ششمین همایش ملی تجارت و اقتصاد الکترونیکی

همایش امنیت و اعتماد

نقش و کاربرد هوش عملیاتی و داده کاوی در کشف تقلب برخط

هادی بنائی، کارشناسی ارشد علوم کامپیوتر، شرکت داده‌کاوان هوشمند توسن

حسام خوش‌نیت، کارشناسی ارشد مهندسی فناوری اطلاعات، شرکت داده‌کاوان هوشمند توسن[1]

چکیده

با گسترش روز افزون استفاده از سامانه‌های مدرن بانکی و افزایش تعداد تراکنش‌های بانکی، سوء استفاده‌های مالی و تقلب در این تراکنش‌ها بیش از پیش نمود پیدا کرده است. این سوء استفاده‌ها علاوه بر از دست دادن منابع مالی هنگفت، منجر به کاهش اعتماد مشتریان به استفاده از سامانه‌های مدرن بانکی و در نتیجه کاهش اثربخشی این سامانه‌ها در مدیریت بهینه‌ی سرمایه و تراکنش‌های مالی می‌شود. هر چند جلوگیری از تقلب بهترین راه کاهش تقلب‌های بانکی است، اما افراد سودجو از راه‌هایی به اهداف خود دست پیدا می‌کنند. بنابراین روش‌هایی مورد نیاز است تا تراکنش‌های مشکوک به صورت برخط شناسایی و از انجام آنها ممانعت به عمل آید. در سال‌های اخیر سازوکارهای هوش عملیاتی و روش‌های داده‌کاوی توانسته‌اند با موفقیت در جلوگیری از پول‌شویی، تشخیص تقلب کارت‌های سپرده و اعتباری به کار گرفته شود. در این مقاله سامانه‌ای برای شناسایی برخط تقلب در تراکنش‌های کارت‌های سپرده‌ی بانکی معرفی شده است. این سامانه که بر بستر هوش عملیاتی و با به کارگیری روش‌های هوش مصنوعی، داده‌کاوی، الگوریتم‌های یادگیری ماشین و تحلیل گراف پیاده‌سازی شده است، قادر است به صورت برخط و بدون ایجاد تاخیر در خدمات‌رسانی تراکنش بانکی، میزان مشکوک بودن تراکنش به تقلب و سوء استفاده را تشخیص داده و آن را به سوئیچ بانک گزارش کند.

کلمات کلیدی: هوش عملیاتی، داده کاوی، کشف تقلب برخط، سوء استفاده بانکی، هوش مصنوعی.

۱. مقدمه

امروزه بهره‌گیری از فنّاوری‌های نوین در مدیریت تراکنش‌های بانکی رشد چشم‌گیری داشته است. بانک‌ها و موسسات مالی و اعتباری برای خدمات‌رسانی موثر، ناگزیر از مهاجرت از بانکداری سنتی به بانکداری مدرن و برخط شده‌اند. هر چند استفاده از این سامانه‌ها باعث مدیریت بهتر فرایندهای مالی و افزایش کارایی و سرعت خدمات‌رسانی به مشتریان این موسسات شده، اما مشکلات و مخاطراتی نیز به همراه داشته است. تقلب و سوء استفاده‌های مالی یکی از مشکلاتی است که این سازمان‌ها در پی پیشگیری از آنها و کاهش اثرات آنها بوده‌اند. تقلب[2] به معنای به دست‌آوردن مال یا کالا یا خدمات از راه‌های غیر اخلاقی و غیر غیر قانونی است که امروزه در سراسر دنیا رو به گسترش است. سوء استفاده‌هایی که در تراکنش‌های مالی رخ می‌دهد، علاوه بر

[1] نویسنده عهده‌دار مکاتبات: تلفن: ۵۰۲۶-۸۲۱۹۸۰۰-۰۲۱ آدرس: تهران-خیابان دستگردی- پلاک ۳۴۴- شرکت داده‌کاوان هوشمند توسن رایانامه: khoshniat@tosanidm.com

[2] Fraud

آنکه باعث از دست دادن منابع مالی هنگفتی می‌شود، منجر به کاهش اعتماد مشتریان به استفاده از سامانه‌های مدرن بانکی و در نتیجه کاهش اثربخشی این سامانه‌ها در مدیریت بهینه‌ی تراکنش‌های مالی می‌شود.

در سال‌های اخیر کشور ایران رشد چشمگیری در تعداد تراکنش‌های بانکی داشته است. شکل ۱ مقایسه‌ای از تعداد تراکنش‌های کارت‌های بانکی بر روی دستگاه‌های خودپرداز و پایانه‌های فروش را از سال ۱۳۸۷ تا ۱۳۸۹ نشان می‌دهد. همان‌طور که از شکل پیداست تعداد این تراکنش‌ها در سال ۱۳۸۹ نسبت به سال ۱۳۸۸ رشدی برابر ۶۳/۵۵ درصد داشته است.

