مروری بر روش‌های کشف تقلب بانکی با استفاده از هوش مصنوعی

سیدمجید اکبرالسادات [۱]، بابک اسماعیل پور قوچانی [۲]،

۱- گروه هوش مصنوعی و رباتیکز ، واحد قوچان ، دانشگاه آزاد اسلامی قوچان ، ایران
Majid.akbarosadat@gmail.com
۲- گروه هوش مصنوعی و رباتیکز ، واحد قوچان ، دانشگاه آزاد اسلامی قوچان ، ایران
نویسنده مسئول: babakesma@yahoo.com

خلاصه

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 [۱]. در طول رکود اقتصادی ناشی از همه‌گیری COVID-۱۹ به دلیل افزایش تعداد افراد بیکار در طول دوره، شاهد افزایش فعالیت‌های کلاه‌برداری بوده‌ایم [۲]. هزاران نفر بیکار شدند، حقوق‌ها کاهش یافت و نرخ بیکاری افزایش یافت. به‌طور 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 [۳]. به گفته محققین،

کلاه‌برداری زمانی اتفاق می‌افتد که فردی فشار و انگیزه غیرقابل تقسیم برای ارتکاب کلاه‌برداری ایجاد کند.

در بیشتر موارد، مرتکب تقلب نیازی برآورده نشده اما با منابع محدود دارد. نیازهای برآورده نشده بی‌پایان است و برای افراد مختلف متفاوت می‌باشد. این 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 افتضاح 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 [۴]. 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 [۵].

کشف دقیق تقلب (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 یک مسئله حائز اهمیت تبدیل شده است [۶]. بررسی شاخص ایمنی رایانه مایکروسافت نشان داد که تأثیر سالانه فیشینگ و اشکال مختلف سرقت هویت می‌تواند 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 [۷].

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[۱۱] (۲۰۲۰) Ciobanu مطالعه‌ای را با ۱۰۰۰ مصرف‌کننده بزرگسال انجام داد و در آن متوجه شد حدود ۲۵٪ 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 .[۸] 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 [۹] در دسترس بودن اطلاعات دقیق در مورد سیستم کشف تقلب بسیار محدود است زیرا صنعت بانکداری به‌ندرت آمار کشف تقلب را منتشر می‌کند. به‌ویژه، 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[۱۰] . 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 [۱۱]. 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.

در دو دهه گذشته، با تکامل فناوری مورد استفاده در بخش بانکداری مالی، طرح‌های تقلب مورد استفاده کلاه‌برداران نیز تغییر کرد [۱۲]. 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[۱۲] (۲۰۲۰) [۱۳] بر افزایش هدف قرار دادن بازرگانان 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 [۱۴]. حدود ۵۶ درصد از اروپایی‌ها نگران این هستند که قربانی تقلب شوند [۱۵]. در سال ۲۰۱۹، ۲۶ درصد از جمعیت اتحادیه اروپا گزارش دادند که پیام‌های جعلی دریافت کرده‌اند، از جمله پیام‌های مربوط به اعتبار بانکداری الکترونیک [۱۶]. خانواده‌های مختلف بدافزار آسیب‌های مختلفی به مصرف‌کننده، زیرساخت‌های حیاتی، مؤسسات

مالی و بانکی وارد کرده‌اند و به اهداف مورد نظر تبدیل شده‌اند [۱۷]. 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 [۱۳].

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. پیشینه تحقیق

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[۱۸] و همکاران (۲۰۱۸)، در دهه گذشته شیوه‌های مختلفی را که می‌توان از چنین جزئیاتی برای کشف و پیشگیری از تقلب استفاده کرد، تحلیل کرده‌اند [۱۸]. نتیجه‌گیری‌های آن‌ها منجر به رویکردهای مختلفی برای الگوریتم‌های کشف تقلب، با تأکید بر روش‌های یادگیری ماشین شده است [۱۹]، زیرا تجمیع داده‌ها و تحلیل تاریخی داده‌ها می‌تواند به یافتن الگوهای تقلب کمک کند [۲۰].

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 [۲۱]، هم برای الگوریتم‌های کارآگاهی و هم برای الگوریتم‌های پیش‌بینی تمرکز کرده‌اند.

با توجه به انواع الگوریتم‌های تشخیص تقلب پیشنهاد شده در مقالات تحقیقاتی منتشر شده در سه سال گذشته (۲۰۱۸ تا 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 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 بانکیGDPR	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

از سال ۲۰۱۸، مطالعات مربوط به سیاست حفظ حریم خصوصی داده‌های شخصی[1] (سیاست GDPR) شروع به ظاهر شدن کردند. استالا بوردیلون[2] و همکاران (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 [۲۲]. علاوه بر این، در ارتباط با این موضوع، مطالعاتی در مورد میزان نفوذ در درخواست داده‌های شخصی شروع شده است. در این رابطه هوراک[3] و همکاران (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 [۲۳].

