
بسم الله الرحمن الرحیم

دانشگاه پیام نور

دانشکده فنّی و مهندسی

گروه علمی مهندسی فنّاوری اطلاعات و ارتباطات

ارائه مدلی جهت کشف رفتارهای مشکوک در بانکداری الکترونیکی با استفاده از الگوریتم‌های درخت تصمیم‌گیری

پایان نامه

برای دریافت درجه کارشناسی ارشد

در رشته مدیریت فنّاوری اطّلاعات

استاد راهنما:

آقای دکتر نصراله مقدّم

استاد مشاور:

آقای مهندس داود وحدت

نگارش:

روح الله کوثری لنگری

تابستان ۱۳۸۹

تقدیم به

پدر و مادر عزیزم

تقدیر و تشکر

بدینوسیله مراتب امتنان خود را به تمام عزیزانی تقدیم می‌نمایم که مرا در انجام این پایان نامه یاری نموده و از محبت و پشتیبانی خود در تکمیل این پژوهش دریغ ننموده‌اند.

راهنمایی‌های روشنگرانه و بی‌بدیل اساتید محترم خصوصاً آقای دکتر مقدم و آقای مهندس وحدت و همچنین سعه صدر مهندسین و خبرگان اینترنتی بانک ملّت، سامان و پارسیان سزاوار عالی‌ترین مراتب قدردانی است.

از درگاه خداوند منان سلامتی و کامیابی این بزرگواران را آرزومندم.

چکیده تحقیق

دگرگونی‌های جهان به واسطه گسترش فنآوری اطلاعات و اینترنت، رقابتی دانش‌محور را در عرصه تجارت الکترونیکی به وجود آورده که منجر به افزایش توان رقابتی بین سازمان‌ها شده است. در این شرایط با افزایش نرخ مبادلات تجاری، گسترش تضمینی، همراه با سرعت و کیفیت بالا، منوط به تحقق سیستم پویای بانکداری الکترونیکی است تا با بکارگیری فنآوری‌های نوین، فرایند تجارت الکترونیکی را تسهیل بخشد. بانکداری اینترنتی به عنوان یک فرصت بالقوه که رکن اساسی و تعیین کننده بانکداری الکترونیکی محسوب می‌شود در فضای سایبر با موانع و تهدیدهای گوناگونی مواجه است که یکی از این چالش‌ها، عدم قطعیت کامل در تضمین امنیت تراکنش‌های مالی و همچنین وجود رفتارهای غیرمتعارف و مشکوک از سوی اخلال‌گران الکترونیکی در جهت سوءاستفاده‌های مالی است. تاکنون سیستم‌های مختلفی بواسطه روش‌های ماشینی هوشمند و تکنیک داده‌کاوی جهت شناسایی جرم در رفتار کاربران طراحی شده و در صنایع مختلفی همچون بیمه، پزشکی و بانکداری بکارگرفته شده است. هدف این پایان‌نامه تشخیص رفتارهای غیرمعمول کاربران در سامانه بانکداری اینترنتی است لذا تشخیص رفتار کاربران و دسته‌بندی الگوهای بوجود آمده، شرایطی را در جهت پیش‌بینی نفوذ غیرمجاز و تشخیص رفتارهای مشکوک و تقلب‌آمیز مهیا می‌سازد. از آنجایی که تشخیص رفتار کاربران در سامانه اینترنتی با عدم قطعیت همراه بوده و سوابق بجای مانده از تراکنش‌ها می‌تواند در جهت درک این حرکات راهگشا باشد و همچنین با توجه به اینکه در بین روش‌های ماشینی، تکنیک درخت تصمیم‌گیری متداول‌ترین ابزار برای دسته‌بندی و پیش‌بینی محسوب می‌شود، لذا در این پایان‌نامه ابتدا متغیرهای موثر و وزن هرکدام در تولید رفتار، تعیین شده و در ادامه ترکیبی از حالت‌های رفتاری مختلف به صورت مدلی از قوانین استنتاجی استخراج گردید تا قابلیت تشخیص رفتارهای مختلف فراهم گردد. نهایتا روند چهار الگوریتم Chaid، ex_Chaid، C4.5 و C5.0 جهت دسته‌بندی و کشف الگوهای موجود روی تراکنش‌های واقعی یک بانک خصوصی، مورد مقایسه و ارزیابی قرار گرفت.

