Industrial Management Perspective

Print ISSN: 2251-9874, Online ISSN: 2645-4165
Volume 11, Serial Number 43, Autumn 2021, Pages 45-67 (Original Article)
DOI: 10.52547/JIMP.11.3.45

Identifying the Impact of Fraud on Corporate Customers' Credit Scoring by Data Mining Approaches

Seyed Mahdi Sadat Rasoul*, Omid Mahdi Ebadati, Mahsa Sadat Bakhtiari*****

Abstract

Credit risk is one of the most important risks which banks and financial organizations face. It is related to unpaid and delayed installments. Banks evaluate their customers' credit in order to prevent this hazard. Development banks, which are the focus of this research, fund facilities based on working capital, so customers sometimes do fraud in declaring working capital. Considering fraud consequences and making a credit scoring model with sensitivity to fraud are the main aims of this research. The statistical population of this research includes companies who have referred to branches of an Iranian Bank. This research includes 55 financial and non-financial variables based on the credit scoring model. In the first step, fraudulent companies have been realized. Finally, in order to offer an optimized and sustainable model through merging machine learning methods and reporting performance evaluation indicators, the impacts of fraud have been considered.

Keywords: Credit Scoring; Credit Risk; Fraud; Working Capital; Data Mining.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

Received: Apr. 17, 2020; Accepted: Jun. 22, 2021.
* Assistant Professor, Kharazmi University (Corresponding Author).
Email: msadatrasoul@khu.ac.ir
** Assistant Professor, Kharazmi University.
*** M.Sc., Kharazmi University.

چشم‌انداز مدیریت صنعتی
شاپای چاپی: ۹۸۷۴-۲۲۵۱، شاپای الکترونیکی: ۴۱۶۵-۲۶۴۵
سال یازدهم، شماره ۴۳، پاییز ۱۴۰۰، ص‌ص ۴۵ – ۶۷ (**نوع مقاله: پژوهشی**)
DOI: 10.52547/JIMP.11.3.45

شناسایی آثار دست‌کاری احتمالی در داده‌های اعتبارسنجی بر مدل‌های اعتبارسنجی مشتریان حقوقی با استفاده از داده‌کاوی

سیدمهدی سادات‌رسول*، امید مهدی عبادتی، مهساسادات بختیاری*****

چکیده

بانک‌ها برای پیشگیری از زیان‌های ناشی از ریسک اعتباری به دلیل عدم بازپرداخت اقساط تسهیلات بانکی، به اعتبارسنجی مشتریان خود می‌پردازند. بانک‌های توسعه‌ای به‌دلیل اعطای تسهیلات با شرایط تسهیل‌شده در نرخ، وثیقه و بازه زمانی بلندمدت تسهیلات صرفاً بر اساس میزان سرمایه در گردش موردنیاز مشتریان، به آن‌ها تسهیلات اعطا می‌کنند. این موضوع باعث می‌شود تا برخی از مشتریان اقداماتی را در صورت‌های مالی و سایر داده‌های خود صورت دهند که بتوانند از تسهیلات بیشتر از نیاز خود برخوردار شوند. هدف پژوهش حاضر، بررسی آثار این نوع تقلب بر مدل‌های اعتبارسنجی و ایجاد یک مدل اعتبارسنجی پایدار و حساس به تقلب است. جامعه آماری این پژوهش شامل شرکت‌هایی است که به‌منظور اخذ تسهیلات به شعب بانک توسعه صادرات مراجعه کرده‌اند. متغیرهای این پژوهش شامل ۵۵ متغیر مالی و غیرمالی است که بر اساس آن‌ها مدل اعتبارسنجی ایجادشده و به کارگرفته شده است. ابتدا شرکت‌هایی که احتمالاً در صورت‌های مالی دست‌کاری احتمالی در سرمایه در گردش داشتند، شناسایی و برچسب‌گذاری شدند و درنهایت با استفاده از ترکیبی از روش‌های داده‌کاوی و گزارش شاخص‌های ارزیابی عملکرد به بررسی آثار این نوع تقلب بر مدل‌های اعتبارسنجی پرداخته شد تا بتوان به مدل بهینه و پایدار یک بانک توسعه‌ای دست یافت.

کلیدواژه‌ها: اعتبارسنجی؛ ریسک اعتباری؛ تقلب؛ سرمایه در گردش؛ داده‌کاوی.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

تاریخ دریافت مقاله: ۱۳۹۹/۰۱/۲۹، تاریخ پذیرش مقاله: ۱۴۰۰/۰۴/۰۱.

* استادیار، دانشگاه خوارزمی (نویسنده مسئول).
Email: msadatrasoul@khu.ac.ir
** استادیار، دانشگاه خوارزمی.
*** کارشناسی ارشد، دانشگاه خوارزمی.

۱. مقدمه

فرایند جهانی‌شدن و به‌تبع آن تشدید رقابت مؤسسه‌های مالی و اقتصادی در بازارهای محلی و جهانی، ضرورت تقویت و اصلاحات سیستمی را در بنگاه‌های مالی و سرمایه‌ای در سراسر دنیا افزایش می‌دهد. به دلیل اینکه سودآوری و متناسب‌سازی آن با ریسک یکی از مهم‌ترین اهداف در فعالیت بنگاه‌های اقتصادی محسوب می‌شود و ارتباط تنگاتنگی با میزان انواع ریسک فعالیت‌های متنوع مالی و بانکی دارد، مدیریت ریسک اعتباری به‌عنوان یک اولویت در اصلاح و بهینه‌سازی سازوکارهای بانکی در تمام دنیا به‌شمار می‌رود [۳، ۲۳].

بنگاه‌های اقتصادی عمدتاً تمایل دارند برای تأمین بخشی از سرمایه در گردش موردنظر خود از تسهیلات بانک‌ها استفاده کنند و در این میان می‌کوشند تا بخش عمده‌ای از این تسهیلات را از بانک‌های توسعه‌ای (در ایران مانند بانک توسعه صادرات، بانک صنعت و معدن و بانک کشاورزی) دریافت کنند [۳، ۱۷،۲۹]. از وجوه تمایز بانک‌های توسعه‌ای، به‌خصوص در ایران، با سایر بانک‌های تجاری (مانند بانک ملت، پاسارگاد و غیره)، پایین‌بودن نرخ تسهیلات، طولانی‌تر-بودن دوره بازپرداخت، درنظرگرفتن شاخص‌هایی مانند اشتغال ایجادشده و رفع نیاز کشور در ارزیابی طرح‌ها و وثایق تسهیل‌شده آن‌ها نسبت به بانک‌های تجاری است. درواقع بانک‌های توسعه‌ای، بانک‌هایی هستند که به‌منظور تسریع و تسهیل توسعه اقتصادی یک کشور، از طریق تجهیز و تخصیص منابع میان‌مدت و بلندمدت و تشویق و ترغیب بخش خصوصی در کمک به توسعه اقتصادی، نقش‌آفرینی می‌کنند. وجه تمایز این بانک‌ها با بانک‌های تجاری در این است که بانک‌های توسعه‌ای معمولاً در طرح‌های دارای بازدهی و ریسک بالاتر و یا گاهاً پایین‌تر (منظور هر دو شاخص به‌صورت توأم است) و دارای دوره بازدهی طولانی اقدام به اعطای تسهیلات و حمایت می‌کنند که بانک‌های تجاری تمایل چندانی به این نوع تأمین مالی ندارند [۲، ۱۷].

به دلیل اینکه عمدتاً اخذ تسهیلات بانک‌های توسعه‌ای آسان‌تر است، این بانک‌ها اقدام به محاسبه میزان تسهیلات موردنیاز بنگاه‌های اقتصادی به‌صورت دقیق می‌کنند تا به میزان موردنیاز تسهیلات در اختیار آن‌ها قرار داده و مابقی منابع محدود خود را در اختیار سایر بنگاه‌های اقتصادی کشور که به آن نیاز دارند، قرار دهند. گاهی بنگاه‌های اقتصادی در زمان دریافت تسهیلات تمایل دارند تا با تغییر داده‌های مالی و غیرمالی میزان سرمایه در گردش موردنیاز سازمان را بیشتر از میزان واقعی به بانک گزارش کنند؛ درنتیجه برای رسیدن به بیشترین سرمایه در گردش موردنیاز، گاهی اقدام به دست‌کاری در بعضی از اقلام اثرگذار صورت‌های مالی و اقلام قراردادی اثرگذار بر تعیین سرمایه در گردش می‌کنند [۸]. این تغییرات عمدتاً به دریافت تسهیلات بیش از نیاز بنگاه منجر شده و مبلغ اضافه بر نیاز، صرف فعالیت‌های اقتصادی غیرمولد می‌شود؛ همچنین به‌دلیل محدودبودن منابع مالی بانک‌های توسعه‌ای باعث می‌شود تا این بانک

ها نتوانند به کلیه بنگاه‌های اقتصادی که برای رشد به این منابع نیاز دارند، منابع لازم موردنیازشان برای رشد را تخصیص دهند و از این طریق کشور دچار تأخیر در فرایند رشد می‌شود. ورود به فعالیت‌های غیر مولد عمدتاً با بالابودن ذخایر و تسهیلات اعطایی سوخت‌شده و یا معوق بانک‌های کشور به دلایل مختلف ازجمله عدم‌تمایل به بازپرداخت به بانک توسعه‌ای به دلیل نرخ ترجیحی، خروج مازاد منابع موردنیاز از کشور ایران و غیره همراه می‌شود که این امر گویای نبود مدل‌های مناسب اندازه‌گیری ریسک اعتباری و سیستم‌های مدیریت ریسک در شبکه نظام بانکی توسعه‌ای است [۱۸].

