Iran J. Insur. Res., 8(1): 27-40, Winter 2019, Serial #27

Iranian Journal of Insurance Research
(IJIR)
Homepage: https://ijir.irc.ac.ir/?lang=en

Insurance Research Center (IRC)

ORIGINAL RESEARCH PAPER

Customer credit scoring using data envelopment analysis and discriminant analysis in a fuzzy environment (case study: a leasing company affiliated with a private bank)

A. Alinezhad*, S. Kashanifar

Department of Industrial Engineering, Faculty of Industrial and Mechanical Engineering, Islamic Azad University, Qazvin Branch, Iran

ARTICLE INFO

Article History
Received: 17 May 2018
Revised: 15 June 2018
Accepted: 18 December 2018

Keywords
Data Envelopment Analysis/ Discriminant Analysis; Fuzzy Data; Credit Risk; Credit Scoring.

ABSTRACT

In this research, in order to manage and control the credit risk of customers, from the combination of two models of discriminant analysis and data coverage analysis to detect the presence or absence of an overlap between two groups by means of a separating hyperscreen and assuming the existence of each observation with independent characteristics with The presence of fuzzy data, the observations were categorized into two categories of good customers and bad customers. The variables of this research were selected from the 6C method, and out of the 17 selected indicators, using the Delphi method, 8 influential indicators were included in the research model. These indicators were used for 83 real customers of a leasing company who received facilities during 2013 and 2014. The results show that each of the observations is definitively placed in the category of good customers and bad customers, and with the arrival of each new observation, its credit status is predicted.

***Corresponding Author:**
Email: *alalinezhad@gmail.com*
DOI: 10.22056/ijir.2019.01.03

نشریه علمی پژوهشنامه بیمه

سایت نشریه: https://ijir.irc.ac.ir/?lang=fa

مقاله علمی

امتیازدهی اعتباری مشتریان با استفاده از تحلیل پوششی داده‌ها و تحلیل ممیز در محیط فازی (مطالعهٔ موردی: یک شرکت لیزینگ وابسته به یک بانک خصوصی)

علیرضا علی‌نژاد*، سعید کاشانی‌فر

گروه مهندسی صنایع، دانشکدهٔ مهندسی صنایع و مکانیک، دانشگاه آزاد اسلامی، واحد قزوین، ایران

اطلاعات مقاله

تاریخ دریافت: ۲۷ اردیبهشت ۱۳۹۷

تاریخ داوری: ۲۵ خرداد ۱۳۹۷

تاریخ پذیرش: ۲۷ آذر ۱۳۹۷

کلمات کلیدی

تحلیل پوششی داده‌ها/ تحلیل ممیز

داده‌های فازی

ریسک اعتباری

امتیازدهی اعتباری

چکیده:

در این تحقیق به منظور مدیریت و کنترل ریسک اعتباری مشتریان، از تلفیق دو مدل تحلیل ممیز و تحلیل پوششی داده‌ها برای تشخیص وجود و یا عدم وجود یک هم‌پوشانی میان دو گروه به‌وسیله یک ابرصفحه جداکننده و با فرض وجود مشاهده هریک با مشخصه مستقل با حضور داده‌های فازی، مشاهدات به دو رده مشتریان خوش‌حساب و مشتریان بدحساب دسته‌بندی شدند. متغیرهای این تحقیق از روش C۶ انتخاب و از تعداد ۱۷ شاخص منتخب، با استفاده از روش دلفی، تعداد ۸ شاخص تأثیرگذار وارد مدل تحقیق شده است. این شاخص‌ها برای ۸۳ نفر از مشتریان حقیقی یک شرکت لیزینگ که طی سال‌های ۱۳۹۳ و ۱۳۹۴ تسهیلات دریافت کرده‌اند، مورد استفاده قرار گرفت. نتایج نشان می‌دهند که هریک از مشاهدات به طور قطعی در ردهٔ مشتریان خوش‌حساب و مشتریان بدحساب قرار گرفته و با ورود هر مشاهده جدید وضعیت اعتباری آن پیش‌بینی می‌شود.

***نویسنده مسئول:**

ایمیل: *alalinezhad@gmail.com*

DOI: 10.22056/ijir.2019.01.03

مقدمه

امروزه اکثر مؤسسات مالی اقدام به طراحی سیستمهای امتیازدهی اعتبار عینی و معتبری مبتنی بر الگوها و مدلهای علمی کرده‌اند. لیزینگ به‌عنوان یکی از مهم‌ترین روشهای تأمین اعتبار در سطح اقتصاد داخلی و بین‌المللی جایگاه خود را تثبیت کرده است و بخش عمده‌ای از عملیات مالی، تسهیلات و سرمایه‌گذاری و تأمین مالی در جهان، با استفاده از روشهای متنوع لیزینگ صورت می‌گیرد. رشد اقتصادی سال ۲۰۰۵، صنعت لیزینگ را به‌عنوان یک محصول مالی مکمل محصولات بانکی، معرفی کرده است. با توجه به اهمیت صنعت لیزینگ در دنیا، بحث شناسایی بازار لیزینگ و ایجاد روند بهبود در آن یکی از بحثهای مهم در این صنعت است.

در گذشته بسیاری از مؤسسات مالی، ریسک اعتباری[1] خود را فقط برای دوره‌ای طولانی و با توجه به ارزش اعتباری مشتری مدیریت می‌کردند. همچنین فرایند تصمیم‌گیری اعتباری بسیار ابتدایی بود و اغلب اوقات تصمیمها بر اساس اعطا و یا عدم اعطای تسهیلات قرار داشت. حتی اگر مشتریان ورشکسته هم می‌شدند، زیانهای آنها اغلب به‌وسیلهٔ وثیقه جبران می‌شد ولی بعد از مدتی هزینه‌های اعتباری در مؤسسات مالی به علت افزایش تعداد مشتریان ورشکسته و کاهش ارزش وثیقه‌ها به طور معنی‌داری افزایش یافت. به همین سبب کنترل ریسک اعتباری مهم‌ترین نیازمندی مدیریت مؤثر شناخته می‌شود. لذا توسعهٔ روشهای اعتبارسنجی و همچنین فرهنگ‌سازی در استفاده از این روشها بخشی از زیرساختهای موردنیاز مؤسسات مالی و اعتباری هستند. از طرف دیگر، عدم بازپرداخت به موقع تسهیلات، بیانگر این است که دریافت‌کنندهٔ تسهیلات در بهره‌برداری از آن موفقیت چندانی نداشته است. بنابراین تسهیلات مذکور به مطالبات معوق تبدیل می‌شوند. برای هر مؤسسهٔ اعطاکنندهٔ اعتبار و تسهیلات، توانایی تشخیص مشتریان بدحساب از خوش‌حساب بسیار مهم و ضروری است (طالبی و شیرزادی، ۱۳۹۰).

با توجه به توضیحات ذکرشده، در این مقاله به دنبال پاسخ به سئوالات زیر هستیم:

چگونه می‌توان با استفاده از روش تحلیل پوششی داده‌ها[2]/تحلیل ممیز[3] (DEA/DA)، یک جایگاه اعتباری (درجهٔ عضویت) برای مشتریان تعیین کرد؟

چگونه می‌توان با استفاده از روش تحلیل پوششی داده‌ها/تحلیل ممیز، یک تابع تمایز بین مشتریان خوش‌حساب و مشتریان بدحساب تعیین کرد؟

آیا روش تحلیل پوششی داده‌ها/تحلیل ممیز از کارایی لازم برای امتیازدهی اعتباری مشتریان در صنعت لیزینگ برخوردارست؟

اهمیت این مسئله باعث شده است تا مقالهٔ حاضر به امتیازدهی اعتباری مشتریان یکی از شرکتهای لیزینگ فعال کشور با استفاده از تحلیل پوششی داده‌ها و تحلیل ممیز در محیط فازی با درنظرگرفتن شاخصهای مؤثر در امتیازدهی مشتریان بپردازد.

