Studies of ethics and behavior in accounting and auditing
Volume 5, Issue 1, 2025

Detecting Fraudulent Financial Statements Using Deep Learning Algorithms

Javad Sharafkhani[1], Ali Ismailzadeh Moqri[2], Mohammad Ali Bidari[3]

[1] PhD., Student, Department of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran.
Drjavadsharafkhani@gmail.com
[2] Professor, Department of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran
(Corresponding author). aIies35091@gmail.com
[3] Assistant Professor, Department of Accounting, Central Tehran Branch, Islamic Azad University, Tehran, Iran.
bidari@ut.ac.ir

Abstract

Objective: The purpose of this study is to detect fraudulent financial statements in companies listed on the Tehran Stock Exchange using deep learning algorithms.

Method: The study analyzed 1,800 company-years (150 companies over 12 years) of annual financial reports from companies listed on the Tehran Stock Exchange, covering the period from 2012 to 2023. Three deep learning algorithms were employed in this research: support vector machine (SVM), convolutional neural network (CNN), and recurrent neural network (RNN). Additionally, the "two-sample mean comparison test" method was employed to select the final variables for the study and to develop the model.

Findings: The findings indicate that the overall accuracy of the deep learning algorithms is as follows: Support Vector Machine (SVM) at 88.4%, Convolutional Neural Network (CNN) at 81.3%, and Recurrent Neural Network (RNN) at 94.4%. This suggests that the RNN algorithm demonstrates the best performance, while the CNN algorithm exhibits the lowest performance in detecting companies with fraudulent financial statements. In other words, the results indicate that the recurrent neural network (RNN) algorithm is more efficient than other deep learning algorithms. Therefore, among the three algorithms—Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN)—used in companies listed on the Tehran Stock Exchange, the RNN algorithm offers the most effective model for detecting fraudulent financial statements.

Conclusion: The findings of this study offer valuable insights for shareholders, creditors, analysts, and other participants in the capital markets. These insights can enhance the prediction of fraudulent financial statements, minimize errors in decision-making based on financial information, improve the evaluation of company performance, facilitate the adoption of optimal trading strategies, and help identify suitable opportunities for buying and selling stocks based on financial statement data.

Keywords: Fraudulent Financial Statements, Fraud Detection, Deep Learning, Recurrent Neural Network Algorithms.

http://sebaa.journal.qom-iau.ac.ir/

https://doi.org/10.71907/sebaa.2025.10201196032

Received: 2024/12/06 ; **Revised:** 2025/01/17 ; **Accepted:** 2025/02/13 ; **Published online:** 2025/04/07

Publisher: Qom Islamic Azad University **Article type:** Research Article © The Author(s).

مطالعات اخلاق و رفتار در حسابداری و حسابرسی
دوره ۵، شماره ۱، ۱۴۰۴

کشف صورت‌های مالی متقلبانه براساس الگوریتم‌های یادگیری عمیق

جواد شرف خانی[۱]، علی اسماعیل‌زاده مقری[۲]، محمدعلی بیداری[۳]

[۱] دانشجوی دکتری، گروه حسابداری، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران. Drjavadsharafkhani@gmail.com
[۲] استاد، گروه حسابداری، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران **(نویسنده مسئول)**. aIies35091@gmail.com
[۳] استادیار، گروه حسابداری، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران. bidari@ut.ac.ir

چکیده

هدف: هدف پژوهش حاضر کشف صورت‌های مالی متقلبانه براساس الگوریتم‌های یادگیری عمیق در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران می‌باشد.

روش: ۱۸۰۰ سال–شرکت (۱۵۰ شرکت برای ۱۲ سال) از گزارش‌های مالی سالیانه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در طی دوره زمانی ۱۳۹۱ تا ۱۴۰۲ مورد آزمون قرار گرفتند. در پژوهش حاضر از سه الگوریتم یادگیری عمیق (شامل الگوریتم‌های ماشین بردار پشتیبان (SVM)، شبکه عصبی پیچشی (CNN) و شبکه عصبی بازگشتی (RNN)) استفاده شد. همچنین جهت انتخاب متغیرهای نهایی پژوهش از روش «آزمون مقایسه میانگین دو نمونه» جهت ایجاد مدل استفاده شده است.

یافته‌ها: نتایج حاصل از الگوریتم‌های یادگیری عمیق نشان می‌دهد که صحّت کلی الگوریتم‌های SVM، CNN و RNN به ترتیب 88.4٪، 81.3٪ و 94.4٪ می‌باشد که نشان‌دهنده این است که الگوریتم RNN بهترین عملکرد و الگوریتم CNN بدترین عملکرد را در کشف شرکت‌های دارای صورت‌های مالی متقلبانه دارد. به عبارت دیگر، نتایج نشان‌دهنده کاراً بودن الگوریتم شبکه عصبی بازگشتی (RNN) نسبت به سایر الگوریتم‌های یادگیری عمیق است. بنابراین، در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران از بین سه الگوریتم SVM، CNN و RNN، الگوریتم RNN کاراترین مدل را برای کشف صورت‌های مالی متقلبانه فراهم می‌کند.

نتیجه‌گیری: یافته‌های پژوهش حاضر می‌تواند برای سهامداران، اعتباردهندگان، تحلیلگران و سایر مشارکت‌کنندگان در بازار سرمایه، اطلاعات سودمندی را در بهبود پیش‌بینی صورت‌های مالی متقلبانه و کاهش خطا در اتخاذ تصمیمات مبتنی بر اطلاعات صورت‌های مالی، ارزیابی بهتر عملکرد شرکت بر مبنای اطلاعات، اتخاذ استراتژی‌های معاملاتی بهینه و شناسایی فرصت‌های مناسب خرید و فروش سهام بر مبنای اطلاعات صورت‌های مالی فراهم کند.

کلیدواژه‌ها: صورت‌های مالی متقلبانه، کشف تقلب، یادگیری عمیق، الگوریتم شبکه عصبی بازگشتی.

http://sebaa.journal.qom-iau.ac.ir/

استناد به این مقاله: شرف‌خانی، جواد؛ اسماعیل‌زاده مقری، علی؛ بیداری، محمدعلی (۱۴۰۴). کشف صورت‌های مالی متقلبانه براساس الگوریتم‌های یادگیری عمیق. *مطالعات اخلاق و رفتار در حسابداری و حسابرسی*، ۵(۱)، ص۱۱۷-۱۴۸.
https://doi.org/10.71907/sebaa.2025.10201196032

تاریخ دریافت: ۱۴۰۳/۰۹/۱۶ ؛ **تاریخ اصلاح:** ۱۴۰۳/۱۰/۲۸ ؛ **تاریخ پذیرش:** ۱۴۰۳/۱۱/۲۵ ؛ **تاریخ انتشار آنلاین:** ۱۴۰۴/۰۱/۱۸

ناشر: دانشگاه آزاد اسلامی واحد قم **نوع مقاله:** پژوهشی © نویسندگان.

۱. مقدمه

تقلب در گزارشگری مالی در سال‌های اخیر رشد قابل توجهی داشته است، در این راستا، صورت‌های مالی هرچه بیشتر مورد تحریف قرار گرفته‌اند. در مواقعی که صورت‌های مالی حاوی تحریفی با اهمیت باشد، به گونه‌ای که عناصر تشکیل‌دهنده آن صورت مالی بیانگر واقعیت نباشد، گفته می‌شود که تقلب صورت گرفته است (اسپاتیس[1]، ۲۰۰۲). لذا، توجه به اهمیت کشف تقلب در صورت‌های مالی در جهت حمایت از منافع سرمایه‌گذاران، امری بسیار مهم است. متأسفانه تقلب از هر نوع آن و با هر قصدی، تأثیر نامطلوبی بر شرکت خواهد داشت. از این‌رو شناخت و پیشگیری از آن امری ضروری برای شرکت‌ها است. اعمال متقلبانه به زیان‌های مالی بزرگی برای اقتصادها و واحدهای تجاری در سطح جهان منجر شده است. بنابراین، در زمان کنونی، کشف و مبارزه با تقلب، جهت جبران زیان‌های وارده بسیار با اهمیت است. صورت‌های مالی متقلبانه به عنوان حذفیات عمده یا ارائه نادرست در صورت‌های مالی که به دلیل عدم گزارش عمدی داده‌های مالی مطابق با استانداردهای حسابداری پذیرفته شده است، مشخص می‌شود. صورت‌های مالی متقلبانه می‌تواند تأثیرات قابل توجهی بر ذینفعان شرکت‌های متقلب و شرکت‌های غیرمتقلب در صورت عدم شناسایی و پیشگیری به موقع ایجاد کند (ال-بانانی و همکاران[2]، ۲۰۲۰). متأسفانه تشخیص صورت‌های مالی متقلبانه آسان نیست. علاوه‌بر این، حتی در صورت کشف، آسیب قابل توجهی به طور کلی قبلاً رخ داده است. در نتیجه، سیاست‌گذاران، حسابرسان و سرمایه‌گذاران از تکنیک‌های کارآمدتر و مؤثرتری که می‌توانند صورت‌های مالی متقلبانه را شناسایی کنند، سود زیادی خواهند برد. انجمن بازرسان خبره تقلب[3] بیان می‌کند که تقلب در صورت‌های مالی عبارت است از ارائه نادرست عمدی وضعیت مالی یک شرکت از طریق تحریف یا حذف عمدی مبالغ یا افشاء در صورت‌های مالی برای گمراه کردن استفاده‌کنندگان از صورت‌های مالی. براساس گزارش مرکز کیفیت حسابرسی[4]، افراد یا شرکت‌ها به دلایل مختلفی ازجمله مزایای پولی، نیاز به تحقق اهداف مالی کوتاه‌مدت یا سرپوش گذاشتن بر اخبار بد، به دنبال دستکاری صورت‌های مالی هستند. استفاده‌کنندگان خارجی و داخلی صورت‌های مالی، دائماً صورت‌های مالی را زیر سؤال می‌برند و نهادهای نظارتی نمی‌توانند با اطمینان بگویند که صورت‌های مالی معتبر هستند و مطابق با دستورات

1. Spathis
2. El-Bannany
3. The Association of Certified Fraud Examiners (ACFE)
4. The Center for Audit Quality (CAQ)

http://sebaa.journal.qom-iau.ac.ir/

قانونی و اخلاقی رویه‌های حسابداران و حسابرسان تهیه شده‌اند (کولیکووا و ساتداروا[1]، ۲۰۱۶؛ دیبیاک و همکاران[2]، ۲۰۲۲). در نتیجه، کشف تقلب یا فریب به منظور اطمینان از صحّت صورت‌های مالی مهم است. در این زمینه، مطالعه حاضر از اهمیت عملی برای مشاغل و حسابرسان برخوردار است؛ زیرا بازار جهانی شاهد افزایش تقلب در حسابداری مالی است که میلیاردها دلار در سال برای کسب‌وکارها هزینه دارد. آشفتگی مالی تأثیر قابل توجهی بر کسب‌وکارها و اعتباردهندگان یک کشور و در نتیجه بر اقتصاد آن دارد (ال-بانانی و همکاران، ۲۰۲۰؛ اسریدهاران و همکاران[3]، ۲۰۲۰؛ کومار و تریپاتی[4]، ۲۰۲۰). در نتیجه، کشف و پیش‌بینی تقلب حسابداری مالی به موضوعی نوظهور برای مطالعات دانشگاهی و متخصصان صنعت تبدیل می‌شود.

در محیط تجاری مدرن امروزی با پیچیده‌تر شدن سیستم‌ها و فعالیت‌ها و ازدیاد اطلاعات فریبکارانه، فناوری‌های مورد استفاده برای جلوگیری و کشف تقلب نیز به‌روز شده است، بنابراین، توسعه و رشد روش‌های مربوط به داده‌های متنوع ازجمله داده‌کاوی، برای تشخیص تقلب‌های مالی در اولویت قرار دارد. در ایران به موضوع کشف تقلب صورت‌های مالی توجه زیادی نشده است. مسئله گزارشگری مالی متقلبانه در ایران از اهمیتی ویژه برخوردار است. افزایش تعداد شرکت‌های پذیرفته شده در بورس که به منظور جذب منابع مالی به انتشار اوراق بهادار اقدام می‌کنند، تلاش به منظور کاهش مالیات بر درآمد و... ازجمله دلایل اهمیت این موضوع است. بنابراین، در حوزه کشف تقلب، هرچقدر حسابرسان و ذینفعان شرکت به تکنیک‌ها و رویکردهای کارآمدتری ازجمله الگوریتم‌های یادگیری عمیق و رویکردهای کارآمدی چون تکنیک یادگیری عمیق شبکه عصبی بازگشتی دسترسی داشته باشند، قاعدتاً، توانایی و کارایی آنها در حوزه تائیدپذیری اطلاعات صورت‌های مالی را می‌تواند بهبود بخشد.

