دوره ۲، شماره ۲، سال ۱۴۰۵

هوشمندسازی در حسابداری، اقتصاد و مدیریت مالی

بررسی نقش هوش مصنوعی توضیح‌پذیر در افزایش اعتماد مدیران مالی به خروجی‌های الگوریتمی و تصمیمات سرمایه‌گذاری

۱. مهدی طاهری نیا: گروه حسابداری، دانشگاه تهران، تهران، ایران

۲. ابراهیم رویانی: گروه حسابداری، دانشگاه تهران، تهران، ایران

۳. سید رضا مرادیان*: گروه حسابداری، دانشگاه تهران، تهران، ایران (نویسنده مسئول). seyedreza.moradian.acc@gmail.com

چکیده

هدف پژوهش حاضر تبیین نقش هوش مصنوعی توضیح‌پذیر در شکل‌گیری اعتماد مدیران مالی به خروجی‌های الگوریتمی و نحوه اثرگذاری آن بر تصمیمات سرمایه‌گذاری در سازمان‌های مالی و غیرمالی شهر تهران بود. این پژوهش با رویکرد کیفی و روش تحلیل مضمون انجام شد. مشارکت‌کنندگان شامل ۲۰ نفر از مدیران مالی، مدیران سرمایه‌گذاری، تحلیلگران ارشد مالی، مدیران ریسک و مدیران فناوری مالی فعال در شرکت‌ها و مؤسسات مالی مستقر در تهران بودند که با نمونه‌گیری هدفمند و گلوله‌برفی انتخاب شدند. داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافته گردآوری شد و فرایند نمونه‌گیری تا دستیابی به اشباع نظری ادامه یافت؛ اشباع اولیه در مصاحبه هجدهم حاصل شد و دو مصاحبه تکمیلی برای اطمینان از کفایت داده‌ها انجام گرفت. مصاحبه‌ها پس از اخذ رضایت آگاهانه ضبط، پیاده‌سازی و با استفاده از نرم‌افزار NVivo تحلیل شدند. تحلیل داده‌ها بر اساس مراحل کدگذاری باز، طبقه‌بندی مفهومی و استخراج مضامین اصلی انجام شد و برای افزایش اعتبار یافته‌ها از بازبینی مشارکت‌کنندگان، بررسی همتایان، مسیر ممیزی و مقایسه مستمر داده‌ها استفاده گردید. تحلیل داده‌ها به استخراج پنج مقوله اصلی منجر شد: «کاهش ابهام شناختی از طریق فهم منطق الگوریتم»، «تقویت اعتبارسنجی حرفه‌ای و کنترل انسانی»، «افزایش پاسخگویی، شفافیت ریسک و حاکمیت داده»، «طراحی تعاملی و زمینه‌مند توضیحات الگوریتمی» و «تبدیل اعتماد کورکورانه به اعتماد مشروط و قابل پایش». مشارکت‌کنندگان تأکید کردند که توضیح‌پذیری زمانی اعتماد ایجاد می‌کند که فراتر از نمایش فنی متغیرها باشد و بتواند ارتباط میان داده‌ها، منطق مدل، سطح عدم‌قطعیت و پیامد تصمیم سرمایه‌گذاری را برای مدیر مالی روشن کند. نتایج نشان داد هوش مصنوعی توضیح‌پذیر نه صرفاً یک ابزار فنی، بلکه سازوکاری مدیریتی برای کاهش فاصله میان تحلیل الگوریتمی و قضاوت حرفه‌ای است. اعتماد مدیران مالی به خروجی‌های الگوریتمی زمانی افزایش می‌یابد که توضیحات ارائه‌شده دقیق، قابل فهم، متناسب با زمینه تصمیم، قابل ممیزی و همراه با امکان مداخله انسانی باشد. بنابراین، پیاده‌سازی موفق XAI در تصمیمات سرمایه‌گذاری مستلزم ترکیب شفافیت فنی، حاکمیت داده، آموزش مدیران مالی و چارچوب‌های پاسخگویی سازمانی است.

کلیدواژه‌ها: هوش مصنوعی توضیح‌پذیر؛ اعتماد الگوریتمی؛ مدیران مالی؛ تصمیمات سرمایه‌گذاری؛ تحلیل مضمون؛ حاکمیت داده؛ خروجی‌های الگوریتمی

شیوه استناددهی: طاهری نیا، مهدی.، رویانی، ابراهیم.، و مرادیان، سید رضا. (۱۴۰۵). بررسی نقش هوش مصنوعی توضیح‌پذیر در افزایش اعتماد مدیران مالی به خروجی‌های الگوریتمی و تصمیمات سرمایه‌گذاری. *هوشمندسازی در حسابداری، اقتصاد و مدیریت مالی*، ۲(۲)، ۲۶-۱۴.

© ۱۴۰۵ تمامی حقوق انتشار این مقاله متعلق به نویسنده است.

Volume 2, Issue 2, Year 2026

Smartification in Accounting, Economics, and Financial Management

The Role of Explainable Artificial Intelligence in Enhancing Financial Managers' Trust in Algorithmic Outputs and Investment Decisions

1. Mehdi Taheri Nia: Department of Accounting, University of Tehran, Tehran, Iran
2. Ebrahim Royani: Department of Accounting, University of Tehran, Tehran, Iran (Corresponding Author). seyedreza.moradian.acc@gmail.com
3. Seyed Reza Moradian*: Department of Accounting, University of Tehran, Tehran, Iran

Abstract

This study aimed to explain how explainable artificial intelligence contributes to financial managers' trust in algorithmic outputs and influences investment decision-making in organizations located in Tehran. This qualitative study was conducted using thematic analysis. The participants consisted of 20 financial managers, investment managers, senior financial analysts, risk managers, and financial technology managers working in financial and non-financial organizations in Tehran. Participants were selected through purposive and snowball sampling. Data were collected through semi-structured interviews and continued until theoretical saturation was achieved. Initial saturation occurred at the eighteenth interview, and two additional interviews were conducted to confirm data adequacy. Interviews were recorded with informed consent, transcribed verbatim, and analyzed using NVivo software. Data analysis involved open coding, conceptual categorization, and extraction of main themes. Credibility and trustworthiness were enhanced through member checking, peer review, audit trail, and constant comparison. The analysis resulted in five main categories: "reducing cognitive ambiguity through understanding algorithmic logic," "strengthening professional validation and human control," "enhancing accountability, risk transparency, and data governance," "interactive and context-sensitive design of algorithmic explanations," and "transforming blind trust into conditional and monitorable trust." Participants emphasized that explainability builds trust only when it goes beyond the technical display of variables and clarifies the relationship among data, model logic, uncertainty level, and the consequences of investment decisions. The findings indicated that explainable artificial intelligence is not merely a technical tool but a managerial mechanism for bridging the gap between algorithmic analysis and professional judgment. Financial managers' trust in algorithmic outputs increases when explanations are accurate, understandable, context-relevant, auditable, and accompanied by human intervention mechanisms. Accordingly, successful implementation of XAI in investment decision-making requires the integration of technical transparency, data governance, financial managers' training, and organizational accountability frameworks.

