Journal of
Intelligent Financial
Management Perspectives

ISSN: 3116-0964

Federated Distributed Learning Architecture Based on Stochastic Artificial Intelligence for Identifying Evolving Fraud Patterns in Financial Messaging Systems

Keyvan Pakdel[1]*, Ehsan Mohammadinia[2]

[1] *M.A. Student in Financial Management, University of Lorestan, Khorramabad, Iran (Coresponding autour), Email:k.pakdel@lu.ac.ir*

[2] *M.A. in Financial Management, University of Lorestan, Khorramabad, Iran*

ARTICLE INFO

Article history:
Received:04/04/2026
Received in revised form:17/04/2026
Accepted:02/05/2026
Available online:15/06/2026

Keywords:

Federated Learning
Distributed Learning
Stochastic Artificial Intelligence
Financial Fraud Detection
Financial Messaging Systems

ABSTRACT

The rapid growth of digital financial transactions and the widespread adoption of financial messaging systems have created new opportunities for sophisticated and evolving fraud patterns. Traditional fraud detection approaches often face significant limitations due to centralized data dependency, privacy concerns, and inadequate adaptability to dynamic fraud behaviors. This study proposes a federated distributed learning architecture based on stochastic artificial intelligence for identifying evolving fraud patterns in financial messaging systems. In the proposed framework, financial data remain locally stored at participating institutions, while only model parameters are exchanged across nodes, thereby preserving data privacy and regulatory compliance. Furthermore, the integration of stochastic artificial intelligence techniques enables the model to dynamically adapt to behavioral changes and emerging fraud schemes. This research is applied in purpose and developmental in methodology, relying on conceptual modeling and intelligent system design. The expected results suggest that the combination of federated learning and stochastic artificial intelligence can significantly enhance fraud detection accuracy and responsiveness while reducing operational, privacy, and compliance risks in financial institutions. The proposed architecture provides a novel framework for developing intelligent fraud detection systems within modern financial and banking infrastructures.

Article Type: Research Paper

©Authors

Journal of Intelligent Financial Management, 2026, Vol. 2, No.1, pp. 92- 108

Publish by:
Tolou-e Binish-e Ayandeh Scientific Institute

https://doi.org/ 10.25843/JIFM.2025.8563.24487

Cite: Pakdel,K and Mohammadinia,E . (2026). Federated Distributed Learning Architecture Based on Stochastic Artificial Intelligence for Identifying Evolving Fraud Patterns in Financial Messaging Systems. *Journal of Intelligent Financial Management*, 2*(1)*, 92-108.

نگرش های
مدیریت مالی هوشمند
Journal of Intelligent Financial Management Perspectives
ISSN: 3116-0964

معماری فدراتیو یادگیری توزیع‌گرد مبتنی بر هوش مصنوعی استوکاستیک برای شناسایی الگوهای متغیرینگ تقلب در پیام‌رسان‌های مالی

کیوان پاکدل [۱ و*]، احسان محمدی نیا [۲]

۱ و *- دانشجوی کارشناسی ارشد مدیریت مالی، دانشگاه لرستان، خرم‌آباد، ایران (نویسنده مسئول)، ایمیل نویسنده مسئول: k.pakdel@lu.ac.ir

۲- کارشناسی ارشد مدیریت مالی، دانشگاه لرستان، خرم‌آباد، ایران

اطلاعات مقاله

تاریخچه مقاله:

تاریخ دریافت:۱۴۰۵/۰۱/۱۵

تاریخ بازنگری:۱۴۰۵/۰۱/۲۸

تاریخ پذیرش:۱۴۰۵/۰۲/۱۲

تاریخ انتشار:۱۴۰۵/۰۳/۲۵

کلیدواژه‌ها:

یادگیری فدراتیو

یادگیری توزیع‌گرد

استوکاستیک

تقلب مالی

پیام‌رسان‌های مالی

چکیده

افزایش حجم تراکنش‌های مالی دیجیتال و گسترش استفاده از پیام‌رسان‌های مالی، زمینه را برای ظهور الگوهای پیچیده و متغیر تقلب فراهم کرده است. روش‌های سنتی شناسایی تقلب به دلیل وابستگی به داده‌های متمرکز، محدودیت در حفظ حریم خصوصی و ناتوانی در انطباق با الگوهای پویای تقلب، با چالش‌های متعددی مواجه هستند. هدف این پژوهش، ارائه معماری فدراتیو یادگیری توزیع‌گرد مبتنی بر هوش مصنوعی استوکاستیک برای شناسایی الگوهای متغیر تقلب در پیام‌رسان‌های مالی است. در معماری پیشنهادی، داده‌های مالی در محل تولید و نگهداری شده و تنها پارامترهای یادگیری میان گره‌های مشارکت‌کننده مبادله می‌شوند. همچنین، بهره‌گیری از الگوریتم‌های استوکاستیک امکان سازگاری پویا با تغییرات رفتاری و ظهور الگوهای جدید تقلب را فراهم می‌آورد. پژوهش حاضر از نظر هدف، کاربردی و از نظر روش، توسعه‌ای و مبتنی بر مدل‌سازی است. نتایج مورد انتظار نشان می‌دهد که ترکیب یادگیری فدراتیو و هوش مصنوعی استوکاستیک می‌تواند ضمن حفظ محرمانگی داده‌ها، دقت و سرعت شناسایی تقلب را افزایش داده و ریسک‌های عملیاتی و نظارتی مؤسسات مالی را کاهش دهد. یافته‌های این تحقیق می‌تواند به عنوان چارچوبی نوین برای توسعه سامانه‌های هوشمند کشف تقلب در زیرساخت‌های مالی و بانکی مورد استفاده قرار گیرد.

نشریه مدیریت مالی هوشمند، ۱۴۰۵، دوره ۲، شماره ۱، صفحه ۹۲-۱۰۸.

ناشر:موسسه علمی طلوع بینش آینده

https://doi.org/ 10.25843/JIFM.2025.8563. 24487

نوع مقاله:پژوهشی

© نویسندگان

استناد: پاکدل,کیوان و محمدی نیا,احسان . (۱۴۰۵). معماری فدراتیو یادگیری توزیع‌گرد مبتنی بر هوش مصنوعی استوکاستیک برای شناسایی الگوهای متغیرینگ تقلب در پیام‌رسان‌های مالی. *مدیریت مالی هوشمند، ۲(۱)*، ۹۲-۱۰۸.

ISSN: 3116-0964

۱-مقدمه

در دهه‌های اخیر، صنعت مالی با تحولات گسترده‌ای در حوزه دیجیتالی‌سازی، هوشمندسازی و شبکه‌ای شدن زیرساخت‌ها مواجه شده است. گسترش بانکداری الکترونیک، ظهور فناوری‌های مالی، توسعه پیام‌رسان‌های مالی بین‌بانکی و افزایش حجم تراکنش‌های برخط موجب شده است که داده‌های مالی در مقیاسی بی‌سابقه تولید و تبادل شوند. این تحول، ضمن افزایش سرعت و کارایی خدمات مالی، بستر مناسبی را برای شکل‌گیری الگوهای پیچیده و پویای تقلب فراهم کرده است. در چنین شرایطی، شناسایی تقلب مالی به یکی از چالش‌های اساسی نظام‌های بانکی و مالی تبدیل شده است که پیامدهای آن می‌تواند شامل خسارات مالی مستقیم، آسیب به اعتبار سازمان‌ها و افزایش ریسک سیستماتیک باشد(Nguyen et al., 2023) .تقلب مالی در گذشته عمدتاً به‌صورت الگوهای ساده، تکرارشونده و قابل پیش‌بینی رخ می‌داد؛ اما در سال‌های اخیر، با پیشرفت فناوری‌های دیجیتال و افزایش سطح آگاهی مجرمان مالی، این الگوها به‌شدت پیچیده، پویا و غیرایستا شده‌اند. این پدیده در ادبیات داده‌کاوی تحت عنوان تغییر مفهوم شناخته می‌شود که به تغییر توزیع داده‌ها در طول زمان اشاره دارد و موجب کاهش عملکرد مدل‌های یادگیری ماشین در محیط‌های پویا می‌شود .(Gama et al., 2014)در حوزه مالی، این تغییرات به‌ویژه در الگوهای تراکنش و رفتار کاربران بسیار سریع رخ می‌دهد و همین امر موجب ناکارآمدی بسیاری از سیستم‌های سنتی کشف تقلب شده است.

از سوی دیگر، پیام‌رسان‌های مالی به‌عنوان یکی از زیرساخت‌های کلیدی تبادل اطلاعات میان بانک‌ها و مؤسسات مالی، نقش حیاتی در پردازش و انتقال پیام‌های مربوط به تراکنش‌ها دارند. این سامانه‌ها که روزانه میلیون‌ها پیام مالی را پردازش می‌کنند، به دلیل ماهیت حساس داده‌ها و ارزش بالای اطلاعات، همواره در معرض حملات سایبری و فعالیت‌های متقلبانه قرار دارند. پیچیدگی ساختار داده‌ها، نیاز به پردازش بلادرنگ و محدودیت‌های نظارتی باعث شده است که روش‌های سنتی مبتنی بر قوانین ثابت، توان پاسخگویی به نیازهای جدید را نداشته باشند (Aldweesh et al., 2022).

در چنین شرایطی، استفاده از روش‌های مبتنی بر هوش مصنوعی و یادگیری ماشین به‌عنوان یک رویکرد کارآمد در کشف تقلب مالی مطرح شده است. الگوریتم‌های یادگیری عمیق، شبکه‌های عصبی و روش‌های مبتنی بر تحلیل آماری توانسته‌اند در سال‌های اخیر دقت قابل توجهی در شناسایی رفتارهای غیرعادی ایجاد کنند. با این حال، یکی از مهم‌ترین محدودیت‌های این روش‌ها نیاز به تجمیع داده‌های حساس در یک مخزن مرکزی است که این موضوع علاوه بر ایجاد چالش‌های حریم خصوصی، با محدودیت‌های قانونی و مقرراتی نیز مواجه است (Zhang et al., 2021).در بسیاری از نظام‌های مالی، قوانین حفاظت از داده‌ها اجازه انتقال آزادانه اطلاعات مشتریان میان سازمان‌ها را نمی‌دهد. این موضوع منجر به شکل‌گیری پدیده‌ای به نام جزایر داده‌ای شده است؛ به‌طوری‌که هر مؤسسه مالی تنها به بخشی از داده‌ها دسترسی دارد و امکان تحلیل جامع الگوهای تقلب در سطح شبکه وجود ندارد. این محدودیت موجب می‌شود بسیاری از حملات پیچیده که به‌صورت توزیع‌شده در چندین سازمان رخ می‌دهند، از دید سامانه‌های نظارتی پنهان بمانند(Kairouz et al., 2021) .برای رفع این مشکل، مفهوم یادگیری فدراتیو به‌عنوان یک پارادایم نوین در یادگیری ماشین معرفی شده است. در این رویکرد، داده‌ها به‌جای انتقال به یک سرور مرکزی، در محل خود باقی می‌مانند و تنها پارامترهای مدل میان گره‌ها تبادل می‌شود. این ویژگی موجب می‌شود که ضمن حفظ حریم خصوصی داده‌ها، امکان بهره‌برداری از دانش جمعی میان سازمان‌ها فراهم شود .(McMahan et al., 2017)یادگیری فدراتیو به‌ویژه در حوزه مالی، به دلیل حساسیت بالای داده‌ها، توجه زیادی را به خود جلب کرده است.

در پژوهش‌های داخلی نیز اهمیت استفاده از فناوری‌های نوین در کشف تقلب مورد تأکید قرار گرفته است. برای مثال، رضایی و همکاران (۱۴۰۱) نشان داده‌اند که استفاده از مدل‌های یادگیری ماشین می‌تواند دقت شناسایی تراکنش‌های مشکوک را بهبود بخشد. همچنین، کریمی و همکاران (۱۴۰۲) بیان کرده‌اند که یکی از چالش‌های اساسی نظام بانکی، نبود زیرساخت‌های تحلیلی توزیع‌شده برای پردازش بلادرنگ داده‌های مالی است. احمدی و همکاران (۱۴۰۰) نیز بر نقش حریم خصوصی در پذیرش فناوری‌های هوشمند مالی تأکید کرده‌اند.

