مجله‌ی بررسی‌های آمار رسمی ایران
سال۲۳، شماره‌ی ۲، پاییز و زمستان ۱۳۹۱، صص ۱۴۱-۱۵۷

مدل‌سازی روی‌گردانی مشتریان مؤسسه‌ی مالی و اعتباری مهر با استفاده از تکنیک‌های داده‌کاوی (رویکرد: هیبریدی فازی – ماشین بردارهای پشتیبانی)

حمیدرضا احمدی خالدی* و علی‌محمد احمدوند

دانشگاه امام حسین (ع)

چکیده: در سال‌های اخیر با گسترش روزافزون داده‌ها و پایگاه داده‌ها روبرو شده‌ایم. به موازات این امر شاهد پیشرفت تکنولوژی و علوم مختلف می‌باشیم تا بتوان از این خیل عظیم داده‌ها نهایت بهره را برد. در دنیای امروز و در مبحث مدیریت ارتباط با مشتری یکی از مزیت‌های رقابتی برای شرکت‌ها، سازمان‌ها و نهادها استفاده‌ی بهینه از داده‌ها است که در صورت وقوع چنین امری سازمان می‌تواند گامی بزرگ در راستای اهداف خود از جمله حفظ مشتریان قدیمی و جذب مشتریان جدید بردارد. از صنعت‌هایی که امروزه با داده‌های انبوه مشتریان سر و کار دارد صنعت بانکداری است. با توجه به این امر در این مقاله مسئله‌ی روی‌گردانی مشتریان مؤسسه‌ی مالی و اعتباری مهر مورد بررسی قرار گرفته است و پس از مرور ادبیات و مطالعات نظری در این زمینه و جمع‌آوری داده‌های مورد نیاز، مدل‌سازی روی‌گردانی مشتریان با رویکرد هیبریدی انجام گرفته است که در ابتدای این رویکرد گروه مشتریان با کمک قوانین فازی تعیین شده است. بر این اساس مشتریان در سه گروه فعال، متوسط و ضعیف قرار گرفته‌اند. در مرحله‌ی دوم نیز مدل‌سازی با کمک روش ماشین بردارهای پشتیبانی صورت گرفته است. نتایج حاصل از این رویکرد با رویکرد منطقی و تک‌متغیره مقایسه شده است و بنا به نتایج به دست آمده رویکرد هیبریدی فازی–ماشین‌های بردارهای پشتیبانی به‌منظور مدل‌سازی روی‌گردانی مشتریان مؤسسه‌ی مالی و اعتباری مهر معرفی شده است.

واژگان کلیدی: داده‌کاوی؛ مدیریت ارتباط با مشتری؛ روی‌گردانی؛ دسته‌بندی؛ ماشین بردارهای پشتیبانی.

---

* نویسنده‌ی عهده‌دار مکاتبات

دریافت: ۱۳۹۱/۶/۲۶، پذیرش: ۱۳۹۲/۹/۴.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

۱- مقدمه

در سال‌های اخیر، استفاده‌ی مؤثر از داده‌های مشتریان در سازمان‌ها و مؤسسات فعال با اهداف بازرگانی و تجاری، آموزشی و پرسنلی می‌تواند به بهترین نحو به مدیران، برنامه‌ریزان و پژوهشگران در اتخاذ تصمیم‌های راهبردی و توصیف و شناسایی وضعیت موجود یاری رساند و باعث انجام و اجرای علمی‌تر و دقیق‌تر فعالیت‌های شغلی برای این دسته از افراد شود. این امر به خصوص با افزایش سطح پوشش پایگاه‌های داده در سازمان‌های دولتی یا خصوصی در کشور اهمیتی دوچندان پیدا کرده است. در این میان از جمله صنایعی که در آن استفاده از داده‌های مشتریان باعث افزایش دقت تصمیم‌گیری و کاهش هزینه‌های مادی و معنوی تصمیم‌گیری می‌شود، صنعت بانکداری است. امروزه با استفاده از تکنیک‌های گوناگون سعی در استفاده‌ی بهینه از داده‌های عظیم در این مجموعه‌ها شده است. لذا در این مقاله سعی خواهد شد تا با بهره‌گیری از تکنیک‌های داده‌کاوی در صنعت بانکداری (پیش‌بینی روی‌گردانی[1] مشتریان بانکی) گامی اساسی برداشته شود. هدف از انجام این مقاله، تهیه‌ی مدلی به‌منظور تحقق این امر می‌باشد. به این منظور در ابتدا مروری مختصر بر ادبیات موضوع خواهیم داشت. پس از آن مدل را معرفی خواهیم کرد و در بخش بعدی نیز نتایج و یافته‌های حاصل از تحقیق را خواهیم آورد. در پایان نیز نتیجه‌گیری نهایی از مقاله را ارایه خواهیم داد.

۲- داده‌کاوی

داده‌کاوی[2] و کشف دانش در پایگاه داده‌ها از جمله موضوع‌هایی هستند که هم‌زمان با ایجاد و استفاده از پایگاه داده‌ها در اوایل دهه‌ی ۸۰ برای جستجوی دانش در داده‌ها شکل گرفت. شاید بتوان لوول را اولین شخصی دانست که گزارشی در مورد داده‌کاوی تحت عنوان «شبیه‌سازی فعالیت داده‌کاوی» ارایه نمود. هم‌زمان با او پژوهشگران و متخصصان علوم رایانه، آمار، هوش مصنوعی، یادگیری ماشین و ... نیز به پژوهش در این زمینه و زمینه‌های مرتبط با آن پرداختند. دانش داده‌کاوی، یکی از ده دانش در حال توسعه‌ای است که دهه‌ی آینده را با انقلاب تکنولوژیک مواجه خواهد ساخت و از این رو در سال‌های اخیر گسترش فوق‌العاده سریعی داشته است. پژوهش‌های جدی بر روی داده‌کاوی از اوایل دهه‌ی ۹۰ شروع شده و از آن پس مطالعات زیادی در این زمینه