Keywords: *Explainable Artificial Intelligence; Algorithmic Trust; Financial Managers; Investment Decisions; Thematic Analysis; Data Governance; Algorithmic Outputs*

How to cite: Taheri Nia, M., Royani, E., & Moradian, S. R. (2026). The Role of Explainable Artificial Intelligence in Enhancing Financial Managers' Trust in Algorithmic Outputs and Investment Decisions. *Smartification in Accounting, Economics, and Financial Management, 2*(2), 14-26.

© 2026 the authors.

مقدمه

تحول دیجیتال در سال‌های اخیر، ماهیت تصمیم‌گیری مالی را به‌طور بنیادین دگرگون کرده است. سازمان‌ها دیگر صرفاً بر گزارش‌های مالی سنتی، تجربه مدیران و تحلیل‌های خطی متکی نیستند، بلکه به‌صورت فزاینده از الگوریتم‌های یادگیری ماشین، مدل‌های پیش‌بینی، سامانه‌های رتبه‌بندی ریسک، تحلیل داده‌های کلان و ابزارهای پشتیبان تصمیم برای ارزیابی فرصت‌های سرمایه‌گذاری استفاده می‌کنند. در این بستر، تصمیمات سرمایه‌گذاری از حالت مبتنی بر قضاوت فردی صرف خارج شده و به فرایندی داده‌محور، الگوریتم‌پشتیبان و چندمنبعی تبدیل شده است. بانک انگلستان و نهاد ناظر مالی بریتانیا در پیمایش سال ۲۰۲۴ گزارش کردند که ۷۵ درصد شرکت‌های مالی مورد بررسی از هوش مصنوعی استفاده می‌کنند و ۱۰ درصد دیگر نیز قصد استفاده از آن را در سه سال آینده دارند؛ همچنین مدل‌های بنیادین ۱۷ درصد موارد استفاده از هوش مصنوعی را تشکیل داده‌اند، که نشان‌دهنده ورود مدل‌های پیچیده‌تر به فضای خدمات مالی است . این روند نشان می‌دهد که مسئله اصلی سازمان‌ها دیگر صرفاً «استفاده یا عدم استفاده» از هوش مصنوعی نیست، بلکه چگونگی اعتماد، نظارت و تفسیر خروجی‌های الگوریتمی در تصمیمات حساس مالی است.

با وجود مزایای هوش مصنوعی در افزایش سرعت تحلیل، پردازش حجم عظیم داده‌ها، شناسایی الگوهای پنهان و کاهش برخی خطاهای انسانی، استفاده از الگوریتم‌ها در تصمیمات سرمایه‌گذاری با یک چالش اساسی مواجه است: بسیاری از مدل‌های قدرتمند یادگیری ماشین از منظر کاربران نهایی ماهیتی جعبه‌سیاه دارند. مدل‌هایی مانند شبکه‌های عصبی عمیق، جنگل تصادفی، گرادیان بوستینگ و سامانه‌های ترکیبی پیش‌بینی ممکن است از نظر دقت عملکردی برتر باشند، اما منطق درونی آن‌ها برای مدیران مالی، هیئت‌مدیره، حسابرسان، ناظران و حتی گاهی متخصصان داده به‌راحتی قابل تبیین نیست. ادبیات هوش مصنوعی توضیح‌پذیر نشان می‌دهد که مسئله توضیح‌پذیری دقیقاً از همین شکاف میان «قدرت پیش‌بینی» و «قابلیت فهم» پدید آمده است. آریتا و همکاران، XAIرا حوزه‌ای می‌دانند که تلاش می‌کند مدل‌های یادگیری ماشین را از طریق شفافیت، قابلیت فهم، پاسخگویی و تفسیرپذیری برای کاربران انسانی قابل استفاده‌تر کند (Barredo Arrieta et al., ۲۰۲۰) . آدادی و برادا نیز تأکید می‌کنند که نبود شفافیت در سامانه‌های هوشمند، مانعی مهم در پذیرش اجتماعی و سازمانی آن‌هاست (Adadi & Berrada, ۲۰۱۸) .

در حوزه مالی، این مسئله اهمیت بیشتری می‌یابد؛ زیرا خروجی‌های الگوریتمی معمولاً به تصمیماتی منجر می‌شوند که پیامدهای اقتصادی، حقوقی، اعتباری و راهبردی دارند. تصمیم درباره خرید یا فروش یک دارایی، ورود به یک پروژه سرمایه‌گذاری، تخصیص منابع به پرتفوی، ارزیابی ریسک نکول، تعیین حد اعتباری یا خروج از بازار، صرفاً یک پیش‌بینی آماری نیست؛ بلکه تصمیمی است که می‌تواند بر سودآوری شرکت، نقدینگی، ارزش سهام، اعتبار مدیران و اعتماد ذی‌نفعان اثر بگذارد. از این رو، مدیر مالی برای پذیرش خروجی الگوریتم نیاز دارد بداند مدل بر اساس چه داده‌هایی به نتیجه رسیده، کدام متغیرها نقش تعیین‌کننده داشته‌اند، سطح اطمینان پیش‌بینی چقدر است، آیا داده‌ها

سوگیرانه یا ناقص بوده‌اند، و در صورت خطای مدل چه کسی پاسخگو خواهد بود. اصول هوش مصنوعی سازمان همکاری و توسعه اقتصادی نیز بر شفافیت، توضیح‌پذیری، آگاهی کاربران از تعامل با سامانه هوشمند، فهم عوامل و منطق منتهی به خروجی و امکان اعتراض به نتایج تأکید می‌کند . بنابراین، توضیح‌پذیری در تصمیمات مالی نه یک ویژگی تزئینی، بلکه بخشی از حکمرانی ریسک و پاسخگویی حرفه‌ای است.

