

## Transparency and algorithmic bias on the acceptance of AI-based fraud detection systems with trust mediation using a structural equation approach

Parastoo Bafghi

Department of Business Management, ShQ.C.,  
Islamic Azad University, Shahr-e Qods, Iran.

Mahmoud Ahmadi Sharif\*

Department of Business Management, ShQ.C.,  
Islamic Azad University, Shahr-e Qods, Iran.

Morteza shafati

Assistant Professor , Department of accounting  
and Management , Payame Noor  
University, Tehran, Iran.

### Abstract

In recent years, the use of AI systems in the field of financial and banking services, especially in the field of fraud detection, has been considered as one of the key tools in increasing security and reducing financial risk. However, the acceptance of these new technologies by bank employees and customers requires the provision of prerequisites such as transparency of algorithm performance and reducing their possible biases. The present study aimed to investigate the effect of transparency and algorithmic bias on the acceptance of AI-based fraud detection systems with the role of trust mediation and using a structural equation approach. This study is applied in terms of purpose and descriptive-correlational in terms of method. The statistical population of the study included: 653 employees and experts of Bank Melli branches were selected in Karaj. The minimum sample size was estimated to be 242 people according to the Cochran formula and was analyzed. In this study, a standard questionnaire was used to collect data. Reliability was examined using Cronbach's alpha coefficient, and face and content validity were examined and confirmed by the supervisor and experts. The data obtained were also analyzed using SPSS and lisrel statistical software. The findings of the study showed that algorithmic transparency has a positive and significant effect on trust and also on the acceptance of fraud detection systems. In addition, the results indicate that the lack of algorithmic bias also directly affects the level of employee trust in the system, thereby increasing the acceptance of artificial intelligence technology. The mediating role of trust in this relationship was confirmed and it was determined that trust, as a mediator, strengthens the intensity of the effect of transparency and impartiality of algorithms on the acceptance of artificial intelligence-based systems. Based on the Sobel test and model fit indices, the desired relationships had appropriate statistical validity.

**Keywords:** Transparency, algorithmic unbiasedness, acceptance of AI-based fraud detection systems, trust.

**How to Cite:** Bafghi, P. , Ahmadi Sharif, M. and Shafati, M. (2025). Transparency and algorithmic bias on the acceptance of AI-based fraud detection systems with trust mediation using a structural equation approach.. Journal of Intelligent Strategic Management .4(4), 313-346.  
doi: 10.87453/bumara.2026.373601.4851



Intelligent Strategic Management (JISM) in Development and Evolution is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.

© Authors

\* Corresponding Author : m.ahmadisharif@godsiau.ac.ir

## شفافیت و عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی با میانجی‌گیری اعتماد با رویکرد معادلات ساختاری

پرستو بافتی

گروه مدیریت بازرگانی، واحد شهر قدس، دانشگاه آزاد اسلامی، شهر قدس، ایران.

محمود احمدی  
شریف\*

گروه مدیریت بازرگانی، واحد شهر قدس، دانشگاه آزاد اسلامی، شهر قدس، ایران.

مرتضی شفعتی

استادیار، گروه مدیریت و حسابداری، دانشگاه پیام نور، تهران، ایران.

### چکیده

در سال‌های اخیر، استفاده از سیستم‌های هوش مصنوعی در حوزه خدمات مالی و بانکی به‌ویژه در زمینه تشخیص تقلب، به‌عنوان یکی از ابزارهای کلیدی در افزایش امنیت و کاهش ریسک مالی مورد توجه قرار گرفته است. با این حال، پذیرش این فناوری‌های نوین از سوی کارکنان و مشتریان بانک‌ها مستلزم فراهم شدن پیش‌شرط‌هایی همچون شفافیت عملکرد الگوریتم‌ها و کاهش سوگیری‌های احتمالی آن‌هاست. پژوهش حاضر با هدف بررسی تأثیر شفافیت و عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی با نقش میانجی اعتماد و با بهره‌گیری از رویکرد معادلات ساختاری انجام شد. این پژوهش از لحاظ هدف کاربردی و از نظر روش، توصیفی - همبستگی است. جامعه آماری پژوهش شامل: کارکنان و کارشناسان شعب بانک ملی به تعداد ۶۵۳ نفر در شهر کرج انتخاب شدند. حداقل حجم نمونه طبق فرمول کوکران ۲۴۲ نفر برآورد گردید و مورد تجزیه و تحلیل قرار گرفت. در این پژوهش برای گردآوری اطلاعات از پرسشنامه استاندارد استفاده شده است. پایایی با استفاده از ضریب الفای کرونباخ مورد بررسی قرار گرفت و روایی صوری و محتوا با نظر استاد راهنما و خبرگان مورد بررسی و تأیید قرار گرفت. همچنین تجزیه و تحلیل داده‌های به‌دست آمده با استفاده از نرم‌افزار آماری SPSS و LISREL صورت گرفته است. یافته‌های پژوهش نشان داد که شفافیت الگوریتمی تأثیر مثبت و معناداری بر اعتماد و همچنین بر پذیرش سیستم‌های تشخیص تقلب دارد. علاوه بر این، نتایج بیانگر آن است که عدم سوگیری الگوریتمی نیز به‌طور مستقیم بر سطح اعتماد کارکنان نسبت به سیستم اثرگذار بوده و از این طریق پذیرش فناوری هوش مصنوعی را افزایش می‌دهد. نقش میانجی اعتماد در این رابطه تأیید شد و مشخص گردید که اعتماد به‌عنوان حلقه واسط، شدت اثرگذاری شفافیت و بی‌طرفی الگوریتم‌ها را بر پذیرش سیستم‌های مبتنی بر هوش مصنوعی تقویت می‌کند. بر اساس آزمون سوبیل و شاخص‌های برازش مدل، روابط مورد نظر از اعتبار آماری مناسبی برخوردار بودند.

**کلیدواژه‌ها:** شفافیت، عدم سوگیری الگوریتمی، پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی، اعتماد

**استناد به این مقاله:** بافتی، پرستو و احمدی شریف، محمود و شفعتی، مرتضی. (۱۴۰۴). شفافیت و عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی با میانجی‌گیری اعتماد با رویکرد معادلات ساختاری. مدیریت استراتژیک هوشمند، ۴(۴)، ۳۴۶-۳۱۳.



مدیریت استراتژیک هوشمند (JISM) در توسعه و تکامل تحت مجوز بین‌المللی کرییتیو کامنز با شرایط انتساب-غیرتجاری ۴٫۰ منتشر می‌شود.

©نویسندگان

## مقدمه

در عصر تحول دیجیتال و گسترش روزافزون فناوری‌های هوشمند، هوش مصنوعی به یکی از ارکان کلیدی در طراحی و پیاده‌سازی سیستم‌های پیشرفته در حوزه‌های مختلف تبدیل شده است (یاسین<sup>۱</sup> و همکاران، ۲۰۲۵). یکی از کاربردهای حیاتی و رو به رشد این فناوری، در سیستم‌های تشخیص تقلب نمود یافته است. این سیستم‌ها با تحلیل داده‌ها، به صورت بلادرنگ الگوهای غیرمعمول را شناسایی کرده و به کشف فعالیت‌های مشکوک در زمینه‌های مالی، بانکی، بیمه و تجارت الکترونیک کمک می‌کنند (اسلام<sup>۲</sup> و همکاران، ۲۰۲۳)، با این حال، با وجود مزایای بالقوه این ابزارها، یکی از چالش‌های جدی پیش‌روی استفاده و پذیرش آن‌ها، مربوط به مسائل مرتبط با شفافیت الگوریتمی و سوگیری‌های الگوریتمی است که می‌تواند به شکل معناداری بر اعتماد کاربران و در نتیجه پذیرش نهایی این سیستم‌ها تأثیرگذار باشد (دایلان<sup>۳</sup> و همکاران، ۲۰۲۵). در واقع، تصمیم‌گیری‌های هوش مصنوعی، به‌ویژه در حوزه‌هایی که پیامدهای حساس مالی، حقوقی یا اجتماعی دارند، نیازمند سطح بالایی از شفافیت و عدالت هستند؛ در صورتی که کاربران نهایی (از جمله: کارشناسان ضد تقلب، تحلیل‌گران داده، مدیران فناوری اطلاعات، و مشتریان نهایی) احساس کنند که تصمیمات سیستم بر پایه منطق نامشخص، داده‌های ناقص، یا الگوریتم‌های دارای سوگیری اتخاذ می‌شود، احتمال بی‌اعتمادی و مقاومت در برابر استفاده از این فناوری افزایش می‌یابد. این در حالی است که عدم پذیرش یا استفاده ناکارآمد از چنین سیستم‌هایی می‌تواند موجب شکست در کشف تقلب و تحمل زیان‌های هنگفت برای سازمان‌ها شود (بصری<sup>۴</sup> و همکاران، ۲۰۲۳). همچنین شفافیت الگوریتمی اشاره به میزان توانایی کاربران در درک فرآیند تصمیم‌گیری سیستم‌های هوش مصنوعی دارد. در سیستم‌های سنتی، فرآیند تصمیم‌گیری معمولاً مبتنی بر قواعد مشخص و قابل پیگیری است، اما در بسیاری از مدل‌های پیچیده یادگیری ماشین و یادگیری عمیق، این فرآیند به صورت «جعبه سیاه» عمل می‌کند که درک عملکرد درونی آن برای کاربران و حتی طراحان آن نیز دشوار است (احمدی<sup>۵</sup> و همکاران، ۲۰۲۲). این عدم شفافیت می‌تواند به بروز حس عدم اطمینان، ترس از اشتباهات سیستم و حتی نگرانی‌های اخلاقی

1 Yaseen

2 Islam

3 Dayalan

4 Basri

5 Ahmadi

منجر شود، در چنین شرایطی، کاربران ممکن است دچار تردید شوند که آیا سیستم واقعاً قادر به تشخیص صحیح تقلب است یا خیر و در نتیجه سطح اعتماد آنان کاهش می‌یابد. از سوی دیگر، سوگیری الگوریتمی که ناشی از داده‌های آموزشی مغرضانه یا طراحی ناعادلانه الگوریتم‌ها است، می‌تواند به نتایجی منجر شود که ناعادلانه، تبعیض‌آمیز یا نادرست باشند. برای مثال، اگر داده‌های آموزشی سیستم تنها از یک جامعه خاص جمع‌آوری شده باشد یا شامل تعصب‌های انسانی باشد، مدل آموزش‌دیده نیز این سوگیری‌ها را بازتولید خواهد کرد. چنین سوگیری‌هایی ممکن است منجر به شناسایی نادرست افراد یا تراکنش‌های سالم به‌عنوان تقلبی شود و اعتماد کاربران را به‌صورت جدی زیر سؤال ببرد، بنابراین، پرداختن به سوگیری در طراحی و توسعه سیستم‌های مبتنی بر هوش مصنوعی یک دغدغه اخلاقی و فنی محسوب می‌شود (اوزیکو و همکاران، ۲۰۲۴).

در این میان، اعتماد به سیستم به‌عنوان یکی از سازه‌های کلیدی رفتاری و روان‌شناختی در پذیرش فناوری‌های نوین، نقش میانجی حیاتی ایفا می‌کند. ادبیات نظری موجود، اعتماد را به‌عنوان عاملی تسهیل‌کننده در مواجهه با فناوری‌های پیچیده و نامطمئن معرفی می‌کند که می‌تواند اثر مثبت شفافیت و بی‌طرفی سیستم را تقویت کرده و مقاومت کاربران در برابر فناوری‌های جدید را کاهش دهد (آیینی و همکاران، ۲۰۲۴). کاربران زمانی که به سیستم اعتماد می‌کنند، حتی در شرایطی که به‌طور کامل عملکرد درونی آن‌را درک نکنند، تمایل بیشتری به استفاده از آن دارند، برعکس، فقدان اعتماد می‌تواند حتی در صورت عملکرد مطلوب فنی، منجر به رد سیستم شود. بنابراین، شناخت دقیق نقش واسطه‌ای اعتماد، می‌تواند به تبیین بهتر ساز و کارهای روانی و رفتاری در مسیر پذیرش سیستم‌های تشخیص تقلب کمک کند (عبدالرحمان و همکاران، ۲۰۱۹).