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 به‌منظور 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 [۲۴]، تشخیص تقلب در کارت اعتباری با استفاده از روش‌های Pipeling و Ensemble Learning [۲۵]، و تشخیص تقلب در کارت اعتباری با استفاده از شبکه عصبی مصنوعی [۲۶].

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 [۲۷].

اوگرک و همکاران (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 (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 [۲۸]. همچنین، لی و همکاران (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 [۲۹].

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، اما نتایج مطالعه نشان داد که می‌تواند مجموعه داده مفیدی در سراسر جهان ارائه کند [۱۸]. چن و همکاران (۲۰۱۹) نیز روشی را برای تشخیص تراکنش‌ها بر اساس رفتار افراد ایجاد کردند. مدلی که در مقاله آنها استفاده شد؛ هایپر کره نام دارد. این مطالعه به این نتیجه رسید که اثر توصیف رفتار انسان با فراوانی معاملات آنها مرتبط است [۳۰]. علاوه بر این، وانگ و وانگ (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 برتری دارد [۳۱].

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 [۳۲]. مصرا و همکاران (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 [۳۳]. اولووکره و همکاران (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 [۳۴].

مطالعات دیگر بر روی تشخیص‌های خاص‌تر متمرکزشده‌اند. آماراسینگه[□□□□] و همکاران (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 [۳۵]. دونگ[□□□] و همکاران (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 تجزیه‌وتحلیل ۱۲۰۰۰ برنامه مشکوک، FraudDroid ۳۳۵ مورد تقلب را شناسایی کرد و این نتیجه تأیید می‌کند که روش مفیدی برای کشف این نوع جرم است [۳۶]. علاوه بر این، سودارسان[□□] و همکاران (۲۰۱۹) 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 [۳۷].

۳. چالش‌های بانکداری آنلاین

هر ساله تقلب در اشکال مختلف افزایش می‌یابد که منجر به خسارات مالی قابل‌توجهی می‌شود [۳۸]. 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 [۷]. 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 [۳۹] از سال ۲۰۱۶ بالغ بر ۱۳ درصد افزایش در تقلب رخ داده است. IC۳ یک منبع ارزشمند برای قربانیان جرائم اینترنتی و مجریان قانون در شناسایی، رسیدگی و تعقیب جرائم است [۴۰]. IC۳ 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 [۴۱] 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 [۳۸, ۴۴, ۴۳, ۱۰, ۴۲].

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 [۹, ۴۳]. 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□□□ (۲۰۰) نیز [۱۰] 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 (۲۰۱۰) [۱۱] تأکید می‌کنند که کشف تقلب با استفاده از تکنیک‌های داده‌کاوی در صنعت بسیار رایج است.

۴. 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.

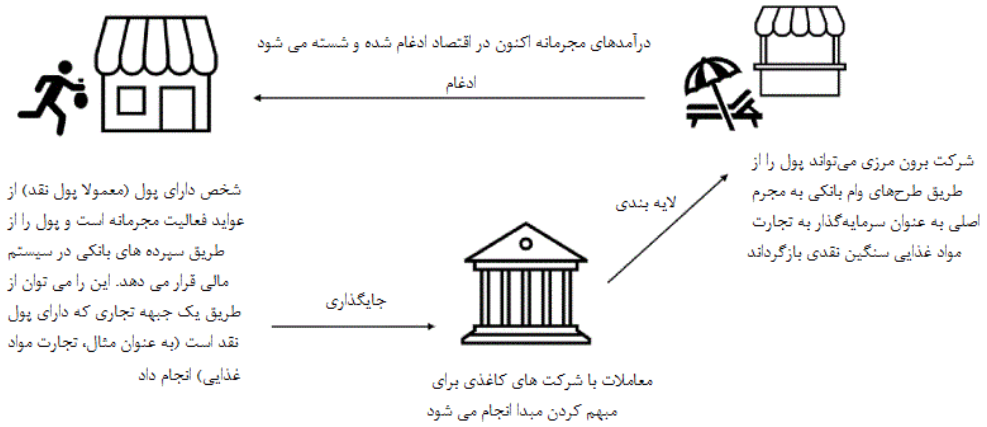
دستگیری این بازیگران کلاه‌بردار آسان نیست و معمولاً به نیروی کار زیادی برای جمع‌آوری شواهد در مدت زمان طولانی نیاز دارد، به‌ویژه در موارد تجارت داخلی. با این حال، پیشرفت‌های اخیر در کاربردها و تکنیک‌های یادگیری ماشین به شناسایی بازیگران بد به روشی کارآمدتر و سریع‌تر کمک می‌کند. برخی از روش‌های مورد استفاده توسط افرادی که فعالیت سرمایه‌گذاری متقلبانه انجام می‌دهند عبارتند از: دست‌کاری بازار، تجارت داخلی، پول‌شویی، تأمین مالی تروریسم و بسیاری موارد دیگر. دست‌کاری بازار، 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.

۵. 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□□□□□ (CPS) بریتانیا [۴۵] 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. نمونه‌ای از پول‌شویی با استفاده از قرار دادن، لایه‌بندی و ادغام را می‌توان در شکل ۱ مشاهده کرد.



