مقاله پژوهشی

مروری بر روش‌های کشف تقلب بانکی با استفاده از هوش مصنوعی

Doi: 10.30508/kdip.2022.349333.1044

سیدمجید اکبرالسادات (نویسنده مسئول)[۱] | بابک اسماعیل پور[۲]

۱- گروه هوش مصنوعی و رباتیکز، دانشگاه آزاد اسلامی (واحد قوچان)، قوچان، ایران
۲- استادیار گروه کامپیوتر، دانشگاه آزاد اسلامی (واحد مشهد)، مشهد، ایران

تاریخ دریافت: ۱۴۰۱/۰۴/۰۶
تاریخ پذیرش: ۱۴۰۱/۱۰/۰۶
صفحه: ۴۱ - ۲۰

چکیده:

تقلب بانکی با توجه به تأثیرات مخرب آن، یکی از تهدیدآمیزترین مشکلاتی است که هر جامعه بشری با آن دست‌وپنجه 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 به‌صورت 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.

کلمات کلیدی: تقلب بانکی، بانکداری الکترونیک، کلاه‌برداری، معاملات متقلبانه، تراکنش‌های جعلی، کشف تقلب بانکی.

۱- مقدمه

تقلب با توجه به تأثیرات مخرب آن، یکی از تهدیدآمیزترین مشکلاتی است که هر جامعه بشری با آن دست‌وپنجه نرم می‌کند. این عمل به استفاده عمدی از اطلاعات نادرست برای کلاه‌برداری از پول یا دارایی یک فرد یا سازمان دیگر اشاره دارد (ویلیام[1]، ۲۰۲۲). در طول رکود اقتصادی ناشی از همه‌گیری کرونا، به دلیل افزایش تعداد افراد بیکار در طول دوره، شاهد افزایش فعالیت‌های کلاه‌برداری بوده‌ایم (شیهم بیتسا[2]، ۲۰۲۱). هزاران نفر بیکار شدند، حقوق‌ها کاهش یافت و نرخ بیکاری افزایش یافت. به طور طبیعی، افراد بیشتری منابع کمتری برای بقای خود در اختیار دارند، که این انگیزه‌ای برای افزایش فعالیت‌های کلاه‌برداری می‌باشد، زیرا مردم در تلاش برای زنده ماندن در دوران سخت اقتصادی هستند. کارشناسان تقلب استدلال می‌کنند که تقلب از سه عنصر که شامل فشار، فرصت و منطقی سازی است؛ نشأت می‌گیرد (کارتیس[3]، ۲۰۱۲). به گفته محققین، کلاه‌برداری زمانی اتفاق می‌افتد که فردی فشار و انگیزه غیرقابل تقسیم برای ارتکاب کلاه‌برداری ایجاد کند. در بیشتر موارد، برای مرتکب تقلب، نیازهای برآورده نشده بی‌پایان است و برای افراد مختلف متفاوت می‌باشد. که می‌تواند شامل؛ افزایش صورت‌حساب پزشکی، کاهش درآمد در خانواده یا بدهی قمار باشد. هنگامی‌که فرد نیازهای برآورده نشده داشته باشد و منابع محدودی داشته باشد، فرصت ارتکاب تقلب را شناسایی می‌کند. ادراک فرصت‌ها ممکن است مدیریت بی‌پروا یا فقدان کنترل‌های داخلی در یک واحد تجاری باشد که تقلب را به یک فعالیت آسان تبدیل می‌کند. درنهایت، فرد تصمیم خود برای ارتکاب تقلب با متقاعد کردن خود مبنی بر اینکه به پول بیشتری نیاز دارد یا درنهایت آن را پس می‌دهد، منطقی می‌کند. با توجه به شرایط اقتصادی، فشارها و انگیزه‌هایی برای ارتکاب کلاه‌برداری افزایش‌یافته است که این امر باعث می‌شود کلاه‌برداران به‌راحتی اقدامات خود را منطقی کنند. تقلب در صنعت بانکداری رایج است و شامل؛ فیشینگ ایمیل، کلاه‌برداری از کارت اعتباری، پول‌شویی، تقلب در درخواست وام، تقلب در صورت‌های مالی و کلاه‌برداری سایبری می‌شود. با ظهور بانکداری دیجیتال، کلاه‌برداری دیجیتال نیز در این بخش رایج‌تر شده است. بنابراین، مهم است که اذعان کنیم که مدیریت تقلب در صنعت بانکداری و بازرگانی ضروری شده است، که مسلماً فرآیندی طاقت‌فرسا است. کلاه‌برداران در کشف حفره‌ها ماهر شده‌اند و تکنیک‌های مؤثری مانند فیشینگ برای افراد ناآگاه و کلاه‌برداری خلاقانه از آنها ایجاد کرده‌اند (ماریوت[4]، ۲۰۲۱). بنابراین، روش‌های کشف کلاه‌برداری باید به‌طور مداوم تکامل یابند زیرا کلاه‌برداران در طراحی تکنیک‌هایی مؤثرتر می‌شوند که سیستم‌های امنیتی سفت و سخت بانکی را دور می‌زنند و یاد می‌گیرند که چگونه افراد ناآگاه را متقاعد کنند که پول خود را در اختیار آنها بگذارند. تقلب مربوط به پرداخت یکی از جنبه‌های کلیدی آژانس‌های جرائم سایبری است و تحقیقات اخیر نشان داده است که استفاده از تکنیک‌های یادگیری ماشین برای شناسایی تراکنش‌های جعلی در مقادیر زیادی از داده‌های پرداخت مؤثر است. چنین تکنیک‌هایی توانایی شناسایی معاملات متقلبانه‌ای را دارند که حسابرسان انسانی ممکن است نتوانند آن‌ها را شناسایی کنند و همچنین این کار را به‌صورت بلادرنگ انجام دهند (نیکولاس، کوپا و لیخاک[5]، ۲۰۲۱). کشف دقیق تقلب[6] (FD) می‌تواند اعتماد مشتریان بانک را در انجام امور بانکی‌شان به‌صورت آنلاین افزایش دهد. با پیشرفت تکنولوژی، تقلب به‌طور چشمگیری افزایش می‌یابد، که منجر به زیان‌های قابل‌توجهی برای شرکت‌ها می‌شود، بنابراین شناسایی تقلب به مسئله‌ای حائز اهمیت تبدیل شده است (کو، لو، سیوانگتان و هانگ[7]، ۲۰۰۴). بررسی شاخص ایمنی رایانه مایکروسافت نشان داد که تأثیر سالانه فیشینگ و اشکال مختلف سرقت هویت می‌تواند

1- Williams
2- Shihembetsa
3- CURTIS
4- Maruti
5- Nicholls, Kuppa, & Le-Khac
6- Fraud detection
7- Kou, Lu, Sirwongwattana, & Huang

به ۵ میلیارد دلار آمریکا برسد، درحالی‌که هزینه ترمیم آسیب به شهرت افراد آنلاین می‌تواند تا ۶ میلیارد دلار یا بیشتر باشد. به‌طور متوسط برای هر ضرر ۶۳۲ دلار آمریکا تخمین زده می‌شود (مارکیان و لیم[1]، ۲۰۱۴). روش‌های سنتی تشخیص تقلب در صنعت بانکداری مبتنی بر قانون بوده است که در آن انسان‌ها قوانین را تعریف می‌کنند. ۹۰ درصد مؤسسات مالی و بانکی بر این روش‌ها تکیه می‌کنند. درحالی‌که افراد بیشتری از فناوری‌های جدید استفاده می‌کنند، سناریوهای تقلب بیشتری ممکن است؛ اتفاق بیفتد و این روش‌های مبتنی بر قوانین را در آینده غیرقابل مهروموم و ناپایدار می‌سازد. علاوه بر این، مثبت کاذب (یعنی تراکنش‌های غیر متقلبانه که به‌عنوان جعلی فهرست‌بندی شده‌اند) باعث ضرر میلیون‌ها دلاری در معاملات و شکایات مشتریان در صنعت بانکداری می‌شود. روش‌های مبتنی بر قانون کمک زیادی به این نتایج می‌کنند. کیوبانو[2] (۲۰۲۰) مطالعه‌ای را با ۱۰۰۰ مصرف‌کننده بزرگسال انجام داد و در آن متوجه شد حدود ۲۵ درصد از آنها که تراکنش‌هایشان رد شده بود، به تجارت با رقبا تمایل داشتند. این میزان روی آوردن به رقبا برای مصرف‌کنندگان ۱۸ تا ۲۴ ساله به ۳۶ درصد افزایش‌یافته است. همچنین برای افراد بین ۲۵ تا ۳۴ سال به ۳۱ درصد افزایش‌یافته است. این نتایج مطالعه نشان‌دهنده نیاز مبرم به روش‌های دقیق‌تر و مدرن‌تر کشف تقلب است. به جای جلوگیری از تراکنش غیرمجاز، سیستم‌های بانکداری اینترنتی کشف تقلب باید بلافاصله کلاهبرداری‌ها را در یک حساب در معرض خطر شناسایی کنند (ریچاردز[3]، ۲۰۰۳). رویکرد مبتنی بر قوانین متعارف برای کسب دانش (KA) برای یک تجارت بسیار کند، کار فشرده و پرهزینه‌ای است. شکنندگی نیز یکی از کاستی‌های پایه‌های قوانین مرسوم بود و این سیستم‌ها همیشه سعی می‌کنند پاسخی بدهند، حتی ممکن است نادرست باشد. شکنندگی به نرم‌افزاری اطلاق می‌شود که در مواجهه با یک موقعیت غیرقابل‌پیش‌بینی ممکن است اشتباه کند. بنابراین، دقت کمتری دارد، زیرا همیشه به دانش فعلی خود اعتماد دارد؛ حتی برای مواردی که دانش کافی نیست.

کشف تقلب برای بانکداری آنلاین یک زمینه مطالعاتی بسیار مهم است، زیرا مجرمان سایبری حملات جدید کلاه‌برداری پیچیده‌ای را به‌صورت روزانه طراحی می‌کنند، بنابراین این امر مستلزم آن است که محققان تکنیک‌های جدید کشف تقلب را به‌طور مداوم توسعه دهند. طبق نظر (ماراتونا[4]، ۲۰۱۳) در دسترس بودن اطلاعات دقیق در مورد سیستم کشف تقلب بسیار محدود است زیرا صنعت بانکداری به‌ندرت آمار کشف تقلب را منتشر می‌کند. به ویژه، تأمین‌کنندگان امنیت مؤسسات مالی، شرکت‌های شخص ثالثی هستند که از مالکیت معنوی خود در برابر رقبا محافظت می‌کنند. بنابراین، هم بانک‌ها و هم آژانس‌های فناوری اطلاعات؛ اطلاعات سیستم‌های امنیتی خود را منتشر نمی‌کنند (بولتون و هند[5]، ۲۰۰۲). همچنین، توسعه روش‌های جدید برای شناسایی تقلب دشوار است زیرا تبادل افکار در مورد شناسایی تقلب بسیار محدود است، اما نویسندگان بر این مفهوم تأکید می‌کنند که تکنیک‌های کشف تقلب نباید به‌طور عمومی بیان شود. در غیر این صورت مجرمان ممکن است از همان داده‌ها بهره بگیرند (فوآ، لی، اسمیت و گایلر[6]، ۲۰۱۰). اغلب دو انتقاد اصلی از داده‌کاوی برای شناسایی تقلب وجود دارد: کمبود اطلاعات تجربی واقعی در دسترس عموم، و فقدان تکنیک‌ها و روش‌های به‌خوبی تبلیغ‌شده است.