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

صورت گرفته است [۲]. فن داده‌کاوی سازمان‌ها و شرکت‌ها را قادر می‌سازد تا از سرمایه‌ی داده‌های خود بهره‌برداری نموده و از آن به‌عنوان ابزاری برای پشتیبانی فرایند تصمیم‌گیری استفاده کنند. کاربردهای این علم در زمینه‌های مختلف مورد توجه دانشمندان قرار گرفته و نتایج بدست آمده از آن باعث تشویق پژوهشگران به استفاده‌ی گسترده‌تر از علم داده‌کاوی شده است. داده‌کاوی مرحله‌ای از فرایند اکتشاف دانش است که به تشخیص رابطه بین داده‌ها با استفاده از تکنیک‌های آماری کمک می‌کند [۹]. تعاریف متعددی در زمینه‌ی داده‌کاوی ارایه شده است که از آن جمله می‌توان به موارد زیر اشاره کرد:
داده‌کاوی عبارت است از استخراج اطلاعات و دانش و کشف الگوی پنهان از یک پایگاه داده‌های بسیار بزرگ که این الگوها و دانش‌ها معمولاً مستتر در داده می‌باشند [۱]. داده‌کاوی فرایند کشف رابطه‌ها، الگوها و روندهای جدید معناداری است که به بررسی حجم وسیعی از اطلاعات ذخیره‌شده در انبارهای داده با فناوری‌های تشخیص الگو (مانند ریاضی و آمار) می‌پردازد (سایت spss.com).

همان‌طور که در تعاریف بالا مشخص گردید داده‌کاوی به کشف الگوی پنهان و مستتر داده‌ها می‌پردازد.

از داده‌کاوی می‌توان جهت دسته‌بندی، تخمین، پیش‌بینی، خوشه‌بندی و همچنین یافتن روابط بین داده‌ها استفاده کرد [۹]. مسئله‌ی روی‌گردانی مشتریان زیرمجموعه‌ای از مسائل مربوط به حیطه‌ی دسته‌بندی است. دسته‌بندی عبارت است از جداسازی یا مرتب کردن اشیاء در کلاس‌های از قبل تعیین‌شده. دسته‌بندی نوعی یادگیری با نظارت می‌باشد که در طی فرایند آن سیستم، آموزش می‌بیند به‌طوری که اگر یک شی جدید به سیستم آموزش‌یافته ارایه شود شی را در یکی از کلاس‌های موجود قرار دهد [۱]. به‌منظور دسته‌بندی از روش‌های گوناگونی استفاده می‌گردد که در زیر سه نمونه از آن‌ها مورد بررسی قرار می‌گیرد.

ماشین بردار پشتیبانی

در هنگام تلاش برای کشف الگوها و مدل‌های طبقه‌بندی، یادگیری ماشین می‌تواند یک ابزار قوی به شمار رود. بیش‌ترین استفاده از تکنیک‌های داده‌کاوی، طبقه‌بندی مسائل و مثال‌های جدید درون کلاس‌های خاص است. این تکنیک به‌صورت موفق برای مشکلاتی که برای طبقه‌بندی الگوهای مختلف وجود دارد، به کار گرفته می‌شود [۱۱]. این تکنیک

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

به این صورت عمل می‌کند: این مدل از یک نگاشت خطی برای تبدیل داده‌های اصلی به ابعاد بالاتر استفاده می‌کند که در بعد جدید ابرصفحه جداساز بهینه را پیدا می‌کند (این ابرصفحه «تصمیم‌گیری دوتایی[۳]» است که دو کلاس خاص را از هم جدا می‌کند) [۹].

رگرسیون لوژستیک

این تکنیک مانند رگرسیون خطی سعی در پیدا کردن بهترین منحنی برای داده‌های مشاهده‌ای دارد. به جای یک خط از تابعی استفاده می‌کند که تابع لوژستیک[۴] نام دارد. هدف رگرسیون لوژستیک تعیین احتمال شرطی مربوط به مشاهده‌های مشخص یک کلاس با توجه به مقادیر متغیرهای مستقل است. به‌عبارت ساده با گرفتن متغیر ورودی مقدار متغیر وابسته به آن را پیش‌بینی می‌کند. دقت نسبتاً بالا و انعطاف‌پذیری از خصوصیات این مدل است [۴].

شبکه‌های عصبی مصنوعی

این سیستم از سه لایه تشکیل شده است: لایه‌ی ورودی، لایه‌ی میانی، لایه‌ی خارجی. لایه‌ی ورودی اطلاعات را از یک منبع خارجی دریافت می‌کند و لایه‌ی خارجی نیز اطلاعات را به سیگنال‌هایی برای استفاده منبع خارجی تبدیل می‌کند. لایه‌ی میانی نیز پلی بین لایه‌های ورودی و خروجی است و در حقیقت فرایند پردازش داده را انجام می‌دهد. این شبکه نیازی به دانستن توزیع داده‌ها ندارد و خود توزیع مربوط را به دست می‌آورد و کاربرد زیادی در فرایندهای ناپایدار دارد. این شبکه‌ها ابزارهایی هستند که در موارد پیش‌بینی، طبقه‌بندی، خوشه‌بندی و تخمین کاربرد دارند [۵].