اعتماد مدیران مالی به خروجی‌های الگوریتمی را نمی‌توان صرفاً با دقت آماری مدل توضیح داد. اعتماد در این زمینه پدیده‌ای چندبعدی است که از تعامل میان کیفیت داده، اعتبار مدل، تجربه قبلی مدیر، میزان ریسک تصمیم، شفافیت فرایند، فرهنگ سازمانی و امکان کنترل انسانی شکل می‌گیرد. پژوهش‌های مربوط به «گریز از الگوریتم» نشان داده‌اند که کاربران پس از مشاهده خطای الگوریتم، حتی اگر مدل در مجموع عملکرد بهتری از انسان داشته باشد، ممکن است از آن فاصله بگیرند .(Dietvorst et al., ۲۰۱۵) این پدیده در مالی تشدید می‌شود، زیرا زیان ناشی از خطای الگوریتمی می‌تواند ملموس، سریع و پرهزینه باشد. در مقابل، برخی مطالعات نشان می‌دهند که توضیحات مناسب می‌توانند پذیرش مشاوران الگوریتمی مالی را افزایش دهند. بن‌دیوید، رشِف و ترون در مطالعه تجربی خود درباره مشاوران الگوریتمی مالی نشان دادند که توضیحات مبتنی بر دقت و ویژگی‌ها می‌تواند پذیرش، اعتماد و آمادگی پرداخت برای مشاوره الگوریتمی را افزایش دهد، به‌ویژه زمانی که مدل با شکست یا خطا مواجه می‌شود . (Ben David et al., ۲۰۲۱) این یافته برای مدیران مالی اهمیت دارد، زیرا در محیط واقعی، هیچ الگوریتمی همیشه صحیح نیست و اعتماد پایدار زمانی شکل می‌گیرد که مدیر بتواند خطا، عدم‌قطعیت و حدود کاربرد مدل را درک کند.

از منظر فنی، هوش مصنوعی توضیح‌پذیر ابزارها و رویکردهای متنوعی را برای فهم مدل‌ها فراهم می‌کند. روش LIME که توسط ریبیرو، سینگ و گسترین معرفی شد، پیش‌بینی هر طبقه‌بند را از طریق ساخت یک مدل محلی قابل تفسیر در اطراف همان پیش‌بینی توضیح می‌دهد و هدف آن کمک به کاربر برای ارزیابی اعتماد به پیش‌بینی است . (Ribeiro et al., ۲۰۱۶) روش SHAP نیز که توسط لوندبرگ و لی ارائه شد، بر اساس مقادیر شاپلی تلاش می‌کند سهم هر ویژگی را در خروجی مدل مشخص کند و چارچوبی یکپارچه برای تفسیر پیش‌بینی‌ها فراهم آورد . (Lundberg & Lee, ۲۰۱۷)با این حال، کاربرد این روش‌ها در سطح مدیریتی نیازمند ترجمه سازمانی است. یک نمودار SHAP یا فهرست اهمیت متغیرها ممکن است برای متخصص داده معنادار باشد، اما برای مدیر مالی تنها زمانی ارزش تصمیم‌ساز دارد که به زبان ریسک، بازده، نقدینگی، محدودیت سرمایه، سناریوی بازار و مسئولیت تصمیم ترجمه شود. بنابراین، توضیح‌پذیری مؤثر در مدیریت مالی صرفاً ارائه خروجی فنی نیست، بلکه فرایند معناسازی بین مدل و تصمیم است.

ادبیات اجتماعی و شناختی توضیح نیز نشان می‌دهد که انسان‌ها توضیح را صرفاً به‌عنوان بازنمایی کامل فرایند علی نمی‌خواهند؛ بلکه معمولاً به توضیحاتی نیاز دارند که انتخاب‌گر، زمینه‌مند، مقایسه‌ای و متناسب با سؤال آن‌ها باشد. میلر با مرور بینش‌های علوم اجتماعی در XAI تأکید می‌کند که توضیحات باید با نحوه فهم انسانی از علیت، تضادها و انتظارات اجتماعی هماهنگ باشند . (Miller, ۲۰۱۹) این موضوع در تصمیمات سرمایه‌گذاری کاملاً مشهود است. مدیر مالی ممکن است نپرسد «کل ساختار ریاضی مدل چیست؟»، بلکه بپرسد «چرا این پروژه نسبت به پروژه مشابه رد شد؟»، «کدام متغیر بیشترین اثر را بر توصیه فروش داشت؟»، «اگر نرخ ارز ۱۰ درصد تغییر کند، پیشنهاد مدل چه

می‌شود؟» یا «آیا مدل ریسک نقدشوندگی را لحاظ کرده است؟». چنین پرسش‌هایی نشان می‌دهد که توضیح‌پذیری باید به پرسش‌های تصمیم‌گیرنده پاسخ دهد، نه صرفاً به نیازهای فنی طراح مدل.

از سوی دیگر، توضیح‌پذیری همیشه به‌طور خودکار به اعتماد بهتر منجر نمی‌شود. برخی پژوهش‌ها هشدار داده‌اند که ابزارهای تفسیرپذیری می‌توانند به سوءبرداشت، اعتماد بیش از حد یا بار شناختی منجر شوند. مطالعه پورسبزی-سنگده و همکاران نشان داد که شفاف‌تر شدن ظاهری مدل لزوماً باعث پیروی مناسب‌تر کاربران از پیش‌بینی نمی‌شود و حتی ممکن است توانایی آنان در تشخیص خطاهای مدل را کاهش دهد . (Poursabzi-Sangdeh et al., ۲۰۲۱) همچنین پژوهش کائور و همکاران نشان داد که حتی متخصصان داده نیز ممکن است ابزارهای تفسیرپذیری را به شیوه‌های متفاوت و گاه نادرست به کار گیرند . (Kaur et al., ۲۰۲۰) بنابراین، هدف XAI نباید افزایش اعتماد مطلق باشد، بلکه باید شکل‌گیری «اعتماد کالیبره‌شده» یا «اعتماد متناسب با شواهد» باشد؛ یعنی مدیر مالی بداند چه زمانی به مدل اتکا کند، چه زمانی احتیاط کند و چه زمانی خروجی را به چالش بکشد.

در زمینه سازمانی ایران و به‌ویژه تهران، که بخش مهمی از شرکت‌های بزرگ، مؤسسات مالی، شرکت‌های سرمایه‌گذاری، بانک‌ها، کارگزاری‌ها و شرکت‌های فناوری مالی در آن متمرکز هستند، مسئله اعتماد به خروجی‌های الگوریتمی ابعاد ویژه‌ای دارد. مدیران مالی با محیطی مواجه‌اند که در آن نوسانات اقتصادی، عدم‌قطعیت‌های سیاستی، تغییرات نرخ ارز، محدودیت‌های اطلاعاتی، کیفیت نامتوازن داده‌ها و حساسیت تصمیمات سرمایه‌گذاری بسیار بالاست. در چنین فضایی، حتی اگر الگوریتم‌ها بتوانند الگوهای پیچیده را تحلیل کنند، پذیرش آن‌ها بدون توضیح‌پذیری دشوار خواهد بود. مدیر مالی باید بتواند خروجی مدل را در جلسات کمیته سرمایه‌گذاری، هیئت‌مدیره یا کمیته ریسک دفاع کند. اگر خروجی الگوریتم صرفاً به‌صورت عدد، امتیاز، رتبه یا توصیه خرید و فروش ارائه شود، مدیر ممکن است آن را به‌عنوان ابزار کمکی بپذیرد اما مسئولیت تصمیم نهایی را بر اساس آن نپذیرد. بنابراین، توضیح‌پذیری به پلی میان تحلیل الگوریتمی و پاسخگویی مدیریتی تبدیل می‌شود.

