Journal of Accounting Knowledge
Shahid Bahonar University of Kerman
Iranian Accounting Association
Print ISSN 2008-8914
Online ISSN 476-292X

New Approach to Predicting and Detecting Financial Statement Fraud, Using the Bee Colony

Elaheh Tashdidi[*]
Sahar Sepasi (Ph.D)[**]
Hossein Etemadi (Ph.D)[***]
Adel Azar (Ph.D)[****]

Abstract

Objective: Considering complex financial plans to conceal fraud in financial statements, the development of fraud detection methods can be regarded as solution for this problem. The present study uses the bee algorithm to develop methods for fraud detection in financial statements.

Method: Three methods of bee algorithm, genetic algorithm and logistic regression have been used to study the subject. The statistical sample consists of 120 companies accepted in the Tehran Stock Exchange (60 companies are suspected of fraud and 60 ones are not suspected) for the period 1396-1385. The companies were suspected of fraud, based on 1) revised audit opinion after unacceptable expression, 2) existence of significant annual revisions, and revised financial statements for inventories and other assets; 3) existence of tax disputes with the tax area, according to notes on income tax filing, general tax filings and conditioned clauses in audit reports. Following the use of cross-entropy, 16 financial ratios were introduced as the potential predictors of fraudulent financial reporting.

Result: The results showed that the bee algorithm method with prediction accuracy of 82.5% has better performance in identifying suspicious companies in fraudulent financial statements than the other two methods.

Conclusion: The results of the research indicate that the proposed method of this study compared to other methods has higher rate of prediction accuracy, lower error rate and relatively good speed.

Keywords: *Detecting of Fraud in Financial Statements, Bee Algorithm, Mutual Entropy, Financial Ratios*.

Citition: Tashdidi, E., Sepasi, S., Etemadi, H., Azar, A. (2019). Provide a new approach to predict and detect financial statements fraud by using the bee colony. *Journal of Accounting Knowledge*, 10(3), 139-167.

[*] Ph.D Student in Accounting, Tarbiat Modares University, Tehran, Iran.
[**] Associate Professor of Accounting, Tarbiat Modares University, Tehran, Iran.
[***] Professor of Accounting, Tarbiat Modares University, Tehran, Iran.
[****] Professor of Industrial management, Tarbiat Modares University, Tehran, Iran.
Corresponding Author: Sahar Sepasi (Email: sepasi@modares.ac.ir).
Submitted: 17 February 2019 Accepted: 5 May 2019
DOI: 10.22103/jak.2019.13616.2927

www.SID.ir

Archi

مجله دانش حسابداری
شاپا چاپی ۲۰۰۸-۸۹۱۴
شاپا الکترونیکی ۲۴۷۶-۲۹۲X
دانشگاه شهید باهنر کرمان
انجمن حسابداری ایران

دانشکدهٔ مدیریت و اقتصاد دانشگاه شهید باهنر کرمان
انجمن حسابداری ایران
مجلهٔ دانش حسابداری
دورهٔ دهم، شمارهٔ ۳
پاییز ۱۳۹۸، پیاپی ۳۸
صص. ۱۳۹ تا ۱۶۷

ارائه رویکردی نوین در پیش‌بینی و کشف تقلب صورت‌های مالی با استفاده از الگوریتم زنبور عسل

الهه تشدیدی *
دکتر سحر سپاسی **
دکترحسین اعتمادی ***
دکتر عادل آذر ****

چکیده

هدف: با توجه به اینکه برای کتمان تقلب در صورت های مالی از طرحهای پیچیده و سازمان یافته استفاده می‌شود، توسعه روش‌های کشف تقلب می‌تواند به عنوان راهکاری مورد توجه قرار گیرد. لذا، پژوهش حاضر با استفاده از الگوریتم زنبورعسل به توسعه روشهای کشف تقلب در صورت های مالی پرداخته است.

روش: برای بررسی موضوع سه روش الگوریتم زنبور عسل، الگوریتم ژنتیک و رگرسیون لوجستیک به کارگرفته شده است. نمونه آماری این مطالعه متشکل از ۱۲۰ شرکت پذیرفته شده در بورس اوراق بهادار (۶۰ شرکت مشکوک به تقلب و ۶۰ شرکت غیر متقلب) برای دوره زمانی ۱۳۹۶-۱۳۸۵

* دانشجوی دکترای حسابداری، دانشگاه تربیت مدرس، تهران، ایران.
** دانشیار حسابداری، دانشگاه تربیت مدرس، تهران، ایران.
*** استاد حسابداری، دانشگاه تربیت مدرس، تهران، ایران.
**** استاد مدیریت صنعتی، دانشگاه تربیت مدرس، تهران، ایران.
نویسندهٔ مسئول مقاله: سحر سپاسی (رایانامه: sepasi@modares.ac.ir).

تاریخ دریافت: ۹۷/۱۱/۲۸ تاریخ پذیرش: ۹۸/۲/۱۵
DOI: 10.22103/jak.2019.13616.2927

www.SID.ir

Archive of SID

است. شرکت‌های مشکوک به تقلب بر مبنای ۱) اظهارنظر تعدیل شده حسابرسی، ۲) وجود تعدیلات سنواتی با اهمیت و صورت های مالی تجدید ارائه شده در مورد موجودی‌ها و سایر دارایی‌ها و... ۳) وجود اختلافات مالیاتی با حوزه مالیاتی طبق یادداشت ذخیره مالیات بر درآمد و پرونده مالیاتی و بند شرط گزارش حسابرسی انتخاب شدند. پس از استفاده از آنتروپی متقابل، ۱۶ نسبت مالی به عنوان پیش بینی‌کننده‌های بالقوه گزارشگری مالی متقلبانه معرفی شدند.

یافته‌ها: یافته‌های پژوهش نشان داد که روش الگوریتم زنبور عسل با دقت پیش‌بینی ۸۲/۵ درصد نسبت به دو روش الگوریتم ژنتیک با دقت ۷۷/۵ درصد و رگرسیون لوجستیک با دقت ۷۲/۵ درصد، از عملکرد بهتری جهت شناسایی شرکت‌های مشکوک به تقلب در صورت‌های مالی برخوردار است.

نتیجه‌گیری: نتایج پژوهش حاکی از آن است، روش پیشنهادی این پژوهش در مقایسه با دیگر روش‌های تکاملی، از دقت پیش‌بینی بالاتر، درصد خطای کمتر و سرعت نسبتاً خوبی برخوردار است.

واژه‌های کلیدی: *کشف تقلب در صورت‌های مالی، الگوریتم زنبور عسل، آنتروپی متقابل، نسبت‌های مالی.*

استناد: تشدیدی، الهه؛ سپاسی، سحر؛ اعتمادی، حسین؛ آذر، عادل. (۱۳۹۸). ارائه رویکردی نوین در پیش‌بینی و کشف تقلب صورت‌های مالی با استفاده از الگوریتم زنبور عسل، *دانش حسابداری*، ۱۰(۳)، ۱۶۷-۱۳۹.

مقدمه

حسابداری، سیستمی اطلاعاتی است که محصول آن صورت‌های مالی است. بیانیه شماره یک چارچوب نظری هیئت تدوین استانداردهای حسابداری، تاکید ویژه‌ای بر سرمایه گذاران به عنوان استفاده کنندگان از صورت‌های مالی دارد. سرمایه گذاران با کسب اطلاعات از طریق صورت‌های مالی در تحلیل خود تجدید نظر نموده و اقدام به

www.SID.ir

Archive of SID

تصمیم‌گیری می‌نمایند. لذا، اعتماد به صحت اطلاعات، سنگ بنای بازارهای سرمایه است و گزارشگری مالی متقلبانه این اعتماد را خدشه دار می‌کند (رضایی و رایلی[1]، ۲۰۱۰).

آثار زیان باری که رسوایی‌های مالی در سال‌های اخیر به وجود آورده است، توجه به بحث تقلب و به ویژه تقلب در صورت‌های مالی را اجتناب ناپذیر کرده است. رسوایی‌های مالی نه تنها هزینه‌ها و زیان‌های اقتصادی مانند متضرر کردن اعتباردهندگان، سرمایه گذاران و سهامداران داشته، بلکه هزینه‌های سیاسی، قضایی و ... را نیز به بار آورده است. رضایی و رایلی (۲۰۱۰) معتقدند گستردگی و وسعت صورت‌های مالی متقلبانه، نگرانی‌هایی را در سطح عمومی پدید آورده است و اعتماد عمومی به فرآیند گزارشگری مالی و عملکرد حسابرسی را خدشه دار کرده است.

در کشور ما نیز با توجه به افزایش تعداد شرکت‌های پذیرفته شده در بورس اوراق بهادار، روند خصوصی سازی، رشد سرمایه و ضرورت انتشار عمومی صورت‌های مالی (اعتمادی و زلقی، ۱۳۹۲) پیش‌بینی احتمال ارتکاب تقلب در اطلاعات مالی انتشار یافته یک ضرورت است و از اهمیت ویژه‌ای برخوردار است. علی‌رغم تاکید استاندارد ۲۴۰ بر توجه حسابرس به مفهوم تقلب، این استاندارد الگوی جامعی در خصوص پیش‌بینی تقلب ارائه نکرده است. این استاندارد با اشاره به عواملی مانند به کارگیری قضاوت، استفاده از رسیدگی نمونه‌ای، محدودیت‌های ذاتی سیستم کنترل داخلی و این واقعیت که بیشتر شواهد در دسترس حسابرس ماهیت متقاعدکننده دارد نه قطعی، کشف همه تحریفهای با اهمیت را تضمین نمی‌کند و از آنجا که برای کتمان تقلب از طرحهای پیچیده و سازمان یافته استفاده می‌شود، خطر کشف نشدن تقلب افزایش می‌یابد. در نتیجه توسعه روش‌های کشف تقلب می‌تواند به عنوان راهکاری جهت کاستن از محدودیت‌های موجود مورد توجه قرار گیرد. پژوهش حاضر در تلاش است با تحلیل روش‌های کمی (رگرسیون لوجستیک، الگوریتم ژنتیک و الگوریتم زنبور عسل)، کارایی روش‌های کشف تقلب در صورت‌های مالی را ارزیابی نماید و به سؤال زیر پاسخ دهد:

www.SID.ir

Archive of SID

«کدام روش برای کشف تقلب در ارائه صورت‌های مالی مناسب‌تر است؟»

لذا بر اساس سؤال پژوهش می‌توان فرضیه زیر را مطرح کرد:

به منظور کشف احتمال تقلب در صورت های مالی، الگوریتم زنبورعسل در مقایسه با روش رگرسیون لوجستیک و الگوریتم ژنتیک اثر بخش تر است.