مروری بر پیشینهٔ پژوهش

به رغم توسعهٔ روش تحلیل پوششی داده‌ها، تاکنون کمتر مطالعه‌ای به تعیین جایگاه اعتباری مشتریان در صنعت لیزینگ با استفاده از این روش پرداخته است. از این رو، به دلیل شباهت فعالیت صنعت لیزینگ با بخشی از صنعت بانکداری، مطالعات انجام‌شده در خصوص رتبه‌بندی اعتباری و بررسی عوامل مؤثر بر ریسک اعتباری مشتریان در بانکداری با استفاده از روش تحلیل پوششی داده‌ها و سایر روشهای موجود همچون تحلیل ممیزی، رگرسیون ساده و لجستیک و شبکه‌های عصبی مورد بررسی قرار گرفت که در ادامه به برخی از این مطالعات اشاره شده است:

در اواخر سال ۱۹۹۰، روش تحلیل پوششی داده‌ها برای تحلیل رتبه‌بندی اعتباری از سویترات[4] و همکاران، معرفی شد که این روش صرفاً نیازمند اطلاعات ثبتی پیشین یعنی مجموعهٔ مشاهده‌شدهٔ ورودیها و خروجیها برای محاسبهٔ رتبهٔ اعتباری مشتریان بود. بنابراین این امر افق جدیدی برای رتبه‌بندی اعتباری باز کرد (عریانی، ۱۳۸۴).

---

1. Credit Risk
2. Data Envelopment Analysis
3. Discriminant Analysis
4. Trout

مین و لی[1] (۲۰۰۸) در مقاله‌ای با هدف رتبه‌بندی اعتباری، از روش تحلیلی پوششی داده‌ها استفاده کردند. به این منظور محققان از داده‌های مالی حسابرسی‌شدهٔ تعدادی از شرکتهای تولیدی استفاده کردند. در این مقاله ورودیها، نسبتهای مالی مختلف در نظر گرفته شده است.

مارسین[2](۲۰۰۹) در مقاله‌ای تحت عنوان «کاربردی از تحلیل پوششی داده‌ها در رتبه‌بندی اعتباری» به رتبه‌بندی اعتباری با استفاده از روش تحلیل پوششی داده‌ها در پنج مرحله پرداخت. در ابتدا هفت نسبت مالی انتخاب و سپس ۱،۴۰۸ شرکت دارای اطلاعات کامل و مناسب از بین تعداد ۲۶۴،۱۰۳ شرکت انتخاب شدند. در مرحلهٔ بعد ماتریس همبستگی میان نسبتها محاسبه و با استفاده از مدل CCR کارایی شرکتها محاسبه و مشخص شد که فقط ۲۷ شرکت کاملاً کارا بوده‌اند. در این مقاله به منظور اعتباربخشی رتبه‌های حاصل از تحلیل پوششی داده‌ها، از روشهای تحلیل رگرسیون و تحلیل ممیزی استفاده شده است.

وربراکن[3] و همکاران (۲۰۱۴) به توسعهٔ مدل تعیین رتبهٔ اعتباری مصرف‌کنندگان با استفاده از پارامترهای مبتنی بر سود پرداختند. نتایج این مدل نشان می‌دهد که مدل ارئه‌شده در مجموعهٔ مورد آزمون، از نظر دقت اندازه‌گیری نتایج دقیق‌تر از سایر مدلهای موجود است و استقرار آن در سیستم سهل‌تر است.

نعمتی کوتنایی و همکاران (۲۰۱۵)، یک مدل داده‌کاوی ترکیبی برای امتیازدهی اعتباری گروه‌های یادگیری با استفاده از الگوریتمهای مختلف ارائه کردند. نتایج این مدل نشان داد که الگوریتم تحلیل مؤلفه‌های اصلی، بهترین الگوریتم انتخاب ویژگی است. نتایج طبقه‌بندی نیز نشان داد که دقت طبقه‌بندی در شبکه‌های عصبی تطبیقی بالاتر است.

عیسی‌زاده و عریانی (۱۳۸۹)، با هدف رتبه‌بندی مشتریان حقوقی بانک کشاورزی بر حسب ریسک اعتباری، به بررسی کاربرد تحلیل پوشش داده‌ها در این حوزه پرداختند. در این تحقیق با استفاده از روش نمونه‌گیری خوشه‌ای از مناطق شرق و غرب بانک کشاورزی استان تهران، ۲۸۶ شرکت تسهیلات‌گیرنده مورد بررسی قرار گرفته و پس از خارج‌کردن داده‌های نامناسب تنها ۷۵ شرکتی که با استفاده از عقد فروش اقساطی ۲۴ ماهه با سررسید پایان اردیبهشت سال ۱۳۸۴ تسهیلات دریافت کرده بودند، برای تحلیل نهایی مورد استفاده قرارگرفتند. درنهایت، با استفاده از روش تحلیل پوششی داده‌ها، کاراییهای فنی محاسبه و شرکتها رتبه‌بندی شدند. نتایج به‌دست‌آمده نشان داد که ۱۵ شرکت روی مرزکارایی قرار داشته و کاملاً کارا قلمداد شده‌اند. همچنین میانگین کارایی فنی معادل ۷۸ درصد بوده است، به این معنا که در مجموع شرکتهای مورد بررسی، ۲۲ درصد بیش از میزان مورد نیاز، ورودیها و عوامل تولید را مورد استفاده قرار داده و دارای سودآوری پایینی بوده‌اند.

شورورزی و همکاران (۱۳۹۱) به ارزیابی ریسک اعتباری مشتریان حقوقی بانک صادرات بر مبنای دو مدل شبکه‌های عصبی مصنوعی و لوجیت و مقایسهٔ آنها با هم پرداختند. در این پژوهش با استفاده از دو مدل رگرسیون لوژستیکی و شبکه‌های عصبی پرسپترون چند لایهٔ یک نمونه ۱۲۰تایی (۷۱ مشتری خوش‌حساب و ۴۹ مشتری بدحساب) که از بانک صادرات خراسان رضوی تسهیلات اعتباری دریافت کرده‌اند، بررسی شد. ابتدا ۲۶ متغیر مستقل شامل متغیرهای کیفی و مالی شناسایی و درنهایت ۱۳ متغیر را که دارای اثر معنی‌داری بر ریسک اعتباری و تفکیک بین دو رده از مشتریان بودند، انتخاب و مدل نهایی به‌وسیلهٔ آنها برازش شد. به منظور سنجش کارایی مدل شبکهٔ عصبی در مقایسهٔ با رگرسیون لوژستیکی، نتایج حاصل از رگرسیون لوژستیکی و شبکهٔ عصبی با یکدیگر مقایسه و بررسی نتایج نشان داد که هر دو مدل در ارزیابی ریسک اعتباری مشتریان حقوقی بانک از کارایی بالا و قابلیت مشابهی برخوردارند.