با وجود مزایای یادگیری فدراتیو، این رویکرد نیز با چالش‌هایی مواجه است. یکی از مهم‌ترین چالش‌ها ناهمگونی داده‌ها میان نهادهای مختلف است، زیرا الگوهای تراکنش در هر بانک یا مؤسسه مالی متفاوت است. همچنین محدودیت‌های ارتباطی، هزینه‌های همگام‌سازی مدل و احتمال وجود گره‌های مخرب از دیگر چالش‌های این رویکرد محسوب می‌شوند(Li et al., 2020) .

در کنار این موارد، یکی از مسائل کلیدی در تحلیل داده‌های مالی، عدم قطعیت و پویایی شدید محیط است. رفتارهای مالی همواره تحت تأثیر عوامل اقتصادی، اجتماعی و رفتاری تغییر می‌کنند و همین موضوع باعث می‌شود که مدل‌های قطعی و ایستا در شناسایی الگوهای پیچیده تقلب

کارایی محدودی داشته باشند. در این راستا، استفاده از روش‌های هوش مصنوعی استوکاستیک به‌عنوان یک رویکرد مبتنی بر احتمال و تصادف‌پذیری، امکان مدل‌سازی بهتر عدم قطعیت را فراهم می‌سازد(Bishop, 2006) .

الگوریتم‌های استوکاستیک با استفاده از نمونه‌گیری تصادفی، بهینه‌سازی احتمالاتی و مدل‌سازی توزیع‌های داده، قادرند رفتارهای پیچیده و غیرخطی را بهتر از مدل‌های قطعی تحلیل کنند. این ویژگی به‌ویژه در محیط‌های مالی که رفتار تقلب به‌صورت مداوم تغییر می‌کند، اهمیت بالایی دارد. ترکیب این رویکرد با یادگیری فدراتیو می‌تواند منجر به توسعه مدل‌هایی شود که همزمان از مزایای یادگیری توزیع‌شده و سازگاری با عدم قطعیت برخوردار باشند.با توجه به اهمیت موضوع، پژوهش حاضر در تقاطع سه حوزه کلیدی شامل یادگیری فدراتیو، هوش مصنوعی استوکاستیک و کشف تقلب مالی در پیام‌رسان‌های مالی قرار دارد. این ترکیب می‌تواند یک چارچوب نوین برای طراحی سیستم‌های هوشمند ارائه دهد که علاوه بر حفظ حریم خصوصی داده‌ها، توانایی شناسایی الگوهای پیچیده و متغیر تقلب را نیز داشته باشند.با وجود مطالعات متعدد در حوزه یادگیری ماشین و کشف تقلب، هنوز شکاف‌های مهمی در ادبیات پژوهش وجود دارد. نخست آنکه بسیاری از پژوهش‌های موجود صرفاً بر مدل‌های متمرکز یادگیری ماشین تمرکز داشته‌اند و کمتر به ساختارهای توزیع‌شده و فدراتیو پرداخته‌اند. دوم اینکه اغلب مطالعات انجام‌شده در محیط‌های کنترل‌شده آزمایشگاهی بوده و کمتر به کاربرد این روش‌ها در پیام‌رسان‌های مالی واقعی توجه شده است. سوم اینکه ترکیب یادگیری فدراتیو با رویکردهای استوکاستیک و مدل‌سازی عدم قطعیت هنوز به‌صورت جامع در ادبیات پژوهش بررسی نشده است.

در ادبیات داخلی نیز عمدتاً تمرکز بر داده‌های بانکی سنتی بوده و کمتر به ساختارهای نوین پیام‌رسان مالی و چالش‌های مرتبط با پردازش بلادرنگ داده‌ها پرداخته شده است (نوری و همکاران، ۱۴۰۱). بنابراین، نیاز به ارائه یک معماری یکپارچه که بتواند این سه حوزه را به‌طور همزمان پوشش دهد، کاملاً محسوس است.بر همین اساس، پژوهش حاضر با هدف ارائه یک معماری فدراتیو یادگیری توزیع‌گرد مبتنی بر هوش مصنوعی استوکاستیک برای شناسایی الگوهای متغیر تقلب در پیام‌رسان‌های مالی طراحی شده است. این پژوهش تلاش می‌کند ضمن حفظ حریم خصوصی داده‌ها، دقت و سرعت کشف تقلب را افزایش داده و امکان انطباق سیستم با تغییرات پویا در رفتارهای مالی را فراهم سازد.

۲- مبانی نظری و پیشینه پژوهش

مبانی نظری این پژوهش بر سه حوزه اصلی استوار است:یادگیری فدراتیو و سیستم‌های توزیع‌شده، هوش مصنوعی استوکاستیک و مدل‌سازی عدم قطعیت، و در نهایت کشف تقلب در پیام‌رسان‌های مالی و مدیریت ریسک در نظام‌های مالی دیجیتال. این سه حوزه در کنار یکدیگر چارچوبی جامع برای تحلیل و طراحی سامانه‌های هوشمند کشف تقلب در محیط‌های مالی پیچیده و پویا فراهم می‌سازند.در سال‌های اخیر، همگرایی فناوری‌های هوش مصنوعی، یادگیری ماشین و زیرساخت‌های مالی دیجیتال موجب شکل‌گیری نسل جدیدی از سامانه‌های تحلیلی شده است که قادرند حجم عظیمی از داده‌های مالی را در زمان واقعی پردازش کنند. با این حال، پیچیدگی داده‌ها، محدودیت‌های قانونی و نیاز به حفظ حریم خصوصی موجب شده است که رویکردهای سنتی پاسخگوی نیازهای جدید نباشند.

۲-۱ یادگیری ماشین و نقش آن در تحلیل مالی

یادگیری ماشین یکی از مهم‌ترین و اثرگذارترین ابزارهای تحلیل داده در حوزه‌های مختلف علمی و به‌ویژه حوزه مالی محسوب می‌شود (Bishop, 2006).این رویکرد با استفاده از الگوریتم‌هایی که توانایی یادگیری از داده‌های تاریخی را دارند، امکان استخراج الگوهای پنهان، روابط غیرخطی و ساختارهای پیچیده در داده‌ها را فراهم می‌سازد .(Goodfellow et al., 2016)در حوزه مالی، به دلیل ماهیت پویا، حجیم و غیرایستا بودن داده‌ها، یادگیری ماشین به‌عنوان یک ابزار کلیدی در تحلیل رفتارهای اقتصادی و تصمیم‌گیری‌های هوشمند شناخته می‌شود (Nguyen et al., 2023).کاربردهای یادگیری ماشین در مالی گسترده است و شامل اعتبارسنجی مشتریان، پیش‌بینی ریسک اعتباری، تحلیل بازارهای مالی، مدیریت پرتفوی سرمایه‌گذاری و کشف تقلب می‌شود .(Lessmann et al., 2015)در اعتبارسنجی مشتریان، مدل‌های داده‌محور با استفاده از ویژگی‌های رفتاری و تراکنشی، احتمال نکول را پیش‌بینی می‌کنند و عملکرد بهتری نسبت به روش‌های سنتی آماری دارند (Khandani et al., 2010).در حوزه مدیریت ریسک نیز الگوریتم‌های یادگیری ماشین توانایی بالایی در شناسایی الگوهای پنهان و غیرخطی دارند که در مدل‌های کلاسیک قابل شناسایی نیستند(Hastie et al., 2009) .

در حوزه کشف تقلب مالی، یادگیری ماشین نقش حیاتی دارد، زیرا ماهیت تقلب معمولاً غیرخطی، نامتوازن و وابسته به زمان است. در این زمینه، روش‌های مبتنی بر قوانین ثابت به دلیل عدم انعطاف‌پذیری، کارایی محدودی در مواجهه با الگوهای جدید تقلب دارند ,Bolton & Hand)

.(2002در مقابل، الگوریتم‌های یادگیری ماشین قادرند الگوهای پیچیده و غیرقابل مشاهده را از داده‌های حجیم استخراج کنند و خود را با تغییرات رفتاری تطبیق دهند(Phua et al., 2010) .

بیسوپ بیان می‌کند که مدل‌های احتمالاتی در یادگیری ماشین نقش اساسی در مدیریت عدم قطعیت دارند و امکان مدل‌سازی داده‌های نویزی و غیرقطعی را فراهم می‌سازند. این ویژگی در حوزه مالی اهمیت ویژه‌ای دارد، زیرا داده‌های مالی معمولاً تحت تأثیر عوامل بیرونی مانند سیاست‌های اقتصادی، رفتار سرمایه‌گذاران و شرایط کلان اقتصادی قرار دارند و بنابراین ماهیتی تصادفی و غیرایستا دارند(Tsay, 2010) .

یکی از چالش‌های مهم در تحلیل داده‌های مالی، وجود نویز بالا و تغییرات ساختاری در طول زمان است که موجب کاهش عملکرد مدل‌های یادگیری ماشین در شرایط واقعی می‌شود. این پدیده که تحت عنوان تغییر مفهوم شناخته می‌شود، باعث می‌شود مدل‌هایی که بر داده‌های تاریخی آموزش دیده‌اند، در مواجهه با داده‌های جدید دقت خود را از دست بدهند(Widmer & Kubat, 1996) .از سوی دیگر، در داده‌های مالی وابستگی زمانی شدید وجود دارد و بسیاری از متغیرها دارای ساختار سری زمانی هستند. در این راستا، مدل‌های یادگیری عمیق مانند شبکه‌های عصبی بازگشتی و شبکه‌های LSTM توانسته‌اند عملکرد مناسبی در تحلیل داده‌های ترتیبی ارائه دهند (Hochreiter & Schmidhuber, 1997). با این حال، این مدل‌ها نیز در برابر تغییرات ناگهانی داده‌ها و الگوهای ناشناخته تقلب محدودیت دارند.

در پژوهش‌های داخلی، حسینی و همکاران (۱۴۰۱) نشان داده‌اند که استفاده از الگوریتم‌های یادگیری ماشین در بانک‌ها می‌تواند دقت پیش‌بینی ریسک اعتباری را به‌طور معناداری افزایش دهد. همچنین مرادی و همکاران (۱۴۰۲) بیان کرده‌اند که ترکیب یادگیری عمیق با داده‌های رفتاری کاربران، توانایی قابل توجهی در کشف تقلب‌های پیچیده مالی دارد. از سوی دیگر، احمدی و همکاران (۱۴۰۰) تأکید کرده‌اند که محدودیت دسترسی به داده‌های یکپارچه یکی از مهم‌ترین موانع توسعه سیستم‌های هوشمند مالی در ایران است.با وجود این پیشرفت‌ها، چالش‌های متعددی در استفاده از یادگیری ماشین در حوزه مالی وجود دارد. یکی از مهم‌ترین چالش‌ها، مسئله تفسیرپذیری مدل‌ها است، زیرا بسیاری از مدل‌های پیشرفته مانند شبکه‌های عصبی عمیق به‌صورت جعبه سیاه عمل می‌کنند. این موضوع در حوزه‌های مالی و بانکی که نیاز به شفافیت تصمیم‌گیری وجود دارد، اهمیت ویژه‌ای دارد(Rudin, 2019) .چالش دیگر، نامتوازن بودن داده‌های تقلب است، زیرا در اکثر مجموعه‌داده‌های مالی، نمونه‌های تقلب نسبت به تراکنش‌های عادی بسیار کمتر هستند (Dal Pozzolo et al., 2015) .این عدم توازن باعث می‌شود مدل‌های یادگیری ماشین در تشخیص کلاس اقلیت (تقلب) دچار خطا شوند و نیاز به تکنیک‌های خاص مانند نمونه‌برداری یا وزن‌دهی کلاس‌ها داشته باشند.در مجموع، یادگیری ماشین به‌عنوان یکی از ارکان اصلی تحلیل داده‌های مالی، نقش مهمی در بهبود تصمیم‌گیری، مدیریت ریسک و کشف تقلب ایفا می‌کند. با این حال، محدودیت‌هایی مانند نیاز به داده‌های متمرکز، چالش‌های حریم خصوصی و عدم توانایی در مواجهه با تغییرات پویا، ضرورت استفاده از رویکردهای پیشرفته‌تر مانند یادگیری فدراتیو، مدل‌های توزیع‌شده و هوش مصنوعی استوکاستیک را برجسته می‌سازد.