کشف تقلب در صورت‌های مالی به چند دلیل دشوار است؛ نخست، اعضای گروه مدیریت که عمدتاً مرتکبین اصلی تقلب هستند، به طور معمول تلاش زیادی برای پنهان کردن تقلب در صورت‌های مالی می‌کنند؛ دوم، اعضای گروه مدیریت از اعتماد قابل توجهی برخوردار هستند، به راحتی کنترل‌های داخلی کلیدی را نقض کرده و در نتیجه تقلب در صورت‌های مالی آسان‌تر و کشف آن سخت‌تر می‌شود؛ سوم، مرتکبین تقلب در صورت‌های مالی، اغلب از تبانی و سندسازی برای

1. Kulikova and Satdarova
2. Deebak
3. Sreedharan
4. Kumar

http://sebaa.journal.qom-iau.ac.ir/

ارتکاب و پنهان کردن تقلب استفاده می‌کنند و بیشتر رویه‌های حسابرسی مستقل برای کشف چنین طرح‌هایی از تقلب طراحی نشده‌اند؛ چهارم، داده‌های صورت‌های مالی تا حد زیادی خلاصه و تجمیع شده‌اند که پنهان کردن تقلب را ساده‌تر و کشف آن را با استفاده از روش‌های مدل‌سازی آماری و تحلیلی سخت‌تر می‌کند (وایتینگ و همکاران[1]، ۲۰۱۲). دو جریان متفاوت در متون پژوهشی، پیشینه و مبنایی برای این پژوهش فراهم می‌کند؛ یکی پژوهش‌های انجام شده در رابطه با گزارشگری مالی متقلبانه؛ و دیگری متون مربوط به الگوریتم‌های یادگیری عمیق است. این دو جریان قلمروهای پژوهشی گسترده‌ای هستند که پژوهش‌های متقابل بسیار اندکی در ایران، در این زمینه انجام شده است. پژوهش‌های صورت گرفته نیز با توجه به محدودیت در زمینه دسترسی به داده‌های تقلب، با یکدیگر متفاوت هستند؛ از این‌رو، پژوهش حاضر در تلاش است تا با استفاده از تکنیک یادگیری عمیق (ماشین بردار پشتیبان، شبکه عصبی پیچشی و شبکه عصبی بازگشتی) و با بکارگیری جامع‌ترین متغیرهای اثرگذار بر تقلب در صورت‌های مالی (با بررسی جامع پژوهش داخلی و خارجی انجام شده در حوزه تقلب)، مدلی برای کشف صورت‌های مالی متقلبانه با دقت و صحّت بالا ارائه دهد، که این موضوع در نوع خود منحصربه‌فرد می‌باشد.

۲. مبانی نظری و پیشینه پژوهش

۲-۱. مبانی نظری پژوهش

در سطح جهانی، هزینه صورت‌های مالی متقلبانه، طی سال‌ها به افزایش خود ادامه داده است. مستقیم‌ترین قربانیان، سرمایه‌گذاران و سرمایه‌دارانی هستند که به بهانه‌های واهی به شرکت‌ها وجوه می‌دهند. هزینه‌های بیشتری برای جامعه در سطح کلان وجود دارد، مانند از دست دادن اعتماد به سیستم‌های مالی، حتی اگر تأثیر اقتصادی در سطح شرکت می‌تواند بسیار متغیر باشد. صورت‌های مالی متقلبانه آسیب مالی قابل توجهی به سرمایه‌گذاران، سهامداران و جامعه وارد می‌کند. اکثر تحقیقات موجود در ادبیات صورت‌های مالی متقلبانه از تحلیل رگرسیون سنّتی (اندرو و رابین[2]، ۲۰۲۲؛ پآولو سرگئو گومس ماسدا و ویئیرا[3]، ۲۰۲۲؛ وادوا و کومار[4]، ۲۰۲۰؛ آمر و جربوئی[5]، ۲۰۱۳؛ آلسینقلاوی و آلماری[6]، ۲۰۲۱) استفاده کرده‌اند. یادگیری ماشین که زیرمجموعه هوش مصنوعی است، به علت

1. Whiting
2. Andrew & Robin
3. Paulo Sérgio Gomes Macedo & Vieira
4. Wadhwa & Kumar
5. Ama and Jarboui
6. Alsinglaw & Almari

http://sebaa.journal.qom-iau.ac.ir/

خاصیت اکتشافی خود، بدون هیچ فرض اولیه‌ای شروع به الگوسازی رفتار داده‌ها کرده و به مرور زمان و با جلو رفتن الگوریتم، الگو پررنگ‌تر خواهد شد؛ ساختار غیرخطی و مقاوم این روش، توانایی شبیه‌سازی رفتار محیط واقعی را فراهم کرده است. در یادگیری ماشینی از طریق به‌کارگیری یک‌سری الگوریتم‌های خاص، سعی می‌شود الگوهای پنهان بین داده‌ها کشف شود؛ در صورتی که تعداد متغیرهای مستقل و وابسته زیاد باشد و بین آنها ارتباط خطی برقرار نباشد، این روش ازجمله بهترین گزینه‌ها برای پیش‌بینی است. یادگیری عمیق، زیرشاخه‌ای از یادگیری ماشینی و از حوزه‌های نوین این بخش است. یادگیری عمیق معرف سلسله‌ای از الگوریتم‌هاست که در آن از چندین لایه پردازش اطلاعات به ویژه اطلاعات غیرخطی استفاده می‌شود، تا بهترین ویژگی‌های مناسب، از ورودی خام استخراج شود. از میان تکنیک‌های مختلف یادگیری عمیق که در علوم مختلف کاربردهای فراوانی دارد، پژوهشگران حوزه مالی برای پیش‌بینی تقلب، اغلب از الگوریتم‌های به‌خصوصی نظیر شبکه عصبی بازگشتی[1] و شبکه عصبی پیچشی استفاده کرده‌اند (الغفیلی و همکاران[2]، ۲۰۲۰).

شبکه‌های عصبی مصنوعی به دلیل ماهیت همه‌کاره آن، به طور گسترده برای حل بسیاری از مشکلات استفاده شده است. یک رویکرد ترکیبی، یعنی ترکیبی از متغیرهای تحلیل بنیادی و تکنیکی شاخص‌های بازار سهام، برای پیش‌بینی قیمت سهام با هوش مصنوعی در آینده جهت بهبود روش‌های موجود ارائه شده و حرکت شاخص قیمت سهام با استفاده از دو مدل مبتنی بر شبکه عصبی مصنوعی و ماشین بردار پشتیبان مورد بحث قرار گرفته است. پس از مقایسه عملکرد هر دو مدل، نتیجه این شد که میانگین عملکرد مدل ANN به طور قابل‌توجهی بهتر از مدل SVM است و اینکه شبکه عصبی مصنوعی تناسب بهتری با داده‌ها نسبت به روش‌های مرسوم فراهم می‌کند. این شبکه‌های عصبی به گونه‌ای توسعه یافته‌اند که می‌توانند الگوهایی را از داده‌های نویزدار ا استخراج کنند. ANN ابتدا یک سیستم را با استفاده از نمونه بزرگی از داده‌ها به نام فاز آموزشی، آموزش می‌دهد، سپس شبکه را با داده‌هایی آشنا می‌کند که در مرحله آموزش گنجانده نشده است، این مرحله به عنوان مرحله اعتبارسنجی یا پیش‌بینی شناخته می‌شود. تنها انگیزه این روش، پیش‌بینی نتایج جدید است. این ایده یادگیری از آموزش و سپس پیش‌بینی نتایج در ANN از مغز انسان می‌آید که می‌تواند یاد بگیرد و پاسخ دهد. بنابراین، ANN در بسیاری از حوزه‌ها مورد استفاده قرار گرفته و در اجرای توابع پیچیده در زمینه‌های مختلف، عملکرد موفقیت‌آمیزی داشته است (جی و همکاران[3]، ۲۰۲۰).

1. RNN
2. Alghofaili
3. Jay

http://sebaa.journal.qom-iau.ac.ir/

در حالت کلی، به دلیل پیچیدگی مسئله تقلب و حجم بالای اطلاعات مورد پردازش، اغلب استفاده از یک سیستم ساده برای پیش‌بینی نتایج خوبی به همراه ندارد. به همین دلیل محققان با ارائه مدلی ترکیبی، سعی در ارائه سیستمی با پیچیدگی کمتر و کارایی و دقت بیشتر کرده‌اند. امروزه از الگوهای مختلفی مانند: تکنیک‌های آماری (تحلیل تشخیصی، لوجیت و آنالیز فاکتوری[1]) و تکنیک‌های هوش مصنوعی (شبکه‌های عصبی، درخت تصمیم‌گیری، استدلال مبتنی بر موضوع، الگوریتم ژنتیک، مجموعه‌های سخت، ماشین بردار پشتیبان و منطق فازی) و یا ترکیبی از این دو تکنیک، برای پیش‌بینی تقلب استفاده می‌شود. در اکثر مدل‌های پیش‌بینی‌کننده، سیستم فقط با استفاده از اطلاعات یک شاخص به پیش‌بینی می‌پردازد. در سال‌های اخیر، تعدادی از کارشناسان و محققین سعی کرده‌اند از روش‌های یادگیری ماشین و داده‌کاوی برای انجام تحقیقات در این زمینه به عنوان راهی برای کاهش خطاهای تشخیص استفاده کنند. در حالی که بسیاری از استراتژی‌های تجاری، مبتنی بر دقت صورت‌های مالی هستند، و منابع کافی برای تجزیه و تحلیل همه آنها به طور کامل وجود ندارد. از آنجایی که حسابرسان در چندین نمونه از صورت‌های مالی متقلبانه مقصر شناخته شده‌اند، مدل‌های کشف تقلب مالی معتبر باید به روشی آسان برای حسابرسان، سرمایه‌گذاران، سیاست‌گذاران و سایر ذینفعان ارائه شوند. به دلیل اتکای ذاتی به مفروضات توزیعی محدود، تکنیک‌های پارامتریک مانند رگرسیون لجستیک[2] فاقد کاربرد کلی است که ممکن است رویکردهای غیرپارامتری ارائه کنند (اسچرییبر-گریگوری و بادر[3]، ۲۰۱۸). از سوی دیگر، تحقیقات موجود در مورد روش‌های کمّی کشف تقلب مالی، عمدتاً بر صنعت بانکداری و خدمات مالی متمرکز شده است، که مبتنی بر روی کشف تقلب بیمه و کارت اعتباری است. الگوریتم‌های کمّی برای کشف و پیش‌بینی صورت‌های مالی متقلبانه، باید در پژوهش‌های دانشگاهی مورد توجه قرار گیرد. ادبیات علمی و آکادمیک فعلی در حال حاضر به دنبال قوانین یا طبقه‌بندی‌های بیشتر از داده‌های قبلی، برای دستیابی به هدف پیش‌بینی یا تشخیص است. یادگیری ماشینی با مقدار مناسب داده می‌تواند نتایج دقیق‌تری را در مقایسه با رویکرد سنّتی، پیش‌بینی و طبقه‌بندی کند.

کشف تقلب در صورت‌های مالی در چند دهه گذشته با تاکید بر ناهنجاری‌های حسابداری به طور کلی و به طور خاص بر تقلب در صورت‌های مالی، مورد توجه قرار گرفته است. در حالی که تحقیقات اولیه از تکنیک‌های آماری یا سنّتی استفاده می‌کرد که هم زمان‌بر و هم پرهزینه هستند.

1. Diagnostic analysis, logit and factor analysis
2. LR
3. Schreiber-Gregory and Bader

http://sebaa.journal.qom-iau.ac.ir/

اخیراً با ظهور کلان‌داده و یادگیری ماشین، تمرکز بر کشف تقلب با الگوریم‌های یادگیری عمیق بیشتر شده است (محمدی و همکاران[1]، ۲۰۲۰). رویکردهای آماری بر روش‌های ریاضی سنّتی متمرکز می‌باشند، اما روش‌های مربوط به یادگیری ماشین بر یادگیری و هوش متمرکز هستند. در حالی‌که هر دو دسته شباهت‌های زیادی دارند، تفاوت اصلی آنها این است که روش‌های آماری سخت‌تر و غیرمنعطف هستند، در حالی‌که روش‌های یادگیری ماشین قادر به یادگیری و انعطاف‌پذیر می‌باشند (هامفریس و همکاران[2]، ۲۰۱۱).