احمدیان (۱۳۹۲) به بررسی عوامل مؤثر بر تعیین رتبهٔ اعتباری مشتریان سیستم بانکی در ایران پرداخت. شاخصهای مالی، شاخصهای اعتباری، فاکتورهای مدیریتی صنعت، استمرار فعالیت و بازارهای هدف متغیرهایی است که در این پژوهش مورد بررسی قرار گرفت. یافته‌های پژوهش نشانگراین بود که شاخصهای مالی، وضعیت اعتباری مشتری نزد سیستم بانکی، فاکتورهای مدیریتی و استمرار فعالیت بر رتبهٔ اعتباری متقاضی اثرگذار است و همچنین بین صنعت و بازارهای هدف اصلی بنگاه و رتبهٔ اعتباری آن رابطهٔ معنی‌داری وجود ندارد.

علیزاده (۱۳۹۳) به منظور تحلیل بهتر مشتریان و شناسایی نیازهای بالفعل و بالقوه و تعیین اعتبار آنها، با استفاده از روشهای داده‌کاوی به ارائهٔ یک چارچوب بهینه به منظور افزایش دقت رتبه‌بندی ریسک اعتباری مشتریان بانکی پرداخت. در این تحقیق پس از بررسی برخی روشهای پرکاربرد داده‌کاوی برای رتبه‌بندی ریسک اعتباری مشتریان در بانکها و مؤسسات مالی و ساخت مدلهای مربوط به هریک از آنها، با ارائهٔ یک چارچوب بهینه در ترکیب این مدلها میزان دقت صحت رده‌بندی ارتقا یافت.

---

[1]. Min and Lee
[2]. Marcin
[3]. Verbraken

ریسک و انواع آن در صنعت لیزینگ

لیزینگ یکی از پیچیده‌ترین شیوه‌های تأمین مالی در جهان امروزی است و اتخاذ تصمیمهای مؤثر و موفق در این صنعت نیازمند تجربهٔ فراوان است. در طول مدت یک اجاره، ریسک و بازدهٔ آن و همچنین عوامل مؤثر بر این ریسک و بازده تغییر می‌کنند.

ریسکهای لیزینگ تحت سه دستهٔ کلی ریسک تجهیزات (ریسک عدم‌تطابق ارزش تجهیزات با میزان پیش‌بینی‌شده)، ریسک اعتباری (ریسک عدم‌تحقق سود مورد انتظار از مجرای درآمد اجاره‌بها) و ریسک مالیاتی (ریسک عدم‌تحقق منافع مالیاتی پیش‌بینی‌شده) و ریسک اعتباری (عدم‌اطمینان موجود در مورد اینکه آیا مستأجر، پرداختهای باقی‌ماندهٔ اجاره‌بهای خود را انجام خواهد داد یا خیر) تقسیم‌بندی خواهند شد.

در دنیای مدلهای آماری امتیازدهی اعتباری، مدلهای گوناگونی وجود دارد که هر کدام مشخصات، مزایا و معایب خاص خود را دارند. این مدلها به مدلهای قضاوتی و آماری تقسیم‌بندی و خود مدلهای آماری نیز به چند بخش زیر تقسیم می‌شوند:

روشهای آماری مبتنی بر نسبتهای حسابداری[1]، مدلهای ساختاری[2]، مدلهای کاهش‌یافته[3]، و مدلهای ترکیبی[4].

گفتن این مطلب که کدام‌یک از این مدلها توانایی بهتری در پیش‌بینی نکول دارند، دشوار است. هر کدام معایب و مزایای خاص خود را دارند و انتخاب هر کدام از این مدلها به شرایط تجاری متقاضیان و همچنین پورتفوی خاص لیزینگ بستگی دارد.

اما در این میان یکی از معروف‌ترین روشهایی که در سالهای اخیر مورد توجه بسیاری از صاحب‌نظران و تحلیلگران قرار گرفته، روش تحلیل پوششی داده‌هاست. این روش یک روش ناپارامتری و بر اساس روشهای برنامه‌ریزی خطی است که برای تعیین کارایی بنگاه‌های اقتصادی مورد استفاده قرار می‌گیرد.

روش‌شناسی پژوهش

در این مقاله با استفاده از روش تحلیل پوششی داده‌ها/تحلیل ممیز، به امتیازدهی اعتباری مشتریان می‌پردازیم. در روش تحلیل ممیز بر خلاف روش تحلیل پوششی داده‌ها، گروه‌ها از قبل مشخص هستند تا بر اساس آن پیش‌بینی شود که مشاهدات جدید به کدام رده تعلق دارند. به بیان دیگر، تحلیل ممیز روشی برای پیش‌بینی عضویت مشاهدهٔ نمونه‌گیری‌شده جدید در گروه‌های تعیین شده است.

ابتدا فرض می‌شود که اعضای دو گروه مشخص است. سپس یک مجموعه از وزنها و یک مقدار آستانه، برای یک ابرصفحه تعریف می‌شود که دو رده را از هم جدا می‌کند. از این ابرصفحهٔ جداکننده برای پیش‌بینی عضویت مشاهدات جدید می‌توان استفاده کرد. فرض کنیم n مشاهده هر کدام با k مشخصهٔ مستقل موجود هستند و z_{ij} عامل $i(=1,\ldots,k)$ ام از مشاهدهٔ $j(=1,\ldots,n)$ ام است. فرض می‌کنیم n_1 مشاهده در ردهٔ اول (G_1) و n_2 مشاهده در ردهٔ دوم (G_2) قرار دارد به‌نحوی که $n_1+n_2=n$ و $G_1 \cup G_2 = G$.

بر همین اساس، در غیاب هم‌پوشانی، می‌توان مجموعهٔ وزنهای α_i و یک مقدار آستانهٔ d برای ابرصفحه‌ای جداکننده بین دو رده به فرم $p^t z = d$ (p بردار نرمال) یافت به‌طوری که این ابرصفحه سعی می‌کند مشاهدات را به دو رده به صورت

$$\sum_{i=1}^{k} \alpha_i z_{ij} \geq d \quad , \quad j \in G_1 \qquad j \in G_1 = \left\{ j \middle| z_j = (x_j, y_j) \in G_1 \right\},$$

$$\sum_{i=1}^{k} \alpha_i z_{ij} < d \quad , \quad j \in G_2 \qquad j \in G_2 = \left\{ j \middle| z_j = (x_j, y_j) \in G_2 \right\},$$

افراز کند. یافتن ابرصفحه با روشهای مختلفی انجام می‌گیرد که در این تحقیق به دلیل اهمیت کاهش انحرافات داده‌ها، از روش مینیمم‌کردن مجموع انحرافات[5] (MSD) استفاده می‌کنیم.

---

1. Accounting – Based Statistical Methods
2. Structural Models
3. Reduced form Models
4. Hybrid Models
5. Minimizing the Sum of Deviations

با توجه به فرض خطی‌بودن ابرصفحهٔ جداساز در روش تحلیل ممیز، ابرصفحهٔ به‌دست‌آمده ممکن است دو رده را از هم جدا کند یا اینکه بعضی از مشاهدات ردهٔ G_1 را در ردهٔ G_2 و بعضی از مشاهدات ردهٔ G_2 را در ردهٔ G_1 طبقه‌بندی کند. تعدادی از مشاهدات که به طور صریح رده‌بندی نمی‌شوند، در یک مجموعه تحت عنوان هم‌پوشانی قرار می‌گیرند که این مجموعه نشان‌دهندهٔ طبقه‌بندی نادرست روش تحلیل ممیز است. در روش تحلیل ممیز، ابرصفحهٔ جداکننده توسط یک تابع خطی بیان شده است. چون روش تحلیل پوششی داده‌ها تشکیل ابرصفحه‌های قطعه‌ای خطی را می‌دهد، سوئیشی[1] (۱۹۹۹) از این فرض استفاده کرد و روشی جدید را ارائه کرد.