با توجه به رشد روزافزون استفاده از فناوری‌های مبتنی بر هوش مصنوعی در سازمان‌های مالی، بیمه‌ای، بانک‌ها و شرکت‌های فین‌تک، درک عمیق از عوامل تأثیرگذار بر پذیرش این فناوری‌ها بیش از پیش اهمیت یافته است، در حالی که مطالعات متعددی به بررسی عوامل فنی و اقتصادی پذیرش فناوری پرداخته‌اند، بررسی ابعاد رفتاری و ادراکی مانند: شفافیت، سوگیری و اعتماد هنوز به‌صورت جامع و ترکیبی کمتر مورد توجه قرار گرفته است؛ به‌ویژه در زمینه سیستم‌های تشخیص تقلب که با اطلاعات حساس و پیامدهای حیاتی

---

1 Ozioko

2 Ayeni

3 Abdulrahman

سروکار دارند. شناسایی الگوهای پذیرش رفتاری می‌تواند به توسعه راهکارهای اثربخش جهت طراحی سیستم‌های قابل اعتمادتر و مقبول‌تر منجر شود و با گسترش کاربرد هوش مصنوعی در صنعت بانکداری، سیستم‌های تشخیص تقلب به ابزارهای کلیدی برای کاهش ریسک‌های مالی و ارتقای امنیت تبدیل شده‌اند. با این حال، پذیرش این سیستم‌ها از سوی کارکنان بانکی به‌ویژه در محیط‌های عملیاتی مانند، شعب بانک، همچنان با چالش‌هایی مواجه است. دو عامل اساسی که می‌توانند بر نگرش کاربران نسبت به این سیستم‌ها تأثیرگذار باشند، شفافیت الگوریتمی و عدم سوگیری آن‌هاست؛ چراکه عملکرد نامفهوم یا متعصبانه الگوریتم‌ها ممکن است به کاهش اعتماد کاربران منجر شده و پذیرش فناوری را تضعیف کند. در این میان، اعتماد به‌عنوان یک متغیر میانجی، نقشی حیاتی در انتقال اثرات ویژگی‌های فنی سیستم به سطح ادراک و رفتار کاربران دارد. بنابراین، به‌دلیل شکاف تحقیقاتی و نیاز درک بهتر این موضوع، تحقیق حاضر با هدف پاسخگویی به این سؤال شکل گرفت:

- تأثیر شفافیت و عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی با میانجی‌گیری اعتماد با رویکرد معادلات ساختاری در شعب بانک ملی شهر کرج چگونه است؟

## ادبیات نظری پژوهش

### هوش مصنوعی

هوش مصنوعی<sup>۱</sup> (AI) شاخه‌ای از علوم کامپیوتر است که به توسعه سیستم‌ها و الگوریتم‌هایی می‌پردازد که قادر به انجام وظایفی هستند که معمولاً نیازمند هوش انسانی می‌باشند. این وظایف شامل: یادگیری، استدلال، تصمیم‌گیری، پردازش زبان طبیعی و تشخیص الگوها است (گودفلو، بنگیو و کورویل<sup>۲</sup>، ۲۰۱۸). هوش مصنوعی با استفاده از روش‌هایی مانند: یادگیری ماشین<sup>۳</sup>، یادگیری عمیق<sup>۴</sup> و شبکه‌های عصبی، به سیستم‌ها امکان می‌دهد از داده‌ها بیاموزند و عملکرد خود را بهبود بخشند (لکان، بنگیو و هیتون<sup>۵</sup>، ۲۰۲۰). در سال‌های اخیر، با افزایش چشمگیر داده‌ها و پیشرفت‌های سخت‌افزاری، هوش مصنوعی به حوزه‌ای کلیدی در صنایع مختلف تبدیل شده است. یکی از کاربردهای مهم هوش

1 Artificial Intelligence

2 Goodfellow, Bengio, & Courville

3 Machine Learning

4 Deep Learning

5 LeCun, Bengio, & Hinton

مصنوعی، توسعه سیستم‌های تشخیص تقلب است که نقش حیاتی در حوزه‌های مالی، بانکی، بیمه و تجارت الکترونیک دارد (نگای، هو، وانگ، چن، ۲۰۲۱).

### سیستم‌های تشخیص تقلب

تقلب به عنوان هرگونه رفتار نادرست یا فریبکارانه که منجر به زیان مالی یا اعتباری شود، در سیستم‌های مالی و تجاری رایج است. سیستم‌های تشخیص تقلب به منظور شناسایی و پیشگیری از این رفتارها طراحی شده‌اند و هوش مصنوعی نقش کلیدی در افزایش دقت و سرعت این سیستم‌ها دارد. روش‌های سنتی تشخیص تقلب عمدتاً مبتنی بر قواعد و تحلیل‌های آماری ساده بودند که توانایی مقابله با حجم بالای داده‌ها و الگوهای پیچیده را نداشتند، اما با به کارگیری الگوریتم‌های یادگیری ماشین و یادگیری عمیق، سیستم‌های تشخیص تقلب توانسته‌اند الگوهای پنهان و تغییرات رفتاری در داده‌ها را به صورت خودکار شناسایی کنند (جورگوسکی و همکاران<sup>۳</sup>، ۲۰۱۸)؛ برای مثال، الگوریتم‌های جنگل تصادفی<sup>۴</sup>، ماشین بردار پشتیبان و شبکه‌های عصبی عمیق<sup>۵</sup> در تحلیل تراکش‌های مالی برای تشخیص تقلب کاربرد فراوان دارند، همچنین، الگوریتم‌های یادگیری تقویتی و شبکه‌های گراف برای مدل‌سازی ارتباطات پیچیده بین عوامل مختلف در تشخیص تقلب مؤثر واقع شده‌اند (ژانگ و همکاران<sup>۶</sup>، ۲۰۲۰). یکی از چالش‌های اصلی در سیستم‌های تشخیص تقلب، نابرابری داده‌ها است؛ یعنی داده‌های تقلبی نسبت به داده‌های سالم بسیار کمتر هستند که باعث دشواری در آموزش مدل‌های یادگیری ماشین می‌شود، به همین دلیل، تکنیک‌هایی مانند: نمونه‌برداری مجدد<sup>۷</sup>، الگوریتم‌های تشخیص ناهنجاری<sup>۸</sup> و یادگیری نیمه‌نظارتی<sup>۹</sup> به کار گرفته می‌شوند (بانسن، آثوآدا، استویانوویچ و اوترستن، ۲۰۲۰)؛ در نهایت، سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی به دلیل توانایی بالای خود در شناسایی الگوهای پیچیده و یادگیری خودکار، ابزارهای حیاتی برای کاهش خسارات مالی و افزایش امنیت سیستم‌های مالی محسوب می‌شوند. با این حال، نیاز به به‌روزرسانی مداوم

---

1 Ngai, Hu, Wong, Chen  
 2 Fraud Detection Systems  
 3 Jurgovsky et al  
 4 Random Forest  
 5 Deep Neural Networks  
 6 Zhang et al  
 7 Resampling  
 8 Anomaly Detection  
 9 Semi-supervised Learning

و تحلیل داده‌های جدید برای مقابله با روش‌های جدید تقلب از اهمیت بالایی برخوردار است (نگوین، تران، نگو و فونگ، ۲۰۲۱).

### مفهوم شفافیت الگوریتمی

شفافیت الگوریتمی<sup>۲</sup> به معنای قابلیت مشاهده، درک و توضیح عملکرد و تصمیمات الگوریتم‌های کامپیوتری است. با پیشرفت سریع استفاده از الگوریتم‌ها در حوزه‌های مختلف از جمله: خدمات مالی، سلامت، عدالت قضایی و سیستم‌های توصیه‌گر، اهمیت شفافیت در الگوریتم‌ها به‌طور قابل توجهی افزایش یافته است، شفافیت الگوریتمی نه تنها موجب اعتمادسازی کاربران به فناوری می‌شود، بلکه به توسعه‌دهندگان، قانون‌گذاران و جامعه کمک می‌کند تا عملکرد سیستم‌ها را ارزیابی و کنترل کنند (دوشی-ولز و کیم<sup>۳</sup>، ۲۰۱۹).

### \* از جنبه فنی، شفافیت الگوریتمی به دو بخش اصلی تقسیم می‌شود:

(۱) شفافیت مدل<sup>۴</sup>: شفافیت مدل اشاره دارد به قابلیت درک ساختار و عملکرد الگوریتم، به گونه‌ای که یک فرد یا تیم بتواند نحوه تصمیم‌گیری مدل را به‌صورت دقیق شرح دهد. الگوریتم‌های ساده‌تر مانند، درخت‌های تصمیم یا رگرسیون خطی معمولاً شفافیت بیشتری دارند، اما الگوریتم‌های پیچیده و غیرخطی مانند شبکه‌های عصبی عمیق، به دلیل ماهیت سیاه‌جعبه‌ای خود، شفافیت کمتری دارند و توضیح تصمیمات آن‌ها دشوار است (لیپتن<sup>۵</sup>، ۲۰۱۸).

(۲) شفافیت فرآیند<sup>۶</sup>: شفافیت فرآیند به قابلیت ردیابی نحوه جمع‌آوری داده‌ها، مراحل پیش‌پردازش، آموزش مدل و اعمال تصمیمات اشاره دارد. این بخش از شفافیت، برای تضمین اخلاقی بودن، قانونی بودن و عادلانه بودن استفاده از الگوریتم‌ها اهمیت دارد (ویلر<sup>۷</sup>، ۲۰۱۹). اهمیت شفافیت الگوریتمی در جلوگیری از تبعیض و افزایش مسئولیت‌پذیری سیستم‌های مبتنی بر هوش مصنوعی برجسته است. الگوریتم‌هایی که فاقد شفافیت هستند، ممکن است به‌طور ناخواسته تعصبات موجود در داده‌ها را بازتولید کنند یا تصمیماتی اتخاذ کنند که برای کاربران ناعادلانه و غیرقابل فهم باشد، به همین دلیل، پژوهشگران و سازمان‌ها به دنبال روش‌هایی برای افزایش شفافیت الگوریتم‌ها و ارائه توضیحات قابل فهم برای

1 Nguyen, Tran, Ngo, & Phung

2 Algorithmic Transparency

3 Doshi-Velez & Kim

4 Model Transparency

5 Lipton

6 Process Transparency

7 Weller

کاربران هستند ( باروکاس، هاردت و نارایانان<sup>۱</sup>، ۲۰۱۹). روش‌های متعددی برای افزایش شفافیت الگوریتمی وجود دارد. برخی رویکردها بر توسعه مدل‌های قابل تفسیر تمرکز دارند که ساختار ساده و قابل فهمی دارند. برخی دیگر به استفاده از تکنیک‌های توضیح‌پذیری<sup>۲</sup> می‌پردازند که به صورت پسینی، تصمیمات مدل‌های پیچیده را تبیین می‌کنند؛ از جمله روش‌هایی مانند: LIME<sup>۳</sup> و SHAP<sup>۴</sup> که به کاربران امکان می‌دهند دلایل تصمیم یک مدل را به صورت محلی درک کنند. با این حال، تعادل بین دقت مدل و شفافیت یکی از چالش‌های مهم در توسعه الگوریتم‌ها است. مدل‌های ساده‌تر معمولاً شفاف‌تر ولی دقت کمتری دارند، در حالی که مدل‌های پیچیده‌تر دقت بالاتری دارند اما شفافیت آن‌ها کمتر است (دوشی-ولز و کیم<sup>۵</sup>، ۲۰۱۹)، از این رو، در بسیاری از کاربردها، انتخاب مدل به نوع مسأله، اهمیت شفافیت و نیاز به دقت بستگی دارد، در نهایت، شفافیت الگوریتمی نه تنها یک چالش فنی بلکه یک ضرورت اخلاقی و قانونی در دنیای متکی بر هوش مصنوعی است. افزایش شفافیت می‌تواند به بهبود اعتماد عمومی، تسهیل نظارت قانونی و تضمین عدالت در تصمیم‌گیری‌های خودکار کمک کند (ویل، ون کلک و بینز<sup>۶</sup>، ۲۰۲۰).