از عمده دلایل این امر می‌توان به عواملی همچون مسائل فرهنگی، اقتصادی، آموزش، فقدان بانک‌های اطلاعاتی متمرکز، خلأ شبکه تبادل اطلاعات قوی و کارآمد، عدم‌وضع قوانین و مقررات کافی و مسائل سیاسی اشاره کرد [۱۸]؛ بنابراین باید تدابیری اندیشیده شود تا هم منابع مالی موردنیاز متقاضیان تأمین شود و هم بانک اصلی‌ترین وظیفه خود، یعنی اعطای تسهیلات را با حداقل ریسک ممکن انجام دهد؛ زیرا در شرایط متحول امروز، موفقیت هر بنگاه به تسلط آن بر ریسک‌ها و نوع مدیریتی که بر انواع ریسک‌ها اعمال می‌کند، بستگی دارد [۸].

در سال‌های اخیر به‌دلیل رقابت شدید در صنعت اعتباری، اعتبارسنجی[1] به یکی از موضوعات حیاتی و چالشی تبدیل شده است که بنا بر نیاز بانک‌ها مدل‌های متنوعی نیز برای آن ایجاد و توسعه داده شده است. اغلب مدیران اعتبارسنجی بر اساس تجربه‌های خود به ارزیابی اعتبار متقاضیان می‌پردازند که باعث افزایش ریسک اعتباری می‌شود [۱۸]. درواقع بحران‌های مشاهده‌شده در نظام بانکی کشورها عمدتاً ناشی از ناکارایی در مدیریت ریسک اعتباری است [۴]. حجم زیادی از تسهیلات اعطایی سوخت‌شده یا معوقه بانک‌ها، گویای فقدان مدل‌های مناسب اندازه‌گیری ریسک اعتباری در شبکه بانکی است و یکی از مهم‌ترین ابزارهایی که بانک‌ها برای مدیریت و کنترل ریسک اعتباری به آن نیاز دارند، سیستم امتیازدهی اعتباری مشتریان به‌وسیله یک سیستم خبره است و می‌توان آن را گامی بزرگ در جهت تحقق بانکداری الکترونیکی دانست که یکی از مهم‌ترین کارها در تحقق‌بخشیدن به اهداف شهر الکترونیک است [۲۲].

اعتبارسنجی مشتریان ازجمله موضوع‌های مهمی است که در دنیای امروز موردبحث و توجه بانک‌ها قرار گرفته است. در سیستم‌های بانکداری سنتی، مدیران اعتباری اغلب میزان اعتبار مشتریان را با توجه به تجربه خود می‌سنجند؛ ولی نظام بانکداری نوین، با محدودیت زمانی و تعداد فزاینده مشتریان مواجه است [۴]. یکی از اقداماتی که موجب اشتباه در اعتبارسنجی

1. Credit Scoring

مشتریان می‌شود، تقلب است. تقلب بدین معنا که مشتریان با دست‌کاری اطلاعات خود، سعی می‌کنند سرمایه در گردش بیشتری را دریافت کنند.

به‌طورکلی تقلب به سه شکل زیر انجام می‌شود [۱]:

۱. فساد مالی: به‌عنوان تقلبی تعریف می‌شود که در آن متقلبان از نفوذ خود در یک تراکنش مالی به‌منظور کارپردازی برای منافع شخصی خود یا شخصی دیگر، به‌غلط استفاده می‌کنند؛ مانند پذیرش پورسانت و درگیرشدن در تضاد منافع، رشوه‌خواری (مانند حق و حساب و دست‌کاری قیمت‌ها در مزایده و مناقصه)، باج‌گیری و زورگیری اقتصادی.

۲. سوءاستفاده از دارایی‌ها: شامل دزدی یا استفاده نادرست از دارایی‌های یک سازمان است؛ مانند ترفندهای نقددزدی، کالادزدی، دزدی سایر دارایی‌ها و استفاده غیرقانونی از دارایی‌ها به‌عنوان دارایی شخصی.

۳. تقلب گزارشگری مالی: عبارت است از: تحریف عمدی در نتایج صورت‌های مالی برای ارائه تصویر نادرست از شرکت؛ مانند بیش‌نمایی دارایی‌ها و کم‌نمایی هزینه‌ها، بیش‌نمایی درآمدها و بیش‌نمایی دارایی‌ها، کم‌نمایی دارایی‌های دزدیده‌شده.

به همین دلیل کاهش و کنترل ریسک به‌عنوان یکی از عوامل مؤثر بر بهبود فرآیند اعطای اعتبار[1] بر عملکرد بانک‌ها مطرح است و نقش اساسی در تداوم ارائه تسهیلات و بقای بانک و مؤسسه‌های مالی دارد [۴]. آنچه برای بانک اهمیت دارد، این است که قبل از اعطای تسهیلات، در درجه اول از هرگونه تقلب[2] توسط درخواست‌کننده‌های تسهیلات جلوگیری کند و در مرحله دوم احتمال عدم‌بازپرداخت از سوی وام‌گیرندگان را ارزیابی کرده و گروهی را انتخاب کند که از نظر بررسی صورت مالی و ترازنامه در صحت کامل هستند و هم از ادای دین آن‌ها در موعد مقرر مطمئن است. انجام این کار از طریق یک سیستم جامع، ساختار و معیار مناسب امکان‌پذیر است [۴].

با توجه مطالب عنوان‌شده، در این پژوهش که با استفاده از داده های بانک اطلاعاتی یکی از بانک های توسعه ایی کشور جمع آوری شده است، تلاش می‌شود با توجه به شکافی که در پژوهش‌های پیشین در شناسایی این نوع تقلب وجود داشته است و همچنین نیاز کشور ایران به بررسی دقیق‌تر این نوع تقلب، آثار تقلب بر مدل‌های اعتبارسنجی بانک‌های توسعه‌ای به‌منظور اخذ تسهیلات با شرایط تسهیل‌شده شناسایی شود. درنهایت این پژوهش به دنبال ایجاد مدل‌های اعتبارسنجی مناسب و حساس نسبت به تقلب با درنظرگرفتن متغیرهای قابل‌تغییر در اعتبارسنجی برای بانک‌های توسعه‌ای، به‌خصوص بانک توسعه صادرات است. نتایج پژوهش نشان می دهد که مدل‌های اعتبارسنجی مورداستفاده در بانک‌های تجاری نمی‌توانند برای

1. Credit
2. Fraud

بانک‌های توسعه‌ای مورداستفاده قرار گیرند و بانک‌های توسعه‌ای باید مدل‌های اعتبارسنجی خاص خود را با درنظرگرفتن اثر احتمالی دست‌کاری و تقلب ایجاد کنند که این مدل شامل متغیرهای جدید ارزیابی اعتباری، ازجمله متغیرهای سرمایه در گردش موردنیاز دارایی جاری منهای بدهی جاری، مجموع اسناد پرداختی تجاری و مجموعه اسناد دریافتنی تجاری است که باید توسط الگوریتم‌های جدید آموزش و توسعه داده شده باشند.

۲. مبانی نظری و پیشینه پژوهش

علیزاده‌فرد و همکاران (۱۳۹۵)، یک مدل ترکیبی برای ارزیابی اعتبار مشتریان بانک با استفاده از الگوریتم‌های داده‌کاوی ارائه کردند [۳]. در این پژوهش ۹ مدل برای امتیازدهی اعتبار با ترکیب روش‌های انتخاب ویژگی و کاهش بعد با الگوریتم‌های رده‌بندی ایجاد شد. مدل‌ها بر روی یک مجموعه داده که از UCI انتخاب شده بودند (مجموعه داده آلمان) اجرا شدند و سپس دقت آن‌ها مقایسه شد. سه تکنیک رده‌بندی داده‌کاوی استفاده‌شده عبارت‌اند از: درخت تصمیم، نزدیک‌ترین همسایگی و بیز ساده. همچنین از تحلیل مؤلفه‌های اصلی[1] (PCA) و وزن‌دهی متغیرها با استفاده از شاخص جینی به‌عنوان تکنیک‌های کاهش بعد و انتخاب ویژگی استفاده شده است. نتایج نشان داد که ترکیب استفاده از وزن‌دهی متغیرها با استفاده از شاخص جینی با الگوریتم بیز ساده نتیجه بهتری را نسبت به ترکیب روش‌های دیگر به‌دنبال دارد.

پیرایش و همکاران (۱۳۹۵)، به اعتبارسنجی مشتریان حقیقی در «بانک صادرات» استان زنجان با روش رگرسیون لجستیک پرداختند. نتایج نشان داد که بین جنسیت، تأهل، سن، تحصیلات، مدت وام، مبلغ تسهیلات، نوع وثیقه، نوع کسب‌وکار (محل مصرف وام) و نوع تسهیلات پرداختی با وضعیت اعتباری مشتریان، رابطه معناداری وجود دارد؛ همچنین نتایج حاصل از مدل رگرسیونی نشان داد که جنسیت، تأهل، سن، تحصیلات، مدت وام، مبلغ تسهیلات، وثیقه، نوع کسب‌وکار (محل مصرف وام) و نوع تسهیلات پرداختی بر وضعیت اعتباری مشتریان تأثیر دارند [۲۳].