در دو دهه گذشته، با تکامل فناوری مورد استفاده در بخش بانکداری مالی، طرح‌های تقلب مورد استفاده کلاه‌برداران نیز تغییر کرد (استوجنا[7]، ۲۰۱۹). دو حوزه اصلی که در آن تقلب صورت می‌گیرد، شامل؛ برنامه‌های کاربردی وب یا موبایل بانکداری اینترنتی و پرداخت‌های ATM، POS یا تاجر آنلاین با استفاده از کارت‌های بانکی می‌باشد. گزارش نیلسون[8] (۲۰۲۰). بر افزایش هدف قرار دادن بازرگانان

1- Marican, & Lim
2- Ciobanu
3- Richards
4- Maruatona
5- Bolton, & Hand
6- Phua, Lee, Smith, & Gayler
7- Stojanov
8- Nilson Report

توسط سازمان‌های جرائم مالی سازمان‌یافته برای ارتکاب کلاه‌برداری، با توسعه فناوری اطلاعات کشور و بازرگان که بر توانایی تاجر برای جلوگیری از تقلب تأثیر می‌گذارد، تأکید کرده است (گابینو، بریک، ماری، میهایی و اسچو[1]، ۲۰۲۱؛ ناتان، ویکتور، گان و کات[2]، ۲۰۱۹). حدود ۵۶ درصد از اروپایی‌ها نگران این هستند که قربانی تقلب شوند (گابینونا، و همکاران، ۲۰۲۱). در سال ۲۰۱۹، ۲۶ درصد از جمعیت اتحادیه اروپا گزارش دادند که پیام‌های جعلی دریافت کرده‌اند، از جمله پیام‌های مربوط به اعتبار بانکداری الکترونیک (اروستا و کامپوننت[3]، ۲۰۲۰). خانواده‌های مختلف بدافزار آسیب‌های مختلفی به مصرف‌کننده، زیرساخت‌های حیاتی، مؤسسات مالی و بانکی وارد کرده‌اند و به اهداف مورد نظر تبدیل شده‌اند (اسچو، گافتیا، آچیم، و کوتاک[4]، ۲۰۲۰). مکانیسم‌های تقلب از نظر بانکداری اینترنتی و تراکنش‌های کارت دارای ویژگی‌هایی هستند که می‌تواند به کشف تقلب کمک کند، مانند مکان شروع پرداخت، جزئیات گیرنده پرداخت، مهر زمانی پرداخت.

برای افزودن به چالش سیستم سنتی مبتنی بر قوانین، کلاه‌برداران فاقد الگوهای خاص هستند و دائماً رفتار خود را در طول زمان تغییر می‌دهند و سیستم‌ها را دست‌وپاگیر و به سرعت منسوخ می‌کنند. نیاز آشکار به تغییر رویکرد در سیستم‌های امنیتی در سیستم‌های بانکی وجود دارد. طبق گزارش نیلسون (۲۰۲۰)، پیش‌بینی می‌شد که کلاه‌برداری مربوط به کارت‌ها تا سال ۲۰۲۰ تنها به مبلغ حیرت‌انگیز ۳۰ میلیارد دلار در سطح جهان برسد. علاوه بر این، با اختلال فناوری در بخش بانکی به دلیل وجود کانال‌های پرداخت متعدد مانند کارت‌های اعتباری و نقدی، گوشی‌های هوشمند، نرخ تراکنش‌ها طی چند سال گذشته به‌طور تصاعدی افزایش‌یافته است. کلاه‌برداران همچنین تاکتیک‌های کلاه‌برداری بسیار مؤثری را توسعه داده‌اند. با توجه به این وضعیت، نیاز به توسعه رویکردهای سخت و قوی‌تر کشف تقلب در بانک‌ها وجود دارد. عملی‌ترین گزینه الگوریتم‌های یادگیری ماشین نصب شده در سیستم‌های بانکی هستند (گابینو و همکاران، ۲۰۲۱). 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 شناسایی وجود دارد که بتوانند با افزایش سریع تقلب‌ها و موارد پول‌شویی مقابله کنند. لذا این مقاله با هدف بررسی روش‌های کشف تقلب بانکی، به مرور ادبیات اخیر مرتبط با این موضوع می‌پردازد.

۲- مبانی نظری

ادبیات تخصصی و راه‌حل‌های موجود در بازار در دهه گذشته بر جمع‌آوری مقادیر قابل‌توجهی از داده‌ها در مورد تراکنش‌ها (و رفتار کاربر) و اصلاح الگوریتم‌های مورد استفاده برای شناسایی تقلب متمرکز شده‌اند. در همین راستا، قوانین اتحادیه اروپا (به‌عنوان‌مثال، PSD۲) به‌منظور تحمیل تعهدات به ذینفعان برای شناسایی تقلب به تصویب رسیده است. با این حال، از یک‌سو قانون تشریح سطح بالایی از این الزام قانونی ارائه می‌کند و از سوی دیگر، راه‌حل‌های موجود در بازار از نظر داده‌های جمع‌آوری‌شده و به‌ویژه تلاش برای تجمیع داده‌ها به‌منظور تولید متنوع می‌باشد. نتایج دقیق‌تر منجر به مبحثی می‌شود که هنوز در ادبیات تخصصی یا توسط قانون‌گذاران به‌طور عمیق مورد تجزیه‌وتحلیل قرار نگرفته است. محققانی مانند کارمیناتی و همکاران[5] (۲۰۱۸)، در دهه گذشته شیوه‌های مختلفی را که می‌توان از چنین جزئیاتی برای کشف و پیشگیری از تقلب استفاده کرد، تحلیل کرده‌اند (کارمینیات، و همکاران[6]، ۲۰۱۸). نتیجه‌گیری‌های آن‌ها منجر به رویکردهای مختلفی برای الگوریتم‌های کشف تقلب، با تأکید بر روش‌های یادگیری ماشین شده است (بائو، هیلاری، و کی[7]، ۲۰۲۲)؛ زیرا تجمیع داده‌ها و تحلیل

1- Gbudeanu, Brici, Mare, Mihai, & cheau
2- Nathan,Victor, Gan, & Kot
3- Eurostat, & Components
4- cheau,Gaftea, Achim, & Cotoc
5- Carminati
6- Carminati, et al
7- Bao, Hilary, & Ke

تاریخی داده‌ها می‌توانند به یافتن الگوهای تقلب کمک کند (پولیتو، آلیپس، و پاتساکیتس[1]، ۲۰۱۹). این الگوهای تقلب می‌توانند به شناسایی یا پیش‌بینی تراکنش‌های تقلبی احتمالی کمک کنند. الگوریتم‌های تشخیص، ویژگی‌های موجود تقلب‌ها را با تراکنش‌های فعلی در حال تجزیه‌وتحلیل مطابقت می‌دهند، درحالی‌که الگوریتم‌های پیش‌بینی تلاش می‌کنند تا کلاه‌برداری‌هایی را شناسایی کنند که ویژگی‌های متفاوتی نسبت به تقلب‌های تاریخی دارند. محققان عموماً بر دقت نتایج و افزایش جمع‌آوری داده‌ها و تجمیع داده‌ها برای دستیابی به این هدف، هم برای الگوریتم‌های کارآگاهی و هم برای الگوریتم‌های پیش‌بینی تمرکز کرده‌اند (جها، گالین، و وستلند[2]، ۲۰۱۲). با توجه به انواع الگوریتم‌های تشخیص تقلب پیشنهاد شده در مقالات تحقیقاتی منتشر شده در سه سال گذشته (۲۰۱۸ - ۲۰۲۱)، موجود در پنج پایگاه داده تحقیقاتی (ACM، Science Direct، Emerald Full text، IEEE Transactions، ژورنال‌های Springer-Link)، بر اساس کلیدواژه‌های خاصی با هدف شناسایی جنبه‌های حریم خصوصی در نظر گرفته‌شده توسط این الگوریتم‌ها ("تشخیص تقلب بانکی GDPR"، "تشخیص تقلب بانکی حریم خصوصی"، "تکنیک‌های تشخیص تقلب بانکی تجمعی"، "تکنیک‌های تشخیص تقلب بانکی")؛ انواع الگوریتم‌های تشخیص تقلب بانکی در جدول شماره (۱) ارائه شده است.

جدول ۱- تعداد مقالات منتشر شده در سه سال اخیر در خصوص انواع الگوریتم‌های تشخیص تقلب بانکی

پایگاه داده / عبارت کلیدی	ژورنال‌های Springer-Link	IEEE	Emerald	Science Direct	کتابخانه دیجیتال ACM
تشخیص تقلب بانکی GDPR	۱۶۷	۷۲	۲۱	۱۱۳	۵۵۲
تشخیص تقلب در بانکداری حریم خصوصی	۸۹۱	۵۵۴	۱۹۳	۲۵۶	۵۹۱
تکنیک‌های تشخیص تقلب بانکی تجمیعی	۲۸۸	۱۸۱	۴۴	۹۰	۷۷۶

1- Politou, Alepis, & Patsakis
2- Jha, Guillen, & Westland

از سال ۲۰۱۸، مطالعات مربوط به سیاست حفظ حریم خصوصی داده‌های شخصی[1] (سیاست GDPR) شروع به ظاهر شدن کردند. استالابوردیلون، پیارس و تسکالاسکی[2] (۲۰۱۸) یک تحلیل بین‌رشته‌ای از GDPR در زمینه طرح‌های شناسایی الکترونیکی انجام داد. این مطالعه در بریتانیا انجام شد و یک نمای کلی از نحوه تفسیر این موقعیت‌ها ارائه کرد. در این مطالعه، نویسنده یک مبنای قانونی را پیشنهاد می‌کند که می‌تواند به مدیریت خوب حریم خصوصی توسط هر دو طرف درگیر کمک کند. علاوه بر این، در ارتباط با این موضوع، مطالعاتی در مورد میزان نفوذ در درخواست داده‌های شخصی شروع شده است. در این رابطه هوراک، استاپکا و هوساک[3] (۲۰۱۹) به مسائل مربوط به تأثیر GDPR بر نرم‌افزار امنیت سایبری، از نظر عملیاتی پیشگیری از حادثه و رسیدگی به حوادث توجه کردند. خطرات نقض محرمانه بودن و اصل به حداقل رساندن داده‌ها با درخواست داده‌های شخصی مورد بررسی قرار گرفت، و همچنین این واقعیت که به اشتراک‌گذاری این اطلاعات توسط مشتری می‌تواند نگرانی‌های مشابهی را ایجاد کند. ارزیابی تأثیر حریم خصوصی داده‌ها[4] (DPIA) که برای این سناریو انجام شد، نشان داد که با توجه به مکانیسم‌های کاهش خطر خاص، ریسک‌ها بالا نیستند. روش مورد استفاده در این مقاله به درک ریسک‌های موجود و مدیریت آسان‌تر آنها کمک کرد.

این مسائل مزاحمت حریم خصوصی نیز ارتباط نزدیکی با جلوگیری از تقلب کارت اعتباری دارد. تعداد قابل‌توجهی از نویسندگان به موضوع کشف تقلب پرداخته‌اند و الگوریتم‌های مورد استفاده به منظور ترکیب مکانیسم‌ها و طرح‌های جدید مورد استفاده توسط کلاه‌برداران به‌طور مداوم در حال بهبود هستند. تعدادی از مطالعات وجود دارد که تقلب کارت را از دیدگاه‌های مختلف تجزیه‌وتحلیل می‌کند: تشخیص تقلب کارت اعتباری با استفاده از الگوریتم‌های یادگیری ماشین، تشخیص تقلب در کارت اعتباری با استفاده از روش‌های Pipeling و Ensemble Learning (باگا، گویال، گوپتا و گویال[5]، ۲۰۲۰)، و تشخیص تقلب در کارت اعتباری با استفاده از شبکه عصبی مصنوعی (پینتو، و سابریرو[6]، ۲۰۲۲). ابزارهای مورد استفاده برای شناسایی فعالیت‌های متقلبانه برای الگوهای تقلب شناسایی‌شده، سال‌به‌سال پیچیده‌تر و کارآمدتر می‌شوند و الگوریتم‌های یادگیری ماشین یکی از بهترین گزینه‌ها در این زمینه هستند، زیرا پیشرفت‌های فناوری از مرزها عبور می‌کنند و موارد بیشتری در دسترس قرار می‌گیرند (بیلگین، مارکو و دمیر[7]، ۲۰۱۲). اوگرک. ام، اوگرک، ای. و باهتیر[8] (۲۰۱۹) سعی کردند روش‌های تقلب کارت اعتباری را با استفاده از مدلی با مجموعه داده Kaggle ارزیابی کنند. روش انتخاب شده شبکه‌های عصبی مصنوعی چند لایه[9] (MANN) بود. برای شناسایی تقلب، این محققین از ویژگی‌هایی مانند: نقدی/خروجی، بدهی، پرداخت، مبلغ تراکنش یا ارز محلی استفاده کرده‌اند. نتایج مطالعه نشان داد که روش انتخاب شده مؤثر است. همچنین، لی، هانگ، لیو و جیانگ[10] (۲۰۲۱)، با استفاده از مجموعه داده Kaggle، ایده یک روش ترکیبی را برای کاهش عدم تعادل طبقاتی با همپوشانی بر اساس ایده تقسیم و فتح ارائه کردند. برای این منظور از معیار ارزیابی آنتروپی وزن‌دار پویا استفاده شد. این آزمایش حتی از آزمایش‌های قبلی موفق‌تر بود