با وجود رشد مطالعات جهانی درباره XAI ، بخش قابل توجهی از پژوهش‌ها بر جنبه‌های فنی، دقت مدل، مقایسه روش‌های تفسیرپذیری یا آزمایش‌های کاربری عمومی تمرکز داشته‌اند. در مقابل، تجربه زیسته مدیران مالی در مواجهه با خروجی‌های الگوریتمی و انتظارات آنان از توضیح‌پذیری کمتر بررسی شده است. شکاف اصلی این است که هنوز روشن نیست مدیران مالی در عمل چه نوع توضیحی را اعتمادآفرین می‌دانند، چه عواملی باعث تردید آنان نسبت به مدل می‌شود، چگونه میان قضاوت حرفه‌ای و خروجی الگوریتمی تعادل برقرار می‌کنند، و XAI چگونه می‌تواند تصمیمات سرمایه‌گذاری را از حالت اتکای کورکورانه یا رد کامل الگوریتم به سمت تصمیم‌گیری تلفیقی و مسئولانه هدایت کند. از این رو، پژوهش حاضر با رویکرد کیفی تلاش می‌کند نقش هوش مصنوعی توضیح‌پذیر را در افزایش اعتماد مدیران مالی به خروجی‌های الگوریتمی و تصمیمات سرمایه‌گذاری تبیین کند. پرسش اصلی پژوهش این است: هوش مصنوعی توضیح‌پذیر از دیدگاه مدیران مالی چگونه می‌تواند اعتماد به خروجی‌های الگوریتمی و کیفیت تصمیمات سرمایه‌گذاری را افزایش دهد؟

مواد و روش‌ها

پژوهش حاضر از نظر هدف، کاربردی و از نظر رویکرد، کیفی است و با روش تحلیل مضمون انجام شد. انتخاب رویکرد کیفی به این دلیل بود که اعتماد مدیران مالی به خروجی‌های الگوریتمی پدیده‌ای زمینه‌مند، تفسیری و وابسته به تجربه حرفه‌ای است و نمی‌توان آن را صرفاً از طریق شاخص‌های کمی یا پرسشنامه‌های استاندارد به‌طور عمیق فهم کرد. تحلیل مضمون به دلیل انعطاف‌پذیری نظری و توانایی آن در شناسایی الگوهای معنایی در داده‌های مصاحبه انتخاب شد. براون و کلارک تحلیل مضمون را روشی منعطف برای شناسایی، تحلیل و گزارش الگوها در داده‌های کیفی معرفی کرده‌اند. (Braun & Clarke, ۲۰۰۶)

مشارکت‌کنندگان پژوهش شامل ۲۰ نفر از مدیران و متخصصان مالی فعال در شهر تهران بودند. ترکیب نمونه شامل ۷ مدیر مالی شرکت‌های بزرگ و متوسط، ۴ مدیر سرمایه‌گذاری، ۳ تحلیلگر ارشد مالی، ۳ مدیر یا کارشناس ارشد ریسک، ۲ مدیر فناوری مالی و ۱ عضو کمیته سرمایه‌گذاری بود. معیارهای ورود به پژوهش عبارت بودند از: داشتن حداقل پنج سال سابقه فعالیت حرفه‌ای در حوزه مالی، مشارکت مستقیم یا غیرمستقیم در تصمیمات سرمایه‌گذاری، آشنایی عملی با سامانه‌های تحلیلی یا الگوریتمی، اشتغال در سازمان‌های مستقر در تهران و تمایل به شرکت در مصاحبه. معیار خروج نیز ناتمام گذاشتن مصاحبه یا عدم تمایل به ضبط گفت‌وگو بود. نمونه‌گیری به‌صورت هدفمند و با رعایت اصل تنوع حداکثری از نظر نوع سازمان، سابقه کاری و نقش شغلی انجام شد و در مراحل بعد از روش گلوله‌برفی برای دسترسی به مشارکت‌کنندگان متخصص‌تر استفاده شد.

گردآوری داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافته انجام شد. این روش به پژوهشگر امکان داد ضمن حفظ محورهای اصلی، مسیر مصاحبه را بر اساس تجربه و پاسخ‌های هر مشارکت‌کننده گسترش دهد. محورهای راهنمای مصاحبه شامل تجربه استفاده از ابزارهای الگوریتمی در تصمیمات مالی، برداشت مشارکت‌کننده از اعتماد به خروجی مدل، ویژگی‌های توضیح قابل قبول، تجربه مواجهه با خطای الگوریتم، نقش کنترل انسانی، ملاحظات ریسک، داده و پاسخگویی، و اثر توضیح‌پذیری بر تصمیمات سرمایه‌گذاری بود. نمونه پرسش‌ها عبارت بودند از: «چه زمانی به خروجی یک مدل تحلیلی یا الگوریتمی اعتماد می‌کنید؟»، «چه نوع توضیحی باعث می‌شود توصیه الگوریتم برای شما قابل دفاع باشد؟»، «اگر الگوریتم با قضاوت حرفه‌ای شما تعارض داشته باشد چه می‌کنید؟» و «چه عواملی باعث می‌شود خروجی مدل را در کمیته سرمایه‌گذاری قابل ارائه بدانید؟«

مصاحبه‌ها بین ۴۵ تا ۷۵ دقیقه طول کشید و با رضایت آگاهانه مشارکت‌کنندگان ضبط شد. تمامی مصاحبه‌ها به‌صورت کلمه‌به‌کلمه پیاده‌سازی شدند. برای رعایت ملاحظات اخلاقی، نام افراد و سازمان‌ها حذف و کدهایی مانند «مشارکت‌کننده ۱» تا «مشارکت‌کننده ۲۰» جایگزین شد. فرایند مصاحبه تا دستیابی به اشباع نظری ادامه یافت. اشباع اولیه در مصاحبه هجدهم حاصل شد و دو مصاحبه دیگر برای تأیید کفایت داده‌ها انجام گرفت. در پژوهش‌های کیفی، اشباع معمولاً زمانی رخ می‌دهد که داده‌های جدید مضمون یا بینش تازه‌ای به تحلیل اضافه نکنند؛ این مبنا در ادبیات روش‌شناسی کیفی نیز مورد تأکید قرار گرفته است. (Guest et al., ۲۰۰۶)