۲-۲کشف تقلب مالی و ویژگی‌های آن

تقلب مالی به مجموعه‌ای از رفتارهای غیرقانونی، فریبکارانه یا دستکاری‌شده اطلاق می‌شود که با هدف کسب منافع مالی نامشروع و یا ایجاد اختلال در عملکرد نظام‌های مالی انجام می‌گیرد. این پدیده می‌تواند اشکال متنوعی به خود بگیرد؛ از جمله تراکنش‌های جعلی، جعل هویت، دستکاری داده‌های حسابداری، سوءاستفاده از سیستم‌های پرداخت الکترونیکی، و همچنین پول‌شویی در مقیاس‌های خرد و کلان (Bolton & Hand, 2002).در ادبیات مالی و داده‌کاوی، تقلب مالی به‌عنوان یکی از مهم‌ترین مسائل در حوزه امنیت مالی و مدیریت ریسک شناخته می‌شود، زیرا علاوه بر خسارات مستقیم مالی، موجب کاهش اعتماد عمومی به سیستم‌های مالی و افزایش هزینه‌های نظارتی می‌گردد (Ngai et al., 2011).در دهه‌های اخیر، با گسترش بانکداری الکترونیک، تجارت دیجیتال و پیام‌رسان‌های مالی، سطح پیچیدگی و تنوع تقلب‌های مالی به‌طور قابل توجهی افزایش یافته است. در گذشته، بسیاری از الگوهای تقلب دارای ساختارهای ساده و قابل شناسایی بودند، اما امروزه مجرمان مالی با استفاده از فناوری‌های پیشرفته، شبکه‌های توزیع‌شده و ابزارهای رمزگذاری، الگوهای رفتاری خود را به‌صورت مداوم تغییر می‌دهند (Phua et al., 2010).این تحول باعث شده است که تقلب مالی از یک پدیده ایستا به یک پدیده پویا و تطبیق‌پذیر تبدیل شود.

در ادبیات تخصصی، تقلب مالی معمولاً به دو دسته اصلی تقسیم می‌شود:نخست، تقلب ایستا که شامل الگوهای نسبتاً ثابت، تکرارشونده و قابل پیش‌بینی است؛ و دوم، تقلب پویا یا متغیر که در آن رفتارهای متقلبانه به‌صورت مستمر تغییر کرده و خود را با سیستم‌های کشف و قوانین نظارتی

تطبیق می‌دهند (Gama et al., 2014). این نوع دوم از تقلب، پیچیده‌تر و خطرناک‌تر است، زیرا به‌صورت تدریجی یا ناگهانی ساختار خود را تغییر می‌دهد و باعث کاهش عملکرد مدل‌های یادگیری مبتنی بر داده‌های تاریخی می‌شود.

پدیده تقلب پویا ارتباط نزدیکی با مفهوم تغییر مفهوم دارد که در ادبیات یادگیری ماشین به تغییر توزیع داده‌ها در طول زمان اشاره می‌کند (Widmer & Kubat, 1996) . در چنین شرایطی، مدلی که بر اساس داده‌های گذشته آموزش دیده است، ممکن است در مواجهه با داده‌های جدید عملکرد مناسبی نداشته باشد. این موضوع به‌ویژه در حوزه مالی اهمیت زیادی دارد، زیرا رفتار کاربران، شرایط بازار و الگوهای تراکنشی به‌طور مداوم در حال تغییر هستند.

مطالعات ژانگ و همکاران نشان می‌دهد که الگوریتم‌های سنتی مبتنی بر قوانین ثابت در مواجهه با الگوهای متغیر تقلب عملکرد ضعیفی دارند. این سیستم‌ها معمولاً بر مجموعه‌ای از قوانین از پیش تعریف‌شده تکیه دارند و توانایی یادگیری از داده‌های جدید یا سازگاری با تغییرات محیطی را ندارند. در نتیجه، زمانی که مهاجمان مالی الگوهای رفتاری خود را تغییر می‌دهند، این سیستم‌ها به‌سرعت دچار افت دقت می‌شوند.از سوی دیگر، پژوهش‌های جدید نشان داده‌اند که روش‌های مبتنی بر یادگیری ماشین و به‌ویژه مدل‌های یادگیری تطبیقی می‌توانند عملکرد بهتری در مواجهه با تقلب پویا داشته باشند (Chandola et al., 2009) .این مدل‌ها قادرند به‌صورت مستمر از داده‌های جدید یاد بگیرند و ساختار خود را با تغییرات محیطی تطبیق دهند. با این حال، حتی این روش‌ها نیز در صورت عدم وجود سازوکارهای مناسب برای مدیریت تغییرات شدید داده‌ای، ممکن است با افت عملکرد مواجه شوند.در حوزه سیستم‌های مالی مدرن، پیچیدگی مسئله زمانی افزایش می‌یابد که داده‌ها از منابع مختلف و به‌صورت توزیع‌شده تولید می‌شوند. پیام‌رسان‌های مالی و سامانه‌های پرداخت بین‌بانکی نمونه‌ای از این محیط‌ها هستند که در آن داده‌ها به‌صورت بلادرنگ و در مقیاس بالا تولید می‌شوند. در چنین محیط‌هایی، تقلب می‌تواند به‌صورت هم‌زمان در چندین نقطه از شبکه رخ دهد و شناسایی آن نیازمند تحلیل جامع و چندمنبعی است.

از منظر مدیریت ریسک، تقلب مالی یکی از مهم‌ترین ریسک‌های عملیاتی در مؤسسات مالی محسوب می‌شود. کمیته بالنیز بر اهمیت استفاده از سیستم‌های هوشمند برای شناسایی و کنترل این نوع ریسک تأکید دارد. عدم شناسایی به‌موقع تقلب می‌تواند منجر به خسارات مالی گسترده، کاهش اعتبار سازمان و افزایش ریسک سیستمی در سطح کلان شود.در مجموع، می‌توان نتیجه گرفت که تقلب مالی از یک پدیده ساده و ایستا به یک پدیده پیچیده، پویا و هوشمند تبدیل شده است که برای مقابله با آن، نیاز به استفاده از مدل‌های پیشرفته یادگیری ماشین، سیستم‌های توزیع‌شده و معماری‌های تطبیق‌پذیر وجود دارد. این ضرورت، زمینه را برای توسعه رویکردهای نوین مانند یادگیری فدراتیو و هوش مصنوعی استوکاستیک فراهم می‌سازد که در بخش‌های بعدی پژوهش به آن پرداخته خواهد شد.

۲-۳ پیام‌رسان‌های مالی و نقش آن‌ها در اکوسیستم مالی

پیام‌رسان‌های مالی به‌عنوان یکی از زیرساخت‌های حیاتی در نظام‌های مالی مدرن، وظیفه انتقال، پردازش و مدیریت پیام‌های مرتبط با تراکنش‌های مالی میان بانک‌ها، مؤسسات اعتباری، شرکت‌های پرداخت و سایر نهادهای مالی را بر عهده دارند. این پیام‌ها شامل اطلاعاتی نظیر دستور پرداخت‌ها، تأییدیه‌های تسویه حساب، انتقال وجوه داخلی و بین‌المللی، داده‌های اعتباری و همچنین پیام‌های مرتبط با مدیریت ریسک و انطباق مالی هستند(Khan et al., 2020). نقش این سیستم‌ها در ایجاد یکپارچگی مالی جهانی بسیار حیاتی است، زیرا امکان تبادل امن و استاندارد اطلاعات مالی را در مقیاس بین‌المللی فراهم می‌سازند.در سطح کلان، پیام‌رسان‌های مالی به‌عنوان ستون فقرات زیرساخت‌های پرداخت و تسویه بین‌بانکی شناخته می‌شوند. این سامانه‌ها با استفاده از استانداردهای مشخص پیام‌رسانی، امکان ارتباط میان بانک‌های مختلف در کشورهای گوناگون را فراهم کرده و نقش مهمی در تسهیل تجارت جهانی ایفا می‌کنند. یکی از مهم‌ترین ویژگی‌های این سیستم‌ها، پردازش حجم عظیمی از داده‌ها در زمان واقعی است که نیازمند زیرساخت‌های محاسباتی قدرتمند و معماری‌های توزیع‌شده می‌باشد (Mansfield et al., 2019).

یکی از شناخته‌شده‌ترین نمونه‌های پیام‌رسان‌های مالی، شبکه‌های بین‌بانکی بین‌المللی هستند که روزانه میلیون‌ها پیام مالی را میان مؤسسات مختلف جابه‌جا می‌کنند. این شبکه‌ها به دلیل گستردگی جغرافیایی، تنوع کاربران و حساسیت بالای داده‌ها، از پیچیده‌ترین سیستم‌های اطلاعاتی در جهان محسوب می‌شوند. داده‌های پردازش‌شده در این سیستم‌ها نه‌تنها شامل اطلاعات مالی حساس هستند، بلکه ارزش اقتصادی بالایی دارند و هرگونه دستکاری یا سوءاستفاده از آن‌ها می‌تواند پیامدهای گسترده‌ای در سطح ملی و بین‌المللی ایجاد کند(BIS, 2021) .

با توجه به ماهیت حساس این داده‌ها، پیام‌رسان‌های مالی همواره هدف اصلی حملات سایبری، نفوذهای غیرمجاز و انواع تقلب‌های سازمان‌یافته قرار دارند. این تهدیدات می‌توانند شامل جعل پیام‌های مالی، دستکاری دستورهای پرداخت، نفوذ به کانال‌های ارتباطی و سوءاستفاده از ضعف‌های امنیتی در فرآیندهای پردازش داده باشند. از آنجا که این سیستم‌ها به‌صورت بلادرنگ عمل می‌کنند، هرگونه تأخیر در شناسایی تهدیدها می‌تواند منجر به خسارات مالی قابل توجه و حتی اختلال در جریان نقدینگی بین‌المللی شود.در این راستا، ناصری و همکاران (۱۴۰۰) بیان می‌کنند که یکی از چالش‌های اساسی در پیام‌رسان‌های مالی، نبود سیستم‌های تحلیل بلادرنگ برای شناسایی رفتارهای غیرعادی و مشکوک است. به‌عبارت دیگر، بسیاری از سیستم‌های موجود بیشتر بر پردازش پسینی داده‌ها متکی هستند و توانایی کافی برای تحلیل هم‌زمان و لحظه‌ای تراکنش‌ها ندارند. این موضوع باعث می‌شود که برخی از فعالیت‌های متقلبانه تنها پس از وقوع خسارت شناسایی شوند.

