب عمل نمایند، گستره وسیعی از روشها در ادبیات تحقیق مورد بررسی قرار گرفته‌اند. لذا روش موجود با گردآوری تعداد مناسبی داده و چندین بار اجرا در نرم‌افزار و آموزش، مورد اعتبارسنجی واقع شد. آنچه که نقطه مزیت این تحقیق در مقابل سایر تحقیقات مشابه می‌تواند قلمداد شود را می‌توان استفاده ترکیبی از سه روش دانست که گرچه از نگرش‌های مختلفی همچون تصمیم‌گیری چندمعیاری، داده‌کاوی برای استخراج دانش و نیز الگوریتم بهینه‌سازی ترکیبی بهره می‌برد اما توانست با استفاده از یک رویکرد مناسب از مزایای روش‌ها بهره برده تا نقاط ضعف یکدیگر را پوشش دهند. نتایج حاکی از این است که می‌توان از این روش جهت ارائه تخمین‌هایی مناسب برای آینده و نیز ارزیابی وضعیت فعلی بهره برد. آنچه که از ماحصل خروجی‌ها بدست می‌آید، به شرح ذیل می‌باشد:

در بخش اول از این پژوهش با استفاده از تکنیکهای تصمیم‌گیری چندمعیاره و قضاوت خبرگان در ترکیب با روشهای آماری ناپارامتری فریدمنو الگوریتم Topsis ابتدا ۱۷ شاخص را رتبه‌بندی نمودیم. سپس در مرحله بعد یک شاخص غیر مالی و دو شاخص مالی را برای استفاده در سیستم استنتاج فازی انتخاب نمودیم. "سه شاخص "سن"، "سابقه ارتباط با بانک (مدت حساب)" و "میزان اعتبار" انتخاب شدند.

در مرحله بعد، سیستم استنتاج انطباقی فازی اولیه با استفاده از پارامترهای انتخاب شده و آنالیزهای آماری پیاده‌سازی شد. در این مرحله تابع‌های عضویت با استفاده از تکنیکهای پایه آماری و عموماً با رگرسیون و همبستگی بین فضای داده‌های ورودی و خروجی طراحی گردید. با استفاده از سیستم استنتاج انطباقی فازی شبکه عصبی طراحی گردید، بعد از آن از حدود ۷۰ درصد از داده‌ها (۱۰۰۰ رکورد مشتریان) بصورت داده‌های آموزش استفاده شده، ۱۵ درصد از مابقی داده‌ها به عنوان داده‌های اعتبارسنجی استفاده شدند و مجدداً پیکربندی تجدید شد، به کمک همان داده‌ها مجدداً رفتار سیستم ارزیابی شد و هنگامی که ضریب تشخیص شبکه به بیش از ۸۹ درصد رسید، شبکه آماده فرض شد. در نهایت به کمک ۱۵ درصد باقی مانده، رفتار سیستم مورد آزمایش قرار گرفت و در نهایت عدد ۸۹/۶۷ درصد برای ضریب تشخیص نمایشگر این واقعیت بود که ۸۹/۶۷ درصد از مواقع این سیستم می‌تواند تخمین درستی نسبت به رتبه اعتباری مشتریان ارائه دهد. مرحله بعدی با استفاده از الگوریتم کلونی مورچگان مجدداً

ارزش وزنی قوانین استنتاج بهینه شد و مدل با استفاده از همان ترکیب داده‌های آزمایش مورد ارزیابی قرار گرفت و بیشینه ضریب تشخیص به دست آمده ۹۰/۵ درصد می‌باشد. لذا می‌توان سیستم حاصل را به عنوان ابزار مناسبی جهت پشتیبانی از تصمیم‌گیری در سطح اطمینان حدود ۹۰ درصدی عنوان نمود که در تعیین میزان نکول وام در سیستم بانکداری کشور می‌تواند مؤثر عمل نماید. به دیگر سخن می‌توان این ابزار را نه فقط برای صاحبان حساب در بانک‌های مختلف بلکه برای تمامی متقاضیان جدیدی که هیچ سابقه حساب نزد بانک‌ها ندارند را مؤثر دانست. با ترکیب الگوریتم شبکه عصبی فازی انطباقی با کلونی مورچگان می‌توان ضریب تشخیص مدل را تا تقریباً ۱ درصد بهبود دهیم و چنین بیان کنیم عملکرد مدل با استفاده از این روش ترکیبی به میزان معنی داری افزایش پیدا می‌کند.