شکل ۱. مقایسه‌ی تعداد تراکنش‌های بر روی پایانه‌های بانکی کشور[۱]

با در نظر گرفتن افزایش بهره‌گیری از سامانه‌های مدرن بانکی، احتمال رخداد تقلب و سوء استفاده‌های مالی نیز بیشتر خواهد بود. بنابراین روش‌هایی مورد نیاز است تا تراکنش‌های مشکوک به صورت برخط شناسایی شده و از انجام آنها ممانعت به عمل آید. در سال‌های اخیر سازوکارهای هوش عملیاتی[۱] و روش‌های داده‌کاوی[۲] توانسته‌اند با موفقیت در جلوگیری از پول‌شویی، تشخیص تقلب کارت‌های سپرده و اعتباری به کار گرفته شود. با توجه به بالا بودن تعداد این تراکنش‌ها امکان نظارت بر خط بر آنها توسط ناظرین انسانی وجود ندارد و بهره‌گیری از روش‌های هوشمند برای نظارت بر این تراکنش‌ها اجتناب‌ناپذیر است. امروزه روش‌های داده‌کاوی به عنوان بهترین راهکار برای شناسایی خودکار تقلب در حوزه‌های مختلف شناخته شده‌اند. داده‌کاوی به عنوان فرایند کشف و استخراج الگوهای پنهان از حجم بالایی از داده‌ها تعریف می‌شود[۳، ۶ و ۱۶]. در سامانه‌های بسیاری از روش‌های داده‌کاوی برای شناسایی و کشف تقلب و سوء استفاده مالی استفاده شده است[۱۳].

از سوی دیگر صنعت بانکداری در ایران هم از منظر به کارگیری فناوری‌های پیشرفته در مدیریت تراکنش و هم از دیدگاه ساختار تعامل بین بانکی تمایز چشمگیری با سایر کشورها دارد. فارغ از قیمت بالای سامانه‌های تولید شده در کشورهای دیگر، استفاده از این سامانه‌ها به دلیل ساختار منحصربه‌فرد شبکه‌ی بانکی کشور امکان‌پذیر نیست. فناوری استفاده شده در بسیاری از بانک‌های کشور بسیار قدیمی و غیر متمرکز است و تعدادی از بانک‌ها فاقد سامانه‌های مدرن بانکداری متمرکز هستند که این امر آسیب‌پذیری شبکه‌ی بانک در مقابل تقلب را بالاتر می‌برد.

---

[۱] Operational Intelligence

[۲] Data Mining

از طرف دیگر وجود شبکه‌ی شتاب به عنوان شبکه‌ی تبادل اطلاعات بین بانکی تفاوت عمده‌ای بین شبکه‌ی بانکی کشور با سایر کشورها ایجاد کرده که امکان استفاده از سامانه‌های هوشمند تشخیص تقلب تولید سایر کشورها را تقریباً غیر ممکن ساخته است. به همین دلیل نیاز به تولید سامانه‌ی هوشمند تشخیص تقلب برخط تراکنش‌های بانکی و سازگار با ساختار شبکه‌ی بانکی کشور به شدت احساس می‌شود.

در این مقاله یک سامانه‌ی هوشمند برای شناسایی برخط تقلب در تراکنش‌های کارت‌های سپرده‌ی بانکی معرفی شده است. این سامانه که بر بستر هوش عملیاتی و با به کارگیری روش‌های هوش مصنوعی، داده‌کاوی، الگوریتم‌های یادگیری ماشین و تحلیل گراف پیاده‌سازی شده است، قادر است به صورت برخط و بدون ایجاد تاخیر در خدمات‌رسانی تراکنش بانکی، میزان مشکوک بودن تراکنش به تقلب و سوء استفاده را تشخیص داده و آن را به سوئیچ بانک یا دیگر سامانه‌هایی که از سامانه‌ی مذکور سرویس بگیرند، گزارش کند. در این سامانه مدل‌های مختلفی از رفتار کارت و پایانه ساخته شده و سوئیچ با دریافت هر تراکنش و به کمک این مدل‌های آموزش دیده شده، می‌تواند میزان مشکوک‌بودن رفتار کارت و پایانه را تشخیص داده و در صورت لزوم اجرای تراکنش را متوقف کند.

ساختار مقاله بدین صورت است: در بخش ۲ مروری بر پژوهش‌های مرتبط با حوزه‌ی تشخیص تقلب ارائه می‌شود. در بخش ۳ معماری سامانه‌ی پیشنهادی تشریح شده است. در انتها نیز جمع‌بندی و نتیجه‌گیری مقاله در بخش ۴، آورده شده است.