### سوگیری الگوریتمی<sup>۷</sup>

سوگیری الگوریتمی<sup>۸</sup> به خطاهای سیستماتیک در عملکرد الگوریتم‌های هوش مصنوعی و یادگیری ماشین اشاره دارد که باعث می‌شود این سیستم‌ها به گروه‌های خاصی از افراد به طور ناعادلانه یا تبعیض‌آمیز برخورد کنند. این سوگیری‌ها معمولاً از داده‌های آموزشی نامتناسب یا ناقص، طراحی مدل‌های غیرمنصفانه و یا روش‌های یادگیری معیوب ناشی می‌شوند و می‌توانند نابرابری‌های اجتماعی و ساختاری موجود را بازتولید کنند (بینز<sup>۹</sup>، ۲۰۱۸). در چند سال اخیر، گسترش کاربردهای الگوریتم‌های هوش مصنوعی در حوزه‌هایی مانند: عدالت کیفری، سلامت، آموزش و استخدام موجب شده تا نگرانی‌ها در مورد پیامدهای سوگیری الگوریتمی افزایش یابد، به عنوان مثال، الگوریتم‌های ارزیابی ریسک پیش‌محاکمه در سیستم قضایی آمریکا نشان داده‌اند که ممکن است به دلیل

1 Barocas, Hardt & Narayanan,

2 Explainability

3 Local Interpretable Model-agnostic Explanations

4 SHapley Additive exPlanations

5 Doshi-Velez & Kim

6 Veale, Van Kleek, & Binns

7 Algorithmic Bias

8 Algorithmic Bias

9 Binns

سوگیری‌های ضمنی داده‌ها، در تخمین ریسک متهمان اقلیت‌های نژادی نادرست عمل کنند. این مشکلات در حوزه سلامت نیز مشاهده شده است؛ جایی که الگوریتم‌های تشخیص بیماری به دلیل داده‌های ناقص، نتایجی تبعیض‌آمیز برای بیماران اقلیت ارائه داده‌اند (چن و همکاران، ۲۰۲۰).

### پیشینه پژوهش

#### تحقیقات داخلی

\* ولیپور (۱۴۰۴) در پژوهش خود به بررسی «تشخیص تقلب مبتنی بر هوش مصنوعی و پیش‌بینی بازار انرژی» پرداخت. در آینده بازارهای مالی مدرن با ادغام هوش مصنوعی دچار تغییرات فراوانی خواهد شد و به‌طور افزایشی هم می‌توان در آن تقلب کرد و هم می‌توان تقلب را شناسایی کرد. هوش مصنوعی می‌تواند دقت، بهره‌وری و کارایی خدمات مالی را افزایش دهد و روش‌های مؤثرتری برای مدیریت مالی، مدیریت ریسک، امنیت و به‌روزرسانی بازار ارائه دهد. با استفاده از الگوریتم‌های پیشرفته و تکنیک‌های تجزیه و تحلیل داده‌ها، هوش مصنوعی می‌تواند فعالیت‌های متقلبانه را با سرعت و دقت بی‌سابقه شناسایی و از آن جلوگیری کند. شناسایی تقلب و کلاهبرداری توسط هوش مصنوعی یکی از مهمترین برنامه‌های مالی آینده می‌باشد. شناسایی الگوهای تقلب و کلاهبرداری به‌صورت الگوی غیرعادی و پیش‌بینی تهدیدهای موجود قبل از اتفاق آن توسط هوش مصنوعی خطر زیان‌های مالی را کاهش می‌دهد. امنیت را در سیستم‌های مالی افزایش می‌دهد. توانایی تجزیه و تحلیل حجم وسیعی از داده‌ها در زمان بلادرنگ به مؤسسات مالی اجازه می‌دهد تا به‌طور فعال به طرح‌های کلاهبرداری در حال تکامل واکنش نشان دهند و دفاع قوی‌تری در برابر فعالیت‌های مجرمانه سایبری پیچیده ارائه دهند.

\* محمدی آریا وسط چی و یارمحمدی (۱۴۰۴) در پژوهش خود به بررسی «هوش مصنوعی: چالش‌های اخلاقی عصر حاضر» پرداختند. هوش مصنوعی (AI) با تأثیر گسترده در حوزه‌های مختلف، چالش‌های اخلاقی متعددی از جمله سوگیری الگوریتمی، نقض حریم خصوصی، عدم شفافیت در تصمیم‌گیری، تأثیر بر اشتغال و مخاطرات امنیتی را به همراه داشته است. این پژوهش با هدف بررسی این چالش‌ها و ارائه راهکارهایی برای کاهش پیامدهای منفی آن‌ها انجام شده است. نتایج نشان می‌دهد که سوگیری در داده‌های آموزشی منجر به تصمیم‌گیری‌های ناعادلانه شده و عدم شفافیت الگوریتم‌ها موجب

کاهش اعتماد عمومی به هوش مصنوعی می‌شود و همچنین، جمع‌آوری گسترده داده‌های کاربران نگرانی‌های امنیتی ایجاد کرده و اتوماسیون مشاغل برخی از فرصت‌های شغلی را تهدید می‌کند. برای مقابله با این چالش‌ها، راهکارهایی مانند: بهبود کیفیت داده‌ها، افزایش شفافیت الگوریتم‌ها، نظارت انسانی بر تصمیمات هوش مصنوعی، تدوین قوانین سخت‌گیرانه برای حفاظت از حریم خصوصی و حمایت از نیروی کار پیشنهاد شده است. این پژوهش بر لزوم توسعه هوش مصنوعی اخلاقی و مسئولیت‌پذیر برای کاهش اثرات منفی آن تأکید دارد.

\* حسینی و ارسطو (۱۴۰۳) در پژوهش خود به بررسی «اثرات اعتماد شناختی و اعتماد اجتماعی بر قصد پذیرش عامل دیجیتال هوش مصنوعی با نقش میانجی اعتماد عاطفی در بین مشتریان بیمه سامان» پرداختند. یافته‌های پژوهش نشان داد؛ اعتماد اجتماعی بر اعتماد عاطفی، اعتماد اجتماعی بر قصد پذیرش عامل دیجیتال هوش مصنوعی، اعتماد شناختی بر اعتماد عاطفی، اعتماد شناختی بر قصد پذیرش عامل دیجیتال هوش مصنوعی، و اعتماد عاطفی بر قصد پذیرش عامل دیجیتال هوش مصنوعی تأثیر مثبت و معناداری دارد و همچنین اعتماد عاطفی در رابطه بین اعتماد اجتماعی با قصد پذیرش عامل دیجیتال هوش مصنوعی و اعتماد شناختی با قصد پذیرش عامل دیجیتال هوش مصنوعی نقش مثبت و معناداری دارد.

\* دارائی (۱۴۰۳) در پژوهش خود به بررسی «هوش مصنوعی توضیح‌پذیر: گامی به سوی شفافیت و اعتماد در هوش مصنوعی» پرداخت. به‌عنوان یکی از مفاهیم کلیدی در حوزه هوش مصنوعی، تلاش می‌کند تا شکاف میان پیچیدگی مدل‌های یادگیری ماشین و نیاز به شفافیت در تصمیم‌گیری‌ها را کاهش دهد. این مفهوم به کاربران کمک می‌کند تا نحوه تصمیم‌گیری مدل‌ها را درک کرده و اعتماد بیشتری به خروجی‌های آن‌ها داشته باشند. هوش مصنوعی توضیح‌پذیر نه تنها درک بهتری از مدل‌های پیچیده مانند شبکه‌های عصبی عمیق فراهم می‌کند، بلکه به توسعه‌دهندگان اجازه می‌دهد تا نقاط ضعف مدل‌ها را شناسایی کرده و عملکرد آن‌ها را بهبود بخشند. در این مقاله، روش‌های مختلفی از جمله: SHAP، LIME و DeepLIFT که از متداول‌ترین تکنیک‌های هوش مصنوعی توضیح‌پذیر هستند، مورد بررسی قرار گرفته‌اند و همچنین به کاربردهای گسترده هوش مصنوعی توضیح‌پذیر در صنایع مختلف مانند: پزشکی، مالی، خودروهای خودران و سیستم‌های قضایی پرداخته شده است. این روش‌ها می‌توانند تعصب‌های موجود در داده‌ها را کاهش

داده، شفافیت در تصمیم‌گیری‌ها را افزایش دهند و به سازمان‌ها در رعایت مقررات و قوانین کمک کنند. علاوه بر مزایای ذکر شده، چالش‌هایی مانند: پیچیدگی فنی، نبود استانداردهای یکپارچه و نیاز به منابع محاسباتی مورد بحث قرار گرفته‌اند. با این حال، توسعه و گسترش هوش مصنوعی توضیح‌پذیر می‌تواند نقش مهمی در بهبود کیفیت زندگی، افزایش بهره‌وری سیستم‌ها و ایجاد اعتماد بیشتر در استفاده از فناوری‌های هوش مصنوعی ایفا کند. در نهایت، هوش مصنوعی توضیح-پذیر به عنوان یک ابزار اساسی برای مسئولیت-پذیری هوش مصنوعی شناخته می‌شود.

\* روزبھانی و نظری (۱۴۰۳) در پژوهش خود به بررسی «بررسی سیستم‌های تشخیص نفوذ مبتنی بر هوش مصنوعی» پرداختند. امروزه با وجود برخورد با حجم عظیم اطلاعات همواره با مخاطرات امنیتی، حملات هکری، مخرب، نفوذگرانه و نیز یافتن راه کارهایی برای توسعه ابزارهای بهینه در رفع و تشخیص بهتر آن‌ها روبرو هستیم. می‌توان بیان نمود که برای شناسایی برخی از حملات و حفظ امنیت می‌توان از روش تشخیص نفوذ مبتنی بر امضا بهره جست. حال با آمدن هوش مصنوعی، الگوریتم‌های یادگیری عمیق، یادگیری ماشین چنین قلمداد شده که نتایج رو به رشدی در حفاظت از داده‌های حاصل شده است، سعی شده در این مقاله با بهره‌مندی از هوش مصنوعی در شناسایی بهتر حملات با توجه به مباحث مرتبط با آن پردازیم. محققان بر این نظر هستند که استفاده از هوش مصنوعی بر بهبود عملکرد برای شناسایی حملات مفید و تمرکز را بر روی آن قرار داده‌اند، می‌خواهیم مکانیسم‌های را مبنی بر استفاده از هوش مصنوعی در تشخیص نفوذ و نیز بینشی را برای محققان آینده در راستای درک بیشتر از چالش‌های پیش روی ارائه داده شود.

\* اکبرالسادات و اسماعیل پور (۱۴۰۱) در پژوهش خود به بررسی «مروری بر روش‌های کشف تقلب بانکی با استفاده از هوش مصنوعی است» پرداختند. تقلب بانکی با توجه به تأثیرات مخرب آن، یکی از تهدیدآمیزترین مشکلاتی است که هر جامعه بشری با آن دست و پنجه نرم می‌کند. این عمل به استفاده عمدی از اطلاعات نادرست برای کلاهبرداری از پول یا دارایی فرد یا سازمان دیگری اشاره دارد. صنعت بانکداری برای چندین دهه از سیستم‌های مبتنی بر قوانین برای شناسایی تقلب و بررسی انسانی تراکنش‌ها استفاده کرده است. سیستم‌های مبتنی بر قانون شامل الگوریتم‌هایی هستند که انواع مختلفی از اقدامات شناسایی را انجام می‌دهند که به صورت دستی توسط متخصصان تقلب نوشته

شده‌اند. این سیستم‌ها به تنظیم دستی سناریوها نیاز دارند، که در تشخیص ضمنی همبستگی‌های معاملاتی که به تقلب اشاره می‌کنند، دچار چالش می‌شوند. با توجه به ضعف‌های ذاتی رویکرد تشخیص تقلب مبتنی بر قانون در بانک‌ها و داده‌های محدودی که بر الگوریتم‌های یادگیری ماشین تحت نظارت متداول استفاده می‌شود، نیاز مبرمی به تکنیک‌ها یا سیستم‌های جدید تشخیص تقلب وجود دارد که بتوانند با افزایش سریع تقلب‌ها و موارد پولشویی مقابله کنند. این تحقیق با استفاده از رویکرد مروری و با هدف بررسی روش‌های کشف تقلب بانکی به بررسی ادبیات تحقیق و توصیف نتایج مرتبط می‌پردازد.

