فصل‌نامه مطالعاتی در مدیریت بانکی و بانکداری اسلامی

مؤسسه عالی آموزش بانکداری ایران
تابستان ۱۴۰۳، دوره ۱۰، شماره ۲۷، صفحه ۳۸-۱

به‌کارگیری روش شبکه‌های هوشمند عصبی در شناسایی و رتبه‌بندی رفتاری مشتریان مظنون به پول‌شویی (مورد مطالعه: بانک ملت)

عبد المهدی ارجمند نژاد[۱]
رضا حبیبی[۲]
مسعود بابائی کیاپی*[۳]

چکیده

پول‌شویی به‌عنوان پدیده‌ای مجرمانه می‌تواند ضمن ایجاد اختلال در عملکرد بانک‌ها، هزینه‌های بسیاری بر آن‌ها تحمیل نماید. استراتژی کنونی اکثر بانک‌های ایران شناسایی این دسته از مشتریان با استفاده از قوانین عمومی می‌باشد که دارای پاسخ‌های مثبت کاذب فراون بوده و کشف عملیات پول‌شویی را با مشکل مواجه می‌نماید. هدف از این پژوهش فراهم نمودن معیارهایی جهت تشخیص مشتریان بانک با بالاترین احتمال ظن به پول‌شویی با استفاده از الگوریتم‌های هوشمند می‌باشد. در این تحقیق از یک الگو و مدل دو مرحله‌ای برای تجزیه‌وتحلیل رفتار مشتریان با استفاده از پایگاه داده‌ای در خصوص ویژگی‌های مشتریان و داده‌های مالی آن‌ها در بازه زمانی شش ماهه دوم منتهی به اسفند ۱۴۰۱ و از تکنیک‌هایی همچون نگاشت خود سازمان دهنده (SOM) و قوانین وابستگی برای ایجاد پروفایل جامعی از مشتریان، استفاده شده است. متغیرهای تحقیق شامل داده‌های جمعیت شناختی، سرویس‌های بانکی، تراکنش‌های مالی مشتریان می‌باشند که از پایگاه داده‌ای بانک ملت استخراج شده است. نتایج نشان می‌دهد که مدل پژوهشی به‌خوبی قادر به شناسایی و تحلیل مشتریان مظنون به پول‌شویی می‌باشد که قبلاً توسط بانک عامل، بررسی شده است. با استفاده از هر دو مدل پیشنهادی، نتایج، تفاوت معناداری را نشان نمی‌دهند.

واژه‌های کلیدی: شبکه‌های عصبی مصنوعی، نقشه خودسازمان‌دهنده، رتبه‌بندی رفتاری، رتبه‌بندی مشتریان بر مبنای ریسک پول‌شویی.

طبقه‌بندی JEL: E42، M15، G21 و C88.

---

۱. مدرس، موسسه عالی آموزش بانکداری ایران، تهران، ایران؛ am.arjmand@cbi.ir
۲. استادیار، مؤسسه عالی آموزش بانکداری ایران، تهران، ایران؛ r_habibi@ibi.ac.ir
۳. دکتری، بانک ملت، تهران، ایران (نویسنده مسئول)؛ masoudbabaee1355@gmail.com

مقدمه

واژه پول‌شویی بیشتر متوجه تطهیر درآمدهای ناشی از اختلاس، رشوه، دزدی، تقلب در بانک‌ها، بیمه‌ها و بازارهای سرمایه، دایر کردن قمارخانه و خانه‌های فساد، قاچاق انسان و کالا، تجارت اسلحه، شرکت در شبکه‌های کلاهبرداری و تأمین مالی تروریسم و غیره می‌باشد که همه این موارد توسعه اجتماعی را هدف قرار داده‌اند (چن و ماتا[1]، ۲۰۱۱).

یکی از جرم‌هایی که منجر به ایجاد اختلال در حوزه‌های اقتصادی و نهادهای مالی به‌ویژه در عملکرد بانک‌ها می‌شود، جرم پول‌شویی است که هزینه‌های بسیاری بر بانک‌ها تحمیل می‌نماید (لوکارانا[2]، ۲۰۱۹). ازاین‌رو فرایند پول‌شویی می‌تواند آثار منفی زیادی بر فعالیت‌های مالی و اقتصادی داشته باشد. بنابراین شناسایی و پیش‌بینی آن دارای اهمیت بسیار زیادی می‌باشد (پلکیسی و همکاران[3]، ۲۰۱۸).

در ایران تابه‌حال اکثر تحقیقات انجام شده در زمینه رتبه‌بندی مشتریان مظنون به پول‌شویی کشف قوانین عمومی بوده است. مطالعات بسیار کمی از نقطه نظر رتبه‌بندی رفتاری این دسته از مشتریان در بانک صورت پذیرفته است (کوثری، ۱۳۹۲).

در این زمینه، بانک مرکزی جمهوری اسلامی ایران به‌عنوان مرجع رسمی و ناظر بر عملکرد بانک‌ها در خصوص مبارزه با پول‌شویی، اقدام به ابلاغ دستورالعمل‌های مبارزه با پول‌شویی و پیاده‌سازی قواعد ۲۸ گانه‌ای نموده است که در عمل، مستلزم پردازش حجم وسیعی از داده ها و اطلاعات در هر روز می باشد. به عنوان مثال، در بانک ملت و به‌طور متوسط، هر یک از این قواعد با حجم تراکنشی در حدود ۵۰ میلیون در هر روز، دارای خروجی ۴۵۰ الی ۵۰۰ تراکنش مظنون به پول‌شویی می‌باشد که تعداد آن بالغ بر ۱۲،۶۰۰ گزارش (۴۵۰*۲۸) می‌شود. این گزارش ها باید در هر روز در واحد مبارزه با پول‌شویی بانک توسط کارشناسان مربوطه مورد بررسی قرار گرفته و زنجیره عملکرد مالی این دسته از مشتریان از سامانه‌های مختلف استخراج شده و بررسی گردد.

از سوی دیگر، محدودیت تعداد کارشناسان خبره، کندی سیستم‌ها، تراکم پاسخ‌های مثبت کاذب فراوان در هر روز، فرسایشی و پیچیده بودن عملیات شناسایی رفتار مشتریان مظنون به پول‌شویی و طولانی بودن تهیه گزارشات مربوط به این اشخاص و ارسال آن به مقامات ذی‌ربط، ضرورت استفاده از تکنیک‌های جایگزین مانند الگوریتم‌های هوشمند را نمایان می‌سازد.

---

1. Chen & Mathe
2. Lokanan
3. Plaksiy et al

در این تحقیق از یک الگو و مدل دو مرحله‌ای برای تجزیه‌وتحلیل رفتار مشتریان با استفاده از پایگاه داده‌ای درخصوص ویژگی‌های مشتریان و داده‌های مالی آن‌ها استفاده شده است.

مبانی نظری پژوهش

الگوریتم‌ها و روش‌های یادگیری ماشین برای تشخیص تراکنش‌های مشکوک، به‌عنوان یک راه‌حل برای شناسایی، تحلیل ارتباطات، مدل سازی رفتار، امتیاز ریسک و تشخیص ناهنجاری مورد استفاده قرار می‌گیرد (ماگومدو و پاولیو[1]، ۲۰۱۸). رویکرد ریسک محور مبارزه با پولشوئی، دارای سناریویی مبتنی بر طبقه‌بندی ریسک است. بدین صورت که افراد، مناطق، مشاغل، زمان‌ها و غیره، شناسایی و طبقه‌بندی می‌شوند و بر مبنای تحلیل‌های صورت گرفته سناریوهای کنترلی، تعیین یا بازبینی می‌شوند (روچا- سالازار و همکاران[2]، ۲۰۲۱).

شرایط موجود سیستم ضد پول‌شویی

درحال‌حاضر رایج‌ترین روش‌های مورد استفاده مؤسسات مالی ایران برای شناسایی و گزارش‌دهی موارد مشکوک مبتنی‌بر سیستم‌های قانون محور می‌باشند که از سوی مراجع نظارتی مانند بانک مرکزی ج.ا.ا و **FIU** (واحد اطلاعات مالی) مستقر در وزارت اقتصاد تدوین می‌گردد.

سیستم مبتنی بر قانون (رول‌های ۲۸ گانه)، مجموعه‌ای از قوانین از پیش تعریف شده می‌باشند که معاملات و یا تراکنش‌هایی که با شرایط تعریف شده مطابقت دارند، هشدارهایی را ایجاد می‌کند که باعث تحقیقات و بررسی‌های بیشتر توسط کارشناسان ادارات مبارزه با پول‌شویی بانک‌ها می‌شود.

یک اشکال عمده از سیستم‌های قانون محور، تولید حجم قابل توجهی از هشدارهای مثبت اشتباه می‌باشد که از نظر زمان و منابع برای ردیابی موارد، نیاز به بررسی و دارای هزینه هستند (گواورا و همکاران[3]، ۲۰۱۹).

بر مبنای تخمین از هشدارهای اشتباه صادره در سیستم‌هایی بانکی ایران که معمولاً توسط FIU و از طریق پورتال محرمانه اعلام می‌شوند بیش از ۸۰ درصد از کل هشدارهای صادره توسط این سیستم‌های قانون محور اشتباه می‌باشند. در نتیجه، همیشه افزایش تمایل به توسعه ابزارهای پیشرفته‌تر برای

1. Magomedov & Pavelyev
2. Rocha-Salazar et al
3. Guevara et al

تشخیص دقیق‌تر پول‌شویی در معاملات وجود داشته تا در کمترین زمان و با بیشترین دقت تراکنش‌های مشکوک به پول‌شویی شناسایی گردند.

رتبه‌بندی ریسک محور مشتریان از منظر پول‌شویی (RBA) در بانک ملت

رویکرد ریسک محور مبارزه با پول‌شویی در بانک ملت، دارای سناریویی مبتنی‌بر طبقه‌بندی ریسک مشتریان در سه خوشه یا طبقه (۱- دارای ریسک بالا، ۲- دارای ریسک متوسط و ۳- دارای ریسک پایین) می‌باشد. بدین صورت که مشتریان بر مبنای شاخص‌های متعدد، امتیاز ریسک متناسب با آن را دریافت خواهند نمود و با توجه به جمع امتیازهای مربوطه در یکی از طبقات ریسک اشاره شده قرار خواهند گرفت. نحوه توزیع امتیاز به‌گونه‌ای است که اگر دامنه امتیاز به‌دست آمده تا ۳۳ باشد، مشتریان عادی و با ریسک کم قلمداد می‌گردند و درصورتی‌که جمع امتیازهای به‌دست آمده در دامنه ۳۳ تا ۶۶ باشد جزء مشتریان با ریسک متوسط بوده و از ۶۶ به بالا جزء مشتریان پر خطر محسوب می‌گردند. در نهایت بر مبنای تحلیل‌های صورت گرفته، سناریوهای کنترلی تعیین و استراتژی‌های نحوه تعامل با دسته‌های مختلف مشتریان اعمال می‌شوند.

جدول۱: رتبه بندی مشتریان مظنون به پول‌شویی در بانک ملت

رتبه بندی مشتری	امتیاز	رتبه
	بین ۱ تا ۳۳	کم ریسک
	بین ۳۳ تا ۶۶	ریسک متوسط
	بالاتر از ۶۶	ریسک بالا

داده‌کاوی و روش‌های کشف دانش

داده‌کاوی عبارت است از "استخراج اطلاعات و دانش و کشف الگوهای پنهان از یک پایگاه داده‌های بسیار بزرگ". این دانش و الگوها معمولاً مستتر در داده می‌باشند. از داده‌کاوی می‌توان برای انجام کارهایی مثل دسته‌بندی[1]، پیش‌بینی[2]، تخمین[3] و خوشه‌بندی[4] داده‌ها استفاده نمود. برای انجام این کارها تکنیک‌هایی توسعه یافته‌اند که با توجه به پیشرفت کامپیوتر و این علم همه روزه بر تعداد و

1. Classification
2. Prediction
3. Estimation
4. Clustering

کیفیت این تکنیک‌ها افزوده می‌شود. تعدادی از معروف‌ترین این تکنیک‌ها عبارتند از: الگوریتم‌های خوشه‌بندی، شبکه‌های عصبی[1]، الگوریتم‌های ژنتیک[2]، نزدیک‌ترین همسایگی[3] و درخت تصمیم‌گیری[4] (کیم[5]، ۲۰۰۷). ویژگی‌های داده‌کاوی سبب شده که امروزه در تجارت و کسب‌وکار از آن بسیار استفاده شود. شرکت‌های مخابراتی، بانک‌ها، شرکت‌های بازاریابی و تبلیغاتی و کلیه شرکت‌هایی که از بانک‌های اطلاعاتی بزرگی برخوردار بوده و از اهمیت اطلاعات و آگاهی در بازار خبر دارند می‌توانند از داده‌کاوی به‌عنوان ابزار بسیار قوی برای رشد و پیشرفت در بازار استفاده کنند (چن[6]، ۲۰۱۸).