۲. پیشینه‌ی پژوهش

در سال ۱۹۹۴ گاش و رِیلی[۱] روشی مبتنی بر شبکه‌های عصبی مصنوعی[۲] برای شناسایی تقلب بر روی کارت‌های اعتباری ارائه کردند. این شبکه پس از آموزش با حجم عظیمی از تراکنش‌های نمونه‌ی برچسب زده شده، قادر بود تراکنش‌های مشکوک به تقلب را شناسایی کند[۱۲]. در پژوهشی دیگر از شبکه‌ی بیز[۳]، درخت تصمیم C4.5 و شبکه‌ی عصبی پس‌انتشار[۴] به عنوان دسته‌بندهای[۵] پایه و اَبَردسته‌بند[۶] برای انتخاب دسته‌بند مناسب بر اساس چولگی داده‌ها استفاده شده است[۵]. مدل مخفی مارکوف[۷] نیز برای مدل‌سازی ترتیب فعالیت‌ها در پردازش تراکنش‌های کارت اعتباری به کار برده شده است. بر اساس نتایج ارائه شده، مدل مخفی مارکوف که با داده‌های رفتار طبیعی صاحب کارت آموزش داده شده، می‌تواند تا حد زیادی تراکنش مشکوک به تقلب را تشخیص دهد[۱۴]. رویکرد نظریه‌ی بازی‌ها نیز در فرایند تشخیص تقلب و سوء استفاده از کارت‌های اعتباری به کار گرفته شده است. وَستا[۸] در روش خود تعامل میان حمله‌کننده و سامانه‌ی تشخیص تقلب را به عنوان یک بازی چندمرحله‌ای بین دو بازیکن مدل کرده است که هر یک سعی در بیشینه‌سازی سود خود از بازی دارد[۱۵]. در سامانه‌ی

---

۱ Ghosh and Reilly
۲ Artificial Neural Network(ANN)
۳ Naive Bayesian
۴ Back-Propagation
۵ Classifier
۶ Meta-Classifier
۷ Hidden Markov Model(HMM)
۸ Vasta

دیگری برای تشخیص خودکار تقلب بر روی کارت اعتباری از شبکه‌های خودسازمانده[۱] استفاده شده است. در این سامانه این الگوریتم برای مدل‌سازی رفتار صاحب کارت و میزان تخطی تراکنش وی از رفتار مدل‌شده و در نتیجه یافتن تراکنش‌های مشکوک استفاده شده است[۸]. در مدل دیگری، از این شبکه برای شناسایی بی‌درنگ تراکنش مشکوک استفاده شده است[۱۳]. در پژوهش دیگری کارایی چندین روش داده کاوی شامل نزدیکترین K همسایه[۲]، تخمین تابع لاجیستیک[۳]، تحلیل جدا کننده[۴]، شبکه‌ی بیز، شبکه‌ی عصبی و درخت تصمیم[۵] در پیش‌بینی پرداخت‌های متداول مشتری استفاده شده است[۱۷]. مروری بر پژوهش‌های مرتبط با استفاده از روش‌های داده‌کاوی در شناسایی تقلب‌های مالی از سال ۱۹۹۷ تا ۲۰۰۸ در مرجع [۱۷] آورده شده است. یافته‌های پژوهشگران نشان می‌دهد که روش‌های داده‌کاوی در شناسایی تقلب در صنعت بیمه نیز با موفقیت به کار گرفته شده و بهره‌گیری از این الگوریتم‌ها در شناسایی سوءاستفاده‌های مالی و اعتباری در سال‌های اخیر رشد فزاینده‌ای داشته است.

در ادامه معماری سامانه‌ی پیشنهادی برای شناسایی بر خط تراکنش‌های مشکوک بانکی تشریح خواهد شد.