شکل ۱- مراحل پول‌شویی

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 [۴۶].

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 [۱۳].

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 [۱]. 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.

۶. 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، Mule recruitment، کلاه‌برداری در سایت خرید و حراجی، کلاه‌برداری، هرزنامه، و سرقت هویت [۴۷] انواع مختلفی از کلاه‌برداری‌های آنلاین هستند. 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 [۴۸] به‌عنوان "اعمال غیرقانونی انجام شده توسط یک فرد یا گروهی از افراد برای به دست آوردن مزیت مالی یا حرفه‌ای" تعریف شده است. انگیزه اصلی در چنین جنایاتی، منفعت اقتصادی است. مجمع جهانی اقتصاد، جرائم مالی را صنعتی تریلیون دلاری می‌داند [۴۹]. 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□□□□ بیشتر طول می‌کشد [۵۰]. 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٪ از مشتریان پس از تجربه کلاه‌برداری، بانک را تغییر می‌دهند [۵۱]. این تعداد مشتریانی که از یک بانک خاص خارج می‌شوند، به‌طور قابل‌توجهی برای عملیات تجاری یک بانک مضر است، به‌خصوص 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□□□ (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.

۶-۱. سیستم FICO

سیستم FICO یکی از پرکاربردترین سیستم‌های کشف تقلب مورد استفاده توسط بانک‌ها در سطح جهان است [۵۲, ۵۳]. 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 [۵۴].

۶-۲. سیستم PRM

یکی دیگر از سیستم‌های کشف تقلب پرکاربرد PRM است [۵۵]. 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 [۵۶].

۶-۳. سیستم SAS Fraud Manager

سیستم SAS Fraud Manager [۵۷] عمدتاً راه‌حل کشف تقلب برای کارت‌های بدهی و اعتباری است و قابلیت بلادرنگ را دارد. این سیستم کشف تقلب از تکنیک تشخیص ناهنجاری□□□□ (AD) استفاده می‌کند.

۶-۴. سیستم Kount

Kount یکی دیگر از سیستم‌های پیشگیری از تقلب (Kount, ۲۰۰۶) است که از هر دو مدل ML□□□□□ 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 [۵۸, ۵۹, ۶۰].

۷. 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٪ مؤثر نیست. [۶۱, ۵۶, ۶۲] 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.

۸. سیستم‌های مبتنی بر قانون (RBS)

سیستم‌های مبتنی بر قانون□□□□□□ بخشی از گروه بزرگی از رویکردها هستند که سعی در مدل‌سازی تخصص انسان به نام سیستم‌های مبتنی بر دانش□□□□ (KBS) دارند. KBS سنتی از جمله RBS دارای مسائل مشترک KA، شکنندگی□□□ و یادگیری افزایشی□□□□ است. اصطلاح شکنندگی توسط لنات□□□□□ [۶۳] 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 [۶۴].

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 [۶۵].

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 [۶۵]. 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.

۹. 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□□□□□ و همکاران (۲۰۱۹) نمونه‌ای را توصیف می‌کند که در آن یک نمونه اولیه یادگیری ماشین بر روی مجموعه داده‌ای برنامه‌ریزی شد 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. به‌عنوان‌مثال، یکی از اشکالات این است که ML برای دستیابی به دقت، به مقدار قابل‌توجهی 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.

۱۰. سیستم‌های تشخیص نفوذ (IDS)

با توجه به [۶۶] سیستم‌های تشخیص نفوذ□□□□□ (IDS) می‌توانند میزان ایمنی بیشتری را ارائه دهند، اما نسبت به شدت ایمنی، به منابع محاسباتی بسیار بیشتری نیاز دارند. 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. محققین، [۶۷] 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□□□□ قبلی برای حمله به سیستم استفاده می‌کنند. 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□□□□□ 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 توسط [۶۷] برای IDS، عمدتاً بر شناسایی رویدادهای بالقوه، ثبت داده‌ها و تلاش‌های نظارتی تمرکز دارند.

۱۱. 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□□□□□□، تشخیص تازگی□□□□□□□، تشخیص نویز□□□□□ و تشخیص انحراف□□. تشخیص

ناهنجاری فرآیند تجزیه‌وتحلیل مجموعه داده 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، [۶۸] 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 [۶۹] و تشخیص خطای حسگر [۷۰].

محققین، [۷۱] یک تکنیک تشخیص نفوذ مبتنی بر قانون (ID) پیشنهاد و اذعان می‌کنند که AD یکی از اولین رویکردها برای ID است. سایر محققین [۷۲] 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 (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 [۷۲]. برابازون[□□□□] و همکاران (۲۰۱۰) همچنین استفاده از یک رویکرد مبتنی بر AIS برای AD را برای کاهش تقلب در کارت اعتباری توصیف می‌کنند [۵۲].

تجزیه‌وتحلیل غیرعادی برای داده‌های گزارش جریانی توسط [۷۳] انجام شده است. 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 توسط هاروتونیان[□□□□□] و همکاران (۲۰۱۴) [۷۳] در درجه اول مبتنی بر تجزیه‌وتحلیل غیرعادی داده‌های رویداد از فایل‌های گزارش و به دست آوردن اطلاعات مفید از منبع و انواع رویدادها است.