فعالیت‌های پایگاه داده‌ها

۱- دریافت داده: سیستم‌های انبار داده مختلف، نیاز به دریافت داده در بازه‌های زمانی متفاوتی را دارند ولی معمولاً داده‌های کل مجموعه هر ماه یا هر چندماه یکبار برای انبار فرستاده می‌شوند.

۲- پاک‌سازی: پاک‌سازی داده‌ها عبارت است از شناسایی و حذف خطاها و ناسازگاری‌های داده‌ای به‌منظور دستیابی به داده‌هایی با کیفیت بالاتر.

۳- یکپارچه‌سازی: این فاز شامل ترکیب داده‌های دریافتی از منابع اطلاعاتی مختلف، استفاده ازمتاداده‌ها برای شناسایی و حذف افزونگی داده‌ها، تشخیص و رفع برخوردهای داده‌ای می‌باشد.

٤- تبدیل داده‌ها: در این فاز، داده‌های ورودی طی مراحل زیر به شکلی که مناسب عمل داده‌کاوی باشند، در می‌آیند:

- هموارسازی[7]: منظور از داده‌های نویزی، داده‌هایی هستند که در خارج از بازه مورد نظر قرار می‌گیرند.
- ادغام[8]: ادغام داده‌ها به معنی به‌دست آوردن اطلاعات جدید از ترکیب داده‌های موجود می‌باشد.
- کلی‌سازی[9]: کلی‌سازی به معنی دسته‌بندی داده‌های موجود براساس ماهیت و نوع آن‌ها است.

---

1. Neural Network
2. Genetic Algorithm
3. Nearest Neighboring
4. Decision Tree
5. kim
6. Chen
7. Smoothing
8. Aggregation
9. Generalization

– نرمال‌سازی[1]: منظور از نرمال‌سازی، تغییر مقیاس داده‌ها است. به‌عنوان مثالی از نرمال‌سازی، می‌توان به تغییر بازه یک فیلد از مقادیر موجود به بازه ۰ تا ۱ اشاره کرد.

– افزودن فیلدهای جدید: گاهی اوقات برای سهولت عمل داده‌کاوی می‌توان فیلدهایی به مجموعه فیلدهای موجود اضافه کرد (چن و وانگ، ۲۰۱۸).

معرفی بعضی از روش‌های داده‌کاوی

هرگاه داده‌ها دارای خصیصه‌ای خاص باشند که مستقیماً از دیگر خصایص بوجود نیامده باشند اما بین آن مشخصه و دیگر ابعاد رابطه وابستگی وجود داشته باشد، در این صورت می‌توان با کشف مُدلی بر مبنای دیگر مشخصه‌ها، آن بعد مذکور (که نشان‌دهنده دسته خاصی از داده‌ها است) را شناسایی نمود (هسو[2]، ۲۰۰۸).

خوشه‌بندی

از نظر مفهومی خوشه‌بندی یعنی گروه‌بندی یک‌سری موجودیت در گروه‌های مختلف به‌طوری که این گروه‌ها نشان دهنده مفهوم یا معنی خاصی باشند و یا به‌عبارت ساده‌تر به یکدیگر شبیه باشند. از دیدگاهی علمی‌تر خوشه‌بندی یعنی دسته‌بندی داده‌ها به **k** گروه مختلف به‌طوری که داده‌هایی که در یک دسته قرار می‌گیرند به یکدیگر شبیه باشند و داده‌های دسته‌های مختلف با یکدیگر تفاوت داشته باشند (کیم، ۲۰۰۶).

قوانین وابستگی[3]

قوانین وابستگی داده‌کاوی که ابتدا توسط آگراول[4] (۱۹۹۳) توسعه داده شدند، یکی از معروف‌ترین تکنیک‌های داده‌کاوی بوده و کاربردهای وسیعی در کسب‌وکار سنتی و کسب‌وکارهای الکترونیکی پیدا کرده‌اند. فروش، طراحی کاتالوگ، تعیین چیدمان فروشگاه و بخش‌بندی مشتریان و همچنین، بهبود و شخصی‌سازی صفحات وب سایت، نمونه‌هایی از کاربرد قوانین وابستگی داده‌کاوی در کسب‌وکارهای سنتی و الکترونیکی می‌باشند (چن، ۲۰۱۸).

---

1. Normalization
2. Hsu
1. Association Rules
4. Agrowal

شبکه‌های عصبی مصنوعی

شبکه عصبی مصنوعی یک سیستم پردازش اطلاعات است که ویژگی‌های عملکردی مشابه شبکه عصبی بیولوژیکی دارد. شبکه‌های عصبی مصنوعی به‌عنوان مدل‌های ریاضی تعمیم یافته از نورون‌های بیولوژیکی یا شناخت انسان براساس مفروضات زیر توسعه یافته است:

- پردازش اطلاعات در واحدهای بسیار ساده‌ای که نورون نامیده می‌شوند، رخ می‌دهد.
- سیگنال‌های نورون‌ها از طریق اتصالات، انتقال می‌یابد.
- هر ارتباط وزنی دارد که در یک شبکه عصبی متداول، در سیگنال ورودی ضرب می‌شود.
- هر نورون یک تابع فعال‌سازی (که عمدتاً غیرخطی است) دارد (چن ، ۲۰۱۸).

اجزا و ساختار شبکه‌های عصبی مصنوعی

یک شبکه از عناصر پردازشی تشکیل می‌شود. این عناصر پردازشی را می‌توان به طرق مختلف یا در معماری‌های گوناگون سازماندهی کرد.

عناصر پردازشی: یک شبکه عصبی مصنوعی متشکل از نورون‌های مصنوعی (که به‌طور خلاصه به آن‌ها نورون می‌گویند) می‌باشد که در واقع همان عناصر پردازشی (PE[1]) هستند. هرکدام از این نورون‌ها، درون‌داده‌ها را دریافت و پردازش می‌کنند و در نهایت، یک برون‌داد از آن ارائه می‌دهند.

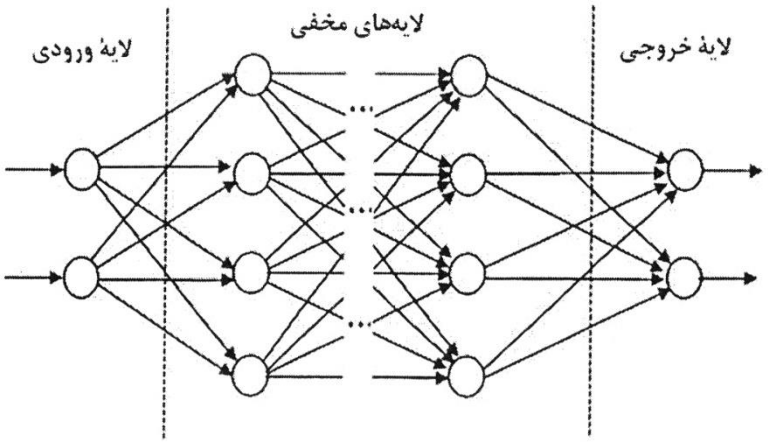
درون‌داد می‌تواند داده‌های خام یا برون‌داد دیگر عناصر پردازشی باشد. برون‌داد می‌تواند محصول نهایی یا درون‌دادی برای یک نورون دیگر باشد.

شبکه: هر ANN از مجموعه نورون‌هایی تشکیل می‌شود که در قالب لایه‌هایی گروه‌بندی می‌شوند. شبکه متشکل از سه لایه می‌باشد؛ لایه ورودی، لایه میانی (که به آن لایه پنهان نیز گفته می‌شود) و لایه خروجی.

متغیرها با وزن و اهمیت خود به‌عنوان محرک‌های ورودی از طریق لایه ورودی وارد شبکه شده و درصورتی‌که نورون‌های عصبی به آستانه تحریک برسند، پیام عصبی ایجاد شده از طریق نورون‌های خروجی یا به‌عنوان محصول نهایی و یا به‌عنوان محرکی برای نورون‌های عصبی دیگر عمل خواهند نمود.

---

1. Processing Element



شکل ۰۱: ساختار لایه‌های شبکه عصبی مصنوعی

ساختار شبکه: همانند شبکه‌های بیولوژیکی، هر کدام ازANN ها را نیز می توان با توپولوژی‌های[1] مختلف سازماندهی کرد؛ به عبارت دیگر، نورون‌ها می‌توانند به طرق مختلف به هم متصل شوند و با هم ارتباط داشته باشند. بنابراین ANN ها در اشکال گوناگون پدیدار می‌شوند. بسیاری از عناصر پردازشی اطلاعات، محاسباتشان را به‌طور همزمان انجام می‌دهند. این پردازش موازی۲ شبیه روشی است که مغز عمل می‌کند و در مغایرت با پردازش متوالی[3] در محاسبات سنتی است (کیم[4]، ۲۰۰۶).

تابع مجموع (جمع‌کننده)

از طریق تابع مجموع می‌توان میانگین موزون تمامی عناصر درون‌داد را در هر یک از عناصر پردازشی محاسبه کرد. یک تابع مجموع، ارزش‌های مربوط به درون‌دادها (**Xi**) را در وزن‌های (**Ws**) مربوطه ضرب و سپس مجموع آن‌ها (**Y**) را محاسبه می‌کند. به‌عنوان مثال برای **n** درون‌داد **i** که به عنصر پردازشی**j** وارد می‌شوند، رابطه ۱ برقرار است:

1. Topologies
2. Parallel Processing
3. Serial Processing
4. Kim

$$Yi = \sum_{i}^{n} Xi * Wij$$ رابطه ۱)

تابع تبدیل (انتقالی)

تابع جمع‌کننده، تحریک داخلی یا سطح فعال شدن داخلی از یک نورون را محاسبه می‌کند. (گاهی اوقات به تابع جمع‌کننده، تابع فعال‌سازی می‌گویند) بر مبنای این سطح، نورون ممکن است یک برون‌داد تولید کند یا برون‌دادی تولید نکند. رابطه بین سطح فعال شدن داخلی و برون‌داد، ممکن است خطی یا غیرخطی باشد. چنین رابطه‌ای توسط تابعی بنام تابع تبدیل، توصیف می‌شود. این تابع انواع مختلفی دارد. انتخاب نوع خاصی از این تابع است که عملیات شبکه را تعیین می‌کند. یکی از توابع غیر خطی بسیار مشهور، تابع سیگموئیدی[1] نام دارد:

$$Yt = \frac{1}{1 + e^{-y}}$$ رابطه ۲)

که Yt عبارت از ارزش تبدیل‌شده (نرمال‌شده Y) است (چن، ۲۰۱۸).

یادگیری

یک شبکه عصبی مصنوعی از اشتباهاتش، یادگیری به‌عمل می‌آورد. معمولاً فرآیند یادگیری یا آموزش شامل سه وظیفه (فعالیت) می‌باشد: ۱. محاسبه کردن برون‌دادها ۲. مقایسه برون‌دادها با پاسخ‌های مطلوب (مورد نظر) ۳. تعدیل وزن‌ها و تکرار فرآیند

معمولاً فرآیند یادگیری با انتخاب تصادفی وزن‌ها آغاز می‌شود. تفاوت بین برون‌داد واقعی (Yt یا Y) و برون‌داد مورد نظر (z)، دلتا (Δ) نامیده می‌شود. در اینجا هدف حداقل کردن دلتا (یا حتی بهتر، به صفر رساندن آن) می‌باشد. کاهش دلتا از طریق ایجاد تغییرات تدریجی در وزن‌ها انجام می‌گردد. پردازش اطلاعات با استفاده از شبکه عصبی مصنوعی شامل تجزیه‌وتحلیل الگوهای فعالیت (شناسایی الگو) می‌باشد.

1. Sigmoid function

یادگیری با ناظر

در یادگیری با ناظر به قانون یادگیری مجموعه‌ای از زوج داده‌ها به نام (داده‌های یادگیری) $\{(\underline{p}^i, \underline{t}^i); i = 1,2,\ldots, L\}$ می‌دهند که در آن $\underline{p}^i$ ورودی به شبکه و $\underline{t}^i$ خروجی مطلوب شبکه برای ورودی $\underline{p}^i$ است. پس از اعمال ورودی به شبکه عصبی در خروجی شبکه $\underline{a}^i$ با $\underline{t}^i$ مقایسه شده و سپس خطای یادگیری محاسبه و از آن جهت تنظیم پارامترهای شبکه استفاده می‌شود. به‌گونه‌ای که اگر دفعه بعد به شبکه همان ورودی $\underline{p}^i$ اعمال شود خروجی شبکه به $\underline{t}^i$ نزدیک‌تر می‌گردد. پارامترهای شبکه عصبی توسط دو سیگنال ورودی و خطا تنظیم می‌شوند. به این صورت که پس از چند تکرار، الگوریتم یادگیری به پارامترهایی در فضای پارامترهای شبکه، همگرا می‌شوند (یانگ، ۲۰۱۹).