حاجی‌زاده و طورانی (۱۳۹۶)، به داده‌کاوی در حوزه بانکداری با نگرشی بر اعتبارسنجی مشتریان پرداختند. سه روش داده‌کاوی استفاده‌شده در پژوهش شامل کارت امتیازی اعتباری، درخت تصمیم‌گیری و رگرسیون لجستیک بود [۹]. طلوعی و همکاران (۱۳۸۸)، به امتیازدهی اعتباری متقاضیان کارت‌های اعتباری پرداختند [۲۲]. در این پژوهش مدل ترکیبی ماشین بردار پشتیبان مبتنی بر دو استراتژی زیر برای امتیازدهی اعتباری متقاضیان کارت‌های اعتباری به‌کار رفته است:

1. Principal Copmonent Analysis (PCA)

۱. استفاده از تکنیک F-Score برای انتخاب ویژگی‌ها؛
۲. استفاده از تکنیک Grid search برای بهینه‌کردن مقادیر پارامترها.

تهرانی و فلاح شمس (۱۳۸۴)، در طراحی و تبیین مدل ریسک اعتباری در نظام بانکی کشور با استفاده از داده‌های اعتباری ۳۱۶ مشتری حقوقی بانک‌های کشور و با استفاده از مدل‌های احتمالی خطی، لجستیک و شبکه‌های عصبی مصنوعی به طراحی و آزمون کارایی مدل ریسک اعتباری پرداختند. نتایج نشان داد که ارتباط بین متغیرها در مدل پیش‌بینی ریسک اعتباری به‌صورت خطی نیست و تابع‌های نمایی و سیگموید، مناسب‌ترین مدل‌های پیش‌بینی ریسک اعتباری هستند و بیشترین کارایی برای پیش‌بینی ریسک اعتباری به‌ترتیب مربوط به شبکه‌های عصبی مصنوعی و مدل لجستیک است [۲۸].

کاشانی و همکاران (۱۳۸۴)، به تشخیص تقلب در بانکداری الکترونیکی با استفاده از داده‌کاوی پرداختند [۳۱]. در این پژوهش، روش‌های داده‌کاوی به‌منظور تشخیص تقلب در بانکداری الکترونیک موردبررسی قرار گرفت. بدین منظور از یک روش ترکیبی شامل خوشه‌بندی برای تفکیک مشتریان و رده‌بندی به‌منظور ساخت مدلی برای کشف تقلب استفاده شد. اسماعیلی و نورانی (۱۳۹۶)، ارائه یک چارچوب داده‌کاوی کارآمد برای پیش‌بینی رفتار مشتری را بررسی کردند؛ همچنین ارتباط بین سازمان‌ها و مشتریان را در دنیای کسب‌وکار پیشرفته‌ی امروزی مدیریت می‌کند. محمدی والا و فرهودی‌نژاد (۱۳۹۴)، به شناسایی تقلب در تراکنش‌های بانکی با استفاده از داده‌کاوی (مطالعه موردی شناسایی در تراکنش‌های بانک مهر اقتصاد) پرداختند [۹].

کیم و سون[1] (۲۰۰۴)، پژوهشی با عنوان «مدل ارزیابی ریسک اعتباری برای مدیریت تقاضای وام با استفاده از الگوهای کلاس‌بندی اشتباه» ارائه کردند [۶]. در این پژوهش بعد از انجام پیش‌بینی با استفاده از مدل شبکه عصبی پرسپترون چندلایه با استفاده از الگوریتم یادگیری پساانتشار خطا، از ماتریس اغتشاش[2] استفاده شد و گروه‌های خوب و بد از متقاضیان به چهار زیرگروه تقسیم شدند. شین و همکاران[3] (۲۰۰۵)، با به‌کارگیری ماشین بردار پشتیبان، مدلی را برای پیش‌بینی ورشکستگی ارائه کردند [۲۷]. لیو و همکاران[4] (۲۰۰۵)، ویژگی‌های برتر برای امتیازدادن به اعتبار مدل‌ها را بررسی کردند [۱۴]. کیرکاس و همکاران[5] (۲۰۰۷)، کارایی درخت‌های تصمیم، شبکه‌های عصبی و شبکه‌های بیزی در ردیابی اظهارنامه‌های مالی متقلبانه را بررسی کردند و نشان دادند که مدل شبکه بیزی بهترین نتیجه را دارا است [۱۲]. این پژوهش به بررسی تجربی ۴ مدل انتخاب خصوصیت یادگیری ماشین می‌پردازد. این مدل‌ها داده‌کاوی

1. Kim & Sohn
2. Confusion Matrix
3. Shin, et al.
4. Liu, et al
5. Kirkos, et al

خودکار را برای کاهش فضای خصوصیت انجام می‌دهند. با پیشرفت فناوری اطلاعات و بهبود کانال‌های ارتباطی، تقلب در سرتاسر جهان گسترش یافته و منجر به ضررهای مالی گسترده‌ای شده است. اگرچه سازوکارهای جلوگیری از تقلب نظیر CHIP & PIN ایجاد شده‌اند، اما این سازوکارها از بیشتر انواع تقلب، نظیر استفاده از کارت اعتباری تقلبی در پایانه‌های فروش یا سفارش‌های مکاتباتی، جلوگیری نمی‌کنند؛ درنتیجه ردیابی تقلب به ابزاری اساسی و شاید به بهترین روش برای جلوگیری از چنین تقلب‌هایی تبدیل شده است.

نگای و همکاران[1] (۲۰۱۱)، استفاده از تکنیک‌های داده‌کاوی در شناسایی تقلب مالی را بررسی کردند. آن‌ها به بررسی و طبقه‌بندی آثار موجود در زمینه استفاده از روش‌های داده‌کاوی برای ردیابی تقلب مالی پرداختند [۲۲]. در این پژوهش نخستین اثر علمی سیستماتیک، جامع و قابل‌شناسایی در زمینه‌ی روش‌های داده‌کاوی در FFD ارائه شده است. راویسانکار و همکاران[2] (۲۰۱۱)، به بررسی تشخیص تقلب صورت‌حساب مالی و انتخاب ویژگی پرداختند [۲۳]. لی و قربانی[3] (۲۰۱۲) از روش‌های داده‌کاوی، نظیر شبکه عصبی فید فروارد چندلایه، ماشین‌های بردار پشتیبان، برنامه‌نویسی ژنتیک، روش گروهی کار با داده‌ها، رگرسیون لجستیک و شبکه عصبی احتمالاتی برای شناسایی شرکت‌هایی استفاده کرند که دست به تقلب در اظهارنامه‌های مالی زده‌اند. لی و قربانی (۲۰۱۲) به بررسی شبکه‌های بی‌سیم و تشخیص تقلب بهبودیافته پرداختند [۱۳] در این پژوهش، دو الگوریتم خوشه‌بندی جدید به نام‌های «شبکه یادگیری رقابتی بهبودیافته» و «شبکه یادگیری رقابتی بهبودیافته با نظارت برای ردیابی تقلب و ردیابی نفوذ در شبکه» پیشنهاد شده است. ساهین[4] و همکاران (۲۰۱۳)، به بررسی روشی در مورد تصمیم‌گیری هزینه‌ای برای شناسایی تقلب پرداختند [۲۶]. شاهین و دومان (۲۰۱۳)، به تشخیص تقلب کارت اعتباری توسط درخت تصمیم و ماشین بردار پشتیبان پرداختند [۲۶].

عبدالحمید[5] و همکاران (۲۰۱۴)، به تشخیص خودکار خطای بانک با استفاده از ماشین‌های بردار پشتیبانی پرداختند. آن‌ها تلاش کردند تا اشکال مختلف تقلب و ابزارهای داده‌کاوی را تحلیل کنند و درنهایت استفاده از روش‌های یادگیری ماشین به نام «ماشین‌های بردار پشتیبان»، به‌منظور ساخت مدل‌هایی برای نمایش رفتار عادی و غیرعادی مشتریان و سپس استفاده از آن برای بررسی تراکنش‌های جدید را پیشنهاد کردند [۱]. بائی و کیم[6] (۲۰۱۵)، یک مدل پیش‌بینی اعتباری شخصی با استفاده از داده‌کاوی در محیط‌های هوشمند ارائه کردند [۶]. در این پژوهش روشی به نام «داده‌کاوی فراگیر هوشمند» ارائه شده است که مدل‌های همگن را در یک محیط