کلمات کلیدی: بانکداری الکترونیکی، بانکداری اینترنتی، امنیت، رفتار مشکوک، درخت تصمیم‌گیری

فهرست مطالب

فهرست جداول

فهرست اشکال

فصل اول

کلّیات تحقیق

❋ شامل موضوعات ذیل :

❋ مقدمه

❋ بیان مساله تحقیق

❋ معرفی موضوع تحقیق، سابقه و ضرورت انجام آن

❋ اهداف تحقیق

❋ قلمرو تحقیق

❋ بیان فرضیه

❋ کاربردهای این تحقیق

❋ ساختار تحقیق

❋ محدودیت‌های تحقیق

❋ روش تحقیق

❋ تعریف واژگان

۱-۱- مقدّمه

ورود به قرن بیست و یکم و عصر اطّلاعات و دانایی، با چالش‌ها و نگرانی‌های بسیار جدّی همراه بوده است بطوریکه هیچ یک از برنامه‌های توسعه‌ای طراحی شده و فنّاوری‌های نوین قرن بیستم نتوانسته‌اند تاثیر قاطعی بر حل این مسائل و تبعات ناشی از بروز آن بر جای گذارند. امّا به نظر بسیاری از محقّقین و صاحب‌نظران، انقلاب فنّاوری اطّلاعات می‌تواند نقش اساسی در مواجهه با این چالش‌ها را داشته باشد[۲۴]. دوره‌ای از زمان که در آن بسر می‌بریم دوران الکترونیکی شدن همه چیز[۱] نامیده شده است که از جمله آنها می‌توان به تجارت الکترونیکی[۲]، بانکداری الکترونیکی[۳]، یادگیری الکترونیکی[۴]، تدارکات الکترونیکی[۵]، شهروندالکترونیکی[۶] دولت الکترونیکی[۷] و در مجموع زندگی الکترونیکی[۸] اشاره کرد. با رشد روزافزون معاملات تجاری در سطح جهان و ظهور پدیده تجارت الکترونیک و نیاز ساختار تجارت به حضور فعال و قدرتمند بانک‌ها جهت نقل و انتقال منابع مالی، بانکداری الکترونیک را به عنوان بخش تفکیک‌ناپذیر از تجارت الکترونیک مطرح می‌کند. پدید آمدن دو مفهوم جدید با عنوان پول الکترونیک[۹] و انتقال الکترونیکی منابع، اساس شکل‌گیری بانکداری الکترونیک شد. در واقع بانکداری الکترونیک اوج استفاده از فناوری اطّلاعات و ارتباطات در جهت حذف دو قید زمان و مکان از خدمات بانکی شده است. تعاریف بسیاری از بانکداری الکترونیک ارائه شده اما بنا به تعریف ارائه شده از موسّسات مالی ایالت ایندیانا، بانکداری الکترونیکی یا EFT[10] استفاده آسان از ابزار الکترونیکی به مفهوم خدمات ۲۴ ساعته جهت دسترسی به انواع حساب‌ها، انتقال مستقیم وجه از یک حساب به حساب دیگر به جای پول نقد یا چک، فارغ از قید زمان و مکان تعریف شده است.

ضرورت یک نظام بانکی کارامد برای حضور در بازارهای جهانی و عضویت در عرصه رقابتی و سازمان‌هایی نظیر سازمان تجارت جهانی[۱۱] ایجاب می‌کند تا بانکداری الکترونیک نه به عنوان یک انتخاب بلکه ضرورت مطرح شود. با توجّه به پیشرفت‌های سریع تجارت الکترونیک، استفاده از کارت‌های اعتباری[۱۲] برای خرید ضرورت یافته و تراکنش‌های کارت‌های اعتباری معیار

1. e-Every thing
2. e-Commerce
3. e-Banking
4. e-Learning
5. e-Logistics
6. e-Citizen
7. e-Government
8. e-Life
9. e-Money
10. Electronic funds transfer (EFT)
11. World Trade Organization (WTO)
12. Credit Card

غیررسمی تجارت الکترونیک بر پایه‌ی وب و اینترنت محسوب می‌شود[۱]. طی گزارش دولت ایالات متّحده، کارت‌های اعتباری تقریبا ۱۳ هزار میلیارد دلار از خریدهای اینترنتی این کشور را در سال ۲۰۰۰ به خود اختصاص داده است که این عدد رشد سریع سالانه را نوید می‌دهد[۵].

۱-۲- بیان مساله تحقیق

در حال حاضر بسیاری از سازمان‌ها سند چشم انداز تجارت الکترونیکی و بازاریابی خود را در ایجاد وب‌سایت طراحی می‌کنند. هم اکنون ۶۳ درصد کسب و کارها در کشور انگلستان از بستر وب‌سایت صورت می‌پذیرد که بی‌تردید این داد و ستد بدون دخالت شبکه گسترده بانکی میسر نخواهد بود[۵۵].