یادگیری بدون ناظر

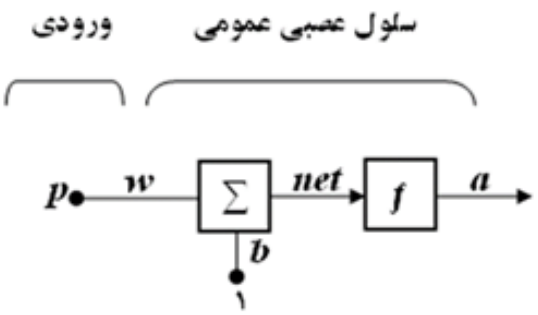
در یادگیری بدون ناظر یا یادگیری خود سازمانده[1] پارامترهای شبکه تنها توسط پاسخ سیستم اصلاح و تنظیم می‌شوند. به عبارتی دیگر تنها اطلاعات دریافتی از محیط به شبکه را بردارهای ورودی تشکیل می‌دهند و بردار جواب مطلوب به شبکه اعمال نمی‌شود. در این حالت، شبکه وزن‌های خود را طوری تنظیم می‌کند که داده‌هایی که بیشترین شباهت را نسبت به هم دارند به خروجی یا طبقه یکسانی اختصاص داده می‌شوند. شبکه‌های خود سازمانده نمونه‌ای از این یادگیری را انجام می‌دهد که معروف‌ترین آن‌ها شبکه‌های کوهونن[2] هستند (فوست[3]، ۲۰۱۹).

1. Self -Organized
2. Kohonen
3. Fausett

مدل ریاضی نورون

- مدل تک ورودی

اسکالرهای **p** و**a** به ترتیب ورودی و خروجی می‌باشند.



شکل ۲: مدل سلول عصبی تک ورودی

میزان تأثیر p روی a بوسیله مقدار اسکالرw تعیین می‌شود. ورودی دیگر که مقدارش ثابت 1 است در جمله تورش b ضرب می‌شود و سپس با wp جمع می‌شود، این حاصل جمع ورودی خالص n برای تابع محرک (تابع تبدیل) f خواهد بود. بدین ترتیب خروجی نورون با معادله ۳ تعریف می‌شود.

$$a = f(wp+b)$$ رابطه ۳)

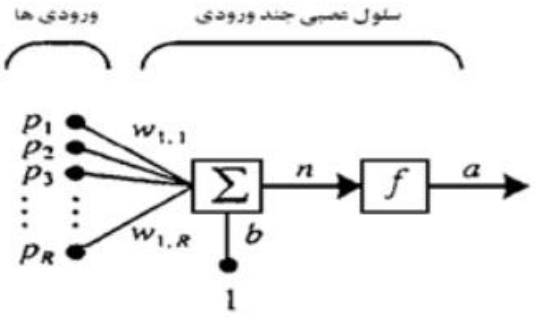
با مقایسه این نورون تک ورودی با یک نورون بیولوژیکی عملا w معادل شدت سیناپس[1] جمع‌کننده و تابع محرک معادل هسته سلول و سیگنال خروجی نورون a معادل سیگنال گذرنده از اکسون خواهند بود. نکته‌ای که باید به آن توجه شود، اهمیت و تأثیر جمله تورش b است. این جمله را می‌توان مانند وزنه w در نظر گرفت با این تصور که میزان تأثیر ورودی ثابت 1 را روی نورون منعکس می‌سازد. اهمیت جمله b بیشتر جهت افزایش محدوده پیش‌بینی نورون‌ها به‌کار می‌رود. باید توجه داشت که پارامترهای b و w قابل تنظیم هستند و تابع محرک f نیز توسط طراح انتخاب می‌شود. بر مبنای انتخاب f و نوع

1. Synapse

الگوریتم یادگیری پارامترهای b و w تنظیم می‌شوند. یادگیری بدین معنی است که b وw طوری تغییر می‌کنند تا رابطه ورودی و خروجی نورون با هدف خاصی مطابقت نماید.

- مدل چند ورودی

عموماً یک نورون، بیش از یک ورودی دارد شکل زیر یک مدل نورون با **R** ورودی را نمایش می‌دهد (چونگ، ۲۰۱۷).



شکل ۳: مدل سلول عصبی چند ورودی

- نگاشت خودسازمانده کوهونن

شبکه‌های خود سازمانده به‌طور گسترده‌ای در داده‌کاوی، نمایش فضاهای پیچیده و خوشه‌بندی فضاهای با ابعاد بالا و به‌طور خاص در پردازش تصویر، کنترل فرآیند، مدیریت پروژه، آنالیزهای مالی و تشخیص‌های صنعتی و پزشکی مورد استفاده قرار می‌گیرند (چن، ۲۰۲۰).

این شبکه از انواع شبکه‌های عصبی با قابلیت یادگیری با ناظر و بدون ناظر[1] می‌باشد که در تحلیل فضاهای پیچیده، توانایی زیادی دارد. این مدل از شبکه‌های عصبی اولین بار توسط کوهونن (۱۹۸۱) و با الگوبرداری از عصب‌های شبکیه چشم، معرفی و در سال ۱۹۸۴ برای تشخیص صدا به‌طور عملی مورد استفاده قرار گرفت (کوهونن، ۱۹۹۸).

قابلیت این شبکه در ارائه خروجی در قالب نقشه‌های گرافیکی گویا و قابل فهم برای مدیران سازمان‌ها، توانایی نمایش روابط خطی و غیرخطی بین متغیرها، قدرت بالا در دسته‌بندی دادگان، عدم حساسیت

1. Unsupervised

شبکه به تعداد دادگان تعلیم و حساسیت کم این نوع شبکه‌ها به وجود نویز در دادگان تعلیم، این ابزار را به ابزاری کارآمد جهت استفاده در مدل‌سازی تبدیل نموده است. (یانگ و چونگ، ۲۰۱۸)

یک نقشهٔ خود سازمان‌دهنده نوعی‌شبکه عصبی مصنوعی است که با استفاده از یادگیری هدایت‌نشده یا هدایت‌شده برای تولید نمایش‌گر با ابعاد کم نمونه‌های آموزشی استفاده می‌شود، درحالی‌که ویژگی‌های توپولوژیکی فضای ورودی را حفظ می‌کند. ازاین‌رو، نگاشت خود سازمان‌دهنده برای به تصویر کشیدن مناظری با ابعاد کم از داده‌های با ابعاد زیاد جهت مقیاس‌گذاری چند بعدی مفید است. این نقشه‌ها می‌توانند طبقه‌بندی بردارهای ورودی را یاد بگیرند به‌گونه‌ای‌که آن‌ها را در فضای ورودی گروه‌بندی نمایند. نورون‌های همسایه در نقشه خود سازمانده تشخیص بخش‌های همسایهٔ فضای ورودی را یاد می‌گیرند. بنابراین، نقشه‌های خودسازمانده، هم توزیع و هم توپولوژی بردارهای ورودی را در زمان آموزش یاد می‌گیرند (یانگ، ۲۰۱۸).

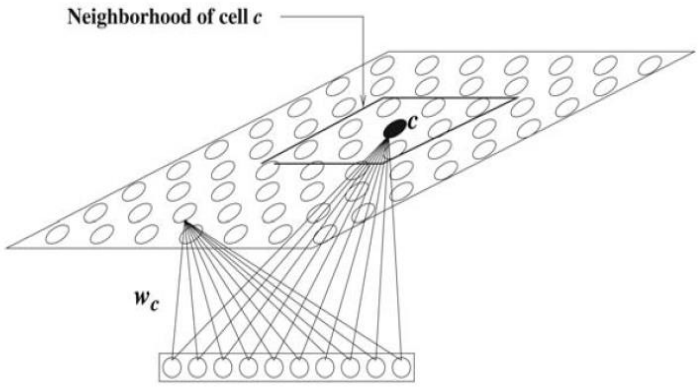
نورون‌ها در لایه‌ای از یک نگاشت خود سازمان‌دهنده اصولاً در موقعیت‌های فیزیکی براساس یک تابع توپولوژی آرایش داده می‌شوند. نورون‌ها در توپولوژی‌های شطرنجی، شش ضلعی یا تصادفی آرایش می‌یابند. فواصل بین نورون‌ها از موقعیت‌هایشان با یک تابع فاصله محاسبه می‌شوند.

شبکهٔ نگاشت خود سازمان‌دهنده معمولاً دو لایه از گروه‌ها دارد: یک لایهٔ ورودی و یک لایهٔ خروجی. نورون‌ها در لایهٔ خروجی در یک شبکه مرتب می‌شوند و از همسایه‌هایشان در این شبکه تأثیرپذیر هستند. هدف خوشه‌بندی خودکار نمونه‌های ورودی به روشی است که نمونه‌های مشابه توسط نورون‌ها خروجی مشابهی نشان داده می‌شوند (چونگ[1]، ۲۰۱۸).

ازآن‌جاکه هر یک از متغیرهای مشخصه به هر نورون خروجی توسط یک اتصال وزنی مرتبط می‌شود، هر نورون خروجی $j\ (j=1,...,K)$ تعداد مشابهی وزن w_j هم بسته به تعداد متغیرهای ورودی دارد که با وزن‌های اولیهٔ تصادفی شروع کرده و تعدیل وزن‌هایش را براساس نمونه‌های ورودی به‌صورت ذیل یاد می‌گیرد. زمانی که یک نمونهٔ ورودی، x_i به شبکهٔ نگاشت خود سازمان‌دهنده وارد می‌شود، نورون‌ها فاصلهٔ بین بردارهای وزنی $w_j=(w_{j1}, w_{j2},..., w_{jp},..., w_{jp})$ و ورودی $x=(x_{i1}, x_{i2},..., x_{ip},..., x_{ip})$ را محاسبه می‌کنند. سپس، نورون با حداقل فاصله به نام برند با استفاد از فرمول ۳ تعیین می‌شود. نمونه‌ای از معماری یک شبکه نگاشت خود سازمان‌دهنده در شکل ۴ نشان

1. Chong

داده شده است، در این شکل اجزا و ساختار آن به‌صورت شبکه دو لایه‌ای با اسکالرهای ورودی و خروجی نشان داده شده است.



شکل ٤: معماری نمونه یک شبکه نگاشت خودسازمان‌دهنده (چونگ،۲۰۱۸)

$$Min\ D_j = \sqrt{\sum_{p=1}^{p} \left[x_{ip} - w_{jp}\right]^2} \qquad \text{رابطه ۳)}$$

که w_{jp} وزن j امین نورون پیوند داده شده به P امین متغیر است (یانگ[1]، ۲۰۱۸).

پیشینه پژوهش‌های خارجی

چوا[2] (۲۰۱۹) برای شناسایی تراکنش‌های مشکوک منجر به پول‌شویی از روش خوشه‌بندی بر درخت پوشای مینیمم بهبود یافته که براساس معیار عدم شباهت ساخته می‌شود، استفاده کرد. در این تحقیق از یک مجموعه داده تجاری جهت پیاده‌سازی نتایج استفاده شده است. مزیت این روش ارائه یک پارامتر نامتجانس جهت اندازه‌گیری تفاوت درجات پرت و سپس طراحی الگوریتم کشف پول‌شویی است.

1. Yang
2 Chao

جولوم و همکاران[1] (۲۰۲۰) با به‌کارگیری الگوریتم‌های یادگیری ماشین جهت شناسایی فعالیت‌های پول‌شویی با استفاده از الگوریتم خوشه‌بندی (k-mean و BIRCH) توانسته‌اند به‌صورت کارآمدتر تراکنش‌های مشکوک را شناسایی کنند. بدین ترتیب که برای هر داده مالی از گره تا برگ را جستجو کرده و در برگ، درباره اینکه در گروه بماند یا خیر تصمیم گرفته می‌شود که البته روی دامنه محدودی قابل اجراست. در این تحقیق از یک مجموعه داده خاص جهت پیاده‌سازی و آزمایشات استفاده شده که به نام و منبع آن اشاره‌ای نشده است.

در مقاله لو[2] (۲۰۱۴) با استفاده از الگوهای تکرار و قواعد انجمنی برای آشکارسازی فعالیت‌های ضد پول‌شویی (AML) و تمرکز روی کشف تراکنش‌های مشکوک در جریان تراکنش‌های تجاری کار شده است. در این روش، به تراکنش‌های تجاری به‌عنوان جریانی از داده‌ها توجه شده و روش به‌دست آوردن پویایی برای آشکارسازی الگوهای مشکوک روی جریان تراکنش‌ها به‌کار برده است؛ مخصوصاً یک الگوریتم دسته‌بندی (کلاس‌بندی) براساس قوانین ارتباط چند کلاسی روی جریان‌های داده‌ها با ایجاد یک درخت الگوی تکرار[3] برای بهبود زمان و فضای کارآمد و سپس کاهش تکرار قوانین بوسیله حد هاؤفدینگ[4] (محدود کردن جمع تعدادی متغیر تصادفی مستقل کران‌دار) روی جریان‌های پویای داده‌ها پیشنهاد شده است. از مزایای این روش، مقیاس‌پذیری در مجموعه داده‌های بزرگ می‌باشد. در این تحقیق، به مجموعه داده استفاده شده در پیاده‌سازی الگوریتم اشاره نشده است. ضمن آن که این روش در مقیاس پایین، از دقت مناسبی برخوردار نیست.