مطالعات قبلی کارایی برتر رویکردهای یادگیری عمیق را نسبت به رویکردهای آماری مرسوم نشان داده‌اند (وست و همکاران[3]، ۲۰۱۴). با توجه به ادبیات موجود، هیچ استراتژی واحدی برای شناسایی تقلب در صورت‌های مالی وجود ندارد (هامال و سنوار[4]، ۲۰۲۱). یافته‌های پیشین (کراجا و همکاران[5]، ۲۰۲۰) نشان می‌دهد که مدل‌هایی که با استفاده از رویکردهای یادگیری ماشین ساخته شده‌اند، می‌توانند به طور مؤثر تقلب در صورت‌های مالی را شناسایی کنند، با تکامل مستمر رفتار متقلبانه گزارشگری مالی همگام باشند و با به‌روزترین فناوری پاسخ دهند. در پژوهشی گوپتا و مهتا[6] (۲۰۲۱)، نشان دادند که مدل‌های تشخیص توسعه‌یافته با استفاده از رویکردهای یادگیری ماشین، دقیق‌تر از روش‌های سنّتی هستند. بسیاری از تحقیقات تشخیص صورت‌های مالی متقلبانه را به عنوان یک مسئله طبقه‌بندی باینری، گاهی به عنوان یک مسئله چندکلاسه و گاهی اوقات به عنوان یک مشکل خوشه‌بندی، فرموله کرده‌اند. محققان هر دو تجزیه و تحلیل کمّی و کیفی صورت‌های مالی متقلبانه را انجام داده‌اند. متن‌کاوی به طور گسترده برای تحقیقات کیفی استفاده شده است. در پژوهش حاضر روی مقالاتی تمرکز شده که تجزیه و تحلیل کمّی را با استفاده از الگوریتم‌های یادگیری ماشین انجام می‌دهند. در مراحل اولیه، تحقیقات عمدتاً شامل رویکردهای شبکه عصبی[7] (سچینی و همکاران[8]، ۲۰۱۰؛ پای و همکاران[9]، ۲۰۱۱؛ الفیض و فاتی[10]، ۲۰۲۲؛ استرلسنیا و پراکونویت[11]، ۲۰۲۳؛ کومار

1. Mohammadi
2. Humpherys
3. West
4. Hamaland Senvar
5. Craja
6. Gupta & Mehta
7. NN
8. Cecchini
9. Pai
10. Alfaiz & Fati
11. Strelcenia & Prakoonwit

http://sebaa.journal.qom-iau.ac.ir/

و همکاران[۱]، ۲۰۲۲)، رگرسیون لجستیک،[۲] درخت تصمیم،[۳] ماشین بردار پشتیبان،[۴] تحلیل متمایز[۵] و شبکه باور بیزی[۶] بود. مطالعاتی که در حوزه یادگیری ماشین در ایالات متحده آمریکا، چین، تایوان و اسپانیا انجام شده است، نشان می‌دهد که ۶۵ درصد از این مطالعات از تکنیک‌های یادگیری نظارت شده برای تجزیه و تحلیل داده‌ها استفاده کرده‌اند (الباشراوی[۷]، ۲۰۱۶). تعداد قابل توجهی از مطالعات، عملکرد طبقه‌بندی‌کننده‌ها را در تشخیص صورت‌های مالی متقلبانه تجزیه و تحلیل کرده و نشان می‌دهند که ماشین بردار پشتیبان (سچینی و همکاران، ۲۰۱۰؛ پای و همکاران، ۲۰۱۱؛ پرولس[۸]، ۲۰۱۱؛ آسیمیت و همکاران[۹]، ۲۰۲۲؛ موئیپیا و همکاران[۱۰]، ۲۰۱۴؛ یائو و همکاران[۱۱]، ۲۰۱۹)، شبکه عصبی (هان[۱۲]، ۲۰۱۷؛ لین و همکاران[۱۳]، ۲۰۱۵؛ راویسانکار و همکاران[۱۴]، ۲۰۱۱؛ ریزکی و همکاران[۱۵]، ۲۰۱۷؛ موروروونخر و همکاران[۱۶]، ۲۰۲۲؛ پیرز لوپز و همکاران[۱۷]، ۲۰۱۹) و درخت تصمیم (گوپتا و گیل[۱۸]، ۲۰۱۲؛ چن[۱۹]، ۲۰۱۶) در کشف/پیش‌بینی صورت‌های مالی متقلبانه، عملکرد خوبی دارند.

۲-۲. پیشینه تجربی پژوهش

تانگ و کریم[۲۰] (۲۰۱۸) در پژوهشی با عنوان «کشف تقلب مالی و تجزیه و تحلیل داده‌های بزرگ: پیامدهای مورد استفاده حسابرسان از نشست‌های ایده‌پردازی»، نشان دادند که روش‌های

1. Kumar
2. LR
3. DT
4. SVM
5. DA
6. Bayesian Belief Network
7. Albashrawi
8. Perols
9. Asimit
10. Moepya
11. Yao
12. Han
13. Lin
14. Ravisankar
15. Rizki
16. Murorunkwere
17. Pérez López
18. Gupta & Gill
19. Chen
20. Tang & Karim

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

http://sebaa.journal.qom-iau.ac.ir/

حسابرسی موجود با هدف کشف تقلب، نیاز به بهبود و اصلاح دارند؛ چراکه توجه به خرد جمعی باعث ایجاد نگرانی و گمراهی خواهد شد. بنابراین، باید از داده‌های بزرگ‌تر در هر مرحله از جمع‌آوری داده‌ها و شناسایی شاخص‌های تقلب و مستندات بیشتر و جلسات متعدد استفاده شود.

الغفیلی و همکاران (۲۰۲۰) در پژوهشی با عنوان «مدل کشف تقلب مالی براساس تکنیک یادگیری عمیق LSTM»، برای ارزیابی مدل پیشنهادی، از یک مجموعه داده واقعی از تقلب‌های کارت اعتباری استفاده کردند و نتایج را با یک مدل یادگیری عمیق موجود به نام مدل Auto-encoder و برخی تکنیک‌های یادگیری ماشین دیگر مقایسه کردند. نتایج تجربی، یک عملکرد عالی LSTM را نشان می‌دهد که در آن ۹۹/۹۵% دقت را در کمتر از یک دقیقه به دست می‌آورد.

شیوگو و شنگ یونگ[1] (۲۰۲۲) در پژوهشی با عنوان «تحلیلی در مورد کشف تقلب در صورت‌های مالی برای شرکت‌های پذیرفته شده چینی با استفاده از یادگیری عمیق»، ابتدا سیستم شاخص مالی را شامل شاخص‌های مالی و غیرمالی تدوین کردند که تحقیقات قبلی معمولاً آنها را حذف می‌کردند. سپس ویژگی‌های متنی در بخش بحث و تحلیل مدیریت گزارش‌های سالانه شرکت‌های پذیرفته شده چینی با استفاده از بردار کلمه استخراج شدند. پس از آن، از مدل‌های یادگیری عمیق قدرتمند استفاده شده و عملکرد آنها به ترتیب با داده‌های عددی، متنی و ترکیبی از آنها مقایسه گردید. نتایج تجربی، بهبود عملکرد عالی روش‌های یادگیری عمیق را در برابر روش‌های یادگیری ماشین سنّتی نشان می‌دهد که بیانگر این است که میزان دقت الگوریتم‌های LSTM و GRU در کشف تقلب مالی به ترتیب ۹۴/۹۸% و ۹۴/۶۲% می‌باشد.

علی و همکاران[2] (۲۰۲۳) در پژوهشی از تکنیک‌های مختلف یادگیری ماشین در محیط پایتون برای پیش‌بینی صورت‌های مالی متقلبانه استفاده کردند و یافته‌های تجربی آنها نشان می‌دهد که الگوریتم XGBoost از دیگر الگوریتم‌های این مطالعه، یعنی رگرسیون لجستیک،[3] درخت تصمیم،[4] ماشین بردار پشتیبان،[5] جنگل تصادفی[6] و آدابوست بهتر عمل کرده است. همچنین الگوریتم XGBoost را برای بدست آوردن بهترین نتیجه با دقت نهایی ۹۶/۰۵ درصد در تشخیص صورت‌های مالی متقلبانه، بهینه کردند.

1. Xiuguo & Shengyong
2. Ali
3. LR
4. DT
5. SVM
6. RF

http://sebaa.journal.qom-iau.ac.ir/

زهیر و همکاران[1] (۲۰۲۳) در پژوهشی با استفاده از داده‌های شاخص ترکیبی شانگهای و روش‌های CNN، RNN، LSTM، CNN-RNN و CNN-LSTM نشان دادند که CNN بدترین عملکرد را دارد، LSTM بهتر از CNN-LSTM، CNN-RNN بهتر از CNN-LSTM و CNN-RNN بهتر از LSTM عمل می‌کند، و مدل RNN تک‌لایه پیشنهادی، بر همه مدل‌های اشاره شده برتری دارد. نتایج تجربی، اثربخشی مدل پیشنهادی را تایید می‌کند که به سرمایه‌گذاران در افزایش سود خود با تصمیم‌گیری خوب کمک می‌نماید.

یولیانتی و همکاران (۲۰۲۴) در پژوهشی با عنوان «کشف تقلب صورت‌های مالی: رویکرد مدل تقلب شش ضلعی» نشان دادند که ثبات مالی بر تقلب صورت‌های مالی تأثیر مثبت دارد. با این حال، نظارت ناکارآمد، تغییر حسابرس، تغییر مدیران، دوره تصدی مدیرعامل و ارتباط با پروژه‌های دولتی، تأثیری بر تقلب در صورت‌های مالی ندارد.

احمدی و همکاران (۱۳۹۹) در پژوهشی با عنوان «ارائه مدلی برای پیش‌بینی صورت‌های مالی متقلبانه و مقایسه صورت‌ها و نسبت‌های مالی با قانون بنفورد»، با استفاده از روش رگرسیون لجستیک نشان دادند که با توجه به نرخ دقت ۶۴/۶ درصدی، این مدل نقش اثربخشی در کشف تقلب صورت‌های مالی دارد. همچنین نتایج آزمون T و آزمون لون در ۳۵ متغیر مستقل بررسی شده نشان داد که در ۲۰ متغیر، تفاوت معناداری در دو گروه متقلب و غیرمتقلب وجود دارد. به‌علاوه، تطابق‌پذیری و انحراف از قانون بنفورد در چهار حالت مختلف بررسی شد. توجه به ارقام صورت سود و زیان و ترازنامه در شرکت‌های غیرمتقلب نشان داد توزیع بنفورد شرکت‌های غیرمتقلب را به درستی تشخیص داده، ولی شرکت‌های متقلب را به صورت نادرست غیرمتقلب تشخیص داده است. توجه به نسبت مالی کل دارایی‌ها به فروش، و دوره پرداخت حساب‌های پرداختنی در شرکت‌های متقلب نشان داد که توزیع بنفورد، این شرکت‌ها را متقلب ارزیابی و به درستی دسته‌بندی کرده، ولی شرکت‌های غیرمتقلب را به صورت نادرست متقلب ارزیابی کرده است.

ملکی کاکلر و همکاران (۱۴۰۰) در پژوهشی با عنوان «کارایی مدل‌های آماری و الگوهای یادگیری ماشین در پیش‌بینی گزارشگری مالی متقلبانه» با استفاده از ۲۰ متغیر در قالب الگوی پنج ضلعی تقلب با تاکید بر ساختار کنترل‌های داخلی در ۱۶۶ شرکت فعال در بورس اوراق بهادار تهران طی سال‌های ۱۳۸۸ الی ۱۳۹۷ و مقایسه بین مدل‌های مورد بررسی، با کمک آزمون مقایسه نسبت‌ها، نشان دادند که به لحاظ آماری، مدل‌های یادگیری ماشین در پیش‌بینی گزارشگری مالی متقلبانه،

1. Zaheer

http://sebaa.journal.qom-iau.ac.ir/

نسبت به مدل‌های آماری، کارایی و دقت بیشتری دارند. ترکیب الگوریتم درخت تصمیم‌گیری CHAID، C5 و C&R بالاترین دقت در پیش‌بینی گزارشگری مالی متقلبانه را با دقت بالای ۹۲/۶۱ درصد در پیش‌بینی تقلب نشان می‌دهد.