۳- مدیریت ارتباط با مشتری[۵]: ابقا یا رویگردانی

برای حفظ رقابت، **سازمان‌ها** نیاز به تدوین استراتژی‌های تمرکز با مشتری[۶]، مشتری محور بودن[۷] و مشتری‌مداری[۸] دارند. همه‌ی این موارد خواسته‌های سازمان‌ها را در راستای ارتباط با مشتریان تعریف می‌کند. مدیریت ارتباط با مشتری راه حلی است که این تلاش‌ها را برای سازمان‌ها و همچنین مشتریان محقق می‌سازد [۳]. مدیریت ارتباط با

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

مشتری شامل چهار بُعد است: شناخت مشتری، جذب مشتری، حفظ مشتری و توسعه‌ی مشتری [۱۴].

مدیریت ارتباط با مشتری شامل محصول تولیدشده برای فروش، به چه مشتری‌هایی و از طریق چه کانال‌هایی می‌باشد. در فروش، کارخانه‌ها از مدیریت تهاجمی[۹] برای افزایش تأثیر در بازار استفاده می‌کنند. سرانجام، مدیریت ارتباط با مشتری قصد نگه‌داری مشتریانش از طریق خدماتی مانند مرکز تماس و پشتیبانی و کمک را دارد [۱۲].

یک تعریف پذیرفته‌شده‌ی دیگر برای مدیریت ارتباط با مشتری عبارت است از: «رویکرد یکپارچه برای درک و تأثیرگذاری بر رفتار مشتری از طریق ارتباطات معین با هدف بهبود کسب مشتری، حفظ مشتری، وفاداری مشتری و سودآوری مشتری با استفاده‌ی استراتژیک از اطلاعات، فرایندها، تکنولوژی و افراد برای مدیریت ارتباط مشتریان با شرکت در طول چرخه حیات مشتری» [۱۴].

اگر چه تعاریف بالا مقداری با یکدیگر متفاوت‌اند، اما همه‌ی آن‌ها به این نکته اشاره می‌کنند که هسته‌ی مدیریت ارتباط با مشتری حول تمرکز بر روی روابط خریدار و فروشنده بوده و تمرکز اصلی در این مفهوم، مشتریان و نیازهای آنان می‌باشد. از جمله کاربردهای داده‌کاوی در حوزه‌ی مدیریت ارتباط با مشتری رتبه‌بندی رویگردانی مشتری به‌منظور ارایه‌ی برنامه‌های مؤثر حفظ مشتری است [۳].

بقای مشتری مسئله‌ی بسیار مهمی برای انواع شرکت‌ها و صنایع است. تعیین روی‌گردانی و ابقا یکی از کاربردهای مهم داده‌کاوی در مدیریت ارتباط با مشتری است. روی‌گردانی مهم است به‌دلیل این‌که مشتریان از دست داده‌شده باید با مشتریان جدید جایگزین گردند و جذب مشتریان جدید هزینه‌ی بیش‌تری را برای شرکت صرف می‌کند همچنین درآمد حاصل از آن‌ها در کوتاه مدت بسیار کم‌تر از مشتریان قبلی است. این مطلب در بسیاری از صنایع صدق می‌کند مخصوصاً زمانی که بازار اشباع شده باشد، هر شخصی مایل به انتخاب محصول و یا خدمتی است که نیازهای او را بهتر برآورده کند و همچنین منبع اصلی جذب مشتری برای یک شرکت، افرادی هستند که شرکت رقیب را ترک کرده‌اند.

انواع مختلف روی‌گردانی عبارت است از: روی‌گردانی ارادی[۱۰]، روی‌گردانی غیرارادی[۱۱] و روی‌گردانی پیش‌بینی‌شده[۱۲] [۴]. البته لازم به ذکر است که با توجه به دلیل پیچیدگی موضوع معمولاً محققان دو نوع رویگردانی اول و دوم را شبیه به هم فرض کرده و در

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

دسته‌بندی تفکیکی بین آن‌ها قائل نمی‌شوند. در حوزه‌ی رویگردانی مشتریان بانکی مقالات مختلفی تدوین شده است که در زیر نتایج برخی از آن‌ها آورده شده است.

وی و چیو از تکنیک درخت تصمیم‌گیری برای روی‌گردانی مشتریان در صنعت مخابرات استفاده کرده‌اند و مورد کاوی آن‌ها در کشور تایوان بوده است. در این مقاله دسته‌بندی در نظر گرفته‌شده برای مشتریان ۲ گروه فعال و غیر فعال بوده و این خروجی‌ها با استفاده از متغیرهای تماس و مشخصات عمومی مشتریان همانند مدت زمان مکالمه‌ی مشتریان، مدت زمان خدمت‌رسانی به مشتری و ... ساخته شده است. طبق استنادی که در مقاله به آن اشاره شده در هر ماه در صنعت مخابرات، نرخ روی‌گردانی مشتریان ٪۲/۲ می‌باشد [۷].

پندهارکر با رویکردی هیبریدی بر پایه‌ی الگوریتم ژنتیک و شبکه‌های عصبی مصنوعی مدل‌سازی برای مشتریان در مرکز خدماتی بی‌سیم را انجام داده است که ابتدا از این دو روش برای پیش‌بینی رویگردانی کمک گرفته شده است و پس از آن سعی شده تا دقت این پیش‌بینی افزایش یابد [۱۵].