1. Ngai, et al
2. Ravisankar et al
3. Lei and Ghorbani
4. Sahin
5. Abdelhamid
6. Bae and kim

رایانش فراگیر هوشمند ادغام می‌کند. رپوسیس[1] (۲۰۱۶)، از مدل BENEIZ برای شناسایی گزارش‌های مالی شرکت‌های بزرگ در بریتانیا استفاده کردند [۲۵]. مدل دست‌کاری ردیابی درآمد می‌کوشد تا اقدامات غیرقانونی یا حداقل غیراخلاقی را افشا کند. زکریازاد و دومان[2] (۲۰۱۶)، شبکه عصبی مصنوعی با استفاده از برنامه‌های کاربردی خریدوفروش اینترنتی را بررسی کردند [۳۰]. هدف اصلی پژوهش طبقه‌بندی صحیح نمونه‌های سودآور به‌گونه‌ای است که تا حد امکان دقت ردیابی سایر نمونه‌هایی که سودآوری کمتری دارند،کاهش پیدا نکند. کیم[3] و همکاران (۲۰۱۶)، به تشخیص نقض تحریم‌های مالی با تقلب با استفاده از چند کلاس یادگیری حساس هزینه پرداختند [۱۰]. بایتا[4] و همکاران (۲۰۱۵)، به بررسی نمره اعتبار با استفاده از تکنیک‌های یادگیری ماشین پرداختند [۴]. مارتینز و تووباک[5] (۲۰۱۷)، نمره اعتباری خرده‌فروشی را با استفاده از داده‌های پرداخت‌نشده بررسی کردند [۱۶]. روش‌های استفاده‌شده در این پژوهش ادغام شبکه‌های عصبی و درخت تصمیم و بیزین به‌وسیله روش انسمبل بود. چن و کیشنیک[6] (۲۰۱۷)، اعتبار بانکی و مدیریت سرمایه در گردش شرکت را موردبررسی قرار دادند [۷]. در این پژوهش تأثیر تغییرات دسترس‌پذیری اعتبار بانکی بر مدیریت سرمایه در گردش توسط شرکت‌های دولتی بررسی شد. نامی و شجری[7] (۲۰۱۸) به تشخیص تقلب کارت اعتباری پرداختند[۲۰]. تقلب کارت اعتباری به ضررهای مالی سالانه سنگینی در سرتاسر دنیا منجر شده است و بنابراین نیاز به بهبود سیستم‌های ردیابی تقلب برای بانک‌ها و مؤسسه‌های مالی احساس می‌شود. به دلیل فضای خاص اقتصادی و وجود تسهیلات با نرخ‌های ترجیحی در اقتصاد ایران، هیچ‌یک از مطالعات خارجی به این موضوع نپرداخته‌اند. در پژوهش‌های داخلی نیز به این موضوع مهم و اهمیت آن پرداخته نشده است.

۳. روش‌شناسی پژوهش

از منظر گردآوری داده‌ها این پژوهش جزو پژوهش‌ها توصیفی ـ پیمایشی است. در روش پژوهش پیمایشی نمونه‌ای از کل جامعه موردپژوهش انتخاب شده و با بررسی و تحلیل نمونه یک نتیجه کلی حاصل می‌شود. در این پژوهش، اطلاعات گردآوری‌شده با استفاده از نرم‌افزارهای موجود برای داده‌کاوی بر مبنای مدل‌های مختلف در این حوزه بررسی و تحلیل خواهد شد. پایه اصلی پژوهش حاضر بر کشف دانش از پایگاه داده‌های بانک موردمطالعه نهاده شده است؛ ازاین‌رو برای

1. Repousis
2. Zakaryazad & Duman
3. Kim
4. Bhaita
5. Martens & Tobback
6. Chen & Kieschnick
7. Nami & Shajari

انجام فرآیند پژوهش از مراحلی شامل درک مسئله، آماده‌سازی داده‌ها، مدل‌سازی و ارزیابی نتایج استفاده شده است.

جامعه آماری این پژوهش در قسمتی از مراحل کار شامل بنگاه‌های اقتصادی در حوزه‌های مختلف صنعت است که درخواست‌کننده دریافت تسهیلات در خلال ۴ سال از ۱۳۸۶ تا ۱۳۹۰ هستند و اطلاعات مربوط به درخواست آن‌ها توسط بانک بررسی می‌شود و در قسمتی دیگر از مراحل کار، جامعه آماری متشکل از سه نفر از خبرگان حوزه اعتبارات بانکی است.

در ادامه مراحل انجام پژوهش حاضر به‌صورت زیر است:

در این پژوهش دو فرآیند کلی انجام شده است که هر یک از این دو فرآیند از تعدادی گام تشکیل شده‌اند؛ بنابراین گام‌های هر یک از این دو فرایند به‌صورت مجزا بیان می‌شود. این دو فرآیند عبارت‌اند از:

الف) کشف روابط علّی و معلولی متغیرهای مالی سرمایه در گردش؛

ب) کشف تقلب در متغیرهای مالی سرمایه در گردش برای دریافت تسهیلات بیشتر؛

در ادامه به‌تفصیل گام‌های هر یک از فرایندها به‌صورت زیر است:

فرآیند الف) کشف روابط علّی و معلولی متغیرهای مالی سرمایه در گردش: در این مرحله برای کشف روابط علّی و معلولی متغیرهای سرمایه در گردش گام‌های زیر صورت گرفته است:

الف-۱) شناسایی متغیرهای مالی تشکیل‌دهنده سرمایه در گردش با استفاده از اخذ نظر خبرگان و مطالعات کتابخانه‌ای و مقالات؛

الف-۲) جمع‌آوری نظرات خبرگان توسط پرسشنامه دیمتل؛

الف-۳) تشکیل ماتریس‌های مربوطه و رسم و تحلیل ماتریس تحلیل تأثیرات متقابل با استفاده از نتایج حاصل از پرسشنامه‌ها؛

الف-۴) استخراج نتایج از ماتریس‌های مرحله قبل و ترسیم نقاط به‌دست‌آمده در محور مختصات؛

الف-۵) ترسیم نمودار علّی و معلولی.

فرآیند ب) کشف تقلب در اثر دست‌کاری در متغیرهای مالی سرمایه در گردش برای دریافت تسهیلات بیشتر: در این مرحله برای کشف تقلب از الگوریتم‌های دسته‌بند مناسب برای داده‌های غیربالانس رتبه‌بندی اعتباری استفاده شده و مراحل به شرح زیر است:

ب-۱) جمع‌آوری و تنظیف داده‌ها؛

ب-۲) کشف الگوی تقلب و ایجاد برچسب بر داده‌ها؛

ب-۳) شناسایی متغیرهای مستقل و وابسته پژوهش؛

ب-۴) اجرای الگوریتم‌ها؛

ب-۵) ثبت نتایج بر اساس معیارهای سنجش عملکرد مدل؛
ب-۶) اجرای آزمون ویلکاکسون بر روی نتایج و ثبت گزارش این آزمون؛
ب-۷) بررسی قبول/ رد بودن فرضیه‌های پژوهش با استفاده از نتایج مرحله قبل.

در فرایند انجام پژوهش از معیار ورشکستگی آلتمن استفاده شد که در ادامه به‌اختصار توضیح داده شده است.

معیار تعیین ورشکستگی آلتمن. آلتمن[1] (2016) نخستین شخصی است که مدل‌های پیش‌بینی چند متغیره را عرضه کرد. وی با به‌کارگیری روش تحلیل تمایزی چندگانه و استفاده از نسبت‌های مالی به‌عنوان متغیرهای مستقل به دنبال پیش‌بینی ورشکستگی بنگاه‌ها بود. او مدل معروف خود را با عنوان «مدل امتیاز Z»[2] ارائه داد که در پیش‌بینی ورشکستگی معروف است. او در این روش از بین ۲۲ نسبت مالی که به نظر وی بهترین نسبت‌ها برای پیش‌بینی ورشکستگی بود، پنج نسبت را انتخاب کرد. در سال‌های بعد انتقاداتی همچون قابلیت کاربردی انحصاری این مدل برای مؤسسه‌های عمومی از سوی تحلیل‌گران و مدیران عنوان شد که آلتمن برای رفع این‌گونه انتقادات، موفق به رفع اشکالات مدل اولیه شد و مدل جدیدی را با عنوان 'Z به‌صورت زیر عرضه کرد:

رابطه (۱): مدل آلتمن $Z' = 0.717x_1 + 0.847x_2 + 3.107x_3 + 0.420x_4 + 0.998x_5$

Z' : شاخص کل
x_1 : نسبت سرمایه در گردش به‌کل دارایی
x_2 : نسبت سود انباشته به‌کل دارایی
x_3 : نسبت سود قبل از بهره و مالیات به‌کل دارایی‌ها
x_4 : نسبت ارزش دفتری سهام شرکت به ارزش دفتری کل بدهی‌ها
x_5 : نسبت فروش به‌کل دارایی‌ها

در این مدل هرچه 'Z پایین باشد، درجه بحران مالی شرکت بیشتر است؛ به‌طوری‌که شرکت‌ها با امتیاز 'Z بالاتر از ۲/۹ وارد طبقه شرکت‌های سالم شده و با امتیاز 'Z کمتر از ۱/۲۳ به‌عنوان شرکت‌های ورشکسته طبقه‌بندی می‌شوند و 'Z بین ۱/۲۳ و ۲/۹ به‌عنوان منطقه تردید تلقی می‌شود. ناحیه مزبور باید با احتیاط بررسی شود. آلتمن با این روش به ۹۴ درصد پیش‌بینی صحیح دست‌یافت [۲۵].

اهلسون[3] (۲۰۱۷)، نیز توانست برای تعیین ورشکستگی شرکت‌ها از مدل رگرسیون لجستیک استفاده کند. وی توانست ورشکستگی شرکت‌ها را برای سال اول تا سوم به‌ترتیب با دقت ۸۵/۱

1. Altman
2. Z-Score
3. Ohlson

درصد، ۸۷/۶ درصد و ۸۲/۶ درصد پیش‌بینی کند و متغیرهای نسبت بدهی کل به دارایی کل و نسبت سود خالص به مجموع دارایی بهترین نسبت‌ها تفکیک‌کننده در مدل وی بودند.

$$Z_i = -1.32 - 0.407x_1 + 6.03x_2 - 1.43x_3 + 0.0757x_4 - 2.37x_5 - 1.83x_6 + 0.285x_7 - 1.72x_8 - 0.521x_9$$

رابطه (۲): مدل رگرسیون لجستیک

Z_i: شاخص نصبی برای محاسبه تابع احتمال

X_1: شاخص نسبی برای محاسبه تابع احتمال

X_2: لگاریتم (نسبت کل دارایی‌ها به شاخص ریالی تولید ناخالص ملّی[1])

X_3: نسبت سرمایه در گردش به‌کل دارایی‌ها

X_4: نسبت بدهی‌های جاری به دارایی‌های جاری

X_5: متغیر مجازی که اگر کل بدهی‌ها بزرگ‌تر یا مساوی کل دارایی‌ها باشد به آن عدد یک و در غیر این‌صورت عدد صفر تخصیص داده می‌شود.