در این مقاله با استفاده از یک روش جدید تلفیقی از مدل جمعی تحلیل پوششی داده‌ها و تحلیل ممیز **Error! Bookmark not defined.** با معیار مینیمم‌کردن مجموع انحرافات، البته با حضور داده‌های فازی، یک فرایند محاسباتی دومرحله‌ای را در پیش می‌گیریم:

مرحلهٔ اول: رده‌بندی مجموعهٔ داده‌ها در دو ردهٔ G_1 و G_2 و یا هم‌پوشی $G_1 \cap G_2$؛

مرحلهٔ دوم: رده‌بندی مجدد داده‌های مشترک در هریک از رده‌های G_1 و G_2.

در غیاب هم‌پوشانی، الگوریتم فوق می‌تواند در فرایندی تک‌مرحله‌ای دو رده را جدا کند. هنگامی که بین دو رده هم‌پوشانی وجود داشته باشد، یک الگوریتم چندمرحله‌ای مورد نیاز است.

توسعهٔ مدل تحلیل پوششی داده‌ها/تحلیل ممیز

توسعهٔ مدل تحلیل پوششی داده‌ها/تحلیل ممیز، یک روش ناپارامتری دومرحله‌ای است که تابع جداسازی خطی (به جای ابرصفحهٔ جداساز) را ارائه می‌دهد و با هریک از مقادیر وزنی منفی و مثبت مرتبط با هر عامل مشاهده سروکار دارد.

رتبه‌بندی فازی اعداد مثلثی

یکی از کارآمدترین رویکردهای مرتب‌سازی مؤلفه‌های $F(\Re)$، تعریف تابع رتبه‌بندی $\tau : F(\Re) \to \Re$ است که هر عدد فازی را به صورت یک عدد قطعی ترسیم می‌کند. رتبه‌ها در $F(\Re)$ به صورت

$\tau(a) \geq \tau(b)$ اگر و فقط اگر $\tilde{a} \succeq \tilde{b}$,

$\tau(a) > \tau(b)$ اگر و فقط اگر $\tilde{a} \succ b$,

$\tau(a) = \tau(b)$ اگر و فقط اگر $\tilde{a} \approx \tilde{b}$,

تعریف می‌شوند، که در آن، $\tilde{a}\tilde{b}$ در $F(\Re)$ وجود دارد.

قضیه: اگر τ یک تابع رتبه‌بندی خطی در نظر گرفته شود، آنگاه:

۱. $\tilde{a} \succeq \tilde{b}$ اگر و فقط اگر $-\tilde{b} \succeq -\tilde{a}$؛

۲. $\tilde{a} \succeq \tilde{b}$ و $\tilde{c} - \tilde{d} \succeq 0$ اگر و فقط اگر $\tilde{a} + \tilde{c} \succeq \tilde{b} + \tilde{d}$.

در اینجا، تابع رتبه‌بندی τ در شرایط زیر صدق می‌کند:

$\tau(k\tilde{a} + \tilde{b}) = k\tau(\tilde{a}) + \tau(\tilde{b})$ به ازای هر $\tilde{a}$ و $\tilde{b}$ متعلق به $\tau(\Re)$، و هر $k \in \Re$ برقرار است.

به ازای عدد فازی $\tilde{a} = (\underline{a}(r), \overline{a}(r))$، تابع رتبه‌بندی

[1]. Sueyoshi

$$\tau(\tilde{a})=\frac{1}{2}\int_0^1\left(\underline{a}(r),\bar{a}(r)\right)dr,$$

به کار گرفته می‌شود.

$$\tau(\tilde{a})=\frac{1}{2}\left(a^m+\frac{1}{2}\left(a^l+a^u\right)\right)$$

این تابع به کاهش می‌یابد.

برای اعداد فازی مثلثی $\tilde{a}=\left(a^m,a^l,a^u\right)$ و $\tilde{b}=\left(b^m,b^l,b^u\right)$، خواهیم داشت:

$$\left(\tilde{a}\succeq\tilde{b}\right)\Leftrightarrow\left(a^m+\frac{1}{2}\left(a^l+a^u\right)\geq b^m+\frac{1}{2}\left(b^l+b^u\right)\right).$$

مدل تحقیق

فرض کنید n مشاهدهٔ فازی وجود دارد که توسط $\tilde{z}_j\ \ (j=1,\ldots,n)$ نمایش داده می‌شود و هر یک با k عامل مستقل به صورت $\tilde{z}_{ij}\ \ (i=1,\ldots,k)$ برای شاخص j ام نشان داده می‌شود. فرض بر آن است که مشاهدات در دو ردهٔ G_1 و G_2، به‌ترتیب هرکدام با n_1 و n_2 مشاهده رده‌بندی می‌شوند. $G_1\cup G_2=G$ و $n_1+n_2=n$ برقرار است. عضویت هریک از مشاهدات G معین است. هدف، یافتن قاعده‌ای برای رده‌بندی صحیح‌تر مشاهدات نادرست است. فرایند یافتن این قاعده در دومرحله صورت می‌گیرد. با تلاش برای حداقل‌سازی انحراف کل در مشاهداتی که به‌نادرستی رده‌بندی شده است، مرحلهٔ اول به تابع ناپارامتری دست می‌یابد که مشاهدات را مجدداً رده‌بندی می‌کند. مرحلهٔ دوم نحوهٔ تعیین عضویت مشاهداتی را مورد بحث و بررسی قرار می‌دهد که در مرحلهٔ اول با موفقیت رده‌بندی نشده بوده‌اند. از منظر ریاضی، برای $i=1,\ldots,k$ و $j=1,\ldots,n$ فرمول‌بندی این دو مرحله به صورت مینیمم‌کردن

$$\sum_{j\in G_1}s_{1j}^{+}+\sum_{j\in G_2}s_{2j}^{-},$$

با شرط‌های

$$\begin{aligned}
&\sum_{i=1}^{k}\left(\lambda_i^{+}-\lambda_i^{-}\right)\tilde{z}_{ij}+s_{1j}^{+}-s_{1j}^{-}\approx d+1, && j\in G_1,\\
&\sum_{i=1}^{k}\left(\lambda_i^{+}-\lambda_i^{-}\right)\tilde{z}_{ij}+s_{2j}^{+}-s_{2j}^{-}\approx d, && j\in G_2,\\
&\sum_{i=1}^{k}\left(\lambda_i^{+}+\lambda_i^{-}\right)=1,\\
&s_{1j}^{+},s_{1j}^{-},s_{2j}^{+},s_{2j}^{-}\geq 0,\\
&\lambda_i^{+},\lambda_i^{-}\geq 0,
\end{aligned}\qquad (۱)$$

بیان می‌شود که مدلی فازی است. به منظور استخراج مدل قطعی متناظر، تابع رتبه‌بندی به کار گرفته می‌شود. با توجه به ویژگی «$A\approx B$ اگر و فقط اگر $\tau(A)=\tau(B)$»، می‌توان دو شرط نخست مدل تحقیق در رابطهٔ (۱) را به صورت

$$\tau\left(\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\tilde{z}_{ij}+s_{1j}^+-s_{1j}^-\right)=\tau(d+1),\qquad j\in G_1,$$
$$\tau\left(\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\tilde{z}_{ij}+s_{2j}^+-s_{2j}^-\right)=\tau(d),\qquad j\in G_2,$$