مبانی نظری

تقلب و انواع آن

طبق بند ۴ بخش ۲۴۰ استانداردهای حسابرسی ایران، تقلب عبارت است از هرگونه اقدام عمدی فریب کارانه یک یا چند نفر از مدیران، کارکنان یا اشخاص ثالث، برای برخورداری از یک مزیت ناروا یا غیرقانونی و در بند ۹ همین استاندارد تقلب و اشتباه را از یکدیگر تفکیک و عنوان می‌کند که تنها ویژگی متمایزکننده تقلب از اشتباه، قصد و نیت است و این که اشتباهات، ناشی از اعمال غیرعمدی (سهوی) است (کمیته تدوین استانداردهای حسابرسی، ۱۳۸۸). همچنین بر اساس گزارش انجمن بازرسان رسمی تقلب [۲]، تقلب‌های حرفه‌ای در سه مقوله فساد مالی، سوء استفاده از دارایی‌ها و تقلب در صورت‌های مالی طبقه‌بندی می‌شود (صدیقی کمال، ۱۳۹۲).

۱. فساد مالی تقلبی است که در آن کارکنان یک واحد از موقعیت و نفوذ خود به منظور دستیابی به منافع مستقیم یا غیر مستقیم سوء استفاده می‌کنند مانند: رشوه، انعام غیرقانونی و ... (صدیقی کمال، ۱۳۹۲).

۲. سوء استفاده از دارایی‌ها، شامل دزدی یا استفاده نادرست کارکنان از موجودی‌ها و سایر دارایی‌های یک سازمان است (صدیقی کمال، ۱۳۹۲).

۳. تقلب در صورت‌های مالی عبارت است از ارائه نادرست، حذف اقلام و افشا نکردن کافی اطلاعات به منظور فریب استفاده کنندگان صورت‌های مالی، به خصوص سرمایه گذاران و اعتبار دهندگان (صدیقی کمال، ۱۳۹۲).

لازم به ذکر است براساس دیدگاه انجمن بررسی کنندگان رسمی تقلب [۳] (2012)، سوء استفاده از دارایی‌ها شایع‌ترین نوع تقلب بوده است. اما با وجود اینکه میزان تقلب در

www.SID.ir

Archive of SID

صورت های مالی، نسبت به سوء استفاده از دارایی‌ها و فساد مالی کمتر است، زیان ناشی از این نوع تقلب بیشتر از دو نوع دیگر است. این نوع رسوایی مالی باعث گمراه کردن استفاده کنندگان به ویژه سرمایه گذاران و اعتباردهندگان می‌شود، این در حالی است که هدف اصلی صورت‌های مالی ارائه اطلاعات مفید برای طیف گسترده‌ای از استفاده کنندگان، به ویژه سرمایه گذاران و اعتباردهندگان به منظور گرفتن تصمیم‌های اقتصادی است (انجمن بررسی کنندگان رسمی تقلب، ۲۰۱۲). در نتیجه توجه بیشتر به بحث تقلب به ویژه تقلب صورت‌های مالی، به عنوان زیان بارترین نوع تقلب در سازمانها و شرکت‌ها از اهمیت بالایی برخوردار است. علی‌رغم قوانین سختگیرانه، تقلب صورت‌های مالی در حال رشد است. متقلبین امروزه بسیار نوآورانه، هوشمندانه و سریع عمل می‌نمایند (سادگالی و همکاران[۴]، ۲۰۱۹). لذا، این مقاله به دنبال تجزیه و تحلیل تطبیقی تکنیک‌های شناسایی تقلب صورت‌های مالی و ارائه روشی بهتر است.

نظریه‌های کشف تقلب

نظریه های غالب در کشف تقلب عبارتند از: نظریه قله کوه یخی[۵] و نظریه سوسک حمام[۶]. در نظریه قله کوه یخی، تقلب مانند کوه یخی است که روی آب قرار گرفته، وجوه ساختاری بالای سطح آب و در دید همگان قرار دارد، اما جنبه‌های رفتاری زیر سطح آب پنهان است. طبق این نظریه که یکی از مهم‌ترین نظریه‌های مطرح در موضوع تقلب است، باید به سراغ این نوع تقلب‌ها با نگاه شک‌آمیزی رفت چراکه ممکن است موارد افشا شده مانند قله نوک کوه یخی باشد که بخش کوچکی از کل پیکره تقلب هستند؛ اما بخش اصلی و مهم که حداقل ۹۵ درصد است در زیر آب و همچنان پنهان باشد. چون همانطوری که همه متخصصان امر می‌دانند این نوع تقلب‌ها که به دلیل ضعف قانون‌گذاری یا حفره‌های قانون‌گذاری رخ می‌دهند فقط محدود به یک فرد یا گروه خاصی نمی‌شود و شامل افراد و جریان‌های دیگری هم می‌تواند باشد (مهدوی و قهرمانی، ۱۳۹۶).

نظریه سوسک حمام به وسیله بلکلی (۲۰۰۹) مطرح شد. این نظریه بر تکرارپذیری تقلب تأکید دارد. به عبارتی وجود یک سوسک نشانه وجود سوسک های دیگر است. از

www.SID.ir

Archive of SID

این روش در داده کاوی استفاده می‌شود و آماره‌های سوسک‌ها را بر اساس ۱۰۰ درصد داده‌ها رسیدگی می‌کنند.

روش‌های پیش‌بینی و کشف تقلب در صورت‌های مالی

سازمان‌ها هرگز نمی‌توانند خطر تقلب را از میان بردارند. همواره افرادی یافت می‌شوند که دارای انگیزهٔ تقلب و مستعد انجام آن هستند و افزون بر این زمینهٔ لازم نیز برای تخطی از کنترل‌های لازم فراهم است یا از طریق تبانی با سایرین آنها می‌توانند از کنترلهای داخلی عدول کنند. از این رو احتمال انجام تقلب در همهٔ سازمان‌ها وجود دارد، بنابراین «از دیدگاه فزونی منافع بر منابع[۷]» نباید تلاش کرد تا همهٔ خطرهای تقلب را از میان برداشت. در نتیجه این سؤال مطرح است آیا راهی وجود داردکه از طریق آن بتوان تقلب در صورت‌های مالی را کشف کرد؟ لذا در ادامه روشهای کشف تقلب در صورت‌های مالی تشریح می‌گردد.

روش رگرسیون لوجستیک

رگرسیون لجستیک یک روش آماری است که به محققین امکان ایجاد الگوهای پیش بینی را می‌دهد. این تکنیک برای درک تأثیر چندین متغیر مستقل بر یک متغیر وابسته باینری به کارگرفته می‌شود. این روش تلاش می‌کند تا متغیر وابسته را بر اساس مجموعه متغیرهای مستقل پیش‌بینی کند، اما اگر پژوهشگران متغیرهای مستقل نادرست را در این روش استفاده ننمایند، الگو ارزش پیش‌بینی بسیار کمی خواهد داشت. هم چنین روش مذکور برای داده‌های پیوسته و خطی قابلیت پیش‌بینی بالاتری نسبت به داده‌های گسسته و غیر خطی دارد (رانگاناتان و همکاران[۸]، ۲۰۱۷). لذا، با توجه به اینکه در موضوع تقلب صورت‌های مالی، متغیرهای مستقل تاثیرگذار، دقیقاً مشخص نمی‌باشند امکان پیش‌بینی درست شرکت‌های متقلب نیز کاهش می‌یابد و این روش به مرور زمان، توسط روش‌های پیشرفته دیگر جایگزین شد.

www.SID.ir

Archive of SID

الگوریتم‌های تکاملی[۹] و الگوریتم ژنتیک[۱۰]

با توجه به محدودیت‌های روشهای خطی در موضوع کشف تقلب، پژوهشگران به دنبال روشهای پیشرفته‌تر و دقیق‌تر هستند. الگوریتم تکاملی الگوریتمی است که جنبه‌های انتخاب طبیعی و تداوم هماهنگی را ترکیب می‌کند. الگوریتم ژنتیک نوع خاصی از الگوریتم‌های تکاملی است که از تکنیک‌های زیست‌شناسی فراگشتی مانند وراثت، جهش زیست‌شناسی و اصول انتخابی داروین برای یافتن فرمول بهینه جهت پیش‌بینی یا تطبیق الگو استفاده می‌کند. در این الگوریتم کاندیداهای بهتر برای ایجاد نسل بعدی انتخاب می‌شوند و این فرایند تا رسیدن به جواب بهینه یا پایان یافتن زمان اجرا ادامه می‌یابد (تقوی و نوبری، ۱۳۸۵). با وجود این، الگوریتم ژنتیک روش غیر قطعی است. بدان معنی که ممکن است با اجرای هر الگوریتم برای یک نمونه مشابه، جواب‌های متفاوتی ارائه گردد. در واقع چندین راه حل برای یک مسئله ارائه می‌دهد. لذا، کیفیت نتایج به شدت به موارد زیر بستگی دارد: الف: جمعیت اولیه. ب: اپراتورهای ژنتیکی (متقاطع، انتخاب، جهش)، ج: احتمال وقوع و جهش (شورمان و همکاران[۱۱]، ۲۰۱۵).

