Business Intelligence Management Studies
Vol. 13, No.50, Winter 2025; P 99-142
ims.atu.ac.ir
DOI: 10.22054/ims.2024.78018.2434

ATU PRESS

Original Research
Accepted: 06 May 2024
Revised: 04 April 2024
Received: 30 January 2024
eISSN: 2821-0816
ISSN: 2821-0964

Detecting Financial Fraud in Public Companies Using Financial and Non-Financial Metrics with a Machine Learning Approach

Abbas Bagherian Kasgari –
PhD student in Information Technology Management, Allameh Tabatabai University, Tehran, Iran

Iman Raeesi Vanani
Associate Professor, Department of Industrial Management, Faculty of Management, Allameh Tabatabai University

Maghsoud Amiri
Full Professor, Department of Industrial Management, Faculty of Management and Accounting, Allameh Tabatabai University, Tehran, Iran

Saeid Homayoun
Associate Professor, Department of Accounting, Gävle University, Gävle, Sweden.

rbstr act

Most traditional fraud detection systems primarily focus on financial criteria to identify financial fraud, often overlooking the potential for fraudulent companies to engage in various types of non-financial misconduct. Recent studies have predominantly highlighted the significance of financial data as the sole indicator of fraud, neglecting the exploration of non-financial or Environmental, Social, and Governance (ESG) metrics as supplementary predictors. This research aims to enhance fraud prediction by integrating financial and ESG data through sophisticated machine learning and deep learning models. It examines the effectiveness of supervised machine learning and deep learning algorithms in detecting financial fraud over a 10-

– Corresponding Author: AbbasBagherian@yahoo.com

How to Cite: Bagherian Kasgari, A., Raeesi Vanani, I., Amiri, M., Homayoun, S. (2025). Detecting Financial Fraud in Public Companies Using Financial and Non-Financial Metrics with a Machine Learning Approach, *Journal of Business Intelligence Management Studies*, 13(50), 99-142. DOI: 10.22054/ims.2024.78018.2434

year period ending in 1401. This study innovatively demonstrates that a hybrid model, which combines financial and non-financial criteria, yields superior predictive accuracy for financial fraud than models based solely on financial data**.** The results of this study, addressing the first research question, indicate that among various machine learning and deep learning algorithms, the classification or bagging algorithm demonstrated superior efficiency. Furthermore, in response to the second research question, it was found that the dataset encompassing all features—integrating both financial and non-financial data—outperformed those datasets limited to either financial or non-financial data alone. The research results indicated that the bagging machine learning algorithms act the best with combined feature set including financial and ESG metrics combined. The adoption of the proposed model significantly improves the accuracy and effectiveness of fraud detection systems.

Introduction

In an era marked by rapid advancements in data analytics and increasing corporate accountability, the detection of financial fraud has become a priority for stakeholders across the global business landscape. Traditional fraud detection systems have primarily focused on analyzing financial data, often at the expense of overlooking non-financial metrics that may equally signal fraudulent activities. This oversight is significant considering the growing evidence suggesting that non-financial indicators, particularly Environmental, Social, and Governance (ESG) metrics, can provide critical insights into the operational integrity of organizations.

Literature Review

Recent scholarly works and industry reports have highlighted a significant shift towards integrating ESG metrics with financial data to enhance the predictive accuracy of fraud detection systems. This integration reflects an expanded understanding of what constitutes corporate transparency and accountability, extending beyond mere financial disclosures to include broader sustainability and governance factors. Indeed, the integration of these diverse data sources promises a more holistic approach to fraud detection, aligning with contemporary demands for corporate responsibility and ethical business practices. The research presented in this paper builds on this

foundation by employing advanced machine learning (ML) and deep learning (DL) algorithms to analyze a combination of financial and non-financial metrics. The study's innovative approach leverages a decade's worth of data from over 6000 public companies, utilizing a variety of ML and DL models to explore the efficacy of integrated datasets in predicting fraudulent activities more effectively than traditional methods. The findings aim to contribute not only to academic discourse but also to practical applications in corporate governance, offering valuable insights for regulators, investors, and policymakers committed to upholding the highest standards of corporate ethics and governance. By synthesizing complex data sets and applying sophisticated analytical techniques, this research underscores the potential of ML and DL models to revolutionize fraud detection, setting a new standard for both the scope and depth of fraud analysis.

Objective

The primary goal of this research is to improve financial fraud detection in public enterprises by integrating Environmental, Social, and Governance (ESG) metrics with traditional financial data, using machine learning (ML) and deep learning (DL) techniques. This approach addresses the limitations of traditional systems that focus mainly on financial indicators, often missing non-financial signs of fraud. This study rigorously tests various ML and DL models trained on ESG-enriched datasets against those using only financial data, exploring whether a holistic approach can enhance fraud predictiveness. The research aims to offer a broader view of company operations, in line with sustainable practices, potentially shifting how data science is applied in fraud detection. Ultimately, this study seeks to enrich discussions on integrating financial and non-financial data in fraud detection, influencing future corporate risk and governance strategies, and improving fraud prediction accuracy in line with emerging standards of corporate accountability and transparency.

Method

This study employs a sophisticated analytical approach using machine learning (ML) and deep learning (DL) to enhance financial fraud detection, leveraging a robust dataset that includes both traditional financial indicators and Environmental, Social, and Governance

(ESG) metrics from over 6000 public companies worldwide. These metrics, sourced from reputable databases such as Thomson Reuters ASSET4, are crucial for advanced analyses. The methodology involves thorough data preprocessing, including handling missing values, normalizing data, and encoding categorical variables, with a focus on balancing the dataset using oversampling techniques to counter class imbalance and improve model generalization for detecting rare fraudulent cases.

The research rigorously evaluates various ML and DL models like Decision Trees, Naive Bayes, SVM, CNN, LSTM, and ensemble methods such as Bagging, Extra Trees, and Random Forests. The models are trained and tested on divided datasets to assess their effectiveness using metrics like accuracy, precision, recall, F1-score, and the Matthews Correlation Coefficient (MCC), with extensive validation techniques including cross-validation to ensure stability and prevent overfitting. The models' performance is compared with baseline models that use only financial data, highlighting the benefits of integrating ESG metrics for deeper insights and enhanced predictiveness in fraud detection.

Results

This study evaluates the integration of Environmental, Social, and Governance (ESG) metrics with traditional financial data in detecting financial fraud using various machine learning (ML) and deep learning (DL) algorithms. Results highlight the enhanced performance of fraud detection models when using combinations of financial and ESG metrics. Notably, the Extra Tree classifier and bagging algorithms excelled, particularly when analyzing balanced datasets that included both types of metrics. The use of oversampling techniques proved crucial in improving detection rates for rare fraudulent cases, thus balancing the dataset and reducing bias.

Models integrating both financial and ESG data consistently outperformed those using only one data type, enhancing accuracy, precision, recall, and F1 score. This underscores the value of a multidimensional approach in fraud detection. Advanced metrics like the Matthews Correlation Coefficient (MCC) and the Area Under the ROC Curve (AUC) provided a nuanced assessment of model performance, with higher MCC and AUC values indicating greater

effectiveness in identifying fraudulent activities. The integration of ESG metrics was particularly effective in identifying potential fraud in companies that might appear financially sound but engage in unethical practices**.**

The findings recommend that companies, regulatory bodies, and technology developers adopt integrated approaches that encompass both financial and ESG data to improve fraud detection. Future research could focus on real-time data integration and more complex models like hybrid deep learning frameworks to further boost detection capabilities. The study demonstrates that using ESG metrics alongside financial data with advanced ML techniques significantly improves the accuracy and reliability of fraud detection systems, aligning with sustainable business practices and setting the stage for future innovations in fraud detection. This comprehensive approach not only yields superior performance but also enhances the model's capabilities, emphasizing the effectiveness of combining financial and non-financial data.

Conclusion

This research significantly advances the use of machine learning (ML) and deep learning (DL) in detecting financial fraud, highlighting the integration of Environmental, Social, and Governance (ESG) metrics with traditional financial data to enrich datasets and enhance model predictive power. Models trained on datasets combining financial and ESG metrics show superior performance in accuracy, precision, recall, and F1 score, improving anomaly detection and fraud prediction. The use of oversampling techniques addresses class imbalance issues, enhancing sensitivity to rare fraudulent cases and boosting the performance of ensemble methods like the Extra Tree classifier**.**

The findings highlight the critical role of ESG metrics in enhancing corporate governance and risk management, providing deeper insights into non-financial behaviors that indicate potential risks, which supports more informed decision-making and boosts transparency. Future research should investigate real-time fraud detection systems and the use of unsupervised and semi-supervised models to adapt to evolving fraud tactics. Practitioners are encouraged to adopt advanced machine learning (ML) and deep learning (DL) techniques, incorporating ESG metrics to improve fraud detection

systems' accuracy and reliability, aligning with sustainable business practices and setting new standards in fraud detection technology.

Keywords: Fraud Detection Intelligent Systems, Deep Learning, Machine Learning, Financial Metrics, Non-Financial Metrics (ESG).

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقاله پژوهشی تاریخ ارسال: ۱۴۰۲/۱۱/۱۰ تاریخ بازنگری: ۱۴۰۳/۰۱/۱۶ تاریخ پذیرش: ۱۴۰۳/۰۲/۱۷
ISSN: 2821-0964 eISSN: 2821-0816

مطالعات مدیریت کسب و کار هوشمند
سال سیزدهم، شماره ۵۰، زمستان ۱۴۰۳، ص ۹۹ تا ۱۴۲
ims.atu.ac.ir
DOI: 10.22054/ims.2024.78018.2434

شناسایی تقلب مالی در شرکت‌های سهامی عام با استفاده معیارهای مالی و غیرمالی با رویکرد یادگیری ماشین

عباس باقریان کاسگری * دانشجوی دکتری مدیریت فناوری اطلاعات، دانشگاه علامه طباطبائی، تهران، ایران

ایمان رئیسی وانانی دانشیار گروه مدیریت صنعتی، دانشکده مدیریت، دانشگاه علامه طباطبائی، تهران، ایران

مقصود امیری استاد گروه مدیریت صنعتی، دانشکده مدیریت و حسابداری، دانشگاه علامه طباطبائی، تهران، ایران

سعید همایون دانشیار گروه حسابداری، دانشگاه گوله، گوله، سوئد

چکیده

بیشتر سیستم‌های تشخیص تقلب سنتی برای شناسایی تقلب مالی، صرفاً معیارهای مالی را در نظر می‌گیرند. این در حالی است که احتمال دارد شرکت‌های متقلب علاوه بر تقلب‌های مالی، مرتکب سایر انواع تقلب‌های غیرمالی نیز شوند. اگرچه تحقیقات اخیر به‌طور مکرر بر اهمیت داده‌های مالی به‌عنوان تنها فاکتور پیش‌بینی‌کننده تقلب تأکید کرده‌اند، هیچ مطالعه‌ای بر معیارهای غیرمالی یا ESG به‌عنوان یک فاکتور کمکی برای پیش‌بینی تقلب انجام نشده است. لذا هدف این تحقیق، بررسی امکان بهبود پیش‌بینی تقلب‌ها بر اساس ترکیبی از داده‌های مالی و ESG با ارائه یک مدل یادگیری ماشینی پیشرفته و یادگیری عمیق است. در این پژوهش، با استفاده از الگوریتم‌های نظارت‌شده یادگیری ماشین و یادگیری عمیق، به بررسی چگونگی شناسایی تقلب مالی در بازه زمانی ۱۰ ساله منتهی به سال ۱۴۰۱ پرداخته می‌شود. این تحقیق به‌طور نوآورانه‌ای نشان می‌دهد که استفاده از مدل ترکیبی مبتنی بر داده‌های مالی و معیارهای غیرمالی، قدرت پیش‌بینی‌کننده‌ای بهتری برای تقلب‌های مالی نسبت به تکیه صرف بر داده‌های مالی دارد.

* نویسنده مسئول : AbbasBagherian@yahoo.com

مطابق یافته‌های این تحقیق، در پاسخ به پرسش اول، بالاترین کارایی در میان الگوریتم‌های یادگیری ماشین و یادگیری عمیق، مربوط به الگوریتم طبقه‌بندی Bagging مشاهده گردید. همچنین یافته‌های این تحقیق در خصوص سؤال دوم نشان می‌دهد که مجموعه داده‌ای شامل همه ویژگی‌ها (مدل ترکیب داده‌های مالی و غیرمالی) کارایی بهتری در مقایسه با مجموعه داده‌های مالی و غیرمالی (به‌تنهایی) داشته است.