۳. معماری سامانه‌ی پیشنهادی

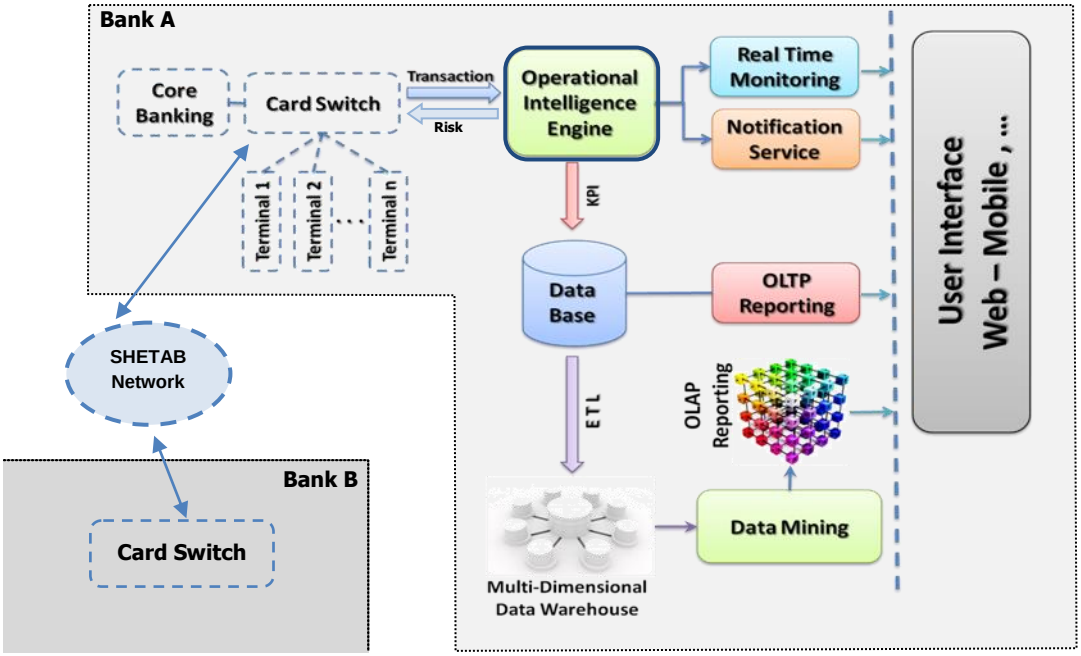
هدف از ارائه‌ی این سامانه، استفاده از زیرساخت هوش‌عملیاتی در صنعت بانکداری کشور با محور کشف تقلب بر خط در تراکنش‌های کارت‌های سپرده و اعتباری است. هر چند تمرکز طراحی این سامانه بر روی تعیین میزان مشکوک‌بودن تراکنش کارت است، اما معماری پیشنهادی قابلیت نصب بر روی سامانه‌های بانکداری مدرن و سامانه‌ی متمرکز بانکداری را نیز داراست. معماری خارجی و نحوه‌ی تعامل این سامانه با سایر نهادهای درگیر در شکل ۲ نشان داده شده است. در این معماری سوئیچ کارت(و یا هر سامانه‌ی فراخواننده‌ی دیگر) تراکنش دریافتی از پایانه‌ها یا شبکه‌ی شتاب را از طریق وب‌سرویس و بر روی پروتکل امن SSL به هسته‌ی سامانه‌ی تشخیص تقلب ارسال می‌کند. این هسته با فراخوانی موتور استنتاج سامانه، میزان مشکوک بودن تراکنش به تقلب را مشخص کرده و نتایج را علاوه بر ارسال به سامانه‌ی فراخواننده، در پایگاه‌داده نیز ذخیره می‌کند. هم‌چنین این هسته با اتصال به سایر سامانه‌های نظارتی، امکان نظارت و اطلاع‌رسانی بی‌درنگ را فراهم می‌آورد. به‌علاوه با ایجاد انبارداده[۶] بر روی پایگاه داده‌ای -که میزان مشکوک بودن هر تراکنش را در خود نگه می‌دارد- می‌توان گزارش‌های تحلیلی OLAP از پایانه‌ها، کارت‌ها و در کل نهادهای مشکوک را استخراج کرد.

هسته‌ی سامانه‌ی معرفی شده در این سامانه قادر است به صورت بی‌درنگ تمامی تراکنش‌های ارسال شده از سمت سوئیچ را پردازش کرده و بدون ایجاد تاخیر در فرایند پردازشی سوئیچ، ریسک تراکنش را محاسبه و گزارش کند. از آنجا که دریافت تراکنش و ارسال مقادیر ریسک در این هسته بر بستر وب‌سرویس انجام می‌شود، قابلیت اتصال به هر سامانه‌ی دیگری را

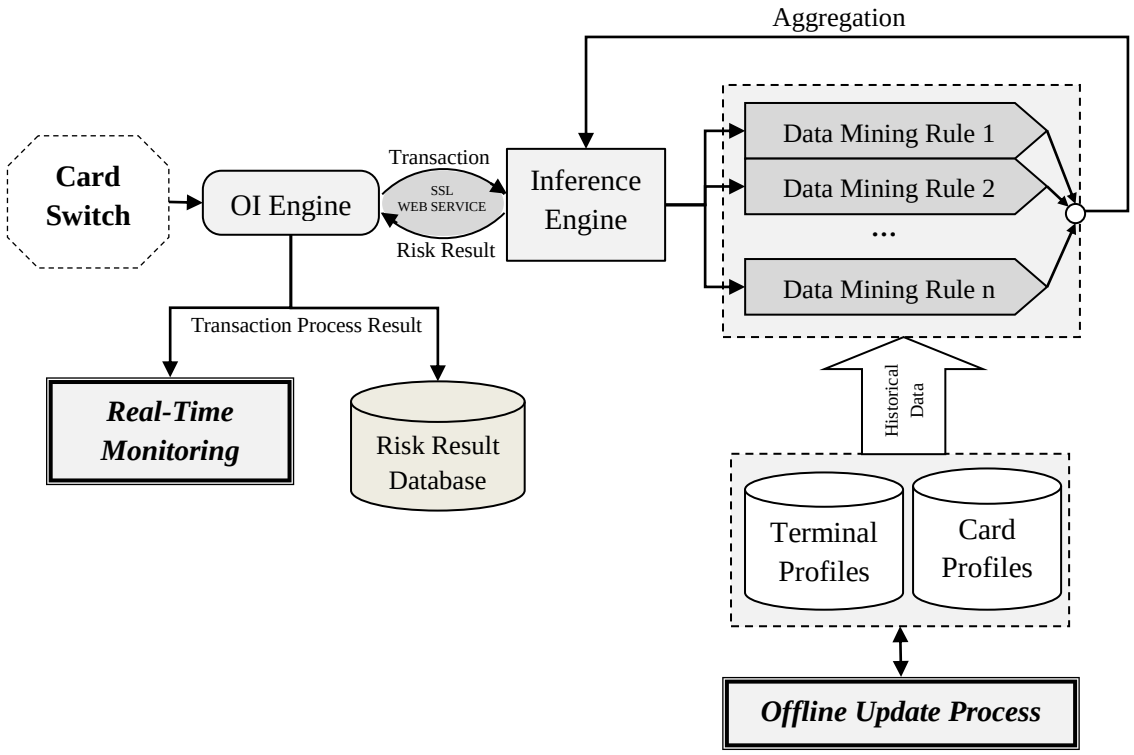
---

۱ Self-Organizing Map
۲ K-Nearest Neighbor(KNN)
۳ Logistic Regression
۴ Discriminant Analysis
۵ Decision Tree
۶ Data Warehouse