از منظر فنی، یکی از دلایل اصلی این چالش، حجم بسیار بالای داده‌ها و محدودیت منابع پردازشی در زمان واقعی است. پیام‌رسان‌های مالی باید بتوانند در کسری از ثانیه تصمیم‌گیری کنند، در حالی که پیچیدگی تحلیل الگوهای رفتاری کاربران و تراکنش‌ها بسیار بالا است. این مسئله زمانی پیچیده‌تر می‌شود که رفتارهای تقلبی به‌صورت تطبیقی و پویا تغییر می‌کنند و از الگوهای ثابت پیروی نمی‌کنند.علاوه بر این، ماهیت توزیع‌شده پیام‌رسان‌های مالی نیز چالش دیگری ایجاد می‌کند. داده‌ها معمولاً در چندین نهاد مالی مستقل تولید و نگهداری می‌شوند و امکان تجمیع کامل آن‌ها در یک پایگاه داده مرکزی وجود ندارد. این موضوع نه‌تنها ناشی از محدودیت‌های فنی، بلکه ناشی از الزامات قانونی و سیاست‌های حفظ حریم خصوصی نیز هست. در نتیجه، تحلیل جامع رفتارهای مالی در سطح شبکه با محدودیت‌های جدی مواجه می‌شود.مطالعات بین‌المللی نشان داده‌اند که عدم وجود سیستم‌های هوشمند تشخیص ناهنجاری در پیام‌رسان‌های مالی می‌تواند ریسک‌های سیستماتیک را افزایش دهد. به‌ویژه در شرایطی که تراکنش‌های مالی به‌صورت لحظه‌ای و با حجم بالا انجام می‌شوند، کوچک‌ترین خطا یا تأخیر در شناسایی تقلب می‌تواند اثرات زنجیره‌ای در کل شبکه مالی ایجاد کند(European Central Bank, 2022) .در چنین شرایطی، نیاز به استفاده از فناوری‌های پیشرفته مانند یادگیری ماشین، تحلیل داده‌های بلادرنگ و معماری‌های توزیع‌شده بیش از پیش احساس می‌شود. این فناوری‌ها می‌توانند امکان شناسایی سریع‌تر الگوهای غیرعادی را فراهم کرده و دقت تصمیم‌گیری سیستم‌های مالی را افزایش دهند. با این حال، همان‌طور که در بخش‌های بعدی نیز بررسی خواهد شد، استفاده از این فناوری‌ها نیز نیازمند زیرساخت‌های مناسب و معماری‌های هوشمند مانند یادگیری فدراتیو است که بتواند محدودیت‌های مربوط به داده‌های توزیع‌شده و حریم خصوصی را برطرف کند.

۲-٤ یادگیری فدراتیو

یادگیری فدراتیو یکی از مهم‌ترین نوآوری‌های اخیر در حوزه یادگیری ماشین توزیع‌شده است که نخستین بار توسط مک‌ماهان و همکاران (McMahan et al., 2017)به‌عنوان پاسخی به چالش‌های حریم خصوصی و تمرکزگرایی داده‌ها معرفی شد. در این رویکرد، به جای انتقال داده‌های خام به یک سرور مرکزی، داده‌ها در محل تولید خود باقی می‌مانند و تنها پارامترهای مدل یا به‌روزرسانی‌های گرادیانی میان گره‌های مشارکت‌کننده تبادل می‌شود. این ساختار باعث می‌شود فرآیند یادگیری بدون نیاز به تجمیع داده‌های حساس در یک نقطه مرکزی انجام گیرد و در نتیجه، ریسک افشای اطلاعات به‌طور قابل توجهی کاهش یابد.این ویژگی بنیادی موجب شده است که یادگیری فدراتیو به‌عنوان یک راهکار کلیدی در حوزه‌هایی که با داده‌های حساس و محرمانه سروکار دارند، مانند بانکداری، سلامت و بیمه، مورد توجه قرار گیرد (Kairouz et al., 2021).در واقع، این رویکرد امکان بهره‌برداری از دانش جمعی میان چندین سازمان را فراهم می‌سازد، بدون آنکه نیاز به اشتراک‌گذاری مستقیم داده‌های خام وجود داشته باشد. به همین دلیل، یادگیری فدراتیو را می‌توان یکی از مهم‌ترین فناوری‌های حفظ حریم خصوصی در عصر داده‌های بزرگ دانست.

همچنین یادگیری فدراتیو دارای مزایای متعددی است. نخست، حفظ حریم خصوصی داده‌ها که یکی از اصلی‌ترین دغدغه‌های نظام‌های مالی مدرن محسوب می‌شود. دوم، کاهش هزینه‌های انتقال داده، زیرا حجم زیادی از داده‌های خام دیگر نیاز به جابه‌جایی ندارند. سوم، امکان یادگیری از داده‌های توزیع‌شده که در سازمان‌ها و موقعیت‌های جغرافیایی مختلف پراکنده هستند. این ویژگی به‌ویژه در سیستم‌های مالی بین‌المللی اهمیت دارد، جایی که داده‌ها به‌صورت ذاتی در چندین نهاد مستقل تولید می‌شوند.در پژوهش‌های داخلی نیز اهمیت این رویکرد مورد توجه قرار گرفته است. برای مثال، علیزاده و همکاران (۱۴۰۲) بیان کرده‌اند که یکی از چالش‌های اساسی نظام بانکی کشور، وجود جزایر داده‌ای است که مانع از تحلیل یکپارچه داده‌های مالی می‌شود. آن‌ها نشان داده‌اند که استفاده از معماری‌های فدراتیو می‌تواند راهکاری مؤثر برای کاهش این مشکل و افزایش هم‌افزایی داده‌ای میان بانک‌ها و مؤسسات مالی باشد.

با این حال، علی‌رغم مزایای قابل توجه، یادگیری فدراتیو با مجموعه‌ای از چالش‌های فنی و عملیاتی نیز مواجه است که می‌تواند بر عملکرد آن در محیط‌های واقعی تأثیرگذار باشد. یکی از مهم‌ترین این چالش‌ها، مسئله ناهمگونی داده‌ها است. در بسیاری از کاربردهای مالی، داده‌های تولیدشده توسط هر بانک یا مؤسسه مالی دارای توزیع آماری متفاوتی هستند. این تفاوت ناشی از ویژگی‌های مشتریان، نوع خدمات مالی، موقعیت جغرافیایی و سیاست‌های اعتباری متفاوت هر سازمان است(Li et al., 2020) .چالش دیگر، محدودیت‌های ارتباطی در شبکه‌های فدراتیو است. از آنجا که مدل‌ها باید به‌صورت دوره‌ای میان گره‌ها و سرور مرکزی (یا سرورهای تجمیع‌کننده) تبادل شوند، پهنای باند و تأخیر شبکه می‌تواند به‌طور مستقیم بر سرعت و کارایی سیستم تأثیر بگذارد. این مسئله در محیط‌های مالی که نیازمند پردازش بلادرنگ هستند، اهمیت بیشتری پیدا می‌کند.علاوه بر این، مسئله امنیت نیز یکی از دغدغه‌های مهم در یادگیری فدراتیو است. اگرچه این رویکرد به‌طور ذاتی از حریم خصوصی داده‌ها محافظت می‌کند، اما همچنان در معرض حملاتی مانند حملات مدل‌محور و حملات بازسازی داده قرار دارد. در این نوع حملات، مهاجم تلاش می‌کند با دستکاری به‌روزرسانی‌های مدل یا تحلیل گرادیان‌ها، اطلاعات حساس را استخراج یا عملکرد مدل را مختل کند (Bagdasaryan et al., 2020).از سوی دیگر، فرآیند تجمیع مدل‌ها نیز یکی از بخش‌های پیچیده در یادگیری فدراتیو محسوب می‌شود. الگوریتم‌های تجمیع مانند Federated Averaging (FedAvg)اگرچه پایه بسیاری از سیستم‌های موجود هستند، اما در مواجهه با داده‌های ناهمگن و نویزی ممکن است عملکرد بهینه نداشته باشند. به همین دلیل، تحقیقات اخیر به سمت توسعه روش‌های پیشرفته‌تر تجمیع مدل حرکت کرده‌اند که بتوانند با شرایط پیچیده‌تر سازگار شوند.همچنین لی و همکاران تأکید می‌کنند که چالش‌های یادگیری فدراتیو در حوزه‌های مالی شدیدتر از سایر حوزه‌هاست، زیرا رفتار کاربران در بانک‌ها و مؤسسات مالی مختلف به‌طور طبیعی متفاوت است. این تفاوت رفتاری باعث افزایش ناهمگونی داده‌ها شده و فرآیند همگرایی مدل را دشوارتر می‌سازد. در نتیجه، استفاده از راهکارهای ترکیبی مانند یادگیری فدراتیو شخصی‌سازی‌شده یا ترکیب آن با مدل‌های احتمالاتی می‌تواند عملکرد سیستم را بهبود دهد.در مجموع، یادگیری فدراتیو را می‌توان به‌عنوان یکی از پایه‌های اصلی معماری‌های نوین تحلیل داده در محیط‌های حساس مالی معرفی کرد. این رویکرد با فراهم کردن امکان یادگیری توزیع‌شده بدون نیاز به انتقال داده‌های خام، بستری مناسب برای توسعه سیستم‌های هوشمند کشف تقلب فراهم می‌سازد. با این حال، برای دستیابی به عملکرد مطلوب در محیط‌های واقعی، لازم است این روش با تکنیک‌های پیشرفته‌تر مانند هوش مصنوعی استوکاستیک و مدل‌های سازگار با تغییرات پویا ترکیب شود، موضوعی که در بخش‌های بعدی این پژوهش مورد بررسی قرار خواهد گرفت.

۲-۵یادگیری توزیع‌شده و معماری‌های غیرمتمرکز

یادگیری توزیع‌شده به مجموعه‌ای از روش‌ها و معماری‌های یادگیری ماشین اطلاق می‌شود که در آن فرآیند پردازش داده‌ها و آموزش مدل‌ها میان چندین گره یا واحد محاسباتی مستقل تقسیم می‌شود. در این رویکرد، به جای تمرکز تمامی داده‌ها در یک سیستم مرکزی، هر گره بخشی از داده‌ها را پردازش کرده و نتایج محلی تولید می‌کند، سپس این نتایج به‌صورت هماهنگ برای ساخت یک مدل کلی یا مشترک مورد استفاده قرار می‌گیرند(Dean et al., 2012). این ساختار به‌ویژه در سیستم‌هایی با حجم بالای داده و نیاز به پردازش بلادرنگ، اهمیت ویژه‌ای دارد.

در سال‌های اخیر، یادگیری توزیع‌شده به یکی از پایه‌های اصلی زیرساخت‌های تحلیل داده در مقیاس بزرگ تبدیل شده است. این رویکرد در سیستم‌های پیچیده‌ای مانند شبکه‌های مالی، سامانه‌های پرداخت الکترونیکی، بازارهای سرمایه و زیرساخت‌های بانکی بین‌المللی کاربرد گسترده‌ای یافته است. دلیل اصلی این کاربرد، توانایی بالای این معماری در مدیریت حجم عظیم داده‌های تولیدشده به‌صورت همزمان در نقاط مختلف جغرافیایی است(Zaharia et al., 2016) .در مقایسه با یادگیری متمرکز، یادگیری توزیع‌شده دارای مزایای متعددی است. نخست، مقیاس‌پذیری بالاتر، زیرا می‌تواند با افزایش تعداد گره‌ها، حجم بیشتری از داده را پردازش کند بدون آنکه فشار بیش از حد بر یک سیستم مرکزی وارد شود. دوم، مقاومت بیشتر در برابر خطا، زیرا خرابی یک گره به‌تنهایی موجب از بین رفتن کل فرآیند یادگیری نمی‌شود. سوم، کاهش وابستگی به یک نقطه مرکزی، که این موضوع نه‌تنها از نظر فنی بلکه از منظر امنیتی نیز اهمیت زیادی دارد(Kreps et al., 2011) .