X_6: نسبت سود خالص به مجموع دارایی‌ها

X_7: نسبت وجوه حاصل از عملیات به‌کل بدهی‌ها

X_8: متغیر مجازی که اگر سود خالص برای دو سال متوالی قبل منفی باشد، عدد یک و در غیر این‌صورت عدد صفر اختصاص می‌یابد.

X_9: تغییرات در سود خالص به‌صورت:

$$x_9 = (NI_T - NI_{T-1})/(|NI_T| + |NI_{T-1}|)$$

رابطه (۳): تغییرات در سود خالص

که NI همان سود خالص و T ارقام سال جاری و T-1 ارقام سال قبل است.

طبق تحلیل لوجیت[2]، و همکاران (۲۰۱۰)، مقدار Z به‌دست‌آمده برای هر مورد در معادله $P(z) = 1/(1 + e^{-z})$ قرار داده شده و احتمال شرطی ورشکستگی آن محاسبه می‌شود. درصورتی‌که P(z) کمتر از ۰/۵ باشد، شرکت به‌عنوان ورشکسته و درصورتی‌که P(z) بزرگ‌تر از ۰/۵ باشد، شرکت سالم طبقه‌بندی می‌شود و نقطه بحرانی در اینجا P(z)=۰/۵ خواهد بود [۲۷]. در بخش چهارم نتایج آزمایش‌ها و بحث ارائه شده است.

آزمایش‌ها و بحث. در این قسمت گام‌های اجرا و یافته‌ها و نتایج پژوهش بررسی شده است.

1. Gross National Production (GNP)
2. Logit Analysis

فرآیند الف) کشف روابط علّی و معلولی متغیرهای مالی سرمایه در گردش. در این مرحله برای کشف روابط علّی و معلولی متغیرهای سرمایه در گردش گام‌های زیر صورت گرفته است:

الف-۱) شناسایی متغیرهای مالی تشکیل‌دهنده سرمایه در گردش با استفاده از اخذ نظر خبرگان و مطالعات کتابخانه‌ای و مقالات: پس از اخذ نظر خبرگان و مطالعات کتابخانه‌ای دو عنصر اصلی دخیل در محاسبات سرمایه در گردش شناسایی شدند. این دو عنصر اصلی عبارت‌اند از: دارایی‌های جاری و بدهی‌های جاری. عمدتاً مبنای محاسبات بانک‌ها برای اعطای سرمایه در گردش، تفاضل بدهی‌های جاری از دارایی‌های جاری است. هر یک از این دو عنصر خود متشکل از حاصل‌جمع مؤلفه‌های دیگری هستند. دارایی جاری حاصل‌جمع موجودی نقد+حساب‌های دریافتی+ سایر حساب‌های دریافتی و گاهی اسناد دریافتی+ موجودی کالا و انبار+ پیش‌پرداخت‌ها است. بدهی‌های جاری نیز شامل حساب‌های پرداختنی+ سایر حساب‌های پرداختی (شامل تسهیلات و انواع بدهی‌های دیگر به تأمین‌کنندگان، جاری شرکا و غیره) است.

بنگاه‌های اقتصادی که به دنبال بیشتر نشان دادن سرمایه در گردش موردنیاز هستند، این مؤلفه‌های مالی را حتی‌الامکان تغییر می‌دهند. تغییر یکی از این متغیرها عمدتاً به تغییر متغیرهای دیگر در صورت‌های مالی منجر می‌شود. به‌منظور شناسایی این اثر نظرهای خبرگان با استفاده از تکنیک دیمتل مورداستفاده قرار گرفته شده است.

الف-۲) جمع‌آوری نظرهای خبرگان توسط پرسشنامه دیمتل[1]: مراحل زیر برای ۵ متغیر وجه نقد موجود (موجودی نقد)، حساب‌های دریافتی، سایر حساب‌های دریافتی، موجودی کالا و انبار و نهایتاً پیش‌پرداخت‌ها بررسی شد. همچنین ازآنجاکه عناصر مالی بدهی جاری عملکرد یکسانی دارند و لزوماً رابطه‌ای با دارایی‌های جاری ندارند در این محاسبات علّی و معلولی در نظر گرفته نشده‌اند.

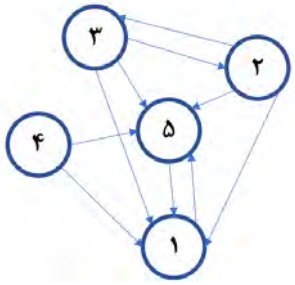
الف-۳) تشکیل ماتریس‌های مربوطه و رسم و تحلیل ماتریس تحلیل تأثیرات متقابل با استفاده از نتایج حاصل از پرسشنامه‌ها؛

الف-۴) استخراج نتایج از ماتریس‌های مرحله قبل و ترسیم نقاط به‌دست‌آمده در محور مختصات؛

الف-۵) ترسیم نمودار علّی و معلولی.

شکل ۱، نمودار علت و معلولی میان این ۵ متغیر را نشان می‌دهد.

۱. با توجه به روتین‌بودن این تکنیک دیمتل در صورت تمایل برای اطلاع از ارائه مراحل آن در این پژوهش خودداری شده است و برای بررسی بیشتر می‌توان به پایان‌نامه‌ای که مقاله از آن مستخرج شده است، مراجعه کرد.



شکل ۱. نمودار علت و معلولی عناصر مالی مداخله‌گر در تقلب و وابستگی آن‌ها به یکدیگر

همان‌گونه که از شکل مشخص است، عناصر یک و پنج که موجودی نقد و پیش‌پرداخت‌ها هستند، بیشترین رابطه را با سایر عوامل دارند و بیشترین تغییرات در صورت تغییر سایر عناصر شامل حال آن‌ها می‌شود.

فرآیند ب) کشف تقلب در اثر دست‌کاری در متغیرهای مالی سرمایه در گردش برای دریافت تسهیلات بیشتر: در این مرحله برای کشف تقلب از درخت تصمیم C4.5 استفاده شده است. مراحل به شرح زیر است:

ب-۱) جمع‌آوری و تنظیف داده‌ها: در این مرحله داده‌های اعتباری شرکت‌های متقاضی وام که به‌منظور اخذ وام به شعب بانک توسعه صادرات کل کشور مراجعه کرده‌اند، جمع‌آوری شد. در ابتدا این پایگاه داده شامل اطلاعات ۱۱۰۹ شرکت متقاضی وام بوده است که پاک‌سازی‌شده‌اند. در مرحله تنظیف و پاک‌سازی، داده‌های زائد، پرت و موارد مفقود توسط نرم‌افزار SPSS Statistics 23 شناسایی و حذف شدند و این عملیات به حذف داده‌های ۳۸۶ شرکت منجر شد؛ همچنین ویژگی‌هایی که به‌صورت کیفی بوده‌اند، مجازی‌سازی شدند. درنهایت ۵۵ ویژگی مربوط به ۷۲۳ شرکت باقیمانده بررسی شد. این ویژگی‌ها به‌عنوان عوامل تعیین‌کننده ریسک اعتباری مشتریان تعیین شده‌اند و شامل متغیرهای کیفی و مالی هستند.

جدول ۱، به‌خوبی نمایانگر مشخصات داده‌های موردبررسی قبل و بعد از پاک‌سازی است.

جدول ۱. مشخصات داده‌های قبل و بعد از پاک‌سازی

وضعیت	اندازه جامعه	متغیرهای پژوهش		
		کل	پیوسته	رسته‌ای
قبل از پاک‌سازی	۱۱۰۹	۵۱	۴۳	۸
بعد از پاک‌سازی	۷۲۳	۵۵	۳۶	۱۹

متغیرهای مسئله را می‌توان به تفکیک در جدول ۲، مشاهده کرد.