۱۲. 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 [۳۸, ۹]. بیشترین رویکردهای مورد استفاده برای کشف تقلب روش‌های نظارت شده، بدون نظارت، نیمه نظارت شده و ترکیبی هستند [۴۲, ۱۰, ۷۴, ۴۴].

کواچ و روجیرو[۷۵] (۲۰۱۱) [۷۵] یک سیستم کشف تقلب را برای بانکداری آنلاین با تمرکز بر تجزیه‌وتحلیل محلی و جهانی رفتار کاربران پیشنهاد 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 مجزایی که هر دستگاه به آن دسترسی دارد، تمرکز دارد.

روشی برای شناسایی تقلب در کارت اعتباری توسط دومان و اوزچلیک[۷۶] (۲۰۱۱) [۷۶] 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[۷۵] و همکاران (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 [۶]. یک استراتژی کشف تقلب توسط هرلند، خوش‌گفتار، و بادر[۷۵] (۲۰۱۸) [۷۷] برای تقلب مدیکر[۷۵] 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[۷۵] (۲۰۱۳) Bai 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) [۷۵] 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) [۷۶] از رویکردهای فراابتکاری استفاده می‌کند. با این حال، نویسندگان روش‌هایی برای برخورد با اطلاعات و اطلاعات بزرگ در زمان واقعی ارائه نکرده‌اند.

وی[۷۵] و همکاران (۲۰۱۳) Wei 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 [۶].

چن و همکاران [۷۸] 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 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 [۷۹] 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 (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 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.

۱-۱۲. رگرسیون لجستیک برای تشخیص تقلب

رگرسیون لجستیک 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.

۲-۱۲. درخت تصمیم برای تشخیص تقلب

الگوریتم‌های درخت تصمیم[۱۰] 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 می‌کند.

۳-۱۲. جنگل تصادفی برای تشخیص تقلب

جنگل تصادفی[۱۱] از ترکیبی از درختان تصمیم برای بهبود نتایج استفاده می‌کند. هر درخت معاملات را برای شرایط مختلف ارزیابی می‌کند [۸۰]. 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.

۴-۱۲. شبکه‌های عصبی برای تشخیص تقلب

شبکه‌های عصبی[۱۲] 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 [۸۱]. 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.

۱۳. نتیجه‌گیری

فعالیت‌های متقلبانه در طول دوره COVID-۱۹ افزایش‌یافته 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.

۱۴. مراجع

۱ Acfe.com. ۲۰۲۱. Association of Certified Fraud Examiners - Fraud ۱۰۱. [online] Available at: <https://www.acfe.com/fraud-۱۰۱.aspx> [Accessed ۷ March ۲۰۲۱].

۲ COLVIN, G., ۲۰۲۰. The pandemic may be the greatest environment for business fraud in decades. [online] Fortune. Available at: <https://fortune.com/۲۰۲۰/۱۱/۱۲/pandemic-corporate-fraud¬scams/#:-:text=Fraud%۲۰experts%۲۰say%۲۰every%۲۰corporate,so%۲۰they%۲۰re

۳ Littman, A., ۲۰۱۱. The Fraud Triangle: Fraudulent Executives, Complicit Auditors and Intolerable Public Injury. CreateSpace Independent Publishing Platform.

٤ Maruti Techlabs. ۲۰۲۱. How Machine Leaming Facilitates Fraud Detection?. [online] Available at:<https ://marutitech.com/machine-leaming-fraud-

٥ J. Nicholls, A. Kuppa and N. -A. Le-Khac, "Financial Cybercrime: A Comprehensive Survey of Deep Learning Approaches to Tackle the Evolving Financial Crime Landscape," in IEEE Access, vol. ۹, pp. ١٦٣٩٦٥-١٦٣٩٨٦, ۲۰۲۱, doi: ۱۰,۱۱۰۹/ACCESS.۲۰۲۱,۳۱۳٤۰۷٦.

٦ Kou, Y., Lu, C.-T., Sirwongwattana, S., & Huang, Y.-P. (۲۰۰٤). Survey of fraud detection techniques. Paper presented at the IEEE International Conference on Networking, Sensing and Control, ۲۰۰٤.

۷ Marican, L., & Lim, S. (۲۰۱٤). Microsoft Consumer Safety Index reveals impact of poor online safety behaviours in Singapore. Retrieved from https://news.microsoft.com/en¬s g/۲۰۱٤/۰۲/ ۱۱ /microsoft-consumer-safety-index-reveals-impact-of-poor-online-safety

۸ Richards, D. (۲۰۰۳). Knowledge-based system explanation: The ripple-down rules

alternative. Knowledge and Information Systems, ٥(١), ٢-٢٥. doi:١٠,١٠٠٧/s١٠١١٥-٠٠٢- ٠٠٧٦-٣

٩ Maruatona, ٠. (٢٠١٣). Internet banking fraud detection using prudent analysis. (PHD), University of Ballarat.

١٠ Bolton, R. J., & Hand, D. J. (٢٠٠٢). Statistical Fraud Detection: A Review. Statistical Science, ١٧(٣), ٢٣٥-٢٥٥.

١١ Phua, C., Lee, V., Smith, K., & Gayler, R. (٢٠١٠). A comprehensive survey of data mining–based fraud detection research. arXiv preprint arXiv.I ٠٠٩,٦١١٩.