تحلیل داده‌ها با استفاده از نرم‌افزار NVivo انجام شد. ابتدا پژوهشگر چندین بار متن مصاحبه‌ها را مطالعه کرد تا با کلیت داده‌ها آشنا شود. سپس کدگذاری اولیه به‌صورت باز انجام گرفت و عبارات معنادار مرتبط با اعتماد، توضیح‌پذیری، تصمیم‌گیری سرمایه‌گذاری، کنترل انسانی و ریسک استخراج شد. در مرحله بعد، کدهای مشابه در خوشه‌های مفهومی قرار گرفتند و مضامین فرعی شکل گرفتند. سپس مضامین فرعی با یکدیگر مقایسه و در قالب مقوله‌های اصلی سازمان‌دهی شدند. در نهایت، پنج مقوله اصلی استخراج شد. برای افزایش اعتبار پژوهش از بازبینی مشارکت‌کنندگان، بررسی همتایان، نگهداری مسیر ممیزی، یادداشت‌برداری تحلیلی و مقایسه مستمر داده‌ها استفاده شد. معیارهای اعتبارپذیری، انتقال‌پذیری، اتکاپذیری و تأییدپذیری بر اساس رویکرد لینکلن و گوبا مدنظر قرار گرفت .(Lincoln & Guba, ۱۹۸۵) همچنین برای گزارش شفاف پژوهش، اصول کلی COREQ در زمینه گزارش مصاحبه‌ها و تحلیل کیفی مورد توجه قرار گرفت؛ COREQ یک چک‌لیست ۳۲موردی برای گزارش مطالعات کیفی مبتنی بر مصاحبه و گروه متمرکز است. (Tong et al., ۲۰۰۷)

یافته‌ها

از میان ۲۰ مشارکت‌کننده پژوهش، ۱۴ نفر مرد و ۶ نفر زن بودند. از نظر سطح تحصیلات، ۳ نفر دارای مدرک کارشناسی، ۱۳ نفر دارای مدرک کارشناسی ارشد و ۴ نفر دارای مدرک دکتری بودند. دامنه سنی مشارکت‌کنندگان بین ۳۲ تا ۵۶ سال بود؛ ۵ نفر در گروه سنی ۳۰ تا ۳۹ سال، ۱۰ نفر در گروه ۴۰ تا ۴۹ سال و ۵ نفر در گروه ۵۰ سال و بالاتر قرار داشتند. از نظر سابقه کاری، ۴ نفر بین ۵ تا ۱۰ سال، ۹ نفر بین ۱۱ تا ۲۰ سال و ۷ نفر بیش از ۲۰ سال تجربه حرفه‌ای داشتند. همچنین ۸ نفر در شرکت‌های سرمایه‌گذاری و هلدینگ‌ها، ۵ نفر در بانک‌ها و مؤسسات مالی، ۴ نفر در شرکت‌های تولیدی و خدماتی بزرگ و ۳ نفر در شرکت‌های فناوری مالی فعالیت داشتند. این ترکیب نشان داد که داده‌ها از طیفی متنوع از تجربه‌های مالی، سرمایه‌گذاری، ریسک و فناوری گردآوری شده است.

مقوله اول: کاهش ابهام شناختی از طریق فهم منطق الگوریتم .نخستین مقوله نشان داد که مدیران مالی زمانی به خروجی الگوریتم اعتماد بیشتری پیدا می‌کنند که بتوانند منطق کلی تولید خروجی را بفهمند. منظور آنان از فهم، دانستن جزئیات ریاضی یا کدنویسی مدل نبود، بلکه آگاهی از متغیرهای اثرگذار، جهت اثر متغیرها، مفروضات اصلی، منبع داده‌ها و حدود کاربرد مدل بود. مشارکت‌کنندگان تأکید کردند که خروجی‌هایی مانند «امتیاز ریسک»، «پیشنهاد خرید» یا «رتبه جذابیت سرمایه‌گذاری» بدون توضیح کافی، برای تصمیم‌گیری مدیریتی قابل اتکا نیستند. یکی از مدیران مالی بیان کرد: «وقتی سیستم فقط می‌گوید این پروژه ریسک بالاست، برای من کافی نیست. باید بدانم ریسک از کجا آمده؛ از جریان نقدی، بدهی شرکت، نوسان بازار یا داده‌ای که شاید اصلاً به‌روز نباشد» (مشارکت‌کننده ۷). مشارکت‌کننده دیگری گفت: «من لازم نیست فرمول مدل را خط‌به‌خط بدانم، اما باید بفهمم چرا نتیجه امروز با تحلیل ماه قبل فرق کرده است» (مشارکت‌کننده ۳). این مقوله نشان می‌دهد XAI با تبدیل خروجی خام الگوریتم به توضیح قابل فهم، ابهام شناختی مدیر را کاهش می‌دهد و امکان ورود خروجی مدل به فرایند تصمیم‌گیری را فراهم می‌کند.

مقوله دوم: تقویت اعتبارسنجی حرفه‌ای و کنترل انسانی .یافته‌ها نشان داد اعتماد مدیران مالی به الگوریتم زمانی افزایش می‌یابد که توضیح‌پذیری به آنان امکان اعتبارسنجی حرفه‌ای بدهد. مشارکت‌کنندگان نمی‌خواستند الگوریتم جایگزین قضاوت آنان شود؛ بلکه انتظار داشتند مدل به‌عنوان ابزار پشتیبان، فرضیه‌ها و تحلیل‌های آنان را تکمیل یا به چالش بکشد. آنان توضیح‌پذیری را ابزاری برای مقایسه میان تجربه انسانی و خروجی ماشینی می‌دانستند. یکی از مدیران سرمایه‌گذاری اظهار کرد: «اگر مدل بگوید سهام یک شرکت مناسب خرید است، من باید بتوانم ببینم این توصیه با چه فرضی درباره سودآوری، نقدشوندگی و ریسک صنعت ساخته شده. آن وقت می‌توانم تصمیم بگیرم قبول کنم یا نه» (مشارکت‌کننده ۱۱). یکی دیگر از مشارکت‌کنندگان گفت: «اعتماد یعنی بتوانم مدل را سؤال‌پیچ کنم، نه اینکه هر چه گفت اجرا کنم» (مشارکت‌کننده ۱۵). این مقوله نشان داد که اعتماد مطلوب از دیدگاه مدیران مالی، اعتماد کورکورانه نیست، بلکه اعتماد مبتنی بر امکان پرسش، بررسی، مقایسه و مداخله انسانی است.