مطالعاتی که رفتار افراد در انجام تراکنش‌ها را به‌عنوان روشی برای شناسایی تراکنش‌های جعلی کارت هدف قرار می‌دهند نیز وجود دارد: کارمیناتی و همکاران (۲۰۱۸) Banksealer را به‌عنوان یک سیستم پشتیبانی تصمیم ارائه کردند. این سیستم، برای هر کاربر یک مدل می‌سازد و سپس آن را بر روی این سیستم مدل می‌کند. در مرحله دوم، استحکام سیستم Banksealer در برابر مجرمان احتمالی بررسی شد. این رویکرد در ایتالیا پیاده‌سازی شد،

1- General data protection legislation
2- Stalla- Bourdillon, Pearce, & Tsakalakis
3- HorakHorák, Stupka, & Husák
4- Data Privacy Impact Assessment
5- Bagga, Goyal, Gupta, & Goyal
6- Pinto, & Sobreiro
7- Bilgin, MARCO LAU, & Demir
8- Örek, M., Örek, E., & Bahtiyar
9- Multi-layered artificial neural networks
10- Li, Huang, Liu, & Jiang

اما نتایج مطالعه نشان داد که می‌تواند مجموعه داده مفیدی در سراسر جهان ارائه کند (کارمینیات، و همکاران[1]، ۲۰۱۸). چن و همکاران[2] (۲۰۱۹) نیز روشی را برای تشخیص تراکنش‌ها بر اساس رفتار افراد ایجاد کردند. مدلی که در مقاله آنها استفاده شد؛ هایپر کره[3] نام دارد. این مطالعه به این نتیجه رسید که اثر توصیف رفتار انسان با فراوانی معاملات آنها مرتبط است (چن و همکاران، ۲۰۱۹). علاوه بر این، وانگ و وانگ[4] (۲۰۱۹) رفتار کاربر را از منظر فاصله زمانی بین درخواست ارزیابی کردند و دریافتند رفتارهای ربات مانند، در سیستم بانکداری آنلاین وجود دارد. نتایج مطالعه نشان داد که پس از مقایسه دو الگوریتم، آنتروپی Renyi در تمایز رفتار ربات از رفتار انسان نسبت به آنتروپی شانون[5] برتری دارد. محققین دیگر از مدل‌های مختلفی به‌منظور شناسایی تراکنش‌های تقلبی استفاده کردند. چن و همکاران (۲۰۲۰) یک مدل امتیازدهی ترکیبی را برای این منظور توسعه داد. این مدل می‌تواند یک امتیاز اعتباری دقیق به دست آورد و حتی ریسک اعتباری را کاهش دهد (چن، یاداک، خان، و ژو[6]، ۲۰۲۰). مصرا، تاکار، گوش و ساها[7] (۲۰۲۰) یک مدل دومرحله‌ای برای شناسایی تراکنش‌های تقلبی پیشنهاد کردند. در مرحله اول از یک رمزگذار خودکار استفاده می‌شود که ویژگی‌های تراکنش را به یک بردار مشخصه کوچک‌تر تبدیل می‌کند. در مرحله دوم از بردار به‌عنوان ورودی طبقه‌بندی کننده استفاده می‌شود. آزمایش بر روی یک مجموعه داده مقایسه‌ای انجام و مشاهده شد که مدل دومرحله‌ای کارآمدتر از سیستم‌هایی است که تنها با یکی از دو مرحله طراحی شده‌اند.

اولووکره و ادوال[8] (۲۰۲۰) چارچوبی را پیشنهاد می‌کنند که تکنیک‌های گروه فرا یادگیری و یادگیری حساس به هزینه را برای کشف تقلب ترکیب می‌کند. نتایج این مطالعه نشان داد که چارچوب مجموعه حساس به هزینه، طبقه‌بندی‌کننده‌های حساس به هزینه را تولید می‌کند که در تشخیص تقلب در پایگاه‌های داده پرداخت‌ها کارآمد هستند. مطالعات دیگر بر روی تشخیص‌های خاص‌تر متمرکزشده‌اند. آماراسینگه، آپونسو، و کریشنارگیا[9] (۲۰۱۸) یادگیری ماشین انتخاب شده و تکنیک‌های تشخیص قبلی را که می‌توانند در یک سیستم تشخیص تراکنش مالی تقلبی ادغام شوند، بررسی کردند. نتیجه‌گیری شد که به‌منظور شناسایی مؤثرتر تقلب بانکی، دانستن الگوریتم‌های خاص (به‌عنوان‌مثال، شبکه‌های بیزی، منطق فازی و غیره) مهم است. دونگ و همکاران[10] (۲۰۱۸) روی حوزه بسیار جالبی از موضوع تقلب، یعنی تقلب‌های تبلیغاتی موبایلی کار کردند و یک رویکرد ترکیبی به نام FraudDroid را برای شناسایی تقلب در برنامه‌های کاربردی در دستگاه‌های اندرویدی پیشنهاد کردند. پس از تجزیه‌و تحلیل ۱۲۰۰۰ برنامه مشکوک، FraudDroid ۳۳۵ مورد تقلب را شناسایی کرد و این نتیجه تأیید می‌کند که روش مفیدی برای کشف این نوع جرم است. علاوه بر این، برخی محققین، تحقق رأی از طریق دستگاه ATM با ارائه احراز هویت بیومتریک یا احراز هویت تشخیص چهره را پیشنهاد کردند. استفاده از برنامه رأی‌گیری ATM در مقایسه با کارت‌های Aadhar برای امنیت و حفظ حریم خصوصی، آسان‌تر است (سادهارسان، کومار، وتکاسان، ساتیاپریا و سارانایا[11]، ۲۰۱۹)

چالش‌های بانکداری آنلاین

هر ساله تقلب در اشکال مختلف افزایش می‌یابد که منجر به خسارات مالی قابل‌توجهی می‌شود (وی، لی، کائو،

1- Carminati, et al
2- Chen
3- Hypersphere
4- Wang and Wang
5- Shannon entropy
6- Chen, Yadav, Khan, & Zhu
7- Misra
8- Olowookere, & Adewale
9- Amarasinghe
10- Dong
11- Sudharsan, Kumar, Venkatesan, Sathyapreiya, & Saranya

اوو و چن[1]، ۲۰۱۳). بر اساس بررسی شاخص ایمنی محاسباتی مایکروسافت (MCSI) در سال ۲۰۱۴، تأثیر سالانه فیشینگ و اشکال مختلف سرقت هویت در سراسر جهان می‌تواند تا ۵ میلیارد دلار آمریکا تخمین زده شود، درحالی‌که هزینه ترمیم آسیب به شهرت آنلاین افراد می‌تواند بیشتر از ۶ میلیارد دلار آمریکا، یا میانگین تخمینی ۶۳۲ دلار آمریکا به ازای هر ضرر باشد (ماریکان و لیم، ۲۰۱۴). مشاوره PwC در نظرسنجی جهانی جرم اقتصادی در سال ۲۰۱۶ اشاره می‌کند که تقریباً یک‌سوم شرکت‌های مورد بررسی گزارش داده‌اند که قربانی نوعی از جرائم سایبری بوده‌اند. همچنین، جرائم سایبری دومین نوع رایج جرائم اقتصادی است که در این نظرسنجی تجزیه‌وتحلیل شده است، درحالی‌که یک نظرسنجی اخیر گزارش می‌دهد (لاویون[2]، ۲۰۱۸). از سال ۲۰۱۶ بالغ بر ۱۳ درصد افزایش در تقلب رخ داده است. IC3 یک منبع ارزشمند برای قربانیان جرائم اینترنتی و مجریان قانون در شناسایی، رسیدگی و تعقیب جرائم است (اف.بی.آی. ۲۰۱۸). IC3 گزارش می‌دهد که ۱۴۴۰۸ شکایت در سال ۲۰۱۸ دریافت کرده که مربوط به کلاه‌برداری پشتیبانی فنی از قربانیان از ۴۸ کشور بود. زیان گزارش شده نشان‌دهنده افزایش ۱۶۱ درصدی زیان نسبت به سال قبل می‌باشد. گزارش نرخ جرائم اقتصادی جهانی توسط (اسکانسی، گانو[3]، ۲۰۱۷)، نشان می‌دهد که خدمات مالی با جرم اقتصادی ۴۸ درصد بیشترین بخش در معرض خطر بوده و دومین جرم اقتصادی گزارش‌شده در جرائم سایبری است. رایج‌ترین دامنه‌های تقلب عبارتند از: بانکداری آنلاین، کارت‌های اعتباری، مخابرات، بیمه مراقبت‌های بهداشتی، بیمه آنلاین، نفوذ کامپیوتر، حراجی آنلاین (اسکانسی، گانو، ۲۰۱۷؛ بولتون و هند، ۲۰۰۲؛ جان، آنل، کندی، اولاجید و کندی[4]، ۲۰۱۶؛ عبدالله، معروف و زینال[5]، ۲۰۱۶؛ وی و همکاران، ۲۰۱۳). بنابراین، انجام تحقیقات کشف تقلب برای بانکداری آنلاین یک کار بسیار مهم است، اما چالش‌هایی در این زمینه وجود دارد که باید برطرف شوند. دانش مکانیزم کشف تقلب بانک‌ها بسیار محدود است و بانک‌ها اغلب آمار سیستم‌های FD را منتشر نمی‌کنند (اسکانسی، گانو[6]، ۲۰۱۷؛ ماراتون، ۲۰۱۳). زیرا بیشتر امنیت توسط شرکت‌های IT شخص ثالث ارائه می‌شود که از مالکیت معنوی در برابر رقبای خود نیز محافظت می‌کنند. بنابراین هم بانک‌ها و هم شرکت‌های امنیتی فناوری اطلاعات بیشتر اطلاعات سیستم‌های امنیتی خود را منتشر نمی‌کنند. بولتون و هند[7] (۲۰۰۲) نیز تأکید می‌کنند که توسعه روش‌های جدید کشف تقلب دشوار است؛ زیرا تبادل نظر در کشف تقلب بسیار محدود است، اما از این ایده حمایت می‌کنند که تکنیک‌های FD نباید به‌طور عمومی با جزئیات توصیف شوند. در غیر این صورت مجرمان نیز ممکن است به آن اطلاعات دسترسی پیدا کنند (کارمیناتی و همکاران، ۲۰۱۵). فوا و همکاران[8] (۲۰۱۰) تأکید می‌کنند که کشف تقلب با استفاده از تکنیک‌های داده‌کاوی در صنعت بسیار رایج است.