بانکداری الکترونیکی و به تبع آن بانکداری اینترنتی، در راستای تحقق اهداف استراتژیک تجارت الکترونیک، علاوه بر مزیت‌های مختلفی همچون رفاه، بهبود کارائی و سرعت خدمات در راستای شکل‌گیری شبکه عمومی اینترنت، با چالش جدیدی پیرامون محرمانگی و امنیت[۱] اطلاعات و جلوگیری از جرائم در فضای سایبر[۲] مواجه است. در بانکداری الکترونیکی، دستیابی به شبکه خدمات برتر، با استفاده از ترکیبی از ابزارهای ارتباطی و مخابراتی و برخی از سامانه‌های بی‌سیم[۳] میسر خواهد بود، حال آن که چگونه این سامانه‌ها و با چه سطحی از دسترسی و امنیت، تراکنش‌های درخواستی را مورد پردازش قرار می‌دهند، درجه‌ای از مخاطره و آسیب‌پذیری را تعریف می‌کند که این درجه خطر، در مورد اینترنت که هویت کاربران آن به آسانی قابل شناسایی نیست بیشتر است[۳۶]. گرایش مردم به خرید برخط[۴] روز به روز در حال افزایش است. مطابق مطالعه ای‌سی‌نیلسن که در سال ۲۰۰۵ اجرا شد، یک دهم مردم جهان به صورت برخط خرید می‌کنند[۱۵]. آلمان و بریتانیا بیشترین خریداران برخط را دارند، و کارت اعتباری معمول‌ترین شیوه پرداخت(۵۹ درصد) است. در حدود ۳۵۰ میلیون تراکنش در هر سال به صورت گزارشی به وسیله بارکلی کارت، بزرگترین شرکت کارت اعتباری در بریتانیا در پایان قرن گذشته انجام شد[۲۴]. بنا به گزارش آنکتاد[۵] تعداد استفاده‌کنندگان از اینترنت در سال ۲۰۰۳ نزدیک به ۶۷۶ میلیون نفر یا به عبارتی ۱۱/۸ درصد از کل جمعیت دنیا را تشکیل می‌دهد که این رقم رشد ۴۹/۵ میلیون نفری یا ۷/۸ درصدی را در مقایسه با آمار پایان سال ۲۰۰۲ نشان می‌دهد. کشورهای در

1. Security
2. Cyber
3. Wireless
4. Online Purchase
5. UNCTAD

حال توسعه بیش از ۳۶ درصد کل کاربران اینترنت را تشکیل می‌دهند و این نسبت از جمعیت کاربران جهانی اینترنت، در بین سال‌های ۲۰۰۰ تا ۲۰۰۳ چیزی حدود ۵۰ درصد رشد داشته است. تعداد وب‌سایت‌های موجود در ژوئن سال ۲۰۰۴ بیش از ۵۱۶۳۵۲۸۴ وب‌سایت بوده است که نسبت به سال قبل ۳۶ درصد افزایش داشته است[۵۴]. بالطبع هرچه شمار کاربران خدمات برخط در پهنه جهان افزایش می‌یابند، فرصت‌ها برای سرقت مهاجمین به جزئیات سوابق عملکرد و اطلاعات مالی مشتریان و سپس ارتکاب به تقلّب[1] در حال افزایش است.