شیخ و همکاران[5] (۲۰۱۸) در مقاله خود مدلی را جهت شناسایی اجتماع و روابط بین تراکنش‌ها و مشتریان با استفاده از تحلیل شبکه‌های اجتماعی ارائه دادند. با کمک این مدل، گروه‌ها و باندهای مافیایی که نقش اصلی در فعالیت‌های پول‌شویی بازی می‌کنند، شناسایی می‌شوند. در این روش یک شبکه اجتماعی با استفاده از گراف شبکه‌ای معنایی و فعالیت‌های شبکه‌هایی از قبیل درجه مرکزیت و خوشه‌بندی و غیره جهت شناسایی مشتریان مشکوک ساخته می‌شود. درجه مرکزیت جهت شناسایی سرشاخه‌ها در گروه‌های اجتماعی کاربرد دارد و هدف از خوشه‌بندی، جداسازی گروه‌هایی با شخصیت‌های مشابه و یا مبتنی بر شاخص‌های ویژه از شاخه‌ها که با دیگر خوشه‌ها نسبت دارند، می‌باشد. شناسایی گروه‌های

---

1. Jullum et al
2. Luo
3. FP-Tree
4. Hoeffding
5. Shaikh et al

مشکوک در شبکه‌های اجتماعی از مزایای این روش تلقی می‌شود ولی این مورد که هیچ ارتباط خاصی در تحلیل بین حساب‌های بانکی و عملیات انجام شده دیده نمی‌شود از نقاط ضعف این روش می‌باشد.

پیشینه پژوهش‌های داخلی

در مقاله کوثری و همکاران (۲۰۱۲) با به‌کارگیری الگوریتم‌های تصمیم‌گیری جهت کشف رفتارهای مشکوک امکان مدل‌سازی رفتار کاربران در پنج دسته مختلف بوجود آمده که با دقت بیشتری نوع رفتار کاربر را پیش‌بینی می‌کند؛ ولی این نظریه در حد تئوریک بوده و عملیاتی نشده است. داده‌های آموزشی این پژوهش به تعداد ۱۰۰،۰۰۰ رکورد اولیه به‌عنوان نمونه واقعی از یک بانک خصوصی دریافت شده است. این مقاله با پژوهش محقق دارای وجوه تمایز و تشابهی به شرح ذیل می باشد:

وجه تشابه: ۱- هردو تحقیق در زمینه کشف رفتارهای مشکوک و تحلیل رفتار آن‌ها در بانک، به‌کار برده شده است. ۲- تعداد داده‌های هردو پژوهش تقریبا یکسان بوده است. ۳- هر دو محقق از تکنیک‌ها و الگوریتم‌های شبکه هوشمند عصبی استفاده نموده‌اند.

وجه تفاوت : ۱- داده‌های به‌کار گرفته شده در تحقیق آقای کوثری آموزشی بوده درحالی‌که داده‌های این پژوهش دارای الگو (مشتریان مظنون به پول‌شویی و معرفی شده از سوی FIU) و مبتنی‌بر عملیات تراکنشی مشتریان بوده‌اند. ۲- الگوریتم به‌کار گرفته شده در این تحقیق، نگاشت خود سازمان‌دهنده از نوع یادگیری با ناظر بوده درحالی‌که الگوریتم به‌کار گرفته شده در تحقیق آقای کوثری و همکاران از نوع درخت تصمیم بوده است. ۳- در تحقیق آقای کوثری داده‌ها در پنج خوشه قرار می‌گیرند، درحالی‌که در این پژوهش در چهار خوشه قرار گرفته است.

در مقاله حیدری‌نیا و همکاران (۲۰۱۴) با مطالعه‌ی پروفایل کاربران، تعدادی از ویژگی‌ها شامل معاملات مالی حجیم در مناطق ریسک‌پذیر از دیدگاه پول‌شویی، فعال شدن حساب راکد با مقادیر قابل توجه و غیره استخراج شدند. آموزش شبکه با استفاده از طراحی یک سیستم فازی، توسعه‌ی یک سیستم استنتاج عصبی – فازی وفقی و اضافه کردن بردار ویژگی‌های کاربران به آن انجام شده است. سپس خروجی شبکه می‌تواند میزان ریسک رفتار کاربر را تعیین نماید. این روش ترکیبی با توجه به صحت و سرعت قابل قبول برای مجموعه داده‌های حجیم انتخاب مناسبی خواهد بود. روش پیشنهادی مبتنی‌بر ANFIS[1] است که یک روش ترکیبی فرا ابتکاری می‌باشد؛ چراکه ویژگی‌ها و قابلیت‌های مناسبی دارد. سیستم استنتاج عصبی- فازی وفقی، الگوریتم‌های منطق فازی و شبکه عصبی را به‌کار می‌گیرد تا بین

1. Adaptive Network –based Fuzzy System

حوزه‌های ورودی و خروجی نگاشت غیر خطی ایجاد کند. با استفاده از قدرت زبان فازی و قدرت عددی شبکه‌های عصبی، ANFIS در مدل‌سازی فرایندهای پیچیده بسیار قدرتمند خواهد بود. سیستم فازی با استفاده از ویژگی‌های مستخرج از داده‌های بانکی و برخی از قواعد فازی، طراحی شده و با استفاده از الگوهای بهینه‌ی ایجاد شده در یک سیستم ANFIS، پول‌شویی هوشمندانه را مشخص می‌کند. دقت ناشی از این سیستم پیشنهادی ۹۶ درصد است. پیاده‌سازی این روش بر روی مجموعه داده‌های یک بانک (به نام آن اشاره‌ای نشده است) صورت گرفته است. از مزایای این روش می‌توان به مناسب بودن برای داده‌های حجیم، سرعت قابل قبول و دقت بالا اشاره نمود.

روش‌شناسی تحقیق:

در این تحقیق ابتدا با مطالعه ادبیات موضوع از چندین مقاله معتبر، متغیرهای مهم برای دسته‌بندی مشتریان مظنون به پول‌شویی تعیین شده است و چهار مجموعه از داده‌ها شامل داده‌های تراکنش‌های مالی و داده‌های جمعیت شناختی مشتریان و نیز مجموعه داده‌های سایر مراجع یعنی واحد اطلاعات مالی(**FIU**)، قواعد ۲۸ گانه بانک مرکزی و سرویس‌های بانکی، جهت خوشه‌بندی مشتریان مورد استفاده قرار گرفته است. در ادامه تحقیق پس از تعیین متغیرها، با استفاده از روش خود سازمان‌دهنده شبکه‌های عصبی، مشتریان طبقه‌بندی شده و با استفاده از قوانین وابستگی، رابطه مشتریان مظنون به پول‌شویی با دیگر گروه‌های مشتریان تعیین می‌گردد.

در رتبه‌بندی به دسته‌بندی مشتریان براساس نرم‌افزار Spss Clementin و ماژول Som[1] پرداخته شده است. پس از تعیین دسته‌های مشتریان و شناسایی مشتریان مظنون به پول‌شویی، استراتژی‌هایی که بانک در برخورد با هر یک از دسته‌ها باید به‌کار برد، مورد بررسی قرار گرفته است. نقشه خود سازمان‌دهنده یک الگوریتم یادگیرنده بدون نظارت و یا باناظر می‌باشد که داده‌های چند بعدی را به‌عنوان بردارهای ورودی مشابه به مکان‌های مشابه از نقشه عصبی ارتباط می‌دهد و قوانین همبستگی را برای پیدا کردن ارتباطات بین اقلام یا ویژگی‌هایی که به‌صورت همزمان دریک پایگاه داده رخ می‌دهند، استفاده می‌کند. از طریق رتبه‌بندی رفتاری مشتریان به پیش‌بینی گروه‌های مشتریان مظنون به پول‌شویی و پیش‌بینی مدل‌های رفتاری آنان پرداخته شده است. پس از شناسایی مشتریان مظنون به پول‌شویی، نماینده هر مورد در هر گروه از مشتریان، خدمات دریافت شده را به‌عنوان ابزاری برای پایه‌گذاری برنامه‌های استراتژیک بانک در نحوه تعامل با این دسته از مشتریان شرح می‌دهد. تخصیص رتبه‌های رفتاری

1. Self-organizing map

به هر مشتری و گروه‌بندی این مشتریان در دسته‌های با ارزش مرتبه‌ای برابر توسط تکنیک‌های فوق صورت می‌پذیرد. در این تحقیق از شاخص های اسمی رفتار تراکنش مالی و شاخص‌های عملکردی برای رتبه‌بندی مشتریان استفاده شده است.

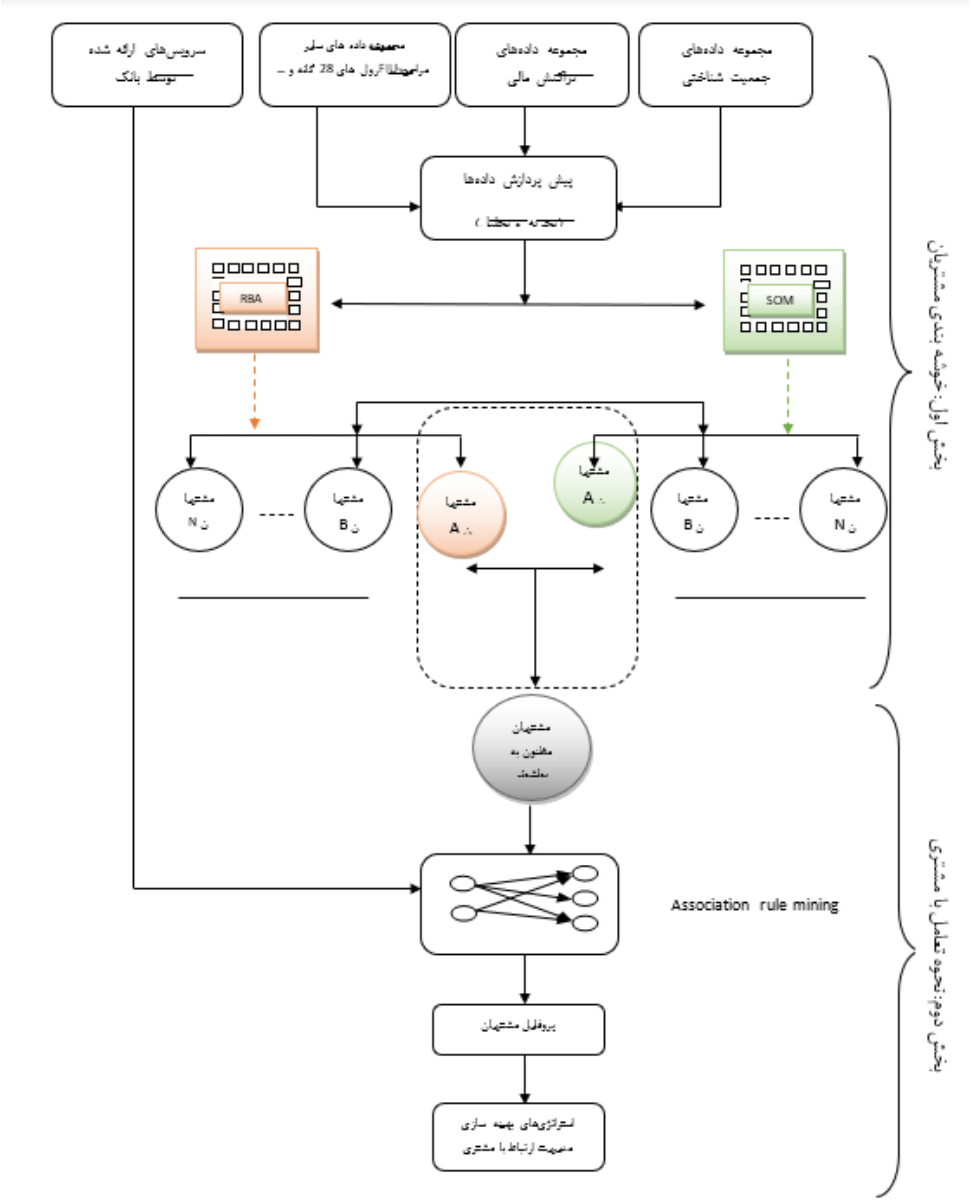
سرانجام با استفاده از شاخص‌های عملکردی تعیین شده می‌توان درستی و مناسب بودن نتایج برای بانک در جهت تعیین استراتژی‌های تعامل با این دسته از مشتریان را مشخص ساخت. این تحقیق قصد دارد از طریق به‌کارگیری مدل‌های رتبه‌بندی رفتاری برای تجزیه‌وتحلیل رفتار مشتریان که شامل ضرورت پیش‌پردازش از مجموعه داده‌های واقعی (داده‌های حساب‌ها و تراکنش‌های مالی) است و همچنین با بکارگیری تکنیک‌های شبکه عصبی مانند نقشه خود سازمان‌دهنده و نیز با به‌کارگیری تکنیک قوانین همبستگی و تشکیل پروفایل مشتریان به رتبه‌بندی مشتریان با ریسک‌های متفاوت از منظر پول‌شویی بانک بپردازد. در نهایت، نتایج به‌دست آمده در روش مبتنی بر شبکه‌های هوشمند عصبی با داده‌های یکسان با شرایط موجود در بانک ملت جهت رتبه‌بندی ریسک محور مشتریان از منظر پول‌شویی (RBA) مورد مقایسه قرار خواهد گرفت.