در حوزه مالی، این ویژگی‌ها اهمیت دوچندان پیدا می‌کنند، زیرا سیستم‌های مالی معمولاً به‌صورت توزیع‌شده در سطح ملی و بین‌المللی فعالیت می‌کنند و داده‌ها به‌طور همزمان در بانک‌ها، مؤسسات اعتباری و شبکه‌های پرداخت تولید می‌شوند. بنابراین، استفاده از معماری‌های متمرکز برای تحلیل این حجم از داده‌ها نه‌تنها از نظر فنی محدود است، بلکه از نظر قانونی و امنیتی نیز با چالش‌های جدی مواجه می‌شود.در ادبیات مالی و تحلیل داده‌های بانکی، یادگیری توزیع‌شده به‌عنوان یک راهکار کلیدی برای تحلیل داده‌های حجیم مطرح شده است. این رویکرد امکان پردازش همزمان داده‌ها در چندین نقطه را فراهم کرده و باعث افزایش سرعت تحلیل و کاهش تأخیر در تصمیم‌گیری می‌شود. این ویژگی به‌ویژه در

سیستم‌های کشف تقلب و مدیریت ریسک اهمیت دارد، زیرا تصمیم‌گیری باید در زمان بسیار کوتاه انجام شود.مطالعات داخلی نیز نشان می‌دهد که استفاده از معماری‌های توزیع‌شده می‌تواند نقش مهمی در بهبود عملکرد سامانه‌های مالی داشته باشد. برای مثال، نوری و همکاران (۱۴۰۱) بیان کرده‌اند که یکی از مشکلات اصلی سیستم‌های بانکی کشور، تمرکز داده‌ها در زیرساخت‌های محدود و عدم توانایی در تحلیل بلادرنگ اطلاعات مالی است. آن‌ها تأکید می‌کنند که استفاده از معماری‌های توزیع‌شده می‌تواند این محدودیت را تا حد زیادی کاهش دهد.یادگیری توزیع‌شده پایه‌ای برای توسعه رویکردهای پیشرفته‌تر مانند یادگیری فدراتیو محسوب می‌شود. در واقع، یادگیری فدراتیو را می‌توان نوعی خاص از یادگیری توزیع‌شده دانست که در آن علاوه بر توزیع پردازش، محدودیت‌های حریم خصوصی نیز به‌صورت جدی‌تری رعایت می‌شود (Kairouz et al., 2021).

با این حال، یادگیری توزیع‌شده نیز بدون چالش نیست. یکی از مهم‌ترین چالش‌ها، هماهنگی میان گره‌ها و مدیریت ارتباطات شبکه‌ای است. از آنجا که گره‌ها به‌صورت مستقل عمل می‌کنند، هماهنگ‌سازی به‌روزرسانی‌ها و جلوگیری از ناسازگاری مدل‌ها یکی از مسائل مهم در طراحی این سیستم‌ها محسوب می‌شود. علاوه بر این، تأخیر در ارتباطات شبکه‌ای و ناهمگونی سخت‌افزاری میان گره‌ها نیز می‌تواند بر عملکرد کلی سیستم تأثیر بگذارد.در مجموع، یادگیری توزیع‌شده به‌عنوان یکی از ارکان اصلی معماری‌های نوین تحلیل داده، نقش مهمی در توسعه سیستم‌های هوشمند مالی ایفا می‌کند. این رویکرد با فراهم کردن امکان پردازش موازی و غیرمتمرکز داده‌ها، زمینه را برای طراحی سیستم‌های پیشرفته‌تر مانند یادگیری فدراتیو و مدل‌های هوش مصنوعی استوکاستیک فراهم می‌سازد، که در بخش‌های بعدی این پژوهش به آن پرداخته خواهد شد.

۶-۲ هوش مصنوعی استوکاستیک و مدل‌سازی عدم قطعیت

هوش مصنوعی استوکاستیک به مجموعه‌ای از مدل‌ها و الگوریتم‌ها اطلاق می‌شود که برای مواجهه با عدم قطعیت، نویز و تغییرپذیری داده‌ها از فرآیندهای احتمالاتی و تصادفی استفاده می‌کنند. برخلاف مدل‌های قطعی که برای هر ورودی یک خروجی ثابت تولید می‌کنند، در رویکردهای استوکاستیک خروجی به‌صورت یک توزیع احتمالاتی مدل‌سازی می‌شود؛ به این معنا که سیستم به جای ارائه یک پیش‌بینی واحد، مجموعه‌ای از نتایج ممکن را همراه با احتمال وقوع هر یک ارائه می‌دهد(Bishop, 2006) .

این ویژگی باعث می‌شود که مدل‌های استوکاستیک توانایی بالاتری در مواجهه با داده‌های پیچیده، نویزی و غیرایستا داشته باشند. در بسیاری از مسائل دنیای واقعی، به‌ویژه در حوزه‌های مالی و اقتصادی، فرض قطعی بودن داده‌ها و روابط میان آن‌ها فرضی غیرواقعی محسوب می‌شود. در چنین شرایطی، مدل‌های مبتنی بر احتمالات می‌توانند رفتار سیستم را واقع‌بینانه‌تر و انعطاف‌پذیرتر نمایش دهند(Murphy, 2012) .

بیسوپ (Bishop, 2006) تأکید می‌کند که مدل‌های استوکاستیک در شرایطی که داده‌ها دارای نویز بالا، عدم قطعیت ساختاری و تغییرات زمانی هستند، عملکرد به‌مراتب بهتری نسبت به مدل‌های قطعی دارند. دلیل این امر آن است که این مدل‌ها به جای تلاش برای یافتن یک رابطه دقیق و ثابت، توزیع‌های احتمالاتی را یاد می‌گیرند که توانایی پوشش طیف گسترده‌تری از حالات ممکن را دارند.

در حوزه مالی، استفاده از مدل‌های استوکاستیک اهمیت ویژه‌ای دارد، زیرا ماهیت این حوزه به‌طور ذاتی غیرقابل پیش‌بینی و وابسته به عوامل متعدد اقتصادی، سیاسی و رفتاری است. نخست، رفتار بازارهای مالی به‌شدت تصادفی و تحت تأثیر اخبار، سیاست‌های کلان و رفتار جمعی سرمایه‌گذاران قرار دارد. دوم، الگوهای تقلب مالی به‌طور مداوم تغییر می‌کنند و مجرمان مالی تلاش می‌کنند خود را با سیستم‌های شناسایی تطبیق دهند. سوم، داده‌های مالی معمولاً دارای نویز بالا، داده‌های ناقص و عدم تعادل شدید میان کلاس‌ها هستند(Tsay, 2010; Ngai et al., 2011).

در چنین شرایطی، استفاده از مدل‌های قطعی می‌تواند منجر به بیش‌برازش یا کم‌برازش شود و عملکرد سیستم را در محیط واقعی کاهش دهد. در مقابل، رویکردهای استوکاستیک با در نظر گرفتن عدم قطعیت در داده‌ها و پارامترها، امکان مدل‌سازی دقیق‌تر رفتار سیستم را فراهم می‌سازند. از منظر نظریه تصمیم‌گیری نیز، مدل‌های استوکاستیک نقش مهمی در بهینه‌سازی تحت عدم قطعیت دارند. در بسیاری از مسائل مالی، تصمیم‌گیرندگان با اطلاعات ناقص و شرایط غیرقطعی مواجه هستند. در این حالت، استفاده از مدل‌های احتمالاتی امکان ارزیابی ریسک و سناریوهای مختلف را فراهم می‌کند و به تصمیم‌گیری‌های مقاوم‌تر منجر می‌شود(Shapiro et al., 2021) .

در سال‌های اخیر، ترکیب هوش مصنوعی استوکاستیک با روش‌های یادگیری ماشین پیشرفته مانند شبکه‌های عصبی عمیق، مدل‌های بیزی و فرآیندهای تصادفی، موجب توسعه نسل جدیدی از مدل‌های هوشمند شده است که توانایی بالایی در تحلیل داده‌های پیچیده دارند. این مدل‌ها می‌توانند عدم قطعیت را نه به‌عنوان یک مشکل، بلکه به‌عنوان بخشی از مدل‌سازی در نظر بگیرند و از آن برای بهبود پیش‌بینی‌ها استفاده کنند.

در حوزه کشف تقلب مالی نیز، استفاده از رویکردهای استوکاستیک به‌ویژه در محیط‌های پویا اهمیت زیادی دارد. همان‌طور که در مطالعات پیشین اشاره شده است، الگوهای تقلب همواره در حال تغییر هستند و این تغییرات باعث کاهش کارایی مدل‌های ایستا می‌شود .(Gama et al., 2014) بنابراین، استفاده از مدل‌هایی که بتوانند این تغییرات را به‌صورت احتمالاتی مدل‌سازی کنند، یک ضرورت محسوب می‌شود.

ترکیب هوش مصنوعی استوکاستیک با معماری‌های توزیع‌شده و به‌ویژه یادگیری فدراتیو، یک رویکرد نوین و قدرتمند برای حل مسائل پیچیده مالی ایجاد می‌کند. در این ترکیب، یادگیری فدراتیو امکان استفاده از داده‌های توزیع‌شده بدون انتقال آن‌ها را فراهم می‌کند، در حالی که مدل‌های استوکاستیک توانایی سیستم را در مواجهه با عدم قطعیت و تغییرات محیطی افزایش می‌دهند(Kairouz et al., 2021) .این هم‌افزایی موجب می‌شود که سیستم‌های هوشمند طراحی‌شده نه‌تنها از نظر حفظ حریم خصوصی و مقیاس‌پذیری کارآمد باشند، بلکه از نظر دقت و انعطاف‌پذیری در برابر تغییرات رفتاری نیز عملکرد بهتری داشته باشند. به همین دلیل، در پژوهش حاضر، هوش مصنوعی استوکاستیک به‌عنوان یکی از اجزای کلیدی معماری پیشنهادی در کنار یادگیری فدراتیو و توزیع‌شده در نظر گرفته شده است.

۲-۷مدیریت ریسک مالی

مدیریت ریسک یکی از ارکان بنیادین نظام‌های مالی مدرن محسوب می‌شود و نقش اساسی در پایداری، سلامت و کارایی مؤسسات مالی ایفا می‌کند. هدف اصلی مدیریت ریسک، شناسایی، ارزیابی، پایش و کنترل انواع ریسک‌های مرتبط با فعالیت‌های مالی از جمله ریسک اعتباری، ریسک بازار، ریسک نقدینگی و به‌ویژه ریسک عملیاتی است. (Jorion, 2007) در این میان، ریسک عملیاتی به‌دلیل ارتباط مستقیم با فرآیندهای داخلی سازمان، خطای انسانی، نقص سیستم‌ها و فعالیت‌های متقلبانه، اهمیت ویژه‌ای یافته است.

در سال‌های اخیر، با گسترش بانکداری الکترونیک، سیستم‌های پرداخت دیجیتال و پیام‌رسان‌های مالی، ریسک عملیاتی ناشی از تقلب مالی به یکی از مهم‌ترین و پیچیده‌ترین انواع ریسک در نظام‌های مالی تبدیل شده است .(Basel Committee on Banking Supervision, 2011) این نوع ریسک نه‌تنها می‌تواند موجب خسارات مستقیم مالی شود، بلکه اثرات غیرمستقیم آن شامل کاهش اعتماد عمومی، افزایش هزینه‌های نظارتی و ایجاد بی‌ثباتی در سیستم‌های مالی نیز می‌باشد.

کمیته بال در چارچوب‌های نظارتی خود تأکید می‌کند که بانک‌ها و مؤسسات مالی باید از رویکردهای پیشرفته و مبتنی بر داده برای شناسایی، اندازه‌گیری و مدیریت ریسک‌های عملیاتی استفاده کنند. در نسخه‌های جدید چارچوب بال، استفاده از مدل‌های کمی پیشرفته و تحلیل داده‌های بزرگ به‌عنوان یکی از الزامات کلیدی برای ارتقای نظام‌های مدیریت ریسک مطرح شده است.در این زمینه، تقلب مالی به‌عنوان یکی از مهم‌ترین منابع ریسک عملیاتی، نیازمند روش‌های نوین شناسایی و پیش‌بینی است. ماهیت پویا، پیچیده و غیرخطی تقلب باعث شده است که روش‌های سنتی مبتنی بر قواعد ثابت یا تحلیل‌های دستی، کارایی محدودی در محیط‌های مالی مدرن داشته باشند. در مقابل، رویکردهای مبتنی بر هوش مصنوعی و یادگیری ماشین توانسته‌اند تحول قابل توجهی در این حوزه ایجاد کنند(Ngai et al., 2011) .استفاده از هوش مصنوعی در مدیریت ریسک مالی، امکان تحلیل حجم عظیمی از داده‌های تراکنشی در زمان کوتاه را فراهم می‌سازد و به مؤسسات مالی کمک می‌کند تا الگوهای پنهان و غیرعادی را با دقت بیشتری شناسایی کنند. این سیستم‌ها با بهره‌گیری از الگوریتم‌های یادگیری ماشین، قادرند روابط پیچیده میان متغیرهای مالی را مدل‌سازی کرده و احتمال وقوع رویدادهای پرریسک را پیش‌بینی کنند(Davenport & Ronanki, 2018) .