لذا این روند عاملی در جهت تهدید امنیت تراکنش‌ها و نقض حریم خصوصی کاربران بشمار آمده و با پیشرفت فن‌آوری، تعداد نقاط آسیب‌پذیر و حملات کاربران غیرمجاز، نگرانی بسیاری را برای کاربران اینترنتی و صنعت بانکداری بوجود می‌آورد[۴۴]. در این راستا کنترل همه جانبه امنیت سامانه‌های پرداخت الکترونیکی، ارتباطات چند سویه و ردیابی آن نیز به نسبت پیچیده‌تر شده است. به طوری که مجموع تقلّب در سال ۲۰۰۶ در ایالات متّحده بیش از ۳ میلیارد دلار برآورد شده است و در خارج از ایالات متّحده به ترتیب ۱,۶ تا ۱,۷ میلیارد دلار اعلام شده که این برآوردها از نوع تقلّب برخط گزارش شده‌اند[۵۵]. شناسایی رفتارهای مشکوک بر پایه تحلیل داده‌های خرید، یک راه محتمل برای کاهش تقلّب‌های موفّق برخط است. هر کاربری می‌تواند مجموعه‌ای از الگوها شامل مکان جغرافیایی، نوع سامانه تراکنش، میزان اشتباه در تولید تراکنش، تعیین گردش حساب، طبقه‌بندی نمونه‌ای خرید، زمان تراکنش، گردش پول‌های خرج شده یا میزان مبلغ حواله و... را خلق کند[۱۴]. انحراف از برخی الگوها تهدید بالقوّه‌ای برای سیستم محسوب می شود. تقلّب در بانکداری الکترونیکی به دو نوع تقسیم شده است: تقلّب غیربرخط[2] و تقلّب برخط[3]. تقلّب غیربرخط با استفاده از یک کارت فیزیکی به سرقت رفته در فروشگاه یا دستگاه خودپرداز انجام می‌شود[۵] که در اکثر موارد، موسّسه صدور کارت می‌تواند کارت سرقتی را قبل از سوءاستفاده قفل کند. تقلّب برخط از طریق سامانه اینترنتی، صفحات وب، خرید از طریق تلفن ثابت، موبایل و به طورکلی با غیبت صاحب کارت صورت می‌گیرد. در این نوع تقلّب فقط جزئیات کارت‌ها مورد نیاز هستند، و امضا یا صحّت وجود کارت در زمان خرید مورد نیاز نیست[۲۶و۵۷].

اخیراً موسّسه SAS[4]، تخمین خود را از هزینه فریب در اقتصاد انگلستان از ۱۴ میلیارد پوند به ۱۸ میلیارد پوند در سال افزایش داده است. این اطّلاعات در حالی است که افت و خیز تقلّب و

1. Fraud
2. Offline Fraud
3. Online Fraud
4. Statements on Auditing Standards

کلاهبرداری به عنوان یک پدیده کلی حدود ۳۰٪ در سال افزایش می‌یابد و هنوز نرخ محکومیت به طور وسیعی طی ۱۰ سال گذشته تغییر نکرده است[۴۱و۱۰].

* با توجّه به موارد فوق سوالات اصلی این تحقیق بدین صورت مطرح می شود :

۱- الگوهای رفتاری غیرمتعارف یا تهدیدآمیز در بانکداری اینترنتی چیست؟

۲- ویژگی‌های مورد نیاز از تراکنش‌های اینترنتی در بانکداری الکترونیکی به چه صورت استخراج می‌شود؟

۳- چگونگی تبیین یا نگاشت الگوهای مرتبط بر روی درخت تصمیم‌ساز به چه صورت است؟

۴- اعتبار الگوریتم درخت تصمیم‌گیری در نزد خبرگان خدمات اینترنتی بانک خصوصی تا چه حدی است؟

۱-۳- معرفی موضوع تحقیق، سابقه و ضرورت انجام آن

بهره‌گیری از مزایای فنآوری نوین، نیازمند شناخت کاملی از ویژگی‌ها، وجود زیرساخت‌ها و امکانات متعدّدی است که در صورت عدم توجّه به آنها نمی‌توان از موفّقیت و اثربخشی رویکرد جدید اطمینان حاصل پیدا کرد. یکی از این زیرساخت‌های مهم در مقبولیت و گسترده‌شدن فرایندهای بانکداری الکترونیکی، افزایش امنّیت و کنترل رفتارهای غیرقانونی در این نوع سیستمها است. امنیّت بانکداری الکترونیکی بیشتر در زمینه بانکداری اینترنتی مطرح بوده و چالش‌هایی را در زمینه اطمینان از تراکنش‌ها بوجود آورده است. هر سیستم بانکداری الکترونیکی باید موضوعاتی مانند تصدیق اصالت، محرمانگی، یکپارچگی و انکارناپذیری و دیگر عوامل امنیتی را در نظر داشته و تضمین کند که فقط افراد مجاز بتوانند به اطلاعات مجاز، محرمانه و حساب‌های مشتریان دسترسی داشته و سوابق معاملات غیرقابل ردیابی و رسیدگی باشند[۴۰]. مسئله اعتماد در محیط بانکداری اینترنتی مهم‌تر از بانکداری در محیط غیربرخط است زیرا ایجاد و پرورش اعتماد وقتی مهم است که عدم اطمینان و ریسک فراگیر باشد[۴۵]. رشد خارق‌العاده شمار تراکنش‌های اینترنتی، به خصوص برای خریدهای برخط، که اخیراً به افزایش ذاتی در فعّالیت‌های متقلّبانه منجر شد، طراحی و بهره‌گیری سیستم شناسایی رفتارهای مشکوک را برای کلیه سرویس‌های اینترنتی موسّسات مالی بخصوص بانک‌ها را در جهت کاهش ریسک و زیان، ضروری می‌سازد.