مقوله سوم: افزایش پاسخگویی، شفافیت ریسک و حاکمیت داده .مشارکت‌کنندگان توضیح‌پذیری را با پاسخگویی سازمانی و مدیریت ریسک پیوند می‌دادند. از دید آنان، خروجی الگوریتمی زمانی قابل اعتماد است که داده‌های ورودی، کیفیت داده، زمان به‌روزرسانی، مالکیت مدل، مسئولیت خطا و سطح عدم‌قطعیت مشخص باشد. مدیران مالی نگران بودند که الگوریتم بر داده‌های ناقص، سوگیرانه یا قدیمی متکی باشد و تصمیمات سرمایه‌گذاری را دچار انحراف کند. یکی از مدیران ریسک گفت: «بزرگ‌ترین نگرانی من خود مدل نیست، داده‌ای است که به مدل داده شده. اگر داده مشکل داشته باشد، توضیح مدل هم ممکن است فقط خطا را شیک‌تر نشان دهد» (مشارکت‌کننده ۹). مشارکت‌کننده دیگری اظهار کرد» :در کمیته سرمایه‌گذاری وقتی پیشنهادی می‌دهم، باید بتوانم بگویم مبنای ریسک چیست و اگر بعداً خطا شد، چه کسی پاسخگوست» (مشارکت‌کننده ۲). این مقوله نشان داد XAI زمانی اعتمادآفرین است که با حاکمیت داده، ثبت مسیر تصمیم، مستندسازی مدل و شفاف‌سازی سطح ریسک همراه باشد.

مقوله چهارم: طراحی تعاملی و زمینه‌مند توضیحات الگوریتمی .یافته‌ها نشان داد مدیران مالی از توضیحاتی استقبال می‌کنند که متناسب با زمینه تصمیم و سطح دانش کاربر طراحی شده باشد. توضیح فنی بسیار پیچیده برای بسیاری از مدیران قابل استفاده نبود، در حالی که توضیح بیش از حد ساده نیز اعتماد ایجاد نمی‌کرد. مشارکت‌کنندگان خواهان داشبوردهایی بودند که امکان مشاهده سناریو، تغییر مفروضات، تحلیل حساسیت و مقایسه گزینه‌های سرمایه‌گذاری را فراهم کند. یکی از تحلیلگران ارشد مالی گفت: «اگر سیستم فقط یک نمودار اهمیت متغیرها بدهد، برای من کافی نیست. باید بتوانم سناریو بسازم؛ مثلاً ببینم اگر نرخ ارز یا نرخ بهره تغییر کند، پیشنهاد مدل چقدر عوض می‌شود» (مشارکت‌کننده ۱۸). مدیر دیگری بیان کرد: «توضیح خوب باید زبان مالی داشته باشد، نه فقط زبان دیتا ساینس» (مشارکت‌کننده ۶). این مقوله نشان می‌دهد طراحی XAI باید تعاملی، سناریومحور و قابل ترجمه به شاخص‌های مالی باشد تا بتواند در تصمیمات سرمایه‌گذاری استفاده شود.

مقوله پنجم: تبدیل اعتماد کورکورانه به اعتماد مشروط و قابل پایش .آخرین مقوله نشان داد که نقش اصلی XAI افزایش اعتماد مطلق نیست، بلکه ایجاد اعتماد مشروط، تدریجی و قابل پایش است. مشارکت‌کنندگان معتقد بودند اعتماد به الگوریتم باید بر اساس عملکرد گذشته، کیفیت توضیح، سازگاری با واقعیت بازار، امکان ممیزی و تجربه استفاده شکل گیرد. آنان بیان کردند که توضیح‌پذیری می‌تواند هم اعتماد را افزایش دهد و هم در مواردی باعث کاهش اعتماد ناموجه شود؛ زیرا مدیر با دیدن منطق مدل ممکن است متوجه ضعف داده یا فرض نادرست شود. یکی از مشارکت‌کنندگان گفت: «گاهی توضیح مدل باعث می‌شود بیشتر اعتماد کنم، گاهی هم برعکس؛ می‌فهمم مدل چیزی را مهم گرفته که از نظر بازار ما اصلاً قابل اتکا نیست» (مشارکت‌کننده ۱۳). مشارکت‌کننده دیگری اظهار کرد: «برای من مدل قابل اعتماد مدلی است که هم دلیل بیاورد، هم خطای خودش را نشان دهد» (مشارکت‌کننده ۲۰). بنابراین، XAIاعتماد را از حالت احساسی یا تقلیدی خارج کرده و به فرایندی تحلیلی، مشروط و قابل بازبینی تبدیل می‌کند.

بحث

یافته نخست پژوهش نشان داد هوش مصنوعی توضیح‌پذیر از طریق کاهش ابهام شناختی، زمینه اعتماد مدیران مالی به خروجی‌های الگوریتمی را فراهم می‌کند. مشارکت‌کنندگان تأکید داشتند که برای پذیرش خروجی مدل، دانستن چرایی و چگونگی کلی تصمیم ضروری است. این یافته با ادبیات XAI همسوست که بیان می‌کند مدل‌های پیچیده یادگیری ماشین به‌رغم دقت بالا، در صورت نبود توضیح، با مانع پذیرش مواجه می‌شوند. آریتا و همکاران توضیح‌پذیری را یکی از شروط اساسی استقرار مسئولانه هوش مصنوعی در سازمان‌ها می‌دانند و آن را با مفاهیمی مانند شفافیت، پاسخگویی، فهم‌پذیری و انصاف پیوند می‌دهند . (Barredo Arrieta et al., ۲۰۲۰) همچنین ریبیرو و همکاران در معرفی LIME تصریح کردند که فهم دلایل پیش‌بینی برای ارزیابی اعتماد به مدل ضروری است . (Ribeiro et al., ۲۰۱۶) در زمینه تصمیمات سرمایه‌گذاری، این مسئله اهمیت مضاعف دارد؛ زیرا مدیر مالی باید بتواند توصیه الگوریتم را نه‌تنها برای خود، بلکه برای کمیته سرمایه‌گذاری، هیئت‌مدیره و سایر ذی‌نفعان توجیه کند. بنابراین، توضیح‌پذیری با کاهش فاصله شناختی میان مدل و مدیر، خروجی الگوریتم را از یک عدد مبهم به یک گزاره قابل بحث تبدیل می‌کند.

یافته دوم نشان داد که XAI اعتماد را از مسیر تقویت اعتبارسنجی حرفه‌ای و کنترل انسانی افزایش می‌دهد. مدیران مالی الگوریتم را جایگزین قضاوت حرفه‌ای نمی‌دانستند، بلکه آن را منبعی برای تکمیل، آزمون یا به چالش کشیدن تحلیل انسانی تلقی می‌کردند. این یافته با مطالعات مربوط به اعتماد کالیبره‌شده سازگار است. بن‌دیوید و همکاران در مطالعه تجربی خود درباره مشاوران الگوریتمی مالی نشان دادند که توضیحات الگوریتمی می‌تواند پذیرش و اعتماد کاربران را افزایش دهد، به‌ویژه زمانی که مدل با خطا یا شکست مواجه می‌شود (Ben David et al., ۲۰۲۱) . این نتیجه نشان می‌دهد توضیح‌پذیری فقط در شرایط عملکرد کامل مدل اهمیت ندارد، بلکه در لحظات ابهام و خطا نقش مهم‌تری دارد. یافته حاضر همچنین با مفهوم گریز از الگوریتم در پژوهش دیتورست و همکاران قابل تبیین است؛ زیرا کاربران پس از مشاهده خطای

الگوریتم ممکن است به‌طور نامتناسب از آن فاصله بگیرند XAI .(Dietvorst et al., ۲۰۱۵) می‌تواند این گریز را کاهش دهد، مشروط بر آنکه به مدیر امکان دهد بفهمد خطا چرا رخ داده، آیا قابل اصلاح است و در چه شرایطی مدل همچنان قابل استفاده است.