جدول ۲. متغیرهای مسئله

متغیر مستقل	۱) وجه نقد موجودی؛ ۲) حساب‌های دریافتنی؛ ۳) سایر حساب‌های دریافتنی؛ ۴) موجودی کالا و انبار؛ ۵) پیش‌پرداخت‌ها؛ ۶) حساب‌های پرداختنی؛ ۷) سایر حساب‌های پرداختنی؛ ۸) دارایی‌های جاری؛ ۹) بدهی‌های جاری؛ ۱۰) دارایی‌های غیر جاری؛ ۱۱) بدهی‌های غیر جاری؛ ۱۲) کل دارایی؛ ۱۳) کل بدهی؛ ۱۴) دارایی مالی درازمدت؛ ۱۵) بدهی مالی درازمدت؛ ۱۶) نوع صنعت؛ ۱۷) سال؛ ۱۸) اظهارنامه مالیاتی؛ ۱۹) سازمان حسابرسی؛ ۲۰) حسابرس معتبر؛ ۲۱) تسهیلات مالی کوتاه‌مدت؛ ۲۲) تسهیلات مالی بلندمدت؛ ۲۳) سرمایه؛ ۲۴) سود یا زیان انباشته؛ ۲۵) حقوق صاحبان سرمایه؛ ۲۶) فروش؛ ۲۷) سود ناخالص؛ ۲۸) هزینه‌های مالی؛ ۲۹) سود خالص؛ ۳۰) فعالیت در بازار داخلی؛ ۳۱) قلمرو بازار خارجی؛ ۳۲) رشد فروش؛ ۳۳) ریسک بازار هدف؛ ۳۴) عامل فصلی؛ ۳۵) سابقه شرکت؛ ۳۶) سابقه مدیران ارشد؛ ۳۷) نوع شرکت (تعاونی، سهامی عام، سهامی خاص، مسئولیت محدود و بورس اوراق بهادار)؛ ۳۸) گزارش حسابرسی؛ ۳۹) سابقه فعالیت با بانک؛ ۴۰) فروش دوره جاری؛ ۴۱) فروش دوره قبل؛ ۴۲) فروش دو دوره قبل؛ ۴۳) دارایی‌های دوره قبل؛ ۴۴) دارایی‌های دو دوره قبل؛ ۴۵) حقوق صاحبان سهام دوره جاری؛ ۴۶) حقوق صاحبان سهام دوره قبل؛ ۴۷) حقوق صاحبان سهام دو دوره قبل؛ ۴۸) گردش بستانکار حساب‌های جاری؛ ۴۹) میانگین وزنی حساب جای؛ ۵۰) میانگین صادرات سه سال گذشته؛ ۵۱) میانگین واردات سه سال گذشته؛ ۵۲) صنعت کشاورزی؛ ۵۳) صنعت پتروشیمی؛ ۵۴) صنعت محصولات شیمیایی؛ ۵۵) سازمان‌های خدماتی
متغیر وابسته	مرحله اول: معوقات یا نکول مرحله دوم: تقلب

ب-۲) کشف الگوی تقلب و ایجاد برچسب بر داده‌ها: در این مرحله باید داده‌ها بر اساس متقلب و غیرمتقلب برچسب‌گذاری شوند. با توجه به اینکه بانک‌های موردبحث این پژوهش، بانک‌های توسعه‌ای هستند که تسهیلات با نرخ ترجیحی ارائه می‌کنند و بنا بر سرمایه در گردش سازمان‌ها اعطای تسهیلات انجام می‌شود؛ بنابراین بعضی از سازمان‌ها و مراکز صنعتی که درخواست‌کننده تسهیلات هستند، می‌کوشند تا سرمایه در گردش موردنیاز خود را بیشتر از میزان واقعی نشان دهند تا بتوانند از مقدار تسهیلات بیشتری بهره‌مند شوند و این کار را با دست‌کاری و یا تقلب در متغیرهای سرمایه در گردش انجام می‌دهند. ولی مقدار سرمایه در گردش حد آستانه‌ای دارد؛ زیرا بر اساس معیارهای تعیین ورشکستگی می‌توان ورشکسته یا سالم‌بودن آن مشتری را تعیین کرد و اگر در اعلام وضعیت، ورشکستگی را اعلام نکرده باشد، لزوماً در صورت مالی تقلب صورت گرفته و به‌اصطلاح سیاه‌نمایی رخ داده است. با استفاده از روش زیر می توان وضعیت انحلا را تشخیص داد.

به‌منظور دسته‌بندی شرکت‌ها به دو گروه ورشکسته و سالم از مفاد ماده ۱۴۱ قانون تجارت استفاده شده است. طبق ماده مزبور اگر بر اثر زیان‌های وارده حداقل نصف سرمایه شرکت از

میان برود، هیئت‌مدیره مکلف است بلافاصله مجمع عمومی فوق‌العاده صاحبان سهام را دعوت کند تا موضوع انحلال یا بقای شرکت موردشور و رأی واقع شود. هرگاه مجمع مزبور رأی به انحلال شرکت ندهد، باید در همان جلسه و با رعایت مقررات ماده ۵ این قانون[۱] سرمایه شرکت به مبلغ سرمایه موجود کاهش یابد. درصورتی‌که هیئت‌مدیره برخلاف این ماده به دعوت مجمع عمومی فوق‌العاده مبادرت نورزد و یا مجمعی که دعوت می‌شود نتواند مطابق مقررات قانونی منعقد شود، هر ذی‌نفع می‌تواند انحلال شرکت را از دادگاه صلاحیت‌دار درخواست کند. به همین دلیل شرکت‌ها موظف هستند در صورت‌وضعیت مالی و ترازنامه خود به‌صورت رسمی اعلام ورشکستگی کنند.

با توجه به روش ذکر شده در فوق و با توجه به نظر خبرگان بخش اعتبارات بانکی این سه معیار رکن اساسی در تعیین وضعیت مشتریان تسهیلات در ایران است. چون هیچ‌یک از ۷۲۳ مشتری موجود در پایگاه داده اعلام ورشکستگی نکرده‌اند، طبق معیار سوم، می‌توان گفت هیچ یک از این شرکت‌ها ورشکسته نیستند و اگر بر اساس معیار ۱ و ۲ در دسته ورشکستگی قرار بگیرند، با توجه به‌دقت دو مدل گفته‌شده، در صورت مالی تغییراتی را در جهت حداقل‌کردن سرمایه در گردش اعمال کرده‌اند. درنتیجه می‌توان داده‌هایی را بر اساس الگوی رفتاری صورت مالی این سازمان‌ها ایجاد کرد. پس از اعمال دو معیار ورشکستگی آلتمن و اهلسون از ۷۲۳ مشتری تسهیلات در پایگاه داده، ۱۵ شرکت ورشکسته اعلام شدند که برای تناسب کلاس متقلبین، بر اساس این الگوی رفتاری تعداد ۳۵۰ داده تولید شد که درنهایت علاوه بر ستون معوقات، ستون تقلب نیز ایجاد شد.

ب-۳) شناسایی متغیرهای وابسته و مستقل پژوهش: با توجه به اینکه داده‌ها شامل ۵۵ ویژگی هستند که ۷ ویژگی متغیرهای سرمایه در گردش و ۴۸ ویژگی دیگر سایر عناصر مالی هستند و ازآنجاکه این پژوهش درصدد شناسایی آثار تقلب بر مدل‌های اعتبارسنجی و ساخت یک مدل بهینه است، گام‌های این مرحله را می‌توان به دو بخش تقسیم کرد:

الف) بررسی داده‌ها با درنظرگرفتن نکول یا معوقات؛

ب) بررسی داده‌ها با درنظرگرفتن تقلب.

۱. طبق این ماده، هنگام تأسیس، سرمایه شرکت‌های سهامی عام از پنج میلیون ریال و سرمایه شرکت‌های سهامی خاص از یک میلیون ریال نباید کمتر باشد. در صورتی که سرمایه شرکت بعد از تأسیس به هر علت از حداقل یادشده در این ماده کمتر شود، باید ظرف مدت یک سال نسبت به افزایش سرمایه تا میزان حداقل مقرر اقدام به‌عمل آید یا شرکت به نوع دیگری از انواع شرکت‌های یادشده در این قانون تغییر شکل یابد؛ وگرنه هر ذی‌نفع می‌تواند انحلال آن را از دادگاه صلاحیت‌دار درخواست کند.

برای تحقق دو هدف الف و ب و نیز بررسی فرضیه‌های پژوهش از الگوریتم‌های داده‌کاوی که در جدول ۳، ذکر شده است، استفاده شد.

روش کار به‌صورت زیر است:

۱. جداسازی داده‌ها بر اساس ویژگی‌ها به دو گروه:

ـ گروه اول: کل داده‌ها بدون درنظرگرفتن ستون متغیرهای سرمایه در گردش؛ درنتیجه ۴۸ ویژگی باقی می‌ماند.

ـ گروه دوم: کل پایگاه داده؛ درنتیجه ۵۵ ویژگی حضور دارند.

۲. تقسیم‌بندی کار با دو ستون متغیر وابسته:

ـ متغیر وابسته اول: نکول یا معوقات

ـ متغیر وابسته دوم: تقلب

ب- ۵ و ۴) اجرای الگوریتم‌ها و ثبت نتایج بر اساس معیارهای سنجش عملکرد مدل.