تغییر داد، که با توجه به ویژگیهای تابع رتبه‌بندی (τ)، به صورت

$$\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\tau\left(\tilde{z}_{ij}\right)+s_{1j}^+-s_{1j}^-=d+1,\qquad j\in G_1,$$
$$\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\tau\left(\tilde{z}_{ij}\right)+s_{2j}^+-s_{2j}^-=d,\qquad j\in G_2,$$

به دست می‌آیند؛ بنابراین، با درنظرگرفتن متناظر محاسبه‌شده در تابع (τ)، فرمول‌بندی مدل نهایی به‌دست‌آمده در مرحلهٔ یک به صورت مینیمم‌کردن

$$\sum_{j\in G_1}s_{1j}^+ + \sum_{j\in G_2}s_{2j}^-,$$

با شرط‌های

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]+s_{1j}^+-s_{1j}^-=d+1,\qquad j\in G_1,\\
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]+s_{2j}^+-s_{2j}^-=d,\qquad j\in G_2,\\
&\sum_{i=1}^{k}\left(\lambda_i^++\lambda_i^-\right)=1,\\
&s_{1j}^+,s_{1j}^-,s_{2j}^+,s_{2j}^-\geq 0,\\
&\lambda_i^+,\lambda_i^-\geq 0,
\end{aligned}\qquad (۲)$$

بیان می‌شود. در اینجا، λ_i^* و d^* راه‌حلهای بهینهٔ مدل فوق به شمار می‌روند. اکنون با حل مدل (۲) اعضای ردهٔ G_1 و اعضای ردهٔ G_2 از طریق روابط

$$\begin{aligned}
&1.\quad \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^{+^*}-\lambda_i^{-^*}\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]\leq d^*,\qquad \Rightarrow z_m\in G_2,\\
&2.\quad \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^{+^*}-\lambda_i^{-^*}\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]\geq d^*+1,\ \Rightarrow z_m\in G_1,
\end{aligned}\qquad (۳)$$

شناسایی می‌شوند. در غیر این صورت، z_m به منطقهٔ هم‌پوشانی تعلق دارد. به منظور رده‌بندی z_m، مرحلهٔ دوم آغاز می‌شود. پیش از آغاز مرحلهٔ دوم، مجموعه‌های

$$R_1=\left\{j\in G:\frac{1}{2}\sum_{i=1}^{k}\lambda_i^*\left(z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right)\geq d^*+1\right\},$$
$$R_2=\left\{j\in G:\frac{1}{2}\sum_{i=1}^{k}\lambda_i^*\left(z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right)\leq d^*\right\},$$
$$R_0=\left\{j\in G:d^*<\frac{1}{2}\sum_{i=1}^{k}\lambda_i^*\left(z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right)<d^*+1\right\},\qquad(۴)$$
$$C_1=\left\{j\in R_1:j\in G_1\right\},\ C_2=\left\{j\in R_2:j\in G_2\right\},G_1'=G_1-C_1,\ G_2'=G_2-C_2,$$

تعریف می‌شوند؛ از این رو، فرمول‌بندی مرحلهٔ دوم به صورت مینیمم‌کردن $\sum_{j\in G_1}s_{1j}^+ + \sum_{j\in G_2}s_{2j}^-$ با شرط‌های

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]\geq d+1, && j\in C_1,\\
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]+s_{1j}^+-s_{1j}^-=c\ , && j\in G_1',\\
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]+s_{2j}^+-s_{2j}^-=c\ , && j\in G_2',\\
&\frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^+-\lambda_i^-\right)\left[z_{ij}^m+\frac{z_{ij}^l+z_{ij}^u}{2}\right]\leq d, && j\in C_2,\\
&\sum_{i=1}^{k}\left(\lambda_i^++\lambda_i^-\right)=1,\ d\leq c\leq d+1,\ s_{1j}^+,s_{1j}^-,s_{2j}^+,s_{2j}^-\geq 0\ ,\ \lambda_i^+,\lambda_i^-\geq 0,
\end{aligned}\qquad(۵)$$

بیان می‌شود. در اینجا، c^* و λ_i^* راه‌حلهای بهینهٔ به‌دست‌آمده از مرحلهٔ دوم در نظر گرفته می‌شوند؛ آنگاه:

$$\begin{aligned}
&1.\quad \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^{+^*}-\lambda_i^{-^*}\right)\left[z_{im}^m+\frac{z_{im}^l+z_{im}^u}{2}\right]\geq c^*,\quad \Rightarrow z_m\in G_1,\\
&2.\quad \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i^{+^*}-\lambda_i^{-^*}\right)\left[z_{im}^m+\frac{z_{im}^l+z_{im}^u}{2}\right]\leq c^*,\quad \Rightarrow z_m\in G_2,
\end{aligned}\qquad(۶)$$

متغیرهای تحقیق

باتوجه به اینکه در این تحقیق به امتیازدهی اعتباری مشتریان در شرکتهای لیزینگ می‌پردازیم و از آنجایی که بخش عمده‌ای از درخواست‌کنندگان دریافت تسهیلات از شرکتهای لیزینگ، اشخاص حقیقی هستند، تعداد ۱۷ متغیر اصلی و مؤثر بر ریسک اعتباری با استفاده از روش 6C انتخاب و مورد ارزیابی قرار گرفته‌اند. درنهایت، با توجه به معیارهای بررسی‌شده و قضاوت خبرگان از روش نظرسنجی دلفی، تعداد ۸ شاخص ($I_1,\ldots,I_8$) به ترتیب سن متقاضی، میزان مبلغ تسهیلات، نرخ سود تسهیلات، دورهٔ بازپرداخت تسهیلات، میزن مبلغ قسط ماهیانه، میزان مطالبات معوق و زمان دیربازپرداخت اقساط انتخاب و تعیین شد.

مقادیر داده‌ها

در این تحقیق، جامعهٔ آماری، مشتریان حقیقی شرکت لیزینگ هستند. به این منظور از میان ۳۶۰ مشتری حقیقی متقاضی اخذ تسهیلات، با استفاده از نمونه‌گیری تصادفی ساده، تعداد ۸۳ مشتری که طی سالهای ۱۳۹۳ و ۱۳۹۴ از این شرکت تسهیلات دریافت کرده‌اند، انتخاب شدند.

تبدیل داده‌های نامطلوب فازی به داده‌های مطلوب فازی

با توجه به حضور داده‌های نامطلوب در این تحقیق، برای مطلوب‌کردن آنها می‌توان از روشهای متفاوتی استفاده کرد که در روش انتخابی، معکوس کلیهٔ مقادیر عامل نامطلوب را با بیشترین مقدار (L,M,U) در بین واحدهای تصمیم‌گیرنده برای آن عامل جمع و سپس مقادیر نهایی را با مقداری ثابت جمع می‌کنیم.

اجرای مدل تحقیق

با توجه به داده‌های این تحقیق، بر اساس میانگین زمان دیربازپرداخت اقساط هر مشتری، مشتریان به دو ردهٔ ذیل رده‌بندی می‌شوند:

مشتریان خوش‌حساب (G_1): آن دسته از مشتریانی که مطالبات آنها در سرفصل مطالبات جاری قرار داشته و میانگین زمان دیربازپرداخت اقساط آنها از ۳۰ روز بیشتر نباشد.

مشتریان بدحساب (G_2): آن دسته از مشتریانی که مطالبات آنها در سرفصل غیرجاری بوده و یا میانگین زمان دیر بازپرداخت اقساط آنها بیش از ۳۰ روز باشد.