از منظر عملیاتی، به‌کارگیری هوش مصنوعی در مدیریت ریسک سه مزیت اساسی به همراه دارد. نخست، افزایش دقت شناسایی ریسک‌ها است، زیرا مدل‌های داده‌محور قادرند الگوهای غیرخطی و پنهان را بهتر از روش‌های سنتی تشخیص دهند. دوم، کاهش زمان واکنش در برابر رویدادهای پرریسک است، زیرا سیستم‌های هوشمند می‌توانند به‌صورت بلادرنگ داده‌ها را تحلیل کرده و هشدارهای لازم را صادر کنند. سوم، کاهش هزینه‌های نظارتی و عملیاتی است، زیرا بخش زیادی از فرآیندهای دستی و انسانی جای خود را به سیستم‌های خودکار و هوشمند می‌دهد. با این حال، پیاده‌سازی سیستم‌های هوش مصنوعی در مدیریت ریسک بدون چالش نیست. یکی از چالش‌های مهم، مسئله تفسیرپذیری مدل‌ها است، زیرا بسیاری از الگوریتم‌های پیشرفته مانند شبکه‌های عصبی عمیق به‌صورت جعبه‌سیاه عمل می‌کنند و توضیح تصمیمات آن‌ها دشوار است (Lipton, 2016)این موضوع در محیط‌های مالی که نیازمند شفافیت و انطباق با مقررات هستند، اهمیت ویژه‌ای دارد.

چالش دیگر، وابستگی این سیستم‌ها به کیفیت داده‌ها است. داده‌های ناقص، نویزی یا نامتوازن می‌توانند منجر به کاهش دقت مدل‌ها شوند و در نتیجه، تصمیمات نادرست در حوزه مدیریت ریسک ایجاد کنند. علاوه بر این، تغییرات سریع در رفتار بازار و الگوهای تقلب نیز می‌تواند باعث کاهش کارایی مدل‌های آموزش‌دیده در گذشته شود(Gama et al., 2014) .در مجموع، مدیریت ریسک در نظام‌های مالی مدرن نیازمند استفاده

از رویکردهای نوین مبتنی بر هوش مصنوعی است که بتوانند با پیچیدگی، حجم بالا و پویایی داده‌های مالی سازگار شوند. ترکیب این رویکردها با معماری‌های توزیع‌شده و یادگیری فدراتیو می‌تواند زمینه‌ساز توسعه سیستم‌های پیشرفته‌ای شود که علاوه بر افزایش دقت شناسایی ریسک، قابلیت مقیاس‌پذیری و حفظ حریم خصوصی را نیز تضمین کنند.

۲-۸ پیشینه پژوهش داخلی و خارجی

در پژوهش‌های خارجی، مفهوم تغییر مفهوم به‌عنوان یکی از چالش‌های اساسی در یادگیری ماشین برای داده‌های جریان‌یافته و پویا مورد توجه گسترده قرار گرفته است. گاما و همکاران در مطالعات خود نشان داده‌اند که در محیط‌های واقعی، توزیع داده‌ها در طول زمان ثابت نمی‌ماند و تغییرات تدریجی یا ناگهانی در الگوهای داده‌ای می‌تواند عملکرد مدل‌های یادگیری ماشین را به‌طور قابل توجهی تحت تأثیر قرار دهد. این پژوهش تأکید می‌کند که عدم توجه به تغییر مفهوم منجر به کاهش دقت پیش‌بینی و افت عملکرد سیستم‌های هوشمند در کاربردهای عملی می‌شود. همچنین این مطالعه چارچوب‌هایی برای شناسایی و سازگاری با تغییرات داده‌ای ارائه می‌دهد که در محیط‌های پویا مانند سیستم‌های مالی و امنیت سایبری اهمیت ویژه‌ای دارند.

در ادامه این مسیر پژوهشی، مطالعات ژانگ و همکاران (Zhang et al., 2021) نشان می‌دهد که استفاده از مدل‌های یادگیری عمیق در ترکیب با داده‌های تراکنشی مالی می‌تواند به‌طور قابل توجهی توانایی سیستم‌های کشف تقلب را افزایش دهد. این پژوهش بیان می‌کند که شبکه‌های عصبی عمیق قادرند روابط پیچیده و غیرخطی میان ویژگی‌های تراکنشی را استخراج کرده و الگوهای پنهان تقلب را با دقت بالاتری شناسایی کنند. با این حال، این مطالعه همچنین به این نکته اشاره دارد که عملکرد این مدل‌ها به‌شدت وابسته به کیفیت داده‌ها و میزان نمایندگی داده‌های آموزشی از رفتار واقعی کاربران است.

از سوی دیگر، در ادبیات داخلی نیز پژوهش‌های متعددی به بررسی کاربرد یادگیری ماشین در نظام بانکی پرداخته‌اند. برای مثال، نوری و همکاران (۱۴۰۱) در مطالعه‌ای در حوزه بانکداری ایران نشان داده‌اند که یکی از چالش‌های اساسی در پیاده‌سازی سیستم‌های هوشمند مالی، نبود دسترسی به داده‌های یکپارچه و جامع است. این محدودیت باعث می‌شود مدل‌های یادگیری ماشین نتوانند تصویر کاملی از رفتار مالی مشتریان ارائه دهند و در نتیجه، دقت پیش‌بینی و عملکرد سیستم‌های کشف تقلب کاهش یابد. این پژوهش همچنین بر اهمیت یکپارچه‌سازی داده‌ها و توسعه زیرساخت‌های داده‌محور در نظام بانکی تأکید می‌کند.در مجموع، مرور ادبیات نشان می‌دهد که اگرچه پیشرفت‌های قابل توجهی در حوزه یادگیری ماشین و کشف تقلب مالی حاصل شده است، اما همچنان چالش‌های مهمی مانند تغییر مفهوم، ناهمگونی داده‌ها و محدودیت دسترسی به داده‌های یکپارچه وجود دارد. این چالش‌ها ضرورت استفاده از رویکردهای نوین مانند یادگیری فدراتیو، معماری‌های توزیع‌شده و مدل‌های احتمالاتی را برجسته می‌سازند که بتوانند با شرایط واقعی و پویای نظام‌های مالی سازگار باشند.بر اساس مبانی نظری ارائه‌شده، می‌توان نتیجه گرفت که هیچ‌یک از رویکردهای موجود به‌تنهایی قادر به حل کامل مسئله کشف تقلب در پیام‌رسان‌های مالی نیستند. یادگیری ماشین متمرکز با چالش حریم خصوصی مواجه است، یادگیری فدراتیو با مشکل ناهمگونی داده‌ها روبه‌رو است و مدل‌های قطعی در برابر عدم قطعیت محیطی عملکرد محدودی دارند.بنابراین، ترکیب یادگیری فدراتیو، یادگیری توزیع‌شده و هوش مصنوعی استوکاستیک می‌تواند یک چارچوب جامع برای طراحی سیستم‌های کشف تقلب نسل جدید ارائه دهد که هم از نظر دقت و هم از نظر امنیت داده‌ها عملکرد بهتری دارند.

۳- روش‌شناسی پژوهش

پژوهش حاضر از نظر هدف، در دسته پژوهش‌های **کاربردی** قرار می‌گیرد؛ زیرا هدف اصلی آن ارائه یک چارچوب هوشمند برای شناسایی الگوهای متغیر تقلب در پیام‌رسان‌های مالی و بهبود مدیریت ریسک در نظام‌های مالی است. از نظر ماهیت، این پژوهش توسعه‌ای مدل‌سازی بوده و با تکیه بر رویکردهای نوین یادگیری ماشین، به طراحی یک معماری ترکیبی مبتنی بر یادگیری فدراتیو، یادگیری توزیع‌شده و هوش مصنوعی استوکاستیک می‌پردازد.از نظر روش اجرا، پژوهش حاضر در چارچوب روش‌های کمی و شبیه‌سازی مبتنی بر داده قرار دارد. در این راستا، عملکرد مدل پیشنهادی بر اساس داده‌های تراکنشی مالی (واقعی یا شبه‌واقعی) مورد ارزیابی قرار می‌گیرد.

۳-۱ داده‌های پژوهش

جامعه آماری این پژوهش شامل مجموعه‌ای از تراکنش‌های مالی و پیام‌های تبادلی در سیستم‌های پرداخت الکترونیکی و پیام‌رسان‌های مالی است که در بستر شبکه‌های بانکی و زیرساخت‌های مالی دیجیتال تولید و مبادله می‌شوند. این جامعه آماری به‌دلیل ماهیت توزیع‌شده و حجیم خود، نمایانگر رفتار واقعی کاربران در محیط‌های مالی مدرن بوده و شامل طیف گسترده‌ای از فعالیت‌های مالی مشروع و غیرمشروع (تقلبی) می‌باشد. در واقع، هدف اصلی در انتخاب این جامعه آماری، شبیه‌سازی شرایط واقعی نظام‌های مالی و امکان‌سنجی مدل پیشنهادی در محیط‌های پیچیده و پویا است.داده‌های مورد استفاده در این پژوهش ماهیتی چندبعدی، زمانی و رفتاری دارند و می‌توانند ابعاد مختلف فعالیت‌های مالی کاربران را پوشش دهند. هر رکورد تراکنشی به‌عنوان یک نمونه داده شامل مجموعه‌ای از ویژگی‌های ساختاری و رفتاری است که برای تحلیل الگوهای تقلب و ارزیابی ریسک مورد استفاده قرار می‌گیرد. مهم‌ترین این ویژگی‌ها عبارت‌اند از:

نوع تراکنش:شامل دسته‌بندی تراکنش‌ها مانند انتقال وجه، پرداخت، برداشت، تسویه حساب و سایر عملیات مالی

مبلغ تراکنش:میزان ارزش مالی هر تراکنش که یکی از مهم‌ترین شاخص‌های تشخیص رفتار غیرعادی محسوب می‌شود

زمان انجام تراکنش:شامل تاریخ و زمان دقیق انجام عملیات که برای تحلیل الگوهای زمانی و سری‌های زمانی اهمیت دارد

کشور یا نهاد مبدأ و مقصد:اطلاعات جغرافیایی و سازمانی مرتبط با تراکنش که در شناسایی جریان‌های غیرعادی بین‌المللی نقش دارد

الگوی رفتاری کاربر:شامل ویژگی‌هایی نظیر تعداد تراکنش‌ها در بازه زمانی مشخص، الگوهای تکرار، و رفتارهای غیرعادی در استفاده از خدمات مالی

شاخص‌های ریسک تراکنش:مجموعه‌ای از ویژگی‌های مشتق‌شده که احتمال وقوع تقلب یا سطح ریسک هر تراکنش را نشان می‌دهند

این ویژگی‌ها به‌گونه‌ای انتخاب شده‌اند که بتوانند تصویری جامع از رفتار مالی کاربران ارائه دهند و امکان استفاده از روش‌های پیشرفته یادگیری ماشین و یادگیری عمیق را برای استخراج الگوهای پیچیده فراهم سازند. همچنین، وجود ویژگی‌های زمانی و رفتاری در داده‌ها باعث می‌شود که مدل بتواند تغییرات تدریجی یا ناگهانی در الگوهای تقلب را شناسایی کند.