کلیدواژه‌ها: سیستم‌های هوشمند کشف تقلب، یادگیری عمیق، یادگیری ماشین، معیارهای مالی، معیارهای غیرمالی (ESG).

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

مقدمه

موضوع شناسایی تقلب در شرکت‌های سهامی عام، یک چالش جهانی بزرگ است که عدم موفقیت در آن تأثیرات منفی بر سهامداران و ذینفعانی مانند سرمایه‌گذاران خرد دارد و خسارات قابل توجهی بر اقتصاد جهانی وارد می‌کند. بر اساس گزارش دفتر مبارزه با مواد مخدر و جرم سازمان ملل متحد (۲۰۲۰)، هزینه تخمینی سالانه جرائم پولشویی به تنهایی بین ۲ تا ۵ درصد از تولید ناخالص داخلی جهانی، معادل بالغ بر ۲ تریلیون دلار است. با این حال، این ارقام خیره‌کننده تنها بخشی از هزینه کلی تقلب را نشان می‌دهند.

هنگام در نظر گرفتن هزینه‌های غیرمستقیم، مانند آسیب به شهرت، افزایش هزینه‌های عملیاتی و از دست دادن کسب و کار به دلیل کاهش اعتماد، تأثیر کل تقلب بسیار عمیق‌تر است. این مسأله می‌تواند ریسک را در صنایع مختلف افزایش دهد و به‌طور بالقوه منجر به ورشکستگی شرکت‌ها شود. با توجه به اینکه تقلب مالی ممکن است اثرات غیرمالی نیز به همراه داشته باشد، در این تحقیق متغیرهای غیرمالی نیز در کنار متغیرهای مالی برای پیش‌بینی ریسک وقوع تقلب مورد بررسی قرار می‌گیرند.

در این تحقیق، ابتدا با بررسی تحقیقات گذشته، به شناسایی چالش‌های موجود که در تحقیقات پیشین پوشش داده نشده‌اند پرداخته شده و سپس چگونگی پوشش این چالش‌ها با استفاده از یادگیری ماشین و یادگیری عمیق مورد بررسی قرار گرفته است. نتایج تحقیق جاری دارای فواید زیادی برای کشور خواهد بود و دریچه‌ای به روی محققان آتی برای گسترش این حوزه باز خواهد کرد.

پیشینه پژوهش

یکی از حوزه‌های مورد مطالعه که مورد توجه محققان در زمینه تشخیص تقلب مالی قرار گرفته است، ارزیابی ریسک شرکت‌ها در نتیجه رویدادها، فعالیت‌ها و اخبار مرتبط با تقلب مالی است. این ارزیابی میزان ریسک‌پذیری هر دارایی، بنگاه، شخص، محصول، بانک و غیره را مشخص می‌کند.

وقتی از ریسک تقلب مالی صحبت می‌شود، این ریسک شامل انواع مختلفی است که به انواع تقلب مالی مرتبط می‌شود، مانند ریسک ورشکستگی، ریسک اعتباری، ریسک رتبه اعتباری شرکت، ریسک وام یا بیمه‌نامه، رتبه‌بندی وثیقه، درخواست وام، از بین رفتن اعتبار مصرف‌کننده، رتبه شرکت، ریسک تصمیم‌گیری برای انتخاب وام، پیش‌بینی اوضاع بد مالی، پیش‌بینی عدم موفقیت در تجارت و سایر ریسک‌های مالی می‌باشد.. (ازبیگلو[1]. و همکاران، ۲۰۲۰)

پیش‌بینی ریسک‌های تقلب برای ارزیابی و شناسایی صحیح وضعیت ریسک شرکت در چنین مواردی بسیار مهم است، زیرا قیمت‌گذاری دارایی شرکت متأثر از این ارزیابی ریسک است. در سالیان گذشته، بحران وام مسکن بر اساس ارزیابی نادرست از ریسک تعویض اعتبار (CDS) بین مؤسسات مالی رخ داد و باعث شد که حباب املاک و مستغلات در سال ۲۰۰۸ ترکیده و منجر به رکود بزرگ شود (مارتین دین بیلی[2] و همکاران، ۲۰۰۸). بسیاری از مطالعات ارزیابی ریسک بر اساس اطلاعات امتیازدهی اعتباری و طبقه‌بندی وضعیت حساب‌های بانکی انجام شده‌اند. همچنین، برخی از این مطالعات به بررسی امکان تشخیص یا پیش‌بینی تقلب در شرکت‌ها با استفاده از اطلاعات مالی پرداخته‌اند.

در تحقیقات مربوط به ارزیابی ریسک تقلب مالی، بیشتر از اطلاعات صورت‌های مالی شرکت، گزارش‌های مالی یا گزارش‌های تحلیلی و همچنین داده‌های خرد و کلان استفاده شده است. در نتیجه، مدل‌هایی که با موفقیت داده‌های مکانی و زمانی را ادغام می‌کنند، احتمالاً عملکرد بهتری دارند. بنابراین، محققان قبلی این حوزه ترجیح داده‌اند از مدل‌های DMLP، LSTM و CNN استفاده کنند.

از آنجایی که ارزیابی ریسک مالی می‌تواند به‌عنوان یک مسأله طبقه‌بندی در نظر گرفته شود، انتخاب مدل‌های DMLP و CNN گزینه‌های منطقی‌تری برای چنین تحقیقاتی بودند. با این حال، به نظر می‌رسد که استفاده از LSTM برای حل مسائل راحت‌تر بوده و به

1. A. M. Ozbayoglu
2. Martin Neil Baily

همین دلیل بیشتر مورد استقبال محققان قبلی قرار گرفته است.

در این میان، هرگاه محقق دسترسی بهتری به داده‌های وسیع داشته، از روش یادگیری عمیق استفاده کرده است. در میان مجموعه داده‌های مورداستفاده، داده‌های صورت‌های مالی، داده‌های قیمت و داده‌های مرتبط با آمار مصرف و فروش شرکت‌ها بیشتر مشاهده می‌شود. تعداد اندکی از محققان که از داده‌های دیگری غیر از ارقام مالی استفاده کرده‌اند، به سراغ داده‌های یادداشت‌های متنی در صورت‌های مالی رفته‌اند که باز هم به نوعی مرتبط با اطلاعات سوابق مالی شرکت است (ازبیگلو و همکاران، ۲۰۲۰).

بیشتر سیستم‌های تشخیص تقلب سنتی برای تجزیه و تحلیل تقلب مالی، صرفاً معیارهای مالی را برای کشف تقلب در نظر می‌گیرند. این رویکرد به نظر می‌رسد که برای بررسی جامع انواع تقلب در شرکت‌ها کافی نیست، زیرا احتمال دارد شرکت‌های آسیب‌پذیرتر علاوه بر تقلب‌های مالی، مرتکب سایر انواع تقلب‌های غیرمالی نیز شوند. اگرچه تحقیقات اخیر به‌طور مکرر بر اهمیت داده‌های مالی به‌عنوان تنها فاکتور پیش‌بینی‌کننده تقلب تأکید کرده‌اند، هیچ مطالعه‌ای بر معیارهای غیرمالی یا ESG به‌عنوان یک فاکتور کمکی برای پیش‌بینی تقلب انجام نشده است. حتی اکثر محققان اخیر نیز برای تشخیص تقلب و ارزیابی ریسک آن، از بررسی صرفاً مبتنی بر داده‌های مالی با روش‌های یادگیری ماشین و یادگیری عمیق استفاده کرده‌اند.

در میان این تحقیقات، جای خالی بررسی تقلب با معیارهای غیرمالی به وضوح احساس می‌شود. بنابراین، در این تحقیق سعی بر آن است که نقاط ابهام مذکور که در تحقیقات قبلی به آن‌ها پاسخ داده نشده است، پوشش داده شوند. هدف این تحقیق، ارائه یک مدل یادگیری ماشینی پیشرفته و یادگیری عمیق برای پیش‌بینی تقلب‌های شرکت‌های سهامی عام بر اساس ترکیبی از داده‌های مالی و ESG است. این تحقیق به دنبال پر کردن این شکاف و کمک به توسعه سیستم‌های پشتیبانی تصمیم برای کشف تقلب مالی می‌باشد.

بررسی کمی با یادگیری عمیق و یادگیری ماشین

هوش محاسباتی در تحقیقات حوزه مدیریت در چند دهه گذشته مورد توجه قرار گرفته

است. مطالعات متعددی منتشر شده است که نتیجه آن مدل‌های مختلفی است. در این راستا، روش‌های یادگیری ماشین واخیراً یادگیری عمیق، به دلیل فراتربودن از مدل‌های کلاسیک در حل مسائل این حوزه، توجه بسیاری را به خود جلب کرده است.

یادگیری عمیق[1] *روشی* برای ساخت مدل‌های یادگیری ماشین[2] است که جهت یادگیری نمایشی سلسله مراتبی از داده‌ها به کاربرد دارند. شبکه‌های عصبی عمیق، اصطلاحی کلی است که به مجموعه‌ای از شبکه‌های عصبی با معماری چند لایه اشاره دارد. این شبکه‌ها قادرند با استفاده از تعداد زیادی لایه، ساختارهای بازنمایی مورد نیاز در یادگیری عمیق را ایجاد کنند. الگوریتم‌های یادگیری ویژگی‌های مختلفی دارند که می‌توانند به‌صورت با ناظر و بدون ناظر عمل نمایند و در تنظیم وزن‌های شبکه‌های یادگیرنده به کار گرفته می‌شوند (مهدی قیومی، ۲۰۲۲).

مطابق سوابق تحقیقات گذشته در حوزه سیستم‌های تشخیص تقلب، الگوریتم‌های یادگیری عمیق مختلفی قابل‌استفاده‌اند، از قبیل CNN یا شبکه کانولوشن، RNN یا شبکه بازگشتی، LSTM یا شبکه عمیق حافظه کوتاه‌مدت ماندگار، ماشین بولتزمن محدودشده[3]، شبکه باور عمیق[4] و شبکه رمزنگارهای خودکار[5]. معمولاً برای بررسی داده‌های مربوط تقلب، به علت اینکه داده‌های موجود دارای برچسب هستند، از الگوریتم‌های طبقه‌بندی استفاده می‌شود. الگوریتم‌های یادگیری تحت نظارت قادرند آنچه را که در گذشته یادگرفته شده را به‌منظور انجام پیش‌بینی رویدادهای آتی، با استفاده از نمونه‌های برچسب‌گذاری شده، برای داده‌های جدید به‌کارگیرند. با آغاز فرایند تجزیه‌وتحلیل مجموعه‌ای از داده‌های شناخته‌شده، الگوریتم یادگیری قادرند با استفاده از یک تابع انتزاعی برای پیش‌بینی مقادیر خروجی اقدام نمایند. سیستم نیز می‌تواند اهداف هر ورودی جدید را پس از آموزش کافی مهیا سازد. الگوریتم یادگیری می‌توانند خروجی خود را با

1. Deep Learning
2. Machine Learning
3. Restricted Boltzmann Machines (RBMs)
4. Deep Belief Networks (DBNs)
5. Autoencoders (AEs)

خروجی صحیحِ در نظر گرفته‌شده، مقایسه نموده و باهدف تغییر مدل، خطای الگوریتم را محاسبه نمایند. در این تحقیق از میان الگوریتم یادگیری ماشین تحت نظارت و همچنین الگوریتم‌های طبقه‌بندی[1] استفاده می‌شود. در این پژوهش از یک شیوه نوآورانه‌ای برای دستیابی به پیش‌بینی با صحت بالاتر استفاده می‌گردد که به آن حل مسئله با رویکرد همجوشی اطلاعات اطلاق می‌گردد. استفاده از روش همجوشی اطلاعات جهت افزایش کارایی تشخیص الگو می‌باشد. ترکیب تصمیم‌گیری جهت طبقه‌بندی تقلب و تشخیص آن با همجوشی چند الگوریتم کشف تقلب و ترکیب اطلاعات موجود از شرکت‌ها محقق می‌شود. در این راه‌کار، الگوریتم‌های طبقه‌بندی‌ها در یک ساختار موازی قرار می‌گیرند و الگوی ورودی به همه طبقه‌بندها اعمال می‌گردد و هر یک از آن‌ها جوابی برای مسئله ارائه می‌نمایند. سپس با تلفیق مناسبی از جواب‌های مذکور، پاسخ نهایی به دست می‌آید. در این حالت، منظور از ترکیب، ارتقای صحت حل مسئله و بالابردن کارایی پیش‌بینی سیستم تشخیص تقلب است. همچنین نتیجه این ترکیب را با الگوریتم یادگیری جمعی[2] جهت اعتبار سنجی مضاعف مقایسه می‌کنیم و به ترکیب بهینه می‌رسیم (سادگالی، ۲۰۱۹).