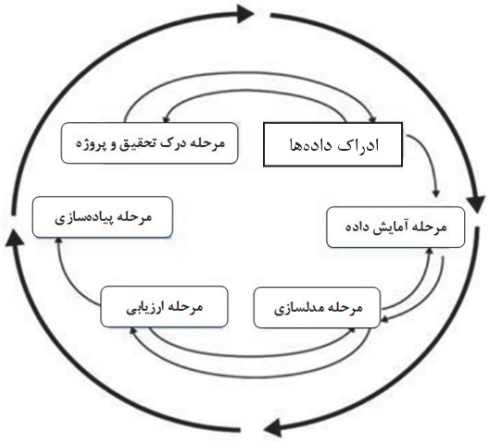
هانگ و دیگران متدولوژی ۳ مرحله‌ای مورد استفاده قرار داده‌اند که عبارت است از: ۱. فراهم‌سازی زیرساخت‌های فناوری اطلاعات به‌منظور تسهیل در تحقیقات که شامل مواردی همچون اطلاعات پایه‌ای و رایج مشتری، خصوصیات وی، تراکنش‌ها، پارامترهای مدل، نتایج مدل و ... می‌شود؛ ۲. فرایند مدل‌سازی به‌منظور کشف روابط پنهان داده‌های مشتریان با استفاده از تکنیک‌های داده‌کاوی؛ ۳. مدل‌سازی با کمک ابزار متفاوتی مانند درخت تصمیم‌گیری و شبکه‌های عصبی. در این مقاله مدل‌سازی بر روی داده‌های مشتریان صنعت خدماتی بی‌سیم در تایوان انجام گرفته است. داده‌ها مربوط به حدود ۱۶۰۰۰۰ مشتری در سال ۲۰۱۱-۲۰۱۲ که شامل اطلاعات عمومی مشتریان، آنالیز پرداخت‌ها و صورت‌حساب‌های آن‌ها، آنالیز جزئیات تماس، آنالیز خدمت مشتری بوده است [۱۰]. لین متغیر روی‌گردانی مشتریان را به سه گروه فعال، روی‌گردانی ارادی و روی‌گردانی غیر ارادی تقسیم کرده است و مدل‌سازی با استفاده از نظریه‌ی مجموعه‌های سخت[۱۳] انجام گرفته است و قوانینی به‌منظور پیش‌بینی روی‌گردانی مشتریان استخراج شده است. کارت‌های اعتباری مشتریان در این مقاله مورد توجه قرار گرفته است و با کمک متغیرهایی همچون سن، وضعیت تأهل، درآمد سالانه و ... روی‌گردانی مشتریانی با دقتی حدود ۹۰٪ محاسبه شده است [۱۳].

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

پوئل و دیگران از تکنیک ماشین بردار پشتیبانی به‌منظور ساخت مدل چرخش مشتری در مشترکین روزنامه استفاده کردند. طبق نوشته‌ی ایشان در فضای رقابتی به وجود آمده بیش‌تر شرکت‌ها به این نکته پی برده‌اند که پایگاه‌های داده‌ای از مهم‌ترین دارایی‌هایشان است. بنابراین شرکت‌ها و کمپانی‌ها تصمیم به انتقال از مدیریت سنتی گرفته‌اند. همچنین طی تحقیقات به عمل آمده حفظ مشتری بسیار سودآورتر از جذب مشتری جدید است [۶]. مقالات دیگر نیز طبق روش‌های مندرج در جدول و در صنعت‌های ذکر شده روی-گردانی مشتریان را انجام داده‌اند.

۴- معرفی مدل

موفق‌ترین پروژه‌های داده‌کاوی، در چارچوب فرایند استانداردی اجرا می‌شود که توسط یک تیم کاری در شرکت SPSS در قالب پروژه‌هایی به نام CRISP-DM ارایه شده است. بر طبق این مدل، یک پروژه‌ی داده‌کاوی معین شامل چرخه‌ی حیات شش مرحله‌ای است که توالی مراحل را نشان می‌دهد. هر مرحله از ترتیب مراحل، اغلب نتیجه‌ی وابستگی مراحل قبل را نیز دارد. در
شکل ۱- توالی مراحل در مدل CRISP-DM را می‌توان ملاحظه کرد [۸].



شکل ۱- توالی مراحل در مدل CRISP-DM

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

مسأله‌ی روی‌گردانی مشتریان در مؤسسات مالی و اعتباری از اهمیت ویژه‌ای برخوردار است. بر این اساس، مدل‌سازی رویگردانی مشتریان به‌عنوان هدف اصلی مقاله مطرح است و با توجه به گام‌های مطرح در روش تصمیم‌گیری سخت (CRISP-DM) انجام می‌شود.

به‌منظور مدل‌سازی روی‌گردانی مشتریان در ابتدا داده‌های مربوط به ۱۸۷۷۷ مشتری مؤسسه‌ی مالی و اعتباری مهر، بالغ بر ۳۱۰۰۰ رکورد ثبت‌شده در سیستم بوده است[۱۴]. خصوصیات داده‌های جمع‌آوری شده برای هر مشتری عبارت‌اند از: نام مشتری، شماره‌ی تماس، نشانی، نوع حساب، منطقه، شماره‌ی مشتری، تعداد برداشت، تعداد واریز، تعداد گردش، تاریخ افتتاح حساب، مبلغ سپرده‌گذاری، سن مدرک تحصیلی و جنسیت. به‌منظور آماده‌سازی داده‌ها برای مدل‌سازی گام‌های زیر برداشته شد.

- ترکیب و تناظر داده‌ها؛
- حذف داده‌های زائد (کمبود و یا نقص اطلاعات)؛
- کدگذاری تحصیلات، جنسیت و وضعیت تأهل؛
- محاسبه‌ی مدت تشکیل حساب و تعداد حساب‌های افتتاح‌شده پس از افتتاح اولین حساب؛
- تفکیک حساب‌های جاری و حساب‌های دارای بازه‌ی زمانی یک سال؛
- تعیین متغیر خروجی.

در نتیجه‌ی آماده‌سازی داده‌ها، متغیرهای اصلی مسئله شناسایی‌شده و در گروه‌بندی جدول ۱ در تدوین مدل‌سازی جای خواهند گرفت.