### تحقیقات خارجی

\* یاسین و الامارنه (۲۰۲۵) در پژوهش خود به بررسی «پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در بانکداری: نقش اعتماد، شفافیت و درک انصاف در مؤسسات مالی در امارات متحده عربی و قطر» پرداختند. این مقاله به بررسی میزان پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی با تمرکز ویژه بر اعتماد، شفافیت و درک از انصاف، در میان مؤسسات مالی در امارات متحده عربی و قطر می‌پردازد. علی‌رغم پتانسیل هوش مصنوعی در شناسایی ناهنجاری‌های مالی، فرآیندهای تصمیم‌گیری مبهم و سوگیری الگوریتمی، پذیرش گسترده آن را به‌ویژه در بخش‌های نظارت بانکی محدود می‌کند. این مطالعه نشان می‌دهد که شفافیت، اعتماد را تا حد زیادی افزایش می‌دهد، که پیش‌بینی‌کننده اصلی پذیرش هوش مصنوعی است. درک از انصاف، اثرات منفی سوگیری الگوریتمی را تعدیل می‌کند و بر نقش مهم آن در ایجاد اعتبار سیستم تأکید دارد. تحلیل گروه‌های مختلف نشان می‌دهد که از نظر حساسیت به اعتماد و انصاف بر اساس شغل و منطقه تفاوت‌هایی وجود دارد و حسابرسان داخلی و افرادی که بیشتر در معرض هوش مصنوعی هستند، آمادگی بیشتری برای پذیرش نشان می‌دهند. انطباق با مقررات نیز به‌عنوان یک فعال‌کننده مثبت برای پذیرش ظاهر می‌شود. این مقاله پیشنهاد عملی برای بانک‌ها، توسعه‌دهندگان و تنظیم‌کننده‌ها به‌منظور همسو کردن استقرار هوش مصنوعی با آرمان‌های اخلاقی و نظارتی ارائه می‌دهد. این مطالعه ابزارهای هوش مصنوعی شفاف، قابل توضیح و حساس به انصاف را به‌عنوان عنصری اساسی برای ترویج پذیرش در

بخش‌های مبتنی بر مقررات توصیه می‌کند. یافته‌ها راهنمایی‌هایی برای اجرای مسئولانه و مبتنی بر اعتماد هوش مصنوعی در شناسایی تقلب ارائه می‌کنند.

\* شلکه و همکاران (۲۰۲۵) در پژوهش خود به بررسی «مطالعه‌ای بر اثربخشی تشخیص تقلب مبتنی بر هوش مصنوعی در بانکداری آنلاین» پرداختند. گسترش سریع بانکداری دیجیتال با ارائه راحتی و کارایی بیشتر، صنعت مالی را متحول کرده است؛ با این وجود، این گذار به پلتفرم‌های دیجیتال، پتانسیل کلاهبرداری را نیز افزایش داده است، زیرا مجرمان سایبری از تکنیک‌های پیچیده‌تری برای سوءاستفاده از خدمات آنلاین استفاده می‌کنند. در پاسخ به این چالش‌ها، مؤسسات مالی به تدریج در حال پذیرش فناوری‌های هوش مصنوعی (AI)، از جمله: یادگیری ماشینی، تجزیه و تحلیل رفتاری و نظارت بلادرنگ، برای تقویت قابلیت‌های تشخیص کلاهبرداری خود هستند. نتایج این نظرسنجی نشان می‌دهد که در حالی که کاربران عموماً از سهم هوش مصنوعی در پیشگیری از کلاهبرداری آگاه هستند، اما همچنان شکاف قابل توجهی در اعتماد و درک وجود دارد. تقریباً نیمی از پاسخ‌دهندگان ابزارهای مبتنی بر هوش مصنوعی را تشخیص دادند، اما بسیاری از آن‌ها نگرانی‌هایی در مورد حریم خصوصی، قابلیت اطمینان سیستم و عدم نظارت انسانی ابراز کردند. تجزیه و تحلیل کای اسکور برای ارزیابی اینکه آیا برداشت کاربران از هوش مصنوعی در مقایسه با روش‌های سنتی به‌طور قابل توجهی بیشتر است یا خیر، انجام شد و هیچ ترجیح آماری قابل-توجهی را نشان نداد. این نتایج نشان می‌دهد که برای دستیابی به اثربخشی و پذیرش بهینه هوش مصنوعی در تشخیص تقلب، بانک‌ها باید افزایش شفافیت، رسیدگی به نگرانی‌های مشتریان و ارائه منابع آموزشی برای تقویت اعتماد کاربران به راه‌حل‌های مبتنی بر هوش مصنوعی را در اولویت قرار دهند.

\* اسکندریانی (۲۰۲۴) در پژوهش خود به بررسی «پذیرش هوش مصنوعی و یادگیری ماشینی در سیستم‌های بانکی: یک بررسی کیفی از هیئت مدیره» پرداخت. هدف این مقاله دو گانه است: اول، بررسی نقش هیئت مدیره در تسهیل پذیرش هوش مصنوعی و یادگیری ماشینی در بخش بانکداری عربستان سعودی؛ دوم، بررسی اثربخشی هوش مصنوعی و یادگیری ماشینی در محافظت از بخش بانکداری عربستان سعودی در برابر حملات سایبری. یک رویکرد تحقیق کیفی با استفاده از مصاحبه‌های عمیق با ۱۷ هیئت مدیره از بانک‌های برجسته عربستان سعودی به کار گرفته شد. مطالعه حاضر، فرصت‌ها و چالش‌های

ادغام فناوری‌های پیشرفته هوش مصنوعی و یادگیری ماشینی را در این صنعت بسیار قانونمند برجسته می‌کند. یافته‌ها نشان می‌دهد که فناوری‌های پیشرفته هوش مصنوعی و یادگیری ماشینی مزایای قابل توجهی، به‌ویژه در زمینه‌هایی مانند تشخیص تهدید، پیشگیری از کلاهبرداری و اتوماسیون فرآیند، ارائه می‌دهند و بانک‌ها را قادر می‌سازند تا استانداردهای نظارتی را رعایت کرده و تهدیدات سایبری را به‌طور مؤثر کاهش دهند. با این حال، این تحقیق همچنین موانع قابل توجهی را شناسایی می‌کند، از جمله زیرساخت‌های محدود فناوری، فقدان استراتژی‌های منسجم هوش مصنوعی و نگرانی‌های اخلاقی در مورد حریم خصوصی داده‌ها و سوگیری الگوریتمی؛ مصاحبه‌شوندگان بر نقش حیاتی هیئت مدیره در ارائه جهت استراتژیک، تأمین منابع و تقویت مشارکت با ارائه‌دهندگان فناوری هوش مصنوعی تأکید کردند. این مطالعه همچنین بر اهمیت همسوسازی ابتکارات هوش مصنوعی و یادگیری ماشینی با اهداف توسعه ملی، مانند: چشم‌انداز ۲۰۳۰ عربستان سعودی، برای تضمین رشد و رقابت‌پذیری پایدار تأکید می‌کند. یافته‌های این مطالعه، پیامدهای ارزشمندی برای سیاست‌گذاران بانکی در جهت‌یابی پیچیدگی‌های پذیرش هوش مصنوعی و یادگیری ماشینی در خدمات مالی، به‌ویژه در بازارهای نوظهور، ارائه می‌دهد.

\* ادیمی و همکاران<sup>۱</sup> (۲۰۲۴) در پژوهش خود به «بررسی نقش هوش مصنوعی در تشخیص و پیشگیری از کلاهبرداری در خدمات مالی» پرداختند. این بررسی به بررسی مدل‌های متنوع هوش مصنوعی مورد استفاده در پیشگیری از کلاهبرداری، از جمله یادگیری نظارت‌شده و بدون نظارت، یادگیری عمیق و پردازش زبان طبیعی، می‌پردازد. تجزیه و تحلیل دقیق، اثربخشی هوش مصنوعی را در شناسایی الگوهای پیچیده نشان‌دهنده رفتار کلاهبرداری نشان می‌دهد و برتری آنرا در تشخیص ناهنجاری‌ها در مجموعه داده‌های گسترده و پویا نشان می‌دهد. علاوه بر این، این بررسی، پیامدهای دنیای واقعی هوش مصنوعی در تشخیص کلاهبرداری را روشن می‌کند و مواردی را که این فناوری با موفقیت طرح‌های کلاهبرداری را خنثی کرده است، برجسته می‌کند. ملاحظات اخلاقی ذاتی در پیشگیری از کلاهبرداری مبتنی بر هوش مصنوعی نیز مورد بررسی قرار گرفته و بر اهمیت شیوه‌های مسئولانه و شفاف برای کاهش سوگیری‌ها و تضمین انصاف در فرآیندهای تصمیم‌گیری تأکید می‌شود. همزمان با گذر چشم‌انداز مالی از دوران تحول

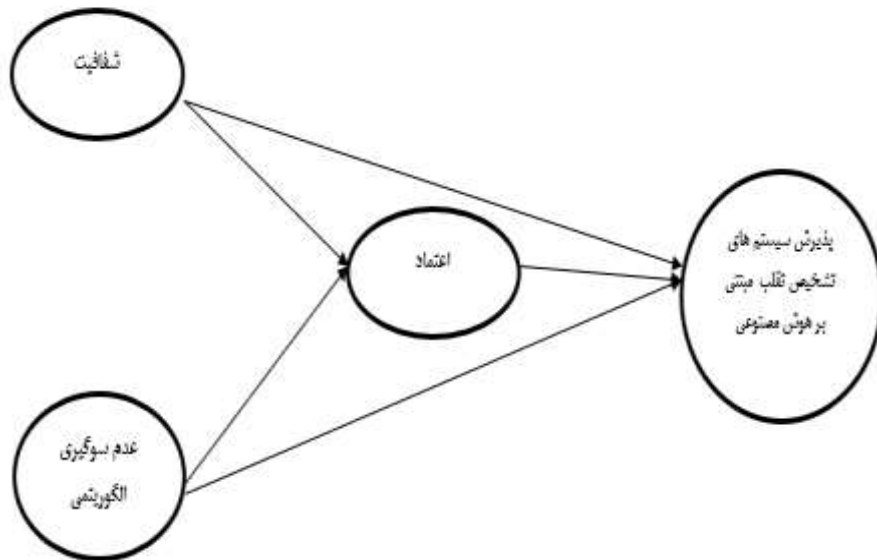
دیجیتال، این بررسی، روندها و نوآوری‌های آینده در تشخیص کلاهبرداری مبتنی بر هوش مصنوعی را روشن می‌کند. پیشرفت‌های پیش‌بینی‌شده شامل: ادغام هوش مصنوعی قابل توضیح (XAI)، یادگیری فدرال و سازگاری مداوم با تهدیدهای نوظهور است. این بحث به تلاش‌های مشترک بین مؤسسات مالی، نهادهای نظارتی و ارائه‌دهندگان فناوری برای ایجاد یک اکوسیستم قوی که قادر به پیشی گرفتن از تاکتیک‌های کلاهبرداری در حال تکامل باشد، گسترش می‌یابد. در نتیجه، این بررسی، چشم‌انداز پویای هوش مصنوعی در تشخیص و پیشگیری از کلاهبرداری در خدمات مالی را خلاصه می‌کند. این تحلیل، تأثیر دگرگون‌کننده هوش مصنوعی را نه تنها در تقویت اقدامات امنیتی، بلکه در پرورش یک رویکرد پیشگیرانه و تطبیقی برای مقابله با ماهیت در حال تکامل کلاهبرداری مالی برجسته می‌کند. تلفیق دیدگاه‌های تاریخی، کاربردهای فعلی و مسیرهای آینده، درک جامعی از چگونگی تغییر شکل الگوی تشخیص کلاهبرداری در حوزه مالی توسط هوش مصنوعی ارائه می‌دهد.

\* عزیز و آندریانسياه<sup>1</sup> (۲۰۲۳) در پژوهش خود به بررسی «نقش هوش مصنوعی در بانکداری مدرن: بررسی رویکردهای مبتنی بر هوش مصنوعی برای پیشگیری پیشرفته از کلاهبرداری، مدیریت ریسک و انطباق با مقررات» پرداختند. تجزیه و تحلیل نمودار با تجسم روابط تراکنشی، که به‌طور بالقوه فعالیت‌های مشکوکی مانند انتقال سریع وجوه که نشان‌دهنده پولشویی است را برجسته می‌کند، دیدگاه منحصر به فردی ارائه می‌دهد. تجزیه و تحلیل پیش‌بینی‌کننده، فراتر از روش‌های سنتی امتیازدهی اعتباری، مجموعه داده‌های متنوعی را در بر می‌گیرد و بینش جامع‌تری از اعتبار مشتری ارائه می‌دهد. این تحقیق همچنین بر اهمیت رابط‌های کاربرپسند مانند: چت‌بات‌های مبتنی بر هوش مصنوعی برای گزارش فوری فعالیت‌های مشکوک و ادغام تأییدهای بیومتریک پیشرفته، از جمله تشخیص چهره و صدا، تأکید می‌کند. تجزیه و تحلیل مکانی و بیومتریک رفتاری به ترتیب با تجزیه و تحلیل مکان‌های تراکنش و الگوهای تعامل کاربر، امنیت را بیشتر تقویت می‌کنند. یکی از مزایای قابل توجه هوش مصنوعی در سازگاری آن نهفته است. سیستم‌های خودآموز تضمین می‌کنند که با تکامل تاکتیک‌های کلاهبرداری، مکانیسم‌های هوش مصنوعی به‌روز می‌مانند و اثربخشی خود را حفظ می‌کنند. این سازگاری به تشخیص فیشینگ، ادغام اینترنت اشیا و تجزیه و تحلیل بین‌کانالی نیز گسترش می‌یابد و یک دفاع

1 Aziz & Andriansyah

جامع در برابر تلاش‌های کلاهبرداری چندوجهی ارائه می‌دهد. علاوه بر این، توانایی هوش مصنوعی در شبیه‌سازی سناریوهای اقتصادی به ریسک پیشگیرانه کمک می‌کند.  
مدل مفهومی پژوهش

روابط مورد پژوهش حاضر در مدل مفهومی زیر نشان داده شده است:



شکل ۱: مدل مفهومی پژوهشگر برگرفته از پژوهش یاسین و همکاران (۲۰۲۵)

### از این رو فرضیه‌های پژوهش حاضر به این صورت مطرح می‌شوند:

اعتماد بر تأثیر شفافیت بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی نقش میانجی دارد؛

اعتماد بر تأثیر عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی نقش میانجی دارد.