یادگیری با نظارت[3]

این دسته از الگوریتم‌ها، یک روش عمومی در یادگیری ماشین هستند که در آن‌ها به یک سیستم، مجموعه‌ای از جفت داده‌های ورودی-خروجی ارائه‌شده و سیستم تلاش می‌کند تا تابعی از ورودی به خروجی را فراگیرد. یادگیری تحت نظارت نیازمند تعدادی داده ورودی به‌منظور آموزش سیستم است. بااین‌حال رده‌ای از مسائل وجود دارند که خروجی مناسب که یک سیستم یادگیری تحت نظارت نیازمند آن است، برای آن‌ها موجود نیست. هدف سیستم یادگیری به دست آوردن فرضیه‌ای است که تابع یا رابطه بین ورودی یا خروجی را حدس بزند به این روش یادگیری با نظارت گفته می‌شود.

---

[1] Classification Algorithms
[2] Ensemble Learning
[3] Supervised Machine Learning

تکنیک‌های یادگیری تحت نظارت، همواره نیاز به مجموعه داده‌ای دارند که به‌عنوان «تقلب» و «غیر تقلب»، برچسب‌گذاری شده است تا به‌عنوان داده‌های یادگیری، مورداستفاده قرار گیرند. مزیت اصلی یادگیری با نظارت این است که تمام کلاس‌های خروجی با الگوریتم این رویکرد دست‌کاری می‌شوند و می‌توانند به‌راحتی برای طبقه‌بندی الگوی تشخیصی و رگرسیون داده‌ها مورداستفاده قرار گیرند. یادگیری با نظارت، محدودیت‌هایی دارد. اولین محدودیت، دشواری در جمع‌آوری داده‌های نظارت‌شده یا برچسب‌گذاری شده است. زمانی که حجم زیادی از داده‌های ورودی وجود دارد، هزینه زیادی صرف تهیه داده‌های مذکور می‌شود. حتی در شرایطی ممکن است جمع‌آوری چنین داده‌هایی با برچسب‌گذاری کل داده‌ها، غیرممکن باشد. دومین مسئله این است که گاهی پیدا کردن علائم متمایز، دشوار است و عدم قطعیت و ابهام در نظارت یا برچسب وجود دارد.

شکل ۱. انواع مدل‌های یادگیری با نظارت (هریم تاتسات و همکاران، ۲۰۲۰)

نمونه الگوریتم‌های یادگیری با نظارت عبارت‌اند از:

- الگوریتم‌های طبقه‌بندی[1] از قبیل شبکه‌های عصبی هوشمند،
- طبقه‌بندی K نزدیک‌ترین همسایگی[2]،
- طبقه‌بندی درخت تصمیم[3]،
- طبقه‌بندی ترکیبی[4]،
- الگوریتم طبقه‌بندی جنگل تصادفی[5]،
- و ماشین‌بردار پشتیبان[6]
- الگوریتم‌های رگرسیون[7]،
- الگوریتم رگرسیون لجستیک[8]،
- رگرسیون خطی[9] و رگرسیون ساده،
- رگرسیون چندجمله‌ای[10]،

و همچنین از دیگر الگوریتم‌های یادگیری با نظارت، می‌توان به شبکه‌های عصبی بازگشتی RNN[11] و کانولوشنی CNN[12] اشاره کرد (هریم تاتسات و همکاران، ۲۰۲۰)[13].

مدل نوآورانه تحقیق

یکی از چالش‌های اصلی تحقیقات گذشته تشخیص تقلب، افزایش کارایی الگوریتم‌های یادگیری عمیق برای شناخت الگوهای تقلب است. بررسی تحقیقات مذکور نشان می‌دهد

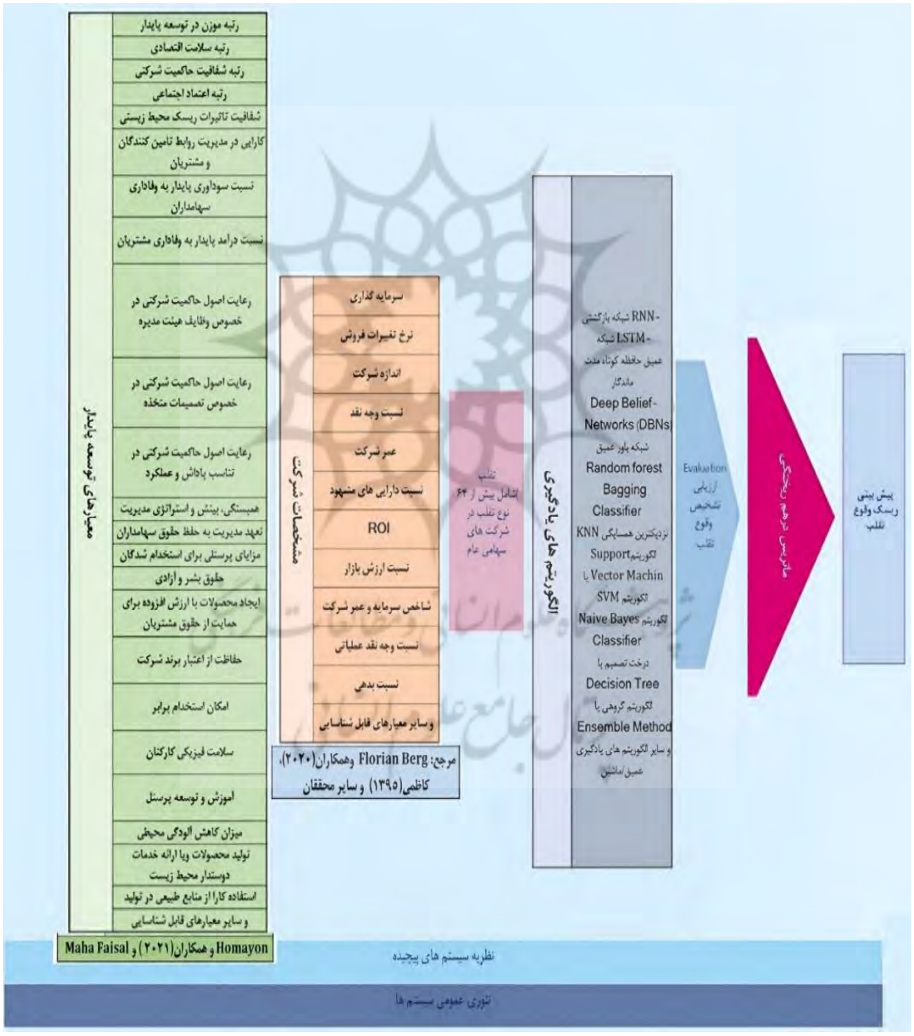
1 Classification
2 K-nearest neighbors algorithm(KNN)
3 Decision Tree
4 Ensemble Classifiers
5 Random Forest
6 Support Vector Machines
7 Regression
8 Logistic Regression
9 Linear Regression
10 Polynomial Regression
11 Recurrent Neural Network
12 Convolutional Neural Network
13 Hariom Tatsat, S. P., Brad Lookabaugh (2020).

کـه محققـان قبلـی داخلـی یـا خـارجی بـه طراحـی سیسـتم هوشـمند شناسـایی تقلـب در شرکت‌های سهامی عام به‌صورت تخصصی و با رویکرد جامع، نپرداخته‌اند (رئیسـی وانـانی و همکاران، ۱۴۰۲). در شرکت‌های سهامی عام، به دلیل گسـتردگی حـوزه تـأثیر فعالیـت و طیـف وسـیع ذینفعـان، اسـتفاده از سیسـتم‌های کـارای هوشـمند شناسـایی تقلـب ضـروری می‌نمایاند. رئیسـی وانـانی و باقریـان کاسـگری و همکـاران (۱۴۰۲) در تحقیـق خـود کلیـه تحقیقـات بـیش از دو دهـه تحقیقـات اخیـر در حـوزه سیسـتم‌های شناسـایی تقلـب مـالی را بررسی نموده‌اند که نتایج آن ضمن تأیید بر تأثیر تحول‌آفرین فناوری‌های یـادگیری ماشـین و یادگیری عمیق، بر ضرورت تغییر مجموعه داده‌های مورداسـتفاده در تحقیقـات شناسـایی تقلب برای افزایش کارایی نتایج تأکید دارد. آن‌ها اسـتدلال نمودنـد کـه ازآنجـایی‌کـه اثـر رویداد تقلب مالی، صرفاً در نتایج گزارش‌های مالی شرکت تأثیرگذار نیست و بـر اعتبـار و عملکرد شرکت و سایر ذینفعان نیز ناگزیر تأثیر می‌گذارد، بنابراین باید در شناسـایی تقلـب متغیرهای غیرمالی نیز در نظر گرفتـه شـود. بـاوجوداین، بررسـی تحقیقـات گذشـته در ایـن حوزه نشان می‌دهد که تحقیقات گذشته صرفاً بر استفاده از داده‌هـای مـالی بـرای پیش‌بینـی تقلب تمرکز کرده‌اند (فلوریان برگ و همکاران، ۲۰۲۰؛ کاظمی و همکاران، ۱۳۹۵ و سایر محققان) و از بررسی قابلیت پیش‌بینی‌گر اطلاعـات غیرمـالی (همـایون و همکـاران، ۲۰۲۱؛ ماها فـای صـال و همکـاران، ۲۰۲۱) غافـل مانده‌انـد. علیـرغم اینکـه ایـن تحقیقـات بـه‌طور روزافزونی بر استفاده از تکنولوژی‌های مبتنی بر هوش مصنوعی و علم داده بـرای پیش‌بینـی تقلب تأکید نموده‌اند، ولی به دلیـل ضـعف در داده‌هـای ورودی و عـدم‌جامعیت متغیرهـا و داده‌ها، هنوز تحقیقات سیستم‌های کشف تقلـب مـالی بـه کمـال خـود نرسـیده‌اند؛ بنـابراین تحقیق جاری به‌طور نوآورانه‌ای قصد دارد تا این ضعف را پوشش داده با به‌کارگیری مـدل ترکیبی با ادغـام داده‌هـای سـنتی مـالی بـا عوامـل غیرمـالی از قبیـل ESG افـزایش کـارایی مدل‌های تشخیص تقلب را بررسی نماید.

به‌عنوان یـک رویکـرد جدیـد در شناسـایی تقلـب مـالی، بررسـی پتانسـیل معیارهـای غیرمالی برای تشخیص تقلب، می‌تواند موجب ایجاد راه‌حل‌های پیچیده‌تر، ولـی کارآمـدتر

و شفاف‌تر گردد؛ بنابراین مدل ترکیب داده‌های مالی با داده‌های غیرمالی به‌عنوان جدیدی تعریف می‌شود که درصورتی‌که موجب افزایش کارایی تشخیص تقلب گردد، دریچه جدیدی بروی محققان آتی بازمی‌نماید. (شکل ۲، مدل مفهومی پیشنهادی برای تحقیقات تشخیص تقلب).

شکل ۲. پیش‌بینی ریسک وقوع تقلب بر اساس الگوریتم‌های یادگیری ماشین و عمیق بر مبنای تئوری سیستم‌های پیچیده (مؤلفین)



پرسش‌های تحقیق

پرسش اصلی که تحقیق جاری به دنبال پاسخ به آن است عبارت است از؛۱) کدام‌یک از الگوریتم‌های یادگیری ماشین و یادگیری عمیق می‌توانند نتیجه اعتبار سنجی بهتری در پیش‌بینی تقلب ارائه دهند. ۲) سیستم هوشمند نظارتی طراحی‌شده مبتنی بر کدام‌یک از عوامل مالی، توسعه پایدار (ESG) یا ترکیب آن‌ها قادر به پیش‌بینی مؤثرتر ریسک تقلب می‌باشند؟ پاسخ گفته خواهد شد.

مجموعه داده‌های تحقیق

داده‌های مورداستفاده در تحقیق جاری شامل داده تقلب در خصوص شرکت‌های خارجی، نمونه آماری شامل بیش از ۶۰۰۰ شرکت سهامی عام شامل دادهای تقلب در ۶۸ کشور در سراسر دنیا می‌باشد. داده‌های موردبررسی در تحقیق شامل داده‌های مالی و غیرمالی می‌باشند. مرجع دسته‌بندی داده‌های تقلب توسط بانک اطلاعاتی Thomson Reuters ESG ۴ASSET است که محققان متعددی از آن برای انجام پژوهش در خصوص شرکت‌ها استفاده نموده‌اند و نتایج تحقیقات ایشان در مجلات معتبر به چاپ رسیده است. در این بانک اطلاعاتی معیارهای رفتاری برای شرکت در نظر گرفته شده است که محققان را قادر به تحلیل شرکت‌های متفاوت در زمان‌ها، مکان‌ها و حاکمیت‌های متفاوت قانونی گرداند.