١٢ European Payments Council. ٢٠١٩. Payment Threats and Fraud Trends Report. Brussels: European Payments Council

١٣ Nilson Report. ٢٠٢٠. Issue ١١٨٧-December ٢٠٢٠. Available online: https://nilsonreport.com/publication_newsletter_archive_issue.php?issue=١١٨٧ (accessed on ٩ April ٢٠٢١)

١٤ Nathan, R.J., V. Victor, and C. L. Gan. ٢٠١٩. Electronic commerce for home-based businesses in emerging and developed economy. Eurasian Business Review ٩: ٤٦٣–٨٣

١٥ Eurobarometer. ٢٠١٥. Special Eurobarometer ٤٢٣, Cybersecurity. Eurobarometer. Available online: https://www.adepp.info/wp-content/uploads/٢٠١٦/٠٧/studio-su-cybercrime.pdf (accessed on ٩ April ٢٠٢١)

١٦ Eurostat. ٢٠٢٠. ICT Usage in Households and by Individuals. Eurostat. Available online: https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm (accessed on ٩ April ٢٠٢١)

١٧ Șcheau, Mircea Cosntantin, Viorel Nicolae Gaftea, Monica Violeta Achim, and Corina-Narcisa Cotoc. ٢٠٢٠. Cyber Security Reactivity in Crisis Times and Critical Infrastructures. Paper presented at ٢٤th International Conference on System Theory, Control and

١٨ Carminati, Michele, Mario Polino, Andrea Continella, Andrea Lanzi, Federico Maggi, and Stefano Zanero. ٢٠١٨. Security Evaluation of a Banking Fraud Analysis System. ACM Transactions on Privacy and Security ٢١: ٣

١٩ Yang, Bao, Hilary Gilles, and Ke Bin. ٢٠٢٠. Artificial Intelligence and Fraud Detection. Innovative Technology at the interface of Finance and Operations. Springer Series in Supply Chain Management. Springer Nature

٢٠ Politou, Eugenia, Efthimios Alepis, and Constantinos Patsakis. ٢٠١٩. Profiling tax and financial behaviour with big data under the GDPR. Computer Law & Security Review ٣٥: ٣٠٦–٢٩

٢١ Jha, Sanjeev, Montserrat Guillen, and J. Christopher Westland. ٢٠١٢. Employing transaction aggregation strategy to detect credit card fraud. Expert Systems with Applications ٣٩: ١٢٦٥٠–٥٧

٢٢ Stalla-Bourdillon, Sophie, Pearce Henry, and Tsakalakis Niko. ٢٠١٨. The GDPR: A game changer for electronic identification schemes? The case study of Gov.UK Verify. Computer Law & Security Review ٣٤: ٧٨٤–٨٠٥

٢٣ Horak, Martin, Václav Stupka, and Martin Husák. ٢٠١٩. GDPR Compliance in Cybersecurity Software: A Case Study of DPIA in Information Sharing Platform. In

ARES '۱۹: Proceedings of the ۱٤th International Conference on Availability, Reliability and Security.

۲٤ Dornadula, Vaishnavi, and Nath S Geetha. ۲۰۱۹. Credit Card Fraud Detection using Machine Learning Algorithms. Procedia Computer Science ۱٦٥: ٦۳۱–٤۱

۲٥ Bagga, Siddhant, Goyal Anish, Gupta Nmaita, and Arvind Goyal. ۲۰۲۰. Credit Card Fraud Detection using Pipeling and Ensemble Learning. Procedia Computer Science ۱۷۳: ۱۰٤–۱۲

۲٦ Asha, R. B., and K. R. Suresh Kumar. ۲۰۲۱. Credit Card Fraud Detection Using Artificial Neural Network, Global Transitions Proceedings. Journal Pre-Proof

۲۷ Mehmet, Huseyin Bilgin, Chi Keung, Marco Lau, and Ender Demir. ۲۰۱۲. Technology Transfer, Finance Channels, and SME Performance: New Evidence from Developing Countries, The Singapore Economic Review (SER). Singapore: World Scientific Publishing Co. Pte. L

۲۸ Öğrek, Mahmut, Öğrek Eyüp, and Bahtiyar Şerif. ۲۰۱۹. A deep learning method for fraud detection in financial systems: Poster. In WiSec '۱۹: Proceedings of the ۱۲th Conference on Security and Privacy in Wireless and Mobile Networks. New York: Association f

۲۹ Li, Zhenchuan, Huang Mian, and Jiang Changjun. ۲۰۲۱. 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 ۱۷٥: ۱۱٤۷٥۰