شفافیت بر اعتماد در شعب بانک ملی تأثیر معناداری دارد؛

عدم سوگیری الگوریتمی بر اعتماد در شعب بانک ملی تأثیر معناداری دارد؛

شفافیت بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی تأثیر معناداری دارد؛

عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی تأثیر معناداری دارد؛

اعتماد بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی تأثیر معناداری دارد.

## روش‌شناسی پژوهش

پژوهش حاضر در زمره پژوهش‌های کاربردی قرار دارد، چنانچه طبقه‌بندی انواع پژوهش - ها بر اساس ماهیت و روش مدنظر قرار گیرد، روش پژوهش حاضر از لحاظ ماهیت در زمره پژوهش‌های توصیفی پیمایشی، قرار دارد. جامعه آماری پژوهش حاضر کلیه کارکنان و کارشناسان شعب بانک ملی در شهر کرج به تعداد ۶۵۳ نفر استفاده می‌کنند. برای تعیین حجم نمونه از فرمول کوکران استفاده شده است که با توجه به جامعه آماری محدود، حجم نمونه موردنیاز پژوهش ۲۴۲ نفر است روش انتخاب نمونه‌های آماری در این پژوهش، با توجه به جامعه آماری نامحدود، روش نمونه‌گیری تصادفی ساده است. و برای جمع‌آوری داده‌های موردنیاز از روش مطالعات کتابخانه‌ای و مطالعات میدانی استفاده شده است. همچنین برای تجزیه و تحلیل داده‌ها از فن مدل‌سازی معادلات ساختاری لیزرل بهره برده‌ایم.

## ابزار جمع‌آوری داده‌ها

جمع‌آوری داده‌ها به صورت پیمایشی با استفاده از ابزار پرسشنامه محقق ساخته انجام خواهد شد. بخش اول پرسشنامه شامل متغیرهای جمعیت شناختی (جنس، میزان تحصیلات) است و بخش دوم حاوی سؤالات مرتبط با متغیرهای پژوهش است که بر اساس مقیاس لیکرت تهیه شد. برای گردآوری شواهد روایی ابزار پژوهش، راه‌های مختلفی وجود دارد که عبارت‌اند از: روایی محتوایی، روایی صوری و روایی سازه (شواهد همگرا، واگرا و همسانی درونی و روایی عاملی). در اولین مرحله از پژوهش، برای بررسی روایی صوری پرسشنامه در اختیار خبرگان در دسترس قرار گرفت و تأکید شد که در ارزیابی کیفی روایی محتوا، موارد رعایت دستور زبان، استفاده از کلمات مناسب، اهمیت سؤالات، قرارگیری سؤالات در جای مناسب خود، زمان تکمیل ابزار طراحی شده را مدنظر قرار دهند. نهایتاً نظرات آنان به صورت تغییراتی جزئی در پرسشنامه اعمال شد. در مرحله دوم از بررسی روایی، برای ارزیابی روایی محتوایی از نظر متخصصان در مورد میزان هماهنگی محتوای ابزار اندازه‌گیری و هدف پژوهش، استفاده شده است. بدین ترتیب که پرسشنامه به ۱۵ نفر از متخصصین رشته مدیریت داده و از آنان درخواست شد که در خصوص سؤالات ابزار به صورت مفید است و مفید نیست پاسخ دهند. پاسخ‌ها بر اساس فرمول CVR محاسبه و با جدول لاوشه انطباق داده شد. اعداد بالاتر از ۰,۴۹ مورد قبول واقع شد. پس از ارزیابی روایی، جهت بررسی پایایی پرسشنامه در این پژوهش از روش ضریب

امگامکدونالد جهت سنجش پایایی ابزار پژوهش استفاده شده است. توزیع سؤالات پرسشنامه و ضریب امگامکدونالد آن‌ها در جدول (۱) بیان شده است.

جدول ۱: ارتباط میان متغیرها و سؤالات پرسشنامه

ردیف	متغیر	سؤالات	امگامکدونالد	منبع
۱	شفافیت	۴-۱	۰/۸۵۵	(ریبیرو و همکاران، ۲۰۱۶)
۲	عدم سوگیری الگورتیمی	۸-۵	۰/۸۲۱	(محرابی و همکاران، ۲۰۲۱)
۳	اعتماد	۱۲-۹	۰/۸۳۳	
۴	پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی	۱۶-۱۳	۰/۸۷۱	(یاسین و العمارنه، ۲۰۲۵)

## تجزیه و تحلیل آماری

### روش مدل معادلات ساختاری (SEM)

در این پژوهش برای بررسی و آزمون مدل مفهومی ارائه شده توسط پژوهشگر از روش معادلات ساختاری و نرم افزار لیزرل استفاده شده است. مدل سازی معادله ساختاری به پژوهشگر یاری می‌رساند تا الگویی نظری را که از اجزای مختلف و متنوعی تشکیل یافته، هم به طور کلی و هم به گونه‌ای جزئی مورد آزمون و واریسی قرار دهد (عاقلی و همکاران، ۲۰۲۳). مدل معادله ساختاری ترکیبی از مدل‌های مسیر (روابط ساختاری) و مدل‌های عاملی تأییدی (روابط اندازه‌گیری) است. در مدل‌های مسیر پژوهشگر تلاش می‌کند تا با مجموعه‌ای از روابط یک‌سویه و دوسویه پدیده یا پدیده‌هایی را تبیین کند در حالی که متغیرهای حاضر در مدل مسیر از نوع مشاهده شده هستند. در مدل‌های عاملی تأییدی نیز پژوهشگر به دنبال تعریف سازه یا سازه‌هایی پنهان بر مبنای مجموعه‌ای از معرف‌هاست.

### مدل اندازه‌گیری

در مرحله اول به دنبال پاسخ به این پرسش هستیم که آیا هر یک از مدل‌های اندازه‌گیری تدوین شده برای سازه‌ها حائز حداقل معیارهای علمی تعریف شده هستند یا خیر، لازم است تا هر یک از مدل‌های اندازه‌گیری را جداگانه مورد تحلیل قرار دهیم. مدل مفهومی پژوهش حاضر دارای دو متغیر است. در این مرحله به بررسی بار عاملی حاصل از تحلیل

1 Ribeiro et al

2 Mehrabi et al

3 Agheli

عاملی تأییدی پرداخته‌ایم که مؤلفه‌هایی با بار عاملی بیشتر از ۰/۵ برای استخراج مدل نهایی مورد نظر قرار گرفته‌اند (عاقلی و آجرلو، ۲۰۱۸). همچنین به منظور بررسی کفایت داده‌ها برای تحلیل عاملی از شاخص KMO و آزمون بارتلت استفاده شده است؛ شاخص KMO شاخصی از کفایت نمونه‌گیری است. این شاخص در دامنه صفر تا یک قرار دارد. اگر مقدار شاخص نزدیک به یک باشد، داده‌های مورد نظر برای تحلیل عاملی مناسب هستند و در غیر این صورت (معمولاً کمتر از ۰/۵)، نتایج تحلیل عاملی برای داده‌های مورد نظر چندان مناسب نیستند (عاقلی و همکاران، ۲۰۲۳). همچنین مقدار شاخص KMO برای تمامی متغیرهای پژوهش بالاتر از ۰/۵ به دست آمد که مناسب و پذیرفتنی است و از طرف دیگر نیز آزمون بارتلت برای تحلیل عاملی تمام متغیرها معنادار است، پس می‌توان گفت که نمونه کفایت مناسب دارد. بار عاملی برای گویه‌های هر متغیر نیز محاسبه شده است، این مقادیر در جدول ۲ آمده است.

جدول ۲: جدول بارهای عاملی و آماره تی

متغیر	گویه	بار عاملی	آماره تی	CR	AVE	MSV	ASV
شفافیت	Q01	۰/۶۱	۱۲/۶۴	۰/۹۵۶	۰/۵۵۶	۰/۴۳۶	۰/۵۲۱
	Q02	۰/۶۳	۱۱/۹۷				
	Q03	۰/۶۷	۱۳/۰۶				
	Q04	۰/۶۱	۱۰/۸۸				
عدم سوگیری الگوریتمی	Q05	۰/۶۶	۱۰/۳۴	۰/۹۸۴	۰/۵۸۲	۰/۴۱۷	۰/۴۷۸
	Q06	۰/۶۴	۱۴/۳۸				
	Q07	۰/۶۴	۱۳/۸۱				
اعتماد	Q08	۰/۶۴	۱۳/۲۸	۰/۹۴۷	۰/۵۲۲	۰/۵۱۴	۰/۴۱۹
	Q09	۰/۶۳	۱۳/۰۶				
	Q10	۰/۵۷	۶/۲۷				
پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی	Q11	۰/۶۱	۱۱/۳۷	۰/۹۲۱	۰/۵۲۲	۰/۴۳۷	۰/۴۷۳
	Q12	۰/۶۳	۱۱/۱۷				
	Q13	۰/۶۹	۱۲/۶۷				
	Q14	۰/۶۶	۱۲/۰۴				
	Q15	۰/۷۱	۱۱/۹۴				
	Q16	۰/۶۴	۱۰/۹۷				

همان‌طور که در جدول بالا مشاهده می‌شود بارهای عاملی که بالای ۰/۵ هستند و در سطح اطمینان ۹۹ درصد معنادار هستند نشان‌دهنده آن است که شاخص‌ها (متغیرهای نشانگر)، متغیرهای مفهومی را به خوبی تبیین می‌کند.

همچنین خوشبختانه مدل اندازه‌گیری انعکاسی ما صاحب پایایی و روایی سازه است.