تقلب در سرمایه‌گذاری سهام و اوراق بهادار

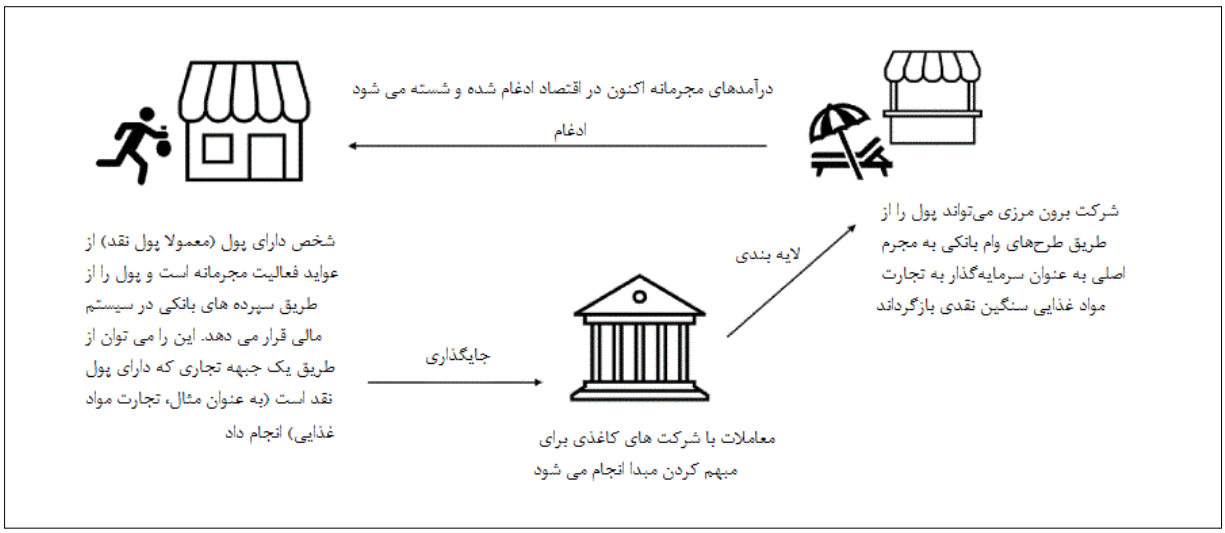
بازار سهام و اوراق بهادار مالی به مردم این امکان را می‌دهد که پول خود را با هدف کسب بازدهی مثبت بر اساس انجام تحقیقات یا فقط یک حدس، سرمایه‌گذاری کنند. با این حال، مشخص است که بخشی از فعالان بازار تقلب می‌کنند و با این کار سودهای کلانی به قیمت سرمایه‌گذاران نهادی و خرد به دست می‌آورند. دستگیری این بازیگران کلاه‌بردار آسان نیست و معمولاً به نیروی کار زیادی برای جمع‌آوری شواهد در مدت زمان طولانی نیاز دارد، به‌ویژه در موارد تجارت داخلی. با این حال، پیشرفت‌های اخیر در کاربردها و تکنیک‌های یادگیری ماشین به شناسایی بازیگران بد به روشی کارآمدتر و سریع‌تر کمک می‌کند. برخی از روش‌های مورد استفاده توسط افرادی که فعالیت سرمایه‌گذاری متقلبانه انجام می‌دهند عبارتند از: دست‌کاری بازار، تجارت داخلی، پول‌شویی، تأمین مالی تروریسم و بسیاری موارد دیگر. دست‌کاری بازار،

1- Wei, Li, Cao, Ou, & Chen
2- Lavion
3- Ocansey, & Ganu
4- John, Anele, Kennedy, Olajide, & Kennedy
5- Abdallah, Maarof, & Zainal
6- Ocansey, & Ganu
7- Bolton & Hand
8- Phua

اقدامی برای فروش یا خرید یک اوراق بهادار مالی با هدف دست‌کاری هدفمند قیمت دارایی یا اوراق بهادار پایه در نظر گرفته می‌شود. تجارت داخلی غیرقانونی زمانی است که «خودی‌ها» یا افرادی که از مطالب شرکت خصوصی و غیردولتی مطلع هستند، از آن اطلاعات قبل از انتشار عمومی برای بهره‌مندی مالی استفاده می‌کنند. این نه تنها شامل معامله اوراق بهادار، بلکه نشت اطلاعات غیر عمومی به اشخاص ثالث نیز می‌شود.

کشف تقلب و مبارزه با پول‌شویی (AML)

اغلب اوقات، بین تقلب بانکی و سرقت از بانک سردرگمی وجود دارد. دزدی از بانک با تقلب بانکی متفاوت است، چرا که سرقت از بانک شامل خشونت می‌شود، درحالی‌که تقلب بانکی اغلب یک فرآیند آهسته و مخفیانه است که تا زمانی که عمل کامل شود، شناسایی نمی‌شود. تقلب بانکی اغلب مستلزم ترغیب مشتریان به ارائه اطلاعات حساس بانکی خود است که ممکن است برای بدست آوردن پول از راه دور استفاده شود. پول‌شویی روشی است که توسط مجرمان و افرادی که دارایی‌های «کثیف» یا به‌طور غیرقانونی از طریق فعالیت‌های مجرمانه به‌دست‌آمده‌اند، برای تبدیل این پول‌ها به یک حالت «پاک» یا مشروع از نظر قانون و دولت‌ها استفاده می‌کنند. سرویس دادستانی سلطنتی[1] (CPS) بریتانیا طرح پول‌شویی را معمولاً شامل سه مرحله تعریف می‌کند: اول، جایگذاری است که فرآیند واریز پول مجرمانه به سیستم مالی است. دوم، لایه‌بندی است که پول را در سیستم مالی از طریق شبکه‌های پیچیده‌تراکنش‌ها با هدف مبهم کردن حرکت می‌دهد. لایه‌بندی معمولاً از طریق شرکت‌های خارج از کشور انجام می‌شود. درنهایت، ادغام، پول مجرمانه‌ای است که از طریق سرمایه‌گذاری‌هایی مانند املاک، خرید سهام و اقلام لوکس در اقتصاد واقعی جذب یا ترکیب می‌شود (هاپتون[2]، ۲۰۱۶). نمونه‌ای از پول‌شویی با استفاده از قرار دادن، لایه‌بندی و ادغام را می‌توان در شکل شماره (۱) مشاهده کرد.



شکل (۱): مراحل پول‌شویی

1- Crown Prosecution Service
2- Hopton

کشف تقلب مستلزم فعالیت‌هایی است که از به دست آوردن پول یا دارایی از طریق جعل نادرست جلوگیری می‌شود. در بخش بانکی، تقلب مستلزم جعل چک و استفاده از کارت اعتباری سرقت شده است. ممکن است متضمن اغراق زیان یا ایجاد رویدادهای ناگوار مانند تصادفات با هدف صرف پرداخت باشد (شیهم بتسا، ۲۰۲۱). بر اساس گزارش نیلسون (۲۰۲۰)، پیش‌بینی می‌شد که کلاه‌برداری مربوط به کارت‌ها تا سال ۲۰۲۰ تنها به مبلغ حیرت‌انگیز ۳۰ میلیارد دلار در سطح جهان برسد. علاوه بر این، اختلال در فناوری در بانکداری و پرداخت‌ها به دلیل افزایش کانال‌های پرداخت مانند اعتبار و کارت‌های نقدی و گوشی‌های هوشمند، تعداد تراکنش‌ها طی سال‌های اخیر افزایش چشمگیری داشته است (گابینو و همکاران، ۲۰۲۱). تلاش‌های کشف تقلب با استفاده از تجزیه‌وتحلیل داده‌ها، نرم‌افزار طراحی‌شده برای کشف تقلب، و ابزارها، برنامه‌های طراحی‌شده برای کشف تقلب، سازمان‌ها را قادر می‌سازد تاکتیک‌های رایج کلاه‌برداری را پیش‌بینی کنند، فرآیند ارجاع متقابل داده‌ها را خودکار کنند، تراکنش‌های خود را به‌طور مستمر در زمان واقعی نظارت کنند، و کشف کنند (ویلیام، ۲۰۲۲). منابع تشخیص و پیشگیری از تقلب مانند نرم‌افزار در دو نسخه اختصاصی و رایگان موجود است. برخی از ویژگی‌های رایج در این نرم‌افزار عبارتند از: داشبورد، واردات و صادرات داده‌ها، تجسم داده‌ها، یکپارچه‌سازی مدیریت ارتباط با مشتری، مدیریت تقویم، زمان‌بندی، بودجه‌بندی و مدیریت رمز عبور. آنها همچنین دارای رابط‌های برنامه‌نویسی کاربردی (API)، صدور صورت‌حساب، احراز هویت دومرحله‌ای و مدیریت پایگاه داده مشتری هستند.

سیستم‌های تشخیص تقلب در بانکداری

کلاه‌برداری یا تقلب آنلاین، تقلب اینترنتی و جرائم سایبری مبنای گسترده‌ای دارند و به روش‌های مختلفی رخ می‌دهند. کلاه‌برداری آنلاین را می‌توان به‌عنوان هر عمل غیرقانونی انجام شده به‌صورت آنلاین توصیف کرد. کلاه‌برداری در بانکداری اینترنتی، بانکداری تلفن همراه، فیشینگ، کلاه‌برداری در سایت خرید و حراجی، کلاه‌برداری، هرزنامه، و سرقت هویت انواع مختلفی از کلاه‌برداری‌های آنلاین هستند (ای.پی.اف. ۲۰۱۸). تقلب بانکداری اینترنتی نوعی کلاه‌برداری است که با استفاده از هر فناوری آنلاین برای برداشت غیرقانونی پول یا انتقال آن به‌حساب بانکی دیگر انجام می‌شود. جرم مالی یا جرائم اقتصادی توسط یوروپل به‌عنوان «اعمال غیرقانونی انجام شده توسط یک فرد یا گروهی از افراد برای به دست آوردن مزیت مالی یا حرفه‌ای» تعریف شده است (نیکولس، و همکاران، ۲۰۲۱). انگیزه اصلی در چنین جنایاتی، منفعت اقتصادی است. مجمع جهانی اقتصاد، جرائم مالی را صنعتی تریلیون دلاری می‌داند (ناندوری، جیا، اوکا، بیوار و لیو[1]، ۲۰۲۰). جرائم مالی در فضای مجازی با استفاده از ابزارهای هک و تکنیک‌های مهندسی اجتماعی که امنیت مؤسسات مالی و شرکتی را دور می‌زند، انجام می‌شود. این امر باعث می‌شود که محققان و شرکت‌ها به جرائم مالی از منظر دیگری نگاه کنند، به‌این‌ترتیب که تمایز بین جرائم مالی، هک و مهندسی اجتماعی برای منافع اقتصادی کمرنگ شده است. اینجاست که نویسندگان اصطلاح جرائم سایبری مالی را معرفی می‌کنند که ترکیبی از جرائم مالی، هک و مهندسی اجتماعی است که در فضای سایبری تنها با هدف کسب سود اقتصادی غیرقانونی انجام می‌شود. کشف تقلب در سازمان‌های مالی سنتی[2] بیشتر طول می‌کشد (پاسکودال، مارچینی، و میللر[3]، ۲۰۱۷). این مقدار زمان برای مشتریان و مؤسسات مالی نامطلوب است؛ زیرا ممکن است چندین فعالیت متقلبانه قبل از شناسایی در چنین چارچوب زمانی رخ دهد. کلاه‌برداری بر بانک‌هایی که با خدمات پرداخت آنلاین سروکار دارند، به‌ویژه در پیشرفت‌های تکنولوژیک معاصر در بخش تجاری، تأثیر منفی می‌گذارد. به‌عنوان‌مثال، حدود ۲۰ درصد از مشتریان پس از تجربه کلاه‌برداری، بانک را تغییر می‌دهند (ساندو[4]، ۲۰۲۱). این تعداد مشتریانی که از یک بانک خاص خارج می‌شوند، به طور قابل توجهی برای عملیات تجاری یک

1- Nanduri, Jia, Oka, Beaver, & Liu
2- brick-and-mortar financial organizations
3- Pascual, Marchini, & Miller
4- Sando

بانک مضر است، به خصوص اگر این روند طی چندین سال ادامه یابد. ضرورت ایجاد رویکردهای مناسب و قوی کشف تقلب در سیستم‌های مؤسسات مالی برای مهار این عمل وجود دارد. دو رویکرد اصلی کشف تقلب وجود دارد که شامل رویکرد مبتنی بر قانون و تشخیص تقلب یادگیری ماشین است. سیستم‌های بانکداری آنلاین[1] (OBS) مکانیسم‌های امنیتی مختلفی برای جلوگیری از دسترسی غیرمجاز کلاه‌برداران دارند. علیرغم همه این پیشرفت‌ها، قربانیان بی‌خبر بازهم اعتبار خود را در برابر کلاه‌برداران فیشینگ و سارقان هویت آنلاین از دست می‌دهند. هنگامی‌که یک کلاه‌بردار به‌حساب بانکی کاربر دسترسی پیدا می‌کند، تشخیص آن فعالیت آسان نیست. بانک‌های مختلف از سیستم‌های کشف تقلب مختلفی استفاده می‌کنند. برخی از سیستم‌های کشف تقلب شناخته شده تجاری مورد استفاده برای بانکداری آنلاین: FICO، PRM، Kount و SAS Fraud Manager هستند.