۳۰ Chen, Ligong, Lijun Yang, Zhaohui Zhang, and Meng Ying. ۲۰۱۹. A Method for Online Transaction Fraud Detection Based on Individual Behavior. In ACM TURC '۱۹: Proceedings of the ACM Turing Celebration Conference—China. New York: ACM, pp. ۱–۸

۳۱ Wang, Yuan, and Liming Wang. ۲۰۱۹. Bot-like Behavior Detection in Online Banking. In ICBDC ۲۰۱۹: Proceedings of the ۲۰۱۹ ٤th International Conference on Big Data and Computing. New York: Association for Computing Machinery, pp. ۱٤۰–٤٤

۳۲ Chen, Keqin, Yadav Amit, and Zhu Kun. ۲۰۲۰. Credit Fraud Detection Based on Hybrid Credit Scoring Model. Procedia Computer Science ۱٦۷: ۲–۸

۳۳ Misra, Sumit, Thakur Soumyadeep, Ghosh Manosij, and Saha Sanjoy Kumar. ۲۰۲۰. An Autoencoder Based Model for Detecting Fraudulent Credit Card Transaction. Procedia Computer Science ۱٦۷: ۲٥٤–٦۲

۳٤ Olowookere, Toluwase, Ayobami Adewale, and Olumide Sunday. ۲۰۲۰. A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach. Scientific African ۸: e۰۰٤٦٤

۳٥ Amarasinghe, Thushara, Achala Aponso, and Naomi Krishnarajah. ۲۰۱۸. Critical Analysis of Machine Learning Based Approaches for Fraud Detection in Financial Transactions. In Paper presented at the ICMLT '۱۸: Proceedings of the ۲۰۱۸ International Conference

۳٦ Dong, Feng, Haoyu Wang, Li Li, Yao Guo, Tegawendé F. Bissyandé, Tianming Liu, Guoai Xu, and Jacques Klein. ۲۰۱۸. FraudDroid: Automated Ad Fraud Detection for Android Apps. In ESEC/FSE ۲۰۱۸: Proceedings of the ۲۰۱۸ ۲٦th ACM Joint Meeting on European Softwa

۳۷ Sudharsan, K., D. Ambeth Kumar, R. Venkatesan, V. Sathyapreiya, and G. Saranya. ۲۰۱۹. Two Three Step Authentication in ATM Machine to Transfer Money and for Voting Application. Procedia Computer Science ۱٦٥: ۳۰۰–٦

۳۸ Wei, W., Li, J., Cao, L., Ou, Y., & Chen, J. (۲۰۱۳). Effective detection of sophisticated online banking fraud on extremely imbalanced data. World Wide Web, ۱٦(٤), ٤٤٩-٤٧٥. doi: ۱۰,۱۰۰۷ /s ۱۱۲۸۰-۰۱۲-۰۱۷۸-۰

۳۹ Lavion, D. (۲۰۱۸). Global Economic Crime and Fraud Survey ۲۰۱۸, ۳۰. Retrieved from Pulling fraud out of the shadows website: https://www.pwc.com/gx/en/forensics/global¬economic-crime-and-fraud-survey-۲۰۱۸.pdf

٤۰ FBI. (۲۰۱۸). Internet Crime Complaint Center, ۲۰۱۸۲۰۱۹(۲۰/۰۸/۲۰۱۹). Retrieved from https ://pdf.i c۳. gov /۲۰۱۸ _ I C۳ Report. pdf

٤۱ PwC. (۲۰۱٦). Global Economic Crime Survey ۲۰۱٦, ٥٦. Retrieved from Adjusting the Lens on Economic Crime website: https://www.pwc.com/gx/en/economic-crime-survey/pdf/GlobalEconomicCrimeSurvey۲۰۱٦.pdf

٤۲ Abdallah, A., Maarof, M. A., & Zainal, A. (۲۰۱٦). Fraud detection system: A survey. Journal of Network and Computer Applications, ٦۸, ۹۰-۱۱۳.

٤۳ Carminati, M., Caron, R., Maggi, F., Epifani, I., & Zanero, S. (۲۰۱٥). BankSealer: A decision support system for online banking fraud analysis and investigation. Computers & Security, ٥۳(C), ۱۷٥-۱۸٦. doi:۱۰,۱۰۱٦/j.cose.۲۰۱٥,۰٤,۰۰۲

٤٤ John, S. N., Kennedy, ۰. ۰., Kennedy, C. G., Anele, C., & Olajide, F. (۲۰۱٦). Rea/time Fraud Detection in the Banking Sector Using Data Mining Techniques/Algorithm. Paper presented at the ۲۰۱٦ International Conference on Computational Science and Computat

٤٥ Money Laundering Offences, Jul. ۲۰۲۱, [online] Available: https://www.cps.gov.U.K./legal-guidance/proceeds-crime-act-۲۰۰۲-part-۷-money-laundering-offences

٤٦ Rouse, M. (۲۰۱۹). What is fraud detection?. Retrieved ٥ January ۲۰۲۰, from https :// searchsecurity. techtarget. com/ definition/fraud-detection