یافته سوم نشان داد که اعتماد به خروجی‌های الگوریتمی در تصمیمات مالی بدون حاکمیت داده، شفافیت ریسک و پاسخگویی سازمانی پایدار نیست. مشارکت‌کنندگان تأکید کردند که کیفیت داده، منبع داده، به‌روزرسانی اطلاعات، مسئولیت خطا و امکان ممیزی، بخش جدایی‌ناپذیر اعتماد به مدل است. این یافته با رویکردهای حکمرانی هوش مصنوعی همسوست. اصول OECD بر ارائه اطلاعات معنادار درباره قابلیت‌ها، محدودیت‌ها، داده‌ها، عوامل و منطق مؤثر بر خروجی تأکید دارد و همچنین پاسخگویی و قابلیت ردیابی را از شروط سامانه‌های هوشمند قابل اعتماد می‌داند . گزارش اداره بانکداری اروپا نیز عناصر اعتماد در کاربردهای کلان‌داده و تحلیل پیشرفته را شامل اخلاق، امنیت، حکمرانی و روش‌شناسی تحلیلی معرفی می‌کند . بنابراین، توضیح‌پذیری اگر از سازوکارهای حاکمیت داده جدا شود، ممکن است به «ظاهر شفافیت» تبدیل شود؛ یعنی مدل توضیحی ارائه کند اما کیفیت داده یا مسئولیت تصمیم همچنان مبهم بماند. در مدیریت مالی، چنین شفافیت ظاهری می‌تواند خطرناک باشد، زیرا ممکن است اعتماد بیش از حد به خروجی‌هایی ایجاد کند که بر داده‌های ناقص یا نادرست بنا شده‌اند.

یافته چهارم بر ضرورت طراحی تعاملی و زمینه‌مند توضیحات الگوریتمی تأکید داشت. مدیران مالی توضیحاتی را ارزشمند می‌دانستند که با زبان تصمیم‌گیری مالی سازگار باشد، امکان تحلیل سناریو فراهم کند و اثر تغییر مفروضات را نشان دهد. این یافته با دیدگاه میلر درباره ماهیت انسانی توضیح همخوان است؛ زیرا توضیح مؤثر باید به نیازها، انتظارات و پرسش‌های مخاطب پاسخ دهد، نه اینکه صرفاً بازنمایی فنی مدل باشد . (Miller, ۲۰۱۹) همچنین لانگر و همکاران تأکید کرده‌اند که رویکردهای XAI باید بر اساس نیازهای ذی‌نفعان و زمینه استفاده طراحی شوند، زیرا گروه‌های مختلف از توضیح‌پذیری انتظارات متفاوتی دارند . (Langer et al., ۲۰۲۱) در تصمیمات سرمایه‌گذاری، این به معنای آن است که توضیح باید به پرسش‌هایی مانند «چرا این گزینه؟»، «چرا اکنون؟»، «در چه سناریویی توصیه تغییر می‌کند؟» و «چه ریسکی بیشترین اثر را دارد؟» پاسخ دهد. بنابراین، طراحی داشبوردهای توضیح‌پذیر باید از حالت گزارش فنی ثابت به سمت تعامل، مقایسه، سناریوسازی و ترجمه مالی حرکت کند.

یافته پنجم نشان داد که XAI باید اعتماد کورکورانه را به اعتماد مشروط و قابل پایش تبدیل کند. مشارکت‌کنندگان بیان کردند که توضیح‌پذیری گاهی اعتماد را افزایش و گاهی کاهش می‌دهد؛ زیرا ممکن است نقاط ضعف مدل را آشکار کند. این یافته از نظر نظری بسیار مهم است، زیرا نشان می‌دهد هدف XAI نباید صرفاً افزایش پذیرش الگوریتم باشد، بلکه باید به تنظیم صحیح سطح اعتماد کمک کند. پژوهش پورسبزی-سنگده و همکاران نشان داد که شفافیت ظاهری و کاهش تعداد ویژگی‌ها لزوماً به پیروی بهتر از مدل منجر نمی‌شود و حتی ممکن است تشخیص خطا را دشوارتر کند . (Poursabzi-Sangdeh et al., ۲۰۲۱) همچنین کائور و همکاران نشان دادند که کاربران حرفه‌ای نیز ممکن است ابزارهای تفسیرپذیری را نادرست درک یا بیش از حد به آن‌ها اعتماد کنند . (Kaur et al., ۲۰۲۰) بنابراین، توضیح‌پذیری باید

همراه با آموزش، ممیزی، پایش عملکرد، شاخص‌های عدم‌قطعیت و سازوکارهای هشدار باشد تا اعتماد مدیران مالی به سطحی متناسب با شواهد و محدودیت‌های مدل برسد.

در مجموع، نتایج پژوهش نشان می‌دهد که نقش XAI در تصمیمات سرمایه‌گذاری صرفاً فنی نیست، بلکه ماهیتی مدیریتی، شناختی و حکمرانی دارد. توضیح‌پذیری زمانی اعتماد ایجاد می‌کند که سه شکاف را هم‌زمان کاهش دهد: شکاف شناختی میان مدل و مدیر، شکاف سازمانی میان خروجی الگوریتم و پاسخگویی تصمیم، و شکاف عملیاتی میان پیش‌بینی آماری و اقدام سرمایه‌گذاری. این نتیجه با تحولات نظارتی جهانی نیز همسوست. قانون هوش مصنوعی اتحادیه اروپا بر رویکرد مبتنی بر ریسک و ضرورت اعتمادپذیری سامانه‌های هوشمند تأکید دارد و تصریح می‌کند که در برخی موارد، فهم چرایی تصمیم یا پیش‌بینی سامانه هوش مصنوعی دشوار است و همین امر ارزیابی پیامدهای ناعادلانه را پیچیده می‌کند . از این منظر، مدیران مالی نه‌تنها کاربران خروجی الگوریتمی، بلکه حاملان مسئولیت حرفه‌ای و سازمانی تصمیم هستند. بنابراین، استقرار XAI باید با سیاست‌های داخلی حاکمیت مدل، مستندسازی، تفکیک مسئولیت، آموزش مدیران و سازوکارهای نظارت انسانی همراه شود.