سیستم FICO: سیستم FICO یکی از پرکاربردترین سیستم‌های کشف تقلب مورد استفاده توسط بانک‌ها در سطح جهان است (اف.آی.سی.او.، ۲۰۱۰؛ باربازون، کاهیل و والش[2]، ۲۰۱۰). این سیستم، از یک شبکه عصبی و یک موتور مبتنی بر قانون استفاده می‌کند. قابلیت‌های بلادرنگ دارد. FICO توسط مؤسسات مالی، بانک‌ها، اتحادیه‌های تولیدی و اعتباری، عمدتاً برای کشف تقلب بدهی، اعتبار، سپرده و پرداخت الکترونیکی استفاده می‌شود (کاپتر[3]، ۲۰۱۹).

سیستم PRM : یکی دیگر از سیستم‌های کشف تقلب پرکاربرد PRM است (آی.سی.آی.، ۲۰۱۱). PRM مانند FICO، از یک شبکه عصبی و یک رویکرد مبتنی بر قوانین نیز استفاده می‌کند. سیستم PRM در بیش از ۴۰ کشور از جمله هشت بانک از ۲۰ بانک برتر جهان استفاده می‌شود. انواع تقلب از کارت بدهی و اعتباری، تقلب حساب و تقلب پول‌شویی توسط PRM شناسایی می‌شود (حفیظ، عقیلی، و زاواراسکی[4]، ۲۰۱۶).

سیستم SAS Fraud Manager: این سیستم عمدتاً راه‌حل کشف تقلب برای کارت‌های بدهی و اعتباری است و قابلیت بلادرنگ را دارد. این سیستم کشف تقلب از تکنیک تشخیص ناهنجاری[5] (AD) استفاده می‌کند (اس. آی.اس. ۲۰۰۷).

سیستم Kount: یکی دیگر از سیستم‌های پیشگیری از تقلب، کانت است که از هر دو مدل ML[6] نظارت شده و بدون نظارتی استفاده می‌کند. برخی از بزرگ‌ترین ارائه‌دهندگان خدمات پرداخت، دروازه‌ها، کیف پول‌ها و پردازنده‌ها از این سیستم استفاده می‌کنند. تعداد کمی از سیستم‌های کشف تقلب و پیشگیری دیگر وجود دارند (پاتم سی، ۲۰۱۵؛ فاندنت، ۱۹۹۷؛ ریسک نت، ۱۹۹۸).

-نقاط ضعف سیستم‌های تشخیص تقلب تجاری

سیستم‌های کشف تقلب تجاری ذکر شده در بالا به‌طور گسترده توسط بانک‌ها و مؤسسات مالی استفاده می‌شوند. با این حال، هیچ سیستم کشف تقلب به‌طور ۱۰۰ درصد مؤثر نیست. برخی از کاستی‌ها در سیستم‌های کشف تقلب تجاری عبارتند از:

این سیستم‌ها فاقد امتیازدهی و ردیابی مکان مصرف‌کننده برای تراکنش دستگاه تلفن همراه هستند.

لایه یکپارچه‌سازی برای این سیستم‌ها برای وارد کردن داده‌ها کامل نیست.

همه سیستم‌ها قابلیت ثبت رمزگذاری داده‌ها را ندارند.

سیستم‌ها مبالغ بیشتری از تراکنش‌ها را در نظر می‌گیرند و عمدتاً مبالغ کمتر را نادیده می‌گیرند (دیون[7]، ۲۰۱۴؛ هرسچل، لیدن، و کارت[8]، ۲۰۱۵).

1- Online Banking Systems
2- Brabazon, Cahill, Keenan, & Walsh
3- Capterra
4- Hafiz, Aghili, & Zavarsky,
5- Anomaly detection
6- Machine learning
7- Dehaven
8- Herschel, Linden, & Kart

سیستم‌های مبتنی بر قانون (RBS)

سیستم‌های مبتنی بر قانون[1] بخشی از گروه بزرگی از رویکردها هستند که سعی در مدل‌سازی تخصص انسان به نام سیستم‌های مبتنی بر دانش[2] (KBS) دارند. KBS سنتی از جمله RBS دارای مسائل مشترک KA، شکنندگی[3] و یادگیری افزایشی[4] است. اصطلاح شکنندگی توسط لیگ (۲۰۰۷) ابداع شده است و به نرم‌افزاری اطلاق می‌شود که احتمالاً در مواجهه با سناریوی غیرمنتظره به نتیجه نادرستی می‌رسد. درحالی‌که یادگیری افزایشی یکی از راه‌حل‌های ممکن برای مشکل مقیاس‌پذیری است، که در آن داده‌ها در بخش‌هایی پردازش می‌شوند و سپس نتایج را برای کاهش استفاده از حافظه ترکیب می‌کنند (سید، لیو، و سانگ[5]، ۱۹۹۹). فعالیت‌های مربوط به تقلب در حوزه مالی با کاوش سیگنال‌های سطحی و واضح قابل شناسایی هستند. با وجود غیرعادی بودن، تراکنش‌های بزرگ و آنهایی که در مکان‌های غیرمعمول اتفاق می‌افتند، باید تأیید قوی‌تری داشته باشند. سیستم‌های مبتنی بر قانون از الگوریتم‌هایی استفاده می‌کنند که موقعیت‌های مختلف کشف تقلب را ارزیابی می‌کنند که به‌صورت دستی توسط تحلیلگران تقلب نوشته‌شده‌اند. در حال حاضر، سیستم‌های حقوقی تقریباً از ۳۰۰ قانون مختلف برای تأیید معاملات استفاده می‌کنند. این توضیح می‌دهد که چرا استفاده از سیستم‌های مبتنی بر قانون یک فرآیند ساده است. این سیستم‌ها نیاز به تنظیم دستی سناریوها و موقعیت‌هایی دارند که قادر به تشخیص همبستگی‌های ضمنی نیستند. علاوه بر این، سیستم‌های مبتنی بر قانون اغلب از نرم‌افزار قدیمی استفاده می‌کنند که به‌سختی جریان داده‌های بلادرنگ را از طریق سیستم‌هایی که در حوزه دیجیتال مهم هستند، پردازش می‌کنند. سیستم‌های مبتنی بر قوانین؛ به قوانین از پیش تعیین شده بستگی دارند تا تغییرات در رفتار را تشخیص دهند. این سیستم‌ها سفت و سخت هستند و قادر به انطباق با صنایعی مانند خدمات مالی نیستند که برای غلبه بر چالش‌های مرتبط با شناسایی تقلب نیاز به پلت‌فرم‌های سبک‌تر و چابک‌تری دارند (مون و کیم[6]، ۲۰۱۷). رویکردهای مبتنی بر قانون زمان بر هستند زیرا به تحلیل‌گران تقلب نیاز دارند تا قوانین مورد استفاده برای کشف تقلب را ایجاد کنند. این رویکردها همچنین به کار دستی و چندین فرآیند تأیید نیاز دارند که مانع از تجربه کاربر می‌شود. رویکردهای مبتنی بر قوانین، الگوهای تقلب آشکار را شناسایی می‌کنند و بنابراین با تغییرات کلاه‌برداری که با تکامل کلاه‌برداران رخ می‌دهد، سازگار نیستند. از سوی دیگر، مدل‌های مبتنی بر قانون برای نگهداری مجموعه داده‌ها پرهزینه‌تر می‌شوند یا اندازه پایگاه مشتری گسترش می‌یابد.

تشخیص تقلب مبتنی بر یادگیری ماشین (ML)

رویدادهای پنهانی در رفتار کاربر وجود دارد که ممکن است کاملاً مشهود نباشد و تراکنش‌های جعلی احتمالی را به نمایش بگذارد. یادگیری ماشین اجازه ایجاد الگوریتم‌هایی را می‌دهد که می‌توانند مجموعه داده‌های بزرگ‌تری را با چندین متغیر پردازش کنند و به شناسایی این همبستگی‌های پنهان بین رفتار اپراتور و احتمال فعالیت‌های متقلبانه کمک می‌کنند. سیستم‌های یادگیری ماشین بهتر از سیستم‌های مبتنی بر قوانین هستند، زیرا در پردازش داده‌ها سریع‌تر هستند و عملیات دستی کمتری دارند. به‌عنوان‌مثال، الگوریتم‌های هوشمند با تجزیه‌وتحلیل رفتار در کاهش مراحل تأیید مورد نیاز مطابقت دارند. شرکت‌هایی که با مقررات خدمات مالی سر و کار دارند در نظارت بر فعالیت‌های تقلبی احتمالی درگیر هستند: آنها باید در مورد فعالیت‌های علامت‌گذاری شده با یکدیگر ارتباط برقرار کنند. ویلالوبوس[7] و همکاران (۲۰۱۹) نمونه‌ای را توصیف می‌کنند که در آن یک نمونه اولیه یادگیری ماشین بر روی مجموعه داده‌ای برنامه‌ریزی شد

1- Rule-based systems
2- Knowledge-Based Systems
3- Brittleness
4- Incremental learning
5- Syed, Liu, & Sung
6- Moon, & Kim
7- Villalobos

که تراکنش‌های آن به‌صورت مجرمانه انجام شده بود. نمونه اولیه مورد استفاده در سیستم مبتنی بر قانون به کشف روابط پنهان بین معاملات و فعالیت‌های مجرمانه کمک کرد. چنین سیستم‌هایی حجم کار را در بانک‌های کوچک‌تر درگیر در نظارت بر تقلب به حداقل می‌رساند. راه‌حل پیشنهادی نشان داد که ۹۹٫۶ درصد از معاملات پول‌شویی و تراکنش‌های گزارش شده از ۳۰ درصد به ۱ درصد کاهش یافته است. ML به الگوریتم‌هایی بستگی دارد که با افزایش اندازه مجموعه داده‌ها مؤثرتر هستند. هر چه داده‌ها بیشتر باشد، نمونه اولیه ML بیشتر بهبود می‌یابد و می‌تواند شباهت‌ها و تفاوت‌ها را در رفتارهای مختلف تشخیص دهد. هرچه مدل ML تفاوت بین تراکنش‌های مشروع و جعلی را بیشتر شناسایی کند، سیستم‌ها در دسته‌بندی مجموعه‌های داده در دسته‌های مختلف کارآمدتر می‌شوند. بنابراین، سیستم‌های ML با رشد پایگاه داده مشتری، مقیاس‌پذیرتر هستند. درحالی‌که الگوریتم‌های ML مزایای متعددی را ارائه می‌کنند، اما دارای معایب قابل‌توجهی هستند که استفاده از آنها را در تشخیص تقلب محدود می‌کند. به عنوان مثال، یکی از اشکالات این است که ML برای دستیابی به دقت، به مقدار قابل توجهی داده نیاز دارد. این حجم داده قابل مدیریت است. با این حال، باید نقاط داده کافی وجود داشته باشد که روابط علّی مشروع را در سازمان‌های کوچک‌تر تشخیص دهد. علاوه بر این، مدل‌های یادگیری ماشین بر روی اعمال، رفتار و فعالیت‌ها عمل می‌کنند. این مدل تمایل دارد اتصالات واضح را نادیده بگیرد، چنین کارتی در دو حساب مختلف استفاده می‌شود، از این رو، فعالیت کشف تقلب را بی‌اثر می‌کند.