پیش‌بینی رفتار کاربران در نظام‌های مالی می‌تواند در شرایط متفاوتی استفاده شود. یکی از جالب‌ترین زمینه‌های پیش‌بینی تقلّب، خطوط اعتباری در محیط برخط است، به خصوص

پرداخت کارت‌های اعتباری برای حجم داده‌های بیش از ۴۰۰ هزار تراکنش در هر روز، یک کاهش ۲,۵ درصدی تقلّب برای ذخیره یک میلیون دلار در هر سال پیش‌بینی می‌شود[۴۷]. به طور حتم، همه تراکنش‌های موجود، مجاز نیستند. با این وجود، تراکنش‌هایی وجود دارند که به طور رسمی معتبر هستند، بنابراین اجتناب از تقلّب به وسیله تراکنش برخط پیش از آنکه غیرقانونی شناخته شود، مهم است[۳].

یکی از روش‌های مورد استفاده در تشخیص رفتارهای مشکوک، استفاده از تکنیک‌های داده-کاوی است که بر تحلیل‌های آماری و کشف رفتار مشتریان و استفاده از الگوهای آموزشی برای شناسایی جرم تمرکز دارد[۳۷]. این روش‌ها مبتنی بر یادگیری هستند. یوفنگ‌کو و همکارانش در سال ۲۰۰۲ به «بررسی تکنیک‌های شناسایی تقلّب» پرداختند. آنها شناسایی تقلّب را در سه حوزه،کارت‌های اعتباری، شناسایی تجاوز کامپیوتری و سامانه اینترنتی بحث کردند و رویکرد شبکه‌های عصبی را ابزاری بسیار معمولی برای شناسایی تقلّب دانستند. اگرچه این تکنیک به علّت فقدان دسترسی به مجموعه داده‌ها برای اجرا سخت است[۵۷]. در تحقیقی از[۸] با استفاده از روش دسته‌بندی در داده‌کاوی و مقایسه الگوریتم‌های درخت تصمیم[1]، شبکه‌های عصبی و شبکه-های بیزین[2] استفاده شده تا در شناسایی متخلفین مالیاتی کمک شود و همچنین در تحقیقی که از[۲] ارائه شده از تکنیک شبکه عصبی برای تشخیص تقلب در بانکداری الکترونیکی استفاده شده و با کمک فرایند یادگیری با نظارت، برای ساختن مدل‌هایی از تراکنش‌های فریب‌آمیز بانکداری اینترنتی بهره برده است. در تحقیقی دیگر که در سال ۲۰۰۵ روی سایت eBay صورت پذیرفت از تکنیک‌های معروف C4.5 و C5 که مربوط به الگوریتم‌های درخت تصمیم‌ساز است در جهت تشخیص کلاهبرداری در حراج الکترونیکی سایت e-Bay استفاده شده که تا ۸۳٪ در ارائه الگوی‌های متناظر با رفتار مشتریان در جهت کشف کلاهبرداری موثر واقع گردید[۷]. البته جیان‌یونگ‌تو و همکارانش نیز در سال ۲۰۰۴ تحقیقی با عنوان «سیستم حفاظت مصنوعی شناسایی رفتارهای مشکوک» انجام دادند. آنها یک سیستم حفاظت مصنوعی ژنتیکی پایه‌ای برای شناسایی تقلّب پیشنهاد دادند که یک سیستم خودسازگار طراحی شده برای شناسایی تقلّب کارت‌های اعتباری است. با استفاده از این مدل یادگیری پایه‌ای و الگوریتم ژنتیکی، سیستم می‌تواند یادگیری برخط را با زمان و هزینه محدود اجرا کند و قابلیت شناسایی ناهنجاری را در پیشرفت سریع تراکنش‌ها و فعّالیت‌های تجاری به روز کند[۲۰].

1. Decision Tree Algorithms
2. Bayesian Belief Networks

علی‌ایحال با توجّه به وجود چالش‌های امنیتی متعدّدی که به عنوان تهدید در مسیر خدمات برخط بانکداری الکترونیکی تعریف می‌شود و افزایش ریسک سوءاستفاده از حجم عظیم اطّلاعات مالی، بانک‌ها را الزام می‌دارد تا جهت سلامت تراکنش‌های انجام شده به بررسی رفتار کاربران بپردازند. در این تحقیق سعی شده تا با استفاده از تکنیک طبقه‌بندی و بکارگیری و مقایسه الگوریتم‌های درخت تصمیم‌گیری، دسته‌بندی دقیقی از رفتار مشتریان بانک تهیه نموده و مدلی را برای تشخیص رفتارهای غیرمتعارف و تقلّب در بانکداری اینترنتی ارائه کرده و در نهایت اعتبارسنجی و میزان دقت آن را مورد تحلیل و بررسی قرار دهیم.