٤۷ APF. (۲۰۱۸). Western Australia Police Force. Retrieved from https://www.police.wa.gov.au/

٤۸ Economic Crime, Jul. ۲۰۲۱, [online] Available: https://www.europol.europa.eu/crime-areas-and-trends/crime-areas/economic-crime

٤۹ LexisNexis Risk Solutions ۲۰۱۷ True Cost of Fraud, ۲۰۱۷, [online] Available: https://risk.lexisnexis.com/-/media/files/financial%۲۰services/research/۲۰۱۸-true-cost-of-fraud-overall-rep%۲۰pdf.pdf?la=en-us

٥۰ Pascual, A., Marchini, K. and Miller, S., ۲۰۱۷. ۲۰۱۷ Identity Fraud: Securing the Connected Life. [online] Javelin. Available at: <https:/ /www.javelinstrategy.com/coverage-area/۲۰۱۷-identity¬fraud-securing-connected-life> [Accessed ۳۰ July ۲۰۲۱].

٥۱ Sando, S., ۲۰۲۱. Consumer Preference Drives Shift in Authentication. [online] Javelin. Available at: <https ://www .j ave I instrate gy. com/ coverage-area/ consumer-preference-drives-shift-authenti cation> [Accessed ۳۰ July ۲۰۲۱].

٥٢ Brabazon, A., Cahill, J., Keenan, P., & Walsh, D. (٢٠١٠). Identifying online credit card fraud using Artificial Immune Systems.

٥٣ FICO. (٢٠١٠). Fico Application Fraud Manager. Retrieved from https ://www.fico.com/en/products/fico-application- fraud-manager

٥٤ Capterra. (٢٠١٩). Fraud Management Software. Retrieved from https://www.capterra.com.au/directory/١٠٠٥٨/financial-fraud-detection/software

٥٥ ACI.(٢٠١١). Proactive Risk Manager. Retrieved from https://www.aciworldwide.com/products/proactive-risk-manager

٥٦ Hafiz, K. T., Aghili, S., & Zavarsky, P. (٢٠١٦). The use of predictive analytics technology to detect credit card fraud in Canada.

٥٧ SAS. (٢٠٠٧). SAS Fraud Management. Retrieved from https://www.sas.com/en_ au/software/fraud-management.html

٥٨ FraudNet. (١٩٩٧). Enterprise Fraud Prevention. Retrieved from https://fraud.net

٥٩ PattemSpy. (٢٠١٥). PattemSpy For Banking - Fraud Management Software. Retrieved from https://www.pattemspy.tech/

٦٠ RiskNet. (١٩٩٨). Fraud Managed Services. Retrieved from https :/ /www.aicorporation.com/products-services/fraud-managed-services/

٦١ Dehaven, V. R. (٢٠١٤). Machine Leaming: Future Capabilities and their Implications.

٦٢ Herschel, G., Linden, A., & Kart, L. (٢٠١٥). Magic quadrant for advanced analytics platforms. Gartner Report G, ٢٧٠٦١٢.

٦٣ Lenat, D. (٢٠٠٦). Creativity vs. Common Sense. In L. Weiss (Ed.). CA, USA: USC ICT.

٦٤ Syed, N. A., Huan, S., Kah, L., & Sung, K. (١٩٩٩). Incremental learning with support vector machines. Citeseer.

٦٥ Moon, W., & Kim, S. ٢٠١٧. Adaptive Fraud Detection Framework for FinTech Based on Machine Leaming. Advanced Science Letters, ٢٣(١٠), ١٠١٦٧-١٠١٧١. doi: ١٠,١١٦٦/asl.٢٠١٧,١٠٤١٢ Omnisci.com. ٢٠٢١.

٦٦ Lee, J.-H., Park, M.-W., Eom, J.-H., & Chung, T.-M. (٢٠١١). Multi-level Intrusion Detection System and log management in Cloud Computing. Paper presented at the ١٣th International Conference on Advanced Communication Technology (ICACT٢٠١١), Seoul.

٦٧ Chen, C.-M., Guan, D. J., Huang, Y.-Z., & Ou, Y.-H. (٢٠١٢). Attack Sequence Detection in Cloud Using Hidden Markov Model. Paper presented at the ٢٠١٢ Seventh Asia Joint Conference on Information Security, Tokyo.