-سیستم‌های تشخیص نفوذ (IDS)

با توجه به سیستم‌های تشخیص نفوذ[1] (IDS) می‌توانند میزان ایمنی بیشتری را ارائه دهند، اما نسبت به شدت ایمنی، به منابع محاسباتی بسیار بیشتری نیاز دارند (لی، پارک، اوم و چانگ[2]، ۲۰۱۱). نویسندگان یک IDS چند سطحی و مدیریت گزارش را برای IDS مؤثر در رایانش ابری پیشنهاد کردند. این امر با معرفی مقدار قابل‌توجهی از مسئولیت ایمنی به مشتریان بسته به نوع روش ناهنجاری، که کاربران را بر اساس نرخ ناهنجاری به سازمان‌های ایمنی متمایز متصل می‌کند، به استفاده کارآمد از منابع کمک می‌کند. برخی محققین، بر این باورند که هکرها می‌توانند یک سری از حملات را به یک سیستم مقصد امن در فضای ابری انجام دهند، به‌عنوان‌مثال، با فرار از یک دستگاه با استفاده آسان مبتنی بر ابر و سپس از درب پشتی[3] قبلی برای حمله به سیستم استفاده می‌کنند (چن، گان، هانگ، و کیو[4]، ۲۰۱۲). سیستم تشخیص پیشنهادی لاگ های مختلف را از ابر برای به دست آوردن معانی فعالیت‌های گزارش تجزیه‌وتحلیل می‌کند. برای تعداد کمی از تخلفات، فعالیت‌های مشکوک اغلب توسط مدیر نادیده گرفته می‌شود. مدل پنهان مارکوف[5] برای مدل‌سازی توالی حملات انجام‌شده توسط هکرها و چنین رخدادهای مخفیانه در یک چارچوب طولانی مدت اجرا می‌شود، زیرا این مدل آگاه از وضعیت می‌باشد. سیستم‌های پیشنهاد شده برای IDS، عمدتاً بر شناسایی رویدادهای بالقوه، ثبت داده‌ها و تلاش‌های نظارتی تمرکز دارند.

تشخیص ناهنجاری (AD)

تشخیص ناهنجاری (AD) شامل استفاده از تکنیک‌های مختلف محاسباتی و ریاضی برای تشخیص نقاط غیرعادی در یک مجموعه داده است. در ادبیات مرتبط، تشخیص ناهنجاری نام‌های مختلفی دارد: تشخیص بیرونی[6]، تشخیص تازگی[7]، تشخیص نویز[8] و تشخیص انحراف[9]. تشخیص ناهنجاری فرآیند تجزیه‌وتحلیل مجموعه داده

1- Intrusion Detection Systems
2- Lee, Park, Eom, & Chung
3- Backdoor
4- Chen, Guan, Huang, & Ou
5- Hidden Markov Model
6- Outlier Detection
7- Novelty
8- Noise
9- Deviation

برای شناسایی موارد انحرافی است و شامل یک یا دو کار زیر است:

(الف) شناسایی داده‌های غیرعادی، به‌عنوان‌مثال، نویز، انحرافات یا نقاط پرت از مجموعه داده اصلی و (ب) کشف نمونه‌های داده جدید بر اساس دانش‌آموخته شده بر اساس مجموعه داده اصلی، تشخیص ناهنجاری را می‌توان برای اهداف مختلفی استفاده کرد (آگراوال[1]، ۲۰۱۷)، مانند تشخیص تقلب، تجزیه‌وتحلیل کیفیت داده‌ها، اسکن امنیتی، نظارت بر فرآیند و سیستم، نظارت تصویر/ویدئو، تشخیص هرزنامه، شناسایی حملات مخرب داخلی، پاک‌سازی داده‌ها قبل از آموزش مدل‌های آماری، تجزیه‌وتحلیل رفتار انسان (چیو، کیم، کانگ، و یوم[2]، ۲۰۲۱) و تشخیص خطای حسگر (درویشی، کیانزو، عید و راسی[3]، ۲۰۲۰).

برخی از محققین، یک تکنیک تشخیص نفوذ مبتنی بر قانون (ID) پیشنهاد و اذعان می‌کنند که AD یکی از اولین رویکردها برای ID است (ایلگان، کیمرر و پاراس[4]، ۱۹۹۵). سایر محققین نشان دادند که سرقت حساب یا سرویس در رایانش ابری خطرناک‌تر است. نویسندگان چارچوبی با تکنیک AD برای نمایه کردن رفتارهای معمولی کاربر پیشنهاد کردند. هنگامی‌که یک نمایه کاربر از اطلاعات جمع‌آوری‌شده کشف می‌شود، هشدارها توسط همه رفتارهای مشکوک شناسایی شده توسط نمایه فعال می‌شوند. هشدارها به پایگاه داده ارسال می‌شود و به صاحب حساب و مدیر ابر اطلاع داده می‌شود. دو جزء اصلی چارچوب وجود دارد: بخش اول جمع‌آوری داده و بخش دوم ماژول یادگیری است که تکنیک‌های مختلف ML را برای استخراج روندهای مکرر از داده‌های آموزشی و به دست آوردن محدودیت رتبه‌بندی مشکوک معرفی کرده است. اگر رتبه‌بندی مشکوک از حد تعیین شده بیشتر شود، تراکنش به‌عنوان یک ناهنجاری توسط طرح گزارش می‌شود چیو و همکاران[5] (۲۰۱۳) پیشنهاد می‌کنند که تکنیک AD اخطار را برای کاربران سیستم ارسال می‌کند و بیشتر متمرکز بر ابر است و برای اطلاعات بزرگ مناسب نیست (چیو، یه و لی[6]، ۲۰۱۳). برابازون و همکاران[7] (۲۰۱۰) همچنین استفاده از یک رویکرد مبتنی بر AIS برای AD را برای کاهش تقلب در کارت اعتباری توصیف می‌کنند. تجزیه و تحلیل غیرعادی برای داده‌های گزارش جریانی توسط محققین انجام شده است (مروستی، پوگاسیان، هارتانیان، و گیروگیان[8]، ۲۰۱۴). نویسندگان مفهوم آستانه استراتژیک داده سری زمانی را به هر نوع اطلاعاتی که جریانی از اسناد و فعالیت‌ها هستند، گسترش می‌دهند. آنها مفهوم نرمال بودن آن جریان‌ها را در روش پیشنهادی خود پیاده‌سازی کردند و مکانیزمی را برای تشخیص ناهنجاری آنها در جریان زمان اجرا ایجاد کردند. تحت محدودیت‌های منحصربه‌فرد در پیچیدگی و مقیاس‌پذیری، آنها ساختار تصمیم‌گیری جدیدی را برای بازیابی اطلاعات از جریان داده‌ها پیاده‌سازی کردند. تکنیک پیشنهاد شده فوق، در درجه اول مبتنی بر تجزیه‌وتحلیل غیرعادی داده‌های رویداد از فایل‌های گزارش و به دست آوردن اطلاعات مفید از منبع و انواع رویدادها است.

الگوریتم‌های تشخیص تقلب بانکی

بسیاری از روش‌های ML برای مبارزه با تقلب توسعه داده شده است. تکنیک‌های متداول کشف تقلب عبارتند از ANN، ES (مبتنی بر دانش (KB))، موتور استنتاج و داده‌کاوی (ماراتون، ۲۰۱۳؛ وی و همکاران، ۲۰۱۳). بیشترین رویکردهای مورد استفاده برای کشف تقلب روش‌های نظارت شده، بدون نظارت، نیمه نظارت شده و ترکیبی هستند (بولتون و هند، ۲۰۰۲؛ جان و همکاران، ۲۰۱۶؛ عبدالله و همکاران، ۲۰۱۶؛ کانادول و بانجری[9]، ۲۰۰۹). کواچ و روجیرو[10] (۲۰۱۱) یک سیستم کشف تقلب را برای بانکداری آنلاین با تمرکز بر تجزیه‌وتحلیل محلی و جهانی رفتار کاربران پیشنهاد

1- Aggarwal
2- Choi, Kim, Kang, & Youm
3- Darvishi, Ciuonzo, Eide, & Rossi
4- Ilgun, Kemmerer, & Porras
5- Chiu
6- Chiu, Yeh, & Lee
7- Brabazon
8- Marvasti, Poghosyan, Harutyunyan, & Grigoryan
9- Chandola, & Banerjee
10- Kovach and Ruggiero

می‌کنند. ارزیابی افتراقی برای به دست آوردن شواهدی از تقلب استفاده می‌شود که در آن تغییر قابل‌توجهی از رفتار عادی یک تقلب بالقوه را نشان می‌دهد. شواهد تقلب بر تعداد دسترسی‌های کاربر و مقدار احتمالی که در طول دوره متفاوت است متمرکز می‌باشد. روش پیشنهادی کشف تقلب آنها بر شناسایی کارآمد دستگاه‌های مورد استفاده برای کنترل حساب‌ها و ارزیابی احتمال کلاه‌برداری با نظارت بر تعداد سوابق مجازیی که هر دستگاه به آن دسترسی دارد، تمرکز دارد. روشی برای شناسایی تقلب در کارت اعتباری توسط دومان و اوزچلیک[1] (۲۰۱۱) ارائه شده است. نویسندگان ترکیبی از دو روش فراابتکاری معروف را برای رتبه‌بندی پیشنهاد کردند که عبارتند از: الگوریتم‌های ژنتیک و جستجوی پراکنده. این تکنیک بر روی داده‌های واقعی پیاده‌سازی شده است و یافته‌های به‌دست‌آمده نسبت به سیستم فعلی در حال استفاده بسیار مؤثر هستند. هر تراکنش با این رویکرد رتبه‌بندی می‌شود و عملیات بسته به نتایج به‌صورت تقلب یا مشروع طبقه‌بندی می‌شود. آنها معتقدند که یک طرح FD بهتر از طرحی است که بسیاری از کلاه‌برداری‌های کم‌خطر را شناسایی می‌کند، که جرم را حتی از نظر مقدار کمتر اما از نظر ارزش بیشتر کشف می‌کند. کو و همکاران[2] (۲۰۰۴) نیز یک مطالعه مبتنی بر داده‌کاوی از تحقیقات کشف تقلب را برای دسته‌بندی تحقیقات مبتنی بر چهار روش اصلی شامل رویکردهای نظارت شده، ترکیبی، نیمه نظارتی و بدون نظارت انجام دادند و همچنین رابطه کشف تقلب را با سایر زمینه‌ها شناسایی کردند. یک استراتژی کشف تقلب توسط هرلند، خوش‌گفتار، و بادر[3] (۲۰۱۸) برای تقلب مدیکر[4] با استفاده از سه مجموعه داده خدمات مدیکر و مدیکر پیشنهاد شده است. روش آنها بر روی مجموعه داده ترکیبی با اتصال چندین مجموعه داده آموزشی عمل می‌کند. نویسندگان از سه طبقه‌بندی کننده استفاده کردند: مدل‌های جنگل تصادفی، تقویت درخت گرادیان و مدل‌های رگرسیون لجستیک و از متریک ناحیه زیر منحنی (ROC) برای اندازه‌گیری عملکرد FD استفاده کردند. آنها به این نتیجه رسیدند که بالاترین خروجی در کشف تقلب روی مجموعه داده ترکیبی است. اندازه مجموعه داده مورد بحث قرار نمی‌گیرد، اما این تکنیک برای مجموعه داده‌های بزرگ که در آن مجموعه داده دیگری با ترکیبی از مجموعه داده‌های اولیه مورد نیاز است، بهینه نیست.

بای (۲۰۱۳) روشی را پیشنهاد می‌کند که در درجه اول یک گزینه جستجوی مؤثر برای اطلاعات بلادرنگ است، اما برای حوزه طبقه‌بندی مناسب نیست، چرا که هر عملیات باید به‌عنوان تقلب یا غیر تقلب طبقه‌بندی شود. کواچ و روجیرو (۲۰۱۱) پیشنهاد می‌کنند که سیستم ردیابی تقلب بر رفتار مشتریان متمرکز است و به شدت به دسترسی دستگاه بستگی دارد که برای اطلاعات بزرگ و در زمان واقعی مناسب نیست. راه‌حل کشف تقلب کارت اعتباری پیشنهاد شده توسط دومان و اوزچلیک (۲۰۱۱) از رویکردهای فراابتکاری استفاده می‌کند. با این حال، نویسندگان روش‌هایی برای برخورد با اطلاعات و اطلاعات بزرگ در زمان واقعی ارائه نکرده‌اند.

وی و همکاران (۲۰۱۳) یک تکنیک کشف تقلب را برای داده‌های بسیار نامتعادل با استفاده از الگوریتم ContrastMiner ارائه می‌دهند که الگوهای تضاد را استخراج می‌کند و تقلب را از رفتار واقعی متمایز می‌کند. کو و همکاران (۲۰۰۴) معتقدند که تحقیقات کشف تقلب عمدتاً از داده‌کاوی، آمار و هوش مصنوعی استفاده می‌کند و تقلب از ناهنجاری‌ها در داده‌ها و الگوها تشخیص داده می‌شود.