اکنون داده‌های جمع‌آوری شده در دو ردهٔ G_1 با تعداد ۵۹ مشاهده و ردهٔ G_2 با تعداد ۲۴ مشاهده که هریک دارای ۸ عامل مستقل ($I_1,\ldots,I_8$) هستند رده‌بندی شده‌اند. همچنین مقادیر z_{ij}^l، z_{ij}^m و z_{ij}^u به ازای $i=1,\ldots,8$ و $j=1,\ldots,83$ با هر مشاهده و بالطبع با هر عامل مشاهده مطابقت دارد.

با حل مدل، مقادیر بهینهٔ به‌دست‌آمده با استفاده از مدل مرحلهٔ یک (رابطهٔ ۲) در جدول ۱ نشان داده شده است.

جدول ۱: مقادیر λ_i^* محاسبه‌شده در مرحلهٔ اول

عامل	۱	۲	۳	۴	۵	۶	۷	۸
λ+*	۰	۰/۴۸۶۰	۰	۰/۰۰۴۳	۰	۰	۰	۰
λ-*	۰/۰۰۸	۰/۴۸۶۰	۰	۰	۰/۰۱۰۷	۰	۰	۰/۰۰۴۹

مقادیر بهینهٔ به‌دست‌آمده با استفاده از مدل مرحلهٔ یک، طبق جدول ۱ محاسبه و همچنین مقدار d^*=۴/۸۵ است. اکنون با استفاده از راه‌حلهای بهینهٔ مدل (رابطهٔ ۳) و دراختیارداشتن λ_i^* و d^*، مقادیر z_m را محاسبه می‌کنیم که مقادیر به‌دست‌آمده به شرح جدول ۲ هستند.

مطابق راه‌حلهای بهینهٔ ارائه‌شده، در صورتی که z_m به هریک از رده‌های G_1 و G_2 تعلق نداشته باشد به منطقهٔ هم‌پوشانی تعلق دارد. مانند مشاهدهٔ دهم که z_m آن معادل ۵/۴۱۷۲ است و از آنجایی که این عدد مابین مقدار عددی d^* و d^*+1 است، به منظور رده‌بندی باید وارد مرحلهٔ دوم شود.

جدول ۲: مقادیر z_m محاسبه‌شده در مرحلهٔ اول

مشاهده	z_m	مشاهده	z_m	مشاهده	z_m	مشاهده	z_m	مشاهده	z_m
۱	۵/۶۵۳۳	۱۸	۵/۸۳۵۳	۳۵	۶/۱۹۳۱	۵۲	۵/۸۷۸۵	۶۹	۴/۴۲۶۹
۲	۵/۹۷۷۳	۱۹	۵/۶۴۰۱	۳۶	۵/۵۶۳۵	۵۳	۵/۹۴۲۱	۷۰	۴/۴۸۰۷
۳	۵/۹۵۷۷	۲۰	۶/۰۲۹۰	۳۷	۶/۰۳۲۰	۵۴	۵/۸۵۱۵	۷۱	۴/۱۲۸۹
۴	۵/۷۶۶۵	۲۱	۵/۹۸۷۹	۳۸	۵/۷۸۷۴	۵۵	۵/۸۴۶۶	۷۲	۴/۶۶۹۱
۵	۵/۹۴۳۴	۲۲	۵/۶۸۷۶	۳۹	۵/۹۳۱۲	۵۶	۵/۹۱۷۴	۷۳	۴/۴۶۲۴
۶	۵/۶۹۷۵	۲۳	۵/۸۶۲۶	۴۰	۵/۷۵۸۸	۵۷	۶/۰۷۲۲	۷۴	۴/۶۶۲۱
۷	۵/۸۱۴۴	۲۴	۵/۷۹۳۰	۴۱	۵/۷۷۸۲	۵۸	۵/۴۸۵۴	۷۵	۴/۵۸۴۰
۸	۵/۷۴۴۹	۲۵	۵/۸۵۷۱	۴۲	۵/۹۱۴۹	۵۹	۵/۹۰۱۱	۷۶	۴/۶۷۰۵
۹	۵/۸۰۹۹	۲۶	۵/۷۶۲۸	۴۳	۵/۸۳۳۲	۶۰	۴/۲۲۱۸	۷۷	۴/۴۰۹۷
۱۰	۵/۴۱۷۲	۲۷	۶/۰۰۰۹	۴۴	۵/۹۵۱۷	۶۱	۴/۵۳۶۲	۷۸	۴/۶۵۸۷
۱۱	۵/۴۱۷۲	۲۸	۵/۷۸۷۰	۴۵	۵/۸۷۴۳	۶۲	۴/۴۵۷۴	۷۹	۴/۷۳۴۹
۱۲	۵/۴۱۷۲	۲۹	۵/۵۶۴۱	۴۶	۵/۸۶۹۳	۶۳	۴/۱۷۰۳	۸۰	۴/۷۵۵۳
۱۳	۵/۴۱۷۲	۳۰	۵/۸۲۶۶	۴۷	۵/۷۷۸۲	۶۴	۴/۵۸۶۹	۸۱	۵/۰۰۰۰
۱۴	۵/۴۱۷۲	۳۱	۶/۰۰۷۱	۴۸	۵/۸۴۲۸	۶۵	۴/۴۳۲۱	۸۲	۴/۶۴۰۵
۱۵	۵/۴۱۸۹	۳۲	۵/۹۵۳۲	۴۹	۵/۷۹۲۱	۶۶	۴/۴۸۰۵	۸۳	۴/۸۸۸۰
۱۶	۵/۹۳۸۶	۳۳	۵/۷۷۶۰	۵۰	۵/۸۱۰۶	۶۷	۴/۵۳۹۳	-	-
۱۷	۵/۸۹۸۵	۳۴	۵/۸۳۸۸	۵۱	۶/۰۹۴۷	۶۸	۴/۶۵۵۴	-	-

اکنون پیش از آغاز مرحلهٔ دوم می‌بایست مطابق مجموعه‌های تعریف‌شده در رابطهٔ (۴)، مشاهده‌های مربوط به هریک از مجموعه‌های R_1، R_2، R_0، C_1، C_2، G_1' و G_2' را مشخص و سپس با استفاده از رابطهٔ (۵)، مقادیر بهینه را محاسبه کنیم که مطابق جدول ۳ به دست می‌آیند.

جدول ۳: مقادیر λ_i^* محاسبه‌شده در مرحلهٔ دوم

عامل	۱	۲	۳	۴	۵	۶	۷	۸
λ+*	۰	۰/۴۶۵۶	۰	۰	۰	۰	۰	۰
λ-*	۰/۰۱۱۸	۰/۴۶۵۴	۰	۰/۰۱۶۸	۰/۰۳۳۹	۰	۰	۰/۰۰۶۷

در اینجا مقادیر c^* = ۰ و d^* = -۰/۲۲۲ است. اکنون با استفاده از راه‌حلهای بهینهٔ مدل در رابطهٔ (۶) و با درنظرگرفتن λ_i^* و c^*، مقادیر z_m مرحلهٔ دوم را محاسبه می‌کنیم که مقادیر به‌دست‌آمده به شرح جدول ۴ هستند.