در نتیجه زمانی که الگوریتم ژنتیک یک راه حل بهینه ارائه دهد، نمی‌توان بهینه بودن دومی را اثبات کرد. هم چنین اجرا کردن این الگوریتم به دلیل بالا بودن تعداد نسل‌ها بسیار طولانی و زمان بر است (شورمان و همکاران، ۲۰۱۵). لذا، با توجه به محدودیت‌های مطرح شده در اجرای این الگوریتم، الگوریتم‌های اقتباس شده از فرآیندهای زیستی[۱۲] ارائه شدند.

الگوریتم‌های اقتباس شده از فرآیندهای زیستی

الگوریتم‌های اقتباس شده از فرآیندهای زیستی بنیان خود را از منابع گوناگونی مانند عملکرد نورون‌ها (شبکه عصبی)، فرآیند تکامل در جهان، فعل و انفعالات بین ارگانیسم‌ها و ... اقتباس نموده‌اند و در حل مسائل موجود از آنها استفاده می‌کنند. از این نوع الگوریتم‌ها بیشتر در حل مسائل پیش‌بینی و طبقه‌بندی استفاده می‌شود و به ویژه در الگو سازی مالی و سیستم‌های تجاری کاربرد دارند (گرجی، ۱۳۸۸). شبکه‌های عصبی

www.SID.ir

Archive of SID

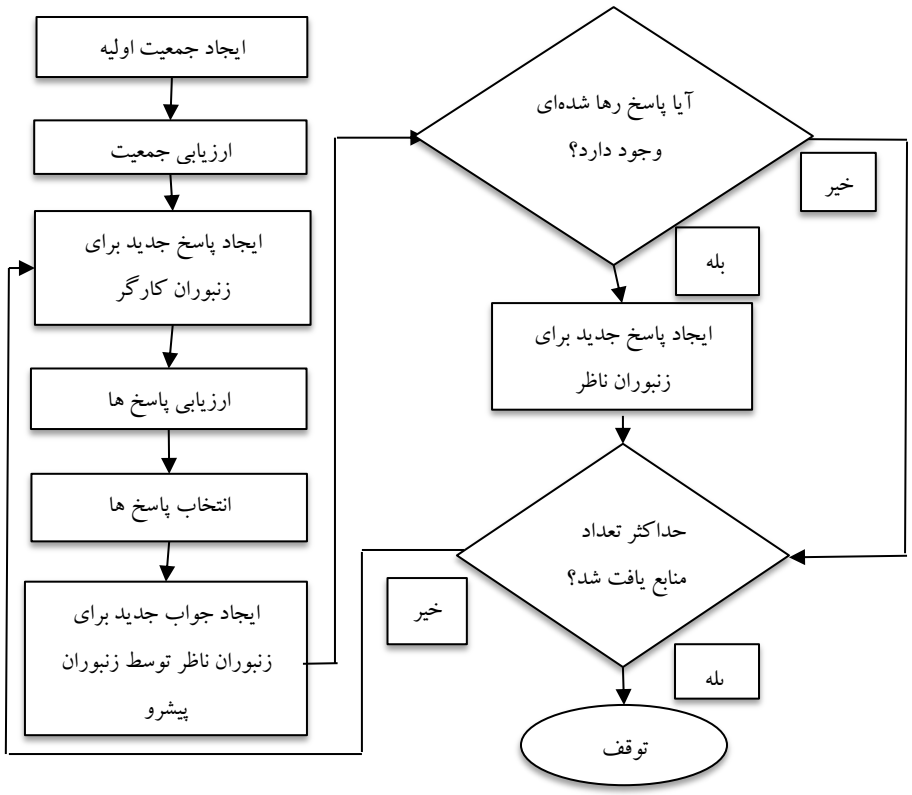
مصنوعی[۱۳]، محاسبات تکاملی[۱۴] و سیستم‌های اجتماعی[۱۵] سه گروه مربوط به الگوریتم‌های اقتباس شده از فرآیندهای زیستی هستند.

الگوهای اجتماعی از رفتار گروهی موجودات زنده اقتباس شده‌اند. سه نوع شناخته شده از این الگوها از رفتار اجتماعی پرندگان، ماهی‌ها و حشرات اجتماعی مانند مورچگان الهام گرفته شده‌اند. این الگوریتم‌ها عبارتند از الگوریتم بهینه یابی دسته ذرات [۱۶] و الگوریتم مورچگان[۱۷]، الگوریتم زنبور عسل[۱۸].

الگوریتم زنبور عسل شامل گروهی مبتنی بر الگوریتم جستجو است که اولین بار در سال ۲۰۰۵ میلادی توسعه یافت. این الگوریتم شبیه‌سازی رفتار جستجوی غذای گروه‌های زنبور عسل است. الگوریتم زنبور عسل تقلید استراتژی جستجوی غذای زنبورعسل به دنبال بهترین راه حل برای مشکل بهینه‌سازی است. هر چرخه جستجو متشکل از پنج روش: استخدام، جستجوی محلی، کوچک شدن محله، متروکه شدن محل و جستجوی کلی است (گونی[۱۹]، ۲۰۱۰). شکل ۱ الگوی مفهومی الگوریتم کلونی زنبور عسل را نشان می‌دهد.

www.SID.ir

Archive of SID



شکل ۱. الگوی مفهومی الگوریتم زنبور عسل (رحیمی و خوانساری، ۱۳۹۳)

پیشینهٔ پژوهش

در ذیل برخی از پژوهش‌های صورت گرفته در زمینه کشف تقلب در صورت‌های مالی ارائه می‌گردد. سادگالی و همکاران (۲۰۱۹) در پژوهشی با عنوان عملکرد تکنیک‌های یادگیری ماشینی در کشف تقلب، به بررسی جامع روش‌های کشف تقلب مالی با استفاده از روش‌های داده کاوی پرداختند. آنها یافتند که روش‌های هوش مصنوعی در مقایسه با روش رگرسیون لوجستیک از عملکرد بهتر و دقیق‌تری در پیش بین تقلب برخوردار است. چن سودان[۲۰](۲۰۱۶) به بررسی کشف تقلب در ارائه گزارشگری مالی با استفاده از رویکرد داده کاوی ترکیبی پرداختند. نتایج پژوهش آنها نشان داد که حدود ۸۷ الی ۹۷ درصد تقلب‌ها از طریق داده کاوی قابل شناسایی است.

www.SID.ir

Archive of SID

اریس و روحنا عثمان (۲۰۱۵) در مطالعه‌ای با عنوان کشف صورت‌های مالی متقلبانه با استفاده از روش‌های آماری نشان دادندکه شرکت‌هایی با کارمندان کمتر از ۱۰۰ نفر نسبت به شرکت‌های بزرگ، موارد تقلب بیشتری را تجربه کرده‌اند. تانگوت وکلکارنی[۲۱] (۲۰۱۵) در مطالعه‌ای با عنوان کشف تقلب صورت‌های مالی با استفاده از روش داده کاوی و تحلیل عملکرد، به این نتیجه رسیدند که اطلاعات موجود در صورت‌های منتشرشده دارای شاخص‌های متقلبانه هستند. همچنین آنها نشان دادند که شبکه عصبی نسبت به الگوریتم‌های دیگر دقت بالاتری در کشف تقلب دارد. کات سیانتیس (۲۰۰۷) عملکرد وکارایی روش‌های الگوریتم مورچگان، رگرسیون لوجستیک، شبکه عصبی مصنوعی، درخت تصمیم‌گیری و شبکه گمانه‌زنی بیزی در کشف تقلب در صورت‌های مالی را بررسی کردند. طبق نتایج الگوریتم مورچگان نتایج عملکرد موفق‌تری نسبت به سایر روش‌ها داشت.

مهدوی و قهرمانی (۱۳۹۶) در پژوهشی به ارائه الگویی برای کشف تقلب به وسیله حسابرسان با استفاده از شبکه عصبی مصنوعی پرداختند. نتایج پژوهش نشان داد که الگوی شبکه عصبی مصنوعی طراحی شده با ۹ نرون در لایه پنهان دارای دقت ۸۶,۹ درصد توانایی شناسایی شرکت‌های متقلب و غیرمتقلب را دارد. هاشمی و حریری (۱۳۹۶) در پژوهشی به ارزیابی توانایی قانون بنفورد در شناسایی و پیش‌بینی کشف تقلب مالی در شرکت‌های پذیرفته شده در بورس پرداختند. نتایج آزمون فرضیه آنها نشان داد، توزیع ارقام صورت‌های مالی از قانون بنفورد تبعیت می‌کند در واقع می‌توان از این قانون برای بررسی سوء جریان در صورت‌های مالی استفاده کرد.

مرادی و همکاران (۱۳۹۳) به روشی توصیفی – پیمایشی به شناسایی عوامل مؤثر بر احتمال وقوع تقلب در گزارشگری مالی پرداختند. نتایج پژوهش نشان می‌دهد که بین عوامل خطر مرتبط با شرایط بازار و صنعت، ویژگی‌های عملیاتی، نقدینگی و ثبات مالی با احتمال وقوع تقلب، رابطهٔ معنادار وجود دارد. فرقاندوست و همکاران (۱۳۹۳) پژوهشی در

www.SID.ir

Archive of SID

زمینه ارتباط مدیریت سود و امکان تقلب در صورت‌های مالی انجام دادند. پژوهش ایشان با استفاده از روش رگرسیون لوجستیک، به این نتیجه دست یافت که در شرکت‌های با سابقه مدیریت سود، احتمال تقلب در صورت‌های مالی وجود دارد. اعتمادی و زلقی (۱۳۹۲) به تقلب در صورت‌های مالی در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از برخی نسبتهای مالی مرتبط پرداختند. آنها ۹ نسبت مالی را به عنوان پیش‌بینی کننده‌های انتخاب و از روش رگرسیون لوجستیک جهت تدوین الگویی برای شناسایی عوامل مرتبط با تقلب استفاده کردند. این الگو در طبقه‌بندی صحیح نمونه مورد نظر در این پژوهش از نرخ دقت ۸۳/۸ درصد برخوردار بوده است.