در این تحقیق از یک الگو و مدل دو مرحله‌ای برای تجزیه و تحلیل رفتار مشتریان با استفاده از پایگاه داده‌ای درخصوص ویژگی‌های مشتریان و داده‌های مالی آن‌ها استفاده شده است.

موضوع بحث شامل: پیش پردازش داده‌ها، مدل‌های رتبه‌بندی رفتاری (نگاشت خود سازمان‌دهنده و رتبه‌بندی مشتریان از منظر ریسک پول‌شویی در بانک ملت)، گام دو مرحله‌ای درخصوص رتبه‌بندی رفتاری مشتریان می‌شود.

ویژگی اصلی از مدل‌های رتبه‌بندی رفتاری دو مرحله‌ای: یکی، مرحله نقشه خود سازمان‌دهنده و دیگری کشف قوانین همبستگی می‌باشد.

شکل ۵ نشان‌دهنده چارچوب مفهومی تحقیق می‌باشد که برای دستیابی به اهداف تحقیق طراحی گردیده است. این شکل نشان می‌دهد که دو جز بخش‌بندی مشتری و تشکیل پروفایل مشتریان، اهداف اصلی تحقیق می‌باشند.



شکل ۵: به‌کارگیری تکنیک‌های شبکه عصبی و رتبه‌بندی مشتریان از منظر ریسک پول‌شویی جهت شناسایی مشتریان هدف

قلمرو زمانی، مکانی و موضوعی تحقیق

قلمرو زمانی تحقیق: پژوهش حاضر مربوط به دوره زمانی شش ماهه ۱۴۰۱/۰۷/۰۱ تا ۱۴۰۱/۱۲/۲۸ می‌باشد. یعنی بررسی پایگاه داده‌های مربوط به حساب‌های کلیه مشتریان حقیقی نزد بانک ملت در مقطع زمانی فوق را شامل می‌شود.

قلمرو مکانی تحقیق: شامل شعب ممتاز بانک ملت در مدیریت شعب تهران و شهرستان ها (مشتمل بر ۵۶ شعبه) می‌باشد. علت انتخاب شعب ممتاز بعنوان قلمرو مکانی تحقیق این است که توزیع پراکندگی مشتریان هدف بانک در این شعب بیشتر می‌باشد. از اینرو تابع هدف قلمرو مکانی محقق شعب ممتاز مناطق تهران و شهرستان‌ها بخصوص مناطق آزاد تجاری و صنعتی و همچنین مناطق پرخطر از سوی مرکز اطلاعات مالی [1](FIU) می‌باشد.

فرضیه‌ها یا سؤال‌های تحقیق

با توجه به اینکه در این پژوهش قصد بر آن است تا با استفاده از تکنیک‌ها و الگوریتم‌های داده‌کاوی و (**RBA**) به تحلیل الگوی رفتاری مشتریان مظنون به پول‌شویی پرداخته شود و با کشف این الگوها و روندهای رفتاری مشتریان، مدلی جهت طبقه‌بندی مشتریان ارائه گردد، لذا این پژوهش فاقد فرضیه می‌باشد. سؤالات تحقیق به قرار زیر می‌باشد:

۱- مدل‌های رتبه‌بندی رفتاری مشتریان چگونه به کشف الگوهای رفتاری مشتریان و شناسایی مشتریان مظنون به پول‌شویی و یا مشتریان با ریسک بالای پول‌شویی کمک می‌کند؟

۲- نگاشت خود سازمان دهنده چگونه می‌تواند، طبقه بندی بهتری را نسبت به مفهوم (RBA) در خصوص رتبه‌بندی این دسته از مشتریان در بانک ملت ارائه نمایند؟

۳- استراتژی‌های مناسب بانک جهت برخورد با هر یک از دسته‌های مشتریان با درجات ریسک پول‌شویی متفاوت (بالا- متوسط و کم) چگونه است؟

نوع روش تحقیق

براساس اهداف پژوهش، تحقیق حاضر از نوع تحقیق کاربردی و براساس نحوه گردآوری داده‌ها، روش مورد استفاده در آن، از نوع توصیفی- تحلیلی می‌باشد.

---

1. Financial Intelligence Unit; FIU

روش گردآوری اطلاعات و داده‌ها

در این تحقیق از مطالعات کتابخانه‌ای به طور گسترده‌ای استفاده شده است.

پایگاه اطلاعاتی مشتریان بانک ملت: در این روش از اطلاعات ذخیره شده در سیستم‌های اطلاعاتی بانک ملت استفاده شده است. بنابراین برای جمع‌آوری این داده‌ها می‌بایست پایگاه داده در سرورهای شرکت بهسازان ملت و همچنین شرکت به پرداخت ملت، مورد تحلیل قرار گیرد.

جامعه آماری

در این تحقیق، تجزیه‌وتحلیل بر روی داده‌های خام پایگاه داده بانک ملت که شامل داده‌های خام مشتریان حقیقی شعب ممتاز (الف و ب) تهران و مناطق استان‌ها می‌باشد، صورت پذیرفت. این داده‌ها دارای حجم بالایی بوده که در بازه تاریخ شش ماه ذخیره گردیده‌اند. در این تحقیق با استفاده از روش نمونه‌گیری تصادفی، داده‌های جمعیت شناختی و تراکنش‌های مالی مشتریان حقیقی بانک که در پایگاه داده‌های بانک موجود می‌باشند، انتخاب شده و مورد بررسی قرار گرفتند.

متغیرهای تحقیق

داده‌های جمعیت شناختی: این متغیرها شامل جنسیت، سن، تحصیلات، وضعیت تاهل، شغل و غیره می‌باشد.

سرویس‌های ارائه شده توسط بانک: سرویس‌های مورد مطالعه در این تحقیق در حوزه‌های مختلف بانکداری مانند (ارزی، اعتباری، سرویس‌های بانکداری الکترونیک و سایر خدمات) می‌باشد. این سرویس‌ها بسترهای تعاملی مناسبی را بین بانک و مشتریان فراهم آورده است. در هرکدام از این کانال‌های ارتباطی، طیف وسیعی از خدمات بانکی ارائه می‌گردد. هر کدام از این سرویس‌ها براساس سطح و کلاس ریسک پول‌شویی مشتری، می‌تواند از سوی مشتری استفاده گردد.

تراکنش‌های مالی مشتریان: این متغیر شامل مبالغ برداشت یا واریز مشتری در بازه زمانی مورد تحقیق است. برای بررسی الگوی رفتاری مشتری و طبقه‌بندی مشتریان بانک در تحقیقات مختلف از روش‌های متنوعی استفاده شده است. ازآنجاکه تراکنش‌های مالی مشتری می‌تواند به‌عنوان شاخص خوبی برای بررسی الگوی رفتاری مشتری باشد، در این تحقیق از این متغیر برای رفتارشناسی مشتری استفاده شده است. این متغیرها در پایگاه اطلاعاتی وجود دارند و از طریق پردازش داده‌های تراکنشی مشتریان محاسبه می‌شوند.

در این پژوهش از مدل Max-Min جهت نرمال‌سازی داده‌ها استفاده شده است. نرمال‌سازی داده‌ها می‌تواند موجب بهتر شدن مدل‌سازی شود و از مشکلات عددی پیشگیری کند. نرمال‌سازی Max-Min، یک الگوریتم مقیاس‌گذاری خطی است؛ به‌طوری که محدودهٔ ورودی اصلی را به یک محدودهٔ داده‌ای جدید (نوعاً ۰–۱) تبدیل می‌کند. در این پژوهش از نرم افزارSpss Clemantin استفاده شده است.

جدول ۲: نمونه متغیرهای تحقیق

نوع	نام فارسی	ستون
smallint	نوع کلاس رتبه مشتری (۱ بالا، ۲ متوسط، ۳ پائین)	CLASSTYPE
smallint	امتیاز مظنونین به پول‌شویی	SUSPECTEDSCR
smallint	امتیاز سهام‌داران بیش از ۵٪	MNGSTOCKSCR
smallint	امتیاز وکیل چند مشتری بطور همزمان	LAWYERSCR
smallint	امتیاز ذینفع واحد	BENUNITSCR
smallint	امتیاز صندوق امانات	SAFEBOXSCR
smallint	امتیاز فرد داخلی در معرض مخاطرات سیاسی	IRNPOLRSKSCR
smallint	امتیاز فرد خارجی در معرض مخاطرات سیاسی	FRNPOLRSKSCR
smallint	امتیاز مشتری مهجور	SEPERATEDSCR
smallint	امتیاز مبدأ واریز وجه داخل کشور	IRNSRCDEPSCR

نمونه‌ای از داه‌های نرمال و وزن داده شده که با استفاده از تکنیک‌های مختلف همچون تلخیص، تجمیع وغیره پیش پردازش‌های لازم بر روی آن‌ها صورت پذیرفته است.

جدول ۳: نمونه‌ای از داده‌های پردازش شده

ID	Birthdate	Age	Gender	Education	Job	Birth-L	BANKAmel_N	HesabMaghsad_N	CTRHoolshodel_N
1	1343/10/14	۵۴	F	L	تعمیرات	تهران	Melli	1308629	1996246
2	1359/11/13	۳۸	F	Elementary	قطعات یدکی/ ابزار	تهران	Sepah	1234567112	1996247
3	1351/01/31	۴۱	F	Diplom	رستوران	شهید	Maskan	1234532001	1996248
4	1342/12/17	۴۰	M	Cycle	سایرخدمات	فراهان	Sadarat	12456964668	1996249
5	1354/07/17	۴۸	F	FL	Unimployer	قم	Dey	7286410	1996250
6	1344/08/15	۵۱	F	Cycle	نظامی وانتظامی	جلیلیار	Tejarat	1265231111	1996251
7	1355/07/19	۴۱	F	Fdip	خودرو	سازی	Refah	12161920	1996252
8	1347/05/29	۵۵	M	Fdip	مشاور املاک وساختمان	اهسل	Shahr	17502000	1996253
9	1359/12/12	۳۴	F	Fdip	حمل و نقل	آباده	Sarmaye	17822568	1996254
10	1353/08/12	۴۹	M	L	Unimployer	زابل	Parsian	7854236	1996255
11	1328/04/27	۷۹	F	Fdip	خودرو	تهران	Pasargad	128640	1996256
12	1350/06/01	۵۳	F	Fdip	مشاور املاک وساختمان	زرند	Keshavarzi	1308629	1996257
13	1332/01/28	۷۰	F	FL	رستوران	تهران	Melli	86624171	1996258
14	1349/09/05	۵۳	M	Fdip	پولی و مالی	الشترخوره آباد	Sina	4714582	1996259
15	1351/10/20	۴۱	F	L	معلم/استاد دانشگاه	دورود	Maskan	12145236	1996260
16	1335/03/30	۴۷	F	FL	پولی و مالی	کاشان	Sadarat	1456 82566	1996261
17	1346/09/13	۵۴	M	L	حمل و نقل	تهران	Tejarat	45482124	1996262
18	1357/07/25	۴۵	F	L	خودرو	دورود	Dey	1232 16548	1996263
19	1352/01/22	۵۰	M	Diplom	پزشکی و دامپزشکی	وشت	Refah	1234786	1996264
20	1360/01/27	۴۲	M	Diplom	وکالت/مشاورخدمات مهندسی	سمیرم	Shahr	11141523	1996265
21	1358/11/18	۴۴	M	Diplom	متخصص	نائین	Sarmaye	145225	1996266
22	1359/07/11	۴۲	M	FL	لوازم الکترونیک	زابل	Parsian	1234568	1996267
23	1345/05/19	۵۷	M	L	Manager	سنقر	Keshavarzi	9215632	1996268
24	1355/05/23	۴۷	M	Elementary	Manager	فارسان	Pasargad	1452382	1996269
25	1358/08/25	۴۴	M	Fdip	پزشک و متخصص	کاشان	Sadarat	12347852	1996270
26	1350/10/19	۵۳	F	FL	شهوری	بهشهر	Dey	46865	1996271
27	1337/01/10	۴۵	M	Diplom	پزشکی و دامپزشکی	گرگان	Tejarat	41042032	1996272
28	1355/06/18	۴۱	M	Cycle	عالی	قم	Refah	7485963	1996273
29	1346/11/12	۵۴	F	FL	قطعات یدکی/ ابزار	نجف آباد	Shahr	7415921	1996274
30	1334/07/26	۴۸	F	Fdip	Unimployer	کاشمر	Sarmaye	45652562	1996275
31	1341/08/22	۴۱	F	Elementary	مشاور املاک وساختمان	سازی	Parsian	74189633	1996276
32	1344/08/01	۵۱	F	Diplom	رستوران	تربت‌حیدریه	Pasargad	4715632	1996277
33	1331/03/05	۷۱	M	Diplom	سایرخدمات	تهران	Sina	1452365	1996278
34	1363/08/20	۱۹	F	Cycle	معلم/استاد دانشگاه	تهران	Melli	4863896211	1996279
35	1360/06/12	۴۲	F	Cycle	متخصص	قائم‌شهر	Sepah	1452 328362	1996280

در این تحقیق مطالعه بر روی داده‌های خام مشتریان حقیقی شعب ممتاز (مشتمل بر ۵۶ شعبه ممتاز الف و ب، با شناسه‌های چهار رقمی و رقم کنترلی واحد) مدیریت شعب تهران و شهرستان‌ها در بازه زمانی شش ماهه صورت پذیرفته است. تعداد کل داده‌های دریافتی از پایگاه داده‌های شرکت‌های تابعه بانک ملت (بهسازان ملت، به‌پرداخت) مشتمل بر ۹۲،۶۸۶ رکورد بوده که پس از به‌کارگیری تکنیک‌های آماده‌سازی داده‌ها همچون (تجمیع، تلخیص، یکپارچه‌سازی، نرمال‌سازی و غیره) تعداد ۵۰،۰۰۰ رکورد اطلاعاتی جهت انجام پژوهش حاضر فراهم گردیده است.