قلمرو زمانی تحقیق عبارت است یک دوره ۱۰ ساله منتهی به سال ۲۰۲۰ است که توسط بانک اطلاعاتی Thomson Reuters ASSET ۴ESG از متغیرهای توسعه پایدار برای شرکت‌های مختلف ارائه شده است. این مجموعه داده، وقوع تقلب در داده‌های تقلب با مقادیر ۱/۰ مشخص‌شده است که ۱ نشان‌دهنده وجود تقلب و ۰ نشان‌دهنده عدم مبادرت به تقلب است. بسیاری از محققان گذشته نیز از این رویکرد استفاده کرده‌اند. داده‌های مورداستفاده در تحقیق جاری شامل ۱۵۰۹۸ سطر داده تقلب مربوط به ۱۱۱ متغیر مالی و غیرمالی است. در این داده ۵۲۳ سطر مربوط به تقلب هستند که نشان‌دهنده نامتوازن بودن داده‌های تقلب دارد.

متغیرهای تحقیق

این قسمت به تشریح متغیرهای مورداستفاده در این تحقیق شامل متغیر مالی و غیرمالی و همچنین متغیرهای وابسته و مستقل و تشریح آن‌ها می‌پردازد. متغیر وابسته تحقیق، تقلب است که بر اساس سوابق پرونده‌های تخلفات مورد شناسایی توسط نهادهای نظارتی مشخص می‌شود. نمونه متغیرهای مستقل غیرمالی مورداستفاده در این تحقیق عبارت‌اند از رتبه موزن در توسعه پایدار T رتبه سلامت اقتصادی و رتبه شفافیت حاکمیت شرکتی (ESG database ۴Thomson Reuters Asset). همچنین نمونه متغیرهای مستقل مالی مورداستفاده در این تحقیق عبارت‌اند از:

۱. سرمایه‌گذاری
۲. نرخ تغییرات فروش
۳. اندازه شرکت
۴. نسبت وجه نقد
۵. عمر شرکت
۶. نسبت دارایی‌های مشهود
۷. ROI
۸. نسبت ارزش بازار
۹. شاخص سرمایه و عمر شرکت
۱۰. نسبت وجه نقد عملیاتی
۱۱. نسبت بدهی

معیار سرمایه‌گذاری، نسبت کل سرمایه‌گذاری به کل دارایی است. نرخ تغییرات فروش همان نرخ رشد فروش سالانه شرکت است. معیار اندازه شرکت، معادل کل دارایی‌های پایان دوره است. نسبت وجه نقد همان نسبت وجوه نقد و معادل‌های نقدی تقسیم بر دارایی‌های خالص است. همچنین ROI، نسبت سود خالص/میانگین کل دارایی‌ها است.

روش تحقیق

همان‌گونه که ذکرشده در این تحقیق سه دسته مجموعه داده در این تحقیق مورداستفاده قرار گرفته است: ۱) داده‌های غیرمالی یا معیارهای ESG[1]، ۲) داده‌های مالی[2] و ترکیب داده‌های مالی و غیرمالی[3]. داده‌های تقلب بر اساس یک پراکسی احتمال تقلب به‌صورت درصد در اختیار قرار گرفتند که با توجه به یک سطح تعیین‌شده به‌عنوان آستانه تقلب (مانند ۶۷٪) سطح ریسک بالاتر معادل تقلب و پایین‌تر از آن معادل عدم تقلب در نظر گرفته شد؛ بنابراین همان‌گونه که در قسمت قبلی گفته‌شده، داده‌های شامل ستون تقلب در تحقیق شامل مقادیر ۱/۰ هستند. (۱۳۹۹ سطر داده ۱ و بقیه شامل ۰ هستند) شایان توجه است که کلیه متغیرها مستقل عددی (عدد اعشاری و یا عدد صحیح) هستند (ازبگلو و همکاران[4]، ۲۰۲۰ و رد ریگس و همکاران[5]، ۲۰۲۲).

روش تفسیر نتایج تحقیق

برای بررسی نتایج کسب‌شده از معیارهای زیر کمک می‌گیریم:

- معیار دقت یا Accuracy ≠
- معیار صحت یا Precision ≠
- معیار پوشش یا Recall ≠
- معیار F1 score ≠
- معیار MCC ≠
- معیار زمان ≠

در زیر هر یک از معیارهای مورداستفاده تعریف‌شده است. برای تعریف این معیارها ابتدا نیاز است که به تعریف ماتریس اغتشاش پرداخته شود زیرا این معیارها بر اساس ماتریس

---

1 Environmental, Social, and Governance (ESG) metrics
2 Financial data
3 Combination of ESG and Financial data
4 A.M. Ozbayoglu, et al.
5 Rodrigues, V. F., et al.

اغتشاش به دست می‌آیند. در **Error! Reference source not found.** ماتریس اغتشاش نشان داده است.

ماتریس اغتشاش جدولی است که اغلب برای توصیف عملکرد مدل طبقه‌بندی بر روی مجموعه‌ای از داده‌های آزمون که مقادیر واقعی آن‌ها مشخص است، استفاده می‌شود (ازبگلو و همکاران[1]، ۲۰۲۰). ماتریس اغتشاش بر اساس مفاهیم زیر بیان می‌شود:

≠ مثبت درست (TP۲): تراکنش‌هایی که الگوریتم آن‌ها را به‌عنوان تراکنش تقلب پیش‌بینی کرده است و در واقعیت نیز این تراکنش‌ها تقلب هستند.

≠ منفی درست (TN۳): تراکنش‌هایی که الگوریتم آن‌ها را به‌عنوان تراکنش نرمال پیش‌بینی کرده است و در واقعیت نیز این تراکنش‌ها نرمال هستند.

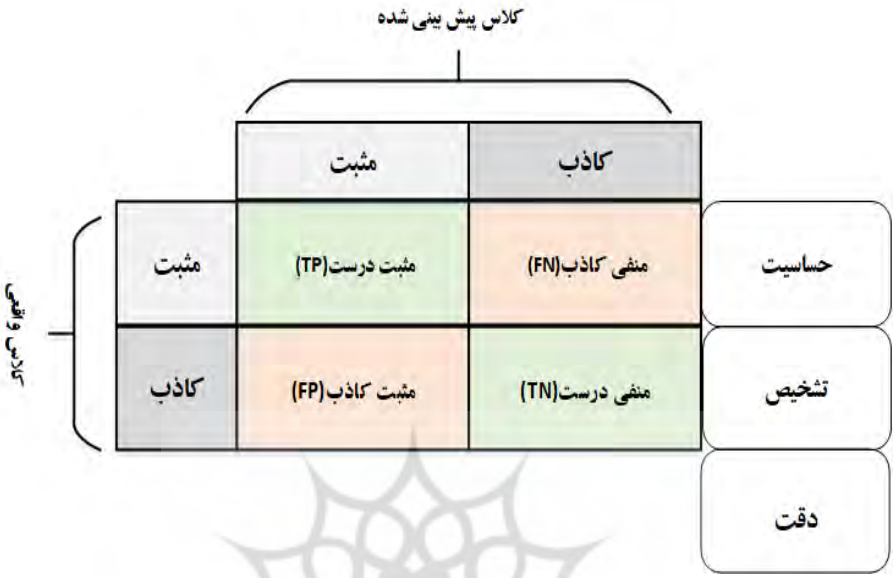
≠ مثبت کاذب (FP۴): تراکنش‌هایی که الگوریتم به‌اشتباه آن‌ها را به‌عنوان تراکنش تقلب پیش‌بینی کرده است ولی در واقعیت این تراکنش‌ها یک تراکنش نرمال می‌باشند (آشتیانی و همکاران[5]، ۲۰۲۲).

≠ منفی کاذب (FN۶): تراکنش‌هایی که الگوریتم به‌اشتباه آن‌ها را به‌عنوان نرمال پیش‌بینی کرده است ولی در واقعیت این تراکنش‌ها یک تراکنش تقلب می‌باشند به‌عبارت‌دیگر برابر است با یک تراکنش تقلب که شناسایی نمی‌شود.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

1 A.M. Ozbayoglu, et al.
2 True Positive
3 True Negative
4 False Positive
5 Ashtiani, M. N.
6 True Negative

شکل ۳. ماتریس اغتشاش و نحوه محاسبه معیارهای اصلی با استفاده از آن



≠ دقت (Accuracy): دقت دسته‌بندی بیشتر بر روی تشخیص درست کلاس مثبت از منفی تمرکز دارد و به تعیین عملکرد کلی طبقه بند می‌پردازد و نشان می‌دهد که چه تعداد از کل موارد آزمایشی به‌درستی توسط طبقه بند، طبقه‌بندی‌شده‌اند. دقت به‌صورت فرمول (۱) بیان می‌شود (احمد راویندر و همکاران[1]، 2023، مندونکا و همکاران[2]، ۲۰۲۴).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (۱)$$

≠ صحت (Precision): این معیار به اندازه‌گیری قابلیت اطمینان طبقه بند می‌پردازد و بیانگر نسبت تقلب‌های واقعی در بین تراکنش‌های هشدار داده‌شده است و به‌صورت فرمول شماره (۲) بیان می‌شود. احمد راویندر و همکاران[3]، 2023 مندونکا و همکاران[4]، ۲۰۲۴.

$$Precision = \frac{TP}{TP + FP} \qquad (۲)$$

1 Ahmed, A. N. S., Ravinder
2 Mendonça, M. O., et al.
3 Ahmed, A. N. S., Ravinder
4 Mendonça, M. O., et al.

≠ پوشش (Recall): پوشش که به‌عنوان حساسیت و نرخ مثبت واقعی نیز شناخته می‌شود، نسبت تقلب‌های واقعی تشخیص داده‌شده توسط مدل را نشان می‌دهد. پوشش به‌صورت فرمول شماره (۳) هست. احمد راویندر و همکاران، 2023 مندونکا و همکاران، ۲۰۲۴).

$$Recall = \frac{TP}{TP + FN} \quad (۳)$$

≠ امتیاز F1 (F1 score): این معیار میانگین وزن‌دار نرخ مثبت واقعی (پوشش) و صحت است. پوشش به‌صورت فرمول شماره (۴) هست. احمد راویندر و همکاران، 2023، 2023

$$F1\ score = 2 . \frac{precision\ .\ recall}{precision + recall} \quad (۴)$$

مندونکا و همکاران، ۲۰۲۴).

• پارامتر دیگری است که برای ارزیابی کارایی الگوریتم‌های یادگیری ماشین از آن استفاده می‌شود. این پارامتر بیان‌گر کیفیت کلاس‌بندی برای یک مجموعه باینری می‌باشد. MCC[1]، سنجه‌ای است که بیان‌گر بستگی مابین مقادیر مشاهده‌شده از کلاس باینری و مقادیر پیش‌بینی‌شده از آن می‌باشد. مقادیر مورد انتظار برای این کمیت در بازه ۱- و ۱ متغیر می‌باشد. مقدار ۱+، نشان‌دهنده پیش‌بینی دقیق و بدون خطای الگوریتم یادگیر از کلاس باینری می‌باشد. مقدار ۰، نشان‌دهنده پیش‌بینی تصادفی الگوریتم یادگیر از کلاس باینری می‌باشد. مقدار ۱-، نشان دهنده عدم تطابق کامل مابین موارد پیش‌بینی شده از کلاس باینری و موارد مشاهده شده از آن می‌باشد. پوشش به‌صورت فرمول شماره (5) هست.(احمد راویندر و همکاران[2]، 2023).

$$MCC = \frac{TP*TN - FP*FN}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(TN+FN)}} \quad (۶)$$

وقتی همه ویژگی‌ها را استفاده می‌کنیم accuracy بالا است ولی در داده‌های غیرمتوازن

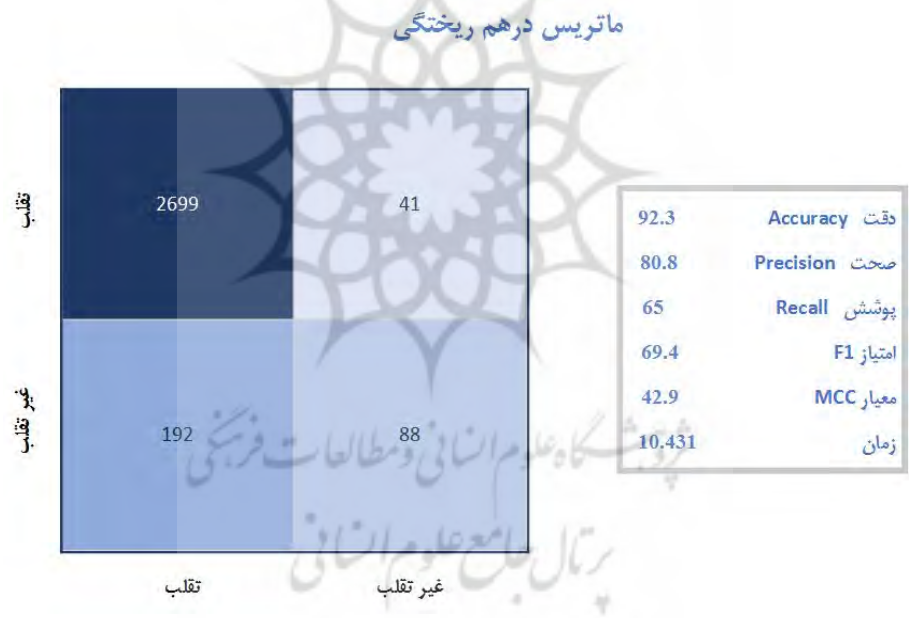
1 Matthews correlation coefficient
2 Ahmed, A. N. S., Ravinder

معیار دقت به‌تنهایی برای ارزیابی مناسب نمی‌باشد و باید معیارهای دیگر مانند F1-score، recall و MCC را نیز بررسی کنیم (ازبیگلو و همکاران، ۲۰۲۲).