یافته‌ها

بررسی متغیرهای جمعیت شناختی

در این بخش به منظور توصیف ویژگی‌های نمونه، ابتدا داده‌های جمع‌آوری شده با استفاده از شاخص‌های آمار توصیفی خلاصه و طبقه‌بندی می‌شود، سپس با استفاده از شاخص‌های آمار استنباطی به تأیید یا رد فرضیات می‌پردازیم. در ضمن کل پاسخ‌دهندگان ۳۰۰ نفر بوده‌اند که از این تعداد ۲۴۲ پرسشنامه درست استخراج شدند. با توجه به تحلیل داده‌ها، اطلاعات زیر پیرامون نمونه آماری پژوهش استخراج شد:

جدول ۳: فراوانی متغیرهای جمعیت شناختی

درصد	فراوانی	
۶۳	۱۵۲	مرد
۳۸	۹۰	زن
۱۵	۳۵	۲۰ تا ۳۰ سال
۳۴	۸۲	۳۱ تا ۴۰ سال
۳۱	۷۵	۴۱ تا ۵۰ سال
۲۰	۵۰	بالای ۵۱ سال
۳۴	۸۲	لیسانس و کمتر
۵۳	۱۳۰	فوق لیسانس
۱۳	۳۰	دکتر
۱۰۰	۲۴۲	جمع

مدل ساختاری

در یک مدل معادله ساختاری به معنای عام، پژوهشگر از طرفی به دنبال آن است که مجموعه‌ای از متغیرهای پنهان را با مجموعه‌ای از معرف‌ها اندازه‌گیری کرده و از طرف دیگر روابط ساختاری بین متغیرهای پنهان را مورد تجزیه و تحلیل قرار دهد، هرچند ممکن است در این میان برخی از متغیرهای حاضر در مدل ساختاری از نوع متغیرهای مشاهده شده باشند. پس می‌توان این‌گونه بیان کرد که یک مدل معادله ساختاری به‌طور معمول ترکیبی از مدل‌های اندازه‌گیری و مدل‌های ساختاری‌اند. (عاقلی و همکاران، ۲۰۲۳). با اجرای

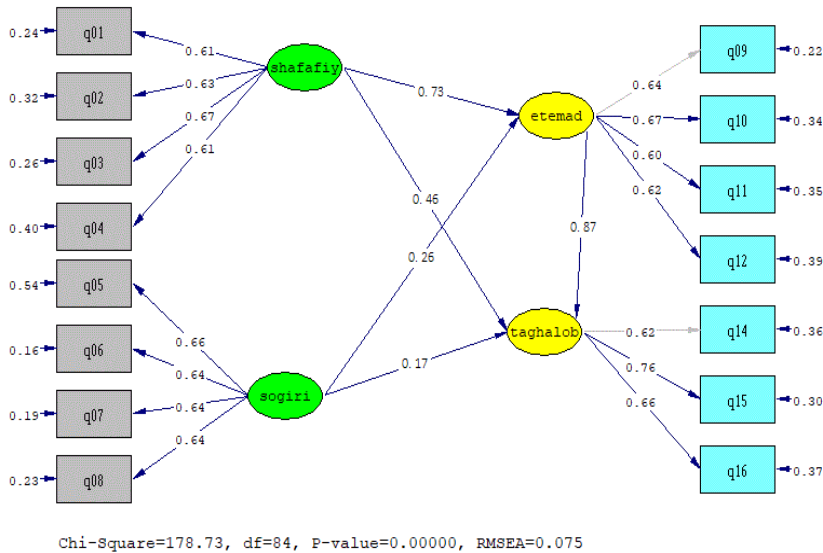
آزمون مدل‌یابی معادلات ساختاری در نرم‌افزار شاخص‌های برازشی ارائه می‌شوند که نشان می‌دهند تا حدی مدل مفهومی پژوهش با داده‌های تجربی برازش می‌شود. برخلاف آزمون‌های مرسوم آماری که با یک آماره تأیید یا رد می‌شوند، در مدل‌یابی معادلات ساختاری دسته‌ای از شاخص‌ها معرفی می‌شوند. وضعیت شاخص‌های برازش برای مدل مفهومی پژوهش در جدول زیر آمده است:

جدول ۴: شاخص‌های برازش در مدل معادلات ساختاری تأییدشده پژوهش

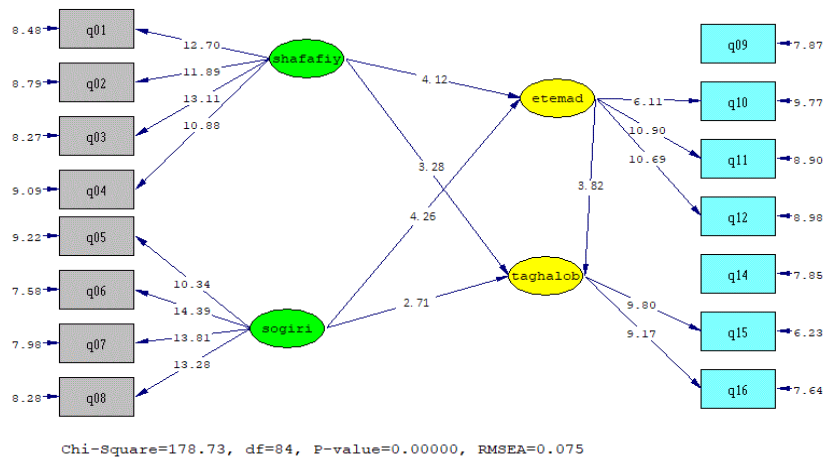
نتیجه	مقدار اولیه	برازش مناسب	نام شاخص	
تأیید	۲/۱۷	$< 3$	نسبت مجذور خی به درجه آزادی ( $\chi^2/df$ )	شاخص‌های برازش مقتصد
		$(\chi^2/df)$		
تأیید	۰/۰۷۵	$< 0/80$	ریشه خطای میانگین مجذورات تقریبی (RMSEA)	
		RMSEA		
تأیید	۰/۷۷	$> 0/5$	شاخص برازش هنجار شده مقتصد (PNFI)	شاخص‌های برازش مطلق
		PNFI		
تأیید	۰/۸۹	$> 0/8$	شاخص نیکویی برازش (GFI)	
		GFI		
تأیید	۰/۸۵	$> 0/8$	شاخص خوبی برازندگی تعدیل شده (AGFI)	شاخص‌های برازش تطبیقی
		AGFI		
تأیید	۰/۹۷	$> 0/9$	شاخص نرم شده برازندگی (NFI)	
		NFI		شاخص‌های برازش تطبیقی
تأیید	۰/۹۸	$> 0/9$	شاخص برازش هنجار نشده (NNFI)	
		NNFI		
تأیید	۰/۹۸	$> 0/9$	شاخص برازندگی تطبیقی (CFI)	
		CFI		
تأیید	۰/۹۶	$> 0/9$	شاخص برازش نسبی (RFI)	
		RFI		
تأیید	۰/۹۸	$> 0/9$	شاخص برازندگی فزاینده (IFI)	
		IFI		

همان‌طور که در جدول مشخص است کلیه شاخص‌های برازش مدل از وضعیت مناسبی برخوردار هستند. به‌عنوان مثال سطح معناداری کای اسکوتر مدل از ۵ درصد بیشتر است و این نشان از آن دارد که داده‌های تجربی به نحو مناسبی از مدل مفهومی پژوهش حمایت می‌کنند. همچنین مقدار ریشه میانگین مربعات خطای برآورد نیز کمتر از ۱۰ درصد

است و این شاخص نیز برازش مدل را تأیید می‌کند. قابل قبول بودن سایر شاخص‌های مدل نیز در خروجی‌های نرم‌افزار دیده می‌شود. به منظور آزمون فرضیه‌های تحقیق، از مدل یابی معادلات ساختاری (SEM) با به کارگیری روش حداکثر درست‌نمایی در نرم‌افزار لیزرل استفاده شد که در طی همه مراحل تحلیل، ماتریس کواریانس مبنای کار قرار گرفت. شکل زیر روابط ساختاری مدل تحقیق را نشان می‌دهد.



شکل ۲: مدل تخمین استاندارد



شکل ۳: مدل ضرایب معناداری

جدول ۵: ضریب مسیر فرضیات

ردیف	فرضیه	A	B	SEA	SEB	Sobel test statistic	Two-tailed probability	نتیجه
۱	شفافیت ← اعتماد ← پذیرش سیستم های تشخیص تقلب مبتنی بر هوش مصنوعی	۰/۷۳	۰/۸۷	۰/۰۷۸	۰/۰۶۳	۲/۵۳۰	۰/۰۰۰	میانجی است
۲	عدم سوگیری الگوریتمی ← اعتماد ← پذیرش سیستم های تشخیص تقلب مبتنی بر هوش مصنوعی	۰/۲۶	۰/۸۷	۰/۰۷۷	۰/۰۶۸	۳/۶۴۳	۰/۰۰۰	میانجی است
ردیف	فرضیه	ضریب استاندارد	آماره تی	نتیجه				
۳	شفافیت ← اعتماد	۰/۷۳	۴/۱۲	تأیید				
۴	عدم سوگیری الگوریتمی ← اعتماد	۰/۲۶	۴/۲۶	تأیید				
۵	شفافیت ← پذیرش سیستم های تشخیص تقلب مبتنی بر هوش مصنوعی	۰/۴۶	۳/۲۸	تأیید				
۶	عدم سوگیری الگوریتمی ← پذیرش سیستم های تشخیص تقلب مبتنی بر هوش مصنوعی	۰/۱۷	۲/۷۱	تأیید				
۷	اعتماد ← پذیرش سیستم های تشخیص تقلب مبتنی بر هوش مصنوعی	۰/۸۷	۳/۸۲	تأیید				

در این مرحله به منظور بررسی معناداری هر کدام از پارامترهای الگو از آماره  $t$  استفاده می‌شود. این آماره از نسبت ضریب هر پارامتر به خطای انحراف معیاران پارامتر به دست می‌آید که باید در آزمون  $t$  بزرگ‌تر از ۱٫۹۶ باشد تا این تخمین‌ها از لحاظ آماری معنادار شود. همچنین به منظور سنجش روابط داخل مدل از ضرایب رگرسیونی استاندارد شده استفاده می‌شود. ضرایب استاندارد شده عددی بین ۱- تا ۱+ است چنانچه ضریب استاندارد به عدد صفر نزدیک باشد یعنی تأثیرگذاری متغیر ناچیز است. اگر ضریب استاندارد به عدد ۱ نزدیک باشد یعنی تأثیرگذاری متغیر مثبت و هم‌جهت با متغیر وابسته است. اگر ضریب استاندارد به عدد ۱- نزدیک باشد یعنی تأثیرگذاری متغیر منفی و معکوس با متغیر وابسته است. شکل ۳ مشخص نموده است که تمامی مسیرهای بین متغیرهای مدل مورد تأیید قرار گرفته است و معنادار می‌باشند.

### بحث، نتیجه‌گیری و پیشنهادها

تحولات سریع در عرصه فناوری‌های نوین، به‌ویژه هوش مصنوعی، فرصت‌ها و چالش‌های متعددی را برای نظام‌های بانکی و مالی به همراه داشته است. یکی از مهم‌ترین کاربردهای این فناوری، توسعه سامانه‌های هوشمند تشخیص تقلب است که می‌تواند نقش بسزایی در ارتقای امنیت مالی و اعتماد مشتریان ایفا کند. با این حال، پذیرش و استفاده گسترده از این سیستم‌ها صرفاً به دقت و توانمندی فنی محدود نمی‌شود، بلکه عوامل رفتاری و ادراکی همچون شفافیت عملکرد و نبود سوگیری الگوریتمی، نقش تعیین‌کننده‌ای در اعتمادسازی و در نتیجه پذیرش این فناوری‌ها دارند. پژوهش حاضر با اتخاذ رویکرد معادلات ساختاری، به بررسی روابط میان این متغیرها پرداخته و می‌کوشد تصویری روشن از اهمیت اعتماد به عنوان عامل میانجی در مسیر پذیرش سامانه‌های تشخیص تقلب مبتنی بر هوش مصنوعی ارائه نماید. در ادامه به بررسی فرضیات پژوهش می‌پردازیم.

\* فرضیه اول: اعتماد بر تأثیر شفافیت بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج نقش میانجی دارد. این فرضیه بر پایه نظریه‌های اعتماد و پذیرش فناوری شکل گرفته است و بیان می‌کند که شفافیت در عملکرد سیستم‌های هوشمند، اعتماد کاربران را افزایش می‌دهد و این اعتماد زمینه پذیرش فناوری را فراهم می‌کند. از نظر آماری، آزمون سوبل مقدار آماره‌ای بالاتر از ۱/۹۶ و احتمال دو دنباله کمتر از ۰/۰۵ را نشان می‌دهد که در سطح اطمینان ۹۹ درصد، نقش میانجی اعتماد را تأیید می‌کند و اثر شفافیت بر پذیرش سیستم را تقویت می‌کند.

- یافته‌های پژوهش‌های داخلی مانند: دارائی (۱۴۰۳) و محمدی آریا وسط چی و یارمحمدی (۱۴۰۴) نیز نشان می‌دهند که شفافیت الگوریتم‌ها و کاهش ابهام تصمیم‌گیری موجب افزایش اعتماد کاربران و تسهیل پذیرش فناوری می‌شود.

- علاوه بر این، پژوهش‌های خارجی مانند: یاسین و الامارنه (۲۰۲۵) و شلکه و همکاران (۲۰۲۵) نیز تأکید دارند که شفافیت الگوریتمی اعتماد کاربران را افزایش داده و این اعتماد، نقش میانجی مؤثری در پذیرش سیستم‌های تشخیص تقلب ایفا می‌کند. به این ترتیب، شفافیت از طریق اعتماد کاربران، تأثیر قابل توجهی بر پذیرش سیستم‌های هوشمند دارد و این رابطه در پژوهش‌های داخلی و خارجی همسو و همراستا تأیید شده است.

\* فرضیه دوم: اعتماد بر تأثیر عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج نقش میانجی دارد.