یافته های پژوهش

نتایج حاصل از خوشه‌بنـدی داده‌های مربوط به مشتریان حقیقی (۵۰،۰۰۰ رکورد اطلاعاتی) براساس روش **RBA** به‌شرح جدول ۴ می‌باشد:

جدول ۴: نتایج حاصل از رتبه‌بندی مشتریان براساس مدل RBA در بانک ملت

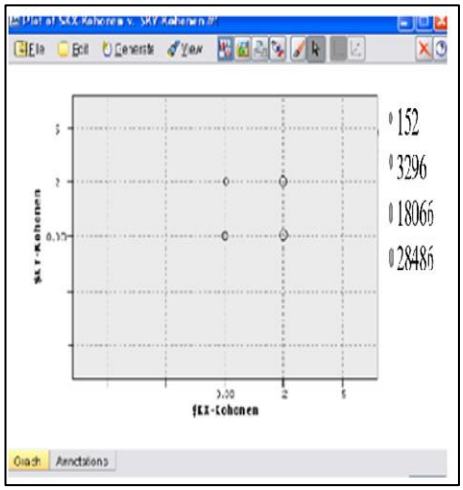
Cluster	Cluster type	Range Of Points	Number of Cases in each Cluster	%Percent
۱	High Risk	۱۰۰-۶۷	۱۶۸	۰/۳
۲	Middle Risk	۶۶-۳۳	۲/۳۹۲	۴/۷
۳	Low Risk	۳۳-۱	۱۱/۰۵۲	۲۲
۴	Free Risk	۰	۳۶،۳۸۸	۷۳
	Valid		۵۰/۰۰۰	۱۰۰/۰۰
	Missing		۰۰۰	

نتایج حاصله در خوشه‌بندی مشتریان در مدل RBA بانک ملت نشان از وجود ۱۶۸ مشتری مظنون به پول‌شویی در کلاس A (۰.۳ درصد) و ۲.۳۹۲ مشتری با ریسک متوسط در کلاس B (۴.۷ درصد)، ۱۱،۰۵۲ مشتریان با ریسک کم در کلاس C (۲۲ درصد) و ۳۶،۳۸۸ مشتری فاقد ریسک در کلاس D (۷۳ درصد) قرار گرفته‌اند. پس از امتیازدهی مشتریان و طبقه‌بندی آن‌ها در چهارکلاس یاد شده، به تحلیل رفتاری مشتریان در کلاس A که مظنون به پول‌شویی می‌باشد پرداخته شد. سپس نتایج حاصل از تحلیل این دسته از مشتریان با تحلیل مشتریان کلاس A که از روش شبکه هوشمند عصبی به‌دست آمده‌اند، مورد مقایسه قرار خواهند گرفت.

در تحقیق حاضر از نگاشت خود سازمان‌دهنده کوهونن به‌علت استفاده از روش یادگیری با ناظر و نمای گرافیکی آن برای خوشه‌بندی استفاده شده است. ورودی شبکه دارای ۲۸۶ متغیر در ۲۵ جدول (Table) می‌باشد. برای انجام خوشه‌بندی از ماژول نگاشت خود سازمان‌دهنده کوهونن و نرم‌افزار SPSS Clementine استفاده شده است. خروجی حاصل از شبکه به‌صورت یک ماتریس ۲*۲ می‌باشد که مشتریان براساس میانگین متغیرهای بدست آمده برای هر کدام از قسمت‌های ماتریس خوشه‌بندی می‌شوند و براساس شباهت خود در این معیارها در خوشه مشابه قرار می‌گیرند. خوشه‌بندی مشتریان در این قسمت کمک می‌کند تا خصوصیات مشتریان مشابه را ارزیابی کرده و ترجیحات هر گروه مشخص شود. خروجی خوشه‌بندی مشتریان که در جدول ۵ آورده شده است نشان‌دهنده ۴ (ماتریس ۲*۲) خوشه از مشتریان می‌باشد که هر خوشه بر روی نمودار در محل تقاطع محور x , y قرار دارد.

براساس جدول ۵ تعداد ۱۵۲ نفر از مشتریان جزء مشتریان کلاس A محسوب می‌شوند. به‌طور کلی با توجه به نگاه کاربردی بانک می‌توان مشتریان بانک را به چهار گروه براساس میزان ریسک پول‌شویی آن‌ها دسته‌بندی نمود.

جدول ۵: نتایج خوشه‌بندی مشتریان براساس روش نگاشت خود سازمان‌دهنده



$KY- Kohonen			
	۲	Records:152	Records: 18066
	۱	Records:3296	Records:28486
	۰	1	2
		$KX- Kohonen	

براساس خروجی مدل شبکه عصبی هوشمند (SOM)، ۱۵۲ مشتری (۰/۳ درصد) جزء کلاس A مظنونین به پول‌شویی می‌باشند. همچنین ۳،۲۹۶ نفر (۶/۵ درصد) جزء کلاس B یا مشتریان دارای ریسک متوسط، ۱۸،۰۶۶ نفر (۳۶درصد) جزء مشتریان کلاس C یا مشتریان با ریسک کم و در نهایت ۲۸،۴۸۶ نفر(۵۷/۲ درصد) مشتریان فاقد هرگونه ریسک در حوزه پول‌شویی می‌باشند. نتایج حاصل از خوشه‌بندی مشتریان بانک با توجه به نگاشت خود سازمان‌دهنده کوهونن که در جداول فوق نشان داده شده است، بیانگر دسته‌بندی مشتریان در ۴ خوشه می‌باشد. تعداد نورون‌های خروجی مدل ۲*۲ با توجه به الگوی رفتاری مشتریان توسط مدل تشخیص داده شده است. خوشه‌بندی مشتریان بدین طریق کمک می‌کند تا مشتریانی که دارای مشابه‌ترین الگوی رفتاری برحسب متغیرهای تراکنشی مالی و متغیرهای جمعیت‌شناختی هستند، در گروه‌های واحدی دسته‌بندی شده و برای هر گروه سرویس‌های مناسب به تفکیک خصوصیات آن‌ها در نظر گرفته شود. در ادامه تحقیق به بررسی الگوی رفتاری و کشف قوانین وابستگی در هر خوشه پرداخته شده است. در این پژوهش برای بررسی الگوی رفتاری و کشف قوانین

وابستگی در خوشه‌های ایجاد شده از مشتریان، تنها یکی از خوشه‌ها برای تحلیل انتخاب گردید. برای این منظور خوشه Y=2 ، X=1 به‌لحاظ اینکه جزء مشتریان با سطح بالاترین ریسک پول‌شویی می‌باشند، انتخاب می‌گردد. و ازسوی‌دیگر جهت مقایسه نیز مشتریان کلاس A دو روش فوق با یکدیگر در نظر گرفته شده است. بررسی و تحلیل رفتاری و کشف قوانین وابستگی بین سرویس‌ها و این دسته از مشتریان به مدیریت ارشد بانک جهت تدوین استراتژی‌های نحوه تعامل با این مشتریان کمک بسیاری می‌نماید. بنابراین خوشه‌ی X=1 ، Y=2 با تعداد ۱۵۲ مشتری و متغیرهای مربوطه آن به‌عنوان خوشه‌ی هدف در نظر گرفته شده است.

بخش دوم: سرویس های بانکی مورد استفاده توسط مشتریان با ریسک بالا

نتایج حاصل از خوشه‌بندی مشتریان به‌عنوان ورودی این مرحله برای شناسایی سرویس‌های مورد علاقه هر خوشه استفاده شده است. در این مرحله با استفاده از الگوریتم **Apriori** که یکی از تکنیک‌های رایج مورد استفاده در قوانین وابستگی است، رابطه میان سرویس‌های مختلف در خوشه مورد نظر شناسایی می‌گردد و در نهایت با استفاده از خصوصیات مشتریان در خوشه **X=1، Y=2** و سرویس‌های مورد علاقه این‌دسته از مشتریان، استراتژی مناسب برای نحوه ادامه همکاری با آن‌ها انتخاب می‌گردد.

کشف قوانین وابستگی بین سرویس‌های هر خوشه

برای شناسایی رابطه میان گروه‌های مختلف مشتریان و سرویس‌های مورد علاقه آن‌ها از قوانین وابستگی استفاده شده است. در این مرحله ابتدا خوشه مورد نظر از مشتریان انتخاب شده و براساس اطلاعات سرویس‌های استفاده شده توسط هرکدام از مشتریان در این خوشه، قوانین وابستگی محاسبه می‌گردد. با شناسایی رابطه سرویس‌ها و خوشه مزبور می‌توان علایق مشتریان هدف را بدست آورد. در الگوریتم‌های موجود برای قوانین وابستگی، **Apriori** یکی از پرکاربردترین الگوریتم‌ها می‌باشد (لیو و همکاران، ۲۰۱۸). در این مطالعه از این تکنیک برای شناسایی رابطه میان سرویس‌های مختلف مورد علاقه در خوشه یاد شده، استفاده شده است.

جدول ٦: قوانین وابستگی در خوشه انتخاب شده

	Antecedent	**Support %**
LC	Presentation of heavy documents - Failure to repay the loan	17.51
Total number of transactions	Above the level of activity- Rules AML28	19.41
RTGS	Unusual transactions- Combined tools	18.21
Current account	spider web- Same zip code- Joint IP	18.91
Dangerous OR free areas	Money transfer network- High currency transfer activity	10.61
List Of FIU	Obtaining AML information	13.21
High Amount Check	Concentration of money-spider web	11.71
The amount of facilities and obligations	Gold transactions-Non-financial jobs	14.71
Internet sales portals	Sudden deactivation of active -Multiple change of address accounts	13.11
Age between 45 and 65	Unusual transactions- High turnover accounts	12.51
Account opening branches	Reactivation of inactive accounts	10.21
Shared account holders	Safe box- High turnover accounts	11.31
Joint attorney	Above the level of activity-spider web	9.61
tax fraud	Communication with employees who have a financial violation case	8.71
Major shareholders	Combined tools- RTGS	10.11
Number of access cards	Unusual transactions-spider web	15.11
Safe box	Multiple withdrawals of significant amounts	10.11
Transaction tools in branches	High currency transfer activity	18.91
Rules AML	High turnover accounts-Joint IP	18.21
Daily limit of interbank remittance	Dangerous OR free areas- Unusual transactions	13.11
Cash transaction report over the limit	Combined tools-Newly opened account	19.51
Combined tools	spider web- Zero time- Joint IP	11.71
Zero time	false statement-spider web	17.14
A real customer with a legal background	Trading precious metals- B to C transactions	12.81

	Antecedent	Support %
Joint IP	Major shareholders- Combined tools	8.91
spider web	Political personality- Combined tools	16.91
Rental accounts	spider web -Joint IP- Combined tools	15.61
Use of children's accounts	Newly opened account-Joint IP	19.21
Newly opened account	Presentation of heavy documents - Failure to repay the loan-tax fraud	12.21
High amounts in limited time returns	spider web-List Of FIU	11.91
settlement of facilities	virtual interlocutors-Joint IP	17.15
Same zip code	Daily limit of interbank remittance	18.31
Contract termination without reason	Rules AML- Rental accounts- Number of access cards	19.41

در این مرحله برای شناسایی رابطه میان سرویس‌های مورد استفاده در خوشه انتخابی از الگوریتم Apriori با حداقل درجه پشتیبانی ۲۰ درصد و حداقل اطمینان ۸۰ درصد استفاده شده است. قوانین ایجاد شده در این خوشه در جدول ۶ آورده شده است، قوانین بدست آمده نشان‌دهنده رابطه میان سرویس‌های مختلف در خوشه مورد بررسی است. برای مثال قانون شماره سه بیان می‌کند افرادی که از طریق تسویه ناخالص آنی (RTGS) انتقال وجه انجام می‌دهند، با ضریب اطمینان ۷۸٪ و پشتیبانی (تکرار شوندگی) ۱۸٪ دارای تراکنش‌های مالی غیر معمول (با مبالغ بسیار بالا و یا خارج از سطح فعالیت مالی مورد انتظار) بوده و همچنین از ابزارهای ترکیبی (مدیاهای مختلف) برای نقل و انتقال وجوه خود استفاده می‌نمایند.