صفرزاده (۱۳۸۹) با استفاده از تحلیل لاجیت در داده‌های مقطعی به بررسی نقش داده‌های حسابداری در ایجاد یک الگو برای کشف عوامل مرتبط با تقلب در گزارشگری مالی پرداخت. نتایج تحقیق حکایت از عملکرد مناسب الگو در طبقه‌بندی شرکت‌های نمونه داشت به گونه‌ای که درصد صحت طبقه‌بندی الگو از ۸۲/۹۸ درصد تجاوز نمود. طبق پژوهش کات سیانتیس (۲۰۰۷) و سادگالی و همکاران(۲۰۱۹)، در مورد پیش‌بینی و کشف تقلب در صورت‌های مالی، استفاده از شبکه عصبی مصنوعی و الگوریتم ژنتیک نتایج بهتری نسبت به سایر روش‌ها ارائه داده‌اند. با پیشرفت تکنولوژی طرح‌های تقلب نیز پیچیده و کشف آنها مشکل‌تر شده است، در نتیجه نیاز به روش‌های جدیدی است که بتواند احتمال خطا در کشف تقلب را کاهش دهند. در این پژوهش سعی بر آن است که با استفاده از الگوریتم کلونی زنبور عسل، روش‌های کشف تقلب را بهبود و روشی نوین را ارائه دهیم.

روش پژوهش

از آنجایی که پژوهش حاضر به ارزیابی روش‌ها و توسعه روشهای کشف تقلب در صورت‌های مالی می‌پردازد، از بعد هدف از نوع پژوهش‌های کاربردی و از بعد ماهیت و روش نیز از نوع پژوهش‌های توصیفی است.

www.SID.ir

Archive of SID

قلمرو پژوهش و جامعه آماری

پژوهش حاضر از لحاظ مکانی شامل شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است و قلمرو زمانی آن از تاریخ ۸۵/۰۱/۰۱ الی ۹۶/۱۲/۲۹ است. جامعه آماری مورد بررسی عبارت است از شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. نمونه آماری شامل شرکت‌هایی است که شرایط زیر را دارا باشند.

۱. از تاریخ ۸۵/۰۱/۰۱ به بعد عضو بورس اوراق بهادار تهران باشند.
۲. سال مالی آنها منتهی به ۱۲/۲۹ هر سال باشد.
۳. در قلمرو زمانی پژوهش، سال مالی خود را تغییر نداده باشند.
۴. اطلاعات آنها قابل تهیه و در دسترس باشد.
۵. جزو شرکت‌های واسطه مالی، سرمایه گذاری و بانک‌ها نباشند.

جهت تشخیص و طبقه‌بندی واحدهای اقتصادی به شرکت‌های متقلب یا غیر متقلب در گزارشگری مالی، با توجه به استانداردهای شماره ۲۴۰ و ۴۵۰ حسابرسی ایران، می‌توان بر مبنای مقدار و محتوای ویژگی‌های داده‌های صورت‌های مالی حسابرسی شده توسط حسابداران رسمی یا سازمان حسابرسی، معیارهایی را برای محسوب نمودن تحریف ارائه نمود (قادری و قادری، ۱۳۹۶). بنابراین، با توجه به پژوهش‌های گذشته (اعتمادی و زلقی، ۱۳۹۲؛ مرادی، و همکاران، ۱۳۹۳؛ صفرزاده، ۱۳۸۹؛ فرقاندوست حقیقی و همکاران، ۱۳۹۳) معیارهایی زیر درنظرگرفته شده است:

۶. اظهارنظر تعدیل شده حسابرسی،
۷. وجود تعدیلات سنواتی با اهمیت و صورت‌های مالی تجدید ارائه شده در مورد موجودی‌ها و سایر دارایی‌ها و... و
۸. وجود اختلافات مالیاتی با حوزه مالیاتی طبق یادداشت ذخیره مالیات بر درآمد و پرونده مالیاتی و بند شرط گزارش حسابرسی.

در مورد تشخیص سطح اهمیت تعدیلات سنواتی، از دستورالعمل نشریه ۱۵۰ سازمان حسابرسی استفاده گردید. لذا، در صورتی که هر سه شرط مذکور در ۳ سال پیاپی برای هر

www.SID.ir

Archive of SID

شرکتی صادق بود، شرکت مذکور جزء شرکت‌های مشکوک به تقلب در ارائه صورت‌های مالی در نظر گرفته می‌شود. لذا با توجه به شرایط بالا ۶۰ شرکت به عنوان شرکت مشکوک به تقلب شناسایی شد. پس از اینکه شرکت‌های متقلب مشخص شد، تعدادی شرکت غیرمتقلب نیز از جامعه آماری انتخاب گردید. شرکت‌های غیرمتقلب، دوره زمانی مشابهی با شرکت‌های متقلب دارند و اطلاعات آنها در دسترس است.

لحاظ کردن موارد به انتخاب نمونه‌ای متشکل از ۶۰ شرکت مشکوک به تقلب و ۶۰ شرکت غیرمتقلب منجر گردید. جدول شماره ۱ ویژگی‌های نمونه آماری تحقیق را نشان می‌دهد. همانگونه که جدول شماره ۱ نشان می‌دهد، میانگین سود خالص شرکت‌های مشکوک به تقلب، ۱۶۳/۷۶۷ میلیون ریال و غیرمتقلب ۲۰۶/۶۱۳ میلیون ریال است. همچنین بین پراکندگی هزینه بهره و سود انباشته شرکت‌های مشکوک به تقلب و شرکت‌های غیر متقلب تفاوت قابل ملاحظه‌ای وجود دارد.

جدول۱. ویژگی‌های نمونه آماری پژوهش(ارقام به میلیون ریال)

ویژگی نمونه آماری شرکت‌های مشکوک به تقلب	میانگین	انحراف معیار	کمینه	بیشینه	میانه
کل دارایی	۲/۳۹۹/۴۰۶	۵/۹۷۲/۱۱۸	۱۷/۹۸۶	۷۶/۹۱۹/۵۶۰	۲/۳۹۹/۴۰۶
حقوق صاحبان سهام	۵۷۵/۷۰۴	۳/۰۶۶/۳۵۹	۹۲۳	۴۱/۴۰۷/۸۱۳	۵۷۵/۷۰۴
سود انباشته	۱۰/۷۷۷	۲۰۷/۹۸۳	۰	۶/۸۹۱/۸۱۴	-۵۲/۶۷۳
فروش	۱/۵۱۱/۲۴۳	۳/۳۴۹/۴۳۴	۹/۴۶۶	۴۳/۳۳۹/۵۴۰	۱/۵۱۱/۲۴۳
سود خالص	۱۶۳/۶۷۶	۱/۱۷۹/۱۲۰	۰	۱۳/۵۷۵/۸۶۳	۱۶۳/۶۷۶
هزینه بهره	۱۳۶/۸۶۷	۴۳۵/۳۵۱	۰	۴/۲۳۷/۱۵۸	۱۳۶/۸۶۷
کل دارایی	۱/۷۸۳/۹۲۳	۳/۲۹۶/۹۵۰	۳۱/۰۶۳	۳۳/۲۵۱/۷۹۴	۱/۷۸۳/۸۲۱
حقوق صاحبان سهام	۶۱۹/۱۷۲	۱/۳۸۱/۶۸۳	۱/۵۹۰	۱۱/۰۴۷/۴۷۳	۶۱۹/۱۷۲
سود انباشته	۲۲۹/۴۵۴	۹۱۵/۶۲۲	۰	۵/۰۴۵/۷۱۲	۵۹/۲۶۵
فروش	۹۱۶/۹۴۲	۱/۵۳۳/۳۵۹	۴	۱۴/۸۵۸/۵۶۴	۹۱۶/۹۴۲
سود خالص	۲۰۶/۶۱۳	۶۳۴/۷۳۹	۰	۶/۵۱۶/۲۵۹	۲۰۶/۶۱۳
هزینه بهره	۶۱/۳۶۴	۱۳۱/۹۳۶	۰	۱/۵۹۰/۶۵۱	۶۱/۳۶۴

www.SID.ir

Archive of SID

متغیرهای پژوهش

متغیر وابسته: متغیر وابسته در این پژوهش، تقلب در گزارشگری مالی است که دارای ماهیت کیفی و مقیاس سنجش اسمی است. دراندازه گیری این متغیر، به شرکت‌های متقلب عدد یک و به شرکت‌های غیرمتقلب، عددصفر تخصیص داده می‌شود.

متغیر مستقل: متغیرهای مستقل مورد استفاده در این مطالعه، نسبت‌هایی است که از اقلام ترازنامه و سود و زیان شرکت‌های نمونه محاسبه شده است. در این تحقیق به منظور انتخاب متغیرهای مستقل از شواهد تجربی موجود در این زمینه استفاده شده است برای این منظور، مطالعات انجام شده قبلی (چن سودان، ۲۰۱۶؛ اریس و روحنا عثمان، ۲۰۱۵؛ تانگوت و کلکارنی، ۲۰۱۵؛ کات سیانتیس، ۲۰۰۷؛ مهدوی و قهرمانی، ۱۳۹۶؛ مرادی و همکاران، ۱۳۹۳؛ فرقاندوست و همکاران، ۱۳۹۳؛ اعتمادی و زلقی، ۱۳۹۲؛ صفرزاده، ۱۳۸۹) دربارهٔ موضوع کشف تقلب در گزارشگری مالی بدقت مورد بررسی قرار گرفت. بررسی اولیه، ۲۳ نسبت مالی (که اطلاعات مرتبط با محاسبه آنها در ایران وجود داشت) را در اختیار ما قرار داد که این نسبت‌ها عبارت بودند از نسبت جاری (CA/CL)، نسبت دارایی جاری به کل دارایی (CA/TA)، نسبت دارایی ثابت به کل دارایی (FA/TA)، نسبت سرمایه درگردش به کل دارایی‌ها (WC/TA)، نسبت کل بدهی به کل دارایی (TD/TA)، نسبت سود انباشته به کل دارایی (RE/TA)، نسبت فروش به کل دارایی (SALE/TA)، نسبت سود خالص به کل دارایی (NI/TA)، نسبت سود ناخالص به کل دارایی‌ها (GP/TA)، نسبت کل بدهی به حقوق صاحبان سهام (TD/TE)، نسبت بدهی بلندمدت به حقوق صاحبان سهام (LTD/TE)، نسبت بهای تمام شده کالای فروش رفته به فروش (CGS/SALE)، نسبت سود ناخالص به فروش (GP/SALE)، نسبت سود خالص به فروش (NI/SALE)، نسبت هزینه‌های عملیاتی به فروش (OE/SALE)، نسبت سود عملیاتی به فروش (OI/SALE)، نسبت سود خالص به حقوق صاحبان سهام (NI/TE)، نسبت هزینه‌های مالی به کل بدهی(IE/TD)، نسبت حسابهای دریافتنی به فروش(REC/SAL)، نسبت موجودیها به

www.SID.ir

Archive of SID

فروش (INV/SAL)، موجودی نقد به کل دارایی‌ها (CASH/TA)، نسبت موجودیها به کل بدهی‌ها (INV/TD)، نسبت سود خالص به سود ناخالص (NP/GP) و نسبت جریان وجه نقدعملیاتی به کل دارایی‌ها (OCF/TA).