رضائی و همکاران (۱۴۰۰) در پژوهشی با عنوان «پیش‌بینی تقلب صورت‌های مالی با استفاده از رویکرد کریسپ»[1] نشان دادند که همه تکنیک‌ها قابلیت کشف تقلب صورت‌های مالی را در سطح نسبتاً بالایی دارند و تکنیک پیشنهادی آدابوست ماشین بردار پشتیبان در مرحله آموزش با نرخ دقت ۸۱٫۶۹% دارای دقت و توان ارزیابی بالاتری نسبت به سایر تکنیک‌ها بوده و این تکنیک در مرحله آزمایش، ۸۲% صورت‌های مالی متقلبانه و غیرمتقلبانه سال ۱۳۹۶ را بدرستی تشخیص داد.

کاظمی و پیری (۱۴۰۱) در پژوهشی با عنوان «پیش‌بینی طرح تقلب در گزارشگری مالی با استفاده از رویکرد یادگیری ماشین در فضای چند کلاسه»، نشان دادند که تفاوت معناداری در عملکرد الگوهای یادگیری ماشین در فضای چند کلاسه وجود دارد، و روش ماشین بردار پشتیبان نسبت به سایر روش‌ها عملکرد بهتری دارد. با تقلیل فضای مسیله به دسته‌بندی دو کلاسه، تفاوت معناداری در عملکرد الگوهای یادگیری ماشین در تشخیص گزارش‌های مالی مشکوک به بیش‌نمایی دارایی، کم‌نمایی بدهی و هزینه، بیش‌نمایی دارایی و کم‌نمایی هزینه» و کم‌نمایی هزینه و بدهی، تایید نشد. با این حال، عملکرد ماشین بردار پشتیبان بر عملکرد روش رگرسیون لجستیک و درخت تصمیم در پیش‌بینی گزارش‌های مالی مشکوک به بیش‌نمایی دارایی و درآمد ارجح است.

احمدی و همکاران (۱۴۰۳) در پژوهشی با عنوان «تکنیک‌های داده‌کاوی و پیش‌بینی تقلب صورت‌های مالی» نشان دادند که روش‌های داده‌کاوی در تمایز صورت‌های مالی متقلبانه از غیر متقلبانه موثر هستند. بدین ترتیب که روش شبکه عصبی ۶۹/۴ درصد، درخت تصمیم ۶۵٫۴ درصد، نزدیکترین همسایگی ۶۴/۴ درصد و ماشین بردار پشتیبان ۷۸ درصد پیش‌بینی صحیح داشته اند.

امیرمعزی و همکاران (۱۴۰۳) در پژوهشی با عنوان «کشف تقلب صورت‌های مالی: قیاس توانایی مدل‌های مبتنی بر متغیرهای حسابداری» نشان دادند که دقت معیار F-SCORE دچو، در تشخیص شرکت‌های دارای احتمال دستکاری و غیر آن، بیش از ۷۰ است که بیانگر توانایی متوسط مدل مذکور در تشخیص تقلب است. همچنین توانایی معیار فوق الذکر با ۷۳٫۱۷ درصد در تشخیص تقلب از مدل بنیش با ۶۹٫۵۱ درصد بالاتر است. از حیث خطای تشخیص نیز، خطای نوع I و II که به ترتیب بیانگر خطای کارایی و اثربخشی مدل است، در مدل دچو و همکاران (۲۰۱۱)

1. CRISP

http://sebaa.journal.qom-iau.ac.ir/

به مراتب کمتر از مدل بنیش است. بنابراین، می‌توان نتیجه گرفت که معیار F-SCORE دچو در موارد تشخیص احتمال دستکاری در صورت‌های مالی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در فاصله سال‌های ۱۳۹۴ تا ۱۳۹۸ بهتر عمل کرده است.

زارعی و همکاران (۱۴۰۳) در پژوهشی با عنوان «ارائه الگوی پیش‌بینی تقلب مبتنی بر هوش مصنوعی (بیز ساده)» نشان دادند که روش مورد بررسی در این پژوهش با ضریب صحّت ۸۴ درصد، ضریب دقت ۸۴ درصد و ضریب بازخوانی ۹۸ درصد، از توانایی بالایی برای پیش‌بینی تقلب براساس متغیرهای موجود در صورت مالی شرکت‌های پذیرفته شده در بازار بورس برخوردار است.

۳. روش پژوهش

۳-۱. جمع‌آوری داده‌ها

این پژوهش از لحاظ هدف کاربردی بوده و از نوع پژوهش‌های شبه‌تجربی و پس‌رویدادی است و با استفاده از اطلاعات تاریخی انجام شده است. داده‌های مورد نیاز از نرم‌افزار «ره‌آورد نوین» و سایت‌های اینترنتی «مدیریت پژوهش، توسعه و مطالعات اسلامی سازمان بورس اوراق بهادار»، «کدال»، بانک مرکزی و مرکز آمار ایران گردآوری شده است.

۳-۲. جامعه آماری، نمونه آماری و بازه زمانی پژوهش

جامعه آماری پژوهش حاضر، شامل کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران است. بازه زمانی نیز یک دوره زمانی ۱۲ ساله براساس صورت‌های مالی سال‌های ۱۳۹۱-۱۴۰۲ می‌باشد.

در پژوهش حاضر برای تعیین نمونه آماری، از روش غربالگری استفاده شده است. بدین منظور آن دسته از شرکت‌های جامعه آماری که شرایط زیر را دارا بودند، به عنوان نمونه آماری انتخاب و مابقی حذف شدند:

۱) سال مالی شرکت منتهی به تاریخ پایان اسفند ماه هر سال باشد.

۲) شرکت طی دوره مورد بررسی، تغییر سال مالی نداده باشد.

۳) شرکت‌های مورد بررسی جزء شرکت‌های سرمایه‌گذاری، هلدینگ، واسطه‌گری مالی و بیمه نباشند.

۴) اطلاعات و داده‌های آنها در دسترس باشد.

با توجه به شرایط و محدودیت‌های فوق، از بین شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران در مجموع ۱۵۰ شرکت به عنوان نمونه آماری پژوهش انتخاب شدند.

http://sebaa.journal.qom-iau.ac.ir/

۳-۳. روش اندازه‌گیری متغیرهای پژوهش

متغیر وابسته (پاسخ): متغیر وابسته در پژوهش حاضر صورت‌های مالی متقلبانه می‌باشد که با استفاده از شاخص F_SCORE تدوین شده توسط دیچو و همکاران[1] (۲۰۱۱) اندازه‌گیری می‌شود که به صورت زیر است:

رابطه (۱):

$$\text{Predicted Value} = -7.893+0.790\times(\text{rsst\_acc})+2.518\times(\text{ch\_rec})+1.191\times(\text{ch\_inv})+1.979\times(\text{soft\_assets})+0.171\times(\text{ch\_cs})-0.932\times(\text{ch\_roa})+1.029\times(\text{issue})$$

که در آن:

rsst_acc: برابر است با (Δfin + Δwc + Δnco) تقسیم بر میانگین کل دارایی‌ها.

wc: برابر دارایی‌های جاری منهای وجه نقد و سرمایه‌گذارهای کوتاه مدت، منهای بدهی‌های جاری است.

Nco: برابر مجموع دارایی‌ها، منهای دارایی جاری، منهای سرمایه‌گذاری‌ها، منهای مجموع بدهی‌ها، به کسر بدهی‌های جاری و بلندمدت است.

Fin: برابر سرمایه‌گذاری کوتاه مدت، به‌علاوه سرمایه‌گذاری بلندمدت، منهای مجموع بدهی‌های بلندمدت، بدهی‌های جاری و سهام ممتاز است.

ch_rec: برابر با تغییر در حساب‌های دریافتنی در طول دوره‌ی جاری تقسیم بر میانگین دارایی‌ها است.

ch_inv: برابر با تغییر در موجودی‌ها در طول دوره‌ی جاری تقسیم بر میانگین دارایی‌ها است.

soft_assets: برابر با مجموع دارایی‌ها به کسر اموال، ماشین‌آلات و تجهیزات و وجه نقد و معادل آن تقسیم بر مجموع دارایی‌ها است.

ch_cs: برابر با درصد تغییر در فروش نقدی است، که از طریق فروش دوره جاری، منهای تغییر در حساب‌های دریافتنی محاسبه می‌شود.

ch_roa: برابر با تغییر در بازده دارایی‌ها است، که از طریق تقسیم سود خالص بر میانگین دارایی‌ها بدست می‌آید.

Issue: متغیری ساختگی است؛ به این ترتیب که برای شرکتی که اوراق بدهی یا اوراق مالکیت منتشر کرده باشد، عدد (۱) و در غیر این صورت (۰) می‌باشد.

برای محاسبه F_SCORE، احتمال پیش‌بینی از طریق تقسیم $e^{PV}/(1+e^{PV})$ بر احتمال غیرشرطی

1. Dechow

http://sebaa.journal.qom-iau.ac.ir/

ارائه نادرست صورت‌های مالی (۰٫۰۰۳۷) بدست می‌آید. که در آن PV ارزش پیش‌بینی بدست آمده از رابطه (۱) است. براساس پژوهش دیچو و همکاران (۲۰۱۱)، در این پژوهش مشاهدات دارای F_SCORE بالاتر از ۱٫۸۵ به عنوان شرکت‌های با ریسك بالای صورت‌های مالی متقلبانه شناسایی می‌شوند.

۳-۴. متغیرهای مستقل (پیش‌بین)

به پیروی از پژوهش‌های پیشین (علی و همکاران، ۲۰۲۳؛ جان[1]، ۲۰۲۱؛ کات سیس و همکاران[2]، ۲۰۱۲؛ پرلوس[3]، ۲۰۱۱؛ رضائی و همکاران، ۱۴۰۰؛ خواجوی و ابراهیمی، ۱۳۹۶؛ اعتمادی و زلقی، ۱۳۹۲؛ صفرزاده، ۱۳۸۹) متغیرهای مستقل پژوهش حاضر به شرح جدول زیر ارائه می‌گردد:

جدول ۱- نحوه اندازه‌گیری متغیرهای مستقل پژوهش

ردیف	متغیر مستقل	نماد	نحوه اندازه‌گیری
۱	حاشیه سود ناخالص	GPM	سود ناخالص تقسیم بر فروش
۲	حاشیه سود خالص	NPM	سود خالص تقسیم بر فروش
۳	حاشیه سود عملیاتی	OPM	سود عملیاتی تقسیم بر فروش
۴	بازده دارایی‌ها	ROA	سود خالص تقسیم بر کل دارایی
۵	بازده حقوق صاحبان سهام	ROE	سود خالص تقسیم بر ارزش دفتری حقوق صاحبان سهام
۶	نسبت سود قبل از بهره و مالیات	EBIT/A	سود قبل از بهره و مالیاتی تقسیم بر کل دارایی
۷	گردش کل دارایی‌ها	ASTRN	فروش خالص تقسیم بر کل دارایی
۸	گردش دارایی‌های ثابت	FXASTRN	فروش خالص تقسیم بر دارایی‌های ثابت
۹	گردش موجودی کالا	INVTRN	بهای تمام شده کالای فروش رفته تقسیم بر متوسط موجودی کالا
۱۰	گردش حساب‌های دریافتنی	ACRECTRN	فروش خالص تقسیم بر متوسط حساب‌های دریافتنی
۱۱	نسبت کل بدهی‌ها به کل دارایی‌ها	LIB/ASST	کل بدهی‌ها تقسیم بر کل دارایی‌ها
۱۲	نسبت کل بدهی‌ها به حقوق صاحبان سهام	LIB/EQT	کل بدهی‌ها تقسیم بر ارزش دفتری حقوق صاحبان سهام
۱۳	نسبت سود انباشته به کل دارایی‌ها	ACUM/ASST	سود (زیان) انباشته تقسیم بر کل دارایی
۱۴	نسبت سود انباشته به سرمایه	ACUM/EQT	سود (زیان) انباشته تقسیم بر سرمایه سهام عادی