- این فرضیه بر اهمیت عدالت و بی‌طرفی الگوریتم‌های هوش مصنوعی تمرکز دارد و بیان می‌کند که الگوریتم‌های بدون سوگیری موجب ایجاد حس انصاف در کاربران شده و اعتماد آنان را افزایش می‌دهند. اعتماد کاربران نقش میانجی بین عدم سوگیری الگوریتمی و پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی را ایفا می‌کند. از نظر آماری، آزمون سوبل نشان می‌دهد که مقدار آماره بالاتر از  $1/96$  و احتمال دو دنباله کمتر از  $0/05$  است، بنابراین در سطح اطمینان ۹۹ درصد، اعتماد میانجی معناداری است که اثر عدم سوگیری بر پذیرش سیستم را منتقل می‌کند. پژوهش‌های داخلی مانند: حسینی و ارسطو (۱۴۰۳) و محمدی آریا وسط چی و یارمحمدی (۱۴۰۴) نشان داده‌اند که سوگیری الگوریتمی اعتماد کاربران را کاهش می‌دهد و در نتیجه پذیرش سیستم کمتر می‌شود، در حالی که الگوریتم‌های بدون سوگیری اعتماد و پذیرش را افزایش می‌دهند. یافته‌های خارجی مانند: یاسین و الامارنه (۲۰۲۵) و اودیمی و همکاران (۲۰۲۴) نیز نشان می‌دهند که الگوریتم‌های بی‌طرف و منصف، اعتماد کاربران را افزایش داده و نقش میانجی آن در پذیرش سیستم‌های تشخیص تقلب حیاتی است؛ بنابراین، اعتماد کاربران واسطه‌ای مهم در انتقال اثر عدم سوگیری الگوریتمی به پذیرش سیستم‌های هوشمند محسوب می‌شود و این رابطه در پیشینه‌های داخلی و خارجی همسو و همراستا تأیید شده است.

\* فرضیه سوم: شفافیت بر اعتماد در شعب بانک ملی شهر کرج تأثیر معناداری دارد. یافته‌های پژوهش نشان می‌دهد که شفافیت با ضریب تأثیر  $0/73$  و شاخص معناداری  $4/12$ ، بر اعتماد مشتریان در شعب بانک ملی شهر کرج تأثیر قوی و معناداری دارد، به این معنا

که هر چه سطح شفافیت در فرآیندها، اطلاعات و ارتباطات بانکی بیشتر باشد، سطح اعتماد مشتریان به بانک نیز افزایش می‌یابد. شفافیت به‌عنوان یکی از اصول کلیدی حکمرانی مطلوب، زمینه درک بهتر مشتریان از تصمیم‌گیری‌ها و عملکرد بانک را فراهم می‌کند و با کاهش ابهام و رفع نگرانی‌ها، موجب تقویت اعتماد می‌شود. نتایج این پژوهش همسو با یافته‌های دارائی (۱۴۰۳) است که در حوزه هوش مصنوعی توضیح‌پذیر نشان داد شفاف‌سازی تصمیمات فناورانه به افزایش اعتماد کاربران منجر می‌شود و همچنین با پژوهش محمدی آریا و وسط‌چی و یارمحمدی (۱۴۰۴) که بر ضرورت شفافیت الگوریتم‌ها برای کاهش سوگیری و تقویت اعتماد عمومی تأکید داشتند، همراستا است. این موضوع نشان می‌دهد شفافیت نه تنها در حوزه بانکی بلکه در تمام خدمات فناورانه و مالی به‌عنوان یک عامل بنیادین برای ایجاد اعتماد شناخته می‌شود. یافته‌های این مطالعه همچنین با پژوهش یاسین و الامارنه (۲۰۲۵) در امارات و قطر همخوانی دارد که نشان داد، شفافیت سیستم‌های هوش مصنوعی در بانکداری، مهم‌ترین پیش‌بینی‌کننده اعتماد و پذیرش این فناوری است و همچنین پژوهش شلکه و همکاران (۲۰۲۵) نیز بر این نکته تأکید کرد که نبود شفافیت در خدمات آنلاین، موجب کاهش اعتماد مشتریان می‌شود و بانک‌ها باید شفافیت را در اولویت قرار دهند. بنابراین، می‌توان گفت نتایج این فرضیه در سطح داخلی و خارجی همسو و همراستا است و نقش کلیدی شفافیت در ارتقای اعتماد مشتریان به‌ویژه در حوزه بانکی، کاملاً تأیید می‌شود.

\* فرضیه چهارم: عدم سوگیری الگوریتمی بر اعتماد در شعب بانک ملی شهر کرج تأثیر معناداری دارد.

- یافته‌های پژوهش نشان می‌دهد که عدم سوگیری الگوریتمی با ضریب تأثیر ۰/۲۶ و شاخص معناداری ۴/۲۶ بر اعتماد مشتریان در شعب بانک ملی شهر کرج اثرگذار است، هرچند این اثر مثبت در مقایسه با سایر متغیرها ضعیف‌تر است، اما معناداری آن نشان می‌دهد که مشتریان نسبت به انصاف و بی‌طرفی سیستم‌های بانکی حساسیت بالایی دارند؛ به‌بیان دیگر، زمانی که مشتریان احساس کنند الگوریتم‌ها و تصمیم‌گیری‌های بانکی فاقد تبعیض و سوگیری هستند، اعتماد آنان به بانک تقویت می‌شود، حتی اگر شدت این اثرگذاری به اندازه شفافیت نباشد. این موضوع اهمیت توجه به طراحی و نظارت بر الگوریتم‌ها را برجسته می‌سازد. نتایج این پژوهش با یافته‌های محمدی آریا و وسط‌چی و یارمحمدی (۱۴۰۴) همسو است؛ آن‌ها نشان دادند که سوگیری در داده‌های آموزشی و

الگوریتمی منجر به تصمیم‌گیری‌های ناعادلانه شده و اعتماد عمومی به هوش مصنوعی را کاهش می‌دهد و همچنین پژوهش‌داری (۱۴۰۳) درباره هوش مصنوعی توضیح‌پذیر تأکید کرد که یکی از مزایای آن کاهش تعصب‌های موجود در داده‌ها و الگوریتم‌ها است که به افزایش اعتماد کاربران می‌انجامد. این همسویی نشان می‌دهد که در بستر بانکی نیز همانند سایر صنایع، بی‌طرفی الگوریتم‌ها یکی از عوامل مهم جلب اعتماد مشتریان است. این یافته با پژوهش یاسین و الامارنه (۲۰۲۵) همراستا است که نشان داد درک مشتریان از انصاف و بی‌طرفی سیستم‌های هوش مصنوعی، اثر منفی سوگیری الگوریتمی را کاهش داده و اعتماد به سیستم‌های بانکی را افزایش می‌دهد؛ همچنین مطالعه اُدیمی و همکاران (۲۰۲۴) نیز بر اهمیت شیوه‌های مسئولانه و منصفانه در استفاده از هوش مصنوعی برای جلوگیری از بی‌عدالتی الگوریتمی تأکید کرده است. در مجموع می‌توان گفت نتایج این فرضیه در سطح داخلی و خارجی همسو و همراستا است. هرچند شدت تأثیر آن ضعیف‌تر از شفافیت است، اما اهمیت راهبردی بالایی دارد زیرا حتی میزان اندک سوگیری الگوریتمی می‌تواند اعتماد مشتریان را به‌طور جدی تضعیف کند.

\* فرضیه پنجم: شفافیت بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج تأثیر معناداری دارد.

- نتایج پژوهش نشان می‌دهد که شفافیت با ضریب تأثیر ۰/۴۶ و شاخص معناداری ۴/۲۸ بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج اثرگذار است. این تأثیر در سطح متوسط و مثبت گزارش شده و نشان می‌دهد هرچه بانک در خصوص منطق تصمیم‌گیری و نحوه عملکرد سیستم‌های هوش مصنوعی شفاف‌تر عمل کند، میزان پذیرش این فناوری از سوی کارکنان و مشتریان افزایش می‌یابد. شفافیت موجب می‌شود کاربران احساس اطمینان بیشتری نسبت به فرآیندهای تصمیم‌گیری ماشینی داشته و مقاومت احتمالی در برابر استفاده از این فناوری کاهش یابد. یافته‌های حاضر با پژوهش محمدی آریاوسط‌چی و یارمحمدی (۱۴۰۴) همسو است؛ آن‌ها نشان دادند که شفاف‌سازی نحوه عملکرد الگوریتم‌های هوش مصنوعی به افزایش اعتماد و پذیرش کاربران منجر می‌شود و همچنین پژوهش‌داری (۱۴۰۳) درباره هوش مصنوعی توضیح‌پذیر بیان می‌کند که تبیین‌پذیری و شفافیت تصمیم‌ها نه تنها سوگیری را کاهش می‌دهد بلکه احتمال پذیرش فناوری‌های نو توسط ذی‌نفعان را افزایش می‌دهد. در بستر بانکی نیز این یافته تأیید می‌کند که شفافیت یکی از عوامل کلیدی در کاهش تردید نسبت

به استفاده از سیستم‌های تشخیص تقلب است. نتایج این فرضیه با پژوهش اُدیمی و همکاران (۲۰۲۴) همراستا است؛ آنان نشان دادند که شفافیت و توضیح‌پذیری سیستم‌های مبتنی بر هوش مصنوعی در بانکداری الکترونیک از مهم‌ترین عوامل افزایش پذیرش و رضایت کاربران است و همچنین مطالعه یاسین و الامارنه (۲۰۲۵) نیز بیان می‌کند که مشتریان زمانی تمایل بیشتری به پذیرش ابزارهای هوش مصنوعی دارند که فرآیندها برای آنان روشن و قابل درک باشد. به‌طور کلی، یافته‌های این فرضیه در سطح داخلی و خارجی همسو و همراستا است و نشان می‌دهد که شفافیت در سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی، پذیرش آن‌ها را به شکل معناداری افزایش می‌دهد؛ موضوعی که برای موفقیت پیاده‌سازی چنین فناوری‌هایی در صنعت بانکداری ضروری است.

\* فرضیه ششم: عدم سوگیری الگوریتمی بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج تأثیر معناداری دارد.

- نتایج پژوهش نشان می‌دهد که عدم سوگیری الگوریتمی با ضریب تأثیر ۰/۱۷ و شاخص معناداری ۲/۷۱ بر پذیرش سیستم‌های تشخیص تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج اثرگذار است، این اثر در سطح ضعیف، مثبت و معنادار گزارش شده است؛ به‌بیان دیگر، اگرچه رفع یا کاهش سوگیری در الگوریتم‌های هوش مصنوعی به پذیرش این سیستم‌ها کمک می‌کند، اما شدت تأثیر آن کمتر از عواملی همچون شفافیت است. این امر نشان می‌دهد که کاربران و مشتریان در وهله نخست بیشتر به وضوح و قابل درک بودن عملکرد سیستم اهمیت می‌دهند و در مرحله بعد به موضوع بی‌طرفی الگوریتم توجه دارند. یافته با پژوهش محمدی آریا و وسط‌چی و یارمحمدی (۱۴۰۴) همسو است؛ آن‌ها نشان دادند که کاهش سوگیری در الگوریتم‌های هوش مصنوعی می‌تواند اعتماد و پذیرش کاربران را افزایش دهد، اما این تأثیر به‌تنهایی قوی نیست و نیازمند همراهی با سایر عوامل مانند شفافیت است و همچنین پژوهش دارائی (۱۴۰۳) نیز بیان می‌کند که تبیین‌پذیری و شفافیت نقش پررنگ‌تری نسبت به رفع سوگیری در پذیرش فناوری‌های نو دارند، هرچند هر دو عامل در تعامل با هم موجب ارتقای اعتماد می‌شوند. نتایج این فرضیه با پژوهش شلکه و همکاران (۲۰۲۵) همخوانی دارد که نشان داد رفع سوگیری الگوریتمی در سیستم‌های هوش مصنوعی به افزایش پذیرش کاربران منجر می‌شود، اما در مقایسه با شفافیت و سهولت استفاده، اثرگذاری کمتری دارد و همچنین مطالعه گال و همکاران (۲۰۲۴) تأکید می‌کند که در حوزه بانکداری دیجیتال، کاربران ابتدا به شفافیت و

توضیح‌پذیری سیستم‌ها توجه دارند و سپس به موضوع عدالت و عدم سوگیری الگوریتمی واکنش نشان می‌دهند. به‌طور کلی، نتایج این فرضیه در سطح داخلی و خارجی همسو و همراستا است و نشان می‌دهد هرچند عدم سوگیری الگوریتمی به پذیرش سیستم‌های تشخیص‌تقلب کمک می‌کند، اما اثر آن در مقایسه با شفافیت ضعیف‌تر است؛ بنابراین برای ارتقای سطح پذیرش، باید هر دو عامل به صورت مکمل مورد توجه قرار گیرند.