داراست. همچنین انتقال اطلاعات به صورت رمز شده انجام می‌شود تا شناسایی تقلب در بستری امـن انجـام پـذیرد. معمـاری درونی این سامانه در شکل ۳ نشان داده شده است.



شکل ۲. معماری خارجی سامانه و تعامل آن با سایر سامانه‌های درگیر



شکل ۳. معماری داخلی سامانه‌ی کشف تقلب برخط

فرایند کشف تقلب با دریافت پیام حاوی تراکنش از سوئیچ شروع می‌شود. این تراکنش به موتور استنتاج ارسال می‌شود. در موتور استنتاج بسته به نوع تراکنش و نوع ترمینال، قواعد داده‌کاوی مربوطه فراخوانی می‌شوند. این قواعد شامل قواعد مبتنی بر شبکه‌ی عصبی، قواعد آماری، قواعد مبتنی بر گراف و قواعد خوشه‌بندی هستند. هر یک از این قواعد نقش یک خبره در موتور استنتاج را بازی می‌کنند که میزان مشکوک بودن یک تراکنش را در حوزه‌ی تخصصی آن قاعده تشخیص می‌دهد. این قواعد برای تصمیم‌گیری از مدل‌ها یا پروفایل‌هایی استفاده می‌کند که به صورت دوره‌ای بر اساس داده‌های موجود از تراکنش‌های کارت و پایانه ساخته می‌شوند. این پروفایل‌ها مدل رفتار کارت یا پایانه را در خود نگه‌داری می‌کنند که بر اساس الگوریتم‌هایی به صورت شبانه یا هفتگی به‌روز می‌شوند. تمامی قواعد موتور استنتاج به صورت موازی فراخوانی شده و پس از اتمام پردازش قواعد، نتایج علاوه بر گزارش به سوئیچ کارت، در پایگاه‌داده‌ای برای گزارش‌گیری‌های آتی ذخیره می‌شود.

یکی از قواعد مبتنی بر گراف، تراکنش حاصل از کارت کپی‌شده را تشخیص می‌دهد. برای تشخیص کارت کپی شده نیاز است موقعیت فیزیکی پایانه‌های تمامی بانک‌ها را در اختیار داشت که عملاً چنین امکانی وجود ندارد. در این سامانه موقعیت پایانه‌ها و نزدیکی آنها نسبت به یکدیگر به کمک یک گراف شبیه‌سازی شده است. منظور از نزدیکی در این مدل الزاماً نزدیکی فیزیکی نیست. از دید سامانه پایانه‌هایی که دارای تراکنش‌های زیاد از افراد یکسان هستند، نزدیک محسوب می‌شوند. در نتیجه در برخی مواقع ترمینال‌هایی که از نظر فیزیکی از هم دورند، در مدل ما به هم نزدیک هستند. به عنوان مثال می‌توان دو دستگاه خودپرداز در دو ایستگاه مترو پر رفت آمد شهر تهران را در نظر گرفت که با وجود اینکه از نظر فیزیکی از هم فاصله دارند، اما افراد زیادی هستند که روی این دو ترمینال تراکنش مالی پشت سرهم انجام می‌دهند. با بررسی رفتار افراد مختلف

در یک بازه زمانی خاص، می‌توان برای هر دو ترمینال دلخواه فاصله (نزدیکی یا دوری) تعریف کرد. با داشتن این فواصل یک گراف حاصل می‌شود که به خوبی این فاصله‌ها را مدل می‌کند. این قاعده با در نظرگرفتن آخرین پایانه‌هایی که فرد بر روی آنها تراکنش داشته و زمان ترانکش می‌تواند میزان مشکوک بودن تراکنش را محاسبه کرده و عدد ریسک را گزارش کند.

قواعد آماری که دسته‌ی دیگری از قواعد سامانه هستند، به تغییر رفتار کارت یا پایانه حساس بوده و به محض مشاهده‌ی رفتار غیرمتعارف، ریسک بالا گزارش می‌کند. این قواعد نیز از سوابق کارت یا پایانه که در پروفایل‌های آنها نگهداری می‌شود برای تصمیم‌گیری و تولید ریسک استفاده می‌کنند.

قاعده‌ی شبکه‌ی عصبی پایانه‌ی بعدی که کاربر بر روی آن تراکنش خواهد داشت را پیش‌بینی می‌کند. این قاعده در ادامه بیشتر تشریح خواهد شد.