1. Jan
2. Katsis
3. Perlos

http://sebaa.journal.qom-iau.ac.ir/

ردیف	متغیر مستقل	نماد	نحوه اندازه‌گیری
۱۵	نسبت پوشش هزینه بهره	INTCOV	هزینه بهره تقسیم بر سود قبل از بهره و مالیات
۱۶	نسبت جاری	CURRAT	دارایی‌های جاری تقسیم بر بدهی‌های جاری
۱۷	نسبت آنی	QUICRAT	(وجه نقد + حساب‌های دریافتنی + سرمایه‌گذاری کوتاه مدت) تقسیم بر بدهی‌های جاری
۱۸	نسبت وجه نقد آزاد به درآمد	FCF/REV	{سود خالص+هزینه استهلاک+هزینه مالی×(۱- مالیات) - سرمایه‌گذاری در دارایی ثابت - سرمایه‌گذاری در سرمایه در گردش} تقسیم بر فروش
۱۹	نسبت مخارج سرمایه‌ای به درآمد	CE/REV	مخارج سرمایه‌ای (تفاوت خالص ارزش دفتری دارایی‌های ثابت در ابتدا و پایان دوره به‌علاوه هزینه استهلاک) تقسیم بر فروش
۲۰	نسبت مخارج سرمایه‌ای به سود خالص	CE/NP	مخارج سرمایه‌ای، تقسیم بر سود خالص
۲۱	نسبت وجه نقد عملیاتی به درآمد	OCF/REV	خالص جریان وجوه نقد عملیاتی حاصل از فعالیت‌های عملیاتی تقسیم بر فروش خالص
۲۲	نسبت وجه نقد عملیاتی به بدهی	OCF/LIB	خالص جریان وجوه نقد عملیاتی حاصل از فعالیت‌های عملیاتی، تقسیم بر کل بدهی‌ها
۲۳	ارزش شرکت	EV	لگاریتم طبیعی (ارزش بازار سهام+ارزش دفتری بدهی‌ها- نقد و معادل نقد)
۲۴	ارزش دفتری شرکت	BV	لگاریتم طبیعی ارزش دفتری حقوق صاحبان سهام

۳-۵. الگوریتم‌های یادگیری عمیق مورد استفاده در پژوهش

ماشین بردار پشتیبان (SVM[1]): ماشین بردار پشتیبان یکی از روش‌های یادگیری با نظارت است از آن برای طبقه‌بندی و رگرسیون استفاده می‌شود. این روش از روش‌های نسبتاً جدیدی بوده که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی نشان داده است. مبنای کاری دسته‌بندی‌کننده SVM دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب گردد که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده‌ها به وسیله روش‌های QP که روش‌های شناخته‌شده‌ای در حل مسائل محدودیت‌دار هستند، صورت می‌گیرد.

قبل از تقسیم خطی، برای اینکه ماشین بتواند داده‌هایی با پیچیدگی بالا را دسته‌بندی کند، داده‌ها با استفاده از تابع phi به فضایی با ابعاد خیلی بالاتر برده می‌شوند. برای اینکه بتوان مسئله ابعاد خیلی

1. Support vector machines

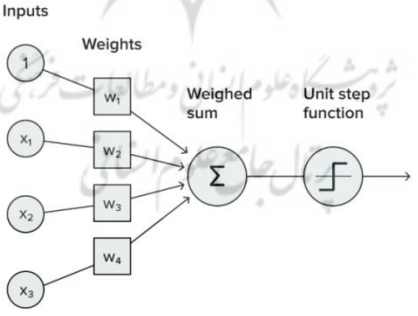
http://sebaa.journal.qom-iau.ac.ir/

بالا را با استفاده از این روش‌ها حل کرد، از قضیه دوگانی لاگرانژ[1] برای تبدیل مسئله مینیمم‌سازی مورد نظر به فرم دوگانی آن که در آن به جای تابع پیچیده phi که ما را به فضایی با ابعاد بالا می‌برد، تابع ساده‌تری به نام تابع هسته که ضرب برداری تابع phi است، استفاده می‌گردد. از توابع هسته مختلفی از جمله هسته‌های نمایی، چندجمله‌ای و سیگموید[2] می‌توان استفاده نمود.

شبکه عصبی پیچشی(CNN[3]): شبکه عصبی پیچشی یا به اختصار CNN که به آن شبکه عصبی کانولوشنی نیز گفته می‌شود، نوعی از شبکه‌های عصبی است که عموماً برای یادگیری بر روی مجموعه داده‌های بصری (مانند تصاویر و عکس‌ها) استفاده می‌شود. از لحاظ مفهوم، این شبکه‌ها مانند شبکه‌های عصبی ساده هستند؛ یعنی از فازهای پیش‌خور[4] و پس‌انتشار خطا[5] استفاده می‌کنند، ولی از لحاظ معماری تفاوت‌هایی با شبکه‌های عصبی ساده دارند. این شبکه‌ها در دسته یادگیری عمیق قرار می‌گیرند؛ زیرا لایه‌های موجود در این شبکه‌ها، زیاد است.

شبکه عصبی پیچشی همانند سایر شبکه‌های عصبی از لایه‌های نورونی با وزن و بایاس با قابلیت یادگیری تشکیل شده است. در هر نورون اتفاقات زیر رخ می‌دهد:

۱. نورون مجموعه‌ای ورودی دریافت می‌کند.
۲. ضرب داخلی بین وزن‌های نورون و ورودی‌ها انجام می‌شود.
۳. حاصل با بایاس جمع می‌شود.
۴. درنهایت، از یک تابع غیرخطی[6] عبور داده می‌شود.



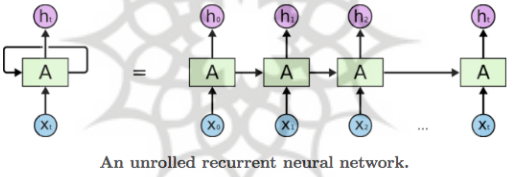
شکل ۱- یک بلوک شبکه عصبی پیچشی (علی و همکاران، ۲۰۲۳)

1. Lagrange's duality theorem
2. Sigmoid
3. Convolutional Neural Networks
4. Feed forward
5. Back propagation of error
6. Activation function

http://sebaa.journal.qom-iau.ac.ir/

فرآیند بالا لایه به لایه انجام می‌شود و درنهایت به لایه خروجی می‌رسیم. لایه خروجی، پیش‌بینی شبکه را تولید می‌کند.

شبکه عصبی بازگشتی (RNN[1]): شبکه عصبی بازگشتی که به آن شبکه عصبی مکرر نیز گفته می‌شود، نوعی از شبکه عصبی مصنوعی است که در تشخیص گفتار، پردازش زبان طبیعی[2] و همچنین در پردازش داده‌های ترتیبی[3] استفاده می‌شود. بسیاری از شبکه‌های عمیق مانند CNN شبکه‌های پیش‌خور[4] هستند؛ یعنی سیگنال در این شبکه‌ها فقط در یک جهت از لایه ورودی، به لایه‌های مخفی و سپس به لایه خروجی حرکت می‌کند و داده‌های قبلی به حافظه سپرده نمی‌شوند. اما شبکه‌های عصبی بازگشتی، یک لایه بازخورد دارند که در آن خروجی شبکه به همراه ورودی بعدی، به شبکه بازگردانده می‌شود. RNN می‌تواند به علت داشتن حافظه داخلی، ورودی قبلی خود را به خاطر بسپارد و از این حافظه برای پردازش دنباله‌ای از ورودی‌ها استفاده کند. به بیان ساده، شبکه‌های عصبی بازگشتی شامل یک حلقه بازگشتی هستند که موجب می‌شود اطلاعاتی را که از لحظات قبلی بدست آمده، از بین نروند و در شبکه باقی بمانند.



شکل ۲- یک بلوک شبکه عصبی بازگشتی (علی و همکاران، ۲۰۲۳)

همانطور که در شکل (۲) مشاهده می‌شود، X_0 را از دنباله ورودی می‌گیرد، خروجی h_0 به همراه X_1، ورودی مرحله بعدی است. در مرحله بعد، خروجی h_1 و X_2 ورودی مرحله بعد است. به این ترتیب، شبکه در هنگام آموزش، قادر به یادآوری ورودی‌های قبلی خواهد بود.

شبکه عصبی بازگشتی می‌تواند کاربردهایی به شرح زیر داشته باشد:

۱. شرح‌نویسی عکس:[5] شبکه عصبی بازگشتی با تحلیل حالت کنونی عکس، برای شرح‌نویسی عکس به‌کار می‌رود.

1. Recurrent Neural Networks
2. NLP
3. Sequential data
4. Feed Forward
5. Image Captioning

http://sebaa.journal.qom-iau.ac.ir/

۲. پیش‌بینی سری‌های زمانی:[1] هر مسئله سری زمانی مانند پیش‌بینی قیمت یک سهام در یک ماه خاص، با RNN قابل انجام است.

۳. پردازش زبان طبیعی:[2] کاوش متن و تحلیل احساسات می‌تواند با استفاده از RNN انجام شود.

۴. ترجمه ماشینی:[3] شبکه RNN می‌تواند ورودی خود را از یک زبان دریافت و آن را به عنوان خروجی به زبان دیگری ترجمه کند.

٤. یافته‌های پژوهش

۴-۱. آمار توصیفی متغیرهای پژوهش

جدول ۲- آمار توصیفی متغیرهای پژوهش

پنل الف: متغیرهای پیوسته							
متغیرها	نماد	تعداد مشاهدات	میانگین	میانه	حداکثر	حداقل	انحراف معیار
حاشیه سود ناخالص	GPM	1800	0.179	0.136	0.616	-0.151	0.196
حاشیه سود خالص	NPM	1800	0.164	0.128	0.566	-0.154	0.183
حاشیه سود عملیاتی	OPM	1800	0.191	0.157	0.528	-0.107	0.172
بازده دارایی‌ها	ROA	1800	0.131	0.108	0.414	-0.089	0.134
بازده حقوق صاحبان سهام	ROE	1800	0.294	0.291	0.749	-0.236	0.260
نسبت سود قبل از بهره و مالیات	EBIT/A	1800	0.187	0.161	0.466	-0.029	0.139
گردش کل دارایی‌ها	ASTRN	1800	0.919	0.795	2.067	0.297	0.478
گردش دارایی‌های ثابت	FXASTRN	1800	6.396	4.479	21.330	0.703	5.679
گردش موجودی کالا	INVTRN	1800	4.008	2.851	12.172	1.007	2.988
گردش حساب‌های دریافتنی	ACRECTRN	1800	5.416	3.545	20.900	0.925	5.168
نسبت کل بدهی‌ها به کل دارایی‌ها	LIB/ASST	1800	0.572	0.579	0.923	0.208	0.200
نسبت کل بدهی‌ها به حقوق صاحبان سهام	LIB/EQT	1800	1.849	1.268	6.786	0.172	1.701
نسبت سود انباشته به کل دارایی‌ها	ACUM/ASST	1800	0.148	0.152	0.495	-0.249	0.193
نسبت سود انباشته به سرمایه	ACUM/EQT	1800	0.332	0.401	0.880	-0.723	0.413
نسبت پوشش هزینه بهره	INTCOV	1800	19.244	4.459	153.122	-0.401	37.205
نسبت جاری	CURRAT	1800	1.526	1.348	3.440	0.599	0.717

1. Time Series Prediction
2. Natural Language Processing
3. Machine Translation

http://sebaa.journal.qom-iau.ac.ir/

پنل الف: متغیرهای پیوسته							
متغیرها	نماد	تعداد مشاهدات	میانگین	میانه	حداکثر	حداقل	انحراف معیار
نسبت آنی	QUICRAT	1800	0.816	0.728	2.067	0.193	0.484
نسبت وجه نقد آزاد به درآمد	FCF/REV	1800	0.054	0.043	0.323	-0.261	0.139
نسبت مخارج سرمایه‌ای به درآمد	CE/REV	1800	0.045	0.012	0.224	-0.019	0.075
نسبت مخارج سرمایه‌ای به سود خالص	CE/NP	1800	0.248	0.070	1.652	-0.519	0.600
نسبت وجه نقد عملیاتی به درآمد	OCF/REV	1800	0.136	0.112	0.473	-0.117	0.153
نسبت وجه نقد عملیاتی به بدهی	OCF/LIB	1800	0.247	0.162	1.145	-0.144	0.316
ارزش شرکت	EV	1800	15.635	15.598	18.757	12.944	1.669
ارزش دفتری شرکت	BV	1800	14.695	14.505	17.970	12.255	1.516

پنل ب: متغیرهای دو ارزشی				
متغیرها	نماد	طبقه	فراوانی	درصد فراوانی
صورت‌های مالی متقلبانه	FSF	1	536	29.8%
		0	1264	70.2%

در جدول (۲)، برخی از مفاهیم آمار توصیفی متغیرها، شامل میانگین، میانه، حداقل مشاهدات، حداکثر مشاهدات و انحراف معیار ارائه شده است. با توجه به پنل «الف»، میانگین متغیر حاشیه سود ناخالص ۰/۱۷۹ می‌باشد که با توجه به انحراف معیار (۰/۱۹۶)، از نوسان‌پذیری بالایی برخوردار است. میانگین متغیر بازده دارایی‌ها ۰/۱۳۱ می‌باشد که با توجه به انحراف معیار (۰/۱۳۴)، از نوسان‌پذیری بالایی برخوردار است. میانگین نسبت کل بدهی‌ها به کل دارایی‌ها ۰/۵۷۲ می‌باشد که نشان‌دهنده این است که به طور متوسط ۵۷/۲% منابع مالی شرکت‌ها از طریق بدهی، تأمین مالی شده است. با توجه به پنل «ب» نیز نتایج نشان می‌دهد که درصد فراوانی صورت‌های مالی متقلبانه ۲۹/۸% می‌باشد که نشان‌دهنده این است که صورت‌های مالی ۲۹/۸% شرکت‌های موجود در نمونه آماری پژوهش، متقلبانه است.