با توجه به اینکه در فرایند الگوسازی و داده کاوی همواره اضافه نمودن متغیرهای جدید راهگشا نبوده، چه بسا با انباشت اطلاعات زاید از دقت و کارایی الگو کاسته شود؛ لذا، پژوهش حاضر از آنتروپی متقابل جهت انتخاب بهترین و موثرترین متغیرهای مستقل استفاده نموده است که در ادامه تشریح خواهد شد.

انتخاب متغیر مستقل

در مطالعات اخیر از رویکرد انتخاب ویژگی [۲۲] جهت حذف متغیرهایی با اطلاعات زاید استفاده می‌شود. در این رویکرد بسته به معیارهای مختلف خطی و غیرخطی ویژگی‌هایی که توصیف مناسب‌تری از هدف مورد بررسی (در اینجا احتمال تقلب) داشته و هزینه حضور آنها در الگو کمتر از هزینه حذف آنها از الگو است، در الگوی نهایی باقی می‌مانند. بر این اساس در مطالعات اخیر به کمک تئوری اطلاعات و نظریه آنتروپی کلودشانون[۲۳] (۱۹۴۸) روابط غیرخطی میان متغیرها نیز در فرآیند انتخاب ویژگی لحاظ می‌شود. در این تئوری اطلاعات حشو[۲۴] و آنتروپی دو مفهوم رایج بشمار می‌رود. حشو نتیجه پیش‌بینی‌پذیری زیاد است؛ در حالی‌که آنتروپی نشانه پیش‌بینی‌پذیری اندک است. هدف این تکنیک انتخاب متغیرهای بهینه و موثرتر است. لذا، در ادامه به تجزیه و تحلیل روش مذکور جهت انتخاب متغیرهای مستقل بهینه پرداخته می‌شود.

رابطه زیر نشان دهده اطلاعات متقابل نرمال شده میان دو سری زمانی X و Y است:

$$R = \frac{H(Y) - H(Y/X)}{H(Y)} \qquad (۱)$$

در رابطه فوق پارامتر R مقداری بین ۰ و ۱ می‌پذیرد و نشان دهندهٔ میزان اطلاعاتی درباره متغیر هدف است که توسط متغیر ورودی توضیح داده می‌شود. بر این اساس چنانچه که متغیر X کلیه اطلاعات ضروری برای پیش‌بینی Yرا دارا باشد، در این صورت $H(Y|X)$

www.SID.ir

Archive of SID

برابر با صفر شده و در نتیجه R برابر ۱ خواهد بود. از سوی دیگر چنانچه متغیر X هیچ اطلاعاتی ازمتغیر Y را در بر نداشته باشد، در این صورت $H(Y|X)$ برابر با H(Y) شده و در نتیجه R برابر با ۰ خواهد بود.

$Min\ P(x_1 \dots x_n)$ همان‌گونه که مشاهده می‌شود در رابطه (۱) همبستگی میان متغیرهای ورودی در نظر گرفته نشده است. لذا، بر اساس اصول حداکثر ارتباط و حداقل حشو مسئله بهینه‌یابی چند هدفهٔ زیر جهت جلوگیری از حشو در انتخاب موثرترین ورودی‌های سیستم استفاده می‌شود (پینگ، لانگ و دینگ[۲۵]، ۲۰۰۵).

$$Max\ V(x_1 \dots x_n) \quad (۲)$$

$$Min\ P(x_1 \dots x_n) \quad (۳)$$

در روابط فوق V و P بصورت زیر محاسبه می‌شوند:

$$V = \frac{1}{n}\sum_{i=1}^{n} I(X_i; Y) \quad (۴)$$

$$P = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{i=1}^{n} I(X_i; X_j) \quad (۵)$$

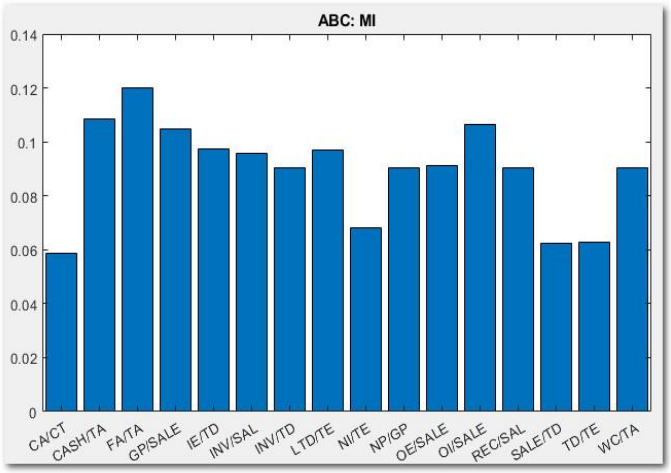
به منظور یافتن پاسخ بهینه مسئله در یک الگوریتم بهینه‌یابی، با فرض $\emptyset = P - V$ می‌توان آن را به‌صورت مسئله حداقل سازی زیر در نظر گرفت:

$$\min \phi(x_1 \dots x_n) \quad (۶)$$

با توجه به آنچه بیان شد در ادامه از میان ۲۳ ویژگی که فضای مسئله آن۲۳! است، به کمک الگوریتم زنبور عسل (ABC) به جستجوی ویژگی‌هایی می‌پردازیم که بالاترین ارتباط و کمترین حشو را ایجاد نمایند. شکل ۲ ویژگی‌های انتخاب شده توسط الگوریتم زنبور عسل را نشان می‌دهد:

www.SID.ir

Archive of SID



شکل ۲. ویژگی‌های انتخاب شده توسط الگوریتم زنبورعسل

همانگونه که نتایج شکل ۲ نشان می‌دهد، از میان ۲۳ ویژگی مورد بررسی ۱۶ ویژگی توانسته‌اند اطلاعات متقابل بالاتری را ایجاد نمایند. لیکن این متغیرها در فرایند الگوسازی نهایی استفاده شدند.

نتایج پژوهش

نتایج رگرسیون لوجستیک

رگرسیون لوجستیک با اختصاص وزن‌هایی به متغیرهای مستقل، رتبه هر یک از شرکت‌های نمونه را پیش‌بینی می‌کند. از این رتبه برای تعیین احتمال عضویت در یک گروه معین (متقلب یا غیرمتقلب) استفاده می‌شود. احتمال موفقیت یا عدم موفقیت (یا هر رتبه کیفی دوگانه دیگر) در این الگو با استفاده از تابع توزیع تجمعی زیر محاسبه می‌شود:

$$p_i = E(Y = 1|X_i) = \left(\frac{1}{1} + e^{-z}\right) \qquad (۷)$$

$$p_i = (1/(1 + e^{-a-\sum_{i=1}^{k} b_i x_i}))$$

در این الگو p_i؛ احتمال وقوع $Y = 1$، X_i؛ متغیر مستقل، α و b_i نیز پارامترهای برآورد الگو هستند. احتمال p_i همواره عددی بین صفر و یک است. اگر z به سمت منفی بی‌نهایت حرکت کند، p_i به سمت صفر میل می‌کند. اگر z به سمت مثبت بی‌نهایت میل کند، p_i به سمت عدد یک میل می‌کند. وقتی که z مساوی با صفر است، احتمال حاصله

www.SID.ir

Archive of SID

مساوی ۰/۵ است. فرم تبعی رابطهٔ فوق را می‌توان با عملیات سادهٔ ریاضی به شکل رابطهٔ زیر بیان کرد:

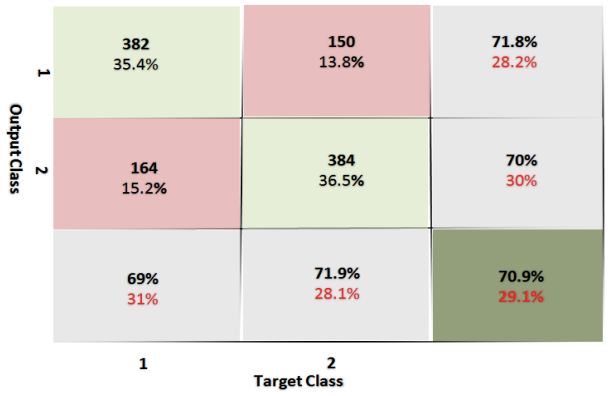
$$Log\left[\frac{p_i}{1-p_i}\right] = Z = a + \sum_{i=1}^{k} b_i x_i \qquad (۸)$$

در این مطالعه متغیر وابسته (z)؛ لگاریتم نسبت احتمال تقلب شرکت‌ها به عدم تقلب آنها است، متغیرهای مستقل x_i نیز مشتمل بر نسبت‌های مالی شرکت‌ها است.