جدول ۴: مقادیر z_m محاسبه‌شده در مرحلهٔ دوم

مشاهده	z_m	مشاهده	z_m	مشاهده	z_m	مشاهده	z_m	مشاهده	z_m
۱	۰/۴۴۵۱	۱۸	۰/۷۰۲۰	۳۵	۱/۱۹۷۹	۵۲	۰/۷۹۷۷	۶۹	-۰/۷۶۸۳
۲	۰/۸۷۲۸	۱۹	۰/۴۴۸۱	۳۶	۰/۰۵۶۸	۵۳	۰/۸۷۰۳	۷۰	-۰/۷۰۴۵
۳	۰/۸۳۷۳	۲۰	۰/۹۸۴۹	۳۷	۰/۹۶۶۶	۵۴	۰/۴۹۱۲	۷۱	-۱/۱۰۳۶
۴	۰/۶۲۴۳	۲۱	۰/۹۲۱۰	۳۸	۰/۵۶۷۸	۵۵	۰/۴۸۳۳	۷۲	-۰/۳۵۹۸
۵	۰/۸۰۴۸	۲۲	۰/۴۴۹۳	۳۹	۰/۸۰۷۲	۵۶	۰/۷۸۲۵	۷۳	-۰/۶۲۳۲
۶	۰/۴۸۱۱	۲۳	۰/۷۳۸۰	۴۰	۰/۶۱۵۷	۵۷	۱/۰۲۲۰	۷۴	-۰/۴۶۵۹
۷	۰/۶۳۷۵	۲۴	۰/۶۲۰۶	۴۱	۰/۶۷۱۲	۵۸	۰/۰۴۳۰	۷۵	-۰/۲۲۳۳
۸	۰/۵۴۷۵	۲۵	۰/۷۲۰۰	۴۲	۰/۸۳۷۷	۵۹	۰/۴۶۷۱	۷۶	-۰/۴۵۲۶
۹	۰/۶۵۰۶	۲۶	۰/۴۹۸۲	۴۳	۰/۷۴۶۲	۶۰	-۱/۰۵۲۲	۷۷	-۰/۸۰۳۵
۱۰	۰/۴۰۲۸	۲۷	۰/۹۰۷۵	۴۴	۰/۸۸۷۴	۶۱	-۰/۶۶۵۱	۷۸	-۰/۷۳۰۳
۱۱	۰/۴۰۲۸	۲۸	۰/۶۲۷۷	۴۵	۰/۸۱۶۳	۶۲	-۰/۷۳۳۸	۷۹	-۰/۶۱۵۶
۱۲	۰/۴۰۲۸	۲۹	۰/۳۰۱۹	۴۶	۰/۷۵۹۸	۶۳	-۱/۱۵۷۹	۸۰	-۰/۵۷۹۷
۱۳	۰/۴۰۲۸	۳۰	۰/۶۹۰۸	۴۷	۰/۶۷۱۲	۶۴	-۰/۵۸۴۱	۸۱	-۱/۱۹۲۹
۱۴	۰/۴۰۲۸	۳۱	۰/۷۲۶۵	۴۸	۰/۷۷۲۸	۶۵	-۰/۷۸۲۱	۸۲	-۰/۷۵۲۴
۱۵	۰/۴۱۵۴	۳۲	۰/۸۵۰۴	۴۹	۰/۶۷۶۱	۶۶	-۰/۱۹۷۸	۸۳	-۰/۳۸۱۶
۱۶	۰/۸۴۵۲	۳۳	۰/۶۵۱۰	۵۰	۰/۶۶۶۸	۶۷	-۰/۶۵۴۸	-	-
۱۷	۰/۷۸۹۸	۳۴	۰/۷۳۵۸	۵۱	۱/۱۰۲۰	۶۸	-۰/۴۸۷۶	-	-

همان‌گونه که مشهود است، هر یک از مشاهدات به طور قطعی در رده‌های G_1 و G_2 قرار گرفته‌اند و اکنون می‌توان با ورود مشاهدهٔ جدید و محاسبهٔ z_m مرحلهٔ اول و در صورت لزوم z_m مرحلهٔ دوم، وضعیت اعتباری آن را پیش‌بینی کرد. به عنوان مثال:
مشتری جدید اول با شاخصهای واقعی در جدول ۵ را در نظر بگیرید:

جدول ۵: شاخص‌های مربوط به مشتری جدید اول

عامل	۱	۲	۳	۴	۵	۶	۷	۸
L	۳۲	۴۵۵،۷۵۰،۰۰۰	۲،۰۰۰،۰۰۰،۰۰۰	۲۲	۳۶	۷۴،۳۹۸،۱۴۸	۷۶،۵۰۹،۸۳۳	۳۷
M	۳۲	۴۵۵،۷۵۰،۰۰۰	۲،۰۰۰،۰۰۰،۰۰۰	۲۲	۳۶	۷۴،۳۹۸،۱۴۸	۷۵،۹۱۰،۵۷۱	۲۷
U	۳۲	۴۵۵،۷۵۰،۰۰۰	۲،۰۰۰،۰۰۰،۰۰۰	۲۲	۳۶	۷۴،۳۹۸،۱۴۸	۵۴،۵۱۲،۲۹۳	۲

با محاسبهٔ مقدار z_m=۶/۲۱ مشاهده می‌شود که مقدار z_m از مقدار d^*+۱=۵/۸۵ بزرگتر است؛ لذا این مشتری در ردهٔ G_1 یا همان مشتریان خوش‌حساب قرار می‌گیرد.

مشتری جدید دوم با شاخصهای مربوط مورد بررسی قرار گرفت. با محاسبهٔ مقدار z_m=۴/۸۲ مشاهده شد که مقدار z_m از مقدار d^*=۴/۸۵ کوچکتر است؛ لذا این مشتری در ردهٔ G_2 یا همان مشتریان بدحساب قرار گرفت.

مشتری جدید سوم نیز با شاخصهای مربوط مورد بررسی و با محاسبهٔ مقدار z_m = ۵/۷۷ مشاهده شد که مقدار z_m بین مقادیر d^* = ۴/۸۵ و d^* +۱ = ۵/۸۵ قرار دارد؛ لذا این مشتری در منطقهٔ هم‌پوشانی قرار دارد. با اجرای مرحلهٔ دوم مدل و محاسبهٔ c =۰ و z_m = ۰/۰۷۴۷، از آنجایی که مقدار z_m از مقدار c بزرگتر است، نتیجه می‌گیریم که این مشتری در ردهٔ G_1 و جزو مشتریان خوش‌حساب رده‌بندی می‌شود.

همان‌گونه که در مثالهای فوق که البته در همهٔ آنها از داده‌های واقعی استفاده شده است مشاهده می‌شود، مدل ارائه‌شده توانایی دارد تا هریک از مشتریان جدید را در هریک از رده‌های تعریف‌شده (خوش‌حساب و بدحساب) رده‌بندی کند.

بررسی اعتبار مدل توسعهٔ تحلیل پوششی داده‌ها/ تحلیل ممیز

به منظور بررسی عملکرد و ارزیابی اعتبار مدل توسعهٔ تحلیل پوششی داده‌ها/تحلیل ممیز، سوئیشی (۲۰۰۶) در مقاله‌ای با استفاده از یک مجموعه از داده‌های واقعی ۱۰۰ بانک در کشور ژاپن و چندین مجموعه دادهٔ تصنعی، به مقایسهٔ عملکرد این مدل با شش روش تحلیل ممیز شامل دو روش تحلیل ممیز خطی (MSD و MSD دومرحله‌ای)، سه روش تحلیل ممیز غیرخطی (MSD لگاریتمی، MSD لگاریتمی دومرحله‌ای و MSD رتبهٔ دوم) و یک روش تحلیل ممیز ناپارامتری (تحلیل پوششی داده‌ها/تحلیل ممیز) پرداخت و نتایج زیر به دست آمد:

توسعهٔ تحلیل پوششی داده‌ها/تحلیل ممیز، به میزان ۸۹/۷٪ و در بخش داده‌های تصنعی به طور میانگین ۹۹/۳٪ رده‌بندی درست داده‌ها را در مقابل شش روش دیگر روش تحلیل ممیز انجام داده است. رده‌بندی درست به‌وسیلهٔ تعداد مشاهداتی که به‌درستی در هر رده از نمونه‌ها رده‌بندی شده، تعیین شده است.