بدین ترتیب الگوریتم‌ها به‌صورت مجزا بر روی داده‌ها با ۴۸ ویژگی و ۵۵ ویژگی و با درنظر گرفتن متغیر وابسته‌ی تقلب و معوقات به‌صورت جداگانه اعمال و گزارش شاخص‌های کلیدی عملکرد و نیز نمایش انتخاب ویژگی در جدول‌های ۳ و ۴ ثبت شد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

جدول ۳. نتایج شاخص‌های کلیدی ارزیابی مدل و گزارش ویژگی‌های خروجی یا درنظرگرفتن معوقات

روش کار	ویژگی خروجی	FPR	TPR	ACC	AUC	متغییر وابسته	متغیر وابسته
LSVM+Greedy	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۱ ۲۲، ۲۳، ۲۵، ۲۷، ۲۸، ۲۹، ۴۳، ۴۴، ۴۹	۰/۱۱	۰/۹۱	۹۱/۱۸	۹۰/۳۶	نکول یا معوقات	۴۸
LSVM+Greedy	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲، ۲۳،۲۵،۲۷،۲۸،۲۹، ۴۳، ۴۴، ۴۹، ۱، ۲	۰/۱۴	۱	۹۲/۵۶	۹۱/۲۲	نکول یا معوقات	۴۸+۷
Random Forest+GA	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۷ ۲۹، ۲۸، ۴۳، ۴۴	۰/۰۷	۰/۹۰	۸۹/۳۷	۸۹/۰۸	نکول یا معوقات	۴۸
Random Forest+GA	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱، ۲۷، ۲۹، ۲۸، ۴۳، ۴۴، ۳، ۷، ۸	۰/۱۰	۰/۹۲	۹۰/۱۴	۸۹/۲۵	نکول یا معوقات	۴۸+۷
KNN+Decision Tree C4.5+SVM+Ensemble	۱۰، ۱۱، ۱۲، ۱۳، ۲۱، ۲۵، ۲۷، ۲۹، ۴۳۴۴، ۴۹	۰/۰۴	۰/۹۵	۹۲/۳۹	۹۳/۶۵	نکول یا معوقات	۴۸
KNN+Decision Tree C4.5+SVM+Ensemble	۱۰، ۱۱، ۱۲، ۱۳، ۲۱، ۲۵، ۲۷، ۲۹، ۴۳، ۴۴، ۴۹، ۱، ۳، ۷	۰/۰۳	۰/۹۶	۹۳/۳۲	۹۳/۸۶	نکول یا معوقات	۴۸+۷
Neural Network+GA	۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲۲۶، ۲۸، ۳۲، ۳۸	۰/۲۵	۰/۷۰	۸۹/۰۱	۸۹/۸۰	نکول یا معوقات	۴۸
Neural Network+GA	۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲ ۲۶، ۲۸، ۳۲، ۳۸، ۲، ۷، ۸	۰/۳۰	۰/۷۷	۸۹/۹۹	۹۰/۸۸	نکول یا معوقات	۴۸+۷

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

جدول ۴. نتایج شاخص‌های کلیدی ارزیابی مدل و گزارش ویژگی‌های خروجی یا درنظرگرفتن تقلب

روش کار	ویژگی خروجی	FPR	TPR	ACC	AUC	متغیر وابسته	تعداد متغیر مستقل
LSVM+Greedy	۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۳، ۲۴، ۲۵، ۲۷، ۲۸، ۴۰، ۴۱، ۴۲، ۴۴، ۴۸	۰/۱۹	۰/۸۷	۸۸/۹۹	۸۹/۰۹	تقلب	۴۸
LSVM+Greedy	۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۳، ۲۴، ۲۵، ۲۷، ۲۸، ۴۰، ۴۱، ۴۲، ۴۴، ۴۸، ۱، ۴	۰/۱۵	۰/۹۰	۸۹/۱۰	۹۰/۸۶	تقلب	۴۸+۷
Random Forest+GA	۱۰، ۱۱، ۱۲، ۱۳، ۲۱، ۲۲، ۲۳، ۲۵، ۲۷، ۴۳، ۴۴، ۴۹	۰/۱۰	۰/۹۴	۹۰/۴۳	۹۰/۰۱	تقلب	۴۸
Random Forest+GA	۱۰، ۱۱، ۱۲، ۱۳، ۲۱، ۲۲، ۲۳، ۲۵، ۲۷، ۴۳، ۴۴، ۴۹، ۴، ۵	۰/۰۸	۰/۹۵	۹۰/۸۷	۹۰/۹۸	تقلب	۴۸+۷
KNN+Decision Tree C4.5+SVM+Ensemble	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۴، ۲۶، ۲۸، ۴۳، ۴۵، ۴۸، ۴۹	۰/۱۵	۰/۸۸	۹۱/۱۸	۹۱/۰۹	تقلب	۴۸
KNN+Decision Tree C4.5+SVM+Ensemble	۸، ۹، ۱۰، ۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۲۴، ۲۶، ۲۸، ۴۳، ۴۵، ۴۸، ۴۹، ۱، ۴، ۵	۰/۰۴	۰/۹۲	۹۲/۰۴	۹۲/۸۹	تقلب	۴۸+۷
Neural Network+GA	۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲، ۲۵، ۲۷، ۲۸، ۴۳، ۴۴، ۴۶، ۴۷	۰/۲۰	۰/۷۳	۸۹/۰۲	۸۹/۹۲	تقلب	۴۸
Neural Network+GA	۱۲، ۱۳، ۱۴، ۱۵، ۲۱، ۲۲، ۲۵، ۲۷، ۲۸، ۴۳، ۴۴، ۴۶، ۴۷، ۲، ۳، ۵	۰/۱۲	۰/۸۲	۸۹/۹۰	۹۰/۰۱	تقلب	۴۸+۷

ب-۶ و ۷) اجرای آزمون ویلکاکسون بر روی نتایج و ثبت گزارش این آزمون و بررسی قبول/ رد بودن فرضیه‌های پژوهش با استفاده از نتایج مرحله قبل. در ادامه به‌منظور اثبات فرضیه‌های پژوهش، نتایج که در جدول ۴، نمایش داده شده‌اند، وارد نرم‌افزار SPSS می‌شوند تا آزمون ویلکاکسون روی آن‌ها صورت پذیرد. به‌دلیل محبوبیت بیشتر آزمون‌های ناپارامتری، در این پژوهش برای آزمون فرضیه اول و دوم از آزمون ویلکاکسون و برای فرضیه سوم از آزمون کروسکال والیس استفاده شده است. ابتدا به‌منظور اثبات فرضیه اول، داده‌هایی که با معوقات سنجیده شده است، وارد می‌شود.

فرضیه اول:

H_0: مدل اعتبارسنجی با وجود متغیرهای سرمایه در گردش از منظر شاخص‌های کلیدی عملکرد تفاوت معناداری با مدل اعتبارسنجی بدون این متغیرها ندارد (با درنظرگرفتن معوقات).

$H_0: \mu_1=\mu_2$

H_1: مدل اعتبارسنجی با وجود متغیرهای سرمایه در گردش از منظر شاخص‌های کلیدی عملکرد تفاوت معناداری با مدل اعتبارسنجی بدون این متغیرها دارد (با در نظر گرفتن معوقات).

$H_0: \mu_1 \mu \mu_2$

در تفسیر نتایج آزمون ویلکاکسون برای پی‌بردن به این نکته که آیا تفاوت معناداری بین مدل اعتبارسنجی بدون وجود متغیرهای سرمایه در گردش با مدل کلی اعتبارسنجی وجود دارد (با درنظرگرفتن نکول یا معوقات) یا خیر، از نتایج جدول 4 استفاده می‌شود. ازآنجاکه در نتایج در قسمت Test Statistic معیار تصمیم (.Sig) مقدار ۰/۰۰۱ به‌دست آمده و این مقدار از ۰/۰۵ کمتر است، فرض صفر رد می‌شود؛ یعنی تفاوت معناداری بین مدل اعتبارسنجی بدون متغیرهای سرمایه در گردش و کل پایگاه داده وجود دارد.

فرضیه دوم:

H_0: مدل اعتبارسنجی با وجود متغیرهای سرمایه در گردش از منظر شاخص‌های کلیدی عملکرد تفاوت معناداری با مدل اعتبارسنجی بدون این متغیرها ندارد (با درنظرگرفتن تقلب).

$H_0: \mu_1=\mu_2$

H_1: مدل اعتبارسنجی با وجود متغیرهای سرمایه در گردش از منظر شاخص‌های کلیدی عملکرد تفاوت معناداری با مدل اعتبارسنجی بدون این متغیرها دارد (با درنظرگرفتن تقلب).

$H_0: \mu_1 \mu \mu_2$

با توجه به نتایج و اینکه معیار تصمیم اهمیت مقدار ۰/۰۱۷ به‌دست آمده و این مقدار از ۰/۰۵ کمتر است، فرضیه صفر رد می‌شود و تفاوت معناداری بین مدل اعتبارسنجی بدون حضور متغیرهای سرمایه در گردش و کل پایگاه داده وجود دارد (با درنظرگرفتن تقلب).

فرضیه سوم:

H_0: مدل‌های ارائه‌شده از منظر شاخص‌های کلیدی عملکرد نتایج یکسان دارند.

$\mu_1= \mu_2= \mu_3= \mu_4$

H_1: مدل‌های ارائه شده از منظر شاخص‌های کلیدی عملکرد نتایج یکسان ندارند.

$\mu_1\neq \mu_2 \neq \mu_3 \neq \mu_4$

برای آزمون این فرضیه ابتدا از آزمون کولموگروف ـ اسمیرنوف استفاده شد تا نرمال‌بودن داده‌ها سنجیده شود. نتایج این آزمون بر نرمال‌نبودن داده‌ها دلالت داشت؛ بنابراین از آزمون کروسکال والیس استفاده شد. ازآنجاکه مقدار معیار تصمیم از ۰/۰۵ بزرگ‌تر است، فرض صفر رد نمی‌شود و مدل‌های ارائه‌شده از نظر شاخص‌های کلیدی عملکرد تفاوت معناداری ندارند و با توجه به میانگین این شاخص‌ها، مدل ساخته‌شده با روش انسمبل از نظر این شاخص‌ها کمی بهتر از بقیه عمل کرده است.

فرضیه اصلی این پژوهش وجود یا عدم‌وجود تأثیر تقلب یا دست‌کاری در متغیرهای سرمایه در گردش برای اخذ تسهیلات بیشتر، بر مدل اعتبار سنجی بود. پس از اعمال الگوریتم‌های دسته‌بند و انتخاب متغیر بر روی دو گروه داده (داده‌هایی با ویژگی‌های کل پایگاه داده، داده‌هایی بدون حضور متغیرهای سرمایه در گردش) دو نوع خروجی حاصل شد:

۱. گزارش شاخص‌های عملکرد مدل، مانند سطح زیر منحنی، صحت، حساسیت و ویژگی (حاصل از الگوریتم‌های دسته‌بند)؛

۲. گزارش ویژگی‌های خروجی (حاصل از نتایج الگوریتم‌های انتخاب متغیر).