نتایج تحقیق

در این قسمت، نتایج تحقیق با الگوریتم‌ها و مجموعه‌های داده‌ای مختلف آمده است. ازنظر معیار F۱ در همه داده‌های موردبررسی الگوریتم جنگل تصادفی نتایج بهتری را نشان می‌دهد که در این میان مجموعه داده ترکیبی (ترکیب داده‌های مالی و غیرمالی) بهترین عملکرد را نشان می‌دهد.

شکل ٤. نتایج ماتریس درهم‌ریختگی برای همه ویژگی‌ها



در استفاده از همه ویژگی‌ها شامل معیارهای مالی، غیرمالی و ترکیب مالی و غیرمالی، الگوریتم Bagging بهترین نتیجه را دارد. شکل ۴، رسم ماتریس درهم‌ریختگی نتایج الگوریتم‌ها را به ازای همهٔ ویژگی‌ها را نشان می‌دهد.

جدول ۱. خلاصه یافته‌های حاصل از تحقیق

معیار / حالت	accuracy	precision	recall	f1 score	MCC
همه ویژگی‌ها	0.923 Bagging	0.906 SVM	0.653 Decision tree	0.694 Bagging	0.429 Bagging
ویژگی‌های غیرمالی ESG	0.909 Extra tree	0.954 Extra tree	0.519 Extra tree	0.513 Extra tree	0.144 Extra tree
ویژگی‌های مالی	0.909 Extra tree	0.788 Voting	0.603 Decision tree	0.6 Decision tree	0.247 Extra tree

مطابق نتایج **Error! Reference source not found.** بالاترین دقت (accuracy) در مجموعه داده‌ای همه ویژگی‌ها (مدل ترکیب داده‌های مالی و غیرمالی) برای الگوریتم طبقه‌بندی Bagging (۰,۹۲۳) مشاهده می‌شود که نشان‌دهنده عملکرد کلی قوی است. در استفاده از ویژگی‌های غیرمالی یاESG بدون بیش نمونه‌برداری، الگوریتم Extra Tree دقت (accuracy) بالای (۰,۹۰۹) و بالاترین precision(۰,۹۵۴) ولی recall پایین‌تری (۰,۵۱۹) را به همراه دارد که نشان‌دهنده اثربخشی متوسط در شناسایی موارد مثبت و منفی واقعی است. با ویژگی‌های مالی، طبقه‌بندی‌کننده درخت اضافی نیز دقت (accuracy) خوبی را نشان می‌دهد (۰,۹۰۹)، اما مدل درخت تصمیم دارای MCC پایین (۰,۲۴۷) است که نشان‌دهنده عملکرد ضعیف‌تر در طبقه‌بندی باینری متعادل است. (**Error! Reference source not found.**) در مقایسه مدل‌ها، طبقه‌بندی‌کننده Extra Tree اغلب ازنظر دقت (accuracy) پیشرو است، اما در معیارهای دیگر مانند MCC ناسازگار است. طبقه‌بندی‌کننده جنگل تصادفی عملکرد متعادل‌تری را در معیارهای مختلف نشان می‌دهد، به‌ویژه با روش‌های انتخاب ویژگی. همچنین روش‌های Decision Tree و Bagging معمولاً معیارهای عملکرد پایین‌تری را نشان می‌دهند که نشان‌دهنده اثربخشی کمتر برای این مجموعه داده خاص است. از منظر کلی ویژگی‌های ESG و مالی هر دو دقت (accuracy) بالایی را ارائه می‌دهند اما در سایر معیارها متفاوت و پایین هستند. باتوجه به نتایج مذکور، داده‌های ترکیبی شامل داده‌های مالی و داده‌های غیرمالی در تمامی معیارها

از دو دسته دیگر شامل داده‌های مالی به‌تنهایی و داده‌های غیرمالی به‌تنهایی، بهتر عمل می‌کنند.

درنهایت می‌توان نتیجه گرفت که اثربخشی مدل‌های یادگیری ماشین، به‌شدت به انتخاب ویژگی‌ها بستگی دارد. درحالی‌که مدل‌های مبتنی بر داده‌های مالی، ازنظر دقت (accuracy) مناسب هستند، در زمینه‌های دیگر مانند precision ، پوشش (recall) یا MCC ضعیف عمل کنند. این تحلیل بر اهمیت در نظر گرفتن و استفاده از داده‌های ترکیبی شامل داده‌های مالی و داده‌های غیرمالی، هنگام شناسایی تقلب مالی، تأکید می‌کند.

محدودیت‌های تحقیق و پیشنهادها برای تحقیقات آتی

نتایج تحقیقات محققین قبلی نشان‌دهنده لزوم طراحی الگوریتم‌هایی است که به تشخیص در لحظه تقلب کمک نماید و این محققین بر لزوم توسعه آن در تحقیقات آینده تأکید نموده‌اند. همچنین جای خالی تحقیقاتی باهدف پیش‌بینی تقلب در مراحل آغازین و ایجاد امکان پیشگیری از آن، در این تحقیقات مشاهده می‌گردد؛ بنابراین باتوجه به اینکه بررسی رویکرد پیشگیری از تقلب در میان تحقیقات قبلی دیده نشده است و نظر به اینکه هدف مهم سازمان‌های نظارتی، پیشگیری از تقلب است، تحقیقات آینده باید برای یافتن راه‌حل برای آن باید توسعه یابند.

علیرغم مزایای ادغام معیارهای غیرمالی با داده‌های مالی در سیستم‌های تشخیص تقلب که در این تحقیق آشکار گردید، استفاده از معیارهای غیرمالی (ESG) و ادغام آن‌ها با داده‌های مالی، خالی از چالش نیست. در درجه اول، در دسترس بودن داده‌های استاندارد و قابل‌اعتماد غیرمالی ESG در حال حاضر محدود است. برخلاف داده‌های مالی که به‌شدت افشاء می‌شوند و حتی افشای اجباری دارند، داده‌های ESG اغلب تحت رژیم‌های افشای داوطلبانه قرار می‌گیرند که منجر به داده‌هایی می‌شود که می‌توانند بعضاً پراکنده و متناقض باشند (Kwamie Dunbar، ۲۰۲۳). علاوه بر این، ماهیت نیمه کیفی برخی از معیارهای ESG، ادغام آن‌ها را در مدل‌های کمی تشخیص تقلب مالی که عمدتاً بر

متغیرهای عددی و طبقه‌ای هستند را پیچیده می‌کند. علاوه بر این، گنجاندن این متغیرهای اضافی می‌تواند منجر به افزایش پیچیدگی مدل شود که نیازمند تکنیک‌های محاسباتی پیچیده‌تر است که به‌طور بالقوه موجب به خطر انداختن تفسیرپذیری مدل می‌شود (Shi، Li و Asal۲۰۲۳). تحقیقات آینده باید راه‌های جدیدی برای مدیریت این چالش‌ها در نظر بگیرند. این راه‌کارها می‌توانند شامل استانداردسازی معیارهای غیرمالی ESG و توسعه الگوریتم‌هایی با قابلیت پردازش انواع داده‌های ناهمگن باشند. (Jiawei Zhang۲۰۲۳، Xunfa Lu۲۰۲۳).

باتوجه به محرمانه بودن داده‌های تقلب و همچنین نادر بودن آن‌ها، مجموعه داده‌های تقلب به‌سختی در دسترس محققان قرار گرفت. در این تحقیق داده‌های تقلب در بیش از ۶۴ کشور موردبررسی قرار گرفت که باتوجه به نادر بودن موارد تقلب، دیتاست موجود بشدت غیرمتوازن بود که تحقیقات آتی می‌توانند اثر راه‌حل‌های رفع عدم توازن را بر پیش‌بینی‌پذیر تقلب بررسی نمایند. همچنین باتوجه به عدم افشای عمومی داده‌های پرونده‌های شرکت‌های متخلف در ایران، در این بررسی، داده‌های کشور خارج از دسترس محققین بوده است. بعلاوه باتوجه به عدم افشای گزارش‌های ESG درکشور، امکان بررسی این تحقیق در خصوص شرکت‌های ایرانی پس از افشای اطلاعات مذکور میسر خواهد بود. همچنین بررسی اثر معکوس به معنای اثر وقوع رویداد تقلب مالی بر متغیرهای غیرمالی نیازمند دسترسی اطلاعات تفضیلی است که مجموعه داده‌های آن در دسترس محققان تحقیق جاری نبوده است.

در این تحقیق بانک اطلاعاتی معیارهای رفتاری برای شرکت در نظر گرفته شده است (بانک اطلاعاتی Thomson Reuters ASSET ESG۴) که محققان را قادر به تحلیل شرکت‌های متفاوت در زمان‌ها، مکان‌ها و حاکمیت‌های متفاوت قانونی گرداند، لکن خصوصیات خاص درونی شرکت‌ها در کشورهای مختلف در تحقیقات آینده می‌تواند موردبررسی قرار گیرد که به دلیل عدم دسترسی به داده‌های دیگر، جزء محدودیت‌های تحقیق فعلی محسوب می‌گردد.

تعارض منافع

این تحقیق فاقد هرگونه تعارض منافع است.

ORCID

Abbas Bagherian Kasgari **https://orcid.org/0000-0003-1630-5207**
Iman Raeesi Vanani **https://orcid.org/0000-0001-8324-9896**
Maghsoud Amiri **https://orcid.org/0000-0002-0650-2584**
Saeid Homayoun **https://orcid.org/0000-0002-2536-0446**

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

منابع

۱. استراوس، انسلم؛ کربین، جولیت. (۱۳۹۴). *مبانی پژوهش کیفی: فنون و مراحل تولید نظریه زمینه‌ای*. ابراهیم افشار، تهران: نشر نی.

۲. باقریان، عباس. (۱۳۸۵). بورس الکترونیک؛ استانداردها و سیستم‌های هوشمند نظارتی. در *اولین همایش سراسری بورس الکترونیک* (ص. ۱-۱۰). قزوین: دانشگاه آزاد اسلامی واحد قزوین. https://www.civilica.com/Paper-SEB01-SEB01_001.html

۳. بولو، قاسم، برزیده، فرخ، الهیاری ابهری، حمید. (۱۳۹۹). الگویی برای ارزیابی خطر تقلب در حسابرسی صورت‌های مالی. *مجله دانش حسابداری*، *۱۱*(۴)، ۲۵-۴۵. doi:10.22103/jak.2020.15880.3254.

۴. بهرامی، آسو، نوروش، ایرج، راد، عباس و محمد ملقرنی، عطاالله. (۱۴۰۰). تقلب در صورت‌های مالی و تکنیک‌های نوین مورداستفاده جهت کشف آن. *مطالعات حسابداری و حسابرسی*، *۱۰*(۳۸)، ۱۰۵-۱۱۸. doi:10.22034/iaas.2021.134547.

۵. ثقفی، دکتر علی و بهار مقدم، دکتر مهدی. (۱۳۸۷). محرک‌های مؤثر بر مدیریت سود. *توسعه و سرمایه*، *۱*(۲)، ۱۰۳-۱۲۵. doi:10.22103/jdc.2008.1894.