جدول ۳- آزمون t در شرکت‌های دارای صورت‌های مالی متقلبانه و بدون صورت‌های مالی متقلبانه

متغیرها	نماد	صورت‌های مالی متقلبانه (تعداد مشاهدات = ٥٣٦)	بدون صورت‌های مالی متقلبانه (تعداد مشاهدات = ١٢٦٤)	آماره t
حاشیه سود ناخالص	GPM	0.0592	0.2292	-18.315***
حاشیه سود خالص	NPM	0.0534	0.2114	-17.911***
حاشیه سود عملیاتی	OPM	0.0739	0.2403	-22.228***

http://sebaa.journal.qom-iau.ac.ir/

متغیرها	نماد	صورت‌های مالی متقلبانه (تعداد مشاهدات = ٥٣٦)	بدون صورت‌های مالی متقلبانه (تعداد مشاهدات = ١٢٦٤)	آماره t
بازده دارایی‌ها	ROA	0.0199	0.1779	-31.566****
بازده حقوق صاحبان سهام	ROE	0.1082	0.3729	-22.301****
نسبت سود قبل از بهره و مالیات	EBIT/A	0.0701	0.2366	-32.817***
گردش کل دارایی‌ها	ASTRN	0.5773	1.0637	-27.466***
گردش دارایی‌های ثابت	FXASTRN	3.5672	7.5956	-16.780****
گردش موجودی کالا	INVTRN	3.622	4.1721	-3.644***
گردش حساب‌های دریافتنی	ACRECTRN	3.5139	6.2228	-11.929***
نسبت کل بدهی‌ها به کل دارایی‌ها	LIB/ASST	0.6622	0.535	12.732***
نسبت کل بدهی‌ها به حقوق صاحبان سهام	LIB/EQT	2.312	1.653	6.873***
نسبت سود انباشته به کل دارایی‌ها	ACUM/ASST	-0.0111	0.2168	-27.260***
نسبت سود انباشته به سرمایه	ACUM/EQT	0.0941	0.433	-15.039***
نسبت پوشش هزینه بهره	INTCOV	5.4553	25.0905	-14.125***
نسبت جاری	CURRAT	1.1201	1.6985	-19.202***
نسبت آنی	QUICRAT	0.6192	0.8993	-13.255***
نسبت وجه نقد آزاد به درآمد	FCF/REV	-0.056	0.1007	-23.443***
نسبت مخارج سرمایه‌ای به درآمد	CE/REV	0.06	0.0381	4.967***
نسبت مخارج سرمایه‌ای به سود خالص	CE/NP	0.3256	0.2151	3.028****
نسبت وجه نقد عملیاتی به درآمد	OCF/REV	0.1137	0.1461	-3.929***
نسبت وجه نقد عملیاتی به بدهی	OCF/LIB	0.0893	0.3146	-19.555***
ارزش شرکت	EV	15.3186	15.7694	-5.280***
ارزش دفتری شرکت	BV	14.5811	14.744	-1.996**

* معناداری در سطح اطمینان ۹۰٪ ، ** معناداری در سطح اطمینان ۹۵٪ و *** معناداری در سطح اطمینان ۹۹٪

با توجه به جدول (۳)، نتایج حاصل از آزمون t نشان می‌دهد که در سطح اطمینان ۹۵%، همه متغیرهای مستقل (پیش‌بین) در شرکت‌های دارای صورت‌های مالی متقلبانه و بدون صورت‌های مالی متقلبانه، تفاوت معناداری با هم دارند و میانگین متغیرهای حاشیه سود ناخالص، حاشیه سود خالص، حاشیه سود عملیاتی، بازده دارایی‌ها، بازده حقوق صاحبان سهام، نسبت سود قبل از بهره و مالیات، گردش کل دارایی‌ها، گردش دارایی‌های ثابت، گردش موجودی کالا، گردش حساب‌های دریافتنی، نسبت سود انباشته به کل دارایی‌ها، نسبت سود انباشته به سرمایه، نسبت پوشش هزینه بهره، نسبت جاری، نسبت آنی، نسبت وجه نقد آزاد به درآمد، نسبت وجه نقد عملیاتی به درآمد، نسبت وجه نقد عملیاتی به بدهی، ارزش شرکت و ارزش دفتری شرکت، در شرکت‌های دارای صورت‌های

http://sebaa.journal.qom-iau.ac.ir/

مالی متقلبانه در مقایسه با شرکت‌های دارای بدون صورت‌های مالی متقلبانه کمتر است؛ در حالی که، میانگین متغیرهای نسبت کل بدهی‌ها به کل دارایی‌ها، نسبت کل بدهی‌ها به حقوق صاحبان سهام، نسبت مخارج سرمایه‌ای به درآمد و نسبت مخارج سرمایه‌ای به سود خالص، در شرکت‌های دارای صورت‌های مالی متقلبانه، در مقایسه با شرکت‌های دارای بدون صورت‌های مالی متقلبانه، بیشتر است.

۴-۲. فرآیند ایجاد مدل‌های پژوهش

به منظور پاسخگویی به سوالات و همچنین دستیابی به اهداف پژوهش، ابتدا مدل‌های مورد نظر شامل مدل‌های ایجادی با ماشین بردار پشتیبان، شبکه عصبی پیچشی و شبکه عصبی بازگشتی، با استفاده از متغیرهای انتخاب شده براساس «آزمون مقایسه میانگین دو نمونه» ایجاد شده و مقایسه نتایج انجام شد.

۴-۲-۱. متغیرهای نهایی پژوهش

به منظور تعیین متغیرهای نهایی (با اهمیت) پژوهش از میان متغیرهای اولیه پژوهش (۲۴ متغیر انتخاب شده براساس مبانی نظری و مطالعات تجربی)، از آزمون مقایسه میانگین دو نمونه مستقل استفاده شد. به منظور بررسی مقایسه میانگین دو نمونه، ابتدا باید واریانس دو نمونه را مقایسه نمود؛ به عبارت دیگر، آزمون تساوی واریانس‌ها مقدم بر آزمون تساوی میانگین‌ها است. جهت تساوی واریانس‌ها از آزمون لوین که مبتنی بر آماره فیشر است، استفاده شد. در این آزمون نیازی نیست که توزیع داده‌ها نرمال باشد.

آماره t جهت آزمون تساوی میانگین دو نمونه، در حالت تساوی و عدم تساوی واریانس دو نمونه مورد نظر، محاسبه می‌شود. در حالت تساوی واریانس‌ها از رابطه (۲) برای محاسبه آماره t استفاده می‌شود که در این حالت درجه آزادی برابر $df=n_1+n_2-2$ است.

رابطه (۲):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (u_1 - u_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_1}}}$$

رابطه (۳):

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

در حالت عدم تساوی واریانس‌ها، آماره t از رابطه (۴) و درجه آزادی از رابطه (۵) محاسبه می‌شود.

http://sebaa.journal.qom-iau.ac.ir/

رابطه (۴):

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (u_1 - u_2)}{\sqrt{\frac{S_2^2}{n_1} + \frac{S_2^2}{n_1}}}$$

رابطه (۵):

$$t = \frac{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{S_2^2}{n_2})^2}{n_2 - 1}}$$

که در آنها:

n_1 و n_2 تعداد اعضای نمونه اول و نمونه دوم؛ و S_1 و S_2 انحراف معیار نمونه اول و دوم هستند.

سپس براساس عدد t محاسبه شده طبق روابط فوق و عدد t طبق جدول، با توجه به سطح خطای ۵ درصد، تفسیر نتایج انجام می‌شود. اگر عدد t محاسبه شده از عدد t طبق جدول کوچک‌تر باشد (p-value<0.05)، معنادار بودن تفاوت میانگین‌های دو گروه، پذیرفته شده و در غیر این صورت رد خواهد شد. با توجه به جدول (۴)، نتایج حاصل از آزمون t نشان می‌دهد که در سطح اطمینان ۹۵%، همه متغیرهای مستقل (پیش‌بین) در شرکت‌های دارای صورت‌های مالی متقلبانه و بدون صورت‌های مالی متقلبانه، تفاوت معناداری با هم دارند. بنابراین، در ادامه به منظور پیاده‌سازی الگوریتم یادگیری عمیق از ۲۴ متغیر پیش‌بین انتخاب شده براساس آزمون t استفاده می‌گردد.

۴-۲-۲. اجرای مدل‌ها

به منظور پیاده‌سازی و ارزیابی الگوریتم‌ها، داده‌های پژوهش باید به دو دسته داده‌های آموزش و آزمایش تقسیم شوند. از مجموعه داده‌های آموزش برای ساخت و از داده‌های آزمایش نیز برای ارزیابی مدل ایجاد شده در مرحله آموزش استفاده می‌شود. تکنیک‌های دسته‌بندی را می‌توان براساس معیارهایی مانند صحّت، سرعت و پایداری با هم مقایسه نمود. صحّت یک روش دسته‌بندی، بستگی به تعداد پیش‌بینی‌های درستی که آن مدل انجام می‌دهد، دارد. سرعت یک روش دسته‌بندی نیز زمان لازم برای ساخت و استفاده از مدل در دسته‌بندی است و پایداری توانایی برخورد مدل در مواجهه با داده‌های غیرمعمول و یا مقادیر مفقود شده را نشان می‌دهد.

به منظور ارزیابی صحّت و پایداری مدل‌ها، داده‌های پژوهش ۵۰ مرتبه به صورت تصادفی به دو دسته داده‌های آموزش و آزمایش تقسیم شده و ایجاد مدل و ارزیابی نتایج حاصل از آنها انجام شد. به این صورت که در هر مرتبه ۷۵ درصد داده‌ها برای آموزش مدل و ۲۵ درصد آن را برای ارزیابی و آزمایش مدل مورد استفاده قرار گرفته و میانگین نتایج حاصل از ۵۰ مرتبه اجرای هر مدل، به عنوان نتیجه نهایی آن در نظر گرفته شده است.

http://sebaa.journal.qom-iau.ac.ir/

جدول ٤- پارامترهای شبیه‌سازی

پارامتر شبیه‌سازی	مقدار
تعداد نمونه‌ها کل دیتاست	۱۸۰۰ سال- شرکت (مشاهده)
تعداد ویژگی هر نمونه	۲۴ ویژگی
تعداد نمونه‌های آموزش مدل	۱۳۵۰
تعداد نمونه‌های تست مدل	۴۵۰

۴-۲-۳. پارامترهای ارزیابی مدل

تشخیص یک روش دسته‌بندی بر روی داده‌ها که تنها دارای دو کلاس می‌باشد، با یکی از موارد مثبت درست (TP)، منفی نادرست (FN)، منفی درست (TN) و مثبت نادرست (FP) بیان می‌شود. بر این اساس می‌توان پارامتر صحت (یا دقت)[1] را تعریف نمود. دقت، استانداردترین متریک برای خلاصه سازی عملکرد تشخیص در تمامی‌کلاس‌ها می‌باشد که بصورت فرمول زیر محاسبه می‌شود. رابطه (۶):

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

همانگونه که در جدول (۵) نشان داده شده است، برای محاسبه دقت دسته‌بندی، ابتدا باید ماتریس اغتشاش[2] محاسبه شود. داده‌های روی قطر اصلی، تشخیص‌های درست روش دسته‌بندی استفاده شده است.