۳,۱. شبکه‌ی عصبی مصنوعی

در طیّ دهۀ اخیر توجه گسترده‌ای به تحلیلهای غیرالگوریتمی و پردازش داده‌ها به کمک فنّاوریهای نرم‌رایانش[۱] مانند سیستمهای خبره، شبکه‌های عصبی مصنوعی، منطق فازی و الگوریتم ژنتیک معطوف شده است. این الگوهای پردازشی بهترین راه حل را با روشهایی همانند آنچه یک انسان معمولی انجام می‌دهد، جستجو می‌کنند. هر گاه پیچیدگی الگوریتمی برا حل مساله‌ای، بسیار زیاد و یا به سختی قابل تعیین باشد این رهیافتها مورد توجه قرار می‌گیرند[۱۱]. برخی از پیشرفت‌هایی که در زمینه‌ی توسعه‌ی سامانه‌های هوشمند صورت گرفته از شبکه‌ی عصبی زنده (زیستی) الهام گرفته شده است. محققانِ زمینه‌های مختلف علمی برای حل مسائل متنوعی در بازشناسی الگو، پیش‌بینی، بهینه‌سازی، حافظه‌ی انجمنی و کنترل از شبکه‌های عصبی مصنوعی استفاده می‌کنند. گرچه راه‌حل‌های مرسوم نیز برای حل این مسائل پیشنهاد شده است اما این راه-حل‌ها فقط در حیطۀ خاص خود استفاده می‌شوند و هیچکدام انعطاف‌پذیری لازم را برای عمل درخارج از محدوده خود را ندارند در حالی که شبکه‌های عصبی مصنوعی راه‌حلهای جدیدی را ارائه می‌کند و در بسیاری از مسائل می‌توان از آنها بهره جست[۷]. رایانه‌ها محاسبات عددی را سریعتر و مطمئن‌تر از انسان انجام می‌دهند در حالی که حیطه‌ی وسیعی از پردازش اطلاعات باقی می‌ماند که انسان از ماشین به آسانی پیش می‌افتد. بنابراین نکتۀ مهمی درباره نوع محاسباتی که توسط مغز انجام می‌گیرد وجود دارد که با دانستن اینکه اجزای تشکیل‌دهندۀ مغز انسان از لحاظ سرعت و قابلیت اطمینان در مقایسه با قطعات الکترونیکی جدید در مرتبۀ پایین‌تری قرار دارند، قوّت می‌گیرد[۴]. با الهام از شبکه‌ی عصبی زیستی، شبکه‌ی عصبی مصنوعی (ANN)[۲] یک سیستم محاسباتی موازی شامل تعداد زیادی پردازشگرهای ساده یا گره‌ها با پیوندهای بین آنهاست[۷]. به تعبیر دیگر شبکه‌ی عصبی مصنوعی شبکه‌ای از عناصر (معمولاً قابل تطبیق) ساده و سلسله مراتبی است که به طور موازی به هم متصل شده‌اند و همانند سیستم عصبی زنده با دنیای واقعی ارتباط برقرار می‌کنند[۱۸]. مدل‌های شبکه‌ی عصبی مصنوعی سعی در تقلید از قوانین سازمان یافته‌ای دارند که در مغز انسان به کار رفته است[۷].

---

[۱] Soft-computing

[۲] Artificial Neural Network (ANN)

شبکه‌های عصبی مصنوعی به جای انجام روندی از دستورالعمل‌ها، فرضیات بسیار زیادی را به طور همزمان با استفاده از شبکه‌های موازی کاوش می‌کنند. عناصر محاسباتی یا گره‌ها که در شبکه‌های عصبی مورد استفاده قرار می‌گیرند غیرخطی هستند[۹]. قدرت محاسباتی در یک شبکه‌ی عصبی از پیچیدگی هر عنصر پردازشی، مانند آنچه در رایانه‌های سنتی وجود دارد، ناشی نمی‌شود بلکه به انبوهی و پیچیدگی پیوندها مربوط می‌شود. از طرفی شبکه‌های عصبی مصنوعی دارای قدرت تحمل خطای بیشتری نسبت به رایانه‌های معمولی هستند بدین معنا که با ایجاد مشکل برای تعدادی ازگره‌ها یا پیوندها، عملکرد شبکه دچار مشکل نمی‌شود[۹]. جذّابیت شبکه‌های عصبی مصنوعی از توانایی چشمگیر آنها در پردازش اطلاعات همانند سیستمهای زنده ناشی می‌شود که این تواناییها شامل غیرخطی بودن، پردازش موازی، قوام[۱]، تحمل خطا و نویز، یادگیری، توانمندی کار با داده‌های مبهم و قابلیت تعمیم (در مواجهه با الگویی جدید) است[۲]. شبکه‌های عصبی برای حل بسیاری از مسائل مانند پردازش سیگنال، بازشناسی الگو[۲]، بازشناسی گفتار[۳]، رباتیک، شناسایی سیستم، خوشه‌بندی[۴]، تحلیل سری‌های زمانی، کاربردهای مالی، پیش‌بینی و کنترل بسیار کارا بوده‌اند[۲].