۶. جلال جمالی، علی اصغر متقی و احمد محمدی. (۱۴۰۰). مطالعه مقایسه‌ای الگوهای پیش‌بینی ورشکستگی و ارائه الگوی بهینه برای محیط اقتصادی ایران. *توسعه و سرمایه*، *۶*(۲)، ۱۱۱-۱۳۴. doi:10.22103/jdc.2022.18728.1187.

۷. جهانگیر یاری. (۱۳۸۷). بررسی موانع ایجاد سازمان یادگیرنده در مرکز آموزش ایران خودرو. *سال اول، شماره ۴، زمستان*.

۸. حسن حکمی. (۱۳۹۲). رویکرد شناختی؛ تاریخچه، بینایی و توجه، علوم شناختی و هوش ماشین.

۹. دانیل رایف، استفن لیسی و فریدریک جی. فیکو. (۱۳۸۵). *چاپ دوم، تحلیل پیام‌های رسانه‌ای (کاربرد تحلیل محتوای کمی در تحقیق)*. ترجمه مهدخت بروجردی علوی، تهران: سروش.

۱۰. رضائیان، علی. (۱۳۷۷). *تجزیه و تحلیل و طراحی سیستم*. تهران: سمت.

۱۱. رئیسی وانانی، ایمان، باقریان کاسگری، عباس، امیری، مقصود و همایون، سعید. (۱۴۰۲). تحلیل محتوای دو دهه پژوهش‌های سیستم‌های هوشمند نظارتی برای شناسایی تقلب مالی. *توسعه و سرمایه*. doi:10.22103/jdc.2023.22263.1426.

۱۲. سرمد، زهره، بازرگان، عباس و حجازی، الهه. (۱۳۹۰). *روش‌های تحقیق در علوم رفتاری*. تهران: نشر آگه.

۱۳. سنگه، پیتر. (۱۳۷۷). *پنجمین فرمان*. ترجمه حافظ کمال هدایت و محمد روشن. تهران: نشر سازمان مدیریت صنعتی.

۱۴. سیاوش شایان، سید فاطمه نوربخش و همکاران. (۱۳۹۵). تئوری پیچیدگی و رویکرد کلاژیسم در سیستم‌های ژئومورفیک. *مطالعات جغرافیایی مناطق خشک*، *۶*(۲۰)، بهار.

۱۵. شکوهی فرد، سیامک، ابوالحسنی، اصغر و فرهنگ، امیرعلی. (۱۴۰۰). اثرات فساد بر شکنندگی مالی در ایران: رهیافت رگرسیون کوانتایل. *توسعه و سرمایه*، *۶*(۲)، ۹۳-۱۱۰. doi:10.22103/jdc.2021.18460.1169.

۱۶. فاضلی، ن.ا. (۱۳۷۶). آموزش، تحقیق و ترویج (تحلیل محتوای نامه علوم اجتماعی). *نمایه پژوهش*، *۱*(۱)، ۱-۱۶.

۱۷. قائدی، محمدرضا و همکاران. (۱۳۹۵). روش تحلیل محتوا: از کمی‌گرایی تا کیفی‌گرایی. *فصلنامه علمی-پژوهشی «روش‌ها و مدل‌های روانشناختی»*، *۷*(۲۳)، ۵۷-۸۲.

۱۸. قربانیان، امیر، عبدلی، محمدرضا، ولیان، حسن و بودلائی، حسن. (۱۴۰۲). ارزیابی کارکردهای حسابرسی داخلی شهروندِ شرکتی. *توسعه و سرمایه*، *۸*(۱)، ۱۴۳-۱۶۵. doi:10.22103/jdc.2022.19858.1273.

۱۹. قیومی، مهدی. (۲۰۲۲). *یادگیری عمیق در عمل*. دانشگاه کرنل. ISBN 9780367458621.

۲۰. کاظمی، توحید. (۱۳۹۵). انتخاب سبد سهام بهینه از بین سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از الگوریتم مورچگان. پایان‌نامه کارشناسی ارشد، رشته حسابداری، دانشگاه آزاد اسلامی واحد تهران مرکز.

۲۱. کردستانی، دکتر غلامرضا و آشتاب، علی. (۱۳۸۸). پیش‌بینی مدیریت سود بر مبنای تعدیل سود هر سهم. *توسعه و سرمایه*، *۲*(۲)، ۱۴۱-۱۵۸. doi:10.22103/jdc.2009.1912.

۲۲. کمیته تدوین استانداردهای حسابداری. (n.d.). *استانداردهای حسابداری*. سازمان حسابرسی.

۲۳. کودره، دیوید. (۱۳۹۱). *پیشگیری و کشف تقلب به کمک رایانه*. ترجمه امیر پوریانسب و آیدا پوریا نسب. تهران: انتشارات هوشیار ممیز.

۲۴. گل ایری، تحفه، کرد، صفوره و خاری، رضا. (۱۳۹۹). مسئولیت حسابرس و کشف تقلب. *شباک*، *۶*(۶)، ۱۳۵-۱۴۲. SID. https://sid.ir/paper/525408/fa.

۲۵. گلدمن، پیتر و کافمن، هیلتون. (۱۳۹۵). *راهنمای کاربردی ریسک‌های تقلب و کنترل‌های ضد تقلب*. ترجمه امیر پوریانسب و محسن قاسمی. تهران: انتشارات حساب افزار ایرانیان.

۲۶. مدرس، دکتر احمد و افلاطونی، عباس. (۱۳۸۸). مدیریت سود در شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران. *توسعه و سرمایه*، *۲*(۲)، ۵۱-۷۲. doi:10.22103/jdc.2009.190.

۲۷. محمدرضا قدوسی و همکاران. (۱۳۹۸). ارائه چارچوب حکمرانی دانش در شبکه‌های اجتماعی. *دو فصلنامه علمی مدیریت اطلاعات*، *۵*(۱).

۲۸. مشایخی، بیتا و همکاران. (۱۳۹۲). تدوین مدل کیفیت حسابرسی. *فصلنامه بورس و اوراق بهادار*، *۶*(۲۳)، ۱۰۳-۱۳۷.

۲۹. منصور صادقی ما لامیری. (۱۳۹۳). رفتارهای سه‌گانه بخل، میانه‌روی و اسراف از دیدگاه سیستمی. *دو فصلنامه علمی پژوهشی مدیریت اسلامی*، *۲۲*(۱)، ۱۴۱-۱۶۶.

۳۰. میریعقوب سیدرضایی. (۱۳۹۵). تنوع‌بخشی و پیچیده‌سازی سیستم اداری به منزله الگویی برای اداره بهینه پایتخت. *فصلنامه تحقیقات جغرافیایی*، *۳۱*(۴)، زمستان.

۳۱. مؤمنی راد، اکبر؛ علی آبادی، خدیجه؛ فردانش، هاشم؛ و مزینی، ناصر. (۱۳۹۲). تحلیل محتوای کیفی در آیین پژوهش: ماهیت، مراحل و اعتبار نتایج. *اندازه گیری تربیتی*، *۴*(۱۴).

۳۲. مهرانی، دکتر ساسان، گنجی، حمیدرضا، تحریری، آرش و عسکری، محمدرضا. (۱۳۸۸). ارزیابی رتبه‌بندی شرکت‌ها بر اساس اطلاعات حسابداری و غیرحسابداری و مقایسه آن با رتبه‌بندی شرکت‌ها در بورس اوراق بهادار تهران. *توسعه و سرمایه*، *۲*(۱)، ۷-۳۲. doi:10.22103/jdc.2009.1899.

۳۳. نمازی، دکتر محمد و ناظمی، امین. (۱۳۸۷). مروری بر پژوهش‌های حسابداری انجام شده در بورس اوراق بهادار تهران. *توسعه و سرمایه*، *۱*(۲)، ۹-۴۸. doi:10.22103/jdc.2008.1891.

۳۴. نویدی عباس پور، ابراهیم و واعظی، جمیله. (۱۴۰۱). عوامل تعیین‌کننده توانایی حسابرسی برای کشف تقلب: عوامل داخلی و خارجی. مطالعه موردی: حسابرسان مؤسسات حسابرسی عضو جامعه حسابداران رسمی ایران. *دوازدهمین کنفرانس بین‌المللی پژوهش‌های نوین در مدیریت، اقتصاد، حسابداری و بانکداری*. https://civilica.com/doc/1566301.

۳۵. هولستی، آل. آر. (۱۳۷۳). *تحلیل محتوا در علوم اجتماعی و انسانی*. ترجمه نادر سالارزاده امیری. تهران: انتشارات دانشگاه علامه طباطبایی (تاریخ ترجمه اثر به زبان اصلی ۱۹۶۹).

References

36. Fredrick. (2018). The extents of neuroscience and neuropsychology in the study of artificial intelligence. *IRA-International Journal of Applied Sciences*, *13*(3).
37. Á. Alonso, et al. (2019). An identity framework for providing access to FIWARE OAuth 2.0-based services according to the eIDAS European Regulation. *IEEE Access*, *7*, 88435-88449. https://doi.org/10.1109/ACCESS.2019.2926556
A. M. Ozbayoglu, M. U. Gudelek, & O. B. Sezer. (2020). Deep learning for financial applications: A survey. *Applied Soft Computing Journal*.
38. Abakarim, Y., Lahby, M., & Attioui, A. (2018, October). An efficient real-time model for credit card fraud detection based on deep learning. In *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications* (pp. 1-7).
39. Abdulghani, A. Q., et al. (2021). Credit card fraud detection system using machine learning algorithms and fuzzy membership.
40. Abhimanyu Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams, & P. Beling. (2018). Deep learning detecting fraud in credit card transactions. In *2018 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE.
41. Abhishek Nair, D. Reckien, & M. F. A. M. van Maarseveen. (2019). A generalised fuzzy cognitive mapping approach for modelling complex systems. *Applied Soft Computing*, *84*, 105754. https://doi.org/10.1016/j.asoc.2019.105754
42. Accounting Standards Drafting Committee. (2018). *Accounting standards*. Auditing Organization [In Persian].
43. Ahmadpour Kasgari, A., Bagherian Kasgari, R., & Bagherian Kasgari, A. (2009). The feasibility of eXtensible Business Reporting Language for

listed companies on Tehran Stock Exchange. *Quarterly Journal of Securities Exchange, 2*(6).

44. Ahmed, A. N. S., & Ravinder. (2023). A survey on detection of fraudulent credit card transactions using machine learning algorithms. *3rd International Conference on Intelligent Communication and Computational Techniques, ICCT 2023*. Jaipur.
45. Aihua Li, J. W., & Zhidong Liu. (2017). Market manipulation detection based on classification methods. *Elsevier Procedia Computer Science, 122*, 788-795.
46. Alan Hevner, S. C. (2010). Design research in information systems: Theory and practice. *Springer Science & Business Media*.
47. Allan Inocencio de Souza Costa & Luis Silva. (2016). Sequence classification of the limit order book using recurrent neural networks.
48. Amr Kayid. (2020). The role of artificial intelligence in future technology. *Artificial General Intelligence*. https://doi.org/10.13140/RG.2.2.12799.23201
49. Anuj Sharma & Prabin Kumar Panigrahi. (2012). A review of financial accounting fraud detection based on data mining techniques. *International Journal of Computer Applications, 39*(1), 37–47.
50. Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Massachusetts: Addison-Wesley Publishing Company.
51. Arwin Datumaya Wahyudi Sumari & Adang Suwandi Ahmad. (2018). Intelligent system, cognitive artificial intelligence: Concept and applications for humankind. *Kasetsart University*, IEEE. ISBN: 978-1-78923-607-1.
52. Ashtiani, M. N., & Raahemi, B. (2022). Intelligent fraud detection in financial statements using machine learning and data mining: A systematic literature review. *IEEE Access, 10*, 72504-72525.
53. Aslam, F., Hunjra, A. I., Ftiti, Z., Louhichi, W., & Shams, T. (2022). Insurance fraud detection: Evidence from artificial intelligence and machine learning. *Research in International Business and Finance, 62*. https://doi.org/10.1016/j.ribaf.2022.101744
54. Athanasios Drigas & Maria Karyotaki. (2019). Attention and its role: Theories and models. *International Journal of Emerging Technologies in Learning (iJET)*, June.
55. Bagherian Kasgari, A. (2007). Electronic stock exchange, standards and intelligent surveillance networks. *Bourse Journal, Securities and Exchange Organization of Iran Monthly Journal, 1*(61), 40-61.
56. Bagherian Kasgari, A. (2009). Company performance analysis using XBRL, intelligent visual fraud, and XBRL surveillance and analysis

desktop. *Prize Winners of Ninth Global XBRL Academic Competition 2008-2009 Results*, Bryant University, USA.