پیش از الگوسازی الگوی لوجستیک، داده‌ها به دو دسته آموزشی و آزمون (تست) تقسیم شدند تا پس از یافتن الگوی نهایی بر اساس داده‌های آموزشی به قابلیت کاربرد آن در آینده به کمک داده‌های تست نیز پرداخته شود. بدین منظور از مجموع کل مشاهدات که ۱۴۴۰ مشاهده است، ۱۰۸۰ مشاهده معادل ۷۵ درصد کل به عنوان داده‌های آموزشی و تعداد ۳۶۰ مشاهده معادل ۲۵درصد کل مشاهدات به عنوان داده‌های تست استفاده می‌شود. جداول زیر نتایج برآورد و پیش‌بینی احتمال تقلب در داده‌های آموزشی و آزمون را نشان می‌دهد:

جدول شماره ۲. نتایج پیش‌بینی احتمال تقلب در داده‌های آموزشی با استفاده از رگرسیون لوجستیک

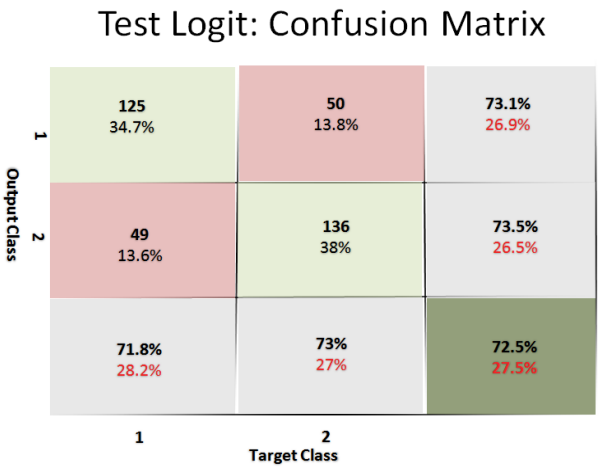
Training Logit: Confusion Matrix



www.SID.ir

Archive of SID

جدول شماره ۳. نتایج پیش‌بینی احتمال تقلب در داده‌های آزمایشی با استفاده از رگرسیون لوجستیک



همانگونه که جداول شماره ۲ و ۳ نشان می‌دهد الگوی لوجستیک شرکت‌های متقلب داده‌های آموزشی را با دقت ۷۰/۹ درصد و شرکت‌های متقلب در داده‌های تست را با دقت ۷۲/۵ درصد صحیح دسته‌بندی کرده است. در مورد داده‌های تست، توزیع‌های تجمعی نشان می‌دهد که ۷۳/۱ درصد شرکت‌ها به عنوان غیرمتقلب پیش‌بینی شده‌اند، در حالی که در واقع این نسبت ۷۱/۸ درصد است. همچنین ۷۳/۵ درصد شرکت‌ها متقلب پیش‌بینی شده‌اند، در حالی که ۷۳ درصد آنها در واقع متقلب بوده‌اند.

نتایج الگوریتم ژنتیک

در این بخش با استفاده از الگوریتم ژنتیک به آموزش دسته‌بندی شرکت‌ها در داده‌های آموزشی پرداخته می‌شود. در این مرحله تابع هدفی که الگوریتم ژنتیک به دنبال بهینه‌یابی آن است، یک تابع کلاس‌بندی کننده لوجستیک به صورت زیر است:

$$L(B.X) = \frac{1}{1+e^{-XB^t}}$$
$$\Gamma = (B.X) = \begin{cases} 1 & if \quad L(B.X) > 0.5 \\ 0 & if \quad L(B.X) \leq 0.5 \end{cases} \qquad (۹)$$
$$\Omega = (B.X.Y) = \left\{\sum_{i=1}^{n} (Y - \Gamma_i(B.X))\right\}/N$$

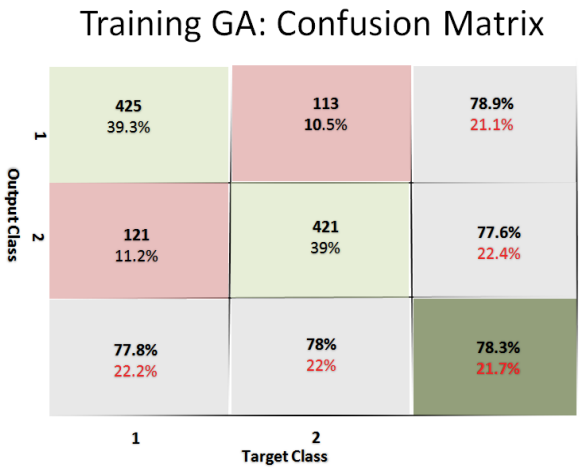
www.SID.ir

Archive of SID

در رابطه ۹ تابع L همان تابع لوجستیک است که به صورت ماتریسی نوشته شده است. همچنین تابع گاما تعیین می‌کند که خروجی به کدام دسته صفر یا یک تعلق دارد و در نهایت تابع اومگا تابع هزینه‌ای است که الگوریتم به دنبال بهینه‌یابی آن است. بنابراین، الگوریتم ژنتیک با تابع هزینه فوق به دنبال کاهش مجموع اختلاف و خطای دسته بندی در داده‌های آموزشی خواهد بود که آن را شایستگی تابع الگوریتم ژنتیک تعریف می‌کنیم.

پس از بررسی و آزمون حالات مختلف الگوریتم با پارامترهای گوناگون پارامترهای نهایی آن به این صورت تعیین شد که تعداد جمعیت ۲۰۰، نرخ همبری[۲۶] ۸۰ درصد و نرخ جهش[۲۷] ۳۰ درصد در نظر گرفته شد، چرا که الگوریتم ژنتیک در این حالت کمترین میزان تابع هزینه و بالاترین دقت پیش‌بینی را ارائه می‌دهد. شکل‌های زیر نتایج برآورد و پیش‌بینی احتمال تقلب در داده‌های آموزشی و آزمون را با استفاده از الگوریتم ژنتیک نشان می‌دهد:

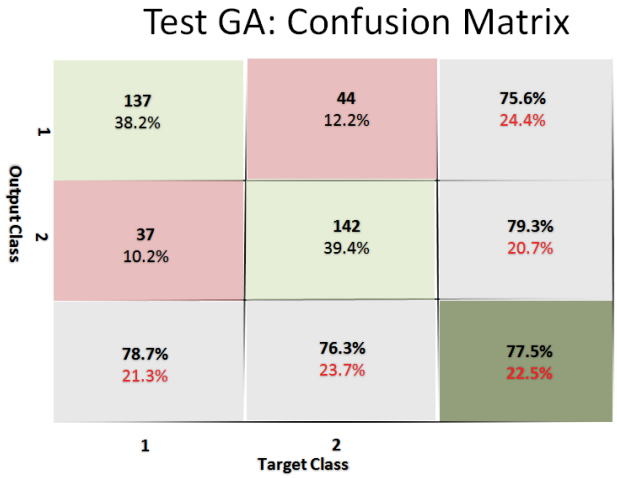
جدول شماره ۴. نتایج پیش‌بینی احتمال تقلب در داده‌های آموزشی با استفاده از الگوریتم ژنتیک



www.SID.ir

Archive of SID

جدول شماره ۵. نتایج پیش‌بینی احتمال تقلب در داده‌های آزمایشی با استفاده از الگوریتم ژنتیک



همان‌گونه که جداول شماره ۵ و ۴ نشان می‌دهد، الگوی الگوریتم ژنتیک شرکت‌های متقلب در داده‌های آموزشی را با دقت ۷۸/۳ درصد و شرکت‌های متقلب در داده‌های تست را با دقت ۷۷/۵ درصد صحیح دسته‌بندی کرده است. در مورد داده‌های تست نیز ۱۳۷ شرکت غیرمتقلب و ۱۴۲ شرکت متقلب را به درستی در دسته‌های خود پیش‌بینی کرده است.

نتایج الگوریتم کلونی زنبور عسل

درالگوریتم کلونی زنبور عسل، موقعیت یک منبع غذایی، یک راه حل مسئله بهینه سازی و مقدار شهد از منبع غذایی، مقدار شایستگی راه حل را نشان می‌دهد. تعداد زنبورهای کارگر برابر با تعداد راه حل‌ها در جامعه هستند. دراین الگو، تولید موقعیت منبع جدید غذا نیز بر اساس یک فرآیند مقایسه موقعیت منبع غذایی است. آنها به طور تصادفی یک موقعیت منبع غذایی را انتخاب می‌کنند و تغییراتی را بر روی یکی از منابع موجود در حافظه خود ایجاد می‌کند. به شرطی که مقدار شهد منبع جدید بیشتر از منبع قبلی حفظ شده در حافظه زنبور عسل باشد، موقعیت جدید را حفظ کرده، موقعیت قبلی را فراموش می‌کند. درغیراین صورت او موضع قبلی را نگه می‌دارد. پس از اینکه فرایند جستجوی

www.SID.ir

Archive of SID

تمام زنبورهای کارگر تکمیل گردید، آنها اطلاعات شهد ازمنابع غذایی (راه حل) و اطلاعات مربوط به موقعیت خود را با زنبورهای تماشاچی در محدوده رقص به اشتراک می‌گذارند. یک زنبور تماشاچی اطلاعات شهد گرفته شده از همه زنبورهای کارگر را ارزیابی می‌کند و یک منبع غذایی با احتمال مربوط به مقدار شهد آن انتخاب می‌شود. همین‌طور در مورد زنبور کارگر، تولید تغییراتی در موقعیت (راه حل) موجود در حافظه خود و مقدار شهد از منبع انتخابی (راه حل) را چک می‌کند و آن شهدی که بیشتر از قبلی باشد را ارائه می‌دهد. زنبورعسل موقعیت جدید را حفظ می‌کند و قبلی را فراموش می‌کند. زنبور تماشاچی یک منبع غذایی را با توجه به مقدار احتمال مرتبط با آن منبع غذایی انتخاب می‌کند، pi ، که با عبارت زیر محاسبه می‌شود:

$$p_i = \frac{\mathrm{fit_i}}{\sum_{\mathrm{n=1}}^{\mathrm{sn}} fit_n} \qquad (۱۰)$$

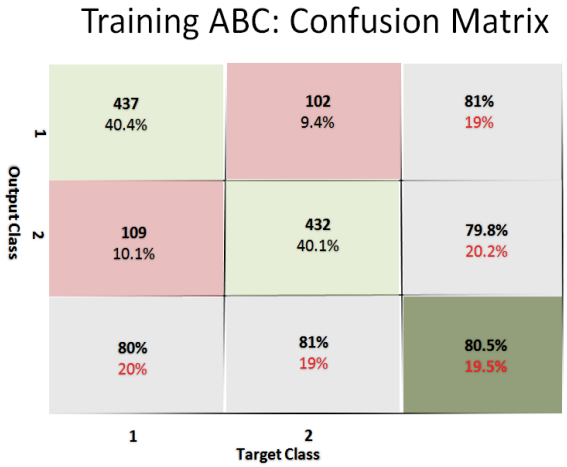
که در آن $\mathrm{fit_i}$ میزان شایستگی از راه حل i که توسط زنبور کارگر ارزیابی شده است که این ارزیابی متناسب با مقدار شهد منبع غذایی در موقعیت i است و SN تعداد منابع غذایی که برابر با تعداد زنبورهای کارگر (BN) است. در این روش، زنبورهای کارگر اطلاعات خود را با زنبورهای تماشاچی تبادل می‌کنند.