در مرور ادبیات انجام‌شده مشخص شد که در دسته‌بندی داده‌ها، متغیری به‌عنوان تعیین‌کننده‌ی دسته و یا گروه داده‌ها در نظر گرفته می‌شود که این متغیر گروه‌ها را از یکدیگر تفکیک می‌کند و مابقی متغیرها به‌عنوان مشخصات هر یک از داده‌ها در دسته‌های موجود مد نظر قرار می‌گیرند. در این مسئله نیز تعداد ۷ متغیر به‌عنوان مشخصات داده‌ها (متغیر ورودی) در نظر گرفته شده است و سه متغیر نیز به‌عنوان خروجی یا تعیین‌کننده‌ی گروه مشتریان در نظر گرفته خواهد شد (البته لازم به ذکر است که متغیر روی‌گردانی نیز از نوع خروجی بوده و در تعیین گروه‌ها کاربرد دارد. این متغیر با استفاده از مدل‌های استنتاج فازی با توجه به سه متغیر خروجی (Y_1، Y_2 و Y_3) دیگر ساخته خواهد شد).

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

جدول ۱ - متغیرهای مدل‌سازی

ردیف	عنوان	نوع	نشانه
۱	نوع حساب	ورودی	X_1
۲	منطقه‌ی حساب	ورودی	X_2
۳	مبلغ سپرده	ورودی	X_3
۴	سن	ورودی	X_4
۵	مدرک	ورودی	X_5
۶	جنسیت	ورودی	X_6
۷	وضعیت تأهل	ورودی	X_7
۸	تعداد برداشت	خروجی	Y_1
۹	تعداد واریز	خروجی	Y_2
۱۰	تعداد حساب	خروجی	Y_3
۱۱	روی‌گردانی	خروجی	Y_T

با توجه به متغیرهای موجود در بالا مدل‌سازی مسئله با سه رویکرد مختلف انجام خواهد گرفت:

- رویکرد تک متغیره: که هر یک از متغیرهای خروجی به‌عنوان کلاس دسته‌بندی برای تعیین روی‌گردانی مشتریان استفاده می‌شود و با استفاده از رابطه‌ی متغیرهای ورودی و خروجی به تعیین گروه‌های مشتریان پرداخته می‌شود.
- رویکرد هیبریدی- منطقی: در این رویکرد، متغیرهای خروجی با هم ترکیب شده و در نهایت یک متغیر (روی‌گردانی مشتریان) به‌عنوان متغیر نهایی در دسته‌بندی به کار خواهد رفت. در این حالت، این متغیرها به‌صورت منطقی (ضرایب وزنی) با هم ترکیب می‌شوند.
- رویکرد هیبریدی- فازی: در این رویکرد، متغیرهای خروجی با یکدیگر ترکیب شده و در نهایت یک متغیر (روی‌گردانی مشتریان) به‌عنوان متغیر نهایی در دسته‌بندی مشتریان به کار خواهد رفت. در این حالت، که رویکردی ابتکاری است، این متغیرها با کمک قوانین فازی، که در مدل‌های استنتاج فازی مطرح‌اند، با هم ترکیب خواهند شد تا متغیر نهایی تشکیل شود.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

با توجه به این مسئله که بیش‌تر مشتریان این مؤسسه از بستن حساب خود صرف نظر می‌کنند و کاربری آن‌ها منوط به میزان فعال بودن آن‌ها در مؤسسه و تراکنش‌های آن‌ها می‌باشد لذا دسته‌بندی ۳تایی برای متغیرهای خروجی در نظر گرفته شده است که با توجه به این دسته‌بندی مشتریان در یکی از کلاس‌های ضعیف، معمولی و فعال قرار می‌گیرند. به‌منظور تعیین حدود و نحوه‌ی دسته‌بندی، پرسش‌نامه‌هایی در اختیار کارشناسان بانک قرار گرفته که نتایج آن در جدول ۲ قابل ملاحظه است.

جدول ۲- محدوده متغیرهای خروجی در تعیین روی‌گردانی مشتریان

نوع گردش	ضعیف		معمولی		فعال	
(متغیر خروجی)	حد پایین	حد بالا	حد پایین	حد بالا	حد پایین	حد بالا
تعداد واریز	۱	۴/۸	۴/۸	۲/۲۵	۲/۲۵	۵/۱۴
تعداد برداشت	۰	۶/۴	۶/۴	۸/۲۰	۸/۲۰	۱۸۰
تعداد حساب	۰	۶/۰	۶/۰	۶/۲	۶/۲	۵

رجوع به جدول بالا نوع روی‌گردانی هر یک از مشتریان را مشخص خواهد ساخت. به‌عنوان مثال اگر با رویکرد تک‌متغیره مسئله را بنگریم و متغیر خروجی را تعداد واریز فرض کنیم و مشتری خاصی تعداد واریزهای حسابش عدد ۹ می‌باشد این مشتریان در کلاس مشتریان معمولی قرار خواهد گرفت. در این مقاله به دنبال آن هستیم تا بهترین نوع دسته‌بندی را با توجه به میزان صحت آن انتخاب کنیم لذا نتایج رویکردهای مختلف در بخش بعدی آورده خواهد شد.