جدول ۵- ماتریس اغتشاش

True Positives (TP)	False Positives (FP)
False Negatives (FN)	True Negatives (TN)

با توجه به هدف پژوهش حاضر که کشف صورت‌های مالی متقلبانه است، پارامترهای ماتریس اغتشاش به صورت زیر تعریف می‌شوند:

TP: درصد تشخیص درست شرکت‌های دارای صورت‌های مالی متقلبانه

TN: درصد تشخیص درست شرکت‌های بدون صورت‌های مالی متقلبانه

FP: درصد تشخیص غلط شرکت‌های دارای صورت‌های مالی متقلبانه به عنوان شرکت‌های بدون صورت‌های مالی متقلبانه

FN: درصد تشخیص غلط شرکت‌های بدون صورت‌های مالی متقلبانه به عنوان شرکت‌های دارای صورت‌های مالی متقلبانه.

1. Accuracy
2. Confusion Matrix

http://sebaa.journal.qom-iau.ac.ir/

در ادامه ماتریس اغتشاش هر روش دسته‌بندی در کشف صورت‌های مالی متقلبانه نشان داده شده است.

نتایج حاصل از الگوریتم ماشین بردار پیشتیبان که در قالب ماتریس اغتشاش و در جدول (۶) آمده است، نشان می‌دهد که میزان خطای مدل در شناسایی شرکت‌های دارای صورت‌های مالی متقلبانه و شرکت‌های بدون صورت‌های مالی متقلبانه به ترتیب 10.9% و 12.4% می‌باشد. در حالت کلی، نتایج حاصل از الگوریتم SVM نشان می‌دهد که میزان دقت این مدل در کشف صورت‌های مالی متقلبانه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، تقریباً 88.4% است، این موضوع نشان‌دهنده آن است که با توجه به متغیرهای انتخاب شده به روش T-test و الگوریتم ماشین بردار پیشتیبان، می‌توان با احتمال 88.4% شرکت‌های دارای صورت‌های مالی متقلبانه را شناسایی کرد.

جدول ٦- نتایج ماتریس اغتشاش در کشف صورت‌های مالی متقلبانه (الگوریتم ماشین بردار پشتیبان)

SVM & T-test		
	Positive	**Negative**
Positive	0.891	0.124
Negative	0.109	0.876
Accuracy	0.884	

نتایج حاصل از الگوریتم شبکه عصبی پیچشی که در قالب ماتریس اغتشاش و در جدول (۷) ارائه شده است، نشان می‌دهد که میزان خطای مدل در شناسایی شرکت‌های دارای صورت‌های مالی متقلبانه و شرکت‌های بدون صورت‌های مالی متقلبانه به ترتیب ۱۸/۳% و ۱۹/۱% می‌باشد. در حالت کلی، نتایج حاصل از الگوریتم CNN نشان می‌دهد که میزان دقت این مدل در کشف صورت‌های مالی متقلبانه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، تقریباً ۸۱/۳% است، این موضوع نشان‌دهنده آن است که با توجه به متغیرهای انتخاب شده به روش T-test و الگوریتم شبکه عصبی پیچشی، می‌توان با احتمال ۸۱/۳% شرکت‌های دارای صورت‌های مالی متقلبانه را شناسایی کرد.

جدول ۷- نتایج ماتریس اغتشاش در کشف صورت‌های مالی متقلبانه (الگوریتم شبکه عصبی پیچشی)

CNN & T-test		
	Positive	**Negative**
Positive	0.817	0.191
Negative	0.183	0.809
Accuracy	0.813	

http://sebaa.journal.qom-iau.ac.ir/

نتایج حاصل از الگوریتم شبکه عصبی بازگشتی که در قالب ماتریس اغتشاش و در جدول (۸) ارائه شده است، نشان می‌دهد که میزان خطای مدل در شناسایی شرکت‌های دارای صورت‌های مالی متقلبانه و شرکت‌های بدون صورت‌های مالی متقلبانه به ترتیب 5.2% و 6.1% می‌باشد. در حالت کلی، نتایج حاصل از الگوریتم RNN نشان می‌دهد که میزان دقت این مدل در کشف صورت‌های مالی متقلبانه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، تقریباً ۹۴/۴% است. این موضوع نشان‌دهنده آن است که با توجه به متغیرهای انتخاب شده به روش T-test و الگوریتم شبکه عصبی بازگشتی، می‌توان با احتمال ۹۴/۴% شرکت‌های دارای صورت‌های مالی متقلبانه را شناسایی کـرد.

جدول ۸- نتایج ماتریس اغتشاش در کشف صورت‌های مالی متقلبانه (الگوریتم شبکه عصبی بازگشتی)

RNN & T-test		
	Positive	**Negative**
Positive	0.948	0.061
Negative	0.052	0.939
Accuracy	0.944	

جدول ۹- خلاصه نتایج الگوریتم LSTM و سایر الگوریتم‌های یادگیری عمیق در کشف صورت‌های مالی متقلبانه

FN	Accuracy	الگوریتم
0.109	0.884	SVM
0.183	0.813	CNN
0.052	0.944	RNN

با توجه به جدول (۹)، به طور خلاصه نتایج حاصل از الگوریتم‌های یادگیری عمیق نشان می‌دهد که صحّت کلی الگوریتم‌های SVM، CNN و RNN به ترتیب 88.4%، 81.3% و 94.4% می‌باشد که نشان‌دهنده این است که الگوریتم RNN بهترین عملکرد و الگوریتم CNN بدترین عملکرد را در کشف شرکت‌های دارای صورت‌های مالی متقلبانه دارد. به عبارت دیگر، نتایج نشان‌دهنده کارآ بودن الگوریتم شبکه عصبی بازگشتی نسبت به سایر الگوریتم‌های یادگیری عمیق می‌باشد. بنابراین، در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، از بین سه الگوریتم SVM، CNN و RNN، الگوریتم RNN کاراترین مدل را برای کشف صورت‌های مالی متقلبانه فراهم می‌کند.

۵. نتیجه‌گیری

در محیط تجاری مدرن امروزی با پیچیده‌تر شدن سیستم‌ها و فعالیت‌ها و ازدیاد اطلاعات

http://sebaa.journal.qom-iau.ac.ir/

فریبکارانه، فناوری‌های مورد استفاده برای جلوگیری و کشف تقلب نیز به‌روز شده است. بنابراین، توسعه و رشد روش‌های مربوط به داده‌های متنوع ازجمله داده‌کاوی، برای تشخیص تقلب‌های مالی در اولویت است. در ایران به موضوع کشف تقلب صورت‌های مالی توجه زیادی نشده است. مسئله گزارشگری مالی متقلبانه در ایران از اهمیتی ویژه برخوردار است. افزایش تعداد شرکت‌های پذیرفته شده در بورس که به منظور جذب منابع مالی به انتشار اوراق بهادار اقدام می‌کنند، تلاش به منظور کاهش مالیات بر درآمد و... ازجمله دلایل اهمیت این موضوع است. بنابراین، در حوزه کشف تقلب هرچقدر حسابرسان و ذینفعان شرکت به تکنیک‌ها و رویکردهای کارآمدتری ازجمله الگوریتم‌های یادگیری عمیق و رویکردهای کارآمدی چون شبکه عصبی و ماشین بردار پشتیبان دسترسی داشته باشند، قاعدتاً، توانایی و کارایی آنها در حوزه تاییدپذیری اطلاعات صورت‌های مالی بالاتر خواهد رفت. کشف تقلب در صورت‌های مالی در چند دهه گذشته با تاکید بر ناهنجاری‌های حسابداری به طور کلی و به طور خاص بر تقلب در صورت‌های مالی مورد توجه قرار گرفته است.

در این راستا، پژوهش حاضر با استفاده از سه الگوریتم یادگیری عمیق (شامل الگوریتم‌های ماشین بردار پشتیبان، شبکه عصبی پیچشی و شبکه عصبی بازگشتی) و با بکارگیری جامع‌ترین متغیرهای اثرگذار بر تقلب در صورت‌های مالی (با بررسی جامع پژوهش داخلی و خارجی انجام شده در حوزه تقلب)، مدلی برای کشف صورت‌های مالی متقلبانه با دقت و کارائی بالا ارائه کرده است که این موضوع در نوع خود منحصربه‌فرد می‌باشد. از طریق تحلیل اطلاعات مربوط به ۱۵۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران طی سال‌های ۱۳۹۱ تا ۱۴۰۲، نتایج حاصل از الگوریتم‌های یادگیری عمیق نشان می‌دهد که صحّت کلی الگوریتم‌های SVM، CNN و RNN به ترتیب 88.4%، 81.3% و 94.4% می‌باشد که نشان‌دهنده این است الگوریتم RNN بهترین عملکرد و الگوریتم CNN بدترین عملکرد را در کشف شرکت‌های دارای صورت‌های مالی متقلبانه دارد. به عبارت دیگر، نتایج نشان‌دهنده کارآ بودن الگوریتم شبکه عصبی بازگشتی نسبت به سایر الگوریتم‌های یادگیری عمیق می‌باشد. بنابراین، در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران از بین سه الگوریتم SVM، CNN و RNN، الگوریتم RNN کاراترین مدل را برای کشف صورت‌های مالی متقلبانه فراهم می‌کند. این نتایج تا حدودی با یافته‌های پژوهش کاظمی و پیری (۱۴۰۱)، بهشتی مسئله‌گو و همکاران (۱۴۰۱)، سیاهکارزاده (۱۴۰۰)، زهیر و همکاران (۲۰۲۳)، علی و همکاران (۲۰۲۳) همخوانی دارد.

نوآوری و دستاوردهای علمی پژوهش حاضر به شرح زیر است: اول، از دیدگاه نظری، با توجه به بررسی پژوهش‌های تجربی صورت گرفته در ایران، این پژوهش اولین موردی است که با استفاده از

http://sebaa.journal.qom-iau.ac.ir/

الگوریتم حافظه طولانی کوتاه‌مدت، به کشف صورت‌های مالی متقلبانه در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران پرداخته است. دوم، با توجه به اینکه کیفیت و درستی صورت‌های مالی شرکت یکی از متغیرهای کلیدی در تصمیمات سرمایه‌گذاران است، این پژوهش با کشف صورت‌های مالی متقلبانه، با بکارگیری الگوریتم حافظه طولانی کوتاه مدت و سایر الگوریتم‌های یادگیری عمیق، سهامداران، تحلیلگران و اعتباردهندگان را در اتخاذ تصمیمات اثربخش و کارآمد یاری می‌کند و در نتیجه باعث افزایش سطح کارآیی بازار در بلندمدت می‌گردد. سوم، برای نهادهای نظارتی و تدوین‌کنندگان قوانین و استانداردها، مسئله تقلب به ویژه صورت‌های مالی متقلبانه، همواره در کانون توجه بوده است. یافته‌های این پژوهش می‌تواند به سازمان حسابرسی، سازمان بورس و اوراق بهادار و سایر نهادهای نظارتی، رهنمودهای لازم را در زمینه تدوین قوانین و استانداردها ارائه دهد. چهارماً، نتایج این پژوهش می‌تواند حسابرسان را در زمینه گردآوری شواهد و ارزیابی شواهد یاری کند. در نهایت، یافته‌های این پژوهش می‌تواند ایده‌های جدیدی در اختیار پژوهشگران و دانشگاهیان قرار دهد.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

http://sebaa.journal.qom-iau.ac.ir/

منـابـع

احمدی، سید جلال؛ فغانی ماکرانی، خسرو؛ فاضلی، نقی (۱۳۹۹). ارائه مدلی برای پیش‌بینی صورت‌های مالی متقلبانه و مقایسه صورت‌ها و نسبت‌های مالی با قانون بنفورد. *دانش حسابداری و حسابرسی مدیریت*، ۹(۳۵)، ص۲۲۱-۲۳۷.