چن و همکاران (۲۰۱۹) یک سیستم تشخیص تقلب مبتنی بر گراف را برای بیمه تجارت الکترونیکی به نام InfDetect مورد بحث قرار دادند. این آزمایش بر روی داده‌های شرکت بیمه خصوصی از جمله بیمه سپرده تضمینی و بیمه باربری برگشت انجام شد. مدل آنها از ترکیبی از گراف‌ها و ویژگی‌های خام به‌عنوان داده ورودی، و ساخت ویژگی از طریق مدل یادگیری گراف تحت نظارت و بدون نظارت به نام DeepWalk استفاده می‌کند. رمزگذار خودکار حذف نویز و پردازش ویژگی نیز پیاده‌سازی شده

1- Duman and Ozcelik
2- Kou
3- Berland, Khoshgoftaar, and Bauder
4- Medicare

است. آنها سپس یک درخت تصمیم‌گیری مبتنی بر گرادیان تقویت‌شده مبتنی بر سرور به نام PSMART را برای خروجی احتمال تقلب اعمال می‌کنند. محققان ادعا کرده‌اند که طرح پیشنهادی آنها به صرفه‌جویی میلیون‌ها دلار در سال برای این شرکت‌های تجارت الکترونیک کمک می‌کنند. یکی از ویژگی‌های کلیدی مهندسی شده توسط محققان، استفاده از اختصاص امتیازهای bin برای مبالغ تراکنش بود که به بهبود عملکرد کمک کرد.

آراجو و همکاران (۲۰۱۷) الگوریتم "BreachRadar" را توسعه داده‌اند، یک الگوریتم متناوب توزیع شده که احتمال در معرض خطر قرار گرفتن کارت بانکی را به مکان‌های مختلف ممکن برای استفاده از کارت اختصاص می‌دهد. این نقاط سازش[1] (POC) نقطه شروع برای تقلب در تراکنش‌های بانکی است. یک مثال می‌تواند نقض داده‌ها باشد که منجر به کسب اطلاعات مشتریان بانکی می‌شود، از جمله اطلاعات کارت اعتباری آنها که سپس به‌صورت آنلاین از طریق سایت‌های DarkNet شسته می‌شود یا شخصاً توسط مجرمان استفاده می‌شود. مثال دیگر می‌تواند اسکیمینگ[2] کارت باشد، که در آن جزئیات کارت مشتری با استفاده از دستگاهی که عمداً برای به دست آوردن غیرقانونی این اطلاعات تغییر داده شده، کپی یا شبیه‌سازی می‌شود. هدف از شناسایی POC جلوگیری از استفاده جعلی از داده‌های مشتری است. محققان مدلی با عملکرد بالا با دقت ۹۰ درصد و یادآوری در برابر مجموعه داده‌ای با ۱۰ درصد کارت‌های اعتباری مشتریان را ارائه کردند. این تکنیک شامل تشکیل یک مدل گراف دوبخشی برای نشان دادن همه کارت‌ها و مکان آنها است. با اجرای یک الگوریتم در حافظه که امکان به‌روزرسانی احتمالات POC را فراهم می‌کند، محققان اولین روش توزیع شده را تولید کرده‌اند که قادر به تشخیص خودکار POC ها می‌باشد.

۳- یافته‌های تحقیق

رگرسیون لجستیک برای تشخیص تقلب

رگرسیون لجستیک[3] به یک روش یادگیری نظارت شده اشاره دارد که با تصمیمات قطعی استفاده می‌شود. به این معنی که نتایج به‌دست‌آمده در صورت انجام معامله، تقلب یا غیر تقلب محسوب می‌شود. این رویکرد از یک رابطه علت و معلولی برای توسعه مجموعه داده‌های سازمان‌یافته استفاده می‌کند. تکنیک تحلیل رگرسیون زمانی که در تشخیص تقلب مورد استفاده قرار می‌گیرد به دلیل چندین متغیر و اندازه مجموعه داده پیچیده‌تر است. این مدل (الگوریتم) پیش‌بینی می‌کند که آیا تراکنش‌های جدید به‌عنوان تقلبی علامت‌گذاری می‌شوند یا خیر. این مدل‌ها معمولاً برای مشتریان خود از تجار بزرگ‌تر دقیق هستند، اما معمولاً مدل‌های عمومی همچنان قابل اجرا هستند.

درخت تصمیم برای تشخیص تقلب

الگوریتم‌های درخت تصمیم[4] برای طبقه‌بندی فعالیت‌های غیر معمول در یک تراکنش از یک کاربر مجاز استفاده می‌شود. این الگوریتم‌ها دارای محدودیت‌هایی هستند که در مجموعه داده‌ها در طبقه‌بندی تقلب استفاده می‌شوند. الگوریتم‌های درخت تصمیم در طبقه‌بندی یا مشکلات مدل‌سازی برون‌یابی رگرسیون استفاده می‌شوند. آنها اساساً مجموعه قوانینی هستند که برای استفاده از موارد کلاه‌برداری شامل مشتریان مهارت دارند. ایجاد یک درخت تصمیم ویژگی‌های نامرتبط را نادیده می‌گیرد و نیازی به عادی‌سازی داده‌های گسترده ندارد. هنگامی‌که یک درخت تحت بازرسی قرار می‌گیرد، دلیل اینکه چرا یک تصمیم خاص بسته به لیست قوانین فعال شده توسط یک مشتری خاص گرفته‌شده است، درک می‌شود. خروجی الگوریتم یادگیری ماشین ممکن است مدلی باشد که میمون درخت تصمیم است. این یک امتیاز احتمالی تقلب را بر اساس شرایط قبلی تعیین

1- Points-of-Compromise
2- Skimming
3- Logistic Regression
4- Decision Tree

می‌کند.

جنگل تصادفی برای تشخیص تقلب

جنگل تصادفی[1] از ترکیبی از درختان تصمیم برای بهبود نتایج استفاده می‌کند. هر درخت معاملات را برای شرایط مختلف ارزیابی می‌کند (آیدیوارا، کی.؛ آیدیوارا، وی.، ۲۰۸)، همچنین مجموعه داده‌های تصادفی آموزش داده می‌شوند. بسته به آموزش درخت تصمیم، هر درخت یک تراکنش را به‌عنوان تقلبی یا غیر متقلبانه طبقه‌بندی می‌کند. سپس از مدل برای پیش‌بینی نتیجه استفاده می‌شود. این به آشکارسازهای تقلب اجازه می‌دهد تا خطای موجود در درخت را یکسان کنند. دقت مدل عملکرد کلی را افزایش می‌دهد و در عین حال توانایی تفسیر نتایج را حفظ می‌کند و امتیازات قابل توضیحی را برای کاربران ما ارائه می‌دهد.

زمان‌های اجرا جنگل تصادفی سریع هستند و داده‌هایی را که گم شده یا نامتعادل هستند مدیریت می‌کنند. ML های جنگل تصادفی دارای نقاط ضعفی هستند، مثلاً هنگامی‌که در رگرسیون استفاده می‌شوند، قادر به پیش‌بینی فراتر از تنوع در آموزش داده‌ها نیستند و ممکن است مجموعه‌های داده‌ای که نویز در نظر گرفته می‌شوند بیش از حد برازش کنند.

شبکه‌های عصبی برای تشخیص تقلب

شبکه‌های عصبی[2] مبتنی بر مغز انسان هستند. آنها از لایه‌های محاسباتی مختلفی استفاده می‌کنند. آنها از محاسبات شناختی استفاده می‌کنند که به ساخت ماشین‌هایی کمک می‌کند که می‌توانند از الگوریتم‌های خودآموزی که شامل داده‌کاوی، تشخیص الگوها و پردازش زبان طبیعی است استفاده کنند (گراپ، ۲۰۱۶). شبکه‌های عصبی برای فرآیند آموزش داده‌ها چندین لایه را پشت سر می‌گذارند. نتایج دقیق‌تری را در مقایسه با مدل‌های دیگر ارائه می‌دهد؛ زیرا از محاسبات شناختی استفاده می‌کند و از الگوهای رفتار مجاز می‌آموزد. بنابراین، بین معاملات «تقلبی» و غیر تقلبی تمایز قائل می‌شود. شبکه‌های عصبی کاملاً سازگار هستند و از مجموعه الگوهای رفتار مشروع درس می‌گیرند. اینها با تغییر در رفتار آنچه که تراکنش‌های استاندارد تلقی می‌شوند سازگار می‌شوند و اشکال تراکنش‌های تقلب را شناسایی می‌کنند. فرآیند شبکه‌های عصبی سریع است و در زمان واقعی کار می‌کند.

۴- نتیجه‌گیری

فعالیت‌های متقلبانه در طول دوره کرونا افزایش یافته است و سازمان‌هایی با روش‌های سنتی کشف تقلب ثابت کرده‌اند که نتایج چشمگیری داشته‌اند؛ زیرا چشم انسان همیشه ناهنجاری‌ها را در سیستم بانکی تشخیص نمی‌دهد. کلاه‌برداری مالی یکی از اصلی‌ترین مشکلاتی است که اعتماد مشتریان را تضعیف می‌کند و همچنین زیان اقتصادی به بانک‌ها و مؤسسات مالی وارد می‌کند. در سال‌های اخیر، هم‌زمان با گسترش تقلب، مؤسسات مالی به دنبال راه‌حلی برای یافتن راه‌حل مناسب در مبارزه با کلاه‌برداری بودند. با توجه به تغییرات پیشرفته و متنوع در روش‌های تقلب، تحقیقات گسترده‌ای برای کشف تقلب انجام شده است. تقلب پیچیده بانکداری آنلاین نشان‌دهنده سوء استفاده یکپارچه از منابع در دنیای اجتماعی، سایبری و فیزیکی است. با این حال، اطلاعات بسیار محدودی برای تشخیص کلاه‌برداری پویا از رفتار مشتری واقعی در چنین محیط داده‌ای بسیار کم و نامتعادل در دسترس است، که باعث می‌شود تشخیص فوری و مؤثر مهم و چالش‌برانگیز شود. بانک‌ها به دنبال کاهش زیان‌های هنگفت از طریق سیستم‌های تشخیص و پیشگیری از تقلب هستند. بسیاری از فن‌آوری‌های پیشرفته کلاه‌برداری برای شناسایی و پیشگیری از تراکنش‌های بانکی اینترنتی تقلبی استفاده می‌شوند. با این حال، آنها هیچ مکانیسم شناسایی مؤثری برای شناسایی کاربران قانونی و ردیابی فعالیت‌های غیرقانونی آنها ندارند. توجه به نتایج به‌دست‌آمده از این مطالعه، توصیه می‌شود که از الگوریتم‌های یادگیری ماشین برای کشف تقلب استفاده شود تا جایگزین سیستم‌های تشخیص تقلب مبتنی بر قانون نادرست شود

1- Random Forest
2- Neural Networks

منابع:

5-Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, *68*, 90-113.
6-ACI.(2011). Proactive Risk Manager. Retrieved from https://www.aciworldwide.com/products/proactive-risk-manager.
7-Aggarwal, C. C. (2017). Outlier Analysis. 2nd. *Cham, Switzerland: Springer*.
8-Amarasinghe, T., Aponso, A., & Krishnarajah, N. (2018, May). Critical analysis of machine learning based approaches for fraud detection in financial transactions. In *Proceedings of the 2018 International Conference on Machine Learning Technologies* (pp. 12-17).
9-APF. (2018). Western Australia Police Force. Retrieved from https://www.police.wa.gov.au/
10-Araujo, M., Almeida, M., Ferreira, J., Silva, L., & Bizarro, P. (2017, June). Breachradar: Automatic detection of points-of-compromise. In *Proceedings of the 2017 SIAM International Conference on Data Mining* (pp. 561-569). Society for Industrial and Applied Mathematics.
11-Ayyadevara, V. K., & Ayyadevara, V. K. (2018). Gradient boosting machine. *Pro machine learning algorithms: A hands-on approach to implementing algorithms in python and R*, 117-134.
12-Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. *Innovative Technology at the Interface of Finance and Operations: Volume I*, 223-247.
13-Bilgin, M. H., MARCO LAU, C. K., & Demir, E. (2012). Technology transfer, finance channels, and SME performance: New evidence from developing countries. *The Singapore Economic Review*, *57*(03),