با توجه به قوانین ایجاد شده مشخص شد که مشتریان مظنون به پول‌شویی در میان سرویس‌های بانکی از خدماتی چون: حساب‌های اجاره‌ای، IPهای مشترک، ابزارهای نوع ترکیبی و نقل و انتقال شبکه‌ای، بیشتر از سایر سرویس‌ها استفاده خواهند نمود.

تحلیل سرویس‌های مورد استفاده مظنونین موجود، جهت پیش بینی و شناسایی مظنونین احتمالی

با شناسایی سرویس‌های مورد علاقه مظنونین در خوشه مشتریان با ریسک پول‌شویی بالا می‌توان با تحلیل، مشتریانی که به تازگی وارد شبکه بانک شده‌اند و رفتار مشابهی مانند این دسته از مشتریان در استفاده از این سرویس‌ها و خدمات بانکی دارند، را شناسایی نمود. در قسمت قبل، با ضریب اطمینان مشخص و با میزان خاصی از تکرارشوندگی، سرویس‌های مورد علاقه خوشه مورد نظر از مشتریان، شناسایی شدند. در ادامه با تحلیل اختلاف میان محصولات و سرویس‌های مورد استفاده توسط هر مشتری در خوشه مشتریان مظنون به پول‌شویی و علاقه مشتریان جدید به استفاده از همان محصولات و سرویس‌ها، می‌توان به شناخت جامع‌تری از مشتری دست یافت. به‌گونه‌ای که مشتریانی که خصوصیات مشابه با خوشه مورد بررسی را دارند، ممکن است به سرویس‌های مورد نیاز این خوشه علاقه‌مندی نشان دهند. با استفاده از این تحلیل می‌توان مشتریان جدید را نیز مورد بررسی قرار داده و متناسب با خصوصیات آن‌ها، سرویس‌های مورد استفاده ایشان را شناسایی نمود. همان‌گونه که اشاره شد، با استفاده از قوانین وابستگی، سرویس‌های مورد علاقه خوشه مظنونین مشخص گردیده است. در این مرحله برای شناسایی مشتریانی که رفتار مشابهی در استفاده از محصولات، خدمات و همچنین سرویس‌های مورد استفاده مشتریان مظنون به پول‌شویی داشتند، شناسایی می‌شود و می‌توان با تحلیل اختلاف میان سرویس‌های مورد علاقه و سرویس‌های مورد استفاده، پیشنهادات سیستم را استخراج کرد. به‌گونه‌ای که ممکن است یک مشتری مظنون به عملیات پول‌شویی در استفاده از سرویس‌ها، رفتار خلاقانه‌ی متفاوتی ارائه نماید که در نهایت منجر به کشف قانون جدیدی گردد که می بایستی به قوانین قبلی اضافه گردد.

با استفاده از حروف T و F نشان داده شده است که آیا مشتری مورد نظر از سرویس‌ها استفاده می‌کند یا خیر؛ بدین صورت که برای سرویس‌های مورد استفاده مشتری حرف T در نظر گرفته شده است. N-ID یک New ID می‌باشد که رفتار مشابهی با ID مربوطه دارد.

جدول ۷: تحلیل رفتار مشتریان مظنون به پول‌شویی

Services	*ID*	*Rental Accounts*	*spider web*	*Combined Tools*	*Safe box*	*RTGS*	*LC*	*Internet Sales Portals*	*Presentation of heavy documents*
Real	۲۸۹	T	T	T	F	T	F	T	T
N-ID		•	•	•	•	•	•	•	
Real	۱۲۱۱	T	T	T	T	F	F	T	F
N-ID		•	•	•		•		•	
Real	۱۲۰۵۴	F	T	F	F	T	T	F	T
N-ID		•	•		•	•		•	•
Real	۱۸۲۶۸	T	T	F	T	F	T	F	T
N-ID			•	•		•		•	•
Real	۲۲۹۶۶	T	F	F	T	T	T	T	F
N-ID		•	•		•		•	•	•
Real	۲۸۲۱۳	F	F	F	T	T	T	F	T
N-ID			•	•	•	•			•
Real	۳۱۰۶۵	F	T	F	F	T	F	T	F
N-ID		•		•		•	•	•	•
Real	۳۵۹۱۷	T	F	F	T	T	F	F	T
N-ID			•	•	•		•		•
Real	۳۹۸۵۲	T	F	T	T	F	F	T	T
N-ID		•	•	•		•		•	•
Real	۴۰۱۵۳	F	T	T	F	F	F	T	F
N-ID		•		•	•	•	•	•	
Real	۴۴۲۳۷	T	F	T	T	T	F	T	T
N-ID				•		•		•	•
Real	۴۹۷۵۳	T	T	T	F	T	F	F	F
N-ID			•		•			•	

تشکیل پروفایلی از مشتریان هدف جهت تسهیل در تدوین استراتژی‌های مناسب

در نظام بانکداری، با توجه به ریسک شهرتی که مشتریان پول‌شویی به بانک تحمیل می‌نمایند، شناسایی مشتریان مظنون به پول‌شویی از اهمیت بسیار بالائی برخوردار می باشد.

در این پژوهش با تحلیل مشتریان هدف (کلاس A– مشتریان مظنون به پول‌شویی در بانک ملت) می‌توان به اطلاعات ارزشمندی دست یافت. بنابراین مهمترین دسته از نظر الگوی رفتاری جهت تحلیل رفتار و تشکیل پروفایل جامعی از اطلاعات، مشتریان کلاس A می‌باشند. بنابراین بانک باید هزینه‌های بیشتری جهت شناسایی و تحلیل رفتار این دسته از مشتریان جهت جلوگیری از هزینه‌های پیش‌بینی نشده آتی، بپردازد. با توجه به تعداد کم این مشتریان نسبت به کل مشتریان (تعداد ۱۶۸ نفر مشتری مظنون به پول‌شویی براساس روش RBA و ۱۵۲ نفر براساس روش (SOM))، بانک می‌تواند، نسبت به شناسایی و تحلیل رفتار جامع آنان اقدام نماید تا در دوره‌های بعد، مشتریانی با الگوهای رفتاری مشابه را شناسایی نماید.

ضمناً مشتریانی با الگوی رفتاری فوق در گروه مشتریان کلاس A در بین خدمات بانکی، به یکسری خاصی از خدمات، از جمله: استفاده از حساب‌های اجاره‌ای، استفاده از IPهای مشترک، استفاده از ابزارهای نوع ترکیبی و نقل و انتقال شبکه‌ای، اشتیاق خاصی پیدا نموده‌اند. نکته قابل توجه در تحلیل رفتاری این گروه از مشتریان این است که تمایل ایشان به استفاده از خدمات در حوزه بانکداری اینترنتی رو به تزاید می‌باشد.

مقایسه نتایج حاصل از رتبه‌بندی مشتریان از طریق RBA و SOM

در جداول ذیل تعداد و درصد مشتریان در روش‌های مختلف محاسبه شده است. در تحلیل رفتار مشتریان با توجه به اینکه مشتریان کلاس‌های **C** و **D** (مشتریان با ریسک کم و مشتریان فاقد ریسک) دارای رفتار نزدیک به هم می‌باشند، می‌توان آن‌ها را با هم در نظر گرفت. در نهایت می‌توان بیان نمود که نتایج حاصل، از یک همگرائی جامعی نسبت به هر دو روش مطرح شده در پژوهش حاضر برخوردار می‌باشد. خلاصه نتایج به شرح جدول ۸ می‌باشد.

جدول ۸: مقایسه تعداد مشتریان مظنون به پول‌شویی در دور روشRBA و SOM

Ranking	Cluster	Number of RBA	Number of SOM	Total gaps
۱	۱	۱۶۸	۱۵۲	۱۶
۲	۲	۲۳۹۲	۳۲۹۶	۹۰۴
۳	۳	۱۱۰۵۲	۱۸۰۶۶	۷۰۱۴
۴	۴	۳۶۳۸۸	۲۸۴۸۶	۷۹۰۲
جمع		۵۰۰۰۰	۵۰۰۰۰	۱۵۸۳۶

نتایج جدول فوق نشان‌دهنده این است که مشتریان مظنون به پول‌شویی (کلاس A) دارای کمترین درصد (در حدود ۰/۳ درصد) و مشتریان فاقد ریسک و یا دارای ریسک کم دارای بیشترین میزان فراوانی (در حدود بیش از ۹۰ درصد) می‌باشند.

همچنین، نتایج حاصل در این پژوهش در دو روش (RBA و SOM) در شناسایی اشخاص مظنون به پول‌شویی، از یک همگرائی خاصی برخوردار می‌باشد. در این پژوهش جهت ارزیابی صحت نتایج به‌دست آمده از توزیع مشتریان در کلاس‌های مختلف، از آزمون ANOVA استفاده شده است. براساس نتایج آزمون مقدار P-Value نشان‌دهنده معنی‌دار بودن نتایج آزمون بوده و همچنین مقدار Chi-squared نشان می‌دهد توزیع فراوانی مورد انتظار و مورد مشاهده یکسان می‌باشند.

پاسخ سوال اول

مدل‌های رتبه‌بندی رفتاری مشتریان چگونه به کشف الگوهای رفتاری مشتریان و شناسایی مشتریان مظنون به پول‌شویی و یا مشتریان با ریسک بالای پول‌شویی کمک می‌کند؟

با به‌کارگیری تکنیک‌های هوشمند داده‌کاوی، الگوریتم (SOM) و رتبه‌بندی مشتریان بر پایه ریسک پول‌شویی (RBA) از یک پایگاه داده‌ای مشترک که از پایگاه داده‌ای بانک ملت استخراج گردید به این سوال که طبقه‌بندی مشتریان بانک براساس روش‌های هوشمند و شرایط موجود در بانک ملت، چگونه است؛ بدین طریق پاسخ داده شده است:

۱- در مدل شبکه عصبی هوشمند (ANN):

با استفاده از متغیرهای پژوهش برای بررسی الگوی رفتاری مشتریان و همچنین با استفاده از شبکه عصبی کوهونن، ماتریس نورون‌های خروجی شبکه ۲*۲ تشخیص داده شده و مشتریان در ۴ خوشه

دسته‌بندی گردیدند. این طبقات هرکدام با توجه به بیشترین شباهت در الگوی رفتاری دسته‌بندی گردیدند. از میان خوشه‌های ایجاد شده خوشه با مختصات (X=1 ، Y=2) با تعداد ۱۵۲ مشتری برای انجام تحلیل بیشتر و تشخیص قوانین انجمنی بین سرویس‌های این گروه انتخاب شده‌اند.

۲- با استفاده از مدل موجود در بانک ملت (RBA)

از سوی دیگر با به‌کارگیری تکنیک رتبه‌بندی مشتریان مبتنی‌بر ریسک پول‌شویی و با استفاده از همان متغیرها و براساس مدل امتیازدهی، به تحلیل خوشه‌بندی مشتریان در ۴ دسته پرداخته شد.

مشتریان فاقد ریسک (بدون امتیاز) که دامنه امتیازی آن‌ها (بین ۱ تا ۳۳) بوده، مشتریان با ریسک کم و مشتریانی که دامنه امتیازی آن‌ها (بین ۳۳ تا ۶۶) بوده مشتریان با ریسک متوسط و مشتریانی که جمع امتیازهای آن‌ها بالاتر از ۶۶ بوده مشتریان با ریسک بالا می‌باشند. تعداد ۱۶۸ نفر از مشتریان براساس این مدل در گروه کلاس (A) قرار گرفته‌اند.

سپس براساس میزان امتیاز بدست آمده، هر یک از دسته‌های مشتریان تعیین شده است. خوشه با ریسک بالا (کلاس A) شناسائی شده و با استفاده از قوانین انجمنی بین سرویس‌های بانکی به تحلیل رفتار این دسته از مشتریان پرداخته شد.

پاسخ سوال دوم

نگاشت خود سازمان دهنده (**SOM**) چگونه می‌تواند، طبقه‌بندی بهتری را نسبت به مفهوم (**RBA**) در خصوص رتبه‌بندی این دسته از مشتریان در بانک ملت ارائه نمایند؟

در این پژوهش جهت ارزیابی صحت نتایج بدست آمده از توزیع مشتریان در کلاس‌های مختلف، از آزمون ANOVA استفاده شده است. براساس نتایج آزمون مقدار P-Value نشان‌دهنده معنی‌دار بودن نتایج آزمون بوده و همچنین مقدار Chi-squared نشان می‌دهد توزیع فراوانی مورد انتظار و مورد مشاهده یکسان می‌باشند.