در ادامه نحوه‌ی کاربرد شبکه‌ی عصبی مصنوعی به عنوان ابزاری برای پیش‌بینی پایانه‌ی بعدی تراکنش و تشخیص مشکوک‌بودن تراکنش ارائه می‌شود.

۳,۲. پیش‌بینی نوع پایانه به کمک شبکه‌ی عصبی

یکی از پارامترهایی که می‌تواند به کشف تقلب برخط کمک کند، پیش‌بینی رفتار بعدی یک کارت بر اساس الگوی رفتاری آن در تراکنش‌های قبلی است. در واقع با نگاه به بعضی از مولفه‌های تراکنش‌هایی که قبلا یک کارت بر روی پایانه‌های مختلف انجام داده است، می‌توان پیش‌بینی کرد که تراکنش جاری این کارت بر روی چه نوع پایانه‌ای انجام خواهد پذیرفت.

برای انجام این پیش‌بینی از یک مدل شبکه‌ی عصبی به نام پرسپترون چند لایه برای پیش‌بینی نوع پایانه‌ی استفاده شده است. پرسپترون چند لایه[۵] (MLP)، یک مدل از شبکه‌های عصبی پیش‌رو[۶] است که مجموعه‌ای از داده‌های ورودی را بر روی مجموعه ای از داده‌های خروجی نگاشت می‌نماید. MLP شامل چندین لایه از گره‌ها در یک گراف جهت‌دار است، به‌طوری که هر گره در هر لایه به تمام گره‌های لایه‌ی بعدی متصل است. به جز گره‌های لایه‌ی ورودی، هر گره یک نرون[۷] یا عنصر پردازش با یک تابع فعالیت غیر خطی است.

شبکه‌ی عصبی MLP دارای سه لایه از نرون‌ها است. لایه‌ی ورودی که اطلاعات خام را از ورودی دریافت می‌کند؛ لایه‌های پنهان که عملکرد آن‌ها به وسیله ورودی‌ها و وزن‌های ارتباطی بین آن‌ها و لایه‌های پنهان دیگر تعیین می‌شود و در نهایت

---

۱ robustness
۲ Pattern Recognition
۳ Speech Recognition
۴ clustering
۵ Multilayer Perceptron
۶ Feedforward
۷ Neuron

لایه‌ی خروجی که عملکرد واحد خروجی را بسته به فعالیت لایه‌های پنهان و وزن‌های ارتباطی بین لایه‌های پنهان و لایه‌ی خروجی در بر دارد. شبکه‌ی عصبی MLP از یک تکنیک یادگیری با نظارت به‌نام پس‌انتشار[1] برای آموزش شبکه بهره می‌گیرد. در تکنیک پس‌انتشار سعی بر آن است که مربع خطای بین خروجی های شبکه و تابع هدف کمینه شود.

همان‌گونه که شرح داده شد، هدف پیش‌بینی نوع پایانه‌ای است که یک کارت مشخص در تراکنش جاری از آن استفاده می‌کند. در زیر تعدای از مولفه‌های تراکنش‌های قبلی هر کارت که برای این پیش‌بینی مورد استفاده قرار گرفته‌اند، آورده می‌شوند.

- نوع پایانه‌هایی که در سه تراکنش آخر منتهی به تراکنش جاری توسط این کارت مورد استفاده قرار گرفته‌اند.
- تعداد تراکنش‌هایی انجام شده بر روی هر نوع از پایانه‌ها در طول دوره‌ی استفاده از کارت.
- درصد تعداد تراکنش‌های انجام شده بر روی هر نوع پایانه در ساعات مختلف شبانه روز.
- درصد تعداد تراکنش‌های انجام شده بر روی هر نوع پایانه در روزهای مختلف ماه.
- درصد تعداد تراکنش‌های انجام شده بر روی هر نوع پایانه در روزهای مختلف هفته.
- ساعت انجام تراکنش جاری.
- روز انجام تراکنش جاری در ماه.
- روز انجام تراکنش جاری در هفته.

این مولفه‌ها به همراه نوع پایانه‌ی تراکنش‌های شبیه‌سازی شده‌ای که در این مقاله مورد استفاده قرار گرفته‌اند، به عنوان ورودی به یک شبکه‌ی عصبی MLP با یک لایه‌ی پنهان داده شده است. تعداد نرون‌ها در لایه‌ی ورودی به اندازه‌ی طول بردار ورودی است. تعداد نرون‌های لایه‌ی پنهان برابر ۹ و تعداد نرون‌های لایه‌ی خروجی برابر با تعداد انواع پایانه‌های موجود در نظر گرفته شده است. از مجموعه‌ی داده‌های موجود، بخشی برای یادگیری شبکه استفاده شده‌اند و بخشی برای تست که خطای کمتر از ۱۰ درصد را در بر داشته است. بنابر این شبکه‌ی مورد نظر به عنوان یک مدل مناسب داده‌کاوی برای پیش‌بینی نوع پایانه‌ی تراکنش دریافتی در مدل کشف تقلب جای گرفته است.