57. Bagherian Kasgari, A. (2009). Intelligent visual fraud: Supporting fraud detection efforts of exchange regulators using visual modeling. *1st International Conference on Digital Forensics and Cyber Crime*.
58. Bagherian Kasgari, A. (2010). Discussant speech on the paper "Enhancing the regulation of Chinese securities companies, an analysis of the function of information disclosure." *8th International Workshop for Young Scholars (WISH)*, organized by European Law Journal and ESSCA, Angers, Loire Valley, France.
59. Bagherian Kasgari, A. (2012). The next generation of financial mobile networks. *The 24th XBRL International Conference*.
60. Bagherian Kasgari, A. (2013). XBRL decision cloud. *US XBRL Challenge*.
61. Bagherian Kasgari, A., & Sheykhi, K. (2014). Stock price fluctuations and corporate forecasts disclosure: Evidence from an emerging market. *International Journal of Management & Information Technology*, *9*(1), 1481-1495.
62. Bagherian Kasgari, A. (2014). The relationship between independent auditors and receiving financial facilities. *International Journal of Management Science*, *1*(4), 58-64.
63. Kasgari, A. B. (2016). A new prototype for intelligent visual fraud detection in agent-based auditing framework. *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)*, *4*(2), 070-076.
64. Bagherian Kasgari, A. (2016). Is it possible to proactively detect frauds with zero cost? *Journal of Engineering and Innovative Technology (IJEIT)*, *5*(7).
65. Bagherian Kasgari, A. (2018). Anti-money laundering compliance recommendation and risk level assessment. *Anti-Money Laundering Journal of Securities and Exchange Organization*, *2*(1), 1-33.
66. Bagherian Kasgari, A., Mousavi, H., & Homayoun, S. (2020). Extensible visual business intelligence for analyzing XBRL big data on blockchain. In *Eurofiling Online Conferences 2020 TECTONIC SHIFT, Eurofiling Innovation Day*.
67. Bagherian Kasgari, A., Golchin Kharazi, S., & M. T. T. (2019). Price manipulation fraud detection by intelligent visual fraud surveillance system. *IEEE*, 1646-1651.
68. Bagherian Kasgari, A. (2007). Intelligent pyramid continues auditing agent framework for softening SOX burden on small business. *X*

International Congress on Innovations in Teaching Accounting and Business Research (IAAER).

69. Bagherian Kasgari, A. (2007). Software services in OTC markets. *Bourse Journal, Securities and Exchange Organization of Iran Monthly Journal*.
70. Bagherian Kasgari, A., Ranjbarzadeh, R., Caputo, A., Baseri Saadi, S., & Bendechache, M. (2023). Brain tumor segmentation based on Zernike moments, enhanced ant lion optimization, and convolutional neural network in MRI images. In *Metaheuristics and Optimization in Computer and Electrical Engineering: Volume 2: Hybrid and Improved Algorithms* (pp. 345-366). Cham: Springer International Publishing.
71. Bahrami, A., Norush, I., Rad, A., & Mohammad Mulqarni, A. (2021). Fraud in financial statements and new techniques used to detect it. *Accounting and Auditing Studies, 10*(38), 105-118. doi:10.22034/iaas.2021.134547 [In Persian].
72. Bakhshi, K., et al. (2021). Fraud detection system in online ride-hailing services.
73. Barbara Kitchenham, S. C. (2007). Guidelines for performing systematic literature reviews in software engineering. *Technical report, EBSE Technical Report EBSE-2007-01*, 1-57.
74. Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal, 55*(5), 24-36.
75. Beneish, M. D., & Vorst, P. (2020). The cost of fraud prediction errors. *Kelley School of Business Research Paper No. 2020-55*. Available at SSRN: https://ssrn.com/abstract=3529662 or http://dx.doi.org/10.2139/ssrn.3529662
76. Bockel-Rickermann, C., Verdonck, T., & Verbeke, W. (2023). Fraud analytics: A decade of research: Organizing challenges and solutions in the field. *Expert Systems with Applications, 232*, 120605. https://doi.org/10.1016/j.eswa.2023.120605
77. Boluo, G., Barzideh, F., & Alahyari Abhari, H. (2019). A model for assessing the risk of fraud in the audit of financial statements. *Journal of Accounting Knowledge, 11*(4), 25-45. doi:10.22103/jak.2020.15880.3254 [In Persian].
78. Brown, T., Majors, T. M., & Peecher, M. E. (2020). Evidence on how different interventions affect juror assessment of auditor legal culpability and responsibility for damages after auditor failure to detect fraud. *Accounting, Organizations and Society, 87*, 101172. https://doi.org/10.1016/j.aos.2020.101172

79. Calamaro, N., et al. (2021). An energy-fraud detection system capable of distinguishing frauds from other energy flow anomalies in an urban environment. *Sustainability (Switzerland), 13*(19).
80. Chen, Y., & Wu, Z. (2023). Financial fraud detection of listed companies in China: A machine learning approach. *Sustainability, 15*(1), 105. https://doi.org/10.3390/su15010105
81. Dang, T. K., et al. (2021). Machine learning based on resampling approaches and deep reinforcement learning for credit card fraud detection systems. *Applied Sciences (Switzerland), 11*(21).
82. Daniel Rife, S. Lacy, & F. J. FICO. (2006). Analysis of media messages (application of quantitative content analysis in research). Translated by Mahdekht Borujerdi Alavi. Tehran: Soroush [In Persian].
83. Fazli, N. A. (2008). Education, research and promotion (content analysis of social sciences letters). *Research Profile, 1*(1).
84. Ghorbin, A., Abdoli, M. R., Valian, H., & Bodlai, H. (2023). Evaluation of the internal audit functions of the corporate citizen. *Development and Capital, 8*(1), 143-165. doi:10.22103/jdc.2022.19858.1273 [In Persian].
85. Gol Iri, T., Kurd, S., & Khari, R. (2019). Auditor's responsibility and detection of fraud. *Shabak, 6*(6 (series 57)), 135-142. SID. https://sid.ir/paper/525408/fa [In Persian].
86. Goldman, P., & Kaufman, H. (2015). *Practical guide to fraud risks and anti-fraud controls*. Translated by Amir Pouria Nasab and Mohsen Ghasemi. Tehran: Iranian Accounting Publishing House [In Persian].
87. Gupta, P. (2023). Leveraging machine learning and artificial intelligence for fraud prevention. *SSRG International Journal of Computer Science and Engineering, 10*(5), 47-52. https://doi.org/10.14445/23488387/IJCSE-V10I5P107
88. Habibpour, M., Gharoun, H., Mehdipour, M. R., Tajally, A. R., Asgharnezhad, H., Shamsi, A., Khosravi, A., & Nahavandi, S. (2023). Uncertainty-aware credit card fraud detection using deep learning. *Engineering Applications of Artificial Intelligence, 123*, 106248. https://doi.org/10.1016/j.engappai.2023.106248
89. Hariom Tatsat, S. P., & Lookabaugh, B. (2020). *Machine learning and data science blueprints for finance*. O'Reilly Media, Inc.
90. Hassan Hekami. (2012). Cognitive approach; history, vision and attention, cognitive science and machine intelligence [In Persian].
91. Hilal, W., et al. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications, 193*, 34.

92. Holsti, L. R. (1994). *Content analysis in social and human sciences*. Translated by Nader Salarzadeh Amiri. Tehran: Allameh Tabatabai University Publications (date of translation of the original work in 1969) [In Persian].
93. Iris H-Y Chiu. (2023). An institutional account of responsiveness in financial regulation: Examining the fallacy and limits of 'same activity, same risks, same rules' as the answer to financial innovation and regulatory arbitrage. *Computer Law & Security Review, 51*. https://doi.org/10.1016/j.clsr.2023.105868
94. Jahangir Yari. (2017). Investigation of obstacles to creating a learning organization in Iran Khodro Training Center. *First year, number 4* [In Persian].
95. Jalal Jamali, A. Asghar Motaghi, & A. Mohammadi. (2021). Comparative study of bankruptcy prediction models and providing an optimal model for Iran's economic environment. *Development and Capital, 6*(2), 111-134. doi:10.22103/jdc.2022.18728.1187 [In Persian].
96. Josephine Isabella, S., et al. (2020). An efficient study of fraud detection system using ML techniques. *Lecture Notes in Networks and Systems, 118*, 59-67.
97. Kalbande, D., et al. (2021). A fraud detection system using machine learning. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. https://doi.org/10.1109/ICCCNT51525.2021.9580102
98. Kanika & J. Singla. (2019). Online banking fraud detection system: A review. *International Journal of Advanced Trends in Computer Science and Engineering, 8*(3), 959-962.
99. Kanika & J. Singla. (2020). A survey of deep learning based online transactions fraud detection systems. In *2020 International Conference on Intelligent Engineering and Management (ICIEM)*. IEEE. https://doi.org/10.1109/ICIEM48762.2020.9160200
100. Kanika & J. Singla. (2022). A novel framework for online transaction fraud detection system based on deep neural network. *Journal of Intelligent and Fuzzy Systems, 43*(1), 927-937.
101. Kanika, Singla, J., & Nikita. (2021). Comparing ROC curve based thresholding methods in online transactions fraud detection system using deep learning. In *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. https://doi.org/10.1109/icccis51004.2021.9397167
102. Kazemi, T. (2015). Choosing the optimal stock portfolio from among the stocks of companies admitted to the Tehran Stock Exchange

using the ant algorithm. *Master's thesis, accounting field, Islamic Azad University, Tehran branch* [In Persian].

103. Khattri, V., et al. (2020). Plastic card circumvention: An infirmity of authenticity and authorization. *Journal of Financial Crime, 27*(3), 959-975.
104. Kodreh, D. (2013). Prevention and detection of fraud with the help of computer. Translated by Amir Pourianesb and Aida Pouria Nasb. Tehran: Hoshiar Momis Publications [In Persian].
105. Krippendorff, K. (2022). *Content analysis: An introduction to its methodology*. Sage Publications. https://doi.org/10.4135/9781071878781
106. Kurdestani, G. R., & Ashtab, A. (2009). Profit management forecast based on profit per share adjustment. *Development and Capital, 2*(2), 141-158. doi:10.22103/jdc.2009.1912 [In Persian].
107. Lee, C. (2022). Deep learning-based detection of tax frauds: An application to property acquisition tax. *Data Technologies and Applications, 56*(3), 329-341. https://doi.org/10.1108/DTA-06-2021-0134
108. Liang, Y., et al. (2021). The application of synthetic data generation and data-driven modelling in the development of a fraud detection system for fuel bunkering. *Measurement: Sensors, 18*.
109. Mansour Sadeghi Ma Lamiri. (2013). The three behaviors of miserliness, moderation, and extravagance from a systemic point of view. *Two Scientific Research Quarterly Journals of Islamic Management, 22*(1), 141-166 [In Persian].
110. Mashayikhi, B., et al. (2013). Compilation of audit quality model. *Stock Exchange and Securities Quarterly, 6*(23), 103-137 [In Persian].
111. Mehrani, S., Ganji, H., Tahriri, A., & Askari, M. R. (2009). Evaluating the ranking of companies based on accounting and non-accounting information and comparing it with the ranking of companies in the Tehran Stock Exchange. *Development and Capital, 2*(1), 7-32. doi:10.22103/jdc.2009.1899 [In Persian].
112. Mendonça, M. O., Netto, S. L., Diniz, P. S., & Theodoridis, S. (2024). Machine learning: Review and trends. *Signal Processing and Machine Learning Theory*, 869-959.
113. Miriyagoub Seyed Rezaei. (2015). Diversification and complexity of the administrative system as a model for the optimal administration of the capital. *Geographical Research Quarterly, 31*(4) [In Persian].
114. Modares, A., & Platoni, A. (2010). Profit management in companies listed on the Tehran Stock Exchange. *Development and Capital, 2*(2), 51-72. doi:10.22103/jdc.2009.190 [In Persian].