۵- پیاده‌سازی مدل

با توجه به مذاکرات به عمل آمده با کارشناسان مؤسسه‌ی مالی و اعتباری مهر، یکی از اهداف تحقیقاتی این مؤسسه فعالیت بر روی داده‌های مشتریان به‌منظور حفظ مشتریان کنونی و جذب مشتریان جدید است. بنابراین، مسئله‌ی روی‌گردانی مشتریان در این مؤسسه مطرح و مورد توافق کارشناسان قرار گرفت. همان‌طور که ذکر شد با سه رویکرد مختلف به مسئله نگاه شده است. رویکرد اول، رویکرد تک‌متغیره است که اثر هر یک از متغیرهای خروجی را به تنهایی در روی‌گردانی مشتریان در نظر می‌گیرد و این مطلب نیز

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

عیب این رویکرد است. به‌دلیل برطرف کردن این عیب با رویکرد هیبریدی – منطقی اثر تعاملی این متغیرها به‌صورت رابطه‌ی خطی و با استفاده از ضرایب وزنی در نظر گرفته شده است. اما با توجه به این که این رویکرد نیز به‌صورت خطی اثر متغیرهای خروجی را بررسی می‌کند لذا به‌منظور برطرف کردن آن از مجموعه‌ی قوانین فازی در رویکرد هیبریدی – فازی استفاده شده است. با توجه به وجود سه ورودی و سه حالته بودن هر یک از آن‌ها، سیستم فازی تدوین‌شده حداکثر دارای ۲۷ قانون خواهد بود. این قوانین، برای حالات مختلف ورودی‌ها، خروجی را معرفی می‌کند. برای مثال «اگر واریز غیر فعال، برداشت غیر فعال و حساب غیر فعال باشد، خروجی بسیار پایین است». این قوانین به تعیین گروه‌های مشتریان در رویکرد هیبریدی فازی کمک خواهد کرد. پس از تعیین متغیر رویگردانی و گروه‌بندی مشتریان مدل‌سازی مسئله را با کمک نرم‌افزار SPSS Clementine در هر سه رویکرد و با سه روش رگرسیون لوژستیک، شبکه‌های عصبی مصنوعی و ماشین بردارهای پشتیبانی انجام شده است که نتایج آن در جدول ۳ آمده است.

جدول ۳- مقایسه صحت رویکردهای به کار رفته در مدلسازی

روش / رویکرد	تک‌متغیره (تعداد برداشت)	تک‌متغیره (تعداد واریز)	تک‌متغیره (تعداد حساب)	هیبریدی- منطقی	هیبریدی - فازی
رگرسیون لوژستیک	۷۰/۹۱	۸۵/۹۳	۷۵/۷	۶۲/۱۳	۶۱/۶
شبکه‌های عصبی مصنوعی	۷۱/۸۷	۸۶/۳۹	۷۵/۵۶	۶۱/۲۴	۵۹/۷۵
SVM	۷۹/۱۸	۸۹/۶۳	۸۳/۰۹	۷۳/۱۱	۷۳/۲۹

همان‌طور که از جدول بالا مشخص است در تکنیک‌های تک‌متغیره صحت به دست آمده از مدل‌سازی بیش‌تر است اما مسئله‌ای که اهمیت دارد در نظر گرفتن متغیرها به‌صورت مستقل و جداگانه در این رویکرد است چرا که همان‌گونه که ذکر شد هر سه متغیر (تعداد حساب، تعداد واریز و تعداد برداشت) در مدل‌سازی مسئله تأثیرگذار هستند. بنابراین رویکردهای هیبریدی نیز به این منظور تدوین گردید. از طرفی صحت به دست آمده در

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

رویکردهای هیبریدی تقریباً برابر می‌باشند اما باید توجه داشت که در رویکرد فازی، مسئله‌ی غیر خطی بودن داده‌ها مورد بررسی قرار گرفته است. لذا صحت این رویکرد از رویکرد منطقی بیش‌تر است. با توجه به این مسئله در قسمت آتی، به تحلیل بروی نتایج به دست آمده از رویکرد هیبریدی – فازی پرداخته خواهد شد.

۶- آنالیز حساسیت

پس از تعیین رویکرد هیبریدی- فازی به‌عنوان رویکرد اصلی مسئله در مدل‌سازی مسئله و همچنین تعیین روش SVM به‌عنوان روش پایه در این بخش با تغییر پارامترهای SVM صحت مدل‌سازی را در حالت‌های مختلف مورد بررسی قرار خواهیم داد. پارامترهای تأثیرگذار در روش SVM عبارت‌اند از:

- **معیار توقف** (Stopping Criteria): در این قسمت محدوده‌ی توقف برای بهینه-سازی الگوریتم مشخص می‌شود. مقادیر آن بین ۱×۱۰^{-۱} تا ۱×۱۰^{-۶} تغییر می‌کند و حالت پیش‌فرض نیز ۱×۱۰^{-۳} است. با کاهش مقدار دقت مدل افزایش می‌یابد اما زمان رسیدن به جواب نیز افزایش خواهد یافت.
- **پارامتر تنظیم** ((Regularization Parameter (C): این پارامتر به افزایش تابع حاشیه‌ای و همچنین کم کردن خطای آزمایش کمک می‌کند. مقادیر آن بین ۱ تا ۱۰ تغییر می‌کند و مقدار پیش‌فرض آن نیز ۱۰ می‌باشد. افزایش آن صحت مدل‌سازی را افزایش می‌دهد.
- **پارامتر کرنل** (Kernel Type): تعیین‌کننده نوع کرنل به کار رفته است و یکی از مقادیر چندجمله‌ای (Polynomial)، حلقوی (Sigmoid)، خطی (Linear) و تابع پایه شعاعی (Radial Basis Function) را به خود می‌گیرد.
- **گامای تابع پایه‌ای شعاعی:** این حالت در صورتی مقداردهی می‌شود که نوع کرنل انتخاب شده تابع پایه‌ای شعاعی باشد و مقدار آن بین $۳/K$ و $۶/K$ تغییر می‌کند (K تعداد ستون‌های مدل‌سازی را نشان می‌دهد).
- **گاما:** این حالت در صورتی مقداردهی می‌شود که نوع کرنل انتخاب‌شده، چندجمله‌ای و حلقوی باشد. افزایش آن موجب بالا رفتن صحت مدل‌سازی می‌شود.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

- **اریب (Bias):** این حالت در صورتی مقداردهی می‌شود که نوع کرنل انتخاب‌شده چندجمله‌ای و حلقوی باشد. مقدار پیش فرض آن صفر است.
- **درجه (Degree):** این حالت در صورتی مقداردهی می‌شود که نوع کرنل انتخاب‌شده چندجمله‌ای باشد و بُعد مسئله را نشان می‌دهد. حداکثر مقدار آن ۱۰ خواهد بود.