H_O: μ1=μ2

H_1: حداقل یکی از میانگین‌ها باهم برابر نیستند

براساس آزمون فوق مقدار Sig=1 بیانگر این است که نمی‌توان فرض یکسان بودن میانگین دو گروه مشتریان با ریسک پول‌شویی بالا را که از دو روش (SOM و RBA) بدست آمده‌اند، رد نمود؛ از طرفی

مقدار F بدست آمده از جدول فوق کمتر از مقدار F با درجه آزادی K-1=3 و N-K=۴۹۹۹۷ می‌باشد (۱.۰۷<۳.۱۸) لذا در سطح معنی‌داری ۰.۰۵ نمی‌توان فرض H_0 را رد نمود.

نتایج حاصل از آزمون ANOVA

در این روش تغییرات کل به دو بخش تجزیه می‌شود :

تغییرات کل = تغییرات بین گروهی + تغییرات داخل گروهی

با استفاده از نسبت تغییرات بین گروهی به تغییرات داخل گروهی معیار F بدست خواهد آمد که مبنای تصمیم‌گیری در مورد فرض صفر و فرض مقابل است. درصورتی‌که میانگین k گروه مورد مقایسه باشد دراین‌صورت می‌توان فرض صفر و فرض مقابل را به‌صورت آماری زیر بیان نمود:

H_0: میانگین k گروه مورد مقایسه با هم برابرند

H_1: حداقل یک جفت از میانگین‌ها باهم برابر نیستند.

برای قضاوت در مورد فرض صفر بالا همان‌گونه که گفته شد از روش تحلیل واریانس استفاده می‌شود.

جدول۹: نتایج حاصل از آزمون ANOVA

منابع تغییرات	مجموع مربعات	درجه آزادی	میانگین مربعات	نسبت F^*	سطح معنی‌داری
بین گروهی	SSB	K-1	MSB	MSB/MSW	$P_r(F^*>F)$
داخل گروهی	SSW	N-K	MSW		
جمع	SST	N-1			

تحلیل واریانس برای مقایسه میانگین K گروه

فرض صفر هنگامی رد می‌شود که مقدار **F** بدست آمده بیشتر از **F** جدول با **k-۱** و **n-k** درجه آزادی باشد یا مقدار سطح معناداری کمتر از ۰.۰۵ باشد.

جهت مقایسه نتایج حاصل از رتبه‌بندی مشتریان در هر دو روش از آزمون فوق استفاده می‌شود.

H_0: $\mu_1 = \mu_2$

H_1: حداقل یکی از میانگین‌ها باهم برابر نیستند

جدول ۱۰: تحلیل واریانس برای مقایسه میانگین ۴ گروه از مشتریان

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	۲۸۵۹۶۸۱۳	۳	۷۱۴۹۲۰۳	۱/۰۷	۱
Within Groups	۲۶۷۱۷۳۱۹	۴۹۹۹۷	۶۶۷۹۳۲۹		
Total	۵۵۳۱۴۱۳۲		۱۳۸۲۸۵۳۲		

نتایج بدست آمده در جدول فوق نشان دهنده اختلاف ناچیز پیشنهادات در دو روش به‌کار گرفته شده می‌باشد. بر این اساس می‌توان استدلال کرد که چارچوب طراحی شده برای ارائه طبقه‌بندی مشتریان در بانک از عملکرد مطلوبی برخوردار است و استفاده از هریک از روش‌های دوگانه فوق تقریباً نتایج یکسانی ارائه خواهد نمود.

نتایج حاصل از پژوهش از اختلاف معنی‌داری بین شرایط موجود و نتایج مطلوب حاصل از تحقیق برخوردار نمی‌باشد. براساس مدل RBA تعداد ۱۶۸ نفر از مشتریان دارای بیشترین امتیاز ریسک پول‌شویی بوده و در گروه مشتریان کلاس A قرار گرفتند. همچنین، براساس مدل شبکه عصبی تعداد ۱۵۲ نفر نسبت به دیگر مشتریان دارای ریسک پول‌شویی بالائی بوده‌اند. با نگاهی به آمار و ارقام فوق و مقایسه آن‌ها با هم، استنباط می‌شود که از دو روش فوق نتایج تقریباً مشابهی بدست آمده است که حاکی از یکسان بودن هر یک از روش‌های فوق در رتبه‌بندی مشتریان بانک می‌باشد و می‌توان ادعا نمود که روش‌های مطرح شده در پژوهش حاضر (غیر از مزیت‌های خاص) طبقه‌بندی بهتری را نسبت به شرایط جاری در بانک ملت ارائه نمی‌نمایند.

پاسخ سوال سوم

استراتژی‌های مناسب بانک جهت برخورد با هر یک از دسته‌های مشتریان با درجات ریسک پول‌شویی متفاوت (بالا- متوسط و کم) چگونه است؟

در ادامه استراتژی مناسب بانک در ارتباط با مشتریان هدف شرح داده شده و در نتیجه جهت تسهیل در تدوین استراتژی نحوه تعامل با این دسته از مشتریان، پیشنهادهای خاصی به‌شرح ذیل به مدیریت ارشد بانک ارائه شده است:

۱- مشتریان کلاس A بدون محدویت در ارائه خدمات بانکی، به‌صورت کاملا کنترل شده، توسط واحد مبارزه با پول‌شویی و ضد تأمین مالی تروریسم، تحت مراقبت ویژه قرار گیرند.

۲- متناسب با نزدیک شدن رفتار مشتریان جدید الورود به رفتار مشتریان مظنون به پول‌شویی، طیفی از محدودیت‌های لازم در نظر گرفته شود.

۳- محدودیت کامل در ارائه خدمات بانکی به این‌دسته از مشتریان، ملحوظ نظر واقع گردد.

۴- معرفی این دسته از مشتریان به‌عنوان خروجی سامانه مظنونین به پول‌شویی به مرکز اطلاعات وزارت امور اقتصادی و دارائی (FIU).

نتیجه‌گیری و پیشنهادات

- ✓ با شناسایی خصوصیات مشتریان در طبقات مختلف، به‌راحتی می‌توان مشتریانی را که دارای خصوصیات مشابهی هستند و یا رفتاری مشابه رفتار مشتریان مظنون به پول‌شویی دارند را شناسایی نمود. با شناسایی این گروه از مشتریان و شناخت الگوهای رفتاری آن‌ها، با اتخاذ استراتژی‌های مناسب، می‌توان ضریب امنیت را به‌شدت بالا برد.
- ✓ با تحلیل یکی از خوشه‌ها که دارای مشتریانی با ریسک پول‌شویی بالا می‌باشد، مشخص گردید که این‌دسته از مشتریان از سرویس‌های حوزه بانکداری الکترونیک، ارائه شده توسط بانک، بیشتر از سایر موارد استفاده می‌کنند و کارشناسان اداره مبارزه با پول‌شویی بانک می‌توانند ضمن کسب مهارت‌های لازم در این حوزه با همکاری واحدهای IT بانک، نسبت به شناسایی رفتار این‌دسته از مشتریان اقدام نمایند و یا اینکه بانک نسبت به برگزاری تخصصی دوره‌های مربوطه برای کارشناسان و مدیران مبارزه با پول‌شویی اقدام نماید.
- ✓ با تحلیل سایر خوشه‌ها با استفاده از الگوی ارائه شده، می‌توان به نتایج مشابهی دست یافت و سرویس‌های مورد علاقه آن‌ها را نیز شناسایی نمود. نتیجه حاصل از تحقیق نشان دهنده کاربرد گسترده و مؤثر تکنیک‌های داده‌کاوی در این ارتباط می‌باشد.
- ✓ پایگاه داده‌های بانک جزیره‌ای بوده و داده‌های بانک در شرکت‌های مختلف (سه شرکت وابسته به بانک، بنام‌های شرکت بهسازان ملت (متولی سوئیچ Core Banking)، شرکت یاس ارغوانی (گزارشات BI مشتریان)، شرکت به‌پرداخت ملت (متولی سوئیچ Pos-Vpos) ذخیره گردیده است. پیشنهاد می‌شود جهت اخذ تصمیمات بهینه و سریع استراتژیکی، بانک پایگاه داده‌ای متمرکزی را ایجاد نماید.

- ✓ پیشنهاد می‌شود به تحلیل رفتاری مشتریان گروه (B) مشتریان با ریسک پول‌شویی متوسط، پرداخته و با تحلیل رفتار مناسب این دسته از مشتریان، مشتریانی که دارای طیف امتیازی بالاتری و یا رفتار مشابه بیشتری نسبت به مشتریان گروه هدف هستند را شناسایی نمود.

پیشنهاد می‌شود با توجه به مزایای روش‌های مطرح شده در پژوهش حاضر (توصیف جامع از جامعه آماری مشتریان) از یک روش ادغامی جهت رتبه‌بندی مشتریان استفاده شود.

منابع و مآخذ

الف. فارسی

خاکی، غلامرضا (۱۳۸۶). روش تحقیق در مدیریت، تهران: مرکز انتشارات دانشگاه آزاد اسلامی.

انگلیسی

Agrawal R., Imielinski T., & Swami A. (1993). ***Mining Association Rules between Sets of Items in Large Databases***. In Proceedings of the ACM SIGMOD Conference on Management of Data, 207-216.

Chao, X., Kou, G., Peng, Y., & Alsaadi, F. (2019) ***Behavior Monitoring Methods for Trade-based Money Laundering Integrating Macro and Micro Prudential Regulation: A case from China***. Technological and Economic Development of Economy 25(6), 1-16.

Chen, Z. (2018). ***Machine Learning Techniques for Anti-money Laundering (AML) Solutions in Suspicious Transaction Detection: A Review***. Knowledge and Information Systems, 57(2), 245-285.

Chen, Z., Van Khoa, L.D., Nazir, A., Teoh, E. N., & Karupiah, E. K. (2014). ***Exploration of the Effectiveness of Expectation Maximization Algorithm for Suspicious Transaction Detection in Anti-money Laundering***. IEEE Conference on Open Systems (ICOS), 145–149. doi:10.1109/icos.2014.7042645

Chen Y-T., & Mathe, J. (2011) ***Fuzzy Computing Applications for Anti-money Laundering and Distributed Storage System Load Monitoring***. World conference on soft computing.

Heidarinia, N., Harounabadi, A., & Sadeghzadeh,M. (2014). ***An Intelligent Anti-Money Laundering Method for Detecting Risky Users in the Banking Systems***. International Journal of Computer Applications, 97(22), 35-39.

Guevara, J., Garcia-Bedoya, O., & Granados, O. (2020). ***Machine Learning Methodologies against Money Laundering in Non-banking Correspondents***. in Applied

Informatics (Communications in Computer and InformationScience), H. Florez and S. Misra, Eds. Ota, Nigeria, 1277, 72_88.

Jullum, M., Løland, A., Huseby, R. B., Ånonsen, G., & Lorentzen, J. (2020). ***Detecting money laundering transactions with machine learning***. Journal of Money Laundering Control, 23(1), 173-186.

Kim, M. (2007). ***An Effective Under-sampling Method for Class Imbalance Data Problem***. in Proceedings of the 8th Symposium on Advanced Intelligent Systems, 825-829.

Kohonen, T. (2013). ***Essentials of the Self-organizing Map***. Neural Networks, 37, 52-65.

Hsu, M-H. (2008). ***A personalized English Learning Recommender System for ESL Students***. Expert Systems with Applications, 34, 683-688.

Magomedov, S., Pavelyev, S., Ivanova, I., Dobrotvorsky, A., Khrestina, M., & Yusubaliev, T. (2018). ***Anomaly Detection with Machine Learning and Graph Databases in Fraud Management***. International Journal of Advanced Computer Science and Applications (IJACSA), 9(11), 33-38.

Plaksiy, K., Nikiforov, A., & Miloslavskaya, N. (2018). ***Applying Big Data Technologies to Detect Cases of Money Laundering and Counter Financing of Terrorism***. In Proceedings – 2018 IEEE 6th International Conference on Future Internet of Things and Cloud Workshops, W-FiCloud 2018 (pp. 70-77). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/W-FiCloud.2018.00017

Rocha-Salazar, J.-d.-J., Segovia-Vargas, M.-J., & Camacho-Miñano, M.-d.-M. (2021). ***Money Laundering and Terrorism Financing Detection using Neural Networks and an Abnormality Indicator***. Expert Systems With Applications, 169, 114470.

Shaikh, A.K., & Nazir, A. (2018). ***A Model for Identifying Relationships of Suspicious Customers in Money Laundering using Social Network Functions***. Proceedings of the World Congress on Engineering, WCE 2018 - London, United Kingdom.

Yang, Y., Lian, B., Li, L., Chen, C., & Li, P. (2019). ***DBSCAN Clustering Algorithm Applied to Identify Suspicious Financial Transactions***. International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, 60-65, https://api.semanticscholar.org/CorpusID:9956948

Luo, X. (2014). ***Suspicious Transaction Detection for Anti-Money Laundering***, International Journal of Security and Its Applications, 8(2), 157-166.