115. Mohammad Reza Qudosi et al. (2018). Presentation of knowledge governance framework in social networks. *Two Scientific Quarterly Journals of Information Management, 5*(1) [In Persian].
116. Momeni Rad, A., Ali Abadi, K., Fardanesh, H., & Mazini, N. (2013). Qualitative content analysis in research: Nature, stages, and validity of results. *Educational Measurement, 4*(14).
117. Moreira, M. Â. L., Rocha Junior, C. de S., Lima Silva, D. F., Castro Junior, M. A. P., Araújo Costa, I. P., Gomes, C. F. S., & Santos, M. (2022). Exploratory analysis and implementation of machine learning techniques for predictive assessment of fraud in banking systems. *Procedia Computer Science, 214*, 117-124. https://doi.org/10.1016/j.procs.2022.11.156
118. Namazi, M., & Nazimi, A. (2008). An overview of the accounting research done in Tehran Stock Exchange. *Development and Capital, 1*(2), 9-48. doi:10.22103/jdc.2008.1891 [In Persian].
119. Navidi Abbaspour, E., & Vaezi, J. (2022). Determinants of auditing ability to detect fraud: Internal and external factors. Case study: Auditors of auditing institutions that are members of the Certified Public Accountants Society of Iran. *12th International Conference on Modern Researches in Management, Economics, Accounting and Banking*. https://civilica.com/doc/1566301 [In Persian].
120. Nazeer, I., Prasad, K. D. V., Bahadur, P., Bapat, V., & Kurian, M. J. (2023). Synchronization of AI and deep learning for credit card fraud detection. *International Journal of Intelligent Systems and Applications in Engineering, 11*(5s), 52-59.
121. Pallavi, C., et al. (2021). A relative investigation of various algorithms for online financial fraud detection techniques. *Advances in Parallel Computing, 39*, 22-32.
122. Qaidi, M. R., et al. (2016). Content analysis method, from quantitative to qualitative. *Scientific-Research Quarterly "Psychological Methods and Models", 7*(23), 57-82 [In Persian].
123. Quintin-John Smith, R., & Valverde, R. (2021). A perceptron-based neural network data analytics architecture for the detection of fraud in credit card transactions in financial legacy systems. *WSEAS Transactions on Systems and Control*. https://doi.org/10.37394/23203.2021.16.31
124. Rahman, M. Y. T. M. H. N. I. K. H. A. (2022). Fraud detection during financial transactions using machine learning and deep learning techniques. *2022 IEEE International Conference on Communications, Computing, Cybersecurity and Informatics, CCCI 2022*.

125. Raneem Khaled, H. A., & Ehab K. A. Mohamed. (2021). The sustainable development goals and corporate sustainability performance: Mapping, extent, and determinants. *Journal of Cleaner Production, 311*.
126. Rezaian, A. (1999). *System analysis and design*. Tehran: Samt [In Persian].
127. Rodrigues, V. F., et al. (2022). Fraud detection and prevention in e-commerce: A systematic literature review. *Electronic Commerce Research and Applications, 101207*.
128. Sadgali, I., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Procedia Computer Science, 148*, 45-54.
129. Sahayasakila, V., et al. (2019). Credit card fraud detection system using SMOTE technique and whale optimization algorithm. *International Journal of Engineering and Advanced Technology, 8*(5), 190-192.
130. Salwa Al Balawi, N. A. (2023). Credit-card fraud detection system using neural networks. *International Arab Journal of Information Technology, 20*(2).
131. Sange, P. (1999). *The Fifth Commandment*. Translated by Hafez Kamal Hedayat, Mohammad Roshan. Tehran: Publication of Industrial Management Organization [In Persian].
132. Sarmad, Z., Bazargan, A., & Hijazi, E. (2018). *Research methods in behavioral sciences*. Tehran: Age Publishing [In Persian].
133. Seera, M., et al. (2021). An intelligent payment card fraud detection system. *Annals of Operations Research*.
134. Sh., Siavash, Nourbakhsh, S. F., et al. (2015). Complexity theory and claggism approach in geomorphic systems. *Geographical Studies of Arid Regions, 6*(23) [In Persian].
135. Shokohi Fard, S., Abolhasani, A., & Farhang, A. A. (2021). Effects of corruption on financial fragility in Iran: Quantile regression approach. *Development and Capital, 6*(2), 93-110. doi:10.22103/jdc.2021.18460.1169 [In Persian].
136. Strauss, A., & Corbin, J. (2014). *Fundamentals of qualitative research: Techniques and stages of field theory production*. Ebrahim Afshar. Tehran: Nei Publishing [In Persian].
137. Sun, G., Li, T., Ai, Y., & Li, Q. (2023). Digital finance and corporate financial fraud. *International Review of Financial Analysis, 87*. https://doi.org/10.1016/j.irfa.2023.102566
138. Thaghafi, A., & Bahar Moghadam, M. (2009). Effective drivers on profit management. *Development and Capital, 1*(2), 103-125. doi:10.22103/jdc.2008.1894 [In Persian].

139. Tyagi, N. K., & Goyal, M. (2022). Two-tier model of exports drawback fraud detection system using intuitionistic fuzzy game theory. *Intelligent Decision Technologies, 16*(2), 299-313.
140. Vanini, P., Rossi, S., Zvizdic, E., et al. (2023). Online payment fraud: From anomaly detection to risk management. *Financ Innov, 9*, 66. https://doi.org/10.1186/s40854-023-00470-w
141. Zhang, G., et al. (2022). EFraudCom: An e-commerce fraud detection system via competitive graph neural networks. *ACM Transactions on Information Systems, 40*(3).
142. Zhu, X., Ao, X., Qin, Z., Chang, Y., Liu, Y., He, Q., & Li, J. (2021). Intelligent financial fraud detection practices in the post-pandemic era. *Volume 2, Issue 4*. https://doi.org/10.1016/j.xinn.2021.100176

References (In Persian)

1. Accounting Standards Committee. (n.d.). *Accounting standards*. Tehran: Organization for Accounting and Auditing. (In Persian)
2. Baqarian, A. (2006). Electronic stock exchange; standards and intelligent supervisory systems. In *Proceedings of the First National Electronic Stock Exchange Conference* (pp. 1-10). Qazvin: Islamic Azad University, Qazvin Branch. (In Persian)
3. Bolu, G., Barzeedeh, F., & Aleyari Abadi, H. (2020). A pattern for evaluating fraud risk in auditing financial statements. *Journal of Accounting Knowledge, 11*(4), 25-45. https://doi.org/10.22103/jak.2020.15880.3254 (In Persian)
4. Bahrami, A., Nouroushe, I., Rad, A., & Molkareni, A. (2021). Financial statement fraud and new techniques used to detect it. *Accounting and Auditing Studies, 10*(38), 105-118. https://doi.org/10.22034/iaas.2021.134547 (In Persian)
5. Fazli, N. A. (1997). Education, research and promotion (content analysis of social sciences letters). *Index of Research, 1*(1), 1-16. (In Persian)
6. Ghaedi, M., & Colleagues. (2016). Content analysis method, from quantitative to qualitative. *Scientific-Research Quarterly "Methods and Models of Psychology, 7*(23), 57-82. (In Persian)
7. Ghorbanian, A., Abdoli, M., Velian, H., & Bodlaei, H. (2023). Evaluation of the performance of corporate citizen internal audit functions. *Development and Capital, 8*(1), 143-165. https://doi.org/10.22103/jdc.2022.19858.1273 (In Persian)

8. Ghodosi, M. R., & Colleagues. (2019). Presenting a knowledge governance framework in social networks. *Journal of Information Management, 5*(1). (In Persian)
9. Goldman, P., & Kaufman, H. (2016). *A practical guide to fraud risks and anti-fraud controls*. Translated by Amir PourianSab and Mohsen Ghasemi. Tehran: Hesab Afzar Iraniyan Publications. (In Persian)
10. Hakimi, H. (2013). *Cognitive approach: History, vision, and attention in cognitive sciences and machine intelligence*. Tehran: Sam Publishing. (In Persian)
11. Holsti, L. R. (1975). *Content analysis for the social sciences and humanities*. Translated by Nader Salarzadeh Amiri. Tehran: Allameh Tabatabai University Press (Translation date of the original work in 1969). (In Persian)
12. Jalali Jamali, A., Motaqi, A., & Mohammadi, A. (2022). Comparative study of bankruptcy prediction models and presentation of an optimal model for the Iranian economic environment. *Development and Capital*, *6*(2), 111-134. https://doi.org/10.22103/jdc.2022.18728.1187 (In Persian)
13. Kazemi, T. (2016). Optimal portfolio selection from among the stocks of companies accepted in Tehran Stock Exchange using the ant colony algorithm. *Master's thesis, Department of Accounting, Islamic Azad University, Tehran Central Branch*. (In Persian)
14. Kodreh, D. (2012). *Computer-assisted fraud prevention and detection*. Translated by Amir PourianSab and Aida PourianSab. Tehran: Hoshiar Momis Publications. (In Persian)
15. Kordestani, G. R., & Ashtyab, A. (2009). Predicting earnings management based on earnings per share adjustment. *Development and Capital, 2*(2), 141-158. https://doi.org/10.22103/jdc.2009.1912 (In Persian)
16. Mehrbani, S., Ganji, H., Taheri, A., & Asgari, M. R. (2009). Evaluation of company ranking based on accounting and non-accounting information and comparison with company ranking in Tehran Stock Exchange. *Development and Capital*, *2*(1), 7-32. https://doi.org/10.22103/jdc.2009.1899 (In Persian)
17. Meshaiki, B., & Colleagues. (2013). Developing an auditing quality model. *Journal of Securities and Stock Exchange, 6*(23), 103-137. (In Persian)
18. Miri Yaqoob Seyedrezaee. (2016). Diversification and complication of the administrative system as a model for optimal management of the capital. *Journal of Geography Research, 31*(4), 123-142. (In Persian)

19. Momeni Rad, A., Ali Abadi, K., Fardanesh, H., & Mazini, N. (2013). Qualitative content analysis in research: Nature, stages, and validity of results. *Educational Measurement, 4*(14). (In Persian)
20. Modares, A., & Aflatoni, A. (2009). Earnings management in companies accepted in Tehran Stock Exchange. *Development and Capital, 2*(2), 51-72. https://doi.org/10.22103/jdc.2009.190 (In Persian)
21. Namazi, M., & Nazimi, A. (2008). A review of accounting research conducted on the Tehran Stock Exchange. *Development and Capital, 1*(2), 9-48. https://doi.org/10.22103/jdc.2008.1891 (In Persian)
22. Novidi Abbaspour, E., & Vaezi, J. (2022). Factors determining the ability of auditing to detect fraud: Internal and external factors. A case study: Auditors of auditing institutions that are members of the Society of Iranian Certified Public Accountants. *12th International Conference on Novel Research in Management, Economics, Accounting and Banking*. https://civilica.com/doc/1566301. (In Persian)
23. Raif, D., Liss, S., & Feick, F. G. (2006). *Analyzing media messages (using quantitative content analysis in research)* (2nd ed.). Translated by Mahdokht Boroujerdi Alavi. Tehran: Soroush. (In Persian)
24. Raeesi Vanani, I., Bagherian Kasgari, A., Amiri, M., & Homayoun, S. (2023). A comprehensive analysis of two decades in intelligent surveillance systems for financial fraud detection research. *Journal of Development and Capital*. https://doi.org/10.22103/jdc.2023.22263.1426 (In Persian)
25. Rezaeiyan, A. (1998). *Analysis and design of systems: Methods and techniques of system analysis and design*. Tehran: Samt. (In Persian)
26. Sarmad, Z., Bazargan, A., & Hejazi, A. (2011). *Research methods in behavioral sciences*. Tehran: Agah Publishing. (In Persian)
27. Senge, P. (1997). *The fifth discipline*. Translated by Hafez Kamal Hedayat, Mohammad Roshan. Tehran: Industrial Management Organization Publications. (In Persian)
28. Shayan, S., Nourbakhsh, S. F., & Colleagues. (2016). Complexity theory and collage approach in geomorphic systems. *Journal of Geography of Arid Regions, 6*(20), 1-14. (In Persian)
29. Shokouhi Fard, S., Abolhasani, A., & Farhang, A. (2021). The effects of corruption on financial fragility in Iran: A quantile regression approach. *Development and Capital, 6*(2), 93-110. https://doi.org/10.22103/jdc.2021.18460.1169 (In Persian)

30. Strauss, A., & Corbin, J. (2015). *Basics of qualitative research: Techniques and procedures for generating grounded theory* (2nd ed.). Sage Publications. (In Persian)
31. Thaghafi, A., & Bahar Moghaddam, M. (2008). Effective factors on earnings management. *Development and Capital, 1*(2), 103-125. https://doi.org/10.22103/jdc.2008.1894 (In Persian)
32. Yari, J. (2008). Investigating the barriers to creating a learning organization in the Iran Khodro Training Center. *Journal of Human Resource Management Studies, 1*(4), 1-10. (In Persian)

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

استناد به این مقاله: باقریان کاسگری، عباس.، رئیسی وانانی، ایمان.، امیری، مقصود.، همایون، سعید. (۱۴۰۳). شناسایی تقلب مالی در شرکت‌های سهامی عام با استفاده معیارهای مالی و غیرمالی با رویکرد یادگیری ماشین، *مطالعات مدیریت کسب وکار هوشمند*، ۱۳(۵۰)، ۹۹-۱۴۲. DOI: 10.22054/ims.2024.78018.2434

Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License..