پس از معرفی پارامترهای مختلف روش ماشین بردارهای پشتیبانی و با تغییر این پارامترها مدل‌سازی را برای حالت‌های مختلف و با کرنل‌های ذکرشده انجام داده که نتایج آن در جدول ۴ آمده است.

جدول ۴- نتایج حاصل از مدل‌سازی با تغییر پارامترهای مختلف برای کرنل‌های مختلف

ردیف	نوع کرنل	نتیجه
۱	تابع پایه‌ای شعاعی	افزایش پارامتر تنظیم (C) منتج به افزایش صحت در صورت ثابت نگه داشتن سایر اجزای مدل می‌شود.
۲	تابع پایه‌ای شعاعی	افزایش پارامتر گاما منتج به افزایش صحت در صورت ثابت نگه داشتن سایر اجزای مدل می‌شود.
۳	خطی	تغییر پارامترها تاثیر چندانی در افزایش صحت مسئله ندارد
۴	حلقوی	تغییر پارامترها تاثیر چندانی در افزایش صحت مسئله ندارد
۵	چندجمله‌ای	افزایش بعد مسئله منتج به افزایش صحت در صورت ثابت نگه داشتن سایر اجزای مدل می‌شود.
۶	چندجمله‌ای	افزایش پارامترهای دیگر در صورت ثابت نگه داشتن سایر پارمترهای مدل تاثیر چندانی در افزایش صحت مدلسازی ندارد.

همچنین در جدول ۵ می‌توان نتایج حاصل از مدل‌سازی کرنل‌های مختلف را با یکدیگر مقایسه کرد (لازم به ذکر است که بهترین نتیجه‌ی مدل‌سازی هر کرنل در این جدول آورده شده است).

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

جدول ۵- مقایسه‌ی نتایج مدل‌سازی برای کرنل‌های مختلف

ردیف	نوع کرنل	صحت مدل‌سازی
۱	تابع پایه‌ای شعاعی	۸۰/۳۹
۲	خطی	۶۱/۴۶
۳	حلقوی	۵۷/۳۷
۴	چندجمله‌ای	۸۲/۰۹

همان‌طور که از جدول ۵ مشخص است بهترین حالت مدل‌سازی برای کرنل‌های تابع پایه‌ای، شعاعی و چندجمله‌ای است و این مسئله نیز به‌دلیل اثر غیر خطی متغیرهای مدل‌سازی می‌باشد. بنابراین رویکردی هیبریدی فازی – ماشین بردارهای پشتیبانی به‌عنوان مدل‌سازی روی‌گردانی مشتریان این مؤسسه انتخاب شده و کرنل‌های تابع پایه‌ای شعاعی و حلقوی نیز به‌عنوان بهترین حالت‌ها در این روش انتخاب می‌شوند.

۷- بحث و نتیجه‌گیری

در این مقاله پس از بررسی مرور ادبیات و مطالعات نظری در حوزه‌ی داده‌کاوی، مدیریت ارتباط مشتریان و روی‌گردانی مشتریان و با تعاملات انجام گرفته با کارشناسان مؤسسه‌ی مالی و اعتباری مهر، مدل‌سازی پیش‌بینی روی‌گردانی مشتریان دارای حساب‌های جاری مؤسسه‌ی مالی و اعتباری مهر به‌عنوان هدف انتخاب گردید. با توجه به این که در مؤسسات مالی و اعتباری و یا بانک‌های ایران تعداد بسیار کمی از مشتریان حساب‌های خود را می‌بندند و اکثراً آن را به‌صورت غیر فعال می‌گذارند لذا برای روی‌گردانی مشتریان مؤسسه‌ی مالی و اعتباری مهر سه نوع متغیر فعال، معمولی و ضعیف در نظر گرفته شد. طبق مصاحبه‌ی انجام‌گرفته با کارشناسان مؤسسه، برای دستیابی به متغیر روی‌گردانی مشتریان به‌صورت مستقیم از سه متغیر تعداد واریز هر مشتری، تعداد برداشت هر مشتری، تعداد حساب‌های باز کرده‌ی هر مشتری استفاده می‌شود. با توجه به تأثیر متغیرهای فوق در تعیین روی‌گردانی مشتریان، از سه رویکرد تک‌متغیره، هیبریدی – منطقی و هیبریدی – فازی برای مدل‌سازی مسئله استفاده شد. مدل‌سازی رویکردهای ذکرشده با سه روش رگرسیون لوژستیک، شبکه‌های عصبی مصنوعی و ماشین بردارهای پشتیبانی انجام گرفت. به‌دلیل ماهیت غیر خطی متغیرها و

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

همچنین تأثیر سه متغیر تعداد حساب، تعداد برداشت و تعداد واریز در تعیین نحوه‌ی روی‌گردانی مشتریان رویکرد هیبریدی – فازی به‌عنوان رویکرد اصلی مسئله معرفی شد. همچنین با توجه به نتایج مسئله و به‌دلیل ماهیت غیر خطی داده‌ها روش ماشین بردارهای پشتیبانی به‌عنوان روش مسئله معرفی شد. در قسمت آخر نیز آنالیز حساسیت بر روی نتایج مدل‌سازی حاصل از رویکرد هیبریدی – فازی و روش ماشین بردارهای پشتیبانی انجام شد و با تغییر پارامترهای مختلف برای کرنل‌های مختلف این نتیجه حاصل شد که دو کرنل تابع پایه‌ای شعاعی و چندجمله‌ای بهترین صحت را در مدل‌سازی مسئله دارند. در نهایت و به‌منظور حفظ و توسعه‌ی گروه مشتریان محقق راهکارهای زیر را برای اجرا در مؤسسه پیشنهاد می‌دهد.