۱-٤- اهداف تحقیق

نهایتاً، از این پروژه انتظار می‌رود که با استفاده و مقایسه الگوریتم‌های درخت تصمیم‌گیری و تولید و ارائه داده‌های آموزشی که از طریق پیش‌پردازش اولیه رکوردهای موجود در پایگاه داده یک بانک خصوصی صورت می‌گیرد، مجموعه الگوهای رفتاری مشکوک و عادی را در بین جریان تراکنش‌ها که حاصل نحوه رفتار کاربران است تولید و دسته‌بندی کرده و در مرحله آخر اعتبارسنجی سیستم و میزان دقت تطابق خروجی درخت تصمیم‌ساز را با نظر ۳۰ نفر از خبرگان و کارشناسان ارشد سامانه اینترنتی چند بانک که به صورت دانش استنتاجی یا تجربی مستتر است از طریق ارائه پرسشنامه جمع‌آوری و مورد ارزیابی، تحلیل و مقایسه قرار داده تا درصد اعتماد به این الگو به عنوان مدل تولیدی درخت تصمیم‌ساز جهت کشف تراکنش‌های مشکوک و رفتارهای غیرمعمول کاربران در سامانه اینترنتی مشخص گردد. احتمال بالای شناسایی درست می‌تواند با احتمال کاهش تقلّب شناسایی نشده و هشدارهای غلط اجرایی گردد[۳۵]. امید آن است که با ارائه طرح پیشنهادی بتوان میزان دقت مدل تولیدی و نتایج را بهینه نمود.

۱-۵- قلمرو تحقیق

۱-۵-۱- قلمرو مکانی تحقیق

جامعه تحقیق ما در این پروژه تراکنش‌های ثبت شده حاصل از میانگین رفتار روزانه مشتریان اینترنتی یک بانک خصوصی در کشور خواهد بود که این داده‌ها از بدو ورود به سامانه اینترنتی و حین فرایند تولید تراکنش تا لحظه خروج از سایت، در سرورهای بانک ثبت می‌گردد.

۱-۵-۲- قلمرو زمانی تحقیق

این تحقیق با بررسی‌های انجام شده در خصوص درخواست اجازه دسترسی به داده‌های آزمایشی بانک و همچنین گرداوری مقالات مربوطه، از مهرماه ۸۸ شروع شده و تخمین زده می-شود بعد از مشاوره با اساتید محترم تا اواخر بهار ۸۹ ادامه یابد.

۱-٦- بیان فرضیه

با توجّه به الزامات امنیّتی و عدم دسترسی به اطّلاعات شخصی کاربران، در فاز پیاده‌سازی این تحقیق، از پارامترهای مفیدی چون: سن، جنسیت، شغل، صلاحیت رفتاری و میزان سطح درامد و اهلیّت مشتری که جهت افزایش دقّت و جامعیّت مدل موثر بوده است استفاده نگردید.

۱-۷- کاربردهای تحقیق

چنانچه بتوان درصد بالای اعتماد و تطابق الگوهای رفتاری تولید شده درخت تصمیم‌ساز را با دانش تجربی خبرگان بانکداری الکترونیکی اثبات نمود از این طریق مدل تولیدی درخت تصمیم-ساز جهت یافتن الگوهای جالب و مرتبط با رفتارهای غیرمتعارف و مشکوک در تراکنش‌های بانکداری الکترونیکی کاربرد فراوانی در جهت کاهش ریسک اعتباری و خدمات اینترنتی خواهد داشت لذا با طراحی سیستم هوشمند تصمیم‌گیری و ارتباط آن با دانش درخت تصمیم می‌توان سیستم اتوماسیون تصمیم‌یار را برای مدیران IT و تصمیم‌گیرندگان اصلی در واحد بازرسی و امنیّت فناوری بانک‌ها، موسّسات مالی و اعتباری علاقمند به کشف رفتارهای مشکوک در راستای موضوع کسب مزیت رقابتی از طریق سامانه اینترنتی استفاده برد.

۱-۸- محدودیت‌های تحقیق

الف- منابع اطلاعاتی موجود برای این تحقیق محدود می‌باشد.

ب- نمونه‌های عملی و با سابقه در بانک‌های ایرانی پیاده‌سازی نشده است.