احمدی، سید جلال؛ فغانی ماکرانی، خسرو؛ فاضلی، نقی (۱۴۰۳). تکنیک‌های داده‌کاوی و پیش‌بینی تقلب صورت‌های مالی. *دانش حسابداری و حسابرسی مدیریت*، ۱۳(۵۲)، ص۱۵-۲۸.

اعتمادی، حسین؛ زلقی، حسن (۱۳۹۲). کاربرد رگرسیون لجستیک در شناسایی گزارشگری مالی متقلبانه. *دانش حسابرسی*، ۱۳(۵۱)، ص۲۶-۱.

امیرمعزی، حسین؛ پورآقاجان، عباسعلی؛ جعفری، علی (۱۴۰۳). کشف تقلب صورت‌های مالی: قیاس توانایی مدل‌های مبتنی بر متغیرهای حسابداری. *دانش حسابداری و حسابرسی مدیریت*، ۱۳(۵۲)، ص۱۷۳-۱۸۸.

بهشتی مسئله‌گو، سیده مژگان؛ افشار کاظمی، محمدعلی؛ حقیقت منفرد، جلال؛ رضاییان، علی (۱۴۰۱). یادگیری عمیق برای پیش‌بینی بازار سهام با استفاده از اطلاعات عددی و متنی) رویکرد الگوریتم حافظه کوتاه مدت ماندگار. *مهندسی مالی و مدیریت اوراق بهادار (پذیرفته شده برای انتشار)*.

خواجوی، شکرالله؛ ابراهیمی، مهرداد (۱۳۹۶). ارائهٔ یک رویکرد محاسباتی نوین برای پیش‌بینی تقلب در صورت‌های مالی، با استفاده از شیوه‌های خوشه‌بندی و طبقه‌بندی (شواهدی از شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران). *پیشرفت‌های حسابداری*، ۹(۲)، ص۱-۳۴.

رضائی، مهدی؛ ناظمی اردکانی، مهدی؛ ناصر صدرآبادی، علیرضا (۱۴۰۰). پیش‌بینی تقلب صورت‌های مالی با استفاده از رویکرد کریسپ (CRISP). *دانش حسابداری و حسابرسی مدیریت*، ۱۰(۴۰)، ص۱۵۰-۱۳۵.

رضائی، مهدی؛ ناظمی اردکانی، مهدی؛ ناصر صدرآبادی، علیرضا (۱۴۰۰). پیش‌بینی تقلب صورت‌های مالی با استفاده از رویکرد کریسپ(CRISP). *دانش حسابداری و حسابرسی مدیریت*، ۱۰(۴۰)، ص۱۳۵-۱۵۰.

زارعی، علی؛ رهنمای رودپشتی، فریدون؛ خان‌محمدی، محمدحامد؛ کردلویی، حمیدرضا (۱۴۰۳). ارائه الگوی پیش‌بینی تقلب مبتنی بر هوش مصنوعی (بیز ساده). *مطالعات اخلاق و رفتار در حسابداری و حسابرسی*، ۴(۴)، ص۷-۲۶.

سیاهکارزاده، علی‌محمد (۱۴۰۰). *پیش‌بینی بازارهای مالی با استفاده از یادگیری عمیق*. پایان‌نامه کارشناسی ارشد، مهندسی کامپیوتر. دانشگاه صنعتی شاهرود.

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

کاظمی، توحید؛ پیری، پرویز (۱۴۰۱). پیش‌بینی طرح تقلب در گزارشگری مالی با استفاده از رویکرد یادگیری ماشین در فضای چند کلاسه. *پژوهش‌های تجربی حسابداری*، ۱۲(۴۶)، ص۲۷۶-۲۵۵.

ملکی کاکلر، حسن؛ بحری ثالث، جمال؛ جبارزاده کنگرلویی، سعید؛ آشتاب، علی (۱۴۰۰). کارایی مدل‌های آماری و الگوهای یادگیری ماشین در پیش‌بینی گزارشگری مالی متقلبانه. *اقتصاد مالی*، ۱۵(۵۴)، ص۲۶۷-۲۹۲.

Albashrawi, M. (2016). Detecting financial fraud using data mining techniques: *J. Data Sci*. no.14, p. 553–569.

Alfaiz, N.S. & Fati, S.M. (2022). Enhanced Credit Card Fraud Detection Model Using Machine Learning. *Electronics*, no.11, p. 662.

Alghofaili, Y., Albattah, A., Rassam, M. & Rassam, M. (2020). A Financial Fraud Detection Model

http://sebaa.journal.qom-iau.ac.ir/

Based on LSTM Deep Learning Technique. *Journal of Applied Security Research*, 15(4), p. 498–516. https://doi.org/10.1080/19361610.2020.1815491.

Ali, A.A., Khedr, A.M., El-Bannany, M. & Kanakkayil, S. (2023). A Powerful Predicting Model for Financial Statement Fraud Based on Optimized XGBoost Ensemble Learning Technique. *Appl. Sci*.no. 13, p. 2272. https://doi.org/10.3390/ app13042272

Alsinglawi, M.M.A.S.M.A.O. & Almari, M.O.S. (2021). Predicting Fraudulent Financial Statements Using Fraud Detection Models. *Acad. Strateg. Manag*, no. 20, p.1–17.

Amar, I.A.A.B. & Jarboui, A. (2013). Detection of Fraud in Financial Statements: French Companies as a Case Study. *Int. J. Acad. Res. Bus. Soc. Sci*. no. 3, p. 456–472.

Andrew, C. & Robin, C (2022). Detecting Fraudulent of Financial Statements Using Fraud S.C.O.R.E Model and Financial Distress. *Int. J. Econ. Bus. Account. Res.* no. 6, p. 211–222.

Asimit, A.V., Kyriakou, I., Santoni, S., Scognamiglio, S. & Zhu, R. (2022). Robust Classification via Support Vector Machines. *Risks*, no.10, p.154.

Cecchini, M., Aytug, H., Koehler, G.J. & Pathak, P. (2010). Detecting management fraud in public companies. *Manag. Sci*. no. 56, p.1146–1160.

Chen, S. (2016). Detection of fraudulent financial statements using the hybrid data mining approach. *SpringerPlus*, no. 5, p. 1–16.

Craja, P., Kim, A. & Lessmann, S. (2020). Deep learning for detecting financial statement fraud. Decis. *Support Syst*, no.139, p.113421.

Dechow, P., Ge, W., Larson, C. & Sloan, R. (2011). Predicting Material Misstatements. *Contemporary Accounting Research*, 28(1), p. 17-82.

Deebak, B., Memon, F.H., Dev, K., Khowaja, S.A., Wang, W. & Qureshi, N.M.F. (2022).TAB-SAPP: A trust-aware blockchain-based seamless authentication for massive IoT-enabled industrial applications. *IEEE Trans. Ind. Inform*. No. 19, p. 243250.

El-Bannany, M., Sreedharan, M. & Khedr, A.M. (2020). A robust deep learning model for financial distress prediction. *Int. J. Adv. Comput. Sci. Appl*. No. 11, p.170–175.

Gupta, R. & Gill, N.S. (2012). Prevention and detection of financial statement fraud–An implementation of data mining framework. *Editor. Pref*. no 3, p.150–160

Gupta, S. & Mehta, S.K. (2021). Data mining-based financial statement fraud detection: Systematic literature review and meta-analysis to estimate data sample mapping of fraudulent companies against non-fraudulent companies. *Glob. Bus. Rev*., 25(2), p. 1–26.

Hamal, S. & Senvar, Ö. (2021). Comparing performances and effectiveness of machine learning classifiers in detecting financial accounting fraud for Turkish SMEs. *Int. J. Comput. Intell. Syst*. No. 14, p. 769–782.

Han, D. (2017). Researches of Detection of Fraudulent Financial Statements Based on Data Mining. *J. Comput. Theor. Nanosci*., no. 14, p. 32–36.

Humpherys, S.L., Moffitt, K.C., Burns, M.B., Burgoon, J.K. & Felix, W.F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decis. Support Syst.* No. 50, p. 585–594.

http://sebaa.journal.qom-iau.ac.ir/

Jan, Ch. (2021). Detection of Financial Statement Fraud Using Deep Learning for Sustainable Development of Capital Markets under Information Asymmetry. *Sustainability* 13(17), p. 9879.

Katsis, D.Ch. & et al. (2012). Using Ants to Detect Fraudulent Financial Statements. *Journal of Applied Finance & Banking*, 2(6), p. 73-81.

Kulikova, L. & Satdarova, D. (2016). Internal control and compliance-control as effective methods of management, detection and prevention of financial statement fraud. *Acad. Strateg. Manag. J.* no. 15, p. 92.

Kumar, R. & Tripathi, R. (2020). *Secure healthcare framework using blockchain and public key cryptography*. In: Blockchain Cybersecurity, Trust and Privacy; Springer: Cham, Switzerland, p. 185–202.

Kumar, S., Ahmed, R., Bharany, S., Shuaib, M., Ahmad, T., Tag Eldin, E., Rehman, A.U. & Shafiq, M. (2022). Exploitation of Machine Learning Algorithms for Detecting Financial Crimes Based on Customers' Behavior. *Sustainability*, no.14, p. 13875.

Lin, C.C., Chiu, A.A., Huang, S.Y. & Yen, D.C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. *Knowl-Based Syst*, no. 89, p. 459–470.

Moepya, S.O., Akhoury, S.S. & Nelwamondo, F.V. (2014). *Cost-sensitive classification for financial fraud detection under high classimbalance*. In: Proceedings of the 2014 IEEE international conference on data mining workshop, Shenzhen, China, 14–17 December 2014; IEEE: New York, NY, USA, p. 183–192.

Mohammadi, M., Yazdani, S., Khanmohammadi, M.H. & Maham, K. (2020). Financial reporting fraud detection: An analysis of data mining algorithms. *Int. J. Financ. Manag. Account*, no. 4, p. 1–12.

Murorunkwere, B.F., Tuyishimire, O., Haughton, D. & Nzabanita, J. (2022). Fraud Detection Using Neural Networks: A Case Study of Income Tax. *Future Internet*, no.14, p. 168.

Pai, P.F., Hsu, M.F. & Wang, M.C. (2011). A support vector machine-based model for detecting top management fraud. *Knowl-Based Syst*, no. 24, p. 314–321.

Paulo Sérgio Gomes Macedo, H.C.I. & Vieira, E.S. (2022). A model to detect financial statement fraud in Portuguese companies by the auditor. *Contaduría Adm*, no. 67, p. 185–209.

Pérez López, C., Delgado Rodríguez, M. & de Lucas Santos, S. Tax Fraud Detection through Neural Networks: An Application Using a Sample of Personal Income Taxpayers. *Future Internet*, no. 11, p. 86.

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Audit. J. Pract. Theory*, no. 30, p. 19–50.

Perols, J. (2011). Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms. *A Journal of Practice & Theory*, 30(2), p. 19-50.

Ravisankar, P., Ravi, V., Rao, G.R. & Bose, I. (2011). Detection of financial statement fraud and feature selection using data mining techniques. *Decis. Support Syst.* No. 50, p. 491–500.

Rizki, A.A., Surjandari, I. & Wayasti, R.A. (2017). *Data mining application to detect financial fraud*

http://sebaa.journal.qom-iau.ac.ir/

in Indonesia's public companies. In: Proceedings of the 2017 3rd International Conference on Science in Information Technology (ICSITech), Bandung, Indonesia, 25–26 October 2017; IEEE: New York, NY, USA, p. 206–211.

Schreiber-Gregory, D. & Bader, K. (2018). *Logistic and Linear Regression Assumptions: Violation Recognition and Control*. In: Proceedings of the SESUG Conference, St. Pete Beach, FL, USA, p. 1–6.

Spathis, C., Doumpos, M. & Zopounidis, C. (2002). Detecting falsified financial statements: A comparative study using multicriteria analysis and multivariate statistical techniques. *European Accounting Review*, 11(3), p. 509-535.

Sreedharan, M., Khedr, A.M. & El Bannany, M. (2020). A Multi-Layer Perceptron Approach to Financial Distress Prediction with Genetic Algorithm. *Autom. Control. Comput. Sci.* no. 54, p. 475–482.

Zaheer, S., Anjum, N., Hussain, S., Algarni, A.D., Iqbal, J., Bourouis, S. & Ullah, S.S. (2023). A Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model. *Mathematics*, no.11, p. 590. https://doi.org/10.3390/ math11030590

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

http://sebaa.journal.qom-iau.ac.ir/