- در نظر گرفتن شرایطی به‌منظور ترغیب مشتریان فعلی مؤسسه در افتتاح حساب جدید؛
- در نظر گرفتن تسهیلاتی به‌منظور افزایش واریز و برداشت حساب؛
- قدردانی از مشتریان فعال؛
- اطلاع‌رسانی و تبلیغات؛
- ارایه‌ی تسهیلاتی به‌منظور افتتاح ۲ حساب و یا بیش‌تر توسط مشتریان جدید.

از جمله نوآوری‌های این مقاله می‌توان به تعریف دسته‌بندی مشتریان بانکی به‌منظور تعیین روی‌گردانی آن‌ها و همچنین استفاده از روش ابتکاری (رویکرد هیبریدی) با توجه به ماهیت غیر خطی داده‌ها اشاره کرد که صحت مدل‌سازی مسئله نیز با حالات تک‌متغیره (که معمولاً در مقالات به کار می‌رود) برابری می‌کند و به‌دلیل در نظر گرفتن همه‌ی عوامل با یکدیگر نیز بر آن‌ها ارجحیت دارد. با توجه به مشکلات موجود در تهیه‌ی داده‌ها از سوی مؤسسه‌ی مالی و اعتباری مهر می‌توان این مسئله را در آینده برای داده‌های بهتر و حتی مؤسسات مالی و اعتباری دیگر تعمیم داد.

توضیحات

1. Churn Prediction
2. Data Mining
3. Decision Boundary
4. Logistic Function

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

5. Customer Relationship Management (CRM)
6. Customer Focused
7. Customer Driven
8. Customer Centric
9. Campaign
10. Voluntary Churn
11. Involuntary Churn
12. Expected churn
13. Rough Sets

۱۴. علت بیش‌تر بودن تعداد رکوردها این است که برخی از مشتری دو و یا تعداد بیشتری حساب در مؤسسه باز کرده‌اند.

مرجع‌ها

[۱] فرتوت، رقیه؛ شهرابی، جمال (۱۳۸۹). ارایه‌ی مدلی به‌منظور شناسایی تأمین‌کنندگان مناسب با استفاده از داده‌کاوی، چهارمین کنفرانس داده‌کاوی، تهران، ایران.

[۲] فتحی، مهدی (۱۳۸۵). کاربرد تکنیک‌های داده‌کاوی در مدیریت ارتباط با مشتری، پایان‌نامه کارشناسی، دانشگاه امیرکبیر، ایران.

[۳] غضنفری، مهدی؛ علیزاده، سمیه و تیمورپور، بابک (۱۳۸۷). داده‌کاوی و کشف دانش، انتشارات دانشگاه علم و صنعت ایران، تهران.

[4] Berry, M.J.A. and Linoff, G.S. (2004). *Data Mining Techniques for Marketing, Sales and Customer Relationship Management*, 2nd edition, Wiley Publishing Inc, Indiana, USA.

[5] Bigus, J.P. (1996). *Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support Paperback*, McGraw-Hill, Inc. Hightstown, NJ, USA.

[6] Coussement, K. and Poel, D.Z. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques, *Expert Systems with Applications*, **34**, 313–327.

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

[7] Chie-Ping, Wei; I-Tang Chiu, Turning, (2002), Telecommunications call details to churn prediction: a data mining approach, *Expert Systems with Applications*, **23**, 103-112.

[8] Chih-Fong, Tsai and Yu-Hsin, Lu (2009). Customer churn prediction by hybrid neural networks, *Expert Systems with Applications*, **36**, 12547–12553.

[9] Han, Aiwei and Micheline, Kamber (2006). *Data Mining Concepts and Techniques*, second edition, Elsevier Inc, USA.

[10] Hung, S.Y., Yen, D.C. and Wang, H.Y. (2006). Applying data mining to telecom churn management, *Expert Systems with Applications*, **31**, 515–524.

[11] Jana, J. and Matjaz, G. (2011). *Data Mining Techniques for Explaining Social Events*, Knowledge-oriented applications in data mining.

[12] Kantardzic, M. (2003). *Data Mining: Concepts, Models, Methods and Algorithms*, IEEE Press & John Wiley, Indiana, USA.

[13] Lin, C.S., Tzeng, T.Z. and Chin, Y.C. (2011). Combined rough set theory and flow network graph to predict customer churn in credit card accounts; *Expert Systems with Applications*, **38**, 8–15.

[14] Ngai, E.W.T., Xiu, L. and Chau, D.C.K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification, *Expert Systems with Applications*, **36**, 2592-2602.

[15] Pendharkar, P.C. (2009). Genetic algorithm based neural network approaches for predicting churn in cellular wireless network services, *Expert Systems with Applications*, **36**, 6714–6720.

حمیدرضا احمدی خالدی
دانشجوی کارشناسی ارشد رشته‌ی مهندسی سیستم‌های اقتصادی - اجتماعی
تهران، بزرگراه شهید بابایی، بعد از پل لشگرک، دانشگاه جامع امام حسین (ع).
رایانشانی: hr.a.khaledi@gmail.com

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]

علی‌محمد احمدوند
دکتری مدیریت استراتژیک
تهران، بزرگراه شهید بابایی، بعد از پل لشگرک، دانشگاه جامع امام حسین (ع).
رایانشانی: amahmadvand@mahanmail.com

[Downloaded from ijoss.srtc.ac.ir on 2026-05-20]