1250020.
14-Brabazon, A., Cahill, J., Keenan, P., & Walsh, D. (2010, July). Identifying online credit card fraud using artificial immune systems. In *IEEE Congress on Evolutionary Computation* (pp. 1-7). IEEE.
15-Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical science*, *17*(3), 235-255.
16-Capterra. (2019). Fraud Management Software. Retrieved from https://www.capterra.com.au/directory/10058/financial-fraud-detection/software
17-Carminati, M., Polino, M., Continella, A., Lanzi, A., Maggi, F., & Zanero, S. (2018). Security evaluation of a banking fraud analysis system. *ACM Transactions on Privacy and Security (TOPS)*, *21*(3), 1-31.
18-Carminati, M., Caron, R., Maggi, F., Epifani, I., & Zanero, S. (2015). BankSealer: A decision support system for online banking fraud analysis and investigation. *computers & security*, *53*, 175-186Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). Credit card fraud detection using pipeling and ensemble learning. *Procedia Computer Science*, *173*, 104-112.
19-Chandola, V., & Banerjee, A. V., K.(2009). Anomaly detection: A survey. *ACM Computing survey*, *41*.
20-Chen, C., Liang, C., Lin, J., Wang, L., Liu, Z., Yang, X., ... & Qi, Y. (2019, December). InfDetect: A large scale graph-based fraud detection system for E-commerce insurance. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 1765-1773). IEEE.
21-Chen, C. M., Guan, D. J., Huang, Y. Z., & Ou, Y. H. (2012, August). Attack sequence detection in cloud using hidden markov model. I
22-Chen, L., Zhang, Z., Liu, Q., Yang, L., Meng, Y., & Wang, P. (2019, May). A method for online transaction fraud detection based on individual behavior. In *Proceedings of the ACM Turing Celebration Conference-China* (pp. 1-8).
23-Chen, K., Yadav, A., Khan, A., & Zhu, K. (2020). Credit Fraud Detection Based on Hybrid Credit Scoring Model. *Procedia Computer Science*, *167*, 2-8.
24-Chiu, C. Y., Yeh, C. T., & Lee, Y. J. (2013, December). Frequent pattern based user behavior anomaly detection for cloud system. In *2013 Conference on Technologies and Applications of Artificial Intelligence* (pp. 61-66). IEEE.
25-Choi, S., Kim, C., Kang, Y. S., & Youm, S. (2021). Human behavioral pattern analysis-based anomaly detection system in residential space. *The Journal of Supercomputing*, *77*, 9248-9265.
26-CURTIS, C. V. (2012). The Fraud Triangle: Fraudulent Executives, Complicit Auditors and Intolerable Public Injury.
27-Darvishi, H., Ciuonzo, D., Eide, E. R., & Rossi, P. S. (2020). Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture. *IEEE Sensors Journal*, *21*(4), 4827-4838.
28-Dehaven, V. R. (2014). Machine learning: Future capabilities and their implications.
29-Dong, F., Wang, H., Li, L., Guo, Y., Bissyandé, T. F., Liu, T., ... & Klein, J. (2018, October). Frauddroid: Automated ad fraud detection for android apps. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering* (pp. 257-268).
30-Duman, E., & Ozcelik, M. H. (2011). Detecting credit card fraud by genetic algorithm and scatter search. *Expert Systems with Applications*, *38*(10), 13057-13063.
31-Eurostat, G. D. P., & Components, M. (2020). Available online: https://ec. europa. eu/eurostat/databrowser/view. *NAMQ_10_GDP/default/table*.
32-FBI. (2018). Internet Crime Complaint Center, 20182019(20/08/2019). Retrieved from https ://pdf.ic3. gov /2018 _ I C3 Report. Pdf
33-FICO. (2010). Fico Application Fraud Manager. Retrieved from https ://www.fico.com/en/products/fico-application- fraud-manager
34-FraudNet. (1997). Enterprise Fraud Prevention. Retrieved from https://fraud.net
35-Gbudeanu, L., Brici, I., Mare, C., Mihai, I. C., & cheau, M. C. (2021). Privacy intrusiveness in

financial-banking fraud detection. *Risks*, *9*(6), 104.
36-Graupe, D. (2016). *Deep learning neural networks: Design and case studies*. World Scientific Publishing Company.
37-Hafiz, K. T., Aghili, S., & Zavarsky, P. (2016, June). The use of predictive analytics technology to detect credit card fraud in Canada. In *2016 11th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE.
38-Herland, M., Khoshgoftaar, T. M., & Bauder, R. A. (2018). Big data fraud detection using multiple medicare data sources. *Journal of Big Data*, *5*(1), 1-21.
39-Herschel, G., Linden, A., & Kart, L. (2015). Magic quadrant for advanced analytics platforms. *Gartner Report G*, *270612*.
40-Hopton, D. (2016). Proceeds of Crime Act 2002–Part 7: Requirements and Offences. In *Money Laundering* (pp. 53-78). Gower.
41-Horák, M., Stupka, V., & Husák, M. (2019, August). GDPR compliance in cybersecurity software: a case study of DPIA in information sharing platform. In *Proceedings of the 14th international conference on availability, reliability and security* (pp. 1-8).
42-Ilgun, K., Kemmerer, R. A., & Porras, P. A. (1995). State transition analysis: A rule-based intrusion detection approach. *IEEE transactions on software engineering*, *21*(3), 181-199.
43-Jha, S., Guillen, M., & Westland, J. C. (2012). Employing transaction aggregation strategy to detect credit card fraud. *Expert systems with applications*, *39*(16), 12650-12657.
44-John, S. N., Anele, C., Kennedy, O. O., Olajide, F., & Kennedy, C. G. (2016, December). Realtime fraud detection in the banking sector using data mining techniques/algorithm. In *2016 international conference on computational science and computational intelligence (CSCI)* (pp. 1186-1191). IEEE.
45-Kou, Y., Lu, C. T., Sirwongwattana, S., & Huang, Y. P. (2004, March). Survey of fraud detection techniques. In *IEEE International Conference on Networking, Sensing and Control, 2004* (Vol. 2, pp. 749-754). IEEE.
46-Kovach, S., & Ruggiero, W. V. (2011, February). Online banking fraud detection based on local and global behavior. In *Proc. of the Fifth International Conference on Digital Society, Guadeloupe, France* (pp. 166-171).
47-Lavion, D. (2018). Global Economic Crime and Fraud Survey 2018, 30. Retrieved from Pulling fraud out of the shadows website: https://www.pwc.com/gx/en/forensics/global¬economic-crime-and-fraud-survey-
48-Lee, J. H., Park, M. W., Eom, J. H., & Chung, T. M. (2011, February). Multi-level intrusion detection system and log management in cloud computing. In *13th International Conference on Advanced Communication Technology (ICACT2011)* (pp. 552-555). IEEE.
49-Legg, C. (2007). Ontologies on the semantic web. *Annual review of information science and technology*, *41*, 407.
50-Li, Z., Huang, M., Liu, G., & Jiang, C. (2021). A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection. *Expert Systems with Applications*, *175*, 114750.
51-Marican, L., & Lim, S. (2014). Microsoft Consumer Safety Index reveals impact of poor online safety behaviours in Singapore. Retrieved from https://news.microsoft.com/en¬s g/2014/02/ 11 /microsoft-consumer-safety-index-reveals-impact-of-poor-online-safety.
52-Maruatona, O. (2013). *Internet banking fraud detection using prudent analysis* (Doctoral dissertation, University of Ballarat).
53-Maruti T. (2021). How Machine Leaming Facilitates Fraud Detection?. [online] Available at:<https ://marutitech.com/machine-leaming-fraud-
54-Marvasti, M. A., Poghosyan, A. V., Harutyunyan, A. N., & Grigoryan, N. M. (2014). An enterprise dynamic thresholding system. In *11th International Conference on Autonomic Computing ({ICAC} 14)* (pp. 129-135).

55-Misra, S., Thakur, S., Ghosh, M., & Saha, S. K. (2020). An autoencoder based model for detecting fraudulent credit card transaction. *Procedia Computer Science*, *167*, 254-262.
56-Moon, W. Y., & Kim, S. D. (2017). Adaptive fraud detection framework for fintech based on machine learning. *Advanced Science Letters*, *23*(10), 10167-10171.
57-Nanduri, J., Jia, Y., Oka, A., Beaver, J., & Liu, Y. W. (2020). Microsoft uses machine learning and optimization to reduce e-commerce fraud. *INFORMS Journal on Applied Analytics*, *50*(1), 64-79.
58-Nathan, R. J., Victor, V., Gan, C. L., & Kot, S. (2019). Electronic commerce for home-based businesses in emerging and developed economy. *Eurasian Business Review*, *9*, 463-483.
59-Nicholls, J., Kuppa, A., & Le-Khac, N. A. (2021). Financial cybercrime: A comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape. *Ieee Access*, *9*, 163965-163986.
60-Orek, M., Örek, E., & Bahtiyar, . (2019, May). A deep learning method for fraud detection in financial systems: Poster. In *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks* (pp. 298-299).
61-Olowookere, T. A., & Adewale, O. S. (2020). A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach. *Scientific African*, *8*, e00464.
62-Ocansey, E. O. N. D., & Ganu, J. (2017). The role of corporate culture in managing occupational fraud. *Research Journal of Finance and Accounting*, *8*(24).
63-Pascual, A., Marchini, K., & Miller, S. (2017). Identity fraud: securing the connected life. *Retrieved on January*, *20*, 2018.
64-PattemSpy. (2015). PattemSpy For Banking - Fraud Management Software. Retrieved from https:// www.pattemspy.tech/
65-Pinto, S. O., & Sobreiro, V. A. (2022). Literature review: Anomaly detection approaches on digital business financial systems. *Digital Business*, 100038.
66-Politou, E., Alepis, E., & Patsakis, C. (2019). Profiling tax and financial behaviour with big data under the GDPR. *Computer law & security review*, *35*(3), 306-329.
67-Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*.
68-Richards, D. (2003). Knowledge-based system explanation: The ripple-down rules alternative. *Knowledge and Information Systems*, *5*(1), 2-25.
69-RiskNet. (1998). Fraud Managed Services. Retrieved from https :/ /www.aicorporation.com/ products-services/fraud-managed-services/
70-SAS. (2007). SAS Fraud Management. Retrieved from https://www.sas.com/en_ au/software/fraud-management.html
71-Sando, S., 2021. Consumer Preference Drives Shift in Authentication. [online] Javelin. Available at: <https ://www .j ave I instrate gy. com/ coverage-area/
72-cheau, M. C., Gaftea, V. N., Achim, M. V., & Cotoc, C. N. B. (2020, October). Cyber Security Reactivity in Crisis Times and Critical Infrastructures. In *2020 24th International Conference on System Theory, Control and Computing (ICSTCC)* (pp. 691-698). IEEE.
73-Shihembetsa, E. (2021). *Use of artificial intelligence algorithms to enhance fraud detection in the Banking Industry* (Doctoral dissertation, University of Nairobi).
74-Stalla-Bourdillon, S., Pearce, H., & Tsakalakis, N. (2018). The GDPR: A game changer for electronic identification schemes? The case study of Gov. UK Verify. *Computer Law & Security Review*, *34*(4), 784-805.
75-Stojanov, M. (2019). Protection Against Fraud In Electronic Trade Payments. *21*, *9*(1-ENG), 48-66.
76-Sudharsan, K., Kumar, V. A., Venkatesan, R., Sathyapreiya, V., & Saranya, G. (2019). Two three step authentication in ATM machine to transfer money and for voting application. *Procedia Computer Science*, *165*, 300-306.
77-Syed, N. A., Liu, H., & Sung, K. K. (1999, August). Handling concept drifts in incremental learning

with support vector machines. In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 317-321).
78-Wang, Y., & Wang, L. (2019, May). Bot-like Behavior Detection in Online Banking. In *Proceedings of the 4th International Conference on Big Data and Computing* (pp. 140-144).
79-Wei, W., Li, J., Cao, L., Ou, Y., & Chen, J. (2013). Effective detection of sophisticated online banking fraud on extremely imbalanced data. *World Wide Web*, *16*, 449-475.
80-Williams, M. (2022). Elizabeth Holmes and Theranos: A play on more than just ethical failures. *Business Information Review*, *39*(1), 23-31.