محدودیت‌ها :پژوهش حاضر با رویکرد کیفی و بر اساس تجربه ۲۰ نفر از مدیران و متخصصان مالی شهر تهران انجام شد؛ بنابراین، هدف آن تعمیم آماری نتایج به همه مدیران مالی نبود، بلکه تبیین عمیق تجربه و برداشت مشارکت‌کنندگان بود. همچنین داده‌ها بر اساس خوداظهاری مشارکت‌کنندگان گردآوری شد و ممکن است تحت تأثیر برداشت ذهنی، تجربه سازمانی یا سطح آشنایی آنان با فناوری‌های هوش مصنوعی قرار گرفته باشد. محدودیت دیگر آن است که پژوهش حاضر به‌طور مستقیم عملکرد واقعی مدل‌های XAI را در محیط عملیاتی سرمایه‌گذاری ارزیابی نکرد، بلکه بر ادراک و تجربه مدیران مالی تمرکز داشت.

پیشنهادها برای پژوهش‌های آینده :پیشنهاد می‌شود پژوهش‌های بعدی با روش آمیخته انجام شوند و یافته‌های کیفی حاضر در قالب پرسشنامه و مدل مفهومی آزمون شوند. همچنین می‌توان نقش متغیرهایی مانند سواد داده‌ای مدیران مالی، فرهنگ دیجیتال سازمانی، نوع صنعت، سطح ریسک سرمایه‌گذاری و تجربه قبلی شکست الگوریتمی را در رابطه میان XAI و اعتماد الگوریتمی بررسی کرد. انجام مطالعات مقایسه‌ای میان بانک‌ها، شرکت‌های سرمایه‌گذاری، شرکت‌های تولیدی و استارتاپ‌های مالی نیز می‌تواند تفاوت‌های زمینه‌ای را آشکار سازد. علاوه بر این، پژوهش‌های آزمایشی می‌توانند اثر انواع مختلف توضیح، مانند توضیح مبتنی بر ویژگی، سناریو، مقایسه خلاف‌واقع و شاخص عدم‌قطعیت را بر تصمیمات واقعی مدیران مالی بررسی کنند.

پیشنهادها برای عمل :سازمان‌ها برای استفاده مؤثر از هوش مصنوعی توضیح‌پذیر در تصمیمات سرمایه‌گذاری باید داشبوردهای قابل فهم، سناریومحور و متناسب با زبان مالی طراحی کنند. همچنین لازم است کمیته‌های سرمایه‌گذاری از سیاست‌های روشن درباره مالکیت مدل، کیفیت داده، سطح اختیار الگوریتم، مسئولیت خطا و شرایط مداخله انسانی برخوردار باشند. آموزش مدیران مالی در زمینه مفاهیم پایه هوش

مصنوعی، محدودیت‌های مدل، تفسیر خروجی‌های SHAP و LIME و تشخیص اعتماد بیش از حد می‌تواند کیفیت تصمیم‌گیری را افزایش دهد. در نهایت، XAIباید به‌عنوان بخشی از نظام حاکمیت ریسک و نه صرفاً ابزار فنی واحد فناوری اطلاعات در سازمان پیاده‌سازی شود.

نتیجه‌گیری

پژوهش حاضر با هدف تبیین نقش هوش مصنوعی توضیح‌پذیر در افزایش اعتماد مدیران مالی به خروجی‌های الگوریتمی و تصمیمات سرمایه‌گذاری انجام شد. یافته‌ها نشان داد توضیح‌پذیری از طریق کاهش ابهام شناختی، امکان اعتبارسنجی حرفه‌ای، تقویت کنترل انسانی، شفافیت ریسک، حاکمیت داده، طراحی زمینه‌مند توضیحات و ایجاد اعتماد مشروط می‌تواند زمینه پذیرش مسئولانه خروجی‌های الگوریتمی را فراهم کند. مدیران مالی زمانی به الگوریتم اعتماد می‌کنند که خروجی آن قابل فهم، قابل پرسش، قابل ممیزی و قابل ترجمه به پیامدهای مالی باشد. بنابراین، اعتماد به هوش مصنوعی در تصمیمات سرمایه‌گذاری نه از طریق پنهان کردن پیچیدگی مدل، بلکه از طریق آشکارسازی منطق، محدودیت‌ها و شرایط کاربرد آن شکل می‌گیرد. نتیجه اصلی پژوهش آن است که XAI باید به‌عنوان پل ارتباطی میان تحلیل داده‌محور و قضاوت حرفه‌ای مدیران مالی در نظر گرفته شود. چنین رویکردی می‌تواند تصمیمات سرمایه‌گذاری را از دوگانه خطرناک «اتکای کامل به الگوریتم» یا «رد کامل الگوریتم» خارج کرده و به سمت تصمیم‌گیری تلفیقی، پاسخگو و مبتنی بر شواهد هدایت کند.

تشکر و قدردانی

از تمامی کسانی که در انجام این پژوهش ما را همراهی کردند تشکر و قدردانی به عمل می‌آید.

موازین اخلاقی

در این پژوهش اصول اخلاقی مرتبط با نشر و انجام پژوهش رعایت گردیده است.

مشارکت نویسندگان

تمامی نویسندگان نقش یکسانی ایفا کردند.

تعارض منافع

بر اساس اظهار نویسندگان، در این مقاله هیچ‌گونه تضاد منافعی وجود ندارد.

فهرست منابع

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Bank of England, & Financial Conduct Authority. (2024). *Artificial intelligence in UK financial services – 2024*. Bank of England.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Ben David, D., Resheff, Y. S., & Tron, T. (2021). Explainable AI and adoption of financial algorithmic advisors: An experimental study. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. https://doi.org/10.48550/arXiv.2101.02555

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114–126. https://doi.org/10.1037/xge0000033

European Banking Authority. (2020). *EBA report on big data and advanced analytics*. European Banking Authority.

European Commission. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence*. European Union.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2019). A survey of methods for explaining black box models. *ACM Computing Surveys, 51*(5), 1–42. https://doi.org/10.1145/3236009

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H., & Wortman Vaughan, J. (2020). Interpreting interpretability: Understanding data scientists' use of interpretability tools for machine learning. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3313831.3376219

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from explainable artificial intelligence? A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence, 296*, 103473. https://doi.org/10.1016/j.artint.2021.103473

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

OECD. (2024). *OECD AI principles*. Organisation for Economic Co-operation and Development.

Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Wortman Vaughan, J., & Wallach, H. (2021). Manipulating and measuring model interpretability. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–52. https://doi.org/10.1145/3411764.3445315

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research: A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349–357. https://doi.org/10.1093/intqhc/mzm042