٦٨ C. C. Aggarwal, Outlier Analysis, Cham, Switzerland:Springer, ٢٠١٧

٦٩ S. Choi, C. Kim, Y.-S. Kang and S. Youm, "Human behavioral pattern analysis-based anomaly detection system in residential space", J. Supercomput., vol. ٧٧, no. ٨, pp. ٩٢٤٨-٩٢٦٥, Aug. ٢٠٢١

٧٠ H. Darvishi, D. Ciuonzo, E. R. Eide and P. S. Rossi, "Sensor-fault detection isolation and accommodation for digital twins via modular data-driven architecture", IEEE Sensors J., vol. ٢١, no. ٤, pp. ٤٨٢٧-٤٨٣٨, Feb. ٢٠٢١

٧١ Ilgun, K., Kemmerer, R. A., & Porras, P.A. (١٩٩٥). State transition analysis: a rule-

based intrusion detection approach. IEEE Transactions on Software Engineering, ۲۱ (۳), ۱۸۱-۱۹۹. doi: ۱۰,۱۱۰۹/۳۲,۳ ۷۲۱٤٦

۷۲ Chiu, C.-Y., Yeh, C.-T., & Lee, Y.-J. (۲۰۱۳). Frequent Pattern Based User Behavior Anomaly Detection for Cloud System. Paper presented at the ۲۰۱۳ Conference on Technologies and Applications of Artificial Intelligence (TAAI), Taipei, Taiwan.

۷۳ Harutyunyan, A. N., Poghosyan, A. V., Grigoryan, N. M., & Marvasti, M. A. (۲۰۱٤). Abnormality analysis of streamed log data. Paper presented at the ۲۰۱٤ IEEE Network Operations and Management Symposium (NOMS), Krakow.

۷٤ Chandola, V., Banerjee, A., & Kumar, V. (۲۰۰۹). Anomaly detection: A survey. ACM Computing Surveys (CSUR), ٤۱(۳), ۱-۵۸. doi: ۱۰,۱۱٤۵/۱۵٤۱۸۸۰,۱۵٤۱۸۸۲

۷۵ Kovach, S., & Ruggiero, W. V. (۲۰۱۱). Online bankingfraud detection based on local and global behavior. Paper presented at the Proc. of the Fifth International Conference on Digital Society, Guadeloupe, France.

۷٦ Duman, E., & Ozcelik, M. H. (۲۰۱۱). Detecting credit card fraud by genetic algorithm and scatter search. Expert Systems with Applications, ۳۸(۱۰), ۱۳۰۵۷-۱۳۰٦۳. doi:https://doi.org/۱۰,۱۰۱٦/j.eswa.۲۰۱۱,۰٤,۱۱۰

۷۷ Herland, M., Khoshgoftaar, T., & Bauder, R. (۲۰۱۸). Big Data fraud detection using multiple medicare data sources. Journal of Big Data, ۵(۱), ۱-۲۱. doi:۱۰,۱۱۸٦/s٤۰۵۳۷-۰۱۸- ۰۱۳۸-۳

۷۸ C. Chen, C. Liang, J. Lin, L. Wang, Z. Liu, X. Yang, et al., "InfDetect: A large scale graph-based fraud detection system for E-commerce insurance", Proc. IEEE Int. Conf. Big Data (Big Data), pp. ۱۷٦۵-۱۷۷۳, Dec. ۲۰۱۹

۷۹ M. Araujo, M. Almeida, J. Ferreira, L. Silva and P. Bizarro, "BreachRadar: Automatic detection of points-of-compromise", arXiv:۲۰۰۹,۱۱۷۵۱, ۲۰۲۰

۸۰ Ayyadevara, V., ۲۰۱۸. Pro Machine Learning Algorithms: A Hands-On Approach to Implementing Algorithms in Python and R. Apress.

۸۱ Graupe, D., ۲۰۱٦. Deep Learning Neural Networks. Singapore: World Scientific Publishing Company.

---

[i] Fraud detection
[ii] Ciobanu
[iii] Nilson Report
[iv] Carminati
[v] General data protection legislation
[vi] Stalla-Bourdillon
[vii] Horak
[viii] Data Privacy Impact Assessment
[ix] Öğrek
[x] Multi-layered artificial neural networks

[xi] Li
[xii] Chen
[xiii] Hypersphere
[xiv] Wang and Wang
[xv] Shannon entropy
[xvi] Misra
[xvii] Olowookere
[xviii] Amarasinghe
[xix] Dong
[xx] Sudharsan
[xxi] Bolton & Hand
[xxii] Phua
[xxiii] Crown Prosecution Service
[xxiv] brick-and-mortar financial organizations
[xxv] Online banking systems
[xxvi] Anomaly detection
[xxvii] Machine learning
[xxviii] Rule-based systems
[xxix] Knowledge-Based Systems
[xxx] Brittleness
[xxxi] Incremental learning
[xxxii] Lenat
[xxxiii] Villalobos
[xxxiv] Intrusion Detection Systems
[xxxv] Backdoor
[xxxvi] Hidden Markov Model
[xxxvii] outlier detection
[xxxviii] novelty
[xxxix] noise
[xl] deviation
[xli] Chiu
[xlii] Brabazon
[xliii] Harutyunyan
[xliv] Kovach and Ruggiero
[xlv] Duman and Ozcelik
[xlvi] Kou
[xlvii] Berland, Khoshgoftaar, and Bauder
[xlviii] Medicare
[xlix] Bai
[l] Wei
[li] Araujo
[lii] Points-of-Compromise
[liii] Skimming
[liv] Logistic Regression
[lv] Decision Tree
[lvi] Random Forest
[lvii] Neural Networks