ج- با توجّه به الزامات امنیّتی و عدم دسترسی به اطّلاعات شخصی کاربران، در فاز پیاده‌سازی این تحقیق از پارامترهای مفیدی چون: سن، جنسیت، شغل، صلاحیت رفتاری و میزان سطح درامد و اهلیّت مشتری وجود داشته و می‌شد از آن در جهت افزایش دقّت و جامعیّت مدل بهره برد استفاده نگردیده است.

۱-۹- ساختار تحقیق

۱- شناخت بانکداری اینترنتی و عوامل موثر در تعریف و شناسایی رفتار کاربران و دسته‌بندی رفتارها.

۳- جمع‌آوری تراکنش‌های سامانه اینترنتی بانک خصوصی جهت ایجاد یک پایگاه داده محلّی و پیش پردازش اولیه داده‌ها بمنظور تفکیک رفتار کاربران از طریق نرم افزارMS Excell 2007 و SQL Server 2005 .

۴- بکارگیری و مقایسه الگوریتم‌های درخت تصمیم‌گیری جهت ساخت مدلی که نهایتاً به کشف رفتارهای غیرمتعارف و مشکوک ختم شود.

۵- بررسی و بکارگیری ابزار تحلیلی و پیاده سازی داده‌کاوی در راستای الگوریتم درخت تصمیم به نام Spss_Clementine جهت کشف قوانین قابل فهم و ارائه گزارش سبب تصمیم‌گیری مدیران و مسئولین مرتبط، در درک بیشتر و در صورت امکان جلوگیری از ادامه تخلّف.

۷- ارزیابی و اعتبارسنجی نتایج بدست آماده از تست و آزمون الگوریتم‌های درخت تصمیم‌گیری و مقایسه مدل تولیدی با دانش استنتاجی خبرگان بانک خصوصی و تعیین درصد اعتماد به سیستم.

۱-۱۰- روش تحقیق

در این تحقیق منبع اصلی اطّلاعات جهت پیاده‌سازی، رکوردهای ثبت شده در سرورهای بزرگ‌ترین بانک خصوصی در کشور می‌باشد. همچنین در این تحقیق به طور گسترده از دو روش تفکر عمیق و مطالعه پیمایشی که شامل بررسی کتب، نشریات، مقالات فارسی و لاتین، پایان‌نامه-های مرتبط با موضوع و اینترنت است استفاده شده و جهت جمع‌آوری داده‌ها و تکمیل مراحل تست و پیاده‌سازی الگوریتم‌ها، از روش تحقیقی و میدانی که مشمول بر بررسی اسناد، مصاحبه و ارائه پرسشنامه به متخصّصین IT و خبرگان بانکداری الکترونیکی است، استفاده می‌شود.

۱-۱۱- تعاریف کلی

۱-۱۱-۱- بانکداری الکترونیکی

بانکداری الکترونیکی عبارت است از فراهم آوردن امکاناتی برای کارکنان در جهت افزایش سرعت و کارایی آنها در ارائه خدمات بانکی در محل شعبه و فرآیندهای بین‌شعبه‌ای و بین‌بانکی در سراسر دنیا و همچنین ارائه امکانات سخت‌افزاری و نرم‌افزاری مبتنی بر شبکه و مخابرات برای

تبادل منابع و اطلاعات مالی به صورت الکترونیکی است تا مشتریان بتوانند بدون نیاز به حضور فیزیکی در بانک، در هر ساعت از شبانه‌روز(۲۴ساعته) از طریق کانال‌های ارتباطی ایمن و با اطمینان، عملیات بانکی دلخواه خود را انجام دهند[۶۰].

۱-۱۱-۲- بانکداری اینترنتی

بانکداری اینترنتی به معنی انجام تراکنش‌های بانکی و مالی به کمک بستر اینترنت است و تفاوت آن با سایر تراکنش‌های مالی، در نوع شبکه‌ای است که مورد استفاده قرار می‌گیرد، یعنی در بانکداری اینترنتی، فضای اینترنت به عنوان شبکه گسترده جهانی، فرصت و محرکی جهت ایجاد تحولات تجارت الکترونیکی خواهد بود[۶۰].

۱-۱۱-۳- کشف دانش[1]

فرآیند استخراج دانش مفید از میان انبوهی از داده‌ها، کشف دانش نامیده می‌شود. به استخراج دانش از پایگاه داده،داده‌کاوی و استخراج دانش از متن‌ها، متن‌کاوی[2]گویند[۱۸و۱۹].