\* فرضیه هفتم: اعتماد بر پذیرش سیستم‌های تشخیص‌تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج تأثیر معناداری دارد.

نتایج پژوهش نشان می‌دهد که اعتماد با ضریب تأثیر  $0/87$  و شاخص معناداری  $3/82$ ، قوی‌ترین عامل مؤثر بر پذیرش سیستم‌های تشخیص‌تقلب مبتنی بر هوش مصنوعی در شعب بانک ملی شهر کرج است. این تأثیر قوی، مثبت و معنادار ارزیابی می‌شود، به این معنا که مشتریان و کارکنان بانک زمانی تمایل بیشتری به استفاده و پذیرش این سیستم‌ها دارند که اعتماد کافی به عملکرد، امنیت و نتایج آن داشته باشند. اعتماد، نقش یک متغیر بنیادین در پذیرش نوآوری‌های فناورانه دارد و نبود آن می‌تواند موجب مقاومت در برابر تغییر و استفاده از فناوری‌های جدید شود. این یافته با پژوهش دارائی (۱۴۰۳) در زمینه هوش مصنوعی توضیح‌پذیر همسو است که نشان داد اعتماد، متغیر واسطه‌ای کلیدی میان شفافیت و پذیرش فناوری‌های هوش مصنوعی است و همچنین پژوهش محمدی آریاوسط‌چی و یارمحمدی (۱۴۰۴) نیز تأکید می‌کند که اعتماد، پایه اصلی پذیرش و تعامل مثبت با سیستم‌های هوش مصنوعی است و بدون آن، حتی شفاف‌ترین و بی‌طرف‌ترین الگوریتم‌ها نیز با مقاومت کاربران مواجه خواهند شد. نتایج حاضر با پژوهش یاسین و الامارنه (۲۰۲۵) همراستا است که در کشورهای امارات و قطر نشان داد اعتماد مهم‌ترین عامل پذیرش ابزارهای بانکی مبتنی بر هوش مصنوعی است و همچنین مطالعه اُدیمی و همکاران (۲۰۲۴) نیز نشان داد در بانکداری دیجیتال، اعتماد کاربران بیش از هر عامل دیگری بر قصد استفاده و پذیرش سیستم‌های مبتنی بر هوش مصنوعی اثرگذار است. به‌طور کلی، یافته‌های این فرضیه در سطح داخلی و خارجی کاملاً همسو و همراستا است و تأکید می‌کند که اعتماد، زیربنای اصلی پذیرش سیستم‌های تشخیص‌تقلب مبتنی بر هوش مصنوعی است. بنابراین بانک‌ها باید با تقویت شفافیت، امنیت داده‌ها و عدالت الگوریتمی، بستر لازم برای ایجاد و حفظ اعتماد مشتریان و کارکنان را فراهم کنند.

در ادامه پیشنهادهایی را که برگرفته از یافته‌های پژوهش مطرح می‌گردد:

بانک باید شفافیت در عملکرد و خروجی سیستم‌ها را افزایش دهد تا اعتماد کاربران شکل گیرد و این اعتماد پذیرش را تسهیل کند.

سیستم باید نتایج تصمیم‌گیری خود را به شکل قابل فهم نمایش دهد تا کاربران علاوه بر مشاهده شفافیت، اعتمادشان نیز افزایش یابد.

برگزاری دوره‌های آموزشی و کمپین‌های اطلاع‌رسانی برای مشتریان، هم شفافیت را نشان می‌دهد و هم اعتماد لازم برای پذیرش سیستم را ایجاد می‌کند.

بانک باید گزارش‌هایی منتشر کند که نشان دهد تصمیمات سیستم منصفانه است و کاربران با مشاهده شواهد، اعتماد بیشتری پیدا می‌کنند.

جمع‌آوری بازخورد کاربران و اصلاح الگوریتم‌ها در مواردی که تصمیمات ناعادلانه گزارش شده، موجب افزایش اعتماد و پذیرش می‌شود.

ارائه تأییدیه از نهادهای معتبر برای بی‌طرفی سیستم، به کاربران اطمینان می‌دهد که سیستم عادلانه عمل می‌کند و این امر پذیرش را تقویت می‌کند.

بانک ملی می‌تواند گزارش‌های دوره‌ای درباره عملکرد سیستم‌های مالی و امنیتی منتشر کند. این گزارش‌ها باید شامل: شاخص‌های کلیدی، میزان کشف تقلب و اقدامات انجام شده باشند تا مشتریان و کارکنان دید کاملی از عملکرد سیستم داشته باشند و اعتمادشان افزایش یابد.

متون قراردادی و اطلاعیه‌های بانکی باید ساده و روان باشند و از اصطلاحات پیچیده خودداری شود تا مشتریان به راحتی عملکرد سیستم را درک کنند و احساس اطمینان کنند.

راه‌اندازی یک درگاه آنلاین برای پاسخ به سؤالات مشتریان درباره عملکرد سیستم‌های هوش مصنوعی و نحوه تشخیص تقلب، باعث افزایش شفافیت و کاهش نگرانی‌های کاربران خواهد شد.

بانک باید تیم تخصصی برای ارزیابی و کاهش سوگیری‌های احتمالی در داده‌ها و الگوریتم‌ها تشکیل دهد؛ این کار باعث می‌شود کاربران اطمینان پیدا کنند که سیستم به صورت منصفانه عمل می‌کند.

الگوریتم‌ها باید با داده‌های نماینده از تمامی گروه‌های مشتریان آموزش داده شوند تا از تبعیض جلوگیری شود و اعتماد کاربران افزایش یابد.

## منابع

- ولیپور، مهرداد، ۱۴۰۴، تشخیص تقلب مبتنی بر هوش مصنوعی و پیش بینی بازار انرژی، بیست و سومین کنفرانس ملی اقتصاد، مدیریت و حسابداری، شیروان، <https://civilica.com/doc/2266673>.
- محمدی آریا وسط چی، علیرضا و یارمحمدی، مرضیه، ۱۴۰۴، هوش مصنوعی: چالش های اخلاقی عصر حاضر، اولین همایش بین المللی هوش مصنوعی در آموزش و پرورش، روانشناسی، علوم تربیتی و مطالعات دینی، فرهنگی، اجتماعی و مدیریتی در هزاره سوم، بوشهر، <https://civilica.com/doc/2218865>.
- حسینی، الهام و ارسطو، ایمان، ۱۴۰۳، اثرات اعتماد شناختی و اعتماد اجتماعی بر قصد پذیرش عامل دیجیتال هوش مصنوعی با نقش میانجی اعتماد عاطفی در بین مشتریان بیمه سامان، سی و یکمین همایش ملی و دوازدهمین همایش بین المللی بیمه و توسعه: رضایت مندی و اعتماد مردم به صنعت بیمه، تهران، <https://civilica.com/doc/2148932>.
- دارائی، محمد، ۱۴۰۳، هوش مصنوعی توضیح پذیر: گامی به سوی شفافیت و اعتماد در هوش مصنوعی، <https://civilica.com/doc/2155766>.
- روزبهنانی، زهرا و نظری، سارا و صراف زاده جهرمی، مجتبی، ۱۴۰۳، بررسی سیستم های تشخیص نفوذ مبتنی بر هوش مصنوعی، دومین کنفرانس بین المللی پژوهش ها و فناوری های نوین در مهندسی برق، تهران، <https://civilica.com/doc/2083295>.
- اکبرالسادات سیدمجید، & اسماعیل پور بابک. (۱۴۰۱). مروری بر روش های کشف تقلب بانکی با استفاده از هوش مصنوعی، فصلنامه علمی مؤسسه آموزش عالی فردوس، سال دوم | شماره ششم | پاییز ۱۴۰۱.
- شعری آنافیز، صابر و خراسانی، ابوطالب (۱۳۹۶). واکاوی مفهوم تقلب و بررسی آثار به کارگیری استانداردهای حسابرسی در افشای اطلاعات گزارشگری مالی متقلبان. اولین همایش
- Abdulrahman, M. H. (2019). The impact of Artificial Intelligence (AI) in detecting fraud in the UAE. *Electronic Interdisciplinary Miscellaneous Journal*, 17(10), 1-19.
- Agheli, M., & Ajorloo, F. (2018). The Effect of brand journalism on customers' repatronage intention towards local. *Quarterly Journal of Brand Management*, 5(1), 135-168.
- Agheli, M., NikMenesh, S., Rashidi, H., & Jalali, P. (2023). Training on thesis writing and scientific article writing. Tehran: Dibagaran Book Institute. (In Persian)

- Ahmadi, S. (2022). Advancing Fraud Detection in Banking: Real-Time Applications of Explainable AI (XAI). *Journal of Electrical Systems*, 18(4), 141-150.
- Ayeni, T. J., Durotoye, E. O., & Eriabie, S. (2024, April). Adoption of artificial intelligence for fraud detection in deposit money banks in Nigeria. In 2024 international conference on science, engineering and business for driving sustainable development goals (SEB4SDG) (pp. 1-5). IEEE.
- Bahnsen, A. C., Aouada, D., Stojanovic, A., & Ottersten, B. (2020). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51, 134-142. <https://doi.org/10.1016/j.eswa.2019.12.028>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). Fairness and machine learning. [fairmlbook.org](http://fairmlbook.org).
- Basri, W. S., & Almutairi, A. (2023). Enhancing Financial Self-efficacy through Artificial Intelligence (AI) in Banking Sector. *International Journal of Cyber Criminology*, 17(2), 284-311.
- Chen, I. Y., Johansson, F. D., & Sontag, D. (2020). Why is my classifier discriminatory? *Advances in Neural Information Processing Systems*, 32, 3539–3550.
- Dayalan, P., & Sundaramurthy, B. (2025). Exploring the Implementation and Challenges of AI-Based Fraud Detection Systems in Financial Institutions: A Review. *Creating AI Synergy Through Business Technology Transformation*, 25-38.
- Doshi-Velez, F., & Kim, B. (2019). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
- Eskandarany, A. (2024). Adoption of artificial intelligence and machine learning in banking systems: a qualitative survey of board of directors. *Frontiers in Artificial Intelligence*, 7, 1440051.
- Goodfellow, I., Bengio, Y., & Courville, A. (2018). *Deep learning*. MIT Press.
- Islam, M. Z., Shil, S. K., & Buiya, M. R. (2023). AI-driven fraud detection in the US financial sector: Enhancing security and trust. *International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence*, 14(1), 775-797.
- Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. *Expert Systems with Applications*, 100, 234-245. <https://doi.org/10.1016/j.eswa.2018.01.034>
- LeCun, Y., Bengio, Y., & Hinton, G. (2020). Deep learning. *Nature*, 521(7553), 436-444.

- Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM*, 61(10), 36-43. <https://doi.org/10.1145/3233231>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2021). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 140, 113428. <https://doi.org/10.1016/j.dss.2020.113428>
- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2021). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 140, 113428. <https://doi.org/10.1016/j.dss.2020.113428>
- Nguyen, G. H., Tran, N. H., Ngo, T. D., & Phung, D. (2021). An effective approach to credit card fraud detection using CNN and LSTM. *IEEE Access*, 9, 15323-15333. <https://doi.org/10.1109/ACCESS.2021.3050887>
- Ozioko, A. C. (2024). The Use of Artificial Intelligence in Detecting Financial Fraud: Legal and Ethical Considerations. *Multi-Disciplinary Research and Development Journals Int'l*, 5(1), 66-85.
- 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>
- Shelke, P., Suganthiya, Dr. M. S., Sharma, Prof. Dr. B. (2025). A Study on the Effectiveness of Artificial Intelligence Based Fraud Detection in Online Banking. *International Journal of Research Publication and Reviews*, Vol (6), Issue (4), April (2025), Page – 15303-15308.
- Veale, M., & Binns, R. (2020). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717743530>
- Weller, A. (2019). Transparency: Motivations and challenges. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 17–23. <https://doi.org/10.1145/3306618.3314238>
- Yaseen, H., & Al-Amarneh, A. A. (2025). Adoption of artificial Intelligence-driven fraud detection in banking: the role of trust, transparency, and fairness perception in financial institutions in the

United Arab Emirates and Qatar. *Journal of Risk and Financial Management*, 18(4), 217.

Zhang, B., Dafoe, A., & Chen, J. (2022). Transparency and trust in AI: The role of bias mitigation. *Annual Review of Sociology*, 48, 205-222. <https://doi.org/10.1146/annurev-soc-123521-102931>