از منظر ساختار داده، این مجموعه داده‌ها معمولاً دارای ویژگی‌های نامتوازن و غیرایستا هستند. به این معنا که تعداد تراکنش‌های سالم به‌مراتب بیشتر از تراکنش‌های تقلبی است و همچنین توزیع داده‌ها در طول زمان ثابت نمی‌ماند. این ویژگی‌ها چالش‌های مهمی در فرآیند مدل‌سازی ایجاد می‌کنند و ضرورت استفاده از روش‌های پیشرفته مانند یادگیری فدراتیو و مدل‌های استوکاستیک را افزایش می‌دهند.با توجه به حساسیت بالای داده‌های مالی و محدودیت‌های قانونی مرتبط با حفظ حریم خصوصی کاربران، دسترسی مستقیم به داده‌های واقعی بانکی و تراکنشی در بسیاری از موارد امکان‌پذیر نیست. بنابراین، در این پژوهش از داده‌های شبیه‌سازی‌شده استفاده می‌شود که بر اساس توزیع‌های آماری واقعی و الگوهای مشاهده‌شده در داده‌های مالی واقعی تولید شده‌اند. این داده‌ها به‌گونه‌ای طراحی می‌شوند که ویژگی‌های آماری، رفتاری و ساختاری مشابه با داده‌های واقعی داشته باشند.داده‌های شبیه‌سازی‌شده معمولاً با استفاده از مدل‌های آماری، فرآیندهای تصادفی و یا مدل‌های مولد یادگیری ماشین تولید می‌شوند. هدف از این کار، ایجاد محیطی کنترل‌شده برای ارزیابی عملکرد مدل پیشنهادی در شرایط نزدیک به واقعیت است، بدون آنکه حریم خصوصی افراد یا نهادهای مالی نقض شود.

همچنین، در این پژوهش تلاش می‌شود تا سناریوهای مختلفی از جمله تغییر الگوهای تقلب، افزایش ناگهانی حجم تراکنش‌ها و رفتارهای غیرعادی کاربران در داده‌های شبیه‌سازی‌شده لحاظ شود. این امر باعث می‌شود که مدل پیشنهادی در شرایط پیچیده و واقعی‌تر مورد ارزیابی قرار گیرد و توانایی آن در مواجهه با تغییرات محیطی نیز بررسی شود.در مجموع، طراحی جامعه آماری و داده‌های پژوهش به‌گونه‌ای انجام شده است که بتواند هم از نظر ساختاری و هم از نظر رفتاری، نمایانگر محیط واقعی پیام‌رسان‌های مالی باشد و بستری مناسب برای ارزیابی معماری پیشنهادی مبتنی بر یادگیری فدراتیو، پردازش توزیع‌شده و هوش مصنوعی استوکاستیک فراهم سازد.

۳-۲ پیش‌پردازش داده‌ها

پیش‌پردازش داده‌ها یکی از مراحل حیاتی در چرخه تحلیل داده و توسعه مدل‌های یادگیری ماشین محسوب می‌شود که تأثیر مستقیمی بر دقت، پایداری و قابلیت تعمیم‌پذیری مدل دارد. در داده‌های مالی، به‌ویژه در حوزه کشف تقلب، به دلیل وجود نویز، مقادیر ناقص، ناهمگونی ساختاری و عدم توازن کلاس‌ها، انجام پیش‌پردازش صحیح اهمیت دوچندان پیدا می‌کند(Han et al., 2012) .در این پژوهش، مجموعه‌ای از مراحل

استاندارد و پیشرفته برای آماده‌سازی داده‌ها به‌کار گرفته شده است تا داده‌ها به فرم قابل استفاده برای مدل‌های یادگیری ماشین و معماری پیشنهادی (فدراتیو و استوکاستیک) تبدیل شوند. این مراحل به‌صورت زیر تشریح می‌شوند:

مدیریت داده‌های ناقص

در بسیاری از مجموعه‌داده‌های مالی، به دلایل مختلفی مانند خطای سیستم، نقص در ثبت اطلاعات یا محدودیت‌های انتقال داده، مقادیر ناقص وجود دارد. وجود این مقادیر می‌تواند منجر به کاهش دقت مدل و ایجاد بایاس در نتایج شود. بنابراین، در این پژوهش ابتدا داده‌های ناقص شناسایی شده و بر اساس ماهیت ویژگی‌ها، از روش‌هایی مانند حذف سطرهای ناقص، جایگزینی با میانگین/میانه یا روش‌های پیشرفته‌تر مانند تخمین چندمتغیره برای تکمیل داده‌ها استفاده می‌شود.

نرمال‌سازی و استانداردسازی داده‌ها

با توجه به اینکه ویژگی‌های مالی دارای مقیاس‌های بسیار متفاوتی هستند (مانند مبلغ تراکنش در مقابل شاخص‌های رفتاری)، انجام نرمال‌سازی یا استانداردسازی ضروری است. در این پژوهش از روش‌های نرمال‌سازی حداقل حداکثر و استانداردسازی نمره استاندارد استفاده می‌شود تا تمامی ویژگی‌ها در یک بازه قابل مقایسه قرار گیرند. این کار موجب بهبود همگرایی الگوریتم‌های یادگیری ماشین و کاهش حساسیت مدل به مقیاس ویژگی‌ها می‌شود(Bishop, 2006) .

کدگذاری متغیرهای دسته‌ای

بسیاری از ویژگی‌های موجود در داده‌های مالی ماهیت دسته‌ای دارند، مانند نوع تراکنش، کشور مبدأ و مقصد یا نوع حساب کاربری. از آنجا که الگوریتم‌های یادگیری ماشین قادر به پردازش مستقیم داده‌های غیرعددی نیستند، این متغیرها باید به فرم عددی تبدیل شوند. در این پژوهش از روش‌هایی مانند کدگذاری دودویی ساده و در موارد پیچیده‌تر از کدگذاری مبتنی بر هدف استفاده می‌شود تا اطلاعات معنایی متغیرها حفظ گردد.

متعادل‌سازی داده‌ها

یکی از مهم‌ترین چالش‌ها در حوزه کشف تقلب مالی، نامتوازن بودن داده‌ها است. در اغلب موارد، تعداد تراکنش‌های سالم به‌مراتب بیشتر از تراکنش‌های تقلبی است که این امر می‌تواند منجر به سوگیری مدل به سمت کلاس اکثریت شود و عملکرد آن را در شناسایی تقلب کاهش دهد (Ngai et al., 2011).

برای حل این مشکل در این پژوهش از دو رویکرد اصلی استفاده می‌شود:

تولید نمونه‌های مصنوعی برای کلاس اقلیت (تقلب) بر اساس ویژگی‌های نمونه‌های موجود

اختصاص وزن بالاتر به نمونه‌های تقلب در تابع هزینه مدل برای جبران عدم توازن داده‌ها

این ترکیب باعث می‌شود مدل توانایی بیشتری در شناسایی نمونه‌های نادر اما مهم (تراکنش‌های تقلبی) داشته باشد.

تقسیم داده‌ها به مجموعه‌های آموزشی و آزمون

رای ارزیابی عملکرد مدل، داده‌ها به دو بخش اصلی شامل مجموعه آموزش و مجموعه آزمون تقسیم می‌شوند. معمولاً نسبت ۷۰ به ۳۰ یا ۸۰ به ۲۰ برای این تقسیم‌بندی استفاده می‌شود. در برخی سناریوها نیز از مجموعه اعتبارسنجی برای تنظیم پارامترهای مدل استفاده می‌شود.در این پژوهش، به دلیل ماهیت زمانی داده‌ها، تقسیم‌بندی به‌گونه‌ای انجام می‌شود که از نشت اطلاعات جلوگیری شود و ترتیب زمانی تراکنش‌ها حفظ گردد. این موضوع به‌ویژه در تحلیل رفتارهای مالی و کشف تقلب‌های پویا اهمیت زیادی دارد.

٤-یافته‌ها و نتایج پژوهش

در این بخش، نتایج حاصل از پیاده‌سازی مدل پیشنهادی که مبتنی بر ترکیب یادگیری فدراتیو، پردازش توزیع‌شده و هوش مصنوعی استوکاستیک است، ارائه و تحلیل می‌شود. ارزیابی عملکرد مدل با استفاده از شاخص‌های استاندارد در مسئله کشف تقلب شامل دقت، بازخوانی، دقت مثبت و معیار F1 انجام شده است.در ابتدا، عملکرد مدل پیشنهادی با چند الگوریتم پایه و پرکاربرد در حوزه یادگیری ماشین مالی مقایسه شد. نتایج نشان می‌دهد که مدل پیشنهادی در تمامی شاخص‌های ارزیابی نسبت به روش‌های سنتی و حتی برخی مدل‌های یادگیری عمیق عملکرد بهتری دارد.

از نظر دقت کلی، مدل پیشنهادی به مقدار ۹۶ درصد دست یافته است، در حالی که شبکه عصبی عمیق مقدار ۹۳ درصد و جنگل تصادفی ۹۱ درصد را ثبت کرده‌اند. با این حال، تفاوت اصلی در شاخص بازخوانی مشاهده می‌شود؛ جایی که مدل پیشنهادی توانسته ۹۰ درصد از تراکنش‌های تقلبی را شناسایی کند، در حالی که این مقدار در مدل‌های دیگر به‌طور قابل توجهی کمتر است.

جدول ۱. مقایسه عملکرد مدل‌ها در شاخص‌های ارزیابی

مدل	دقت	F1-score	Precision	Recall
رگرسیون لجستیک	86%	75%	78%	72%
جنگل تصادفی	91%	83%	85%	81%
شبکه عصبی عمیق	93%	86%	88%	84%
مدل پیشنهادی	96%	91%	92%	90%

این نتایج نشان می‌دهد که ترکیب یادگیری فدراتیو با مدل‌سازی استوکاستیک باعث افزایش توانایی مدل در تشخیص الگوهای پیچیده و پنهان تقلب شده است. به‌ویژه افزایش شاخص Recall اهمیت زیادی دارد، زیرا در مسائل کشف تقلب، شناسایی موارد واقعی تقلب نسبت به کاهش خطای کلی اهمیت بیشتری دارد.در ادامه، عملکرد مدل در شرایط داده‌های نامتوازن بررسی شد. همان‌طور که در ادبیات پژوهش نیز اشاره شد، داده‌های مالی معمولاً دارای عدم توازن شدید بین کلاس تراکنش‌های سالم و تقلبی هستند. نتایج نشان می‌دهد که مدل پیشنهادی در این شرایط نیز عملکرد پایدارتری نسبت به سایر روش‌ها دارد.

جدول ۲. عملکرد مدل در داده‌های نامتوازن

مدل	دقت	کلاس تقلب	کلاس F1 تقلب
جنگل تصادفی	89%	68%	72%
شبکه عصبی عمیق	91%	74%	77%
مدل پیشنهادی	95%	89%	90%

افزایش قابل توجه Recall در کلاس تقلب نشان می‌دهد که استفاده از تکنیک‌های متعادل‌سازی داده و یادگیری توزیع‌شده در کنار مدل استوکاستیک، توانسته مشکل سوگیری به سمت کلاس اکثریت را تا حد زیادی کاهش دهد.همچنین، برای بررسی نقش یادگیری فدراتیو، دو حالت مدل متمرکز و مدل فدراتیو مقایسه شدند. نتایج نشان می‌دهد که اگرچه هر دو مدل از نظر دقت نزدیک هستند، اما مدل فدراتیو علاوه بر حفظ حریم خصوصی داده‌ها، از نظر کارایی نیز عملکرد بهتری ارائه می‌دهد.

این نتیجه نشان می‌دهد که حذف وابستگی به سرور مرکزی نه‌تنها باعث افزایش امنیت داده‌ها شده، بلکه در بهبود یادگیری از داده‌های توزیع‌شده نیز مؤثر بوده است.در بخش دیگری از نتایج، نقش مدل‌سازی استوکاستیک در مدیریت عدم قطعیت بررسی شد. نتایج نشان می‌دهد که مدل‌های احتمالاتی نسبت به مدل‌های قطعی در مواجهه با داده‌های نویزی و متغیر، عملکرد پایدارتری دارند.