دراین مرحله پس از بررسی و آزمون حالات مختلف الگوریتم با پارامترهای گوناگون پارامترهای نهایی آن به این صورت تعیین شد که؛ تعداد منابع غذایی ۲۰۰، تعداد حد فراموشی[۲۸] ۱۷۰۰ و حداکثر ضریب شتاب[۲۹] ۰/۵ در نظر گرفته شد، چرا که الگوریتم کلونی زنبور عسل مصنوعی در این حالت کمترین میزان تابع هزینه و بالاترین دقت پیش‌بینی را ارائه می‌دهد. جداول زیر نتایج برآود و پیش‌بینی احتمال تقلب در داده‌های آموزشی و آزمون را با استفاده از الگوریتم کلونی زنبور عسل مصنوعی نشان می‌دهد:

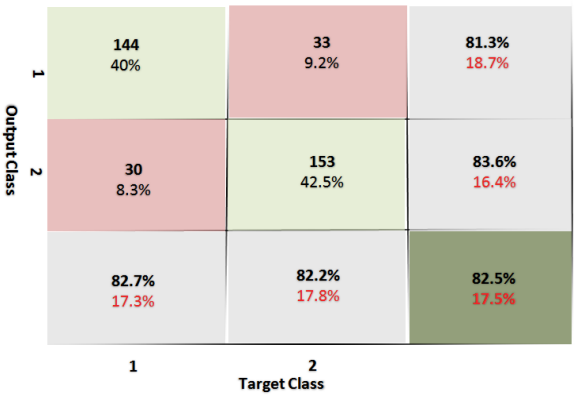
www.SID.ir

Archive of SID

جدول شماره ۶. نتایج پیش‌بینی احتمال تقلب در داده‌های آموزشی با استفاده از الگوریتم زنبورعسل



جدول شماره ۷. نتایج پیش‌بینی احتمال تقلب در داده‌های آزمایشی با استفاده از الگوریتم زنبورعسل



همانگونه که جداول شماره ۷ و ۶ نشان می‌دهد، الگوی الگوریتم کلونی زنبور عسل شرکت‌های متقلب در داده‌های آموزشی را با دقت ۸۰/۵ درصد و شرکت‌های متقلب در داده‌های تست را با دقت ۸۲/۵ درصد صحیح دسته‌بندی کرده است. در مورد داده‌های تست نیز ۱۴۴ شرکت غیرمتقلب و ۱۵۳ شرکت متقلب به درستی در گروه خود پیش‌بینی شده‌اند.

www.SID.ir

Archive of SID

خلاصه و نتیجه‌گیری

همان‌طور که قبلاً بیان شد، سازمان‌ها هرگز نمی‌توانند خطر تقلب را از میان بردارند. همواره فرد یا افرادی یافت می‌شوند که دارای انگیزهٔ تقلب و مستعد انجام آن هستند. ازاین رو احتمال انجام تقلب در همهٔ سازمان‌ها وجود دارد. لذا، با توجه به پیشرفت فناوری اطلاعات و پیچیده‌تر شدن طرح‌های تقلب، لزوم توسعه روش‌های کشف تقلب در صورت‌های مالی نیز حائز اهمیت است. در این پژوهش، پس از جمع آوری اطلاعات شرکت‌های مشکوک به تقلب و شرکت‌های غیر متقلب، سه روش مذکور در نرم افزار متلب پیاده سازی و اجرا شد. جدول شماره ۸ خلاصه نتایج هر یک از روش‌های رگرسیون لجستیک، الگوریتم ژنتیک و الگوریتم زنبور عسل را نشان می‌دهد:

جدول شماره ۸. خلاصه نتایج به کارگیری سه روش کشف تقلب در صورت‌های مالی

روش به کارگرفته شده	الگوریتم زنبور عسل	الگوریتم ژنتیک	رگرسیون لوجستیک
دقت پیش‌بینی در داده‌های آزمایشی	۸۰/۵ درصد	۷۸/۳ درصد	۷۰/۹ درصد
دقت پیش‌بینی در داده‌های آموزشی	۸۲/۵ درصد	۷۷/۵ درصد	۷۲/۵ درصد

با توجه به یافته‌های پژوهش مشخص شد که روش الگوریتم زنبور عسل با دقت پیش‌بینی ۸۲/۵ درصد داده‌های آزمایشی نسبت به دو روش الگوریتم ژنتیک و رگرسیون لوجستیک، از عملکرد بهتری جهت شناسایی شرکت‌های مشکوک به تقلب در صورت‌های مالی برخوردار است. در نتیجه فرضیه پژوهش در مورد اثربخش‌تر بودن روش الگوریتم زنبور عسل، قابل پذیرش است. نتایج مقایسه نشان می‌دهد، روش پیشنهادی این پروژه در مقایسه با دیگر روش‌های تکاملی، از دقت پیش‌بینی بالاتر و درصد خطای کمتری برخوردار است. با توجه به نتایج ارائه شده، الگوریتم زنبور عسل در حل مسئله پیش‌بینی احتمال تقلب درصورت‌های مالی با توجه به مجموعه داده‌های معرفی شده دارای مزایای ذیل است:

www.SID.ir

Archive of SID

۱. با توجه به اجراهای مختلف، شماره نسل همگرایی در رسیدن به جواب‌ها پایین بوده است، یعنی الگوریتم فضای جستجو را به درستی پیمایش نموده و حرکت آن به سمت جواب بهینه با سرعت مطلوبی بوده است.
۲. الگوریتم زنبورعسل نسبت به دو روش دیگر، از سرعت همگرایی بیشتری برای رسیدن به نقطه بهینه برخوردار است.
۳. به دست آوردن دقتی معادل۸۲/۵ درصد برای روش پیشنهادی، نتیجه نسبتاً خوبی تلقی می‌گردد. در مقایسه با الگوریتم ژنتیک با دقت ۷۷/۵ درصد و رگرسیون لجستیک با دقت ۷۲/۵ درصد، می‌توان نتیجه گرفت الگوریتم زنبور عسل برای پیش‌بینی احتمال تقلب در صورت‌های مالی بسیار مناسب است.
۴. روش پیشنهادی مبتنی بر رفتار موجودات زنده بوده و قابلیت استناد با آنچه در واقعیت اتفاق می افتد را دارد.

پژوهش حاضر علاوه بر توسعه روش‌های کشف تقلب در صورت‌های مالی (با معرفی روش الگوریتم زنبور عسل)، از تئوری اطلاعات به منظور انتخاب ویژگی‌های شرکت‌های مشکوک به تقلب از بین ۲۳ ویژگی استخراج شده بهره برده است. در پژوهش‌های پیشین از روش تابع همبستگی متقابل جهت انتخاب موثرترین ویژگی‌ها استفاده شده که این روش تنها بر اساس روابط خطی محاسبه شده و لذا در مواجه با سیستم‌های غیرخطی دارای کاستی‌هایی است، در نتیجه در پژوهش حاضر به منظور رفع کاستی‌های روش تابع هبستگی از تئوری اطلاعات استفاده گردیده است.

پیشنهادهای پژوهش

با توجه به اینکه یکی از عوامل هدر رفت منابع و عدم دستیابی به اهداف سازمانی، تقلب و اشکال متنوع آن در سازمان‌ها است و هم چنین موجب کاهش اعتماد سرمایه گذاران در بلندمدت می‌گردد، به پژوهشگران آتی جهت کمک به سازمان‌های ایرانی و سرمایه گذاران به منظور جلوگیری از ضرر و زیان‌های ناشی از تقلب پیشنهادهای زیر ارائه می‌گردد:

www.SID.ir

Archive of SID

۱. با توجه به اینکه بحث تقلب در مؤسسات مالی و بانک‌های ایرانی در سال‌های اخیر مورد توجه قرار گرفته است، بررسی و شناسایی روش‌های ارتکاب تقلب و ارائه الگوی پیشگیری از آن در مؤسسات مالی از موضوعات حائز اهمیت و قابل بررسی است. لذا، می‌توان با بررسی سیستم‌های کنترلی مؤسسات مالی درگیر تقلب، نقص‌های کنترلی را شناسایی و الگویی جهت پوشش و بهبود این نقص‌ها ارائه نمود.

۲. در همه پژوهش‌های صورت گرفته در زمینه کشف تقلب در صورت‌های مالی، از نسبت‌های مالی به منظور تعیین ویژگی شرکت‌های مشکوک به تقلب استفاده گردیده است. لذا، پیشنهاد می‌گردد پس از بررسی شرکت‌های درگیر تقلب، ویژگی‌های غیرمالی (ویژگی‌های مرتبط با صنعت، روندهای بلندمدت و...) را که می‌تواند نشان‌دهندهٔ حضور تقلب باشد، شناسایی و تأثیر آنها در پژوهش‌های کشف تقلب بررسی شود.

۳. پیشنهاد می‌گردد در جهت توسعه روشهای کشف تقلب، کلیه روش‌های فراابتکاری و اقتباس شده از فرایندهای زیستی در شرایط یکسان با داده‌های مشابه مورد بررسی قرار گیرد تا بهترین روش جهت پیش‌بینی شناسایی گردد.