توسعهٔ تحلیل پوششی داده‌ها/تحلیل ممیز به میزان ۸۸/۵٪ پیش‌بینی درست داده‌ها را در مقابل شش روش دیگر تحلیل ممیز داشته است. پیش‌بینی درست به‌وسیلهٔ تعداد مشاهداتی که به‌درستی در هر رده از نمونه‌های نگه‌داشته‌شده پیشگویی شده، اندازه‌گیری شده است.

توسعهٔ تحلیل پوششی داده‌ها/تحلیل ممیز، ۶۸۹ ثانیه زمان CPU را استفاده و محاسبات مربوط را انجام داده است و این روش بین سایر روشها بهترین زمان اجرا در پیش‌بینی درست و انجام محاسبات را داشته است.

بر اساس نتایج به‌دست‌آمده مذکور، می‌توان گفت که توسعهٔ تحلیل پوششی داده‌ها/تحلیل ممیز در مقابل شش روش دیگر تحلیل ممیز بهتر، دقیق‌تر و سریع‌تر عمل می‌کند. نتیجهٔ مهم دیگری که می‌توان گرفت این است که تمام روشهای پارامتری خطی و غیرخطی در بین این هفت روش، وضعیت نشدنی زیادی ایجاد می‌کنند. ضمن اینکه دو روش ناپارامتری تحلیل پوششی داده‌ها/ تحلیل ممیز و توسعهٔتحلیل پوششی داده‌ها/ تحلیل ممیز برای هر مجموعهٔ داده‌شده، یک مدل قابل حل هستند و این نتیجه یک مزیت روش ناپارامتری را نشان می‌دهد.

توسعهٔ تحلیل پوششی داده‌ها/ تحلیل ممیز از خطای مشخصه جلوگیری می‌کند؛ بنابراین می‌تواند هریک از مجموعه داده‌ها را پوشش دهد.

نتایج و بحث

در این تحقیق با استفاده از یک روش جدید تلفیقی، با روش تحلیل پوششی داده‌ها/ تحلیل ممیز و با معیار مینیمم‌کردن مجموع انحرافات و همچنین با حضور داده‌های فازی یک فرایند محاسباتی دومرحله‌ای در پیش گرفته شد. هرچند روشهای تلفیقی تحلیل پوششی داده‌ها/ تحلیل ممیز توسط سوئیشی و کریهارا[1] (۱۹۹۸) ارائه شده است، ولی مدل به‌کاررفته در این تحقیق مدل قدرتمندتری نسبت به مدلهای پیشین است. نتایج به‌دست‌آمده در مرحلهٔ اول قادر به رده‌بندی مشاهدات در هریک از رده‌های خوش‌حساب و بدحساب بوده و در صورت عدم تشخیص ردهٔ مناسب برای هر کدام از مشاهدات (هم‌پوشانی)، مرحلهٔ دوم قادر به رده‌بندی مشاهده به صورت دقیق است. با استفاده از نتایج این مدل می‌توان با ورود یک متقاضی جدید جایگاه آن را در هریک از رده‌های اعتباری تشخیص داده و در خصوص رفتار اعتباری مشتری جدید پیش‌بینی‌های لازم را اتخاذ کرد.

جمع‌بندی و پیشنهادها

پیشنهادهای کاربردی این تحقیق، عبارت‌اند از:

به‌کارگیری و تحلیل اطلاعات به‌دست‌آمده برای تدوین برنامه‌ها و تصمیم‌گیریهای مربوط برای مشتریان جدید؛

تغییر کاهشی مدت معیار جداسازی گروه‌ها و مقایسهٔ نتایج به‌دست‌آمده با نتایج فعلی برای بررسی ریسک اعتباری مشتری جدید؛

تعریف شاخصهای جدید (مالی و غیرمالی) متناسب با صنعت مورد مطالعه؛

تعیین و تثبیت شاخصها و متغیرهای مؤثر بر رتبه‌بندی متقاضیان حقوقی دریافت تسهیلات.

[1]. Kirihara

منابع و ماخذ

احمدیان، ع.ا.، (۱۳۹۲). عوامل مؤثر بر تعیین رتبهٔ اعتباری مشتریان سیستم بانکی در ایران. پنجمین کنفرانس نظام تأمین مالی در ایران، تهران.

شورورزی، م.ر، مسیح‌آبادی، ا. غیاثی شهرکی، م.، (۱۳۹۱). ارزیابی ریسک اعتباری مشتریان حقوقی بانک صادرات بر مبنای دو مدل شبکه‌های عصبی مصنوعی و لوجیت و مقایسهٔ آنها. اولین کنفرانس ملی حسابداری مدیریت مالی و سرمایه‌گذاری، گرگان.

عریانی، ب.، (۱۳۸۴). رتبه‌بندی ریسک اعتباری مشتریان حقوقی بانکها به روش تحلیل پوششی داده‌ها (مطالعهٔ موردی: شعب مختلف بانک کشاورزی استان تهران سال ۱۳۸۰). پایان‌نامهٔ کارشناسی ارشد، دانشگاه بوعلی سینا.

علیزاده، ف.، (۱۳۹۳)، ارائهٔ یک چارچوب بهینه به منظور افزایش دقت رتبه‌بندی ریسک اعتباری مشتریان بانکی با استفاده از روشهای داده‌کاوی. پایان‌نامهٔ کارشناسی ارشد، دانشگاه پیام نور، تهران.

عیسی‌زاده، س.، عریانی، ب.، (۱۳۸۹). رتبه‌بندی مشتریان حقوقی بانکها بر حسب ریسک اعتباری به روش تحلیل پوششی داده‌ها (مطالعهٔ موردی: شعب بانک کشاورزی). فصلنامهٔ سیاستها و پژوهشهای اقتصادی، سال هجدهم، شمارهٔ ۵۵، صص ۵۹-۸۶.

طالبی م. شیرزادی ن.، (۱۳۹۰). ریسک اعتباری: اندازه‌گیری و مدیریت. سازمان چاپ و انتشارات اوقاف. چاپ اول.

Marcin, T., (2009). Application of Data Envelopment Analysis in Credit Scoring. Master's Thesis in Financial Mathematics, Technical Report, Krzysztofik School.

Min, J.H.; Lee, Y.C., (2008). A practical approach to credit scoring. Expert Systems with Applications, 35(4), pp. 1762-1770.

Nemati Koutanaei, F.; Sajedi H.; Khanbabaei, M., (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. Journal of Retailing and Consumer Services, 27, pp. 11-23.

Sueyoshi, T., (1999). DEA-discriminant analysis in the view of goal programming. European Journal of Operational Research, 115(3), pp. 564-582.

Sueyoshi, T., (2006). DEA-discriminant analysis: methodological comparison among eight discriminant analysis approaches. European Journal of Operational Research, 169(1), pp. 247-272.

Sueyoshi, T.; Kirihara, Y., (1998). Efficiency measurement and strategic classification of Japanese banking institutions. International Journal of Systems Science, 29(11), pp. 1249-1263.

Verberaken, T.; Bravo, C.; Weber, R.; Baesence, B., (2014). Development and application of consumer credit scoring models using profit-based classification measures. European Journal of Operational Research, 238, pp. 505-513.