در هنگام فراخوانی قاعده‌ی پیش‌بینی نوع پایانه، به ازای هر تراکنش دریافتی از سیستم ، یک بردار ورودی ساخته شده و به شبکه‌ی عصبی مورد نظر داده می‌شود. عددهای ظاهر شده در لایه‌ی خروجی شبکه، احتمال رخداد تراکنش بر روی نوع پایانه‌ی متناظر با هر نرون خروجی را نشان می‌دهد.

۴. جمع‌بندی

در این مقاله سامانه‌ی کشف تقلب در تراکنش‌های بانکی به صورت برخط معرفی شد که در آن از قواعد آماری و داده‌کاوی برای تشخیص میزان مشکوک بودن تراکنش به تقلب بهره گرفته شده است. این سامانه که بر بستر هوش عملیاتی و با به کارگیری روش‌های هوش مصنوعی، داده‌کاوی، الگوریتم‌های یادگیری ماشین و تحلیل گراف پیاده‌سازی شده است، قادر است

[1] Backpropagation

به صورت برخط و بدون ایجاد تاخیر در خدمات‌رسانی تراکنش بانکی، میزان مشکوک بودن تراکنش به تقلب و سوء استفاده را تشخیص داده و آن را به سوئیچ بانک گزارش کند. ارزیابی عملکرد سامانه با داده‌های آزمایشی نشانگر قابلیت و کارایی سامانه برای به کارگیری در محیط عملیاتی واقعی و کشف تقلب در تراکنش‌های بانکی به صورت برخط است.

مراجع

[۱] بانک مرکزی ایران. ۱۳۹۰. "گزارش آمار و داده‌های عملکرد سامانه‌های پرداخت کارتی کشور". *پرتال بانک مرکزی ایران*. دسترسی در مهرماه ۱۳۹۰، http://www.cbi.ir/simplelist/2546.aspx

[2] Basheer, I. A.; Hajmeer, M. 2000. "Artificial neural networks: fundamentals, computing, design and application". *Journal of Microbiological Methods*, 43**,** 3-31.

[3] Bose, I.; Mahapatra, R.K. .2001 . "Business Data Mining – a Machine Learning Perspective". *Information System*, 39, 3, 211-225.

[4] Bressloff, P. C. ; Weir, D. J. . 1991. "Neural networks". *GEC Journal of Research*, 8**,** 151-169.

[5] Ghosh, S.; Reilly, D.L. 1994. "Credit Card Fraud Detection with Neural Network", *Proc. 27th Hawaii Int'l Conf. System Sciences: Information System: Decision Support and Knowledge –Based System*, 3, 621-630.

[6] Han, J.; Kamber, M. . 2006 "Data Mining, Concepts and Techniques", *Morgan Kaufmann publisher*, second edition.

[7] Jain, A. K. ; Mao, J. .1996. "Artificial neural networks :a tutorial". *Computer*, 29**,** 31-44.

[8] Kohonen, T. .1988. "An introduction to neural computing". *Neural networks*, 1**,** 3-16.

[9] Lippmann, R. P. . 1987 . "An introduction to computing with neural nets". *IEEE ASSP Magazine*, 4**,** 4-22.

[10] Ngai, E.W.T.; Hu, Y.; Wong, Y.H.; Chen, Y; Sun, X. . 2011 . "The Application of Data Mining Techniques in Financial Fraud Detection: A Classification Framework and Academic Review of Literature". *Decision Support Systems*, 50, 559-569.

[11] Piuri, V. ; Alippi, C. .1998. "Artificial neural networks". *Journal of system architecture*, 44**,** 565-567.

[12] Phua, C.; Alahakoon, D.; Lee, V. . 2004. "Minority Report in Fraud Detection: Classification of Skew Data", *ACM SIGKDD Exploration Newsletter*, 6, 1, 50-59.

[13] Quah, J.T.S; Sriganesh, M.; 2008. "Real-Time credit card fraud detection using computational intelligence" . *Expert System with Applications*, 35 (4), 1721-1732.

[14] Srivantava, A.; Kundu, A.; Sural, S.; 2008. "Credit Card Fraud Detection Using Hidden Markov Model". *IEEE Transaction on Dependable and Secure Computing*, 5, 1, 35-42.

[15] Vatsa, V.; Sural, S.; Majumadar, A.K.; 2005. "A Game Theoretic Approach to Credit Card Fraud Detection". *Proc. First Int'l Conf. Information Systems Security,* 263-276.

[16] Frawley, W.j.; Piatetsky-Shapiro, G.; Matheus, C.J; 1992. "Knowledge Discovery in databases: an Overview", *AI Magazine*,13, 3, 57-70.

[17] Yeh, I.; Lien, C.; 2008. "the comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients", *Expert Systems with Applications*, 36 (2), 2473-2480.

[18] Zaslavsky, V.; Strizhak, A.; 2006. "Credit Card Fraud Detection Using Self-Organizing Maps". *Information and Security*, 18, 48-63.