یادداشت‌ها

1. Rezaee and Riley
2. The Association of Fraud Examiners (ACFE)
3. Association of Certified Fraud Examiners
4. I. Sadgalia, Saela and Benabbo
5. Cost-Benefit
6. Priya, Ranganathan, Pramesh and Rakesh Aggarwa
7. Iceberg Theory
8. Cockroach Theory
9. Evolutionary Algorithms
10. Genetic Algorithm
11. Samer Shorman and Sakinah Ali Pitchay
12. Biologically Inspired Algorithms
13. Artificial Neural Network
14. Evolutionary Computing
15. Social Systems
16. Particle Swarm Optimization
17. Ants Algorithm
18. Bee Algorithm
19. Guney and Onay
20. Chen Suduan
21. Tangod and Kulkarni
22. Feature Selection
23. Claude Shannon entropy
24. Redundancy
25. Peng, Long and Ding
26. Crossover Percentage
27. Mutation Percentage
28. Abandonment Limit
29. Acceleration Coefficient Upper

www.SID.ir

Archive of SID

منابع

اعتمادی، حسین؛ زلقی، حسن. (۱۳۹۲). کاربرد رگرسیون لجستیک درشناسایی گزارشگری مالی متقلبانه. *دانش حسابرسی*، ۵۱، ۲۵-۱.

تقوی، مصطفی ؛ نوبری، نازک. (۱۳۸۵). کاربرد الگوریتم‌های تکاملی در داده کاوی. *اولین همایش بین المللی روش‌های تحقیق در فنون مهندسی*، ۱۵-۲۳.

خواجوی، شکراله؛ ابراهیمی، مهرداد. (۱۳۹۶). مدلسازی متغیرهای اثرگذار برای کشف تقلب در صورت‌های مالی با استفاده از تکنیک داده کاوی، *حسابداری مالی*،۳۳، ۲۵-۲۴.

رحیمی، امیرمسعود؛ رمضانی خوانساری، احسان. (۱۳۹۳). توسعهٔ الگوریتم غذایابی کندوی زنبور عسل برای حل مسئله مسیریابی خودرو، *مهندسی حمل و نقل*، ۶(۱)، ۱۱۵-۱۰۲.

صدیقی کمال، لیلا. (۱۳۹۲). تقلب در صورت‌های مالی براساس گزارش انجمن بازرسان رسمی تقلب، *حسابرس*، ۶۴، ۱۱۸-۱۱۶.

صفرزاده، محمدحسین. (۱۳۸۹). توانایی نسبت‌های مالی در کشف تقلب در گزارشگری مالی:تحلیل لاجیت، *دانش حسابداری*، ۱(۱)،۱۶۳-۱۳۷.

قادری، کاوه؛ قادری، صلاح‌الدین. (۱۳۹۶). تحلیل بیش‌اطمینانی مدیران از عملکرد خود در شرکت‌های متقلب، *بررسی‌های حسابداری و حسابرسی*، ۲(۲۴)، ۲۶۲-۲۴۳.

فرقاندوست حقیقی، کامبیز؛ هاشمی، سیدعباس؛ فروغی دهکردی، امین. (۱۳۹۳). مطالعهٔ رابطهٔ مدیریت سود و امکان تقلب در صورت‌های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، *دانش و پژوهش حسابداری*، ۵۶، ۶۸-۴۷.

کمیته تدوین استانداردهای حسابرسی. (۱۳۸۴). *استانداردهای حسابرسی، سایر خدمات اطمینان بخشی و خدمات مرتبط*، سازمان حسابرسی.

گرجی، عطاالله. (۱۳۸۸). بررسی موانع به کارگیری الگوریتم ژنتیک در انتخاب سبد سرمایه‌گذاری توسط سرمایه‌گذاران در بورس اوراق بهادار تهران، *پایان‌نامهٔ کارشناسی ارشد*، دانشگاه آزاد اسلامی واحد تهران مرکزی.

مرادی، جواد؛ رستمی، راحله؛ زارع، رضا. (۱۳۹۳). شناسایی عوامل خطر مؤثر بر احتمال وقوع تقلب در گزارشگری مالی از دید حسابرسان و بررسی تأثیر آنها بر عملکرد مالی شرکت، *مجله پیشرفت‌های*، ۱، ۱۷۳-۱۴۱.

مهدوی، غلامحسین؛ قهرمانی، علیرضا. (۱۳۹۶). ارائهٔ الگویی برای کشف تقلب به وسیله حسابرسان با استفاده از شبکهٔ عصبی مصنوعی، *دانش حسابرسی*، ۶۷، ۶۰-۵۰.

www.SID.ir

Archive of SID

هاشمی، سیدعباس؛ حریری، امیرسینا. (۱۳۹۶). ارزیابی توانایی قانون بنفورد در شناسایی و پیش‌بینی کشف تقلب مالی، *بررسی‌های حسابداری و حسابرسی*، ۲(۲۴)، ۳۰۲-۲۸۳.

References

Abdul Aris, N., Mohd A., Siti Maznah, O.R., Mohamad Zain, M. (2015). Fraudulent Financial statement detection using statistical techniques: the case of small medium automotive enterprise. *The Journal of Applied Business Research*, 31(4), 115-120.

Association of Certified Fraud Examiners (ACFE). (2012). *Gobal Fraud Study*.

Auditing Standards Setting Committee (2005). *Standard on Auditing, Other Assurance Engagmens, and Related Services*, Audit Organization [In Persian].

Blakley, M. (2009). Fraud detection using a database platform. Available online at: http://www.slideshare.net/mblakley, 24 April 2013.

Dorminey, J.W., Fleming, A.S., Kranacher, M., Riley, R.A. (2012). Financial fraud: A new prespective to an old problem. *The CPA Journal*, 8(3), 61-65.

El-Dyasty, M. (2002). Combining belief functions and neural networks to assess the likelihood of fraud: The case of commercial bank audits, *Working Paper*, Manscmra University, 1-45.

Etemadi, H., Zelaghi, H. (2014). Application of logistic regression in identifying fraud financial reporting. *Auditing Knowledge*, 51, 1-25 [In Persian].

Forghandust Haghighi, K., Hashemi, S.A., Foroughi Dehkordi, A. (2014). Studying the relationship between earnings management and the possibility of fraud in the financial statements of listed companies in Tehran Stock Exchange. *Knowledge and Research Accounting*, 47, 56-68 [In Persian].

Ghaderi, K., Ghadei, S. (2017). The analysis of the executive overconfidence in fraudulent companies. *Journal of Accounting and Auditing Review*, 24(2), 243-262 [In Persian].

Gorji, A. (2009). Investigating the Obstacles Using Genetic Algorithm in Selecting Investment Basis by Investors in Tehran Stock Exchange, *Master's Thesis*, Islamic Azad University, Tehran Central Branch [In Persian].

Guney, K., Onay, M. (2010). Bees algorithm for interference suppression of linear antenna arrays by controlling the phase-only and both the amplitude and phase. *Expert Systems with Applications*, 37, 3129–3135.

Hashemi, S.A., Hariri, A.S. (2017). The analysis of bedford's law ability to identify and predict financial fraud detection., *Journal of Accounting and Auditing Review*, 24(2), 283-302 [In Persian].

Kotsiantis, S. (2007). Forecasting fraudulent financial statements using data mining. *World Enformatika Soc.*, 12, 283-288.

Karaboga, D., Celal, O. (2011). A novel clustering approach: Artificial bee colony (ABC) algorithm, *Applied Soft Computing*, 11, 652- 657.

www.SID.ir

Archive of SID

Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization, *Technical Report*, Computer Engineering Department, Engineering Faculty, Erciyes University.

Khajavi, S., Rezaei, G. (2017). Modeling effective variables for fraud detection in financial statements using data mining techniques. *Journal of Financial Accounting*, 33, 24-26 [In Persian].

Mahdavi, G.H., Ghahremani, A. (2017). Providing a model for fraud investigation by auditors using artificial neural network. *Journal of Audit Science*, 67, 50-60 [In Persian].

Moradi, J., Rostami, R., Zare, R. (2014). Identifying the risk factors affecting the likelihood of fraud in financial reporting from the auditor's point of view and investigating their impact on financial performance. *Journal of Accounting Advancement*, 1, 141-173 [In Persian].

Peng, H., Long, F., Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Machine Intell*, 6, 26-38.

Priya Ranganathan, C.S., Pramesh, L., Rakesh A. (2017). Common pitfalls in statistical analysis: Logistic regression. *Perspect Clin Res*, 8(3), 148–151.

Reddy, Y.V., Sebastian, A. (2009). Parameters for estimation of entropy to study price manipulation in stock markets. Research publication University of Dehli.

Rezaee, Z., Riley, R. (2010). *Financial Statement Fraud: Prevention and Detection*, 2nd edition. Hoboken New Jersey: John Wiley and Sons Inc. xi, 12-284.

Sadgalia, I., Saela, N., Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Second International Conference on Intelligent Computing in Data Sciences* (ICDS 2018), 1986, 92-100.

Safarzadeh, M.H. (2010). The ability of financial ratios to detect fraud in financial reporting: logit analysis. *Accounting Knowledge*, 1(1), 137-163 [In Persian].

Samer, M.S., Sakinah, A.P. (2015). Significance of parameters in genetic algorithm, the strengths, its limitations and challenges in image recovery. *Journal of Engineering and Applied Sciences*, 10(2), 2001-2010.

Sedighi Kamal, L. (2014). Fraud in financial statements according to the report of the association of official insolvency officers. *Auditor*, 64, 116-118 [In Persian].

Stice, J. (1991). Using financial and market information to identify preengagement market factors associated with lawsuits against auditors. *The Accounting Review*, 66(3), 516-533.

Suduan Chen. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. *Springer Plus*, 5(89), 509-535.

Tangod, K., Kulkarni, G. (2015). Detection of financial statement fraud using data mining technique and performance analysis. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(7), 312-320.

www.SID.ir