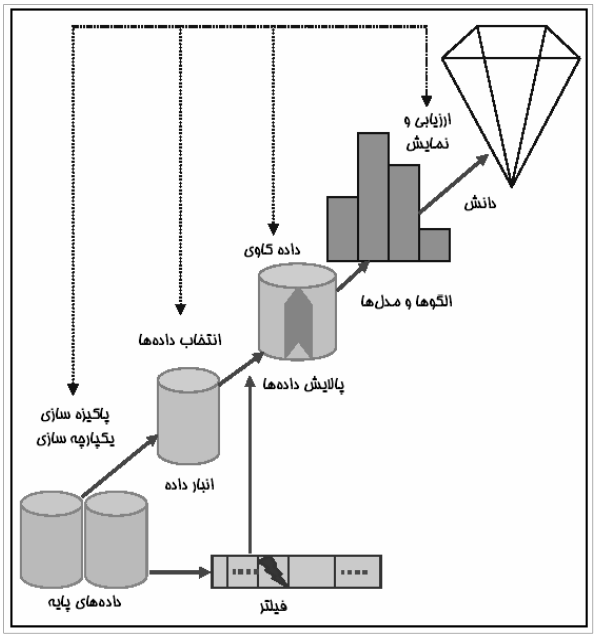
۱-۱۱-۴- سیستم شناسایی تقلّب[3]

به سیستم‌هایی اطلاق می‌شود که به منظور شناسایی رفتار غیرمتعارف و مشکوک کاربران و کشف الگوهای غیرعادی به کار می‌رود.

۱-۱۱-۵- داده‌کاوی[4]

داده‌کاوی یا استخراج دانش از پایگاه های داده، فرآیند مهم شناسایی الگوهای معتبر، جدید و قابل فهم در میان انبوهی از داده‌ها است یا به عبارتی دیگر کشف نیمه‌خودکار دانش، فرآیند تشخیص الگوهای پنهان، معتبر، نو، مفید و نهایتاً قابل درک در داده‌ها است[۱۷]. مطابق با شکل ۱-۱ فرآیند داده‌کاوی شامل شش مرحله می‌باشد که در فاز انتخاب نوع الگوریتم، بسته به اهداف داده‌کاوی(یعنی پیش‌بینی و توصیف) و نوع قواعد مطلوب جهت استخراج دانش، می‌تواند شامل: طبقه بندی(کلاس بندی[5])، خوشه بندی[6]، وابستگی[7]، پیش بینی[8] و غیره باشند.

1. Knowledge Dicovery in Database (KDD)
3. Text mining
3. Fruad detection system
4. Data Mining
5. Classification
6. Clustering
7. Association
8. Prediction



شکل ۱-۱- فرآیند داده‌کاوی[۱۷]

۱-۱۱-۶- طبقه‌بندی(کلاس بندی)[1]

طبقه‌بندی از تکنیک‌های داده‌کاوی بوده که هدف آن یافتن مدلی با تشخیص دسته‌ها و مفاهیم بین آنها است تا اشیاء ناشناخته دیگر را پیش‌بینی کند. طبقه‌بندی یک تابع یادگیری است داده ها را به دسته‌ها نگاشت می‌کند. این داده‌ها به دو بخش آموزش جهت یادگیری قواعد و آزمون که جهت بررسی دقّت طبقه‌بندی و جلوگیری از پیش برازش[2] به کار می‌روند، تقسیم می‌شوند[۱۸].

۱-۱۱-۷- درخت تصمیم‌گیری

درخت یکی از ابزارهای قوی و متداول برای طبقه‌بندی و پیش‌بینی می‌باشد.بطوریکه پیش‌بینی به دست آمده از درخت در قالب یک سری قواعد توضیح داده می‌شود. در حالی که در سایر تکنیک‌ها، تنها نتیجه پیش‌بینی بیان می‌شود و چگونگی به دست آمدن آنها در خود شبکه پنهان می‌ماند. طبقه‌بندی‌هایی که توسط درخت تصمیم‌ساز ایجاد می‌شوند، از روی شباهت داده‌های ذخیره شده که لزوما عددی نیستند از طریق پارامترهای پیش‌بینی کننده، قابل انجام می‌باشد[۱۸].

۱-۱۲- واژگان و اصطلاحات تخصصی طرح

الف : واژگان فارسی:

بانکداری الکترونیکی، بانکداری اینترنتی، امنیت، رفتار مشکوک، درخت تصمیم‌گیری

ب: واژگان لاتین:

Electronic Banking- Internet Banking- Security- Suspicious Behavior- Decision Tree

1. Classification
2. Over fit