#### ۱۰- پی‌نوشت‌ها

1. Credit Scoring
2. Credit Rating
3. Exposure
4. Liquidity, Activity, Profitability, Potential (LAPP)
5. Linear discrimination analysis
6. Classification and Regression Trees
7. Multivariate Adaptive Regression Splines
8. ANN(Artificial Neural Networks)
9. NFS (Neural Fuzzy Systems)
10. ANFIS (Adaptive Neural Fuzzy Inference System)
11. Support Vector Machine
12. Genetic Algorithms

#### ۱۱- مراجع و منابع

- [1] Sinky Jr, Joseph F.; (1992), Commercial Bank Financial Management; 4th Edition; Macmillan, P.65
- [2] Basel Committee on Banking Supervision.; (2000), "Credit rating and complementary sources of credit quality information"; *Working Paper*, No.3, P. 7
- [3] Jamshidi Saedi.; "Customer validation methods"; *Journal of Monetary and*



- Banking Research Center*, pp. 21, 30, 33, 36
- [4] West D.; (2000), "Neural network credit rating models"; *Journal of Computers & Operations Research*, Vol. 27, P. 26
- [5] Lee TS & et al.; (1966), "Credit rating using a hybrid neural discriminant technique"; *Journal of Expert Systems with Applications*, Vol. 23, 2002. P. 43
- [6] Beaver W.; "Financial ratios as predictors of failure"; *Journal of Accounting Research*, Vol. 5, P. 25
- [7] Emel AB. & et al.; (2003), "A credit rating approach for the commercial banking sector"; *Journal of Socio-Economic Planning Sciences*, Vol. 37, P. 16
- [8] Lung Huang C. & et al., (2007), "Credit rating with a data mining approach based on support vector machines"; *Expert System with Application*, Vol.3, 2, P. 165
- [9] Desai. V, Crook. J. and Overstreet. G, 1997; "Credit scoring models in the credit union environment using neural networks and genetic algorithms", *IMA Journal of Mathematics Applied in Business and Industry*, Vol 8, Issue 4, pp. 232-256.
- [10] Cheng, C. B. and Lee, E. S., (1999), "Applying Adaptive network to fuzzy regression analysis.", *Computers and Mathematics with Applications*, Vol. 38, pp. 123-140.
- [11] West. D, (2000), "Neural network credit scoring models", *Computers and Operations Research*, Vol. 27, pp. 1131-1152.
- [12] Boussabaine, A. H. and Wanoous, M." (2000), "A neurofuzzy model for predicting business bankruptcy." In *Business Applications of Neural Networks: The State-of-the-Art of Real-World Applications* (ed.), pp. 55-69.
- [13] Castillo, O. and Melin, P., (2002), "Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic and fractal theory." *IEEE Transactions on Neural Networks*, Vol.13, No. 16.
- [14] Trinkle. B. S, (2006), *Interpretable Credit Model Development Via Artificial Neural Network*, Ph.D. Dissertation, University of Alabama
- [15] Bennell. J.A, Crabbe.D, Thomas. S, Gwilym. O.A, (2006), " Modelling sovereign

- credit ratings: Neural networks versus ordered probit", *Expert Systems With Applications*, Vol. 30, PP. 415-425.
- [16] Tsai C-F, Wu J-W, (2007); "Using neural network ensembles for bankruptcy prediction and credit scoring", *Expert Systems with Applications*, in Press
- [17] Malhorta. Malhorta.D.K, 2002; "Differentiating between good credits and bad credits using neuro-fuzzy systems", *European Journal of Operation Research*, pp. 190-211.
- [18] Huang. Z, Chen. H, Hsu. C. J, Chen. W. H, Wu.S, (2004); "Credit rating analysis with support vector machines and neural networks: A Market Comparative Study", *Decision Support Systems*, 37(4), pp. 543–558.
- [19] Lee. T.S, Chen. I.F, (2005); "A Two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines", *Expert Systems with Applications*, 28(4), pp. 743–752.
- [20] Hsieh. N.C, (2005); "Hybrid Mining approach in the design of credit scoring models", *Expert Systems with Applications*, 28(4), pp. 655–665.
- [21] Laha.A, (2007); "Building Contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring", *Advanced Engineering Informatics*, Vol 21, pp. 281–291.
- [22] Jiao, Y, Syaub. Y-R, Lee.E.S, (2007); "Modelling credit rating by fuzzy adaptive network", *Mathematical and Computer Modelling*, 45, pp. 717-731.
- [23] Hoffmann, F. Baesens, C. and Mues, T., (2007), "Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms", in *European Journal of Operational Research*, Vol. 177.