سپس نتایج شاخص‌های عملکرد برای هر دو گروه داده (۵۵تایی و ۴۸تایی بدون متغیرهای سرمایه در گردش) با آزمون ویلکاکسون مقایسه شد و این نتیجه به‌دست آمد که تفاوت معناداری بین این دو مدل از نظر این شاخص‌ها وجود دارد؛ بنابراین برای اعطای تسهیلات بر مبنای سرمایه در گردش باید از مدل اعتبارسنجی خاص بانک توسعه‌ای استفاده شود و از مدل کلی پیروی نمی‌کند و نیز با استفاده از آزمون کروسکال والیس مشخص شد که تفاوت معناداری بین مدل‌ها از نظر این شاخص‌ها وجود ندارد.

۵. نتیجه‌گیری و پیشنهادها

پژوهش حاضر شکاف پژوهش‌های پیشین در عدم‌بررسی مدل‌های اعتبارسنجی مورداستفاده توسط بانک‌های توسعه‌ای و تأثیر تقلب بر آن‌ها را پوشش می‌دهد. بر این مبنا به‌صورت تخصصی، مدل‌های اعتبارسنجی سرمایه در گردش مختص بانک‌های توسعه‌ای با داده‌های شرکت‌های متقاضی تسهیلات ارائه شده توسط این نوع بانک‌ها که عمدتاً بلندمدت‌تر، دارای نرخ ترجیحی و با وثایق سبک‌تر است، بررسی شد. نتایج آزمایش‌های انجام‌شده با ۵۵ ویژگی در این پژوهش نشان داد که مدل‌های اعتبارسنجی مورداستفاده در بانک‌های تجاری نمی‌توانند برای بانک‌های توسعه‌ای مورداستفاده قرار گیرند و بانک‌های توسعه‌ای باید مدل‌های اعتبارسنجی خاص خود را با درنظرگرفتن اثر احتمالی دست‌کاری و تقلب ایجاد کنند که این مدل شامل متغیرهای جدید ارزیابی اعتباری، ازجمله متغیرهای سرمایه در گردش موردنیاز دارایی جاری منهای بدهی جاری، مجموع اسناد پرداختی تجاری و مجموعه اسناد دریافتنی تجاری است که باید توسط الگوریتم‌های جدید آموزش و توسعه داده شده باشند.

پژوهش‌های آتی می‌تواند بر مسئله رتبه‌بندی رفتاری و رتبه‌بندی وصول مطالبات تسهیلات ارائه‌شده توسط این بانک‌ها متمرکز باشد؛ چراکه به‌دلیل شرایط خاص تسهیلات پرداختی توسط این نهادهای رفتارهای بازپرداخت نیز متفاوت خواهد بود. به این معنا که احتمالاً رفتار بازپرداخت و مدل‌های رتبه‌بندی رفتاری متداول که توسط بانک‌های تجاری استفاده می‌شود، پاسخگوی نوع بازپرداخت در بانک‌های توسعه‌ای نیست؛ همچنین در پژوهش‌های دیگر در حوزه رتبه‌بندی وصول مطالبات و نوع تعامل با انواع دیرکردهای چهارگانه (جاری، سررسید شده، معوق و مشکوک‌الوصول) مدنظر بانک مرکزی چه نوع اقداماتی باید توسط یک بانک توسعه‌ای صورت گیرد و این نوع اقدامات آیا اساساً تفاوتی با اقداماتی که توسط بانک‌های تجاری برای وصول مطالبات صورت می‌گیرند، باید دارای تفاوت معناداری باشند یا خیر؟

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

منابع

1. Abdelhamid, D. Khaoula, S. & Atika, O. (2014). Automatic Bank Fraud Detection Using Support Vector Machines. In *The International Conference on Computing Technology and Information Management (ICCTIM2014)* (pp. 10-17). The Society of Digital Information and Wireless Communication.
2. Abdolli, Gh. Fard Hariri, E. (2016). Refah bank credit scoring model benchmarking, Nazarieh karbordi eghtesad, 1-24. (In Persian)
3. Alizadeh farad, F, Shafiabadi, M.H, & Safara, F. (2015). Introducing a hybrid model for banks credit scoring using datamining. *Third conference on new innovations in engineering and computer*.
4. Arabmazar, M. & Rouientan, P. (2005). Factros effecting bank credit risks: Keshavarzi bank case study, Jastarhaye eghtesadi, 45-80. (In Persian)
5. Amiri, S. & Naderi, N. & Mohammadifar, y. & Rezaie, B. (2018). Designing and Arranging a Model for RevivalStagnant Industrial Enterprises (Case Study: Kermanshah Province). *Chasm andaz modirat-e-sanaati, 36*, 63-88. (In Persian)
6. Bae, J. K. & Kim, J. (2015). A personal credit rating prediction model using data mining in smart ubiquitous environments. *International Journal of Distributed Sensor Networks*, *11*(9), 179060.
7. Chen, C. & Kieschnick, R. (2017). Bank credit and corporate working capital management. *Journal of Corporate Finance, 48, 579-596*.
8. Asadpour, E. & Pouya, A. & Farimani, N. M. (1396). Designing a dynamic balanced scorecard for banks performance management. *Chasm andaz modirat-e-sanaati, 28,* 163-297 (In Persian).
9. Hajizadeh, S. & Torani, Z. (2016). Datamining in banks using customers credit scoring. *4th conference on new studies in computer science and technology*. (In Persian)
10. Kim, Y. J. Baik, B., & Cho, S. (2016). Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert systems with applications*, *62*, 32-43.
11. Kim, Y. S., & Sohn, S. (2004). Managing loan customers using misclassification patterns of credit scoring model. Expert Systems with Applications, 517-573.
12. Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32(4),* 995–1003.
13. Lei, J. Z. & Ghorbani, A. A. (2012). Improved competitive learning neural networks for network intrusion and fraud detection. *Neurocomputing, 75(1),* 135-145.
14. Liu, Y., & Schumann, M. (2005). Data mining feature selection for credit scoring models. *Journal of the Operational Research Society, 56(9),* 1099-1108.
15. Mahmoudi, S., & Sharifi, H. (2013). Study the causes and effective factors in small businesses loan failure at Bank melli Iran. (Case study: Branches of Bank Melli in the Shahre-Kord). *Fist international e confrance on future perspective of Iranian Economy,* 1-20. (In Persian)
16. Martens, D., & Tobback, E. (2017). Retail credit scoring using fine-grained payment data.
17. Martínez, J De L., & Vicente, CL. (2012). Global survey of development banks. 1-41.

18. Mehregan, M. R. & Moradi, Z. (2019) Using the Multi-Stage of Integrating ApproachesData Envelopment Analysis (DEA) and Balanced Scorecard (BSC) for Enhanced Performance Assessment. *The Journal of Industrial mangament perspective, 37,* 143-165 (In Persian).
19. Mohammadi Vala, H. & Farhoudinejad, A. (2014), Fraud detection in banking transactions using datamining: Mehr-e-eghtesad ban case strudy, demand driven researches using latest IT findings, Mashhad, Khavran university.
20. Nami, S. & Shajari, M. (2018). Cost-sensitive payment card fraud detection based on dynamic random forest and k-nearest neighbors. Expert Systems with Applications.
21. Ngai, E. W. Hu, Y. Wong, Y. H. Chen, Y. & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems, 50(3),* 559-569.
22. Tolouie Ashalghi, A. Maghdori Sharibani, F., & Daneshgar, F. (2008). Credit scoring for credit cards using support vectore machines. *2nd confrance of electronic city, ICTRC,* 12-27. (In Persian)
23. Pirayesh, R. Jabari, A., & Bigdeli, Z. (2015). Corporate credit scoring of saderat bank using logistic regression, MSC dissertation, Sofi razi university. (In Persian)
24. Ravisankar, P., Ravi, V., Rao, G. R., & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems, 50(2),* 491-500.
25. Repousis, S. (2016). Using Beneish model to detect corporate financial statement fraud in Greece. *Journal of Financial Crime, 23(4),* 1063-1073.
26. Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications, 40(15),* 5916-5923.
27. Shin, K.-S., Lee, T & Kim, H.-J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert Systems with Applications,* 127-135.
28. Tehrani, R. & Falahshams, M. F. (1384). Desigining and developing a credit risk model in banking sector, Shiraz social and human science, 29-43. (In Persian)
29. Tousi, M. A., Sadatrasoul, S. M., Shafia, S,. (1400), Identifying the Factors That Affect the Customer Experience and Customer Satisfaction Impact on Repurchasing Behaviors in Online Retailers in IRAN, . Chasm andaz modirat-e-sanaati, 41, 271-293.
30. Verikas, A., Kalsyte, Z., Bacauskiene, M., & Gelzinis, A. (2010). Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: a survey.
31. Zakaryazad, A., & Duman, E. (2016). A profit-driven Artificial Neural Network (ANN) with applications to fraud detection and direct marketing. *Neurocomputing, 175,* 121-131.
32. Zaringhalalm, Z. & Nazemi Hrandi, N. (2003). Banking customers credit scoring using datamining, 2th conference on new studies in computer science and technology, 42-53, (in Persia