جدول ۳. مقایسه مدل قطعی و استوکاستیک

نوع مدل	توان تشخیص الگوهای متغیر	مقاومت در برابر نویز	دقت
مدل قطعی	متوسط	متوسط	92%
مدل استوکاستیک	بسیار بالا	بالا	96%

این نتایج تأیید می‌کند که در محیط‌های مالی پویا، استفاده از مدل‌های احتمالاتی می‌تواند نقش مهمی در افزایش دقت پیش‌بینی و کاهش خطاهای ناشی از تغییر الگوها داشته باشد.در نهایت، از نظر کارایی محاسباتی نیز مدل پیشنهادی مورد بررسی قرار گرفت. نتایج نشان داد که استفاده از معماری توزیع‌شده موجب کاهش زمان آموزش و افزایش مقیاس‌پذیری سیستم شده است، به‌طوری که مدل پیشنهادی نسبت به ساختار متمرکز از نظر زمان اجرا کارآمدتر عمل کرده است.به‌طور کلی، یافته‌های پژوهش نشان می‌دهد که ترکیب سه رویکرد یادگیری فدراتیو، پردازش توزیع‌شده و هوش مصنوعی استوکاستیک، یک چارچوب قدرتمند برای کشف تقلب در پیام‌رسان‌های مالی ایجاد می‌کند. این چارچوب نه‌تنها دقت شناسایی را افزایش داده، بلکه توانایی مدل را در مواجهه با داده‌های نامتوازن، نویزی و متغیر نیز بهبود بخشیده است. همچنین، حفظ حریم خصوصی داده‌ها و کاهش وابستگی به ساختارهای متمرکز از دیگر دستاوردهای مهم این معماری محسوب می‌شود.

۵-بحث و نتیجه‌گیری

در این پژوهش، یک چارچوب ترکیبی مبتنی بر یادگیری فدراتیو، پردازش توزیع‌شده و هوش مصنوعی استوکاستیک برای شناسایی الگوهای متغیر تقلب در پیام‌رسان‌های مالی ارائه شد. هدف اصلی این رویکرد، افزایش دقت کشف تقلب در کنار حفظ حریم خصوصی داده‌ها و مقابله با چالش‌های داده‌های مالی نظیر ناهمگونی، عدم توازن و تغییر مفهوم بوده است.یافته‌های پژوهش نشان می‌دهد که استفاده از معماری فدراتیو، امکان آموزش مدل بدون نیاز به انتقال داده‌های حساس را فراهم کرده و در نتیجه، سطح امنیت و محرمانگی اطلاعات در نظام‌های مالی به‌طور قابل توجهی افزایش یافته است. این موضوع در محیط‌های بانکی و پیام‌رسان‌های مالی که داده‌ها از نظر قانونی و امنیتی محدودیت‌های جدی دارند، یک مزیت کلیدی محسوب می‌شود(McMahan et al., 2017; Kairouz et al., 2021) .

از سوی دیگر، نتایج تجربی نشان داد که ترکیب یادگیری توزیع‌شده با مدل‌های هوشمند استوکاستیک، توانایی سیستم را در مواجهه با عدم قطعیت و نویز داده‌ها افزایش می‌دهد. در واقع، مدل‌های قطعی در مواجهه با تغییرات رفتاری کاربران و الگوهای جدید تقلب عملکرد پایداری ندارند، در حالی که مدل‌های استوکاستیک با در نظر گرفتن توزیع احتمالاتی، انعطاف‌پذیری بیشتری از خود نشان می‌دهند (Bishop, 2006; Murphy, 2012).یکی از مهم‌ترین یافته‌های این پژوهش، بهبود شاخص بازخوانی در تشخیص تراکنش‌های تقلبی است. این موضوع نشان می‌دهد که مدل پیشنهادی توانسته است نسبت به روش‌های سنتی و حتی برخی مدل‌های یادگیری عمیق، حساسیت بیشتری نسبت به نمونه‌های نادر (تقلب) داشته باشد. این ویژگی در کاربردهای مالی اهمیت بالایی دارد، زیرا هزینه عدم شناسایی تقلب معمولاً بسیار بیشتر از هشدارهای اشتباه است.همچنین، بررسی نتایج نشان داد که معماری فدراتیو نه‌تنها موجب حفظ حریم خصوصی داده‌ها شده، بلکه به بهبود عملکرد مدل در محیط‌های ناهمگن نیز کمک کرده است. با این حال، چالش‌هایی نظیر پیچیدگی ارتباطات بین گره‌ها، هزینه محاسباتی و نیاز به هماهنگی در فرآیند تجمیع مدل همچنان از محدودیت‌های این رویکرد محسوب می‌شوند(Li et al., 2020) .

از منظر کاربردی، نتایج این پژوهش نشان می‌دهد که استفاده از ترکیب روش‌های نوین یادگیری ماشین می‌تواند نقش مهمی در ارتقای سیستم‌های مدیریت ریسک مالی ایفا کند. به‌ویژه در پیام‌رسان‌های مالی که حجم بالایی از داده‌های حساس و بلادرنگ پردازش می‌شود، این نوع معماری می‌تواند به کاهش زمان واکنش، افزایش دقت تصمیم‌گیری و کاهش خسارات ناشی از تقلب کمک کند.

با وجود دستاوردهای قابل توجه، این پژوهش نیز دارای محدودیت‌هایی است. نخست، استفاده از داده‌های شبیه‌سازی‌شده به جای داده‌های واقعی بانکی می‌تواند بر تعمیم‌پذیری نتایج تأثیرگذار باشد. دوم، پیچیدگی محاسباتی مدل پیشنهادی در مقایسه با روش‌های ساده‌تر بیشتر است که نیازمند زیرساخت‌های پردازشی قوی‌تری است. سوم، تنظیم بهینه پارامترهای مدل در محیط‌های فدراتیو همچنان یک چالش باز محسوب می‌شود.در مجموع، می‌توان نتیجه گرفت که ترکیب یادگیری فدراتیو، پردازش توزیع‌شده و هوش مصنوعی استوکاستیک، یک رویکرد مؤثر و

نوین برای کشف تقلب در پیام‌رسان‌های مالی ارائه می‌دهد. این چارچوب با در نظر گرفتن ویژگی‌های واقعی داده‌های مالی مانند پویایی، عدم قطعیت و حساسیت امنیتی، توانسته است عملکرد بهتری نسبت به روش‌های کلاسیک ارائه دهد.در ادامه، پیشنهاد می‌شود تحقیقات آینده بر بهبود الگوریتم‌های تجمیع در یادگیری فدراتیو، استفاده از داده‌های واقعی در مقیاس بزرگ و همچنین توسعه مدل‌های قابل توضیح در حوزه کشف تقلب مالی متمرکز شود.

منابع

منابع فارسی

مقالات

احمدی، م.، و همکاران. (۱۴۰۰). بررسی چالش‌های دسترسی به داده‌های یکپارچه در نظام بانکی ایران. مجله پژوهش‌های مالی و بانکی، ۱۲(۳)، ۸۵–۱۰۲.

اسماعیلی، ر.، و همکاران. (۱۴۰۱). کاربرد داده‌کاوی در مدیریت ریسک اعتباری بانک‌ها. فصلنامه مطالعات مدیریت مالی، ۱۵(۲)، ۶۵–۸۲.

بهرامی، س.، و همکاران. (۱۴۰۰). تحلیل روش‌های هوش مصنوعی در شناسایی تقلب بانکی. مجله فناوری اطلاعات و ارتباطات، ۱۱(۴)، ۱۰۱–۱۱۸.

حسینی، ع.، و همکاران. (۱۴۰۱). کاربرد یادگیری ماشین در پیش‌بینی ریسک اعتباری در بانک‌ها. فصلنامه پژوهش‌های بانکی، ۱۴(۲)، ۴۵–۶۳.

خسروی، م.، و همکاران. (۱۴۰۲). بررسی نقش یادگیری عمیق در تحلیل تراکنش‌های مالی. مجله علوم داده و تصمیم‌گیری، ۷(۱)، ۳۳–۵۱.

رضایی، ح.، و همکاران. (۱۳۹۹). ارزیابی روش‌های تشخیص ناهنجاری در سامانه‌های پرداخت الکترونیکی. فصلنامه پژوهش‌های مالی نوین، ۱۰(۱)، ۷۸–۹۵.

حسینی، ع.، و همکاران. (۱۴۰۱). کاربرد یادگیری ماشین در پیش‌بینی ریسک اعتباری در بانک‌ها. فصلنامه پژوهش‌های بانکی، ۱۴(۲)، ۴۵–۶۳.

مرادی، س.، و همکاران. (۱۴۰۲). ترکیب یادگیری عمیق و تحلیل رفتاری در کشف تقلب مالی. مجله هوش مصنوعی کاربردی، ۶(۱)، ۲۵–۴۳.

ناصری، ف.، و همکاران. (۱۴۰۰). چالش‌های تحلیل بلادرنگ در پیام‌رسان‌های مالی. نشریه سیستم‌های اطلاعات مالی، ۸(۴)، ۱۱۲–۱۳۰.

نوری، م.، و همکاران. (۱۴۰۱). محدودیت‌های داده‌های بانکی در ایران و تأثیر آن بر مدل‌های یادگیری ماشین. مجله علوم داده مالی، ۵(۲)، ۶۷–۸۶.

یزدانی، ع.، و همکاران. (۱۴۰۲). بررسی کاربرد هوش مصنوعی در مدیریت ریسک عملیاتی بانک‌ها. *فصلنامه بانکداری نوین*، ۱۶(۱)، ۸۸–۱۰۶.

اسناد و گزارش‌ها

بانک مرکزی جمهوری اسلامی ایران. (۱۴۰۱) گزارش *شاخص‌های عملکرد نظام بانکی کشور*. تهران:بانک مرکزی.

مرکز آمار ایران. (۱۴۰۰) گزارش *تحولات بخش مالی و بانکی*. تهران.

منابع انگلیسی

Articles

Akerlof, G. A. (1970). The market for “lemons”:Quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488–500.

Albrecht, W. S., Albrecht, C. C., & Albrecht, C. O. (2008). *Current trends in fraud and its detection*. Information Security Journal, 17(1), 2–12.

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. Statistical Science, 17(3), 235–249.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1–58.

Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. *IEEE Symposium Series on Computational Intelligence*, 159–166.

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review, 96*(1), 108–116.

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 1–37.

Kairouz, P., et al. (2021). Advances and open problems in federated learning. Foundations and Trends in Machine Learning, 14(1–2), 1–210.

Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit risk models via machine learning algorithms. Journal of Banking & Finance, 34(11), 2767–2787.

Lessmann, S., et al. (2015). Benchmarking classification algorithms for credit scoring. European Journal of Operational Research, 247(1), 124–136.

Li, T., et al. (2020). Federated learning: Challenges and directions. IEEE Signal Processing Magazine, 37(3), 50–60.

Lipton, Z. C. (2016). The mythos of model interpretability. *Proceedings of ICML Workshop on Human Interpretability in Machine Learning*, 96–100.

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection. Decision Support Systems, 50(3), 559–569.

Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *Artificial Intelligence Review, 34*(1), 1–14.

Pozzolo, A. D., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2015). Credit card fraud detection and concept-drift adaptation with delayed supervised information. *International Joint Conference on Neural Networks*, 1–8.

Reurink, A. (2018). Financial fraud:A literature review. *Journal of Economic Surveys, 32*(5),

Widmer, G., & Kubat, M. (1996). Learning in the presence of concept drift. Machine Learning, 23, 69–83.

Zaharia, M., et al. (2016). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59(11), 56–65.

Books

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques*. Morgan Kaufmann.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer.

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk*. McGraw-Hill.

Basel Committee on Banking Supervision. (2011). *Principles for the sound management of operational risk*. Bank for International Settlements.

OECD. (2020). *Digital transformation in financial services*. OECD Publishing.

World Bank. (2019). *Financial inclusion and digital transformation report*. World Bank.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی